

# scientific data



OPEN

DATA DESCRIPTOR

## Global 1 km × 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data

Jiandong Chen<sup>1</sup>, Ming Gao<sup>1</sup>✉, Shulei Cheng<sup>1</sup>, Wenxuan Hou<sup>2,3</sup>, Malin Song<sup>4</sup>, Xin Liu<sup>5</sup> & Yu Liu<sup>6,7</sup>✉

As fundamental data, gross domestic product (GDP) and electricity consumption can be used to effectively evaluate economic status and living standards of residents. Some scholars have estimated gridded GDP and electricity consumption. However, such gridded data have shortcomings, including overestimating real GDP growth, ignoring the heterogeneity of the spatiotemporal dynamics of the grid, and limited time-span. Simultaneously, the Defense Meteorological Satellite Program's Operational Linescan System (DMSP/OLS) and National Polar-orbiting Partnership's Visible Infrared Imaging Radiometer (NPP/VIIRS) nighttime light data, adopted in these studies as a proxy tool, still facing shortcomings, such as imperfect matching results, discontinuity in temporal and spatial changes. In this study, we employed a series of methods, such as a particle swarm optimization-back propagation (PSO-BP) algorithm, to unify the scales of DMSP/OLS and NPP/VIIRS images and obtain continuous 1 km × 1 km gridded nighttime light data during 1992–2019. Subsequently, from a revised real growth perspective, we employed a top-down method to calculate global 1 km × 1 km gridded revised real GDP and electricity consumption during 1992–2019 based on our calibrated nighttime light data.

### Background & Summary

Economic activities are essential for human survival and development<sup>1</sup>. Facilitated by social organization and order, humans use labor and other production resources to exchange goods and services to create, transform, and achieve economic value. The economic output in a country or region is its accumulation of created economic values within a certain period; of which, Gross Domestic Product (GDP), is the most prevalent indicator globally. In addition, electricity consumption is adopted as a supplementary indicator that indirectly reflects the economic status and living quality of residents, because it is indispensable for industrialization<sup>1,2</sup> and in daily lives of modern residents. In particular, the extensive consumption of electricity across different industries and sectors, makes it a substantial contributor to economic growth<sup>3,4</sup>. Numerous studies have reported that the availability of electricity for consumption enhances impoverished populations' livelihoods, with consumption patterns reflecting quality of life and lifestyle<sup>5,6</sup>. Therefore, numerous scholars have heeded to the accounts of GDP, electricity consumption and their spatiotemporal change<sup>7–9</sup>. For example, Henderson *et al.*<sup>10</sup> used the Defense Meteorological Satellite Program's Operational Linescan System (DMSP/OLS) nighttime light data to revise real-world countries' real GDP growth.

<sup>1</sup>School of Public Administration, Southwestern University of Finance and Economics, Chengdu, China. <sup>2</sup>School of Finance, Shanghai Lixin University of Accounting and Finance, Shanghai, China. <sup>3</sup>University of Edinburgh Business School, University of Edinburgh, 29 Buccleuch Place, Edinburgh, Scotland. <sup>4</sup>School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, China. <sup>5</sup>Curtin University Sustainability Policy (CUSP) Institute, School of Design and the Built Environment, Curtin University, Perth, Australia. <sup>6</sup>Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China. <sup>7</sup>School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100049, China. ✉e-mail: [minggao1994@163.com](mailto:minggao1994@163.com); [liuyu@casipm.ac.cn](mailto:liuyu@casipm.ac.cn)

Guerrero *et al.*<sup>11</sup> proposed an improved method to revise a single country's real GDP growth-based DMSP/OLS data. Wang *et al.*<sup>12</sup> analyzed the driving forces behind electricity consumption growth in China's industrial sector. However, most studies have been based on specific administrative region, such that detailed spatiotemporal changes in GDP and electricity consumption on a micro-level scale (e.g., small towns, villages and business clusters) are not easily accessible<sup>13</sup>.

Thus, a few studies have estimated gridded GDP and electricity consumption with specific resolutions based on several proxy tools. Gridded population and nighttime light data are the most popular proxy tools, and have been adopted extensively because of their strong correlation with economic output and electricity use. For example, Kummu *et al.*<sup>14</sup> combined sub-national average GDP per capita with gridded population data to estimate gridded GDP data during 1990–2015. Similarly, based on the regional ratio of GDP to nighttime light data, Zhao *et al.*<sup>15</sup> and Wang *et al.*<sup>16</sup> calculated China's 1 km × 1 km gridded GDP during 2000–2015. With the regard to gridded electricity consumption, Shi *et al.*<sup>17</sup> used DMSP/OLS nighttime light data to estimate gridded electricity consumption at 1-km resolution during 1992–2013. Similarly, Chen *et al.*<sup>18</sup> employed a classification regression method to calculate China's electricity consumption in 2015 based on a combination of NPP/VIIRS data and land use and cover change (LUCC) data for 2015.

Considering existing literature, several limitations of estimated gridded GDP and electricity consumption datasets have been noted. (1) In many studies, gridded GDP data are estimated solely based on official GDP statistics; however, the published GDP growth of some countries, especially developing countries (e.g., China<sup>19</sup> and some African countries<sup>20</sup>), may have errors due to poor statistical methods or intentional manipulation<sup>10,21</sup>. Although nighttime light data, as a globally comparable, objective, and highly economical instrumental variable, has been widely used to correct economic growth data<sup>10,11</sup>, it has not been applied in existing research on gridded GDP measurements. Actually, most studies on gridded GDP estimations directly used gridded nighttime light intensity as the distribution weight for allocating a country's official GDP and calculating the raster data<sup>14–16</sup>. Hence, the grid GDP growth rate may be inaccurate in such studies. (2) Most gridded GDP data are based on the hypothesis that the ratio of GDP to nighttime light data are identical in the same countries, ignoring the grid heterogeneity over space. (3) Moreover, due to the gap between DMSP/OLS and NPP/VIIRS products, long time spans of estimated gridded GDP based on the nighttime light data are limited. (4) Furthermore, the time spans of existing datasets are short and outdated, and do not match those of other updated data. Similarly, for gridded electricity consumption data: (1) due to the limited availability of electricity consumption and nighttime light data, the long-span data was only obtained up to 2013; and (2) based on the hypothesis that the ratio of electricity consumption to nighttime light data are identical in regions, the models adopted in the calculations of electricity were simple, thereby failing to capture the grid heterogeneity over space.

Although nighttime light data as a single indicator may ignore factors such as value added or reduced by forestry or desertification, it is still an effective proxy for calibrating economic growth<sup>22</sup>. The influence of neglected factors on total economic output is limited, and night lighting, as satellite data, has advantages that other indicators (e.g., gridded population data) cannot surpass, such as objectivity, wide range, and high correlation with economic indicators<sup>23</sup>. DMSP/OLS (1992–2013)<sup>24</sup> and NPP/VIIRS (2012–)<sup>25</sup> images are widely used nighttime light data, owing to their long timespans, wide space coverage, and ease of obtaining and updating. However, there are apparent gaps between the two sets of nighttime light data, which hinders the wider application of long-term, continuous nighttime light data. Specifically, the gap between the pixel-level values of DMSP/OLS and NPP/VIIRS images in 2013 could primarily be attributed to the inconsistency at the time of observation, different sensors and cloud cover, which caused 'high–low' or 'low–high' problems in the pixel values of the two images<sup>26</sup> (i.e., the pixels in the NPP/VIIRS image have a high [low] DN value in 2013, whereas the pixels in the same place have a low [high] DN value in the DMSP/OLS image). Several studies have attempted to unify the two sets of satellite data at the pixel level<sup>27,28</sup>. However, the matching process has proved difficult, and the results show low fitting and discontinuity in temporal and spatial changes, leaving room for improvements.

Therefore, this study proposed an improved approach to unify the scale of DMSP/OLS and NPP/VIIRS images, and obtained continuous and stable calibrated nighttime light data during 1992–2019, that are better fitting than those in existing literature. Subsequently, from a real growth rate perspective, we estimated the global 1 km × 1 km gridded revised real GDP and electricity consumption based on the top-down method optimized using the Particle Swarm Optimization-Back Propagation (PSO-BP) algorithm. The datasets provided in our study enrich basic data for research on economics, management, and other issues. Simultaneously, considering our gridded GDP growth was revised based on nighttime light data, it is more objective and comparable, and can be applied in research at the micro-level around the world (especially in some countries with poor statistical quality).

## Methods

**Study areas and Data preprocessing.** Given that the estimations were based on the top-down approach, the study areas depended on the countries that provided the available data. The GDP data includes 175 countries (or regions), and the electricity consumption data includes 134 countries (or regions). As such, the research scope covers over 70% of the global land area, and over 90% of the GDP and electricity consumption.

Two sets of nighttime light data were used in this study: Defense Meteorological Satellite Program/Operational Linescan System (DMSP/OLS) and National Polar-Orbiting Partnership's Visible Infrared Imaging Radiometer Suite (NPP/VIIRS) images. Considering the versions of nighttime light data, we selected annual stable DMSP/OLS images after removing noise and monthly NPP/VIIRS images without cloud cover, because they have better fitting effects with economic output and other socioeconomic factors. The DMSP/OLS resolution is approximately 1000 m and it comprises six different DMSP satellites F10 (1992–1994), F12 (1994–1999), F14 (1997–2003), F15 (2000–2007), F16 (2004–2009), and F18 (2010–2013). The geographic coordinate reference system of the DMSP/OLS image is the WGS-84 coordinate system, the acquisition width is 3000 km, and the spatial resolution is 30 arc seconds (approximately 1 km near the equator and 0.8 km at 40° north latitude). The coverage of the image is from –180° to 180° in longitude and from –65° to 75° in latitude (covering all areas of the world where human activities

exist). The spatial resolution of the NPP/VIIRS image data was higher than that of the DMSP/OLS image, which was 413 m. Simultaneously, unlike DMSP/OLS images that only provide relative radiation values in the range of 0–63, NPP/VIIRS images provide absolute radiation values in the unit of Watts/cm<sup>2</sup>/sr. Considering that there are several problems in satellite images, such as saturation, discontinuities, and white noise, these datasets needed to be pre-processed before they could be used further.

