

Article

A Settlement Density Based Allocation Method for Historical Cropland Cover: A Case Study of Jilin Province, China

Zhilei Wu ^{1,2,3}, Xiuqi Fang ^{2,*} and Yu Ye ²¹ College of Geographical Science, Qinghai Normal University, Xining 810008, China² Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China³ Academy of Plateau Science and Sustainability, Xining 810016, China

* Correspondence: xfang@bnu.edu.cn

Abstract: A key focus in research on changes in historical land cover has been to improve existing gridded cropland allocation methods based on land suitability for cultivation to generate credible historical cropland cover data. This study developed a settlement-density-based method for gridded cropland allocation using the locations of settlements to identify the cropland grid and the settlement density as the weight for allocating the cropland area to the grid. This method was applied to allocate the provincial cropland areas in Jilin Province, China, to a 5' × 5' cropland cover at six time points during the last 300 years. The credibility of the reconstruction was assessed using three methods. The following conclusions emerged. First, the settlement density method is founded on the fact of coexistence between rural settlements and cropland. Cropland is only distributed in the grid where the settlements exist, and the cropland area of a grid equals to the cropland area per settlement multiplying by the number of settlements within the grid, without considering differences of settlement size. Second, all three quantitative or qualitative assessments of Jilin Province confirmed the credibility and feasibility of the settlement density method. Therefore, the use of this method to reproduce the temporal and spatial changes in cropland cover in new reclamation regions, such as Jilin Province, is valid. This study provides valuable inputs for enhancing the credibility of historical global land cover data by incorporating human factors into the cropland allocation method.

Keywords: LUCC; historical cropland cover; gridded allocation method; settlement; Jilin Province of China



Citation: Wu, Z.; Fang, X.; Ye, Y. A Settlement Density Based Allocation Method for Historical Cropland Cover: A Case Study of Jilin Province, China. *Land* **2022**, *11*, 1374. <https://doi.org/10.3390/land11081374>

Academic Editor: Francisco Manzano Agugliaro

Received: 19 July 2022

Accepted: 19 August 2022

Published: 22 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the most prevalent types of anthropogenic land cover, cropland not only directly transforms natural land cover types, it also indirectly influences environmental changes at various scales, from regional to global, by inducing changes in the physical conditions and biogeochemical cycles of land surfaces [1–4]. Therefore, the reconstruction of historical changes in cropland is central to research on anthropogenic changes in land cover, and several historical land use/cover datasets have been developed, notably SAGE [5], HYDE [1,6–9], PJ [10], and KK10 [11]. All of these datasets include cropland cover or anthropogenic land cover and have been widely used in studies of global climate and environmental change [12–19].

Despite continuous improvements in the quality of these datasets through the release of updated versions or through improvements in reconstruction methods, uncertainties remain. Not only are there discrepancies among these global cropland datasets [11,20] but their credibility has also been questioned within a large body of regional studies that have used historical records [21–29]. Moreover, even the creators of historical land use/cover datasets, for example, HYDE (version 3.2), have acknowledged that there are some uncertainties pertaining to their data on gridded cropland cover [9].

Factors affecting the credibility of historical cropland cover data relate to the source of historical cropland records, methods of reconstructing the total area of cropland, and

gridded methods of cropland allocation. A large number of available regional historical documents and cropland reconstructions have enabled a continuous reduction in uncertainties caused by the first two factors [11,30]. However, the gridded method of cropland allocation requires further improvement and remains a challenging area in the field of historical land use/cover research.

The gridded allocation of historical cropland is meant to allocate the cropland area within the administrative units to those grids that were not only suitable for cultivation, but had also genuinely been cultivated in the historical period. Existing gridded allocation methods generally entail the use of a comprehensive index of present-day natural factors for assessing the land suitability for cultivation. Following the basic assumption that cultivation is established in the order of “good land first and bad land later,” the cropland area is allocated to each grid according to the land suitability for cultivation in descending order from highly suitable for cultivation to a low level of suitability (Figure 1) [9–11,31–36].

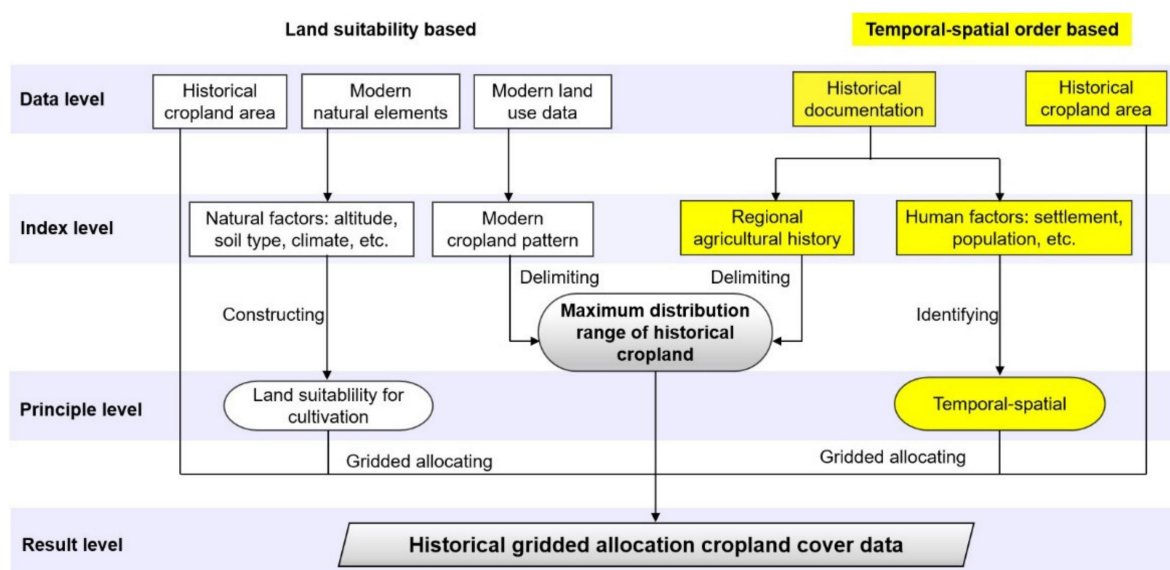


Figure 1. A comparison of two methodological frameworks for gridded allocation of historical cropland cover. The one to the left is widely used, and the one to the right, highlighted in yellow, was introduced in this study.

The rationale for basing gridded allocation on land suitability for cultivation is that the grid with higher land suitability for cultivation can ensure that its priority is to be allocated to cropland area (Figure 1). However, in practice, cultivation activities are influenced by many factors other than land suitability. Given that land suitability for cultivation is only one necessary condition for cultivation, a grid with higher land suitability for cultivation does not necessarily mean it must be cultivated as cropland first. Thus, there may be a discrepancy between allocated historical cropland cover that relies solely on land suitability for cultivation and the actual distribution of historical cropland, which is an important source of the uncertainty relating to gridded historical cropland [37].

To improve the credibility of the allocated historical cropland cover, we sought to design a new method of gridded cropland allocation incorporating human factors that could reflect both the location and the chronological order of human cultivation activities (Figure 1). We first selected settlements on the basis of various human factors associated with the reclamation of cropland. Next, we used settlement density to conduct a method for gridded allocation of historical cropland cover. Finally, we assessed the feasibility of this allocation method in a case study of the allocation of the cropland cover in Jilin Province of China over the last 300 years.

2. Settlement Density Based Allocation Method for Cropland Cover

2.1. The Indication of Settlement to the Distribution of Cropland

Considered from the perspective of its service functions, land can be used as a living space, a productive space, or an ecological space (Figure 2). In the agricultural age, cropland, as the main component of the productive space, was always found concurrently with rural settlements, the living space for residences. Thus, a rural settlement was always coeval with the surrounding cropland [38–41].

