


Improved GDP spatialization approach by combining land-use data and night-time light data: a case study in China's continental coastal area

Qing Chen, Xiyong Hou, Xiaochun Zhang & Chun Ma


To cite this article: Qing Chen, Xiyong Hou, Xiaochun Zhang & Chun Ma (2016) Improved GDP spatialization approach by combining land-use data and night-time light data: a case study in China's continental coastal area, *International Journal of Remote Sensing*, 37:19, 4610-4622, DOI: 10.1080/01431161.2016.1217440



To link to this article: <http://dx.doi.org/10.1080/01431161.2016.1217440>

 View supplementary material 

 Published online: 23 Aug 2016.

 Submit your article to this journal 

 Article views: 131

 View related articles 

 View Crossmark data 



Improved GDP spatialization approach by combining land-use data and night-time light data: a case study in China's continental coastal area

Qing Chen^{a,b,c}, Xiyong Hou^{a,c}, Xiaochun Zhang^d and Chun Ma^d

^aYantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, China; ^bCollege of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China; ^cKey Laboratory of Coastal Environmental Processes and Ecological Remediation, Chinese Academy of Sciences, Yantai, China; ^dCarnegie Department of Global Ecology, Stanford University, Stanford, CA, USA

ABSTRACT

Gross domestic product (GDP) reflects a nation or region's economic growth as a whole, and is the sum of product in the primary, secondary, and tertiary sectors of the economy in the area. However, statistical GDP data is problematic in integrated application with geographical data. The GDP spatialization data, which shows the GDP in grid cells and often is obtained by operating a spatialization model, is more useful than its officially published statistical data recorded by administrative units in both spatial representation and application. Thus, there is a need to improve the GDP spatialization models, and to present these models in a way as clear and transparent as possible. In this article, by taking China's continental coastal area as a case study area, we combined economic census data, land-use data, and night-time light data together, and developed a technique that we call the 'dynamic regionalization' method to improve the GDP spatialization products. We then created GDP spatialization models for three sectors of the economy (i.e. the primary, the secondary, and the tertiary sector) in 2000, 2005, and 2010, respectively. We find the following. (1) Because the 'overflow' effect of night-time light data has a bad influence on spatialization models, we used land-use data to distinguish the distribution plots of the tertiary sector on night-time light images. Compared with setting a threshold merely, land-use data can more effectively remove the 'overflow' effect. (2) Owing to the prominent spatial heterogeneity of GDP distribution in China's continental coastal area, building one spatialization model for the whole area would probably produce the estimated products with poor accuracy, so the 'dynamic regionalization' method was adopted to dynamically divide the whole study area into several subregions, and build separate spatialization models for each subregion. The accuracy assessment showed that the new method improved the accuracy of GDP spatialization data, especially in the area with high spatial heterogeneity.

ARTICLE HISTORY

Received 8 October 2015
Accepted 20 July 2016

CONTACT Xiyong Hou ✉ xyhou@yic.ac.cn 📍 Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003, China

📄 Supplemental data for this article can be accessed [here](#).

© 2016 Informa UK Limited, trading as Taylor & Francis Group

1. Introduction

A country or region's economic growth is typically measured in terms of gross domestic product (GDP), that is, the market value of all final goods and services produced in a given country or region in a given year (Sullivan and Sheffrin 1996). It can be basically divided into a product of three sectors: (1) the primary sector, including agriculture, forestry, animal husbandry, and fishery; (2) the secondary sector, including industry and construction; and (3) the tertiary sector, which is the service sector. For quantitatively assessing interactions between socio-economy and physical environment, it is necessary to integrate socio-economic data with physical environmental data (Doll, Muller, and Morley 2006). However, the spatial representations of socio-economic data and environmental data are usually different. In general, the socio-economic data is collected and compiled into a two-dimensional table from various administrative units – national, state, municipal, and county through census or sampling survey, whereas the physical and environmental data are aggregated in physical and ecological units, such as watersheds, or soil and vegetation zones regardless of administration (Elvidge et al. 2009; Hu et al. 2011). Therefore, socio-economic data sets such as GDP need to be allocated to a regular grid to facilitate their integration with other environmental and physical data sets across administrative boundaries (Doll, Muller, and Morley 2006).

Studies have reported that night-time light data is an effective and low-cost way of estimating the spatial distribution of GDP statistical data at the national or sub-national level (Elvidge et al. 1997; Doll, Muller, and Elvidge 2000; Lo 2002; Doll, Muller, and Morley 2006). But However, application of night-time light data has two problems: one is that the night-time light data is saturated in city centres and other bright areas (Sutton, Ghosh, and Elvidge 2007; Ghosh et al. 2010b), and the other is that the 'overflow' effect of night-time light data, which is caused by non-coherent light that radiates from its source, significantly affects the accuracy of night-time light data in coastal cities (Townsend and Bruce 2010). The pixels around the bright spots would be brighter than they should be, due to the radiation from the bright pixels in the centre. In addition, it has been demonstrated that the spatialization models often underestimate in highly productive areas and overestimate in less-productive areas (Harvey 2002). In other words, the magnification of error in the estimated results would probably occur in spatially heterogeneous areas. Another issue is whether the unlit parts realistically produce zero GDP. For example, waterbody area is usually black in night-time light data, which means that the waterbody produces zero GDP in a spatialization model based on night-time light data only, but it can contribute to the growth of the primary sector with possibility, such as developing aquaculture by the waterbody. Elvidge et al. (2002) provided a comparison of a number of different data products using night-time light data. It could be used to calculate the sum of light intensity values for each administrative unit and to distribute the percentage of the total estimated economic activity not attributed to agriculture (in other words, attributed to commercial/industrial activity) for each administrative unit (Ghosh et al. 2010b). Therefore, it is often more practicable to decompose GDP into agricultural and non-agricultural products (Wu et al. 2013) or into the product of the primary, secondary, and tertiary sectors (Han et al. 2012), and then to model the spatial distribution separately.

