

## Exploring the spatial differentiation of urbanization on two sides of the Hu Huanyong Line – based on nighttime light data and cellular automata



Dongsheng Chen<sup>a,b</sup>, Yatao Zhang<sup>b</sup>, Yao Yao<sup>a,\*</sup>, Ye Hong<sup>c</sup>, Qingfeng Guan<sup>a</sup>, Wei Tu<sup>d,e</sup>

<sup>a</sup> School of Geography and Information Engineering, China University of Geosciences, Wuhan, 430078, Hubei province, China

<sup>b</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, 430079, Hubei province, China

<sup>c</sup> Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, Stefano-Frascini-Platz 5, CH-8093, Zürich, Switzerland

<sup>d</sup> Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, 518060, China

<sup>e</sup> Guangdong Key Laboratory of Urban Informatics, Shenzhen Key Laboratory of Spatial Smart Sensing and Services and Research Institute of Smart Cities, Shenzhen University, Shenzhen, 518060, China

### ARTICLE INFO

#### Keywords:

Urbanization  
Hu huanyong line  
Nighttime light  
Cellular automata  
Belt and road initiative

### ABSTRACT

The Hu Huanyong Line (Hu Line) shows an uneven geographical pattern of urbanization between Southeast and Northwest China. Based on three questions proposed by the Chinese Premier, the balance of urbanization on both sides of the Hu Line and the possibility of breaking through the pattern in the future has provoked substantial interest in academic circles. Based on quantitative analyses and urban simulations, this study proposes a large-scene urbanization quantitative spatial analysis framework via nighttime light and random forests-based cellular automata and explores changes in the urbanization of the Hu Line pattern in the context of the Belt and Road Initiative. This study analyzes the past urbanization status on both sides of the Hu Line and predicts the future distribution and proportion of urban land on both sides of the Hu Line in the context of the Belt and Road Initiative. The new impacts that the Belt and Road Initiative may bring to the domestic regions are concretized and coupled with cellular automata. The results show that the proportion of urban land on both sides of the Hu Line was 93.94:6.06 in 2005 and that it will reach 93.07:6.93 in 2035. A significant difference in urbanization is revealed between the two sides; however, the difference decreases annually. Overall, the distribution of urbanization in China will retain the southeast-northwest pattern of the Hu Line in 2015–2035.

### 1. Introduction

The urbanization level is an important indicator that reflects the social and economic development of a country or a region, and it is closely related to various factors, such as the population composition, economic development, land resources, and social welfare system of the country or the region (Chen, Dadao, & Zhang, 2009). The definition of urbanization in geography is more focused on the process of change in geospatial space (Antrop, 2000). Urbanization is a global concern, especially for developing countries, such as Southeast Asia and South Asia (Dewan & Yamaguchi, 2009; Xu et al., 2019). Since the reform and opening up in 1978, China has been in a rapid phase of urbanization. Studying urbanization will not only provide insights into the dynamic process and spatial pattern of urbanization in China but also help to analyze and regulate the driving forces of urbanization and their mechanisms of action. The results provide an important basis for governments to formulate scientific and rational urban development policies

and urban planning (Gu, 1999; Shi, Chen, & Pan, 2000; Tu et al 2017, 2018; Yan, Mao, & Pu, 2006).

In China, an uneven geographical pattern of urbanization has existed for a long time (Chen, Gong, Li, Lu, & Zhang, 2016; Hu, Wang, Liu, Long, & Peng, 2016; Qi, Liu, Zhao, & Liu, 2016; Zhang, Song, & Zhang, 2015). This pattern is characterized by a dense population and rapid urbanization in the eastern area and sparse population and slow urbanization in the western area (Chen et al., 2016). These two parts are divided by the Hu Line. The Hu Line, also known as the “Hu Huanyong Line”, is an imaginary line stretching from Heihe (a northern city of China located on the Russian border) to Tengchong (a southwestern city of China bordering with Myanmar), which divides the area of China into two roughly equal parts (Hu, 1935). The western part of China contains many important ecological functional areas and areas with harsh environments (such as alpine areas and desert regions), where cities are hard to develop (Zhang et al., 2018). Therefore, the eastern area of the Hu Line has become the main carrier of urbanization (Wang

\* Corresponding author. School of Geography and Information Engineering, China University of Geosciences, Wuhan, 430078, Hubei province, China.

E-mail addresses: [dongsingchen@whu.edu.cn](mailto:dongsingchen@whu.edu.cn) (D. Chen), [yatao@foxmail.com](mailto:yatao@foxmail.com) (Y. Zhang), [yaoy@cug.edu.cn](mailto:yaoy@cug.edu.cn) (Y. Yao), [hongy@student.ethz.ch](mailto:hongy@student.ethz.ch) (Y. Hong), [guanqf@cug.edu.cn](mailto:guanqf@cug.edu.cn) (Q. Guan), [tuwei@szu.edu.cn](mailto:tuwei@szu.edu.cn) (W. Tu).

<https://doi.org/10.1016/j.apgeog.2019.102081>

Received 25 November 2018; Received in revised form 15 August 2019; Accepted 3 September 2019

0143-6228/ © 2019 Elsevier Ltd. All rights reserved.

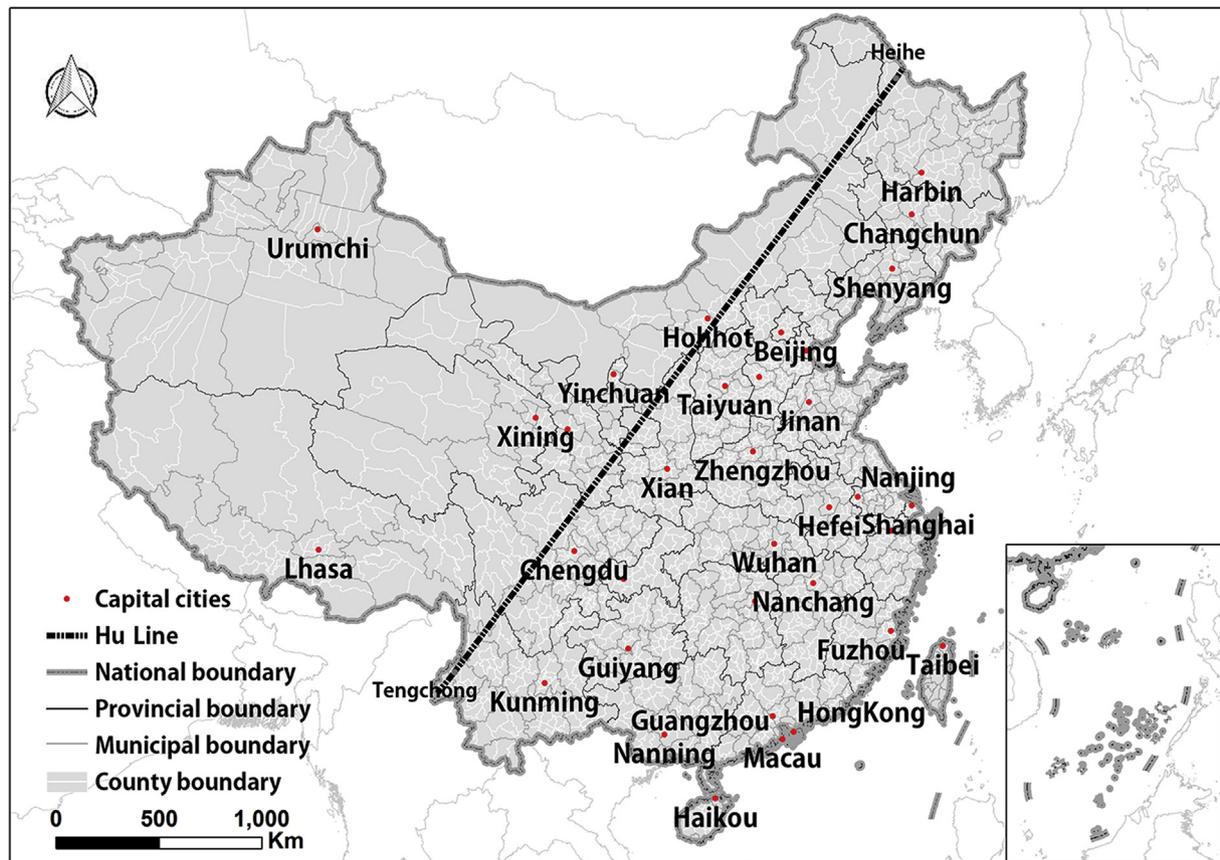


Fig. 1. Administrative boundaries of China at different levels.

& Deng, 2016). This traditional uneven pattern of urbanization is called the Hu Line pattern (Li, 2015; Zhang et al., 2015). A considerable number of previous studies have shown that the Hu Line is China's natural and ecological boundary (Chen et al., 2016; Hu et al., 2016; Qi et al., 2016; Zhang et al., 2015).

In November 2014, Premier Li Keqiang pointed to the Hu Line and stated that this traditional uneven pattern needs to be broken and the central and western areas need to develop to benefit the people in those regions. Three questions about the Hu Line were proposed: "Should we break the line?" "Can we?" And "How?" (Chen et al., 2016; The State Council 2014). These questions are of great importance for advancing China's urbanization and thus have provoked substantial interest within academic circles (Chen et al., 2016). A representative view is that the Hu Line pattern cannot be broken and the uneven pattern is determined by natural and geographical conditions. In the foreseeable future, this population distribution law cannot be changed (Jia, 2014). Another commonly held view is that this pattern is breakable since the new-type of urbanization in China provides an opportunity based on the trends of suburbanization and counterurbanization (Li, 2015).

However, few studies have quantitatively analyzed the spatial differentiation of urbanization on the two sides of the Hu Line and quantitatively simulate future scenarios. Some research has quantitatively explored the relationship between the population distribution in China and the Hu Line pattern (Chen et al., 2016; Ge & Feng, 2010; Qi et al. 2015, 2016; Wu & Wang, 2008), whereas other studies have focused on qualitatively revealing the trend of urbanization on the two sides of the Hu Line and predicting changes of the Hu Line (Chen et al., 2016; Deng & Bai, 2014; Fan, Liu, & Chen, 2013; Li, 2015; Wang & Deng, 2016; Zhang et al., 2015). Only a small number of studies has focused on analyzing China's urbanization on the two sides of the Hu Line quantitatively at a coarse scale without future scenario simulation (Chen, Lu, & Liu, 2010; Hu et al., 2016; Yang, 2014). Thus, a proper

large-scene urbanization quantitative spatial analysis framework is needed.

In addition, some arrangements of national policy will create new development opportunities and help break the Hu Line pattern, such as the "Silk Road Economic Belt" and the "21st-Century Maritime Silk Road", which is called the Belt and Road Initiative for short (National Development & Reform Commission, 2015). These developments will help improve resource rent and location advantages in the western part of the Hu Line and break the Hu-line pattern (Kuang, 2014). This study explores the change in the urbanization of the Hu Line pattern in the context of the Belt and Road Initiative.

Due to the development of remote sensing and GIS technology, we use reliable means for the analysis and prediction of urbanization on the two sides of the Hu Line. NTL (nighttime light) data refer to data obtained by detecting the brightness of the night surface through relevant satellites. The light intensity reflects the temporal and spatial variation characteristics of human activities (Ma et al., 2012); thus, it has been used by many scholars for socioeconomic geography research, such as humanitarian crises monitoring, population density simulation, and international economic and trade activities (Elvidge et al. 2007, 2010; Li & Li, 2014; Zhao & Samson, 2012; Zhuo et al., 2009).

Regarding the study of urbanization, a large number of studies have also explored the correlation between the time series of NTL data and the urbanization level, and they have shown the feasibility of using NTL data for urbanization research (Ma et al. 2012, 2015; Zhang & Seto, 2011).