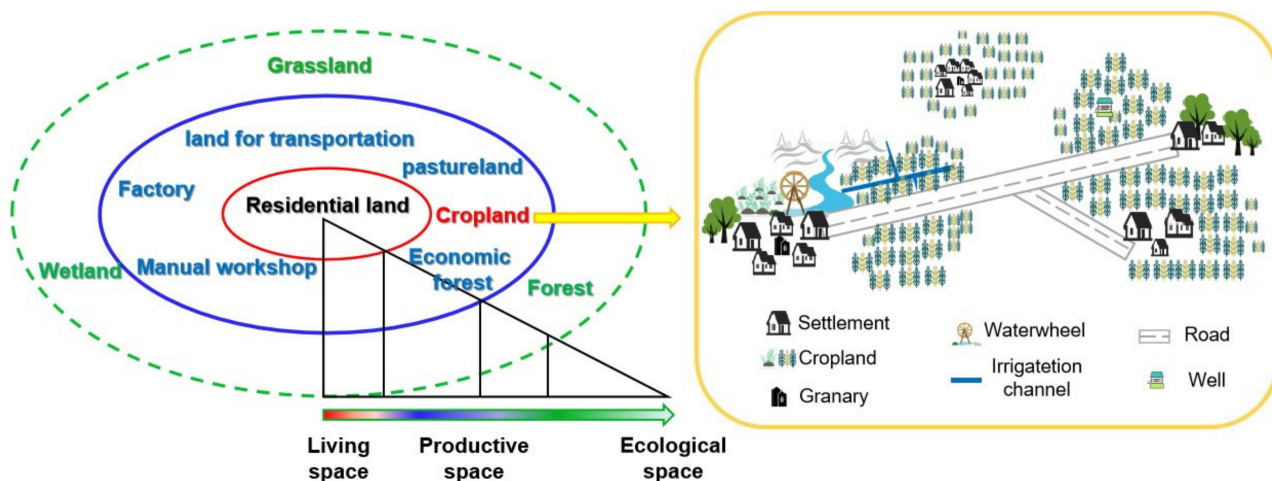


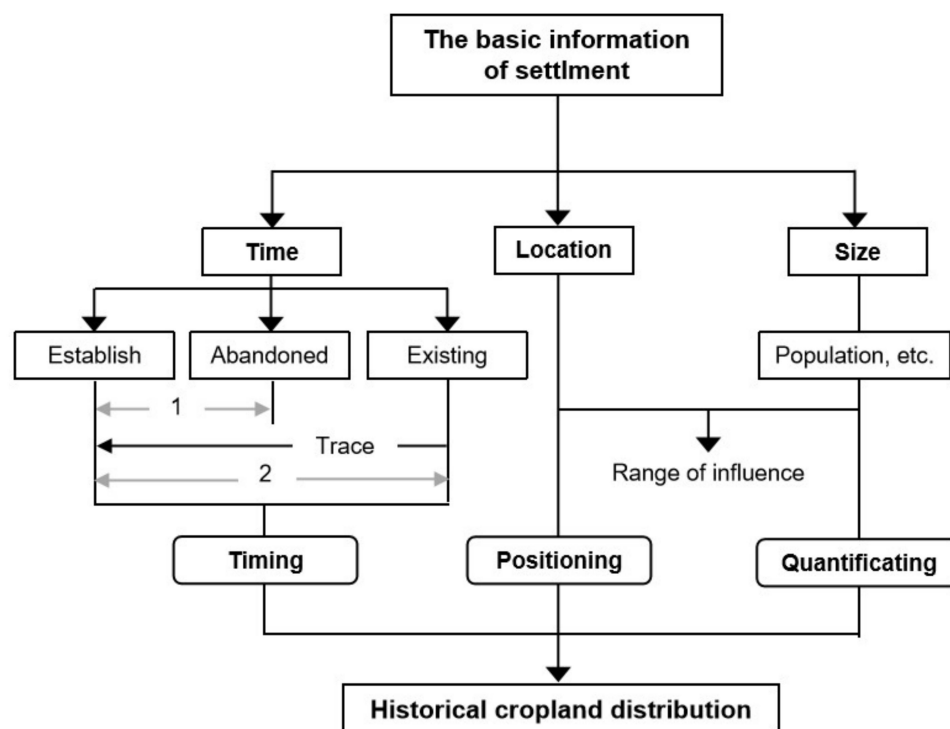
Figure 2. The spatial distribution of land used for production, living, and ecological services and a sketch map of human factors relating to the distribution of cropland.

Agricultural activities are always performed on land suitable for cultivation distributed around settlements. Humans may develop agricultural facilities such as water wheels, irrigation channels, or wells on the cropland to sustain these activities, and they also build granaries to store grain. Moreover, additional facilities, such as roads, may be constructed between settlements to facilitate exchanges of people and goods (Figure 2). These human factors are closely related to the distribution of cropland. The existence of these human factors could indicate that, to some degree, there was cropland nearby during the corresponding historical periods. However, through comparing the characteristics of each human factor (Table 1), we identified the settlement, which, as the living space, was always coeval with the cropland around, and whose information on the location and time period has high accessibility, is more suitable to be an indicator of historical cropland for developing a method for allocating historical cropland cover.

Table 1. Examples of human factors and their usability for indicating the distribution of historical cropland.

Human Factor	Associated Form with Cropland	Information Source	Extractable Indicators	Indicative Significance to Cropland			Geometry
				Time	Location	Scale	
Cropland	Equal	Cropland relics Fossil soil	Area, Crop variety Pollen	✓	✓	✓	Polygon Point
				✓	✓	✓	
Population	Necessary and insufficient conditions	Archaeological site	Grave	✓	✓	✓	Point
		Historical record	Population size and structure	✓	✓	✓	Polygon
Settlement	Sufficient and necessary condition	Archaeological site	Relic building	✓	✓	✓	Point
		Historical record	Name, Location, Site conditions	✓	✓	✓	Point
		Gazetteer	Name, Location, History of settlement	✓	✓	✓	Point

Settlements serve as indicators of the existence of cropland, and specifically, of the time period of use, location, and quantity of cropland (Figure 3). The existing time of settlement means that cropland existed at the same time. The locations of settlements can accurately indicate the distribution of cropland during a historical period [37,42–44]. The number and size of settlements can be used to estimate the total cropland area [45]. Hence, the availability of well-kept records on settlements in a study region can be used to develop a cropland allocation method for credibly reconstructing the spatial distribution of historical cropland cover using the available information on settlements.



1 and 2 represents the period of existence of the settlement

Figure 3. The use of basic information on settlements to indicate the distribution of cropland.

2.2. Settlement Density-Based Cropland Gridded Allocation Method

2.2.1. The Algorithm for Settlement Density-Based Cropland Allocation Method

Settlement density refers to the number of settlements within each spatial unit. We used basic information on settlements, such as the time of existence and location, to develop a method for allocating cropland on a grid that was based on settlement density. Hereinafter, we will refer to this method as the settlement density method.

The core premise underlying the settlement density method is that during a particular historical period, cropland was generally distributed around settlements. Hence, the locations of settlements could also indicate those of cropland, and the total cropland area within a region would comprise the sum of cropland area around each settlement. Before cultivation reached a saturation point, a correlation would have existed between the number of settlements and the amount of cropland area in the region (Figure 4). Assuming that the differences in land suitability for cultivation within the region can be discounted, the spatial unit with a higher settlement density would have a higher cropland fraction for the corresponding historical period.

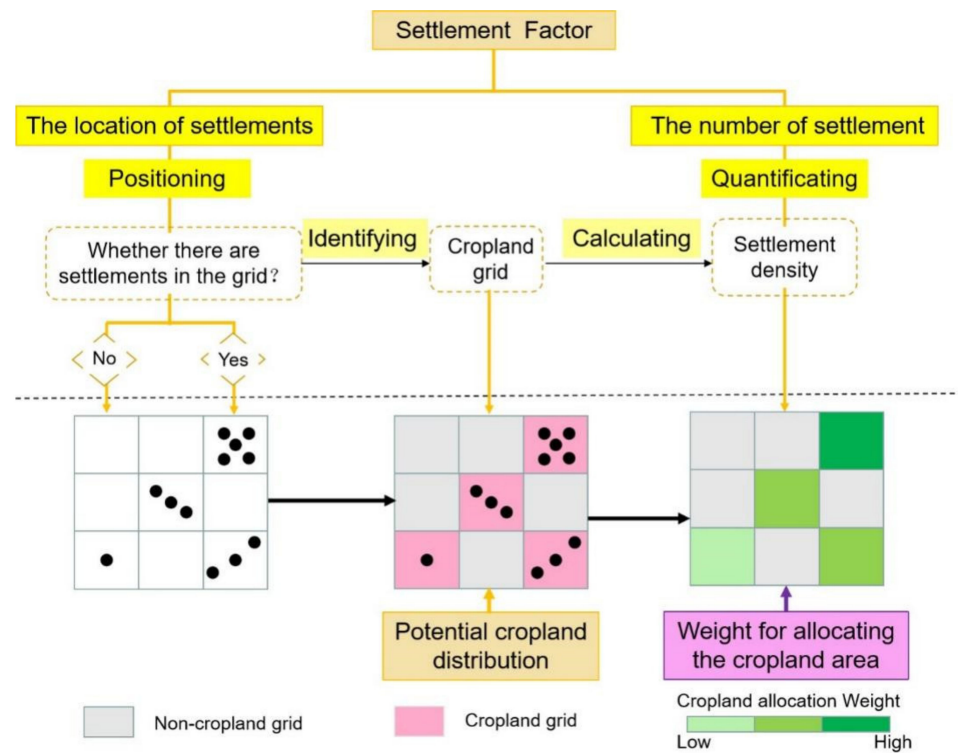


Figure 4. The framework for applying the settlement density-based cropland allocation method.

