



OPEN

A novel classification regression method for gridded electric power consumption estimation in China

Mulin Chen^{1,2}, Hongyan Cai¹✉, Xiaohuan Yang¹ & Cui Jin³

Spatially explicit information on electric power consumption (EPC) is crucial for effective electricity allocation and utilization. Many studies have estimated fine-scale spatial EPC based on remotely sensed nighttime light (NTL). However, the spatial non-stationary relationship between EPC and NTL at prefectural level tends to be overlooked in existing literature. In this study, a classification regression method to estimate the gridded EPC in China based on imaging NTL via a Visible Infrared Imaging Radiometer Suite (VIIRS) was described. In addition, owing to some inherent omissions in the VIIRS NTL data, the study has employed the cubic Hermite interpolation to produce a more appropriate NTL dataset for estimation. The proposed method was compared with ordinary least squares (OLS) and geographically weighted regression (GWR) approaches. The results showed that our proposed method outperformed OLS and GWR in relative error (RE) and mean absolute percentage error (MAPE). The desirable results benefited mainly from a reasonable classification scheme that fully considered the spatial non-stationary relationship between EPC and NTL. Thus, the analysis suggested that the proposed classification regression method would enhance the accuracy of the gridded EPC estimation and provide a valuable reference predictive model for electricity consumption.

Electric power is essential for modern life and society. Having a clear understanding of electric power consumption (EPC) is important for infrastructure planning, energy allocation and environmental protection. However, given the commercial and confidential nature of electricity supply, detailed EPC information is difficult to obtain. Current EPC data are available mainly at administrative units, however, such data rarely meet the needs for fine scale application because of the coarse resolution. Furthermore, the data are also difficult to integrate with spatial data at other scales, and this can limit interdisciplinary applications of EPC. In contrast, gridded EPC data are more efficient in fine scales research. Therefore, developing a convenient and reliable approach for fine scale estimation of EPC is urgently needed.

Remotely sensed nighttime light (NTL) has long been considered as a direct measure of socio-economic development, and has served as primary data for a wide range of human-based research^{1,2}, such as population modelling^{3,4}, GDP estimation^{5,6}, urban expansion mapping^{7,8}, and EPC estimation^{9,10}. Many studies have demonstrated a strong correlation between NTL and EPC at multiple levels, and consequently regression models for such estimation have been built^{1,9,11–15}. However, the relationship between NTL and EPC varies across areas, owing to the local socioeconomic diversity¹⁶. In other words, the relationship between NTL and EPC is spatial non-stationary¹⁷, which tends to be overlooked in existing literature, where global models are often used^{10,18–20}. Thus, from the spatial non-stationary perspective, local regression would be more suitable for EPC estimation. As pointed out by Shi et al.²¹, a single regression model for all administrative units is susceptible to incorrect estimation and an appropriate subdivision of administrative units could improve the estimation accuracy²². Some researchers have divided their study area into several sub-regions based on factors, such as geographic location^{12,23}, income consideration²⁴, and city function¹⁷. Yet, such classifications are limited in terms of accuracy, as they do not fully consider the complexity of interactions between NTL and EPC²⁴. For instance, a classification based on economic factors may not sufficiently distinct the relationship between NTL and EPC, given that cities at the same level of economic development can have different energy consumption patterns and NTL levels depending on their key industries¹⁷. Hence, exploring an appropriate classification basis from the perspective of the relationship between EPC and NTL would provide a better understanding and a new insight for gridded EPC estimation.

¹State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A, Datun Road, Chaoyang District, Beijing 100101, China. ²University of Chinese Academy of Sciences, Beijing, China. ³College of Urban and Environment, Liaoning Normal University, Dalian, China. ✉email: caihy@igsrr.ac.cn

In previous literature, the NTL used for EPC estimation was collected mainly from several sources including the Defense Meteorological Satellite Program Operational Linescan System (DMSP/OLS)^{23,25–27} and the day/night band (DNB) on the Visible Infrared Imaging Radiometer Suite (VIIRS) on-board the Suomi National Polar-orbiting Partnership satellite²¹. Many studies have documented that estimation accuracy would be reduced when using DMSP/OLS NTL, due to the saturation and coarse resolution^{21,28,29}. In contrast, VIIRS/DNB NTL has finer resolution and better data quality than DMSP/OLS^{29–32}, and consequently would provide a more reliable gridded EPC estimation^{21,33}. Even so, some imperfections in the current VIIRS/DNB NTL (details given in the “Materials” section) are undesirable for gridded estimation³⁴. Given these problems, more data processing steps are needed to remove the imperfections before estimation.

Against this background, the aim of this study was to develop a classification regression approach for high accuracy gridded EPC estimation. To achieve this goal, a more suitable NTL dataset for EPC estimation was first produced, based on the original VIIRS/DNB NTL. Secondly, a classification estimation scheme for the gridded EPC estimation was proposed. At the same time, apart from the proposed approach, the gridded EPC was also estimated using ordinary least squares (OLS) and geographically weighted regression (GWR). Finally, the accuracy in estimation of these three methods were compared.

Materials

The VIIRS/DNB NTL data provided by the National Oceanic and Atmospheric Administration’s National Centers for Environmental Information (NOAA/NCEI, <https://ngdc.noaa.gov/eog/download.html>) were used in this study. This dataset has widely used for socioeconomic indicators spatialization owing to its fine resolution^{21,31,35}. However, the data are vulnerable to stray light, resulting in emergence of distorted pixels in the mid and high latitude areas, especially in the summer^{32,34,36,37}. To address this problem, two versions of the NTL are provided by NOAA, denoted by the filenames VCM (VIIRS Cloud Mask) and VCMSL (VIIRS Cloud Mask with Stray Light), respectively. The VCM version eliminates all the pixels that were contaminated by stray light, resulting in numerous missing pixels in the dataset. The VCMSL version corrects the distorted pixels through implementation of a stray light correction procedure³⁸, but the data quality may be reduced in these regions^{29,37,39}. Thus, both the VCM and the VCMSL data have some limitations in respect of data quality. In the present study, a more appropriate NTL dataset was produced through interpolation, which was then used for estimation purposes.

Statistical EPC data of 2015 at the prefectural level were obtained from the China City Statistical Yearbook released by the National Statistics Bureau of China. The statistical data include EPC, industrial electricity consumption (IEC) and household electricity consumption.

The land use and cover change (LUCC) data of 2015 were obtained from the land use database at a scale of 1:100,000, and collected from the Resources and Environmental Scientific Data Center (RESDC) of the Chinese Academy of Sciences (CAS). These data were interpreted by a human–computer interactive method from the Landsat TM/ ETM+ images with an accuracy greater than 90%^{40,41}.

