https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

Open Access Earth System Science Data Discussions

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 Based on historical documents, population data, published results, remote sensing data products, statistical data and survey data, this study reconstructed the cropland area and the spatial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 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 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 dataset of changes in cropland of administrative districts in Northeast China, reconstructed based on improved historical cropland reconstruction methods, significantly enhances time resolution and reliability. Additionally, the dataset shows the 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 gridded dataset.

1 Introduction

Land 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; Shukla et al., 1990). Cropland constitutes one of the primary land use types, being a land category susceptible to human influence and undergoing alterations. 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). Presently, various global historical Land Use and Land Cover (LUCC) datasets, exemplified by

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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. 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; 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). However, there still exists a disparity or uncertainty in the standardization and spatiotemporal accuracy of the aforementioned cropland data. 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 regional agricultural development history. Therefore, precise cropland change data, particularly long-term cropland datasets standardized with high spatiotemporal resolution will not merely improve the accuracy and 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 analyses.

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). Throughout the prolonged agricultural development, the natural vegetation landscape in the Northeast region has undergone notable transformations.

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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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.

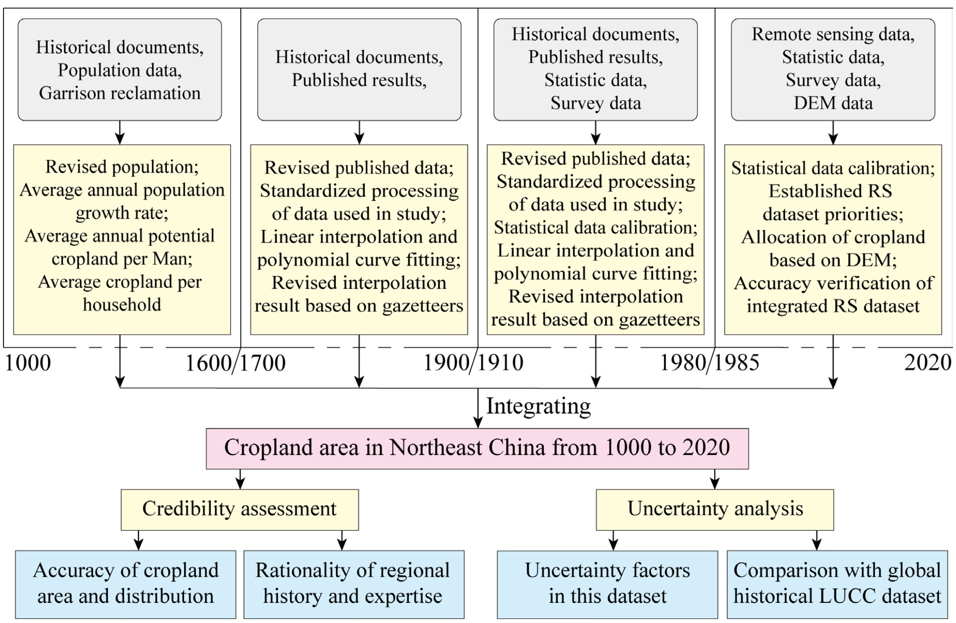


Figure 1: The framework for reconstructing cropland area of Northeast China from 1000 to 2020.

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.

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

Open Access Earth System Science Data Discussions

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. 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. 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). Furthermore, due

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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; 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. 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.,

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

Open Access Earth System Science Data Discussions

2016; Tan, 1982a), the historical maps for the 1500 and 1600 are divided into three provincial districts, and the number of population 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).

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	—	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)	

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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 = 1 : 2	Average per garrison soldier is 46 *Mu* (3.07 hm^2); Regular soldiers and one-third of their dependents is 14 *Mu* (0.93 hm^2)	4875(1500); 5868(1600)
	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 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 Yanyu, are available for the years 1724, 1782, 1868, and 1911 (Tian, 2005).

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:

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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 km^2 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 hm^2) 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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

that time by the number of settlements in the respective county to obtain the cropland area (Appendix A). 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 A).