The basic algorithm applied in the settlement density method is shown in Formulas (1)–(4):

$$CA(w, t) = a_1 + a_2 + a_3 + \dots + a_N \tag{1}$$

$$\bar{a}(w, t) = \frac{(a_1 + a_2 + a_3 + \dots + a_N)}{SN(w, t)} \tag{2}$$

$$CA(w, t) = \bar{a}(w, t) \times SN(w, t) \tag{3}$$

$$CA(i, t) = \bar{a}(w, t) \times SN(i, t) \tag{4}$$

In Formula (1), $CA(w, t)$ denotes the total cropland area in year t and region w , and a_N denotes the cropland area of the N th settlement in year t and region w . N is the serial number of the settlement. In Formula (2), $\bar{a}(w, t)$ denotes the average cropland area per settlement in year t and region w , and $SN(w, t)$ denotes the number of settlements in year t and region w . Formula (3) is derived from Formulas (1) and (2), that is, the total cropland area in a region is equal to the product of the average cropland area per settlement and the number of settlements. At the grid scale, the calculation was performed using Formula (4). Therefore, the cropland area in year t and grid i is equal to the product of the average cropland area per settlement in year t and region w and the number of settlements in grid i and year t .

2.2.2. Steps for Performing the Settlement Density-Based Cropland Allocation Method

Figure 4 shows the framework used for applying the settlement density-based cropland allocation method. We first identified the cropland or the non-cropland grids according to whether or not they contained settlements. Next, we used the settlement density of a grid as a weight for allocating the cropland area to each grid. In this way, according to the settlement density, the total cropland area could be allocated to all cropland grids, usually represented as cropland fraction indicating the proportion of cropland area at the grid scale.

The following formulas were used for the calculation:

$$CA(i, t) = CA(w, t) \times Z(i, t) \tag{5}$$

$$CA(w, t) = \sum_{i=1}^n CA(i, t) \quad (6)$$

$$Z(i, t) = \frac{CA(i, t)}{\sum_{i=1}^n CA(i, t)} \quad (7)$$

In Formulas (5)–(7), $Z(i, t)$ values ranging from 0 to 1 denote the weight for allocating the cropland area in grid i and year t .

The weight of cropland allocation was based on the number of settlements and derived from Formulas (4) and (7), which were simplified into Formula (8) as follows:

$$Z(i, t) = \frac{SN(i, t)}{\sum_{i=1}^n SN(i, t)} \quad (8)$$

$$SD(i, t) = \frac{SN(i, t)}{area(i)} \quad (9)$$

In Formula (9), $SD(i, t)$ denotes the settlement density in grid i and year t , $SN(i, t)$ denotes the number of settlements in grid i and year t , and $area(i)$ is the area of the grid. The weight of cropland allocation, which was based on the settlement density, was derived from Formulas (8) and (9), which were simplified into Formula (10) as follows:

$$Z(i, t) = \frac{SD(i, t)}{\sum_{i=1}^n SD(i, t)} \quad (10)$$

The allocated cropland area was derived from Formulas (5) and (10) as follows:

$$CA(i, t) = CA(w, t) \times \frac{SD(i, t)}{\sum_{i=1}^n SD(i, t)} \quad (11)$$

The fraction of the grid comprising cropland was calculated using Formula (12) as follows:

$$FR(i, t) = \frac{CA(i, t)}{area(i)} \quad (12)$$

In Formula (12), $FR(i, t)$ denotes the cropland fraction in grid i and year t , $CA(i, t)$ denotes the cropland area being allocated in grid i and year t , and $area(i)$ is the area of the grid. Notably, we considered that the grid area varies with latitude.

2.3. An Assessment of the Feasibility of the Settlement Density Method

There are direct and indirect ways of evaluating the credibility of gridded allocation cropland data during the historical period [30]. Direct assessment is an assessment of the accuracy and rationality of reconstructed results. Indirect assessment is an assessment of the credibility and rationality of the gridded allocation method. Direct verification of the credibility of gridded cropland cover during historical periods is the most convincing assessment. If the reconstructed results are found to be consistent with the historical records, this would mean that the gridded allocation method is reasonable, and that the resulting cropland allocation is credible. However, the direct assessment is usually limited by the availability of historical records and can not be applied for regions that lack historical records. In such situations, an indirect assessment has to be carried out to obtain an approximate assessment of the reliability and rationality of the reconstructed results.

This study employed three methods to assess the feasibility and advantages of the settlement density method that we had designed.

2.3.1. Verification of the Gridded Allocation Capacity of the Settlement Density Method Using Modern Cropland Cover Data

As was the case with many gridded allocation methods [33,34], the gridded allocation capacity of the settlement density method that we designed is quantitatively assessed utilizing modern cropland cover data. We first extracted the modern cropland area from the gridded cropland cover data in the region under study. Next, the modern regional cropland area was allocated to the grid using the settlement density method to generate a scenario of gridded modern cropland cover. Finally, we calculated the spatial correlation coefficients between the actual modern cropland cover and the scenario resulting from the gridded cropland allocation. A sufficiently strong correlation indicated that the settlement-density-based gridded allocation method was reasonable, and that the allocation of cropland obtained using this method should, therefore, also be credible.

2.3.2. The Rationality Assessment for the Allocated Result Based on Regional Historical Facts

We performed a qualitative assessment using relevant historical land use/cover evidence to compare the consistency of the allocated historical cropland cover with historical facts relating to land use [30]. The historical land use/cover evidence, including spatio-temporal attributes, was extracted from historical literature, ancient maps, and other data.

2.3.3. The Accuracy Assessment of the Allocated Results Using Sub-Regional Quantitative Reconstructed Cropland Cover Data

We performed a credibility assessment using quantitatively reconstructed historical cropland cover data as the baseline, considered as “true historical values”. To assess the accuracy of our allocation results obtained with the settlement density method, we used the quantitatively reconstructed sub-regional cropland cover data, considered as “true historical values” to calculate the absolute difference between “historically true values” and the sub-regional cropland cover area generated from our allocations.

3. A Case Study of Jilin Province over the Last 300 Years

3.1. An Overview of Jilin Province

The study area is located in Jilin Province in the central part of northeastern China (Figure 5a,b) (40°52' N–46°18' N, 121°38' E–131°19' E), covering 1.87×10^5 km² (Figure 5c). Moving westward from the eastern part of the province, the terrain comprises the eastern Changbai Mountains, the central plain and platform, and the western alluvial plain (Figure 5c). The region has a temperate continental monsoon climate with four distinct seasons. The average temperature during the year is 3–5 °C, and the average annual precipitation ranges between 550 mm and 910 mm [46]. Jilin Province is currently an important commodity grain base in China. The province has a long history of agriculture, especially during the past 300 years [47]. The main crops cultivated are rice, corn, and sorghum, which mature once a year. They are generally sown in April and harvested from mid-August to September.

Jilin Province has experienced a transformation in its land cover over the last 300 years, changing from a region comprised mostly of wilderness into a mature agricultural region. Especially since the lifting of the policy prohibiting reclamation in 1860, cultivation of cropland has been prominent [41,48]. During the study period, cropland area, as well as the number of settlements, rapidly increased and expanded spatially, as substantial flows of migrants entered the Jilin Province to cultivate cropland and establish new settlements (Figure 6). A significant positive correlation (a correlation coefficient of 0.99) is also evident between the number of settlements and cropland area during the past 300 years. Thus, Jilin Province is an ideal area for testing the feasibility of the settlement density based on the gridded allocation method.