The aims of this work are: (1) to build regression models using the dynamic regionalization method (which is specified in [Section 2.3](#)) for estimating the official statistical GDP figures (for three sectors) for individual counties in the years 2000, 2005, and 2010, using land-use data and night-time light images as predictors for each of the three industry sectors; (2) to apply the regression model to each 1 km grid square of China's continental coastal area; and (3) to validate these estimated products by aggregating them up to the county level and comparing the county-level estimated product with the original official statistical figures.

2. Model development

The overall method of GDP spatialization is shown as a flow chart in [Figure 1](#). The spatialized GDP of each sector (using GDP_1 , GDP_2 , and GDP_3 for the production of the primary sector, the secondary sector, and the tertiary sector, respectively) was calculated first and then added together to derive the overall GDP spatialization results. Here, GDP_1 is calculated based on land-use data owing to the fact that the primary sector relates to the natural resource closely. For the calculation of GDP_3 , night-time light data is used because the product of the service industry reflects the economic activities of human beings, which can be captured by night-time light data (Han et al. 2012). As for GDP_2 , we associate it with land-use data considering that night-time light data performs unsatisfactorily in areas of saltern, which is nearly unlit at night.

2.1. Preprocessing land-use data and night-time light data

The primary sector is spatially related with farmland, forest, grassland, inland freshwaters, and mariculture. The secondary industry is spatially related with city, rural settlement, isolated industrial-mining, and salt pan. To facilitate the analysis of relationships among GDP_1 , GDP_2 , and their relating land uses, we developed a model to extract land-use grid images from vector land-use maps. Specifically, we build a fishnet map at 1 km spatial scale, and then overlay it with the land-use map to count the area of certain types of land use in each 1 km cell. Taking farmland as an example, the grid image depicts the distribution of farmland (I_{fa}) in each 1 km cell. Similar calculations are applied to forest, grassland, inland freshwaters, mariculture, city, rural settlement, isolated industrial-mining, and saltern, and we obtain the corresponding land-use grid images, that is, I_{fo} , I_{gr} , I_{wa} , I_{ma} , I_{ci} , I_{ru} , I_{is} , and I_{sa} .

To correct the 'overglow' effect of night-time light data, some studies (Li and Fang 2014; Zhao, Currit, and Samson 2011) subjectively set a threshold at a constant digital number (DN) to identify the distribution region of the tertiary sector. We consider that: (1) the tertiary sector represents the service industry that exists along with human activities, and thus it is necessary to distinguish between a populated area and transportation sites such as airport and harbour, and (2) these aforementioned places are distinctly defined in land-use data: city, rural settlement, and isolated industrial-mining. The distributions of the tertiary sector in a night-time light data are modified by land-use data. Practically, the first step is to create a grid image at the 1 km spatial scale by adding independent land-use maps I_{ci} , I_{ru} , and I_{is} together. The second step is to