(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 the total cropland area: Previous 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 for linear interpolation 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. 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 the CNEC data (Ye et al., 2009) in 1735 for linear interpolation to obtain cropland area data for 1700, 1750, 1800, and 1850 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.

(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.

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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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 A.

(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% (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,

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

1960, 1970, and 1980.

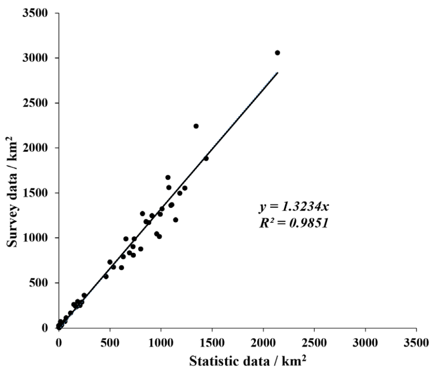


Figure 2: Correlation between the statistical cropland data and survey cropland data of the counties in the Eastern of Inner Mongolia in 1980's.

(4) Linear interpolation and polynomial curve fitting to control the total cropland area: This study selected CNEC (Ye et al., 2009) data in 1908 and 1914 for linear interpolation 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 of cropland area data for each county or district in the Eastern of Inner Mongolia, 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. 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 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),

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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 1985 county-level 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, prefecture-level data from the second national land survey, and county-level data from the third national land survey.

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	2000-2015	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	1985-2020	1.Cropland	https://doi.org/10.5281/zenodo.4417810 [2024/01/10]	(Yang and Huang, 2021)
ESA_WorldCover	Sentinel-1 Sentinel-2	Boolean	10m	2020	40.Cropland	https://viewer.esa-worldcover.org/worldcover/ [2024/01/10]	(Zanaga, 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	1985-2020	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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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. 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=1.256x 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). This indicates a high correlation between the two types of data at the three time points, and the survey data is approximately 19.3% to 25.6% higher than the statistical data at the same period, with an average of about 22%, then corrected cropland area data by 22% for each county in the study area for the 1990, 2000 and 2005.

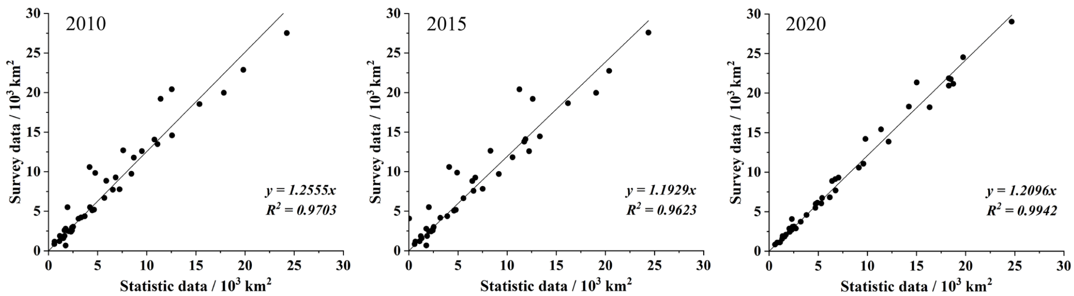


Figure 3: Correlation between the statistical cropland data and survey cropland data of the cities in the Northeast China in 2010, 2015 and 2020.

(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):

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

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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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.