Methods

The methodology flow chart is displayed in Fig. 1. Given that the EPC activities occur mainly on built-up land, the EPC estimation was only performed for built-up land. First, the built-up land was extracted from the LUCC data. Second, a new NTL dataset (hereafter called the interpolated NTL) was obtained by interpolation. The annual total nighttime light (TNL) data were then composited by the interpolated NTL. Third, the gridded EPC was estimated by OLS, classification regression, and GWR, respectively. Finally, the accuracy was evaluated and compared at the prefectural level.

Improved VIIRS/DNB NTL data production. The original monthly VIIRS/DNB NTL data are subject to stray light, and do not have background noise or ephemeral light filtered out, causing emergence of outliers such as negative pixels and abnormally high-valued pixels in the imagery⁴². Such negative effects will reduce the accuracy of estimation²¹. Therefore, an improved VIIRS/DNB NTL data, which has the outliers excluded and the missing data interpolated, was produced prior to estimation.

The threshold method was used to eliminate outliers. Theoretically, the highest level of artificial light should be emitted from the most developed areas⁴³. Thus, the maximal radiation in the most developed cities in China (Beijing, Shanghai, Guangzhou and Shenzhen), which was 264.47 nW cm⁻² sr⁻¹, was determined as the threshold, and pixels with a radiation value larger than 264.47 nW cm⁻² sr⁻¹ were reassigned to that value. Meanwhile, negative-valued pixels were reassigned to zero²¹.

After filtering the outliers, the missing values were interpolated by a temporal interpolation method, as the missing pixels tend to exist over a large area. According to the study of Chen et al.⁴⁴, the cubic Hermite method was used for interpolation. The monthly VCM data from August 2014 to April 2015, and from September 2015 to April 2016 were served as the initial data in interpolation proceeding at the pixel level. Subsequently, the annual TNL was calculated based on the interpolated NTL data at the pixel level (Eq. (1)), that is:

$$TNL_i = \sum_{j=1}^{12} NTL_{ij} \quad (1)$$

where NTL_{ij} is the monthly radiation value of the i th pixel in the j th month, and TNL_i is the annual total radiation value of the i th pixel.

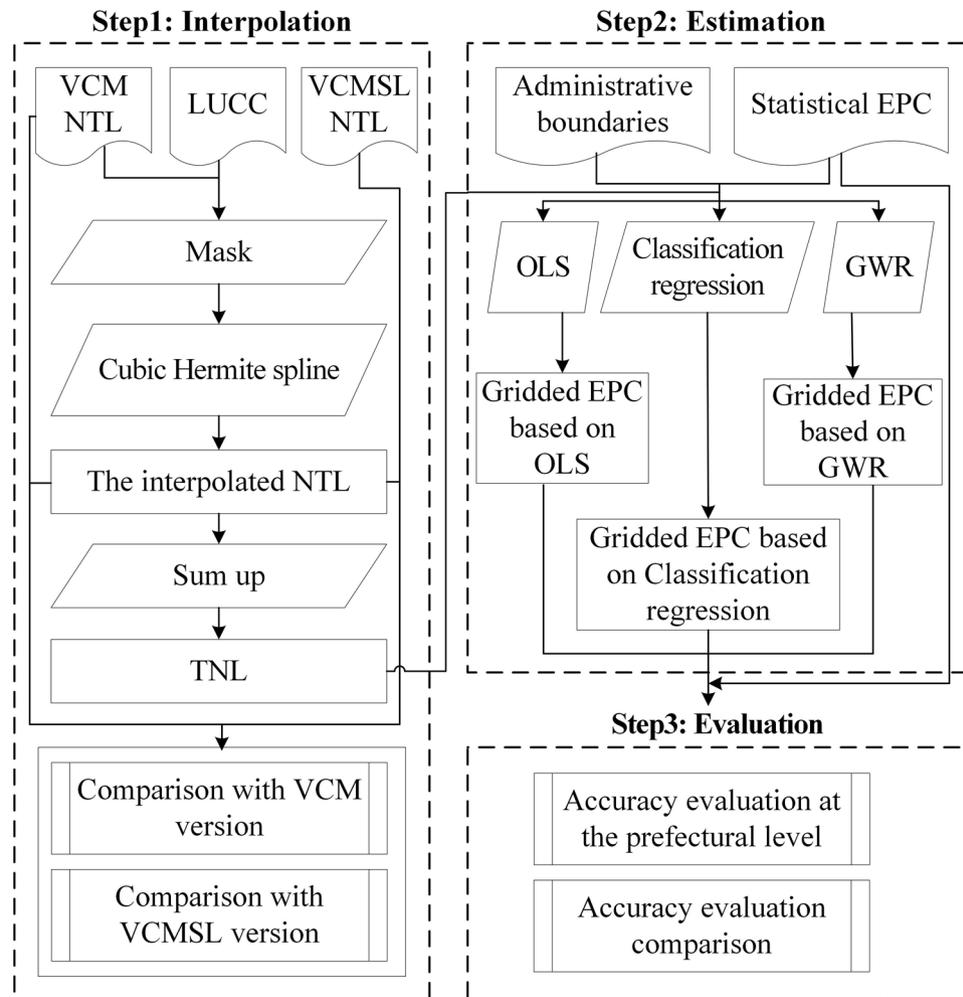


Figure 1. Flow chart for production and evaluation of the gridded EPC.

Regression by OLS, GWR, classification regression. Statistical EPC and TNL were regressed by OLS, GWR, and classification regression, respectively. Due to the lack of EPC statistical data in some cities, 259 cities were used for regression. OLS was given in Eq. (2).

$$EPC_s = a \sum TNL_i + b + e \quad (2)$$

where EPC_s was the statistical data of a certain city, $\sum TNL_i$ was the amount of TNL for all the pixels in the corresponding areas, a was the regression coefficient, b was the intercept, and e was the error term. In previous studies, the intercept was assumed to be zero^{10,18,24,45}. However, since NTL is not fully indicative of EPC, the intercept accounts for non-light emitting EPC. Hence, the intercept was presumed to be non-zero in this study and the F -test was applied to achieve statistical significance.

In consideration of the spatial spillover effect on EPC⁴⁶, GWR was also adopted. GWR is an extension of linear regression model⁴⁷. It uses some distance-based weighting function to permit the coefficients to vary locally within the bandwidth⁴⁸. GWR was performed by ArcGIS 10.2 and the bandwidth was determined by AICc⁴⁹.