With regard to DMSP/OLS images, we projected the images as a Mollweide projection and resampled them at a spatial resolution of 1 km. Next, based on the invariant region method, we adopted the form of a power function to reduce saturation. In light of the power function parameters provided by Shi *et al.*<sup>17</sup>, the images were calibrated. Given that the two sensors both provided images in specific years (e.g., F10-1994 and F12-1994), we averaged them to obtain individual images for each year. As for the discontinuities, annual continuous processing was adopted based on the assumption that the stable DN value of a pixel on the light image in the following year should not be less than the stable DN value of the pixel in the previous year<sup>29</sup>.

For the NPP/VIIRS images, we adopted 0.3 Watts/cm<sup>2</sup>/sr as the threshold to remove the noise, which is consistent with previous studies<sup>26,30</sup>. To avoid the influence of stray light pollution in summer, monthly images from June to August were removed. Next, based on the average monthly data, we estimated the annual NPP/VIIRS images from 2014 to 2019. As for the discontinuities, we also adopted the same annual continuous processing with DMSP/OLS images. Finally, to better match the DMSP/OLS image, we resampled the NPP/VIIRS image from a resolution of 0.5 km × 0.5 km to that of 1 km × 1 km.

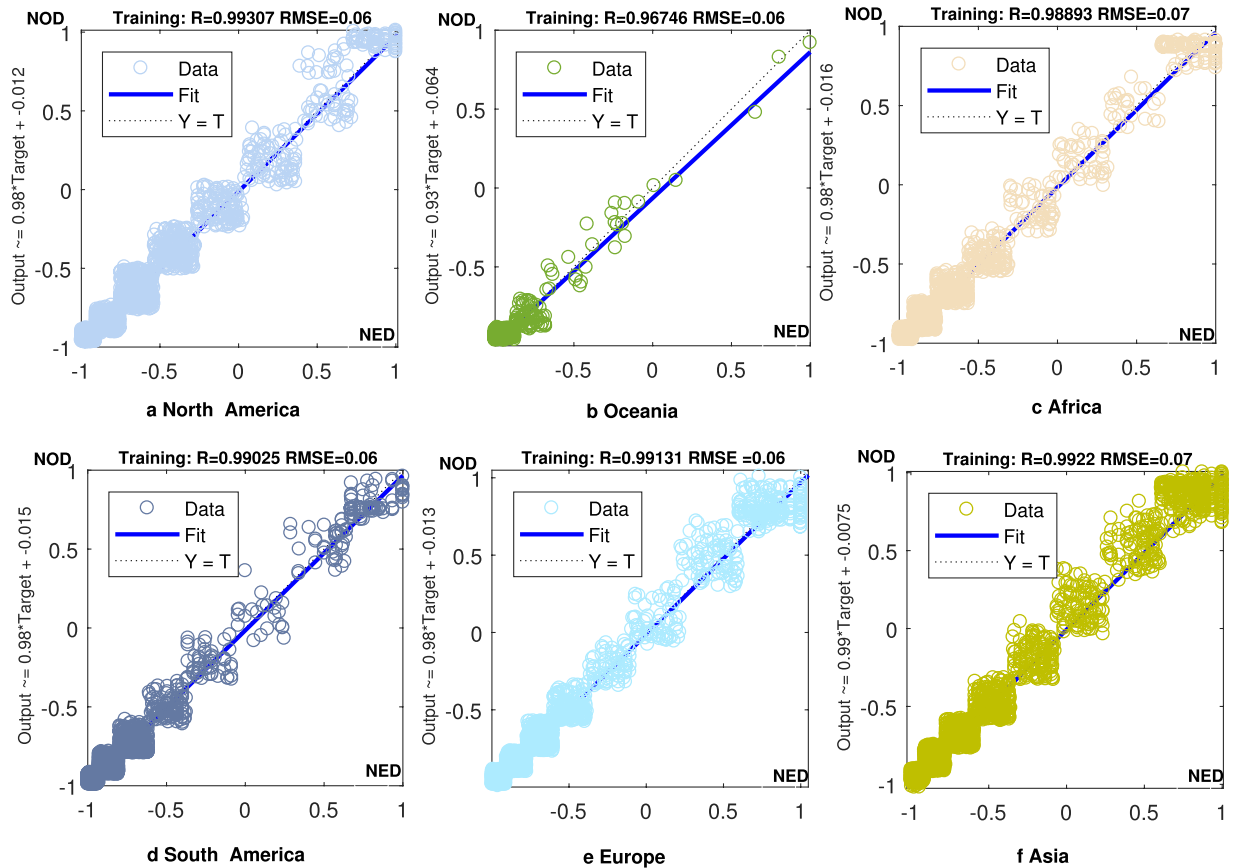
**Matching of the two sets of nighttime light data.** The gap between DMSP/OLS and NPP/VIIRS images is mainly driven by different sensors, spread functions, and spatial and temporal inconsistency<sup>21</sup>. Considering that the relationship between the two data sets is like a “Black Box,” Chen *et al.*<sup>31</sup> used an artificial neural network (ANN) to explore the potential functions on the two data sets, and the matching results proved successful. Based on their study, we also employed a particle swarm optimization-back propagation (PSO-BP) algorithm to unify the scale of DMSP/OLS and NPP/VIIRS images. The initial parameters of the PSO-BP algorithm (i.e., C1 and C2 values were both set to 2.0, and the structure of the model included one hidden layer with five nodes; the maximum iteration number and population size were set to 50 and 20, respectively) were set following Chen *et al.*<sup>31</sup>.

Moreover, because our target is pixel-level matching, errors of hundreds of millions of pixels make the matching effect very poor, even after using machine learning. The difficulty is mainly driven by the ‘high–low’ or ‘low–high’ problems in the pixel DN values of the two images<sup>36</sup> (i.e., the pixels in the NPP/VIIRS image have a high (low) DN value in 2013, whereas the pixels in the same place have a low (high) DN value in the DMSP/OLS image). Therefore, we proposed the principle of ‘high to high’ and ‘low to low’ for the matching job.

Thus, we divided the DMSP/OLS and NPP/VIIRS images into nine categories based on the natural interval method. By matching similar attributes in the two images, we extracted and obtained sampling points that met the principles of ‘high to high’ and ‘low to low’ in the analysis. Subsequently, in line with Chen *et al.*<sup>31</sup> and Li *et al.*<sup>28</sup>, we used the logarithmic form of the pixel DN values in the NPP/VIIRS image in 2013, and their latitude and longitude as input factors. The DN values of the pixels in the DMSP/OLS image in 2013 were selected as the output factors. In addition, according to general practice in machine learning<sup>28,31</sup>, the input and output factors were normalized to avoid the influence of indicators’ units. Considering the continental heterogeneity, we estimated six continental parameters of the PSO-BP neural network (e.g., North America, South America, Oceania, Africa, Asia, and Europe). Antarctica was not considered in this study because the scope of sensors that provided DMSP/OLS and NPP/VIIRS images did not include Antarctica. The matching results based on the training sets (60% of the total samples) are presented in Fig. 1. In particular, the NED in the x-axis represents normalized DN values converted from NPP/VIIRS image scales in 2013 to DMSP/OLS image scales in 2013; the NOD in the y-axis represents normalized DN values of original DMSP/OLS image DN values in 2013.

As shown in Fig. 1, all of the R<sup>2</sup> of the six continents’ training results were > 0.96, indicating that the PSO-BP neural network performed well in identifying the potential relationship between DMSP/OLS and NPP/VIIRS images in 2013. The test results are shown in Fig. 2. The test performances of the six continents can be used to evaluate the prediction effects of the algorithm (i.e., whether the parameters of the PSO-BP algorithm can be employed to convert the pixel normalized DN values of NPP/VIIRS images during 2014–2019 to the scale of normalized DN values in DMSP/OLS images). Except for the fitting result of Oceania (i.e., the R<sup>2</sup> was only 0.91), the fitting results of the other five continents all exceeded 0.98. The poor fitting effect of Oceania may be owing to the lack of light at night in most parts of Oceania. Considering Oceania has fewer stable light sources, its poor prediction results have limited impact on the matching of two sets of night lights at the global scale. Subsequently, the global converted normalized DN values of NPP/VIIRS images during 2013–2019 were denormalized to the original range, which was consistent with the scale of DN values in DMSP/OLS images. Moreover, the final global converted DN values of NPP/VIIRS images in 2013 and the DN values of DMSP/OLS images were compared again to verify the effect of matching: the global coefficient of determination was > 0.98, which was higher than that obtained in previous studies (for example, 0.91 in Zhao *et al.*<sup>32</sup>, 0.84 Lv *et al.*<sup>33</sup>, and 0.87 in Chen *et al.*<sup>34</sup>).

Based on the trained parameters of the neural network, we transformed the scale of the NPP/VIIRS data from 2014 to 2019 to the scale of the DMSP/OLS data. As the generated network was based on “high to high” and “low to low” principle, the same pixels with high DN values in the NPP/VIIRS images can be converted into high DN values at the scale of DMSP/OLS images. However, the matching job was not complete yet. First, there were also certain pixels in NPP/VIIRS images with low DN values transformed into low DN values at DMSP/OLS scales, not matching the high DN values in the same regions of DMSP/OLS in 2013. Second, although the correlation coefficient was close to 1, there were evident and unavoidable discontinuity in some grids during 2013–2014, which also exist in previous studies<sup>32</sup>.



