SCIENCE CHINA
Earth Sciences

•ARTICLE•

September 2024 Vol.67 No.9: 2953–2971
https://doi.org/10.1007/s11430-023-1381-0

Spatially explicit reconstruction of cropland cover in Europe from AD 1800 to 2000

Jun LI[1], Yu YE[1,2*], Xiuqi FANG[1,2], Diyang ZHANG[1] & Zhilei WU[3]

[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] College of Geographical Science, Qinghai Normal University, Xining 810016, China

Received July 15, 2023; revised May 31, 2024; accepted July 12, 2024; published online August 6, 2024

Abstract One of the challenges in global change research is the significant uncertainty in global historical land use and land cover (LUCC) datasets, which are widely used as foundational data. In addition to the regional cropland area reconstructions, improving the grid allocation method is another feasible way to raise the reliability of historical LUCC data. In this study, an integrated reconstruction of the national cropland areas over the past 200 years was developed for 36 European countries. After that, the allocation algorithm was built using physiogeographic variables and historical city sites for accounting for land suitability and cultivation preferences, respectively. Finally, cropland data in Europe with a spatial resolution of 5′×5′ at five time sections from AD 1800 to 2000 were generated using the optimal allocation algorithm in accordance with the stages of the regional history. The results were as follows: (1) The dominant factors governing the distribution of croplands in Europe vary at different agricultural stages, but the results can be merged together. Land suitability was more optimal for allocation during the modern agricultural stage (AD 1950 and 2000); the priority index combined with land suitability and cultivation preference was more reasonable for allocation during the traditional agricultural stage (AD 1800). The average of the allocations by priority index and the land suitability could be adopted as the allocation results during the transitional stage (AD 1850 and 1900) because the grids for absolute differences within ±10 and ±20 percentage points between the results obtained from the above two allocations were above 80% and 95%, respectively, which means the two allocation results could be merged. (2) Over the past 200 years, the total cropland area in Europe first increased to a peak in AD 1900 and then decreased. Spatially, the centre of the higher cropland fraction shifted from the western part of Europe in AD 1800 to the eastern part of the continent after AD 1950. (3) Both the cropland area and the spatial distribution in this study are more reasonable than the global dataset HYDE3.2.

Keywords Land use and land cover change, Historical cropland, Gridding allocation, Europe, Past 200 years

Citation: Li J, Ye Y, Fang X, Zhang D, Wu Z. 2024. Spatially explicit reconstruction of cropland cover in Europe from AD 1800 to 2000. Science China Earth Sciences, 67(9): 2953–2971, https://doi.org/10.1007/s11430-023-1381-0

1. Introduction

Land use and land cover change (LUCC), through influencing biogeophysical and biogeochemical processes can change the climate, which in turn affects biodiversity and ecosystem services (Gaillard et al., 2015; Bai et al., 2017; Marta et al., 2021). Several global historical cropland datasets (Pongratz et al., 2008; Kaplan et al., 2009; Ramankutty and Foley, 2010; Klein Goldewijk et al., 2017a) have been developed to serve global change research, such as input parameters for regional and global climate modelling studies (Kaplan et al., 2011; Brovkin et al., 2013). Currently, available spatially explicit global historical cropland datasets are usually developed by allocating quantitatively reconstructed regional cropland areas (generally on a continental or country scale) to grid units based on an allocation

* Corresponding author (email: yeyuleaffish@bnu.edu.cn)

algorithm constructed according to physiogeographic and socio-economic factors. However, the available global historical LUCC datasets, both in the gridded allocation algorithms and the reconstructed historical cropland areas, remain uncertain, as evidenced by the inconsistency of the global datasets in terms of changes in the amount and distribution of cropland, and numerous studies conducted in regions of varying scales, including China, Europe, Scandinavia, and Germany (Li et al., 2013; Kaplan et al., 2017; Wu et al., 2020; Fang et al., 2020; Wei et al., 2021; Zhang et al., 2021; Zhao et al., 2022).

Improving the credibility of global historical cropland data relies on two fundamental aspects. One is to increase the historical record-based regional reconstruction of cropland areas, and the other is to improve the grid allocation method by building more suitable algorithms that are consistent with the natural and socio-economic features of the region. Through more efficient use of historical empirical data and proxies to directly adopt or indirectly estimate the regional cropland area, regional reconstructions can improve the accuracy of historical cropland data by reducing the uncertainty of the data source (Klein Goldewijk et al., 2017b; Fang et al., 2020; Zhang et al., 2021; He et al., 2023). For example, the History Database of the Global Environment (abbreviation HYDE) has improved its cropland data credibility by steadily increasing the adoption of regional cropland data since its version HYDE3.0 (Ye et al., 2019). Compared with HYDE2.0, where cropland area data were predominantly at the continental scale, except for a few countries, the cropland data in versions from HYDE3.0 to HYDE3.2 have witnessed a remarkable improvement in credibility for countries that have adopted regional cropland reconstruction data at national or even sub-national scales, such as China and Germany, where the errors of cropland distribution in the grid have been confined within the national units (He et al., 2012; Fang et al., 2020; Zhang et al., 2021; Zhao et al., 2022).

A well-designed cropland allocation method should satisfy three key requirements. Where the grids are potentially suitable for cultivation, which is associated with land suitability; when the grids were cultivated, which is associated with the chronological order of cropland expansion; and how many croplands have been cultivated in each grid, that is the amount or fraction of cropland finally allocated to the grid, which is the combined result of land suitability and the chronological order of cultivation.

Land suitability, which indicates the constraints of natural conditions, is revealed by natural factors. Most existing algorithms for land suitability are generally calculated using subjectively selected physiogeographic factors related to cultivation. Land suitability is weighted by multiplying the normalized values of the factors with equal weight, assuming a universal linear relationship between each factor and cultivation intensity within the research region. For example, the HYDE dataset assumed that grids with higher temperatures, lower slopes, and closer proximity to rivers or coasts had higher suitability (Klein Goldewijk et al., 2017a). However, because of the regional differentiation in the natural environment, it is difficult for one physiogeographic factor to affect land suitability with the same rules and degrees all around the world. The influence of physiogeographic factors on land suitability varies from region to region in terms of the numbers, types, and combinations of these factors. This means that there is considerable irrationality in constructing land suitability uniformly on the global and continental scales. Therefore, when doing large-scale allocation, the land suitability-based allocation model should be focused on regional variations in dominant factors (Zhang et al., 2022a). For example, Feng et al. (2014) divided China into the traditional agricultural region, Northeast China, Northwest China, and Tibetan Plateau, and constructed allocation models for each region by selecting the dominant factors influencing cropland distribution within each region. He et al. (2023) calculated the suitability of a provincial unit and constructed regional allocation models of China for eastern China, Xinjiang, and the Tibetan Plateau, respectively, plus limiting cropland ranges by factors such as historical coastlines, northern farming-pastoral ecotone, and military-oriented reclamation area. Zhang et al. (2022b) constructed a land suitability model that was comparable in different regions of the world using the selected dominant natural factors influencing cropland distribution in a 0.5°×0.5° unit.

The chronological order of cultivation indicates the successiveness of cropland expansion in space and the increase in the cropland fraction in a region. Land suitability just provides the possibility of cultivation. Whether there is real cultivation depends on the selection from the possibility by the people named cultivation preference. Population and settlement data are often used to characterise cultivation preference in existing allocation models. Population density was used by Klein Goldewijk et al. (2017a) and Feng et al. (2014) as the weight factor. The larger the grid population density, the more cropland was allocated to the grid. Huo et al. (2020) delimited potential cropland distribution around the settlements. Equidistant decay was used within the distribution radius, with the settlement serving as its centre. The likelihood of cultivation decreased with distance from the settlement. Wu et al. (2022) and Jia et al. (2023) constructed settlement density-based grid allocation models. Zhang et al. (2023) used a kernel density algorithm to quantify the decrease in cropland distribution with distance within an indicated radius of a city.

The cropland area or fraction in each grid at a given time is the combined effect of land suitability and cultivation preference. However, the contributions of these factors varied across the different periods. Generally, cultivation preference exerts a greater influence during periods of expansion. Once a mature agricultural area is formed, the final amount of cropland in each grid is closely related to its suitability. However, such temporal differences are considered less by existing models that employ a uniform algorithm for allocation across both global and historical stages. For instance, the four widely used global historical LUCC datasets, HYDE, SAGE, KK10, and PJ (Pongratz et al., 2008; Kaplan et al., 2009; Ramankutty and Foley, 2010; Klein Goldewijk et al., 2017a), estimated the total cropland area from the historical population, using population as the primary driving factor based on contemporary land use patterns. They allocated cropland using population density as the main indicator, combined with factors such as slope and distance to water bodies or potential vegetation maps/modern agricultural and pastoral boundaries. However, given the distinct differences in social regimes, economic development paths at different historical stages, and natural conditions across different countries and regions, land use patterns have significant spatiotemporal variations. It is difficult for a singular allocation method to capture the spatiotemporal characteristics of land use changes. Therefore, a feasible way to enhance gridding allocation techniques and boost the reliability of historical cropland data is to build models that are appropriate to given regions and periods based on spatiotemporal variations in land use patterns and dominant factors. This is also the scientific issue to replied in this study.

Over the past two centuries, Europe has transformed land use paradigms, shifting from traditional to modern agricultural practices. Different land use patterns at different stages should correspond to different optimal allocation models, which can provide a typical case for how to merge cropland data obtained using different allocation methods. Taking Europe as the study area (Figure 1), this paper reconstructed a 5′×5′ resolution cropland cover of Europe over the past 200 years, which reconstructed the national cropland areas for European countries by supplementing, correcting, and integrating the existing national-scale historical cropland data at first, then constructed an allocation model combining land suitability and cultivation preferences to allocate national cropland area to grid units. Specially, the Kaliningrad Oblast on the Baltic Sea coast, a Russian enclave, has a close relationship with Lithuania in history. Therefore, when allocating cropland, it is incorporated into Lithuania for calculation in this study.

2. Data sources and methods

The research for this paper involved the following steps: First, national cropland area data over a time interval of 50 years, from AD 1800 to 2000, were obtained by making partial modifications to the national cropland area of Europe for the years 1800, 1850, 1900, and 2000, as generated by Ye et al. (2023), and newly reconstructing the national cropland area for AD 1950 by adopting statistical data and published research results by other researchers. Then, considering the data availability of physiogeographic factors and historical settlement, gridding allocation models varying with region and time were constructed to allocate the national cropland area for 36 European countries into grids with a spatial resolution of 5′×5′. Finally, the changes in cropland in Europe over the past 200 years were analysed. A flowchart of this process is shown in Figure 2.

2.1 National cropland area

2.1.1 Data sources for national cropland area

The national cropland areas for AD 1800, 1850, 1900, and 2000 were cited from Ye et al. (2023), but corrections were made to some countries. Based on historical cropland records and proxy data from the FAO, historic statistics, and published research results, Ye et al. (2023) reconstructed the national cropland area of 37 European countries for 4 time slices over the past 200 years, by adopting historical empirical data for 100%, 78%, and 57% of the European countries in AD 1900, 1850, and 1800, respectively. It is much higher than HYDE3.2 (Klein Goldewijk et al., 2017a), which adopted historical empirical data for only 32% of all European countries. Using newly collected historical records, this study corrected the cropland areas of Albania, Belgium, and Portugal reconstructed by Ye et al. (2023). Meanwhile, the cropland areas of European countries in AD 1950 were reconstructed in this study so that the national cropland data with a time resolution of 50 years from AD 1800 to 2000 was obtained. Please refer to Appendix (https://link.springer.com) for the data sources for each country.