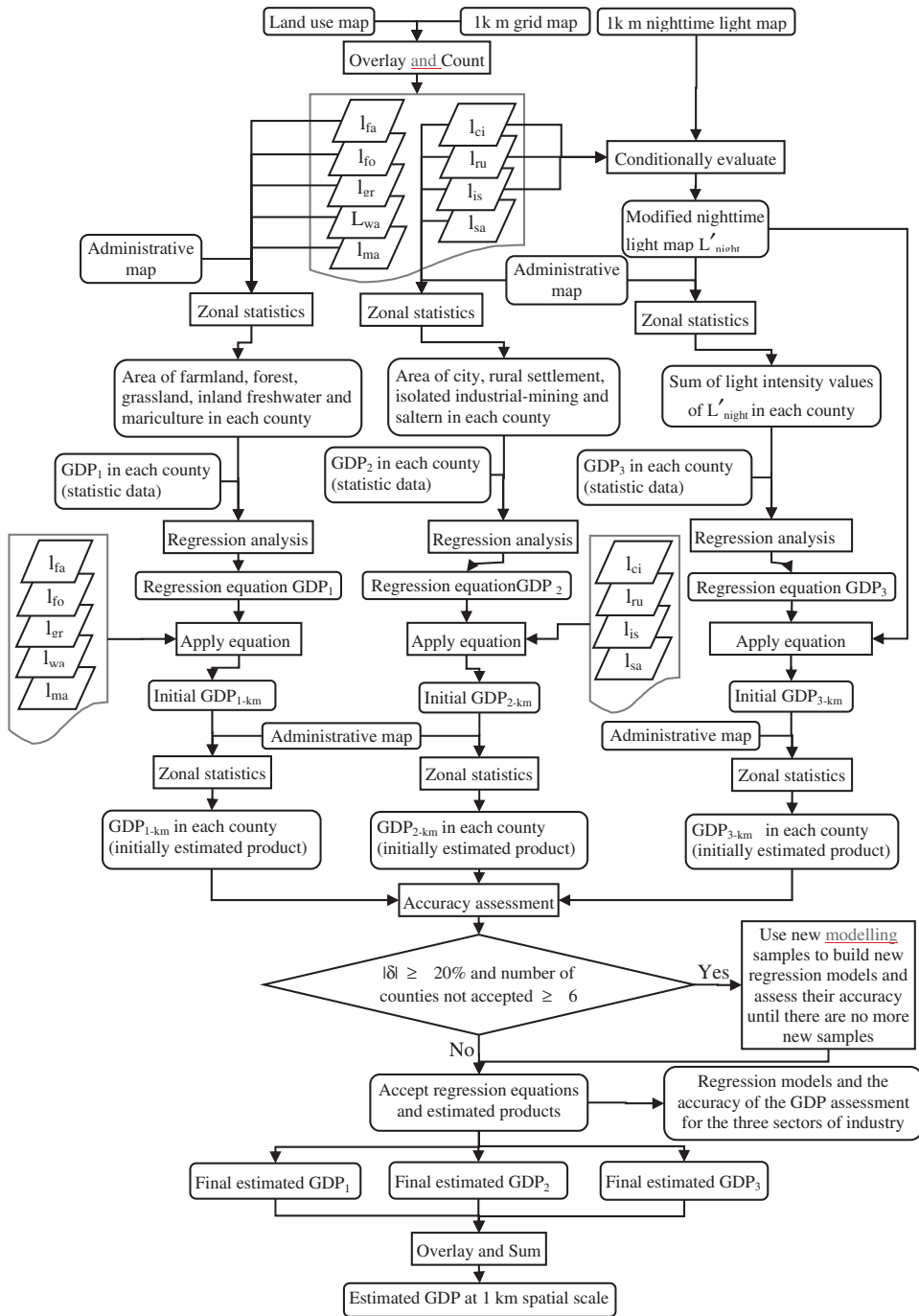


Figure 1. Flow chart of GDP spatialization, here GDP_1 , GDP_2 , and GDP_3 are the production of the primary sector, the secondary sector, and the tertiary sector, respectively; I_{fa} , I_{fo} , I_{gr} , I_{wa} , I_{ma} , I_{ci} , I_{ru} , I_{is} and I_{sa} are the distribution of farmland, forest, grassland, inland freshwaters, mariculture, city, rural settlement, isolated industrial-mining, and saltern in each cell.

calculate the night-time light data (L_{night}) into new data (L'_{night}) with reducing the 'overflow' effect, according to the following formulation:

$$L'_{\text{night}} = \begin{cases} L_{\text{night}}(I_{\text{ci}} + I_{\text{ru}} + I_{\text{is}} > 0) \\ 0(I_{\text{ci}} + I_{\text{ru}} + I_{\text{is}} = 0) \end{cases} \quad (1)$$

2.2. Analysing correlations among data at the county level

We summarized the total area of one type of land use (ΣI) and the sum of light intensity values ($\Sigma L'_{\text{night}}$) within each county using the Zonal Statistic tool in ArcGIS. Then we conducted regression analysis at the county level among GDP_1 , GDP_2 , GDP_3 , and the area of certain types of land-use data or the sum of night-time lights. The coefficient of determination R^2 for the estimated products of each sector (see tables in supplementary material) indicates linear regressions of GDP_1 and GDP_2 are acceptable whereas GDP_3 is relatively unsatisfactory. However, through modification with the dynamic regionalization method (explained in Section 2.3), linear regressions of GDP_3 improved. Thus, linear regressions were adopted for building spatialization models. These linear regression equations are listed as follows:

$$\text{GDP}_1 = k_{\text{fa}} \times \Sigma I_{\text{fa}} + k_{\text{fo}} \times \Sigma I_{\text{fo}} + k_{\text{gr}} \times \Sigma I_{\text{gr}} + k_{\text{wa}} \times \Sigma I_{\text{wa}} + k_{\text{ma}} \times \Sigma I_{\text{ma}} + b_1, \quad (2)$$

$$\text{GDP}_2 = k_{\text{ci}} \times \Sigma I_{\text{ci}} + k_{\text{ru}} \times \Sigma I_{\text{ru}} + k_{\text{is}} \times \Sigma I_{\text{is}} + k_{\text{sa}} \times \Sigma I_{\text{sa}} + b_2, \quad (3)$$

$$\text{GDP}_3 = k_{\text{night}} \times \Sigma L'_{\text{night}} + b_3, \quad (4)$$

where b_1 , b_2 , and b_3 stand for the constants in each equation, ΣI_{fa} means the total area of farmland within each county, k_{fa} represents the regression coefficient of ΣI_{fa} , and other ' Σ 's and ' k 's are defined similarly, and k_{night} is the regression coefficient of $\Sigma L'_{\text{night}}$. Only relevant land uses can contribute to a certain industry, for example, the airport would make no contribution to the primary sector, and Due wise the farmland does not contribute to the secondary or tertiary sector. Therefore, we force the constants in equations to be zero, to reflect the fact that there exist grid squares in the image that produce no product in a certain industry.