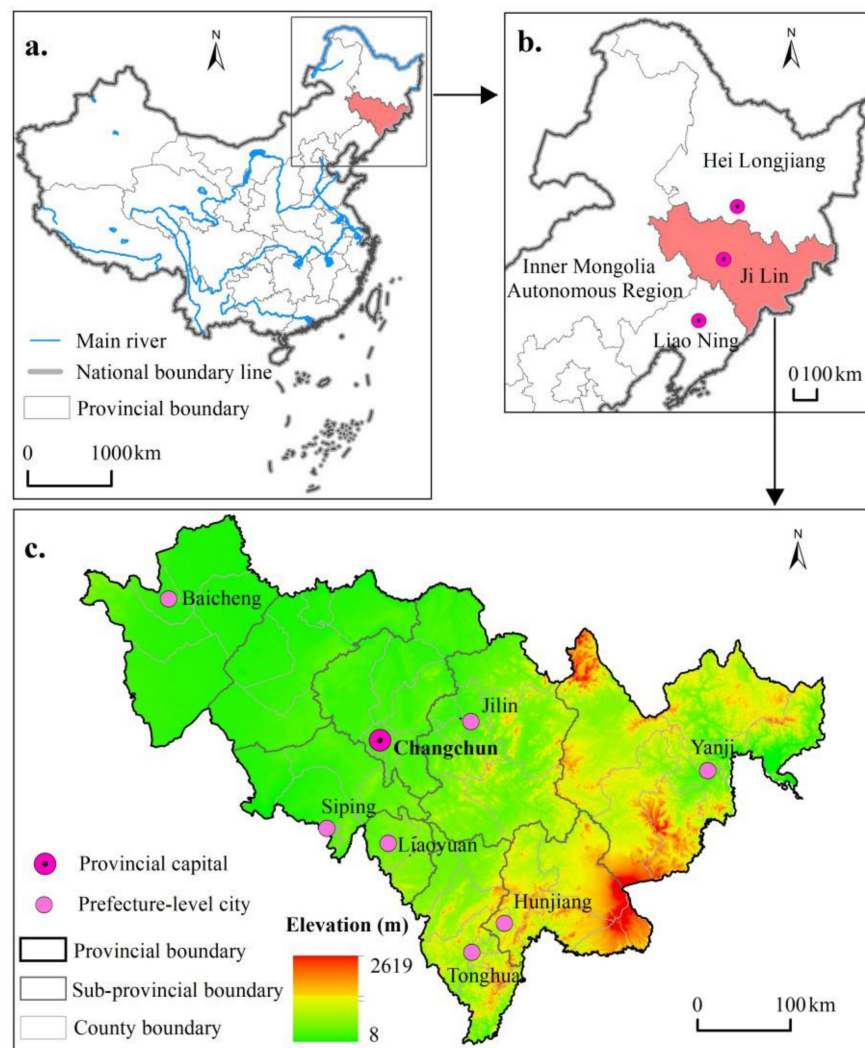


Figure 5. Map of the study area. (a) Geographic location of the study area within China; (b) Geographic location of the study area within northeastern China; (c) Administrative boundaries, cities, and elevations in Jilin Province.

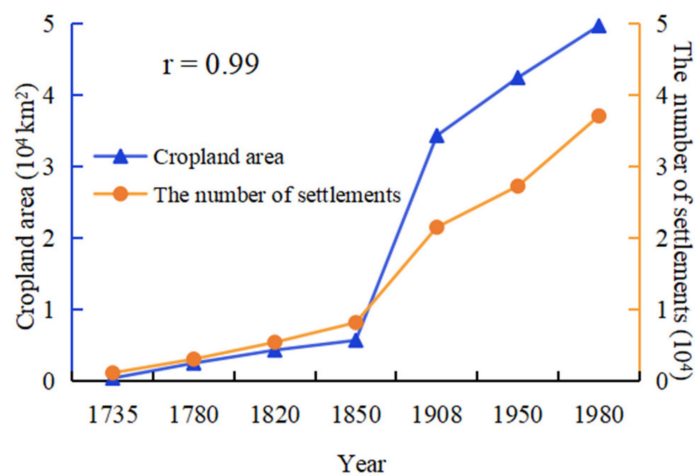


Figure 6. Changes in the cropland area and the number of settlements over the last 300 years in Jilin Province [41,48].

3.2. The Gridded Cropland Cover of Jilin Province over the Last 300 Years Allocated Using the Settlement Density Method

3.2.1. The Source Data Used for the Allocation of Cropland Area

As shown in Table 2, the cropland area in Jilin Province at six time points during the past 300 years were used in this study for allocating cropland. The data on cropland areas for 1735, 1780, 1908, and 1950 at the provincial scale were sourced from Ye et al. [48], while data for 1820 and 1850 were sourced from Fang et al. [49].

Table 2. The data sources, scales, and uses of cropland area in Jilin Province from 1735 to 1950.

Year	Data Source	Administrative Units	Use of Cropland Area in This Study
1735, 1780, 1908, 1950	Ye et al. [48]	Provincial scale County scale	The source data of gridded allocation To assess the accuracy of the allocated results
1820, 1850	Fang et al. [49]	Provincial scale	The source data of gridded allocation

These data on cropland area reconstructed using historical records could objectively reflect the real trend of cropland area and the history of agricultural development in Jilin Province. All of these cropland areas were converted into square kilometers (km²).

3.2.2. The Source Data on Settlements and Its Processing

We used the settlement database on Jilin Province, developed by Zeng et al. [41], in this study. This database includes more than 30,000 settlements in total, with information on the code, name, year of establishment or duration, and other details provided for each settlement. The data were extracted from 21 sets of county-level gazetteers and six sets of district-level gazetteers for Jilin Province and were spatially verified using a topographic map at a scale of 1:50,000 [41]. The settlement data in this database are of high quality, showing spatio-temporal changes in populations and settlements in Jilin Province over the past 300 years.

To match the time periods relating to the cropland area, we extracted settlement data for the years 1735, 1780, 1820, 1850, 1908, 1950, and 1980 from the database.

3.2.3. Allocation of the Cropland Cover of Jilin Province over the Past 300 Years

Using our settlement density method, we respectively allocated cropland area at the provincial scale in Jilin Province to 5' × 5' grid for the years 1735, 1780, 1820, 1850, 1908, and 1950 (Figure 7). The given upper limit of the cropland fraction in the 5' × 5' grid was no greater than 90%, because it is almost impossible for there to be a single dominant land use type. Even in the current era, at least 10–15% of the land is allocated for other land use types [10,50,51]. During historical periods, the cropland fraction in the grids could not have been higher than during the current day.

Figure 7 shows the spatial distribution of cropland cover at a 5' resolution in Jilin Province for six time periods between 1735 and 1950. Overall, it objectively reflects spatial changes in the cropland cover in Jilin Province over the past 300 years. Prior to 1860, the expansion of cropland area and cropland fraction increased slowly; thereafter, rapid increase and spatial expansion of cropland occurred, resulting in the agricultural land cover pattern that currently exists in Jilin Province.

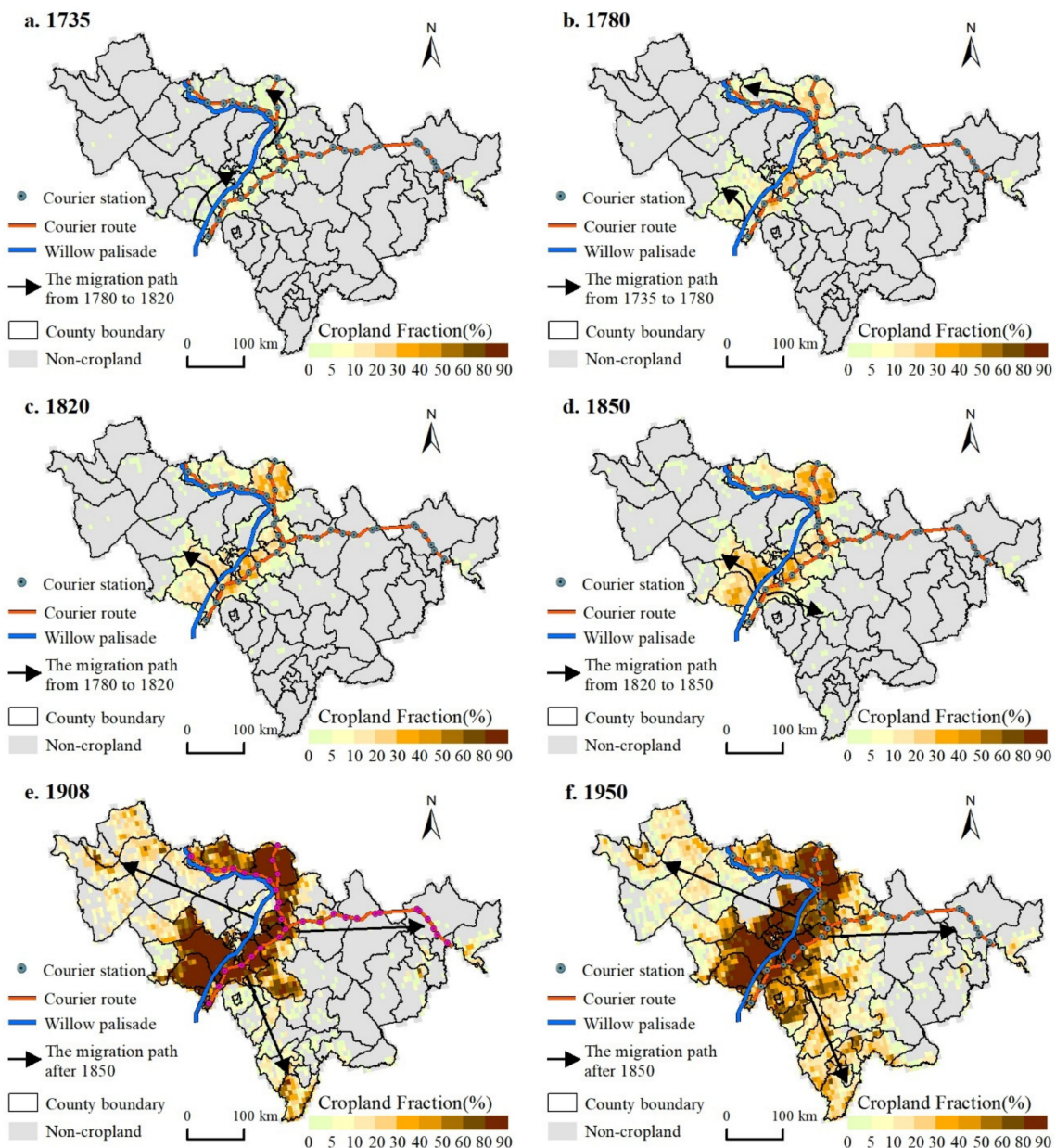


Figure 7. The spatial distribution of 5' × 5' cropland cover in Jilin Province allocated using the settlement density method.