2.1.2 Reconstruction of national cropland area in AD 1950

The data sources used for the reconstruction in AD 1950 are shown in the Appendix. The reconstruction methods used can be classified into two categories. The first category was the record related to the cropland area, which was adopted directly after conversion or revision, so that the definition of cropland was consistent (in line with the FAO[1)] definition of "cropland"). These can be divided into three subcategories according to the type of record used. (1) Data on cropland area or fraction around AD 1950, including 26 countries such as Albania and Bulgaria. (2) The cropland area at a certain time point and the given changing rate of the cropland area

1) FAO data are from https://www.fao.org/faostat/en/#home.

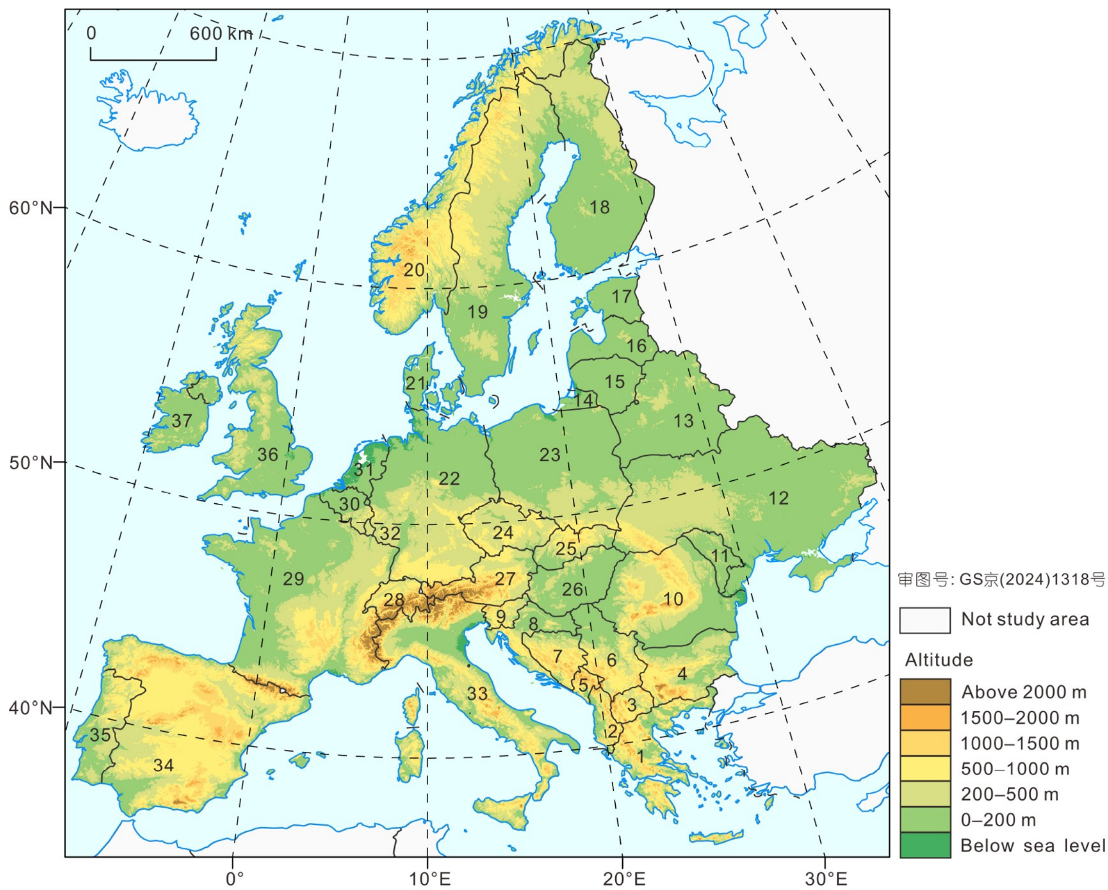

Figure 1 The European countries in this study.

between that time point and AD 1950, including Bosnia and Herzegovina, Greece, Belgium, and the Netherlands. (3) Data on cropland areas that need to be revised, including constituent republics once belonging to former Yugoslavia, except Bosnia and Herzegovina. For the third case, the relevant available data included the total cropland area of Yugoslavia in AD 1961 from the FAO and agricultural census data by country in AD 1960 (Socijalistička Federatina Republika Jugoslavija Savezni Zavod za Statistiku, 1963). The census data for each country is lower than the real values because of concealment (Socijalistička Federatina Republika Jugoslavija Savezni Zavod za Statistiku, 1963). Considering the similarities between political regimes and national identities, the proportion of underreporting was assumed to be consistent across countries. Thus, the total cropland area of Yugoslavia in AD 1961 was downscaled to current individual countries using the proportion of cropland in each country to the total cropland in Yugoslavia in AD 1960. Since the amount of cropland fluctuated in Yugoslavia following the agriculture recovery and reform after the Second World War, simple linear interpolation of cropland areas in the countries is not suitable (Hoffman, 1959; Markovic, 2007). The revised cropland area of the countries in AD 1961 was used directly as the value for AD 1950 in order to eliminate the uncertainty caused by interpolation. The reconstructed cropland fraction for Yugoslavia in AD 1950 differed from the recorded cropland fraction in AD 1953 in the statistical yearbook (Hoffman, 1959) by 2.89 percentage points. This proved the viability of this revision, with the error constrained within the constituent republics of Yugoslavia.

The second category was a lack of records on cropland areas, which have been estimated by interpolation. Slovakia is such a case, whose cropland area in AD 1950 was estimated analogously with the contemporaneous cropland area of Czechia. From AD 1918 to 1992, Czechia and Slovakia were part of the Czechoslovak Republic. Both countries experienced agricultural intensification in the latter half of the 20th century (Jepsen et al., 2015). The socio-economic, demographic, and urban structures of Czechia and Slovakia were very similar till the early 1990s, due to the enhancing technical and economic integration, the unified economic

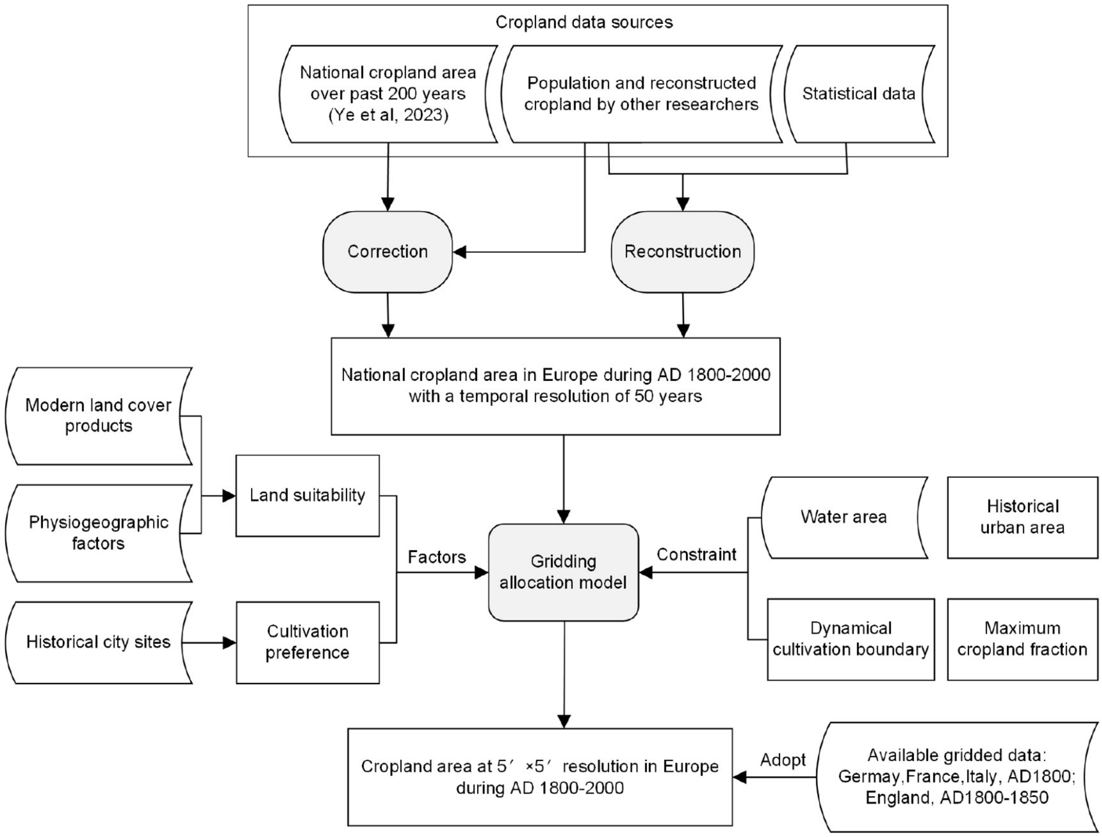

Figure 2 Flowchart of the methodology.

policies, and the adoption of the Soviet Union model after the Second World War (Musil, 1993). Assuming that the proportion of the cropland area between Czechia and Slovakia in AD 1993 remained constant throughout the second half of the 20th century, the cropland area of Slovakia in AD 1950 was calculated using the AD 1950 cropland area of Czechia. This method estimated the cropland fraction in Slovakia in AD 1950 to be 39%. Using the cropland fraction of Slovakia collected since AD 1988 (Izakovičová et al., 2022), the cropland fraction of Slovakia in AD 1950 was 34%, roughly estimated by linear backward regression since the continuous intensification of the economy after the Second World War. This is 5 percentage points different from the fraction of 39% estimated in this study, which proves the feasibility of this estimation method.

2.1.3 Correction of the historical cropland areas in Albania, Portugal, and Belgium

For the national cropland area reconstructed by Ye et al. (2023), we corrected the data for Albania, Portugal, and Belgium by the newly collected relevant research results (Table 1). Ye et al. (2023) pointed out some uncertainties about their reconstructed data in Albania and Portugal. We corrected cropland areas in Albania and Portugal based on the statistical data and created a new proxy for the two countries. For Belgium, the arable land area since the second half of the 19th century, as provided by Dejongh and Vanhaute (1999), was used to correct the data in Ye et al. (2023).

2.2 Gridding allocation model

A gridding allocation model varying with time and space was constructed by integrating land suitability and cultivation preference. Appropriate indicators were selected to measure land suitability and cultivation preference, considering the spatiotemporal variations in the natural environment, human activities, and land use patterns, the heterogeneity of dominant natural factors, and the accessibility of historical human factor records. Among these, land suitability is used to indicate the differences in cropland distribution among grids owing to the heterogeneity in natural conditions and the maximum cultivation intensity within the grid. The cultivation preference is used to identify when the grid should have been tilled and the intensity of human activity within the grid. Furthermore, we constructed a priority index to represent the possibility of being selected for cultivation by integrating land suitability and cultivation preference. The higher the priority index of a grid, the higher the weight of cropland allocation, and the more cropland areas can be allocated to it. In addition, it is necessary, according to the

Table 1 Methods and data sources for correcting the cropland area of Albania, Portugal, and Belgium

Countries	Year (AD)	Cropland fraction (%)		Methods & sources	
		Ye et al. (2023)	This study	Ye et al. (2023)	This study
Albania	1900	17.10	5.95	Bibliographisches Institut, 1909	Jepsen et al., 2015; Lushaj, 2021
	1850	10.69	4.68	Use cropland area per capita in AD 1900	Use growth rate of the land-owning peasant numbers
	1800	8.55	3.75	As above	Use cropland area per capita in AD 1850
Portugal	1900	26.69	49.00	Bibliographisches Institut, 1909	Definition correction; Jones et al., 2011
	1850	20.47	40.95	Portugal Instituto Nacional de Estatística, 1951	
Belgium	1850	38.57	44.08	Use the change rate of cropland fraction in Germany	Use the growth rate of arable land area
	1800	36.00	41.16	As above	Use the change rate of cropland fraction in Germany

natural environment and developmental history, to set a flexible upper-limit area allocated to the grid and to take some additional constraint conditions into consideration for grid cultivation thresholds, including setting the dynamic extent of cultivation and removing non-arable areas like water bodies and historic urban areas.