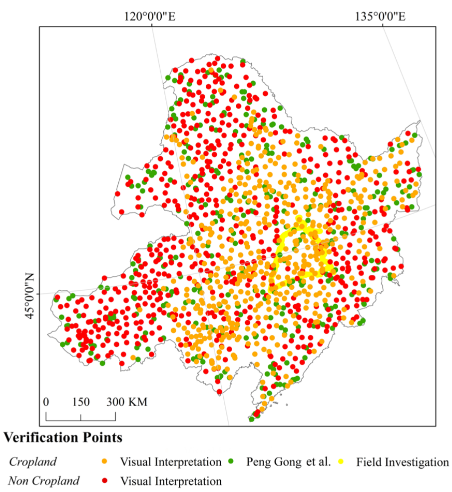
(4) Accuracy assessment and validation of RS products integration results: This study utilizes the confusion matrix was used to assess the accuracy of cropland products. The Producer Accuracy (P.A.) and User Accuracy (U.A.) for each product in 2020 are calculated as two indicators to evaluate the reliability 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 results (Fig. 4): (1) 346 cropland sample points located in the study area from FROM-GLC. (2) 1052 sample points obtained through field investigations conducted by the author 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%. For non-cropland pixels, the producer accuracy is 91.12%, and the user accuracy is 87.32%. The overall accuracy is relatively high.



https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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). 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).

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. 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 × 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).

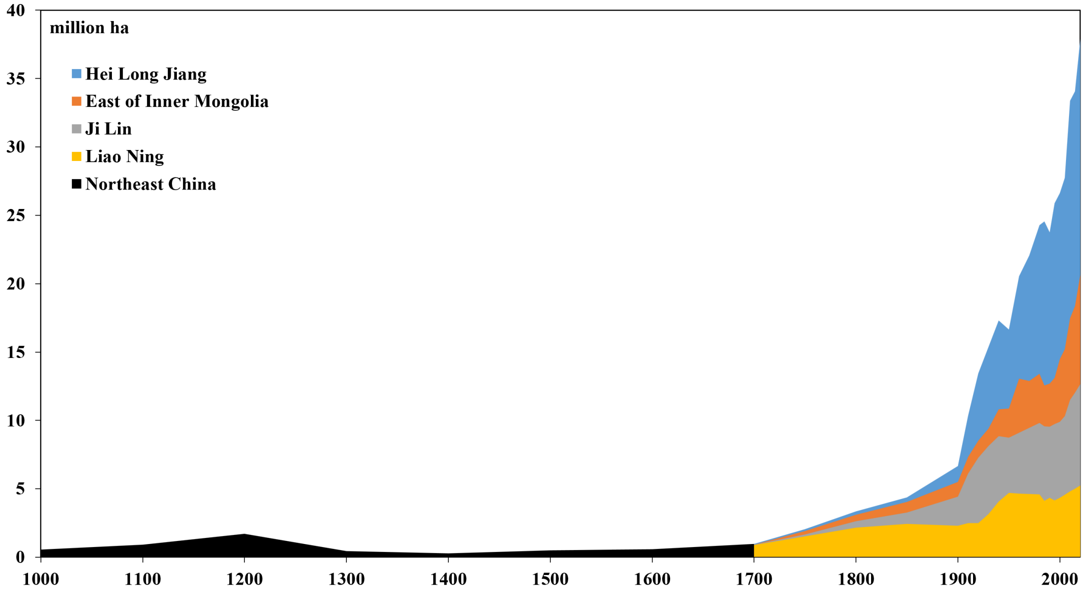


Figure 5: Changes in total cropland area in the Northeast China from 1000 to 2020.

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. 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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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.

