Global and Planetary Change 128 (2015) 90–102

Contents lists available at ScienceDirect

Global and Planetary Change

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

Research on reconstructing spatial distribution of historical cropland over 300 years in traditional cultivated regions of China

Yang Xuhong [a], Jin Xiaobin [a,b,*], Guo Beibei [a], Long Ying [b,c], Zhou Yinkang [a,b]

[a] School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China
[b] Natural Resources Research Center of Nanjing University, Nanjing 210023, China
[c] Beijing Institute of City Planning, Beijing 100045, China

ARTICLE INFO

Article history:
Received 18 June 2014
Received in revised form 19 January 2015
Accepted 17 February 2015
Available online 24 February 2015

Keywords:
China traditional cultivated region
historical cultivated land
spatial distribution
reconstruction
LUCC

ABSTRACT

Constructing a spatially explicit time series of historical cultivated land is of upmost importance for climatic and ecological studies that make use of Land Use and Cover Change (LUCC) data. Some scholars have made efforts to simulate and reconstruct the quantitative information on historical land use at the global or regional level based on "top–down" decision-making behaviors to match overall cropland area to land parcels using land arability and universal parameters. Considering the concentrated distribution of cultivated land and various factors influencing cropland distribution, including environmental and human factors, this study developed a "bottom–up" model of historical cropland based on constrained Cellular Automaton (CA). Our model takes a historical cropland area as an external variable and the cropland distribution in 1980 as the maximum potential scope of historical cropland. We selected elevation, slope, water availability, average annual precipitation, and distance to the nearest rural settlement as the main influencing factors of land use suitability. Then, an available labor force index is used as a proxy for the amount of cropland to inspect and calibrate these spatial patterns. This paper applies the model to a traditional cultivated region in China and reconstructs its spatial distribution of cropland during 6 periods. The results are shown as follows: (1) a constrained CA is well suited for simulating and reconstructing the spatial distribution of cropland in China's traditional cultivated region. (2) Taking the different factors affecting spatial pattern of cropland into consideration, the partitioning of the research area effectively reflected the spatial differences in cropland evolution rules and rates. (3) Compared with "HYDE datasets", this research has formed higher-resolution Boolean spatial distribution datasets of historical cropland with a more definitive concept of spatial pattern in terms of fractional format. We conclude that our reconstruction is closer to the actual change pattern of the traditional cultivated region in China.

1. Introduction

A large amount of research has proven that the Land Use and Cover Change (LUCC) resulting from human activities can have significant effects on climatic and ecological changes at the regional and global scales (Brovkin et al., 2004; Goldewijk and Navin, 2004; Ye and Fu, 2004; Foley et al., 2005). The drastic changes in human demographics and methods of production that were precipitated by the industrial revolution have made these anthropogenic influences much more conspicuous (Lambin et al., 2001; Cao et al., 2013). The study of global change and LUCC throughout the international academic world has continuingly intensified in recent years (Feddema et al., 2005; Shi et al., 2007; Voldoire et al., 2007). In the 1990s, a LUCC research project jointly initiated by the International Geosphere–Biosphere Programme (IGBP) and the International Human Dimensions Programme (IHDP) on Global Environmental Change stressed that the history of land use changes in the past must be reconstructed by all available means. This triggered a wave of research into the land cover changes throughout history. Reconstruction of historical data on land cover, particularly high-precision spatial data, has drawn extensive attention from scholars (Ramankutty and Foley, 1999; Goldewijk, 2001; Zhu et al., 2012; Tian et al., 2014).

Representative achievements in the reconstruction of the spatial pattern of land use include the Sustainability and the Global Environment (SAGE dataset) and the Historical Database of the Global Environment (HYDE dataset) established by Ramankutty and Foley (1999) and Goldewijk (2001), respectively. The SAGE dataset was based on the modern pattern of global land use and reconstructed the global cultivated land distribution in the period from 1700 AD to 1992 AD with a spatial resolution of 0.5° × 0.5°. HYDE dataset has been issued in four versions to date. The latest version (HYDE 3.1) takes into account factors such as population, topographic slope, distance to rivers, urban areas distribution pattern, forest land distribution pattern, and potential vegetation. Compared to previous versions, HYDE 3.1 employs a more complex algorithm to simulate the history of changes experienced by the

* Corresponding author at: School of Geographic and Oceanographic Sciences, Nanjing University, 163 Xianlin Road, Nanjing 210023, China. Tel.: +86 13512541166 (mobile).
E-mail address: jinxb@nju.edu.cn (X. Jin).

http://dx.doi.org/10.1016/j.gloplacha.2015.02.007
0921-8181/

global cultivated land and grassland with a higher spatial resolution of 5′ × 5′ and over a greater time span (past 12,000 years). Some scholars have argued that these two datasets seem highly approximate in terms of arable land amount estimation and spatial distribution pattern when used in regional simulation research. These scholars conclude that the data acquired can only be applied at the global scale and cannot be used as the basis for regional research (Li et al., 2010). Nonetheless, the above-mentioned studies have provided reference data and usable approaches for the related studies. Scholars can use or modify such approaches to perform more intensive data reconstruction. For example, Pongratz et al. (2008) used the land use pattern in AD 1700 as a reference to reconstruct the distribution pattern of the global cultivated land and grassland in the period from AD 800 to AD 1700 with a spatial resolution of 0.5° × 0.5°. Ray and Pijanowski (2010) used artificial neural networks, GIS technologies, and a step-by-step land use conversion method to reconstruct the conversion between built, cultivation, and forest land along Muskegon River in Michigan. Liu and Tian (2010) used high-resolution satellite data and long-term historical survey data to reconstruct the spatial pattern of cropland, forest, and urban land in China during the continuous period from AD 1700 to AD 2005 with a spatial resolution of 10 km × 10 km. Due to the limitation of data, most of the studies are focused on individual regions. For example, Lin et al. (2008) constructed the agricultural population and topography gravity model with population and slope as the factors that affect the cropland distribution pattern in history. They used this model to generate the spatial distribution pattern of cropland in traditional agricultural areas of China during six historical time periods. Their results were designated as the Chinese Historical Cropland Dataset (CHCD). Li et al. (2011) used the MODIS land cover remote sensing data with the historical cultivated land gridding method to construct a dataset of the cultivated land of Yunnan Province of China in 1671 and 1827. Rectifying the amount of population and cultivated land, He et al. (2011) employed the gridding method invented by Lin et al. (2008) to reconstruct the spatial distribution pattern of cropland during the mid-period of the Northern Song Dynasty.

From the perspective of methodology, the above-mentioned studies mostly employ a "top-to-bottom" method that proceeds from quantity reconstruction to spatial pattern reconstruction. Cropland amount in a historical period is obtained from historical archives, the factors that affect spatial distribution of cultivated land are screened and quantified, land arability is determined, and then the historical amounts of cultivated land are spatially matched to areas based on levels of arability. However, the principle of continuous distribution of cultivated land suggests that a piece of arable land surrounded by cropland is more likely to be cultivated, which is consistent with the concept of Cellular Automata (CA) modeling (Long et al., 2014). Based on the modeling concept of the constrained CA (Wu, 1998; Li and Yeh, 1999; Yeh and Li, 2001; Li and Yeh, 2005), we construct a model that treats the natural environment and cultural environment in the region as constraint conditions. This article also takes into account the effects of factors such as cell states and neighborhood statistics to simulate the historical spatial changes in cultivated land.

The traditional agricultural region refers to the district mostly located in the second- and third-level terrain ladders in the terrain pattern of China which is to the east of the Hu Huan-yong population line from Aihui county (Mohe county today) in Heilongjiang Province to Teng Chong county in Yunnan Province and to the south of Yanshan Mountain of Hebei Province. In terms of administration organization system, this region refers to the 17 inland provinces that existed as administrative divisions during the 25th year of the reign of Emperor Jiaqing in the Qing Dynasty (AD 1820). This generally covers south China, southern regions of the Yangtze River, Zhili Province, Shaanxi Province excluding the regions north of the Yellow River, and Fujian Province excluding Taiwan (He et al., 2012). This region was the primary socio-economic area of China during the period and had a relatively high concentration of cultivated land. Over 300 years, despite the dynastic transitions and warfare during which farming had to be discontinued, the peripheral boundaries of cultivated land did not change significantly. Spatial changes of the cultivated land were due to temporary discontinued farming and restorative reclamation. This region is the key area in the research on reconstruction of historical spatial patterns of cultivation in China (Lin et al., 2008; He et al., 2012). This article takes the traditional agricultural area as the study region which provides more relatively detailed historical data compared to other regions. We use the modern cropland spatial pattern, the historical amount of cultivated land, land reclamation suitability, and historical population within a constrained CA model to reconstruct the 1 km × 1 km spatial pattern of cultivated land under the historical partition.

2. Materials and method

2.1. Methodology

Reconstruction of historical cultivated land dataset consists of quantity and spatial pattern reconstruction (Li et al., 2011; Zhu et al., 2012). The former refers to the statistical information on historical amount of cultivated land obtained by reading historical archives, and the latter refers to the process where the historical spatial distribution pattern of the cultivated land is restored on the basis of certain principles, assumptions, or approaches for spatial distribution (Ye et al., 2009; Li et al., 2011; Miao et al., 2012; Cao et al., 2013).