2.2.1 Data sources for allocation

The data used for the grid allocation can be divided into three categories (Table 2). These include modern land cover data and physiogeographic factors used for constructing land suitability, historical city information used for constructing cultivation preference, water body data and other data used to reconstruct historical built-up areas for restricting the extent of cultivation. The vector city data provide information on the location and rank of city sites for Europe from AD 1–2000, at a temporal resolution of 100 years. The relationships between the rankings of cities and their corresponding populations are shown in Table 3.

2.2.2 Data preparation for model factors

(1) Land suitability data based on physiogeographic factors

This study adopted the method of constructing land suitability based on physiogeographic factors by Zhang et al. (2022a). The correlation between modern cultivation intensity and physiogeographic factors was identified for each 0.5°×0.5° grid cell. Then, 5′×5′ global land suitability data was developed based on the integration of all identified factors that were significantly correlated with cultivation intensity.

Considering the cropland cultivation history of Europe, this study modified the assumption for calculating suitability based on the extent of modern cropland distribution and cultivation intensity by Zhang et al. (2022a). In this study, part of the modern grassland was adopted as the distribution of potential cropland for calculating land suitability. In the second half of the 19th century, most Western European countries began to transition from traditional to intensive agriculture. At the same time, due to the large amount of cheap grain from North America and Russia imported into the European market, many countries had shifted from food production-based mixed agriculture to specialised husbandry production, which resulted in the conversion of cropland into the pasture. After the Second World War, Europe experienced sharp intensification and mechanisation, which resulted in agricultural landscapes becoming more extensive and denser farmland. As a result, less profitable farms were shut down, and some croplands were converted to grasslands and woodlands (Jepsen et al., 2015; Bucała-Hrabia, 2017; Culbert et al., 2017). A policy of reverting farmland to forests and grasslands has been in place continuously since the 1970s (Jepsen et al., 2015). These changes suggest that grassland in modern times might have been farmland in the past. Therefore, we adopted grassland as part of the distribution of potential cropland, and calculated land suitability in Europe using the sum of modern cropland and grassland fractions to correlate with physiogeographic factors in 0.5°×0.5° grid cells (Figure 3).

(2) City-site-based cultivation preference

Cultivation preference represents people's attitudes and choices to reclaim a certain land. Cropland as a production space is always accompanied by settlements as a living space. Therefore, settlement sites can be used as indicators of the real existence of croplands during historical periods. This study adopted the method proposed by Zhang et al. (2023), in which the cultivation preference was measured by the attraction radius of a city to the surrounding rural settlements. It was assumed that within the scope of urban attraction, the farther the grid was from the city, the lower the cultivation preference, and the less cropland was allocated.

Table 2 Data sources used for cropland grid allocation in Europe during 1800–2000

Type			Year	Spatial resolution	Data sources
Data for land suitability	Modern land cover products	Cropland	AD 2000	1 km×1 km	Zhang et al., 2019
		Grassland	AD 2000	1 km×1 km	Zheng, 2022
	Physioge-ographic factors	Climate, Terrain, soil variables, Normalized difference vegetation index	–	–	Zhang et al., 2022b
Data for cultivation preference	Historical city		AD 1–2000	Settlement scale	https://www.euratlas.net/shop/maps_gis/index.html
Data for cropland distribution range	Historical urban area	Historical urban area, population, urbanization rate, expansion rate of city walls	Past 1000 years	Settlement scale; national scale	McEvedy and Jones, 1978; Allen, 2000; Hohenberg and Lees, 1995; Clark, 2009; Cesaretti et al., 2016; Zhang et al., 2023; https://www.euratlas.net/shop/maps_gis/index.html; http://demographia.com/; https://www.londononline.co.uk/factfile/historical/
	Water body		AD 2000	300 m×300 m	Bontemps et al., 2013

Table 3 The relationship between the city rank and the corresponding population in Europe

City rank	Population range	Population ratio	Period
4	≥150000	300	After AD 1750
	≥50000	100	Before AD 1750
3	5000–50000	55	–
2	1000–5000	6	–
1	≤1000	1	–

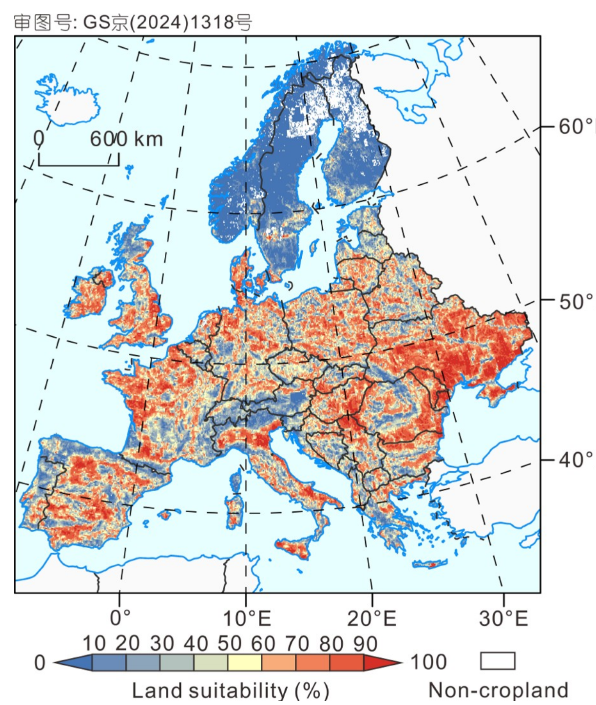

Figure 3 The land suitability with 5′×5′ resolution in Europe.

The key to city-site-based cultivation preference is assigning an attraction radius to a city (i.e., indication radius). For cropland allocation, the major factor limiting the attraction radius is food accessibility to the city, which depends on both human physiological conditions and transportation capacity. Referring to historical research and regional empirical data, Zhang et al. (2023) provided an indication radius of 66 km for Germany, France, and Italy in AD 1800. Considering that transportation capacity changed little throughout Europe until AD 1850 (Jepsen et al., 2015; Pounds, 1990), this study adopted the indication radius of 66 km proposed by Zhang et al. (2023), except for Finland, Norway, and Sweden, because Norway and Sweden had relatively few city sites compared with other European countries and may miss data for different levels. Based on the indication radius of neighbouring countries (Clark, 2009) and the central place theory (Christaller, 2010), the indication radii of Sweden and Norway were increased by 1.5 times of 66 km, or 96 km. Finland had only three city sites in AD 1800, which could not use the city-site-based cultivation method.

Different city levels were used to describe regional differences in cropland demand in this method. In general, regions with higher city rankings (in other words, larger populations) tend to reclaim more croplands. It should be pointed out that the population of the fourth-level cities used the population before AD 1750. This is because when Europe was on the eve of agricultural evolution, with most regions still in traditional agriculture and experiencing little change in production levels, the rapid population growth in fourth-level European cities after AD 1750, was more than three times that before AD 1750 (Table 3), was mainly associated with the Industrial Revolution and was not directly linked to the food supply in the surrounding areas.

(3) Limitations on the allocation extent and maximum cropland fraction

This study limited the extent of potential croplands by excluding non-arable areas and setting the upper limit of the cropland fraction in grids. Most existing cropland allocation models used modern cropland distribution to exclude non-arable areas, assuming that the distribution range of historical and modern croplands was consistent (Klein Goldewijk et al., 2017a; He et al., 2023). However, the extent of cultivated areas changed in the past owing to changes in the natural environment, land reclamation policies, technological advancements, the introduction of new crops, and so on.

Non-arable areas mainly include water bodies, historical urban areas, and regions beyond the boundaries of cultivation. Historical water area data, which is difficult to obtain, was substituted by modern water data. Continuous and complete records of historical built-up areas are not available. Considering that built-up areas generally increased at the same rate as the urban population during the pre-industrial era (Beaujeu-Garnier and Chabot, 1967), the urban population and other relevant proxies were employed to reconstruct the historical built-up areas before AD 1900. Using the area of 173 European cities in AD 1300 reconstructed by Cesaretti et al. (2016), the average urban area for each city rank was estimated. The urban areas from AD 1300–1850 were estimated according to the expansion rates of city walls and urban population growth rates. However, after AD 1900, rapid urbanisation induced by industrialisation in much of Europe resulted in the nonlinear growth of urban expansion and population. It was assumed that the urban core zones were formed by AD 1900, and the subsequent increases in urban area were due to the expansion of peripheral and satellite cities around the core. The built-up area in AD 1950 was regarded as the same as that in AD 2000. Grids in which the proportion of modern urban area to the total grid area exceeded 85% were considered to have developed into cities by AD 1900, which led to the estimation of urban areas in AD 1900. The specific reconstruction steps are presented in Appendix.

The areas beyond the cultivation boundary were not allocated to croplands. Over the past millennium, influenced by climatic fluctuations in warm-cold-warm periods, European agriculture has experienced phases of expansion, retraction, and re-expansion, with cultivation boundaries expanding during warm periods and retracting during cold periods. During the Medieval Warm Period in the 10th–14th centuries, the cultivation boundary was approximately 100–200 m higher than during the cold period of the Little Ice Age from the 15th to the 19th centuries. During the Little Ice Age, the cultivation boundary retreated and settlements at higher altitudes and latitudes decreased. The warm period during AD 1900–2000 resulted in a re-expansion of the cultivation range (Fagan, 2000; Pounds, 1990). In accordance with the above changes, this study used the modern cropland distribution range as the cultivation boundary in AD 2000 and set the historical potential cropland distribution ranges based on the changing vertical limitation and northern boundary (primarily within Finland) during the cold and warm periods in history (Table 4, Figure 4). The specific methods and evidence are provided in Appendix.

Each grid cell had a maximum cropland fraction set. The upper cultivation limit for all grid cells was set at 90% in this study, because, at least 10% of the land is left fallow for other purposes in the case of intensive agricultural utilisation (Haber, 1979).

2.2.3 *Construction of allocation models based on priority index*

The process of allocation is as follows. First, a general formula for the priority index was constructed considering both land suitability and cultivation preference (eq. (1)).

$$Priority(i,t) = [Preference_{\mathrm{nor}}(i,t) + a] \times Suitability_{\mathrm{nor}}(i), \quad (1)$$

where $Priority(i,t)$ is the priority index of grid i at time t; $Preference_{\mathrm{nor}}(i,t)$ is the normalized cultivation preference of grid i at time t; $Suitability_{\mathrm{nor}}(i)$ is the land suitability of grid i; a is an empirical coefficient.