The classification regression was built based on OLS. In order to maximize the stationary aspect of the relationship between EPC and NTL, the classification regression grouped the cities into three patterns based on the 95% confidence interval of the OLS regression coefficients, and then each pattern was regressed by Eq. (2), respectively. The three patterns were: the upper pattern, which consists of cities that were over the 95% confidence interval (72 cities, blue points in Fig. 2b); the middle pattern, which consists of cities that were within the confidence interval (78 cities, green points in Fig. 2b); and the down pattern, which consists of cities that were under the confidence interval (109 cities, red points in Fig. 2b). In general, the upper pattern displayed a low TNL with a relative high EPC, the down pattern featured a high TNL with a relative low EPC, and the middle pattern was in between these two. EPC and TNL spatial distributions of the three patterns were displayed in Fig. 3.

Gridded EPC estimation. The gridded EPC for the three assessing methods was calculated using Eqs. (3)–(5). It was assumed that all the grids in a certain city had a segmented intercept based on their weight. Further-

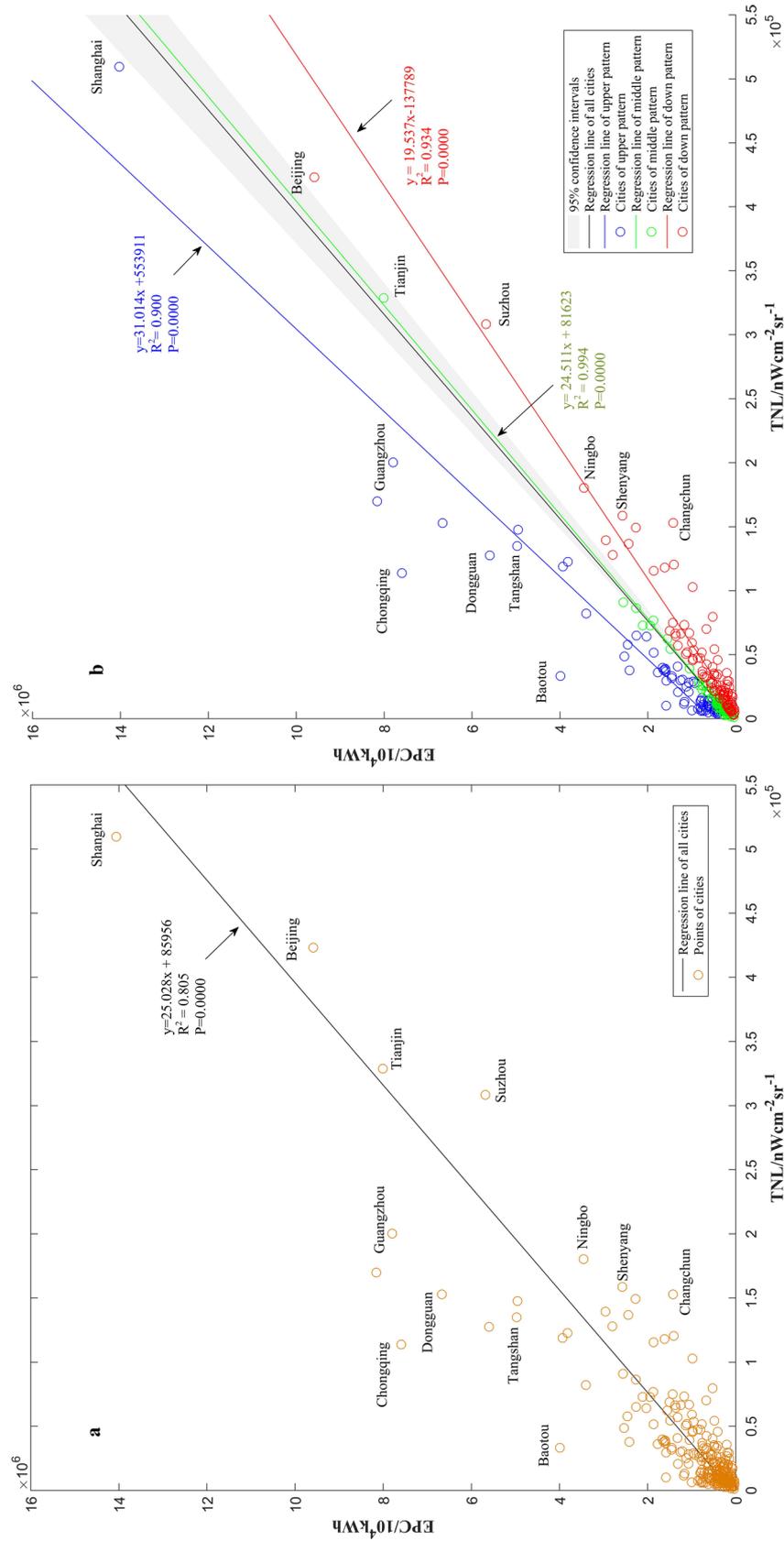


Figure 2. Regression of TNL and EPC by OLS and the proposed classification regression. **(a)** OLS. **(b)** Classification regression.

more, IEC varies significantly between the industrial and non-industrial regions within a city. Thus, this difference was taken into consideration in the estimation.

$$EPC_{Gi} = \begin{cases} aTNL_i + b_1 \frac{TNL_i}{\sum TNL_i} & \text{industrial region} \\ aTNL_i + b_2 \frac{TNL_i}{\sum TNL_i} & \text{non-industrial region} \end{cases} \quad (3)$$

$$b_1 = b \frac{IEC_s}{EPC_s} \quad (4)$$

$$b_2 = b - b_1 \quad (5)$$

where EPC_{Gi} was the estimated EPC of the i th grid in a certain city. The values of a and b were obtained from the three regression methods. IEC_s and EPC_s represented the statistical IEC and EPC of the city. TNL_i was the annual total nighttime light radiation of the i th grid, and $\sum TNL_i$ was the city's aggregated TNL.

Accuracy evaluation. The accuracy at the prefectural level was evaluated using relative error (RE) (Eq. (6)) and mean absolute percentage error (MAPE) (Eq. (7)).

$$RE = \frac{EPC_e - EPC_s}{EPC_s} \times 100\% \quad (6)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{EPC_{ei} - EPC_{si}}{EPC_{si}} \right| \times 100\% \quad (7)$$

where EPC_e represented the estimated EPC and EPC_s represented the statistical EPC. i represented the i th city. n represented the total number of cities and equaled to 259 in present study. The closer RE and MAPE are to zero, the more accurate the estimation is.