As for expansion of cultivated land in China, all Chinese studies have indicated that the amount of cultivated land in China during the period from the Early Qing Dynasty to 1980 increased in an oscillating pattern and reached its peak in 1978 to 1980 (Zhou, 2001; Ge et al., 2003, 2008; Zhang et al., 2003; Feng et al., 2005; Ge and Dai, 2005). Afterwards, the amount of cropland decreased considerably, thus the cultivated land boundary in 1980 can be viewed as the peripheral boundary of its historical distribution range.

Reclamation of arable land is highly dependent on natural conditions and usually proceeds first from land parcels that are easy to reclaim and then to land blocks that are difficult to reclaim. The land units with preferable natural conditions for farming are reclaimed earlier than the ones with unfavorable conditions (He et al., 2003). Only when the population pressure increases or a major disaster happens can the land units with high terrain, steep slopes, and/or low fertility be treated as reclaimable land (Li et al., 2011). Furthermore, the land units surrounding a cultivated land are preferable for reclamation given considerations of return on investment and farming convenience.

The reconstruction of the historical pattern of cultivated land in China's traditional agricultural region was done using the land pattern of 1980 as the maximum potential distribution range for the cultivated land in the past. The model projects past cropland via backward simulation starting with 1980. Based on the land reclamation suitability in a region, distribution of modern cultivated land, and total cultivated land and land boundaries in certain period, the cultivated land transition rules were established and a constrained CA modeling method was used to reconstruct the spatial pattern of cultivated land. A flow chart of the process is shown in Fig. 1.

2.2. Reconstruction model of historical spatial pattern of cultivated land

2.2.1. Assumptions and conception of the model

The constrained CA model used in this model has some basic assumptions that are similar to models used in the literature (Ramankutty and Foley, 1999; Bai et al., 2007; Lin et al., 2008; Pongratz et al., 2008; He et al., 2011; Li et al., 2011). They are as follows: (1) there are similarities between the historical cultivated land spatial pattern and the contemporary spatial pattern; (2) the most unsuitable farming cells were turned into non-cultivated land first (from contemporary to history); (3) cells surrounded by high ratio of non-cultivated land were turned into non-cultivated land first; (4) the range of historical arable land does not

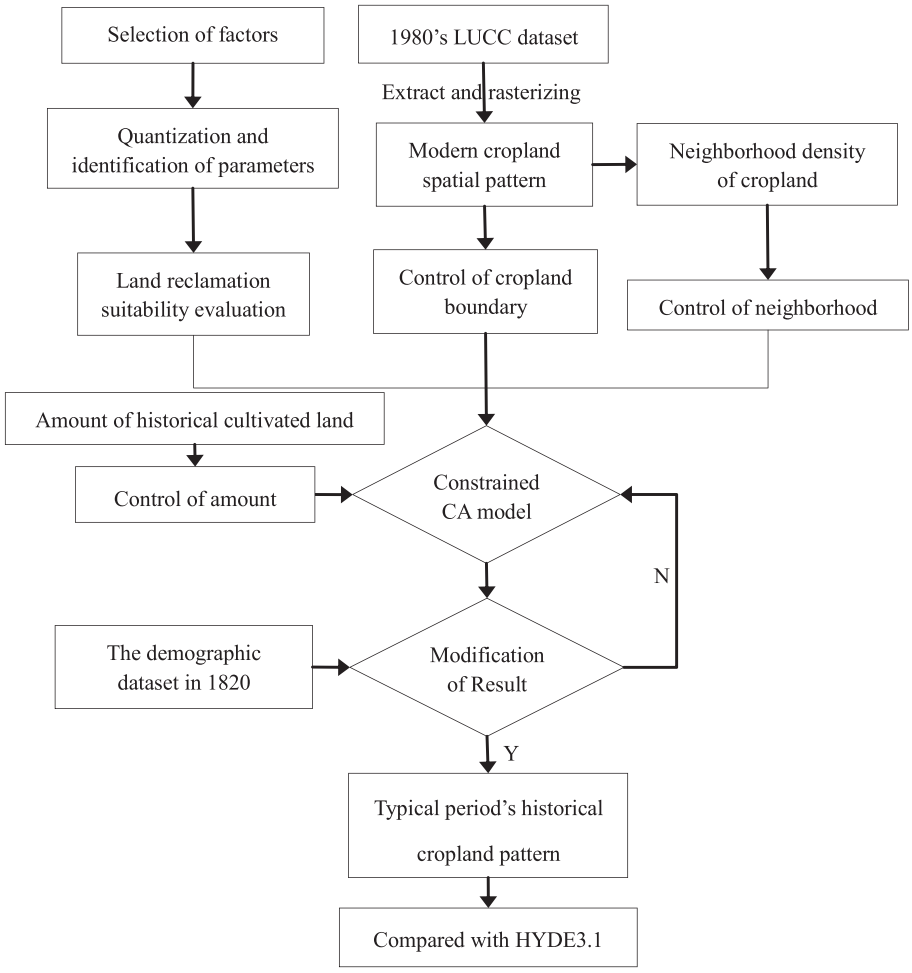


Fig. 1. Flow chart of reconstructing historical spatial pattern of cultivated land based on a constrained CA model.

exceed the scope of contemporary cultivated land, in that it is impossible for an arable land in history not be cultivated land in modern times; and (5) factors influencing land reclamation suitability do not change over time (due to data availability).

Identification of the factors that affect spatial distribution patterns of cultivated land is the basis for building the model. We make use of the results from prior efforts to achieve this aim (Liu et al., 1995; Jiao and Liu, 2004; Lin et al., 2009; Li et al., 2011, 2012; He et al., 2012; Xie et al., 2013). This research concludes two types of factors that affect the spatial distribution of cultivated land: natural factors (terrain elevation, topographic slope, river system, soil quality, etc.) and human factors (population, economic level, agricultural policies, wars, etc.). In addition, the pattern of cultivated land can also be affected by accidental factors or random events. In order to make the simulation result more realistic, a random factor needs to be introduced into the constrained CA model to address uncertainty in the evolution of cultivated land.

Based on the above-mentioned information and the constituents of the CA model (Wu, 1998; Li and Yeh, 1999; Yeh and Li, 2001), the constrained CA model used for reconstruction of the spatial pattern of cultivated land in history can be abstracted as this equation:

$$State_{ij}^{t-1} = f\begin{pmatrix} Natural_{ij}^{t}(Slope, Elevation, \ldots, River, Erosion), \\ Human_{ij}(Population, Policy, \ldots, War), \\ Neighbor_{ij}^{t}, State_{ij}^{t}, Rand_{ij}^{t} \end{pmatrix} \quad (1)$$

where the $State_{ij}^{t-1}$ of grid i of partition j at the time $t-1$ is synthetically determined by natural factors, human factors, cell state, surrounding neighborhood, and random disturbance factors. Herein, the backward deduction is based on the current period.

Natural factors are region-specific and relatively stable features, while some human factors are highly variable and difficult to quantify and spatialize. The research scale and the available data were also constrained. Therefore, this article uses elevation, slope, waterbody accessibility, multi-year average precipitation, and distance to the nearest regional rural settlement as the dominant factors that determine reclamation suitability of land and uses population as the correction factor for spatial distribution of historical cultivated land. Eq. (1) is simplified to Eq. (2).

$$State_{ij}^{t-1} = f\left(Q^{t}{}_{ij}, Neighbor_{ij}^{t}, State_{ij}^{t}, Rand_{ij}^{t}\right) \quad (2)$$

where the $State_{ij}^{t-1}$ of grid i of partition j at the time $t-1$ is synthetically determined by contribution probability of land reclamation suitability (Q_{ij}^{t}) which will be described in detail below, a neighborhood-based variable ($Neighbor_{ij}^{t}$), cell state ($State_{ij}^{t}$), and a random disturbance factor ($Rand_{ij}^{t}$).

According to properties of CA model, the basic elements of this research's constrained CA model are as follows:

(1) Lattices: the entire study area;
(2) Cells: cell size was 1 km × 1 km, spatial cells are rasterized by maximum area method in GIS;
(3) Cell states: V = 1 means cultivated lands, V = 0 means non-cultivated lands;

(4) Transition rules: this will be specifically addressed in the next section, a multi-criteria evaluation (MCE) form;
(5) Neighborhoods: Moore neighborhood, 3 cells × 3 cells, a total of eight neighboring cells;
(6) Termination of iteration: the amount of cropland cell equals to the amount of cultivated land in typical year (exogenous variable);
(7) Constraints: the five aforementioned spatial factors: elevation, slope, waterbody accessibility, multi-year average precipitation and distance to regional rural settlement.

2.2.2. Decision criterion

(1) Spatial criterion of reclamation suitability variable: People prefer to reclaim flat and fertile land parcels. Only when population pressure and demand for agricultural products increases beyond a certain point, will people gradually reclaim high, steep, and less fertile land blocks. Thus, reductions in cultivated land can be modeled by excluding cells which are less suitable for reclamation. Land reclamation suitability (S_{ij}^t) is calculated via a multi-factor synthesis method.

$$S_{ij}^t = a + \gamma_1 \times Elevation_{ij} + \gamma_2 \times Slope_{ij} + \gamma_3 \times Waterbody_{ij} + \gamma_4 \times Precipitation_{ij} + \gamma_5 \times Settlement_{ij} \quad (3)$$

$$Q_i^t = \frac{1}{1 + e^{-s_{ij}^t}} \quad (4)$$

Q_{ij}^t is the contribution probability of land reclamation suitability in grid i of partition j at the time t. $Elevation_{ij}$, $Slope_{ij}$, $Waterbody_{ij}$, $Precipitation_{ij}$, and $Settlement_{ij}$ represent the terrain elevation, slope, waterbody accessibility, multi-year average precipitation, and distance to regional rural settlement in grid i of partition j, respectively. Parameter values a and γ_1–γ_5 are the logistic regression parameters for constant, elevation, slope, waterbody accessibility, multi-year average precipitation, and distance to regional rural settlement, respectively.

