Applied Geography 90 (2018) 239–247

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

# Modeling population density based on nighttime light images and land use data in China

Minghong Tan[a,b,*], Xiubin Li[a,c], Shiji Li[a,c], Liangjie Xin[a], Xue Wang[a], Qian Li[d], Wei Li[a,c], Yuanyuan Li[a,c], Wenli Xiang[a,c]

[a] Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research (IGSNRR), Chinese Academy of Sciences (CAS), Beijing, 100101, China
[b] International College, University of Chinese Academy of Sciences, Beijing, 100190, China
[c] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100190, China
[d] Business School, Shandong University, Weihai, 264209, China

ARTICLE INFO

*Keywords:*
Nighttime light images
Census data
Land use
Population density simulation
China

ABSTRACT

Population change is a key variable that influences climate change, ecological construction, soil and water use, and economic growth. Census data are always point data, whereas planar data are often required in scientific research. By using nighttime light (NTL) images and land use data, combined with the fifth and sixth census data of China at the county level, we carried out spatial matching on the population of each county, respectively, and established population density diagrams of China for 2000 and 2010, which had a spatial resolution of 1 × 1 km. The method proposed in this paper is relatively simple and has a high simulation precision. The results showed that during the first ten years of the 21st century, there are some remarkable characteristics in Chinese population spatial pattern change: 1) the "disappearance" of intermediate-density regions; namely, areas with a population density between 500 and 1500 persons/km$^2$ have decreased by 41% during the ten years; 2) continuous growth of high-density regions; namely, areas with a population density of more than 1500 persons/km$^2$ have increased by 76%; 3) an expansion tendency of low-density regions similar to high-density regions.

Since China's reform and opening up, the country has been entering into a rapid urbanization phase and experienced the largest rural-urban population migration in human history (Zhang & Song, 2003). From 1990 to 2010, the urban population increased from 302 to 671 million; at the same time, the rural population decreased from 841 to 670 million (NBSC, 2013). Urban-rural population migration has greatly changed the production and life styles of humans and exerts a significant impact on the natural environment and the social economy.

Until now, China has carried out six censuses, which provide a solid data foundation to comprehend China's population status and the relationship between humans and their natural environment (Li, Sun, Tan, & Li, 2016). However, census data are usually obtained based on administrative units; however, administrative units at the same hierarchy of China usually show significant differences in their administrative areas in different regions. For instance, some counties have an administrative area of less than 100 km$^2$, while others have a large area (e.g., the area of Ruoqiang County in Xinjiang covers 202,000 km$^2$). Moreover, even in the same county, due to differences in geographical environment and economic development level, population distribution is also very imbalanced. In addition, most studies do not use the administrative unit as the research object (Ryan et al., 2017; Tan et al., 2008; Tritsch & Le Tourneau, 2016; Yang, Jiang, Luo, & Zheng, 2012), but rather consider the basin or geographic unit. Therefore, previous studies conducted spatial simulation on the population distribution using different methods (Anderson, Tuttle, Powell, & Sutton, 2010; Bagan & Yamagata, 2015; Liu et al., 2003b; Lung, Lübker, Ngochoch, & Schaab, 2013; Tian, Yue, Zhu, & Clinton, 2005). For example, Liu et al. (2003b) simulated China's population density by combining digital elevation data, net primary productivity data, transportation infrastructure data, and urban scale data, based on the gravity model method.

Nighttime light data from Defense Meteorological Satellite Program's Operational Linescan System (DMSP/OLS) are important indicators that reflect human agglomeration and energy consumption, which provide a stable and persistent data source (Elvidge et al., 2001; Hsu, Baugh, Ghosh, Zhizhin, & Elvidge, 2015). Currently, these data are widely used to evaluate urban growth (Imhoff, Lawrence, Stutter, & Elvidget, 1997; Lu, Tian, Zhou, & Ge, 2008; Ma, Zhou, Zhou, Haynie,

* Corresponding author. Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China.
*E-mail address:* tanmh@igsnrr.ac.cn (M. Tan).

https://doi.org/10.1016/j.apgeog.2017.12.012
Received 2 January 2017; Received in revised form 13 December 2017; Accepted 13 December 2017
0143-6228/ 

**Table 1**
Comparison of previous studies related to population density simulation in China.

| Authors | Models | Data used | Simulated year | Data format of density simulation |
|---|---|---|---|---|
| Wang, Feng, Yang, and You (2014) | | Census, land area at the county level | 2000, 2010 | Average density at the county level |
| Tian et al. (2005) | Regression method, buffer analysis et al. | Population, road, railway, river, city distribution, rural population, urban population, land use map | 2000 | Grid |
| Liu et al. (2003b) | Gravity model | Data of climate, soil, remote sensing, way, railway, elevation et al. | 2000 | Grid |
| Zhuo et al. (2005) | Regression and allometric models | NTL images, NDVI, non-agricultural population, population | 1998 | Grid |
| Lo (2001) | Regression method | NTL images, non-agricultural population, rural population | 1997 | Average density at the county level |
| Zeng et al. (2011) | Regression method | NTL intensity, land use | 2000 | Grid |
| Dobson, Brlght, Coleman, Durfee, and Worley (2000) | Gravity model et al. | Road proximity, slope, land cover, and nighttime lights. | 1998 | Grid |