Previous studies have pointed out that dynamic urban expansion can be modeled using cellular automata (CA) (Li and Yeh 2000, 2002, pp. 131–152). CA is a discrete grid model that reflects the overall pattern based on the transition rules of the cell state. In the CA model, all cells follow the transition rules, and the states of the cells are changed to simulate the dynamic process, and the results reveal

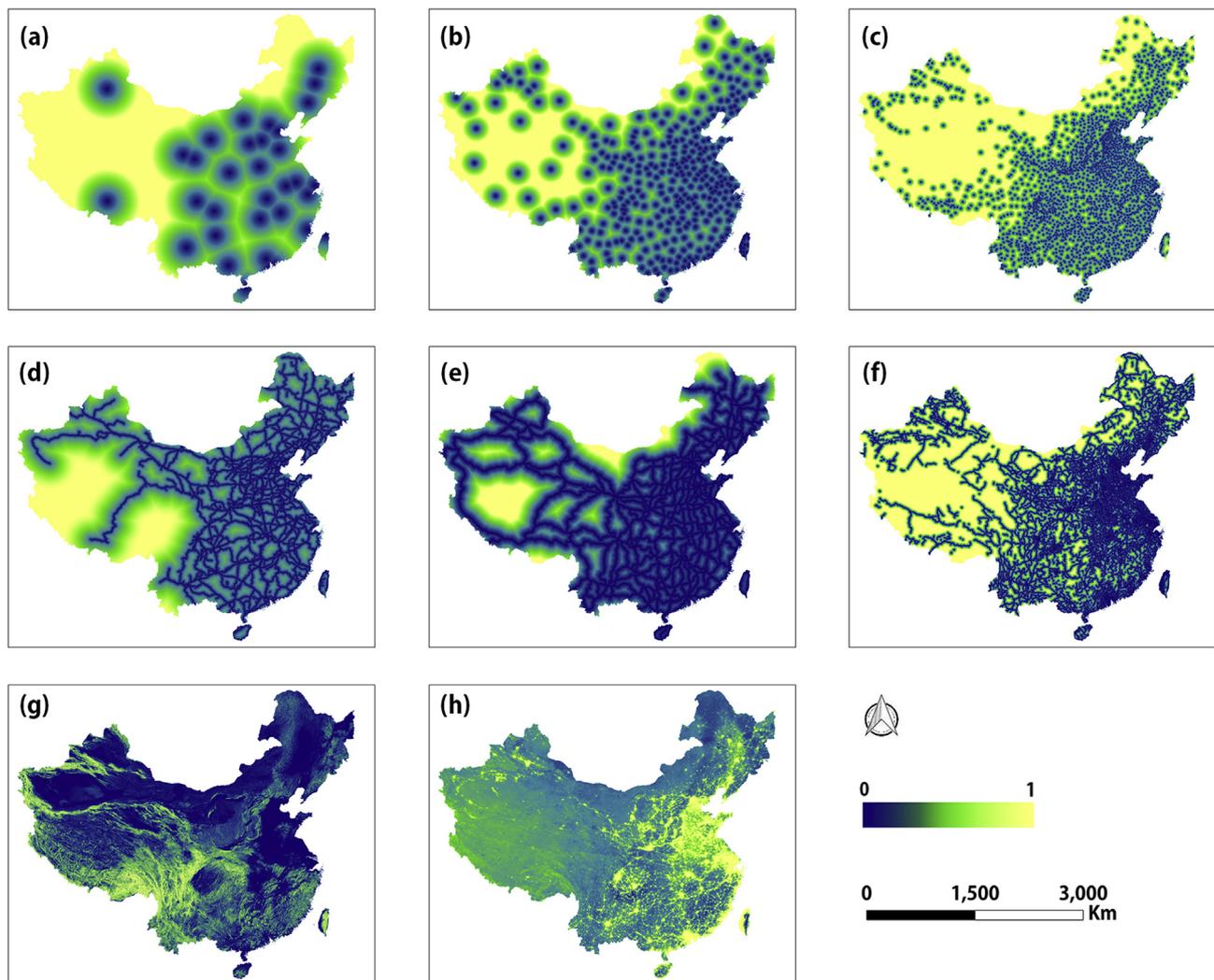


Fig. 2. The spatial variables in mainland China. (a) The distance to the capital cities; (b) the distance to the administrative centers of cities; (c) the distance to the administrative centers of counties; (d) the distance to the major railways; (e) the distance to the main roads; (f) the distance to other roads; (g) terrain slope; (h) nighttime lights.

complex urban expansion (Li et al. 2017a). CA is widely used for urban expansion analysis and has been applied in a large number of studies, both at home (Han et al. 2009; Li and Yeh 2000, 2002, pp. 131–152; Shan, 2008; Yang et al. 2013) and abroad (Arsanjani et al. 2013; Moghadam & Helbich, 2013).

In the context of the Belt and Road (B&R) Initiative, this study focuses on exploring the change in the urbanization of the Hu Line pattern. Based on NTL data and the CA model, we propose a large-scene urbanization quantitative spatial analysis framework to analyze and predict China's urbanization along the Hu Line pattern. The new impacts that the Belt and Road Initiative may bring to domestic regions are concretized and coupled with the CA. This study first analyzes the spatial differentiation of urbanization on the two sides of the Hu Line from 2005 to 2015 and then predicts the future spatial distribution of urbanization in 2015–2035. Based on the results, we aim to identify the trend of urban growth in China and offer new ideas for urban planning by local governments.

## 2. Study area and data

This study selects China as the study area (Fig. 1). The study time ranges from 2005 to 2015 with a time step of 1 year. Since the reform and opening up, China's economy has developed rapidly, and

considerable urban expansion has occurred (Zhang, Wang, & Wang, 2012). The urban built-up area has expanded from 7438 km<sup>2</sup> in 1981 to 32,520.7 km<sup>2</sup> in 2005 (Fang, 2009). The change in urban expansion is significant.

For NTL data, this study uses annual stable light data from the Version 4 DMSP-OLS Nighttime Lights Time Series and the monthly cloud-free product of the Version 1 Suite of Average Radiance Composite Images of VIIRS/DNB (<http://ngdc.noaa.gov/eog>). In the remainder of this study, we use “DMSP data” and “VIIRS data” to represent the two datasets. DMSP data are cloud-free composites with a spatial resolution of 0.008333° and a time step of 1 year, and they span the period from 1992 to 2013. VIIRS data have undergone cloud mask processing, onboard calibration and radiation correction, and they have a spatial resolution of 0.004167° and a time resolution of 1 month. The dataset has been released continuously since April 2012 (Li et al. 2017b; Li et al. 2017c; Li & Zhou, 2017).

DMSP data, whose light intensity and distribution reveal human activities and spatial distribution, represent an important data source for large-scale urbanization research and are widely used in urban studies (Liu et al. 2012; Sutton, Taylor, & Elvidge, 2010; Zhou et al. 2014; Yao et al., 2018). Compared to the DMSP sensor, the VIIRS sensor has three advantages: onboard calibration, higher spatial resolution, and higher radiometric resolution (Li & Zhou, 2017). VIIRS data are a

new generation of NTL data with great research potential (Li et al. 2017b). This study uses land-use data and gas flare coverage to build the mask that eliminates nonurban light sources in the NTL datasets since gas flares and reflective lights on the water surface are nonurban light sources (Li et al. 2017c). In addition, this study averages all monthly products over one year to obtain an annual average product of NPP/VIIRS and remove the outliers caused by the monthly fluctuations.

The update of the DMSP/OLS dataset was discontinued in 2013, and no further years of research have been possible. Since the NPP/VIIRS dataset has been updated since 2012, it can compensate for the shortcomings of the DMSP/OLS dataset (Li et al. 2017b; Li et al. 2017c; Liu et al. 2012). However, there is a difference between the sensors of DMSP/OLS and NPP/VIIRS, and a model needs to be conducted to intercalibrate these two data sets to extend the coverage time of NTL data.

Land-use and land-cover (LULC) data are also important for this study. The national land-use data come from the Institute of Remote Sensing Applications of the Chinese Academy of Sciences and were obtained by manually interpreting Landsat/TM images as well as simultaneously undergoing a uniform quality check (Liu et al. 2010, 2014). To simplify the model, the LULC data are divided into three custom categories, including urban area, nonurban area and limited development area, and the corresponding nine kinds of land-use transformation relationships are generated. Among them, limited development areas are a component of nonurban areas. Limited development areas refer to areas whose probability of urbanization is much lower than that of the other nonurban area because of geographical conditions, and they include wetland, waterbody, tundra and permanent snow/ice-covered areas (Lin & Li, 2014; Yao et al., 2017a).

The transformation of land use is often influenced by a series of driving factors, combinations of adjacent land-use types, and the natural properties of the cells (Batty & Xie, 1994). This study selects eight spatial variables as driving factors for urban land transformation as shown in Fig. 2. Poston and Yaukey (2013) argued that the level of urbanization is directly related to the distance to the transportation network and the distance to important cities, and its effect is particularly pronounced in developing countries (Poston & Yaukey, 2013). The slope of the terrain and the nighttime lights reveal the natural environment and socioeconomic conditions.

Based on the existing rail lines, this study manually drew a rough distribution map of the major planning rail lines, the key rail links, the seaports and the ports of entry (POEs) in China mentioned in the B&R Initiative (Fig. 3) according to the government documents (<http://eng.yidaiyilu.gov.cn>) and the railway planning from the National Development and Reform Commission (<http://en.ndrc.gov.cn/>).

The Silk Road Economic Belt represents 6 corridors for connectivity infrastructure on land. The 21st Century Maritime Silk Road is the ocean route corridor. This road passes through a series of coastal cities in eastern China and travels to the south Pacific, Indian Ocean and Europe.

### 3. Methodology

To explore urbanization in China's past and future, this study will use the following three steps (Fig. 4). 1) We conduct a temporal calibration on the DMSP data and simulate the DSMP data using the VIIRS data, thereby extending the time span of the NTL dataset. 2) Based on the NTL dataset, the time series of the weighted light area (WLA) is constructed and the urbanization level and the urbanization pattern are analyzed via the temporal clustering model and the pattern classification model. 3) Based on the LULC data, we simulate the urbanization process and analyze the future situation using the CA model.

#### 3.1. NTL data preprocessing

Due to the lack of an onboard calibration system for OLS, the time series of DMSP data lacks temporal consistency; thus, a temporal

calibration is required. The second-order regression model based on IR (invariant region) proposed by Elvidge et al. (2009a, b) is currently the most widely used DMSP dataset for temporal analysis (Elvidge et al. (2009a, b); Elvidge et al. (2009a, b); Huang, Schneider, & Friedl, 2016; Li et al. 2016; Zhang, Pandey, & Seto, 2016). This simple and effective method is implemented for DMSP data in this study (Equation (A.1)). At the same time, this study selects Jixi City in Heilongjiang Province, China, which has a stable social and economic status and has not been highly developed over the years, as the invariant region (Liu et al. 2012).

The time series of the DMSP/OLS dataset is conducted for the temporal calibration first since DMSP data lack temporal consistency. After the calibration, the average  $R^2$  of the results is above 0.900, which indicates that the regression model is well performed (Supp. Table 1). In addition, this study draws the temporal curves by summing the total values of all pixels before and after the calibration, which indicates that the consistency of the calibrated DMSP dataset is improved significantly.

To extend the NTL data coverage time, this study uses the intercalibration model to combine the calibrated DMSP/OLS data and the NPP/VIIRS dataset (Li et al. 2017c). This study takes the data of two datasets for the overlapped years (2012 and 2013) as training data to construct the intercalibration model. A nonlinear relationship exists between NPP/VIIRS and DMSP/OLS, which needs to be simulated using Supp. Equation (2) for VIIRS data. After adopting the urban mask and root mean square error (RMSE) (Chai et al. 2014; Wang et al. 2017), we use the global optimization search algorithm to optimize the parameters. In addition, to composite the NTL data for the overlapping years, this study uses the intra-annual composition method (Liu et al. 2012).

After the intercalibration, the VIIRS images are more similar to the calibrated DMSP images (Supp. Fig. 2). The correlation coefficient  $r$  and the RMSE of the intercalibration result are shown in Supp. Table 2. Moreover, to intuitively display the changes before and after intercalibration, this study compares the DN relationship between the DMSP data and the VIIRS data (Li et al. 2017c) (Supp. Fig. 3).

Finally, we conduct the intra-annual composition method on the two data sets and obtain the NTL composite product. The total values of all pixels of the final product are shown in Fig. 5.

#### 3.2. Spatial analysis of the level and pattern of urbanization

This study calculates the WLA of each administrative unit for each year from 2005 to 2015 at the county level and then constructs the time series of WLA. Ma et al. (2012) conducted a correlation analysis with statistical data, such as population, GDP, built-up area and power consumption, and the results indicate that a positive correlation occurs between WLA and the dynamic of urbanization (Ma et al. 2012). Therefore, WLA can be used to measure the urbanization level of the study area.

This study uses the dynamic time warping (DTW) algorithm to denote the similarity of the WLA time series (Petitjean, Ketterlin, & Gançarski, 2011). For any two given temporal sequences  $\mathbf{P} = \{p_1, p_2, \dots, p_m\}$  and  $\mathbf{Q} = \{q_1, q_2, \dots, q_n\}$ , their Euclidean distance matrix  $\mathbf{D}$  ( $m \times n$ ) can be described as follows:

$$d_{i,j} = (p_i - q_j)^2 \quad (1)$$

where  $d_{ij}$  denotes the element in the  $i$ -th row and the  $j$ -th column of the matrix  $\mathbf{D}$ .

