



An extended time-series (2000-2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration

Zuoqi Chen^{1,2}, Bailang Yu^{3,4}, Chengshu Yang^{3,4}, Yuyu Zhou⁵, Xingjian Qian^{3,4}, Congxiao Wang^{3,4}, Bin Wu^{3,4}, Jianping Wu^{3,4}

5 ¹Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education, National & Local Joint Engineering Research Center of Satellite Geospatial Information Technology, Fuzhou University, Fuzhou 35002, China

²The Academy of Digital China, Fuzhou University, Fuzhou 350002, China

³Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China

10 ⁴School of Geographic Sciences, East China Normal University, Shanghai 200241, China

⁵Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA

Correspondence to: Bailang Yu (blyu@geo.ecnu.edu.cn)

Abstract: The nighttime light (NTL) satellite data have been widely used to investigate urbanization process. The Defense Meteorological Satellite Program-Operational Linescan System (DMSP-OLS) stable nighttime light data and Suomi National
15 Polar-Orbiting Partnership-Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) nighttime light data are two widely used NTL datasets. However, the difference of their spatial resolutions and sensor design makes these two datasets can not be directly used together for a long-term analysis of urbanization. To solve this issue, an extended time-series (2000-2018) of NPP-VIIRS-like NTL data were proposed in this study through a cross-sensor calibration from DMSP-OLS NTL data (2000-2012) and a composition of monthly NPP-VIIRS NTL data (2013-2018). Compared with the annual composited NPP-VIIRS
20 NTL data in 2012, our product of extended NPP-VIIRS-like NTL data shows a good consistency at the pixel and city levels with R^2 of 0.87 and 0.95, respectively. We also found that our product has a good accuracy by comparing with DMSP-OLS radiance calibrated NTL (RNTL) data in 2000, 2004, 2006, and 2010. Generally, our extended NPP-VIIRS-like NTL data (2000-2018) have a good spatial pattern and temporal consistency, which are similar to the composited NPP-VIIRS NTL data. In addition, the resulting product could be easily updated and provide a useful proxy to monitor the dynamics of demographic
25 and socio-economic activities for a longer time period compared to existing products. The extended time-series (2000-2018) of nighttime light data are freely accessible at <https://doi.org/10.7910/DVN/YGIVCD> (Chen et al., 2020).

1. Introduction

Along with the artificial electric light has been widely equipped in most buildings and infrastructures, the nighttime light (NTL)
30 remote sensing data have been widely used to investigate human activities (Gaston et al., 2013;Falchi et al., 2011;Elvidge et al., 1997a;Baugh et al., 2013;Li et al., 2018). So far, two NTL data, the Defense Meteorological Satellite Program-Operational



Linescan System (DMSP-OLS) stable nighttime light data and Suomi National Polar-Orbiting Partnership-Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) nighttime light data, are widely used to monitor and diagnose urbanization process. For instance, the population (Sutton et al., 2001;Xu et al., 2015;Elvidge et al., 1997b;Yu et al., 2018) and economic development (Zhao et al., 2017;Lo, 2002;Ma et al., 2012;Yu et al., 2015;Zhao et al., 2019b) has been successfully estimated, 35 the energy consumption (Shi et al., 2018;Shi et al., 2016b) and environment issues (Ou et al., 2013;Shi et al., 2016a;Liu et al., 2018;Jiang et al., 2018) have been well revealed, and the urban area (Shi et al., 2014a;Cao et al., 2009;Zhou et al., 2014;Chen et al., 2019;Zhou et al., 2015) and its spatial structure (Chen et al., 2015;Lu et al., 2018;Chen et al., 2017b;Yu et al., 2014;Wu et al., 2019) have also been effectively detected. Therefore, these two NTL datasets are both good proxies of detecting the 40 dynamics of demographic and socio-economic activities at different spatial scales (Yang et al., 2019).

Unfortunately, these applications were always limited by the two NTL datasets' quality and available time span. The DMSP-OLS NTL annually composited data are available from 1992 but ended in 2013 (Fig. 1). It has three issues, including lack of on-orbit radiance calibration, saturation issue and blooming issue (Letu et al., 2010;Cao et al., 2019;Elvidge et al., 2014;Levin et al., 2020), which limit its potential applications. The NPP-VIIRS NTL data have a better data quality (e.g., higher spatial 45 resolution of ~500 m, etc.) and a superior detection ability, but its short available time span limited its power for long-term analysis and applications. As shown in Fig. 1, the monthly composited data is from April 2012 to present, while the annual NPP-VIIRS NTL data covers only two years: 2015 and 2016. In addition, DMSP-OLS NTL data records the digital number (DN), which is totally different from the radiance value in NPP-VIIRS NTL data. Consequently, these two NTL data are not comparability and could not directly be used together. Therefore, an extended time-series of nighttime light data with an 50 appropriate quality and a better consistency becomes critical for the further temporal nighttime light applications.

(Insert Figure 1 near here)

Until now, literatures are limited for extending NTL data by integrating DMSP-OLS data and NPP-VIIRS NTL data (JESWANI, 2017). Among them, Shao et al. (2014), according to the NPP-VIIRS Day/Night band data and lunar irradiance model, developed a vicarious radiometric calibration for DMSP-OLS daily NTL data. However, this model requires selecting specific 55 events at night as criteria and is not suitable for the annual DMSP-OLS NTL composite data. Zhu et al. (2017) and Li et al. (2017) both attempted to use a power function for integrating the DMSP-OLS NTL data and NPP-VIIRS NTL data. Zhu et al. (2017) fitted the power function by using the cumulated DMSP-OLS and NPP-VIIRS NTL intensity within each province in China from 1992 to 2015. Then this power function was applied to the cumulated NPP-VIIRS NTL intensity to generate simulated DMSP-OLS NTL intensity. But the power function from Li et al. (2017) was fitted from annual DMSP-OLS NTL 60 data and monthly NPP-VIIRS NTL data. This power function was then conducted to inter-calibrate these two NTL data for analyzing Syria's major human settlement loss during a war. These power functions are both heavily relied on the strategy of training samples selection and is not easy to be extended to other world regions or the entire world. Instead of using power



function and traditional DMSP-OLS stable NTL data, Zheng et al. (2019) conducted a geographically weighted regression model to fit the radiance-calibrated DMSP-OLS NTL data and NPP-VIIRS NTL data, then generated the DMSP-like NPP-
65 VIIRS NTL data for the further research. Zhao et al. (2019a) proposed a sigmoid function model with a series of pre-processing procedures to convert NPP-VIIRS NTL data into simulated DMSP-OLS NTL data from 1992 to 2018 in Southeast Asia. Li et al. (2020) provided a global DMSP-OLS-like NTL data, called as harmonized DMSP-OLS NTL data, through a stepwise calibration of DMSP-OLS NTL data and an kernel density-based integration of calibrated DMSP-OLS and NPP-VIIRS NTL data. However, since NPP-VIIRS NTL data has a better quality than DMSP-OLS NTL data, the performance of estimating
70 social-economic index and extracting urban spatial structure from NPP-VIIRS NTL is much higher (Shi et al., 2014b; Chen et al., 2017b). Therefore, simulating an extended time-series of NPP-VIIRS-like NTL dataset other than traditional DMSP-OLS-like NTL data, would be much helpful for the further analysis and applications. Given the difficulties mentioned above, a new approach to cross-sensor calibrating these two NTL data is still challengeable.

Recently, deep learning technologies present a great potential for image process, such as image restoration, image denoising,
75 and target recognition/classification (Goodfellow et al., 2016). An auto-encoder model proposed by Hinton and Zemel (1994) contains a set of recognition weights for encoding the input data and a set of generative weights for reconstructing a similar input data. With convolutional neural networks (CNN), the auto-encoder model becomes more powerful for learning image high-level features and enhancing the input image quality (Wang and Tao, 2016; Jain and Seung, 2009). For instance, Tan and Eswaran (2008) and Vincent et al. (2010) successfully developed a stacked auto-encoder network with CNN to reconstruct and
80 denoise handwritten digital images, respectively. Chen et al. (2017a) applied a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN) to enhance a CT image from low-dense to normal dense. By analogy, the DMSP-OLS NTL data are like the low-dense image with some noises, while the NPP-VIIRS NTL data can be treated as the high-quality image. According to the successful cases from the literatures mentioned above, we deeply believed the auto-encoder model has an ability of converting DMSP-OLS NTL data to NPP-VIIRS NTL data.

85 In this study, we developed an auto-encoder (AE) model including convolutional neural networks to integrate DMSP-OLS NTL and NPP-VIIRS NTL data, and generated an extended time-series of global annual NPP-VIIRS-like NTL data from 2000 to 2018. The Google Earth Engine (GEE) platform (Kumar and Mutanga, 2018) and a parallel computing platform named as Compute Unified Device Architecture (CUDA) of the graphics processing unit (Huang et al., 2015) were used in the proposed framework. The remainder of this paper is organized as follows. Section 2 describes the data involved in this study and the
90 illustration of data preprocessing. The auto-encoder network structure and cross-sensor calibration are presented in Section 3. In Section 4 and 5, the extended time-series of global NPP-VIIRS-like NTL data, accuracy evaluation, spatial pattern and temporal consistency are discussed. The findings are summarized in the final section.



2. Data

This study used three datasets (Table 1). The first one is the enhanced vegetation index-adjusted NTL index (EANTLI) of
95 2000-2013 (Zhuo et al., 2015), as an input data in the AE model, from the annual calibrated DMSP-OLS NTL data and
Enhanced Vegetation Index (EVI) products from Moderate Resolution Imaging Spectroradiometer. The second dataset is the
composited NPP-VIIRS NTL data from 2012 to 2018, which was annually summarized from the monthly NPP-VIIRS NTL
data. This 2012 composited NPP-VIIRS NTL data was used as the label data in AE model while 2013 being the reference for
validation, and 2013-2018 were appended to the NPP-VIIRS-like NTL data (2000-2012) simulated by AE model. The third
100 one is the DMSP-OLS radiance calibrated NTL (RNTL) data, which were for the validation procedure.

(Insert Table 1 near here)

2.1. Enhanced vegetation index-adjusted NTL index (EANTLI)

Due to the mentioned issues in the annual DMSP-OLS NTL data, an enhanced vegetation index-adjusted NTL index (EANTLI)
of 2000-2013 was used as an input in the AE model to simulate NPP-VIIRS-like NTL data. EANTLI was proposed by Zhuo
105 et al. (2015) by fusing the EVI and DMSP-OLS NTL data. The EANTLI not only reduces the saturation issues, but also
enhances the nighttime light intensity variations. It can be expressed mathematically in the equation:

$$\text{EANTLI} = \frac{1+(n\text{NTL}-\text{EVI})}{1-(n\text{NTL}-\text{EVI})} \times \text{NTL} \quad (1)$$

where *EVI* represents the value from the annual average EVI value, *NTL* is the DMSP-OLS NTL intensity, while *nNTL*
indicates the normalized *NTL*.