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

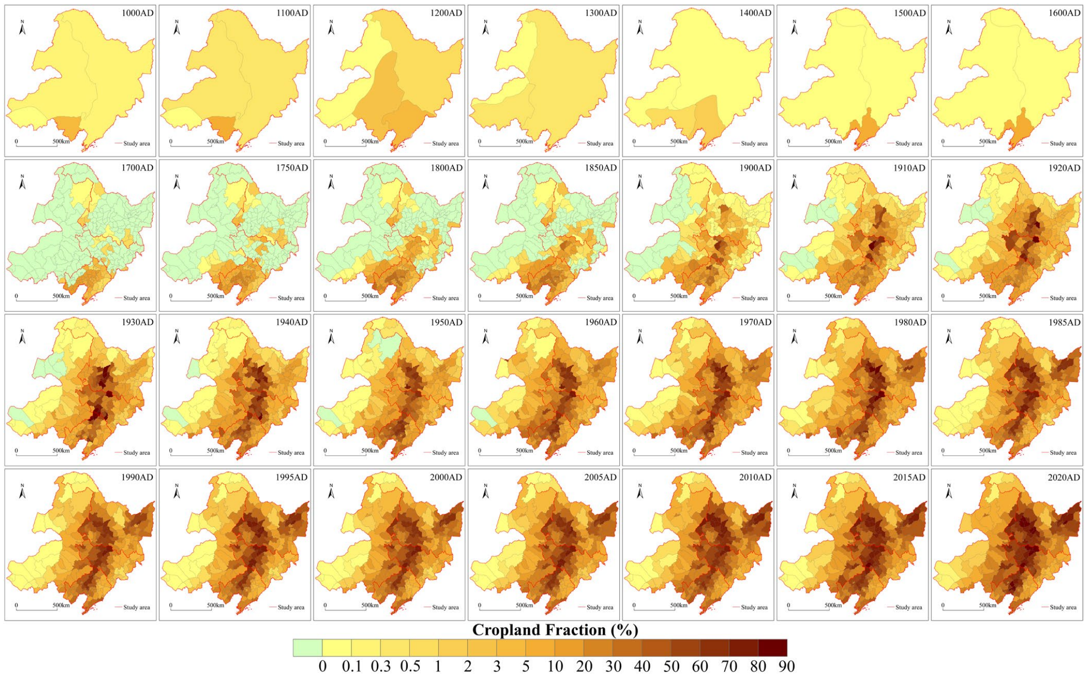


Figure 6: Changes in spatial patterns of cropland in the Northeast China from 1000 to 2020.

4 Discussion

4.1 Comparison with global historical LUCC dataset

To better showcase the achievements of this study, we chose to compare our results with widely used global historical LUCC datasets: 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.

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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.

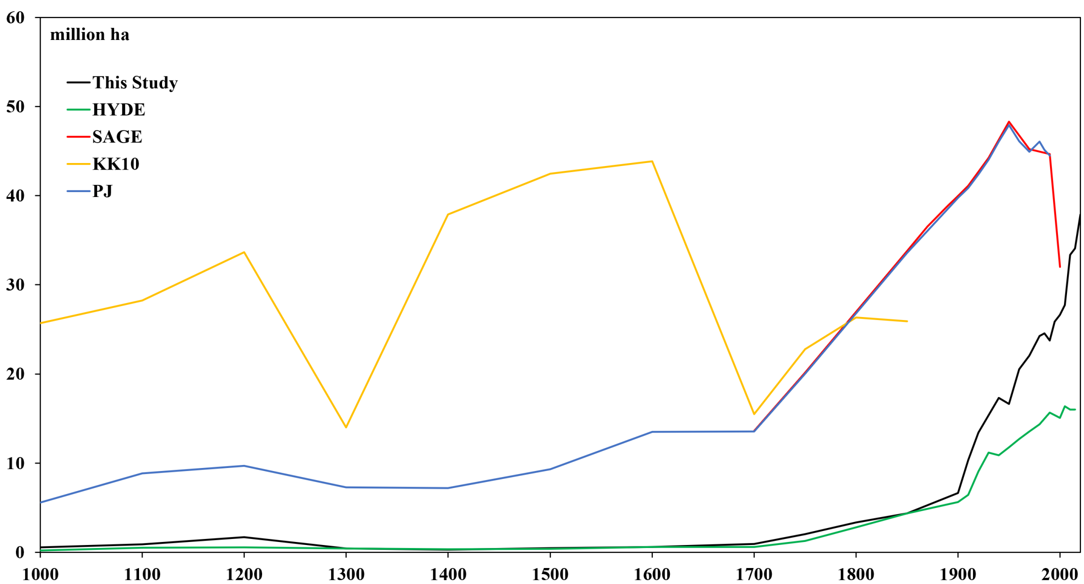


Figure 7: Comparison of total cropland area in the past millennium between HYDE3.2, SAGE, KK10, PJ and this study in the Northeast China.

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. 6~9). 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.

