A 28 time-points cropland area change dataset in Northeast China from 1000 to 2020

Ran Jia[1,3], Xiuqi Fang[1,2], Yundi Yang[1], Masayuki Yokozawa[3], Yu Ye[1,2]

[1]Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

[2]Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing 100875, China

[3]Faculty of Human Sciences, Waseda University, 2-579-15 Mikajima, Tokorozawa 359-1192, Japan

Correspondence to: Yu Ye (yeyuleaffish@bnu.edu.cn)

Abstract. Based on historical documents, population data, published results, remote sensing data products, statistical data and survey data, this study reconstructed the cropland area and the spatial pattern changes at 28 time points from 1000 to 2020 in Northeast China. 1000 to 1600 corresponds to historical provincial-level administrative districts, while 1700 to 2020 corresponds to modern county-level administrative districts. The main findings are as follows: (1) The cropland in Northeast China exhibited phase changes of expansion-reduction-expansion over the past millennium. (2) The the cropland area in Northeast China increased from 0.55×10^4 km^2 in 1000 to 37.90×10^4 km^2 in 2020 and the average cropland fraction increased from 0.37% to 26.27%; (3) From from 1000 to 1200, the cropland area exhibited an increasing trend, peaking in 1200. The scope of land reclamation was comparable to modern times, but the overall cropland fraction remained low. The cropland area significantly decreased between 1300 and 1600, with the main land reclamation area was reduced southward into Liaoning Province. From 1700 to 1850, the cropland area increased slowly, and the agricultural reclamation gradually expanded northward. After 1850, there was almost exponential growth, with the cropland area continuously expanding to the whole study area, and the growth trend persists until 2020; (4) The the dataset of changes in cropland of administrative districts in Northeast China, reconstructed based on multiple data sources and improved historical cropland reconstruction methods, significantly enhances time resolution and reliability. Additionally, the dataset shows relatively better credibility assessment resultsthe changing characteristics of cropland in Northeast China over the past millennium, especially over the past 300 years, which can provide a refined data base for building a historical cropland LUCC gridded dataset reconstruction, carbon emission estimation, climate data construction, etc. The dataset can be downloaded from https://doi.org/10.6084/m9.figshare.25450468.v2 (Jia, 2024).

1 Introduction

Anthropogenic land cover change (ALCC) is a key driver of global change, significantly impacting climate changeLand use and land cover change (LUCC) is not only one of the major manifestations of global change, but also an essential driving factor affecting global environmental change, especially global climate change (Arneth et al., 2017; Dickinson, 1991; Foley et al., 2005; Ito and Hajima, 2020; Ellis et al., 2021; Roberts, 2019)~~; Shukla et al., 1990)~~, over 70% of the Earth's land surface has undergone anthropogenic alterations over the past millennium (Sebastiaan et al., 2014; Shukla et al., 2019; Winkler et al., 2021).~~.~~ Cropland constitutes one of the primary land use types, being a land category susceptible to human influence and undergoing alterations~~. I~~, and it significantly influences food security, soil health, biodiversity, greenhouse gas emissions, and climate change (Friedlingstein et al., 2023; Godfray et al., 2010; Kalnay and Cai, 2003; Poschlod et al., 2005). Additionally, in recent years, croplands cover 12~14% of the global ice-free land (Shukla et al., 2019), Research on the long-term, accurate temporal and spatial changes in cropland are crucial for understanding the carbon budget resulting from human land reclamation, tracking sustainable food production, and other land-based ecosystem functions (Huang et al., 2024; Potapov et al., 2022; Saez-Sandino et al., 2024; Yu and Lu, 2018).

Presently, various global historical Land Use and Land Cover Change (LUCC) datasets, exemplified by the History Database of the Global Environment (HYDE), the Sustainability and the Global Environment (SAGE), the Pongratz Julia (PJ) and the Kaplan and Krumhardt 2010 (KK10) (Goldewijk et al., 2017; Kaplan et al., 2011; Pongratz et al., 2008; Ramankutty et al., 2008; Ramankutty and Foley, 1999), have been extensively employed in global change research. Such as carbon emission and carbon neutrality (Xu et al., 2024), climate data construction (Gortan et al., 2024), ecological footprint (Wang et al., 2024), and biological population assessment (Ye et al., 2024), etc. Furthermore, with the progress of research, historical LUCC study outcomes pertaining to the Northeast China have proliferated from a global scale down to the county level (Bai et al., 2007; Cao et al., 2021; He et al., ~~2022~~2023; Hurtt et al., 2020; Jia et al., 2023; Li et al., 2016; Li et al., 2018; Wu et al., 2020; Wu et al., 2022; Yang et al., 2017; Ye et al., 2009; Ye and Fang, 2012; Yu et al., 2021; Zhang et al., 2014; Zhang et al., 2022; Zeng et al., 2011; Tian, 2005; Jin et al., 2015; Shi, 2015; Zhang, 1991; Zhou, 2001). However, ~~there still exists~~ a disparity or uncertainty persists in the standardization and spatiotemporal accuracy of the aforementioned cropland data~~.~~, leading to ~~The cropland data with higher reliability within the region must be carefully selected across different temporal cross-sections. Additionally,~~ conflicts arise between datasets ~~with high spatiotemporal resolution standardization~~ and historical evidence of regional agricultural development~~ history~~. Consequently, enhancing the accuracy and credibility of historical LUCC datasets remains a focal point in international LUCC research (Gaillard et al., 2018; Yang et al., 2024; Yu et al., 2021). Reconstructing relatively accurate historical cropland cover at the basic-level administrative divisions based on actual historical agricultural development is a primary method for improving historical LUCC datasets (Goldewijk et al., 2017; Yu et al., 2021). For instance, the HYDE dataset demonstrates a boundary effect influenced by modern provincial administrative divisions in Northeast China, resulting in discontinuities in the spatial distribution of cropland in regions within the same historical agricultural development process. Considering the historical evolution of administrative divisions in China (Zhao et al., 2024), the cropland of smallest administrative divisions that can be reconstructed at present is the county-level, which suggests that it is possible to control the error of the gridded allocation to between 0.5° × 0.5° and 1° × 1°. –Therefore, ~~precise cropland change data, particularly~~

long-term precise cropland area change datasets standardized with basic-level administrative divisions and standardized time-points high spatiotemporal resolution will not merely improve the accuracy and credibility reliability of global historical LUCC datasets, but will also play a crucial role in enhancing the precision of climate and environmental simulations and supporting detailed environmental effect analyses in Northeast China.

Northeast China is one of the most important grain bases in China today. Northeast China stands as a representative area for reconstructing historical cropland. During the period between the two land reclamations (eleventh and twelfth centuries; from the nineteenth century to present), there was a prolonged period of nomadism in this area. Agricultural comprehensive development in the area commenced in the late 19th century, transforming it into a vital grain-producing base in China. The grain output constitutes 25.18% of the national total, with corn and soybean contributing 41.64% and 56.20%, respectively (National Bureau Of Statistics, 2023). A study has indicated that the supply centers for China's three major grains (wheat, corn, rice) significantly moved to the Northeast from 2000 to 2020, while the demand centers did not move simultaneously. This shift underscores the rapidly increasing importance of the Northeast China in ensuring China's food security (Xuan et al., 2023). Furthermore, the majority of China's black soil is distributed in Northeast China, which provides an important foundation for the productivity of crops. A study has pointed out that compared to other global black soil regions, the Northeast black soil region's yields of eight major crops (excluding rice) remained in the top three among the world's main black soil distribution countries from 2000 to 2015, with Russia and Ukraine occupying the first two positions (Wang et al., 2024). Additionally, long-term precise cropland area change data reflects the significance for soil and water conservation research in Northeast China, thereby ensuring food security. A typical case study in the Northeast China examined the long-term effects of cultivation on soil carbon, nitrogen, and bacterial community in the Northeast black soil region. The results indicated that prolonged cultivation (e.g., 152 years) led to a negatively and exponentially decline in SOC and total nitrogen (Liu et al., 2024).

The dataset in this study presents a critical update and extension of the former historical cropland cover change in the three provinces of Northeast China over the past 300 years (Ye et al., 2009). Throughout the prolonged agricultural development, the natural vegetation landscape in the Northeast region China has undergone notable transformations. In this study, we used the improved historical cropland reconstruction methods to reconstruct 28 time-points cropland area by assimilating multiple data sources in Northeast China from 1000 to 2020. The mainly new features of this dataset include: (1) Extended the reconstruction period to 1000~2020, aligning with the standard time-points of internationally established global historical LUCC datasets; (2) the reconstruction included the entire East of Inner Mongolia, which area accounts for approximately 45% of the Northeast China. (3) the smallest administrative divisions for the reconstructed cropland are at the provincial-level from 1000 to 1600, and at the county-level from 1700 to 2020. Our main objective is to provide a long-term time series of cropland area change dataset in Northeast China that is close to the historical "truth value" under a unified standard.

To quantitatively reconstruct historical Land Use and Land Cover (LUCC) datasets in this region, providing historical "truth value", enhancing the accuracy and reliability of global historical LUCC datasets, and supporting conservation and innovative research in the Northeast Black Soil region, the aim of the present study: (1) Systematically and comprehensively collect data materials for reconstruction, interpolate missing years in indicator data, and integrate diverse data from various sources and reconstruction methods into a unified set of long-term time series cropland area change datasets with standardized criteria. (2) Evaluate the accuracy of the reconstructed dataset at temporal and spatial scales. Assess the rationality of this dataset based on the historical facts of agricultural development in the Northeast China. Conduct a comparative analysis with global historical LUCC datasets to identify strengths and weaknesses.

2 Data and methods

2.1 The study area and the framework for cropland reconstruction

The definition of Northeast China in this study includes Heilongjiang, Jilin and Liaoning Provinces, Hulunbuir City, Hinggan League, Tongliao City, Chifeng City and Xilin Gol League of Inner Mongolia. Northeast China is located between 38°43' and 53°33' N and between 111°59' and 135°05' E, with a total area of approximately 1.45×10^6 km^2, about 15.1% of the total area of China, and the main part of Northeast China has a temperate continental monsoon climate. In this study, the seven time points from 1000 to 1600 are reconstructed based on the provincial-level administrative districts and derived from the Historical Atlas of China (Tan, 1982a; Tan, 1982b). For the period from 1700 to 2020, twenty-one time points are reconstructed based on the county-level administrative districts using the 1:1,000,000 public version of basic geographical information data released by the National Geomatics Center of China (2021 edition) (https://www.webmap.cn/commres.do?method=result100W, last access: 10 January 2024). For the sake of convenience in research and considering the historical evolution of each region, this study consolidates the administrative districts under each prefecture-level city in the Northeast China into a single administrative unit. Additionally, Nianzishan District is merged into Longjiang County, Bayuquan District into Gaizhou City, Qingmenhe District into Fuxin County, Qinghe District into Kaiyuan City, Zhanqian District into Dashiqiao City, Zhalainuoer District into Manzhouli City, Huolinguole City into Zhalute Banner, and Aershan City into Horqin Right Wing Front Banner.

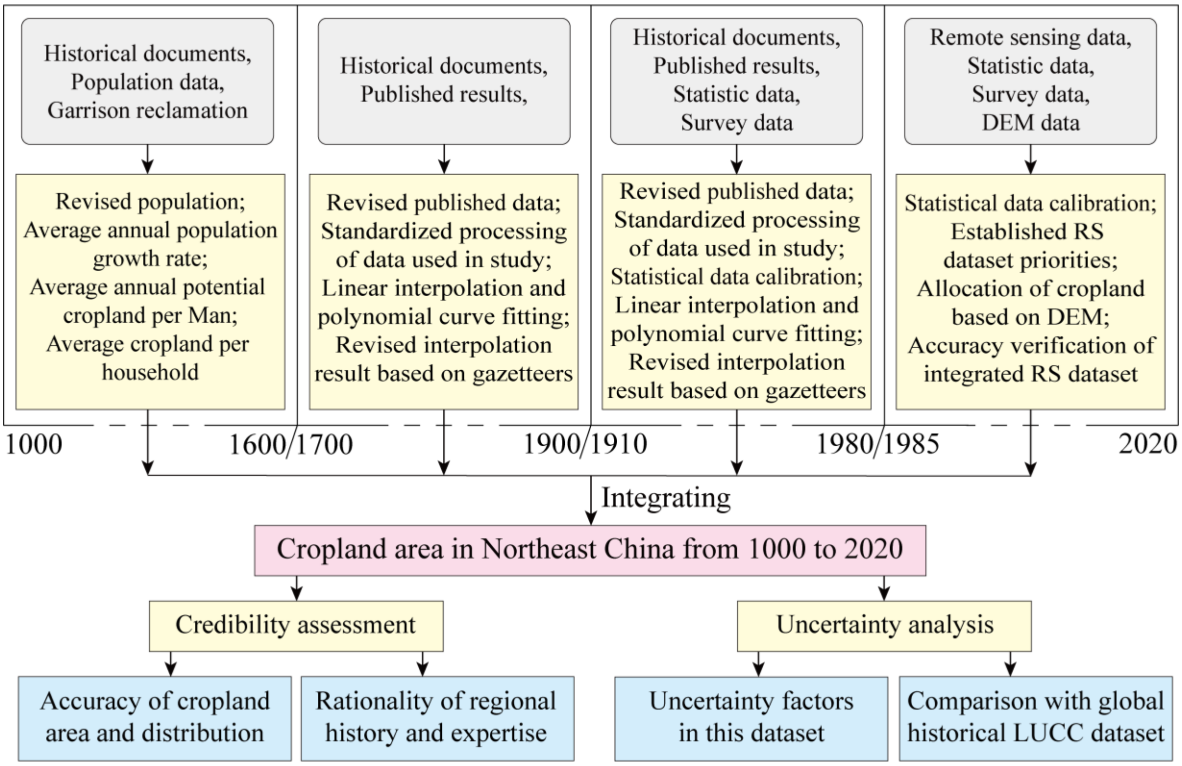


Figure 1: The framework for reconstructing cropland area of Northeast China from 1000 to 2020.

The framework of the cropland data reconstruction process in this study is illustrated in Fig. 1. It is essential to note that, unlike reconstructing historical cropland through simulation or speculation, the data foundation in this study incorporates historical literature, proxy data such as population data, revised published results, statistical data, survey data, and remote sensing data products. Historical period reconstruction primarily relies on population data from historical time points. Population data for adjacent standard time points are calculated using the average annual growth rate, and proxy indicators such as average annual cropland area per Man and average cropland area per household are employed to calculate cropland area. Additionally, after correcting published data and supplementing blank areas through standardized data processing, we used historical facts to interpolate cropland area from nearby time points to standard time points through linear interpolation. Trend extrapolation and total control are achieved through polynomial curve fitting. Finally, errors that may exist in the interpolation are corrected based on local gazetteers of China (https://fz.wanfangdata.com.cn/, last access: 10 January 2024). The reconstruction in the modern period primarily involves analyzing the linear relationship between statistical data and survey data. Survey data sequences established are used to control the cropland pixel data obtained through the regional-scale constrained integration of remote sensing data.

2.2 Data sources and reconstruction methods

2.2.1 Reconstruction of cropland area from 1000 to 1600

This study covers seven standard time points from 1000 to 1600, spanning the Liao, Jin, Yuan, and Ming dynasties. Due to the absence of direct records of cropland area during this period, cropland reconstruction primarily relies on historical documents, population data, and garrison reclamation data corresponding to the provincial-level administrative districts. During the Liao Dynasty period, this study based on the Dynastic History of Liao Dynasty and the History of Population in China (Wu and Ge, 2005a; Toqto'A, 1974) along with other published results (Ge, 2002; Han, 1999; Tan, 1982b), to reconstruct the agricultural and non-agricultural populations within five provincial-level administrative districts in 1111, with an average household size of 6.5 people, 2.08 of whom were Man (a male between the ages of 15 and 50 years in the Liao Dynasty). Population data for the five districts in 1000 and 1100 were calculated based on a 0.5% average annual population growth rate (Wu and Ge, 2005a).

During the Jin Dynasty period, this study is primarily based on the Dynastic History of Jin Dynasty and the History of Population in China (Wu and Ge, 2005a; Toqto'A, 1975) along with other published results (Li et al., 2018; Han, 1999; Jin and Mikami, 1984; Liu, 1994a; Liu, 1994b; Tan, 1982b), to reconstruct the agricultural and non-agricultural populations within five provincial-level administrative districts in 1207, with an average agricultural household size of 5.96 people, 2 of whom were Man (a male between the ages of 17 and 59 years in the Jin Dynasty), while an average non-agricultural household size of 10.59 people. Population data for the five districts in 1200 were calculated based on a 0.9% average annual population growth rate (Toqto'A, 1975).

When calculating cropland area during the Liao and Jin period (1000~~-~~ ~1200), this study primarily involves adjusting the agricultural and non-agricultural population quantities to standard time points. Combining with the constructed method of the average annual cropland area per Man for agricultural population and the average cropland area per household for non-agricultural population during the Liao and Jin Dynasties (Jia et al., 2023), the cropland areas for provincial-level administrative units in the Northeast China in the 1000, 1100, and 1200 are calculated separately (Table 1). The main algorithm applied in the Liao and Jin Dynasties can be found in the supplementary materials. Furthermore, due to the lack of significant technological changes in agricultural production in the Northeast China and the southward shift of the northern boundary of the farming-pastoral ecotone within the study area (He et al., ~~2022~~2023; Han, 2012; Zhang et al., 1997), this study maintains consistency with the Liao and Jin Dynasties. The average annual cropland area per Man for agricultural population is set at 14 *Mu* (0.93 hm^2), and the average cropland area per household for non-agricultural population is set at 2 *Mu* (0.13 hm^2) during the Yuan and Ming Dynasties (1300~1600).

During the Yuan Dynasty, this study primarily based on the Dynastic History of Yuan Dynasty (Song, 1976) to obtain the garrison reclamation area and corresponding number of soldiers in the Northeast China around 1300, and the average cropland area per garrison soldier is 100.1 *Mu* (6.67 hm^2).~~.~~ Additionally, based on the Dynastic History of Yuan Dynasty and the History of Population in China (Wu and Ge, 2005a; Cao and Ge, 2005b; Song, 1976) along with other published results (Cong, 1993a; Zhan, 2017; Xue, 2012; Zhou, 2021)~~-~~, this study reconstructs the number of ordinary households and Mongol households within the three provincial-level administrative districts of the study area during the Yuan Dynasty (Tan, 1982a). Ordinary households are further divided into Han households (agricultural population) and other minority ethnic households (non-agricultural population) in a 7:3 ratio (Cong, 1993b), with an average agricultural household size of 5 people, 1.67 of whom were Man (a male between the ages of 15 and 59 years in the Yuan Dynasty). Population data for garrison soldiers, Han households, minority ethnic households, and Mongol households in the three districts around 1300 are calculated based on different average annual population growth rates ranging from 0.6% to 1.8% during the Yuan Dynasty (Wu and Ge, 2005a). After obtaining the population data, this study subtracts the garrison soldiers and their corresponding households from the ordinary households. Subsequently, the remaining ordinary households are divided into Han households and minority ethnic households in a 7:3 ratio. The cropland area for agricultural population is calculated based on the average annual cropland area per Man for agricultural population, while the cropland area for non-agricultural population, including Mongol households, is calculated using the average cropland area per household for non-agricultural population referring the Liao and Jin Dynasties (Table 1).

