© 2017   Science Press   Springer-Verlag

Reconstruction of cropland spatial patterns and its spatiotemporal changes over the 20th century on the Songnen Plain, Northeast China

ZHANG Lijuan[1], JIANG Lanqi, ZHANG Xuezhen[2,3]

1. Key Laboratory of Remote Sensing Monitoring of Geographic Environment, Harbin Normal University, Harbin 150025, China;
2. Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;
3. Jiangsu Collaborative Innovation Center for Climate Change, Nanjing 210093, China

Abstract: We initially estimated the cropland area at county level using local historical documents for the Songnen Plain (SNP) in the 1910s and 1930s. We then allocated this cropland area to grid cells with a size of 1 km × 1 km, using a range of cultivation possibilities from high to low; this was based on topography and minimum distances to rivers, settlements, and traffic lines. Cropland areas for the 1950s were obtained from the *Land Use Map of Northeast China*, and map vectorization was performed with ArcGIS technology. Cropland areas for the 1970s, 1980s, 1990s, 2000s, and 2010s were retrieved from Landsat images. We found that the cropland areas were 4.92×10^4 km^2 and 7.60×10^4 km^2, accounting for 22.8% and 35.2% of the total area of the SNP in the 1910s and 1930s, respectively, which increased to 13.14×10^4 km^2, accounting for 60.9% in the 2010s. The cropland increased at a rate of 1.18×10^4 km^2 per decade from the 1910s to 1970s while it was merely 0.285×10^4 km^2 per decade from the 1970s to 2010s. From the 1910s to 1930s, new cultivation mainly occurred in the central SNP while, from the 1930s to 1970s, it was mainly over the western and northern parts. This spatially explicit reconstruction could be offered as primary data for studying the effects of changes in human-induced land cover based on climate change over the last century.

Keywords: land use/cover changes; agricultural development; historical documents; gridding data

1   Introduction

In pursuit of human provision throughout history, natural vegetation such as woodland, grassland, and swamps, has been converted into anthropogenic cropland. Due to agricultural development, land-cover changes (LCCs) have modified the surface energy balance, the

Bowen ratio, and wind speed, and have thus had important climatic implications (Notaro *et al.*, 2006; Pielke *et al.*, 2007; Anderson-Teixeira *et al.*, 2012). LCCs have therefore been broadly accepted as one of the driving forces of climate change (Pielke and Niyogi, 2009; Mahmood *et al.*, 2010; Dirmeyer *et al.*, 2010). Because of the large spatial heterogeneity of the predecessors of cropland, the climatic effects of agriculture-induced LCCs vary between regions; they may be positive in some regions and negative in others. To comprehensively understand regional climate change, it is necessary to reveal changes in anthropogenic cropland that indicate human modifications to the land surface.

Relevant studies have reconstructed the historical spatial patterns of cropland for agricultural regions in China (Lin *et al.*, 2009; He *et al.*, 2011; Feng *et al.*, 2014; Jin *et al.*, 2015; Ye *et al.*, 2009; Cao *et al.*, 2013; 2014; Li *et al.*, 2012; 2015; Luo *et al.*, 2014; Zhang *et al.*, 2015; Liu *et al.*, 2010; Wang *et al.*, 2015; Li *et al.*, 2010; Pan *et al.*, 2015). One method is based on using the reclamation rate index to reconstruct the spatial distribution of farmland. For example, Ye *et al.* (2009) used the reclamation rate to reconstruct the spatial distribution of farmland in Northeast China for the previous 300 years with 50–100 year intervals as a county unit. Jin *et al.* (2015) reconstructed China's provincial farmland dataset for the last 300 years (1661–1985) by applying factor correction, citing replacement, linear interpolation, cohesion adjustment, etc. Earlier, Lin *et al.* (2009) designed an empirical model to reconstruct the spatial distribution pattern of cropland in the traditionally cultivated regions of China in 1820 (at 60 km × 60 km). He *et al.* (2011) similarly reconstructed the spatial distribution pattern of cropland during the Northern Song Dynasty using the method of Lin *et al.* (2009). In the same way, Feng *et al.* (2014) also developed a separate method to reconstruct the cropland distribution of China in 1913, 1933, 1950, 1970, 1990, and 2000 at a spatial resolution of 10 km × 10 km. Furthermore, to the best of our knowledge, there are two global datasets. One was created by the Center for Sustainability and the Global Environment (SAGE), University of Wisconsin (hereafter, called the SAGE dataset; Ramankutty *et al.*, 1999; 2010). The other is the History Database of the Global Environment (HYDE) created by the Netherlands Environmental Assessment Agency (hereafter, called the HYDE dataset; Goldewijk *et al.*, 2001; 2011). The SAGE dataset was created through the following approach. National-level cropland areas from inventories in 1992 were used to calibrate the satellite-based discovered dataset to obtain the spatially explicit and accurately quantitative cropland area; then, keeping the relative spatial weights unchanged, the historical national-level cropland area was allocated into grid cells (see Ramankutty *et al.*, 1999, for details). The latest HYDE dataset (v3.1) used the contemporary population and per capita cropland area for the period pre-1960 to estimate national cropland area, and it then allocated the estimated national cropland area into grid cells of 5 minutes (latitude/longitude). The national cropland area was allocated to grid cells using a mix of two weighting maps: a current map, which was constructed from a satellite map of AD 2000 (Klein Goldewijk and van Drecht, 2006), and a historical map, which was constructed based on six rules by considering the effects of population density, temperature, land suitability for crops, distance to water, surface slope, and city development on agricultural activities (for details, see Klein Goldewijk *et al.*, 2011).

After the 1970s, with the development of remote sensing interpretation technology, it was possible to interpret cropland spatial patterns based on remote sensing images. However,

using the arable land reclamation index to reconstruct the spatial distribution of cultivated land has its limitations for precisely reproducing the spatial change in arable land through history. Therefore, too much attention has been paid to the pre-1970 spatial patterns of cropland reconstruction. Pan *et al*. (2015) and Li *et al*. (2010) used integrated suitability indices analysis based on modern land-use patterns to rebuild cropland spatial position patterns in Jiangsu Province (100 m × 100 m) and Yunnan Province (100 m × 100 m) in 1671 and 1827. However, until now, few studies have focused on an extended historical period sequence pattern that is based on the spatial distribution of cultivated land.

Songnen Plain (SNP) forms the main part of the Northeast China Plain, which is a typical agricultural area of China. Since the 19th century, the SNP has been extensively developed amid high levels of immigration (Li and Shi, 1987), and with the impact of large-scale land reclamation in the 1950s, extensive areas of Northeast China were converted to arable use. The area of cropland almost increased exponentially by the year. Northeast China is a typical area for studies on the influence of human activities on the change in land cover (Ye *et al*., 2009). There are three datasets (HYED, SAGE, and Ye *et al*., 2009) that describe the spatial distribution of cropland in the SNP. However, the SAGE dataset is largely different from local historical records (Li *et al*., 2010), and the HYDE dataset for Northeast China still contains great uncertainties because of lack of reliability in the estimated national total cropland area through the use of a constant per capita cropland area (Li *et al*., 2010; He *et al*., 2013). Ye *et al*. (2009) used the proportion of cultivated land area as an index to reconstruct the spatial distribution of cropland; however, there is a limitation in the spatial pattern of cropland change in the SNP over an extended period of time. This paper reconstructs the spatial distribution of cultivated land for the SNP in the 1910s, 1930s, 1950s, 1970s, 1990s, and 2010s. Furthermore, it analyzes cropland spatial pattern changes over an extended time scale using methods involving statistical and spatial analysis. The results not only provide a quantitative basis of the influence of human activities on land-use changes in Northeast China, but also provide models and references for research on long-term sequence cropland spatial patterns.

2   Study area

The study area, lying between 42°30′–51°20′N and 121°40′–128°30′E and located in the Northeast China Plain, covers an area of about 23.75 × 10^4 km^2 (approximately 2.47% of China's total; Figure 1). The SNP is an alluvial plain, which is drained by the Songhua River and Nenjiang River flowing down from the Tianchi Lake of the Changbai Mountains and the Greater Higgnan Mountains. The study area is mostly flat with a few hills and an elevation of about 120–300 m (Zeng *et al*., 2010). Soils are fertile with black soil, meadow soil, and chernozem widely distributed in this region; wetlands and lakes are commonly found in the central plain. The plain supports 559 million ha of arable land, including soybeans, maize, wheat, beets, and potatoes, and it is an important grain commodity base for China.