Results

Evaluation of the interpolated NTL data. The monthly TNL time series for the original VCM, VCMSL and the interpolated NTL data at the national scale were compared and shown in Fig. 4. The TNL data for the VCM version in the summer months were much lower than those in other months, resulting in dramatic fluctuations in the time series curve. In contrast, the VCMSL and interpolated versions were smoother. If a certain region did not suffer from any severe events, such as natural disasters or power outages, but its TNL fluctuated dramatically, it could be inferred that its NTL data did not reflect the real socio-economic conditions^{50,51}, and, therefore, were inappropriate for robust socio-economic research. From this standpoint, the interpolated NTL data are more suitable for estimation of the EPC than the VCM data.

As shown in Fig. 4, there was only a slight difference between the VCMSL version and the interpolation version at the national level. Thus, a further comparison of these two versions was performed at the prefectural level. The results demonstrated that the interpolated NTL data performed better than the VCMSL product in the mid and high latitude areas. The TNL values for some cities in the VCMSL version appeared to be susceptible to sudden sharp drops in certain months, but this feature did not show up in the interpolated NTL data. For reporting purposes, we defined cities where the TNL values suddenly became zero as being atypical cities. A latitudinal statistic was used to calculate the frequency of the atypical cities (Fig. 5). Since a high ratio of 50% came from the 45–50° N (13 cities) and 50–55° N (66.7%, 2 cities) regions, the atypical cities tended to occur in the mid-high latitude regions where stray light shows severe contamination. In comparison, the atypical cities were not evident in the interpolated NTL dataset. This difference in the respective NTL datasets most likely results from the different processes used to produce the two NTL versions. The VCMSL NTL was produced by a stray light correction procedure⁵², and, consequently, the mid-high latitude areas, where there was severe stray light contamination, most likely was susceptible to this problem³⁴, whereas the missing pixels were corrected based on the long, historical, time series data in the interpolated NTL, which could effectively avoid dramatic TNL time series changes.

Evaluation of the gridded EPC estimation. The gridded EPC based on classification regression was shown in Fig. 6. The EPC was concentrated in eastern China. The high EPC areas were centered on megacities and provincial capitals, especially Shanghai, Guangzhou and Beijing. The low TNL areas were mostly distributed in rural regions and less developed satellite cities. Within a city, the highest EPC tended to occur mainly at airports, railway stations and business districts.

The accuracy of the evaluated results at the prefectural level is shown in Table 1. There were 145 cities whose RE values fell into the [−25%, 25%] bin when adopting the classification estimation, while there were only 72 cities in that bin for OLS and 86 for GWR. Furthermore, the number of cities for which the absolute RE values

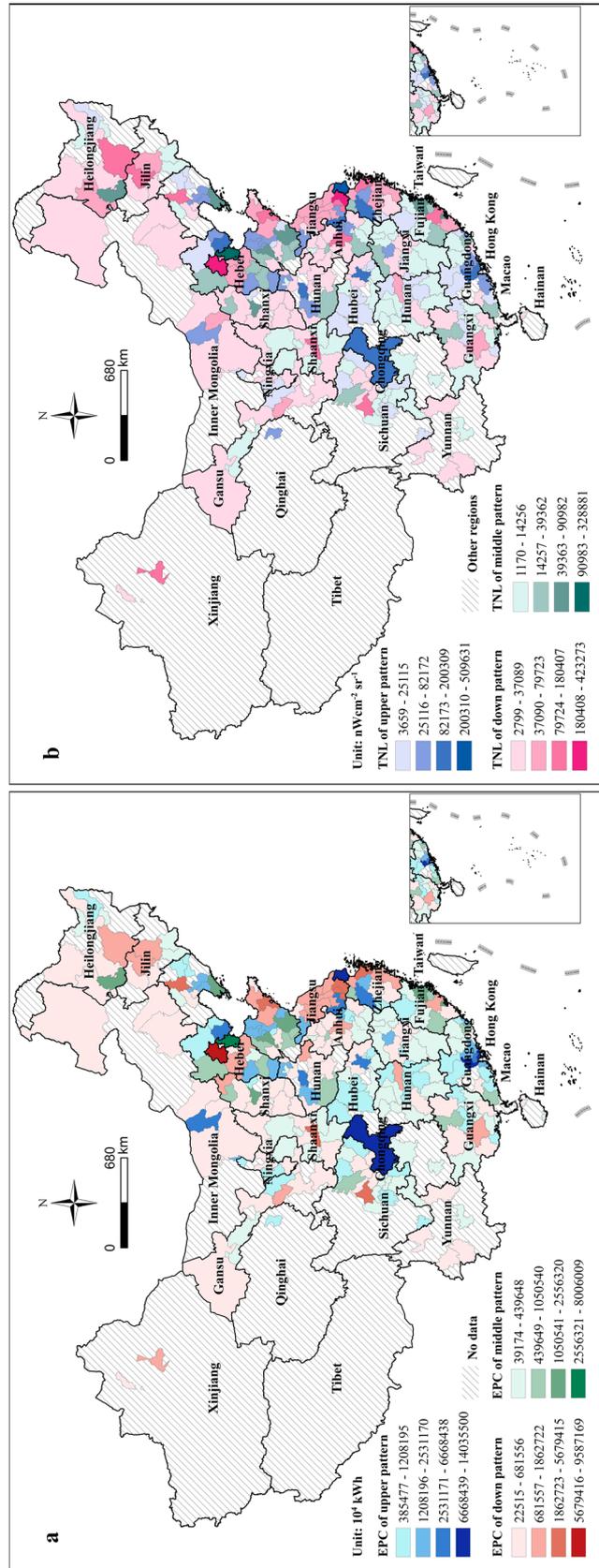


Figure 3. Distributions of the three observed patterns. (a) EPC in the three observed patterns. (b) TNL in the three observed patterns. The map was generated by ArcGIS 10.2 (<https://www.arcgis.com/index.html>).

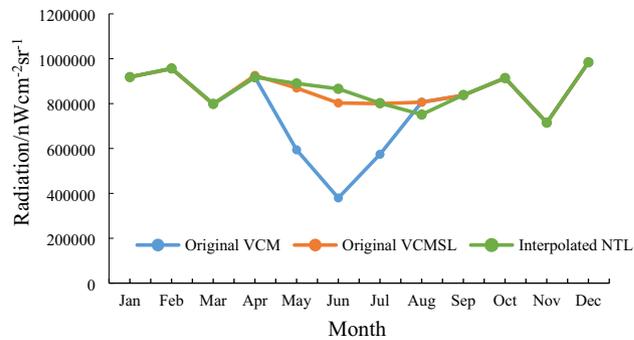


Figure 4. Monthly total nighttime light (TNL) time series at the national level for the original visible infrared imaging radiometer suite cloud mask (VCM) version, the original visible infrared imaging radiometer suite cloud mask with stray light (VCMSL) version and the interpolated nighttime light (NTL) data.

