SCIENCE CHINA
Earth Sciences

•RESEARCH PAPER•

July 2020 Vol.63 No.7: 1013–1025
https://doi.org/10.1007/s11430-019-9555-3

Methodology for credibility assessment of historical global LUCC datasets

Xiuqi FANG[1,2*], Wanyi ZHAO[2], Chengpeng ZHANG[2], Diyang ZHANG[2], Xueqiong WEI[3], Weili QIU[2] & Yu YE[1,2]

[1] *Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing 100875, China;*
[2] *Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China;*
[3] *School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China*

Received May 4, 2019; revised October 18, 2019; accepted November 11, 2019; published online April 1, 2020

Abstract Land use-induced land cover change (LUCC) is an important anthropogenic driving force of global change that has influenced, and is still influencing, many aspects of regional and global environments. Accurate historical global land use/cover datasets are essential for a better understanding of the impacts of LUCC on global change. However, there are not only evident inconsistencies in current historical global land use/cover datasets, but inaccuracies in the data in these global dataset revealed by historical record-based reconstructed regional data throughout the world. A focus in historical LUCC and global change research relates to how the accuracy of historical global land cover datasets can be improved. A methodology for assessing the credibility of existing historical global land cover datasets that addresses temporal as well as spatial changes in the amount and distribution of land cover is therefore needed. Theoretically, the credibility of a global land cover dataset could be assessed by comparing similarities or differences in the data according to actual land cover data (the "true value"). However, it is extremely difficult to obtain historical evidence for assessing the credibility of historical global land cover datasets, which cannot be verified through field sampling like contemporary global land cover datasets. We proposed a methodological framework for assessing the credibility of global land cover datasets. Considering the types and characteristics of the available evidence used for assessments, we outlined four methodological approaches: (1) accuracy assessment based on regional quantitative reconstructed land cover data, (2) rationality assessment based on regional historical facts, (3) rationality assessment based on expertise, and (4) likelihood assessment based on the consistency of multiple datasets. These methods were illustrated through five case studies of credibility assessments of historical cropland cover data. This framework can also be applied in assessments of other land cover types, such as forest and grassland.

Keywords LUCC, Global datasets, Credibility, Assessment methodology

Citation: Fang X, Zhao W, Zhang C, Zhang D, Wei X, Qiu W, Ye Y. 2020. Methodology for credibility assessment of historical global LUCC datasets. Science China Earth Sciences, 63: 1013–1025, https://doi.org/10.1007/s11430-019-9555-3

1. Introduction

Land use changes for meeting developmental needs have altered land cover across the world. Land cover change profoundly induces global and regional climate and environmental changes by replacing or disturbing the natural terrestrial ecosystems and biodiversity, and also by affecting regional physical processes on the land surface and global biogeochemical (i.e. carbon, nitrogen) and water cycles (Ruddiman, 2003; Foley et al., 2005; Goldewijk et al., 2007; Pongratz et al., 2008; Pielke et al., 2011; Zhang et al., 2012; Ellis et al., 2013; Li et al., 2014; Gaillard et al., 2018). Thus, land use-induced land cover change (LUCC) is recognized as an important driving force of global change.

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

Reliable historical global LUCC datasets developed on the basis of an in-depth understanding of land use and land cover changes are essential for quantitatively assessing LUCC impacts on global change. Substantial progress has been made in global and regional land use and land cover reconstruction, beginning with the implementation of international LUCC-related research projects in the 1990s. Many global land use/cover datasets demonstrating high spatiotemporal resolutions and covering an extended period of history have been established. Examples include the History Database of the Global Environment (HYDE) developed by the PBL Netherlands Environmental Assessment Agency (Goldewijk and Battjes, 1997; Goldewijk, 2001; Goldewijk et al., 2011, 2017), the SAGE dataset developed by the Center for Sustainability and the Global Environment at the University of Wisconsin (Ramankutty and Foley, 1999), the PJ dataset established by Pongratz et al. (2008), and the KK10 dataset created by Kaplan et al. (2009, 2010). These datasets have generally been derived from the data on contemporary land cover, historical populations, and climate/land suitability, with limited use of data on regional cropland areas from historical records. These datasets have been widely employed to estimate historical carbon emissions and LUCC impacts on regional or global climates, with the aim of better understanding the contribution of human activities to global change (Ruddiman, 2003; Goldewijk et al., 2007; Pongratz et al., 2008; Pielke et al., 2011; Zhang et al., 2012; Ellis et al., 2013; Gaillard et al., 2018). However, existing historical global land use/cover datasets reveal inconsistencies in the amount and distribution of land cover changes over time and space. Moreover, large amounts of data reconstructed using historical records for regions throughout the world have demonstrated inaccuracies in these data (Kaplan et al., 2009; Li et al., 2010, 2013; Leite et al., 2012; He et al., 2013; Zumkehr and Campbell, 2013; Yuan et al., 2017; Yang X et al., 2019). Thus, an important concern among researchers working on historical LUCC and global change centers on improving the accuracy of the historical global land cover datasets, which is also a goal of the ongoing LandCover6k program (Gaillard et al., 2015, 2018).