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

Open Access Earth System Science Data Discussions

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, 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). Compared to the HYDE3.2 dataset, counties with RD greater than 100% gradually decreased (Fig. 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. 9), indicating that the cropland area in most counties in this study was significantly greater than the HYDE3.2 dataset.

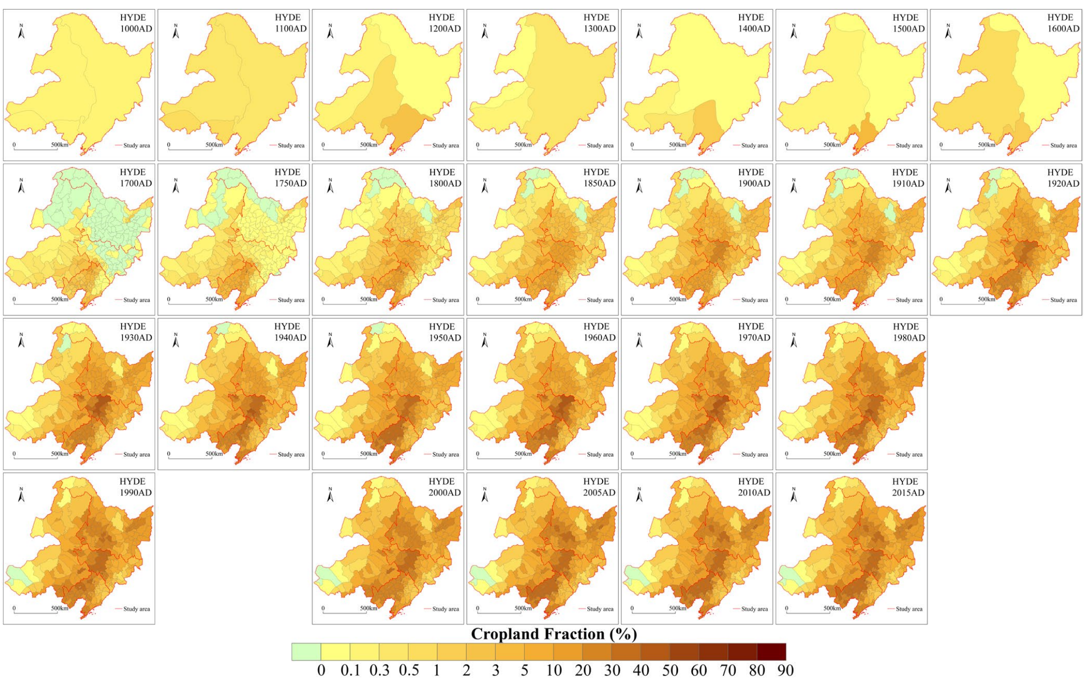


Figure 8: Changes in spatial patterns of cropland of HYDE3.2 dataset in the Northeast China from 1000 to 2015.