Second, grid priority weights were constructed by country. Since the grid area of 5′×5′ decreases with increasing latitude, to ensure that larger grid areas allow for more cropland under the same cultivation priority, it is necessary to adjust the grid priority index when constructing the priority weights by country. Define the ratio of the area of the grid i to the largest grid area in the country it belongs to as the grid area coefficient of grid i (eq. (2)). Then the priority weight of the grid i was calculated as the ratio of the adjusted priority index of the grid i to the sum of the priority indexes of the country to which it belongs (eq. (3)).

$$\varepsilon(i) = \frac{GridArea(i)}{GridArea_{\max}} \times 100\%, \quad (2)$$

$$W_{\mathrm{priority}}(i,t) = \frac{Priority(i,t) \times \varepsilon(i)}{\Sigma_{i=1}^{n}(Priority(i,t) \times \varepsilon(i))}, \quad (3)$$

where $\varepsilon(i)$ is the grid area coefficient of grid i; $GridArea(i)$ is the area of grid i; $GridArea_{\max}$ is the largest grid area in the country the grid i belongs to; $W_{\mathrm{priority}}(i,t)$ is the priority weight of grid i at time t; n is the total number of grids in the country of grid i.

Based on the grid priority weight, the national cropland area was preliminarily allocated (eqs. (4) and (5)).

$$CropArea(i,t) = CropArea(t) \times W_{\mathrm{priority}}(i,t), \quad (4)$$

$$CropFrac(i,t) = \frac{CropArea(i,t)}{GridArea(i)}, \quad (5)$$

where $CropArea(i,t)$ is the cropland area allocated in grid i at

审图号: GS京(2024)1318号

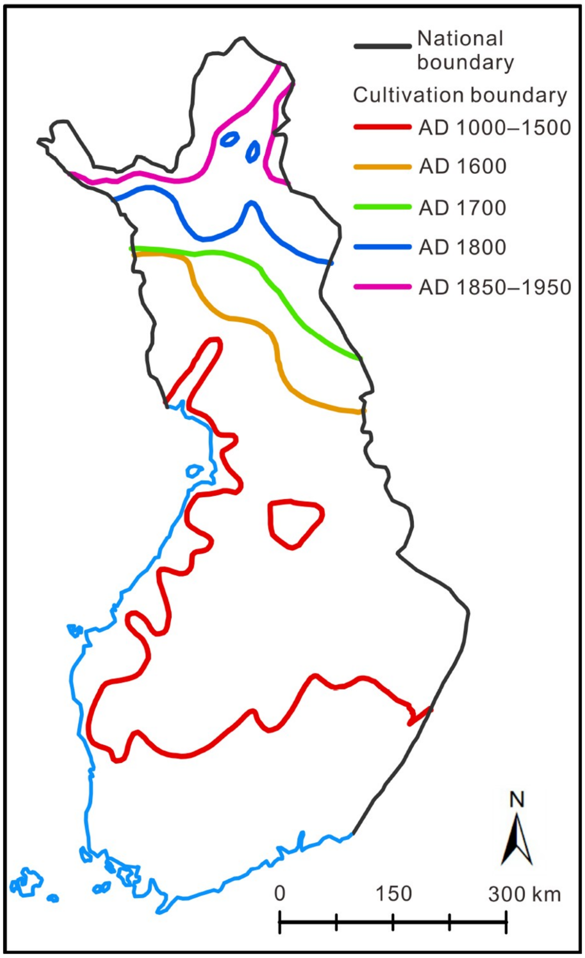

Figure 4 Changes in northern cultivation boundary in Finland over the past millennium (adapted from Sømme, 1960).

time t; $CropArea(t)$ is the total cropland area of the country the grid i belongs to at time t; $CropFrac(i,t)$ is the cropland fraction in grid i at time t.

Finally, the grids with cropland fractions exceeding the maximum values were revised. The maximum cropland value of a grid needs to deduct the non-arable area, such as water and built-up area, from the grid area, and then modified by the upper cultivation limit. The maximum cropland fraction is 0 if the grid is beyond the historical cultivation boundary. In AD 2000, the maximum cropland fraction was 0 if the modern cropland fraction of the gird was 0 (eq. (6)).

$$ArabFrac_{\max}(i,t) = \begin{cases} \dfrac{GridArea(i) - Area_{\text{water}}(i) - Area_{\text{built-up}}(i,t)}{GridArea(i)} \times \delta(Y), \\ \quad i \leqslant \lim(t), \\ 0, i > \lim(t) \text{ or } t = \text{AD } 2000, CropFrac(i) = 0, \end{cases} \tag{6}$$

where $ArabFrac_{\max}(i,t)$ is the maximum cropland fraction in grid i at time t; $Area_{\text{water}}(i)$ is the water area in grid i; $Area_{\text{built-up}}(i,t)$ is the built-up area in grid i at time t; δ is the upper limit of cropland fraction per grid that is assigned to be 90%; $\lim(t)$ is the cultivation boundary at time t.

The preliminarily allocated results $CropFrac(i,t)$ were compared with the corresponding maximum cropland fraction $ArabFrac_{\max}(i,t)$. For the grid with a cropland fraction exceeding the maximum value, $ArabFrac_{\max}(i,t)$ was assigned directly to that grid allocation result (eq. (7)). And the extra cropland area was reallocated to other unsaturated grids until the cropland fraction of all grids was no higher than their corresponding maximum cropland fraction $ArabFrac_{\max}(i,t)$, according to eqs. (4) and (5).

$$CropFrac(i,t) = \begin{cases} ArabFrac_{\max}(i,t), \\ \quad \left(CropFrac(i,t) > ArabFrac_{\max}(i,t)\right), \\ CropFrac(i,t), \\ \quad \left(CropFrac(i,t) \leqslant ArabFrac_{\max}(i,t)\right). \end{cases} \tag{7}$$

2.2.4 *Time-interval-based allocation model utilization*

This study constructed an allocation model using settlement as an indicator of cultivation preference. Settlements and croplands are interdependent. However, the relationship between the two in historical periods existed significant stages. Wu (2021) divided the relationship between settlements and cropland cultivation into five stages. During the rapid development stage, the number of settlements increased significantly with the expansion of settlements, and the cropland area also rapidly increased. With the rapid growth of the population, new settlements increased, and the surrounding land was cultivated. The settlements established in the early period expanded, and more land in their vicinity was tilled, with an increase in the cropland fraction. As

Table 4 The upper cultivation boundaries in the warm and cold periods in Europe

Countries and regions	Cold period AD 1800–1850	Warm period AD 1900–2000	Reference
The UK, Ireland	300 m	400 m	Fagan, 2000
Norway, Sweden, Finland	650 m	750 m	Bele and Norderhaug, 2013; Modern cropland distribution
Rest of Europe	1100 m	1200 m	East, 1967; Pfister, 1983; Bucała-Hrabia, 2017; Modern cropland distribution

agriculture matured, settlement growth subsequently decreased, and until all possible croplands were cultivated, new cropland reclamation activities mostly took place in the vicinity of already existing settlements.

From AD 1800 to 2000, European agriculture underwent significant transformation from traditional agriculture to modernised agriculture (Figure 5), and the relationship between cropland and settlements also shifted from a rapid development stage to a mature stage. During the traditional agricultural period (before AD 1850), Europe relied on mixed farming as its foundation for basic food production, which was characterised by labour-intensive small-scale manor economies. Croplands and settlements were in a rapid development stage, and with the increase in population, the scale and number of settlements continued to increase, corresponding to the spatial expansion of croplands and an increase in the fraction around the settlements. At this stage, settlements could certainly indicate the existence of cropland. The year 1800 belonged to the traditional agricultural period, and the priority index was comprised of land suitability and cultivation preference, as described in eq. (8).

After AD 1900, Europe fully entered industrialisation, and traditional agriculture also transitioned to modern agriculture (Figure 5), which is characterised by more specialised, large-scale mechanised, intensified, and market-oriented agricultural practices. The relationship between croplands and settlements entered a mature stage when the increase in cropland area did not have a close relationship with the number of settlements, as the number of settlements almost no longer increased. AD 1950 and 2000 belonged to the modern agricultural stage. The indication of settlements to the cropland weakened (Zhang et al., 2023). Land suitability was used to streamline the priority because cultivation preference had a minimal impact on the priority index (eq. (9)).

The period from AD 1850 to 1900 marked a transition from traditional agriculture to modern agriculture in Europe when the agricultural and Industrial Revolutions resulted in changes in land use patterns (Figure 5). Meanwhile, the relationship between croplands and settlements shifted from a rapid development stage to a mature stage. In countries influenced by the Industrial Revolution, the paradigm of traditional agriculture has transitioned to modern agriculture. The allocation results used the average of the results obtained by eq. (8) and eq. (9) (eq. (10)). The significance and rationality of this method were discussed in Section 4.2. The detailed allocation method is provided in Appendix.

$$Priority(i,t) = \left[Preference_{\mathrm{nor}}(i,t) + 0.01\right] \times Suitability_{\mathrm{nor}}(i), \tag{8}$$

$$Priority(i,t) = Suitability_{\mathrm{nor}}(i), \tag{9}$$

$$CropArea(i,t) = \frac{CropArea_{\mathrm{city}}(i,t) + CropArea_{\mathrm{suit}}(i,t)}{2}. \tag{10}$$

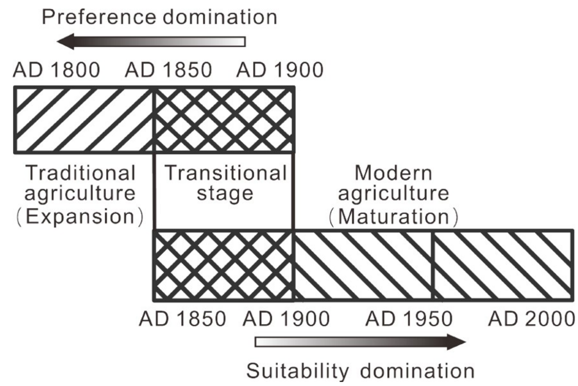

Figure 5 Schematic diagram of allocation method used in different stages.

2.3 Comparison with HYDE3.2

The HYDE (The History Database of the Global Environment) is currently the most widely used historical land cover dataset. The HYDE has undergone continuous updating for improvement since the release of its initial version, HYDE1.1 (Klein Goldewijk and Battjes, 1997) in AD 1997. The version compared in this study was HYDE3.2 (Klein Goldewijk et al., 2017a).

Compared with the global dataset HYDE3.2, this study adopted more historical empirical data with higher reliability to reconstruct national cropland data. In the grid allocation method, we considered the impact of both suitability and preference, and used the corresponding optimal model in different time periods. By comparing HYDE3.2 and this study, we can evaluate the potential and possible paths for improving the reliability of the global cropland dataset.

2.3.1 Cropland data of HYDE3.2

The HYDE3.2 reconstructed the changes in global cropland and grazing land, covering the period from 10000 BC to AD 2015, with a resolution of 5′×5′ (Klein Goldewijk et al., 2017a). The HYDE3.2 cropland data set was based on FAO's cropland data since 1961, as well as limited collected global per capita cropland area data for the past two thousand years. It used concave or bell-shaped curves to fit and estimate the per capita cropland area of representative countries, and reconstructed the cropland area of each country by population data. Among 238 countries and regions, HYDE3.2 updated the historical land use information of 55 countries, covering 23% of the countries in the world (Klein Goldewijk et al., 2017b). However, most of the per capita cropland data derived from historical land use empirical data was concentrated in the 19th century.

HYDE3.2 selected natural factors such as soil quality, distance to rivers, slope, and temperature to quantify land suitability. Then it constructed a grid allocation model based on modern population distribution, natural factors, and some historical population density derived from the regional archaeological data. The grid allocation model was used to calculate the historical weight of cropland distribution. ESA's cropland data was used as the modern weight, and the proportion of historical weight and modern weight was adjusted since AD 1500, which served as the critical point. Meanwhile, built-up areas, protected areas, bare land, and areas with a slope greater than 45° that cannot be cultivated were deducted from the grid. Finally, the historical national cropland areas were input into the grid allocation model to reconstruct the historical cropland grid data (Klein Goldewijk et al., 2017a).