During the Ming Dynasty, this study primarily based on the Dynastic History of Ming Dynasty (Zhang, 1974) to obtain the garrison reclamation area in the Northeast China around 1400. According to historical records and verification, it is determined that each garrison soldier in the Liaodong region possessed 46 *Mu* (3.07 hm^2) of cropland, with the proportion of garrison soldiers among soldiers being approximately 30%, and the number of dependents for each soldier being twice that of soldiers (Cao and Ge, 2005b; Li, 2019; Wang, 2009; Zhang, 1974). Additionally, based on the Dynastic History of Ming Dynasty and the History of Population in China (Cao and Ge, 2005b; Zhang, 1974) along with other published results (Cong, 1985; Kong and Feng, 1989; Li, 2019; Tan, 1982a), this study reconstructs the population of soldiers and their dependents, ordinary households/aborigines, and the population of minority ethnic households and Mongols (non-agricultural population) within the four provincial-level administrative districts in the 1400. Referring to historical records such as refugee migration, the construction of the Great Wall, and supplementary border garrisons (Cao and Ge, 2005b; Kong and Feng, 1989; Liu et al., 2016; Tan, 1982a), the historical maps for the 1500 and 1600 are divided into three provincial districts, and the ~~number of~~ population data for these two time points is obtained based on the aforementioned historical documents. During this period, all regular soldiers in the Dusi of Eastern Liao and one-third of their dependents would operate cropland as farmers. The average agricultural household (ordinary households/aborigines/refugees/migrants) size of 6, 2.25 of whom were Man (a male between the ages of 16 and 60 years in the Ming Dynasty) in the Dusi of Eastern Liao. The average non-agricultural household (minority ethnic households) size of 6, 2 of whom were Man in the Dusi of Nuergan, while size of the Mongol households is 5, 1.67 of whom were Man. Population data for soldiers and their dependents, ordinary households/aborigines/refugees/migrants, minority ethnic households in the Dusi of Nuergan, and Mongol households in the western part of the study area in the three provinces are calculated for the 1500 and 1600 based on average annual population growth rates of 0.8%, 0.5%, 0.2%, and 0.15%, respectively (Cao and Ge, 2005b). After obtaining the population data, we calculated the garrison reclamation area and civilian cropland area within the Dusi of Eastern Liao and the Dusi of Beiping based on the population of soldiers and agricultural population (ordinary households/aborigines) in the 1400. The minority ethnic population in the Dusi of Nuergan and the Mongol population in the Dada are calculated as non-agricultural population referring the Liao and Jin Dynasties (Table 1). For the 1500 and 1600, we calculated the garrison reclamation area and civilian cropland area within the Dusi of Eastern Liao based on the population of soldiers and agricultural population (ordinary households/aborigines/refugees/migrants). The minority ethnic population in the Dusi of Nuergan and the Mongol population in the Dada are calculated as non-agricultural population referring the Liao and Jin Dynasties (Table 1). The main algorithm applied in the Yuan and Ming Dynasties can be found in the supplementary materials.

Table 1: The index of cropland area reconstruction from 1000 to 1600

Period	Population type	Population (10^4)	Proportion of household registration	Corresponding cropland area	Total cropland area (km^2)
1000, 1100	Agricultural population	371(1000); 612(1100)	Average household size: 6.5 people, 2.08 of whom were Man	Average annual cropland area per Man is 14 *Mu* (0.93 hm^2)	5513(1000); 9078(1100)
	Non-agricultural population	140(1000); 231(1100)		Average cropland area per household is 2 *Mu* (0.13 hm^2)	
1200	Agricultural population	587	Average household size: 5.96 people, 2 of whom were Man	Average annual cropland area per Man is 14 *Mu* (0.93 hm^2)	16949
	Non-agricultural population	338	Average household size: 10.59 people	Average cropland area per household is 45.3 *Mu* (3.02 hm^2)	
1300	Garrison soldiers	0.8	Each soldier represents a household	Average per garrison soldier is 100.1 Mu (6.67 hm^2)—	4350
	Agricultural population	111	Average household size: 5 people, 1.67 of whom were Man	Average annual cropland area per Man is 14 *Mu* (0.93 hm^2)	
	Non-agricultural population (Minority ethnic household)	137		Average cropland area per household is 2 *Mu* (0.13 hm^2)	
1400	Soldiers and their dependents	70	Approximately 30% of garrison soldiers; Soldiers : dependents = 1 : 2	Average per garrison soldier is 46 *Mu* (3.07 hm^2)	2790
	Agricultural population (ordinary households/aborigines)	10	Average household size: 6 people, 2.25 of whom were Man	Average annual cropland area per Man is 14 *Mu* (0.93 hm^2)	
	Non-agricultural population (Minority ethnic household, Mongol household)	40	Average minority ethnic household size: 6 people, 2 of whom were Man; Mongol household size: 5, 1.67 of whom were Man	Average cropland area per household is 2 *Mu* (0.13 hm^2)	
1500, 1600	Soldiers and their dependents	25(1500); 12(1600)	Approximately 30% of garrison soldiers; Soldiers : Dependents	Average per garrison soldier is 46 *Mu* (3.07 hm^2); Regular	4875(1500); 5868(1600)

			= 1 : 2	soldiers and one-third of their dependents is 14 *Mu* (0.93 hm^2)	
	Agricultural population (ordinary households/aborigines/refugees/migrants)	83(1500); 137(1600)	Average household size: 6 people, 2.25 of whom were Man	Average annual cropland area per Man is 14 *Mu* (0.93 hm^2)	
	Non-agricultural population (Minority ethnic household, Mongol household)	68(1500); 81(1600)	Same as 1400	Average cropland area per household is 2 *Mu* (0.13 hm^2)	

2.2.2 Reconstruction of cropland area from 1700 to 1900

The reconstruction of cropland in this study at five standard time-points from 1700 to 1900 is primarily based on published results and historical documents. Among them, published results utilize the county-level cropland fraction data (CNEC) reconstructed by Ye (Ye et al., 2009) ~~Yu~~ for the three provinces in Northeast China in 1683, 1735, 1780, and 1908 ~~(Ye et al., 2009)~~. Additionally, data on cropland fraction for 15 counties and districts, including Chifeng City, Balinzuo Banner, Balinyou Banner, Linxi County, Wengniute Banner, Kalaqin Banner, Ningcheng County, Aohan Banner, Kulun Banner, Naiman Banner, Taipusi Banner, Xianghuang Banner, Zhengxiangbai Banner, Zhenglan Banner, and Duolun County, reconstructed by Tian (Tian, 2005) ~~Yanyu~~, are available for the years 1724, 1782, 1868, and 1911 ~~(Tian, 2005)~~. Detailed description of the data and methods for these published results can be found in the supplementary materials.

~~When reconstructing the cropland data for the three provinces in Northeast China during this period, Ye (Ye et al., 2009) primarily utilized historical documents such as General Chorography of Shengjing and statistical data from the late Qing Dynasty's land survey (A, 1997; Li, 1991; Li et al., 2005; Ji et al., 2002; Yang et al., 1990). The methods employed for processing the cropland data include 4 aspects: (1) conversion and standardization of measurement units; (2) correction of the hidden percentage of cropland area in the historical records; (3) estimation of property of Mu used in Northeast China; (4) estimation of cropland area based on population data. Among these, (1) conversion and standardization of measurement units: Converting the diverse measurement units used for different regions and types of cultivated land in the Northeast China during the Qing Dynasty into a unified area measurement unit. Additionally, the conversion of Qing Dynasty area measurement units to km2 was performed based on proportional relationships (Wu, 1984). (2) Correction of the hidden percentage of cropland area in the historical records: Historical literature from the Qing Dynasty and the government's published cropland data may contain concealment or inaccuracies (Buck, 1941; Shi, 2000). Hence, a correction of 20% was applied to the original cropland data. (3) Estimation of property of Mu used in Northeast China: By analyzing the tax system in Northeast China, it is found that the taxation was different in the same area of cropland in high, middle and low productivity. That means the cropland area recorded in the historical document was the real amount, and the problem of tax Mu could be ignored in Northeast China. (4)~~

Estimation of cropland area based on population data: In cases where exact cropland area was unavailable, the minimum requirement of 3 Mu (0.2 hm2) per capita were used to estimate and interpolate cropland area based on population data (Fang et al., 2006). In reconstructing cropland data for the Eastern Mongolian region during this period, Tian (Tian, 2005) primarily relied on historical documents such as local gazetteers and official government records. The methods employed for processing cropland data include 4 aspects: (1) historical data summarization; (2) proportional estimation; (3) population-based estimation; (4) linear interpolation. Among these, (1) historical data summarization: By extensively collecting data, the scrutinized and analyzed data were directly used as the historical cropland area. (2) Proportional estimation: In the absence of exact cropland area, the cropland within a region was proportionally distributed to sub-regions based on the ratio of historical and contemporary cropland area, or the cropland area in some regions was used to estimate the overall cropland area. (3) Population-based estimation: Based on the proportional relationship between the number of farmers and the cropland area in the region during historical periods, cropland area was estimated using population data. (4) Linear interpolation: In cases where the trend of cropland change in the study area did not exhibit significant variations, linear interpolation was applied using the already reconstructed cropland area results for various time sections, ultimately generating cropland areas for multiple time sections.

Before utilizing the published results, this study examined and corrected issues present in the data, unifying it onto the base map used in this study. (1) Correction of published results: CNEC data (Ye et al., 2009) was adjusted based on the historical evolution of administrative boundaries to modern county-level administrative units. In 1908, cropland areas were missing for Qian Gorlos Mongolian Autonomous County, Jiaohe City, Yanji City, Wangqing County, Huichun City, Helong City, and Huinan County in Jilin Province. Wu (Wu, 2021) interpolated these missing values using the principles of geographical proximity and similarity in the regional agricultural development stage. By following the above method, we interpolated data for problematic counties in Jilin Province from CNEC data using settlement names evolution data for the past 300 years (Zeng et al., 2011). It is worth noting that for certain time points, due to the absence of cropland in neighboring counties, this study adopted the approach of multiplying the cropland area owned by unit settlements within Jilin Province at that time by the number of settlements in the respective county to obtain the cropland area (Appendix ATable S1). Furthermore, discrepancies were identified in used CNEC data for some counties in Heilongjiang and Liaoning provinces compared to published data. This study corrected these inconsistencies after verifying historical documents (Appendix ATable S1).

(2) Unified administration boundaries: The CNEC data (Ye et al., 2009) in 1683, 1735, and 1780 corresponds to historical Qing Dynasty administrative districts, and the administrative districts used in 1908, 1914, 1931, 1940, 1950, and 1980 also differed from that of this study. The approach taken in this study involves unifying the cropland fraction within each county or district. The modern county-level administrative vector map used in this study is overlaid with Ye's county-level cropland fraction map. Then we calculated the area of overlap between each county or district in this study and Ye's corresponding

county or district and then calculates the cropland area based on the proportional statistics. Similarly, for the Tian's data (Tian, 2005) used in this study for cropland fraction in 1724, 1782, 1868, 1911, and 1933, the same method is applied to unify them onto the modern map used in this study.

(3) Linear interpolation and polynomial curve fitting to ~~control~~ obtain the ~~total~~ cropland area: Previous studies have used the linear interpolation and polynomial curve fitting to reconstruct cropland areas (He et al., 2017; Jin et al., 2015; Ramankutty and Foley, 1999; Wei et al., 2016; Wei et al., 2021; Ye et al., 2015; Yu, 2019; Fang et al., 2021), and the interpolated data did not reduce the credibility of their datasets. In addition, p~~P~~revious studies have shown that in the process of reclamation in the Northeast China over the past 300 years, 1860 was a dividing point between slow growth and rapid growth, mainly due to the implementation of the immigration and reclamation policy by the Qing government (Fang et al., 2020; Ye et al., 2009; Fang et al., 2005; Kong and Feng, 1989). Therefore, this study selected the CNEC data (Ye et al., 2009) in 1683, 1735, ~~and~~ 1780, 1908 and 1914 for linear interpolation and polynomial curve fitting of cropland area data for each county or district in the three provinces of ~~the~~ Northeast China, obtaining data for 1700, 1750, 1800, ~~and~~ 1850 and 1900. ~~The data for 1908 and 1914 were selected to linearly interpolate the cropland area data for each county or district in the three provinces of the Northeast China to obtain data for 1900. Based on polynomial curve fitting trend extrapolation, the cropland area data at the above time points were obtained at the provincial level, controlling the total cropland area in the three provinces of the Northeast China.~~ In addition, this study selected the data from Tian (Tian, 2005) in 1724, 1782, ~~and~~ 1868, and 1911; ~~and~~ the CNEC data (Ye et al., 2009) in 1735; the data from Ye (Ye and Fang, 2012) in 1916 for linear interpolation and polynomial curve fitting to obtain cropland area data for 1700, 1750, 1800, ~~and~~ 1850, and 1900 in the Eastern of Inner Mongolia. ~~The data for 1911 from Tian (Tian, 2005) and the data for 1916 from Ye (Ye and Fang, 2012) were linearly interpolated to obtain cropland area data for 1900 in the Eastern of Inner Mongolia. Similarly, polynomial curve fitting trend extrapolation was used to obtain the total cropland area at the provincial level for the above time points as a reference.~~ The problems that may be encountered during the operation and the corresponding solutions are as follows:

①Linear interpolation and determination of zero values. The time points involved in this issue include 1700 and 1750 for the three provinces of Northeast China; 1750, 1800, and 1850 for East of Inner Mongolia. For instance, in Northeast China, the cropland area in each county in 1700 is interpolated based on records from 1683 and 1735. At 1700, there are no negative values, but there may be zero values. Specifically, the cropland value in 1683 is 0, while there is definite value in 1735. Our solution involves consulted contemporary county gazetteers to verify the history of land reclamation in 1700. If so, a polynomial curve fitting trend extrapolation was applied to obtain the proportional relationship at the provincial level for adjacent points on the extrapolated trend. Then this proportion was multiplied by the cropland area of the county at the adjacent time-point to obtain the cropland area at that time-point. If the land was not reclaimed, the value at that time point was considered as zero. Similarly, other counties involved in interpolation adopt the same solution when encountering this situation.

②Polynomial curve fitting and correction of negative values. Besides the previously mentioned linear interpolation, polynomial curve fitting based on the least squares method may encounter problems with data points resulting in negative values. First of all, the main reason for this issue is our historical determination that 1860 was a dividing point between slow and rapid growth. Therefore, we use 1860 as a breakpoint and separate interpolated the data for Ye (Ye et al., 2009, Ye and Fang, 2012) and Tian (Tian, 2005) before and after this period. Second, for time points that cannot be directly obtained through linear interpolation, cropland need to be calculated by polynomial fitting backwards (1800 and 1850 in the three provinces of Northeast China; 1900 in East of Inner Mongolia). For instance, in Northeast China, cropland area in each county in 1800 and 1850 are derived from data in 1683, 1735, and 1780 using polynomial curve fitting method. Some counties may show a decline in cropland, potentially resulting in negative values in the extrapolation results. Our solution involves using the proportion of provincial administrative level to multiply by the cropland area in 1780 for correction in the counties' cropland area in 1800 and 1850. Third, for time points that cannot be directly obtained through linear interpolation, cropland need to be calculated by polynomial fitting forwards (1900 in the three provinces of Northeast China; 1700, 1910 in East of Inner Mongolia). For instance, in Northeast China, cropland area in each county in 1900 is derived from data in 1908 and 1914 using polynomial curve fitting method. Due to rapid growth of cropland in some counties from 1908 to 1914, the extrapolation for 1900 may result in negative values. Our solution involves using the proportion of provincial administrative level to multiply by the cropland area in 1908 for correction in the counties' cropland area in 1900.

It should be noted that, considering the historical development process of Northeast China during the Qing Dynasty, war factors, and the encouraging land reclamation policies implemented by the Qing government after 1860, we determined that the cropland area in each county of Northeast China in 1900 would not significantly exceed that of 1908. During this period, in Northeast China, the total cropland area was gradually increasing and was not significantly affected by events such as the Second World War, which led to a notable decrease in cropland area in 1950 compared to 1930 and 1940. Therefore, when the extrapolated value for a county in 1900 exceeds that of 1908, the proportion of provincial administrative level is used to multiply by the cropland area in 1908 for correction in the county's cropland area in 1900.

③The determination of initial cultivation occurred between 1780 and 1908. Few counties in Northeast China where cropland was zero in 1683, 1735, and 1780, but had cropland in 1908. Our solution involves consulted contemporary county gazetteers to verify the history of land reclamation between 1800 and 1900. If local gazetteers indicate the initial cultivation occurred before 1860, this study applies the same method as described in "①Linear interpolation and determination of zero values". If the initial cultivation began after 1860, this study applies the same method as described in "②Polynomial curve fitting and correction of negative values". All the counties where this situation occurs can be found in Table S2.

~~(4) Based on local gazetteers to correct negative or zero values of cropland: After obtaining the interpolation results for the five standard time points from 1700 to 1900 in each county or district in the study area, we found that the cropland area in~~

some counties or districts was negative or zero. For counties or districts with zero values, this study consulted contemporary county gazetteers to verify the history of land reclamation, confirming whether the zero values at certain points are reasonable. For points that have been reclaimed, a polynomial curve fitting trend extrapolation was applied to obtain the proportional relationship at the provincial level for adjacent points on the extrapolated trend. This proportion was multiplied by the cropland area of the county or district at the adjacent point to obtain the cropland area at that point. Similarly, for counties or districts with negative values, the same method was used to estimate the values based on the history of land reclamation. If the land was not reclaimed, the value at that time point was considered as zero.

2.2.3 Reconstruction of cropland area from 1910 to 1980

The reconstruction of cropland at eight standard time points from 1910 to 1980 in this study is mainly based on published results, historical documents, statistical data, and survey data. Among these, the published results include the cropland fraction for the three provinces in Northeast China in 1908, 1914, 1931, 1940, 1950, and 1980 (CNEC) (Ye et al., 2009). As well as the cropland fraction for the farming-pastoral ecotone area reconstructed by Ye in 1916 and 1940 (Ye and Fang, 2012). Additionally, Tian's reconstruction provides cropland fraction for 15 counties in the Eastern of Inner Mongolia in 1911 and 1933 (Tian, 2005). Historical documents include the Summary of county governance in Northeast China (Xiong, 1933) to supplement cropland area data for the Eastern of Inner Mongolia in 1931. Statistical data include Agricultural and Animal Husbandry Production Statistics (Inner Mongolia Provincial Bureau Of Statistics, 1983) to obtain county-level cropland area for the Eastern of Inner Mongolia in 1950, 1960, 1970, and 1980. Survey data include Manchuria Economic Statistics Charts (Office Of The Governor-General Of Kwantung, 1918) to obtain prefecture-level cropland area data for the Eastern of Inner Mongolia in 1917 as a reference. The North Manchuria and East Support Railway (East Branch Railway Administration Of Russia and South Manchuria Railways Co., 1923; East Branch Railway Administration Of Russia and South Manchuria Railways Co., 1923) is used as survey data to supplemented for various counties in the Eastern of Inner Mongolia in 1911 and 1914, which was not covered by existing data from Ye and Tian. Additionally, a digital version of the Manchuria Political Map from this document was used to obtain county-level district maps for Northeast China in the 1920s. Detailed description of the data and methods for these published results can be found in the supplementary materials.