Assessing the credibility of existing global land cover datasets is a precondition for improving data quality. A credibility assessment enables different data qualities to be distinguished, which are data that are credible, data that need to be updated or revised, and data that need to be reconstructed from very beginning. With field validation samples, various methods for testing and improving the credibility of the modern global land cover data have been developed (Chen and Chen, 2018). However, it is difficult to assess the credibility of historical global land cover data applying currently used methods for assessing contemporary data. This is because the actual past land cover data (referred to as the "true value") that serves as the baseline of the credibility assessment is not directly accessible and needs to be reconstructed in most cases. Moreover, historical and natural records available for land cover reconstruction are very limited, and a widely accepted method for such an assessment remains to be developed. Reconstructed regional land cover data derived from historical records are regarded as the baseline in most of the existing studies on credibility assessments of historical global land cover data. The regional reconstructed land cover data are regarded as approximating the "true values", and credibility is assessed in terms of deviations of the data in the global land cover datasets from the reconstructed regional data (Li et al., 2010; He et al., 2012, 2013, 2018; Yuan et al., 2017; Wei et al., 2019; Li et al., 2019). However, the difficulty of obtaining proxy data for the reconstruction makes the process of quantitative reconstruction of regional land cover challenging. The temporal and spatial ranges covered by the existing reconstructed data on regional land cover are limited, and in most cases, the credibility assessments of data within the global land cover datasets has to be done in the absence of any regional reconstructed data. Therefore, there is a need to solve the problem of how to assess the credibility of the data in the global land cover datasets in light of limited available evidence.

Here we propose a methodological framework for credibility assessment of the historical global land use/cover datasets. The four assessment methods discussed in this paper correspond to the different types and characteristics of available evidence or indicators used for historical land cover data assessments. To further illustrate these methods, we present five case studies of credibility assessments. Although these case studies specifically address cropland cover changes, the entailed assessment methods are also applicable to other land cover types, such as forest and grassland cover (He et al., 2019; Yang F et al., 2019).

2. A methodological framework for credibility assessment of historical global land cover data

The credibility of historical global land cover data can be assessed in terms of their spatial and temporal dimensions. The temporal dimension refers to the cropland amount changes over time for a given spatial unit. The spatial dimension refers to the cropland distribution over space for a given time section (Figure 1).

The credibility of historical land cover data can be measured in various ways. Theoretically, the criteria for determining the credibility of historical global land cover data are the extent of similarities or differences of the data to the actual land cover (the "true value"). A smaller deviation from the "true value" corresponds to a higher degree of