110 In this study, to get an EANTLI dataset with a better temporal consistency, the calibrated DMSP-OLS NTL data replaced the
original DMSP-OLS NTL data in Eq. (1). The calibrated DMSP-OLS NTL of 2000–2013 used in this study were derived from
the original annual DMSP-OLS NTL data via a stepwise calibration (Li and Zhou, 2017). The calibrated DMSP-OLS NTL
data have a better temporal consistency and a greater agreement with the NPP-VIIRS NTL data (Li et al., 2020). This calibrated
NTL data covers the same extent (180 to 180 degrees longitude and -65 to 75 degrees latitude) and spatial resolution (30 arc-
115 seconds). The data record the digital number (DN) values with a range from 0 to 63. 0 indicates the no nighttime light intensity,
as background, while 63 indicates the highest nighttime light intensity.

The 16-Day EVI products (MOD13A1) with spatial resolution of 500 m are involved. To mitigate the sensitivity to seasonal
and inter annual fluctuations, the MOD13A1 EVI products were processed as an annual average EVI (Jing et al., 2015). It is
worthy noted that since the MOD13A1 EVI data is only available from 2000 onwards, we only processed the Enhanced
120 vegetation index-adjusted NTL index (EANTLI) of 2000-2013. More details and advantages of calibrated DMSP-OLS NTL
and EANTLI can be found in Li et al. (2020) and Zhuo et al. (2015).



2.2. Composited NPP-VIIRS NTL data

The monthly NPP-VIIRS NTL data were calibrated and aggregated to the annual NPP-VIIRS NTL data from 2012 to 2018. It is worthy to note that the composited NPP-VIIRS NTL data of 2012 was for the validation process, the composited NPP-VIIRS NTL data of 2013 was for the training process in AE model, and the composited NPP-VIIRS NTL data from 2013 to 2018 were appended to the final product as a part of NPP-VIIRS-like NTL data.

The version 1 monthly NPP-VIIRS NTL composite data (vcm version) of April 2012 to December 2018 provided by Colorado School of Mines were used to composite annual NPP-VIIRS NTL data. The monthly NPP-VIIRS NTL data cover the same extent of calibrated DMSP-OLS NTL data with a finer spatial resolution of 15 arc-seconds (approximately 500 m near equator) and a more sensitive sensor with a unit of $\text{nano-Wcm}^{-2}\text{sr}^{-1}$. Since the official annual NPP-VIIRS NTL data are only available in 2015 and 2016 and it requires inaccessible parameters to repeat the official annual composite process for other years (Elvidge et al., 2017), we composited a new annual NPP-VIIRS NTL data by using the median value of 12 NPP-VIIRS NTL monthly composite data per each pixel. To differentiate with the official annual NPP-VIIRS NTL data (2015 and 2016), we named our annual NPP-VIIRS NTL data as composited NPP-VIIRS NTL data in the following sections.

We used the median value, instead of the traditional average or max value because the monthly NPP-VIIRS NTL data were contaminated by stray light in the mid-to-high latitudes region during the entire summer and the contaminated pixels are reassigned as 0 in the official monthly composite data (Elvidge et al., 2017). As the mean value could lower the NTL intensity and the max value could highlight abnormal value, the median value could be more reliable than the mean or max value (Liu et al., 2010). A validation of this median composited NPP-VIIRS NTL was conducted as described in the following section.

According to the correction model of NPP-VIIRS NTL data proposed by Shi et al. (2014b) and Ma et al. (2014), a dark background mask and a nighttime light intensity threshold value are required to remove pixels of unstable and abnormal nighttime light intensity, respectively. First, the dark background mask consists of the EANTLI pixels with value of 0 and the NPP-VIIRS NTL pixels with intensity lower than $1 \text{ nano-Wcm}^{-2}\text{sr}^{-1}$. We then filtered the corresponding annual NPP-VIIRS data by using this dark background mask. Then, according to global city rank from the Globalization and World Cities (GaWC) Research Network (Taylor et al., 2010), Shanghai and Beijing in China and New York City in USA were selected in this study to find out their highest nighttime light intensity as a threshold value for each year. Once an NPP-VIIRS NTL pixel value is higher than this threshold value, this pixel value will be replaced by its maximum NTL intensity within its eight neighbor pixels to eliminate the abnormal value.

2.3. DMSP-OLS radiance calibrated NTL (RNTL) data

DMSP-OLS RNTL data in 2000, 2004, 2006, and 2010 were selected as reference data to evaluate if our product has a good accuracy during the entire time series, since the NPP-VIIRS NTL data is only available after 2012. The DMSP-OLS RNTL



data have a radiance calibration based on pre-flight sensor parameters and are free from sensor saturation (Feng-Chi et al., 2015). However, this dataset is still unitless, because the lacking of on-board calibration system for all DMSP-OLS makes the degradation of sensors over time can not be measured precisely. DMSP-OLS RNTL data has the same extent of the original
155 DMSP-OLS NTL data with the same spatial resolution (30 arc-seconds), but it is only available in specific years. The four selected DMSP-OLS RNTL data were accessed from Colorado School of Mines (https://eogdata.mines.edu/dmsp/download_radcal.html).

3. Methodology

As outlined in Fig. 2, after data preprocess mentioned above, a four-step approach was proposed to generate an extended time-series (2000-2018) of NPP-VIIRS-like NTL data from the EANTLI data and composited NPP-VIIRS NTL data. In step 1, a
160 modified auto-encoder model was developed. In step 2 and 3, the architecture of auto-encoder model with CNN was designed and the cross-sensor calibration model was trained. In step 4, an extended time-series (2000-2018) of NPP-VIIRS-like NTL dataset was generated using the trained model by inputting the 2000-2012 EANTLI data and appending the composited NPP-VIIRS NTL data (2013-2018). Finally, a comprehensive accuracy evaluation was conducted.

165 (Insert Figure 2 near here)

3.1. The auto-encoder (AE) model

The auto-encoder model was trained by setting the 2012 EANTLI as input data and the 2012 composited NPP-VIIRS NTL data as label. Then, the trained model was adopted to simulate the NPP-VIIRS-like NTL data by inputting the 2000-2011 EANTLI data. The auto-encoder model includes two main parts (encoder and decoder) as shown in Fig. 3. Let $\mathbf{X} \in \mathbf{R}^{m \times n}$,
170 $\mathbf{Y} \in \mathbf{R}^{m \times n}$, and $\hat{\mathbf{Y}} \in \mathbf{R}^{m \times n}$ be the annual EANTLI data, composited NPP-VIIRS NTL data, and simulated NPP-VIIRS-like NTL data. The encoder part is to learn a deterministic mapping f_{θ} which could transfer \mathbf{X} into a hidden representation (\mathbf{H}). A typical deterministic mapping function can be expressed as:

$$f_{\theta}(x) = s(Wx + b) \quad (2)$$

where $x \in \mathbf{X}$, and θ represents the parameter set, including weight matrix (W) and offset (b). On the contrary, the traditional
175 decoder part is to reconstruct \mathbf{X} using the high-level features extracted from the hidden representation (\mathbf{H}). This reconstruction $g_{\theta'}$ is called as decoder and can be expressed as:

$$\hat{\mathbf{Y}} = g_{\theta'}(h) = s(W'h + b') \quad (3)$$

where $h \in \mathbf{H}$, and θ' represents the parameter set, including weight matrix (W') and offset (b'). However, in this study the decoder part was modified to map the composited NPP-VIIRS NTL data (\mathbf{Y}), rather than the traditional reconstruction of \mathbf{X} ,



180 which means that the problem can be transformed to build two functions f_{θ} and $g_{\theta'}$ via deep learning technology to minimize a specific loss function (e.g., Mean Square Error):

$$\arg \min_{f,g} \|\hat{\mathbf{Y}} - \mathbf{Y}\|^2 \quad (4)$$

(Insert Figure 3 near here)

3.2. The Auto-encoder with a CNN architecture design

185 AE and CNN have both demonstrated an excellent performance on the image feature extractions. In this study, based on AE and CNN framework, we proposed a 10-layers network architecture as the cross-sensor calibration model (Fig. 4). This model includes five convolutional operations for encoding EANTLI data and then stacks 5 deconvolutional operations for decoding in symmetry. Because our intention is to reconstruct images, instead of to classify targets, the fully-connected layers in traditional encoder and decoder parts were removed from our architecture.

190 (Insert Figure 4 near here)

The kernel size of convolutional and deconvolutional operations adopted in this architecture is 3 by 3 with stride and padding of 1 to maintain the size of output layers the same as its input layer. In the encoder part, the batch normalization (BN) operations were added after each convolution layer to keep from the vanishing or exploding gradient problem (Ioffe and Szegedy, 2015). The rectified linear unit (ReLU) function was applied in this architecture as the activation function after each convolutional and deconvolutional layer, except the last deconvolutional layer. The ReLU function can be formed as:

$$\text{ReLU}(x) = \max(0, x) \quad (5)$$

The traditional CNN structure always contains more than one pooling operation to improve its learning efficiency, but these pooling operations could lose the details of the input images. To keep the image information as much as possible, the pooling processes among all layers were deleted.

200 3.3. Training auto-encoder model with a CNN architecture

Considering the balance between computational limitation and efficiency, the EANTLI data and composited NPP-VIIRS NTL data were severally split into tiles of 256×256 pixels. Since these two NTL data both cover the year of 2013, we built a training set of paired tiles, as $P = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$ by using the 2013 EANTLI data and composited NPP-VIIRS NTL data, respectively, where N indicates the number of tiles. Then this training set was input to train the auto-encoder model designed above by minimizing the loss function \mathcal{L} between composited NPP-VIIRS NTL data (\mathbf{Y}) and simulated NPP-VIIRS-like NTL data ($\hat{\mathbf{Y}}$). The loss function adopted in this study is Mean Square Error function and was then optimized by Adam in each deep learning step (Kingma and Ba, 2014). The loss function can be formed as:



$$\mathcal{L} = \frac{1}{N} \sum_{n=1}^N (\hat{Y}_n - Y_n)^2 \quad (6)$$

3.4. Generating an extended time-series of NPP-VIIRS-like NTL data

210 We generated the extended time-series (2000-2018) of NPP-VIIRS-like NTL data with two components. First, the trained AE model was applied to the 2000-2012 EANTLI data for generating the simulated NPP-VIIRS-like NTL data covering the same period. Second, we composited annual NPP-VIIRS NTL data (2013-2018) from monthly ones. Via appending these two components, the extended time-series (2000-2018) of global NPP-VIIRS-like NTL data were generated.