In the published results used in this study, CNEC data primarily utilized government files or investigation reports, Japanese and Russian survey data, official statistical data and survey data (Committee Of Science And Technology In Northeast China, 1946; National Bureau Of Statistics, 1989; Heilongjiang Provincial Bureau Of Statistics, 1997; Jinlin Provincial Bureau Of Statistics, 1997; Liaoning Provincial Archives et al., 1988; Liaoning Provincial Bureau Of Statistics, 1997; Xin et al., 1999; Xiong, 1933; Ye et al., 2006; Middle East Railway Economic Survey, 1931; Committee Of Integrative Survey Of Natural Resources and Committee Of National Planning Of Chinese Academy Of Sciences, 1989). The methods employed for processing the cropland data include 2 aspects: (1) standardization of multi-sourced data; (2) correlation analysis between statistical data and survey data. (1) Standardization of multi-sourced data: The study area scope of different datasets was determined, measurement units were standardized, and data from different sources during the same period were cross-verified and compared at the county-level. (2) Correlation analysis between statistical data and survey data: Analyzing the correlation between statistical data and survey data for each county in the study area during the same period, then calculating their linear regression equations and explanatory variances. When reconstructing cropland data for the Eastern of Inner Mongolia during the same period, Ye's data (Ye and Fang, 2012) mainly consisted of government statistical reports and Japanese survey reports (Committee Of Science And Technology In Northeast China, 1946; South Manchuria Railways Co., 2015; Ministry Of Agricultural And Commercial, 1919). The methods employed for processing the cropland data include 2 aspects: (1) conversion and standardization of measurement units; (2) correlation analysis between statistical data and survey data. These two data processing methods have been introduced in the previous sections. In the reconstruction of cropland data for the Eastern of Inner Mongolia during this period, Tian (Tian, 2005) primarily used local gazetteers, statistical data and survey data (Cropland Research Group, 1992; Ho, 1988). The methods employed for processing the cropland data mainly involved a comparative analysis of statistical data and survey data: Comparing the obtained statistical data on cropland area at different time points with survey data at the county-level, assessing the accuracy of the data used in the reconstruction.

Before using the published data from this period, this study also assessed and corrected the issues present in the data. Additionally, when supplementing the data using historical documents, statistical data and survey data, this study referred to the data processing methods of the aforementioned published studies. (1) Correction of published results: This study has provided specific explanations for the correction of CNEC data for this period in previous sections, as detailed in Appendix ATable S1.

(2) Standardization of Data: This study adopted the processing method used by Ye (Ye et al., 2006) for the Summary of county governance in Northeast China (Xiong, 1933). It converted the Qing Dynasty's *Mu* unit to the standard unit of measurement, square kilometers (km^2), and made a 10% correction to align this data with the survey data. For the Manchuria Economic Statistics Charts and the North Manchuria and East Support Railway (Office Of The Governor-General Of Kwantung, 1918), this study followed Ye's (Ye et al., 2006) analysis method for similar survey data, treating it as the actual cropland area. Regarding the standardization of administrative boundaries, this study utilized the digitized Manchurian Political Map and employed the method aforementioned to map it onto the modern administrative boundary map used in this study. The standardization of measurement units followed the conversion from the measurement units used in the Japanese survey data to the universal unit of measurement, square kilometers (km^2), as per Weights and Measures in Northeast China (South Manchuria Railways Co., 1927).

(3) Correlation analysis between statistical data and survey data: In this study, we referred the method used by Ye (Ye et al., 2009) in analyzing statistic data for the simultaneous period in the three provinces in Northeast China to process the county-

level cropland area statistical data for the 1950, 1960, 1970, and 1980 in the Eastern of Inner Mongolia (Inner Mongolia Provincial Bureau Of Statistics, 1983). It is found a stronger correlation between the statistical data and land survey data in 1985 (National Bureau Of Statistics, 1989; Committee Of Integrative Survey Of Natural Resources and Committee Of National Planning Of Chinese Academy Of Sciences, 1990), with a linear regression equation of y=1.3234x and R^2=~~98.51%~~0.9851 (Fig. 2). That means the land survey data in the Eastern of Inner Mongolia is approximately 32.34% higher than the corresponding statistical data, then corrected cropland area data by 32.34% for each county in the Eastern of Inner Mongolia for the 1950, 1960, 1970, and 1980.

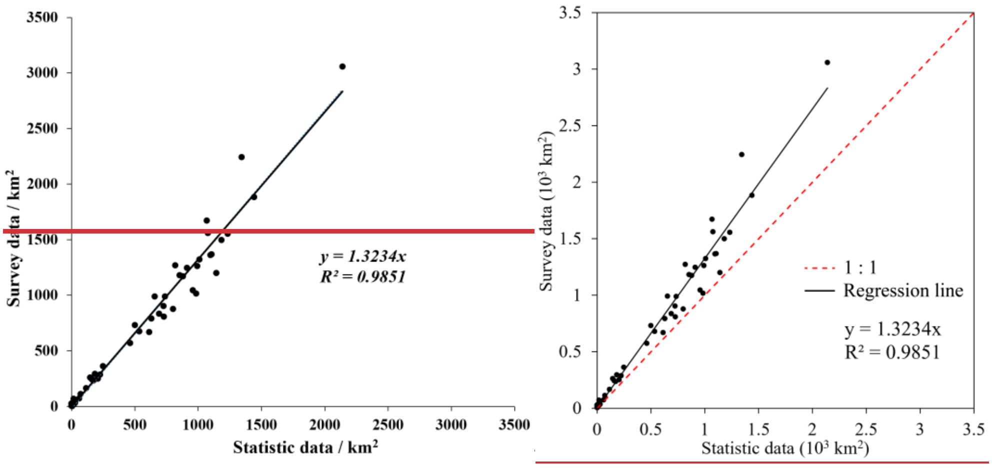


Figure 2: Correlation between the statistical cropland data and survey cropland data of the counties in the Eastern of Inner Mongolia in 1980's.

(4) Linear interpolation and polynomial curve fitting to ~~control~~ obtain the ~~total~~ cropland area: This study selected CNEC (Ye et al., 2009) data in 1908 and 1914 for linear interpolation and polynomial curve fitting of cropland area data for each county or district in the three provinces of ~~the~~ Northeast China, obtaining data for 1910 and 1920. ~~Similarly, polynomial curve fitting trend extrapolation was used to obtain the total cropland area at the provincial level for the above time points as a reference.~~ Additionally, this study selected the data from Tian (Tian, 2005) in 1911 and the data from Ye (Ye and Fang, 2012) in 1916 and 1940, and the corrected data in 1931 from Summary of county governance in Northeast China (Xiong, 1933) for linear interpolation and polynomial curve fitting of cropland area data for each county or district in the Eastern of Inner Mongolia, obtaining data for 1910 and 1920. Since the following operations are the same as 1700~1900, and the problems that may be encountered during the operation and the corresponding solutions have been detailed above, it will not be repeated here.~~Similarly, polynomial curve fitting trend extrapolation was used to obtain the total cropland area at the provincial level for the above time points as a reference.~~

It should be noted that this study considers the corrected data in 1931 in various counties of the Northeast China as data for 1930.

~~(5) Based on local gazetteers to correct negative or zero values of cropland and supplementing data for blank counties in the Eastern of Inner Mongolia in 1940: After obtaining the interpolation results for the eight standard time points from 1910 to 1980 in each county of the study area, we found that the cropland area in some counties or districts was negative or zero in 1910, and we adopted the processing method aforementioned for correction.~~ In addition, the cropland area data for the year 1940 mainly based on the corrected published results. For the missing data in single-digit ~~few~~ counties of the Eastern of Inner Mongolia, this study uses data recorded in local gazetteers to fill in the gaps.

2.2.4 Reconstruction of cropland area from 1985 to 2020

The reconstruction of cropland in this study from 1985 to 2020 at eight standard time points is primarily based on remote sensing data products, statistical data, survey data, and DEM data. Among these, eight sets of remote sensing data products were used (Table 2): AGLC (Xu et al., 2021), CLDC (Yang and Huang, 2021), ESA_WorldCover (Zanaga, 2021), Esri_LandCover (Karra et al., 2021), FROM_GLC (Gong et al., 2013), GFSAD30 (Thenkabail et al., 2021), GLC_FCS30 (Zhang et al., 2023), GlobeLand30 (Chen et al., 2015). It is worth mentioning that we conducted research on ESA_WorldCover and Esri_LandCover after resampling them to a resolution of 30 meters. Survey data includes the year 1985 county-level first general land investigation~~land survey data~~ (Committee Of Integrative Survey Of Natural Resources and Committee Of National Planning Of Chinese Academy Of Sciences, 1989), provincial-level data from the first national land survey (Li, 2000), prefecture-level data from the second national land survey, and county-level data from the third national land survey (https://gtdc.mnr.gov.cn/Share#/, last access: 10 January 2024).

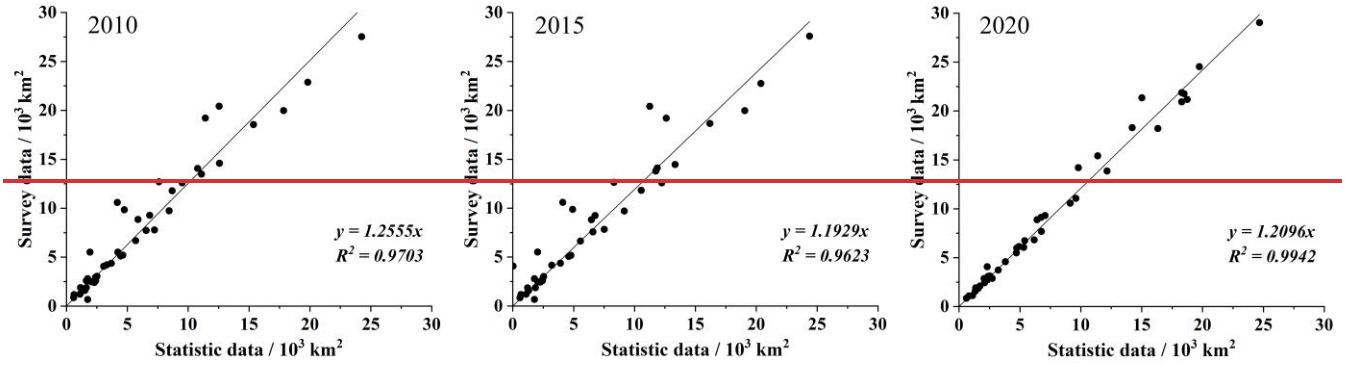
Table 2: Characteristics of the eight RS products

Product	Satellite Sensor	Type	Resolution	Year	Cropland Classes	URL	Reference
AGLC	Landsat 5 TM Landsat 7 ETM+ Landsat 8 OLI	Boolean	30m	2000, 2005, 2010, ~~0-~~ 2015	10.Cropland	https://code.earthengine.google.com/?asset=users/xxc/GLC_2000_2015 [2024/01/10]	(Xu et al., 2021)
CLDC	Landsat 8 OLI TM ETM+	Boolean	30m	1985, 1990, 1995, 2000, 2005, 2010, 2015, 2019~~-2020~~	1.Cropland	https://doi.org/10.5281/zenodo.4417810 [2024/01/10]	(Yang and Huang, 2021)
ESA_World	Sentinel-1	Boolean	10m	2020	40.Cropland	https://viewer.esa-	(Zanaga,

Cover	Sentinel-2					worldcover.org/worldcover/ [2024/01/10]	2021)
Esri_LandCover	Sentinel-2	Boolean	10m	2020	5.Crops	https://livingatlas.arcgis.com/landcover/ [2024/01/10]	(Karra et al., 2021)
FROM_GLC	Landsat TM, ETM+, OLI	Boolean	30m	2010, 2015	10.Cropland	https://data-starcloud.pcl.ac.cn/zh [2024/01/10]	(Gong et al., 2013)
GFSAD30	Landsat ETM+ OLI	Boolean	30m	2015	2.Cropland	https://lpdaac.usgs.gov/products/gfsad30aunzcnmocev001/ [2024/01/10]	(Thenkabail, 2021)
GLC_FCS30D	Landsat OLI	Boolean	30m	1985, 1990, 1995, 2000, 2005, 2010, 2015, ~~-~~2020	10.Rainfed cropland 11.Herbaceous cover 12.Tree or shrub cover (Orchard) 20.Irrigated cropland	https://zenodo.org/records/8239305 [2024/01/10]	(Zhang et al., 2023)
GlobeLand30	Landsat TM/ETM+, HJ-1	Boolean	30m	2000, 2010, 2020	10.Cropland	http://www.webmap.cn/mapDataAction.do?method=globalLandCover [2024/01/10]	(Chen et al., 2015)

In this study, based on remote sensing data products, statistical data, survey data, and DEM data, we have developed a constrained integration method that combines multisource cropland cover products with survey data. (1) Correlation analysis between statistical data and survey data: This study obtained cropland survey data at the county-level in 1985, at the provincial-level in 1996, at the prefecture-level in 2010 and 2015, and at the county-level in 2020. It should be noted that the 2015 cropland survey data was obtained through the annual land change survey based on the second NLS, which is relatively less accurate than the cropland areas from the standard time-points data after the nationwide surveys (e.g., 2010 and 2020). According to the Ministry of Natural Resources of the PR China, the annual land change survey is based on the results of the nationwide survey and the previous year's land change survey, examining the current status and changes in the land use class, location, area, and distribution of various urban and rural lands across the country at the end of each year. For the missing years 1990, 2000, and 2005, this study referred to the correlation analysis between modern survey data and statistical data (Ye et al., 2009; Cropland Research Group, 1992). This study selected survey data and statistical data from 2010, 2015, and 2020 within the study area, respectively, and established linear regression equation between them. The results showed that the linear regression equation was y=1.2824x in 2010, and R^2=0.9689; y=1.1572x in 2015, and R^2=0.9743; y=1.1455x in 2020, and R^2=0.9714 ~~y=1.256x in 2010, and R^2=97.03%; y=1.193x in 2015, and R^2=96.23%; y=1.210x in 2020, and R^2=99.42%~~ (Fig. 3, Table S4). This indicates a high correlation between the two types of data at the three time points, and the survey data is approximately 14.6% to 28.2%~~19.3% to 25.6%~~ higher than the statistical data at the same period, with an average of about

~~22%~~20%, then corrected cropland area data by ~~22%~~20% for each county in the study area for the 1990, 2000 and 2005.



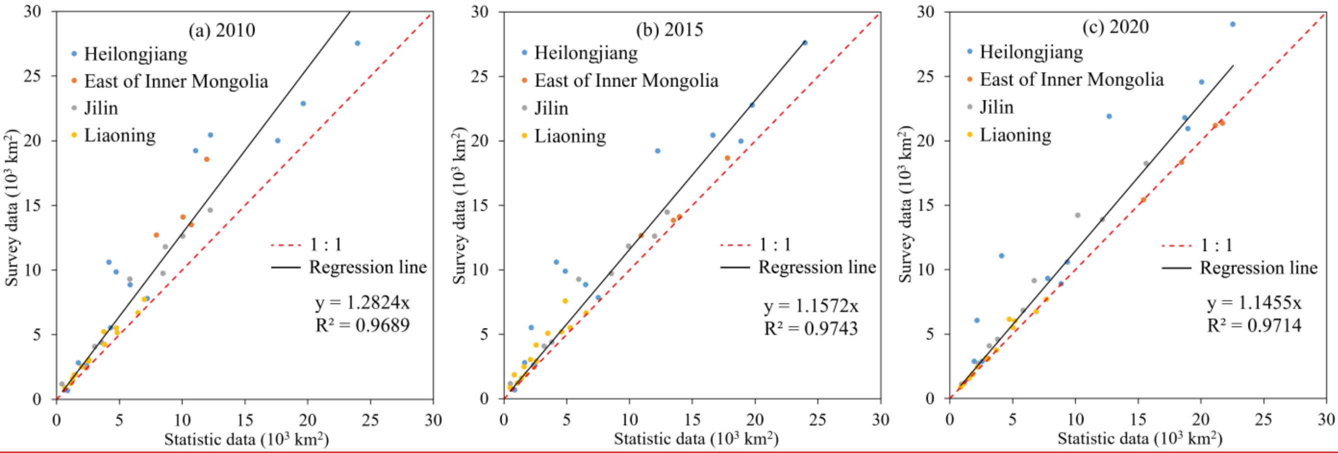


Figure 3: Correlation between the statistical cropland data and survey cropland data ~~of~~ at the prefecture-level ~~cities~~ in the Northeast China in 2010, 2015 and 2020.

(2) Establishing Dataset Priorities: After obtaining the modern land survey data levels for each province in the study area at five-year intervals from 1985 to 2020, the difference between the cropland area in dataset i and the survey data on cropland area, denoted as $D_{i,j}$, was calculated to evaluate the accuracy of the dataset, as shown in Equation (1):

$$D_{i,j}=abs\left(\frac{A_{s,j}-a_{i,j}}{A_{s,j}}\right) , \qquad (1)$$

where $A_{s,j}$ represents the survey data on cropland area in Northeast China for year j, and $a_{i,j}$ represents the cropland area in the i-th subset of the land cover product for year j. The value of $D_{i,j}$ is lower when the consistency with survey data is higher, indicating a higher priority for the input dataset. It should be noted that in this study, based on the priority and overlap of remote sensing data products used at different time points, pixels in the study area are ranked. Pixels belonging to high-priority products with high overlap will be prioritized as cropland.

(3) Allocation of cropland pixels based on DEM data: The survey data includes detailed slope classification, and the slopes were categorized into five classes: <2°, 2~6°, 6~15°, 15~25°, and >25°, and the corresponding cropland areas for each slope class were recorded. In this study, we selected NASA and METI's DEM data jointly released in 2019: ASTER Global Digital Elevation Model V003 30m. The ASTER Global Digital Elevation Model V003 can be downloaded from the NASA

EARTHDATA website (https://www.earthdata.nasa.gov/, last access: 10 January 2024). Pixels prioritized as cropland were allocated to the cropland area corresponding to each slope level in the survey data. The distribution results were controlled by provincial-level cropland area survey data at different time points, resulting in the integration of cropland data at 30m resolution for the Northeast China at 8 time points from 1985 to 2020.