2.3. Dynamic regionalization

Regression coefficients in equations were calculated at the county level and applied to estimate the product of primary ($\text{GDP}_{1\text{-km}}$), secondary ($\text{GDP}_{2\text{-km}}$), and tertiary ($\text{GDP}_{3\text{-km}}$) sectors for each 1 km grid square. This transformation can be expressed as follows:

$$\text{GDP}_{1\text{-km}} = k_{\text{fa}} \times I_{\text{fa}} + k_{\text{fo}} \times I_{\text{fo}} + k_{\text{gr}} \times I_{\text{gr}} + k_{\text{wa}} \times I_{\text{wa}} + k_{\text{ma}} \times I_{\text{ma}}. \quad (5)$$

$$\text{GDP}_{2\text{-km}} = k_{\text{ci}} \times I_{\text{ci}} + k_{\text{ru}} \times I_{\text{ru}} + k_{\text{is}} \times I_{\text{is}} + k_{\text{sa}} \times I_{\text{sa}}. \quad (6)$$

$$\text{GDP}_{3\text{-km}} = k_{\text{night}} \times L'_{\text{night}}. \quad (7)$$

As discussed before, spatialization data in highly developed or developing areas are prone to underestimation or overestimation, respectively. Therefore, a dynamic

regionalization modelling method is designed to build regression models for each sub-area that are divided dynamically. First, we computed the initial estimated product (i.e. initial GDP_{1-km}, GDP_{2-km}, and GDP_{3-km}) by conducting the steps mentioned in Sections 2.1 and 2.2, and then assessed the accuracy of the estimated products at the county level. This is carried out by aggregating the estimated products for each grid square back up to the county level and comparing that with the official statistical product using the relative error (δ), defined as

$$\delta = \frac{P_m - P_s}{P_s} \times 100\%, \quad (8)$$

where P_m is the estimated product and P_s is the official statistical product.

Based on the absolute value of relative error ($|\delta|$) of the estimated products in each county, the spatialization process is proceeded as follows: the estimation results in counties with $|\delta| < 20\%$ were retained. However, for counties with $|\delta| > 20\%$, calculations were carried out separately in the overestimated counties (where $\delta > 20\%$) and the underestimated counties (where $\delta < -20\%$). For each side, we have two different steps to take depending on the number of counties within it. If the number is no more than 6, considering the maximum of regression coefficients (in the model GDP₁) is 5, we retain those estimated products as well, because the counties are too few to rebuild the regression model; otherwise we dismiss those estimated products, and instead rebuild a new regression model for these counties and reassess the accuracy of the new estimated products coming from the new model until the number of either overestimated or underestimated counties is no more than 6. When the estimated GDP₁, GDP₂, and GDP₃, which are represented as grid images at the 1 km spatial scale, have been finalized, each sector's grid images in the same year were added together to obtain the estimated GDP grid image.

3. Case study area and data

3.1. Study area

China's continental coastal area is energetic in economic activities. Simultaneously, it is subject to frequent natural hazards (Zhang et al. 2012). Here, study area is defined as the coastal cities of China's mainland, including Shanghai and Tianjin municipalities, and the coastal cities in Liaoning, Hebei, Shandong, Jiangsu, Zhejiang, Fujian, Guangdong, and Guangxi provinces (Figure 2). In total, the study area contains 301 counties, which are the smallest available units for GDP statistical data. The Beijing-Tianjin-Hebei Metropolis Circle (BTH), the Yangtze River Delta Economic Zone (YRD), and the Pearl River Delta Economic Zone (PRD) are metropolitan areas for the highest density of population and wealth in China's continental coastal area.

Owing to the wide range of China's continental coastal area and its complicated natural condition, as well as the priority of governmental finance to certain coastal economic zones such as YRD and PRD, the variation among regions within the study area is very remarkable. According to China's Statistical Yearbook for Regional Economy in 2011, the densities of GDP in YRD and PRD, the most developed areas, were more



Figure 2. Location map of China's coastal area.

than 0.34×10^8 CNY km^{-2} (where CNY represents Chinese Yuan, the domestic currency in China), whereas that in the Beibu Gulf Economic Zone was only 0.04×10^8 CNY km^{-2} .

3.2. GDP statistical data

Official statistical data of product in the primary, secondary, and tertiary sectors in the years 2000, 2005, and 2010 are obtained from China's Statistical Yearbook for Regional Economy published in the years 2001, 2006, and 2011, respectively. Each book records the regional economy of the last year. All official statistical data are expressed in 10,000 CNY.