2.3.2 Method for data comparison

We conducted the comparison using absolute and relative differences (eqs. (11) and (12)).

$$D_i = V_{Ri} - V_i, \quad (11)$$

$$RD_i = \frac{V_{Ri} - V_i}{V_i} \times 100\%, \quad (12)$$

where D_i and RD_i denote absolute difference and relative difference ratio in period i; V_{Ri} denotes the value of cropland fraction or cropland area in HYDE3.2 in period i; and V_i denotes the value of cropland fraction or cropland area in this study in period i.

3 Results

The cropland areas in European countries over the past 200 years, as supplemented and corrected for in this study, are shown in Figure 6. In terms of national cropland area, except for the newly added data in AD 1950, this study has a very small difference from the data of Ye et al. (2023). The total amount of cropland increased, reaching a peak in AD 1900, and then declined from AD 1900 to 2000 (Table 5), which is similar to the results reported by Ye et al. (2023). This study focused on changes in the spatial distribution of croplands, which manifested as the abandonment of croplands in marginal mountainous areas and a shift in high cultivation centres from west to east (Figure 7a).

Table 5 Comparison of total cropland area in Europe reconstructed in this study and HYDE3.2

Year (AD)	Cropland area in this study (10^4 km^2)	Cropland area in HYDE3.2 (10^4 km^2)	RD (%)
1800	134.14	111.74	−16.70
1850	170.46	138.74	−18.61
1900	202.37	176.39	−12.84
1950	195.46	195.82	0.19
2000	178.40	177.46	−0.53

3.1 The traditional agricultural period from AD 1800 to 1850

Between AD 1800 and 1850, the cropland area in Europe increased by 36.32×10^4 km^2, along with fewer changes in cultivation extent. Change of cropland distribution was primarily characterised by expansion in northern Finland and retreat in the height of the Alps and Pyrenees. However, the overall cultivation intensity significantly increased. The average cropland fraction of the cropland grids increased from 24% to 30% (Figure 8a). The proportion of grids with a lower cropland fraction (<30%) decreased, while the proportion of grids with cropland fractions>50% more than doubled (Figure 8a). Particularly, grids with the highest cropland fraction (70%–90%) increased from 2% to 6% (Figure 8a).

In AD 1800, grids with cropland fractions >50% were mainly distributed in the plains and hilly regions of Italy, France, Belgium, Germany, and Czechia; the regions along the Danube River in Austria, Slovakia, and Hungary; large parts of Moldova; the southern region of Portugal; and the scattered areas in the plains and hilly regions of Ukraine, Poland, England, and Denmark. Approximately 79% of these grids (with cropland fractions >50%) were located on the western side of the European continent. By AD 1850, the grids with cropland fractions >50% increased significantly in Spain, Poland, western Czechia, Ukraine, and Denmark, forming a high cropland fraction belt stretching from west to east across the European continent (Figure 7a).

3.2 The transitional period from AD 1850 to 1900

From AD 1850 to 1900, the cropland area in Europe reached its peak of 202.37×10^4 km^2. During this period, the cultivation area expanded in the Scandinavian Peninsula, Great Britain, mountainous regions of the Balkan Peninsula, and the Carpathian Mountains, leading to a decrease in the proportion of non-cropland grids from 14% to 11% (Figure 8a). The average cropland fraction of cropland grids increased to 35% (Figure 8a). The proportion of grids with lower cropland fractions (<30%) continued to decrease, while grids with cropland fractions >50% continued to increase, and the proportion of grids with the highest cropland fractions (70%–90%) increased to 11% (Figure 8a). The belt of the high cropland fraction expanded southward, with newly expanded areas mainly in Germany, the Hungarian Great Plain, the middle and lower reaches of the Danube River plain, and Ukraine (Figure 7a).

3.3 The modern agricultural period from AD 1950 to 2000

In AD 1950, the cropland area in Europe decreased by

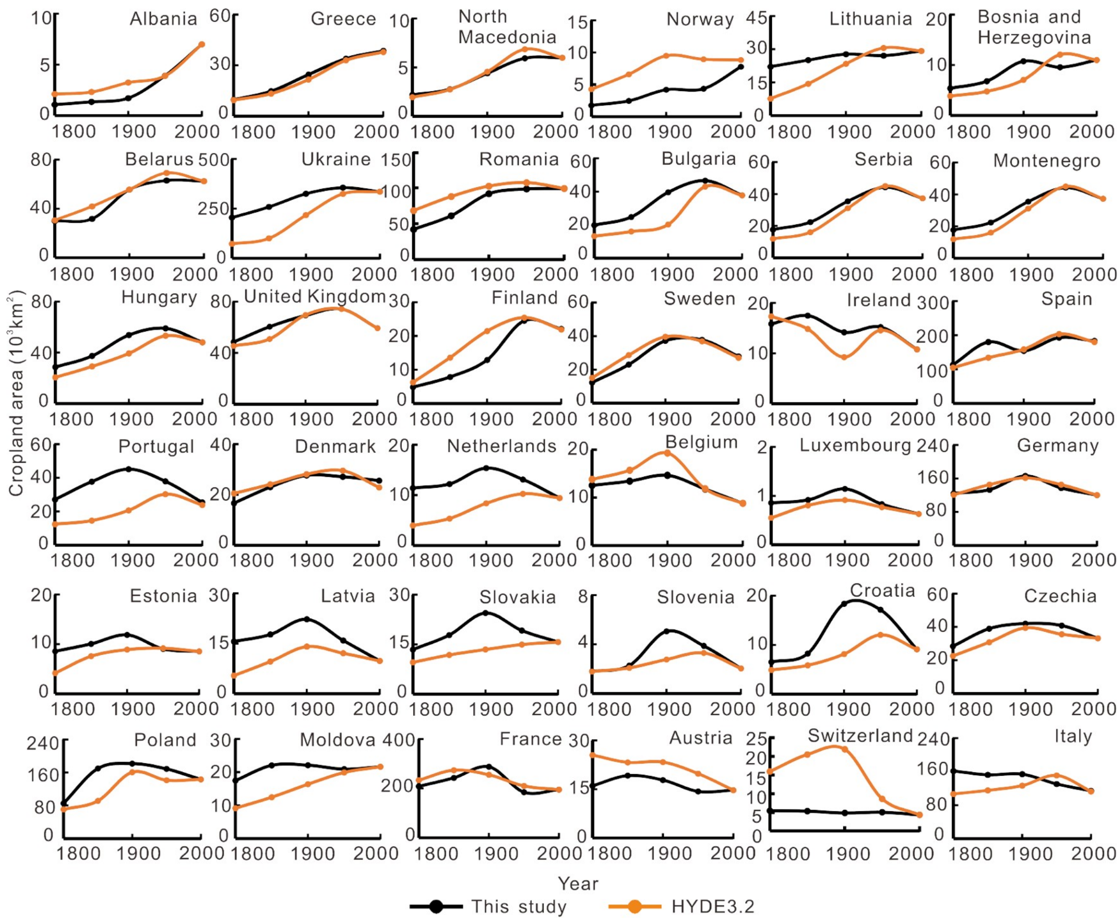

Figure 6 National cropland area in European countries over the past 200 years. Note: Because Serbia and Montenegro were once part of the Federal Republic of Yugoslavia in 2000, these two countries were combined for the calculation. However, they are still displayed according to the modern boundaries. The cropland areas shown in the figure for Serbia and Montenegro are the total areas of the two countries.

6.91×10^4 km^2 compared with AD 1900. The cultivation area remained relatively stable, but there was a significant change in the spatial distribution of the cultivation intensity with the coexistence of cropland grids of abandonment and increase. In AD 1950, the average cropland fraction of the cropland grids was 34%, which was slightly lower than that in AD 1900 (Figure 8a). However, the proportion of grids with the highest cultivation intensity (70%–90%) decreased from 11% to 8%, whereas grids with cropland fractions in the ranges of 10%–30% and 30%–50% increased from 19% to 21% (Figure 8a). The areas experiencing cropland abandonment were mainly in France, Germany, and Italy, as well as parts of Portugal, Austria, Slovakia, Croatia, and the Baltic Countries. In contrast, Finland, Hungary, and most countries in the Balkan Peninsula, Ukraine, and Spain witnessed further increases in cultivation intensity. The high cultivation centre shifted from western to eastern Europe. In the eastern region of Europe, which includes Poland, Czechia, Hungary, Moldova, and Ukraine, 70% of the grids have the highest cropland fractions (70%–90%) (Figure 7a).

From AD 1950 to 2000, Europe experienced further cropland abandonment, resulting in a reduction of 17.06×10^4 km^2 in cropland area. The proportion of non-cropland grids increased from 11% to 19%, with newly increased non-cropland grids mainly in barren lands and high-altitude mountainous regions of the Scandinavian Peninsula, northern Finland, the United Kingdom, and Ireland (Figures 7a, 8a). The average cropland fractions of the cropland grids remained unchanged. However, the proportion of grids with the lowest cropland fractions (0%–10%) decreased from 21% to 15%, while the number of grids with cropland fractions in the range of 10%–70% slightly increased (Figure 8a). Conversely, the proportion of grids with the highest cultivation intensity (70%–90%) decreased from 8% to 4% (Figure 8a). The high cultivation centre has shifted even further towards Eastern Europe, with roughly 67% of the grids with the highest cropland fractions (70%–90%) concentrated in Ukraine and Moldova (Figure 7a).

