Estimation of GDP Using Deep Learning With NPP-VIIRS Imagery and Land Cover Data at the County Level in CONUS

Jie Sun, Liping Di, *Member, IEEE*, Ziheng Sun, Jieyong Wang, and Yingdan Wu

***Abstract*—Accurate estimation of gross domestic product (GDP) at small geographies is of great significance to evaluate the distribution and dynamics of socio-economic development. Nighttime light (NTL) data is becoming increasingly important in socio-economic data estimation. However, previous research has found that using NTL alone is insufficient to accurately measure the GDP at small geographies, and the contribution of NTL for time-series GDP estimation is unreliable. This article proposed a deep learning method for the Contiguous United States (CONUS) time-series (2012–2015) GDP estimation at county level. The model is developed by combining the NTL data from the visible infrared imaging radiometer suite day/night band and the MODIS land cover data. The proposed method can improve the existing methods mainly in two ways. First, by taking advantage of the great computing power of the Google Earth Engine, a histogram-based feature regulation method was employed, which not only keeps more information over regions but also provides dimension-reduced tensors from mass remote sensing data. Second, a multi-inputs convolutional neural network-based model was proposed instead of the traditional linear regression model for multisource feature exploration and learning. The proposed method was evaluated by leave-one-year-out cross-validation with the time-series (2012–2015) data. The results show that the R^2 between the actual and estimated GDP are 0.81, 0.83, 0.83, and 0.83 for years from 2012 to 2015, indicating a good predictive power of the proposed model. Given that the data employed are globally and publicly available, the proposed method would also be applicable in other countries or regions where socio-economic survey data is hard to obtain.**

***Index Terms*—CNN, GDP, Google Earth Engine, remote sensing, time-series.**

Manuscript received January 10, 2020; revised February 29, 2020; accepted March 20, 2020. Date of publication March 30, 2020; date of current version April 21, 2020. This work was supported in part by the Philosophy and Social Science Foundation of Hubei Province under Grant 19Q062, and in part by the China and Natural Science Foundation of Hubei Province, China under Grant 2018CFB540. *(Corresponding author: Jie Sun.)*

Jie Sun is with the School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China, and also with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA 22030 USA (e-mail: jiesun@cug.edu.cn).

Liping Di and Ziheng Sun are with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA 22030 USA (e-mail: ldi@gmu.edu; zsun@gmu.edu).

Jieyong Wang is with the Institute of Geographic Sciences and Resources Research, Chinese Academy Sciences, Beijing 110101, China (e-mail: wjy@igsnrr.ac.cn).

Yingdan Wu is with the School of Science, Hubei University of Technology, Wuhan 430074, China (e-mail: yd_wu2010@hbut.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2020.2983331

I. Introduction

GROSS domestic product (GDP), one of the most important socio-economic indicator, is a monetary measure of the market value of all the final goods and services produced in a period of time, often annually. Accurate estimation of GDP is of great significance to quantitatively measure the socio-economic activities and formulate the corresponding economy policies for a country or region. Comparing to traditional statistics/sampling-based methods, remote sensing has the advantages of objective measurement, regular updating, and complete spatial coverage. With these features, analysis of remote sensing data becomes one of the key methodologies in economics and other applied scientific research. However, the remote sensing method was not introduced to estimate GDP until Elvidge *et al.* [1] found a strong correlation between illuminated area, which was detected by the US Air Force Defense Meteorological Satellite Program (DMSP), Operational Line scan System (OLS), and GDP (both expressed in logarithms) for 21 countries. As a successor to DMSP-OLS, the visible and infrared image radiometer suite (VIIRS) on board the Suomi National Polar-Orbiting Partnership (Suomi NPP) was launched on October 28, 2011. The NPP-VIIRS contains a day/night band (DNB), which has a higher radiometric sensitivity than OLS DNB. The VIIRS-DNB is a panchromatic channel covering the wavelengths from 500 to 900 nm and is sensitive to visible and near-infrared light during daylight hours as well as the low-level radiation observed at night, mainly from nighttime light (NTL).

A number of studies have confirmed that NTL provides a useful proxy for socio-economic development and have great potential for measuring national and subnational GDP [2]–[7]. Forbes found that in the United States, total night lights (TNL) measured by remote sensing and GDP are more strongly correlated at metropolitan scale than at state scale [8]. Chen and Nordhaus found a similar result that high-resolution VIIRS lights provide a better prediction for metropolitan statistical areas (MSA) GDP than for state GDP, which suggests that lights may be more closely related to urban sectors than rural sectors due to the modifiable areal unit problem (MAUP) [9]. The predictive power of NTL at smaller geographies is also worthy of further research, which is useful for understanding the interactions between local progress or failures and those at bigger geographical scales [10], [11]. In China, Li *et al.* found

that TNL from NPP-VIIRS correlated with the county-level GDP at R^2 value of 0.8544. The finding means that TNL from NPP-VIIRS can predict county-level GDP more accurately than that from the DMSP-OLS [12]. Shi *et al.* [13] demonstrated a similar result but at the prefectural level. Besides, Zhao *et al.* also confirmed the strong correlation between the TNL from NPP-VIIRS and county-level GDP in the US. However, Bickenbach, *et al.* [14] found that the relationship between NTL growth and observed GDP growth varies significantly in both statistical and economic terms across regions. Nordhaus and Chen [15] found that for time-series analysis, the contribution of DMSP-OLS is either unreliable or very small for middle and low-income countries. Recently, they also found the VIIRS lights are more useful in predicting cross-sectional GDP than predicting time-series GDP. The most likely reason for the weak time-series relationship is that the variation in GDP growth rates is so much smaller than the cross-sectional GDP variation [9]. It appears that NTL data show a strong predictive ability for cross-sectional GDP at multiscales and the NPP-VIIRS data are shown to be more strongly correlated with economic output than DMSP-OLS data, but all their performance for time-series GDP prediction is unsatisfied.

Meanwhile, many studies also demonstrated that using NTL alone is insufficient to accurately measure the GDP at subnational levels. Limiting predictors to NTL poses a shortcoming because they cannot account for areas without observable NTL. In other words, it cannot be assumed that all types of economic activities emit more NTL as they grow, such as agriculture and forestry [16]. It is suggested that other data should also be taken into account to compensate the shortcoming of overdependence on NTL for GDP estimation. Keola *et al.* [16] presented a model for estimating agricultural and nonagricultural economic growth on national and subnational levels using land cover data and NTL. Yue *et al.* [17] provide a dasymetric approach for GDP estimation at a finer scale by combining DMSP-OLS NTL, enhanced vegetation index (EVI), and land cover data. Faisal *et al.* [18] conducted a regression analysis between GDP and built-up areas by both the normalized difference built-up index (NDBI) and normalized difference vegetation index (NDVI) extracted from Landsat imagery. In addition, weather data, topography data, and agricultural land use are all have been used for economic analysis [19]–[21].

Despite most of the previous work on economic parameters estimation has proven the validity of the linear regression model, machine learning, especially deep learning (DL), has emerged together with big data technologies and high-performance computing to create new opportunities to evaluate and understand economic environments [22]. Jean *et al.* [23] proposed a novel DL approach which proves the convolutional neural network (CNN) can be trained to identify Google image features that explain up to 75% of the variation in local-level economic outcomes in five African countries. By contrast, Perez *et al.* [24] trained CNN models on free and publicly available multispectral daytime satellite images of the African Continent from the Landsat 7 satellite, and the results showed that although Landsat 7 images have lower resolution than the zoom-level 16 Google Static Maps used by Jean *et al.*, better accuracy can be obtained by incorporating additional spectral bands beyond the visible bands. These researches indicate that DL has become a new method bridging the remote sensing data with economic parameters, which is also enlightening for GDP estimation. Nevertheless, to the best of our knowledge, there is few related research on GDP estimation concerning NTL data with DL methods. For which, a possible reason is that most of the previous researches always rely on hand-crafted features such as regional mean or sum of the NTL data, which is simple in calculation but will loss detailed information, hence was not suitable for DL method [1], [3], [6], [13], [25]. Besides, DL methods always extract features from raw images by CNN, which may cause huge computing load especially at large scales.

In order to fill the abovementioned research gaps, this study proposed a multi-inputs convolutional neural network (MICNN)-based method, which combines VIIRS-DNB NTL data and MODIS land cover (LC) data for time-series GDP estimation by county level in the Contiguous United States (CONUS). First, for computational efficiency, a Google Earth Engine (GEE)-based framework was built to take its powerful ability of big data management and analysis [26]–[30]. Based on the framework, a histogram-based method was employed for dimensionality reduction and feature extraction. Finally, a proposed MICNN model was trained by historical GDP data for estimation. The main aim of this study was to explore the potential of the proposed method for county-level time-series GDP estimation.

II. Methods

A. Study Area

In this study, 3109 counties of CONUS were selected. The economy of the United States is the largest in the world; previous studies have suggested that luminosity data can be a proxy for economic statistics in CONUS at different scales such as state or metropolitan. However, it is of great significance to reveal the performance of NTL data for GDP estimation at a smaller scale, county level. The CONUS consists of about 3109 counties covering different economic structures, such as agriculture, industry, and service. The large numbers of samples can ensure the deep learning networks learn the relationship thoroughly, and the diversity of economic structure can prevent the modeling from bias. Fig. 1 shows the selected counties in the GEE code editor. These counties vary greatly in areas, ranging from 40 to 52 000 square kilometers, with a median area of 2542 square kilometers.

B. Data

To reveal the relationship between the time-series county-level GDP and remote sensing data, MODIS LC data and VIIRS-DNB data were collected from 2012 to 2015 in the GEE. Furthermore, the US county-real GDP data (2012–2015) was chosen as the label data and the US county boundary data as auxiliary data.