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

Open Access Earth System Science Data Discussions

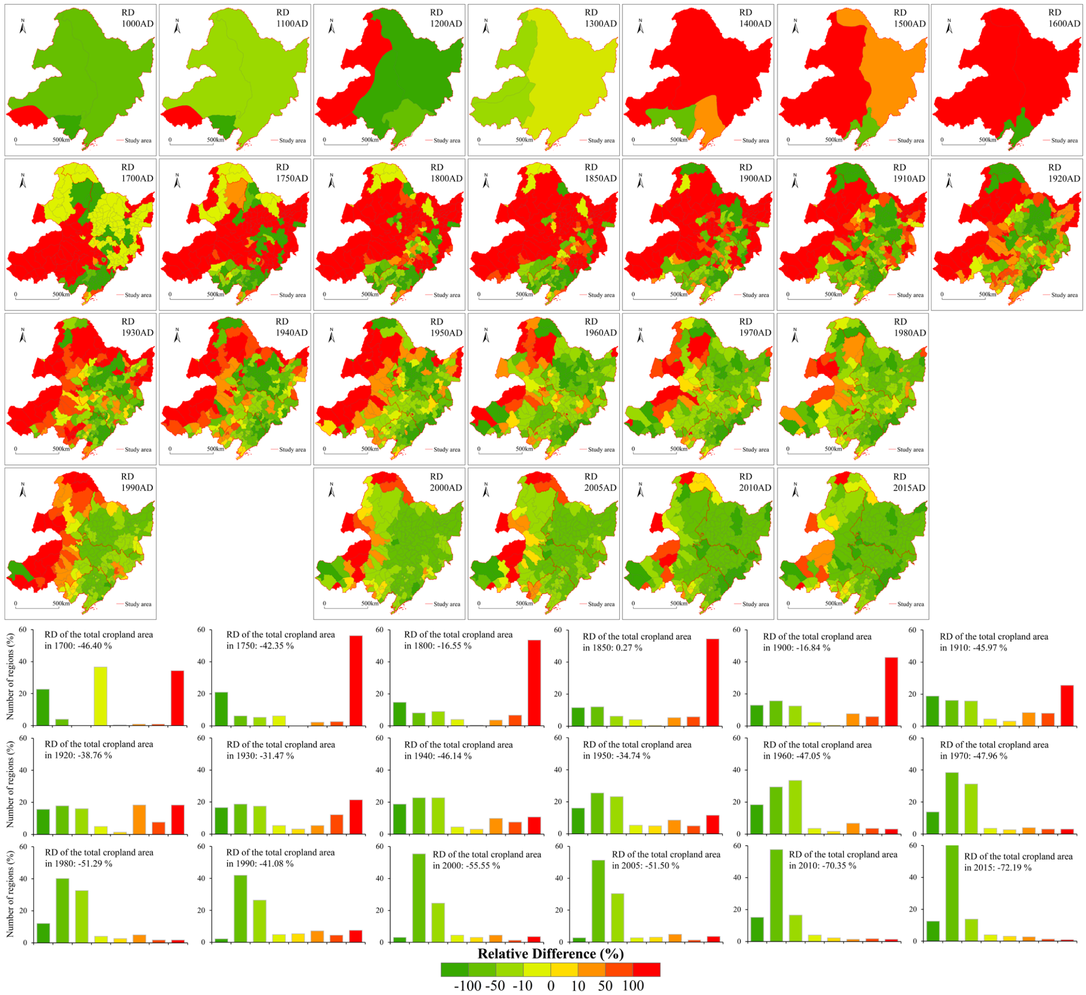


Figure 9: Comparison of the spatial distribution of cropland area between HYDE3.2 and this study in the Northeast China.

4.2 Uncertainty analysis

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. Temporary changes in land use and fallow land during historical periods were not considered, 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. (3) From 1700 to 1980, cropland areas at multiple time points in this study were derived through linear interpolation. This method, compared to data recorded at each specific historical point, may affect the accuracy of the value at those standard time points. (4) 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. The cropland area calculated based on the proportion of overlapping areas between the two may cause minor errors.

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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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 climate-ecosystem modeling and the conservation and utilization of black soil.

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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

		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 settlements and the cropland area owned by unit settlement in this region in 1780
Manchu Autonomous County of Yitong		
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		

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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).

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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.

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.

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.

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.

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, 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.

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.

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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.

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.

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.

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.

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.

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.

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,

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

2017.

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.

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.

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).

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).

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).

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).

Committee Of Science And Technology In Northeast China: Agricultural Statistics of Northeast China, Committee of Science

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

and Technology in Northeast China, 1946 (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).

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

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

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

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

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).

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).

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).

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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

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).

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).

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).

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).

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

Ministry Of Agricultural And Commercial: The Fifth Agricultural and Commercial Statistics Table, Ministry of Agricultural and Commercial, Beijing, China, 1919 (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).

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

https://doi.org/10.5194/essd-2024-94
Preprint. Discussion started: 8 April 2024

Open Access Earth System Science Data Discussions

China, 2017 (in Chinese).

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

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).

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).

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

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).

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