4 Discussion

4.1 Comparison with HYDE3.2

The national cropland areas of Europe in this study showed

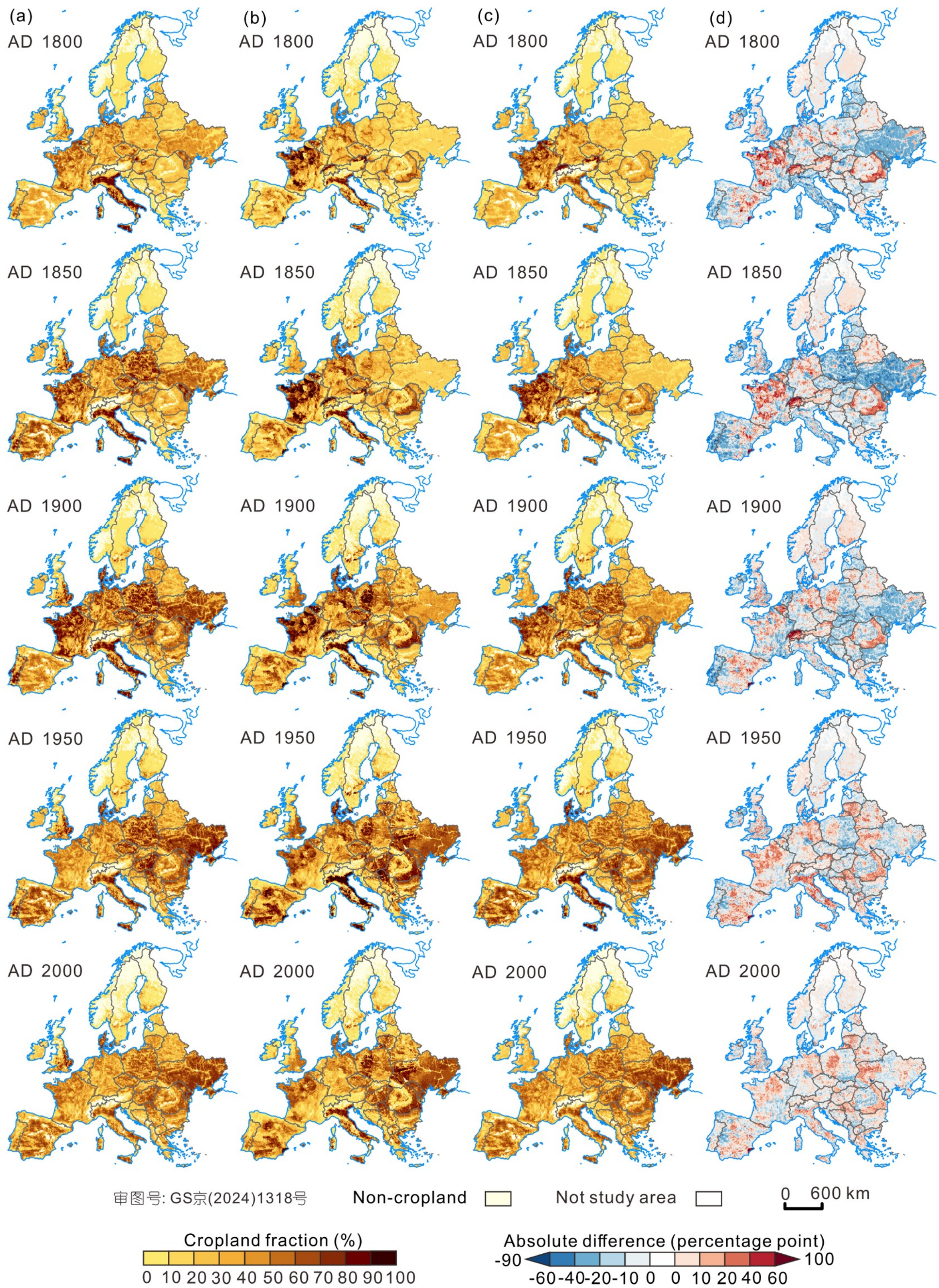

Figure 7 The spatial distribution of cropland fractions and absolute differences in Europe. Column (a) is the spatial pattern in this study; Column (b) is the spatial pattern in HYDE3.2; Column (c) is the spatial pattern obtained by allocating HYDE3.2's cropland area using this study's allocation model; Column (d) is the spatial distribution of absolute differences between HYDE3.2 and this study.

very small differences from the data of Ye et al. (2023), except for the newly added data in AD 1950. Therefore, the conclusions of comparing the European cropland over the past 200 years with HYDE3.2 by Ye et al. (2023) are also applicable to this study. That is, the European croplands in this study and HYDE3.2 both have the same growth trend over the past 200 years, but HYDE3.2 systematically underestimated the total cropland area in Europe, especially from AD 1800 to 1900. On a national scale, the differences in cropland area between HYDE3.2 and this study are partly

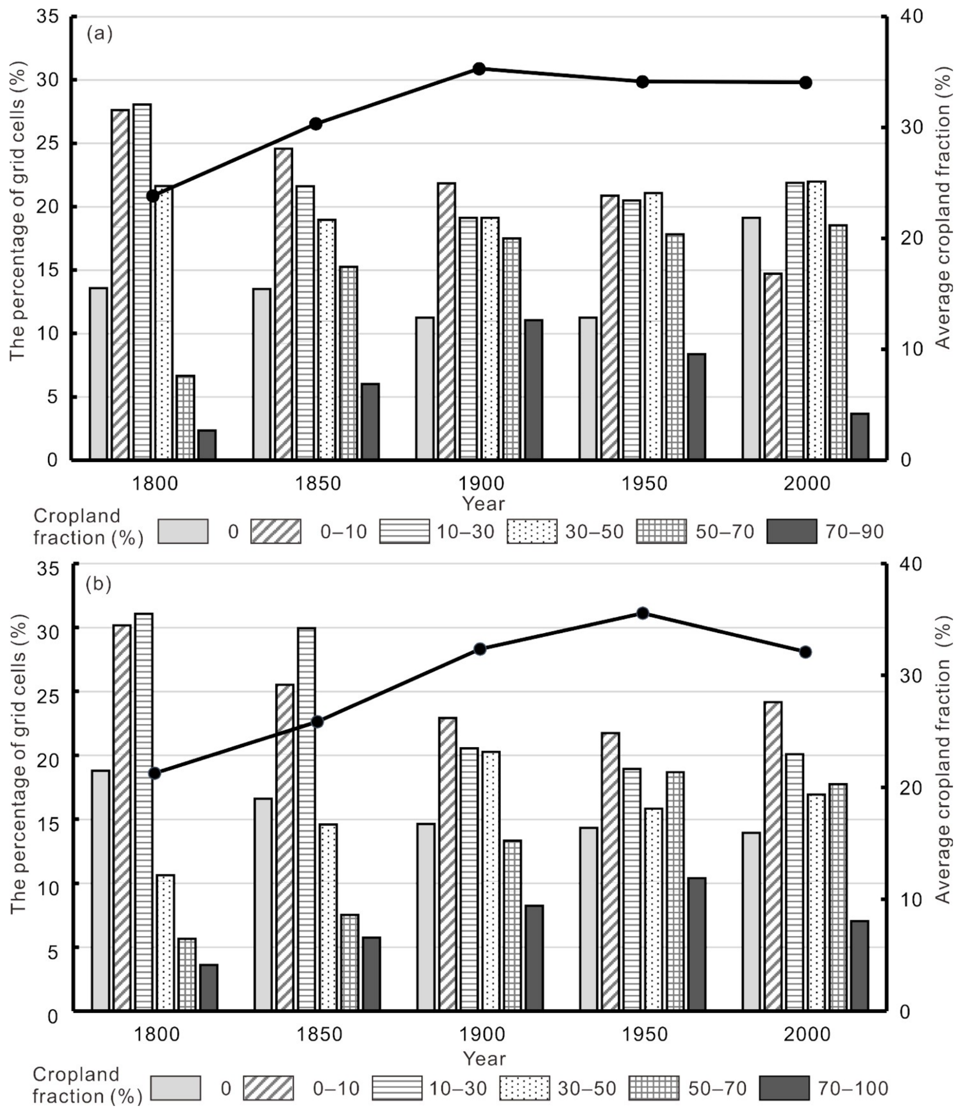

Figure 8 Grid proportion in different cultivation levels and average cropland fraction of cropland grids in Europe from 1800 to 2000 in this study (a) and HYDE3.2 (b).

positive and partly negative. However, the absolute value of the relative difference for most countries decreased continuously from AD 1800 to 1950. This study adopted more historical empirical data and reconstructed cropland areas using multiple indicators and methods. Thus, it has higher credibility than HYDE3.2.

A comparison of the cropland area in AD 1950 with HYDE3.2 was made in this study. The total reconstructed cropland area in AD 1950 was almost the same as HYDE3.2, with a relative difference of only 0.19%. However, regional differences still exist in national cropland areas (Figures 6, 9). For all countries, the absolute differences in cropland fraction fall within ±10 percentage points (PP), and 54.3% of the countries have relative differences within ±10%. There are seven countries with absolute values of relative differences ranging between 10% and 20%, including Czechia, France, Italy, Lithuania, North Macedonia, Poland and Slovenia. Seven other countries have absolute values of relative differences ranging from 20% to 40%, including Austria, Bosnia and Herzegovina, Croatia, Latvia, the Netherlands, Portugal, and Slovakia. The relative differences for Switzerland and Norway were 71% and 106%, respectively, but the absolute differences were only 8.66 and 1.19 PP. Among the 16 countries with absolute values of relative differences greater than 10% in AD 1950, 10 countries (Italy, Portugal, Austria, Czechia, Poland, Latvia, Lithuania, France, Switzerland, and Norway) adopted the statistical cropland data in this study; the cropland areas for Bosnia and Herzegovina, and the Netherlands were estimated by the changing rate of the empirical cropland area; the cropland areas for North Macedonia, Croatia, and Slovenia were the revised cropland area data; only the cropland area for Slovakia was estimated by analogy with the cropland area of Czechia. In contrast, the HYDE3.2 data were all estimated by the per capita cropland

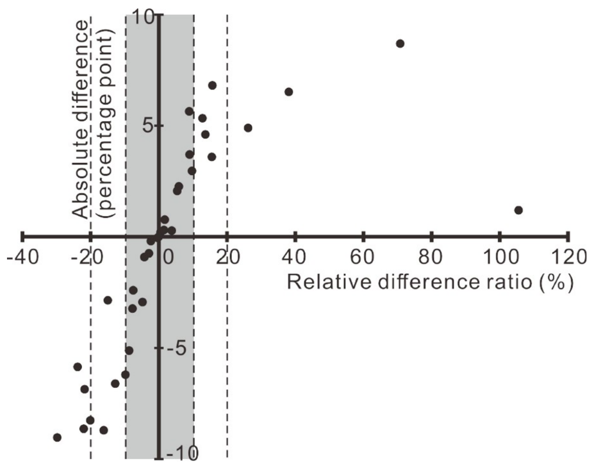

Figure 9 The difference of the national cropland fraction in Europe from 1800 to 2000 between this study and HYDE3.2.

area of the representative countries. Because this study adopted more historical empirical data, it is acceptable that it has a higher credibility compared with HYDE3.2.

In summary, the main differences in the national-scale cropland area between this study and HYDE3.2 were due to the different reconstruction methods and data sources. HYDE3.2 estimated the national cropland area before AD 1960 using the historical population and per capita cropland area. The per capita cropland area from an assumed historical per capita cropland area curve largely depends on the historical records of the representative countries that can be collected (Klein Goldewijk et al., 2017b). To reconstruct the cropland area, this study made use of multiple indicators, methods, and more historical empirical data. Compared with the current HYDE3.2 dataset, which highly relied on the assumption of per capita cropland area, our results are believed to be more reliable.

From the perspective of cropland spatial distribution changes, both HYDE3.2 and this study showed overall similarities but also some differences. In both datasets, the average cropland fraction of the cropland grids experienced an early increase, subsequently followed by a decrease. However, in this study, the peak occurred in AD 1900, with the highest average cropland fraction reaching 35%, while in HYDE3.2, the peak occurred in AD 1950, with the highest value of 36% (Figure 8). The spatial distribution in HYDE3.2 also showed the processes of cropland abandonment and the shift of the high cultivation centre from west to east, but it was slightly delayed compared with this study (Figure 7b). In AD 1800, 64% of the grids with cropland fractions >50% in HYDE3.2 were concentrated in France, Germany, and Italy. The high cropland fraction belt then expanded continuously eastward. By AD 1950, countries had accumulated more than 67% of the grids with cropland fractions >70% across the entire European continent, including Italy, Ukraine, Romania, Spain, France, and Poland. By AD 2000, 65% of the grids with high cropland fractions (>70%) were distributed in Ukraine, Italy, Romania, and Poland, indicating that cropland abandonment occurred in Western Europe and that the cultivation centre shifted towards Eastern Europe.

From the distribution of absolute differences between the two datasets (Figures 7d, 10), AD 1850 is the year with the greatest differences, with 56% and 75% of the grids having absolute differences within ±10 and ±20 PP, respectively. After that, the differences between the two datasets reduced over time, with 70% and 91% of the grids having absolute differences within ±10 and ±20 PP, respectively, by AD 2000.