1) VIIRS-DNB: VIIRS-DNB data was preferred as many studies concluded that VIIRS-DNB has better predictive

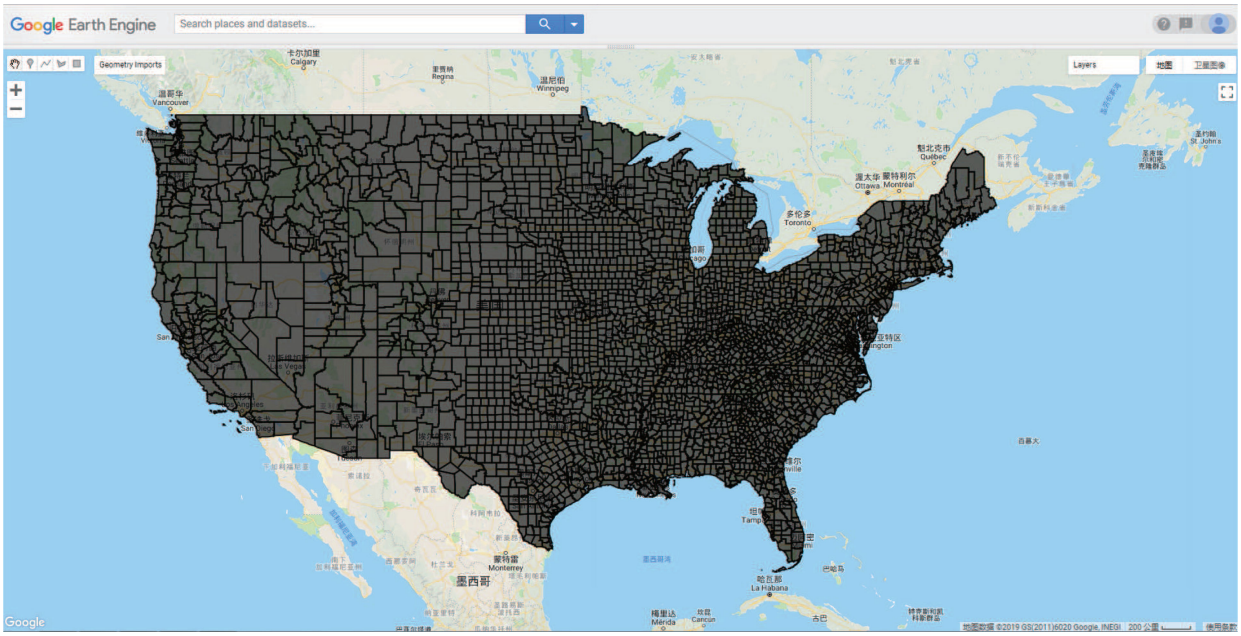


Fig. 1. Study area in the GEE editor (black areas show selected counties).

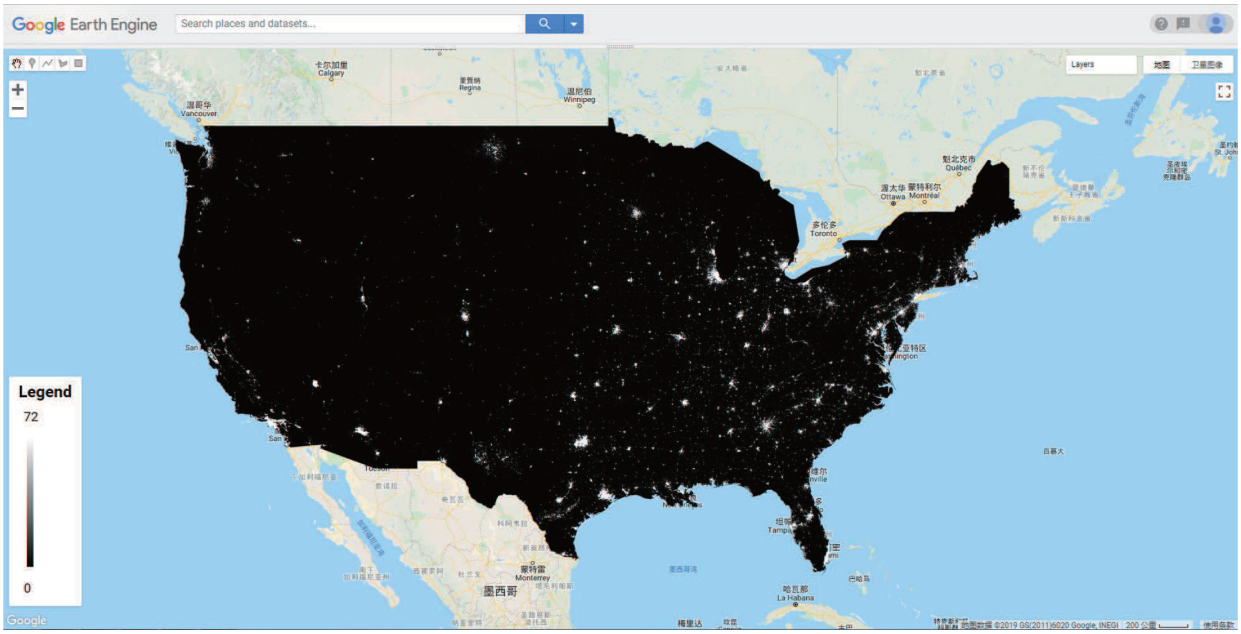


Fig. 2. The mean VIIRS-DNB radiance of CONUS in the GEE (2015).

power for estimating socioeconomic activities than DMSP-OLS data [31], [32]. The version 1 series of global VIIRS nighttime light images were provided by the Earth Observation Group, National Geophysical Data Center (NGDC) at the National Oceanic and Atmospheric Administration (NOAA) [33]. The VIIRS-DNB Cloud-Free composites data are produced monthly and contain spatially gridded nocturnal radiance values across human settlements at a spatial resolution of 15 arc-seconds. Prior to averaging, the DNB data is filtered to exclude the noises impacted by stray light, lightning, lunar illumination, and cloud-cover. Fig. 2 shows the mean radiance of VIIRS-DNB of CONUS in 2015 and Fig. 3 shows the four total distribution of mean VIIRS-DNB data in the study area from 2012 to 2015. The unit is nanoWatts/cm^2/sr and the range of radiance in the study area is from 0 to 72; besides, over 80% of radiance values is 0.

2) Land Cover: The Terra and Aqua combined MODIS LC Type (MCD12Q1) Version 6 data product provides global LC types at yearly intervals with 500-m resolution. The primary land cover scheme, named "LC_Type1," identifies 17 classes defined by the International Geosphere Biosphere Programme (IGBP),

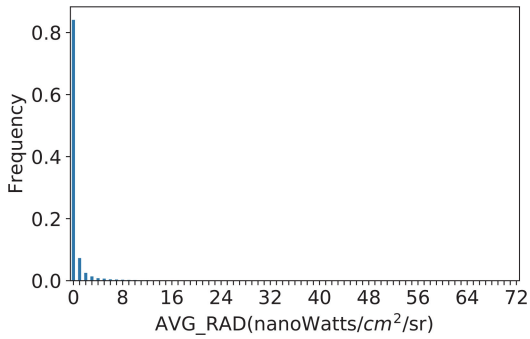


Fig. 3. The distribution of the mean VIIRS-DNB radiance in the study area in 2015.

which include 11 natural vegetation classes, three human-altered classes, and three nonvegetated classes shown in Table A1 of Appendix A. The "LC_Type1" data from 2012 to 2015 was used in the study. The MCD12Q1 Version 6 data product is derived using supervised classifications of MODIS Terra and Aqua reflectance data. The supervised classifications then undergo additional postprocessing that incorporate prior knowledge and ancillary information to further refine specific classes [34]. Fig. 4 shows the map of MCD12Q1 land cover in CONUS 2015. Fig. 5(a) shows the distribution of the land cover types in the study area from 2012 to 2015. The most common land cover types are grasslands, crop land, woody savannas, deciduous broadleaf forests, and open shrub lands, while the class of the urban and built-up lands, which are considered most correlated with the NTL, only account for tiny proportion.

3) US County GDP Data: The real GDP data by county (2012–2015) was released by the US Bureau of Economic Analysis (BEA) on December 12, 2018 [35]. The unit of GDP is thousands of dollars. It is the first time the BEA provides GDP statistics for each county in the United States. The prototype GDP data can address one of the last remaining gaps in economic knowledge, offering policymakers and businesses a new tool to inform their decision-making. Since it may be hard to build a stable relationship for extremely small GDP growth [9], [14], here GDP itself was preferred as an alternative. Moreover, the data was measured logarithmically to correct for heteroskedasticity. Fig. 5(b) shows the distribution of logarithmic GDP (LG) from BEA from 2012 to 2015. The data ranges from 3.44 to 8.82 with the mean of 5.97, and the p-value of 3.93e−130 indicates the data differs from a normal distribution [36]. The BEA LG maps of CONUS from 2012 to 2015 are shown in Fig. A1 of Appendix V. The darker color represents lower LG, and vice versa. It depicts that higher LG values concentrated mainly on the east-central and east of US, as well as the west coast, whereas the lower values are typically found in the west-central US, such as counties in Nebraska, Montana, Idaho, and Oregon. The Los Angeles County, CA and New York County, NY always produce the largest GDP from 2012 to 2015, while the Loving County, TX always has the smallest GDP.

4) US County Boundaries: The county boundaries were collected from TIGER/Line Shapefile, 2015, nation, US, current county, and equivalent national shapefile, and uploaded to the GEE in fusion table format [37]. Fusion tables is an experimental data visualization web application to gather, visualize, and share data tables.

C. Tensor Workflow in GEE

It is a key step to convert images into tensors with specific consistent shape for DL [38]. Most previous studies employed lit area, NTL density, or TNL for economic analysis, and the results suggested that economic parameters are strongly correlated with these features at a certain scale. An assumption can be made from those studies that the economic parameters only depend on the statistic of pixels in a certain grade. Hence, the pixel histogram of NTL over a region, which ignores the position information of individual pixels, can be equivalent to the original regional image for economic parameter estimation. The same goes for the pixels of LC types; the histogram of LC can reveal the economic structure of a region. Consequently, in this study, a histogram-based method was employed to calculate the pixel distribution of both VIIRS-DNB and LC. Unlike the single regional statistic, the histogram can provide high-dimensional features, which include more information. In the method, VIIRS-DNB data and LC data can be transformed int to fixed-bin histograms separately, by which the DL model can work directly with the entire pixel distribution and learn more detailed features. It is a computationally intensive task to calculate histograms for thousands of counties. Thanks to the great computing capacity of GEE, a GEE-based tensor workflow was designed which can export format-compliant tensors efficiently. Fig. 6 depicts the yearly tensor workflow in the GEE [39]. The procedures are described as follows.