**Fig. 1** Different training results for the pixel normalized values on the six continents. (a–f). Training results for the pixel normalized values on (a) North America, (b) Oceania, (c) Africa, (d) South America, (e) Europe, and (f) Asia.

Therefore, inter-annual continuous series correction was adopted for the transformed NPP/VIIRS images from 2014 to 2019. In line with the correction approach, pixels with high DN values in the DMSP/OLS image were maintained in the converted NPP/VIIRS images for the period of 2014–2019. And the potential problem of discontinuity was solved. The equation is as follows:

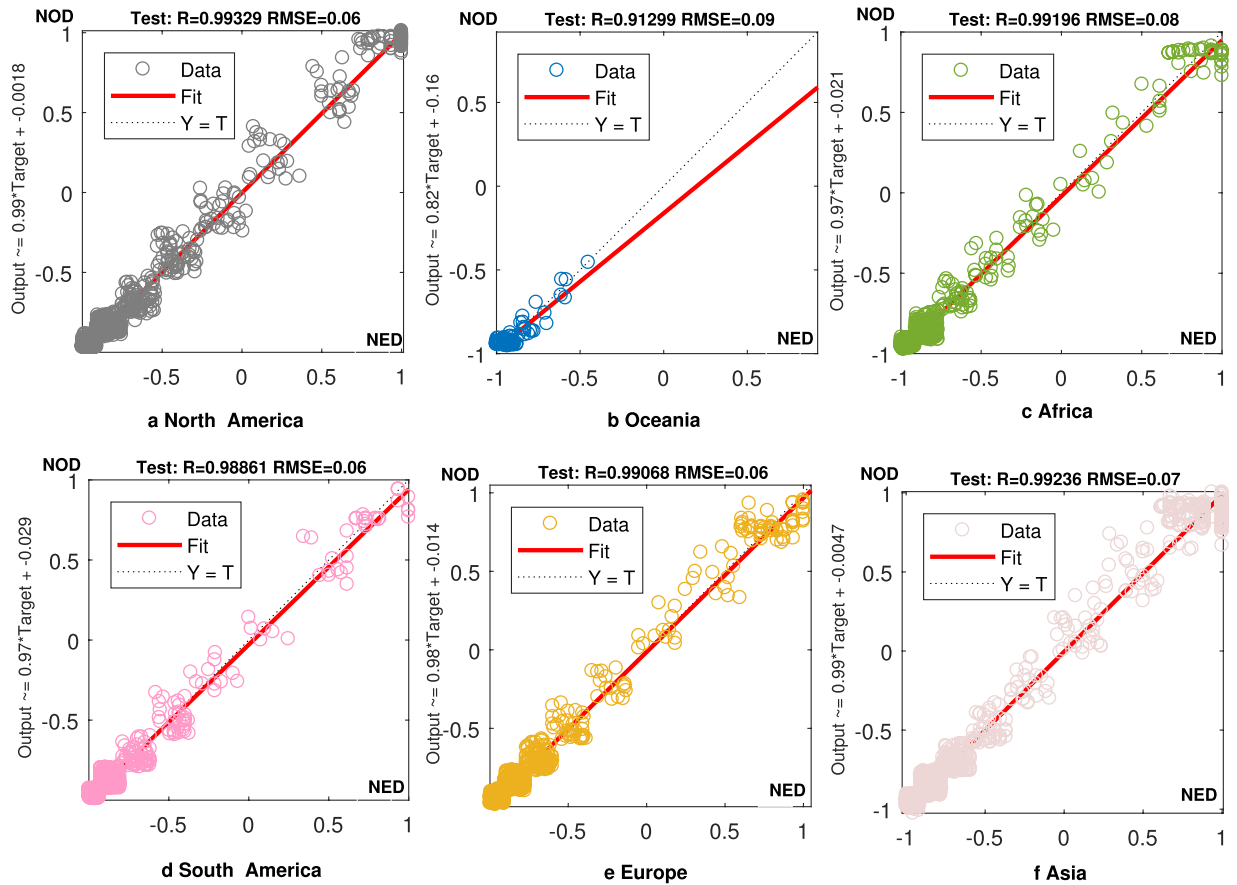
$$DN_{i,t} = \begin{cases} DN_{i,t-1}, & \text{if } DN_{i,t-1} \geq DN_{i,t} \quad (t = 2014, \dots, 2019), \\ DN_{i,t}, & \text{otherwise} \end{cases} \quad (1)$$

In summary, based on the PSO-BP algorithm, we could confidently convert the scale of NPP/VIIRS data from 2013 to 2019 to the scales of DMSP/OLS data and obtain stable and continuous global 1 km × 1 km gridded nighttime light data for the time period of 1992–2019, which laid the foundation for further calculations of global 1 km × 1 km gridded GDP and electricity consumption during the period. Figure 3 presents the spatial distributions of global nighttime light data in 2019.

**Calculation of real GDP and electricity consumption based on growth rate.** Owing to errors in official GDP growth attributed to poor statistical methods or intentional manipulation<sup>10,11,21</sup>, nighttime light data has been employed extensively in revision of official national GDP growth data. Based on the approaches proposed by Henderson *et al.*<sup>10</sup> and Guerrero *et al.*<sup>11</sup>, the revised growth estimate is a composite with different weights of conventionally measured growth and growth predicted from nighttime light data. Considering the approaches of such studies, we employed nighttime light data to revise the real GDP growth rate. In particular, the real GDP growth rate was estimated using Eq. (2).

$$y_{i,t}^* = \rho y_{i,t} + (1 - \rho) y'_{i,t} \quad (2)$$

where  $y_{i,t}^*$  is the  $i^{\text{th}}$  country's real GDP growth in period  $t$ ;  $y_{i,t}$  is the official GDP growth of the  $i^{\text{th}}$  country in period  $t$ ;  $y'_{i,t}$  presents the  $i^{\text{th}}$  country's predicted GDP growth based on the night-time light data in period  $t$ ; and  $(1 - \rho)$  is the optimal weight of predicted growth based on the night-time light data. In the light with the idea proposed by Henderson *et al.*<sup>10</sup>, the optimal value of  $\rho$  was specified to minimize the variance of measurement error in this



**Fig. 2** Different test results for the pixel normalized values on the six continents. **(a–f)**. Test results for the pixel normalized values on **(a)** North America, **(b)** Oceania, **(c)** Africa, **(d)** South America, **(e)** Europe, and **(f)** Asia.

estimate relative to the true value of GDP growth. As long as the optimal weight on  $(1 - \rho)$  is positive, use of night-time light data improves our ability to measure true GDP growth. The variance of this composite GDP growth was estimated by the following equation:

$$\text{var}(y_i^{\wedge*} - y_i^*) = \rho^2 \text{var}(y_i - y_i^*) + (1 - \rho)^2 (y_i' - y_i^*) \quad (3)$$

Following Henderson *et al.*<sup>10</sup>, the relationships between the night-time light data and real GDP growth/official GDP growth were described as the following equations:

$$y_i = y_i^* + \varepsilon_{y,i} \quad (4)$$

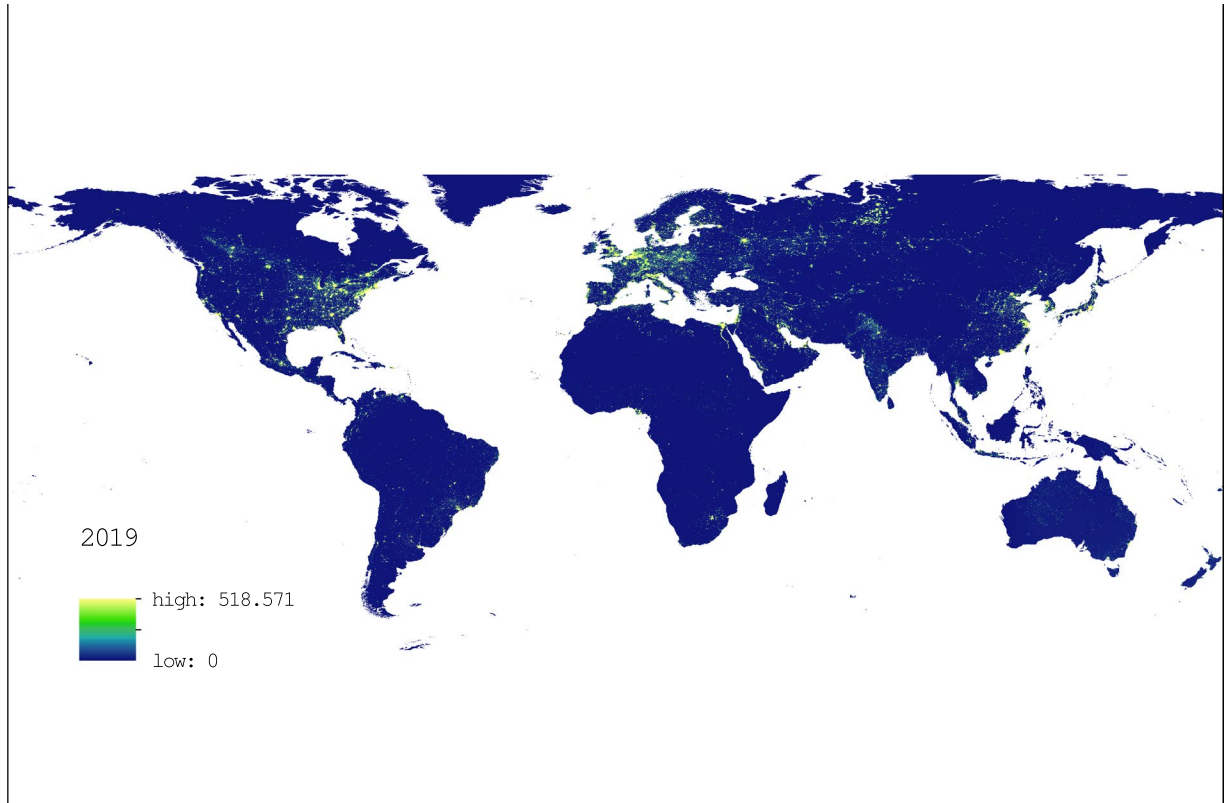
$$\text{sdna}_i = \beta y_i^* + \varepsilon_{\text{sdna},i} \quad (5)$$