3.3. Credibility Assessment of the Allocations

3.3.1. The Capacity of the Settlement Density Method for Reproducing Modern Cropland Cover

We extracted cropland data pertaining to the 1980s from the China Land Use (CLU) database (at a 1 km² spatial resolution) to obtain a realistic representation of the modern cropland cover of Jilin Province (<http://www.geodata.cn/>, accessed on 15 January 2021). We first extracted data on paddy fields and dry land in Jilin Province from the CLU database and subsequently merged them into cropland data. Next, we calculated the cropland area for each 5' × 5' grid cell. Finally, we calculated the cropland fraction in each 5' × 5' grid cell and the total cropland area of Jilin Province.

We conducted a quantitative assessment of the feasibility of using the settlement density method to reconstruct modern cropland cover.

Figure 8a shows the spatial distribution of cropland cover in 1980 according to the CLU database. Figure 8b shows the settlement-density-based allocation of cropland cover

in 1980. Figure 8c shows the absolute difference between the cropland cover derived from each method in each grid, as does the histogram in Figure 8d. As a final step, we calculated the spatial correlation coefficient between the cropland cover obtained from the CLU database and the allocated cropland cover.

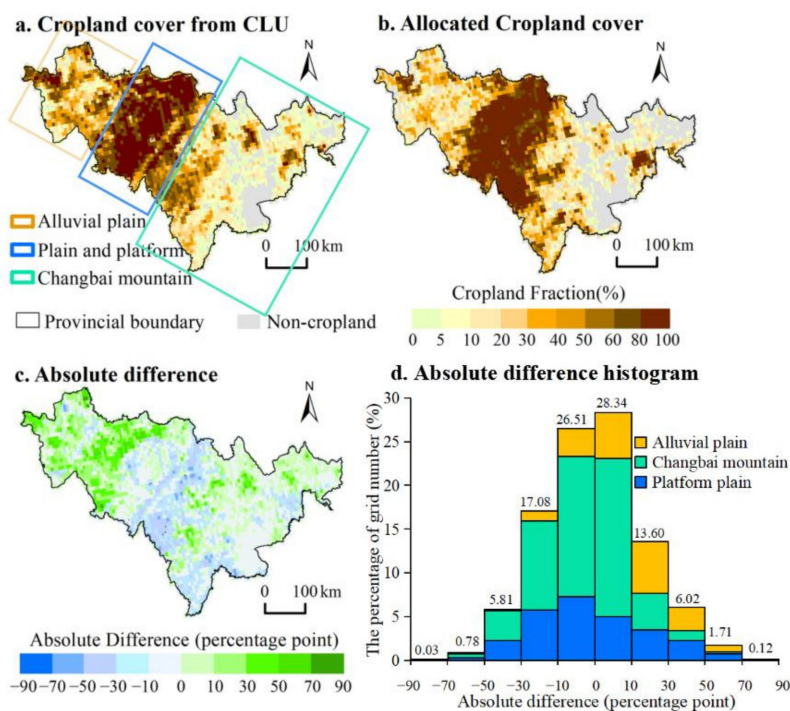


Figure 8. A comparison of cropland cover obtained from the CLU database in 1980 and the cropland cover allocated using the settlement density method.

The cropland cover allocated from the provincial cropland area using the settlement density method was a reasonable reproduction of the spatial distribution of the actual cropland cover in 1980. The correlation coefficient between the allocated cropland cover (Figure 8b) and the cropland cover obtained from the CLU database (Figure 8a) was 0.81, indicating a significantly positive correlation. Moreover, the allocated cropland cover reproduced the main farming region in Jilin Province and its spatial characteristics (Figure 8a,b). Specifically, the cropland grids were contiguously distributed, with a higher cropland fraction (>60%) in the central plain and platform. In the eastern Changbai Mountains, cropland grids with a higher cropland fraction (>50%) were mainly located in the valley basin, presenting a pattern of “string beads” distributed along the river, while those with a higher cropland fraction (>40%) in the western alluvial plain presented a mosaic pattern.

The absolute difference between the two sets of data on cropland cover showed a normal distribution (Figure 8d). The proportion of grids with an absolute difference of −20–20 percentage points was 73.1%, while the proportions of those with an absolute difference of −10–10 percentage points and 0 percentage points were 54.9% and 10.4%, respectively. The proportion of grids for which the absolute difference exceeded 50% was only 2.6%.

3.3.2. The Rationality of the Allocated Results in Relation to Historical Facts Pertaining to This Region

We digitized regional historical information that included temporal and spatial attributes, such as the courier station, courier route and the Willow Palisade, which significantly influenced the historical development of land cover in Jilin Province over the past 300 years. To comprehensively judge the agricultural development of each district

in Jilin Province, we extracted qualitative descriptions of agricultural history from county gazetteers in Jilin Province [52–58].

The distribution of the allocated cropland cover across the six time points was consistent with the history of agricultural development in Jilin Province over the past 300 years (Figure 7). The allocated cropland fractions for this province during the period of 1735–1850 were generally low, with the respective average values for 1735, 1780, 1820, and 1850 being just 0.2%, 1.2%, 2.1%, and 2.8%. Cropland in Jilin Province was mainly concentrated in the central platform and plain, expanding along the courier route from south to north. However, expansion in easterly, southeasterly, and northwesterly directions did not occur. Evidently, the temporal and spatial pattern of allocated cropland cover was highly consistent with regional agricultural development, prior to the lifting of “the policy prohibiting reclamation” in 1860. During this period, the migrants and human activities in Jilin Province were mainly distributed along the courier route and were significantly limited by the Willow Palisade (Figure 7a–d).

Following the lifting of “the policy prohibiting reclamation” in 1860, and especially after the onset of the 20th century, cropland began to expand rapidly from the east and west directions of the courier route (Figure 7e,f). The average cropland fractions in cropland grids were 16.7% and 20.5% in 1908 and 1950, respectively. By 1950, the spatial pattern of cropland cover in Jilin Province (Figure 7f) was very close to matching the pattern of cropland cover in 1980 (Figure 8a).

3.3.3. A Comparison of the Allocated Cropland Area and the Independently Reconstructed Cropland Area Based on County Scale in 1908 and 1950

The accuracy of the allocated $5' \times 5'$ cropland cover in 1908 and 1950 was assessed at the county scale in Jilin Province. The county-level cropland area in the province in 1908 and 1950 was independently reconstructed using historical records sourced from Ye et al. [48] (Table 2). These values were then converted into cropland fractions considered as the “true historical values” in the assessment. First, we calculated the sum of the allocated cropland area obtained using settlement densities within each county of Jilin Province, which we then converted into a cropland fraction at the county scale. Next, we calculated the absolute difference between the two cropland fractions at the county scale (Ye et al. [48] and this study).

As shown in Figure 9, the pattern of our allocated cropland cover and that of Ye et al. [48] at county scale were consistent, and the values of cropland fractions progressively decreased from the central plain and platform to the western alluvial plain, and finally to the eastern Changbai Mountains. In 1908, the cropland fractions of the central plain and platform were above 40%, those of the western alluvial plain were below 10%, and those of the eastern Changbai Mountains were less than 5%. In 1950, the cropland fractions of the central plain and platform exceeded 40%, while those of the western alluvial plain and the eastern Changbai Mountains were below 15% and 5%, respectively.

As shown in Figure 10, the absolute differences in cropland fractions at the county scale between our allocation results and those of Ye et al. [48] in 1908, as well as in 1950, were acceptable, revealing a normal distribution. The counties with absolute differences ranging between -30 and 30 percentage points in 1908 and 1950, respectively, accounted for 78.3% and 87.0% of the total number of counties. Moreover, the proportions of absolute differences above 50 percentage points were 8.8% and 0 in 1908 and 1950, respectively.