1) First, the yearly data of MODIS LC and VIIRS-DNB from 2012 to 2015 are collected. The MODIS LC data and VIIRS-DNB can be accessed from the GEE dataset directly. The two data products have different cadences; the MODIS LC is annual data, while the VIIRS lights include monthly composites from April 2012 to the present, and the annual composites for 2015, but not for 2012, 2013, and 2014. To generate consistent measure of annual VIIRS lights, we use averaged monthly composites as the annual light [9].
2) US county boundary data was imported from fusion tables as a feature collection in the GEE, which is used in a reducer. Reducer is the way to aggregate data over time, space, bands, arrays, and other data structures in GEE. In our method, the reducer is employed for histogram statistics over individual counties.
3) Each county has two different types of data, VIIRS-DNB and MODIS LC. So a fixed-bin normalized histogram transformation was performed on the two types of yearly data in the GEE separately. A crucial parameter for the transformation is the number of bins. As for the VIIRS-DNB data, according to the actual distribution shown in Fig. 3, the radiance value ranges from 0 to 72. In order to eliminate the blooming phenomenon, the new limits are set to 1 and 72 [40], [41]. Besides, the date type of the NTL is float, more bins can represent the distribution

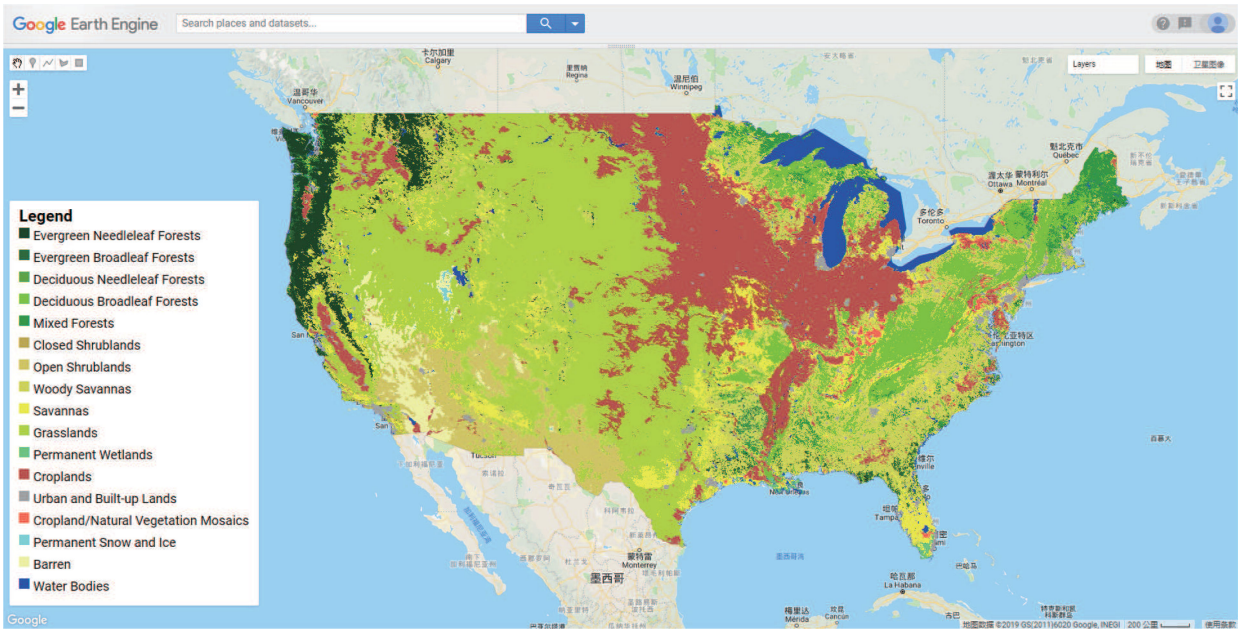
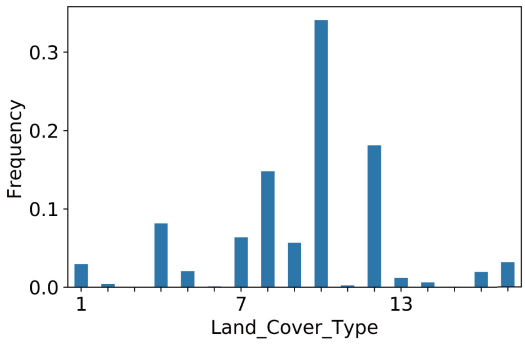
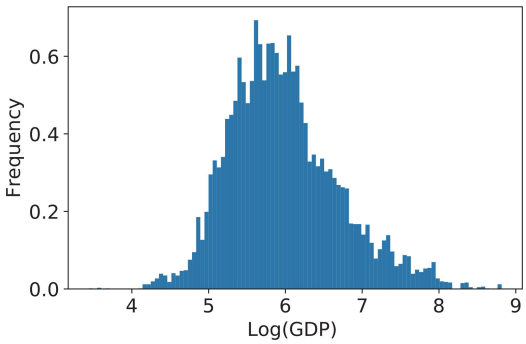


Fig. 4. The 2015 land cover of CONUS in the GEE.



(a)



(b)

Fig. 5. (a) The distribution of land cover type. (b) The distribution of log (GDP).

more accurately in a limited range, so the bin number was set to 288 at an interval of about 0.25. Then, each yearly VIIRS-DNB data can be transformed into county-level tensors with the shape of $1 \times 288 \times 1$ (height$\times$ width $\times$ band). It should be noted that the range of "LC_Type1" in MCD12Q1 is from 1 to 17 representing 17 different land covers, and as a result, the LC data has a bin number of 17 and the shape of the corresponding tensor is $1 \times 17 \times 1$ (height $\times$ width $\times$ band).

4) Before feeding to the training model, the tensor should be normalized, and the tensor of the yearly county *i* is defined as

$$H_i = [N_1, N_2, \ldots, N_{bn-1}, N_{bn}] \tag{1}$$

where i denotes the ID of a county, bn is the bin number, which is set to 288 for VIIRS and 17 for LC, and N represents the pixel quantity in the certain bin. Since different bins have different maximums crossing counties, to avoid too small value, the normalization was performed by each bin respectively. Specifically, the normalized tensor is defined in (2) and (3)

$$H_i^{\text{norm}} = [N_1^{\text{norm}}, N_2^{\text{norm}}, \ldots, N_{bn-1}^{\text{norm}}, N_{bn}^{\text{norm}}] \tag{2}$$

$$N_K^{\text{norm}} = \frac{N_k}{\text{Max}(N_k)}, 1 \le k \le bn \tag{3}$$

where N_K^{norm} is the normalized bin value, and $\text{Max}(N_k)$ represents the maximum of the kth bin in all the counties for the entire time-series.

5) The processed LG data will be assigned to each county as a label in terms of the date and US Federal Information Processing Standard (FIPS) code. If the GDP data of a county was not released in that year, the tensor of the county will be ignored. It is found that data for a few counties mainly in South Dakota and Virginia have not been released.

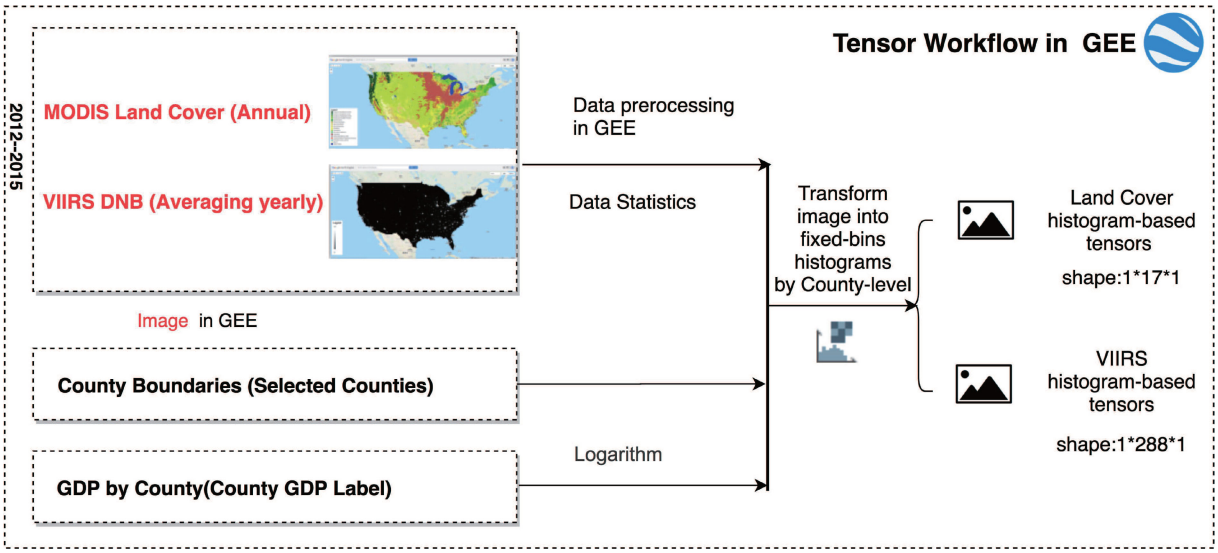


Fig. 6. The GEE-based feature extraction workflow.

Based on the workflow in Fig. 6, the yearly composite was transformed into fixed-bin normalized histograms at county level in GEE. Before the tensors are fed into the regression network, a difference confirmation could be made between tensors, which is thought to be helpful to understand the process. According to the GDP data, several representative counties, including Catron, NW; Los Angeles, CA; Harney, OR; and Dallas, TX were selected for comparison. The LG of the four counties is 4.82, 8.82, 5.35, and 8.39, respectively, in 2015. Fig. 7 depicts the normalized tensors for selected counties in 2015. There are similarities between Catron and Harney that the frequency of the DN value of VIIRS-DNB is extremely low, indicating the two counties have little VIIRS-DNB pixels with DN value over 1, and almost all of the DN value in these counties concentrates at the first few bins which contain low radiance pixels. Besides, most of the land cover type in these counties are type 9 or 10 representing savannas or grasslands. On the contrary, the VIIRS-DNB data of Los Angeles and Dallas both have a wider distribution and higher frequency, which means that there are plenty of pixels having relatively high radiance. Moreover, there is a diversity of LC in these counties, and the LC has the highest peaks at the type 13 (built-up lands) in both of the counties. Given a huge number of samples, it is expected that all the spatial distribution features can be learned by the CNN model for GDP prediction.

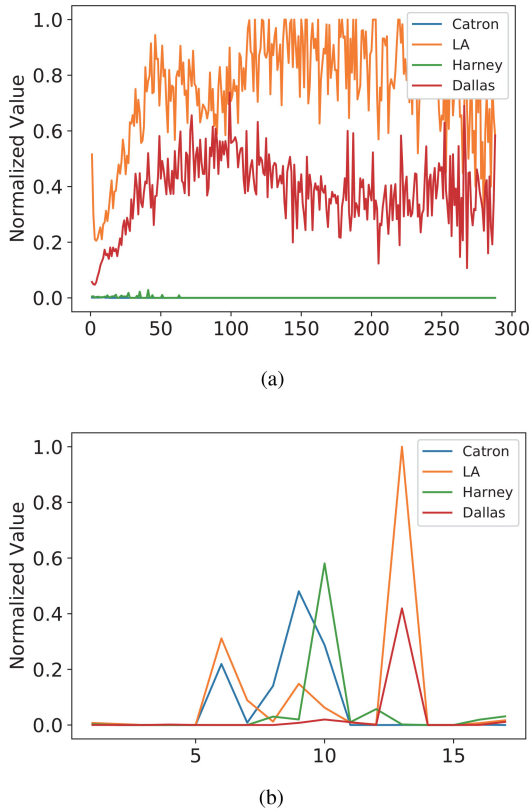


Fig. 7. The comparison of tensors in 2015 (Catron, NM; Los Angeles, CA; Harney, OR; and Dallas, TX). (a) The VIIRS tensor of different counties in 2015. (b) The LC tensor of different counties in 2015.