The SNP has a temperate continental monsoon climate, characterized by significant winds and four seasons with a hot, rainy summer and a cold, dry winter. The average monthly temperature is about –16 to –26℃ in January and precipitation can reach 10 to 24 mm in winter.
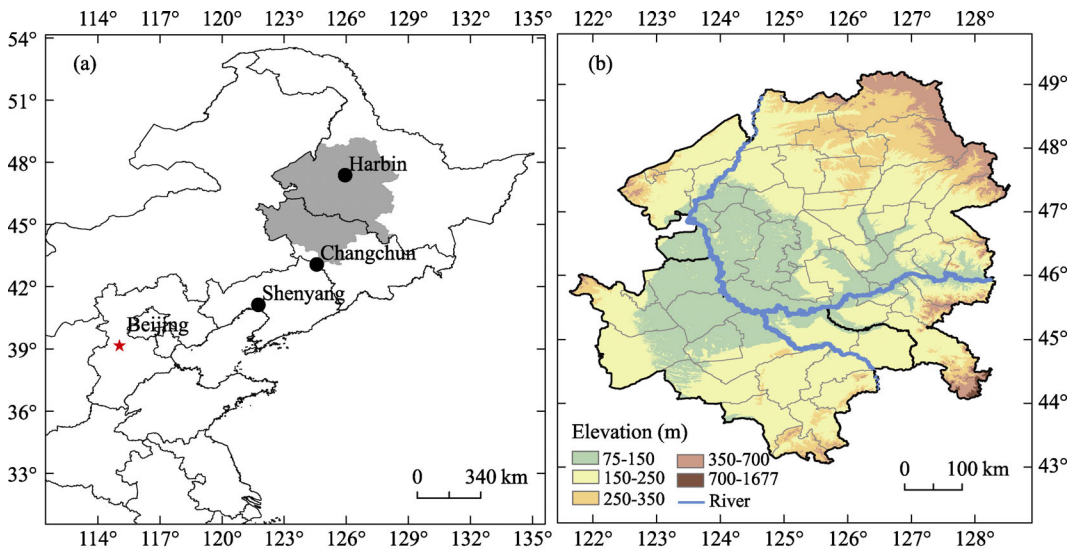
Figure 1   Location of the Songnen Plain (a) and the distribution of counties (b) in Northeast China

3   Data sources

3.1   Cropland data

This study used cropland data from multiple sources. For the 1910s, the cropland area data were compiled from a *Survey Report of Manchuria* (LPA, 2008) and the *General Conditions of Manchuria-Mongolia* (CSMR, 1923). The *Survey Report of Manchuria* was based on a field survey that was conducted by the Survey Department of the South Manchuria Railway Company between 1909 and 1925. The fifth volume of this report records the conditions of society including the sub-provincial level of the cropland area of Heilongjiang Province and Jilin Province for the 34th year of the Emperor Guangxu (i.e. 1908). The *General Conditions of Manchuria-Mongolia*, which was published in 1918, records natural resources including sub-provincial levels of cropland area for Heilongjiang and Jilin for the early 20th century. The two sources overlapped and could therefore verify each other.

For the 1930s, the cropland area data were mainly taken from the *Survey Report of the Northeast Region in the Republic of China* (Xiong, 2009). This field survey was conducted by the Republic of China in 1929 and covered 41 counties in Jilin and 41 counties in Heilongjiang. This report records county-level cropland areas, with some counties having no data. To complement the missing data, we used cropland area records from the *History of Agricultural Development in Heilongjiang Province* (Xin *et al.*, 1999) and the *Gazette of Jilin Province* (CCCJP, 1992). One section of the *History of Agricultural Development in Heilongjiang Province* records natural resources including county-level cropland area for the 1930s. The *Gazette of Jilin Province* records county-level cropland area in Jilin for the 1930s.

To exhibit the temporal-spatial changes in cropland, the current study also used cropland data for the 1950s and thereafter. For the 1950s, since there were no remote sensing images, the cropland area data were taken from the *Land Use Map of Northeast China* (1:300,000; Sun *et al.*, 1959). We registered the scanned map to a relative coordinate, and then vector-

ized the map using ArcGIS technology. For the 1970s, we retrieved the cropland area from Landsat MSS images; for the 1980s, 1990s, 2000s, and 2010s, cropland areas were retrieved from Landsat ETM$^+$ images. The Landsat MSS/ETM$^+$ images underwent manual visual interpretation. All satellite images are available on the USGS Global Visualization Viewer.

3.2   Population and settlement data

The current study also used population data of the early 20th century, which were taken from the *Local Gazette of Heilongjiang Province* (Wan *et al*., 1992) and the *Local Gazette of Jilin Province* (CCCJP, 1992). *The Local Gazette of Heilongjiang Province* contains population figures for the 39 county-level units in Heilongjiang from the 33rd year of the Emperor Guangxu (i.e. 1907) to the 7th year of the Republic of China (i.e. 1918). The *Local Gazette of Jilin Province* contains data of the population covering county-level units of Jilin from the 10th year of the Emperor Shunzhi (i.e. 1653) to 1985 including the time slice of 1910.

We also used historical settlement data. For the 1910s (Figure 2a), the spatial distribution of settlements was taken from the eighth volume of the *Atlas of Historical Geography* (Tan, 1987). For the 1930s (Figure 2b), data were taken from the *New Atlas of the Republic of China*, which was created by Ding *et al*. (1934). This atlas contains thematic maps of political units and traffic lines with a scale of 1:200,000 in eastern China and 1:500,000 in western China. The spatial distribution of settlements is included in the thematic map of political units.
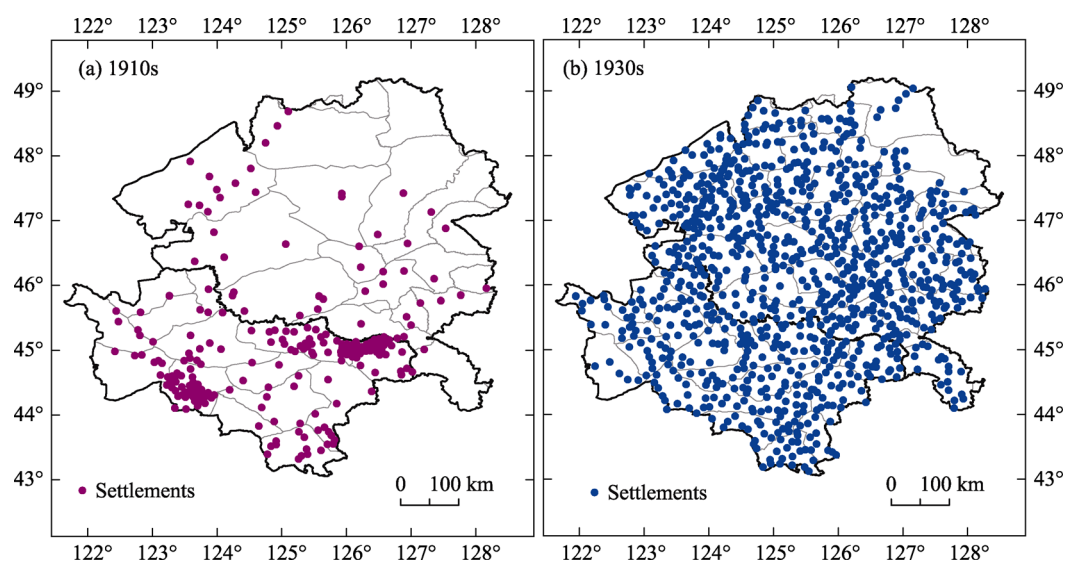

Figure 2   Spatial distribution of settlements in the Songnen Plain, Northeast China in the 1910s (a) and 1930s (b)