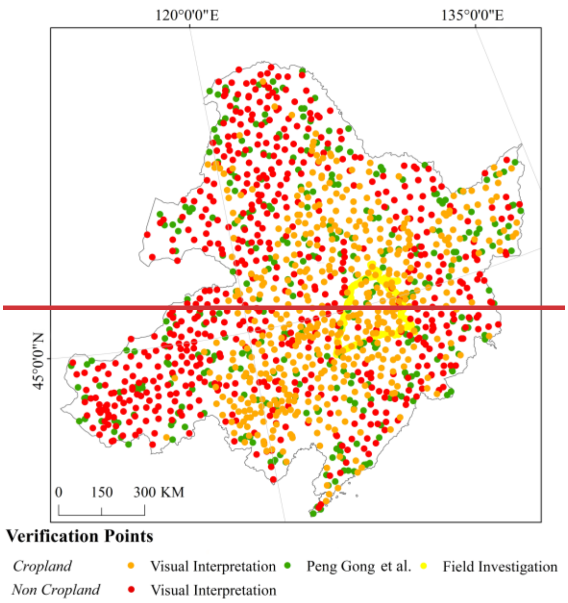
(4) Accuracy assessment and validation of RS products integration results: This study utilizes the confusion matrix was used to assess the accuracy of cropland products. The Producer Accuracy (P.A.) and User Accuracy (U.A.) for each product in 2020 are calculated as two indicators to evaluate the reliability of the spatial distribution of the cropland dataset. The calculation methods are as follows:

$$P.A = \frac{X}{Ni} \times 100\% \ , \tag{2}$$

$$U.A = \frac{X}{Nj} \times 100\% \ , \tag{3}$$

where *X* represents the number of correctly classified samples, *Ni* represents the total number of verification samples, and *Nj* represents the total number of samples in the classified result.

This study used three types of verification points for the verification of the integration result in year 2020~~s~~ (Fig. ~~4~~S1): (1) 346 cropland sample points located in the study area from FROM-GLC. (2) 1052 sample points obtained through field investigations conducted by the authors in April 2023 within the study area. (3) A total of 1200 random sample points were generated within the study area. Using high-resolution imagery from Google Earth captured in 2020, the sample points were visually interpreted and validated indoors through image comparison. The results show that the producer accuracy for cropland pixels is 94.85%, and the user accuracy is 96.49% in year 2020. For non-cropland pixels, the producer accuracy is 91.12%, and the user accuracy is 87.32%. The overall accuracy is relatively high.



~~**Figure 4: Spatial distribution of verification points.**~~

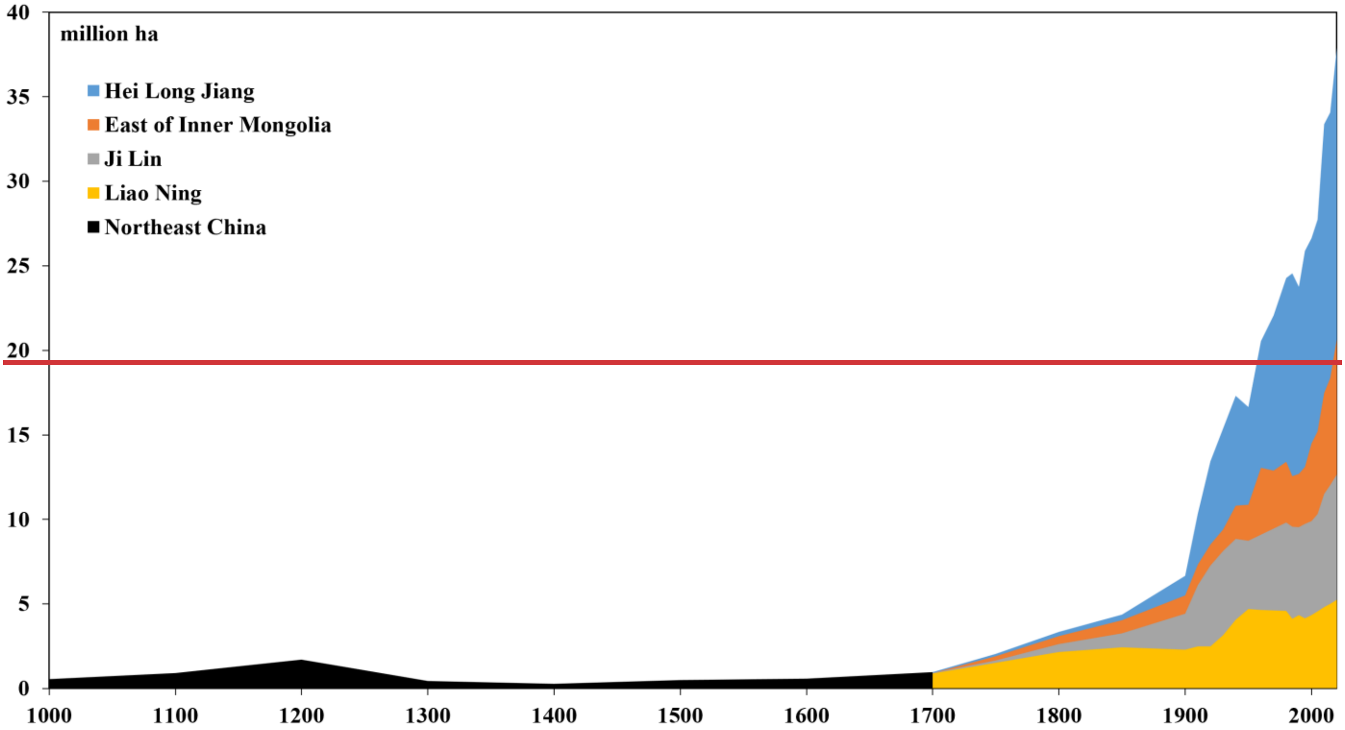
3 Results

The cropland in Northeast China exhibited phase changes of expansion-reduction-expansion over the past millennium. The cropland area in Northeast China increased from 0.55×10^4 km^2 in 1000 to 37.90×10^4 km^2 in 2020 and the average cropland fraction increased from 0.37% to 26.27% (Fig. ~~5~~4). Our results clearly show on the map the process of agricultural reclamation in Northeast China and the expansion of cropland in the Songnen and Sanjiang Plains (Fig. ~~6~~5).

3.1 Changes in the historical cropland area in Northeast China over the past millennium

The changes in cropland area in the Northeast China over the past millennium are illustrated in Figure ~~5~~4. Overall, the proportion of cropland area in the study area from 1000 to 1600 ranged from 0.74% to 4.5% of the total in 2020. During this period, from 1000 to 1200, the cropland area showed a growing trend, with an average annual growth rate of 0.56%. In 1200, it peaked at 1.69×10^4 km^2, with an overall cropland fraction of 1.17%, although the cropland fraction across the region was relatively low. From 1300 to 1600, the cropland area significantly decreased. In 1400, it reached the lowest point in the past millennium, at 0.28×10^4 km^2, with an overall cropland fraction of only 0.19%. The average annual growth rate from 1400 to 1600 was 0.37%. From 1600 to 1850, the cropland area grew slowly, with an average annual growth rate of 0.81%. During this period, the proportion of cropland area in the study area increased from 1.55% to 11.52% of the total in 2020. After 1850, the cropland area exhibited almost exponential growth. The agricultural area continued to expand northward, and this growth trend persisted until 2020, with an average annual growth rate of 1.28%.

~~At the provincial-level, from 1700 to 2020, the cropland area in Liaoning Province increased from 0.87×10^4 km^2 to $5.2^4 \times 10^4$ km^2. The cropland fraction within the region increased from 5.94% to 35.63%, with an average annual growth rate of 0.56%. However, the proportion of cropland area in the entire region showed a significant declining trend, decreasing from 91.28% to 13.81%. During the same period, in Jilin Province, the cropland area increased from 0.03×10^4 km^2 to 7.43×10^4 km^2. The cropland fraction within the region increased from 0.18% to 38.89%, with an average annual growth rate of 1.69%. The proportion of cropland in the entire region first increased, then decreased, and increased again, rising from 3.60% to 19.60%. In Heilongjiang Province, the cropland area increased from 0.04×10^4 km^2 to 17.25×10^4 km^2. The cropland fraction within the region increased from 0.09% to 38.11%, with an average annual growth rate of 1.91%. The proportion of cropland in the entire region exhibited a noticeable upward trend, increasing from 4.30% to 45.53%. In the Eastern of Inner Mongolia, the cropland area increased from 0.01×10^4 km^2 to 7.98×10^4 km^2. The cropland fraction within the region increased from 0.01% to 12.21%, with an average annual growth rate of 2.19%. The proportion of cropland in the entire region showed a fluctuating upward trend, increasing from 0.82% to 21.06% (Fig. 5).~~



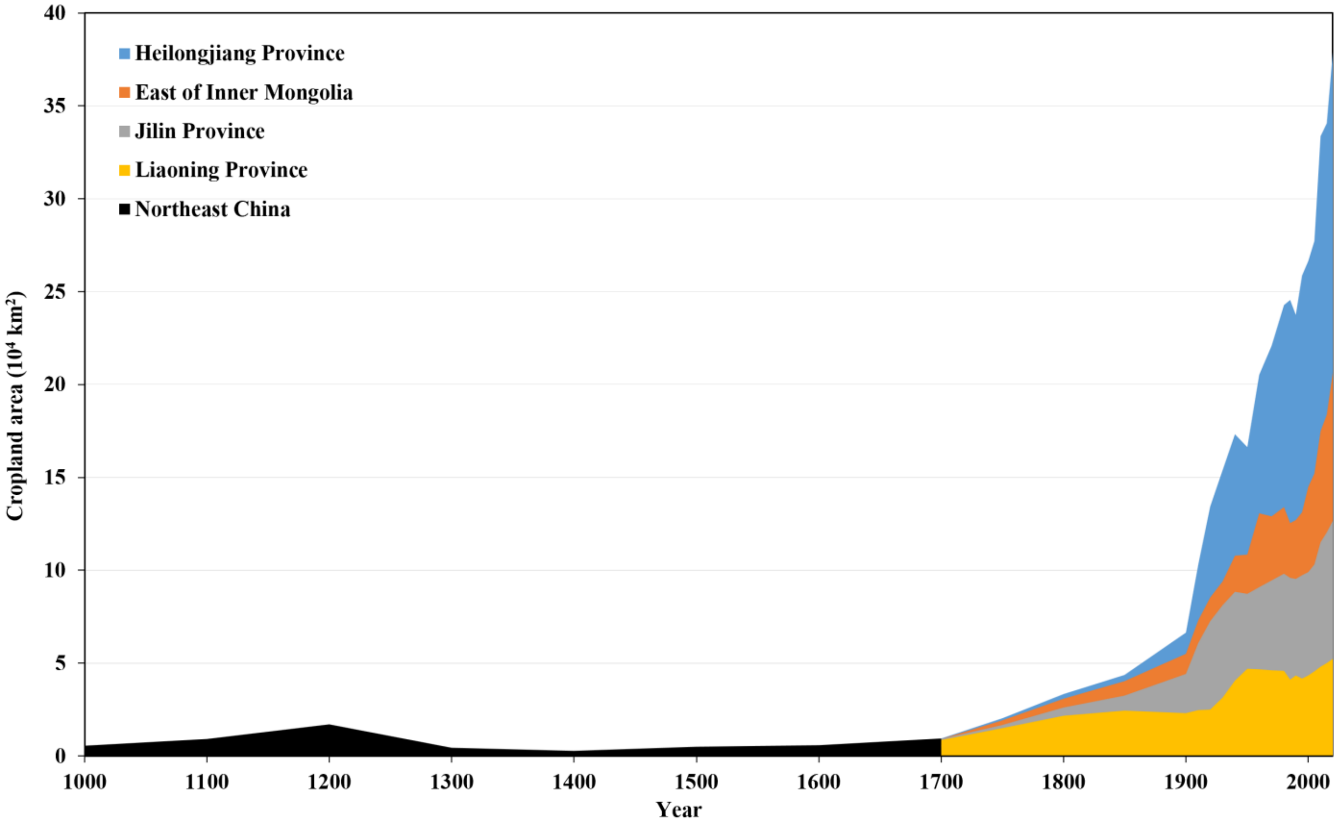


Figure ~~5~~4: Changes in total cropland area in the Northeast China from 1000 to 2020.

3.2 Spatial patterns of cropland distribution in Northeast China over the past millennium

The changes in pattern of cropland in the Northeast China over the past millennium are shown in Figure ~~6~~5. From 1000 to 1200, cropland in the study area had already reached a certain scale in spatial extent, mainly distributed in the Songliao Plain, especially in the southern part of the Liaohe Plain. The extent of cropland was roughly equivalent to the modern era. From 1300 to 1600, the main cultivation areas of cropland gradually receded southward to within the boundaries of Liaoning Province. From 1700 to 1850, cropland was mainly concentrated in the Liaoning Province. With the Qing government establishing military garrisons in the northern part of the Northeast China, farming areas was formed around these garrisons, and the farming area showed a trend of expanding northward. Due to the Qing government abandoning reclamation restrictive policies, from 1900 to 1950, the farming area gradually expanded to cover the entire region. Meanwhile, the cultivation intensity in the Hulunbuir City and Xilin Gol League of Inner Mongolia remained relatively low, influenced by war, leading to a slight decrease in the overall cropland fraction in 1950. After 1950, the farming area expanded rapidly and gradually formed a high cropland fraction agricultural zone with the Liaohe Plain, Songnen Plain, and Sanjiang Plain as its core.

~~At the provincial level, over the past millennium, the Liaohe River Basin has generally maintained a certain scale of agricultural reclamation. Particularly, agricultural reclamation activities in Liaoning Province have been continuous since 1000, with croplands mainly concentrated in the western part of Liaoning Province and sporadically distributed in the east from 1000 to 1200. From 1300 to 1700, cropland gradually concentrated in the area south of the Ming Great Wall. After 1700, the south region has consistently maintained agricultural reclamation activities, gradually forming a development trend with the Liaohe River Basin as a high cropland fraction agricultural zone. In Jilin Province, from 1000 to 1200, cropland was mainly concentrated in the Songnen Plain within its borders. From 1300 to 1600, cropland cultivation showed a declining trend. From 1700 to 1850, cropland mainly concentrated in the areas around Fuyu City in the Bodune Assistant Governorate Jurisdiction and around Jilin City and Changchun City in the Jilin Assistant Governorate Jurisdiction, gradually expanding to the surrounding areas. After 1850, with abandoning reclamation restrictive policies, Jilin Province has consistently maintained agricultural reclamation activities, gradually forming a development trend with the Songnen Plain as a high cropland fraction agricultural zone. In Heilongjiang Province, from 1000 to 1600, there were sporadic croplands, but they did not form a significant scale. From 1700 to 1850, cropland mainly concentrated in the areas around Acheng District in the Alechuka Assistant Governorate Jurisdiction, around Ning'an City in the Ninguta Assistant Governorate Jurisdiction, around Yilan County in the Sanxing Assistant Governorate Jurisdiction, around Qiqihar City in the Qiqihar Assistant Governorate Jurisdiction, and around Nenjiang City in the Moergen Assistant Governorate Jurisdiction, gradually expanding to the surrounding areas. After 1850, with abandoning reclamation restrictive policies, Heilongjiang Province has consistently maintained agricultural reclamation activities, gradually forming a development trend with the Songnen Plain and Sanjiang Plain as high cropland fraction agricultural zones. In the Eastern of Inner Mongolia, from 1000 to 1200, a small-scale agricultural reclamation area was formed in the Xilamulen River Basin. From 1300 to 1600, cropland cultivation showed a declining trend. From 1700 to 1900, cropland mainly concentrated in the northern part of the area, particularly in the Moergen Assistant Governorate Jurisdiction and the Qiqihar Assistant Governorate Jurisdiction, covering the present-day Oroqen~~

~~Autonomous Banner and Daur Autonomous Banner of Morin Dawa. In the central and southern parts, expanding westward and northward from cities like Chifeng and Tongliao, the cultivation range remained mostly within the boundaries of the farming-pastoral ecotone zone.~~

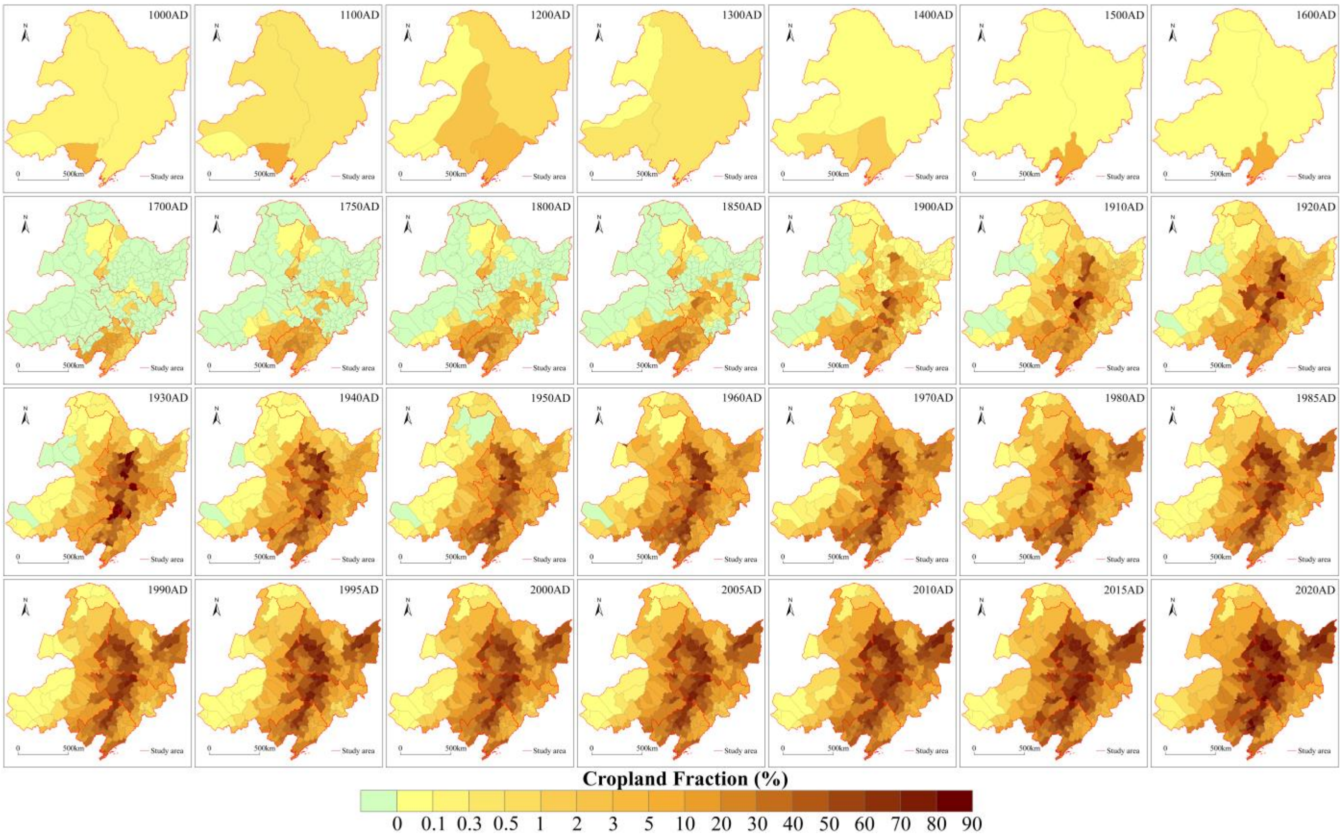


Figure ~~6~~5: Changes in spatial patterns of cropland in the Northeast China from 1000 to 2020.