The split tiles of 2000-2012 EANTLI data with the same size (256 by 256) were input into the trained auto-encoder model to
 215 simulate the NPP-VIIRS-like NTL data. Due to the fluctuation of EVI data, the input EANTLI data have several abnormal pixels which make the output NTL pixels unreasonable. Thus, a post process is required for the simulated NPP-VIIRS-like data, which includes three parts. Firstly, the EANTLI pixels with DN value of 0 were extracted as a dark background mask for each year and overlaid with the simulated NPP-VIIRS-like data to assign the pixels in the same locations as 0, because it is not reasonable to simulate lights at where there is no stable light source. Meanwhile, because of the NPP-VIIRS sensor's
 220 detection limitation, the simulated NTL intensity lower than 1 nano-Wcm⁻²sr⁻¹ was also assigned as 0 (Ma et al., 2014). The third post process was designed to make sure the simulated NPP-VIIRS NTL data have the same temporal change as the calibrated DMSP-OLS NTL images. The performance equation was:

$$\text{SNTL}_{(year,i)} = \begin{cases} \text{SNTL}_{(year+1,i)} & \text{NTL}_{(year,i)} > \text{NTL}_{(year+1,i)} \cap \text{SNTL}_{(year,i)} < \text{SNTL}_{(year+1,i)} \\ \text{SNTL}_{(year+1,i)} & \text{NTL}_{(year,i)} < \text{NTL}_{(year+1,i)} \cap \text{SNTL}_{(year,i)} > \text{SNTL}_{(year+1,i)} \\ \text{SNTL}_{(year,i)} & \text{otherwise} \end{cases} \quad (7)$$

225 where $\text{SNTL}_{(year,i)}$ and $\text{NTL}_{(year,i)}$ indicates the simulated NPP-VIIRS-like intensity and calibrated DMSP-OLS NTL intensity of i th pixel in $year$ (from 2000 to 2012).

4. Results

4.1. Training of the auto-encoder model with CNN

In the training process of AE model, the learning rate in this study was initialized as 1e-4, and optimized according to Adam algorithm proposed by Kingma and Ba (2014). For weight initialization, the method proposed by He et al. (2015) was adopted
 230 in this study, instead of the traditional random weights from Gaussian distribution. In Fig. 5, the reconstruction loss of each epoch is illustrated. We can see that with the increase of epoch, the loss value decreases and becomes stable near the loss value of 200. This means that once the epoch is larger than some kind of threshold, keeping iteratively training this model is not helpful to improve the accuracy. Therefore, the epoch number is fixed as 4000 based on this trial-and-error experiments.

(Insert Figure 5 near here)



235 **4.2. Accuracy evaluation**

According to the pixel-level and city-level validation between extended NPP-VIIRS-like NTL data and composited NPP-VIIRS NTL data of 2012 (Fig. 6), our result is close to the composited NPP-VIIRS NTL data at both spatial scales. At pixel level, more than 1 hundred thousand random pixels were selected as validation points and the coefficient of determination (R^2) between our results and composited NPP-VIIRS NTL data of 2012 was 0.87 with Root-Mean-Squared-Error (RMSE) of 2.96 nano- Wcm^2sr^{-1} , at the significant level of 0.05 (Fig. 6(a)). For city-level validation, the total NTL intensity of 40000 selected cities were adopted as variables and the results shown that the extended NPP-VIIRS-like NTL data has a better performance with R^2 of 0.94 and RMSE of 3024.62 nano- Wcm^2sr^{-1} (Fig. 6(b)). Most of the dots in these two scatter plots are around the nonbias (1:1) line (red line in Fig. 6), which implies that the extended NPP-VIIRS-like NTL data have a positive 1:1 relationship with the composited NPP-VIIRS NTL data.

245 (Insert Figure 6 near here)

Among the global validation points, we also selected six sub-sets within six countries (United States, Italy, China, Brazil, South Africa, and Australia) and found out that the accuracy in these countries are all acceptable and has no significant spatial variation (Fig. 7). Brazil has the highest accuracy ($R^2 = 0.86$), followed by the United States ($R^2 = 0.84$). The rest countries all have an accuracy higher than 0.7. Australia has the R^2 of 0.79, while Italy and China have the R^2 of 0.76 and 0.72, respectively. South Africa also has a good R^2 of 0.70. Meanwhile, it is worthy of note that in these six sample countries, the RMSE of each country is very small (from 1.67 to 5.72), which means our results are very similar to the calibrated NPP-VIIRS NTL intensity at pixel level.

(Insert Figure 7 near here)

Fig. 8 shown that our extended NPP-VIIRS-like NTL data have a strong agreement with the DMSP-OLS RNTL data in four years (2000, 2004, 2006, and 2010), which implies that the AE model is suitable for simulating NPP-VIIRS-like NTL data during the entire period. Before 2012, the composited NPP-VIIRS NTL data is not available for the validation, but the DMSP-OLS RNTL data is accessible in some separated years. To validate our results before 2012, we have to use the DMSP-OLS RNTL data as the reference data. The DMSP-OLS RNTL data was calibrated by using pre-flight sensor calibration, but has no actual radiance value. Thus, this validation was conducted at city level and the total NPP-VIIRS-like NTL intensity and DMSP-OLS RNTL intensity of each city were calculated and scattered (Fig. 8). In these four years, all the accuracies (R^2) are higher than 0.75 and demonstrate our model does work for this entire time series.

260 (Insert Figure 8 near here)

Finally, the evaluation of the composited NPP-VIIRS NTL data indicates that it is close to the official annual NPP-VIIRS NTL data in 2015, as shown in Fig. 9, according to a comparison between these two data in 2015 based on 5000 random validation points. In other words, using the median value to composite annual NPP-VIIRS NTL data is reasonable and appropriate. In



Fig. 9(a), the similar distribution of green bars and blue bars in these two histogram plots indicated that the pixel number of composited and official annual NPP-VIIRS NTL data, is close to each other, within each bin of 10 nano-Wcm²sr⁻¹. The result illustrates that these two datasets have a significant statistic similarity. Meanwhile, the scatter plot between these two NTL intensity at pixel level (Fig. 9(b)) showed that our composited NPP-VIIRS NTL data has a strong linear relationship ($R^2=0.85$ and slope is close to 1) with the official annual NPP-VIIRS NTL data. The two validation results both proved that the
270 composited annual NPP-VIIRS NTL data generation model based on median value is a reasonable solution.

(Insert Figure 9 near here)

4.3. The extended time-series (2000-2018) of global NPP-VIIRS-like NTL data

Fig. 10 shows the global spatial distribution of extended NPP-VIIRS-like NTL in 2012, including three enlarged subplots of
275 New York, Rome, and Shanghai. From each enlarged subplot (Fig. 10(b)-(d)), our results could provide more information for urbanization evaluation, such as the road network and urban spatial structure. The comparison among these three cities shown that New York have much higher NTL intensity than the other two cities and the pixels with high NTL intensity of Rome and Shanghai were located in downtown area.

The dynamics (2000-2018) of extended NTL intensity along with latitude and longitude were also plotted in Fig. 10(e) and (f).
280 At the global scale, the region between 20°N to 45°N concentrates the most NTL intensity, while the NTL intensity is much lower in the southern hemisphere. From 2000 to 2010, the region between 30°N to 45°N has no significant change of NTL intensity, but has a great enhancement after 2010. For the region between 15°N to 30°N, the NTL intensity has been increasing during the 15 years, which was mostly caused by the China's development. In the longitudinal direction, one of the NTL intensity peaks within western hemisphere region was mostly located in the USA (from 70°W to 100°W). In eastern hemisphere
285 region, there are three significant peaks in Europe, India, and China (from west to east). The temporal changes of NTL intensity between 2000 and 2005 are generally slighter than that between 2005 to 2010. But from 2010 to 2015, the NTL intensity has a great intensification, even almost all over the world.

(Insert Figure 10 near here)

Fig. 11 reports the trend of NTL intensity from our extended NPP-VIIRS-like NTL data within each continent. Globally, the
290 NTL intensity has more than doubled during the last 18 years, from 60 million to almost 150 million nano-Wcm⁻²sr⁻¹. At the regional scale, Asia and North America have the most intense NTL intensity increase and also appear to be the main contributors to the global urbanization. Europe and South America as the second group both have a stable but slow increase. In Oceania and Africa, the NTL intensity has no obvious increasement, especially before 2010.

(Insert Figure 11 near here)



295 **5. Discussion**

5.1. Evaluation of spatial patterns of extended NPP-VIIRS-like NTL data in 2012

The extended NPP-VIIRS-like NTL data (Fig. 12(IV)), as well as the composited NPP-VIIRS NTL data (Fig. 12(III)), shows a much obvious spatial variations of NTL intensity and less saturation and blooming issue than the calibrated DMSP-OLS data and EANTLI data (Fig. 12(I) and (II)). In Fig. 12, the calibrated DMSP-OLS data have severe saturation issue and blooming issue in all three selected cities: Shanghai, China, Los Angeles, USA, and Cape Town, South Africa. Through fusing the calibrated DMSP-OLS data and EVI data, the quality of the produced EANTLI data have a significant improvement, but it is still suffering a blooming issue. For example, in the periphery of urban area, there are many pixels should be dark but with low NTL intensity, which makes the lit areas to be much larger than the truth. Compared to the EANTLI data, the composited NPP-VIIRS NTL data and our extended NPP-VIIRS-like NTL data have already got a great improvement. Firstly, the main road network can be easily found from the composited or extended NPP-VIIRS NTL data, which indicates that these two datasets can provide more spatial details of the NTL intensity. Secondly, the urban hierarchy structure is much more clearly than before and each lit area was separated, due to the effective elimination of blooming issue.

(Insert Figure 12 near here)

The fluctuation of extended NPP-VIIRS-like NTL data (solid or dashed line) agrees well with that of the composited NPP-VIIRS NTL data (grey part), from the six profiles cross Shanghai, Los Angeles, and Cape Town in Fig. 13. In these three cities, the general trends of extended NPP-VIIRS-like NTL data have a pretty good agreement with the trends of composited NPP-VIIRS NTL data, especially within the urban core area, even though these extended NPP-VIIRS-like NTL profiles showed a little overestimation. Cross Shanghai, parts of the extended NPP-VIIRS-like NTL profiles have an underestimation by comparing with the composited NPP-VIIRS NTL data. This situation mostly appeared outside the urban core area and within the periphery region (e.g., ID: 80 – 90 in Fig. 13(c)).

(Insert Figure 13 near here)

5.2. Evaluation of temporal consistency of the time-series (2000-2018) of NPP-VIIRS-like NTL data

The extended time-series of NPP-VIIRS-like NTL data has a consistent temporal trend by comparing with the time-series of census data and analyzing the range of NTL intensity change nearby the junction point (year of 2012) at global scale and country scale.

Firstly, our NPP-VIIRS-like NTL data have a good agreement with the population trend from 2000 to 2018, as shown in Fig. 14. The census data of the entire world and seven countries were collected from World Bank Open Data (<https://data.worldbank.org/>). A linear regression model was conducted to compare the log-transformed population and total NTL intensity. The R^2 at global scale is 0.84 and the R^2 of seven selected countries ranges from 0.60 (in United States) to 0.94



325 (in China). Totally, this result illustrated that the NPP-VIIRS-like NTL data (2000-2018) have a reasonable temporal trend.
And in China, Afghanistan, and Thailand, the extended NTL data have a better performance (R^2 is larger than 0.8) of population
estimation.