$$y_i = \gamma \text{sdna}_i + e_i \quad (6)$$

$$\sigma_y^2 = \varepsilon_{y,i}^2 \quad (7)$$

$$\sigma_{\text{sdna}}^2 = \varepsilon_{\text{sdna},i}^2 \quad (8)$$

where  $\text{sdna}_i$  is the growth of the sum of DN values per area;  $\varepsilon_{y,i}$ ,  $\varepsilon_{\text{sdna},i}$  and  $e_i$  are the errors;  $\beta$  was the elasticity of lights growth with respect to real GDP growth;  $\gamma$  was the elasticity of official GDP growth with respect to lights growth;  $\sigma_y^2$  and  $\sigma_{\text{sdna}}^2$  are the variance of errors. Based on the assumption that the degree of measurement error in GDP growth has no effect on the estimated value of the parameter, there is  $\text{cov}(\varepsilon_y, \varepsilon_{\text{sdna}}) = 0$ . Thus, there were further derived equations as follows:



**Fig. 3** The 1 km × 1 km gridded nighttime light data in 2019.

$$\text{var}(sdna) = \beta^2\sigma_{y^*}^2 + \sigma_{sdna}^2 \tag{9}$$

$$\text{cov}(sdna, y) = \text{cov}(y^*, sdna) = \beta\sigma_{y^*}^2 \tag{10}$$

$$\text{var}(y) = \sigma_{y^*}^2 + \sigma_y^2 \tag{11}$$

Then, the relationship between  $\hat{\gamma}$  and the structural parameter  $\beta$  is as follows:

$$\text{Plim}(\hat{\gamma}) = \frac{\text{cov}(sdna, y)}{\text{var}(sdna)} = \frac{1}{\beta} \left( \frac{\beta^2\sigma_{y^*}^2}{\beta^2\sigma_{y^*}^2 + \sigma_{sdna}^2} \right) \tag{12}$$

Thus, the Eq. (3) can be rewritten as follows:

$$\text{var}(y_i^{\wedge*} - y_i^*) = \rho^2\sigma_y^2 + (1 - \rho)^2 \frac{\sigma_{sdna}^2\sigma_{y^*}^2}{\beta^2\sigma_{y^*}^2 + \sigma_{sdna}^2} \tag{13}$$

From Eq. (13), we solve for the weight  $\rho$  which minimizes this variance:

$$\rho^* = \frac{\sigma_{sdna}^2\sigma_{y^*}^2}{\sigma_y^2(\beta^2\sigma_{y^*}^2 + \sigma_{sdna}^2) + \sigma_{sdna}^2\sigma_{y^*}^2} \tag{14}$$

Furthermore, following Henderson *et al.*<sup>10</sup>,  $\rho$  is further classified based on countries with good- and bad-quality data:  $\rho_{i,good}$  and  $\rho_{i,bad}$ . Therefore, the Eq. (11) becomes two Eqs. (15, 16).

$$\text{var}(y_{good}) = \sigma_{y^*}^2 + \sigma_{y,good}^2 \tag{15}$$

$$\text{var}(y_{bad}) = \sigma_{y^*}^2 + \sigma_{y,bad}^2 \tag{16}$$

And the ratio of signal to total variance in official GDP growth for countries with good quality of statistics was estimated. A higher ratio of signal to total variance indicates more reliable GDP growth. The calculation equation was presented as follows:

$$\phi = \frac{\sigma_{y^*}^2}{\sigma_{y^*}^2 + \sigma_{y,good}^2}, \quad (17)$$

where  $\phi$  was set to 0.9 based on Henderson *et al.*<sup>10</sup> and Guerrero *et al.*<sup>11</sup>. Therefore,  $\rho_{i,good}$  and  $\rho_{i,bad}$  can be determined with the following equations:

$$\rho_{i,good} = \frac{\sigma_{sdna}^2 \sigma_{y^*}^2}{\sigma_{y,good}^2 (\beta \sigma_{y^*}^2 + \sigma_{SDNA}^2) + \sigma_{SDNA}^2 \sigma_{y^*}^2} \quad (18)$$

$$\rho_{i,bad} = \frac{\sigma_{SDNA}^2 \sigma_{y^*}^2}{\sigma_{y,bad}^2 (\beta \sigma_{y^*}^2 + \sigma_{SDNA}^2) + \sigma_{SDNA}^2 \sigma_{y^*}^2} \quad (19)$$

Considering the conclusions of previous studies<sup>10,21,35</sup>, statistics from developed countries always have better quality, while those from developing countries are less reliable. Therefore, we characterized the quality of a country's data based on whether it is a developed country. In additions, the weights applied during growth prediction from nighttime light data (i.e.,  $(1 - \rho)$ ) were different between developed and developing countries, which is consistent with Henderson *et al.*<sup>10</sup>. The classification into developed and developing countries was based on that of the United Nations (Statistics Division) provided by the World Bank<sup>36</sup>. Based on the above equations, we obtained the optimal weights of the official GDP growth rate in developed and developing countries (i.e.,  $\rho_{good} = 0.94$  and  $\rho_{bad} = 0.66$ ).

Furthermore, each grid's real GDP growth rate during 1993–2019 can be estimated using the following equation:

$$gy_{ij,t}^* = \begin{cases} \rho_{gb} \times y_{i,t} + (1 - \rho_{gb}) \times \left( \frac{DN_{ij,t} - DN_{ij,t-1}}{DN_{ij,t-1}} \right) \times \alpha, & \text{if } DN_{ij,t-1} \neq 0 \\ y_{i,t}, & \text{if } DN_{ij,t-1} = 0 \end{cases}, \quad (20)$$

where  $gy_{ij,t}^*$  denotes the  $j^{\text{th}}$  grid in the  $i^{\text{th}}$  country's real GDP growth;  $gb = good, bad$ ;  $\alpha$  represents the elasticity of the nighttime light data to GDP (i.e., 0.45 based on the regression results), which was obtained by Eq. (6).

Next, based on the gridded real GDP growth rate during 1993–2019, the gridded GDP data in 1992 or 2019 were estimated as basic values to obtain the gridded real GDP data in other years. Since the DN values in newly built-up areas were zero in 1992, these areas' basic GDP values in 1992 were also zero, thereby leading to values of zero in subsequent years. Thus, the gridded GDP data in 2019 was selected as the basic value, which was calculated based on the top-down method.

Finally, the gridded real GDP based on the real growth rate can be calculated using Eq. (21).

$$RGY_{ij,t}^* = \begin{cases} RGY_{ij,t+1}^*, & \text{if } DN_{ij,t} \neq 0 \\ 1 + gy_{ij,t}^*, & \\ 0, & \text{if } DN_{ij,t} = 0 \end{cases}, \quad (21)$$

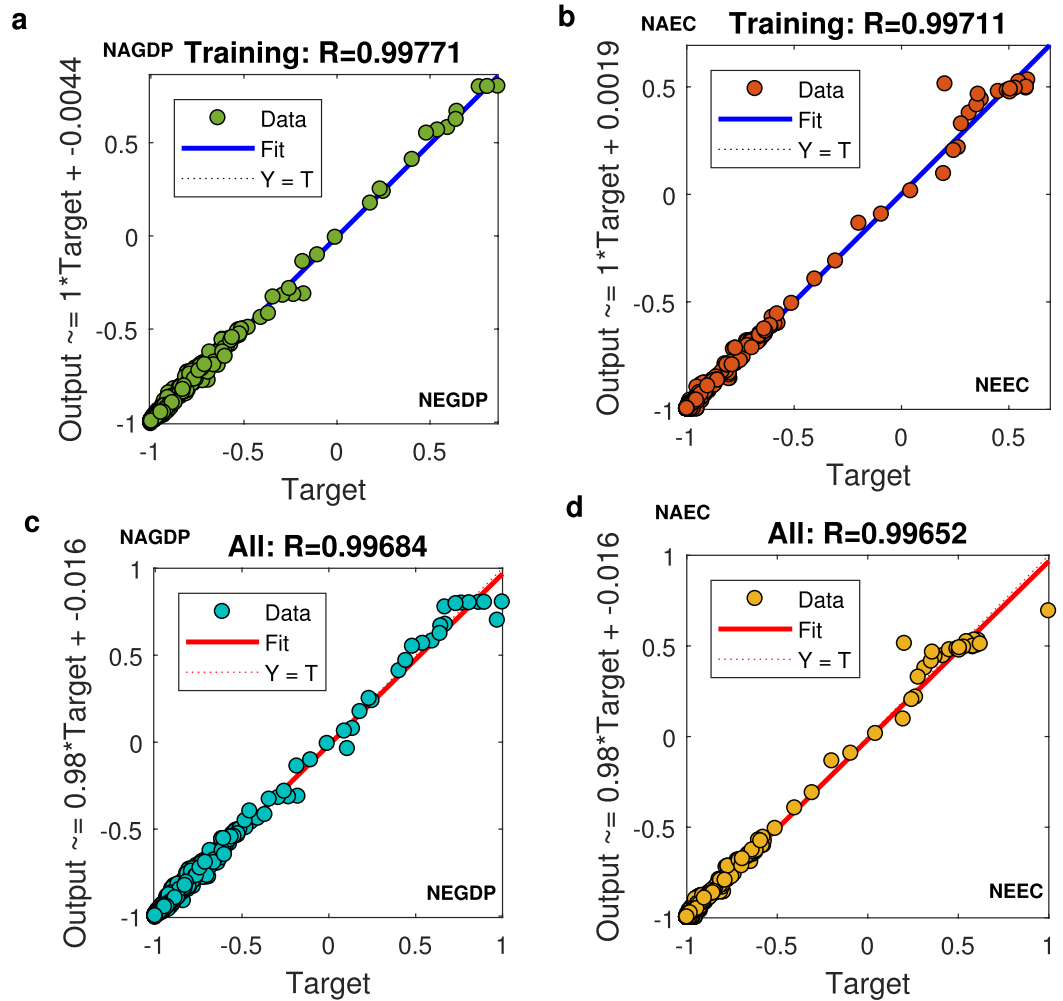
where  $RGY_{ij,t}^*$  denotes the  $j^{\text{th}}$  grid in the  $i^{\text{th}}$  country's real GDP in the period of  $t$  based on the revised real growth rate. The calculations were based on the hypothesis that there is no GDP when the DN value is zero, which is consistent with Shi *et al.*<sup>17</sup> and Wang *et al.*<sup>16</sup>.