3.3. Land-use data

Land-use maps at the scale of 1:100,000 in 2000, 2005, and 2010 are interpreted based on Landsat Thematic Mapper/Enhanced Thematic Mapper Plus (TM/ETM+) images. Landsat TM/ETM+ images were all captured in the growing season, with a good quality (total cloud cover <5%). The classification system of land use includes farmland, forest, grassland, city,

rural settlement, isolated industrial-mining, inland freshwaters, coastal saltwater, saltern, mariculture, unused land, and shallow water. This classification system considers both land and sea regions in China's coastal zone, and emphasizes on the wetland subdivisions (Di, Hou, and Wu 2014; Di et al., 2015).

3.4. Night-time light data

The US Air Force Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) has a unique capability for the global mapping of artificial lighting present at the Earth's surface. DMSP/OLS basically detects sources of night-time light on the Earth's surface, such as city lights, forest fires, gas flare burn-off, and heavily lit fishing boats, all produced by human activities (Elvidge et al. 1997). The annual composites (from 1992 to 2012) were filtered to remove background noise and fires. The stable light composite represents the average intensity of night-time light in terms of DN values ranging from 0 to 63. These stable light composites were downloaded free from the National Geophysical Data Center (NGDC) website (<http://ngdc.noaa.gov/eog/dmsp/downloadV4-composites.html>). The images are geo-located to 30 arc-second grids, equivalent to approximately 1 km². Night-time light data presently lend statistically significant abilities to estimate population densities (Townsend and Bruce 2010; Azar et al. 2010), constructed surface area/anthropogenic impervious surface area (Sutton et al. 2011, 2012), and economic figures (Sutton, Ghosh, and Elvidge 2007; Ghosh et al. 2010a).

4. Results

4.1. Estimated products in the three sectors

Regression model parameters and the coefficients of determination R^2 among GDP₁, GDP₂, GDP₃, and the related factors are given in Table S1–S7. Mariculture in the primary sector mostly exists in seaside of the study area. Similarly, saltern in the secondary sector only exists in the coastal area. Therefore, the study area is divided into two main parts at first for developing the regression models of GDP₁ and GDP₂ individually: one is the sub-area within mariculture for the primary sector (or saltern for the secondary sector) and the other is the sub-area without mariculture (or saltern).

In addition, the rural settlement is one factor in building the regression model for GDP₂ in 2010, but not in 2000 and 2005. That is because the regression coefficients of rural settlement in modelling in 2000 and 2005 are always negative, and we believe this is because the secondary industries in China during 2000–2005 were rather undeveloped compared to those around the year 2010, but the rural settlements were already extensive before 2010. At that time, rural settlements were really poor, disordered although extensive, rather than those we are now familiar with, equipped with buzzing factories, convenient shops, etc. It is difficult to be convinced that rural settlements in 2000 and 2005 were of great benefit to the secondary industry. In other words, the positive correlations between the area of rural settlement and the development of secondary the industry did not yet exist in 2000 or 2005. Therefore, we exclude rural settlement as an indicating factor for the secondary sector in these two years.

Based on the relative error (δ) of the initial estimated product, the study area is further subdivided into three parts, one is the retained area, in which $|\delta|$ of the initial estimated product is below 20%, and the others, in which δ of the initial estimated product is either above 20% or below -20% , for which separate regression models need to be rebuilt to obtain better estimates. With the processing of rebuilding the regression models, coefficients of determination R^2 generally become higher and get closer to 1 during each modelling process, which means the fit of the regression models become better. And this especially benefits the models of GDP_3 . Table S7 shows that R^2 in the initial regression models of GDP_3 in 2000, 2005, and 2010 are relative low, as opposed to the values found by other studies (Han et al. 2012; Liang and Xu 2013), indicating a high correlation between night-time light map and GDP_3 . However, after rebuilding several regressive models further in the sub-areas, the R^2 gradually starts becoming higher and higher, that is, the estimated products become increasingly closer to the official statistical data, and thus can reflect the real spatial distribution of GDP_3 .

Comparing the spatial distribution maps of three sectors' product and GDP in the study area in 2000, 2005, and 2010 (Figure 3), we can obtain an overview of the development of the three sectors and their contribution to GDP in the spatial distribution during the period 2000 ~ 2010. For example, the spatial distribution map of GDP_1 is mainly covered by yellow and green, the middle-low-value area in 2000. However, in 2005, the yellow middle-value area increases significantly, and meanwhile several patches of red high-value area can be found in the side close to the sea. By 2010, the red high-value area sprawls out mainly existing in the upper part rather than the lower part of the coastal area, and almost takes over half of the yellow middle-value area in 2005. From the perspective of the replacement, increase and decrease of colours in the map, we can easily recognize that GDP_1 has largely increased from 2000 to 2010.

4.2. Overall accuracy assessment

According to the rules of the dynamic regionalization modelling method, $|\delta|$ of the estimated product in the vast majority of the whole study area will be below 20%. As Table 1 shows, the proportions of counties whose $|\delta|$ of the estimated GDP_1 in 2000, 2005, and 2010 are under 20% are 92%, 92%, and 91%, GDP_2 90%, 90%, and 92%, and GDP_3 92%, 94%, and 92%, respectively. The standard deviations of $|\delta|$ of these models are rather small, especially those of the tertiary sector; thus the distributions of $|\delta|$ are concentrated, which means few of the outliers (points far away from the majority) of $|\delta|$ exist.