(Insert Figure 14 near here)

Secondly, our extended time-series (2000-2018) of NPP-VIIRS-like NTL data have a smooth temporal change, even nearby
330 the temporal junction point (year of 2012) between the simulated NPP-VIIRS-like NTL data and composited NPP-VIIRS NTL
data. In Fig. 15(a), the total NTL intensity and lit pixels (NTL intensity greater than $1 \text{ nano-Wcm}^{-2}\text{sr}^{-1}$) were measured at global
scale from 2000 to 2018. The total NTL intensity and lit pixel increased steadily in the first ten years (2000-2010). From 2010
to 2014, the changes of total NTL intensity and lit pixel both increased much more quickly than before but are still stable, no
sudden jumps were found during the period (the rectangular box in Fig. 15). In the last four years (2014-2018), the increases
335 of these two indices became small and slow. Consequently, our NPP-VIIRS-like NTL data source change from simulated NPP-
VIIRS-like NTL data (2000-2012) to composited NPP-VIIRS NTL data (2013-2018) does not cause any unreasonable change
of NTL intensity during the entire period at the global scale.

Within the six selected countries, as shown in Fig. 15(b)-(g), the number of lit pixels has a stable trend, while the total NTL
intensity has some fluctuations before 2010. From 2007 to 2010, an obvious reduction of NTL intensity was appeared in these
340 six selected countries, even all over the world. After 2010, most countries were recovered and got an increasement of the NTL
intensity. This temporal change of NTL intensity is properly agreement with the worldwide Great Recession (2007-2010). This
result proves that our extended NPP-VIIRS-like NTL data can truly find out the NTL intensity increasement or decrement
within a country. Meanwhile, we also found that each country has different fluctuation range during the Great Recession. For
example, the United States has a great slump, but the decrement in China is slight, as shown in Fig. 15(b) and (f), which is
345 similar to the conclusion (China is able to withstanding the Great Recession) provided by Wen and Wu (2019). It means that
our product has the capability of revealing the details of NTL intensity change, not only telling increase or decrease.

(Insert Figure 15 near here)

5.3. Limitations

The artificial light from the oceans is not a stable light source (e.g., ship, offshore oil well), and could bring about a
350 misunderstanding during the AE model training. Hence, a land area mask was applied to remove the ocean parts, even though
this procedure could reduce the ability of detecting information from oceans, such as fishery (Waluda et al., 2008) and boats
(Elvidge et al., 2015). After a successful population estimation for evaluating the temporal consistency, we believe our NPP-
VIIRS-like NTL data have a good performance in urban research. Meanwhile, more applications are still required to promote
this dataset.

355 The second issue is about the composited NPP-VIIRS NTL data generation based on its median value. We found that some



abnormal pixels with no lit were occurred in the 2012 median annual NPP-VIIRS NTL data. This issue is because the NPP-VIIRS NTL monthly composite data was only available from April, 2012, and in the high latitude regions the NTL data have been impacted (Román and Stokes, 2015; Levin, 2017; Román et al., 2018). In summary, this lacking data (from January to March) and the stray light in summer make the median NTL intensity become bias (lower than the usual level), even become zero. Thus, the cities or pixels with no lit issue were removed in the validation procedure.

The third one is we only focused on the statistical relationship between DMSP-OLS NTL data and NPP-VIIRS NTL data, without considering some physical parameters (e.g., viewing angles), which indeed could influence the cross-sensor calibrations (Li et al., 2019). For the further research, there are still concerns about the time-series of NTL data with better quality and longer time span.

365 6. Data availability

The extended time-series (2000-2018) of nighttime light data are freely accessible at <https://doi.org/10.7910/DVN/YGIVCD> (Chen et al., 2020)

7. Conclusions

An extended time-series (2000-2018) of NPP-VIIRS-like NTL data was produced in this study. This product includes two parts: the simulated NPP-VIIRS-like NTL data (2000-2012) from DMSP-OLS NTL data and the composited annual NPP-VIIRS NTL data (2013-2018). Compared to the composited NPP-VIIRS NTL data in 2012, our extended NPP-VIIRS-like NTL data show a good accuracy, globally, at the pixel (R^2 : 0.87, RMSE: 2.96) and city levels (R^2 : 0.95, RMSE: 3024.62). At the region scale, all countries show an acceptable accuracy. The R^2 ranges from 0.70 to 0.86 and RMSE is lower than 6 nano- $Wcm^{-2}sr^{-1}$. DMSP-OLS RNTL data in 2000, 2004, 2006, and 2010 were compared with the extended NPP-VIIRS-like NTL data. All R^2 are higher than 0.75, which implies that our model is reliable. These evaluations indicate that our extended NPP-VIIRS-like NTL data has a reasonably good and spatially different quality.

Generally, our extended time-series (2000-2018) of NPP-VIIRS-like NTL data shows a similar spatial pattern as the composited NPP-VIIRS NTL data with a good quality regarding the spatial pattern and temporal consistency. The temporal trend agrees with the population change and the global economy event (e.g., the Great Recession). NTL intensity of our product does not fluctuate around neighbouring years between the simulated NPP-VIIRS-like NTL data (2000-2012) and composited NPP-VIIRS NTL data (2013-2018).

The extended NPP-VIIRS-like NTL data from 2000 to 2018 can be used to better evaluate and analyze the dynamics of demography and urbanization. For example, we can investigate the urban spatial structure, even their road network, and its temporal dynamic for a long time period. Our proposed NTL dataset is available until 2018 so far, but it can be extended in



385 the future when the monthly NPP-VIIRS NTL data for the whole year (e.g., 2019) become available.

Author contributions

Zuoqi Chen: Conceptualization, Methodology, Software, Validation, Formal analysis, Visualization, Writing - Original draft.

Bailang Yu: Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing, Funding acquisition.

Chengshu Yang: Conceptualization, Methodology, Validation, Writing - Review & Editing. **Yuyu Zhou:** Writing - Review &

390 Editing, Validation. **Xingjian Qian:** Methodology, Validation. **Congxiao Wang:** Validation, Writing - Review & Editing. **Bin**

Wu: Methodology, Formal analysis. **Jianping Wu:** Formal analysis, Resources, Supervision.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

395 This research was funded by the National Natural Science Foundation of China (No. 41871331 and 41801343), the China Postdoctoral Science Foundation (2020M671921), the Major Program of National Social Science Foundation of China (17ZDA068), and the Fundamental Research Funds for the Central Universities of China.

Reference

- Baugh, K., Hsu, F.-C., Elvidge, C. D., and Zhizhin, M.: Nighttime lights compositing using the VIIRS day-night band: Preliminary results, *Proceedings of the Asia-Pacific Advanced Network*, 35, 70-86, 10.7125/APAN.35.8, 2013.
- 400 Cao, X., Chen, J., Imura, H., and Higashi, O.: A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data, *Remote Sens Environ*, 113, 2205-2209, 10.1016/j.rse.2009.06.001, 2009.
- Cao, X., Hu, Y., Zhu, X., Shi, F., Zhuo, L., and Chen, J.: A simple self-adjusting model for correcting the blooming effects in DMSP-OLS nighttime light images, *Remote Sens Environ*, 224, 401-411, 10.1016/j.rse.2019.02.019, 2019.
- 405 Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J., and Wang, G.: Low-dose CT with a residual encoder-decoder convolutional neural network, *IEEE transactions on medical imaging*, 36, 2524-2535, 2017a.
- Chen, Z., Yu, B., Hu, Y., Huang, C., Shi, K., and Wu, J.: Estimating House Vacancy Rate in Metropolitan Areas Using NPP-VIIRS Nighttime Light Composite Data, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 2188-2197, 10.1109/JSTARS.2015.2418201, 2015.
- 410 Chen, Z., Yu, B., Song, W., Liu, H., Wu, Q., Shi, K., and Wu, J.: A New Approach for Detecting Urban Centers and Their Spatial Structure With Nighttime Light Remote Sensing, *IEEE T Geosci Remote*, 55, 6305-6319, 10.1109/TGRS.2017.2725917, 2017b.
- Chen, Z., Yu, B., Zhou, Y., Liu, H., Yang, C., Shi, K., and Wu, J.: Mapping Global Urban Areas From 2000 to 2012 Using Time-Series Nighttime Light Data and MODIS Products, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, 12, 1143-1153, 10.1109/JSTARS.2019.2900457, 2019.
- 415 Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., and Davis, E. R.: Mapping city lights with nighttime data from the DMSP operational linescan system, *Photogrammetric Engineering and Remote Sensing*, 63, 727-734, 1997a.



- Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R., and Davis, C. W.: Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption, *International Journal of Remote Sensing*, 18, 1373-1379, 10.1080/014311697218485, 1997b.
- 420 Elvidge, C. D., Hsu, F.-C., Baugh, K. E., and Ghosh, T.: National Trends in Satellite-Observed Lighting 1992-2012, *Global Urban Monitoring and Assessment Through Earth Observation*, 97, 10.1201/b17012-9, 2014.
- Elvidge, C. D., Zhizhin, M., Baugh, K., and Hsu, F.-C.: Automatic Boat Identification System for VIIRS Low Light Imaging Data, *Remote Sensing*, 7, 3020-3036, 10.3390/rs70303020, 2015.
- 425 Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C., and Ghosh, T.: VIIRS night-time lights, *International Journal of Remote Sensing*, 1-20, 10.1080/01431161.2017.1342050, 2017.
- Falchi, F., Cinzano, P., Elvidge, C. D., Keith, D. M., and Haim, A.: Limiting the impact of light pollution on human health, environment and stellar visibility, *Journal of Environmental Management*, 92, 2714-2722, 10.1016/j.jenvman.2011.06.029, 2011.
- 430 Feng-Chi, H., Kimberly, B., Tilotama, G., Mikhail, Z., and Christopher, E.: DMSP-OLS Radiance Calibrated Nighttime Lights Time Series with Intercalibration, *Remote Sensing*, 7, 1855-1876, 10.3390/rs70201855, 2015.
- Gaston, K. J., Bennie, J., Davies, T. W., and Hopkins, J.: The ecological impacts of nighttime light pollution: a mechanistic appraisal, *Biological reviews*, 88, 912-927, 10.1111/brv.12036, 2013.
- Goodfellow, I., Bengio, Y., and Courville, A.: *Deep learning*, MIT press, 2016.
- 435 He, K., Zhang, X., Ren, S., and Sun, J.: Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, 2015.
- Hinton, G. E., and Zemel, R. S.: Autoencoders, minimum description length and Helmholtz free energy, *Advances in neural information processing systems*, 1994, 3-10,
- Huang, Y., Chen, Z., Wu, B., Chen, L., Mao, W., Zhao, F., Wu, J., Wu, J., and Yu, B.: Estimating Roof Solar Energy Potential in the Downtown Area Using a GPU-Accelerated Solar Radiation Model and Airborne LiDAR Data, *Remote Sensing*, 7, 15877, 10.3390/rs71215877, 2015.
- 440 Ioffe, S., and Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift, *arXiv preprint arXiv:1502.03167*, 2015.
- Jain, V., and Seung, S.: Natural image denoising with convolutional networks, *Advances in neural information processing systems*, 2009, 769-776,
- 445 JESWANI, R.: Evaluation of the consistency of DMSP-OLS and SNPP-VIIRS Night-time Light Datasets, Thesis, Geo-Information Science and Earth Observation, University of Twente, 2017.
- Jiang, W., He, G., Leng, W., Long, T., Wang, G., Liu, H., Peng, Y., Yin, R., and Guo, H.: Characterizing Light Pollution Trends across Protected Areas in China Using Nighttime Light Remote Sensing Data, *International Journal of Geo-Information*, 7, 243, 10.3390/ijgi7070243, 2018.
- 450 Jing, W., Yang, Y., Yue, X., and Zhao, X.: Mapping Urban Areas with Integration of DMSP/OLS Nighttime Light and MODIS Data Using Machine Learning Techniques, *Remote Sensing*, 7, 12419, 10.3390/rs70912419, 2015.
- Kingma, D. P., and Ba, J.: Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, 2014.
- Kumar, L., and Mutanga, O.: Google Earth Engine Applications Since Inception: Usage, Trends, and Potential, *Remote Sensing*, 10, 1509, 10.3390/rs10101509, 2018.
- 455 Letu, H., Hara, M., Yagi, H., Naoki, K., Tana, G., Nishio, F., and Shuhei, O.: Estimating energy consumption from night-time DMPS/OLS imagery after correcting for saturation effects, *International Journal of Remote Sensing*, 31, 4443-4458, 10.1080/01431160903277464, 2010.
- 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, 10.1016/j.rse.2017.03.003, 2017.
- 460 Levin, N., Kyba, C. C. M., Zhang, Q., Sánchez de Miguel, A., Román, M. O., Li, X., Portnov, B. A., Molthan, A. L., Jechow, A., Miller, S. D., Wang, Z., Shrestha, R. M., and Elvidge, C. D.: Remote sensing of night lights: A review and an outlook for the future, *Remote Sens Environ*, 237, 111443, 10.1016/j.rse.2019.111443, 2020.
- Li, X., Li, D., Xu, H., and Wu, C.: Intercalibration between DMSP/OLS and VIIRS night-time light images to evaluate city