D. Model Architecture

It is well known that CNN has a powerful capacity in spatial feature extraction as well as excellent performance in regression and classification [42]–[44]. Instead of the traditional linear regression model, a CNN-based model was employed for GDP estimation in the study. Furthermore, the spatial features of the NTL and LC data are different and have their respective meaning; hence, a MICNN model was designed for the two types of data separately. The model, which was implemented by using an open-source neural-network library written in Python named Keras [45], mainly consists of three parts, the input layers, middle layers, and a regression layer. Fig. 8 shows the architecture of the model. First, multiple inputs including VIIRS tensor and LC tensor were fed into the middle layers, respectively. Next,

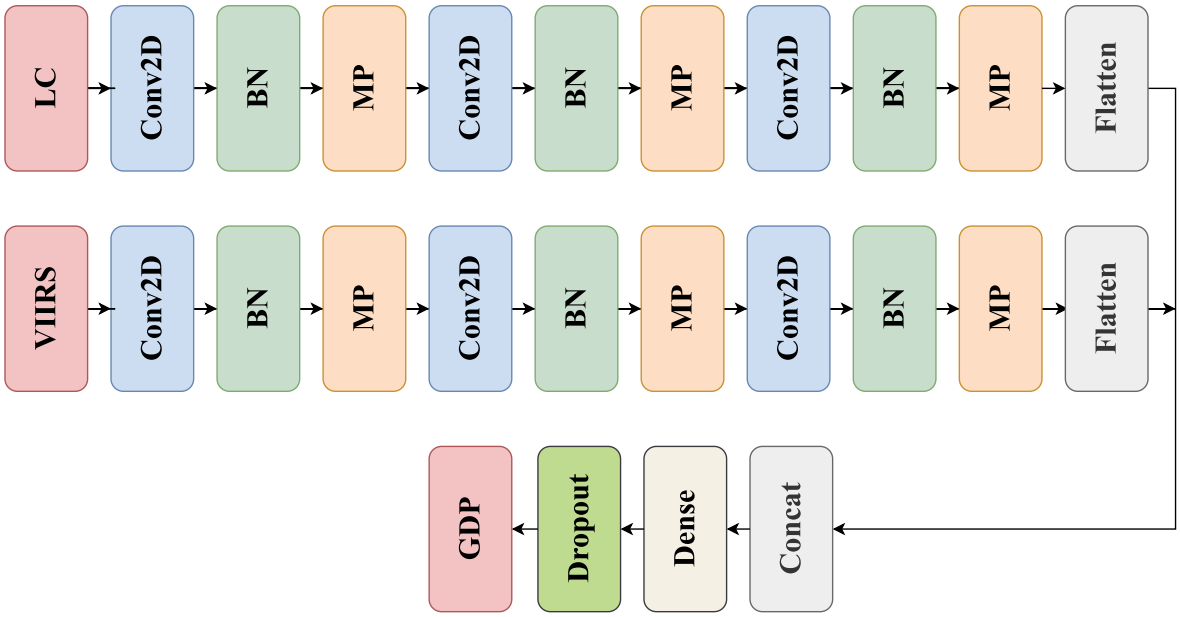


Fig. 8. The architecture of the proposed MICNN-based model. BN: Batch Normalization. MP: Max Pooling. Conv2D: 2-Dimensional Convolutional Neural Network.

each branch of the middle layers explores the features and output the feature map. Finally, the concatenated features were passed into the regression layer for GDP estimation.

The middle layers have extendable branches for multiple inputs. There are two same branches in our model, and each branch consists of a three cascade of 2-D CNN (Conv2D) layers, which is followed by the batch normalization (BN) layer and max-pooling (MP) layer. The role of the convolutional layer is to detect local conjunctions of features from the histograms of VIIRS and land cover data, and the role of the pooling layer is to merge semantically similar features into one. Because the relative positions of the features forming a motif can vary somewhat, reliably detecting the motif can be done by coarse-graining the position of each feature. The BN has two benefits: 1) it can regularize the network, thus reducing overfitting, and 2) it also helps to increase training speed, as gradients are rescaled at each layer during backpropagation, which results in better performance [46]. The MP layer could condense the feature map and preserves the most important features, which could greatly reduce the dimension of the feature maps and avoid overfitting. The feature maps from each branch were imported into a flattened layer, which could reshape the tensor to have the shape that is equal to the number of elements (not including the batch dimension) contained in the tensor. Then, the flattened vectors were concatenated as a single tensor including both VIIRS and LC features and sent into a dense layer, also called fully connected (FC) layer. The fully connected part of the CNN network goes through its own backpropagation process to determine the most accurate weights. Each neuron receives weights that prioritize the most appropriate label. Finally, the neurons vote on each of the labels, and the winner of that vote is the classification or regression decision. A dropout layer can randomly drop neurons out of the network during training, which makes the network less sensitive to the specific weights of neurons. The final layer has one output corresponding to the predicted GDP.

The model parameter details are as follows: the filter number of the three Conv2D layers are 256, 128, and 64 from top to bottom, and the kernel size is 1×3 for the Conv2D, and 1×2 for the MP layer. The unit number of Dense layer is 256, and the rate of dropout is 0.5. The activation function of the Conv2D layers, the dense layer, and the final regression layer are all the rectified linear units (ReLU): $\mathrm{ReLU}(x) = \max(0, x)$. The ReLU activation function is more expressive than the linear function. Its convergence rate is more efficient than that of nonlinear activation functions such as sigmoid and tanh. Besides, since the derivative of ReLU activation function is equal to 1, it can avoid the vanishing gradient problem [47]. Additionally, Adam is chosen as the optimizer [48].

E. Model Training and Evaluation

To avoid biased estimation, the evaluation was performed by using leave-one-year-out cross-validation [49]. A total of 4 years of data were used for developing the estimation models: 3 random years of data were employed as the training set, and the remaining years data as test set. A cross-validation accuracy assessment was conducted so that the above process was repeated for four times with the same model configuration for each case. To keep the impartiality of performance evaluation, 70% of the training set was randomly selected as training samples, and the remaining 30% as the validation samples for performance evaluation. Besides, an early stopping function was employed to stop the training when the "val_loss" quantity has

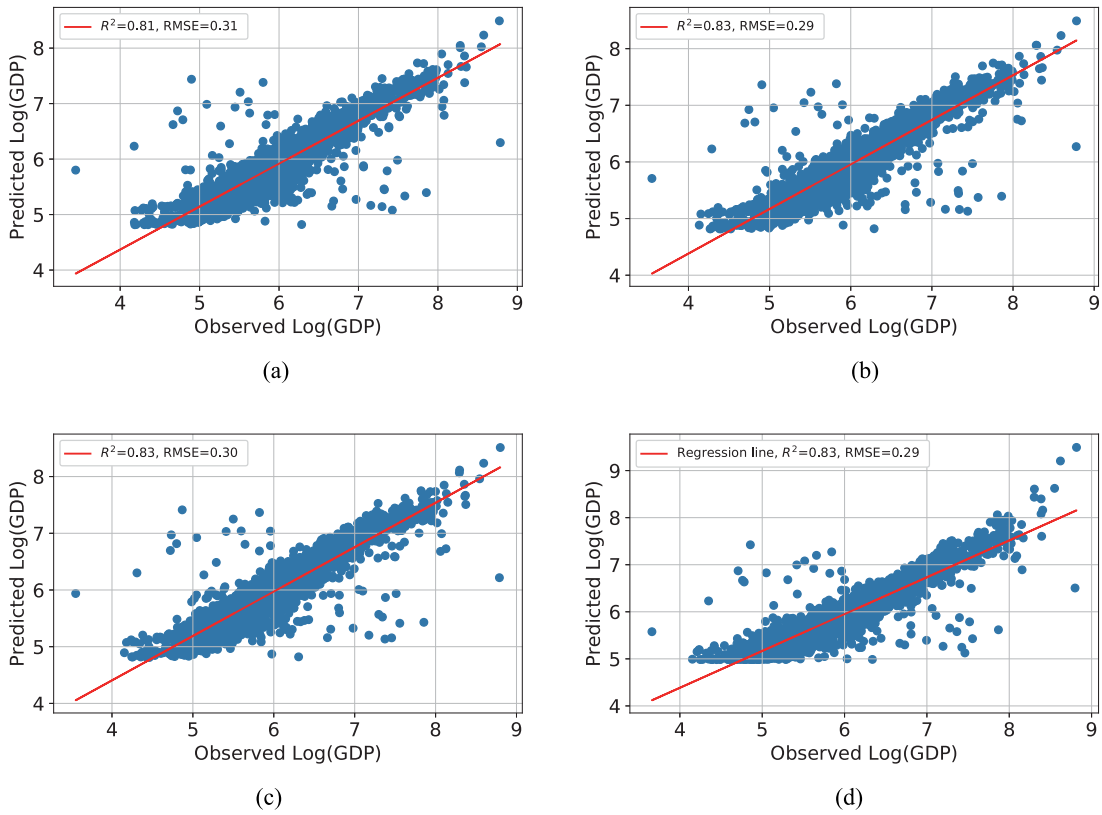


Fig. 9. Observed vs. predicted log(GDP) by the proposed method (2012–2015). (a) 2012. (b) 2013. (c) 2014. (d) 2015.

stopped improving after 20 epochs consecutively. The training batch size was set to 16 and the training epochs was 300. Several metrics were selected for the performance evaluation. Root-mean-squared error (RMSE) was used to measure the model's accuracy of GDP estimation. Percent error (PE) was used as another way to show the difference between predicted and observed data of each county. Formulas of RMSE and PE are presented in (4) and (5), where y_i is the predicted value, $\hat{y}_i$ is the observed value, and n is the number of samples. Besides, the R^2 between the observed and the predicted value was also used to evaluate how well the predicted values can reconstruct the spatial variations of observed GDP. Moreover, the most commonly used log–log linear regression method was used as a baseline to estimate the relationship between the regional TNL and county GDP (both are measured logarithmically to correct for heteroskedasticity) [10], [32], [50]. Similarly, the prediction in each year was also based on the model constructed by the data of other years. Additionally, to investigate the importance of LC data in the GDP estimation, we also compared our results based on the proposed model combining both NTL and LC data (MICNN-VIIRS&LC) with the results excluding LC data (MICNN-VIIRS).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{4}$$

$$PE = \frac{|y_i - \hat{y}_i|}{\hat{y}_i} \cdot 100\%. \tag{5}$$

TABLE I
THE PERFORMANCE COMPARISON OF DIFFERENT MODELS FROM 2012 TO 2015

Year	baseline		MICNN-VIIRS		MICNN-VIIRS&LC	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
2012	0.71	0.39	0.71	0.38	0.81	0.31
2013	0.71	0.38	0.76	0.35	0.83	0.29
2014	0.72	0.37	0.72	0.38	0.83	0.30
2015	0.73	0.37	0.73	0.37	0.83	0.29
Avg	0.72	0.38	0.73	0.37	0.83	0.30

III. RESULTS

A. Performance of the Proposed Method

Fig. 9 plots the observed vs. predicted LG by our method from 2012 to 2015; the RMSEs of the prediction are 0.31, 0.29, 0.30, and 0.29. Besides, the predicted LG can, respectively, explain 81%, 83%, 83%, and 83% of the variance in the observed LG of each year, which indicates a strong correlation between the actual and the estimated LG. The result shows a reasonable performance of the proposed method. Fig. 10 illustrates the map of predicted LG by our method from 2012 to 2015. The brighter color means the higher value, and vice versa. The LG map in each year presents a common pattern that the higher value always occur in Washington, California, Oregon, and New York states, which is similar to the result of BEA in Fig. A1 of Appendix A generally. However, the range of the estimated value (from 4 to