In addition, the sequence of the warping paths  $\mathbf{W} = \{w_1, w_2, \dots, w_k\}$  between  $\mathbf{P}$  and  $\mathbf{Q}$  is built. The sequence  $\mathbf{W}$  must satisfy the following conditions. 1) The origin and destination of the warping paths in  $\mathbf{W}$  must be  $d_{1,1}$  and  $d_{m,n}$  in  $\mathbf{D}$ ; and 2) any two warping paths  $w_i$  and  $w_{i+1}$  in  $\mathbf{W}$  must be continuous and monotonic. The equation is as follows:

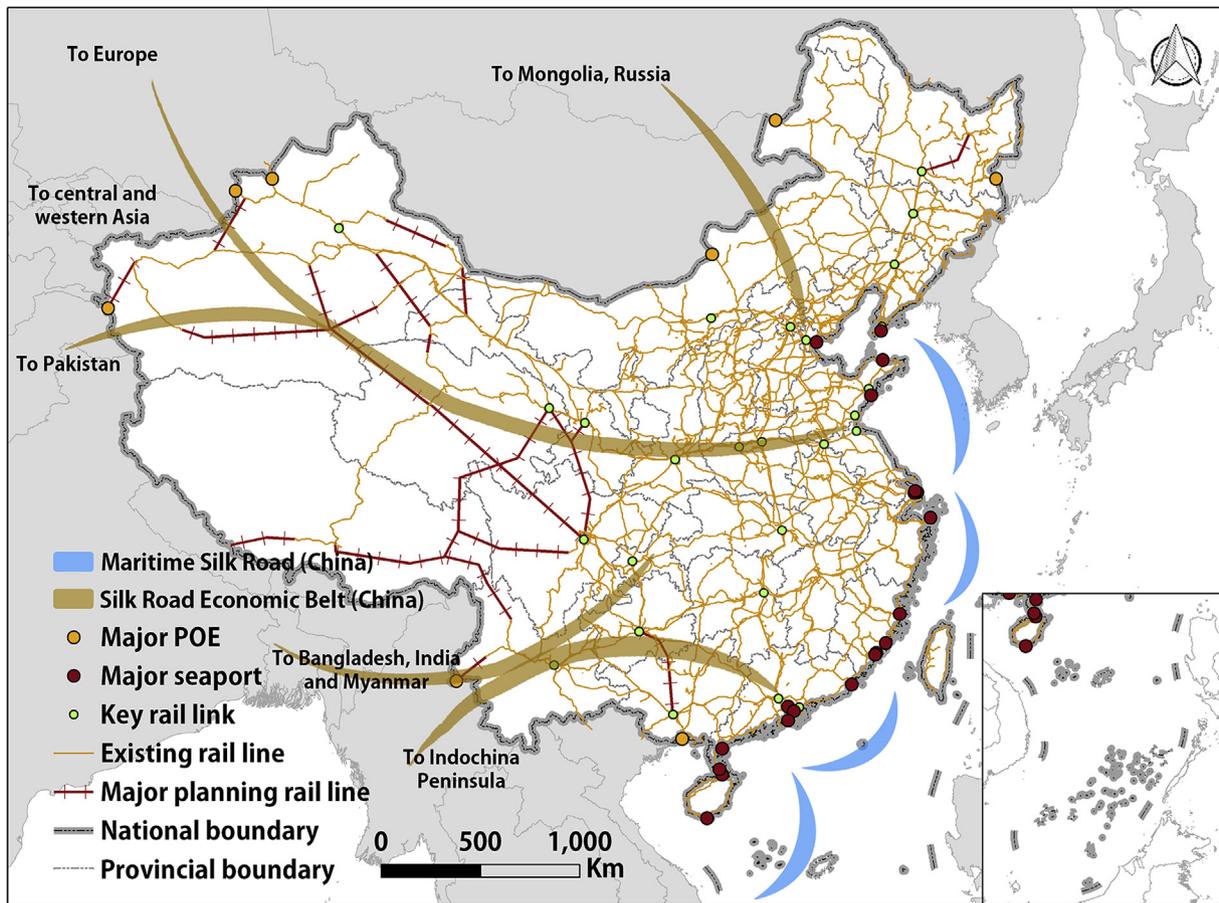


Fig. 3. The core transportation elements of the B&R Initiative, mapped roughly.

$$d_{DTW} = \min \left( \sum_{k=1}^K w_k / K \right) \quad (2)$$

This study uses the k-medoids algorithm to cluster the WLA

sequences. The k-medoids algorithm is an improved version of the k-means clustering algorithm (Hartigan, 1979; Krishnapuram, Joshi, & Yi, 1999). In addition, the silhouette value, which measures the similarity of an object to its own cluster compared to other clusters, is used to

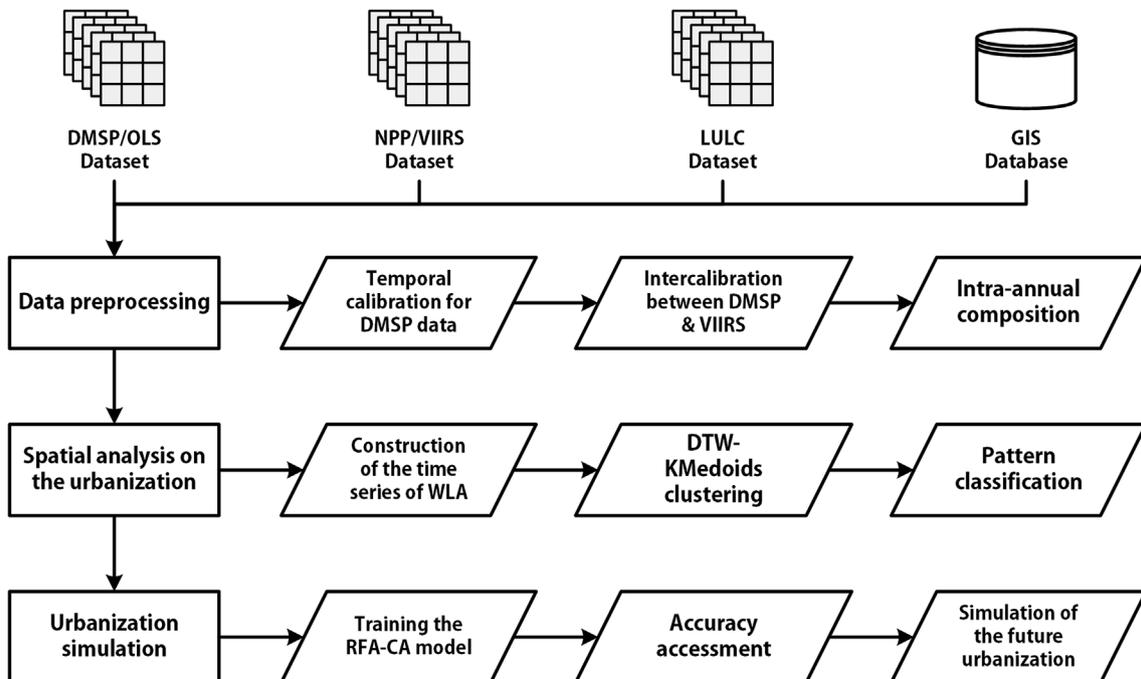


Fig. 4. Flowchart of the analysis of urbanization.

**Table 1**  
Landscape metrics and their descriptions.

Landscape metrics	Description
NP	Number of patches of a particular patch type is a simple measure of the extent of subdivision or fragmentation of the patch type.
CA	Class area is a measure of landscape composition; specifically, how much of the landscape is comprised of a particular patch type.
PLAND	Percentage of landscape quantifies the proportional abundance of each patch type in the landscape.
COHESION	Patch cohesion index measures the physical connectedness of the corresponding patch type.
SHDI	Shannon's diversity index is a measure of diversity in community ecology, applied here to landscapes.
AI	Aggregation index is calculated from an adjacency matrix, which shows the frequency with which different pairs of patch types (including like adjacencies between the same patch type) appear side-by-side on the map.

**Table 2**  
The rates of urban land on both sides of the Hu Line and their proportions.

Year	Rate of the southeast side	Rate of the northwest side	Proportion of the two sides
2005	0.9159%	0.0591%	93.9385:6.0615
2010	1.3884%	0.0806%	94.5133:5.4867
2015	1.5629%	0.1164%	93.0685:6.9315
2020	1.5641%	0.1165%	93.0681:6.9319
2025	1.5669%	0.1167%	93.0684:6.9316
2030	1.5697%	0.1169%	93.0679:6.9321
2035	1.5751%	0.1171%	93.0675:6.9325

interpret and validate the consistency within the clusters of the k-meoids result (Chen et al. 2017; Da Cruz Nassif & Hruschka, 2013; Liu et al. 2017; Rousseeuw, 1987). To improve the silhouette values and the quality of the clusters, a dichotomy method is used for the clustering. First, all the administrative units are clustered into 20 categories. The categories with a silhouette value greater than 0.5 are extracted from the sample population. Then, the remaining units are clustered again, and the clusters with a silhouette value greater than 0.5 are taken out of the sample population. This process continues until all the samples are clustered successfully.

The concavity and convexity of curves can reveal the change pattern, which is widely used in the study of urbanization patterns (Ma et al. 2012; Northam, 1979; Zhang & Seto, 2011). This study derives Equation (4) to analyze the pattern classification model of urbanization in China.

$$y = a + b \times x^c \begin{cases} b > 0, 0 < c < 1, \text{Early Urban Growth} \\ b > 0, c = 1, \text{Constant Urban Growth} \\ b > 0, c > 1, \text{Recent Urban Growth} \\ b < 0, \text{Deurbanization} \end{cases} \quad (3)$$

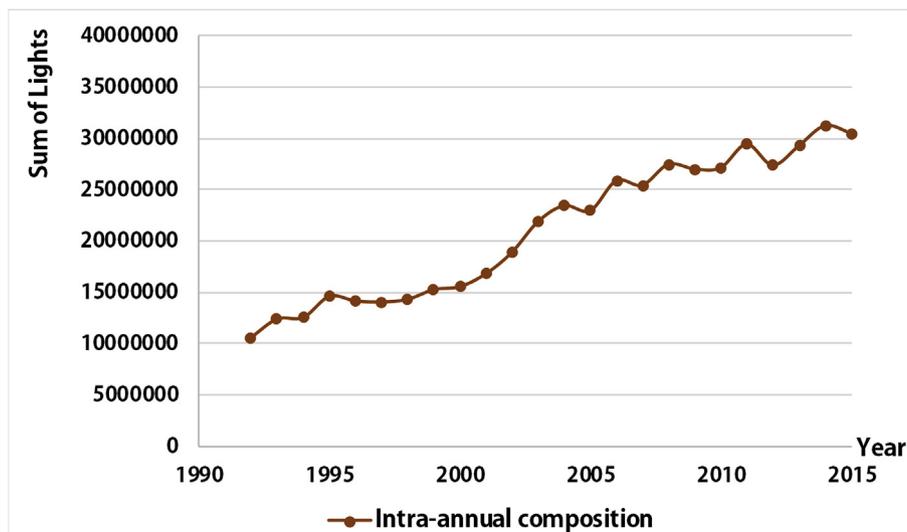


Fig. 5. The curve of the sum of the lights of the NTL composite product.

where  $x$  denotes time,  $y$  denotes WLA, and  $a$ ,  $b$ , and  $c$  are unsolved parameters.

In addition, for the administrative units with WLA of 0 in some years, the urbanization level is considered to be extremely low. Therefore, this study defines these regions as the zero-value areas, which is not included in the power function fitting analysis.

### 3.3. Urbanization simulation at a large scale by integrating random forests and cellular automata

The random forests algorithm (RFA) is a multiclassifier combination model that comprises a large number of decision trees (Breiman, 2001). RFA increases the diversity among the decision tree classifiers by generating different training subdatasets in order to minimize the potential for overfitting during the training process and improve prediction accuracy. Furthermore, a previous study has indicated that 37% of the samples in the original dataset are left as out-of-bag (OOB) data (Biau, 2012). Using OOB data to estimate the performance and accuracy of the RFA classification model is referred to as OOB-estimation, and an error report of OOB-estimation for each decision tree is obtained. The generalization error of the RFA is calculated by averaging the errors of the decision trees via OOB-estimation.