(2) Spatial criterion of the neighborhood-based variable: in terms of return on investment, continuous intensive cultivation, and cultivation convenience, it is more likely that land blocks around cultivated land will be reclaimed. Thus, isolated cultivated land with low reclamation suitability within the modern pattern is more likely to have been reclaimed later than less isolated land with higher reclamation suitability. We believe that such cells should be removed first in the process of reconstructing cultivated land. Following the methods of relevant studies (Long et al., 2009, 2010), the neighborhood-based variable ($Neighbor_{ij}^t$) adopts the ratio of the number of cells with the property of cultivated land in the Moore neighborhood to the total number of neighboring cells (8) at time-step t. See the following equation:

$$Neighbor_{ij}^t = \frac{\sum_{i=1}^{8} cell_{ij}^t}{3 \times 3 - 1}. \quad (5)$$

(3) Compound decision criterion for reconstructing spatial pattern of historical cultivated land: Spatial evolution of historical cultivated land is the result of a compound decision of land reclamation suitability, cell neighborhood, and a random disturbance factor. From the spatial criteria of the above factors and reference to previous studies (Liu et al., 2006; Zhang et al., 2008), we obtained the compound decision criterion for spatial evolution of historical cultivated land:

$$P_{ij}^t = \exp\left[\omega\left(\frac{R_{ij}^t}{R_{gmax,j}^t} - 1\right)\right] \quad (6)$$

$$State_{ij}^{t-1} = \begin{cases} 0 & \text{if} \quad P_{ij}^t < Threshold_j^t \\ 1 & Else \end{cases} \quad (7)$$

where: P_{ij}^t is the final probability for the cell i of partition j to be reduced to non-cultivated land at time-step t. The smaller the value is, the higher the probability that unit cell i may be reduced to non-cultivated land. $R_{gmax,j}^t$ is the maximum value of the set of potential reclamation cells which are involved in the operation at time-step t. ω is a discrete parameter with a value range of 1–10. $State_{ij}^{t-1}$ represents the state at $t-1$ of cell i of partition j involved in the operation at time-step t, i.e., whether it is to be maintained as cultivated land. $Threshold_j^t$ is the threshold of state transition of partition j at time-step t. As the reduction rule of cultivated land is nonlinear and fluctuant, this value is gradually increased at every time step of loop iteration. See the following Eq. (8) for details. R_{ij}^t represents the potential reclamation probability of cell i of partition j under the combined action of contribution probability of land reclamation suitability and cell neighborhood:

$$Threshold_j^t = Threshold_j^{t+1} + \theta \quad (8)$$

$$R_{ij}^t = \left(\alpha \times Q_{ij}^t + \beta \times Neighbor_{ij}^t\right) \times \left[1 + \left(Rand_{ij}^t - 0.5\right)/\varepsilon\right] \quad (9)$$

where: $Threshold_j^{t+1}$ represents the threshold of state transition at $t+1$. θ is the additive constant of the threshold. Q_{ij}^t and $Neighbor_{ij}^t$ are described as above. α and β are the weight value of Q_{ij}^t and $Neighbor_{ij}^t$, respectively. β and $\alpha = 1 - \beta$ are identified by the Monoloop method (Long et al., 2009). Based on equal difference principle, we adjusted β from 0 to β_{max} (maximum weight coefficient, set based on experience), and the weight values obtained under maximum *Kappa* index are assigned to α and β. $Rand_{ij}^t$ is the random disturbance variable from 0 to 1 in the cell i of partition j at time-step t, representing influence of human factors, such as agricultural policy and war, on spatial diffusion of cultivated land. ε is a constant which indicates disturbance degree.

In conducting loop iterations in Python with the above model, the amount of cultivated land in the historical spatial pattern is generated by backward time-steps and decreases progressively. It is possible to identify the pattern of historical cultivated land in a corresponding year if the amount of reconstruction is equal to the amount of cultivated land in aimed year (exogenous variable). This represents the potential distribution of historical cultivated land under the premise of certain social economy, level of productive forces, and amount of cultivated land.

2.3. Model calibration and validation

After generating a potential spatial distribution pattern of cultivated land in typical year with the above model, it is necessary to inspect and correct the reconstructed results to make the simulated result consistent with reality.

Before 1980, agricultural production in China was dominated by small-scale cultivation by farmers with limited natural resources. This entailed family-based cultivation dependent on the local labor force. Thus, the amount of historically cultivated land that is reconstructed in various districts and cities should comply with the local labor supply level. We establish a verification function based on the ability of the labor force to cultivate cropland:

$$\mu_{nj} = \frac{P_{nj} \times m \times lc}{A_{nj}} \quad (10)$$

where: P_{nj} refers to the total population of city j in year n; m refers to the proportion of population engaged in agricultural labor; lc refers to the cultivated land area available for labor force per unit; A_{nj} refers to the

amount of cultivated land in city j in the year n generated by the model; μ_{nj} represents the index of available cultivated land for the labor force in city j in the year n. When $\mu_{nj} > 1$, it indicates that the local labor force level can support the amount of cultivated land to be reconstructed. Otherwise, it is necessary to correct the discrete parameter ε of the above model.

2.4. Model operation

In the state transition rules, the weight of influencing factors is an important component for identifying model parameters. However, the influence mechanisms are different in various cropland evolution regions and thus require independent parameter identification. As the dependent variable for analysis, the state transition of cropland (from cropland to non-cropland the state transition is denoted as 1, while no state transition is denoted as 0) is a binary classification constant that does not meet the conditions of normal distribution, which is a requirement of the logistic regression analysis method. The independent variables correspond to the spatial distribution data of influence factors. The SAMPLE tool of ESRI ArcGIS 9.3 is used for either partial or entire sampling of the spatial data corresponding to the independent and dependent variables. Regression analysis is then carried out in an SPSS environment to obtain the contribution of each influence factor on the state transition of cropland (i.e. weight) to generate the input conditions of the constraint CA model.

Based on the use of historical data for calibrating the model parameters, the model operates as shown below. (1) The environmental variables of the model, cropland quantity in different historical periods, boundary control conditions, influence factors, and weighting coefficients are set. (2) Based on the revised data of the historical cropland in various stages, we calculate the cropland quantity with state transitions in different partitions during different periods. (3) According to the results of the model parameter identification in various partitions, the logistic regression method is used to calculate the cropland suitability and conversion probability of the cropland cell. (4) The spatial backtracking of the historical cropland is carried out in each partition by means of iterating the model until completing the entire simulation process. (5) The backward simulation is completed to obtain the historical cropland in China's traditional cultivated regions.

3. Reconstruction of historical spatial pattern of arable land in China's traditional cultivated region

3.1. Study area

This study aims to investigate the spatiotemporal changes of cropland in China's traditional cultivated regions during the time period of 1661 to 1980. The regional administrative system refers to 17 mainland provinces under the administrative divisions in 1820, generally including south China, southern regions of the Yangtze River, Zhili Province, Shanxi Province excluding the regions north of the Yellow River, and Fujian Province excluding Taiwan. The terrain inclines from southwest to northeast and includes the North China Plain, Middle–Lower Yangtze River Plain, Sichuan Basin, southeast hills, and coastal beaches. The regional climate types are diversified, transitioning between tropical monsoon climate, subtropical monsoon climate, and temperate monsoon climate from south to north. The precipitation gradually decreases from the coast to inland, transitioning from 2526 mm in the south to 293 mm in the north. In this region, the Yangtze River and Yellow River are the two major river systems with complex and large tributary rivers that provide strong water resource support for local agricultural production. The soil types are diverse in the region, including (from south to north) laterite, red soil, yellow soil, yellow-brown soil, brown soil, and cinnamon soil.

In order to maintain the continuity of the data, the provincial grouping method is used in this study to adjust the administrative boundaries of the Qing Dynasty and Republic of China into the modern administrative boundaries (2012), i.e., the current Beijing, Tianjin municipalities, and Hebei Province are merged into the Beijing–Tianjin–Hebei region; Jiangsu Province and Shanghai municipality into the Shanghai–Jiangsu region; Guangdong Province and Hainan Province into the Guangdong–Hainan region; and Sichuan Province and Chongqing municipality into the Sichuan–Chongqing region. Therefore, China's traditional cultivated regions in this study include 17 provinces or regions (excluding sporadic islands), i.e. Anhui, Sichuan–Chongqing region, Fujian, Guangxi, Guizhou, Henan, Hubei, Hunan, Shanghai–Jiangsu region, Jiangxi, the Beijing–Tianjin–Hebei region, Shandong, Shanxi, Shaanxi, Guangdong–Hainan region, Yunnan, and Zhejiang (Fig. 2).

3.2. Data resources and processing

Taking into consideration connecting our model with feasibility of the model in calculation, the scale of relevant research (climate simulation and carbon cycle), as well as convenience of overlays with existing data, the spatial resolution of this research was set as 1 km × 1 km.