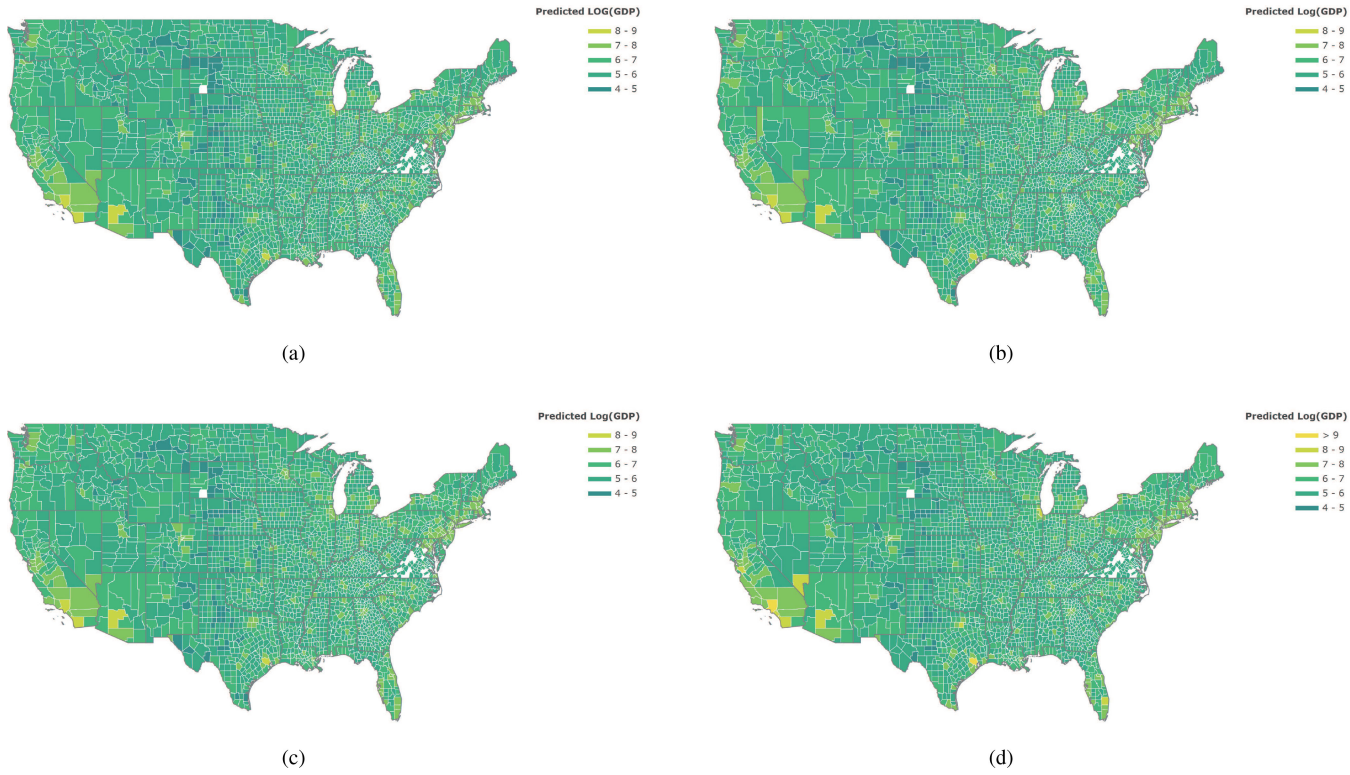


Fig. 10. The prediction log(GDP) map of CONUS at county level by MICNN-VIIRS and LC (2012–2015). (a) 2012. (b) 2013. (c) 2014. (d) 2015.

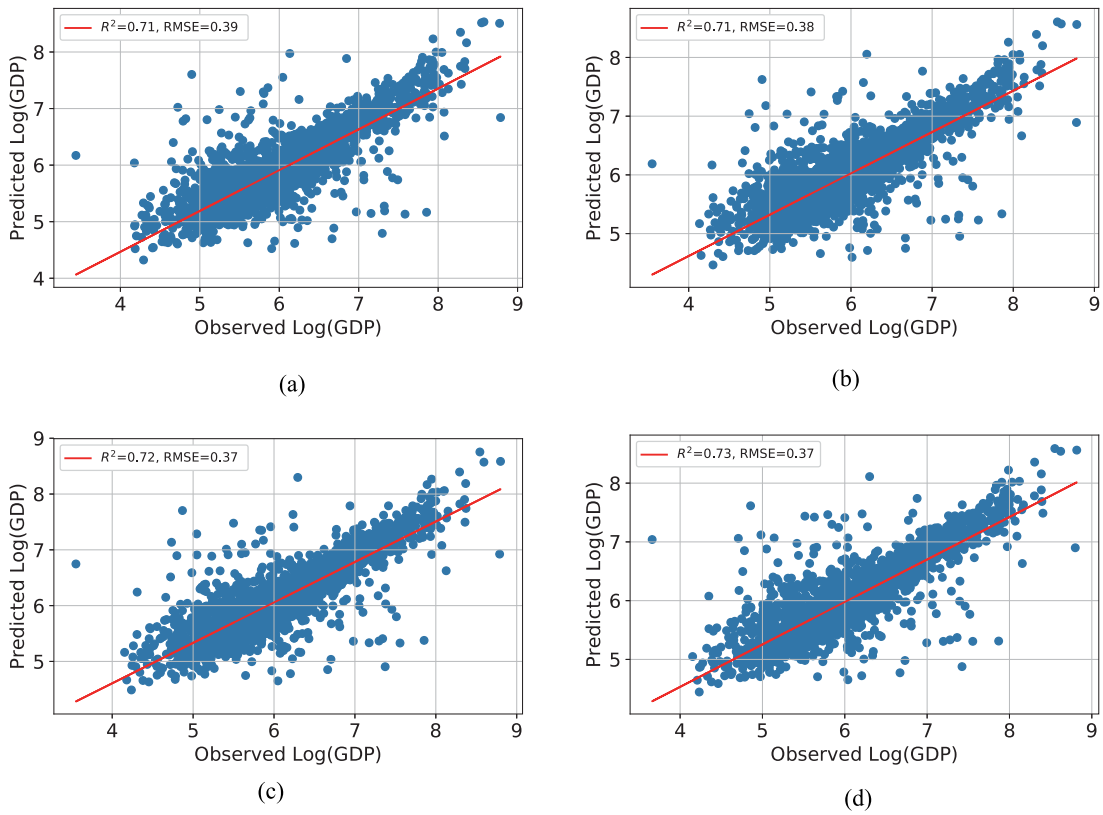


Fig. 11. Observed vs. predicted log(GDP) by baseline (2012–2015). (a) 2012. (b) 2013. (c) 2014. (d) 2015.

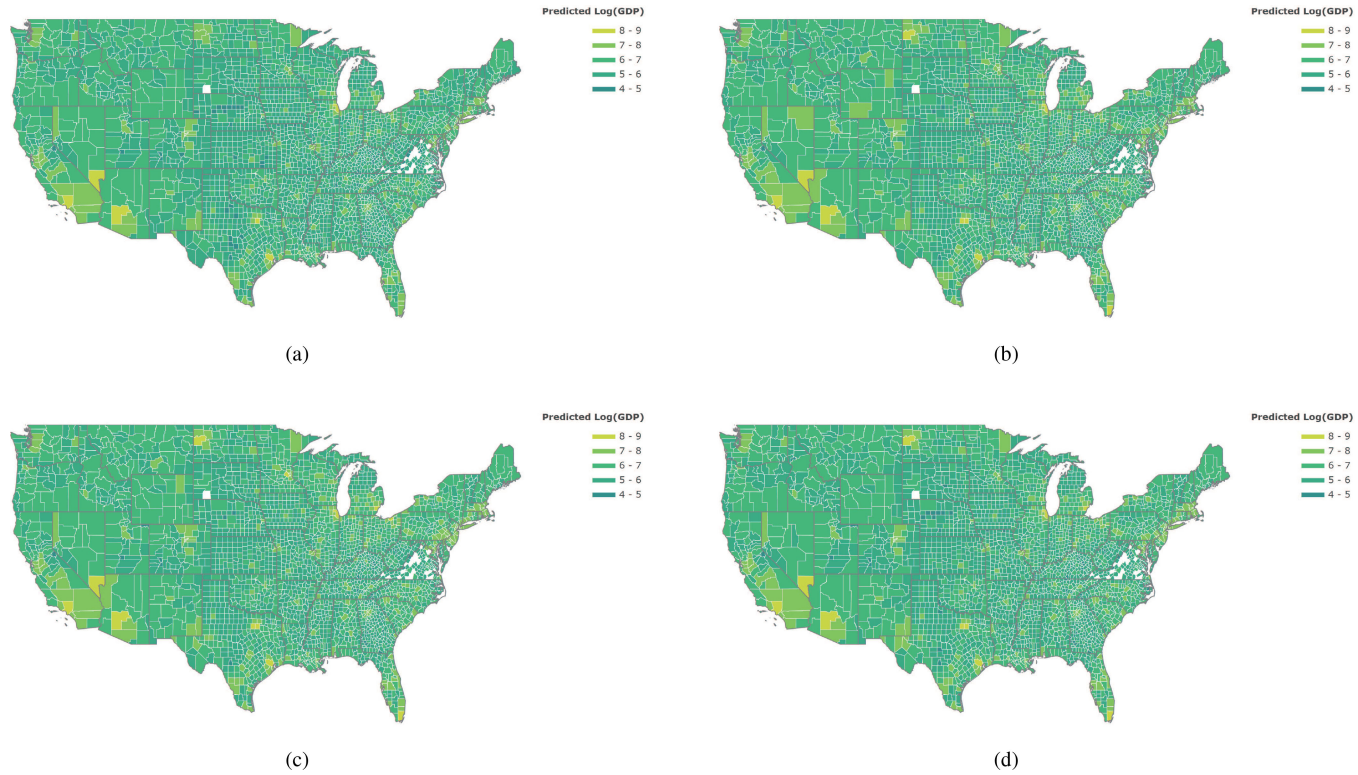


Fig. 12. The log(GDP) prediction map of CONUS at county level by baseline (2012–2015). (a) 2012. (b) 2013. (c) 2014. (d) 2015.

9 above) was narrower than that of the actual LG (from 3 to 9), which means that the LG of some counties was overrated.

B. Performance Comparison

To evaluate the performance of the proposed method, a traditional log(TNL)–log(GDP) linear regression method was employed as the baseline. Besides, the MICNN-VIIRS model was also included to verify the effect of LC data in the prediction. Table I shows the general performance comparison of the three models. The RMSEs of the baseline are 0.39, 0.38, 0.37, and 0.37 from 2012 to 2015, with the average of 0.38. The R^2 of the baseline are 0.71, 0.71, 0.72, and 0.73, with the average of 0.72 (Fig. 11). By contrast, the MICNN-VIIRS has a better performance; the RMSEs are 0.38, 0.35, 0.38, and 0.37 from 2012 to 2015, with the average of 0.37; the corresponding R^2 are 0.71, 0.76, 0.72, and 0.73, with the average of 0.73. The MICNN-VIIRS and LC has the best performance, the average R^2 is 0.83 and RMSE is 0.30, which is a significant improvement. The comparison suggests that the MICNN-VIIRS can improve the prediction at a certain degree for more features from histograms, and the combination of LC data can contribute to greater improvement.