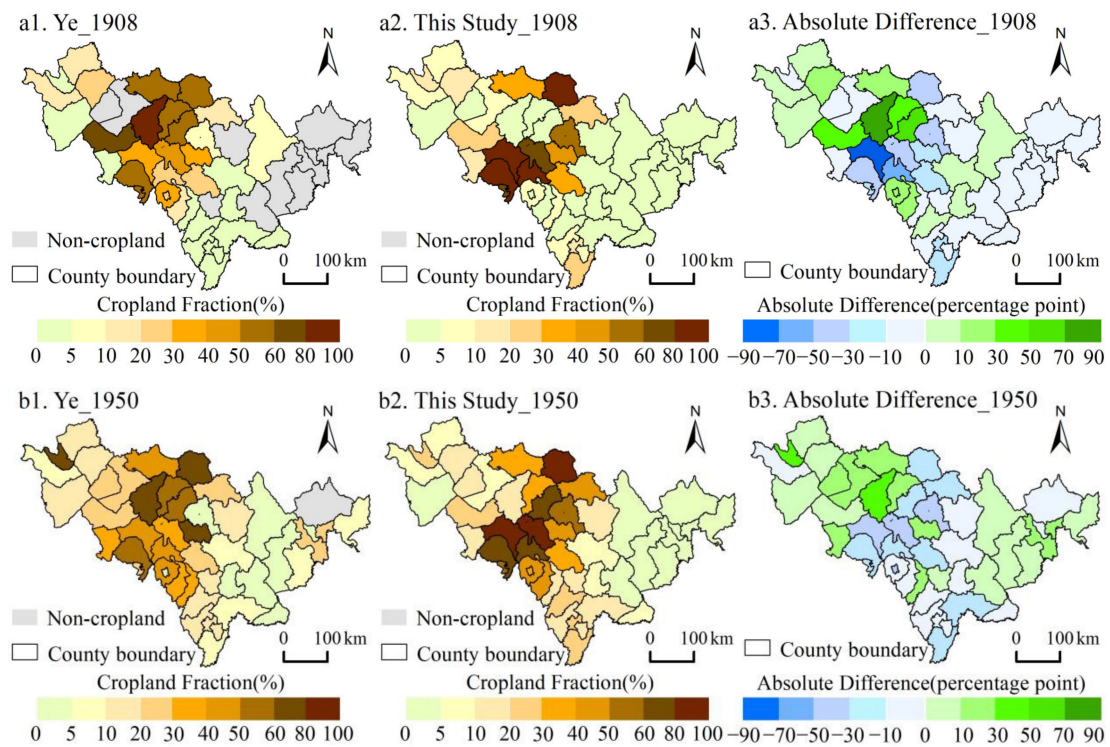


Figure 9. The spatial distribution of cropland fractions derived from Ye et al. [48] (a1,b1) and from this study (a2,b2) at the county scale. The absolute differences between cropland fractions derived from Ye et al. [48] and this study (a3,b3).

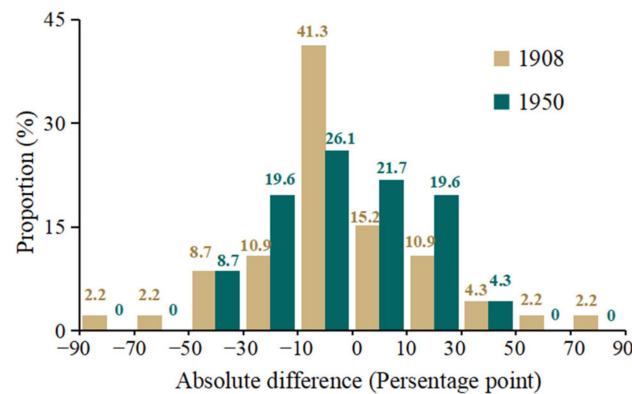


Figure 10. Absolute differences in county-level cropland fractions obtained by Ye et al. [48] and in this study.

4. Discussion: Potential Prospects and Limitations Relating to the Use of Settlement Data

In this study, settlement data from gazetteers were used to develop a historical cropland allocation method, which was successfully applied in the reconstruction of cropland cover over the past 300 years in Jilin Province, China. Archaeological data were also applied as one of the sources of settlement data, used to reconstruct historical cropland cover [37,45]. As places where people gathered together during historical periods, settlement information was often recorded in various historical material. For example, the settlements that were recorded in ancient maps could be processed into our research data through technical means. Abundant settlement data are also available for other regions, apart from China, such as Europe [59,60]. Therefore, this method, which uses settlement data, has considerable potential for application in the reconstruction of historical cropland cover.

4.1. Abundant Sources of Historical Settlement Data

There are abundant records of historical settlements, mainly sourced from gazetteers, ancient maps, and archaeological data (Figure 11). Settlement data from each of these sources have distinctive characteristics. Therefore, the use of settlement data varies and is dependent on practical accessibility, as discussed below.

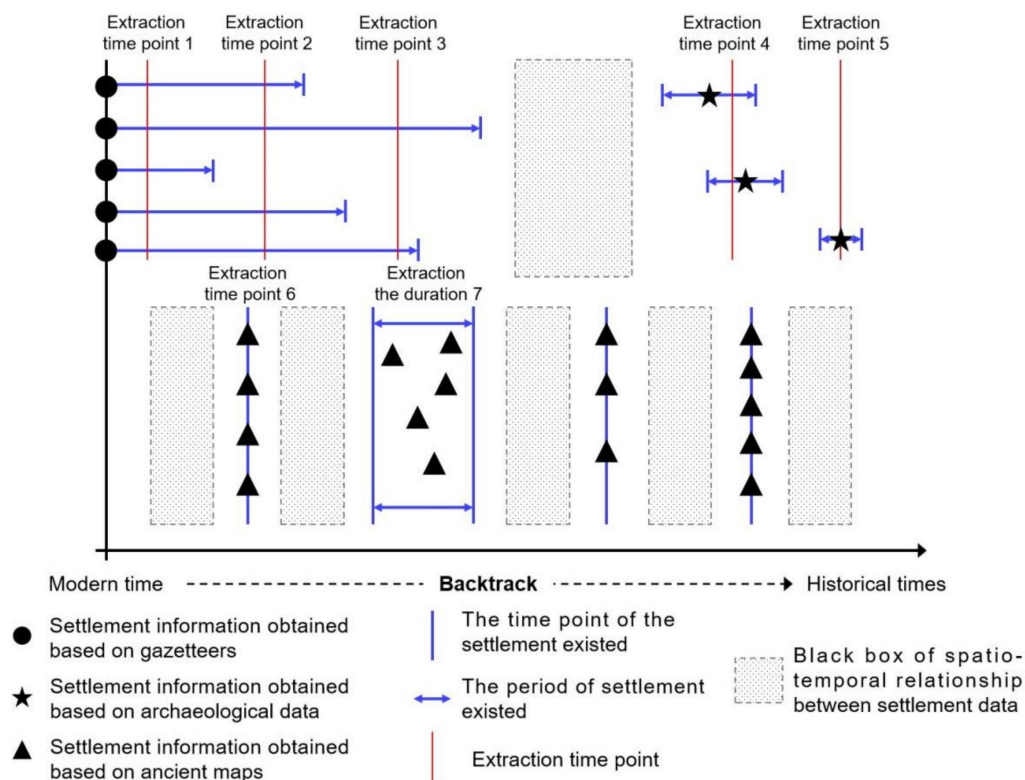


Figure 11. Records and characteristics of the three main kinds of historical settlement information.

Gazetteers, usually in the form of books, contain records of the names, origin, evolution, and geographical locations of modern existing cities, towns, and villages in a particular area. Each settlement has a specific geographical location. The name of a settlement may provide information on the settlement type, physical-geographical environment, and ethnic composition. The recorded history of a settlement could cover the time range, its development, and the changes experienced, which can be used to backtrack the time of its establishment and identify the period of its existence. Settlement data on specific time periods during the settlement’s existence can be extracted. For example, as shown in Figure 11, five, four, and two settlements, respectively, existed during time points 1, 2, and 3.

Ancient maps and historical literature comprise the second key source of settlement data. Settlement records within the historical literature could provide information, such as the name, approximate location, and geographical environment of a settlement. The settlement symbols on ancient maps could provide information, such as the name and specific spatial location of a settlement. The temporal attributes of settlement data obtained from these sources could comprise a time point or duration, which only provides limited temporal information on, for example, the data source. It is unclear whether a relationship of continuation and development exists between the settlements during this time period and those during other time periods. As shown in Figure 11, because of the time limitation relating to the settlement data, only four settlements at time point 6 and five settlements at duration 7 could be identified.

The third source, namely settlement archaeological data mainly comprises information about ancient settlements obtained through archaeological investigations, excavations of these sites and subsequent analysis, which can be accessed from relevant published atlases

containing information on archaeological relics and written reports. According to the location, size, name, and time range of the archaeological settlements, it could be captured on the basic information about the places where human activities had been in the past. Such sources can be used to extract data on settlement sites during a certain time period or time point. As Figure 11 shows, we were able to extract spatial location information for two settlement sites at time point 4 and one settlement site at time point 5, respectively.