(1) The modern cropland data: 1:100,000 land use data of China's traditional cultivated regions in the 1980s were obtained from the Earth System Science Data Sharing Network (http://www.geodata.cn) and were used as the modern cropland data in this study. The maximum area method was applied to divide the cropland area into 1 × 1 km grid cells with a total of 1,225,720 cells in the study area.
(2) Amount of historically cultivated land: Cultivated land from 1661 to 1952 was based on the research of Cao et al. (2013, 2014) combined with the 1933 data from "China's statistical analysis of land issues" (Statistical Bureau of the National Government of China, 1936).
(3) Demographic data: Demographic data for 1820 were adopted from the Chinese Population Geographic Information System, which was taken from the Harvard University CPGIS database (CPGIS, 2007). It is adjusted to current divisions via the administrative area ratio method.
(4) Topographic data: The DEM used in this study was provided by the Data Sharing Infrastructure of Earth System Science which was downloaded from http://www.geodata.cn.
(5) River network and lake data: Basic geographic data of China's river network and lakes were obtained from the National Geomatics Center of China (1:4,000,000). They were divided into grades 1 to 5 according to river level.
(6) Regional rural settlement: Regional rural settlement data for 1820 were obtained from the Chinese Historical Geographic Information System, which was taken from the Harvard University CHGIS database (CHGIS, 2007).
(7) Multi-year average precipitation: The results from the "Eco-environmental background level temperature and humidity data level construction" project are adopted, which was downloaded from http://www.geodata.cn.

3.3. Quantization and identification of parameters

3.3.1. Parameter quantization

Normalization processing is required for all factors to facilitate regression analysis and eliminate dimensional disturbance. Land reclamation suitability is defined as follows: 1 is the most suitable and 0 is the least suitable.

(1) Elevation standardization
Heat and moisture vary with the elevation, which causes a vertical differentiation in land reclamation. When the elevation reaches a certain altitude, the hydrothermal conditions become

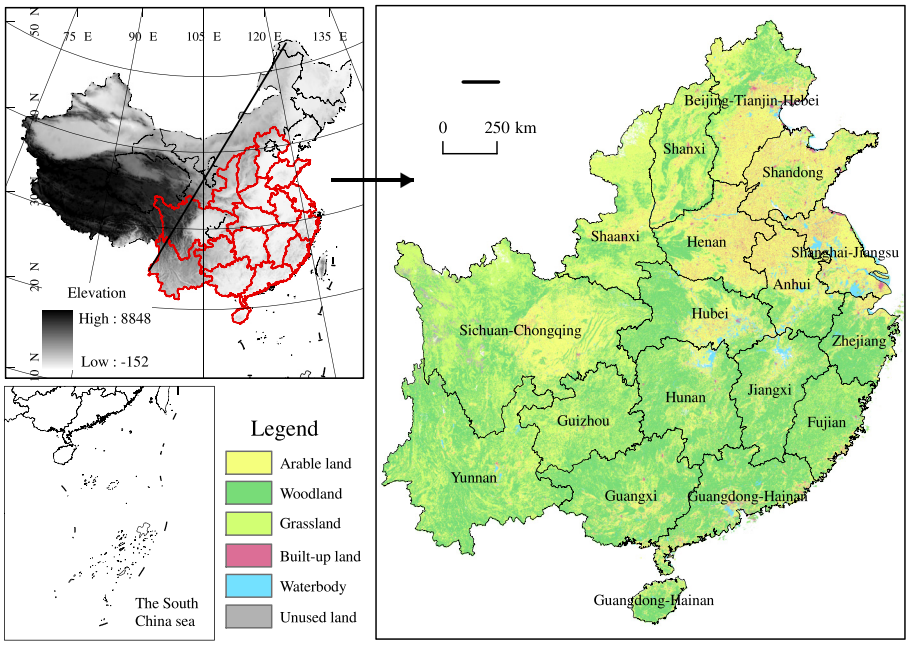


Fig. 2. The location and land cover of traditional cultivation regions in China (1980).

the dominant factor restricting the growth of crops (Lin et al., 2008; Li et al., 2011, 2012). Thus, the distribution of cultivated land presents obvious vertical zonality. The elevation is normalized using the following equation:

$$Elevation_{ij} = \frac{max\left(e_{ij}\right) - e_{ij}}{max\left(e_{ij}\right)} \tag{11}$$

where $Elevation_{ij}$ is the normalization value of elevation, e_{ij} signifies the initial value of grid i of partition j, and $max(e_{ij})$ is the maximum value of elevation within the partition j.

(2) Slope standardization

Terrain slope plays an important role in deciding the land use type. Generally, the higher the slope is, the lower the arability will be. Thus, the slope is normalized using the following equation:

$$Slope_{ij} = \frac{max\left(s_{ij}\right) - s_{ij}}{max\left(s_{ij}\right)} \tag{12}$$

where $Slope_{ij}$ is the normalization value of slope, s_{ij} signifies the initial value of grid i of partition j, and $max\ (s_{ij})$ is the maximum value of slope within the partition j.

(3) Waterbody standardization

Water source is one of the key factors affecting agricultural cultivation and its spatial distribution largely determines the spatial pattern of cultivated land. Plots that are close to river channels have convenient access to water. In the period

Table 1
Result of parameter identification about influence factors of cultivated land.

Provinces/regions	Total prediction accuracies/%	a	γ_1	γ_2	γ_3	γ_4	γ_5
Shaanxi	68.3	−5.680	6.323	1.352	−0.922	−1.845	1.491
Shanxi	65.8	−7.285	6.075	3.433	2.868	−4.499	0.810
Henan	75.1	−11.196	7.847	4.023	1.265	−1.107	0.687
Anhui	74.8	−6.583	6.369	2.228	1.850	−4.604	0.244
Hubei	74.9	−12.771	8.056	2.989	3.160	−0.371	0.202
Sichuan–Chongqing	84.3	−11.929	10.418	2.789	1.865	−3.009	0.833
Jiangxi	70.3	−9.680	8.588	0.693	1.844	−1.273	0.351
Yunnan	76.7	−7.239	3.840	2.197	2.325	−1.977	0.717
Beijing–Tianjin–Hebei	73.0	−8.071	5.574	8.223	1.858	−6.751	−0.844
Shandong	73.7	−8.152	3.486	2.804	1.747	1.082	1.507
Shanghai–Jiangsu	72.4	−4.338	6.405	0.950	−3.507	−0.566	1.479
Zhejiang	78.0	−7.869	8.077	2.408	0.097	−3.417	0.162
Guizhou	65.6	−2.133	−0.249	0.391	1.215	−0.199	0.766
Hunan	69.7	−3.573	4.720	1.039	0.789	−4.339	0.318
Fujian	78.2	−4.006	4.862	1.215	−1.067	−2.117	0.280
Guangxi	74.3	−7.688	4.461	4.394	−0.243	−2.304	0.363
Guangdong–Hainan	69.7	−7.338	6.688	0.773	−0.671	−0.746	1.018

Note: $P < 0.001$.

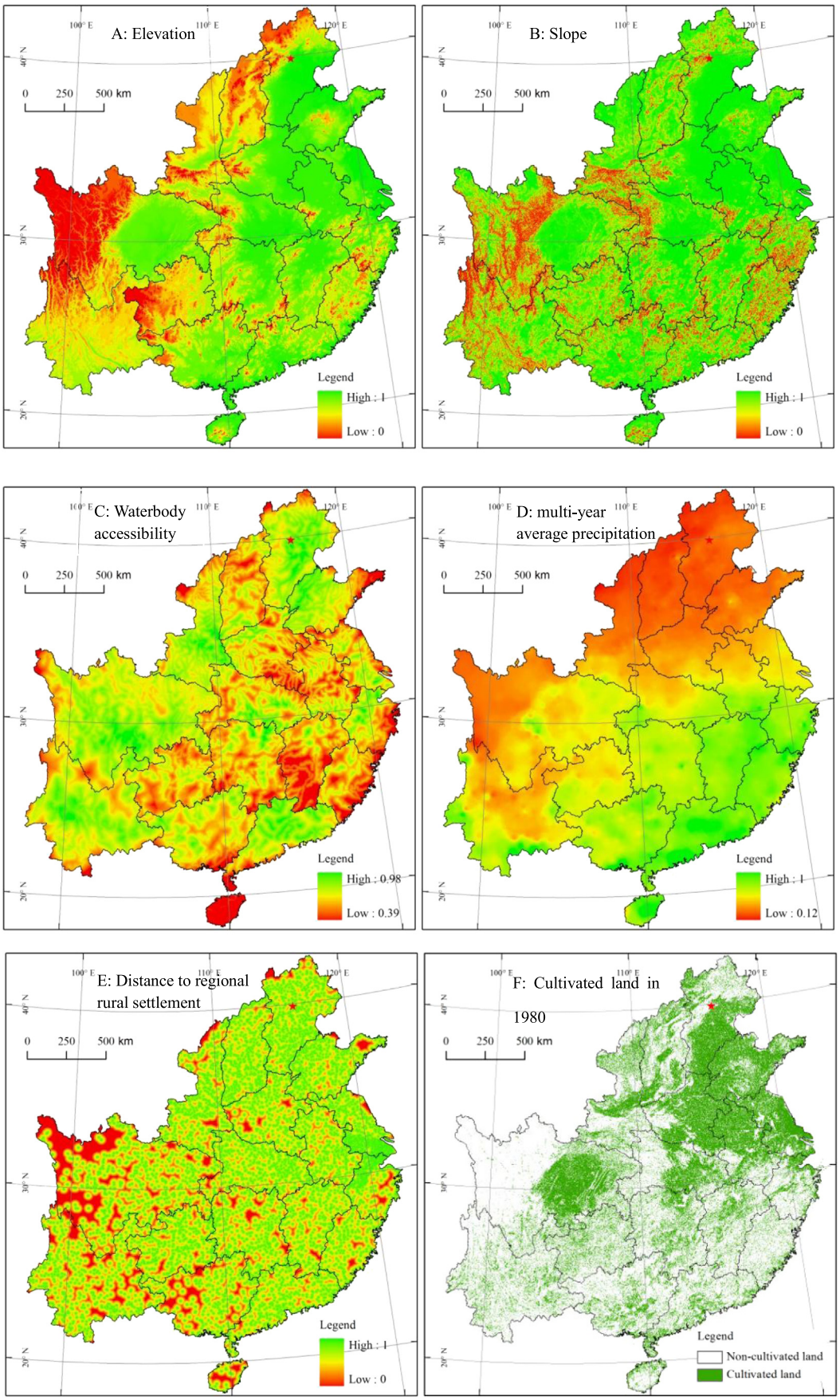


Fig. 3. Spatial pattern of basic data to reconstruct cropland.