4 Discussion

~~4.1 Comparison with global historical LUCC dataset~~

4.1 Credibility assessment

Based on the study of Fang et al. (2020), three methods including accuracy assessment, rationality assessment, and likelihood assessment, can be used to assess the credibility of historical LUCC dataset. Regarding the likelihood assessment, in reconstructing cropland area from 1985 to 2020, we selected eight RS products to assess the consistency. Based on the control of cropland survey data, this study identified high-consistency and high-priority pixels as cropland pixels for this dataset and evaluated and validated the accuracy of the integration results. Theoretically, compared with any single RS products used in this study during this period, the total amount of cropland area in this study is relatively more accurate and the spatial distribution is relatively more reasonable.

4.1.1 Accuracy assessment

The cropland data at lower spatial scales can be used to evaluate the accuracy of reconstructed cropland area. Due to the availability of county-level cropland survey data, we selected the county-level first general land investigation at 1985 and the county-level data from the third NLS at 2020 for comparison. As shown in Fig. 6, the determination coefficients between the cropland area from this study and the cropland area from the survey data for 1985 and 2020 are 0.9582 and 0.9892 respectively. This indicated that the overall accuracy of the reconstructed cropland area at county-level was relatively high, and our constrained integration method that combines multisource cropland cover products with survey data can well match the spatial distribution of cropland cover in Northeast China.

In addition, from 1985 to 2020, the identified high-consistency and high-priority pixels as cropland pixels based on constrained integration method may lead to errors with survey data (Table S3). The relative errors between the cropland area of this study and the cropland survey data for the period 1985 to 2020 as -1.35%, 4.02%, 5.17%, 1.10%, 0.21%, -1.93%, 0.25% and 0.67%, respectively. The vast majority of errors are around 1%, with the larger errors are 4.02% and 5.17%, which indicates that the reconstructed cropland area in this study is relatively accurate from 1985 to 2020.

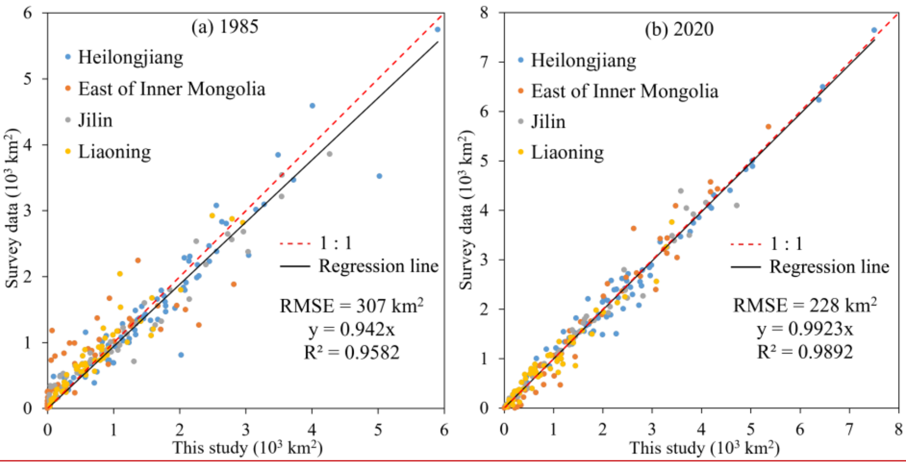


Figure 6: Correlation between the cropland data of this study and survey cropland data at county-level in the Northeast China in 1985 and 2020.

4.1.2 Rationality assessment

Due to the unavailability of actual historical land cover data, we used the actual historical agricultural development of Northeast China as a reference standard for rationality assessment. As one of the cases evaluating the distribution rationality of the HYDE3.2 cropland cover in Northeast China over the past millennium, Fang et al. (2020) analyzed changes in the northern boundary and spatial distribution of settlement relics in the Liao, Jin, Yuan, and Ming periods (916~1644), as well as changes in the cumulative number of towns and spatial distribution of towns in the three provinces of Northeast China during the Qing Dynasty (1644~1911). The unique development history of the Northeast China shown in this case is basically consistent with the process of increase or decrease and spatial distribution of the total cropland area during the same period reconstructed by this study, which reflects the rationality of this dataset.

In addition, this study attempts to briefly summarize the population changes, settlements changes (the settlement relics and the administrative division points derived from Jia et al. (2018) and the Historical Atlas of China (Tan, 1982a; Tan, 1982b)) (Fig. S2), warfare, and land policies that may have influenced land cultivation in Northeast China during the Liao, Jin, Yuan, and Ming periods (1000~1600). The population and settlements in Northeast China from 1000 to 1600 exhibited phase changes of expansion-reduction-expansion, with possible reasons including the Liao and Northern Song Dynasties signed the "*Chanyuan Alliance* (澶渊之盟)" in 1004 after war, the Jin and Southern Song Dynasties signed the "*Shaoxing Peace Treaty* (绍兴和议)" in 1141 after war, the Jin and Southern Song Dynasties signed the "*Longxing Peace Treaty* (隆兴和议)" in 1164 after war. During the three treaties and related wars, both the Liao and Jin dynasties in the north benefited significantly. They not only received reparations but also resettled large numbers of captives to the present-day Northeast China to engage in agricultural and other productive activities. Historical records also indicate that the rulers of the Liao and Jin dynasties during this period both attached much importance to agricultural production (Wu and Ge, 2022; Han, 1999; Toqto'A, 1974; Toqto'A, 1975).

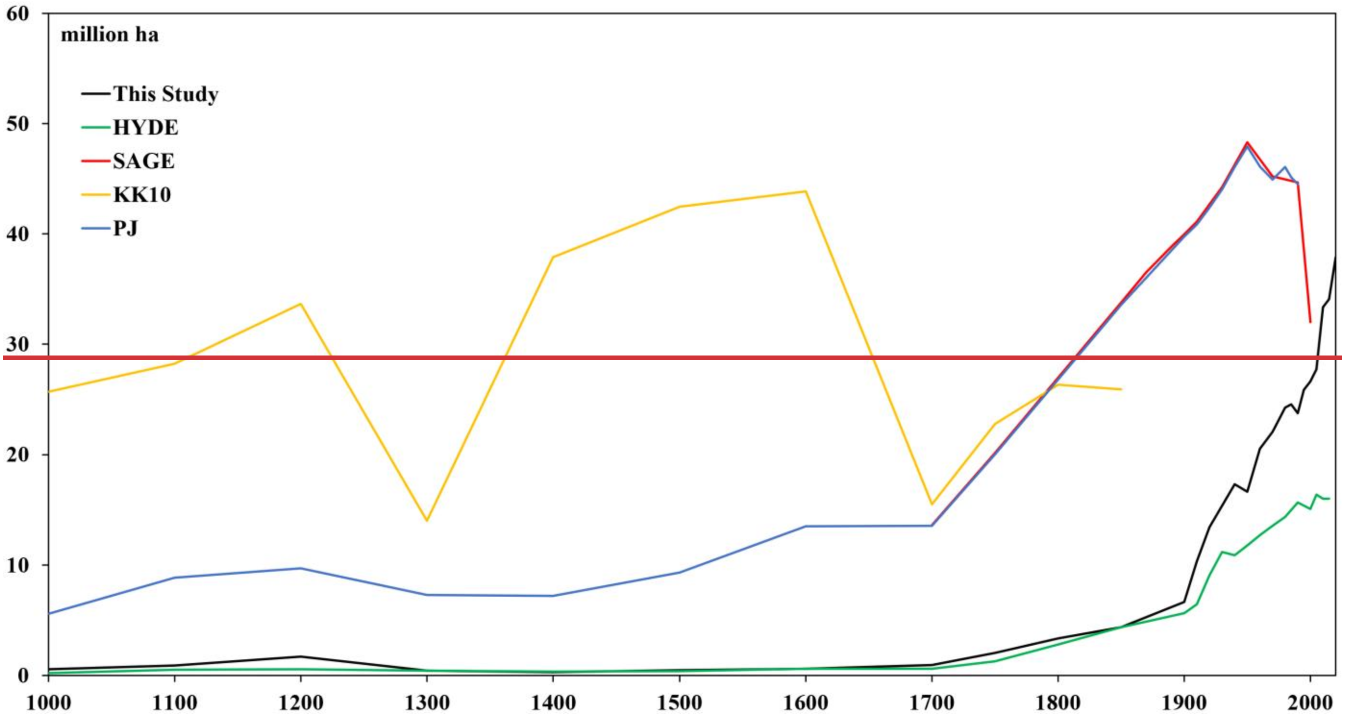
From 1211, when Genghis Khan personally led the Mongol army to attack the Jin Dynasty, until 1233, the Mongols had essentially gained control over the entire Northeast China. Using this region as a base, they also conducted war against Goryeo (present-day Korean Peninsula), which lasted until 1259. From 1259 to 1287, the Mongols made several attempts to establish governing institutions in Northeast China, but faced continuous rebellions. It wasn't until the Yuan Dynasty subdued the rebellions and established the Liaoyang Province in 1287 that effective governance began in the Northeast China. However, during this period, the region suffered from continuous warfare, significant population loss, and severe disruptions to agricultural production (Xue, 2006, 2012). According to the Dynastic History of Yuan Dynasty, from 1294 to 1345, the Yuan government provided relief to Liaoyang Province 40 times. Additionally, rebellions in the Northeast China persisted from 1343 onwards, only being effectively subdued the rebellions by 1362, just six years before the collapse of the Yuan Dynasty in 1368 (Song, 1976; Xue, 2006, 2012).

In 1368, the Ming Dynasty was established, and remnants of the Yuan Dynasty retreated to the northern grassland, known as the Northern Yuan Dynasty (Tatar), which partly within our study area. It wasn't until 1389 that the Ming Dynasty established the "*Uriyangqa three Commanderies* (兀良哈三卫)" in the region from present-day Qiqihar city to Baicheng city, gaining certain practical control over the region. However, from 1399 to 1402, the Ming Dynasty faced the internal strife of the "*Jingnan Campaign* (靖难之役)" weakening its influence over the Northeast China, allowing some ethnic minorities to further occupy territories to the south. In 1409, the Ming Dynasty established the Dusi of Nuergan, reflecting their policy of appeasement and assimilation towards ethnic minorities in the Northeast China. In 1449, the Ming Dynasty experienced the "*Tumu Crisis* (土木之变)", prompting substantial efforts to fortify defensive structures. This also greatly strengthened the defensive capabilities of the Ming Great Wall in the Northeast China and confined the major agricultural population and agricultural areas of the Northeast China within the Dusi of Eastern Liao (south of the Ming Great Wall in the Northeast China). This situation persisted until the Ming Dynasty's collapse in 1644 (Cao and Ge, 2022; Fan, 2015; Cao and Ge, 2005; Zhang, 1974). All these pieces of evidence contribute to the validation of the rationality of our dataset to a certain degree.

4.2 Comparison with global historical LUCC datasets and previous studies

To better showcase the achievements of this study, we chose to compare our results with widely used global historical LUCC datasets (Fig. 7, Table S3): the History Database of the Global Environment (HYDE3.2) (Goldewijk et al., 2017), the Sustainability and the Global Environment (SAGE) (Ramankutty et al., 2008; Ramankutty and Foley, 1999), the Kaplan and Krumhardt 2010 (KK10) (Kaplan et al., 2011), and the Pongratz Julia (PJ) (Pongratz et al., 2008). Overall, the cropland area curve of Northeast China in this study is generally between the HYDE3.2 dataset and the PJ dataset. The SAGE dataset, KK10 dataset, and PJ dataset consistently show significantly higher values than the results of this study throughout the past millennium. It's worth noting that the KK10 dataset provides the combined area of cropland and pastureland, making it notably larger than the results of this study compared to other datasets. The SAGE dataset, which obtained cropland area data using an improved method in 2000, is relatively close to the results of this study. The curve of the PJ dataset is essentially consistent with the SAGE dataset from 1700 to 1990 because the cropland data in the PJ dataset during this period are derived from the SAGE dataset.

From the trend of the curve (Fig. 7), the HYDE3.2 dataset maintains a relatively low level of cropland area from 1000 to 1700. In comparison with this study, it fails to demonstrate the historical fact of cropland cultivation in the study area from 1000 to 1200. The HYDE3.2 dataset shows an increase in cropland area after 1700, with a growth rate similar to this study. The growth rate significantly rises after 1900, but during this period, its growth rate is notably lower than in this study. The SAGE dataset maintains a relatively high total cropland area and growth rate from 1700 to 1950. Subsequently, cropland area starts to decline, approaching the results of this study in the year 2000. However, the total cropland area in the SAGE dataset from 1700 to 2000 is significantly higher than the results of this study. The KK10 dataset exhibits drastic fluctuations from 1000 to 1850, with significant declines in the periods 1200 to 1300 and 1600 to 1700, placing the two points at the trough. For the remaining periods, it maintains a growing trend, and the total area of cropland and pastureland in the KK10 dataset from 1000 to 1850 is significantly higher than the cropland area in this study. The PJ dataset shows a fluctuating upward trend from 1000 to 1700, with trends in growth and decline generally consistent with this study during this period. The minimum cropland point is also around 1400, and after 1700, the total cropland area and growth rate in the PJ dataset are consistent with the SAGE dataset. The cropland area in the PJ dataset is significantly higher than this study from 1000 to 1990.



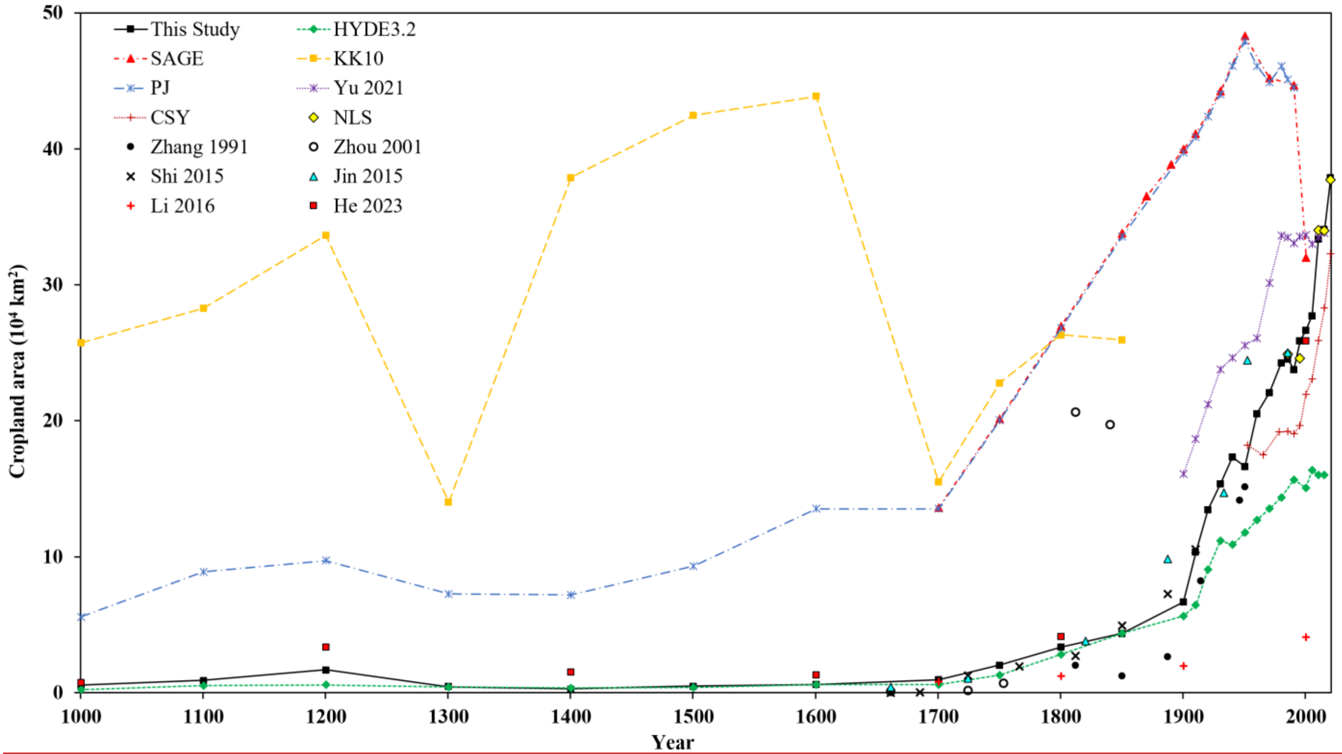


Figure 7: Comparison of total cropland area ~~in the past millennium between HYDE3.2, SAGE, KK10, PJ~~from global historical LUCC datasets, previous studies and this study in the Northeast China. The abbreviations used in the figure are as follows: HYDE3.2 refers to Goldewijk et al. (2017); SAGE refers to Ramankutty et al. (2008) and Ramankutty and Foley. (1999); KK10 refers to Kaplan et al. (2011); PJ refers to Pongratz et al. (2008); Yu 2021 refers to Yu et al. (2021); CSY denotes the Chinese Statistical Yearbook (refer to provincial and prefectural statistical yearbook); NLS denotes the National Land Survey (1985 refer to the first general land investigation; 1995 refers to the first national land survey; 2010 and 2015 refer to the second national land survey; 2020 refers to the third national land survey); Zhang 1991 refers to Zhang (1991); Zhou 2001 refers to Zhou (2001); Shi 2015 refers to Shi (2015); Jin 2015 refers to Jin et al. (2015); Li 2016 refers to Li et al. (2016); He 2023 refers to He et al. (2023).

This study was compared the total cropland area with previous representative published studies in Northeast China (Fig. 7, Table S3).

The data from Shi (2015) for 1661 and 1685 are significantly lower than this study, at these two time-points, Shi (2015) only had the data from Fengtian (roughly equivalent to Liaoning Province). Although the data from Shi (2015) for 1724 included the total area for Heilongjiang, Jilin, and Fengtian, the territorial scope of Heilongjiang and Jilin during this period was larger than that of present-day Heilongjiang and Jilin provinces. We did not exclude the cropland area according to the proportion of these territory outside present-day China. Additionally, there were 15.35, 15.35, and 17.35 million Qing *Mu* (9431 km^2, 9431 km^2, and 10660 km^2) of banner cropland at these three time-points, mainly distributed in Zhili (partly within our study area) and various parts of Northeast China, which could not be accurately divided. Therefore, we did not include the banner cropland for these three time-points. For the Mongolia in 1766, 1812, 1850, 1887, and 1911 from Shi (2015), we converted the data based on the area proportion of the Qing Dynasty Mongolia within our study area, which is 41.58%.

The data from Jin et al. (2015) closely matches this study growth trend. Both studies acknowledge that the 1985 land survey data is relatively accurate, resulting in no significant differences at this point. In Jin et al. (2015), the data in the Inner Mongolia in 1661 is missing, and for subsequent time-points, we calculated the data based on the area proportion of the East of Inner Mongolia (within our study area), which is 55.26%. The data in 1661, 1724, 1820, 1887, 1933, and 1952 from Jin et al. (2015) is similar to this study. The main reason for the difference in cropland area between the two studies may be due to the specific data source and data adjustment methods.