As for electricity consumption, the gridded growth rate of nighttime light data was used to estimate the growth rate of gridded electricity consumption. However, because the growth rate of electricity consumption was mainly driven by the industrial sectors rather than the residential sector<sup>37,38</sup>, the growth rate of the nighttime light data may not comprehensively capture the growth rate of electricity consumption. Thus, we combined the growth of official GDP and nighttime light data to better reveal the gridded growth rate of electricity consumption, which is presented in Eq. (22).

$$\ln EC_{it} = \gamma \ln(SDN_{it}) + \pi \ln(Y_{it}) + c_{it} + \tau_{it}, \quad (22)$$

where  $EC_{it}$  denotes the  $i^{\text{th}}$  country's electricity consumption in the period  $t$ ,  $SDN_{it}$  denotes the  $i^{\text{th}}$  country's sum of DN values in the period  $t$ ,  $c_{it}$  denotes the constant,  $\tau_{it}$  denotes the errors,  $\gamma$  and  $\pi$  denote the coefficients (i.e., 0.22 and 0.71). Then, the gridded growth rate of the electricity consumption  $gEC_{j,t}^*$  was calculated using Eq. (23).

$$gEC_{j,t}^* = \begin{cases} \gamma \times \left( \frac{DN_{ij,t}}{DN_{ij,t-1}} - 1 \right) + \pi \times \left( \frac{Y_{i,t}}{Y_{i,t-1}} - 1 \right), & \text{if } DN_{ij,t-1} \neq 0 \\ 0, & \text{if } DN_{ij,t-1} = 0 \end{cases}, \quad (23)$$



**Fig. 4** Training and all samples' results for the relationship between national normalized actual GDP/ electricity consumption and our estimated GDP/electricity consumption predicted based on the input factors. (a–d). (a) Training results for the relationship between national normalized actual GDP and our estimated GDP predicted based on the input factors; (b) Training results for the relationship between national normalized actual electricity consumption and our estimated electricity consumption predicted based on the input factors; (c) All samples' results for the relationship between national normalized actual GDP and our estimated GDP consumption predicted based on the input factors; (d) All samples' results for the relationship between national normalized actual electricity consumption and our estimated electricity consumption predicted based on the input factors.

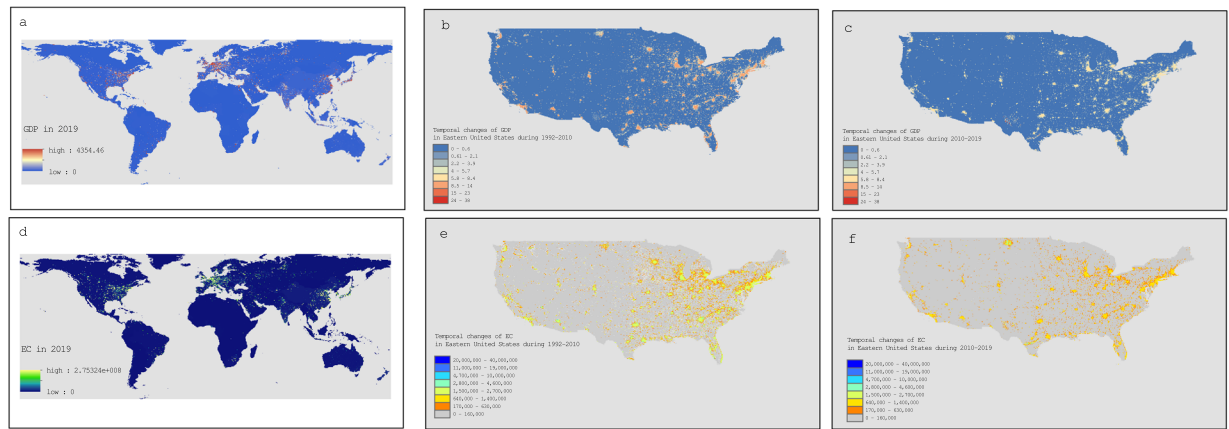
Given that only the worldwide electricity consumption during 1992–2015 was open-access and available freely, we selected the gridded electricity consumption data in 2015 as the basic values. Then, the gridded electricity consumption  $GEC_{j,t}^*$  was calculated using Eq. (24).

$$GEC_{ij,t}^* = \begin{cases} \frac{GEC_{ij,t+1}}{1 + gecg_{ij,t}^*}, & \text{if } DN_{ij,t-1} \neq 0 \\ 0, & \text{if } DN_{ij,t-1} = 0 \end{cases} \quad (24)$$

With regard to the basic values of gridded GDP in 2019 and electricity consumption in 2015, we first established the relationships between national nighttime light data (i.e., the sum of the DN values) and targeted variables (i.e., GDP and electricity consumption) based on the top-down approach, respectively. Thus, the ratios of GDP and electricity to the nighttime light data (i.e., the coefficients of the targeted variables per unit of DN value) can be estimated among different countries (or regions) during 1992–2019, and each 1 km × 1 km grid can be assigned GDP and electricity consumption with the DN value as the weight. Thus, the ratios of GDP or electricity consumption to DN values were estimated using the following equations:

$$Y_{it}^* = \beta_{it} SDN_{it} + \mu_{it} \quad (25)$$





**Fig. 5** Global 1 km  $\times$  1 km gridded GDP and electricity consumption in 2019 and temporal changes of GDP and electricity consumption in Eastern United States during 1992–2010 and 2010–2019 (unit: millions of 2017 US dollar; kwh). **(a–f)**. **(a)** Global 1 km  $\times$  1 km gridded GDP in 2019; **(b)** temporal changes of GDP in Eastern United States during 1992–2010; **(c)** temporal changes of GDP in Eastern United States during 2010–2019; **(d)** Global 1 km  $\times$  1 km gridded electricity consumption in 2019; **(e)** temporal changes of electricity consumption in Eastern United States during 1992–2010; **(f)** temporal changes of electricity consumption in Eastern United States during 2010–2019.

$$EC_{it} = \theta_{it}SDN_{it} + \epsilon_{it}, \quad (26)$$

where  $Y_{it}^*$  represents the  $i^{th}$  country's (or region's) real GDP in the period  $t$ ;  $\beta_{it}$  and  $\theta_{it}$  represent the coefficients of the  $i^{th}$  country's (or region's) in the period  $t$ ;  $\mu_{it}$  and  $\epsilon_{it}$  denote the errors.

Furthermore, in line with Chen *et al.*<sup>31</sup>, we employed the PSO-BP algorithm to fit and train the relationship among real GDP, electricity consumption, and nighttime light data. The real GDP and electricity consumption were selected as the output factors. The sum of DN values, dummy variables of identity and year were used as input parameters. In addition, the other initialized parameters were consistent with those discussed in the earlier section on the inter-calibration. According to the general practice in machine learning<sup>28,31</sup>, the input and output factors were normalized to avoid the influence of indicators' units. The results are shown in Fig. 4. In particular, the NEGDP/NEEC in the x-axis represents our estimated national normalized GDP/electricity consumption predicted based on input factors; the NAGDP and NAEC in the y-axis represent national normalized actual GDP and electricity consumption, respectively.

Notably, the coefficients of determination  $R^2$  of normalized GDP and electricity consumption were over 0.99. Thus, the training and all samples' results showed great fitting effects, which indicated the high effectiveness of the algorithm. Then, based on the top-down method and a DN value-based weighted-average strategy<sup>39–41</sup>, we obtained the 1 km  $\times$  1 km gridded GDP and electricity consumption in 2019 and 2015. Finally, the gridded real GDP and electricity based on the growth rate during 1992–2019 were calculated using Eqs. (21, 24).