3.3   Assistant geographical data

Except for the above-mentioned cropland and population data, the current study also used some geographical data, including maps of forests in Heilongjiang, the spatial distribution of wetlands and lakes for the 1950s, surface elevation data, maps of railways for the 1930s, and the present river systems. The forest map for Heilongjiang exhibits forest types and their spatial distribution for the late 19th century and has a scale of 1:300,000 (Li, 1993). The

spatial distribution of wetlands and lakes for the 1950s was taken from the *Land Use Map of Northeast China* (1:300,000) created by the Department of Economy in the Institute of Geography, Chinese Academy of Sciences (1959). Surface elevation data were taken from the Bureau of Survey and Geoinformation of Heilongjiang Province, which has a spatial resolution of 90 m. The distribution of railways was retrieved from the *Political Map during the Period When Northeast China was Japanese-occupied* (CCCHP, 1999 see Figure 3a). The distribution of rivers was taken from the National Geomatics Center of China (available at http://nfgis. nsdi.gov.cn/nfgis/chinese/c_xz.htm; Figure 3b).
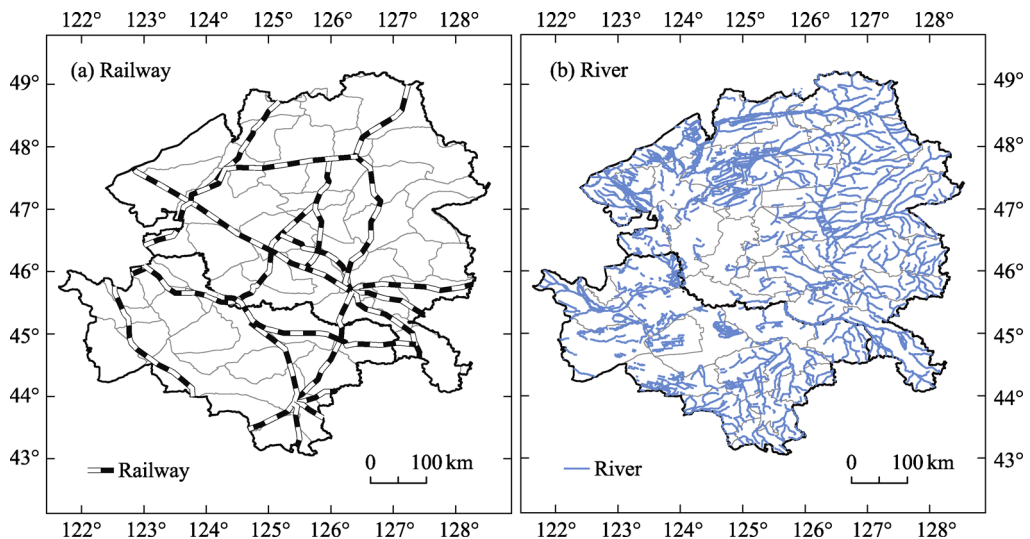

Figure 3 Spatial distribution of railways (a) and rivers (b) in the Songnen Plain, Northeast China in the 1930s

4 Methods

4.1 Method for spatial allocation of cropland for the 1910s and 1930s

To create a spatially explicit cropland dataset for the early 20th century, we estimated the county-level cropland area and subsequently allocated it into spatially explicit grid cells at a size of 1 km × 1 km.

4.1.1 Estimation of county-level unit cropland area

Historical documents generally record cropland area for each county. As the political area changed in the 20th century, it was necessary to locate historical documents based on the distribution of historical and contemporary counties. Figure 4 shows the distribution of counties for the 1910s, 1930s and 1990s, respectively. In the 1910s, there were 34 counties, 20 and 14 of which were governed by Heilongjiang and Jilin, respectively. In the 1930s, there were 49 counties, 32 and 17 of which were governed by Heilongjiang and Jilin, respectively. In the 1990s, there were 51 counties, 35 and 16 of which were governed by Heilongjiang and Jilin, respectively.

The units for measuring cultivated land area differed between historical documents. For the current study, we converted all measurement units into square kilometers (km^2) (Zhang *et al*., 2015).
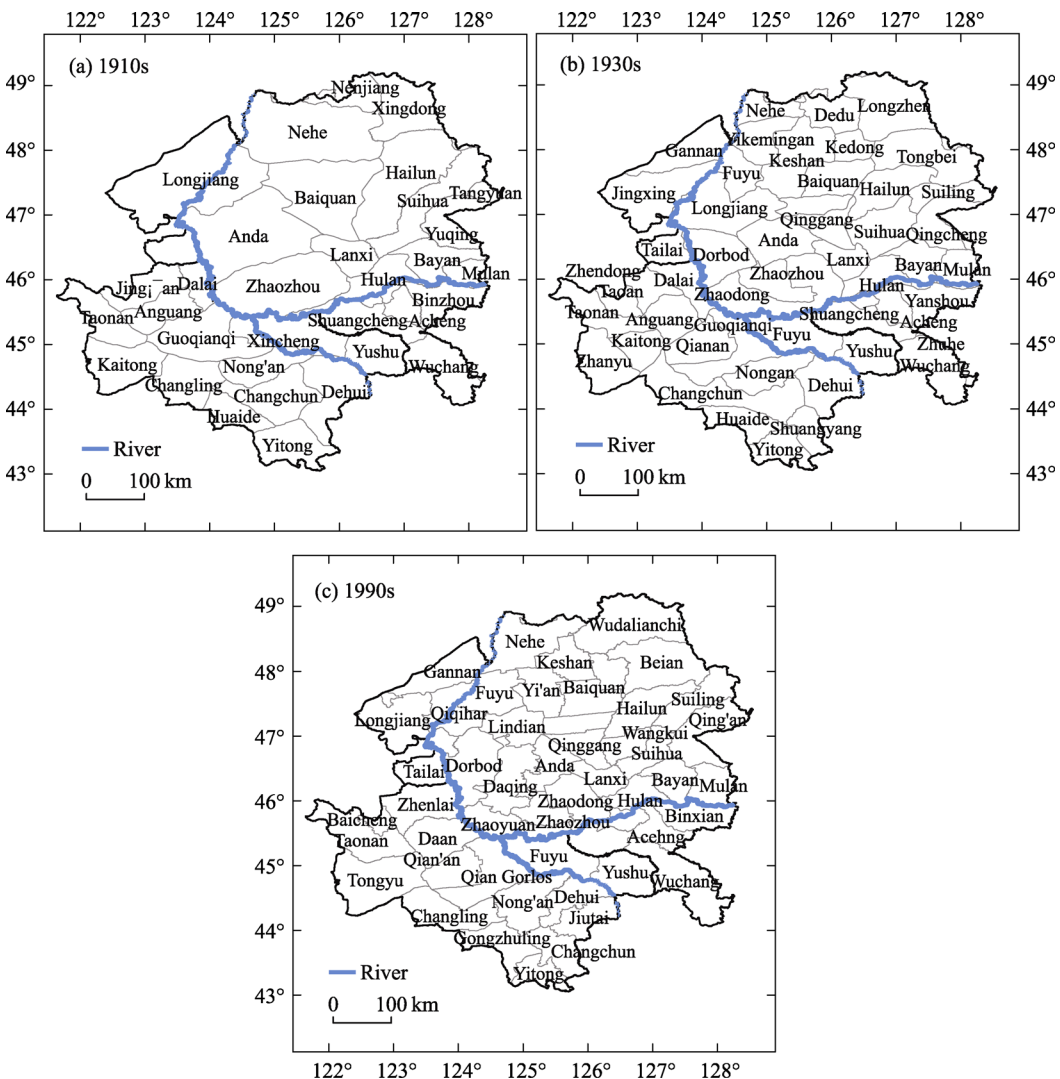
Figure 4   County distribution of the Songnen Plain, Northeast China for the 1910s (a), 1930s (b), and 1990s (c)

Farmland data from 28 counties in the 1910s were directly obtained from *The Survey Report of Mancuria Railway*, which are considered highly credible. The farmland data from four counties were obtained from *The General Conditions of Manchuria-Mongolia* after correction. Calibration used county as a sample to establish regression equations for both sets of farmland data (*The Survey Report of Manchuria Railway* and *The General Conditions of Manchuria-Mongolia*). Calibration equations were established for both Heilongjiang (13 samples) and Jilin (10 samples). There was a high correlation between these two sets of data with explained variances of 99% and 85%, respectively (Table 1).

For those counties without records, we estimated the area of cultivated land according to the population data by establishing regression equations between population and cropland. There was a high correlation between population and arable land, the explained variances of the population data and cropland area in Heilongjiang and Jilin were 84% and 66%, respectively.