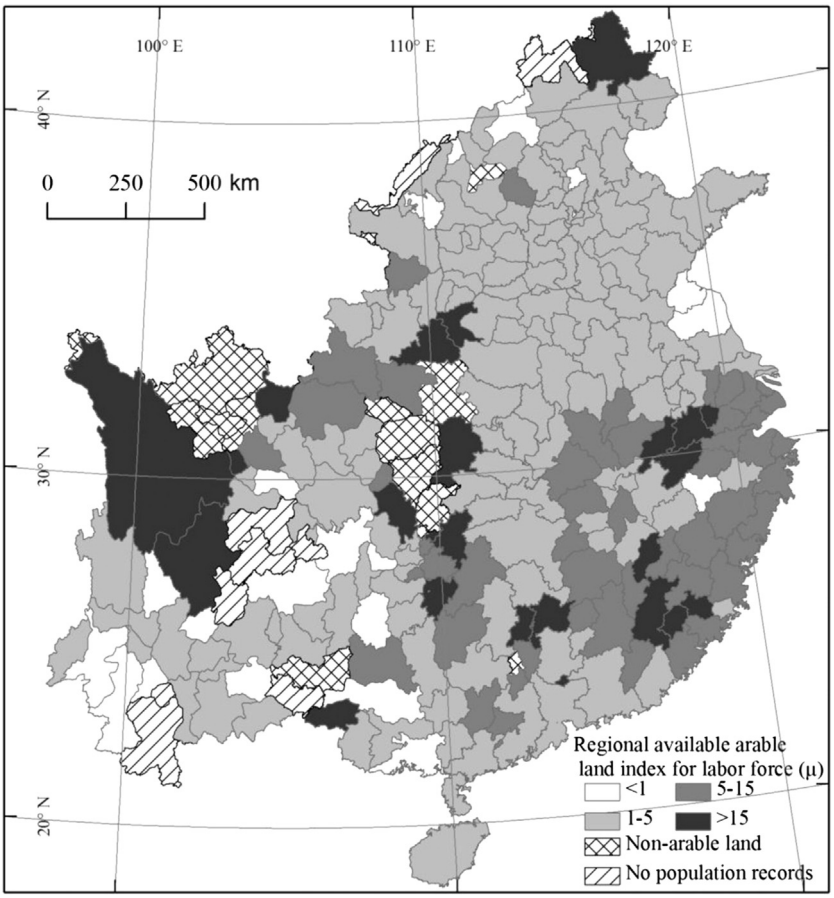


Fig. 4. Spatial distribution of the available cultivated land index based on labor force (1820).

dominated by small-scale farmers with low productivity, such plots are more likely to be reclaimed and maintained as cultivated land due to the lack of large-scale irrigation and water conservation measures (Liu et al., 2011). Waterbody accessibility is subject to an exponentially weighted normalization process based on exponential decay theory of spatial influence distance (Zhang et al., 2008). See the equation below:

$$Waterbody_{ij} = a_1 e^{-b_{1j}d_{1j}} + a_2 e^{-b_{2j}d_{2j}} + a_3 e^{-b_{3j}d_{3j}} + a_4 e^{-b_{4j}d_{4j}} + a_5 e^{-b_{5j}d_{5j}} + a_6 e^{-b_{6j}d_{6j}} \quad (13)$$

where $Waterbody_{ij}$ is the waterbody accessibility and is defined as the spatial distance exponential weight of a river or lake adjacent to grid i of partition j; a_1, a_2, a_3, a_4, a_5, and a_6 are respectively the spatial influence weights of rivers of level 1 to level 5 and adjacent lakes. These weights were obtained via the Analytic Hierarchy Process (AHP) method in conjunction with expert opinion; b_{1j}, b_{2j}, b_{3j}, b_{4j}, b_{5j}, and b_{6j}, are respectively the spatial influence attenuation coefficients of rivers of all levels and lakes, and depend on the spatial influence distance; and d_{1j}, d_{2j}, d_{3j}, d_{4j}, d_{5j}, and d_{6j} are respectively the Euclidean distances between rivers of all levels and lakes to the nearest plot i of partition j.

(4) Precipitation standardization

In addition to considering the distance of water drawing in agricultural plantation, the traditional croplands are located on the fields having benign climate. Atmospheric precipitation is an important water supply for most croplands, thus multi-year average precipitation is introduced into the model as a natural factor, and is quantified and spatialized by the maximum standardization method:

$$Precipitation_{ij} = \frac{p_{ij}}{\max\left(p_{ij}\right)} \quad (14)$$

where $Precipitation_{ij}$ is the standardized value of the multi-year average precipitation in grid i of partition j; p_{ij} is the initial value of grid i of partition j; and $max(p_{ij})$ is the maximum value of the partition j.

(5) Settlement standardization

Distance to regional rural settlement is an important factor in the socio-economic factor set which will affect the farming radius and range. For farming and management convenience, the land blocks located far away from towns and villages have small probabilities for continuous farming, and those located near towns and villages are suitable for continuous farming. Herein, the maximum negative standardized method is used to normalize $Settlement_{ij}$ value, as shown as below:

$$Settllement_{ij} = \frac{max\left(d_{ij}\right) - d_{ij}}{max\left(d_{ij}\right)}. \quad (15)$$

In the equation, d_{ij} is the distance of grid i in partition j from the nearest rural settlement, and $max(d_{ij})$ is the maximum distance of any cell from the rural settlement of the partition j.