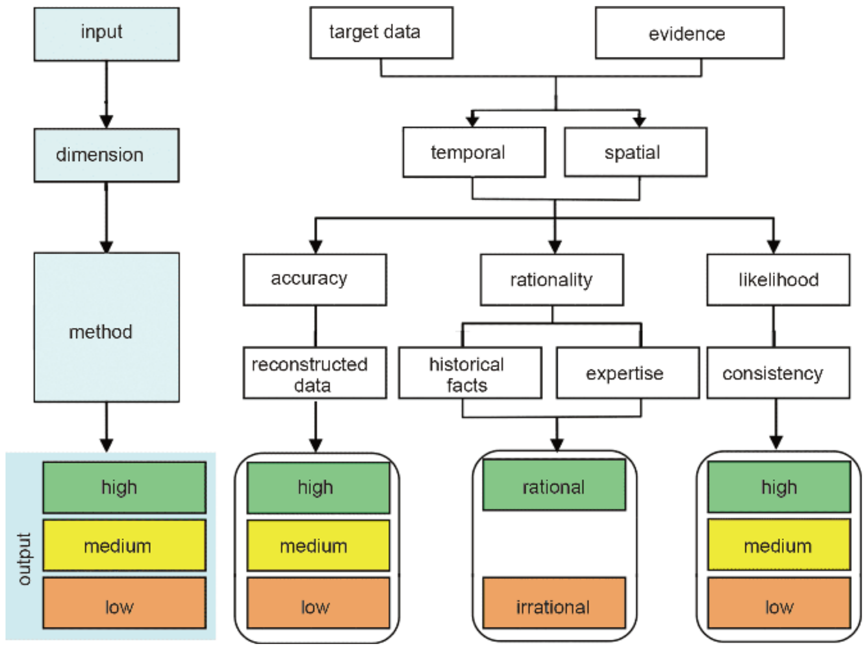

Figure 1 Methodological framework for credibility assessment of historical global land cover data.

credibility. However, the actual land cover ("true value") has already been relegated to history, and its approximation to the true value can only be obtained through a process of reconstruction. Moreover, the limited availability of historical and natural records hinders reconstruction. Because the land cover data to be assessed are reconstructed, the evidence used for credibility assessment should be sufficiently close to the "true value," or it should at least reflect key features of the "true value". In addition, it should be independent of the evidence used in the datasets to be assessed. Quantitative regional reconstructed data that are thought to be a good approximation of the "true value" usually serve as the baseline in quantitative accuracy assessments. When several land cover datasets are available, but it is difficult to distinguish which one is more credible, differences in the consistency of the amount of land cover changes over time or their spatial distribution among these datasets can be used to measure the credibility of the data. If, however, the assessment is solely based on qualitative evidence that partially reflects the key characteristics of the "true value," then a rationality-based approach is used to assess the credibility of the target land cover data. Here, we propose four types of credibility assessment methods (Figure 1) tailored to differences in the data to be assessed and the available evidence. Given restrictions regarding the availability of evidence that can be used for conducting credibility assessments of historical global land cover data, both the goal of the assessment and the characteristics of the available evidence in given regions and periods should be considered when choosing a specific assessment method.

3. Accuracy assessment based on quantitatively reconstructed regional land cover data

3.1 The accuracy assessment method

The accuracy assessment is a quantitative credibility assessment that uses quantitatively reconstructed regional land cover data derived from historical or natural records as the baseline. In this method, the reconstructed regional cropland, forest, or grassland cover data are assumed to approximate the "true value" of the corresponding historical land cover type. The accuracy of the data in the historical global land cover datasets for the study area is assessed by calculating their deviation from quantitatively reconstructed regional land cover data. This method has been employed in most of the existing credibility assessments of historical global land cover data (Li et al., 2010, 2013; He et al., 2012, 2013, 2018; Wei et al., 2016, 2019; Yuan et al., 2017).

A regional historical land cover dataset derived from historical or natural records is a prerequisite for implementing an accuracy assessment. This dataset can be reconstructed from the original historical records, or it can be mined from published papers or existing datasets. The first step for conducting an accuracy assessment entails unifying the time sections and spatial resolutions of the reconstructed regional land cover data and the data in the global land cover datasets for a given area. Subsequently, the accuracy of the data in the global land cover datasets is ascertained by measuring the absolute or relative deviation of the data from the reconstructed regional data. A smaller deviation corresponds to greater degree of accuracy or reliability of the land cover data in the global dataset. As a final step, the data in the global land cover dataset for a given area can then be classified into high, medium, and low levels of accuracy according to the extent of this deviation. However, there are still no commonly accepted criteria or thresholds of accuracy, which vary on a case-by-case basis in existing studies.

The advantage of using the accuracy assessment method is that it provides a quantitative assessment of the credibility of the data in the historical global land cover datasets for a given area. A prerequisite for applying this method is the availability of reconstructed historical regional land cover data for the study area. However, quantitative reconstruction of regional land cover is a challenge that can only be implemented in a limited area where ample historical and natural records required for the reconstruction are available. Thus, this method has limited applicability, given that the land cover data in most regions of the world have not been reconstructed.