The data from He et al. (2023) closely matches this study growth trend. It should be noted that, for clearer comparison with our study, we selected standard time-points every 200 years from 1000 to 2000 on the cropland area curve from He et al. (2023). Similarly, we calculated the cropland area in the East of Inner Mongolia (within our study area) based on the proportion of 55.26%. In He et al. (2023), the data from 1000 to 1800 is slightly higher than this study, possibly because of the different methods for reconstruct the cropland area based on the population and the different proxy indicators used by the two studies during this period.

Similar to the comparison with He et al. (2023), when selecting the CHCD data from Li et al. (2016) for comparison, we chose standard time-points every 100 years from 1700 to 2000 on the cropland area curve for Inner Mongolia from CHCD data, and calculated the area for the East of Inner Mongolia based on 55.26%. The CHCD data for Heilongjiang, Jilin, and Liaoning province is consistent with our study (Ye et al., 2009), however, this study corrected Ye's data as explained earlier (Table S1). The difference in cropland area for the East of Inner Mongolia between the two studies may be due to the calculation of cropland area based on the proportion of 55.26%, which may not align with the actual historical agricultural development of Inner Mongolia.

For the sake of clear comparison, we selected standard time points every decade from 1900 to 1980, and every five years from 1985 to 2015 from Yu et al. (2021). The difference between the two studies in 2015 is minimal, as both studies acknowledge the NLS data is relatively accurate. The main reason for the difference in cropland area between the two studies may be due to the different reconstruction methods. In Yu et al. (2021), the officially released NLS data in 2017 (cropland area in 2016) is used as the benchmark data, and an assumption was made that this most recent data is the most reliable. Then calculated the national cropland area by using the NLS data in 2017 as the baseline and adjusted using the interannual variation information of cropland derived from various sources. Due to the cropland area difference between the national total and the sum of the provincial, a proportional adjustment was applied to match the provincial to the national total. In this study, three times NLS data were adopted, we assumed that the cropland survey data is the most reliable, then corrected the statistical data through the correlation coefficients between the statistical and the survey data in different periods to obtain the cropland area for the time-points without survey data.

The data from Zhang (1991) consistently shows lower values compared to this study across all time points. The differences may arise because the lack of data in Inner Mongolia for all periods except 1949 from Zhang (1991). Both studies agree that national statistical data is reliable for 1950s, the data from Zhang (1991) is slightly underestimates compared to our study, likely due to the calculation of cropland area based on the proportion of 55.26% in the East of Inner Mongolia.

The data from Zhou (2001) shows lower values compares to this study in 1661, 1724, and 1753. The differences may arise because the lack of data in Heilongjiang, Jilin province, and Inner Mongolia in these periods. Conversely, in 1812 and 1840, the data from Zhou (2001) is significantly exceeds to our study. The differences may arise because the assumption of the cropland. In Zhou (2001), an important assumption is that the northern territorial boundaries were much larger than today, then the cropland area of Heilongjiang, Jilin, and Liaoning province in 1952 were used instead of the cropland area in 1812 and 1840. This assumption may contradict the actual historical agricultural development of Northeast China.

4.3 Spatial distribution of cropland cover compared with HYDE3.2 dataset

We acknowledged that there is no more credible cropland area data at the global scale than HYDE up to now. Compared to this study, the HYDE3.2 dataset shows relative differences ratio (RD) in total cropland area for the period 1000 to 1600 as -82.92%, -52.52%, -100.45%, -5.32%, 17.42%, -29.34%, and 0.55%, respectively (Fig. ~~6~~8~9). The relative differences ratio (RD) as shown in Equation (4):

$$\mathrm{RD} = \frac{C_{\mathrm{H}}(y)-C_{\mathrm{T}}(y)}{(C_{\mathrm{H}}(y)+C_{\mathrm{T}}(y))/2} \times 100\%\ , \qquad (4)$$

where $C_{\mathrm{H}}(y)$ represents the total cropland area from HYDE3.2 for year y, and $C_{\mathrm{T}}(y)$ represents the total cropland area from this study for year y.

–Compared to this study, except for the years 1100 and 1300, where the absolute values of RD in most provinces within the study area did not exceed 50%, for other years, most provinces showed relatively large RD. In the years 1000 and 1100, except for certain areas in Xilin Gol League where the HYDE3.2 dataset showed more cropland area, the rest of the regions generally had less cropland area than this study. In 1200, the HYDE3.2 dataset showed more cropland area in the western region, while the opposite was observed in the eastern region. In 1300, the HYDE3.2 dataset indicated less cropland area in the entire region. From 1400 to 1600, the HYDE3.2 dataset showed more cropland area in the northern region. As the scope of the Dusi of Eastern Liao reduced, this study's cropland area in this region significantly exceeded the HYDE3.2 dataset. In 1700, both the HYDE3.2 dataset and this study indicated that most counties in Heilongjiang and Jilin provinces, as well as the northeastern part of Inner Mongolia, had no cropland (Fig. ~~6~~5, Fig. 8). However, the HYDE3.2 dataset showed that during this period, a considerable area of cropland existed in most regions of Inner Mongolia and the Sanjiang Plain, leading to 34.38% of county-level RDs being greater than 100% (Fig. 9). From 1750 to 1850, the HYDE3.2 dataset showed that the expansion of cropland cultivation gradually extended northward to cover the entire region (Fig. 8). This contradicts the areas without cropland caused by the abandoning reclamation restrictive policies of the Qing government during this period. Additionally, during this period, in the counties which both datasets considered with cropland, this study found that, except for a few counties where cropland area was less than the HYDE3.2 dataset, most counties had significantly more cropland area in this study. During this period, over half of the counties in the study area had RDs greater than 100%. From 1900 to 1950, as the abandoning reclamation restrictive policies, this study observed a decreasing trend in cropland fraction from the center to the periphery in the study area (Fig. ~~6~~5). Compared to the HYDE3.2 dataset, counties with RD greater than 100% gradually decreased (Fig. ~~8~~9). Furthermore, during this period, in most areas of the Songnen Plain and the Liaohe Plain, this study's cropland area was significantly greater than the HYDE3.2 dataset. After 1950, the RD for each county in the study area gradually decreased and concentrated in the (-100%, -10] range (Fig. ~~8~~9), indicating that the cropland area in most counties in this study was significantly greater than the HYDE3.2 dataset.

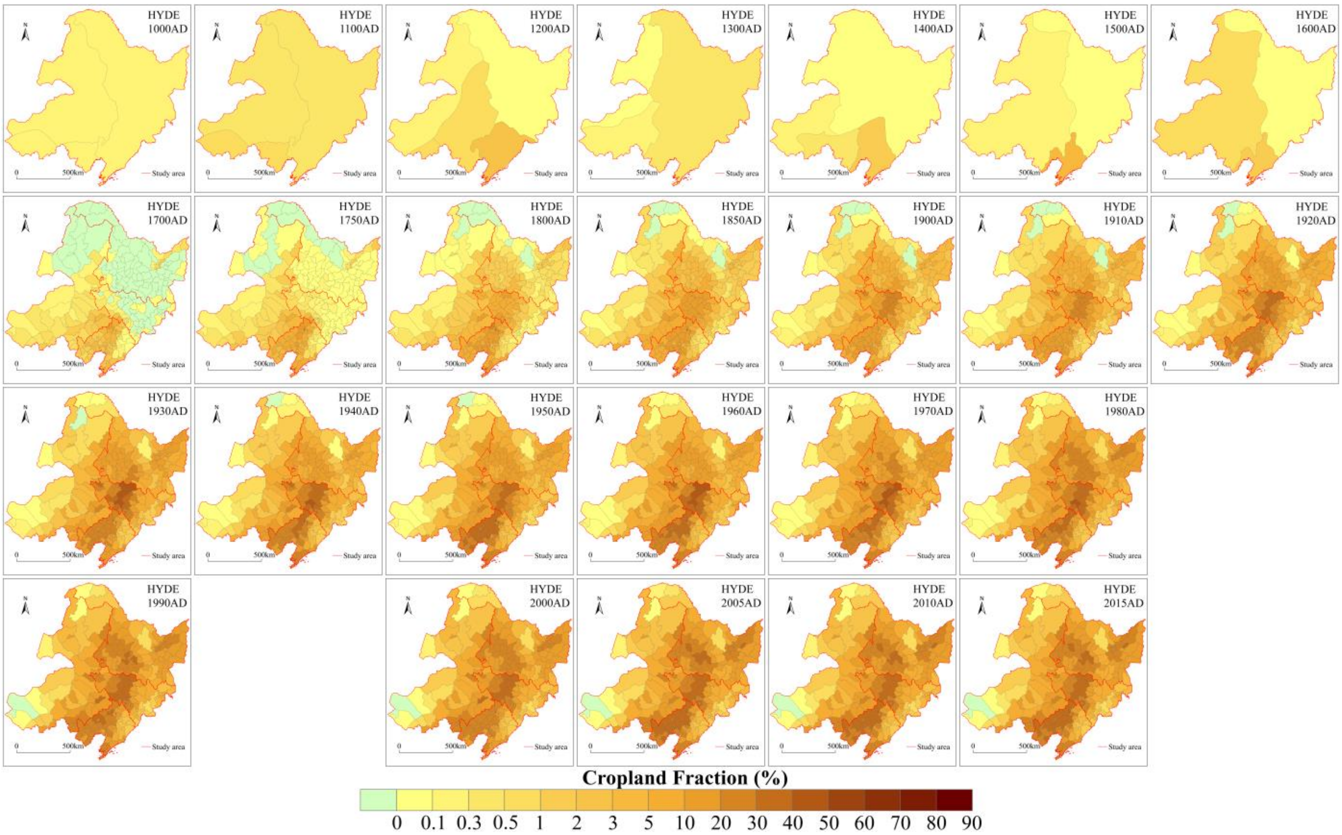


Figure 8: Changes in spatial patterns of cropland of HYDE3.2 dataset in the Northeast China from 1000 to 2015.

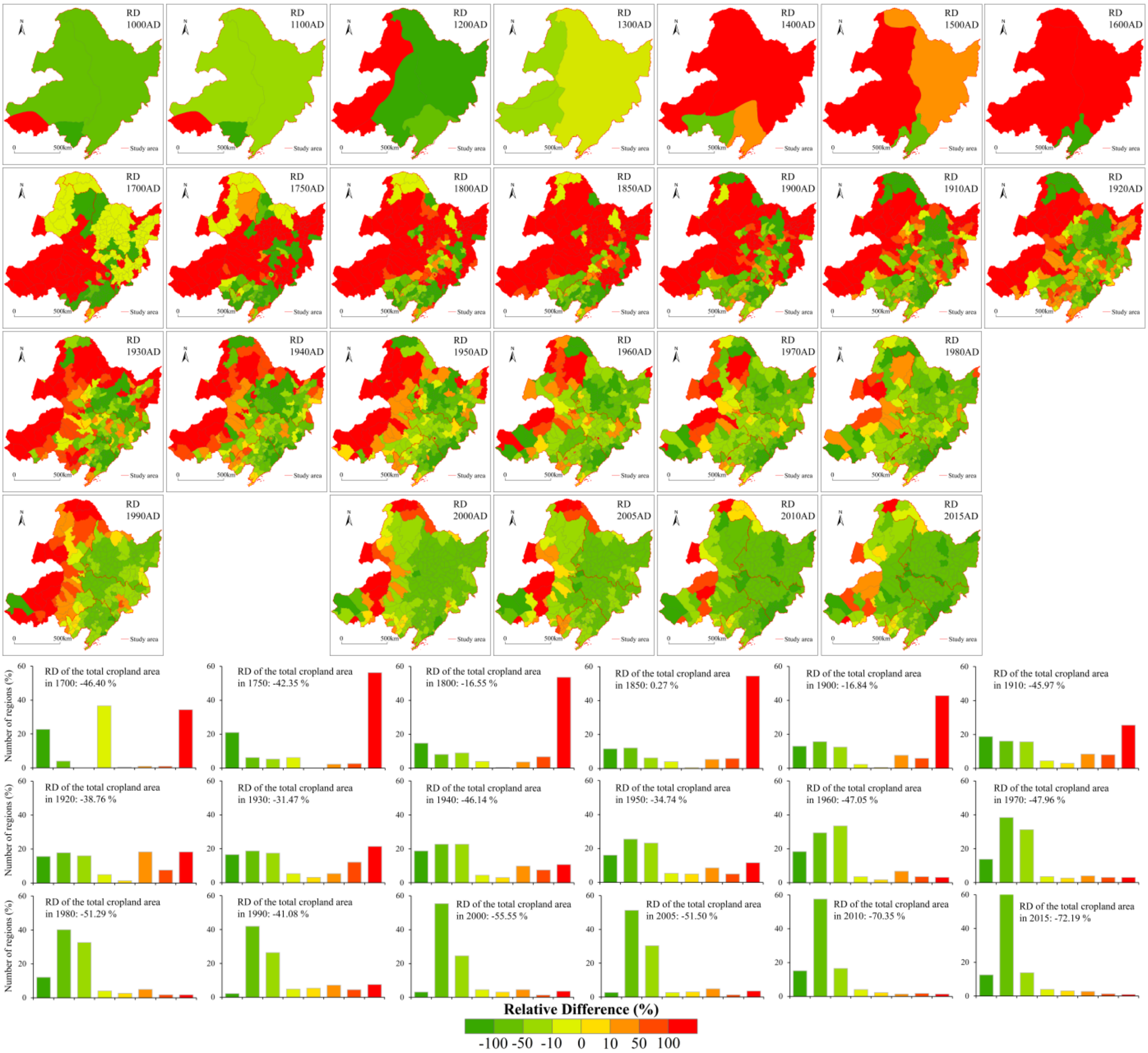


Figure 9: Comparison of the spatial distribution of cropland area between HYDE3.2 and this study in the Northeast China.

4.~~2~~ 4 Uncertainty analysis

In this study, the uncertainty mainly consisted in two aspects: the definition and selection of data, the application of methods. Regarding the data aspect:~~The past millennium cropland area results for the Northeast China reconstructed in this study, can be approximately considered as historical truth value. Comparative analysis with global historical LUCC datasets indicates that the results of this study are relatively credible and more rational. Additionally, various methods were employed during the reconstruction process to ensure the accuracy of the dataset. However, there are still some uncertainties in the reconstruction process:~~ (1) In this study, the definition of cropland before 1950 is: the sum of arable land and land under permanent crops~~. T~~, and the temporary changes in land use and fallow land during historical periods were not considered.~~,~~ The cropland area for

1950 and later are basically consistent with the identification rules in the NLS. Although the temporary changes in land use and fallow land during historical periods, this may still result in our reconstruction slightly less cropland than actual historical period.~~which may affect the accuracy of cropland area.~~

(2) Due to the completeness of historical documents, the reconstruction results of cropland for seven time points from 1000 to 1600 in this study are at the provincial-level, which may not finely reflect the spatiotemporal characteristics of cropland. Especially between 1000 and 1300, the results may lead readers to mistakenly believe that cropland were evenly distributed across the entire Northeast China. However, based on the distribution of settlement relics during this period, cropland may mainly distribute on the Liaohe Plain and on the southern part of the Songnen Plain, then reduced southward into Liaoning Province.

(3) The two proxy indicators of 14 *Mu* (0.93 hm^2, the average annual potential cropland area per Man of the agricultural population) and 2 *Mu* (0.13 hm^2, the average cropland area per household in the nonagricultural population) from 1000 to 1600 may lead to inaccuracies in cropland estimation. The reasons for using population to reconstruct cropland during this period have been detailed in the previous section, necessitating further analysis and clarification of the corresponding cropland-related indicators.

Firstly, the conclusion of 14 *Mu* per Man for agricultural population during the Liao and Jin Dynasties (1000~1200) is primarily derived from historical records in the Jin Dynasty (1200) and the relationship between population and cropland in the early Qing Dynasty (1661~1680) (Jia et al., 2023). There are two reasons why 14 *Mu* was used in the Yuan and Ming Dynasties (1300~1600): one reason is the agricultural household size and the ratios of Man in agricultural household in Northeast China during the Yuan and Ming Dynasties (1300~1600) are closer to those of the Liao and Jin Dynasties (1000~1200) (Table 1). And the per capita cropland area owned by agricultural population in the Liao-Jin-Yuan-Ming periods (1000~1600) consistently ranged between 4 and 5 *Mu* (0.27~0.33 hm^2), slightly higher than the subsistence level of 3 *Mu* per capita in previous studies for the same historical period in this region (Ye et al., 2009; Fang et al., 2006; Shi, 1990), which is relatively reasonable. The second reason is that there were no significant changes in agricultural production technology in Northeast China during the Liao-Jin-Yuan-Ming periods (1000~1600), and the population declined significantly compared with the Liao and Jin Dynasties (1000~1200) due to factors such as warfare. However, considering the social stability at standard time-points during the Yuan and Ming Dynasties (1300~1600), the strong willingness of the agricultural population towards cultivation, and the limitations of individual cultivation capabilities, the cropland from the Liao and Jin Dynasties could be relatively easily inherited and reclaimed by descendants.

Secondly, Similar to the agricultural population, considering the non-agricultural household size, stable agricultural production technology, the historical inheritance of most ethnic groups, this study continues to use 2 *Mu* as the calculation indicator of non-agricultural population in the Yuan and Ming Dynasties (1300~1600) (Cong, 1993a; Cong, 1993b; Wu and

Ge, 2005a; Cao and Ge, 2005b; Liu et al., 2016).

Regarding the method aspect: (1) ~~(3)~~ From 1700 to 1980, cropland areas at multiple time points in this study were derived through linear interpolation and polynomial curve fitting. Although we have fully considered historical facts and other research conclusions (Fang et al., 2020; Ye et al., 2009; Fang et al., 2005) when selecting the interpolation time points, 1860 was chosen as the dividing point between slow growth and rapid growth. This method, compared to data recorded at each specific historical point, may affect the accuracy of the value at those standard time points.

(42) From 1700 to 1980, the county-level administrative boundaries in the published data used in this study differ from the modern county-level administrative boundaries used in this study. Especially in the CNEC data (Ye et al., 2009) in 1683, 1735 and 1780, there is county-level in Liaoning province, Assistant Governorate Jurisdiction (prefecture-level) in Heilongjiang and Jilin province. This would result in counties belonging to different Assistant Governorate Jurisdictions in present-day having the same cropland fraction. This problem is difficult to correct further because the lowest administrative level in Northeast China available in historical data during this period is Assistant Governorate Jurisdiction (prefecture-level).~~The cropland area calculated based on the proportion of overlapping areas between the two may cause minor errors.~~

(3) From 1985 to 2020, the land survey data utilized in this dataset might exhibit uncertainties in early cropland data due to backward technology and other factors. Additionally, the use of a uniform correlation coefficient to correct the cropland statistics data for the entire Northeast China may affect the accuracy of the cropland area in localized areas, this may lead to a lower cropland area at previous time-points. To mitigate the impact of these uncertainties on our dataset during this period, this study mainly adopts two methods: Firstly, this study mainly selects the standard time-points data after the nationwide surveys, avoiding the use of annual land change survey data. For instance, the cropland area in 1985 in this dataset is based on the first general land investigation around 1985; the cropland area in 1995 is based on the first NLS's standard time-point data on October 31, 1996; the cropland area in 2010 is based on the second NLS's standard time-point data on December 31, 2009; and the cropland area in 2020 is based on the third NLS's standard time-point data on December 31, 2019. Secondly, this study uses correlation coefficients to correct the statistical data by category and time point. For instance, the average correlation coefficient of the second and third NLSs with corresponding statistical data is used to correct the statistical data for the 1990, 2000, and 2005; the correlation coefficient of the 1985 first general land investigation with corresponding statistical data is used to correct the statistical data for the 1950~1980.