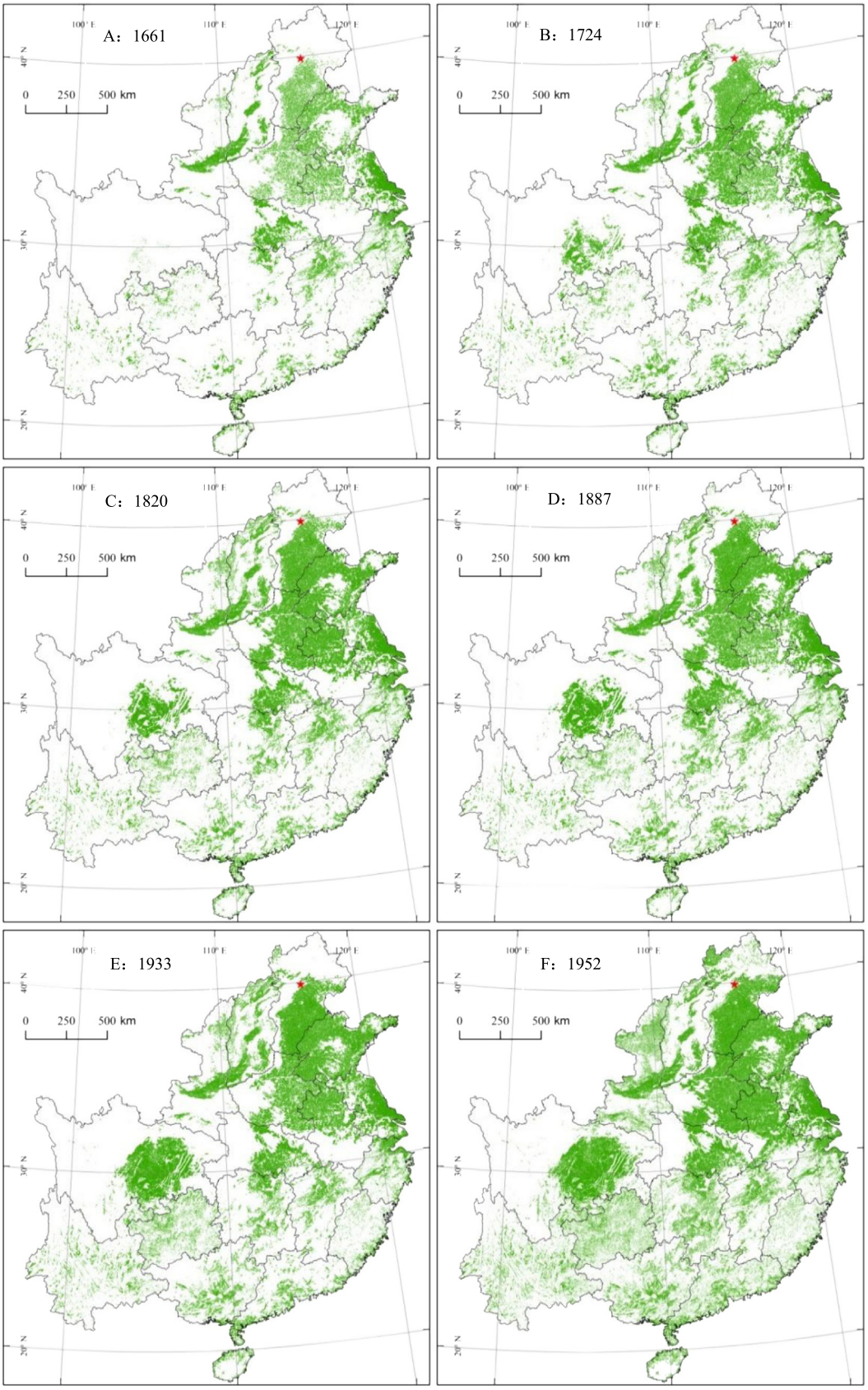
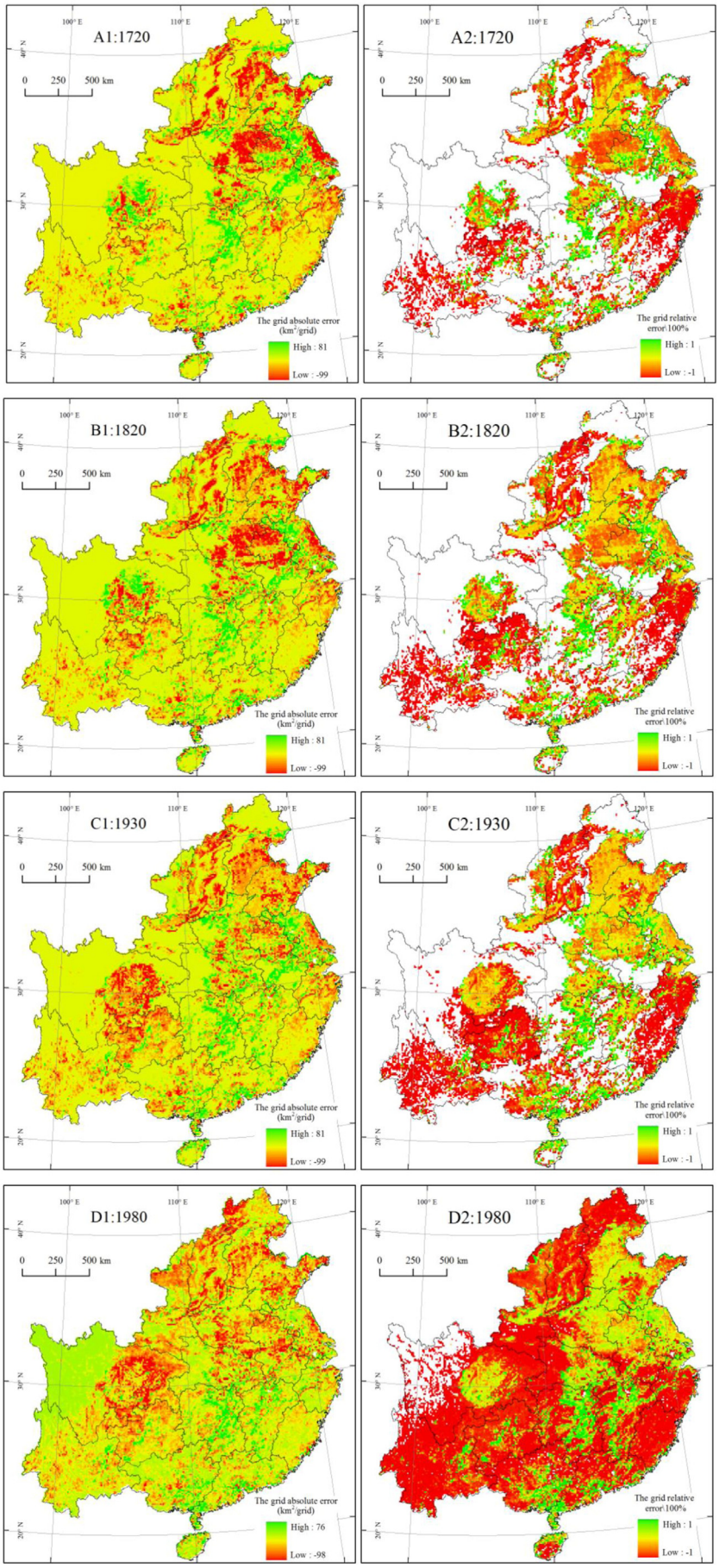


Fig. 5. The reconstruction results of cropland in the traditional cultivated region of China from 1661 to 1952.


A1:1720
A2:1720
B1:1820
B2:1820
C1:1930
C2:1930
D1:1980
D2:1980
100° E
110° E
120° E
40° N
30° N
20° N
0 250 500 km
The grid absolute error (km²/grid)
High : 81
Low : -99
High : 76
Low : -98
The grid relative error\100%
High : 1
Low : -1

3.3.2. Parameter identification

An analysis of modern cultivated land spatial data is required to acquire the relevant parameter value of the state transition rule. In this paper, the cultivated land distribution range of the research area in 1980 is the maximum spatial distribution of historical cultivated land used for parameter identification:

(1) Parameter identification of basic data: the parameter of basic data primarily refers to the *Waterbody* variable. By a combination of expert opinions and AHP, we calculated the spatial influence weights of the a_1, a_2, a_3, a_4, a_5, and a_6 as 0.2996, 0.2267, 0.1799, 0.1039, 0.0841, and 0.1059, respectively. The matrix consistency coefficients were determined to be 0.0403 which are both less than 0.1 and meet the research requirements.

(2) We applied the Binary Logistic module of SPSS 18 to perform a regression of the influence coefficients of 5 land reclamation suitability factors (independent variable) excluding the *Neighbor* variable. The current land use pattern was chosen as the dependent variable. When the grid unit is cropland it is defined as 1, otherwise it is 0. The results are shown in Table 1. It can be seen from the results that all factor values pass the test. The prediction accuracy of the logistic regression model can be measured by precision of regression equation. The total prediction accuracies of the modeling dataset in the various partitions are all greater than 65%, and that in the Sichuan–Chongqing Region reaches 84.30%, indicating higher model prediction accuracy and relatively stable predictive ability. The influence degree and effects of influence factors on the spatial distribution of croplands are not consistent between the various regions. For example, the regression parameters of the slope factors (γ_2) in various provinces (regions) were all positive, indicating a negative correlation between the spatial distribution and slope of cropland, i.e., the larger the slope is, the lower the likelihood that the plot will be cultivated. However, the γ_2 was different in size, indicating differences in the influence degree of the slope in the various regions, e.g., the regression parameters of Henan, the Beijing–Tianjin–Hebei Region, and Guangxi were larger with more significant negative correlations. The regression parameter of the elevation factor (γ_1) was negative only in Guizhou Province (-0.249). The regression parameter of the distance from rural settlement (γ_5) is only negative in the Beijing–Tianjin–Hebei Region (-0.844). The regression parameter of multi-year average precipitation (γ_4) is positive only in Shandong Province (1.082). Thus, the results indicate that in general the elevation, distance from rural settlement, and multi-year average precipitation factors are negatively correlated with the spatial pattern of cropland, while between regions the influence of factors varies.

The identification results of the above parameters are integrated for spatial quantification and visualization of the spatial factors. Larger values of spatial influence factors entail greater land suitability and greater probability of cultivation. The maximum area method is used for to map values onto the spatial pattern of cropland in 1980, in which the cropland and non-cropland are represented by 1 and 0 respectively (Fig. 3).

(3) Identification of α and β values: the starting year of cultivated land conversion is set as 1980. Adjust α and β (the initial value is set as 0.25) values to generate the first cultivated land distribution pattern layer by applying the Monoloop method. Then identify the weight parameters (α and β values) where the maximum *Kappa* index exists (0.80, 0.20).

(4) Result correction: We select the city's demographic dataset in CPGIS and model the amount of cultivated land of each city using population data from 1820 in conjunction with Eq. (10). See references (Cao et al., 2013) for parameter setting and the figure below for results.

The analysis results show that in the traditional 256 cultivated regions, there are 9 regions with no population records, 20 regions with regional available cultivated land index based on labor force (μ) < 1, 141 regions with indices between 1 and 5, 48 regions with indices of >15, and 14 regions with no unallocated cropland. It can be seen that the μ ranges between 1 and 15. Regions with these values account for 73.83% of the total. Taking into account the total available population and total regional population, this study considers this as a reasonable value range for consistent population–cropland relation. The regions with $\mu < 1$ account for 7.81% of the total regions, and are mainly distributed in the Middle–Lower Yangtze River Plain and Sichuan Basin. These areas have flat terrains, abundant water resources, high soil fertility, and have always been concentration centers of cropland in China. However, the Yunnan–Guizhou Plateau in southwestern China has undulating terrain, mature karst landforms, wider distribution of terraced areas, and a lower total population. Thus, the result of $\mu < 1$ is acceptable for this area. Finally, the regions with $\mu > 15$ and unassigned values accounted for 9.38% of the total. In some of these regions, the model lacks explanatory power to a certain extent as the population is too high or it is difficult for the quantity of cropland to be allocated to support the regional population. The above analysis shows that it is reasonable to adjust the discrete parameters to form the spatial pattern of cultivation regions during historical periods (Fig. 4).

3.4. Results and analysis

3.4.1. Results

Based on the arability factor, neighborhood factor, and identification parameter of the land, the spatial distribution pattern of historical cultivated land over 300 years is created by backward modeling of historical cultivated land space. The spatial pattern distribution of historically cultivated land for that time period is then obtained according to the amount of historical cultivated land for each time period after correction (Fig. 5).