- 465 light dynamics of Syria's major human settlement during Syrian Civil War, *International Journal of Remote Sensing*, 1-18, 10.1080/01431161.2017.1331476, 2017.
- Li, X., and Zhou, Y.: A Stepwise Calibration of Global DMSP/OLS Stable Nighttime Light Data (1992–2013), *Remote Sensing*, 9, 637, 10.3390/rs9060637, 2017.
- Li, X., Zhan, C., Tao, J., and Li, L.: Long-Term Monitoring of the Impacts of Disaster on Human Activity Using DMSP/OLS
- 470 Nighttime Light Data: A Case Study of the 2008 Wenchuan, China Earthquake, *Remote Sensing*, 10, 588, 10.3390/rs10040588, 2018.
- Li, X., Ma, R., Zhang, Q., Li, D., Liu, S., He, T., and Zhao, L.: Anisotropic characteristic of artificial light at night – Systematic investigation with VIIRS DNB multi-temporal observations, *Remote Sens Environ*, 233, 111357, 10.1016/j.rse.2019.111357, 2019.
- 475 Li, X., Zhou, Y., Zhao, M., and Zhao, X.: A harmonized global nighttime light dataset 1992–2018, *Scientific Data*, 7, 168, 10.1038/s41597-020-0510-y, 2020.
- Liu, H., Wang, L., Sherman, D., Gao, Y., and Wu, Q.: An object-based conceptual framework and computational method for representing and analyzing coastal morphological changes, *International Journal of Geographical Information Science*, 24, 1015-1041, 10.1080/13658810903270569, 2010.
- 480 Liu, X., Ou, J., Wang, S., Li, X., Yan, Y., Jiao, L., and Liu, Y.: Estimating spatiotemporal variations of city-level energy-related CO₂ emissions: An improved disaggregating model based on vegetation adjusted nighttime light data, *Journal of Cleaner Production*, 177, 101-114, 10.1016/j.jclepro.2017.12.197, 2018.
- Lo, C.: Urban Indicators of China from Radiance-Calibrated Digital DMSP-OLS Nighttime Images, *Annals of the Association of American Geographers*, 92, 225-240, 10.1111/1467-8306.00288, 2002.
- 485 Lu, H., Zhang, C., Liu, G., Ye, X., and Miao, C.: Mapping China's Ghost Cities through the Combination of Nighttime Satellite Data and Daytime Satellite Data, *Remote Sensing*, 10, 1037, 10.3390/rs10071037, 2018.
- Ma, T., Zhou, C., Pei, T., Haynie, S., and Fan, J.: Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities, *Remote Sens Environ*, 124, 99-107, 10.1016/j.rse.2012.04.018, 2012.
- 490 Ma, T., Zhou, C., Pei, T., Haynie, S., and Fan, J.: Responses of Suomi- NPP VIIRS- derived nighttime lights to socioeconomic activity in China's cities, *Remote Sensing Letters*, 5, 165-174, 10.1080/2150704x.2014.890758, 2014.
- Ou, J., Liu, X., Li, X., and Chen, Y.: Quantifying the relationship between urban forms and carbon emissions using panel data analysis, *Landscape Ecology*, 28, 1889-1907, 10.1007/s10980-013-9943-4, 2013.
- Román, M. O., and Stokes, E. C.: Holidays in lights: Tracking cultural patterns in demand for energy services, *Earth's Future*, 3, 182-205, 10.1002/2014EF000285, 2015.
- 495 Román, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., Schultz, L., Bell, J., Stokes, E. C., Pandey, B., Seto, K. C., Hall, D., Oda, T., Wolfe, R. E., Lin, G., Golpayegani, N., Devadiga, S., Davidson, C., Sarkar, S., Praderas, C., Schmaltz, J., Boller, R., Stevens, J., Ramos González, O. M., Padilla, E., Alonso, J., Detrés, Y., Armstrong, R., Miranda, I., Conte, Y., Marrero, N., MacManus, K., Esch, T., and Masuoka, E. J.: NASA's Black Marble nighttime lights product suite, *Remote Sens Environ*, 210, 113-143, 10.1016/j.rse.2018.03.017, 2018.
- 500 Shao, X., Cao, C., Zhang, B., Qiu, S., Elvidge, C., and Von Hendy, M.: Radiometric calibration of DMSP-OLS sensor using VIIRS day/night band, *Earth Observing Missions and Sensors: Development, Implementation, and Characterization III*, 2014, 92640A.
- Shi, K., Huang, C., Yu, B., Yin, B., Huang, Y., and Wu, J.: Evaluation of NPP-VIIRS night-time light composite data for
- 505 extracting built-up urban areas, *Remote Sensing Letters*, 358-366, 10.1080/2150704X.2014.905728, 2014a.
- Shi, K., Yu, B., Huang, Y., Hu, Y., Yin, B., Chen, Z., Chen, L., and Wu, J.: Evaluating the Ability of NPP-VIIRS Nighttime Light Data to Estimate the Gross Domestic Product and the Electric Power Consumption of China at Multiple Scales: A Comparison with DMSP-OLS Data, *Remote Sensing*, 6, 1705-1724, 10.3390/rs6021705, 2014b.
- Shi, K., Chen, Y., Yu, B., Xu, T., Chen, Z., Liu, R., Li, L., and Wu, J.: Modeling spatiotemporal CO₂ (carbon dioxide) emission
- 510 dynamics in China from DMSP-OLS nighttime stable light data using panel data analysis, *Applied Energy*, 168, 523-533, 10.1016/j.apenergy.2015.11.055, 2016a.



- Shi, K., Chen, Y., Yu, B., Xu, T., Yang, C., Li, L., Huang, C., Chen, Z., Liu, R., and Wu, J.: Detecting spatiotemporal dynamics of global electric power consumption using DMSP-OLS nighttime stable light data, *Applied Energy*, 184, 450-463, 10.1016/j.apenergy.2016.10.032, 2016b.
- 515 Shi, K., Yu, B., Huang, C., Wu, J., and Sun, X.: Exploring spatiotemporal patterns of electric power consumption in countries along the Belt and Road, *Energy*, 150, 847-859, 10.1016/j.energy.2018.03.020, 2018.
- Sutton, P., Roberts, D., Elvidge, C., and Baugh, K.: Census from Heaven: an estimate of the global human population using night-time satellite imagery, *International Journal of Remote Sensing*, 22, 3061-3076, 10.1080/01431160010007015, 2001.
- 520 Tan, C. C., and Eswaran, C.: Reconstruction of handwritten digit images using autoencoder neural networks, 2008 Canadian Conference on Electrical and Computer Engineering, 2008, 000465-000470,
- Taylor, P. J., Ni, P., Derudder, B., Hoyler, M., Huang, J., Lu, F., Pain, K., Witlox, F., Yang, X., and Bassens, D.: Measuring the world city network: new results and developments, *ICTs for Mobile and Ubiquitous Urban Infrastructures: Surveillance, Locative Media and Global Networks: Surveillance, Locative Media and Global Networks*, 15, 10.4018/978-1-60960-051-8.ch002, 2010.
- 525 Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., and Manzagol, P.-A. J. J. o. m. l. r.: Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion, *Journal of Machine Learning Research*, 11, 3371-3408, 10.1016/j.mechatronics.2010.09.004, 2010.
- Waluda, C. M., Griffiths, H. J., and Rodhouse, P. G.: Remotely sensed spatial dynamics of the *Illex argentinus* fishery, Southwest Atlantic, *Fisheries Research*, 91, 196-202, 10.1016/j.fishres.2007.11.027, 2008.
- 530 Wang, R., and Tao, D.: Non-local auto-encoder with collaborative stabilization for image restoration, *IEEE Transactions on Image Processing*, 25, 2117-2129, 2016.
- Wen, Y., and Wu, J.: Withstanding the Great Recession Like China, *The Manchester School*, 87, 138-182, 10.1111/manc.12223, 2019.
- 535 Wu, B., Yu, B., Yao, S., Wu, Q., Chen, Z., and Wu, J.: A surface network based method for studying urban hierarchies by night time light remote sensing data, *International Journal of Geographical Information Science*, 33, 1377-1398, 10.1080/13658816.2019.1585540, 2019.
- Xu, H., Yang, H., Li, X., Jin, H., and Li, D.: Multi-Scale Measurement of Regional Inequality in Mainland China during 2005-2010 Using DMSP/OLS Night Light Imagery and Population Density Grid Data, *Sustainability*, 7, 13469, 10.3390/SU71013469, 2015.
- 540 Yang, C., Yu, B., Chen, Z., Song, W., Zhou, Y., Li, X., and Wu, J.: A Spatial-Socioeconomic Urban Development Status Curve from NPP-VIIRS Nighttime Light Data, *Remote Sensing*, 11, 2398, 10.3390/rs11202398, 2019.
- Yu, B., Shu, S., Liu, H., Song, W., Wu, J., Wang, L., and Chen, Z.: Object-based spatial cluster analysis of urban landscape pattern using nighttime light satellite images: a case study of China, *International Journal of Geographical Information Science*, 28, 2328-2355, 10.1080/13658816.2014.922186, 2014.
- 545 Yu, B., Shi, K., Hu, Y., Huang, C., Chen, Z., and Wu, J.: Poverty Evaluation Using NPP-VIIRS Nighttime Light Composite Data at the County Level in China, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 1217 - 1229, 10.1109/JSTARS.2015.2399416, 2015.
- 550 Yu, B., Lian, T., Huang, Y., Yao, S., Ye, X., Chen, Z., Yang, C., and Wu, J.: Integration of nighttime light remote sensing images and taxi GPS tracking data for population surface enhancement, *International Journal of Geographical Information Science*, 33, 687-706, 10.1080/13658816.2018.1555642, 2018.
- Zhao, M., Zhou, Y., Li, X., Zhou, C., Cheng, W., Li, M., and Huang, K.: Building a Series of Consistent Night-Time Light Data (1992-2018) in Southeast Asia by Integrating DMSP-OLS and NPP-VIIRS, *IEEE T Geosci Remote*, 1-14, 10.1109/TGRS.2019.2949797, 2019a.
- 555 Zhao, N., Liu, Y., Cao, G., Samson, E. L., and Zhang, J.: Forecasting China's GDP at the pixel level using nighttime lights time series and population images, *GIScience & Remote Sensing*, 1-19, 10.1080/15481603.2016.1276705, 2017.
- Zhao, X., Yu, B., Liu, Y., Chen, Z., Li, Q., Wang, C., and Wu, J.: Estimation of Poverty Using Random Forest Regression with Multi-Source Data: A Case Study in Bangladesh, *Remote Sensing*, 11, 375, 10.3390/rs11040375, 2019b.