Comparing the estimated GDP_1 , GDP_2 , GDP_3 , and GDP with their official statistical data (Figure S1), we can find that most dots in the figure distribute evenly along the line $y = x$, that is, the estimated products are quite close to the statistical data and the precision of the regression models is proved to be rather high in general as well.

The improvement of precision after using the dynamic regionalization modelling method can be seen by the comparison of the relative error δ between the initial and final estimated products (Figure S2). For the initial estimated product, the δ is bigger and distributes in a more scattered way as a whole. All outliers of δ are positive numbers, indicating that the major modelling problem lies in overestimating. Although for each industry, the δ of the estimated GDP_1 is relatively equally distributed, two or three

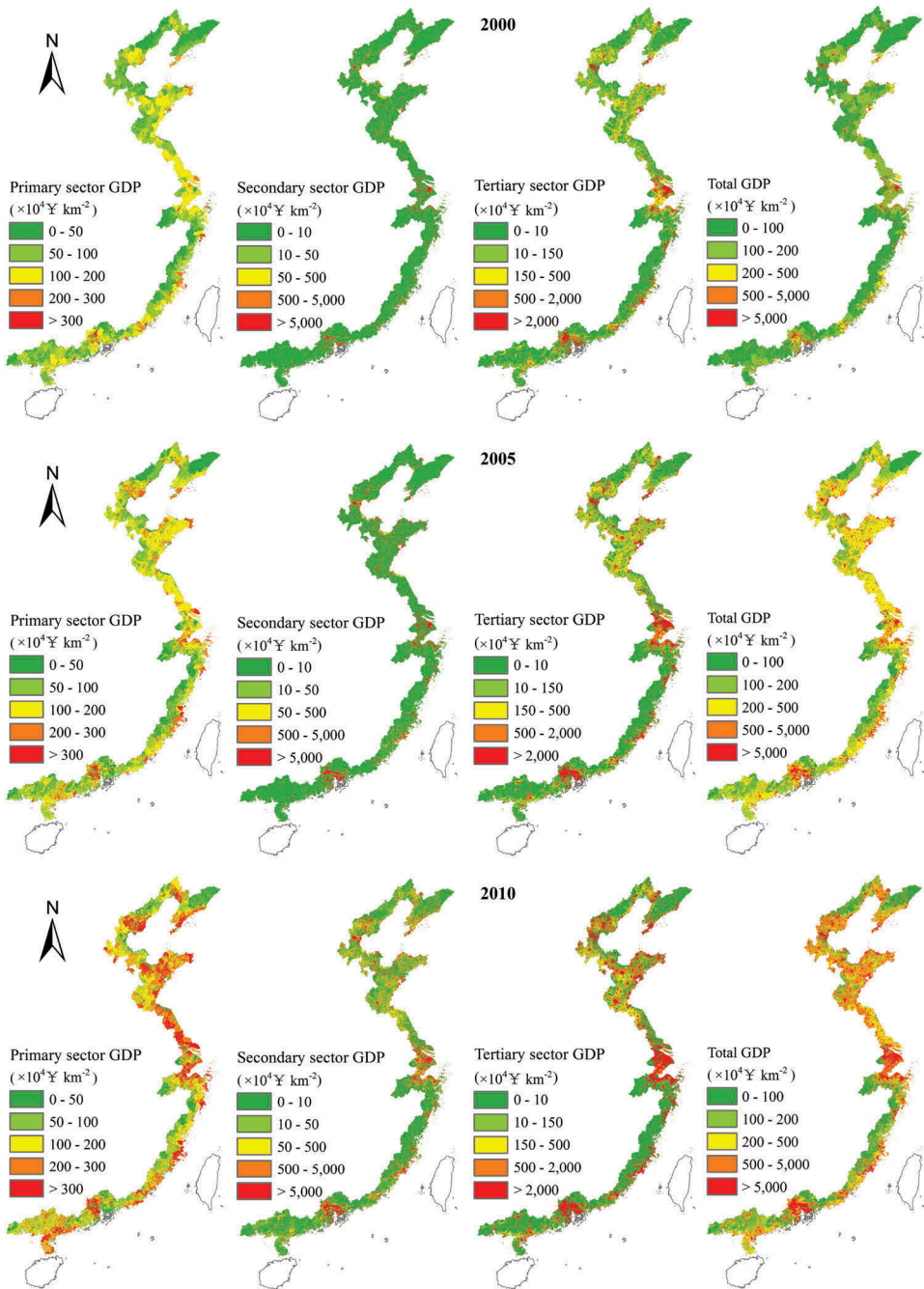


Figure 3. Spatialization map of the product of three sectors and GDP.

positive outliers of δ of the estimated GDP_2 are extraordinarily far away from the others, and most δ of the estimated GDP_3 are above 0; in other words, the regression model based on night-time light map is inclined to overestimate rather than underestimate. On the other hand, when using the dynamic regionalization method, the δ of the final

Table 1. Statistics for the interval distribution of $|\delta|$ of the estimated production of three industrial sectors.