Table 1　Calibration functions of cropland area for the 1910s in the Songnen Plain, Northeast China

	Cropland area from GMM (X)		Population (P)	
	HLJ	JL	HLJ	JL
Calibrated cropland area (Y)	$Y = X+0.6387$	$Y = 0.45*X+1044.7$	$Y = 0.0144*P-78.144$	$Y = 0.0034*P+1060.1$
R^2	0.99	0.85	0.84	0.66
Number of samples	13	10	15	10

Note: GMM represents *The General Conditions of Manchuria-Mongolia* HLJ and JL represent Heilongjiang Province and Jilin Province, respectively.

For the 1930s, the cropland areas of 47 counties were taken directly from the *Survey Report of Northeast Region in the Republic of China*. The cropland areas of eight counties in Heilongjiang were taken from the *History of Agricultural Development in Heilongjiang Province* but with calibration (Table 2). The cropland areas of two counties in Heilongjiang were taken from the *Brief History of Heilongjiang Province* but with calibration (Table 2). The cropland area of one county in Jilin, which was not recorded by the *Survey Report of Northeast Region in Republic of China*, was taken from the *Gazette of Jilin Province* but with calibration (Table 2). These calibrations were performed through regressing the county level of cropland areas from the *Survey Report of Northeast Region in the Republic of China* to those from the other historical documents mentioned above. As shown in Table 3, there were strong correlations with the explained variance from 91% to 99%.

Table 2　Calibration functions of cropland area for the 1930s in the Songnen Plain, Northeast China

	Cropland area from HAD-HLJ (X)	Cropland area from BH-HLJ (X)	Cropland area from GZ-JL (X)
Calibrated cropland area (Y)	$Y = 0.83*X+3.80$	$Y =0.846*X-57.69$	$Y =0.987*X+118.81$
R^2	0.99	0.92	0.91
Number of samples	49	49	12

Note: HAD-HLJ represents the "*History of Agricultural Development in Heilongjiang Province*"; BH-HLJ represents the "*Brief History of Heilongjiang Province*"; GZ-JL represents the "*Gazette of Jilin Province*".

4.1.2　Allocation of county-level cropland area into grid cells

To achieve the spatial distribution of county-level cropland, we estimated the cultivation possibility of each 1 km × 1 km pixel within the potential cultivation area. The potential cultivation area refers to the study area excluding forests, wetlands, rivers, lakes, and mountains. In the current study, we assumed that mountain areas with elevations higher than 200 m and a slope greater than 3° were not suitable for crop production (Lin *et al.*, 2009). Following the order of cultivation possibility from high to low, we assigned the cropland area in pixels of 1 km × 1 km. The assignment of cropland stopped when the county cropland area was completed.

The cultivation possibility is quantified as:

$$R = \frac{1}{a \times \alpha + b \times \beta + c \times \gamma} \tag{1}$$

where R is the cultivation possibility; α, β, and γ represent human dimension factors, water resource factors, and topography complexity factors, respectively. α is quantified by Eq. (2); β is indexed as the normalized minimum distance from rivers; and γ is indexed as the

normalized standard deviation of sub-pixel elevation. *a*, *b*, and *c* are coefficients with a sum of 1.

$$\alpha = d \times X + e \times Y \qquad (2)$$

where *X* and *Y* denote settlement (population) factors and traffic line impact factors (only available for the 1930s), respectively. *X* is the normalized distance from the nearest settlement; *Y* is the normalized minimum distance from the railway; *d* and *e* are coefficients with a sum of 1; particularly, for the 1910s, *d* is 1 and *e* is 0, because there were no railways.

Similarly, we calculated that *d* and *e* were 0.667 and 0.333, respectively. So, Eq. (2) could be represented as:

$$\alpha = 0.667 \times X + 0.333 \times Y \qquad (3)$$

To determine *a*, *b*, *c*, *d*, and *e*, we referred to the previous documents (Zhang *et al*., 2015).

4.2 Method for retrieval of cropland in the mid–late 20th century

We digitized the *Land Use Map of Northeast China* (1:300,000; Sun *et al*., 1959) and converted the vector data to raster data with a resolution of 1 km × 1 km. We used unsupervised classification on the Landsat images for the 1970s, 1980s, 1990s, 2000s, and 2010s to obtain cropland areas at a resolution of 1 km × 1 km. Such retrieved cropland areas were verified through field survey. Among the 400 field survey points, the retrieved 369 points were consistent with the field survey.

5   Results

5.1   County-level cropland area

Figure 5 shows the estimated county-level cropland area for the SNP. There were 4.92×10^4 and 7.60×10^4 km^2 of cropland accounting for 22.8% and 35.2% of the total area of the SNP in the 1910s and 1930s, respectively. Our estimations were close to those of Ye *et al*. (2009; 2011), who estimated 25.6% and 36.5% of the total area of the SNP in the 1910s and 1930s, respectively. The estimated county-level cropland area illustrates that agricultural development in the early 20th century mainly occurred in the eastern part of the SNP. From the 1910s to 1930s, agricultural development extended generally westward.

Table 3 shows changes in total cropland area in the SNP from the 1910s to 2010s; there was an obvious increase (Figure 6). The total cropland area was about 4.92×10^4 km^2, or 22.8% of the total area of the SNP in the 1910s; while it reached 12.65×10^4 km^2 in the 2000s and 13.14×10^4 km^2 in the 2010s, or 58.6% and 60.9% of the total area of the SNP, respectively. The total cropland area increased by about 1.8-fold with a rate of 0.82×10^4 km^2 per decade from the 1910s to 2010s. However, the increase did not last throughout the 20th century. The increasing rate in the first half of the 20th century was much higher than that of the second half. The cropland area increased from about 4.92×10^4 km^2 in the 1910s to 12.0×10^4 km^2 in the 1970s at a rate of 1.18×10^4 km^2 per decade; while the rate of increase from the 1970s to 2010s was merely 0.285×10^4 km^2 per decade.
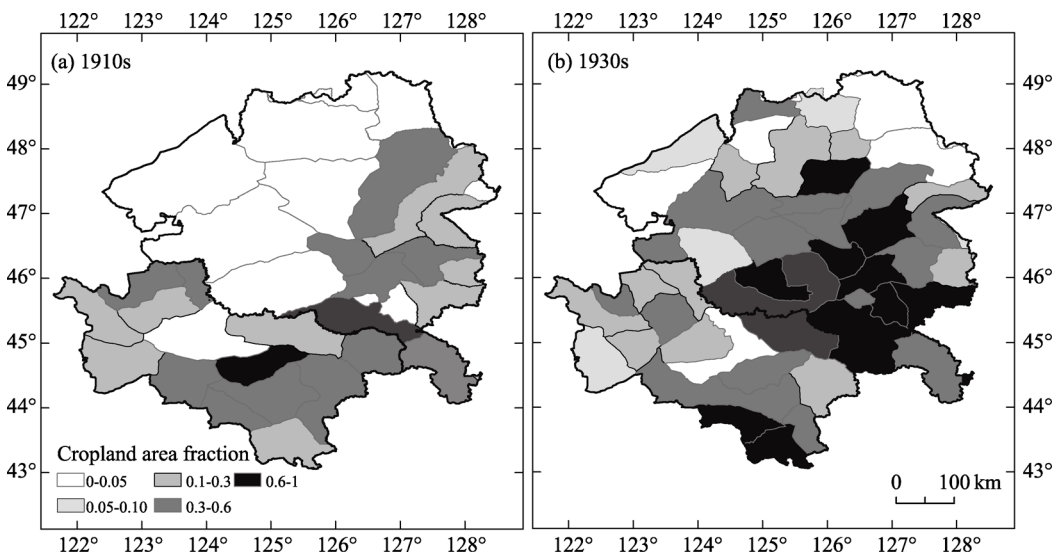
Figure 5 County-level cropland area fraction under county administration in the Songnen Plain, Northeast China in the 1910s (a) and 1930s (b)

Table 3 Total cropland area and cropland area fraction over the Songnen Plain, Northeast China from the 1910s to 2010s