- Zheng, Q., Weng, Q., and Wang, K.: Developing a new cross-sensor calibration model for DMSP-OLS and Suomi-NPP VIIRS night-light imageries, *ISPRS Journal of Photogrammetry and Remote Sensing*, 153, 36-47, 10.1016/j.isprsjprs.2019.04.019, 2019.
- Zhou, Y., Smith, S. J., Elvidge, C. D., Zhao, K., Thomson, A., and Imhoff, M.: A cluster-based method to map urban area from DMSP/OLS nightlights, *Remote Sens Environ*, 147, 173-185, 2014.
- Zhou, Y., Smith, S. J., Zhao, K., Imhoff, M., Thomson, A., Bond-Lamberty, B., Asrar, G. R., Zhang, X., He, C., and Elvidge, C. D.: A global map of urban extent from nightlights, *Environmental Research Letters*, 10, 054011, 10.1088/1748-9326/10/5/054011, 2015.
- Zhu, X., Yang, H., Ge, W., and Ma, M.: Modeling the Spatiotemporal Dynamics of Gross Domestic Product in China Using Extended Temporal Coverage Nighttime Light Data, *Remote Sensing*, 9, 626, 10.3390/rs9060626, 2017.
- Zhuo, L., Zheng, J., Zhang, X., Li, J., and Liu, L.: An improved method of night-time light saturation reduction based on EVI, *International Journal of Remote Sensing*, 36, 4114-4130, 10.1080/01431161.2015.1073861, 2015.



Table 1: The list of used data in this study

Dataset	Source	Role
EANTLI	Calibrated DMSP-OLS NTL data EVI Data	Input Data in AE model (2000-2013)
Composited NPP-VIIRS NTL Data	Monthly NPP-VIIRS NTL data	Reference Data for Validation (2012) Label Data in AE model (2013) part of NPP-VIIRS-like NTL data (2013-2018)
DMSP-OLS RNTL data	F12-F15_20000103-20001229_rad_v4 F14_20040118-20041216_rad_v4 F16_20051128-20061224_rad_v4 F16_20100111-20101209_rad_v4	Reference Data for Validation (2000, 2004, 2006, 2010)

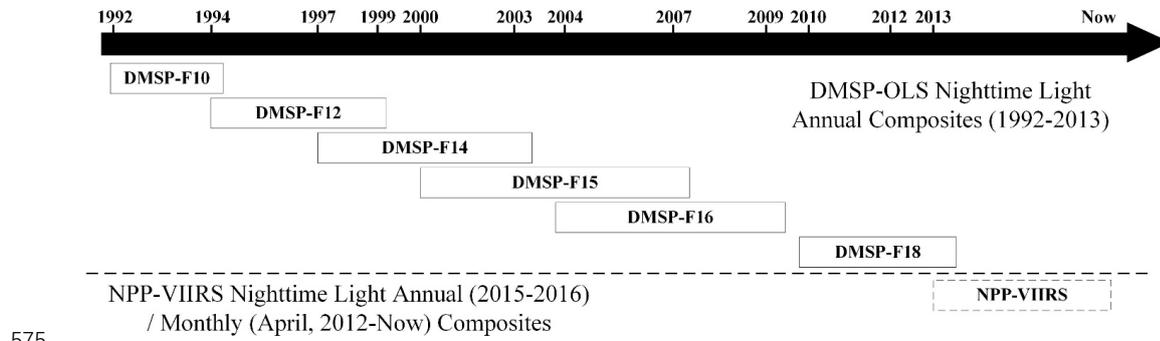


Figure 1: NTL data from DMSP-OLS and NPP-VIIRS

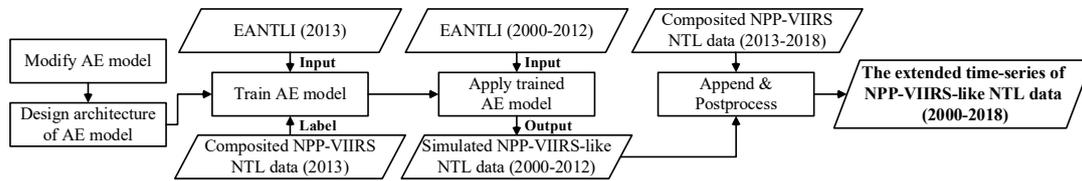


Figure 2: Flow chart of the generation of NPP-VIIRS-like NTL data.

580

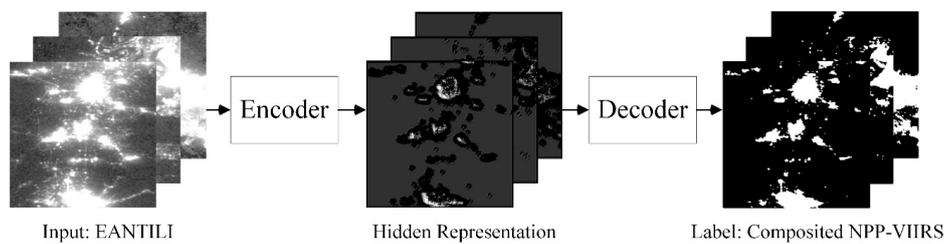
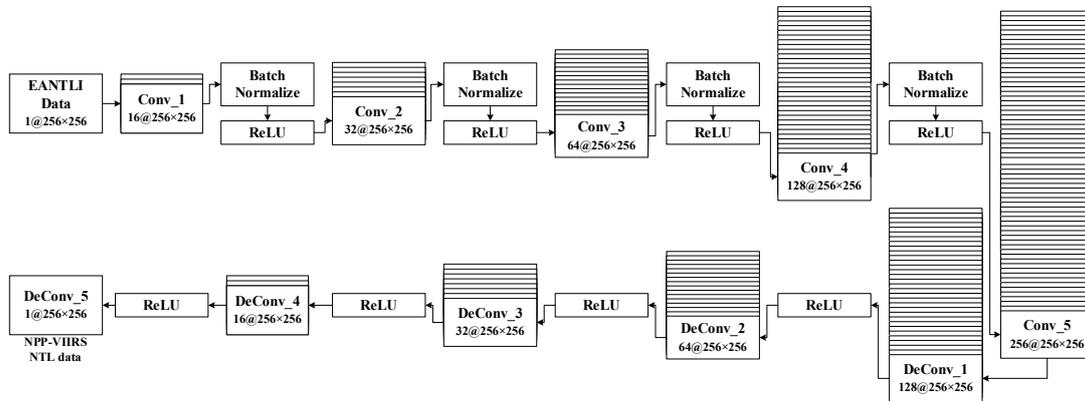


Figure 3: The framework of the auto-encoder model



585 Figure 4: The overall architecture of our proposed auto-encoder with CNN

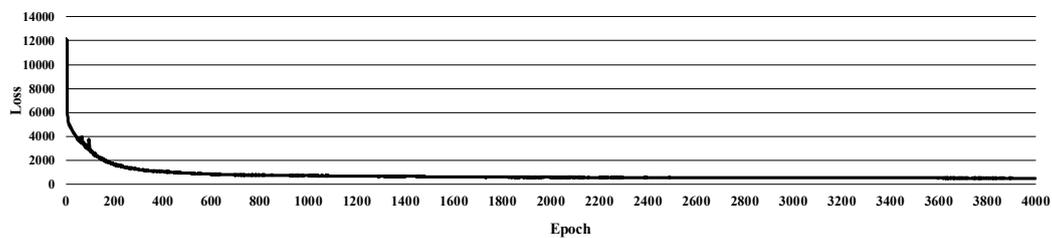
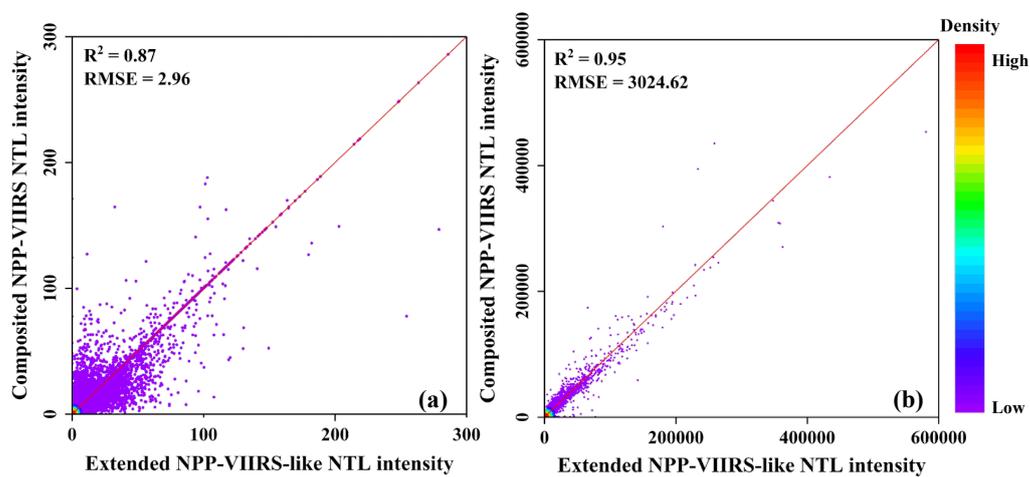
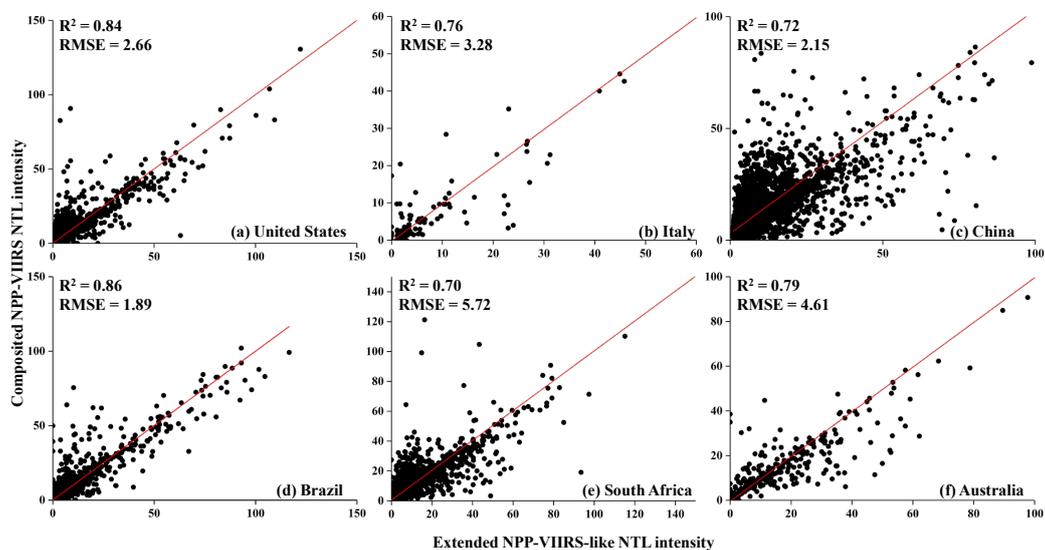