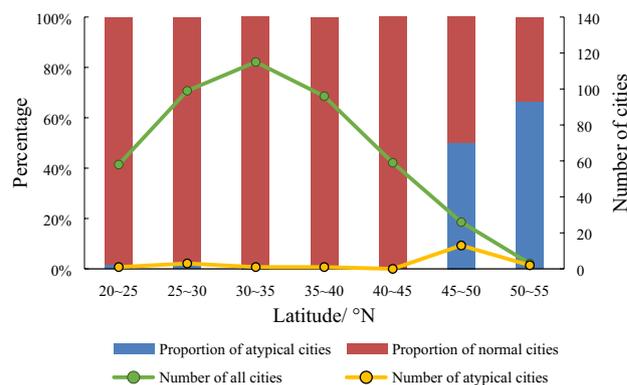


Figure 5. Frequency of occurrence of the sudden TNL drop for the visible infrared imaging radiometer suite cloud mask with stray light (VCMSL) product. Cities that had the sudden drop were called the atypical cities. The “number of all cities” is the total number of cities within the corresponding latitude zone.

exceeded 50% decreased from over a hundred to 62 after classification. The lowest MAPE values for the proposed classification estimation also demonstrated that the approach resulted in a significant improvement in accuracy compared to OLS and GWR.

Discussion

Different from most previous literature, a novel classification regression method that fully considered the spatial non-stationary relationship between the EPC and NTL was described. The results indicated that the proposed method performed best among the three assessing methods (Table 1). Compared to OLS, the relationship between the EPC and NTL was stronger for the classification regression, as revealed by the R^2 values (0.900, 0.994 and 0.934 vs. 0.805). Moreover, a previous study by Shi et al.²¹ also confirmed this finding. In that study, the VIIRS/DNB NTL data were also applied to gridded EPC estimation based on OLS. The RE values in the present work were significantly reduced compared to that study, notably in Anhui, Hebei, Heilongjiang and Shaanxi provinces. In addition, GWR is powerful for spatial non-stationary parameters regression⁵³. However, the Moran’s bivariate spatial autocorrelation index of statistical EPC and TNL was only 0.070. It was evident that the spatial autocorrelation between EPC and TNL was weak, in other words, the distance between cities has little influence on the regression. Thereby, the distance-based weighting scheme was not likely to render a better performance than the proposed method.

The favorable results benefited mainly from a reasonable classification scheme being realized. The classification basis is important to high accuracy estimation. In a study of Li et al.¹⁷, cities were grouped into technology and education cities, industrial cities, and service-oriented cities. However, the classification basis based on city function still have some limitations on ensuring the spatial stationary relationship between EPC and TNL. For example, Changchun, Shenyang, Baotou, Dongguan were the famous industrial cities in China, but differed greatly in the relationship between EPC and TNL (Fig. 2). In part because EPC used for different purposes would emit different levels of light. Industrial electricity consumed for electrolysis, heating and other energy-intensive processes, has a little light emitted²², whereas industrial electricity consumed for oil, natural gas, and coal mineral resources mining, has bright light emitted due to the mining lights and industrial fires²². It leads to different key industries playing a role in the relationship between EPC and TNL¹⁷. The classification basis based on city