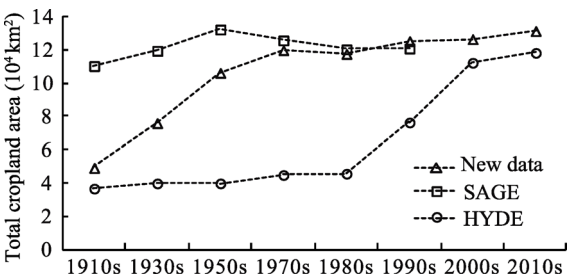|	1910s	1930s	1950s	1970s	1980s	1990s	2000s	2010s
Cropland area (km^2)	49176.0	76013.2	106199.6	119883.6	117626.4	125286.7	126520.5	131380.9
Cropland area fraction (%)	22.78	35.21	49.19	55.52	54.48	58.03	58.60	60.85

Figure 6 Total cropland area in the Songnen Plain, Northeast China from the 1910s to 2010s

5.2 Spatially explicit distribution of cropland

Figure 7 illustrates the reconstructed spatially explicit cropland in the 1910s and 1930s. We found that agricultural development in the 1910s mostly occurred in the eastern SNP while the other areas were less developed. The cropland area fractions ranged from 31.34% in Binxian county to 65.89% in Wangkui county located in the eastern part of the region (including Wangkui, Bayan, Yushu, Hulan, Shuangcheng, and Binxian counties), while the maximum cropland fraction was only 28.64% in Zhenlai county in the west. Such a distribution pattern is consistent with that of historical immigration. It was reported that in the late

Qing Dynasty, immigrants moved northward to the SNP and first settled and developed along the Songhua River because of the rich soil and plentiful water resources (Xu, 1985).
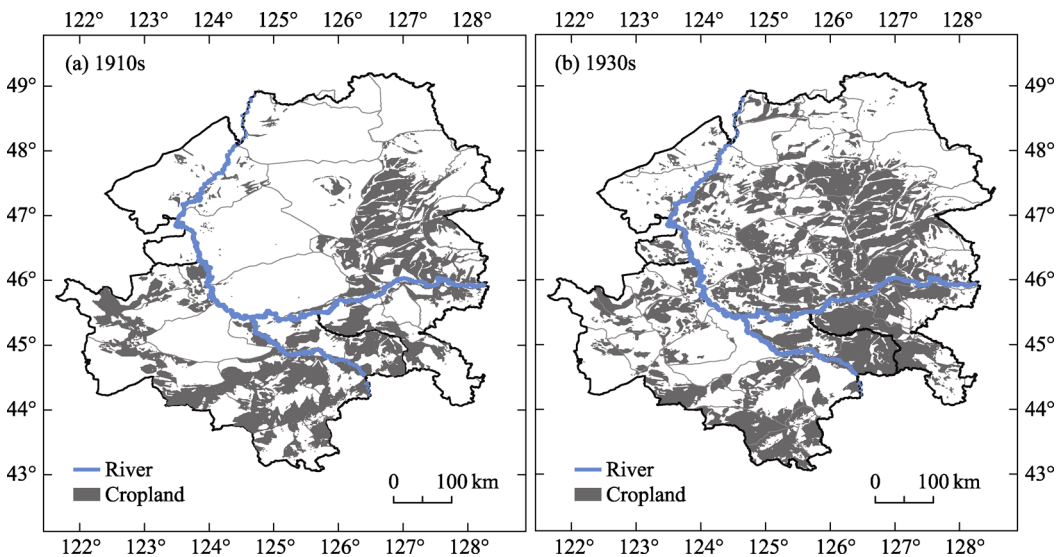

Figure 7   Reconstructed cropland distribution with a pixel size of 1 × 1 km in the 1910s (a) and 1930s (b) in the Songnen Plain, Northeast China

Up to the 1930s, the cropland area increased extensively. In the eastern SNP, cultivation was enhanced. It is estimated that the maximum cropland fraction reached 80.45% in Yushu county. Meanwhile, agricultural development occurred largely within the reach of the Nenjiang River and the central part of the SNP. There was a maximum cropland fraction of 63.9% in Zhaozhou county. Such an increase in cropland fraction was consistent with immigration across the Nenjiang River from the east to the west (Xu, 1985). However, up to the 1930s, little immigration occurred in the northern and western parts of the SNP (Xu, 1985); therefore, there was still little cropland in these areas.

Figure 8 shows pixel-based cropland expansion in the SNP over the last century. From the 1910s to 1930s, newly developed cropland mainly occurred in the central SNP. The largest increase in cropland fraction, from 0.06% to 63.90%, occurred in Zhaozhou county in the central SNP. The second largest increase from 0.02% to 54.00% was in Daqing county, also located in the central SNP. From the 1930s to 1970s, new cropland was mainly established in the northern and southeastern SNP. In the north, the largest increase in cropland fraction, from 0.28% to 10.18%, occurred in Keshan county. In the south, the largest increase in cropland fraction from 1.21% to 7.30% occurred in Yitong county. In total, over the last century, the most extensive increase in cropland occurred in the northwestern SNP. The cropland fraction increased from 1.74% to 61.46%, from 14.46% to 72.71%, and from 41.23% to 77.96%, in Lindian, Yi'an, and Qiqihar counties, respectively. From the 1910s to 1970s, the increasing rate of cropland fraction at county level ranged from 0.44% to 14.11% per decade. The high rates mostly occurred in the central-northern SNP and the largest rate of 14.11% per decade occurred in Keshan county. From the 1970s to 2010s, the highest rate of increase at the county level was only 5.61% in Taonan county, western SNP.
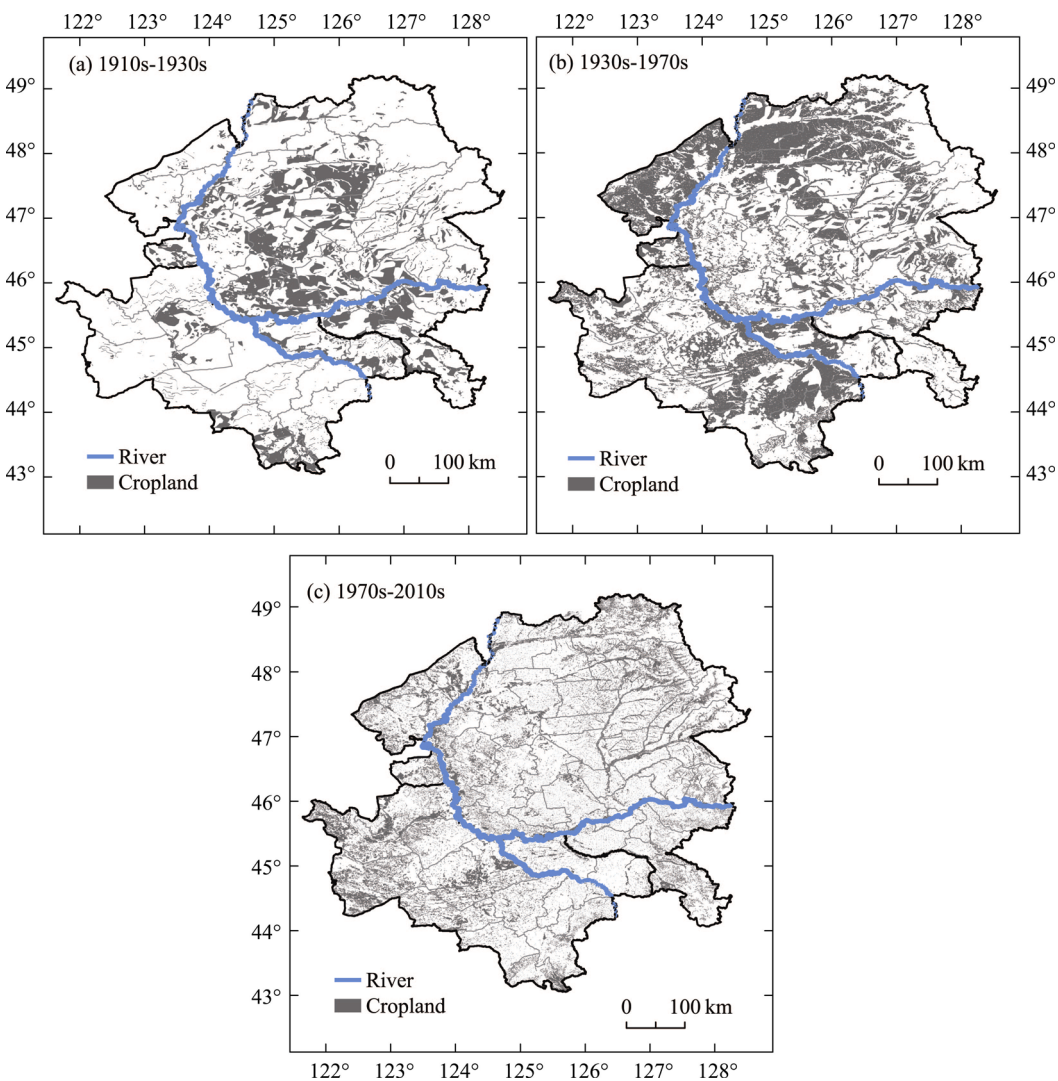
Figure 8 Cropland expansion with a pixel size of 1 km × 1 km in the Songnen Plain, Northeast China from the 1910s to 1930s (a), from the 1930s to 1970s (b), and from the 1970s to 2010s (c)