Furthermore, for county-level analysis, Fig. 12 shows the baseline prediction map of LG in CONUS from 2012 to 2015. There is a small difference between the four years' results. However, compared with the observed results in Fig. A1 of Appendix V or Fig. 10, there is an obvious difference that many counties showing brighter colors are overrated. These counties are mainly distributed discretely in the central and west of CONUS.

To gain more insight, a more detailed comparison of PE map calculated by (5) is shown in Fig. 13, in which, the first column includes the PE maps of the baseline from 2012 to 2015, the second column includes the results of the MICNN-VIIRS, and the last column presents the results of the MICNN-VIIRS and LC. The value of PE was separated into seven grades shown in the PE legend; the higher the value, the brighter the color, vice versa. From the top to the bottom, the results of the linear model show an apparent regular pattern that the relatively higher PE mainly concentrate in the west-central and the northeast CONUS, especially in Montana, North Dakota, Oregon, Nevada, and Texas States. Combined with Fig. 2, it can be found that most of the counties with high PE always have either too low or too high TNL. The results prove that the traditional linear model using only NTL data has a poor performance in the regions with abnormal TNL. In comparison, the number of counties with relatively higher PE decreased significantly in the results of the MICNN-VIIRS and MICNN-VIIRS&LC. These relatively high PE values always happen in Nevada, California, Texas, Vermont, Wyoming, and the west coast.

A few analyses were made to reveal the influence of the enrichment of features and the join of LC data in the estimation. Fig. 14 shows the counties where the results are improved step by step. The column "a" of Fig. 14 shows the counties where the results are improved by using MICNN-VIIRS. It is obvious that most of the improved counties mainly concentrate in the west-central CONUS and the others are scattered all over the nation.

Fig. 13. The PE maps of different models.

A statistical mean was performed on these counties; the result in Fig. 15(a) shows that the improved counties always share relatively low VIIRS-DNB values ranging from 1 to 24. In other words, the feature enrichment would enhance the estimation for the counties with relatively low VIIRS-DNB. Column "b" of Fig. 14 shows the counties where the results are improved from MICNN-VIIRS to MICNN-VIIRS and LC. There is no clear pattern in the distribution from 2012 to 2015. To find out the common feature of these counties, a statistical mean of the LC tensors was calculated and plotted in Fig. 15(b). The results suggest that the LC type 12 (croplands), 4 (deciduous broadleaf forests), and 8 (woody savannas) account the biggest proportion in the accuracy-improved counties by adding the LC data. It is probably because the economy of these counties mainly depends on farming and livestock.

IV. Discussion

This study proposes a MICNN-based model for time-series GDP estimation using VIIRS-DNB NTL and MODIS LC data at county level in CONUS. The R^2 between the predicted LG and the observed LG at county level in CONUS is 0.81, 0.83, 0.83, and 0.83 from 2012 to 2015, indicating a strong correlation between the observed and estimated GDP, and showing a good performance of the MICNN-based model. To evaluate the performance of the proposed method, the traditional log(TNL)–log(GDP) linear regression method was used as the baseline. It is found that the traditional linear regression method has poor performance in the counties with exceptional TNL. There are two possible reasons; first, a simple linear regression method which only relies on the mean value of regional NTL cannot fit the relationship between time-series NTL data and GDP very well, and second, it is inadequate to estimate GDP by using only NTL data, especially in some regions with abnormal TNL values. For example, the TNL of the Highland County, VA is extremely weak; however, the county has a very high regional GDP for its farming and livestock. On the contrary, the TNL of El Paso County, CO is relatively high, but the regional GDP of the county is not as high as expected. To overcome the shortcomings, our method uses the MICNN-based model as the nonlinear method for GDP prediction. By using the histogram-based feature extraction, the model relies on not only the VIIRS-DNB feature but also MODIS LC feature, which is a proxy for the economic structure of a region, and therefore is considered a good supplement to NTL data. The results prove the proposed method can outperform the traditional log(TNL)–log(GDP) linear regression method. Furthermore, we have also added the experiment of random forest (RF) for a comparison. The RMSEs of the RF are 0.40, 0.39, 0.39, and 0.39 from 2012 to 2015; and the R^2 are 0.68, 0.69, 0.71, and 0.70, which suggest that our method can also outperform the RF.

In order to further ascertain the influence of the NTL and land-cover pattern to PE distribution, a statistical analysis of tensors was made in terms of PE grades. The statistical mean was calculated for each bin of all the tensors by selected grades. Fig. 16 shows the means of each bin in the LC tensors by three different grades of PE from 0 to 20% above, with the higher value representing the larger land-cover area of a certain type.

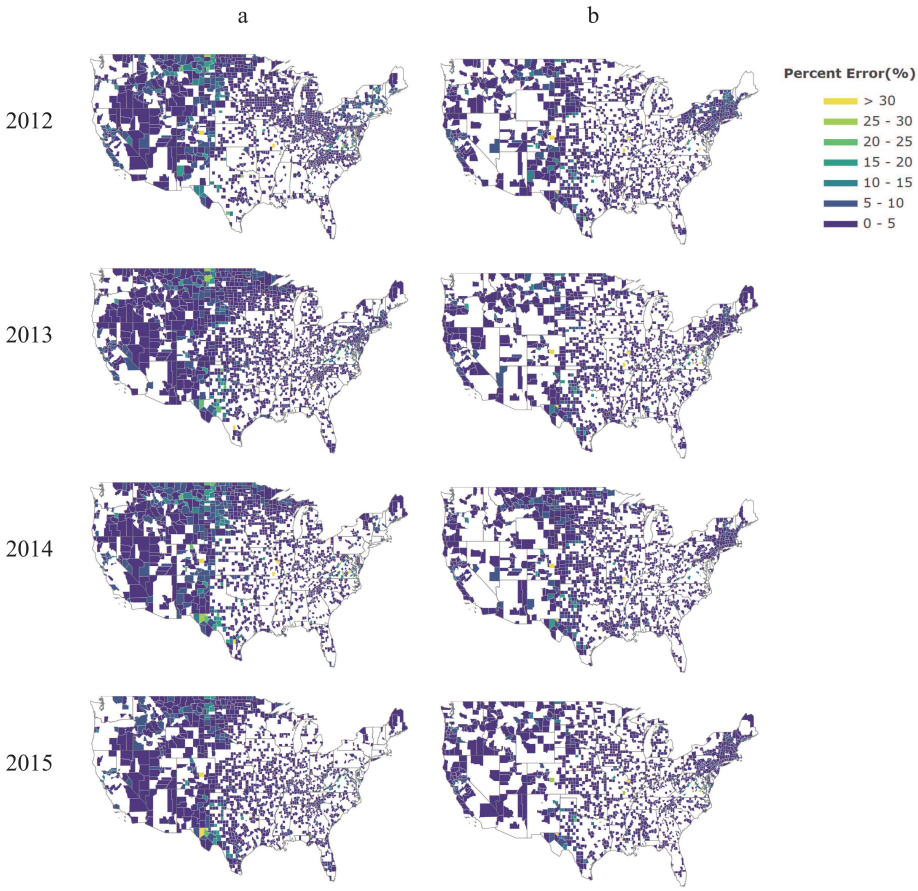


Fig. 14. The maps of counties where the results are improved. (a) The counties where the results are improved from baseline to MICNN-VIIRS. (b) The counties where the results are improved from MICNN-VIIRS to MICNN-VIIRS and LC.

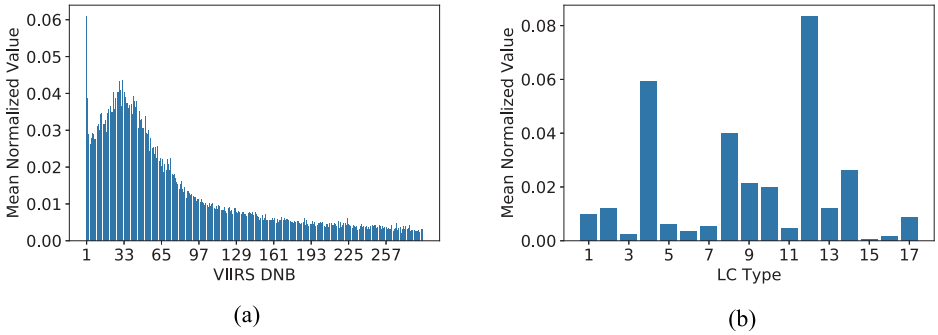


Fig. 15. Statistical mean of tensors of improved counties. (a) VIIRS tensors of counties improved from baseline to MICNN-VIIRS. (b) LC tensors of counties improved from MICNN-VIIRS to MICNN-VIIRS and LC.

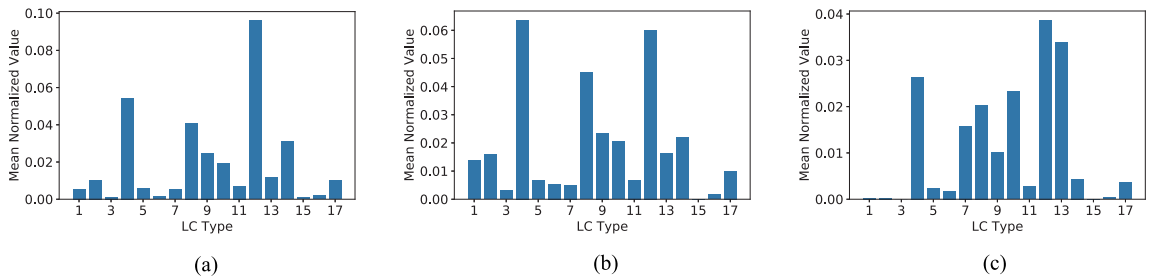


Fig. 16. Statistical mean of LC tensors by grades. (a) LC tensors with PE in [0, 5). (b) LC tensors with PE in [5, 10). (c) LC tensors with PE in [20).

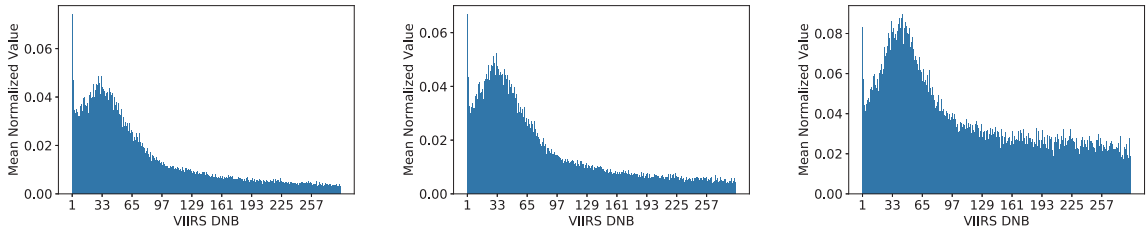


Fig. 17. Statistical mean of VIIRS tensors by grades. (a) VIIRS tensors with PE in [0, 5). (b) VIIRS tensors with PE in [5, 10). (c) VIIRS tensors with PE in [20).