Figure 5: The training of each epoch's loss value

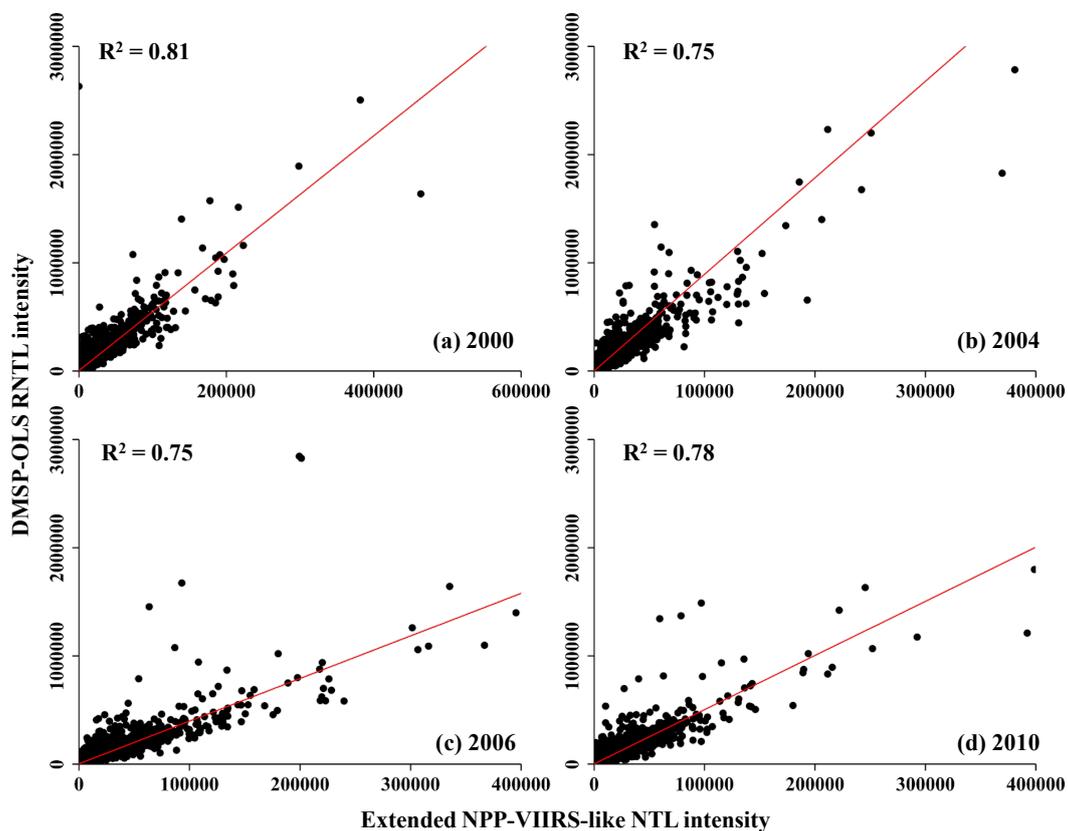


590

Figure 6: The comparison between the global composited NPP-VIIRS NTL data and extended NPP-VIIRS-like NTL data in 2012 (unit: $\text{nano-Wcm}^2\text{sr}^{-1}$): (a) at the pixel and (b) city levels. Solid line denotes the 1:1 line.



595 **Figure 7:** The comparison between composed NPP-VIIRS NTL data and extended NPP-VIIRS-like NTL data in 2012 (unit: nano- $\text{Wcm}^{-2}\text{sr}^{-1}$) at the pixel level in six countries.



600 Figure 8: The comparison between the annual DMSP-OLS RNTL intensity (DN Value) and extended NPP-VIIRS-like NTL intensity (unit: $\text{nano-Wcm}^{-2}\text{sr}^{-1}$) at the city level in (a) 2000, (b) 2004, (c) 2006, and (d) 2010.

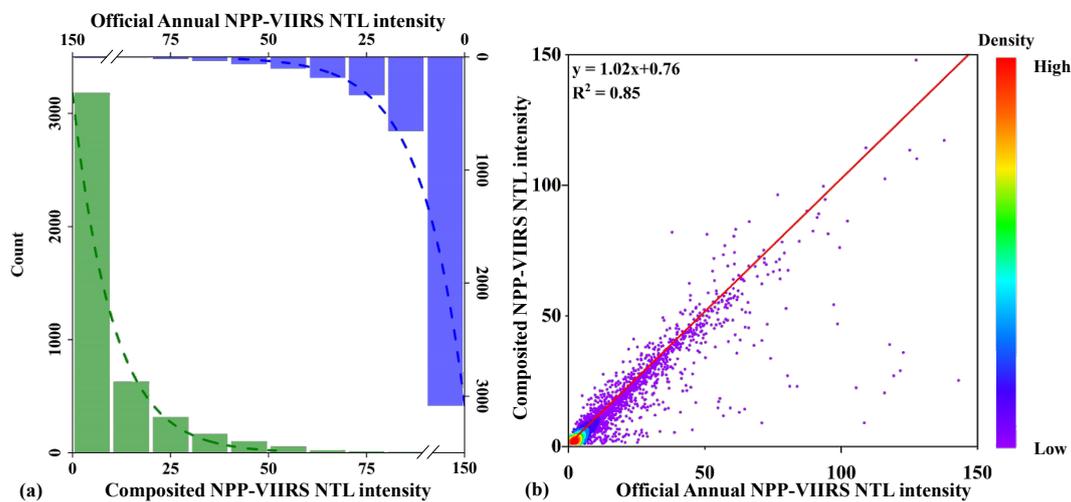
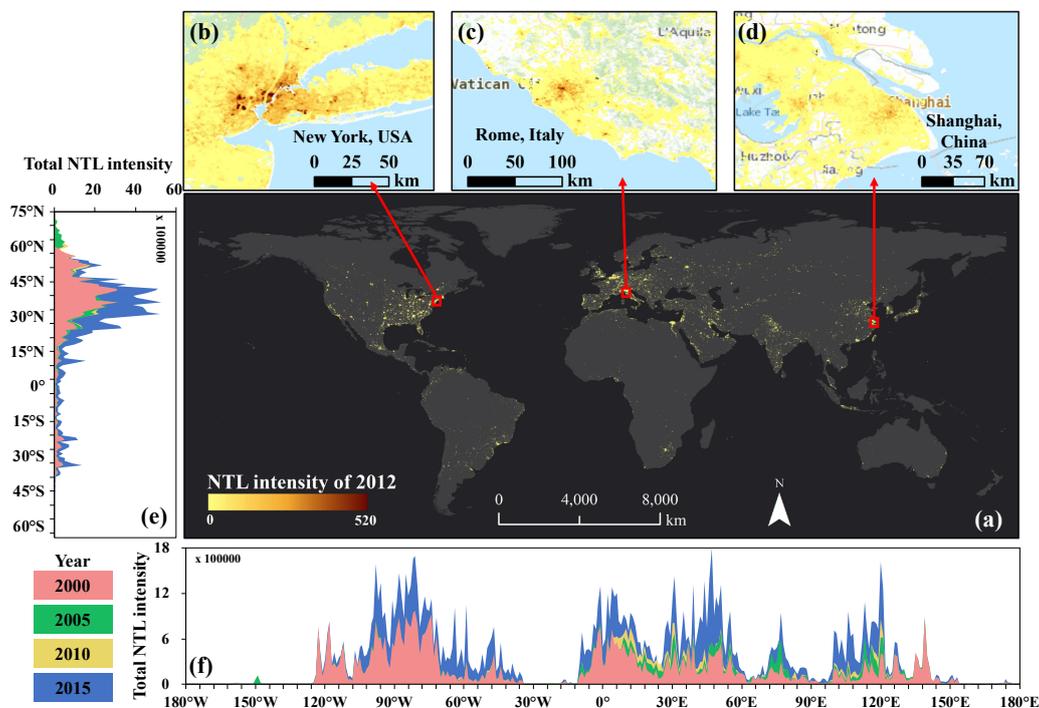
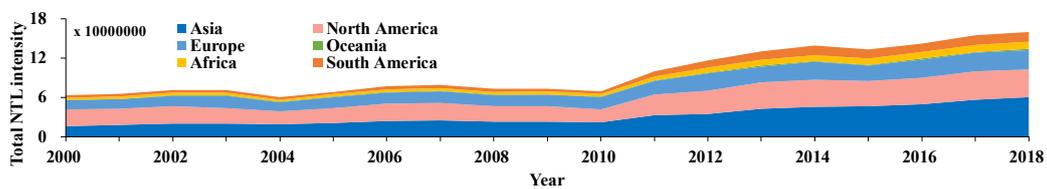


Figure 9: A comparison of composited and official annual NPP-VIIRS NTL data via (a) histograms and (b) scatter density plot.



605

Figure 10: The extended NPP-VIIRS-like NTL data (unit: $\text{nano-Wcm}^{-2}\text{sr}^{-1}$) with three enlarged subplots of New York, Rome, and Shanghai in 2012 and the dynamics of NTL intensity from 2000 to 2015 by longitude and latitude (1-degree bins).