3.2 Case study 1: An accuracy assessment of the quantity of HYDE cropland in Germany over the last 1000 years

Bork et al. (1998) reconstructed Germany's cropland fraction for the period AD600–2000, which was subsequently used as a baseline for conducting an accuracy assessment of the amount of HYDE cropland in Germany over the last 1000 years. The cropland fractions of Germany were extracted from the HYDE 3.1 and 3.2 (Goldewijk et al., 2011, 2017). The relative deviations (RDs) of the cropland fraction extracted from each of these datasets from that by Bork for Germany over the last 1000 years were then calculated (Figure 2).

Both the size and trend of the HYDE 3.1 cropland fraction for Germany over the last 1000 years differed markedly from Bork's cropland fraction (Figure 2), being significantly lower than those of the latter for all time sections. The RD of the former from the latter was more than 20% lower prior to AD1940, more than 50% lower prior to AD1830, and more than 75% lower prior to AD1300. Bork's curve showed evident fluctuations for the cropland fraction in Germany over the last 1000 years, but the HYDE 3.1 cropland fraction did not show this fluctuation. Therefore, the amount of cropland in Germany over the last 1000 years in the HYDE 3.1 is considered a low level of accuracy.

The size and trend of the HYDE 3.2 cropland fraction for Germany over the last 1000 years were similar to those of Bork's cropland fraction (Figure 2). This is because the amount of cropland in Germany was revised in the HYDE 3.2 with reference to the reconstructed data by Bork et al. (1998). The RD of the HYDE 3.2 cropland fraction from that reconstructed by Bork decreased to −13–15%, which could generally be considered to indicate a high level of accuracy. Prior to AD1700, the relative deviation was minor. However, because the temporal resolution of the HYDE 3.2 data is set at 100 years, the time of the appearance of the peaks and valleys of the HYDE 3.2 cropland fraction differ slightly from those of Bork's cropland fraction. In particular, fluctuations in the cropland fraction caused by major historical events with durations of less than 100 years, such as the Thirty Years' War (AD1618–1648), are difficult to discern. During the period AD1700–1940, the RD of the HYDE 3.2 cropland fraction from that reconstructed by Bork fluctuated

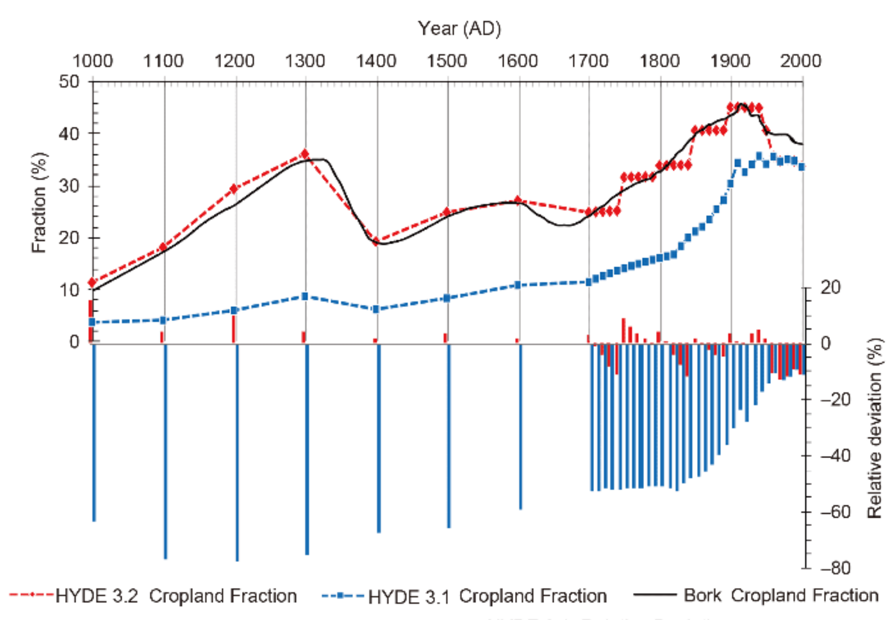

Figure 2 Comparison of the HYDE cropland fractions of Germany over the last 1000 years (Goldewijk et al., 2011, 2017) with the reconstructed regional data (Bork et al., 1998).

by −12–9% over 100-year cycles, because even though a high temporal resolution of a decade was applied, the HYDE 3.2 cropland fractions remained the same during each half century. This result for the HYDE 3.2 cropland fraction for most time sections during the period AD1700–1940 was even less accurate than that was prior to AD1700.