4.2. The Limitations of Settlement Data

Although there are abundant sources of settlement data, their use entails some limitations. On the one hand, in some regions, there are missing historical settlement records, or even no records. Therefore, the settlement density method we designed cannot be applied in studies of regions for which limited or no settlement data are available. On the other hand, settlement data generally contain three basic types of information relating to time, space, and settlement size. Under ideal conditions, historical cropland cover can be reconstructed with settlement data in timing, positioning, and quantitative ways. However, in practice, it is not common to completely obtain the three types of settlement data information. In this study, limited by the settlement data we obtained, all settlements are treated as the same size in the actual operation. Although differences caused by varying sizes of settlements were not significant at the spatial resolution of $5' \times 5'$ used for this study, we will incorporate settlement size into the gridded allocation method in future studies.

5. Conclusions

Due the uncertainty of the existing cropland gridded allocation method, which is based on land suitability for cultivation, this study innovatively employed human factors, namely settlement information, to develop a settlement density method. The settlement density method was then applied to reconstructing cropland cover at a $5' \times 5'$ spatial resolution at six time points over the past 300 years in China's Jilin Province, and its credibility was assessed by three methods. Our major conclusions were as follows.

First, being a living space, a rural settlement is always coeval with the cropland cultivated around it. Therefore, the locations of historical settlements could be regarded as an important indicator of previously existing historical cropland. This study designed a method based on settlement density for the gridded allocation of historical cropland. This method is premised on the coexistence of rural settlements and cropland. The cropland was only distributed in the grid where settlements existed and did not consider differences in settlement size. The cropland area of a grid comprised the cropland area per settlement multiplied by the number of settlements within the grid.

Second, in the case of Jilin Province, which underwent a dramatic transformation from a near wilderness state to a mature agricultural region, changes in provincial settlements and the cropland area were significantly positively correlated ($r = 0.99$). All of the assessments using the three methods confirmed the feasibility of the settlement-density-based cropland allocation that we designed. They demonstrated that the cropland cover allocated from the provincial cropland area using the settlement density method reasonably reproduced the actual cropland cover obtained from the CLU database in 1980. We obtained a correlation coefficient of 0.81, revealing a significantly positive correlation between the two sets of cropland cover. Moreover, the distribution of the allocated grid cropland cover was consistent with the history of agricultural development in Jilin Province over the past 300 years. Differences in the cropland area at the county scale between the sum of the gridded reconstruction allocated from the provincial cropland area and the independent reconstructions developed using historical records for the years 1908 and 1950 were also acceptable.

Our case study confirms that the settlement density method that we designed can objectively and credibly reproduce temporal and spatial changes in cropland cover in new reclamation regions, such as Jilin Province, even though the cropland area to be allocated is at a lower spatial resolution, such as the provincial scale.

This research is a valuable attempt to improve the reliability of historical global land cover data by introducing human factors into the allocation method, which provides a possibility for utilizing other human factors, such as tombs and irrigation channels, into the gridded allocation method in the future. However, there are some shortcomings and limitations that need to be addressed in order to improve this method. First, this study restricted the allocation of cropland to a grid that contained the settlements. Thus, our results may be biased. Even these deviations are also acceptable, according to the credibility assessment of our reconstruction using the three methods. In the future, the influence of settlement distance should be fully considered for reconstructing cropland cover at a finer grid scale (e.g., at 1 km × 1 km grid scale). Second, if the study area is large (e.g., at the global scale) or contains complex geographical units, it is necessary to use land suitability for cultivation constructed by natural factors (e.g., topography and soil) to correct the settlement density method. Third, we used the cropland area at the provincial scale with a low resolution. Consequently, the reconstruction results could be further improved by using cropland area at a higher resolution, such as the county scale.

Author Contributions: Conceptualization, Z.W., X.F. and Y.Y.; Methodology, Z.W., X.F. and Y.Y.; Writing—original draft, Z.W.; Writing—review and editing, Z.W., X.F. and Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the National Key R&D Program of China (2017YFA0603304, 2021YFD1500704).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

- Goldewijk, K.K.; Dreht, G.V.; Bouwman, A.F. Mapping contemporary global cropland and grassland distributions on a 5 × 5 minute resolution. *J. Land Use Sci.* **2007**, *3*, 167–190. [[CrossRef](#)]
- Ellis, E.C.; Kaplan, J.O.; Fuller, D.Q.; Vavrus, S.; Goldewijk, K.K. Used planet: A global history. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7978–7985. [[CrossRef](#)] [[PubMed](#)]
- Ellis, E.C.; Beusen, A.H.W.; Goldewijk, K.K. Anthropogenic biomes: 10,000 BCE to 2015 CE. *Land* **2020**, *5*, 129. [[CrossRef](#)]
- Gaillard, M.J.; Morrison, K.D.; Madella, M.; Whitehouse, N. Past land-use and land-cover change: The challenge of quantification at the subcontinental to global scales. *Pages Mag.* **2018**, *26*, 3.
- Ramankutty, N.; Foley, J.A. Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Glob. Biogeochem. Cycle* **1999**, *13*, 997–1027. [[CrossRef](#)]
- Goldewijk, K.K. Estimating global land use change over the past 300 years: The HYDE Database. *Glob. Biogeochem. Cycle* **2001**, *15*, 417–433. [[CrossRef](#)]
- Goldewijk, K.K.; Beusen, A.; Janssen, P. Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. *Holocene* **2010**, *20*, 565–573. [[CrossRef](#)]
- Goldewijk, K.K.; Beusen, A.; van Drecht, G.; de Vos, M. The HYDE 3.1 spatial explicit database of human-induced global land-use change over the past 12,000 years. *Glob. Ecol. Biogeogr.* **2011**, *20*, 73–86. [[CrossRef](#)]
- Goldewijk, K.K.; Beusen, A.; Doelman, J.; Stehfest, E. Anthropogenic land use estimates for the Holocene-HYDE 3.2. *Earth Syst. Sci. Data* **2017**, *9*, 927–953. [[CrossRef](#)]
- Pongratz, J.; Reick, C.; Raddatz, T.; Claussen, M. A reconstruction of global agricultural areas and land cover for the last millennium. *Glob. Biogeochem. Cycle* **2008**, *22*, 1–16. [[CrossRef](#)]
- Kaplan, J.O.; Krumhardt, K.M.; Ellis, E.C.; Ruddiman, W.F.; Lemmen, C.; Goldewijk, K.K. Holocene carbon emissions as a result of anthropogenic land cover change. *Holocene* **2011**, *21*, 775–791. [[CrossRef](#)]
- Strassmann, K.M.; Joos, F.; Fischer, G. Simulating effects of land use changes on carbon fluxes: Past contributions to atmospheric CO₂ increases and future commitments due to losses of terrestrial sink capacity. *Tellus Ser. B Chem. Phys. Meteorol.* **2008**, *60B*, 583–603. [[CrossRef](#)]
- Steyaert, L.T.; Knox, R.G. Reconstructed historical land cover and biophysical parameters for studies of land-atmosphere interactions within the eastern United States. *J. Geophys. Res.* **2008**, *113*, D02101. [[CrossRef](#)]
- Kaplan, J.O.; Krumhardt, K.M.; Zimmermann, N. The prehistoric and preindustrial deforestation of Europe. *Quat. Sci. Rev.* **2009**, *28*, 3016–3034. [[CrossRef](#)]
- Pongratz, J.; Raddatz, T.; Claussen, M. Effects of anthropogenic land cover change on carbon cycle of the last millennium. *Glob. Biogeochem. Cycle* **2009**, *23*, GB4001. [[CrossRef](#)]