6 Discussion

(1) This spatially explicit reconstruction of cropland was based on essential local histori-cal documents and an estimation of the possibility of local cultivation after considering mul-tiple factors. This reconstruction not only confirmed the county-level estimations from Ye *et al*. (2009), but also presented a high-resolution grid dataset. This dataset could be directly offered as primary data to study the effects of human-induced LCCs on land surface heat/moisture flux and thus on regional climate change over the last century.

(2) This spatially explicit reconstruction of cropland demonstrated that HYDE and SAGE datasets had a low ability to capture temporal changes and spatial patterns of regional crop-land before the age of satellites. As mentioned above, the HYDE and SAGE datasets provided spatially explicit global cropland data. Figure 6 shows the total cropland area in the

SNP throughout the 20th century. We found the HYDE dataset presented a small low bias of cropland area compared with our estimations for the 1910s and 2010s. The underestimations of the HYDE data were about 1.0×10^4 km^2 in the 1910s and 1.2×10^4 km^2 in the 2010s. However, this dataset did not represent the characteristics of the increase in cropland area from the 1910s to 2010s; it presented a small rate of increase before the 1970s and a high rate of increase after the 1970s. Such a pattern is the reverse to that suggested by local historical-document-based estimations, which present a high rate of increase before the 1970s and a low rate of increase after the 1970s. Because of such a reversal, the underestimation of the HYDE dataset reached a peak in the 1970s. The estimated cropland area was about 4.2×10^4 km^2 in the HYDE dataset, while the local historical-document-based estimation was about 12×10^4 km^2 in the 1970s.

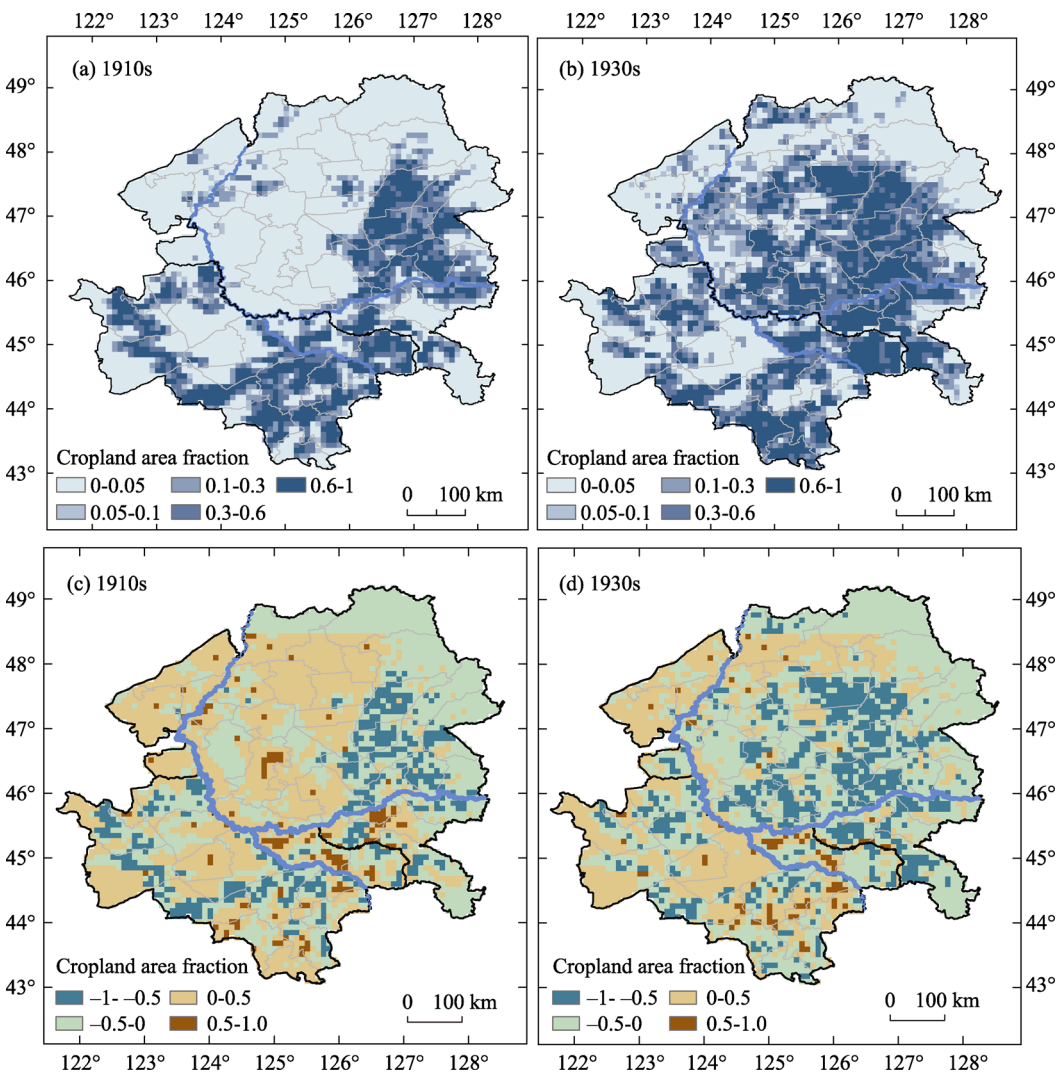Figure 9 illustrates the pixel-based estimation of cropland area fraction of the HYDE

Figure 9　HYDE-reconstructed cropland area fraction (a, b) and its differences compared with the new estimations of this study (c, d) at a cell size of 5 minutes × 5 minutes in the 1910s (a, c) and 1930s (b, d) in the Songnen Plain, Northeast China

dataset in the 1910s and 1930s. The HYDE dataset revealed large areas of cropland in the east and south, and little cropland in the west and north. Such a spatial distribution is similar to our estimation shown in Figure 7. However, we found that in the main body of agricultural development, the cropland fractions of the HYDE data were mostly lower than our estimations. Moreover, there were more pixels underestimated in the 1930s than in the 1910s. This was because the underestimation of total cropland area in the 1930s was much larger than that in the 1910s (Figure 6).