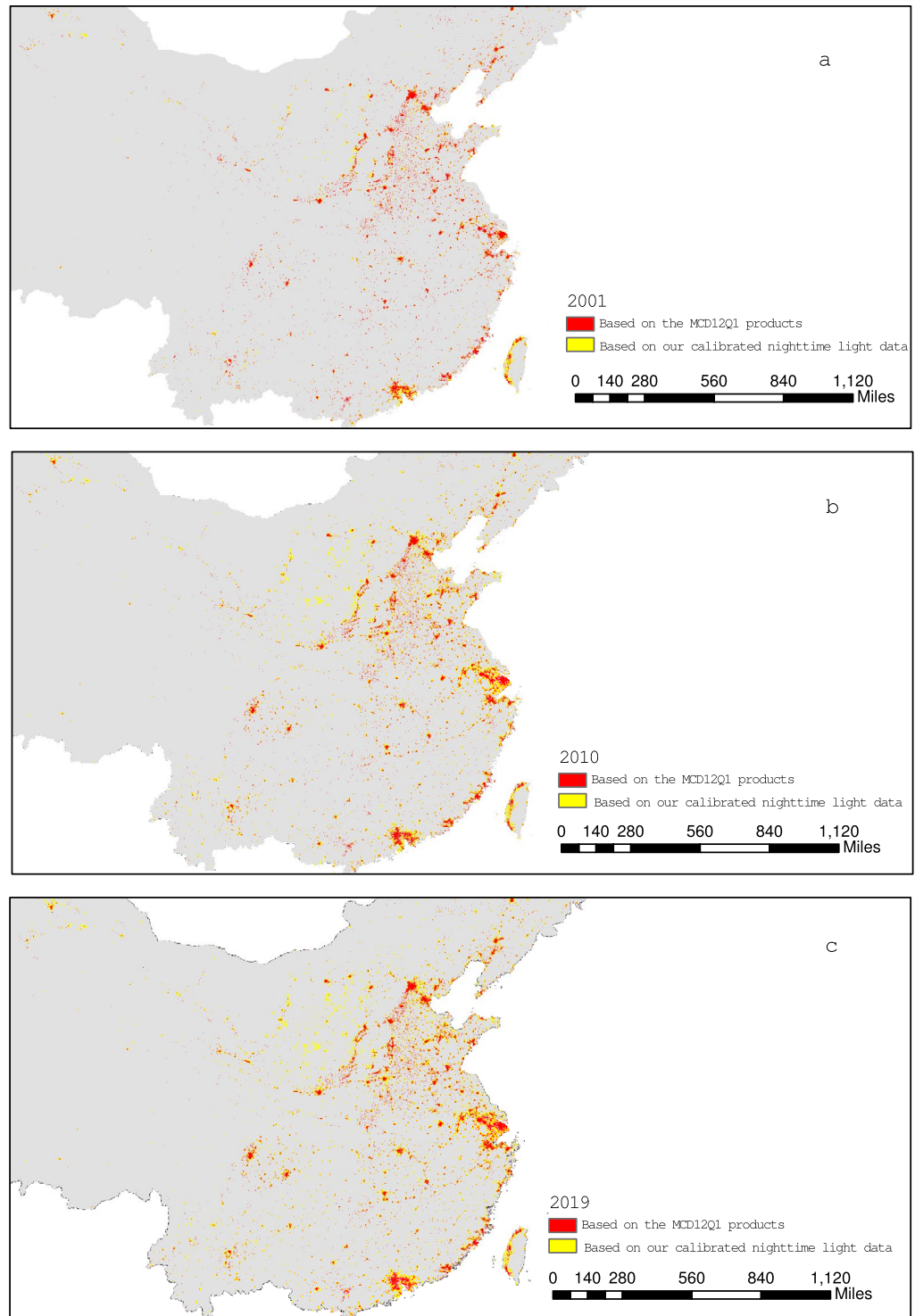
## Data Records

A total of two sets of data records (gridded real GDP and electricity consumption) during 1992–2019 were calculated. The units for the estimated 1 km  $\times$  1 km gridded real GDP and electricity consumption data are millions of 2017 US dollars and kilowatt hours, respectively. The presented datasets and codes are publicly available under Figshare<sup>42,43</sup>. The global 1 km  $\times$  1 km gridded real GDP and electricity consumption emissions in 2019 are presented in Fig. 5(a),(d). To better present the detailed temporal changes of GDP and electricity during 1992–2019, we selected Eastern United States as samples, and Fig. 5(b,c,e,f) show the temporal changes of GDP and electricity consumption in Eastern United States during 1992–2019.

## Technical Validation

**Validity testing for the nighttime light data changes.** Validity testing the spatial patterns of calibrated nighttime light data. Due to the strong relationship between the brightness of area of nighttime light and urbanization, nighttime light data has always been employed to extract built-up areas in urban development. Thus, we used the neighborhood boundary method proposed by Su *et al.*<sup>40</sup> to extract the global urban built-up areas in 2001, 2010 and 2019 based on our calibrated nighttime light data. Next, the urban built-up areas provided by the MCD12Q1 products<sup>44</sup> were selected as a reference to validate our calibrated nighttime light data in spatial patterns. Considering that the areas of urban built-up land in China's central and eastern areas experienced quick increments, we selected their urban built-up areas as samples. Figure 6 presents the comparison of their urban built-up areas derived from the two sets of data.

It is evident that the urban built-up lands based on our calibrated nighttime light data were highly consistent with those derived from the MCD12Q1 products. It should be noted that the urban built-up area based on nighttime light data was relatively larger than that from MCD12Q1 products. Because urban built-up areas from the nighttime light data based on Su *et al.*<sup>40</sup> considered the central urban region and the surrounding urban region, while the urban built-up land from MCD12Q1 products only identified approximately 30% of impervious surface



**Fig. 6** Comparison of urban built-up areas of China's central and eastern regions derived from MCD12Q1 products and our calibrated nighttime light data in 2001, 2010 and 2019. **(a–c)**. **(a)** The comparison of China's central and eastern urban built-up areas derived from MCD12Q1 products and our calibrated nighttime light data in 2001. **(b)** The comparison of China's central and eastern urban built-up areas derived from MCD12Q1 products and our calibrated nighttime light data in 2010. **(c)** The comparison of China's central and eastern urban built-up areas derived from MCD12Q1 products and our calibrated nighttime light data in 2019.

areas (only including building materials, asphalt, and vehicles). The comparison results indicated that our calibrated nighttime light data performed well in the spatial patterns of urbanization.

Validity testing the spatial distributions of nighttime light data among the observed countries (or regions). Because nighttime light data tend to be highly consistent with economic output and electricity consumption,

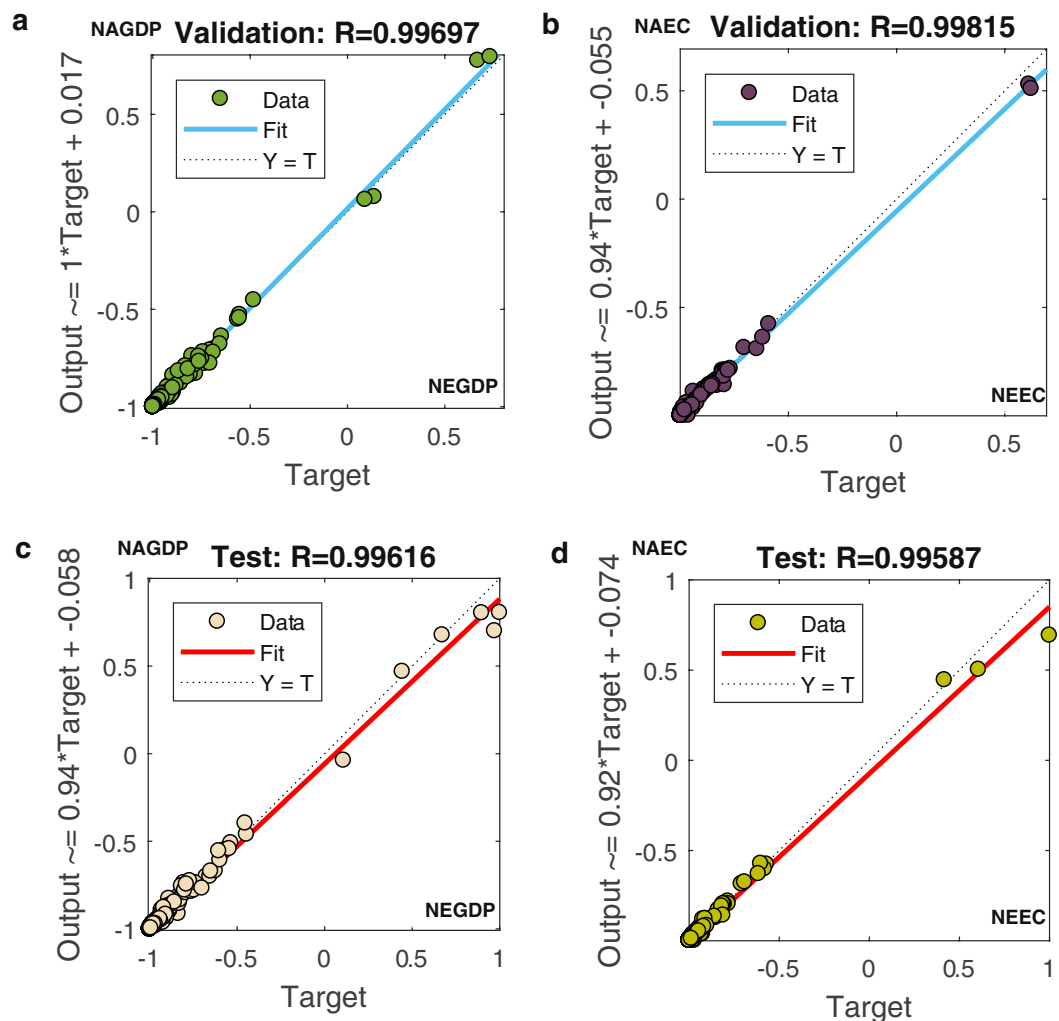
| Year | GDP     |                |           |           | Electricity consumption |                |           |           |
|------|---------|----------------|-----------|-----------|-------------------------|----------------|-----------|-----------|
|      | Slope   | R <sup>2</sup> | AIC value | BIC value | Slope                   | R <sup>2</sup> | AIC value | BIC value |
| 1992 | 0.13*** | 0.82           | 28.71     | 28.75     | 3.75e+07***             | 0.92           | 67.07     | 67.11     |
| 1993 | 0.12*** | 0.82           | 28.78     | 28.82     | 3.53e+07***             | 0.91           | 67.15     | 67.20     |
| 1994 | 0.12*** | 0.81           | 28.87     | 28.91     | 3.40e+07***             | 0.91           | 67.29     | 67.33     |
| 1995 | 0.12*** | 0.81           | 28.94     | 28.98     | 3.39e+07***             | 0.91           | 67.32     | 67.37     |
| 1996 | 0.12*** | 0.81           | 29.02     | 29.06     | 3.39e+07***             | 0.91           | 67.39     | 67.43     |
| 1997 | 0.12*** | 0.82           | 29.07     | 29.1      | 3.39e+07***             | 0.91           | 67.43     | 67.47     |
| 1998 | 0.13*** | 0.82           | 29.09     | 29.13     | 3.39e+07***             | 0.91           | 67.48     | 67.53     |
| 1999 | 0.13*** | 0.83           | 29.13     | 29.17     | 3.48e+07***             | 0.91           | 67.52     | 67.56     |
| 2000 | 0.13*** | 0.84           | 29.19     | 29.23     | 3.55e+07***             | 0.91           | 67.59     | 67.63     |
| 2001 | 0.14*** | 0.83           | 29.24     | 29.28     | 3.45e+07***             | 0.90           | 67.61     | 67.66     |
| 2002 | 0.14*** | 0.84           | 29.28     | 29.32     | 3.53e+07***             | 0.90           | 67.71     | 67.76     |
| 2003 | 0.14*** | 0.84           | 29.32     | 29.36     | 3.56e+07***             | 0.89           | 67.85     | 67.89     |
| 2004 | 0.15*** | 0.85           | 29.38     | 29.42     | 3.63e+07***             | 0.88           | 68.00     | 68.04     |
| 2005 | 0.15*** | 0.84           | 29.49     | 29.53     | 3.76e+07***             | 0.87           | 68.71     | 68.75     |
| 2006 | 0.16*** | 0.84           | 29.59     | 29.63     | 3.81e+07***             | 0.86           | 68.39     | 68.44     |
| 2007 | 0.16*** | 0.83           | 29.71     | 29.74     | 3.90e+07***             | 0.83           | 68.64     | 68.69     |
| 2008 | 0.16*** | 0.83           | 29.74     | 29.77     | 3.91e+07***             | 0.83           | 68.69     | 68.74     |
| 2009 | 0.16*** | 0.81           | 29.86     | 29.9      | 2.56e+07***             | 0.91           | 67.10     | 67.15     |
| 2010 | 0.16*** | 0.81           | 29.97     | 30.01     | 3.68e+07***             | 0.81           | 68.77     | 68.81     |
| 2011 | 0.16*** | 0.82           | 30.03     | 30.07     | 3.85e+07***             | 0.80           | 68.98     | 69.03     |
| 2012 | 0.16*** | 0.82           | 30.07     | 30.11     | 3.91e+07***             | 0.77           | 69.24     | 69.28     |
| 2013 | 0.17*** | 0.81           | 30.17     | 30.21     | 3.93e+07***             | 0.76           | 69.33     | 69.37     |
| 2014 | 0.17*** | 0.8            | 30.33     | 30.36     | 4.05e+07***             | 0.73           | 69.54     | 69.59     |
| 2015 | 0.17*** | 0.79           | 30.43     | 30.46     | 4.10e+07***             | 0.72           | 69.65     | 69.94     |
| 2016 | 0.18*** | 0.79           | 30.53     | 30.57     | —                       | —              | —         | —         |
| 2017 | 0.18*** | 0.78           | 30.61     | 30.64     | —                       | —              | —         | —         |
| 2018 | 0.19*** | 0.79           | 30.63     | 30.67     | —                       | —              | —         | —         |
| 2019 | 0.19*** | 0.78           | 30.71     | 30.74     | —                       | —              | —         | —         |