There are significant trends of certain land cover types changes from the distribution of different grades. The value of the LC types 12 and 14, representing croplands or natural vegetation mosaics, declines with the increase of PE, indicating that the relatively accurate prediction always happens in the counties having more croplands. In contrast, the value of type 13 (urban lands) increases with the increase of the PE, which means that the poor prediction occurs in the counties with larger urban lands. Furthermore, the means of each bin in the VIIRS-DNB tensors by the same grades are also shown in Fig. 17. There is a similarity in the distributions, i.e., each distribution has only one wave crest between the 1st and 97th bin, representing the DN of VIIRS-DNB from 1 to 24. However, the values of the wave crest and trough both increase when the PE is greater than 5%. These results suggest that most of the DN of VIIRS-DNB concentrate in the range from 1 to 24, and the higher PE often happened in the counties having more bright pixels, especially range from 1 to 24.

V. Conclusion

It is the first time that an MICNN-based model, which combined VIIRS DNB NTL data and LC data, was tested for time-series GDP estimation at the county level in CONUS. Compared with the baseline, the proposed method showed better performance with R^2 of 0.81, 0.83, 0.83, and 0.83 from 2012 to 2015. The corresponding RMSE of the proposed method was 0.31, 0.29, 0.30, and 0.29. Besides, there are also several new findings as follows. First, based on the model, the predicted GDP and PE maps were generated from 2012 to 2015, which were compared with the results of the linear model as the baseline. The comparison suggested that the proposed method can outperform the baseline mainly in the counties with exceptional TNL. One possible reason for this is that the linear model is commonly used in economic parameter estimation; however, it cannot perform well in some special cases with complex economic structure. Besides, NTL data may be an effective economic predictor in cities, but it cannot account for the economic parameters well in some counties, for instance, whose main economic activities are farming or livestock. In contrast, the proposed method can overcome the shortcomings by combining the features from both NTL and LC data, and then build the relationship between these features and GDP based on the MICNN. Moreover, the LC data can contribute to more improvement than only feature richness from NTL. Second, based on the result of the proposed method, a statistic analysis of different tensors was made by different PE grades, which indicates that the counties where the PE is relatively high are liable to have more bright NTL data and fewer croplands.

As a preliminary attempt to investigate US county-level GDP estimation by using DL, some limitations still exist. To begin with, copious amounts of data are required to train DL algorithms as they learn progressively. There have been only a few years since the NOAA/NGDC began to publish NPP-VIIRS NTL data in April 2012. Additionally, BEA only has released county-level GDP data for several years. However, longer term data cannot only help the DL network to better understand the relationship between features and economic data, but also supply abundant data for validation. Next, the proposed model only has two input interfaces which can be extended if more information should be integrated such as population, geographical information, and economic/industry structure data. Economic/industry structure is an important factor in GDP; high-tech economy can produce higher GDP with less light/space. This can explain why we have high prediction error in high NTL value areas, because cities have very diverse economic structure, which was not considered in the study. For example, San Francisco area has much higher per capita productivity/GDP because of high-tech than El Paso County, CO, where the main industry is heavy industry. The low NTL area usually has similar farming/forest/livestock economic structure which is considered in LC data. This is why in the low NTL area, the prediction is more accurate. We can include total population and population density as additional inputs. What is more is that, the asymmetric network architecture may also be tested in terms of the different inputs in the future.

Overall, this study introduces an MICNN-based method for GDP estimation using multisource remote sensing data. It is a new innovative application of DL which bridges remote sensing data and time-series GDP directly. Based on the GEE, the framework can transform accessible, passively collected data into histogram-based tensors effectively. Through a deep understanding of various features by CNN, it allows for yearly socio-economic forecasting of small regions objectively. The method is also applicable to other countries, especially for developing countries or low-income regions where socio-economic data are hard to obtain. In the future, given much more data set, the transferred property of the model will be also studied.

APPENDIX

TABLE A1
MCD12Q1 LAND COVER TYPE

Value	Description
1	Evergreen Needleleaf Forests: dominated by evergreen conifer trees (canopy > 2m). Tree cover > 60%.
2	Evergreen Broadleaf Forests: dominated by evergreen broadleaf and palmate trees (canopy > 2m). Tree cover > 60%.
3	Deciduous Needleleaf Forests: dominated by deciduous needleleaf (larch) trees (canopy > 2m). Tree cover > 60%.
4	Deciduous Broadleaf Forests: dominated by deciduous broadleaf trees (canopy > 2m). Tree cover > 60%.
5	Mixed Forests: dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy > 2m). Tree cover > 60%.
6	Closed Shrublands: dominated by woody perennials (1-2m height) > 60% cover.
7	Open Shrublands: dominated by woody perennials (1-2m height) 10-60% cover.
8	Woody Savannas: tree cover 30-60% (canopy > 2m).
9	Savannas: tree cover 10-30% (canopy > 2m).
10	Grasslands: dominated by herbaceous annuals (< 2m).
11	Permanent Wetlands: permanently inundated lands with 30-60% water cover and > 10% vegetated cover.
12	Croplands: at least 60% of area is cultivated cropland.
13	Urban and Built-up Lands: at least 30% impervious surface area including building materials, asphalt and vehicles.
14	Cropland/Natural Vegetation Mosaics: mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation.
15	Permanent Snow and Ice: at least 60% of area is covered by snow and ice for at least 10 months of the year.
16	Barren: at least 60% of area is non-vegetated barren (sand, rock, soil) areas with less than 10% vegetation.
17	Water Bodies: at least 60% of area is covered by permanent water bodies.

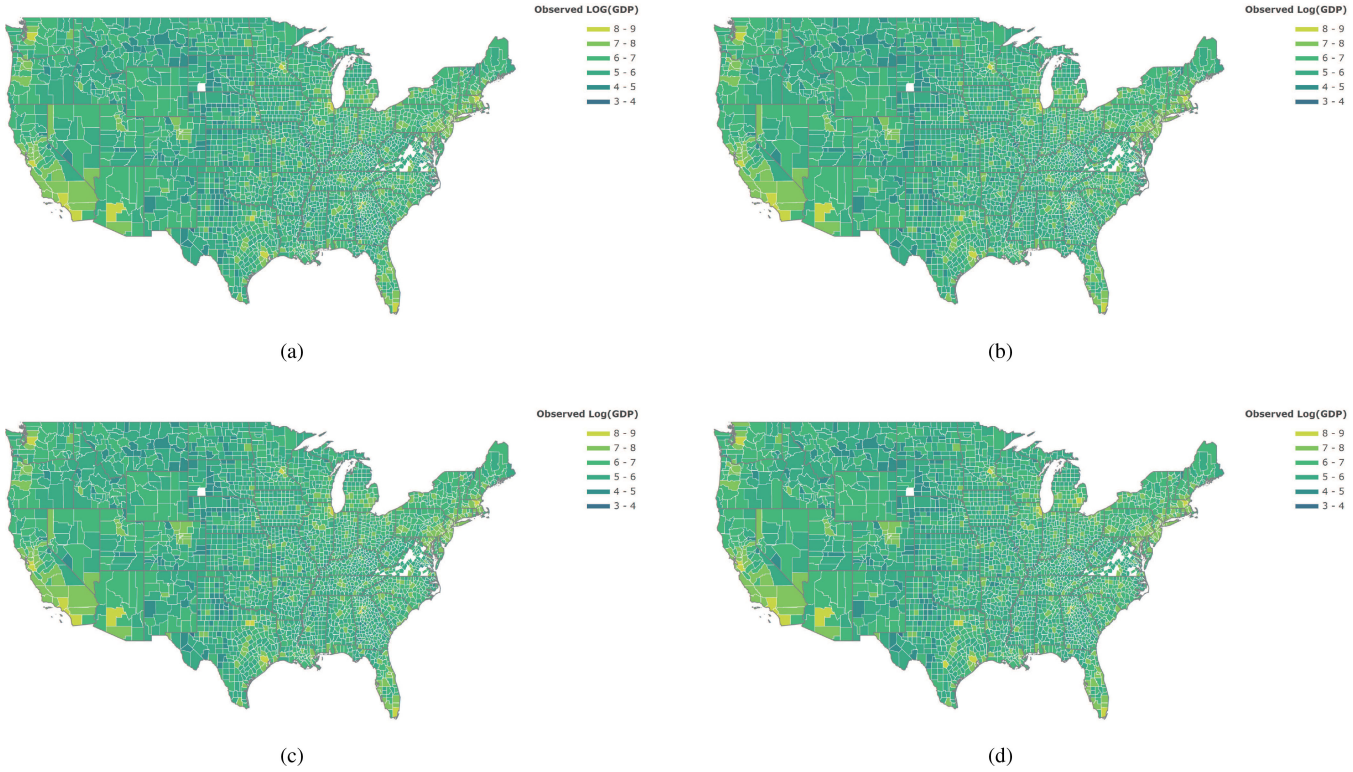


Fig. A1. The BEA log(GDP) map of CONUS at county level (2012–2015).

REFERENCES

[1] C. D. Elvidge, K. E. Baugh, E. A. Kihn, H. W. Kroehl, E. R. Davis, and C. W. Davis, "Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption," *Int. J. Remote Sens.*, vol. 18, no. 6, pp. 1373–1379, 1997.

[2] S. Ebener, C. Murray, A. Tandon, and C. C. Elvidge, "From wealth to health: Modelling the distribution of income per capita at the sub-national level using night-time light imagery," *Int. J. Health Geographics*, vol. 4, no. 1, p. 5, Feb. 2005.

[3] P. Sutton, C. Elvidge, and G. Tilottama, "Estimation of gross domestic product at sub-national scales using nighttime satellite imagery," *Int. J. Ecological Econ. Statist.*, vol. 8, pp. 5–21, Jan. 2007.

[4] X. Chen and W. D. Nordhaus, "Using luminosity data as a proxy for economic statistics," *Proc. Nat. Acad. Sci.*, vol. 108, no. 21, pp. 8589–8594, 2011.

[5] C. Small, F. Pozzi, and C. Elvidge, "Spatial analysis of global urban extent from DMSP-OLS night lights," *Remote Sens. Environ.*, vol. 96, no. 3, pp. 277–291, 2005.

[6] C. N. Doll, J.-P. Muller, and J. G. Morley, "Mapping regional economic activity from night-time light satellite imagery," *Ecological Econ.*, vol. 57, no. 1, pp. 75–92, 2006.

[7] T. Ma, Y. Zhou, Y. Wang, C. Zhou, S. Haynie, and T. Xu, "Diverse relationships between SUOMI-NPP VIIRS night-time light and multi-scale socioeconomic activity," *Remote Sens. Lett.*, vol. 5, no. 7, pp. 652–661, 2014.