Previous studies have pointed out that the RFA is the optimal classifier to avoid overfitting and address nonlinear, high-dimensional problems (Biau, 2010; Breiman, 2001; Fernández-Delgado et al. 2014), and it is widely used in research on land use and urban study. This study conducts a simulation of urban expansion by integrating the RFA and CA models. The RFA-CA model performs well in simulating urban growth with high accuracy (He et al. 2018; Kamusoko & Gamba, 2015; Liang et al. 2018; Yao et al., 2017a).

For the CA model, the equation is as follows (Li & Yeh, 2002):

$$P(k, t + 1) = Pg(k, t) \cdot \Omega(k, t) \cdot RA \tag{4}$$

where  $P$  denotes the urban conversion probability,  $Pg$  represents the development suitability,  $\Omega$  is the neighborhood effect,  $RA$  stands for the random probability,  $k$  is the  $k$ -th cell, and  $t$  denotes the time.

The development suitability  $Pg$  is calculated via the RFA (Kamusoko & Gamba, 2015). The model generates  $N$  sample points to obtain the training data  $D$ . The equation is as follows:

$$D = \begin{bmatrix} x_{1,1}, x_{1,2}, \dots, x_{1,S}, Y_i \\ x_{2,1}, x_{2,2}, \dots, x_{2,S}, Y_i \\ \dots \\ x_{N,1}, x_{N,2}, \dots, x_{N,S}, Y_i \end{bmatrix} \tag{5}$$

where  $S$  is the total number of spatial variables and  $Y_i$  is the  $i$ -th land-use conversion type (e.g., in this study  $i = 1, 2, \dots, 9$ ).

The equation for calculating the development suitability  $Pg$  is as follows:

$$Pg(k, t) = \begin{cases} \frac{I(h(x) = Y_{non-urban})}{M}, & Cell(k, t) = non - urban\ area \\ 1, & Cell(k, t) = urban\ area \\ \frac{I(h(x) = Y_{limited-area})}{M}, & Cell(k, t) = limited - developed\ area \end{cases} \tag{6}$$

where  $I(\cdot)$  denotes an indicator function.

Generally, when the neighboring area is mostly urban, this area is likely to be converted into an urban area. The equation can be expressed as follows:

$$\Omega(k, t) = \frac{\sum_{i=x-w/2}^{x+w/2} \sum_{j=y-w/2}^{y+w/2} P_{Cell(i,j,t)}}{w^2} \tag{7}$$

where

$$P_{Cell(i,j,t)} = \begin{cases} 0, & Cell(i, j, t) \neq urban\ area \\ 1, & Cell(i, j, t) = urban\ area \end{cases} \tag{8}$$

$\Omega(k, t)$  denotes the neighborhood effective probability of the  $k$ -th cell at time  $t$ ,  $(x, y)$  denotes the position of  $k$ -th cell in the  $x$ -th row and the  $y$ -th column, and  $w$  is the size of the neighborhood effect window.

Since it is difficult to reveal comprehensive driving factors with common spatial variables, a random disturbance term is usually added to the CA simulation process to generate highly reasonable results (Kamusoko & Gamba, 2015), which contribute to producing more accurate simulated results and can be described as follows:

$$RA = 1 + (-\ln Y)^\alpha \tag{9}$$

where  $\gamma$  is a uniform random variable within the range of 0–1 and  $\alpha$  is a parameter for controlling the size of the stochastic perturbation.

In this study, the Markov chain was used to predict the future growing numbers of urban cells using the LULC data in 2005, 2010 and 2015 (Arsanjani et al. 2013) and the future development probability was calculated using the RFA-CA model. The development probability of regional urbanization trends is based on the types of urbanization patterns, including the Recent Urban Growth, the Early Urban Growth, the Deurbanization and the Zero-value Area. The larger the probability is, the greater the potential of urbanization.

Recent Urban Growth covers most regions in China, thus representing the overall pattern of China. Therefore, the  $Pg_{trend}$  of these regions is set to 1. The urbanization of the Recent Urban Growth class slows down but is still relatively large. Therefore, the  $Pg_{trend}$  of the regions in the Early Urban Growth class is set to 0.6. Limited changes in urbanization will occur in the regions in the Zero-value Area class and the Deurbanization class due to their harsh natural environment. Thus, the  $Pg_{trend}$  in those regions is set to 0.2.

To simulate future urban expansion, the future development probability must be calculated first. The regional potential of urbanization and future core elements of national policy are taken into account in the calculation of the future development probability. The equation is as

follows:

$$Pg_{future} = Pg_{present} \times \frac{1}{d} \times Pg_{trend} \tag{10}$$

where  $Pg_{future}$ ,  $Pg_{present}$ ,  $d$  and  $Pg_{trend}$  denote the future development probability, the current development probability, the regional trend of urbanization, and the distance to the core elements of national policy, respectively.

### 3.4. Accuracy evaluation

The model calibration of RFA-CA is conducted using the LULC data in 2005, 2010 and 2015. We first mine the urban transformation rules using RFA and the LULC data in 2005, and the cross-validation error of the RFA is 1.75%. Then, based on the model and parameters, the overall development probability is calculated. Next, we simulate the urban land in 2010 and 2015 and evaluate the accuracy of the simulations using the LULC data in 2010 and 2015.

Classic CA models employed the overall accuracy (OA) and Kappa to evaluate the accuracy of the simulation results. However, previous studies revealed that Kappa and OA are not suitable for use in large-scale scenes because a large number of pixels are not converted and the ratio of conversion to nonconversion is extremely low (Du et al. 2012; Pontius Jr et al. 2008; Yao et al., 2017b). In a large-scale scene, the values of OA and Kappa are close to 100%, which are not comparable. Actually, only the pixels that have undergone land-use conversion are addressed. Therefore, we introduce the figure of merit (FoM) to evaluate the simulation results (Pontius Jr et al. 2008). The FoM focuses on the union of the observed change and simulated change to the intersection of the observed change and simulated change of the land use and land cover (Perica & Foufoula Georgiou, 1996).

Moreover, landscape metrics are used to assess the landscape patterns between the two sides of the Hu Line (Chen et al. 2014; McGarigal, Cushman, & Ene, 2012). This study adopted several landscape metrics to measure the landscape patterns of LULC, including the number of patches (NP), class area (CA), percentage of landscape (PLAND), patch cohesion index (COHESION), Shannon's diversity index (SHDI) and aggregation index (AI). These metrics were calculated using Fragstats 4 (McGarigal et al. 2012). The descriptions of the landscape metrics are shown in Table 1.

## 4. Results

### 4.1. Spatial differentiation of the urbanization level in China

The WLA of each administrative unit for each year from 2005 to 2015 in China is calculated using the preprocessed NTL dataset. Based on the DTW distance between the WLA sequences of all units, this study uses the k-medoids algorithm to cluster the regions with a similar process of urbanization.

This study uses the DTW-based k-medoids method for clustering based on the silhouette value (Boultif & Louër, 2004). As shown in Fig. 6, the units are grouped into 9 clusters with silhouette values that are all greater than 0.5, indicating that all clusters have superior quality.

The areas with high urbanization levels are mainly in the North China Plain and the eastern coastal area (Fig. 6(a)–(c)). Most of the provincial capitals are also at a relatively high level of urbanization. Level 1 has the highest level of urbanization. Those regions of level 1 are basically gathered in the southeastern part of the Hu Line, especially the Beijing, Tianjin and Tangshan regions as well as the Yangtze River Delta and the Pearl River Delta. Level 9, which is the lowest urbanization level, covers the northwest side of the Hu Line. It is worth noting that this level is also distributed in the southeast part and concentrated in the mountainous areas, such as the Yunnan-Guizhou Plateau, Changbai Mountains, and the nearby Xing'an Mountains region.

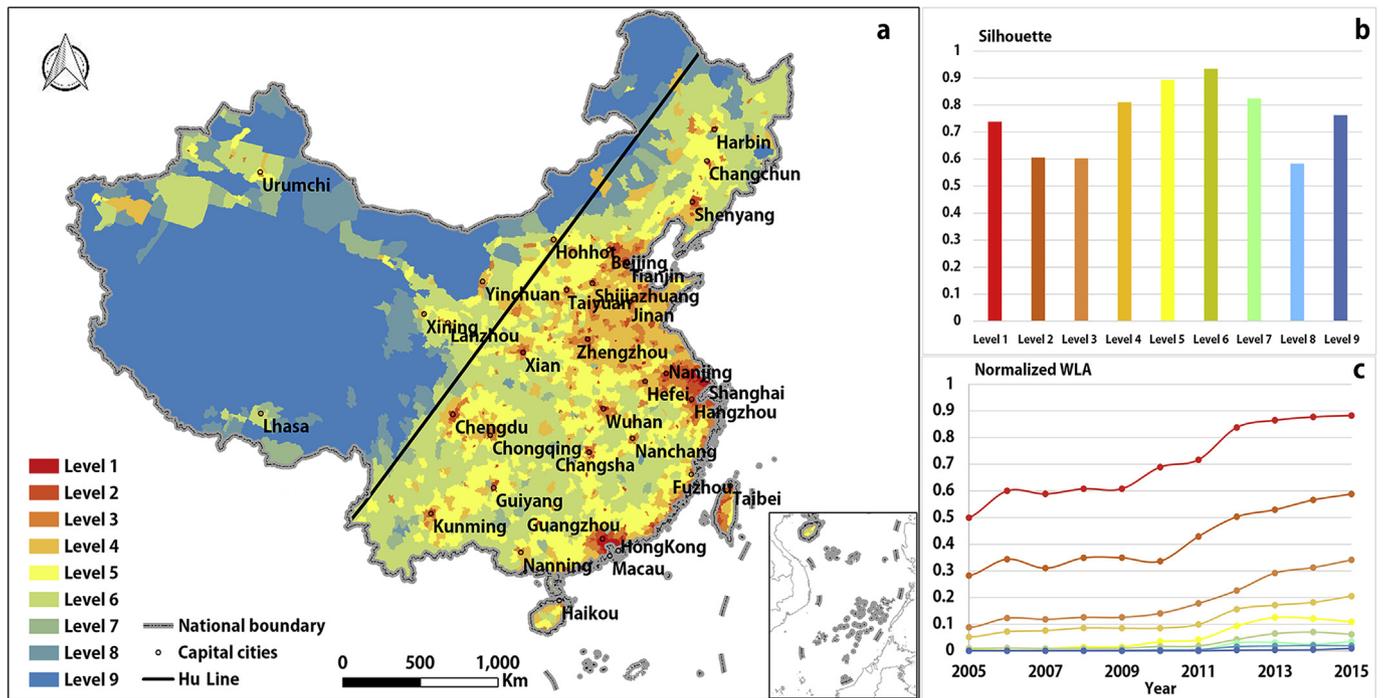


Fig. 6. The clustering result for the urbanization level in China. (a) Spatial distribution of the clustering result; (b) silhouette values for the clusters; (c) normalized WLA curves for each cluster center.

The urbanization level is closely related to the terrain (Batty & Xie, 1994). In the harsh environment areas of the northwest part and some mountainous areas of the southeast part, urbanization levels are very low. The areas with high urbanization levels are mainly distributed in the flat terrain, such as the North China Plain, the middle and lower reaches of the Yangtze River, the Northeast Plain, and the Pearl River Delta.

#### 4.2. Spatial differentiation of the urbanization pattern in China

We conducted a pattern classification analysis on the WLA sequences and obtained the classification result of the spatial differentiation for the urbanization patterns in China (Fig. 7). The urbanization pattern of most areas in China is in the Recent Urban Growth class, which is a process of accelerating growth.

The pattern of several regions is the Early Urban Growth class, which is a state of deceleration, including the Beijing-Tianjin-Tangshan region, the Yangtze River Delta, the Pearl River Delta region, Taichung City, Heilongjiang Province, and several parts of Shanxi Province. The developed areas, including the Beijing-Tianjin-Tangshan region, Yangtze River Delta, and Pearl River Delta regions as well as Taichung City and Heilongjiang Province, have been the leaders in economic development since China's reform and opening up. Therefore, compared with the urban land coverage of other cities, the coverage of these regions is higher and their urbanization status is currently close to saturation.

Special attention should be paid in Shanxi Province, where three urbanization patterns exist at the same time, i.e., Recent Urban Growth, Early Urban Growth and Deurbanization, which indicates that the WLA sequences of Shanxi Province are not stable. The trend of WLA is more similar to Shanxi's economic change compared with the economic statistics from the government. The reason for this instability may be that the government's remediation of the local environment has recently led to a decline in the coal economy (Hu et al. 2016; Zhao et al. 2007).