5 Data availability

All cropland data reconstructed in this study are publicly available at https://doi.org/10.6084/m9.figshare.25450468.v2 (Jia, 2024).

6 Conclusion

Based on historical documents, proxy data such as population data, revised published results, remote sensing data products, statistical data, and survey data, and utilizing a series of data processing methods, as well as accuracy and rationality assessment methods, we established a 28 time-points cropland area dataset in Northeast China at provincial-level and county-level spatial resolutions from 1000 to 2020. Reconstruction results indicate that cropland area in Northeast China grew slowly before 1850 and experienced rapid expansion after 1850, maintaining this growth trend until 2020. ~~At the provincial-level, over the past millennium, Liaoning Province has generally maintained a state of moderate to high-intensity cultivation, with other provinces experiencing a gradual increase in cultivation intensity after 1850. From 1700 to 2020, in terms of the increase in cropland area, Heilongjiang Province leads with a total increase of 17.21×10^4 km^2, followed by the Eastern of Inner Mongolia, Jilin Province is third, and Liaoning Province has the smallest increase at 4.37×10^4 km^2. In terms of the cropland fraction, in 1700, Liaoning Province had the highest at 5.94%, followed by Jilin Province, Heilongjiang Province, and the Eastern of Inner Mongolia with the lowest at 0.01%. In 2020, Jilin Province recorded the highest at 38.89%, followed by Heilongjiang Province, Liaoning Province, and the Eastern of Inner Mongolia with the lowest at 12.21%. In terms of the average annual growth rate of cropland area, the Eastern of Inner Mongolia exhibited the highest at 2.19%, followed by Heilongjiang Province, Jilin Province, and Liaoning Province with the lowest at 0.56%.~~

This dataset illustrates the characteristics of cropland changes in Northeast China over the past millennium, especially in the past 300 years. Between 1000 and 1200, the extent of cropland was roughly equivalent to the modern era. Subsequently, until 1850, the cropland was mainly concentrated in the Liaoning Province. However, with the Qing government establishing military garrisons in the northern part of the Northeast China, farming areas was formed around these garrisons from 1700 to 1850. With the implementation of the immigration and cultivation policy in the latter half of the 19th century, the spatial pattern of cropland coverage in Northeast China changed significantly after 1850, with agricultural zones rapidly expanding across the entire region. After 1950, the expansion of high cropland fraction agricultural zones in Northeast China became more pronounced, gradually forming core areas with high cropland fraction in the Liaohe Plain, Songnen Plain, and Sanjiang Plain. ~~At the provincial-level, from 1700 to 2020, the proportion of cropland area in each province to the entire region underwent significant changes. Liaoning Province decreased from 91.28% to 13.81%, while Jilin Province increased from 3.60% to 19.60%. Heilongjiang Province increased from 4.30% to 45.53%, and the Eastern of Inner Mongolia increased from 0.82% to 21.06%. This indicates a trend in the Northeast China of cropland concentration towards higher latitudes.~~

~~Compared to global historical LUCC datasets such as HYDE3.2, the SAGE dataset, KK10 dataset, and PJ dataset all~~

show significantly larger cropland areas over the past millennium than our dataset. The HYDE3.2 dataset shows a certain degree of consistency in the changing trends over the past millennium compared to our dataset. However, HYDE3.2 fails to reflect the historical reality of cropland cultivation in the study area from 1000 to 1200. Moreover, HYDE3.2 systematically underestimates cropland areas in the study area after 1900. The growth rate of cropland areas during this period is significantly lower than our dataset, and spatially, it fails to depict the formation process of high cropland fraction agricultural zones with the Liaohe Plain, Songnen Plain, and Sanjiang Plain as its core.

Despite the fact that the cropland area change dataset in this study is presented at the provincial-level and county-level, the dataset we reconstructed based on historical records at 28 time points can be approximated as “truth value”. This dataset provides crucial support for the long-term land use changes in the Northeast China. In the future, we will further investigate gridded cropland allocation methods based on the historical cultivation process in the Northeast China, aiming to better serve research such as carbon emission, climate data construction, climate-ecosystem modeling and the conservation and utilization of black soil, etc.

Appendix A: Data records of CNEC

Table A1. Revisions to the CNEC dataset

Counties where cropland is missing or inconsistent	Periods when cropland is missing or inconsistent	Adjacent counties for interpolation or based on the number of settlements
Dashiqiao City	1683	Haicheng City, Yingkou City, Gaizhou City, Panjin City
Tieling County	1683, 1735, 1780	Faku County, Tieling City, Diaobingshan City
Suizhong County	1683	Huludao City, Jianchang County
Xingcheng City		Huludao City, Jianchang County
Liaoyuan City	1683, 1735, 1780, 1908	Dongliao County
Xifeng County	1735, 1780	Changtu County, Siping City, Dongfeng County, Dongliao County
Jiaohe City	1735, 1780, 1914	Interpolated based on Shulan City, Yongji County, and Jilin City in 1735 and 1780; Linear interpolation in 1914 using data from 1908 and 1931
Meihekou City	1735, 1780	Interpolated based on multiplying the number of settlements and the cropland area owned by unit settlement in this region 1735 and 1780.
Fuxin Mongolian Autonomous County	1780, 1914	Interpolated based on Fuxin City and Beipiao City in 1780; Linear interpolation in 1914 using data from 1908 and 1931
Huadian City		Interpolated based on Dunhua City, Jingyu County, and Fusong County in 1780; Linear interpolation in 1914 using data from 1908 and 1931
Shulan City	1780	Jilin City, Changchun City, Jiaohe City, Wuchang City
Panshi County		Huadian City, Dongliao County, Dongfeng County, Meihekou City, Huinan County
Yushu City		Interpolated based on multiplying the number of

Manchu Autonomous County of Yitong		settlements and the cropland area owned by unit settlement in this region in 1780
Songyuan City		
Changling County		
Fuyu City		
Siping City	1780, 1914	Interpolated based on multiplying the number of settlements and the cropland area owned by unit settlement in this region in 1780; Linear interpolation in 1914 using data from 1908 and 1931
Lishu County		
Gongzhuling City		
Shuangliao City		
Jianchang County	1908, 1914	Interpolated based on Lingyuan City, Kalaqinzuoyi Mongolian Autonomous County, Suizhong County and Xingcheng County in 1908; Linear interpolation in 1914 using data from 1908 and 1931
Zhenlai County	1908	Baicheng City, Da'an City, Ulan Hot City
Tailai County		Jalaid Banner, Qiqihar City, Longjiang County
Dorbod Mongolian Autonomous County		Qiqihar City, Daqing City, Lindian County
Fuxin City	1914	Linear interpolation based on cropland fraction in 1908 and 1931
Chaoyang City		
Chaoyang County		
Jianping County		
Kalaqinzuoyi Mongolian Autonomous County		
Beipiao City		
Lingyuan City		
Baishan City		
Fusong County		
Jingyu County		
Korean Autonomous County of Changbai		
Qian Gorlos Mongolian Autonomous County		
Qian'an County		
Antu County		
Mohe City		
Qiqihar City	1940, 1950	Revised according to Qiqihar Agricultural Annals

Author contributions. RJ, XF and Yu Y designed this work. RJ wrote the manuscript. XF and Yu Y provided suggestions on methodology. Yu Y and Yundi Y developed the dataset. All the authors contributed to the review of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Financial support. This research has been supported by the National Key R&D Program of China (grant no.

2021YFD1500704), and the Fund of the China Scholarship Council (grant no. 202306040062).

References

Arneth, A., Sitch, S., Pongratz, J., Stocker, B. D., Ciais, P., Poulter, B., Bayer, A. D., Bondeau, A., Calle, L., Chini, L. P., Gasser, T., Fader, M., Friedlingstein, P., Kato, E., Li, W., Lindeskog, M., Nabel, J. E. M. S., Pugh, T. A. M., Robertson, E., Viovy, N., Yue, C., and Zaehle, S.: Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed, Nat. Geosci., 10, 79-84, https://doi.org/10.1038/NGEO2882, 2017.

Bai, S., Shuwen, Z., and Yangzhen, Z.: Digital rebuilding of lucc spatial-temporal distribution of the last 100 years: taking dorbod mongolian autonomous county in daqing city as an example, Acta Geographica Sinica, 62, 427-436, 2007.

Cao, B., Yu, L., Li, X., Chen, M., Li, X., Hao, P., and Gong, P.: A 1 km global cropland dataset from 10 000 bce to 2100 ce, Earth Syst. Sci. Data, 13, 5403-5421, https://doi.org/10.5194/essd-13-5403-2021, 2021.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., and Mills, J.: Global land cover mapping at 30 m resolution: a pok-based operational approach, Isprs-J. Photogramm. Remote Sens., 103, 7-27, https://doi.org/10.1016/j.isprsjprs.2014.09.002, 2015.

Dickinson, R. E.: Global change and terrestrial hydrology - a review, Tellus Ser. B-Chem. Phys. Meteorol., 43A-B, 176-181, https://doi.org/10.1034/j.1600-0870.1991.00015.x, 1991.

Ellis, E. C., Gauthier, N., Goldewijk, K. K., Bird, R. B., Boivin, N., Díaz, S., Fuller, D. Q., Gill, J. L., Kaplan, J. O., Kingston, N., Locke, H., Mcmichael, C. N. H., Ranco, D., Rick, T. C., Rebecca Shaw, M., Stephens, L., Svenning, J. C., and Watson, J. E. M.: People have shaped most of terrestrial nature for at least 12,000 years, Proceedings of the National Academy of Sciences - PNAS, 118, 1, https://doi.org/10.1073/pnas.2023483118, 2021.

Fang, X., Zhao, W., Zhang, C., Zhang, D., Wei, X., Qiu, W., and Ye, Y.: Methodology for credibility assessment of historical global lucc datasets, Science China Earth Sciences, 1013-1025, https://doi.org/10.1007/s11430-019-9555-3, 2020.

Foley, J. A., Defries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., and Snyder, P. K.: Global consequences of land use, Science, 309, 570-574, https://doi.org/10.1126/science.1111772, 2005.

Friedlingstein, P., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T., Chini, L. P., Dou, X., Enyo, K., Evans, W., Falk, S., Feng, L., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Hurtt, G. C., Iida, Y., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Ma, L., Marland, G., Mayot, N., Mcguire, P. C., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S., Niwa, Y., O'Brien, K. M., Olsen, A., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Rehder, G., Robertson, E., Rosan, T. M., Schwinger, J., Séférian, R., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Sun, Q., Sutton, A. J., Sweeney, C., Tans, P. P., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global carbon budget 2023, Earth Syst. Sci. Data, 15, 5301-5369, https://doi.org/10.5194/essd-15-5301-2023, 2023.

Gaillard, M., Morrison, K. D., Madella, M., and Whitehouse, N.: Past land-use and land-cover change: the challenge of quantification at the subcontinental to global scales, https://doi.org/10.22498/pages.26.1.3, 2018.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C.: Food security: the challenge of feeding 9 billion people, Science, 327, 812-818, https://doi.org/10.1126/science.1185383, 2010.

Goldewijk, K. K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the holocene - hyde 3.2, Earth Syst. Sci. Data, 9, 927-953, https://doi.org/10.5194/essd-9-927-2017, 2017.

Gong, P., Wang, J., Le Yu, Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., and Chen, J.: Finer resolution observation and monitoring of global land cover : first mapping results with landsat tm and etm+ data, Int. J. Remote Sens., 34, 2607-2654, https://doi.org/10.1080/01431161.2012.748992, 2013.

Gortan, M., Testa, L., Fagiolo, G., and Lamperti, F.: A unified dataset for pre-processed climate indicators weighted by gridded economic activity, Sci. Data, 11, 533, https://doi.org/10.1038/s41597-024-03304-1, 2024.

He, F., Yang, F., Zhao, C., Li, S., and Li, M.: Spatially explicit reconstruction of cropland cover for china over the past millennium., Science China Earth Sciences, 66, 111-128, https://doi.org/10.1007/s11430-021-9988-5, ~~2022~~2023.

He, F., Li, M., and Li, S.: Reconstruction of Lu-level cropland areas in the Northern Song Dynasty (AD976-1078), Acta Geographica Sinica, 606-618, https://doi.org/10.1007/s11442-017-1395-3, 2017.

Huang, X., Ibrahim, M. M., Luo, Y., Jiang, L., Chen, J., and Hou, E.: Land Use Change Alters Soil Organic Carbon: Constrained Global Patterns and Predictors, Earth's Future, 12, n/a, https://doi.org/10.1029/2023EF004254, 2024.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Goldewijk, K. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., Vuuren, D. P. V., Zhang, X., and Pacific Northwest National Lab. Pnnl, R. W. U. S.: Harmonization of global land use change and management for the period 850–2100 (luh2) for cmip6, Geosci. Model Dev., 13, 5425-5464, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

Ito, A. and Hajima, T.: Biogeophysical and biogeochemical impacts of land-use change simulated by miroc-es2l, Prog. Earth Planet. Sci., 7, 1-15, https://doi.org/10.1186/s40645-020-00372-w, 2020.

Jia, D., Li, Y., and Fang, X.: Complexity of factors influencing the spatiotemporal distribution of archaeological settlements in northeast China over the past millennium, Quat. Res., 413-424, 2018.

Jia, R., Fang, X., and Ye, Y.: Gridded reconstruction of cropland cover changes in northeast china from ad 1000 to 1200, Reg. Envir. Chang., 23, 128, https://doi.org/10.1007/s10113-023-02118-y, 2023.

Jia, R., Fang, X., Yang, Y., Yokozawa, M., and Ye, Yu.: A 28 time-points cropland area change dataset in Northeast China from 1000 to 2020. figshare. Dataset. https://doi.org/10.6084/m9.figshare.25450468.v2, 2024.

Jin, X., Cao, X., Du, X., Yang, X., Bai, Q., and Zhou, Y.: Farmland dataset reconstruction and farmland change analysis in China during 1661-1985, Acta Geographica Sinica, 1058-1074, 2015.

Kalnay, E. and Cai, M.: Impact of urbanization and land-use change on climate, Nature, 423, 528-531, https://doi.org/10.1038/nature01675, 2003.

Kaplan, J. O., Krumhardt, K. M., Ellis, E. C., Ruddiman, W. F., Lemmen, C., and Goldewijk, K. K.: Holocene carbon emissions as a result of anthropogenic land cover change, The Holocene, 21, 775-791, https://doi.org/10.1177/0959683610386983, 2011.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., and Brumby, S. P.: Global land use/land cover with sentinel 2 and deep learning, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, The Institute of Electrical and Electronics Engineers, Inc. (IEEE), Piscataway, 2021-1-1, 4704-4707, 2021.

Li, M., He, F., Li, S., and Yang, F.: Reconstruction of the cropland cover changes in eastern china between the 10th century and 13th century using historical documents, Sci Rep, 13552, https://doi.org/10.1038/s41598-018-31807-6, 2018.

Li, S., He, F., and Zhang, X.: Spatially explicit reconstruction of cropland cover in china from 1661 to 1996, Reg. Envir. Chang., 16, 417-428, https://doi.org/10.1007/s10113-014-0751-4, 2016.

Liu, J., Liu, B., Liu, H., and Zhang, F.: Long‐term cultivation drives soil carbon, nitrogen, and bacterial community changes in the black soil region of northeastern China, Land Degrad. Dev., 35, 428-441, https://doi.org/10.1002/ldr.4925, 2024.

Pongratz, J., Reick, C., Raddatz, T., and Claussen, M.: A reconstruction of global agricultural areas and land cover for the last millennium, Glob. Biogeochem. Cycle, 22, GB3018, https://doi.org/10.1029/2007GB003153, 2008.

Poschlod, P., Bakker, J. P., and Kahmen, S.: Changing land use and its impact on biodiversity, Basic Appl. Ecol., 6, 93-98, https://doi.org/10.1016/j.baae.2004.12.001, 2005.

Potapov, P., Turubanova, S., Hansen, M. C., Tyukavina, A., Zalles, V., Khan, A., Song, X. P., Pickens, A., Shen, Q., and Cortez, J.: Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century, Nat. Food, 3, 19-28, https://doi.org/10.1038/s43016-021-00429-z, 2022.

Ramankutty, N. and Foley, J. A.: Estimating historical changes in global land cover: croplands from 1700 to 1992, Glob. Biogeochem. Cycle, 13, 997-1027, https://doi.org/10.1029/1999GB900046, 1999.

Ramankutty, N., Evan, A. T., Monfreda, C., and Foley, J. A.: Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, Glob. Biogeochem. Cycle, 22, GB1003, https://doi.org/10.1029/2007GB002952, 2008.

Roberts, N.: How humans changed the face of Earth, Science (American Association for the Advancement of Science), 365, 865-866, https://doi.org/10.1126/science.aay4627, 2019.

Saez-Sandino, T., Maestre, F. T., Berdugo, M., Gallardo, A., Plaza, C., Garcia-Palacios, P., Guirado, E., Zhou, G., Mueller, C. W., Tedersoo, L., Crowther, T. W., and Delgado-Baquerizo, M.: Increasing numbers of global change stressors reduce soil carbon worldwide, Nat. Clim. Chang., https://doi.org/10.1038/s41558-024-02019-w, 2024.

Sebastiaan, L., Mathilde, J., Paul, C. S., Stephan, E., Julia, P., Eric, C., Galina, C., Axel, D., Karlheinz, E., Morgan, F., Bert, G., Thomas, G., Richard, A. H., Katja, K., Alexander, K., Thomas, K., Tobias, K., Tuomas, L., Annalea, L., Denis, L., Matthew, J. M., Patrick, M., Eddy, J. M., Kim, N., Kim, N., Juliane, O., Kim, P., Casimiro, A. P., Serge, R., Corinna, R., James, R., Andrew, E. S., Andrej, V., Martin, W., and Dolman, A. J.: Land management and land-cover change have impacts of similar magnitude on surface temperature, Nat. Clim. Chang., 4, 389-393, https://doi.org/10.1038/nclimate2196, 2014.

Shukla, J., Nobre, C., and Sellers, P.: Amazon deforestation and climate change, Science, 247, 1322-1325, https://doi.org/10.1126/science.247.4948.1322, 1990.

Shukla, P. R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H. O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., Van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., Huntley, E., Kissick, K., Belkacemi, M., and Malley, J.: IPCC, 2019: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, Intergovernmental Panel on Climate Change (IPCC), 2019.

Thenkabail, P. S. T. P.: Global cropland-extent product at 30-m resolution (gcep30) derived from landsat satellite time-series data for the year 2015 using multiple machine-learning algorithms on google earth engine cloud, U.S. Geological Survey Professional Paper, 1868, 63p, https://doi.org/https://doi.org/10.3133/pp1868, 2021.