[8] D. J. Forbes, "Multi-scale analysis of the relationship between economic statistics and DMSP-OLS night light images," *GIScience Remote Sens.*, vol. 50, no. 5, pp. 483–499, 2013.
[9] X. Chen and W. D. Nordhaus, "VIIRS nighttime lights in the estimation of cross-sectional and time-series GDP," *Remote Sens.*, vol. 11, no. 9, 2019. [Online]. Available: http://www.mdpi.com/2072-4292/11/9/1057
[10] M. Zhao, W. Cheng, C. Zhou, M. Li, N. Wang, and Q. Liu, "GDP spatialization and economic differences in South China based on NPP-VIIRS nighttime light imagery," *Remote Sens.*, vol. 9, no. 7, 2017. [Online]. Available: http://www.mdpi.com/2072-4292/9/7/673
[11] B. Yu, K. Shi, Y. Hu, C. Huang, Z. Chen, and J. Wu, "Poverty evaluation using NPP-VIIRS nighttime light composite data at the county level in china," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 3, pp. 1217–1229, Mar. 2015.
[12] X. Li, H. Xu, X. Chen, and C. Li, "Potential of NPP-VIIRS nighttime light imagery for modeling the regional economy of china," *Remote Sens.*, vol. 5, no. 6, pp. 3057–3081, 2013.
[13] K. Shi *et al.*, "Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: A comparison with DMSP-OLS data," *Remote Sens.*, vol. 6, no. 2, pp. 1705–1724, 2014.
[14] F. Bickenbach, E. Bode, P. Nunnenkamp, and M. Soeder, "Night lights and regional GDP," *Rev. World Econ.*, vol. 152, no. 2, pp. 425–447, May 2016.
[15] W. Nordhaus and X. Chen, "A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics," *J. Econ. Geography*, vol. 15, no. 1, pp. 217–246, Jan. 2015.
[16] S. Keola, M. Andersson, and O. Hall, "Monitoring economic development from space: Using nighttime light and land cover data to measure economic growth," *World Develop.*, vol. 66, pp. 322–334, Feb. 2015.
[17] W. Yue, J. Gao, and X. Yang, "Estimation of gross domestic product using multi-sensor remote sensing data: A case study in Zhejiang province, East China," *Remote Sens.*, vol. 6, no. 8, pp. 7260–7275, 2014.
[18] K. Faisal, A. Shaker, and S. Habbani, "Modeling the relationship between the gross domestic product and built-up area using remote sensing and gis data: A case study of seven major cities in canada," *ISPRS Int. J. Geo-Inf.*, vol. 5, no. 3, 2016. [Online]. Available: http://www.mdpi.com/2220-9964/5/3/23
[19] D. Donaldson and A. Storeygard, "The view from above: Applications of satellite data in economics," *J. Econ. Perspectives*, vol. 30, no. 4, pp. 171–98, Nov. 2016.
[20] X. Zhao *et al.*, "Estimation of poverty using random forest regression with multi-source data: A case study in Bangladesh," *Remote Sens.*, vol. 11, no. 4, 2019, Art. no. 375.
[21] Q. Chen, X. Hou, X. Zhang, and C. Ma, "Improved GDP spatialization approach by combining land-use data and night-time light data: A case study in China's continental coastal area," *Int. J. Remote Sens.*, vol. 37, no. 19, pp. 4610–4622, 2016.
[22] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state-of-the-art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.
[23] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, vol. 353, no. 6301, pp. 790–794, 2016.
[24] A. Perez, C. Yeh, G. Azzari, M. Burke, D. Lobell, and S. Ermon, "Poverty prediction with public landsat 7 satellite imagery and machine learning," 2017, *arXiv:1711.03654*.
[25] J. Fan, T. Ma, C. Zhou, Y. Zhou, and T. Xu, "Comparative estimation of urban development in China's cities using socioeconomic and DMSP/OLS night light data," *Remote Sens.*, vol. 6, no. 8, pp. 7840–7856, 2014.
[26] N. Clinton *et al.*, "A global geospatial ecosystem services estimate of urban agriculture," *Earths Future*, vol. 6, no. 1, pp. 40–60, Jan. 2018.
[27] X. Liu *et al.*, "High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform," *Remote Sens. Environ.*, vol. 209, pp. 227–239, May 2018.
[28] A. Agapiou, "Remote sensing heritage in a petabyte-scale: Satellite data and heritage Earth Engine © applications," *Int. J. Digit. Earth*, vol. 10, no. 1, pp. 85–102, Jan. 2017.
[29] G. Azzari and D. B. Lobell, "Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring," *Remote Sens. Environ.*, vol. 202, pp. 64–74, Dec. 1 2017.
[30] J. N. Hird, E. R. DeLancey, G. J. McDermid, and J. Kariyeva, "Google Earth engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping," *Remote Sens.*, vol. 9, no. 12, Dec. 2017, Art. no. 1315.
[31] X. Jing, X. Shao, C. Cao, X. Fu, and L. Yan, "Comparison between the SUOMI-NPP day–night band and DMSP-OLS for correlating socio-economic variables at the provincial level in China," *Remote Sens.*, vol. 8, no. 1, 2016. [Online]. Available: https://www.mdpi.com/2072-4292/8/1/17
[32] Z. Dai, Y. Hu, and G. Zhao, "The suitability of different nighttime light data for GDP estimation at different spatial scales and regional levels," *Sustainability*, vol. 9, no. 2, 2017. [Online]. Available: http://www.mdpi.com/2071-1050/9/2/305
[33] NOAA, "Version 1 VIIRS day/night band nighttime lights," 2019. Accessed Apr. 19, 2019. [Online]. Available: https://ngdc.noaa.gov/eog/viirs/ download_dnb_composites.html
[34] M. Friedl and D. Sulla-Menashe, "MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006," NASA EOSDIS Land Processes DAAC, 2019, doi: 10.5067/MODIS/MCD12Q1.006.
[35] BEA, "Prototype gross domestic product by county, 2012-2015," 2018. [Online]. Available: https://www.bea.gov/news/2018/prototype-gross-domestic-product-county-2012-% 2015
[36] E. S. Pearson, R. B. D'Agostino, and K. O. Bowman, "Tests for departure from normality: Comparison of powers," *Biometrika*, vol. 64, no. 2, pp. 231–246, 1977.
[37] DATA.GOV, "U.S. County Boundaries," 2019. Accessed Apr. 19, 2019. [Online]. Available: https://catalog.data.gov/dataset/u-s-county-boundaries/
[38] N. Cohen, O. Sharir, and A. Shashua, "On the expressive power of deep learning: A tensor analysis," in *Proc. Conf. Learn. Theory*, 2016, pp. 698–728.
[39] GEE, "Google Earth engine," 2019. Accessed Apr. 19, 2019 [Online]. Available: https://developers.google.com/earth-engine
[40] Y. Zhou *et al.*, "A global map of urban extent from nightlights," *Environ. Res. Lett.*, vol. 10, no. 5, 2015, Art. no. 054011.
[41] Y. Liu, T. Delahunty, N. Zhao, and G. Cao, "These lit areas are undeveloped: Delimiting Chinas urban extents from thresholded nighttime light imagery," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 50, pp. 39–50, 2016.
[42] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.
[43] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, May 2017.
[44] S. M. Pandey, T. Agarwal, and N. C. Krishnan, "Multi-task deep learning for predicting poverty from satellite images," in *32nd AAAI Conf. Artif. Intell.*, 2018.
[45] A. Gulli and S. Pal, *Deep Learning with Keras*. Packt Publishing Ltd., Livery Place, 35 Livery Street Birmingham B3 2PB, UK, pp. 7793–7798, 2017.
[46] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.
[47] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2013, pp. 8609–8613.
[48] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[49] J.-W. Ma, C.-H. Nguyen, K. Lee, and J. Heo, "Regional-scale rice-yield estimation using stacked auto-encoder with climatic and MODIS data: A case study of South Korea," *Int. J. Remote Sens.*, vol. 40, no. 1, pp. 51–71, Jan. 2019.
[50] J. Proville, D. Zavala-Araiza, and G. Wagner, "Night-time lights: A global, long term look at links to socio-economic trends," *PLoS ONE*, vol. 12, no. 3, pp. 1–12, Mar. 2017.

Jie Sun received the Ph.D. degree in photogrammetry and remote sensing from the Faculty of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2011.

Since 2011, he has been a Lecturer with the China University of Geosciences, Wuhan, China. He has authored or coauthored more than 20 papers. His research interests include remote sensing application for geology, agriculture, and machine learning for object recognition.

Liping Di received the Ph.D. degree in remote sensing/GIS (geography) from the University of Nebraska Lincoln, Lincoln, NE, USA in 1991.

He is the Founding Director of the Center for Spatial Information Science and Systems, a University Research Center, George Mason University, Fairfax, VA, USA. He is also a Professor with the Department of Geography and Geoinformation Science, George Mason University. He has engaged in geoinformation science research for more than 30 years and has authored or coauthored more than 400 publications. His research interests include geospatial information standards, geospatial cyberinfrastructure, web-based geospatial information and knowledge systems, and geoinformation science applications, particularly in agriculture.

Dr. Di was the recipient of many prestigious awards, such as the Merit Award from International Committee on Information Technology Standards, the Honor Award from the Secretary of US Department of Agriculture, and the R&D Award from the R&D Magazine. He has received more than $57 million research grants from US federal agencies and international organizations as the Principal Investigator. He was the Chair of INCITS/L1, the US national committee responsible for setting US national standards on geographic information and representing the United States at the ISO Technical Committee 211 (ISO TC 211) from 2010 to 2016. He was also the Elected Chair of Data Archive and Distribution Technical Committee (which has been renamed to Earth Science Informatics Technical Committee) of the IEEE Geoscience and Remote Sensing Society from 2005 to 2009.

Ziheng Sun received the B.S. degree in geographic information system and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2009 and 2015, respectively.

Since February 2013, he has been with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA, USA, where he is currently working as a Research Assistant Professor. His research interests include geospatial web services, workflow-based geospatial modeling, machine learning, remote sensing, agricultural land cover classification, and drought monitoring. His research interests include regular object similarity index, parameterless automatic classification, customized vegetation condition index, LSTM-aided crop mapping, web-based supervised classification system, and location-based service for onsite information access.

Jieyong Wang received the Ph.D. degree in human geography from the Chinese Academy of Sciences, Beijing, China, in 2008.

He has been with the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences since 2008, where he is currently working as an Associate Professor. His research interests include rural geography and agricultural land monitoring.

Yingdan Wu received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China.

She became South Lake Scholar in 2015 and is currently an Associate Professor, Masters Tutor with the School of Science, Hubei University of Technology. She mainly engaged in the theory and method of high-accuracy geolocation and spatial information extraction for satellite remote sensing imagery, information coding, and solar radiation prediction research. More than 20 papers in important journals have been published at home and abroad, of which 12 papers are indexed by SCI/EI. As the principal investigator, she has taken charge of more than eight projects supported by the National Natural Science Foundation of China, Natural Science Foundation of Hubei Province, Educational Commission of Hubei Province of China, and so on.

Dr. Wu has also directed students to participate in the China Undergraduate Mathematical Contest in Modeling (CUMCM) and the MCM/ICM Contest, and received the First Prize several times.