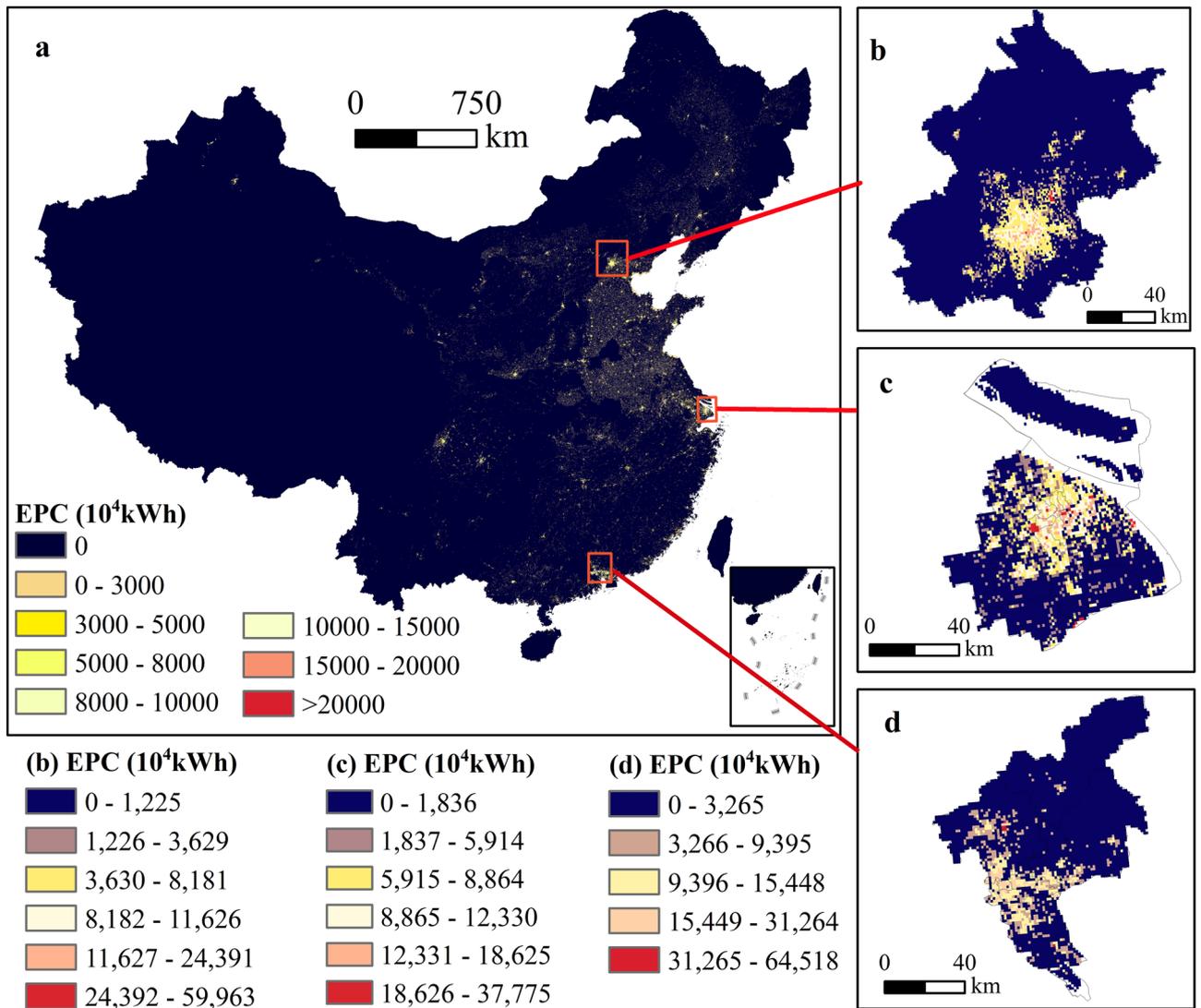


Figure 6. Distribution of the gridded EPC based on the classification regression method. (a) Spatial distribution of gridded EPC in China. Gridded EPC distribution in Beijing, Shanghai and Guangzhou is displayed in (b–d), respectively. The map was generated by ArcGIS 10.2 (<https://www.arcgis.com/index.html>).

	OLS	GWR	Classification regression
$ RE \leq 25\%$	72	86	145
$25\% < RE \leq 50\%$	71	70	52
$50\% < RE $	116	103	62
MAPE	57.1%	73.9%	38.0%

Table 1. Number of cities in the different relative error (RE) bins and mean absolute percentage error (MAPE) values for each of the three estimation methods.

function, thus, was insufficient for spatial stationary classification of EPC and TNL; instead, the classification scheme we proposed effectively avoided ambiguous classification.

The threshold of confidence interval may cause uncertainty of the classification method. Thus, apart from 95% confidence interval, 90% and 99% confidence intervals were also served as the threshold, respectively. As the comparison results shown, whether the regression coefficients, or the MAPE values, the differences were slight (Fig. 7). It was because for the most cities, their patterns did not change as the confidence intervals changed, and, consequently, the classification regression method was seldom affected by the threshold of confidence interval. Thus, from this perspective, the classification regression method was robust.

Inevitably, some limitations existed when estimating the gridded EPC based on the VIIRS/DNB NTL. First, the overpass time of the SNPP satellite is after midnight and near 01:30 am⁵⁴. The outdoor artificial light is low at

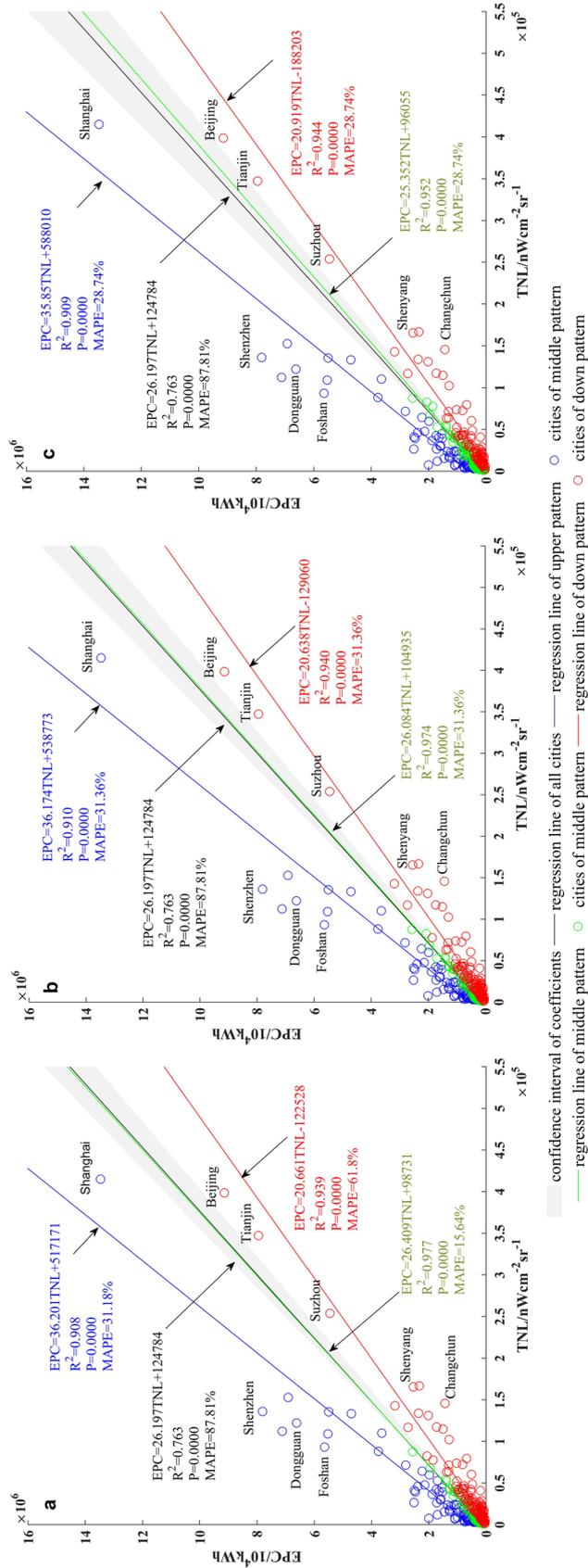


Figure 7. Classification regression based on different confidence intervals. **(a)** 90% confidence interval. **(b)** 95% confidence interval. **(c)** 99% confidence interval.

this time, leading to much EPC that is not reflected in the NTL dataset. Second, monthly changes in vegetation and snow cover are also contained in the current VIIRS/DNB NTL dataset. These physical changes are known to be undesirable for socio-economic studies⁵⁵. Despite this, given that China is mainly located in the low-mid latitude regions, most parts of the country would not be impacted by that seasonal noise⁵⁶, and, consequently, the EPC estimation result is considered reliable. However, for high latitude areas, researchers should give close attention to the physical effects of using the VIIRS/DNB NTL data for EPC estimation and other socio-economic studies. Besides, NTL is just one of components of EPC, other factors such as GDP, population, and urbanization also impact on EPC⁴⁶. However, coarse statistical GDP and population data are unfeasible for fine scale estimation. Furthermore, using estimated gridded GDP and population product would introduce inevitable errors. Thus, GDP, population, and urbanization were not considered when EPC estimation in the present study.