The historical croplands in the traditional cultivated regions reconstructed in this study are mainly distributed in the regions of the North China Plain, Middle–Lower Yangtze River Plain, Guanzhong Plain, Sichuan Basin, Dongting Lake Plain, eastern coastal beach areas, and southeast hilly areas. These regions are characterized by suitable temperature, abundant water resources, fertile soil, and flat terrain. Over time, the spatial pattern of cropland shows an expansion from the main reclamation area to the periphery, and the reclamation rate increases inside the main reclamation area.

3.4.2. Analysis of results

In order to further analyze the reasonableness of the reconstruction results, the results are compared with the current mainstream dataset. The SAGE (2010) dataset has a coarse spatial grid ($0.5° \times 0.5°$) and simple linear increment (decrement) algorithm, and as a result the simulated cropland quantity and spatial distribution are quite different from the actual situation. Although the CHCD dataset has a higher precision, results of spatial reconstruction were not published openly. The HYDE 3.1 dataset (http://themasites.pbl.nl/en/themasites) has a wide array of basic data sources, a higher grid resolution ($5' \times 5'$), and more abundant time span. Thus, the cropland spatial patterns of the years 1720,

Fig. 6. The spatial pattern of differences between the HYDE 3.1 dataset and reconstructed results in this paper. Note: non-colored grids in the figure are non-cropland in the reconstructed results. Cells are defined as 1 if the grid relative errors exceed 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

1820, 1930, and 1980 in the HYDE 3.1 dataset are selected for the grid absolute error and relative error analysis in this paper (Li et al., 2010; He et al., 2012). Absolute error is the difference between HYDE 3.1 dataset and our results. The ratio of absolute error to the reconstructed area of cultivated land in the grid is the relative error value. To facilitate the analysis, the fractional grid data of cropland is used as the data type and the grid resolution is rescaled to 10 × 10 km.

Fig. 6 reveals that, in terms of absolute error, the spatial pattern of negative error is generally stable and mainly reflects the difference in local area. The negative error is focused on the North China Plain, Middle–Lower Yangtze River Plain, Guanzhong Plain, Sichuan Basin, and other main reclamation areas of cropland. The positive error mainly occurs in the surroundings and periphery of the main reclamation area and is distributed more widely and stably in the southeast hilly areas. The trends in positive and negative errors indicate that the negative errors gradually gathered in the North China Plain and Middle–Lower Yangtze River Plain, the Sichuan Basin presents a pattern of positive error transitioning to negative error, and no significant changes occur in the southeast hilly areas. With respect to the actual distribution of cropland in 1980, the absolute errors of the HYDE 3.1 dataset are dominated by negative errors with large magnitudes to the north of the Yangtze River and by negative errors to the south of the Yangtze River (except for local positive errors with smaller magnitudes).

With the increases in land reclamation rate and expansion of reclamation scope over time, the grid cells with relative errors >90% extend from the main reclamation area of cropland to the periphery and migrate towards the southeast hilly areas, Guangdong, and the Guangxi coastal areas. The regions with relative errors <−90% are concentrated in the areas with data missing in the HYDE 3.1 dataset (unallocated cropland). These areas are mainly distributed in the Yunnan–Guizhou Plateau, southeastern areas of Zhejiang and Fujian, and the Taihang mountain areas.

The grids with relative difference rates between the HYDE and constructed dataset exceeding ±90% (>90% or <90%), ±70% (>70% or <70%), ±30% (<30% or >30%), and ±10% (between −10% and 10%) respectively account for 46–58%, 57–68%, 17–25%, and 5–8% (Table 2). The relative error and standard deviation increased every year. For example, the grid proportion with relative difference rates exceeding ±90% (>90% or <90%) jumped from 47% in the period of 1720–1930 to 58% in 1980, and the respective standard deviations increased from 9.40% to 15.62%. In terms of the actual cropland pattern in 1980, the data with relative difference rates exceeding ±90% (>90% or <90%), ±70% (>70% or <70%), ±30% (<30% or >30%), and ±10% (between −10% and 10%) respectively account for 58%, 68%, 17%, and 7.79% in the HYDE 3.1 dataset. This indicates that in the HYDE dataset cropland quantity in 1980 was significantly lower than the actual value and quite different from the actual cropland pattern. Thus, the dataset is unable to reflect the actual cropland pattern in the traditional cultivation regions.

Table 2
The percentage of grid cells with different relative errors between HYDE 3.1 and reconstructed results in this paper.

Relative error (%)	1720	1820	1930	1980
>90	12.68	10.67	9.93	2.67
70 to 90	1.40	1.43	1.39	0.68
50 to 70	1.90	1.63	1.91	0.93
30 to 50	2.93	2.87	3.08	1.37
10 to 30	3.80	3.68	4.02	2.14
−10 to 10	6.88	7.00	7.53	4.69
−30 to −10	8.36	11.36	13.31	10.53
−50 to −30	5.46	6.31	6.14	5.53
−70 to −50	9.36	9.14	7.17	6.92
−90 to −70	12.36	9.84	7.86	9.46
<−90	34.88	36.08	37.64	55.08
Standard deviation (%)	9.40	9.65	10.11	15.62

The reason for these differences is investigated below. The HYDE 3.1 dataset are based on DIScover and GLC2000 modern remote sensing data. Next, the modern population distribution pattern is used to allocate the historical population in the research unit. Then, the spatial allocation method is adopted to allocate the cropland quantity based on land reclamation suitability, the distance from the river, and other factors. On the one hand, DIScover and GLC2000 modern remote sensing data had respective global accuracies of only 66.9% and 68.6% (Li et al., 2010). However, comparing with the regional accuracy (above 95%) of the 1980 remote sensing data from China's traditional cultivated regions used in this study, it is revealed that a greater accuracy difference exists in the reconstructed cropland pattern of the initial year (Liu et al., 2001, 2003). The HYDE 3.1 dataset mainly aggregates the regional results at the country scale with no inspection or correction of provincial cropland data, leading to a significant gap between total cropland and historical fact. This inevitably results in larger absolute and relative errors. In addition, the HYDE 3.1 datasets use uniform rules to allocate cropland quantity in a top–down manner. This is significantly different from the bottom–up reconstruction concept used in this study, which considers cultivation convenience, scale, and partition–synchronization evolution. Thus, the HYDE 3.1 dataset is more inconsistent and unreasonable in terms of the spatial pattern of the cropland.

4. Conclusions and discussion

This study utilizes constrained CA to build a reconstruction model of the spatial pattern of historical cropland. Taking China's traditional cultivated regions as the research object, the relevant influencing factors of cropland spatial distribution were used as the model spatial variables. Parameter identification was carried out based on land use data for 1980. The provincial (regional) administrative areas are used to divide evolutionary partitions for inverse modeling. Then, the simulation results were compared with the international mainstream datasets, and the following conclusions are obtained.

This research deviates from the traditional “top–down” cropland spatial distribution method based on land suitability assessment and introduces certain methodological innovations. By introducing the constrained CA model with partition synchronization and considering the continuity of cropland, we built a set of “bottom–up” spatial pattern reconstruction models of historical cropland. Restricted by modern cropland patterns and historical cropland quantity, the spatial pattern of the historical cropland is then projected backward.

The concept of partition synchronization modeling is introduced. During the simulation of large-scale spatial patterns, multiple evolutionary partitions are divided by means of partition synchronization. Then the influence factor of the cropland spatial pattern is selected to screen the influence weight of each factor. After the training of parameters in each partition is performed, a set of transformation systems is formed to reflect the spatial heterogeneity of evolutionary rules and rates.

This study comprehensively considers the cropland convenience, continuity, scale, and other traditional cultivation characteristics. The regional population data for 1820 were used for the calibration and testing of the model transformation parameters and this can improve the reasonability of reconstructed results to some extent. The analysis of the absolute error and relative error shows that the HYDE 3.1 dataset underestimates total cropland in the traditional cultivated regions. Meanwhile, the simulated cropland pattern is quite different from the actual situation, indirectly verifying the reasonableness of the dataset reconstruction used in this study.

Integrating the patterns of modern cropland, land suitability, neighborhood, and random interference factors, the spatial distribution pattern of the historical cropland is reconstructed using the amount of cropland in partition. This method of simulating a theoretical abstraction of the spatial distribution of historical cropland proved to be efficient.

Restricted by research methods and the availability of historical data, there are several deficiencies present in this study, which require further follow-up.

In this study, the provincial administrative regions are taken as the main partition unit to reconstruct the spatial pattern of the historical cropland with a 1 × 1 km grid. As a result, the relevant reconstruction results are only applicable at the macroscopic scale, and results cannot be interpolated to smaller scales or provide insight into the overall trends in historical croplands.

In this study, the regional population data at the typical time period (1820) are used for the indirect testing of the available cropland index for labor force, which may later be used to enrich the test data and methods. For example, the CORONA, ARGON, and LANYARD images from the US-based USGS platform during the period of 1960–1970 are used for point-to-point validation. The cross-examination based on local history is one possible way for this to be accomplished.

Acknowledgments

We thank the National Basic Research Program of China (no. 2011CB952001) and the National Natural Science Foundation of China (nos. 41340016 and 41201386) for the financial support on this research.