**Table 1.** Validity test results for the relationship between GDP/ electricity consumption and nighttime light data. Notes: \*\*\* significance at the 1% level. The slope represents the coefficients of the linear regression of the national cross-sectional GDP or electricity consumption with the sum of the DN values. The AIC denotes the Akaike information criterion. BIC represents the Bayesian information criterion.

many scholars<sup>45,46</sup> test the spatial distributions of region's DN values by using the national cross-sectional GDP and electricity consumption to individually perform linear regressions with the sum of DN values (SDN). Following them, we also made regressions during 1992–2019. The results are shown in Table 1. It was evident that there was a significant positive relationship between national cross-sectional GDP and SDN during 1992–2019. The average R<sup>2</sup> value was approximately 0.82 and 0.87. The results imply that our calibrated nighttime light data performed better in fitting the relationship between nighttime light data and GDP/electricity consumption. In addition, the AIC and BIC values were small, thereby implying that the inter-calibrated nighttime light data characterized the economic output well.

**Validity Testing the PSO-BO algorithm for estimated GDP and electricity consumption.** With regard to the validity of the PSO-BO algorithm for predicted GDP and electricity consumption based on nighttime light data, we first validate the effectiveness of the PSO-BO algorithm. Figure 7 shows the validation and testing results of the parameters in the PSO-BO algorithm. In particular, the NSDN in the x-axis represents national normalized sum of DN values; the NGDP and NEC in the y-axis represents national normalized GDP and electricity consumption, respectively. It is evident that all coefficients of determination R<sup>2</sup> of GDP and electricity consumption are over 0.99, respectively. The results indicate that the employed PSO-BO algorithm was effective in our study.

Next, following Shi *et al.*<sup>17</sup> and Chen *et al.*<sup>31</sup>, we utilized the original actual GDP<sup>47</sup> and electricity consumption<sup>17</sup> based on existing literature to conduct a comparison with the summary of our estimated GDP and electricity consumption. The results are shown in Fig. 8. Panels (a) and (b) in Fig. 8 individually show the scatter plots of our simulated national GDP and electricity with the data based on existing literature from 1992 to 2019. The results in each graph were highly consistent, indicating that there are no outliers in aggregated grid GDP and electricity consumption.



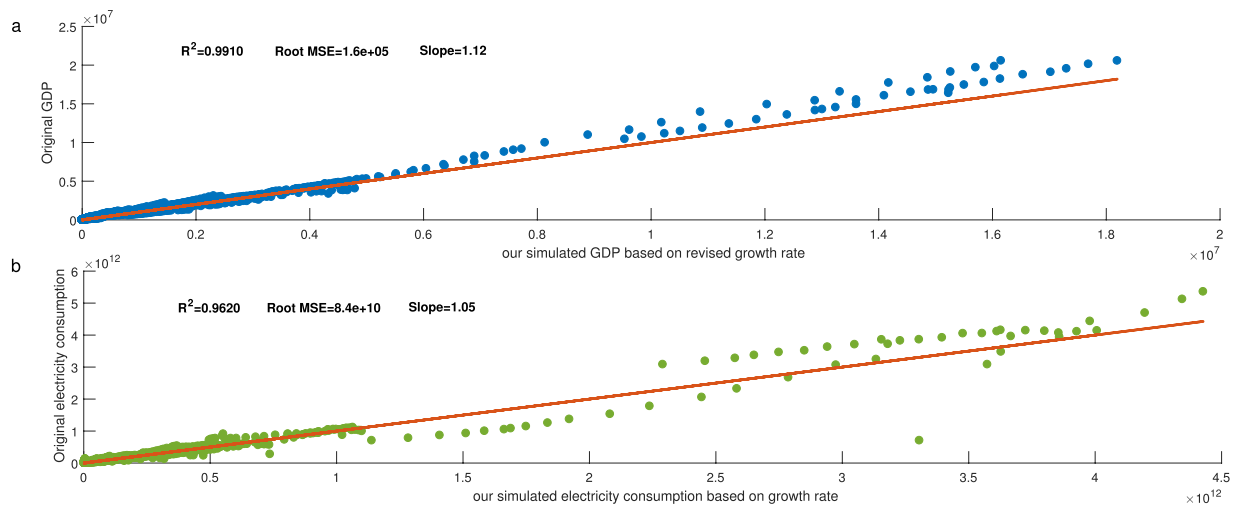
**Fig. 7** Validation and testing results for the relationship between national normalized actual GDP/ electricity consumption and our estimated GDP/ electricity consumption predicted based on the input factors. **(a–d)**. **(a)** Validation results for the relationship between national normalized actual GDP and our estimated GDP predicted based on the input factors; **(b)** Validation results for the relationship between national normalized actual electricity consumption and our estimated electricity consumption predicted based on the input factors; **(c)** Testing results for the relationship between national normalized actual GDP and our estimated GDP consumption predicted based on the input factors; **(d)** Testing results for the relationship between national normalized actual electricity consumption and our estimated electricity consumption predicted based on the input factors.

### Usage Notes

Since the  $1 \text{ km} \times 1 \text{ km}$  gridded GDP and electricity consumption were estimated based on the top-down idea, the scope of the gridded data did not involve all of the worldwide countries (175 countries' or regions' GDP data; 134 countries' or regions' electricity consumption). Thus, the values of the area beyond the scope of our study were set to 0. The official GDP data were derived from the Penn World Table, and the electricity consumption data were obtained from the World Bank. In addition, the projected coordinate system of all images was set as the Mollweide coordinate.

The provided datasets have the advantages of wide coverage and a long-time span. The datasets can help fill the existing data gaps and can be further used in future research. For example, gridded GDP data growth was revised based on nighttime light data, which is more objective and comparable, and more appropriate for use in research at a more micro-level in countries with poor quality statistics; at the same time, considering that our estimated electricity consumption data is more based on objective satellite data, it will be less affected by administrative intervention. Thus, the comparison between published electricity consumption and our estimated gridded electricity consumption based on satellite data can be employed to reveal failed/poorly governed states.

In addition, it should also be noted that our estimated gridded electricity consumption were based on the assumption that a more developed area generally has brighter lights and higher electricity consumption<sup>15–17</sup>. Thus, shocks like huge price fluctuations or other special events during a particular period were not considered in this study.