Conclusions

Artificial light is closely related to the EPC, therefore the NTL is a superior indicator of the EPC and is widely used for gridded EPC estimation^{9,10}. However, the current VIIRS/DNB monthly NTL data did not meet high estimation accuracy requirements due to the impact of missing pixels and ephemeral light⁵⁷. Thus, in this study, a more suitable interpolated NTL dataset was first proposed based on the cubic Hermite method, then the data were subjected to regression analysis. Moreover, few studies took the spatial non-stationary relationship between EPC and NTL into account in the estimation. In contrast to previous research^{12,58}, this study has developed a classification regression method that based on the relationship between EPC and NTL.

Generally, the proposed approach resulted in a significant improvement in accuracy over OLS and GWR. The R^2 values for the proposed method were greater than the OLS (0.900, 0.994 and 0.934 vs. 0.805). Furthermore, with regard to the RE values, the proposed method had more cities (145 vs. 72 and 86) that fell into the [−25%, 25%] bin than OLS and GWR. In addition, by comparing estimation results with previous literature, results for this study showed a higher accuracy in many provinces. In conclusion, the comparative results for this study indicate that the approach used gives new insights into the explicit spatial distribution estimation of EPC based on the NTL. In addition to gridded EPC estimation, the proposed method also has a great potential for other socio-economic indicators estimation (e.g., GDP and carbon dioxide emission).

Received: 6 January 2020; Accepted: 16 October 2020

Published online: 29 October 2020

References

1. Proville, J., Zavala-Araiza, D. & Wagner, G. Night-time lights: A global, long term look at links to socio-economic trends. *PLoS ONE* **12**, e0174610 (2017).
2. Kyba, C. *et al.* High-resolution imagery of earth at night: New sources, opportunities and challenges. *Remote Sens.* **7**, 1–23 (2015).
3. Tan, M. *et al.* Modeling population density based on nighttime light images and land use data in China. *Appl. Geogr.* **90**, 239–247 (2018).
4. Wang, L. *et al.* Mapping population density in China between 1990 and 2010 using remote sensing. *Remote Sens. Environ.* **210**, 269–281 (2018).
5. Zhao, M. *et al.* GDP spatialization and economic differences in south China based on NPP-VIIRS nighttime light imagery. *Remote Sens.* **9**, 673 (2017).
6. Zhu, X., Ma, M., Yang, H. & Ge, W. Modeling the spatiotemporal dynamics of gross domestic product in China using extended temporal coverage nighttime light data. *Remote Sens.* **9**, 626 (2017).
7. Huang, X., Schneider, A. & Friedl, M. A. Mapping sub-pixel urban expansion in China using MODIS and DMSP/OLS nighttime lights. *Remote Sens. Environ.* **175**, 92–108 (2016).
8. Xie, Y. & Weng, Q. Updating urban extents with nighttime light imagery by using an object-based thresholding method. *Remote Sens. Environ.* **187**, 1–13 (2016).
9. Dugoua, E., Kennedy, R. & Urpelainen, J. Satellite data for the social sciences: Measuring rural electrification with night-time lights. *Int. J. Remote Sens.* **39**, 2690–2701 (2018).
10. Pan, J. & Li, J. Spatiotemporal dynamics of electricity consumption in China. *Appl. Spatial Anal. Policy* **12**, 1–28 (2017).
11. Shi, K. *et al.* Detecting spatiotemporal dynamics of global electric power consumption using DMSP-OLS nighttime stable light data. *Appl. Energy* **184**, 450–463 (2016).
12. Xie, Y. & Weng, Q. Detecting urban-scale dynamics of electricity consumption at Chinese cities using time-series DMSP-OLS (Defense Meteorological Satellite Program-Operational Linescan System) nighttime light imageries. *Energy* **100**, 177–189 (2016).
13. Shi, K., Yu, B., Huang, C., Wu, J. & Sun, X. Exploring spatiotemporal patterns of electric power consumption in countries along the Belt and Road. *Energy* **150**, 847–859 (2018).
14. Fobi, S., Deshpande, V., Ondiek, S., Modi, V. & Taneja, J. A longitudinal study of electricity consumption growth in Kenya. *Energy Policy* **123**, 569–578 (2018).
15. Xiao, H. *et al.* Spatio-temporal simulation of energy consumption in China's provinces based on satellite night-time light data. *Appl. Energy* **231**, 1070–1078 (2018).
16. Hu, T. & Huang, X. A novel locally adaptive method for modeling the spatiotemporal dynamics of global electric power consumption based on DMSP-OLS nighttime stable light data. *Appl. Energy* **240**, 778–792 (2019).
17. Li, S. *et al.* City type-oriented modeling electric power consumption in China using NPP-VIIRS nighttime stable light data. *Energy* **189**, 116040 (2019).
18. Cao, X., Wang, J., Chen, J. & Shi, F. Spatialization of electricity consumption of China using saturation-corrected DMSP-OLS data. *Int. J. Appl. Earth Obs. Geoinf.* **28**, 193–200 (2014).
19. Chi, Z., Zhou, K., Yang, S. & Zhen, S. On electricity consumption and economic growth in China. *Renew. Sustain. Energy Rev.* **76**, 353–368 (2017).
20. Elvidge, C. D. *et al.* Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *Int. J. Remote Sens.* **18**, 1373–1379 (1997).
21. Shi, K. *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.* **6**, 1705–1724 (2014).
22. Dai, Z., Hu, Y. & Zhao, G. The suitability of different nighttime light data for GDP estimation at different spatial scales and regional levels. *Sustainability* **9**, 305 (2017).