The spatial differences between the two datasets mainly arise from differences in the national cropland areas and the allocation algorithm. To better distinguish the main causes of the differences, we used the allocation method constructed in this study to allocate the cropland area of HYDE3.2 (Figure 7c). The differences between it and the results of this study are attributed to the different cropland areas of each country, and the differences between it and the results of HYDE3.2 are attributed to the different allocation methods. The relative rationality between HYDE3.2 and this study is judged further by comparing with the historical facts of regional development. Taking the year 1800 as an example, the differences in Ukraine, Moldova, the three Baltic countries, Italy, Portugal, the Netherlands, Denmark, and Slovakia, were mainly due to the different national amounts of cropland. The differences in Romania, Hungary, France, Germany, Poland, and the Balkan Peninsula were mainly due to different allocation methods. In Switzerland and Austria, the contributions of differences in the total amount of cropland and method were both significant.

As shown in Figure 7b, the cropland distribution of HYDE3.2 was mainly based on the modern cropland distribution pattern, which concentrated on flat plains and along rivers. For example, in HYDE3.2, the high cultivation centre in Romania in AD 1800 was distributed in Wallachia and Moldavia, located south of the Carpathians. Particularly, the most grids of the high cropland fraction were distributed in Wallachia, where the Danube Plain and Dobrogea Hills have a flat topography and fertile soil. However, according to the agricultural history in Romania, before the signing of the Treaty of Adrianople in AD 1829, it was difficult to utilize the fertile Danube Plain because of frequent attacks by the Turks. It was not until AD 1829, when the Black Sea opened to the international market and the Danube River was opened up for food transportation, that Wallachia and Moldavia experienced a steady increase in cropland area and grain production for the agricultural expansion along with population explosion (Georgescu, 1990; Fraser and Stringer, 2009). Therefore, the cropland distribution pattern of HYDE3.2 in AD 1800 was not reasonable because the centre

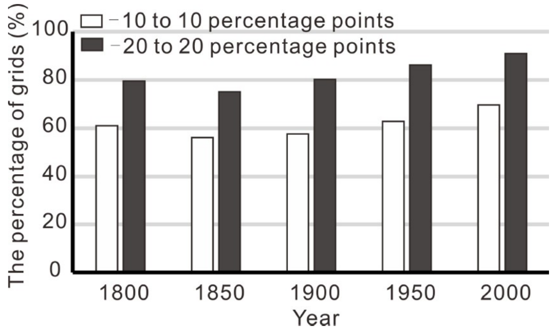

Figure 10 The percentage of the grids with absolute differences between HYDE3.2 and this study for the whole of Europe.

of high cultivation in Romania should not appear in Wallachia at this time. For another example, in HYDE3.2, the cropland in Hungary was distributed along the Danube and Tisza rivers in both AD 1800 and 1850, owing to the use of the distance to the river as an allocation factor. However, it was only after AD 1846 that Hungary started to harness the river by draining wetlands into croplands, especially in the Tisza River (Pinke, 2014; Lieskovský et al., 2018).

In summary, because of the choice of the allocation model factors and the limitations of allocation assumptions, the cropland distribution of HYDE3.2 is more strongly influenced by modern cropland patterns and natural factors. Such an allocation model tends to allocate more cropland to plains for all historical periods but less to account for the role of agricultural history. Particularly in earlier periods of agricultural development, when the amount of cropland had not yet been saturated, and cultivation was more influenced by human choices, it could cause significant deviations from the real distribution of cropland in HYDE3.2. On the contrary, by adopting the cultivation preference indicated by settlements to reflect human activities, this study could make cropland distribution more credible, which is more consistent with the historical cultivation process.

4.2 Influence of allocation algorithms and dynamic cultivation boundary on the allocation results

In this study, different allocation algorithms were used for the traditional agricultural and modern stages, respectively. For countries in the transitional period from AD 1850 to 1900, the grid results involved merging two groups of allocation data. The solution in this study was to use the average of the results because there were acceptable differences between the results of the two methods. In AD 1850, for the countries excluding Finland, Eastern Europe, and the Balkan Peninsula, the absolute difference in the results obtained using the two algorithms of eqs. (8) and (9) were acceptable, that is, within ±10 PP for 92% of the grids and ±20 PP for 98% of the grids. The absolute difference showed that there was good consistency between the two results obtained purely based on suitability and a combination of suitability and human preference, indicating that cultivation preference for cropland indicators has declined and the two methods can be effectively merged. Similarly, in AD 1900, it was also acceptable for the countries in Eastern Europe and the Balkan Peninsula, to use the average of both methods as the final result, because the proportion of grids with absolute differences within ±10 PP was 81% and within ±20 PP was 95%.

One of the differences between this study and Zhang et al.'s (2022a)suitability-based allocation method, Zhang et al.'s (2023) priority method combining suitability and preference is the addition of a dynamic cultivation boundary to limit the historical cropland distribution range, which is expected to improve the rationality of cropland distribution. The extent to which this affected the credibility of the results is discussed by setting different scenarios for comparison.

Scenario 1: Setting or not setting the highest upper cultivation boundary. For the European continent (excluding Sweden, Norway, Finland, the UK, and Ireland), comparing the results obtained without setting a cultivation boundary with the results of this study at five time sections, the grid proportions with absolute differences within ±5 PP and ±10 PP are above 94% and 96%, respectively.

Scenario 2: Setting different heights of regional upper cultivation boundaries. For Norway and Sweden, the upper cultivation boundaries in this study were set at 650 m during the cold period and 750 m during the warm period. Suppose the upper limit is modified to 750 m in the southern part and 500 m in the northern part during the cold period, and 850 m in the southern part and 600 m in the northern part during the warm period. In that case, the proportions of grids with absolute differences within ±5 PP are above 99% for all periods when compared with the results of this study.

Scenario 3: Setting or not setting a latitudinal cultivation boundary. For Finland, comparing the results obtained without setting a northern cultivation boundary with the results of this study, the proportions of grids with absolute differences within ±5 PP are above 99.60% for all five time sections.

Scenario 4: Using or not using the cultivation range in AD 2000 to restrict the cropland distribution. If the restriction of modern real cropland distribution is not used, the proportions of the grids with an absolute difference within ±5 PP are 99% when compared with the results of this study.

In summary, the effects of cultivation boundaries and modern cropland distribution on the allocation results are limited for the whole of Europe. However, it could have a greater impact on certain regions. For example, in the mountainous countries of Switzerland and Austria, the proportions of grids with an absolute difference within ±10 PP at five time sections are between 57% and 71% for the two distribution results of setting or without setting an upper cultivation boundary. However, the maximum total cropland area in these two countries accounts for only 2% of the total cropland in Europe.

5 Conclusions

This study reconstructed spatially explicit cropland data for Europe from AD 1800 to 2000, with a temporal resolution of 50 years and a spatial resolution of 5′×5′[2]. The cropland dataset was developed by employing optimal allocation models according to the regional agricultural history to allocate the revised historical cropland areas of the countries based on modern administrative boundaries. The allocation model was constructed by a priority index combining land suitability based on physiogeographic factors and cultivation preference based on historical city sites. The main findings were as follows.

(1) This study employed different allocation models based on land suitability by physiogeographic factors and the cultivation preferences of historical city sites for countries at different agricultural stages. Croplands in the countries in the traditional agricultural stage were allocated by a combination of land suitability and cultivation preferences, whereas that in the countries in the modernised agricultural stage were allocated based on land suitability alone. For the countries in transition between the two stages, the cropland was allocated by using the average of the two allocation results because there is a good coherence with the absolute differences in cropland fraction within ±10 PP for over 80% of the grids and within ±20 PP for over 95% of the grids between the two methods. Dynamically limiting the historical cropland distribution range can enhance the rationality of cropland distribution from AD 1800 to 2000 in Europe, but has a limited improvement in accuracy.

(2) Over the past 200 years, the total amount of cropland in Europe increased, reaching a peak in AD 1900, and then declined. The spatial distribution of croplands showed abandonment in marginal mountainous areas and a shift in the agricultural cultivation centre from the west to the east. In AD 1800, 79% of the grids with cropland fractions >50% were located in the western part of Europe. The high cultivation belt stretched from west to east across Europe by AD 1850 and expanded southward by AD 1900. By AD 1950, the high cropland fraction centre had shifted towards the east, when approximately 70% of grids with the highest cropland fractions (70%–90%) were concentrated in the eastern part of Europe, meanwhile cropland abandonment occurred in the western regions. In AD 2000, the cultivation centre further shifted towards Eastern Europe, when 67% of the grids with the highest cropland fractions (70%–90%) were distributed in Ukraine and Moldova.

(3) The total amount of cropland in HYDE3.2 also showed a trend of first increasing and then decreasing, but its peak occurred in AD 1950. Spatially, HYDE3.2 exhibited cropland abandonment and a shift in the high cropland fraction centre from west to east by AD 2000. Both differences in the national total cropland area and allocation methods could contribute to the distribution differences between HYDE3.2 and this study. Owing to the limitations of the allocation method of HYDE3.2, the earlier the allocation period, the greater the deviation between allocation results and development history. Compared with HYDE3.2, the cropland area and the spatial distribution in this study are believed to be more reliable and reasonable because they are more consistent with the regional development history.

Acknowledgements *This work was supported by the National Key Research and Development Program of China (Grant No. 2017YFA0603304).*

Conflict of interest The authors declare that they have no conflict of interest.

References

Allen R C. 2000. Economic structure and agricultural productivity in Europe, 1300–1800. Eur Rev Econ Hist, 4: 1–25

Bai Y, Wang M, Li H, Huang S F, Alatalo J M. 2017. Ecosystem service supply and demand: Theory and management application (in Chinese). Acta Ecol Sin, 37: 5846–5852

Beaujeu-Garnier J, Chabot G. 1967. Urban Geography. London and Harlow: Longmans, Green and Co. Ltd. 488

Bele B, Norderhaug A. 2013. Traditional land use of the boreal forest landscape: Examples from Lierne, Nord-Trøndelag, Norway. Norw J Geol, 67: 12–23

Bibliographisches Institut. 1909. Meyers Großes Konversations-Lexikon. 6th ed. Leipzig und Wien: Bibliographisches Institut

Bontemps S, Defourny P, Radoux J, Van Bogaert E, Lamarche C, Achard F, Mayaux P, Boettcher M, Brockmann C, Kirches G, Zulkhe M, Kalogirou V, Seifert F M, Arino O. 2013. Consistent global land cover maps for climate modelling communities: Current achievements of the ESA′ land cover CCI. ESA Living Planet Symposium. 1–6

Brovkin V, Boysen L, Arora V K, Boisier J P, Cadule P, Chini L, Claussen M, Friedlingstein P, Gayler V, Van Den Hurk B J J M, Hurtt G C, Jones C D, Kato E, De Noblet-Ducoudré N, Pacifico F, Pongratz J, Weiss M. 2013. Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the Twenty-First Century. J Clim, 26: 6859–6881

Bucała-Hrabia A. 2017. Long-term impact of socio-economic changes on agricultural land use in the Polish Carpathians. Land Use Policy, 64: 391–404

Cesaretti R, Lobo J, Bettencourt L M, Ortman S G, Smith M E. 2016. Population-area relationship for medieval European cities. PLoS One, 11: E0162678, DOI: 10.1371/journal.pone.0162678

Christaller W. 2010. Central Places in Southern Germany (in Chinese). Beijing: The Commercial Press. 456

Clark P. 2009. European Cities and Towns: 400–2000. Oxford: Oxford

2) The gridded cropland data in this study will be shared in the National Meteorological Science Data Center. http://dcpzx.cma.cn/#/data-collect/data?subId=cf08990f4f02080968738fb55a50f5e1&status=5&level=2.