The SAGE dataset presented generally consistent results in cropland area using satellite-based estimations in the 1980s and 1990s (Figure 10). However, the SAGE dataset presented a much higher estimation in the 1910s; it estimated the cropland area to be about 11.0×10^4 km^2 while historical-document-based estimation was only about 4.92×10^4 km^2. Due to overestimation in the 1910s, the SAGE dataset did not reveal the high rate of increase in
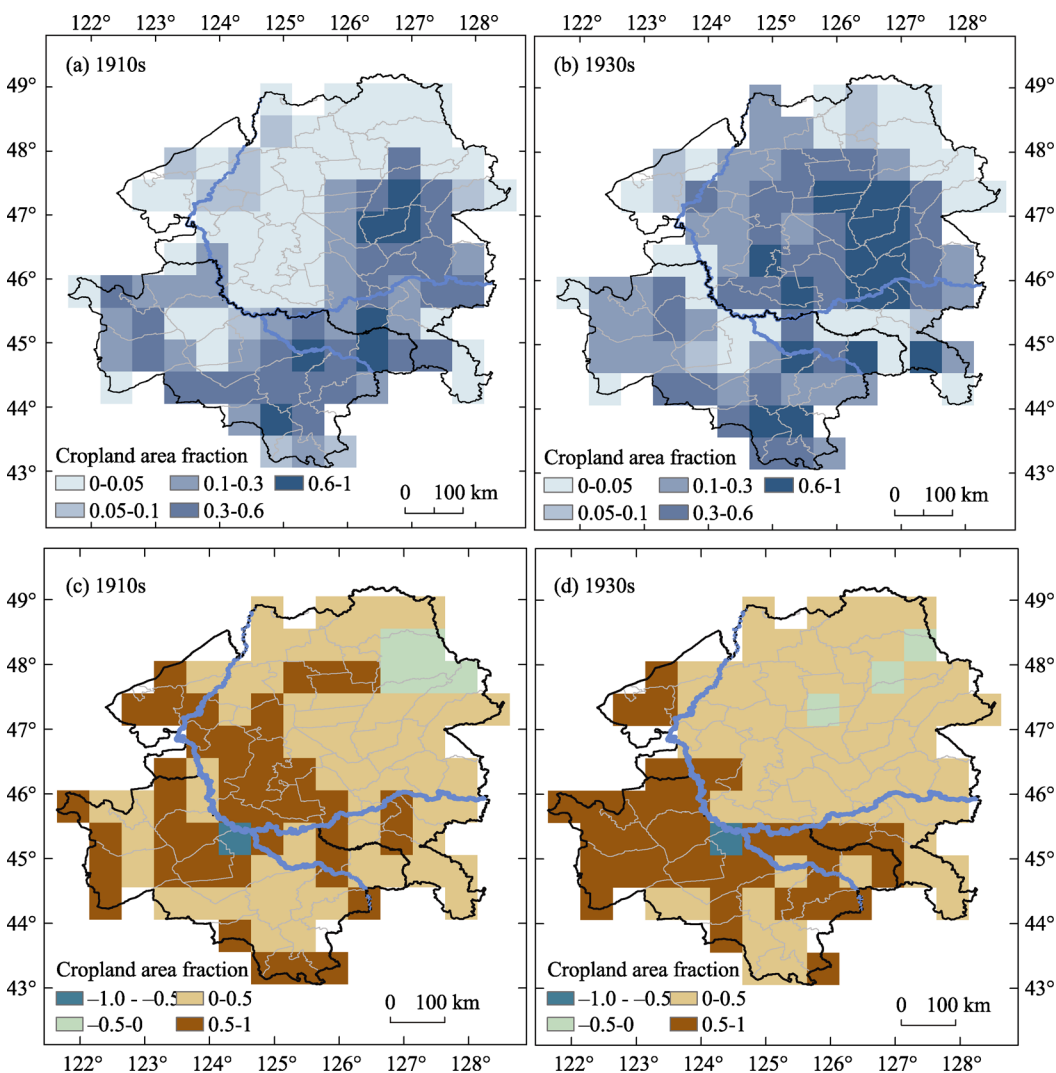

Figure 10   SAGE-reconstructed cropland area fraction (a, b) and its differences with the new estimations of the current study (c, d) at a cell size of 0.5°×0.5° in the 1910s (a, c) and 1930s (b, d) in the Songnen Plain, Northeast China

cropland in the 20th century. The SAGE data also showed a more extensive cropland area fraction over the SNP than the historical-document-based estimation in the 1910s and 1930s.

(3) The SNP's cropland area has increased notably for nearly a century. With the 1970s as the turning point, the increase in cropland area in the SNP can be divided into two periods of rapid development and slow growth (Figure 6). Farmland expansion can have different driving forces at different times, and the driving forces are listed below.

(i) Policies. In 1895–1914, because of the financial crisis during the war, the Qing Government proposed a "ban-reclaiming cancel policy" in Northeast China, which resulted in most of the wasteland in this region being reclaimed by the 1910s. Such a large-scale and fast land development has rarely been seen worldwide (Li *et al*., 2005; Dong, 2012). In the 1950s, to feed the country's population, the Chinese Government mobilized one hundred thousand demobilized military personnel and hundreds of thousands of urban youths to cultivate wastelands in Northeast China. Subsequently, cropland areas in the SNP expanded rapidly due to government policies.

(ii) Natural disasters. During the Qing Dynasty, natural disasters caused severe damage to the society in the lower reaches of the Yellow River. To eke out an existence in the old society, many people (from Shandong Province) braved the journey to Northeast China. This migration from Shanhaiguan to the northeast involved the population of so many provinces and lasted for such a long time that it became known as "the largest population movement in human history." The number of migrants reached over five million from the hinterland to Northeast China during 1923–1930 (Zhu, 2006). With a burgeoning population, there was a need for cropland areas to increase dramatically.

(iii) Transportation development. With the construction of the Manchuria Railway (1903), the pace of population migration from Shanhaiguan to the Northeast accelerated. According to statistics, as there were no railroads in Northeast China during 1894–1895, there were about 8 million, 2–2.5 million, and 1 million Chinese living in Liaoning, Jilin, and Heilongjiang provinces, respectively. However, 35 years later (1895–1930), Northeast China's population rose from more than 10 million to about 32 million. Among them, the population of Liaoning more than doubled and the populations of Jilin and Heilongjiang tripled (Zhu, 2006).

(iv) Other factors. Warfare, markets, and economic benefits have also had an impact on the development of agriculture in the SNP of Northeast China.

(4) There is no record of county population data to estimate the cultivated area, so the current study used population data mainly from the *Gazette of Heilongjiang Province* (Wan *et al*., 1992) and the *Local Gazette of Jilin Province* (CCCJP, 1992); these data sources provided accurate records of population data. This study estimated the cultivated areas of three counties (Dehui county, Shuangyang county, and Nenjiang county) using a sample of 24 counties to establish a population and arable land regression equation. This had little effect on the overall accuracy since population data were only used to estimate cultivated areas in three counties. A settlement distribution map, which was based on figures from the *Atlas of Historical Geography* (1910s) and *New Atlas of the Republic of China* (1930s), was obtained through vectorization. Because the two settlement maps were from two data sources, they were not comparable; the settlement distribution is a point on a large-scale map, so it inevitably resulted in space error during the process of vectorization.

Several relevant limitations of this study should be noted. The conversion of area units from old units to current ones is a difficult task in the field of historical geography. In this study, we used conversion ratios from Ye *et al*. (2009) and Li *et al*. (2010). Even though historical evidence was used by Ye *et al*. (2009) and Li *et al*. (2010), more evidence would be needed to verify these conversion ratios. Additionally, the potential cultivation area was estimated by excluding forests, wetlands, rivers, lakes, and mountains. Since these geographical factors were not temporal contemporaries with the historical cropland, the estimated cultivation area might not be accurate. Furthermore, it should be noted that we used 2 as the decision scale in the AHP model. This value will be validated or improved upon by collecting experience from more experts in the future.

7   Conclusions

This study used multiple historical documents to reconstruct the spatial distribution of cropland areas in the SNP, Northeast China, in the 1910s and 1930s. The original records from different historical documents were calibrated with each other. We subsequently defined an agricultural suitability index model and distributed the spatial cropland area into 1 km × 1 km size pixels using this model. The spatial distributions of cropland of the 1950s, 1970s, 1980s, 1990s, 2000s, and 2010s for the Songnen Plain were reconstructed based on the map of cultivated land distribution vectorization and remote sensing image interpretation. This study achieved a research goal of detecting a regional scale spatial pattern of cropland change over an extended time scale. It provided a surface drive parameter to simulate regional climate change research and also provided basic data for the study of human activities and the coupling of ecological and environmental effects.

The above findings demonstrated that the SNP was extensively developed during the last century with the cropland area fraction increasing from 22.8% in the 1910s to 60.9% in the 2010s. During this time, agricultural development exhibited large temporal and spatial variability. There was a turning point in the growth rate of cropland area in the 1970s; from the 1910s to 1970s, the cropland area fraction increased at a rate of 5.5% per decade, while from the 1970s to 2010s, the increase was only 1.3% per decade. Newly developed cropland was mostly situated in the central and northwestern parts of the SNP, since the eastern part of the SNP had been developed by the 1910s. From the 1910s to 1930s, new cultivation mainly occurred in the mid-SNP, while from the 1930s to 1970s new cultivation mainly spread over the western and northern parts.

References

Anderson-Teixeira K J, Snyder P K, Twine T E *et al*., 2012. Climate regulation services of natural and agricultural eco-regions of the Americas. *Nature Climate Change*, 2: 177–181.