610 Figure 11: The dynamics of total NTL intensity from 2000 to 2018 in each continent.

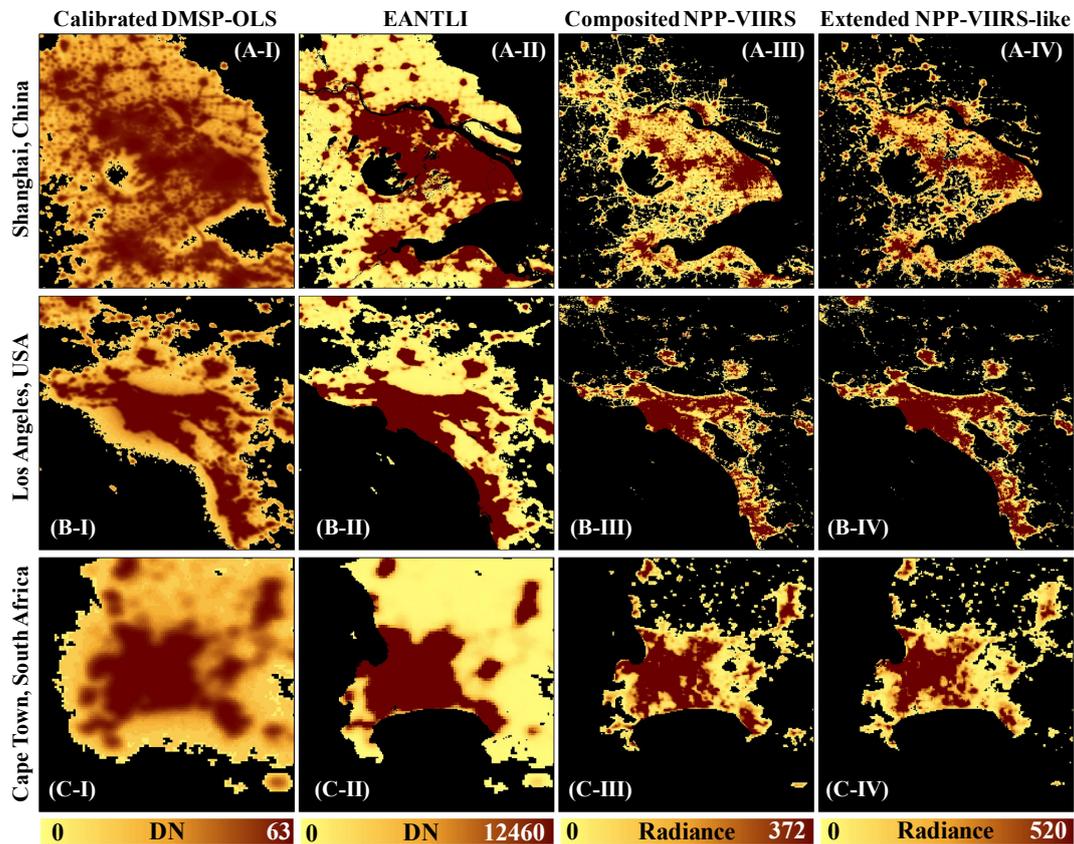


Figure 12: Spatial patterns of NTL intensity in 2012 of (I) calibrated DMSP-OLS data, (II) EANTLI, (III) composited NPP-VIIRS NTL data, and (IV) our extended NPP-VIIRS-like NTL data in three cities (A) Shanghai, China, (B) Los Angeles, USA, and (C) Cape Town, South Africa.

615

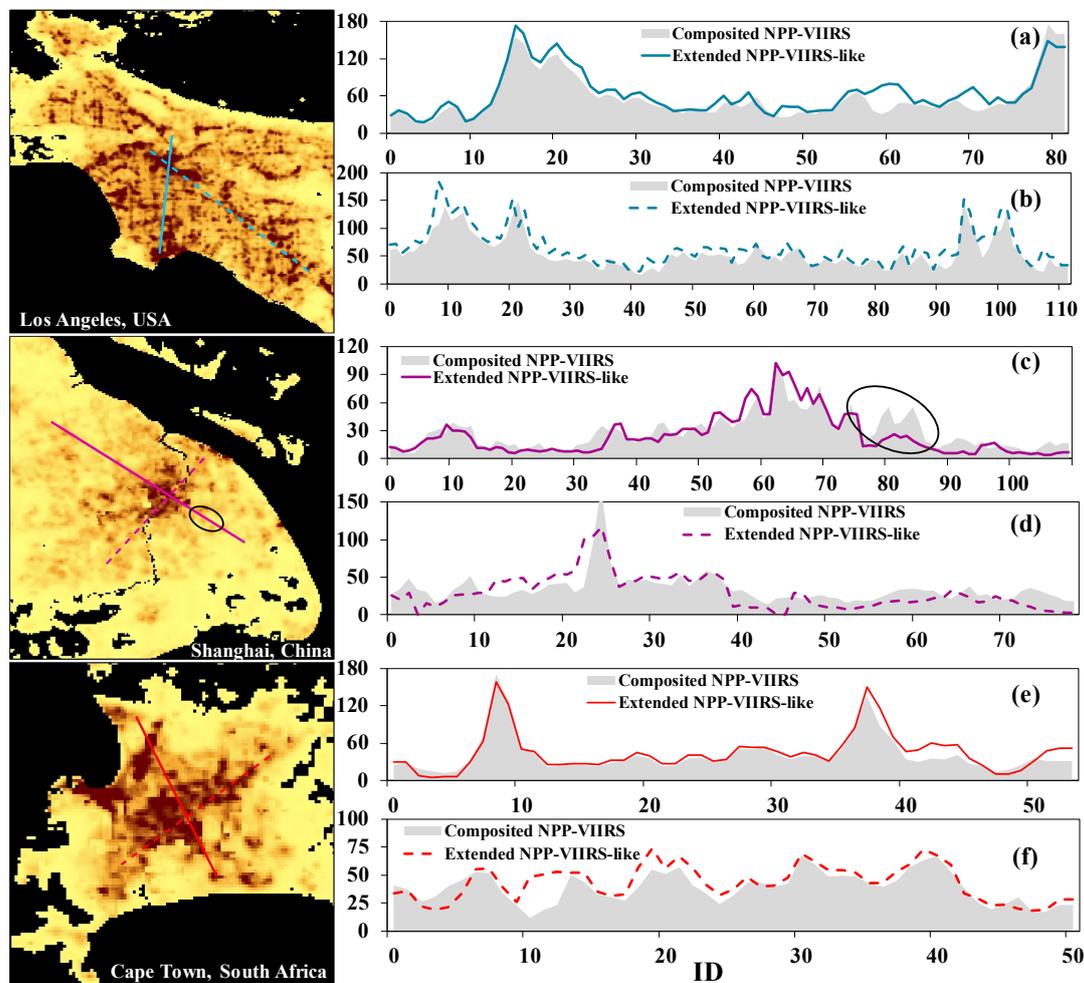


Figure 13: Profiles of composited NPP-VIIRS NTL data and extended NPP-VIIRS-like NTL intensity (unit: $\text{nano-Wcm}^{-2}\text{sr}^{-1}$) across (a and b) Los Angeles, USA, (c and d) Shanghai, China, and (e and f) Cape Town, South Africa.

620

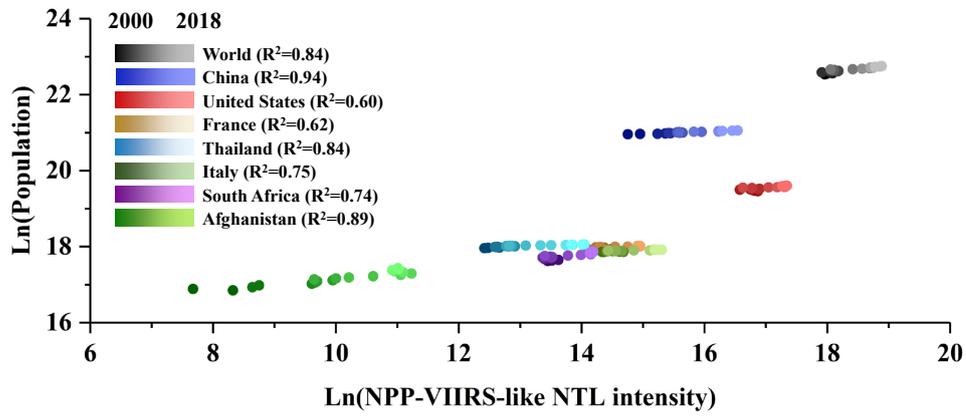
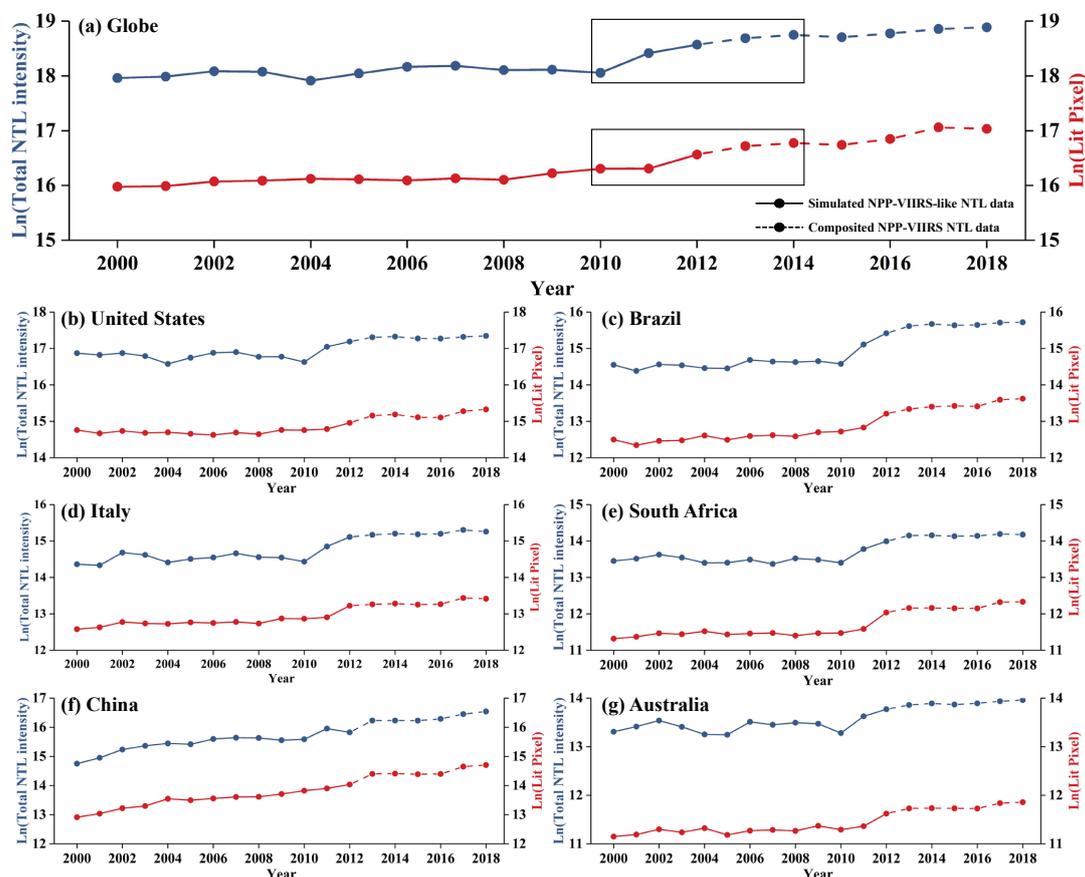


Figure 14: Comparison of log-transformed urban population with log-transformed total NTL intensity.



625 **Figure 15:** Temporal trend of extended time-series (2000-2018) NPP-VIIRS-like NTL data for the log-transformed total NTL intensity and lit pixels at (a) the global scale and within six selected countries: (b) United States, (c) Brazil, (d) Italy, (e) South Africa, (f) China, and (g) Australia.