University Press. 432
Culbert P D, Dorresteijn I, Loos J, Clayton M K, Fischer J, Kuemmerle T. 2017. Legacy effects of past land use on current biodiversity in a low-intensity farming landscape in Transylvania (Romania). Landscape Ecol, 32: 429–444
Dejongh G, Vanhaute E. 1999. Arable productivity in Belgian agriculture, c.1800-c.1950. In: Thoen E, Van Bavel B J P. Land Productivity and Agro-systems in the North Sea Area (Middle Ages–20th Century). Elements for Comparison. Turnhout: Brepols. 65–83
East G. 1967. An Historical Geography of Europe. London: Methuen & Co. Ltd. 492
Fagan B. 2000. The Little Ice Age: How Climate Made History, 1300–1850. New York: Basic Books. 246
Fang X, Zhao W, Zhang C, Zhang D, Wei X, Qiu W, Ye Y. 2020. Methodology for credibility assessment of historical global LUCC datasets. Sci China Earth Sci, 63: 1013–1025
Feng Y, Zhang S, He F, Zhou Z. 2014. Separate reconstruction of Chinese cropland grid data in the 20th century (in Chinese). Prog Geogr, 33: 1546–1555
Fraser E D G, Stringer L C. 2009. Explaining agricultural collapse: Macro-forces, micro-crises and the emergence of land use vulnerability in southern Romania. Glob Environ Change, 19: 45–53
Gaillard M J, Morrison K, Whitehouse N. 2015. Past anthropogenic land use and land cover change at the global scale for climate modelling studies: PAGES LandCover6k Working Group. Quat Perspect, 22: 25–27
Georgescu V. 1990. The Romanians: A History. Columbus: Ohio State University Press. 357
Haber W. 1979. Raumordnungs-konzepte aus der Sicht der Ökosystemforschung. Forschungs- und Sitzungsberichte Akademie f. Raumforschung und Landesplanung, 131: 12–24
He F, Li S, Zhang X, Ge Q, Dai J. 2012. Comparisons of reconstructed cropland area from multiple datasets for the traditional cultivated region of China in the last 300 years (in Chinese). Acta Geogr Sin, 67: 1190–1200
He F, Yang F, Zhao C, Li S, Li M. 2023. Spatially explicit reconstruction of cropland cover for China over the past millennium. Sci China Earth Sci, 66: 111–128
Hoffman G W. 1959. Yugoslavia: Changing character of rural life and rural economy. Am Slavic East Eur Rev, 18: 555–578
Hohenberg P M., Lees L H. 1995. The Making of Urban Europe, 1000–1994. Revised ed. Cambridge: Harvard University Press. 448
Huo R, Yang Y, Man Z. 2020. Gridded reconstruction of spatiotemporal evolution of cropland in the Zhangjiu River Basin in Yunnan Province from 1700 to 1978 (in Chinese). Acta Geogr Sin, 75: 1966–1982
Izakovičová Z, Špulerová J, Raniak A. 2022. The development of the Slovak agricultural landscape in a changing world. Front Sustain Food Syst, 6: 862451
Jones N, de Graaff J, Rodrigo I, Duarte F. 2011. Historical review of land use changes in Portugal (before and after EU integration in 1986) and their implications for land degradation and conservation, with a focus on Centro and Alentejo regions. Appl Geography, 31: 1036–1048
Jepsen M R, Kuemmerle T, Müller D, Erb K, Verburg P H, Haberl H, Vesterager J P, Andrič M, Antrop M, Austrheim G, Björn I, Bondeau A, Bürgi M, Bryson J, Caspar G, Cassar L F, Conrad E, Chromý P, Daugirdas V, Van Eetvelde V, Elena-Rosselló R, Gimmi U, Izakovicova Z, Jančák V, Jansson U, Kladnik D, Kozak J, Konkoly-Gyuró E, Krausmann F, Mander Ü, McDonagh J, Pärn J, Niedertscheider M, Nikodemus O, Ostapowicz K, Pérez-Soba M, Pinto-Correia T, Ribokas G, Rounsevell M, Schistou D, Schmit C, Terkenli T S, Tretvik A M, Trzepacz P, Vadineanu A, Walz A, Zhllima E, Reenberg A. 2015. Transitions in European land-management regimes between 1800 and 2010. Land Use Policy, 49: 53–64
Jia R, Fang X Q, Ye Y. 2023. Gridded reconstruction of cropland cover changes in Northeast China from AD 1000 to 1200. Reg Environ Change, 23: 128
Kaplan J O, Krumhardt K M, Zimmermann N. 2009. The prehistoric and preindustrial deforestation of Europe. Quat Sci Rev, 28: 3016–3034
Kaplan J O, Krumhardt K M, Zimmermann N E. 2011. The effects of land use and climate change on the carbon cycle of Europe over the past 500 years. Glob Change Biol, 18: 902–914
Kaplan J O, Krumhardt K M, Gaillard M J, Sugita S, Trondman A K, Fyfe R, Marquer L, Mazier F, Nielsen A. 2017. Constraining the deforestation history of Europe: Evaluation of historical land use scenarios with pollen-based land cover reconstructions. Land, 6: 91
Klein Goldewijk C G M, Battjes J J. 1997. A hundred year (1890–1990) database for integrated environmental assessments (HYDE, version 1.1). Bilthoven: National Institute of Public Health and the Environment. 109
Klein Goldewijk K. 2001. Estimating global land use change over the past 300 years: The HYDE Database. Glob Biogeochem Cycle, 15: 417–433
Klein Goldewijk K, Beusen A, Doelman J, Stehfest E. 2017a. Anthropogenic land use estimates for the Holocene—HYDE 3.2. Earth Syst Sci Data, 9: 927–953
Klein Goldewijk K, Dekker S C, Van Zanden J L. 2017b. Per-capita estimations of long-term historical land use and the consequences for global change research. J Land Use Sci, 12: 313–337
Li B B, Jansson U, Ye Y, Widgren M. 2013. The spatial and temporal change of cropland in the Scandinavian Peninsula during 1875–1999. Reg Environ Change, 13: 1325–1336
Lieskovský J, Kaim D, Balázs P, Boltižiar M, Chmiel M, Grabska E, Király G, Konkoly-Gyuró É, Kozak J, Antalová K, Kuchma T, Mackovčin P, Mojses M, Munteanu C, Ostafin K, Ostapowicz K, Shandra O, Stych P, Radeloff V C. 2018. Historical land use dataset of the Carpathian region (1819–1980). J Maps, 14: 644–651
Lushaj S. 2021. Improving the governance and administration of agricultural land in Albania. Annu Rev Territ Gov West Balk, III: 58–70
Mark B T. 2010. Agriculture in World History. London: Routledge. 208
Markovic D. 2007. Country study Bosnia and Herzegovina. In: EI Moujabber M, EI Bitar L, Raeli M, eds. Study of the Organic and Safety Agriculture in the Adriatic Cross-border Region and of Training Needs. Bari: CIHEAM. 39–57
Marta S, Brunetti M, Manenti R, Provenzale A, Ficetola G F. 2021. Climate and land-use changes drive biodiversity turnover in arthropod assemblages over 150 years. Nat Ecol Evol, 5: 1291–1300
McEvedy C, Jones R. 1978. Atlas of World Population History. Hammondsworth: Penguin Books Ltd. and Allen Lane. 368
Musil J. 1993. Czech and Slovak society: Outline of a comparative study. Czech Soc Rev, 1: 5–21
Pfister C. 1983. Changes in stability and carrying capacity of lowland and highland agro-systems in Switzerland in the historical past. Mountain Res Dev, 3: 291–297
Pinke Z. 2014. Modernization and decline: An eco-historical perspective on regulation of the Tisza Valley, Hungary. J Historical Geogr, 45: 92–105
Pongratz J, Reick C, Raddatz T, Claussen M. 2008. A reconstruction of global agricultural areas and land cover for the last millennium. Glob Biogeochem Cycle, 22: GB3018
Portugal Instituto Nacional De Estatística. 1951. Estatística Agícola. Lisboa
Pounds N J G. 1990. An Historical Geography of Europe. Cambridge: Cambridge University Press. 500
Ramankutty N, Foley J A. 2010. ISLSCP II Historical Croplands Cover, 1700–1992. In: Hall F G, Collatz G, Meeson B, Los S, Brown de Colstoun E, Landis D, eds. ISLSCP Initiative II Collection. Data Set. Tennessee: Oak Ridge National Laboratory Distributed Active Archive Center
Socijalistička Federatina Republika Jugoslavija Savezni Zavod Za Statistiku. 1963. Popis Poljoprivrede 1960. Beograd: Savezni Zavod za Statistiku. 418
Sømme A. 1960. A Geography of Norden: Denmark, Finland, Iceland, Norway, Sweden. Oslo: Cappelen. 363
Wei X Q, Widgren M, Li B B, Ye Y, Fang X Q, Zhang C P, Chen T X. 2021. Dataset of 1 km cropland cover from 1690 to 1999 in Scandinavia. Earth Syst Sci Data, 13: 3035–3056
Wu Z L. 2021. Development of gridded allocation method for historical

cropland based on settlement information and its application (in Chinese). Doctoral Dissertation. Beijing: Beijing Normal University. 1–169

Wu Z L, Fang X Q, Jia D, Zhao W Y. 2020. Reconstruction of cropland cover using historical literature and settlement relics in farming areas of Shangjing Dao during the Liao Dynasty, China, around 1100 AD. Holocene, 30: 1516–1527

Wu Z L, Fang X Q, Ye Y. 2022. A settlement density based allocation method for historical cropland cover: A case study of Jilin Province, China. Land, 11: 1374

Ye Y, Li J, Lu Y, Fang X. 2019. Comparing data sources in six global data sets of historical land cover (in Chinese). J Beijing Norm Univ (Nat Sci), 55: 786–795

Ye Y, Li J, Fang X Q, Zhang D Y, Zhao Z L, Wu Z L, Lu Y J, Li B B. 2023. Reconstruction of cropland change in European countries using integrated multisource data since AD1800. Boreas, 52: 60–77

Zhang C P, Ye Y, Fang X Q, Li H S B, Wei X Q. 2019. Synergistic modern global 1 km cropland dataset derived from multi-sets of land cover products. Remote Sens, 11: 2250

Zhang C P, Fang X Q, Ye Y, Tang C C, Wu Z L, Zheng X, Zhang D Y, Jiang C, Li J, Li Y K, Zhao Z L. 2022a. A spatially explicit reconstruction of cropland cover in China around 1850 C.E. employing new land suitability based gridded allocational gorithm. Quat Int, 641: 62–73

Zhang C P, Ye Y, Fang X Q. 2022b. Development of a global land suitability dataset for cultivation based on physiogeographic factors. J Glob Change Data Discov, 6: 386–394

Zhang D Y, Fang X Q, Yang L E. 2021. Comparison of the HYDE cropland data over the past millennium with regional historical evidence from Germany. Reg Environ Change, 21: 15

Zhang D Y, Lu Y J, Fang X Q, Ye Y, Zhang C P, Zheng X. 2023. Identifying a scenario for preindustrial cropland cover using cultivation data: A case study of France, Germany and Italy. Anthropocene, 43: 100388

Zhao C S, He F N, Yang F, Li S C. 2022. Uncertainties of global historical land use scenarios in past-millennium cropland reconstruction in China. Quat Int, 641: 87–96

Zheng X. 2022. Reconstruction of forest/woodland and grassland cover in Europe over the past millennium (in Chinese). Doctoral Dissertation. Beijing: Beijing Normal University. 1–112

(Editorial handling: Xin LI)