Pei, & Xu, 2015; Pandey, Joshi, & Seto, 2013; Shao & Liu, 2014; Small, Pozzi, & Elvidge, 2005; Su et al., 2015; Wei, Liu, Song, Yu, & Xiu, 2014; Xiao, Wang, Feng, Zhang, & Yang, 2014; Yi et al., 2014; Yue, Zhang, & Liu, 2016; Zhang & Su, 2016; Zhou et al., 2014), economic development (Chen & Nordhausa, 2011; Elvidge et al., 2009; Keola, Andersson, & Hall, 2015; Levin & Duke, 2012; Propastin & Kappas, 2012; Wu, Wang, Li, & Peng, 2013), and urban transition (Saksena et al., 2014; Tan, 2015). Some scholars have analyzed the relationship between population distribution and light intensity in different regions, showing a significant correlation between them (Anderson et al., 2010; Bagan & Yamagata, 2015). In addition, some authors have conducted spatial matching on the demographic data by using light image data (Amaral, Câmara, Monteiro, Quintanilha, & Elvidge, 2005; Sutton, 1997). For example, Zhuo et al. (2005) conducted simulation on the population distribution of China in 1998 by using vegetation coverage data, light image data, and non-agricultural population and rural population data, and obtained good simulation results.

To sum up, the existing researches related to population density simulation are relatively complex and involve various and abundant data sets (Table 1). These complex simulation processes may reduce the spatial matching accuracy of population distribution. In addition, such studies usually divide a country into several areas (Zeng, Zhou, Wang, Yan, & Zhao, 2011) to establish the relationship between population density and light intensity, and then conduct spatial simulation on the population density at the regional scale. Furthermore, until now, the spatial simulation has been ignored based on the fifth (2000) and sixth census (2010) data of China, using the same method, and comparison of spatial changes in population density is also not enough, especially at the grid level across the country.

In this paper, we used a simpler and more direct simulation model to conduct spatial matching and simulation on the demographic data for each county in China, respectively. Specifically, based on the fifth and sixth census data, light image and land use data, we analyzed the relationships between population density and light intensity, cultivated land distribution, and settlement distribution at the county level after removing unused land and water areas (Fig. 1). On the basis of these relationships, we conducted spatial matching on the demographic data of each county, respectively, and obtained spatial distribution diagrams (with a resolution of 1 × 1 km) of population density in 2000 and 2010, which were then used to analyze the variation of population density. The simulated population density data can be used as the input data to explore the influences of population changes on the natural environment, ecological footprint, and climate change.

## 1. Data sources

Land use data in 2000 were obtained from the land use database with a scale of 1:100,000, provided by the Resources and Environment Data Center, CAS. The data were manually interpreted mainly based on the Landsat TM data. A large amount of field investigation proved that land use classification had an accuracy of more than 90% (Liu, Liu, Deng, Zhuang, & Zhang, 2002, 2003a).

Nighttime light data were obtained from the time series data of DMSP/OLS (http://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html accessed October 15, 2014). The NTL images used in this study in 2000 and 2010 were obtained from satellites F14 and F18, respectively. The data were in the grid format, with grid values between 0 and 63 (Keola et al., 2015), in which the larger values represent higher light intensity. Across China, the grid values varied between 1 and 63. Since the light image data were obtained from different satellites, in order to conduct comparisons for time series data, it is necessary to calibrate these data to endow the light image data in different years with comparability. In this study, on the basis of respectively simulating the relationships among light image intensity (DN value), residential land proportion, and population density in 2000 and 2010, we carried out spatial matching on the census data. Since time series comparison for the light image intensity was not conducted here, we did not calibrate the light image data in different years.

Demographic data in 2000 and 2010 were obtained from the database of the fifth and sixth census at the county level. Demographic data in the 2005 were from China County Statistical Yearbook in 2006 (NBSC, 2006). The demographic data used in this study were the data of permanent population defined by the census (NBSC, 2002; 2012). According to the definition, migrants were recorded at their current abode rather than at their place of registration if they lived at the abode for over 6 months/year. These migrants are regarded as permanent population.

## 2. Methods

According to China's land use map of the year 2000, water areas and unused land were removed as uninhabited areas and the population density values of these regions were set as 0. In this study, we firstly analyzed the relationships between the light images, cultivated land distribution, settlement distribution, and the population density at the county level. Subsequently, we conducted spatial matching and simulation on the census data at the county level. Finally, the simulation results were verified by the township-level census data.

As previously introduced, light images are closely related to human activities, which can reflect the population density to a large extent. Therefore, the premise of model simulation is to determine the quantitative relation between light image intensity and population density. Due to the lack of demographic data, it is extremely difficult to describe the quantitative relationship at the grid level. Tian et al. (2005) analyzed the relationship between the population and different land use