**Fig. 8** Scatter plots of our simulated national GDP and electricity consumption with the original data based on existing literature during 1992–2019. **(a,b)** Scatter plots of our simulated **(a)** national GDP compared with the GDP and **(b)** national electricity consumption compared with the electricity consumption based on existing literature during 1992–2019.

### Code availability

The programs used to generate all the results were MATLAB (R2017b) and ArcGIS (10.5). The PSO-BP codes for modelling the relationships between the national GDP and nighttime light data are publicly available under Figshare<sup>43</sup>.

Received: 19 November 2021; Accepted: 8 April 2022;

Published online: 12 May 2022

### References

- Shi, Y., Ren, X., Guo, K., Zhou, Y. & Wang, J. Research on the economic development pattern of Chinese counties based on electricity consumption. *Energy Policy*. **147**, 111881 (2020).
- Cui, W., Li, J., Xu, W. & Güneralp, B. Industrial electricity consumption and economic growth: A spatio-temporal analysis across prefecture-level cities in China from 1999 to 2014. *Energy*. **222**, 119932 (2021).
- Tang, C. F. & Shahbaz, M. Sectoral analysis of the causal relationship between electricity consumption and real output in Pakistan. *Energy Policy*. **60**, 885–891 (2013).
- Long, P. D. The relationship between foreign direct investment, electricity consumption and economic growth in Vietnam. *Int. J. Energy Econ. Policy*. **8**, 267–274 (2018).
- Niu, S. W. *et al.* Does electricity consumption improve residential living status in less developed regions? An empirical analysis using the quantile regression approach. *Energy*. **95**, 550–560 (2016).
- Karatasou, S. & Santamouris, M. Socio-economic status and residential energy consumption: A latent variable approach. *Energy Buildings*. **198**, 100–105 (2019).
- Rawski, T. G. What is happening to China's GDP statistics? *China Econ. Rev.* **12**, 347–354 (2001).
- Ranjan, M. & Jain, V. K. Modelling of electrical energy consumption in Delhi. *Energy*. **24**, 351–361 (1999).
- Tso, G. K. F. & Yau, K. K. W. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy*. **32**, 1761–1768 (2007).
- Henderson, J., Vernon, A. S. & Weil, D. N. Measuring economic growth from outer space. *Am. Econ. Rev.* **102**, 994–1028 (2012).
- Guerrero, V. M. & Mendoza, J. A. On measuring economic growth from outer space: a single country approach. *Empir. Econ* **57**, 971–990 (2019).
- Wang, W., Mu, H., Kang, X., Song, R. & Ning, Y. Changes in industrial electricity consumption in China from 1998 to 2007. *Energy Policy*. **38**, 3684–3690 (2010).
- Cao, X. *et al.* Spatialization of electricity consumption of China using saturation-corrected DMSP-OLS data. *Int. J. Appl. Earth Obs. Geoinf.* **28**, 193–200 (2014).
- Kummu, M., Taka, M. & Guillaume, J. H. Gridded global datasets for gross domestic product and Human Development Index over 1990–2015. *Sci. Data*. **5**, 1–15 (2018).
- Zhao, N., Liu, Y., Cao, G., Samson, E. L. & Zhang, J. Forecasting China's GDP at the pixel level using nighttime lights time series and population images. *GISci. Remote Sens.* **54**(3), 407–425 (2017).
- Wang, X., Sutton, P. C. & Qi, B. Global mapping of GDP at 1 km<sup>2</sup> using VIIRS nighttime satellite imagery. *ISPRS Int J Geoinf.* **8**(12), 580 (2019).
- Shi, K. F. *et al.* Detecting spatiotemporal dynamics of global electric power consumption using DMSP-OLS nighttime stable light data. *Appl. Energy*. **184**, 450–463 (2016).
- Chen, M. L. *et al.* A novel classification regression method for gridded electric power consumption estimation in China. *Sci. Rep.* **10**, 1–12 (2020).
- Holz, C. A. The quality of China's GDP statistics. *China Econ. Rev.* **30**, 309–338 (2014).
- Sandefur, J. & Glassman, A. The political economy of bad data: Evidence from African survey and administrative statistics. *J. Dev. Stud.* **51**, 116–132 (2015).
- Chen, X. & Nordhaus, W. D. Using luminosity data as a proxy for economic statistics. *PNAS*. **108**, 8589–8594 (2011).
- Hu, Y. & Yao, J. Illuminating economic growth. *J. Econom.* <https://doi.org/10.1016/j.jeconom.2021.05.007> (2021).
- Keola, S., Andersson, M. & Hall, O. Monitoring economic development from space: Using nighttime light and land cover data to measure economic growth. *World Dev.* **66**, 322–334. <https://doi.org/10.1016/j.worlddev.2014.08.017> (2015).
- National Centers for Environmental Information <https://ngdc.noaa.gov/eog/download.html>.
- National Centers for Environmental Information [https://ngdc.noaa.gov/eog/dmsp/download\\_radcal.html](https://ngdc.noaa.gov/eog/dmsp/download_radcal.html) (2006).

26. 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, <https://doi.org/10.1080/2150704X.2014.890758> (2014).
27. Chen, Z. *et al.* An extended time series (2000–2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration. *Earth Syst. Sci. Data*. **13**, 889–906 (2021).
28. Li, X., Zhou, Y., Zhao, M. & Zhao, X. A harmonized global nighttime light dataset 1992–2018. *Sci. Data*. **7**, 1–9, <https://doi.org/10.1038/s41597-020-0510-y> (2020).
29. 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**(1), 0–72 (2012).
30. Hillger, D. *et al.* First-light imagery from Suomi NPP VIIRS. *Bull. Am. Meteorol. Soc.* **94**, 1019–1029 (2013).
31. Chen, J. *et al.* County-level CO<sub>2</sub> emissions and sequestration in China during 1997–2017. *Sci. Data*. **7**, 1–12 (2020).
32. Zhao, J. *et al.* Spatio-temporal dynamics of urban residential CO<sub>2</sub> emissions and their driving forces in China using the integrated two nighttime light datasets. *Appl. Energy*. **235**, 612–624 (2019).
33. Lv, Q., Liu, H., Wang, J., Liu, H. & Shang, Y. Multiscale analysis on spatiotemporal dynamics of energy consumption CO<sub>2</sub> emissions in China: Utilizing the integrated of DMSP-OLS and NPP-VIIRS nighttime light datasets. *Sci. Total Environ.* **703**, 134394, <https://doi.org/10.1016/j.scitotenv.2019.134394> (2020).
34. Chen, Z. *et al.* An Extended Time Series (2000–2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration. *Earth Syst. Sci. Data* **13**, 889–906 (2021).
35. Stecklov, G., Weinreb, A. & Carletto, C. Can incentives improve survey data quality in developing countries?: Results from a field experiment in India. *J. R. Stat. Soc. A* **181**, 1033–1056, <https://doi.org/10.1111/rssa.12333> (2018).
36. Fantom, N. & Serajuddin, U. *The World Bank's classification of countries by income*. (The World Bank, 2016).
37. Zhang, C., Zhou, K., Yang, S. & Shao, Z. On electricity consumption and economic growth in China. *Renew. Sust. Energy Rev.* **76**, 353–368 (2017).
38. Chen, H. B., Pei, L. L. & Zhao, Y. F. Forecasting seasonal variations in electricity consumption and electricity usage efficiency of industrial sectors using a grey modeling approach. *Energy*. **222**, 119952 (2021).
39. Meng, L., Graus, W., Worrell, E. & Huang, B. Estimating CO<sub>2</sub> (carbon dioxide) emissions at urban scales by DMSP/OLS (Defense Meteorological Satellite Program's Operational Linescan System) nighttime light imagery: Methodological challenges and a case study for China. *Energy* **71**, 468–478 (2014).
40. Su, Y. *et al.* China's 19-year city-level carbon emissions of energy consumptions, driving forces and regionalized mitigation guidelines. *Renew. Sust. Energy Rev.* **35**, 231–243 (2014).
41. Yang, D., Luan, W., Qiao, L. & Pratama, M. Modeling and spatio-temporal analysis of city-level carbon emissions based on nighttime light satellite imagery. *Appl. Energy* **268**, 114696 (2020).
42. Chen, J. D. & Gao, M. Global 1 km × 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. *figshare. Dataset*. <https://doi.org/10.6084/m9.figshare.17004523.v1> (2021).
43. Gao, M. & Chen, J. PSO-BP codes for modelling the relationships between the national GDP and nighttime light data. *figshare. Dataset*. <https://doi.org/10.6084/m9.figshare.19517272.v1> (2022).
44. Sulla-Menashe, D. & Friedl, M. A. User Guide to Collection 6 MODIS Land Cover (MCD12Q1) Product. *NASA EOSDIS Land Processes DAAC: Sioux Falls, SD, USA* (2018).
45. 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).
46. Elvidge, C. D. *et al.* Night-Time Lights of the World: 1994–1995. *ISPRS J. Photogramm.* **56**, 81–99 (2001).
47. Boussemart, J. P., Leleu, H. & Shen, Z. Y. Worldwide carbon shadow prices during 1990–2011. *Energ. Policy*. **109**, 288–296 (2017).

## Acknowledgements

This work was supported by the National Key Social Science Foundation of China [Grant number 21ATJ008] and the National Natural Science Foundation of China [72125010, 71974186].

## Author contributions

C.J. and G.M. led the project and provided ideas for this research. G.M. collected the raw data, calculated and assembled the data, and prepared the initial manuscript. C.S. collected the raw data and participated in the database construction. H.W. and S.M. revised the manuscript. L.Y. designed the study. L.X. provided the technical support.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.G. or Y.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022