Industrial sector	Year	Number of counties in different intervals of distribution of $ \delta $				Standard deviation of $ \delta $
		0%–10%	10%–20%	20%–40%	>40%	
Primary	2000	164	114	11	12	0.18
	2005	183	94	15	9	0.16
	2010	178	95	15	13	0.18
Secondary	2000	169	101	21	10	0.17
	2005	189	81	23	8	0.14
	2010	190	87	17	7	0.17
Tertiary	2000	155	122	20	4	0.15
	2005	186	98	13	4	0.14
	2010	152	126	17	6	0.14

estimated products have dropped a lot and the values of outliers are far less than those in the initial models.

5. Discussion and conclusions

Based on land-use data and night-time light image, we built regression models for estimating the GDP of three economy sectors using an improved method called dynamic regionalization. The regression equations were used to obtain the estimates of GDP in the years 2000, 2005, and 2010 at 1 km grid square of China's continental coastal area. Then the accuracy of these estimated products was assessed by comparing the estimated values aggregated at the county level with the original numbers reported in county statistical yearbooks. The main conclusions are as follows.

- (1) Building regression models for three sectors' product rather than GDP as a whole can demonstrate the spatial distribution feature of different economy fractions. Our results demonstrate that the primary sector correlates closely with farmland, forest, grassland, inland freshwaters, and mariculture, and the secondary sector has high correlations with city, rural settlement, isolated industrial-mining site, and saltern, and the tertiary sector can be related to night-time light data.
- (2) Night-time light data is considered as one of the best data sources to estimate sub-national GDP distribution, but its 'overflow' effect can cause significant error in application. Setting a threshold to draw out a specific area such as urban and suburban areas is a common but subjective method to address the 'overflow' effect. In this study, we use the land-use data to extract the city, rural settlement, and isolated industrial-mining site as the distribution plots of the tertiary sector, which can effectively alleviate the disturbance coming from the 'overflow' effect of night-time light data in a more objective manner.
- (3) The spatial heterogeneity of GDP distribution in China's continental coastal area is prominent, including variation between the North and the South, urban and suburban area, etc. Hence one regression model for the whole entire area would obtain estimated products with poor accuracy, particularly in counties whose GDP density is rather high or low. This study adopted the dynamic regionalization method, which includes building regression models for each sub-area that is divided dynamically from the study area. The essence of this

method is to divide the entire area, which has much spatially heterogeneous GDP distribution, into several sub-areas with somewhat homogeneous distribution of GDP internally. It can solve the problem of spatial heterogeneity simply, as well as assuage the adverseness of 'pixel saturation' in night-time light data.

- (4) With the dynamic regionalization method, we found that separate regression models were usually required for the highly developed and developing areas. We further consider dividing the whole study area into several sub-areas at first based on reducing the spatial discrepancy in sub-areas. The challenge is dividing the study area into several sub-areas, and simultaneously meeting these two requirements: (1) ensuring each sub-area has weak spatial discrepancy internally; and (2) minimizing the number of insignificant sub-areas.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Belmont Forum Collaborative Research Action on Scenarios of biodiversity and ecosystem services, NSFC-BF/IGFA, No.31461143032), and the Key Research Programme of the Chinese Academy of Sciences (Grant No. KZZD-EW-14 and KZZD-EW-TZ-15). We are grateful to Prof Hua Zhang for his helpful comments, suggestions, and language polish.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Natural Science Foundation of China [31461143032]; Key Research Programme of the Chinese Academy of Sciences [KZZD-EW-14]; [KZZD-EW-TZ-15].

References

- Azar, D., J. Graesser, R. Engstrom, J. Comenetz, R. M. Leddy, N. G. Schechtman, and T. Andrews. 2010. "Spatial Refinement of Census Population Distribution Using Remotely Sensed Estimates of Impervious Surfaces in Haiti." *International Journal of Remote Sensing* 31: 5635–5655. doi:10.1080/01431161.2010.496799.
- Di, X., X. Hou, Y. Wang, and L. Wu. 2015. "Spatial-Temporal Characteristics of Land Use Intensity of Coastal Zone in China during 2000–2010." *Chinese Geographical Science* 25: 51–61. doi:10.1007/s11769-014-0707-0.
- Di, X., X. Hou, and L. Wu. 2014. "Land Use Classification System for China's Coastal Zone Based on Remote Sensing." *Resources Science* 36: 463–472. (in Chinese)
- Doll, C. N. H., J. P. Muller, and C. D. Elvidge. 2000. "Nighttime Imagery as a Tool for Global Mapping of Socioeconomic Parameters and Greenhouse Gas Emissions." *AMBIO: A Journal of the Human Environment* 29: 157–162. doi:10.1579/0044-7447-29.3.157.
- Doll, C. N. H., J. P. Muller, and J. G. Morley. 2006. "Mapping Regional Economic Activity from Night-Time Light Satellite Imagery." *Ecological Economics* 57: 75–92. doi:10.1016/j.ecolecon.2005.03.007.
- Elvidge, C. D., K. E. Baugh, E. A. Kihn, H. W. Kroehl, E. R. Davis, and C. W. Davis. 1997. "Relation between Satellite Observed Visible-Near Infrared Emissions, Population, Economic Activity and