3.3 Case study 2: The accuracy assessment of the spatial distribution of the HYDE 3.2 cropland cover in the North China Plain over the last 300 years

Wei et al. (2019) reconstructed a gridded (10 km × 10 km) cropland dataset for the North China Plain (NP dataset) covering the last 300 years on the basis of historical records. This dataset has seven time sections, namely the late 17th century, the 18th century, the 19th century, AD1916, AD1933, the 1980s, and AD2008. The historical records used for cropland cover reconstruction, entailing a county-level spatial resolution, were collected from different sources. The data for the period of the Qing Dynasty (AD1644–1911) were extracted from gazetteers, while those for the Republic of China era (AD1912–1949) were extracted from gazetteers and statistical records. Data for the period of the People's Republic of China (AD1949–present) were sourced from survey and statistical records. To enable comparison with the reconstructed NP dataset, the cropland data in the North China Plain, extracted from the HYDE 3.2 (Goldewijk et al., 2017) in seven time sections, namely AD1700, AD1750, AD1850, AD1910, AD1930, AD1980, and AD2008. The cropland fractions' RDs of the HYDE 3.2 from the NP dataset in the North China Plain were calculated for 10 km×10 km grids. The accuracy of the HYDE 3.2 cropland cover in the North China Plain was then assessed on the basis of the RD values (Figure 3).

From version 3.1, the HYDE took the reconstructed cropland data of China over the last 300 years by Ge et al. (2008) as reference (Goldewijk et al., 2011). Thus, differences in the amounts of cropland in North China between the HYDE 3.2 and the NP dataset were small. The RD between the two datasets varied between −10% and 13% in six of the seven time sections. The biggest RD was 16.60% for the AD1980 (Figure 3). Therefore, on the whole, the amount of HYDE 3.2 cropland in the North China Plain is considered a high level of accuracy.

Differences in the spatial distribution of the cropland fractions in the HYDE 3.2 and the NP datasets were, however, significant, even though both datasets indicated that over the last 300 years, the cropland in North China was mainly distributed in the plain area, with a smaller proportion distributed in the surrounding mountainous areas. In the five time sections between AD1700 and AD1930, the grids with absolute RD values above 50% accounted for 49–57% of the total grids, whereas the grids with an RD of −10–10% accounted for just 8–12% of the total grids. In terms of the spatial distribution of the RD in the study area, the HYDE 3.2 shows a systematic bias, with the cropland fractions generally being lower in the northern part and higher in the southern part compared with these fractions in the NP dataset. Moreover, 76–86% of the grids with an RD above 50% were located south to the Yellow River, whereas 63–67% of the grids with an RD below −50% were located north to the Yellow River. In addition, the HYDE 3.2 cropland fraction was significantly higher in locations along the major rivers like the Yangtze, Yellow, and Huaihe rivers. For the two modern-day time sections, namely AD1980 and AD2008, the differences in the spatial distribution between the two datasets were lower than those for the earlier historical periods. The grids with a lower RD (between −10% and 10%) increased to 13–21% of the total grids. Conversely, the grids with an absolute RD value above 50% decreased to 29–35% of the total grids. In terms of the spatial distribution of the RD in the study area, the HYDE 3.2 cropland fraction was generally smaller in the plain area but larger in the surrounding mountainous and plateau areas compared with the corresponding fractions in the NP dataset. In the plain area, 31–66% of the grids had an RD lower than −10%, whereas 48–70% of the grids had an RD greater than 10% in the surrounding mountains and plateaus.

Of the various factors that could account for the above described spatial differences, the following ones are ascertainable. The gridded cropland cover data of the NP dataset reconstructed by Wei et al. (2019) demonstrated less bias in terms of spatial allocation compared with the HYDE 3.2 because it is derived from source data with a county-level rather than a province-level resolution, as reconstructed by Ge et al. (2008) and taken as reference in the HYDE 3.2. The assumption in the HYDE 3.2 is that land located near the river is preferentially reclaimed, whereas the river factor was not considered in the spatial allocation of the NP dataset. Thus, more cropland was allocated to the grids near the river in HYDE 3.2, but this is evidently inconsistent with the actual agricultural history of the North China Plain.