The Zero-value Areas are mainly distributed in the northwest part of the Hu Line and the mountainous areas in the southeast part. The distribution is very similar to that of the clustering result, which is related

to the definition of the Zero-value Area. The Zero-value Area, which is defined as areas with a low urbanization level where the WLA value is 0 in some years, is used to eliminate the unstable sequences of WLA and to avoid misjudgment in the process of pattern classification.

#### 4.3. Model calibration and analysis of the simulation result

The calibrated model has FoMs of 0.184 (2010) and 0.185 (2015), with producer accuracies of 23.06% (2010) and 23.12% (2015) and user accuracies of 18.36% (2010) and 18.38% (2015). The FoM values of the results of 2010 and 2015 are close since no large changes took place in China's regional economic development plan and urbanization was developing steadily based on the analysis of the WLA during the period of 2010–2015. The RFA-CA calibration has a good performance. Previous studies have noted that the FoM value generally tends to be low, while CA models are applied to large-scale regions and areas with small amounts of observed change (Lin & Li, 2016; Pontius Jr et al. 2008).

The simulation results of urban expansion are compared with the LULC data (Fig. 8). The overall distribution and contour of simulated urban land are generally consistent with those of the urban land of the LULC data. In detail, the degree of aggregation of simulated urban land is higher than that of the LULC and the urban cells are combined into patches. The phenomenon of “salt and pepper” is reduced, and the result is more similar to the real urban development pattern.

The contributions of the driving factors can be analyzed after considering the transition rules using the RFA (Fig. 9). We found that the driving factors of the top 4 contributions are the NTL, the distance to other roads, the distance to high-speed railways and the DEM slopes. These spatial variables represent three related factors that directly affect urbanization: socioeconomic activities, accessibility and natural environment. Among them, the DEM slope reveals the terrain and represents the natural environment, which is the basic condition for human economic activities and urban development (Gu, Hu, & Cook, 2017).

The distance to other roads and the distance to high-speed railways represent the accessibility of the interior and exterior of the city,

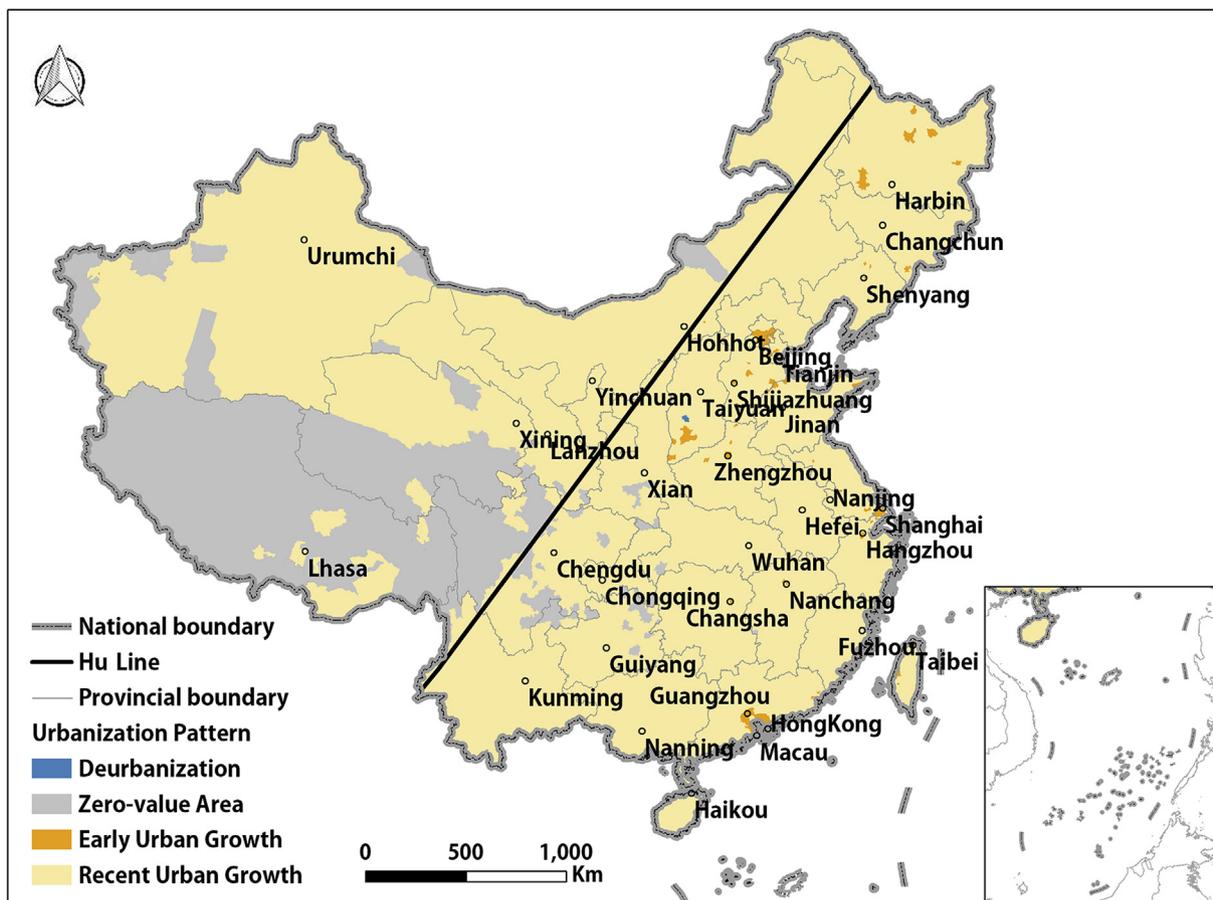


Fig. 7. The classification result for the urbanization pattern in China.

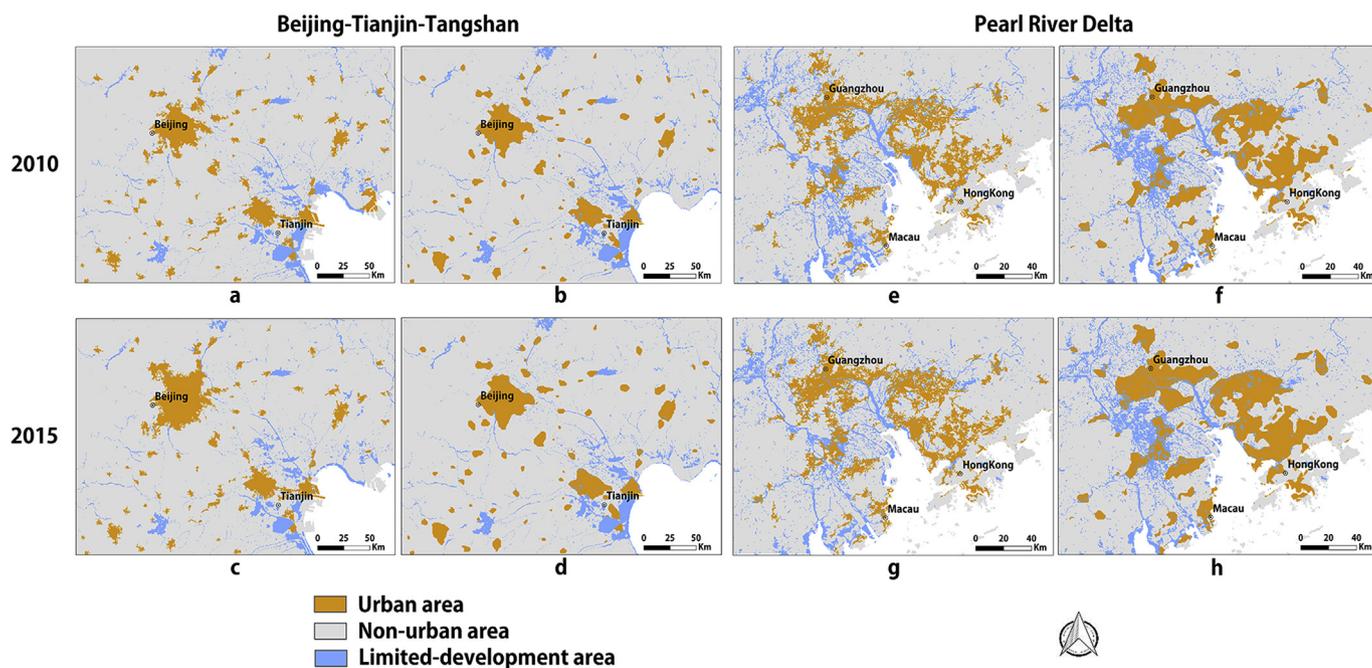


Fig. 8. Details of the RFA-CA simulation results. (a)–(d) are the LULC data in 2010, the simulation results in 2010, the LULC data in 2015, and the simulation results in 2015 in the Beijing-Tianjin-Tangshan region. (e)–(h) are the LULC data in 2010, the simulation results in 2010, the LULC data in 2015, and the simulation results in 2015 in the Pearl River Delta region.

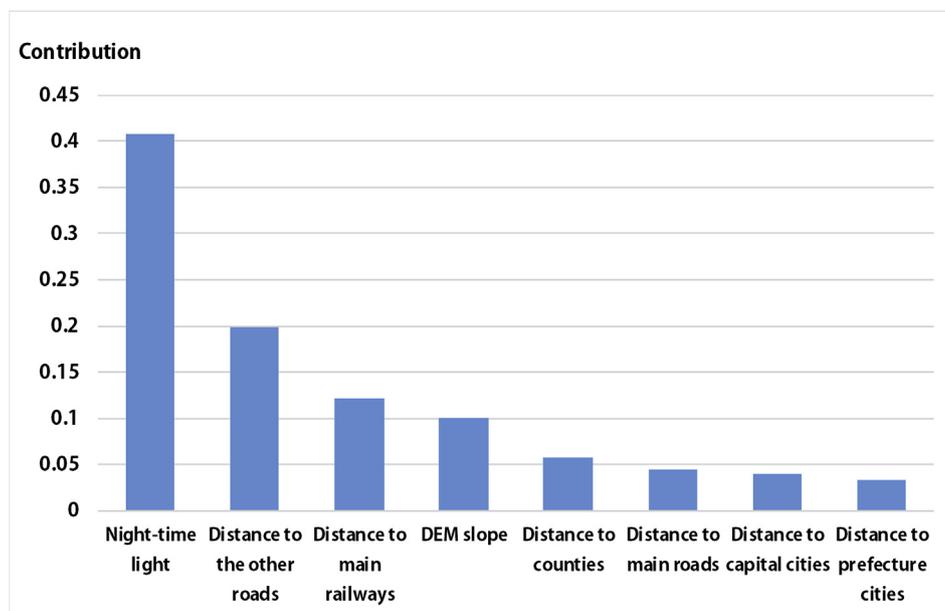


Fig. 9. Contribution of spatial driving factors.

respectively. A convenient transportation infrastructure can improve residents' efficiency and quality of life. The expansion of urban land will be carried out in the most economical way. Therefore, urbanization is closely related to the accessibility of the transportation infrastructure (Yao et al., 2017a). Previous studies have indicated that the NTL is significantly correlated with demographic and socioeconomic variables and able to represent socioeconomic activities (Ma et al. 2012). China's urbanization level is driven by economic activity (Gu et al. 2017). It is worth noting that the contribution of the NTL exceeds 40%, which is significantly higher than the other driving factors, which is because the NTL is related to multiple urbanization indicators (e.g., population, GDP, built-up area and electric power consumption), and it is a comprehensive indicator that is highly correlated with urbanization (Ma et al. 2012).

We calculated the ratio of the urban land to the total land in each county unit. As shown in Fig. 10, the distribution of the urban-land ratio simulated based on the RFA-CA model is consistent with the LULC data. With the Hu Line as the boundary, the uneven distribution of urban land on the east and west sides of China can be clearly seen. The areas with a high proportion of urban land ( $> 4.35\%$ ) are mainly distributed in the northern China plain and the southeastern coastal area. The northwest side covers most of the areas with a low proportion of urban land ( $< 0.04\%$ ).

#### 4.4. Analysis of the possibility of China's urbanization breaking through the Hu Line pattern in the future

Previous studies have suggested that China's economy will embrace a new pattern of opening to the west and developing inland in the context of the B&R Initiative (Chen et al. 2016; Yuan & Qin, 2016; Zhang et al. 2015). On one hand, the B&R Initiative bridges the influence of domestic regional policies on the development on both sides of the Hu Line (Liu, 2015). The initiative will promote regional economic development and break through the bottleneck through the creation of win-win cooperation and joint prosperity (Huang et al. 2016). On the other hand, the transportation system is an important basis for the implementation of national policies, shapes the economic spatial pattern, and likely plays a fundamental role in fostering regional cooperation and development, especially at the early stage of the initiative (Dong et al. 2016; Huang, 2016). Therefore, from the perspective of transportation, we concretize the impact of the B&R

Initiative on domestic urbanization using the distance to the core transportation elements that the B&R Initiative mentioned.