- Electric Power Consumption." *International Journal of Remote Sensing* 18: 1373–1379. doi:10.1080/014311697218485.
- Elvidge, C. D., V. R. Hobson, I. L. Nelson, J. M. Safran, B. T. Tuttle, K. E. Baugh, and J. B. Dietz. 2002. "Global Observations of Urban Areas Based on Nocturnal Lighting." *LUCC Newsletter* 8: 10–12.
- Elvidge, C. D., P. C. Sutton, T. Ghosh, B. T. Tuttle, K. E. Baugh, B. Bhaduri, and E. Bright. 2009. "A Global Poverty Map Derived from Satellite Data." *Computers & Geosciences* 35: 1652–1660. doi:10.1016/j.cageo.2009.01.009.
- Ghosh, T., R. Powell, S. Anderson, P. C. Sutton, and C. D. Elvidge. 2010a. "Informal Economy and Remittance Estimates of India Using Nighttime Imagery." *International Journal of Ecological & Statistics* 17: 16–50.
- Ghosh, T., R. Powell, C. D. Elvidge, K. E. Baugh, P. C. Sutton, and S. Anderson. 2010b. "Shedding Light on the Global Distribution of Economic Activity." *The Open Geography Journal* 3: 147–160. doi:10.2174/1874923201003010147.
- Han, X., Y. Zhou, S. Wang, R. Liu, and Y. Yao. 2012. "GDP Spatialization in China Based on DMSP/OLS Data and Land Use Data." *Remote Sensing Technology and Application* 27: 396–405. (in Chinese)
- Harvey, J. T. 2002. "Estimating Census District Populations from Satellite Imagery: Some Approaches and Limitations." *International Journal of Remote Sensing* 23: 2071–2095. doi:10.1080/01431160110075901.
- Hu, Y., Q. Wang, Y. Liu, J. Li, and W. Ren. 2011. "Index System and Transferring Methods to Build the National Society and Economy Grid Database." *Journal of Geo-information Science* 13: 573–578. (in Chinese). doi:10.3724/SP.J.1047.2011.00573.
- Li, G., and C. Fang. 2014. "Global Mapping and Estimation of Ecosystem Services Values and Gross Domestic Product: A Spatially Explicit Integration of National 'Green GDP' Accounting." *Ecological Indicators* 46: 293–314. doi:10.1016/j.ecolind.2014.05.020.
- Liang, Y., and Z. Xu. 2013. "Modeling the Spatial Distribution of GDP Based on Night Light Radiation: A Case Study in Ganzhou District, Zhangye Municipality." *Journal of Glaciology and Geocryology* 35: 249–254. (in Chinese)
- Lo, C. P. 2002. "Urban Indicators of China from Radiance Calibrated Digital DMSP-OLS Nighttime Images." *Annals of the Association of American Geographers* 92: 225–240. doi:10.1111/1467-8306.00288.
- Sullivan, A., and S. M. Sheffrin. 1996. *Economics: Principles in Action*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Sutton, P. C., S. J. Anderson, B. T. Tuttle, and L. Morse. 2012. "The Real Wealth of Nations: Mapping and Monetizing the Human Ecological Footprint." *Ecological Indicators* 16: 11–22. doi:10.1016/j.ecolind.2011.03.008.
- Sutton, P. C., C. D. Elvidge, K. Baugh, and D. Ziskin. 2011. "Mapping the Constructed Surface Area Density for China." *Proceedings of the Asia-Pacific Advanced Network* 31: 69–78. doi:10.7125/APAN.31.
- Sutton, P. C., T. Ghosh, and C. D. Elvidge. 2007. "Estimation of Gross Domestic Product at Sub-National Scales Using Nighttime Satellite Imagery." *International Journal of Ecological Economics & Statistics* 8: 5–21.
- Townsend, A. C., and D. A. Bruce. 2010. "The Use of Night-Time Lights Satellite Imagery as a Measure of Australia's Regional Electricity Consumption and Population Distribution." *International Journal of Remote Sensing* 31: 4459–4480. doi:10.1080/01431160903261005.
- Wu, J., Z. Wang, W. Li, and J. Peng. 2013. "Exploring Factors Affecting the Relationship between Light Consumption and GDP Based on DMSP/OLS Nighttime Satellite Imagery." *Remote Sensing of Environment* 134: 111–119. doi:10.1016/j.rse.2013.03.001.
- Zhang, X., C. Ma, S. Zhan, and W. Chen. 2012. "Evaluation and Simulation for Ecological Risk Based on Emery Analysis and Pressure-State-Response Model in a Coastal City, China." *Procedia Environmental Sciences* 13: 221–231. doi:10.1016/j.proenv.2012.01.021.
- Zhao, N., N. Currit, and E. Samson. 2011. "Net Primary Production and Gross Domestic Product in China Derived from Satellite Imagery." *Ecological Economics* 70: 921–928. doi:10.1016/j.ecolecon.2010.12.023.