Thenkabail, P. S., Teluguntla, P. G., Xiong, J., Oliphant, A., Congalton, R. G., Ozdogan, M., Gumma, M. K., Tilton, J. C., Giri, C., Milesi, C., Phalke, A., Massey, R., Yadav, K., Sankey, T., Zhong, Y., Aneece, I., and Foley, D.: Global cropland-extent product at 30-m resolution (gcep30) derived from landsat satellite time-series data for the year 2015 using multiple machine-learning algorithms on google earth engine cloud, U.S. Geological Survey Professional Paper, 2021, 1-63, https://doi.org/10.3133/pp1868, 2021.

Wang, S., Fang, C., Chen, X., Liang, J., Liu, K., Feng, K., Hubacek, K., and Wang, J.: China's ecological footprint via biomass import and consumption is increasing, Commun. Earth Environ., 5, 212-244, https://doi.org/10.1038/s43247-024-01399-3, 2024.

Wang, Y., Yu, H., Wang, S., Li, H., and Wang, Y.: Unveiling trends and environmental impacts of global black soil crop production: a comprehensive assessment, Resources, Conservation and Recycling, 208, 107717, https://doi.org/10.1016/j.resconrec.2024.107717, 2024.

Wei, X., Ye, Y., Zhang, Q., and Fang, X.: Reconstruction of cropland change over the past 300 years in the Jing-Jin-Ji area, China, Reg. Environ. Change, 2097-2109, 2016.

Wei, X., Widgren, M., Li, B., Ye, Y., Fang, X., Zhang, C., and Chen, T.: Dataset of 1 km cropland cover from 1690 to 1999 in Scandinavia., Earth Syst. Sci. Data, 3035-3056, 2021.

Winkler, K., Fuchs, R., Rounsevell, M., and Herold, M.: Global land use changes are four times greater than previously estimated, Nat. Commun., 12, 2501, https://doi.org/10.1038/s41467-021-22702-2, 2021.

Wu, Z., Fang, X., Jia, D., and Zhao, W.: Reconstruction of cropland cover using historical literature and settlement relics in farming areas of shangjing dao during the liao dynasty, china, around 1100 ad, The Holocene, 30, 1516-1527, https://doi.org/10.1177/0959683620941293, 2020.

Wu, Z., Fang, X., and Ye, Y.: A settlement density based allocation method for historical cropland cover: a case study of jilin province, china, Land, 11, 1374, https://doi.org/10.3390/land11081374, 2022.

Xu, M., Zhang, Z., Yue, C., Zhao, J., Zhang, P., Wang, M., Wang, J., Zhao, H., Liu, J., Tang, X., and He, J.: Contributions of China's terrestrial ecosystem carbon uptakes to offsetting CO2 emissions under different scenarios over 2001–2060, Glob. Planet. Change, 238, 104485, https://doi.org/10.1016/j.gloplacha.2024.104485, 2024.

Xuan, X., Zhang, F., Deng, X., and Bai, Y.: Measurement and spatio-temporal transfer of greenhouse gas emissions from agricultural sources in China: a food trade perspective, Resources, Conservation and Recycling, 197, 107100, https://doi.org/10.1016/j.resconrec.2023.107100, 2023.

Yang, F., Dong, G., Wu, P., and He, F.: Stylized facts of past 1000-year of China's cropland changes, Land Use Policy, 144, 107258, https://doi.org/10.1016/j.landusepol.2024.107258, 2024.

Yang, J. and Huang, X.: The 30 m annual land cover dataset and its dynamics in china from 1990 to 2019, Earth Syst. Sci. Data, 13, 3907-3925, https://doi.org/10.5194/essd-13-3907-2021, 2021.

Yang, Y., Zhang, S., Liu, Y., Xing, X., and De Sherbinin, A.: Analyzing historical land use changes using a historical land use reconstruction model: a case study in zhenlai county, northeastern china, Sci Rep, 7, 41275, https://doi.org/10.1038/srep41275, 2017.

Ye, X., Garber, P. A., Li, M., and Zhao, X.: Climate and anthropogenic activities threaten two langur species irrespective of their range size, Diversity & Distributions, 30, n/a, https://doi.org/10.1111/ddi.13841, 2024.

Ye, Y. Wei, X. Li, F. and Fang, X.: Reconstruction of cropland cover changes in the Shandong Province over the past 300 years, Sci. Rep., 13642, https://doi.org/10.1038/srep13642, 2015.

Ye, Y. and Fang, X.: Expansion of cropland area and formation of the eastern farming-pastoral ecotone in northern china during the twentieth century, Reg. Envir. Chang., 923-934, https://doi.org/10.1007/s10113-012-0306-5, 2012.

Ye, Y., Fang, X., Ren, Y., Zhang, X., and Chen, L.: Cropland cover change in northeast china during the past 300 years, Science China Earth Sciences, 52, 1172-1182, https://doi.org/10.1007/s11430-009-0118-8, 2009.

Yu, Z., Jin, X., Miao, L., and Yang, X.: A historical reconstruction of cropland in china from 1900 to 2016, Earth Syst. Sci. Data, 13, 3203-3218, https://doi.org/10.5194/essd-13-3203-2021, 2021.

Yu, Z. A. L. C.: Largely underestimated carbon emission from land use and land cover change in the conterminous United States., Glob. Change Biol., 3741-3752, 2019.

Yu, Z. and Lu, C.: Historical cropland expansion and abandonment in the continental U.S. During 1850 to 2016, Glob. Ecol. Biogeogr., 27, 322-333, https://doi.org/10.1111/geb.12697, 2018.

Zanaga, D. V. D. K.: Esa worldcover 10m 2020 v100, https://doi.org/https://doi.org/10.5281/zenodo.5571936, 2021.

Zhang, C., Fang, X., Ye, Y., Tang, C., Wu, Z., Zheng, X., Zhang, D., Jiang, C., Li, J., Li, Y., and Zhao, Z.: A spatially explicit reconstruction of cropland cover in china around 1850 c.e. Employing new land suitability based gridded allocation algorithm, Quat. Int., 641, 62-73, https://doi.org/10.1016/j.quaint.2022.06.001, 2022.

Zhang, L., Jiang, L., Zhang, X., Zhang, A., and Jiang, C.: Reconstruction of cropland over heilongjiang province in the late 19th century, Acta Geographica Sinica, 69, 448-458, https://doi.org/10.11821/dlxb201404002, 2014.

Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X., and Liu, L.: Glc_fcs30d: the first global 30-m land-cover dynamic monitoring product with a fine classification system from 1985 to 2022 using dense time-series landsat imagery and continuous change-detection method, Earth Syst. Sci. Data, 2023, 1-32, https://doi.org/10.5194/essd-2023-320, 2023.

Zhao, Y., Wang, K., Zhao, B., and Wang, F.: Spatio-temporal process and pattern of the establishment of county-level administrative divisions in China in the past 2200 years, Acta Geographica Sinica, 79, 890-908, https://doi.org/10.11821/dlxb202404005, 2024.

East Branch Railway Administration Of Russia and South Manchuria Railways Co.: North manchuria and east branch railway, Volume 1, South Manchuria Railways Co., Dalian, China, 1923 (in Japanese).

East Branch Railway Administration Of Russia and South Manchuria Railways Co.: North manchuria and east branch railway, Volume 2, South Manchuria Railways Co., Dalian, China, 1923 (in Japanese).

~~Middle East Railway Economic Survey: Economic map of North Manchuria, Middle East Railway Economic Survey, Harbin, China, 1931 (in Japanese).~~

Office Of The Governor-General Of Kwantung: Economic Statistics of Manchuria and Mongolia, Office of the Governor-

General of Kwantung, Dalian, China, 1918 (in Japanese).

~~A, G.: General Chorography of Shengjing (ed. in 1736), Liaohai Publishing House, Shenyang, China, 1997 (in Chinese).~~

~~Buck, J. L.: Land use in China, Cheng City Press, Chengdu, China, 1941 (in Chinese).~~

Cao, S. and Ge, J.: The History of Chinese Population, Volume 4, Fudan University Press, Shanghai, China, 2005b (in Chinese).

Cao, S. and Ge, J.: The History of Chinese Migration, Volume 5, Fudan University Press, Shanghai, China, 2022 (in Chinese).

~~Committee Of Science And Technology In Northeast China: Agricultural Statistics of Northeast China, Committee of Science and Technology in Northeast China, 1946 (in Chinese).~~

Committee Of Integrative Survey Of Natural Resources and Committee Of National Planning Of Chinese Academy Of Sciences: Data Compilation of National Land Resources, Volume 1, Committee of National Planning of Chinese Academy of Sciences, 1989 (in Chinese).

Committee Of Integrative Survey Of Natural Resources and Committee Of National Planning Of Chinese Academy Of Sciences: Data Compilation of National Land Resources, Volume 3, Committee of National Planning of Chinese Academy of Sciences, Beijing, China, 1990 (in Chinese).

Cong, P.: Garrison Reclamation in Liaodong in the Ming Dynasty, Journal of Chinese Historical Studies, 93-107, 1985 (in Chinese).

Cong, P.: Agriculture in Liaoyang Province in the Yuan Dynasty, Northern Cultural Relics, 78-88, https://doi.org/10.16422/j.cnki.1001-0483.1993.01.019, 1993a (in Chinese).

Cong, P.: Khitan, Goryeo, Semu and Mongolian in Liaoyang Province in the Yuan Dynasty, Collected Papers of History Studies, 7-14, 1993b (in Chinese).

Cropland Research Group: Quantitative Economic Analysis of Cropland Decline in China, Economic Science Press, Beijing, China, 1992 (in Chinese).

Fan, X.: Research on the Mechanism of the Layout of Ming Great Wall Military Defense System, Ph.D. thesis, Tianjin University, 324 pp., 2015 (in Chinese).

Fang, X., He, F., Wu, Z., and Zheng, J.: General characteristics of the agricultural area and fractional cropland cover changes in China for the past 2000 years, Acta Geographica Sinica, 1732-1746, https://doi.org/10.11821/dlxb202107012, 2021 (in Chinese).

Fang, X., Ye, Y., Ge, Q., and Zheng, J.: Land Exploitation in the Northeast China during the Qing Dynasty Inferred from the Development of Town System, Scientia Geographica Sinica, 25, 129-134, https://doi.org/10.3969/j.issn.1000-0690.2005.02.001, 2005 (in Chinese).

Fang, X., Ye, Y., and Zeng, Z.: Interaction of extreme climate event—cultivation—policy management, Science China Earth Sciences, 36, 680-688, https://doi.org/10.3969/j.issn.1674-7240.2006.07.009, 2006 (in Chinese).

Ge, J.: The History of Chinese Population, Volume 1, Fudan University Press, Shanghai, China, 2002 (in Chinese).

Han, M.: Historical agricultural geography of China, Peking University Press, Beijing, China, 2012 (in Chinese).

Han, M.: Agricultural Geography of Liao and Jin Dynasties, Social Sciences Academic Press, Beijing, China, 1999 (in Chinese).

~~Ho, P.: Interpretation and evaluation of ancient and modern land quantity in China, China Social Sciences Press, Beijing, China, 1988 (in Chinese).~~

~~Heilongjiang Provincial Bureau Of Statistics: Heilongjiang Statistical Yearbook, China Statistics Press, Beijing, China, 1997 (in Chinese).~~

Inner Mongolia Provincial Bureau Of Statistics: Agricultural and animal husbandry production statistics, Inner Mongolia Provincial Bureau of Statistics, Hohhot, China, 1983 (in Chinese).

~~Ji, J., Li, X., Zhao, D., and Zhou, Y.: New Volume of General Chorography of Baqi (ed. in 1796), Literary History Press of Jilin, Changchun, China, 2002 (in Chinese).~~

~~Jinlin Provincial Bureau Of Statistics: Jinlin Statistical Yearbook, China Statistics Press, Beijing, China, 1997 (in Chinese).~~

Jin, Q. and Mikami, T.: Research on Jurchen in Jin Dynasty, Heilongjiang People's Publishing House, Harbin, China, 1984 (in Chinese).

Kong, J. and Feng, Y.: Historical Geography of Northeast China, Heilongjiang People's Publishing House, Harbin, China, 1989 (in Chinese).

~~Li, L.: Restoration and expansion of Northeast China farming area in the early Qing Dynasty, Journal of Chinese Historical Geography, 111-132, 1991 (in Chinese).~~

Li, M.: Reconstruction of cropland change data for eastern Asia and its spatial-temporal characteristics analysis in the last millennium, Ph.D. thesis, University of Chinese Academy of Sciences, China, 176 pp., 2019 (in Chinese).

~~Li, W., Zhang, P., and Song, Y.: Analysis on Land Development and Causes in Northeast China during Qing Dynasty, Scientia Geographica Sinica, 7-16, https://doi.org/10.3969/j.issn.1000-0690.2005.01.002, 2005 (in Chinese).~~Li, Y.: Land Resources of China, China Land Press, Beijing, China, 2000 (in Chinese).

~~Liaoning Provincial Archives, Jilin Provincial Archives, and Heilongjiang Provincial Archives: Compilation of Documents of Financial Economy History in Northeast China, Heilongjiang People's Press, Harbin, China, 1988 (in Chinese).~~

~~Liaoning Provincial Bureau Of Statistics: Liaoning Economic Statistics Yearbook, China Statistics Press, Beijing, China, 1997 (in Chinese).~~

Liu, P.: A Study of registered residence in Jin Dynasty, Study of Chinese History, 86-96, 1994a (in Chinese).

Liu, P.: A Study on the population status of Mengan Mouke in Jin Dynasty, Ethno-National Studies, 81-89, 1994b (in Chinese).

Liu, Z., Han, G., and Liu, H.: Quan Liao Zhi (ed. in 1566), Science Press, Beijing, China, 2016 (in Chinese).

~~Ministry Of Agricultural And Commercial: The Fifth Agricultural and Commercial Statistics Table, Ministry of Agricultural~~

and Commercial, Beijing, China, 1919 (in Chinese).

National Bureau Of Statistics: China Statistical Yearbook, China Statistics Press, Beijing, China, 2023 (in Chinese).

National Bureau Of Statistics: The Summary of Rural Social Economic Statistics on County Level in China, Chinese Statistical Press, Beijing, China, 1989 (in Chinese).

Shi, F.: History Draft of Chinese Population Migration, Heilongjiang People's Publishing House, Harbin, China, 1990 (in Chinese).

Shi, Z.: An Estimate of Agricultural Economic Indicators in the Qing Dynasty, Researches in Chinese Economic History, 5-30, 2015 (in Chinese).

South Manchuria Railways Co.: Weights and Measures in Northeast China, South Manchuria Railways Co., Dalian, China, 1927 (in Japanese).

South Manchuria Railways Co.: Encyclopedia of Manchuria and Mongolia (ed. in 1923), Inner Mongolia University Press, Hohhot, China, 2015 (in Chinese).

Shi, Z.: Analysis on Variable of the Population and Land in the Area of Gan, Ning and Qing of the Late Qing Dynasty, Agricultural History of China, 19, 72-79, 2000 (in Chinese).

Song, L.: Dynastic History of Yuan Dynasty, Zhonghua Publishing House, Beijing, China, 1976 (in Chinese).

Tan, Q.: The Historical Atlas of China, Volume 7, SinoMaps Press, Beijing, China, 1982a (in Chinese).

Tan, Q.: The Historical Atlas of China, Volume 6, SinoMaps Press, Beijing, China, 1982b (in Chinese).

Tian, Y.: The land use and land cover change in Re-Cha-Sui area in recent 300 years, Master. thesis, Graduate School of Chinese Academy of Sciences, 166 pp., 2005 (in Chinese).

Toqto'A: Dynastic History of Liao Dynasty, Zhonghua Publishing House, Beijing, China, 1974 (in Chinese).

Toqto'A: Dynastic History of Jin Dynasty, Zhonghua Publishing House, Beijing, China, 1975 (in Chinese).

Wang, Y.: Garrison Reclamation of the Ming Dynasty, Zhonghua Publishing House, Beijing, China, 2009 (in Chinese).

Wu, C.: History of Chinese Weights and Measures, The Commercial Press, Beijing, China, 1984 (in Chinese).

Wu, S. and Ge, J.: The History of Chinese Population, Volume 3, Fudan University Press, Shanghai, China, 2005a (in Chinese).

Wu, S. and Ge, J.: The History of Chinese Migration, Volume 4, Fudan University Press, Shanghai, China, 2022 (in Chinese).

Wu, Z.: Development of Gridded Allocation Method for Historical Cropland Based on Settlement Information and Its Application, Ph.D. thesis, Beijing Normal University, China, 184 pp., 2021 (in Chinese).

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

Xiong, Z.: Summary of county governance in Northeast China, Lida Publishing House, 1933 (in Chinese).

Xu, X., Li, B., Liu, X., Li, X., and Shi, Q.: Mapping annual global land cover changes at a 30 m resolution from 2000 to 2015, National Remote Sensing Bulletin, 1896-1916, https://doi.org/10.11834/jrs.20211261, 2021 (in Chinese).

Xue, L.: Research of the Governance Over the Northeast Region in the Yuan Dynasty, Social Sciences Academic Press, Beijing, China, 2012 (in Chinese).

Xue, L.: Research on the Liaoyang Province and Governance over the Northeast China of the Yuan Dynasty, Ph.D. thesis, Nankai University, 261 pp., 2006 (in Chinese).

~~Yang, Y., Wang, G., and Zhang, Y.: Northeast China history of the Qing Dynasty, Liaoning Education Publishing House, Shenyang, China, 1990 (in Chinese).~~

Ye, Y., Fang, X., Dai, Y., Zeng, Z., and Zhang, X.: Calibration of cropland data and reconstruction of rate of reclamation in Northeast China during the period of Republic of China, Progress in Natural Science, 16, 1419-1427, 2006 (in Chinese).

Zeng, Z., Fang, X., and Ye, Y.: The Process of Land Cultivation Based on Settlement Names in Jilin Province in the Past 300 Years, Acta Geographica Sinica, 985-993, 2011 (in Chinese).

Zhan, J.: Heilongjiang History of Land Reclamation in border Areas, Volume 1, Social Sciences Academic Press, Beijing, China, 2017 (in Chinese).

Zhang, L., Fang, X., Ren, G., and Suo, X.: Environmental Changes in the North China Farming Grazing Transitional Zone, Earth Science Frontiers, 127-136, https://doi.org/10.3321/j.issn:1005-2321.1997.01.015, 1997 (in Chinese).

Zhang, T.: Dynastic History of Ming Dynasty, Zhonghua Publishing House, Beijing, China, 1974 (in Chinese).

Zhang, Y.: Re-estimation on China's population and cropland in modern times, Researches in Chinese Economic History, 20-30, 1991 (in Chinese).

Zhou, S.: Study of Jurchen in Liaoyang Province in the Yuan Dynasty, Shanghai Jiaotong University Press, Shanghai, China, 2021 (in Chinese).

Zhou, R.: A General Inspection and Re-appraise on Area Under Cultivation in the Early Period of the Qing, Journal of Chinese Social and Economic History, 39-49, https://doi.org/10.3969/j.issn.1000-422X.2001.03.004, 2001 (in Chinese).