The geographical differences of urbanization on the two sides of the Hu Line will continue to exist in 2015–2035, which reveals the large gap of urbanization between the southeast part and the northwest part in China. Compared with the regions in 2010 and 2015, more regions on both sides of the Hu Line are in red, which means that their ratios of urban land will continue to increase (Figs. 10 and 11). However, spatial differentiation exists between the two parts of the Hu Line, which can be clearly seen from Fig. 11.

To further analyze the distribution of urban land, the ratios of urban land on both sides to their total land use and the proportions of the east and west sides are calculated (Table 2). The ratio of the northwest side doubles from 0.06% in 2005 to 0.12% in 2035. The proportion of the two sides of the Hu Line gradually decreases from 93.94:6.06 in 2005 to 93.07:6.93 in 2035. The differentiation between the two sides is narrowing. However, the proportion of the two sides remains large, which means that a very large gap will exist in 2015–2035. Therefore, urbanization on both sides of the Hu Line is in an accelerating pattern, although the spatial distribution of the urban land in China will maintain the Hu Line pattern in 2015–2035.

Table 3 shows a comparison of the landscape metrics between the eastern side and western side of the Hu Line. Here, CA measures the total area of the urban land class. The order of magnitude difference between the CAs of the two sides shows a large gap between the urban land bases of the two sides. The eastern side of the Hu Line has a large urban land base and more development resources, and it covers major economic development areas in China. This finding indicates that a small change in the proportion corresponds to very large development growth. Compared with the eastern side, the western side of the Hu Line covers a large number of regions with harsh natural conditions and has a small urban land base.

The SHDI measures the overall diversity of the landscape pattern of the two sides. The values of the western side are much smaller than those on the eastern side because of the large number of nonurban developable areas on the western side. The values of COHESION on both sides are increasing year by year, which means that the connectivity of the urban patches on both sides is increasing. Here, AI measures the aggregation index of the urban land class. The values of the two sides are similar, indicating that their urban agglomerations are similar. However, it is worth noting that the trend of the AI of the

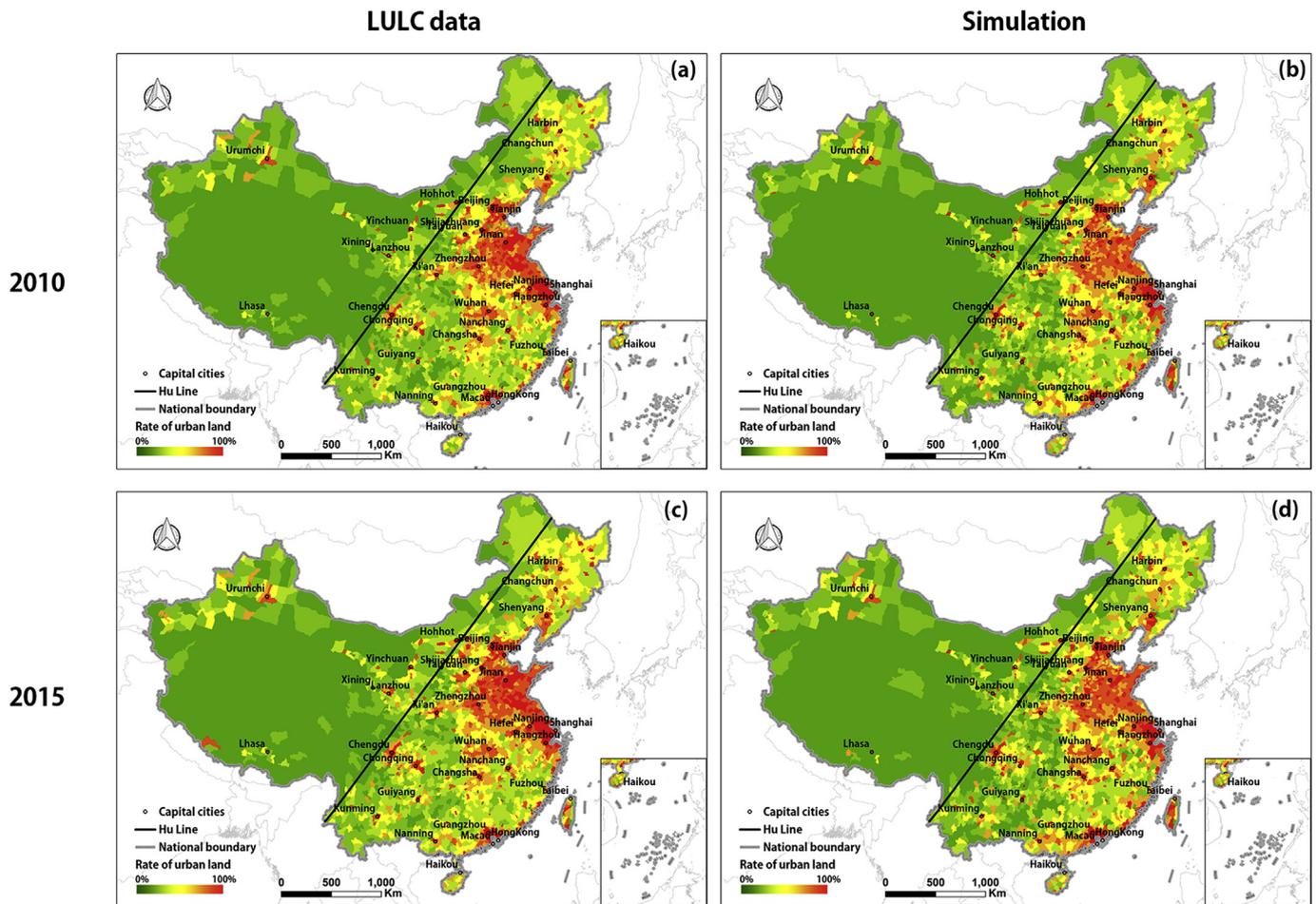


Fig. 10. The ratio of urban land for comparison between the LULC data and simulation results. (a) LULC data in 2010; (b) the simulation results in 2010; (c) LULC data in 2015; (d) the simulation results in 2015.

eastern side is on the rise while that of the western side is on the decline. The values of PLAND on both sides are increasing year by year, indicating that both sides are urbanizing. However, as seen from the NP, the number of urban patches on the eastern side decreases while that on the western side increases annually.

The urbanization on the eastern side is mainly based on the internal expansion of the urban agglomeration, while that on the western side is mainly based on outward expansion from year 2020 to year 2035. Considering the trend of AI, PLAND and NP, we can identify the difference between the urbanization patterns of the two sides. On the eastern side, the expansion of urban land mainly occurs in the connection between cities and the urban patches are clustered together, such as the connection between Guangzhou and Foshan (two adjacent cities in the Pearl River Delta (Fig. 8)). Thus, the NPs decrease as the AIs increase. On the western side, urbanization mainly develops new urban land in nonurban developable areas. With increasing numbers of new urban areas, the overall pattern of the urban land class is increasingly fragmented. Therefore, the NPs are increasing while the AIs are decreasing.

## 5. Discussion

This study is the first to analyze long-term urbanization using NTL data coupled with DMSP/OLS and NPP/VIIRS datasets. In previous studies, the research time of the long-term analysis of nighttime lighting has to stop in 2013 because few updates have been performed for the DMSP/OLS dataset (Elvidge et al. (2009a, b); Huang et al. 2016; Li et al. 2016; Liu et al. 2012; Zhang et al. 2016). However, as time passes, the

demand for NTL data after 2013 is growing. The NPP/VIIRS dataset is a new version of the NTL data from 2012 and cannot be directly integrated into the DMSP/OLS dataset. In this paper, the intercalibrated VIIRS data are used as a continuation of the DMSP/OLS dataset, and it breaks the limitation of research time and is successfully used for the analysis of long-term urbanization in China.

This paper is the first to attempt to concretize the new regional impacts that the B&R Initiative may cause and successfully couple it with the CA urban land simulation. On one hand, although the B&R Initiative is a diplomatic initiative for economic globalization, domestic development is greatly affected by the international environment. With the impact of globalization, countries with substantial inward and domestic investment have made remarkable progress, thus reducing poverty and improving the lives of millions of people (Sheppard et al. 2016). The B&R Initiative will promote regional economic development and help break through the bottleneck of China through the creation of win-win cooperation and joint prosperity (Huang et al. 2016). On the other hand, traffic infrastructure development is a crucial element of the B&R Initiative, which affects infrastructure connectivity, unimpeded trade, and people-to-people exchange and likely plays a fundamental role in fostering regional cooperation and development, especially at the early stage of the initiative (Dong et al. 2016; Huang, 2016). Moreover, our experiments suggest that transportation infrastructure is an important driving factor in urbanization (Fig. 9). Therefore, from the perspective of transportation infrastructure and connectivity, we analyze the impact of the B&R Initiative on regional differences in domestic urbanization using the distance to the core transportation elements that the B&R Initiative mentioned.

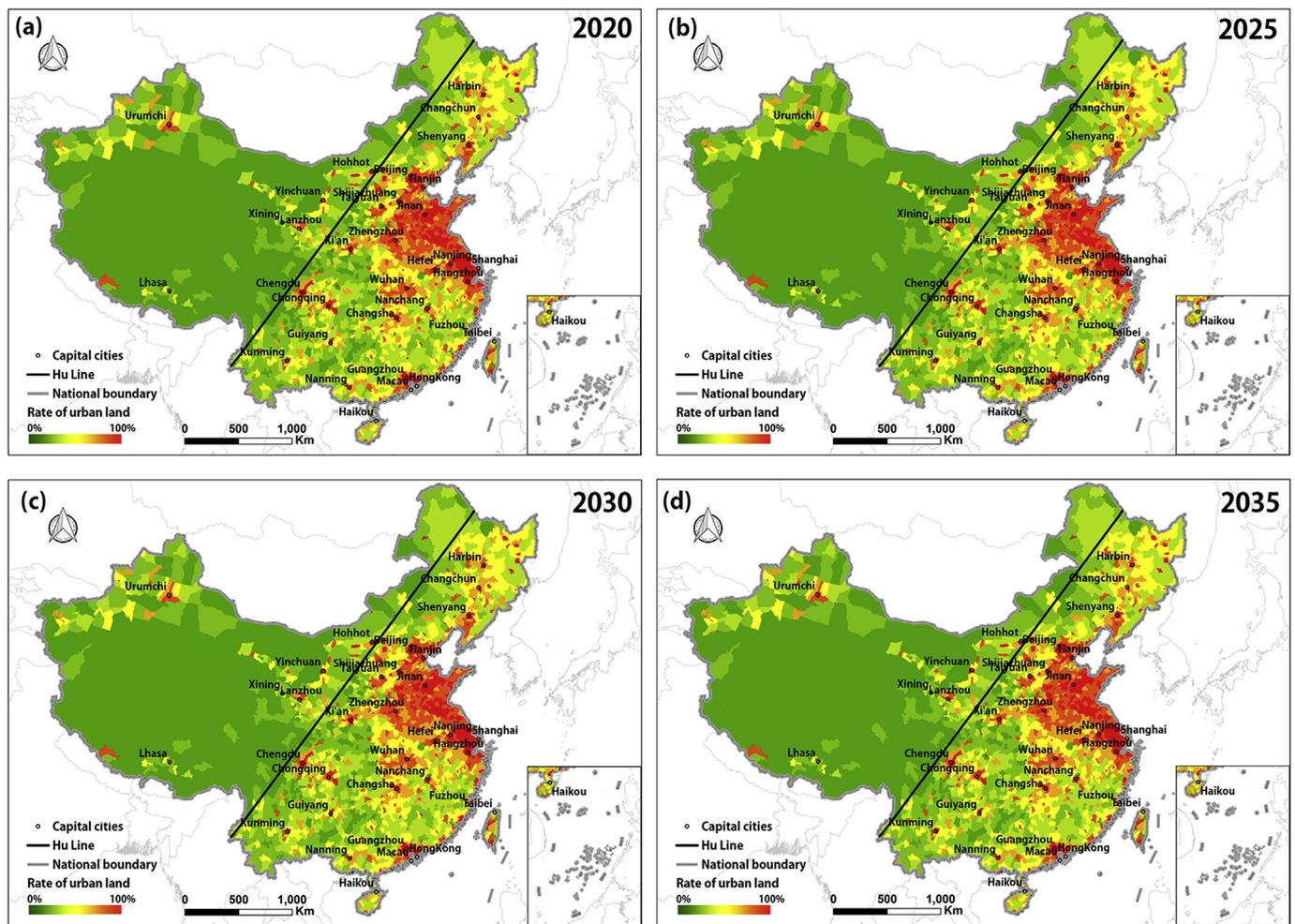


Fig. 11. The simulation results of the future urban land from 2020 to 2035 via RFA-CA model.