16. Pongratz, J.; Raddatz, T.; Reick, C. Radiative forcing from anthropogenic land cover change since A.D. 800. *Geophys. Res. Lett.* **2009**, *36*, L02709. [[CrossRef](#)]
17. Sterling, M.S.; Agnes, D.; Polcher, J. The impact of global land-cover change on the terrestrial water cycle. *Nat. Clim. Change* **2013**, *3*, 385–390. [[CrossRef](#)]
18. Arneeth, A.; Sitch, S.; Pongratz, J.; Stocker, B.D.; Ciais, P.; Poulter, B.; Bayer, A.D.; Bondeau, A.; Calle, L.; Chini, L.P.; et al. Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed. *Nat. Geosci.* **2017**, *10*, 79–85. [[CrossRef](#)]
19. Harrison, S.P.; Gaillard, M.J.; Stocker, B.D.; Vander, L.M.; Goldewijk, K.K.; Boles, O.; Braconnot, P.; Dawson, A.; Fluet-Chouinard, E.; Kaplan, J.O.; et al. Development and testing scenarios for implementing land use and land cover changes during the Holocene in Earth system model experiments. *Geosci. Model Dev.* **2020**, *13*, 805–824. [[CrossRef](#)]
20. Goldewijk, K.K.; Verburg, P.H. Uncertainties in global-scale reconstructions of historical land use: An illustration using the HYDE data set. *Landsc. Ecol.* **2013**, *28*, 861–877. [[CrossRef](#)]
21. Leite, C.C.; Costa, M.H.; Soares-Filho, B.S.; Hissa, L.D.B.V. Historical land use change and associated carbon emissions in Brazil from 1940 to 1995. *Glob. Biogeochem. Cycle* **2012**, *26*, 1–13. [[CrossRef](#)]
22. Yu, Z.; Lu, C.Q. Historical cropland expansion and abandonment in the continental U.S. during 1850 to 2016. *Glob. Ecol. Biogeogr.* **2018**, *27*, 322–333. [[CrossRef](#)]
23. Liu, M.L.; Tian, H.Q. China's Land cover and land use change from 1700 to 2005: Estimations from high-resolution satellite data and historical archives. *Glob. Biogeochem. Cycle* **2010**, *24*, GB3003. [[CrossRef](#)]
24. Zhang, X.Z.; He, F.N.; Li, S.C. Reconstructed cropland in the mid-eleventh century in the traditional agricultural area of China: Implications of comparisons among datasets. *Reg. Environ. Change* **2013**, *13*, 969–977. [[CrossRef](#)]
25. Zumkehr, A.; Campbell, J.E. Historical U.S. cropland areas and the potential for bioenergy production on abandoned croplands. *Environ. Sci. Technol.* **2013**, *47*, 3840–3847. [[CrossRef](#)]
26. Li, M.J.; He, F.N.; Li, S.C.; Yang, F. Reconstruction of the cropland cover changes in eastern China between the 10th century and 13th century using historical documents. *Sci. Rep.* **2018**, *8*, 13352. [[CrossRef](#)]
27. Liu, X.; Li, S.C.; He, F.N.; Hua, L. Reconstruction of cropland areas for South Asia from AD 640 to 2016. *Reg. Environ. Change* **2022**, *22*, 47. [[CrossRef](#)]
28. Zhao, C.S.; He, F.N.; Yang, F.; Li, S.C. Uncertainties of global historical land use scenarios in past-millennium cropland reconstruction in China. *Quat. Int.* **2022**, *in press*. [[CrossRef](#)]
29. Li, S.C.; He, F.N.; Zhang, X.Z.; Zhou, T.Y. Evaluation of global historical land use scenarios based on regional datasets on the Qinghai-Tibet Area. *Sci. Total Environ.* **2018**, *657*, 1615–1628. [[CrossRef](#)]
30. Fang, X.Q.; Zhao, W.Y.; Zhang, C.P.; Zhang, D.Y.; Wei, X.Q.; Ye, Y. Methodology for credibility assessment of historical global LUCC datasets. *Sci. China Earth Sci.* **2020**, *63*, 1013–1025. [[CrossRef](#)]
31. Lin, S.S.; Zheng, J.Y.; He, F.N. Gridding cropland data reconstruction over the agricultural region of China in 1820. *J. Geogr. Sci.* **2009**, *19*, 36–48. [[CrossRef](#)]
32. Li, K.; He, F.N.; Zhang, X.Z. An approach to reconstruction spatial distribution of historical cropland with grid-boxes by utilizing MODIS land cover datasets: A case study of Yunnan Province in the Qing Dynasty. *Geogr. Res.* **2011**, *30*, 2281–2288.
33. Luo, J.; Zhang, Y.L.; Liu, F.G.; Chen, Q.; Zhou, Q.; Zhang, H.F. Reconstruction of cropland spatial patterns for 1726 on Yellow River-Huangshui River Valley in northeast Qinghai-Tibet Plateau. *Geogr. Res.* **2014**, *33*, 1285–1296.
34. Li, S.C.; He, F.N.; Zhang, X.Z. A spatially explicit reconstruction of cropland cover in China from 1661 to 1996. *Reg. Environ. Change* **2016**, *16*, 417–428. [[CrossRef](#)]
35. Yang, X.H.; Jin, X.B.; Du, X.D.; Xiang, X.M.; Han, J.; Shan, W.; Fan, Y.T.; Zhou, Y.K. Multi-agent model-based historical cropland spatial pattern reconstruction for 1661–1952, Shandong Province, China. *Glob. Planet. Change* **2016**, *143*, 175–188. [[CrossRef](#)]
36. Cao, B.W.; Yu, L.; Li, X.C.; Chen, M.; Li, X.; Hao, P.Y.; Gong, P. A 1km global cropland dataset from 10000 BCE to 2100 CE. *Earth Syst. Sci. Data* **2021**, *13*, 5403–5421. [[CrossRef](#)]
37. Wu, Z.L.; Fang, X.Q.; Jia, D.; Zhao, W.Y. Reconstruction of Cropland Cover Using Historical Literature and Settlement Relics in Farming Areas of Shangjing Dao during the Liao Dynasty around 1100 AD. *Holocene* **2020**, *30*, 1516–1527. [[CrossRef](#)]
38. Darby, H.C. *A New Historical Geography of England before 1600*; Cambridge University Press: London, UK, 1976.
39. Darby, H.C. *A New Historical Geography of England after 1600*; Cambridge University Press: London, UK, 1976.
40. Entwisle, B.; Rindfuss, R.R.; Walsh, S.J.; Page, P.H. Population growth and its spatial distribution as factors in the deforestation of Nang Rong, Thailand. *Geoforum* **2008**, *39*, 879–897. [[CrossRef](#)]
41. Zeng, Z.Z.; Fang, X.Q.; Ye, Y. The process of land cultivation based on settlement names in Jilin Province in the past 300 years. *Acta Geogr. Sin.* **2011**, *66*, 985–993.
42. Han, M.L. Distribution of settlements and environmental choices in the Xila Mulun River Valley and its neighborhood during the Liao Dynasty. *Acta Geogr. Sin.* **2004**, *59*, 543–549.
43. Huo, R.L.; Yang, Y.D.; Man, Z.M. Gridded reconstruction of spatiotemporal evolution of cropland in the Zhangjiu River Basin in Yunnan Province from 1700 to 1987. *Acta Geogr. Sin.* **2020**, *9*, 1966–1982.
44. Li, Y.K.; Ye, Y.; Fang, X.Q.; Liu, Y.C. Reconstruction of agriculture-driven deforestation in Western Hunan Province of China during the 18th century. *Land* **2022**, *11*, 181. [[CrossRef](#)]
45. Yu, Y.Y.; Wu, H.B.; Finke, P.A.; Guo, Z.T. Spatial and temporal changes of prehistoric human land use in the Wei River valley, northern China. *Holocene* **2016**, *26*, 1788–1801. [[CrossRef](#)]

46. Jilin Provincial Local Records Compilation Committee. *Jilin Provincial Records: Physical Geograph*; Jilin People's Publishing House: Changchun, China, 1992.
47. Jilin Provincial Local Records Compilation Committee. *Jilin Provincial Records: Agriculture Records—Relationship of Rural Production*; Jilin People's Publishing House: Changchun, China, 1999.
48. Ye, Y.; Fang, X.Q.; Ren, Y.Y.; Zhang, X.Z.; Chen, L. Changes of cropland cover in Northeast China in the past 300 years. *Sci. China D Ser.* **2009**, *39*, 340–350. [[CrossRef](#)]
49. Fang, X.Q.; He, F.N.; Wu, Z.L.; Zheng, J.Y. General characteristics of the agricultural area and fractional cropland cover changes in China for the past 2000 years. *Acta Geogr. Sin.* **2021**, *76*, 1732–1746.
50. Li, H.G.; Yue, T.X. An application of Geo-Informatic graphic analysis (Tupu) in modelling regional sustainable development. *Geo-Inf. Sci.* **2000**, *1*, 48–52.
51. Li, S.C.; He, F.N.; Chen, Y.S. Gridding Reconstruction of cropland spatial pattern in southwest China in the Qing Dynasty. *Prog. Geogr.* **2012**, *31*, 1196–1203.
52. Antu County Local Records Compilation Committee. *Antu County Records*; Jilin Wen Shi Publishing House: Changchun, China, 1993.
53. Dehui County (City) Local Records Compilation Committee. *Dehui County Records*; Changchun Publishing House: Changchun, China, 2001.
54. Fusong County Local Records Compilation Committee. Fusong Archives. In *Fusong County Records*; Jilin University Publishing House: Changchun, China, 2009.
55. Tumen City of Jilin Province Local Records Compilation Committee. *Tumen City Records*; Jilin Wen Shi Publishing House: Changchun, China, 2006.
56. Longjing County Local Records Compilation Committee. *Longjing County Records*; Northeast Korea National Education Press: Yanji, China, 1989.
57. Nongan County Local Records Compilation Committee. *Nongan County Records*; Jilin Wen Shi Publishing House: Changchun, China, 1993.
58. Qianan County Local Records Compilation Committee. *Qianan County Records*; Jilin People's Publishing House: Changchun, China, 2008.
59. Mead, W.R. An atlas of settlement in sixteenth-century Finland. *J. Hist. Geogr.* **1975**, *1*, 17–20. [[CrossRef](#)]
60. Darby, H.C.; Versey, G.R. *Domesday Gazetteer*; Cambridge University Press: London, UK, 1988.