References

Bai, S.Y., Zhang, S.W., Zhang, Y.Z., 2007. Digital rebuilding of LUCC spatial–temporal distribution of the last 100 years, taking Dorbod Mongolian Autonomous County in Daqing City as an example. Acta Geograph. Sin. 62 (4), 427–436.
Brovkin, V., et al., 2004. Role of land cover changes for atmospheric CO_2 increase and climate change during the last 150 years. Glob. Chang. Biol. 10 (8), 1253–1266.
Cao, X., Jin, X.B., Zhou, Y.K., 2013. Research on cropland data recovery and reconstruction in the Qing Dynasty, method and case study. Acta Geograph. Sin. 68 (02), 45–256.
Cao, X., Jin, X.B., Wang, J.S., et al., 2014. Reconstruction and change analysis of cropland data of China in recent 300 years. Acta Geograph. Sin. 69 (07), 896–906.
CHGIS, 2007. version4. Harvard Yenching Institute, Cambridge.
CPGIS, 2007. version4. Harvard Yenching Institute, Cambridge.
Feddema, J.J., Oleson, K.W., Bonan, G.B., et al., 2005. The importance of land-cover change in simulating future climates. Science 310 (5754), 1674–1678.
Feng, Z.M., Liu, B.Q., Yang, Y.Z., 2005. A study of the changing trend of Chinese cultivated land amount and data reconstructing, 1949–2003. J. Nat. Resour. 20 (1), 35–43.
Foley, J.A., DeFries, R., Asner, G.P., et al., 2005. Global consequence of land use. Science 309 (5734), 570–574.
Ge, Q.S., Dai, J.H., 2005. Analysis of Chinese agricultural and forest land use change and its driving force in early and mid stage of 20th century. Sci. China (Ser. D) 35 (1), 54–63.
Ge, Q.S., Dai, J.H., He, F.N., et al., 2003. Cultivated land amount change and driving forces analysis of some provinces of China in past 300 years. Adv. Nat. Sci. 13 (8), 825–832.
Ge, Q.S., Dai, J.H., He, F.N., et al., 2008. Land use and land cover change of the past 300 years in China and carbon cycle research. Sci. China Ser. D Earth Sci. 38 (2), 197–210.
Goldewijk, K.K., 2001. Estimating global land use change over the past 300 years, the HYDE database. Glob. Biogeochem. Cycles 15 (2), 417–433.
Goldewijk, K.K., Navin, R., 2004. Land cover change over the last three centuries due to human activities, the availability of new global data sets. GeoJournal 61 (4), 335–344.
He, F.N., Tian, Y.Y., Ge, Q.S., 2003. Spatial–temporal characteristics of land reclamation in Guanzhong region in the Qing Dynasty. Geogr. Res. 22 (6), 687–697.
He, F.N., Li, S.C., Zhang, X.Z., 2011. Reconstruction of cropland area and spatial distribution in the mid-Northern Song Dynasty (AD1004–1085). J. Geogr. Sci. 22 (2), 359–370.
He, F.N., Li, S.C., Zhang, X.Z., et al., 2012. Comparisons of reconstructed cropland area from multiple datasets for the traditional cultivated region of China in the last 300 years. Acta Geograph. Sin. 67 (009), 1190–1200.
Jiao, L.M., Liu, Y.L., 2004. Application of fuzzy neural networks to land suitability evaluation. Geomatics Inf. Sci. Wuhan Univ. 29 (6), 513–516.
Lambin, E.F., et al., 2001. The causes of land-use and land-cover change: moving beyond the myths. Glob. Environ. Chang. Hum. Policy Dimens. 11 (4), 261–269.
Li, X., Yeh, A.G.O., 1999. Constrained cellular automata for modeling sustainable urban forms. Acta Geograph. Sin. 54 (04), 289–298.
Li, X., Yeh, A.G.O., 2005. Cellular automata for simulating complex land use systems using neural networks. Geogr. Res. 24 (1), 19–27.
Li, B.B., Fang, X.Q., Ye, Y., Zhang, X.Z., 2010. Regional accuracy evaluation of global land use datasets on northeastern parts of China. Sci. China Ser. D Earth Sci. 40 (08), 1048–1059.
Li, K., He, F.N., Zhang, X.Z., 2011. An approach to reconstructing spatial distribution of historical cropland with grid-boxes by utilizing MODIS land cover dataset, a case study of Yunnan Province in the Qing Dynasty. Geogr. Res. 30 (12), 2281–2288.
Li, S.C., He, F.N., Chen, Y.S., 2012. Gridding reconstruction of cropland spatial patterns in southwest china in the Qing Dynasty. Prog. Geogr. 31 (9), 1196–1203.
Lin, S.S., Zheng, J.Y., He, F.N., 2008. The approach for gridding data derived from historical cropland records of the traditional cultivated region in China. Acta Geograph. Sin. 63 (01), 83–92.
Lin, S.S., Zheng, J.Y., He, F.N., 2009. Gridding cropland data reconstruction over the agricultural region of China in 1820. J. Geogr. Sci. 01 (19), 36–48.
Liu, M.L., Tian, H.Q., 2010. China's land cover and land use change from 1700 to 2005, estimations from high-resolution satellite data and historical archives. Glob. Biogeochem. Cycles 24 (3), 1–18.
Liu, Y.L., Liu, Y.F., Xia, Z.F., 1995. Land suitability evaluation based on fuzzy comprehensive judgment. J. Wuhan Tech. Univ. Surv. Mapp. 20 (1), 71–75.
Liu, M.L., Tang, X.M., Liu, J.Y., et al., 2001. Research on scaling effect based on 1 km grid cell data. J. Remote. Sens. 5 (3), 183–190.
Liu, J.Y., Zhang, Z.X., Zhuang, D.F., et al., 2003. A study on the spatial–temporal dynamic changes of land-use and driving forces analyses of China in the 1990s. Geogr. Res. 22 (1), 1–12.
Liu, X.P., Li, X., Yeh, A.G., 2006. Multi-agent systems for simulating spatial decision behaviors and land use dynamics. Sci. China (Ser. D) 36 (11), 1027–1036.
Liu, M.H., Dai, Z.Y., Qiu, D.C., et al., 2011. Influencing factors analysis and rational distribution on rural settlements in mountains region. Econ. Geogr. 31 (3), 476–482.
Long, Y., Han, H.Y., Mao, Q.Z., 2009. Establishing urban growth boundaries using constrained CA. Acta Geograph. Sin. 64 (8), 999–1008.
Long, Y., Shen, Z.J., Mao, Q.Z., et al., 2010. Form scenario analysis using constrained cellular automata. Acta Geograph. Sin. 65 (006), 643–655.
Long, Y., Jin, X.B., Yang, X.H., et al., 2014. Reconstruction of historical arable land use patterns using constrained cellular automata: a case study of Jiangsu, China. Appl. Geogr. 52, 67–77.
Miao, L.J., Zhu, F., He, B., et al., 2012. Synthesis of China's land use in the past 300 years. Glob. Planet. Chang. 100 (2013), 224–233.
Pongratz, J., Reick, C., Raddatz, T., et al., 2008. A reconstruction of global agricultural areas and land cover for the last millennium. Glob. Biogeochem. Cycles 22 (3), 1–16.
Ramankutty, N., Foley, J.A., 1999. Estimating historical changes in global land cover, croplands from 1700 to 1992. Biogeochem. Cycles 13 (4), 997–1027.
Ray, D.K., Pijanowski, B.C., 2010. A backcast land use change model to generate past land use maps, application and validation at the Muskegon River watershed of Michigan, USA. J. Land Use Sci. 5 (1), 1–29.
Shi, Z.G., Yan, X.D., Yin, C.H., et al., 2007. Effects of historical land cover changes on climate. China Sci. Bull. 52 (12), 1436–1444.
Statistics Bureau of the National Government, 1936. Statistical Analysis of Chinese Land Issues. Zhongzheng Publishing House, Nanjing.
Tian, H.Q., Banger, K., Bo, T., et al., 2014. History of land use in India during 1880–2010: large-scale land transformations reconstructed from satellite data and historical archives. Glob. Planet. Chang. 121, 78–88.
Voldoire, A., Eickhout, B., Schaeffer, M., et al., 2007. Climate simulation of the twenty-first century with interactive land-use changes. Clim. Dyn. 29 (2–3), 2–3.
Wu, F., 1998. SimLand, a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. Int. J. Geogr. Inf. Sci. 12 (1), 63–82.
Xie, Y.W., Wang, X.Q., Wang, G.S., et al., 2013. Cultivated land distribution simulation based on grid in middle reaches of Heihe river basin in the historical periods. Adv. Earth Sci. 28 (1), 71–78.
Ye, D.Z., Fu, C.B., 2004. Some advance in global change science study. Bull. Chin. Acad. Sci. 19 (05), 336–341.
Ye, Y., Fang, X.Q., Ren, Y.Y., et al., 2009. Northeast cultivated land cover changes over the past 300 years. Sci. China Ser. D Earth Sci. 39 (3), 340–350.
Yeh, A.G.O., Li, X., 2001. A constrained CA model for the simulation and planning of sustainable urban forms by using GIS. Environ. Plan. B Plan. Des. 28 (5), 733–753.
Zhang, G.P., Liu, J.Y., Zhang, Z.X., 2003. Spatial–temporal changes of cropland in china for the past 10 years based on remote sensing. Acta Geograph. Sin. 58 (3), 323–332.
Zhang, H.H., Zeng, Y.N., Jin, X.B., et al., 2008. Urban land expansion model based on multiagent system and application. Acta Geograph. Sin. 63 (8), 869–881.
Zhou, R., 2001. A general inspection and re-appraisal on area under cultivation in the early period of Qing Dynasty. Jianghan Tribune (09), 57–61.
Zhu, F., Cui, X.F., Miao, L.J., 2012. China's spatially-explicit historical land-use data and its reconstruction methodology. Prog. Geogr. 31 (12), 563–1573.