**Table 3**  
Comparison of landscape metrics between the eastern side (E) and western side (W).

Year		2020	2025	2030	2035
NP	E	9448	9408	9386	9298
	W	980	987	995	1013
CA	E	3858175	5649150	6433525	6583050
	W	311650	439150	491550	613950
PLAND	E	19.69%	28.83%	32.83%	33.59%
	W	1.59%	2.24%	2.51%	3.13%
COHESION	E	93.615	93.640	93.634	93.699
	W	91.254	91.262	91.273	91.281
SHDI	E	0.674	0.674	0.674	0.675
	W	0.0296	0.0297	0.0297	0.0298
AI	E	74.205	74.210	74.213	74.215
	W	74.139	74.137	74.096	74.017

Using the NTL data and CA model, we provide a quantitative answer to the question of whether the Hu Line pattern of urbanization in China will change, and the results offer new ideas and a reliable numerical analysis for government decision-making and urban planning studies. According to the simulation results, changes in the proportion of urban land on both sides are less than 0.001 in the period of 20 years, and the geographical differences of urbanization on the two sides of the Hu Line will continue to exist in 2015–2035. However, we cannot ignore a large change in the difference between the two sides of urban growth that the small change in the proportion means. The quantity changes of urban land between the two sides are very different. There is an order of

magnitude difference between the rates of urban land on the two sides (Table 3). However, the proportions of the two sides are still decreasing. Since 2015, the rate of urban growth on the western side (96.95%) has exceeded that on the eastern side (70.64%), which indicates that the urbanization on the western side is accelerating to catch up with that on the eastern side (Fig. 12). According to the World Urbanization Prospects from the United Nations, China will be an important growth pole for urbanization in the world in the next 30–40 years (UN, 2015). With the expansion of research time, we believe that the proportion of the two sides of the Hu Line will increasingly change.

Although reliable results are obtained, some deficiencies are observed in this study. First, this paper does not consider transport and telecommunications among countries. Different countries have different degrees of connection with each other. In addition, different ports of entry into China adopt the telecommunications of different countries. Although the location of the seaports and the ports of entry are considered in this paper, such data are not sufficient to conduct further research. Second, this paper only considered the traffic infrastructure as the impacts that the B&R Initiative may bring to diplomatic regions. Although traffic infrastructure plays a fundamental role in the early stage of the initiative, the scope of the B&R Initiative is more comprehensive and covers policy dialogue, financial support and so on. If we perform further research to predict longer-term urbanization, more factors need to be concretized. However, the purpose of this paper is to propose a framework to analyze urbanization change under the background of the traffic infrastructure at the early stage of the initiative. This paper has successfully found that the geographical differences of urbanization on the two sides of the Hu Line will continue to

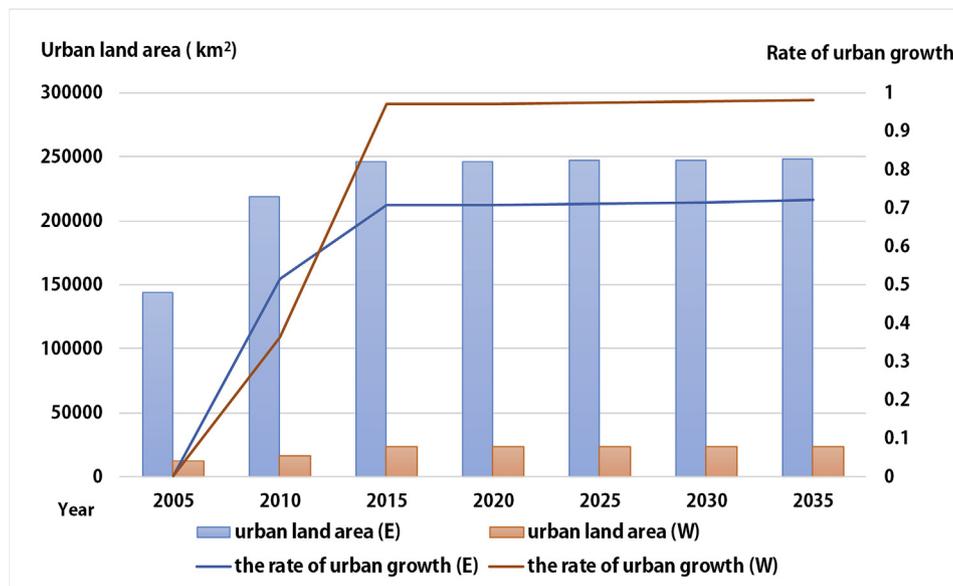


Fig. 12. Comparison of numbers of urban pixels and rates of urban growth between the two sides of the Hu Line.

exist in 2015–2035, which reveals the large gap of urbanization between the southeast part and the northwest part in China. Moreover, the results offer new insights for government decision-making and urban planning studies.

## 6. Conclusions

This study focuses on the issue of urbanization of the Hu Line pattern in the past and future in the context of the B&R Initiative. A large-scene urbanization quantitative spatial analysis framework via nighttime light and RFA-based CA is proposed. We first analyze the urbanization level and pattern on both sides of the Hu Line from 2005 to 2015. Next, we consider the impacts that the B&R Initiative brings and couple them with the urban land simulation model. Finally, we analyze the differences in the urban land pattern between the two sides of the Hu Line from 2015 to 2035.

We found that spatial differentiation occurs in the urbanization level on both sides of the Hu Line, with high-urbanization-level areas concentrated on the eastern side and low-urbanization-level areas concentrated on the western side. Moreover, the geographical differences of urbanization on the two sides of the Hu Line, which is called the Hu Line pattern, will continue to exist from 2015 to 2035. However, the proportion of urban land on both sides of the Hu Line gradually decreased from 93.94:6.06 in 2005 to 93.07:6.93 in 2035. Although a large gap still exists, the difference between the two sides is narrowing since the urbanization on the western side is accelerating to catch up with that on the eastern side. In addition, the trends of future urbanization on the two sides are not the same, with urbanization on the eastern side mainly the internal expansion of urban agglomeration while that on the western side mainly the outward expansion.

In summary, this study analyzes the spatial differentiation of current and future urbanization on both sides of the Hu Line based on NTL datasets and CA models under the background of the traffic infrastructure at the early stage of the B&R Initiative. The results can provide new ideas and a reliable numerical analysis for government decision-making and urban planning studies.

## Funding

This work was supported by the National Natural Science Foundation of China [41801306, 41671408, 71961137003]; Fundamental Research Funds for National University, China

University of Geosciences(Wuhan) [CUG190606]; Open Fund of State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University [18S01]; National Key R&D Program of China [2017YFB0503804]; the Natural Science Foundation of Hubei Province [2017CFA041]; the Basic Research Program of Shenzhen Science and Technology Innovation Committee [JCYJ20180305125113883].

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.apgeog.2019.102081>.

## References

- Antrop, M. (2000). Changing patterns in the urbanized countryside of Western Europe. *Landscape Ecology*, 15(3), 257–270.
- Arsanjani, J. J., Helbich, M., Kainz, W., & Boloorani, A. D. (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation*, 21, 265–275.
- Batty, M., & Xie, Y. (1994). From cells to cities. *Environment and Planning B: Planning and Design*, 21(7), 31.
- Biau, G. (2010). Analysis of a random forests model. *Journal of Machine Learning Research*, 13(2), 1063–1095.
- Biau, G. (2012). Analysis of a random forests model. *Journal of Machine Learning Research*, 13, 1063–1095 Apr.
- Boultif, A., & Louër, D. (2004). Powder pattern indexing with the dichotomy method. *Journal of Applied Crystallography*, 37(5), 724–731.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250.
- Chen, M., Dadao, L. U., & Zhang, H. (2009). Comprehensive evaluation and the driving factors of China's urbanization. *Acta Geographica Sinica*, 64(4), 387–398.
- Chen, Y., Li, X., Liu, X., & Ai, B. (2014). Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy. *International Journal of Geographical Information Science*, 28(2), 234–255.
- Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., et al. (2017). Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method. *Landscape and Urban Planning*, 160, 48–60.
- Chen, M., Gong, Y., Li, Y., Lu, D., & Zhang, H. (2016). Population distribution and urbanization on both sides of the Hu Huanyong Line: Answering the Premier's question. *Journal of Geographical Sciences*, 26(11), 1593–1610.
- Chen, M., Lu, D., & Liu, H. (2010). The provincial pattern of the relationship between China's urbanization and economic development. *Acta Geographica Sinica*, 65(12), 1443–1453.
- Da Cruz Nassif, L. F., & Hruschka, E. R. (2013). Document clustering for forensic analysis: An approach for improving computer inspection. *IEEE Transactions on Information Forensics and Security*, 8(1), 46–54.