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 Geographica Sinica*, 68(2): 245–256. (in Chinese)

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 Geographica Sinica*, 69(7): 896–906. (in Chinese)

Compiling Committee of Chorography in Heilongjiang Province (CCCHP), 1999. The Chorography of Heilongjiang Province. Harbin: Heilongjiang People's Publishing House. (in Chinese)

Compiling Committee of Chorography in Jilin Province (CCCJP), 1992. The Chorography of Jilin Province. Changchun: Jilin People's Publishing House. (in Chinese)

Course on the Survey of Manchurian Railway (CSMR), 1923. Local Gazette of Manchu-Mongolian. Manchuria Riri Press. (in Chinese)

Ding W J, Weng W H, Zeng S Y, 1934. New Atlas of China. Shanghai: Museum Returns Building. (in Chinese)

Dirmeyer P A, Niyogi D, Noblet-Ducoudre N *et al*., 2010. Impacts of land use change on climate. *International Journal of Climatology*, 30(13): 1905–1907.

Dong Q J, 2012. Research of the ban-reclaiming cancel and northeast agricultural development of the late Qing Dynasty [D]. Qiqihar: Qiqihar University (in Chinese).

Feng Y H, Zhang S H, He F N *et al*., 2014. Separate reconstruction of Chinese cropland grid data in the 20th century. *Progress in Geography*, 33(11): 1546–1555. (in Chinese)

Goldewijk K K, 2001. Estimating global land use change over the past 300 years: The HYDE database. *Global Biogeochemical Cycle*, 15 (2): 417–433.

Goldewijk K K, Beusen A, Drecht G V *et al*., 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12000 years. *Global Ecology and Biogeography*, 20(1): 73–86.

Goldewijk K K, van Drecht G, 2006. HYDE 3: Current and historical population and land cover. In: Bouwman A F, Kram T, Goldewijk K K (eds.). Integrated Modeling of Global Environmental Change. An Overview of IMAGE 2.4. Bilthoven, The Netherlands: PBL, 93–111.

He F N, Li S C, Zhang X Z, 2011. The reconstruction of cropland area and its spatial distribution pattern in the mid-Northern Song Dynasty. *Acta Geographica Sinica*, 66(11): 1531–1539. (in Chinese)

He F N, Li S C, Zhang X Z *et al*., 2013. Comparisons of cropland area from multiple datasets over the past 300 years in the traditional cultivated region of China. *Journal of Geographical Sciences*, 23(6): 978–990.

Jin X B, Cao X, Du X D *et al*., 2015. Farmland dataset reconstruction and farmland change analysis in China during 1661–1985. *Journal of Geographical Sciences*, 25(9): 1058–1074.

Liaoning Provincial Archives (LPA), 2008. Survey of Manchurian Railway (Vol. 3). Nanning: Guangxi Normal University Press. (in Chinese)

Li B B, Fang X Q, Ye Y *et al*., 2010. Accuracy assessment of global historical cropland datasets based on regional reconstructed historical data: A case study in Northeast China. *Science China Earth Sciences*, 53(11): 1689–1699.

Li D B, Shi F, 1987. A General Study of Immigration in Heilongjiang. Harbin: Heilongjiang People's Publishing House. (in Chinese)

Li J W, 1993. Forests in Heilongjiang. Beijing: China Forestry Publishing House. (in Chinese)

Li Q K, Wang S M, 2013. Systematic development of agriculture in Northeast China from a historical perspective. *Journal of Arid Land Resources and Environment*, 27(2): 11–17. (in Chinese)

Li S C, He F N, Chen Y S, 2012. Gridding reconstruction of cropland spatial patterns in Southwest China in the Qing Dynasty. *Progress in Geography*, 31(9): 1196–1203. (in Chinese)

Li S C, Zhang Y L, He F N, 2015. Reconstruction of cropland distribution in Qinghai and Tibet for the past one hundred years and its spatiotemporal changes. *Progress in Geography*, 34(2): 197–206. (in Chinese)

Li W, Zhang P, Song Y, 2005. Analysis on land development and causes in Northeast China during Qing Dynasty. *Scientia Geographica Sinica*, 25(1): 7–16. (in Chinese)

Lin S S, Zheng J Y, He F N, 2009. Gridding cropland data reconstruction over the agricultural region of China in 1820. *Journal of Geographical Sciences*, 19(1): 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. *Global Biogeochemical Cycles*, 24, GB3003.

Luo J, Zhang Y L, Liu G F *et al*., 2014. Reconstruction of cropland spatial patterns for 1726 on Yellow River-Huangshui River Valley in northeast Qinghai-Tibet Plateau. *Geographical Research*, 33(7): 1285–1296. (in Chinese)

Mahmood R, Pielke R A, Hubbard K G *et al*., 2010. Impacts of land use/land cover change on climate and future research priorities. *Bulletin of the American Meteorological Society*, 91: 37–46.

Notaro M, Liu Z, Williams J W, 2006. Observed vegetation-climate feedbacks in the United States. *Journal of Climate*, 19: 763–786.

Pan Q, Jin Y B, Zhou Y K, 2015. Gridding reconstruction of land use pattern in Jiangsu Province in the mid-Qing Dynasty. *Acta Geographica Sinica*, 70(19): 1449–1462. (in Chinese)

Pielke R A, Adegoke J, Beltran-Przekurat A *et al*., 2007. An overview of regional land-use and land-cover impacts on rainfall. *Tellus Teries B-Chemical and Physical Meteorology*, 59(3): 587–601.

Pielke R A, Niyogi D, 2009. The role of landscape processes within the climate system. In: Otto J C, Dikaum R (eds.). Landform-Structure, Evolution, Process Control: Proceedings of the International Symposium on Landforms organised by the Research Training Group 437, 115. Berlin: Springer, 67–85.

Ramankutty N, Foley J A, 2010. ISLSCP II Historical Croplands Cover, 1700–1992. Dataset. Available from: http://daac.ornl.gov/.

Ramankutty N, Foley J A, 1999. Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Glob. Biogeochem. Cycle*, 13(4): 997–1027.

Sun J Z, 1959. The Regional Economic Geography Science in Northeast China. Beijing: Science Press. (in Chinese)

Tan Q X, 1987. Historical Atlas of China: The Eighth Book. Beijing: China Cartographic Publishing House, 12–15. (in Chinese)

Wan F L, Zhang B Y, Cui C Q *et al*., 1992. The Chorography of Heilongjiang Province. Harbin: Heilongjiang People's Publishing House. (in Chinese)

Wang Y K, Tao J P, Liu F G *et al*., 2015. Reconstruction of cropland spatial pattern in 1830 in the middle reaches of Yarlung Zangbo River Valley. *Geographical Research*, 34(12), 2355–2367. (in Chinese)

Xin P L, Zhang F M, Gao X Y, 1999. The Development History of Heilongjiang. Harbin: Heilongjiang People's Publishing House. (in Chinese)

Xiong Z B, 2009. Summary of County Jurisdiction of Northeast Region in the Republic of China. Beijing: China National Microfilming Center for Library Resources. (in Chinese)

Xu J H, 2009. Mathematical Methods in Contemporary Geography. Beijing: Higher Education Press. (in Chinese)

Xu Z L, 1985. The Summary of Heilongjiang. Harbin: Heilongjiang People's Publishing House. (in Chinese)

Ye Y, Fang X Q, Ren Y Y *et al*., 2009. Cropland cover change in Northeast China during the past 300 years. *Science China*: *Earth Science*, 39(3): 340–350. (in Chinese)

Zeng L H, Song K S, Zhang B *et al*., 2010. Analysis of spatiotemporal variations in evapotranspiration and its influencing factors over the Songnen Plain in the growing season during the period 2000–2008. *Resources Science*, 32(12): 2305–2315. (in Chinese)

Zeng Z Z, Fang X Q, Ye Y, 2011. The process of land cultivation based on settlement names in Jilin Province in the past 300 years. *Acta Geographica Sinica*, 66(7): 985–993. (in Chinese)

Zhang L J, Jiang L Q, Zhang X Z, 2015. Spatially precise reconstruction of cropland areas in Heilongjiang Province, Northeast China during the late Qing Dynasty (1910s). *Journal of Geographical Sciences*, 25(5): 627–637.

Zhu Y G, Dai A G, 2006. Modern China: Social and Economic Research. Shanghai: Fudan University Press. (in Chinese)