23. He, C., Ma, Q., Li, T., Yang, Y. & Liu, Z. Spatiotemporal dynamics of electric power consumption in Chinese mainland from 1995 to 2008 modeled using DMSP/OLS stable nighttime lights data. *J. Geogr. Sci.* **22**, 125–136 (2012).
24. Xie, Y. & Weng, Q. World energy consumption pattern as revealed by DMSP-OLS nighttime light imagery. *GISci. Remote Sens.* **53**, 265–282 (2016).
25. 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 Plann.* **106**, 62–72 (2012).
26. Zhao, N., Ghosh, T. & Samson, E. L. Mapping spatio-temporal changes of Chinese electric power consumption using night-time imagery. *Int. J. Remote Sens.* **33**, 6304–6320 (2012).
27. 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.* **31**, 4459–4480 (2010).
28. Shi, K. *et al.* Evaluating spatiotemporal patterns of urban electricity consumption within different spatial boundaries: A case study of Chongqing, China. *Energy* **167**, 641–653 (2019).
29. Zhao, M. *et al.* Applications of satellite remote sensing of nighttime light observations: Advances, challenges, and perspectives. *Remote Sens.* **11**, 1971 (2019).
30. Baugh, K. E., Hsu, F. C., Elvidge, C. D. & Zhizhin, M. Nighttime lights compositing using the VIIRS day-night band: Preliminary results. *Proc. Asia-Pac. Adv. Netw.* **35**, 70–86 (2013).
31. Elvidge, C. D., Baugh, K. E., Zhizhin, M. N. & Hsu, F. C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proc. Asia-Pac. Adv. Netw.* **35**, 62–69 (2013).
32. Elvidge, C. D., Baugh, K. E., Zhizhin, M. N., Hsu, F. C. & Ghosh, T. VIIRS night-time lights. *Int. J. Remote Sens.* **38**, 5860–5879 (2017).
33. Jing, X., Shao, X., Cao, C., Fu, X. & Yan, L. Comparison between the Suomi-NPP day-night band and DMSP-OLS for correlating socio-economic variables at the provincial level in China. *Remote Sens.* **8**, 17 (2016).
34. Lee, S. & Cao, C. Suomi NPP VIIRS day/night band stray light characterization and correction using calibration view data. *Remote Sens.* **8**, 138 (2016).
35. Cao, J., Chen, Y., Wilson, J. P., Tan, H. & Xu, Z. Modeling China's prefecture-level economy using VIIRS imagery and spatial methods. *Remote Sens.* **12**, 839 (2020).
36. Liang, C. K., Mills, S., Hauss, B. I. & Miller, S. D. Improved VIIRS day/night band imagery with near-constant contrast. *IEEE Trans. Geosci. Remote Sens.* **52**, 6964–6971 (2014).
37. Zhao, X., Li, D., Li, X., Zhao, L. & Wu, C. Spatial and seasonal patterns of night-time lights in global ocean derived from VIIRS DNB images. *Int. J. Remote Sens.* **39**, 8151–8181 (2018).
38. Chen, H., Xiong, X., Sun, C., Chen, X. & Chiang, K. Suomi-NPP VIIRS day-night band on-orbit calibration and performance. *J. Appl. Remote Sens.* **11**, 036019 (2017).
39. Chen, M. L. & Cai, H. Y. VIIRS/DNB monthly and yearly nighttime light dataset in Beijing-Tianjin-Hebei region (2013–2018). *J. Glob. Change Data Discov.* **3**, 376–381 (2019).
40. Liu, W., Liu, J., Kuang, W. & Ning, J. Examining the influence of the implementation of major function-oriented zones on built-up area expansion in China. *J. Geogr. Sci.* **27**, 643–660 (2017).
41. Jiyuan, L. *et al.* The land use and land cover change database and its relative studies in China. *J. Geogr. Sci.* **12**, 275–282 (2002).
42. Ma, T., Zhou, C., Pei, T., Haynie, S. & Fan, J. Responses of Suomi-NPP VIIRS-derived nighttime lights to socioeconomic activity in China's cities. *Remote Sens. Lett.* **5**, 165–174 (2014).
43. Ma, W. & Li, P. An object similarity-based thresholding method for urban area mapping from visible infrared imaging radiometer suite day/night band (VIIRS DNB) data. *Remote Sens.* **10**, 263 (2018).
44. Chen, M. L. & Cai, H. Y. Interpolation methods comparison of VIIRS/DNB nighttime light monthly composites: A case study of Beijing. *Prog. Geogr.* **38**, 126–138 (2019).
45. Letu, H., Nakajima, T. Y. & Nishio, F. Regional-scale estimation of electric power and power plant CO₂ emissions using defense meteorological satellite program operational linescan system nighttime satellite data. *Environ. Sci. Technol. Lett.* **1**, 259–265 (2014).
46. He, Y. & Huang, H. Energy intensity in Guangdong of China (2006–2015): A spatial dynamic general equilibrium econometric model. *Nat. Resour. Modell.* **22**, 1–14 (2019).
47. Fotheringham, A. S., Charlton, M. E. & Brunson, C. Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environ. Plan. A* **30**, 1905–1927 (1998).
48. Griffith, D. A. Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environ. Plan. A* **40**, 2751–2769 (2008).
49. Geniaux, G. & Martinetti, D. A new method for dealing simultaneously with spatial autocorrelation and spatial heterogeneity in regression models. *Region. Urban Econ.* **72**, 74–85 (2017).
50. Li, X., Zhang, R., Huang, C. & Li, D. Detecting 2014 Northern Iraq insurgency using night-time light imagery. *Int. J. Remote Sens.* **36**, 3446–3458 (2015).
51. Li, X., Li, D., Xu, H. & Wu, C. Intercalibration between DMSP/OLS and VIIRS night-time light images to evaluate city light dynamics of Syria's major human settlement during Syrian Civil War. *Int. J. Remote Sens.* **38**, 5934–5951 (2017).
52. Krikigianni, E., Tsiakos, C. & Chalkias, C. Estimating the relationship between touristic activities and night light emissions. *Eur. J. Remote Sens.* **52**, 233–246 (2019).
53. Gao, J. & Li, S. Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using Geographically Weighted Regression. *Appl. Geogr.* **31**, 292–302 (2011).
54. Miller, S. D. *et al.* Illuminating the capabilities of the suomi national polar-orbiting partnership (NPP) visible infrared imaging radiometer suite (VIIRS) day/night band. *Remote Sens.* **5**, 6717–6766 (2013).
55. Levin, N. & Zhang, Q. A global analysis of factors controlling VIIRS nighttime light levels from densely populated areas. *Remote Sens. Environ.* **190**, 366–382 (2017).
56. Levin, N. The impact of seasonal changes on observed nighttime brightness from 2014 to 2015 monthly VIIRS DNB composites. *Remote Sens. Environ.* **193**, 150–164 (2017).
57. Wu, R., Yang, D., Dong, J., Zhang, L. & Xia, F. Regional inequality in China based on NPP-VIIRS night-time light imagery. *Remote Sens.* **10**, 240 (2018).
58. Tripathy, B. R. *et al.* Modeling of electric demand for sustainable energy and management in India using spatio-temporal DMSP-OLS night-time data. *Environ. Manage.* **61**, 615–623 (2018).

Acknowledgements

We would like to thank the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences and University of Chinese Academy of Sciences. This work is supported by the Strategic Priority Research Program of Chinese Academy of Sciences [No. XDA20010203]; the National Key R&D Program of China [No.2018YFC1800103]; the National Natural Science Foundation of China [No. 41771460].

Author contributions

M.C. conceived the proposed method, generated all the GIS maps and figures, and wrote the manuscript, under the guidance of H.C. and X.Y., C.J. polished the manuscript. All authors discussed the results and commented on the manuscript at all stages.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020