- Deng, X., & Bai, X. (2014). Sustainable urbanization in western China. *Environment Science and Policy for Sustainable Development*, 56(3), 12–24.
- Dewan, A. M., & Yamaguchi, Y. (2009). Land use and land cover change in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. *Applied Geography*, 29(3), 390–401.
- Dong, S., Cheng, H., Guo, P., Li, F., Yu, Li, Li, Z., et al. (2016). Transportation industry patterns and strategy of the belt and road. *Bulletin of Chinese Academy of Sciences*, 6, 663–670.
- Du, Y., Ge, Y., Lakhani, V. C., Sun, Y., & Cao, F. (2012). Comparison between CBR and CA methods for estimating land use change in Dongguan, China. *Journal of Geographical Sciences*, 22(4), 716–736.
- Elvidge, C., Tuttle, B., Sutton, P., Baugh, K., Howard, A., Miles, C., et al. (2007). Global distribution and density of constructed impervious surfaces. *Sensors*, 7(9), 1962–1979.
- Elvidge, C., Ziskin, D., Baugh, K., Tuttle, B., Ghosh, T., Pack, D., et al. (2009a). A fifteen year record of global natural gas flaring derived from satellite data. *Energies*, 2(3), 595–622.
- Elvidge, C. D., Sutton, P. C., Ghosh, T., Tuttle, B. T., Baugh, K. E., Bhaduri, B., et al. (2009b). A global poverty map derived from satellite data. *Computers & Geosciences*, 35(8), 1652–1660.
- Elvidge, C., Baugh, K., Hobson, V., Kihn, E., Kroehl, H., Davis, E., et al. (1997). Satellite inventory of human settlements using nocturnal radiation emissions: a contribution for the global toolchest. *Global Change Biology*, 3(5), 387–395.
- Fang, C. (2009). The urbanization and urban development in China after the reform and opening-up. *Economic Geography*, 29(1), 19–25.
- Fan, J., Liu, Y., & Chen, T. (2013). The key strategies and innovative thinking for optimization on spatial pattern of urbanization in China. *Bulletin of Chinese Academy of Sciences*, 28(1), 20–27.
- Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research*, 15(1), 3133–3181.
- Ge, M., & Feng, Z. (2010). Classification of densities and characteristics of curve of population centers in China by GIS. *Journal of Geographical Sciences*, 20(4), 628–640.
- Gu, C. (1999). Study on phenomena and mechanism of land use/cover change in Beijing. *Journal of Natural Resources*, 14(4), 307–312.
- Gu, C., Hu, L., & Cook, I. G. (2017). China's urbanization in 1949–2015: Processes and driving forces. *Chinese Geographical Science*, 27(6), 847–859.
- Han, J., Hayashi, Y., Cao, X., & Imura, H. (2009). Application of an integrated system dynamics and cellular automata model for urban growth assessment: A case study of Shanghai, China. *Landscape and urban planning*, 91(3), 133–141.
- Hartigan, J. A. (1979). A K-means clustering algorithm. *Appl Stat*, 28(1), 100–108.
- He, J., Li, X., Yao, Y., Hong, Y., & Jinbao, Z. (2018). Mining transition rules of cellular automata for simulating urban expansion by using the deep learning techniques. *International Journal of Geographical Information Science*, 32(10), 2076–2097.
- Hu, H. (1935). Distribution of China's population. *Acta Geographica Sinica*, 2(2), 33–74.
- Huang, Y. (2016). Understanding China's Belt and Road initiative: Motivation, framework and assessment. *China Economic Review*, 40, 314–321.
- Huang, X., Schneider, A., & Friedl, M. A. (2016). Mapping sub-pixel urban expansion in China using MODIS and DMSP/OLS nighttime lights. *Remote Sensing of Environment*, 175, 92–108.
- Hu, W., Liu, J., Li, M., & Zhu, L. (2016). Exploration of substitute industry for shanxi's coal economy: on the path of key scenic spots propelling regional economic development. *China Population Resources & Environment*, 26(04), 168–176.
- Hu, Z., Wang, Y., Liu, Y., Long, H., & Peng, J. (2016). Spatio-temporal patterns of urban-rural development and transformation in east of the “Hu Huanyong Line”, China. *ISPRS International Journal of Geo-Information*, 5(3), 24.
- Jia, S. (2014). *Premier Li, it is unworkable to break Hu line!*. 2014-12-14 <http://blog.sciencenet.cn/blog-267937-850831.html>.
- Kamusoko, C., & Gamba, J. (2015). Simulating urban growth using a random forest-cellular automata (RF-CA) model. *ISPRS International Journal of Geo-Information*, 4(2), 447–470.
- Krishnapuram, R., Joshi, A., & Yi, L. (1999). A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering. *Fuzzy systems conference proceedings, 1999. FUZZ-IEEE '99*. 1999 (pp. 1281–1286). IEEE International.
- Kuang, X. (2014). One Belt one road: A way to break Hu line. *Shenzhen Special Zone Daily*. [http://sztqb.sznews.com/html/2014-12/01/content\\_3079606.htm,2014-12-01](http://sztqb.sznews.com/html/2014-12/01/content_3079606.htm,2014-12-01).
- Li, P. (2015). *New urbanization and the break of Hu line*. *People's Daily*. 2015-01-18 <http://theory.people.com.cn/n/2015/0108/c40531-26347255.html>.
- Liang, X., Liu, X., Li, D., Zhao, H., & Chen, G. (2018). Urban growth simulation by incorporating planning policies into a CA-based future land-use simulation model. *International Journal of Geographical Information Science*, 32(11), 2294–2316.
- Li, Q., Lu, L., Weng, Q., Xie, Y., & Guo, H. (2016). Monitoring urban dynamics in the southeast USA using time-series DMSP/OLS nightlight imagery. *Remote Sensing*, 8(7), 578.
- Li, X., Chen, G., Liu, X., Liang, X., Wang, S., Chen, Y., et al. (2017a). A new global land-use and land-cover change product at a 1-km resolution for 2010 to 2100 based on human-environment interactions. *Annals of the American Association of Geographers*, 107(5), 1040–1059.
- Li, X., Elvidge, C., Zhou, Y., Cao, C., & Warner, T. (2017b). Remote sensing of night-time light. *International Journal of Remote Sensing*, 38(21), 5855–5859.
- Li, X., Li, D., Xu, H., & Wu, C. (2017c). Intercalibration between DMSP/OLS and VIIRS night-time light images to evaluate city light dynamics of Syria's major human settlement during Syrian Civil War. *International Journal of Remote Sensing*, 38(21), 5934–5951.
- Li, X., & Li, D. (2014). Can night-time light images play a role in evaluating the Syrian Crisis? *International Journal of Remote Sensing*, 35(18), 6648–6661.
- Lin, J., & Li, X. (2014). MFOZ planning of Dongguan based on spatial autocorrelation by using genetic algorithms. *Geographical Research*, 33(2), 349–357.
- Lin, J., & Li, X. (2016). Knowledge transfer for large-scale urban growth modeling based on formal concept analysis. *Transactions in GIS*, 20(5), 684–700.
- Liu, W. (2015). Scientific understanding of the Belt and Road Initiative of China and related research themes. *Progress in Geography*, 34(5), 538–544.
- Liu, J., Zhang, Z., Xu, X., Kuang, W., Zhou, W., Zhang, S., et al. (2010). Spatial patterns and driving forces of land use change in China during the early 21st century. *Journal of Geographical Sciences*, 20(4), 483–494.
- Liu, Z., He, C., Zhang, Q., Huang, Q., & Yang, Y. (2012). Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landscape and Urban Planning*, 106(1), 62–72.
- Liu, J., Kuang, W., Zhang, Z., Xu, X., Qin, Y., Ning, J., et al. (2014). Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s. *Journal of Geographical Sciences*, 24(2), 195–210.
- Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., et al. (2017). Classifying urban land use by integrating remote sensing and social media data. *International Journal of Geographical Information Science*, 31(8), 1675–1696.
- Li, X., & Yeh, A. G. (2000). *Modelling sustainable urban development by the integration of constrained cellular automata and GIS*. Vol. 14 2.
- Li, X., & Yeh, A. G. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16(4), 323–343.
- Li, X., & Zhou, Y. (2017). Urban mapping using DMSP/OLS stable night-time light: A review. *International Journal of Remote Sensing*, 38(21), 6030–6046.
- Ma, T., Zhou, C., Pei, T., Haynie, S., & Fan, J. (2012). Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities. *Remote Sensing of Environment*, 124, 99–107.
- Ma, T., Zhou, Y., Zhou, C., Haynie, S., Pei, T., & Xu, T. (2015). Night-time light derived estimation of spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data. *Remote Sensing of Environment*, 158, 453–464.
- McGarigal, K., Cushman, S. A., & Ene, E. (2012). FRAGSTATS v4: Spatial pattern analysis program for categorical and continuous maps. Computer software program produced by the authors at the university of Massachusetts. Amherst. Available at the following web site <http://www.umass.edu/landeco/research/fragstats/fragstats.html>.
- Moghadam, H. S., & Helbich, M. (2013). Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model. *Applied Geography*, 40, 140–149.
- National Development & Reform Commission (2015). *Vision and actions on jointly building silk road economic belt and 21st-century maritime silk road*. URL: [http://en.ndrc.gov.cn/newsrelease/201503/t20150330\\_669367.html](http://en.ndrc.gov.cn/newsrelease/201503/t20150330_669367.html).
- Northam, R. M. (1979). Urban geography. *Routledge*, 23(2), 430–444.
- Perica, S., & Foufloula Georgiou, E. (1996). Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. *Journal of Geophysical Research: Atmosphere*, 101(D21), 26347–26361.
- Petitjean, F., Ketterlin, A., & Gançarski, P. (2011). A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition*, 44(3), 678–693.
- Pontius, R. G., Boersma, W., Castella, J. C., Clarke, K., Nijs, T., Dietzel, C., et al. (2008). Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science*, 42(1), 11–37.
- Poston, D. L., Jr., & Yaukey, D. (2013). The population of modern China. *Springer Science & Business Media*.
- Qi, W., Liu, S., & Zhao, M. (2015). Study on the stability of Hu Line and different spatial patterns of population growth on its both sides. *Acta Geographica Sinica*, 10(10), 139–150.
- Qi, W., Liu, S., Zhao, M., & Liu, Z. (2016). China's different spatial patterns of population growth based on the “Hu Line”. *Journal of Geographical Sciences*, 26(11), 1611–1625.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(2), 53–65.
- Shan, J. (2008). Fuzzy inference guided cellular automata urban-growth modelling using multi-temporal satellite images. *International Journal of Geographical Information Science*, 22(11–12), 1271–1293.
- Sheppard, E. (2016). *Limits to globalization: The disruptive geographies of capitalist development*. Oxford University Press.
- Shi, P., Chen, J., & Pan, Y. (2000). Landuse change mechanism in shenzhen city. *Acta Geographica Sinica*, 55(2), 151–160.
- Sutton, P. C., Taylor, M. J., & Elvidge, C. D. (2010). *Using DMSP OLS imagery to characterize urban populations in developed and developing countries*. Dordrecht: Springer.
- The State Council. Premier promotes urbanization. 2014. URL: [http://english.gov.cn/premier/news/2014/11/28/content\\_281475016489266.htm/Accessed 28 November 2018](http://english.gov.cn/premier/news/2014/11/28/content_281475016489266.htm/Accessed 28 November 2018).
- Tu, W., Cao, J., Yue, Y., Shaw, S. L., Zhou, M., Wang, Z., et al. (2017). Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. *International Journal of Geographical Information Science*, 31(12), 2331–2358.
- Tu, W., Hu, Z., Li, L., Cao, J., Jiang, J., Li, Q., et al. (2018). Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sensing*, 10(1), 141.
- UN, D. (2015). *World urbanization prospects: The 2014 revision*. New York, NY, USA: United Nations Department of Economics and Social Affairs, Population Division.
- Wang, K., & Deng, Y. (2016). Can new urbanization break through the Hu Huanyong line: further discussion on the geographical connotations of the Hu Huanyong line. *Geographical Research*, 35(5), 825–835.
- Wang, M., He, G., Zhang, Z., Wang, G., Zhang, Z., Cao, X., et al. (2017). Comparison of spatial interpolation and regression analysis models for an estimation of monthly near surface air temperature in China. *Remote Sensing*, 9(12), 1278.

- Wu, J., & Wang, Z. (2008). Agent-based simulation on the evolution of population geography of China during the past 2000 years. *Acta Geographica Sinica*, 63(2), 185.
- Xu, G., Jiao, L., Liu, J., Shi, Z., Zeng, C., & Liu, Y. (2019). Understanding urban expansion combining macro patterns and micro dynamics in three Southeast Asian megacities. *Science of the Total Environment*, 660, 375–383.
- Yang, B. (2014). Study on population distribution pattern at the county level of China. *Northwest Population Journal*, 03, 33–36.
- Yang, J., Xie, P., Xi, J., Ge, Q., Li, X., & Kong, F. (2013). Spatiotemporal simulation of tourist town growth based on the cellular automata model: the case of Sanpo Town in Hebei Province. *Abstract and Applied Analysis*, 2013, 975359 7 pages.
- Yan, X., Mao, J., & Pu, J. (2006). Research on the human dimensions of land use changes in the mega-urban region: A case study of the Pearl River Delta. *Acta Geographica Sinica*, 61(6), 613–623.
- Yao, Y., Liu, X., Zhang, D., Liang, Z., & Zhang, Y. (2017a). Simulation of Urban Expansion and Farmland Loss in China by Integrating Cellular Automata and Random Forest. *arXiv*. 1705.05651.
- Yao, Y., Liu, X., Li, X., Liu, P., Hong, Y., Zhang, Y., et al. (2017b). Simulating urban land-use changes at a large scale by integrating dynamic land parcel subdivision and vector-based cellular automata. *International Journal of Geographical Information Science*, 31(12), 2452–2479.
- Yao, Y., Chen, D., Chen, L., Wang, H., & Guan, Q. (2018). A time series of urban extent in China using DMSP/OLS nighttime light data. *PLoS one*, 13(5), e0198189.
- Yuan, M., & Qin, F. (2016). The future of Chinese cities and urban planning with "one Belt, one road" strategy. *Planners*, 32(02), 5–10.
- Zhang, Q., Pandey, B., & Seto, K. C. (2016). A robust method to generate a consistent time series from DMSP/OLS nighttime light data. *IEEE Trans. Geoscience and Remote Sensing*, 54(10), 5821–5831.
- Zhang, Q., & Seto, K. C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment*, 115(9), 2320–2329.
- Zhang, Y. Y., Song, Y. J., & Zhang, C. Y. (2015). The new urbanization and the possibility of breaking through the "HU line". *Journal of East China Normal University*, (02), 101–112.
- Zhang, J., Wang, L., & Wang, S. (2012). Financial development and economic growth: Recent evidence from China. *Journal of Comparative Economics*, 40(3), 393–412.
- Zhang, Z., Wang, C., Zhang, H., Tang, Y., & Liu, X. (2018). Analysis of permafrost region coherence variation in the Qinghai-Tibet Plateau with a high-resolution TerraSAR-X image. *Remote Sensing*, 10(2), 298.
- Zhao, X., Wu, D., Rong, X., & He, Y. (2007). Study on industry transformation development of Resource-Based City: A Case Study of Jining. *Geography and Geo-Information Science*, 23(6), 87–91.
- Zhao, N., & Samson, E. L. (2012). Estimation of virtual water contained in international trade products using nighttime imagery. *International Journal of Applied Earth Observation and Geoinformation*, 18(18), 243–250.
- Zhou, Y., Smith, S. J., Elvidge, C. D., Zhao, K., Thomson, A., & Imhoff, M. (2014). A cluster-based method to map urban area from DMSP/OLS nightlights. *Remote Sensing of Environment*, 147, 173–185.
- Zhuo, L., Ichinose, T., Zheng, J., Chen, J., Shi, P. J., & Li, X. (2009). Modelling the population density of China at the pixel level based on DMSP/OLS non-radiance-calibrated night-time light images. *International Journal of Remote Sensing*, 30(4), 1003–1018.