4. Rationality assessment based on regional history and expertise

4.1 The rationality assessment method

The rationality assessment is a qualitative assessment. The assumption of this method is that the credible land cover data should firstly be the rational data. The rationality of specific land cover data can be assessed in two ways. The first entails an assessment of whether or not the data can be verified through relevant historical evidence of the region. The second entails a determination of whether there are favorable natural environments and socioeconomic conditions for the

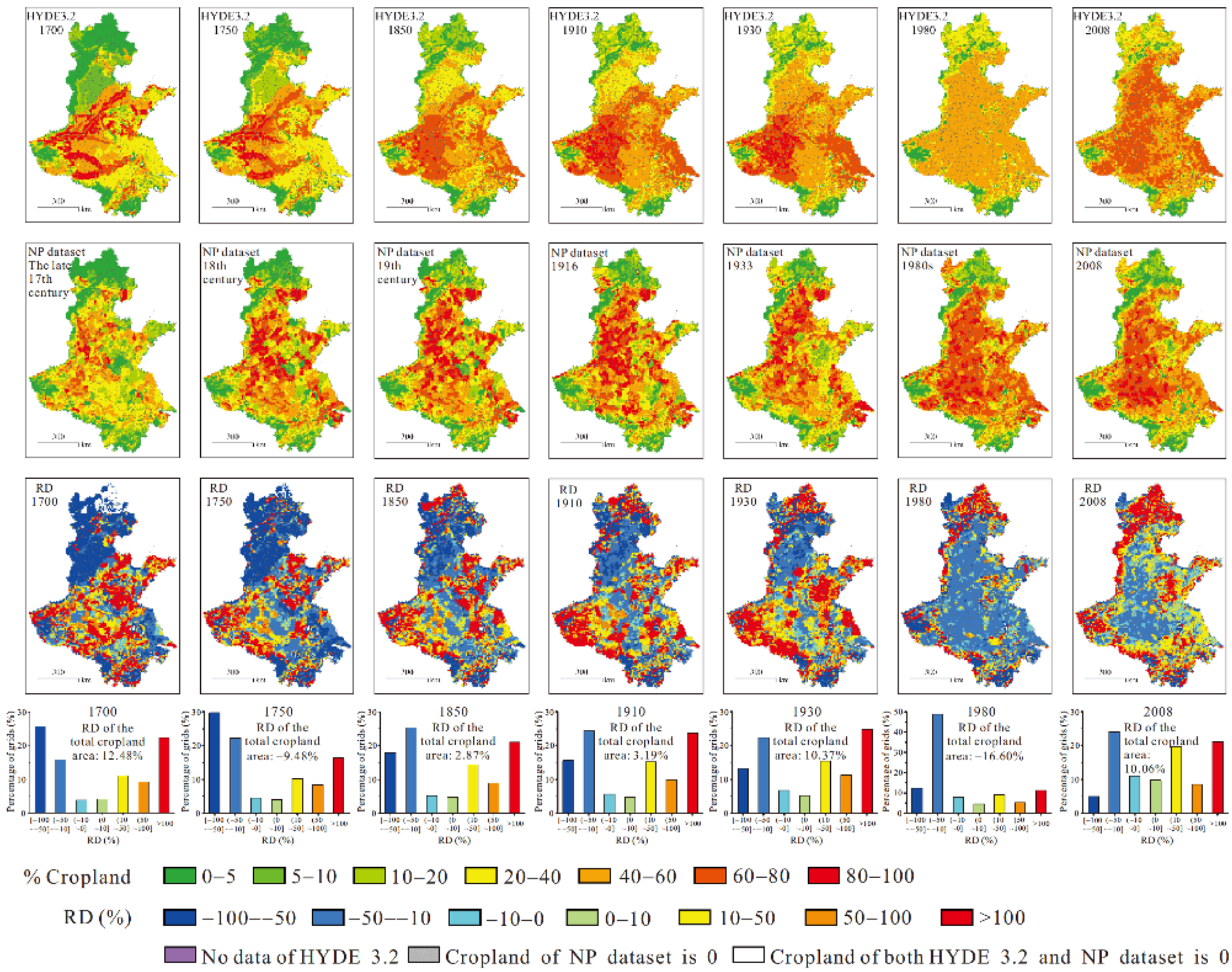

Figure 3 Distribution of cropland fractions in North China over the last 300 years based on the HYDE 3.2 (Goldewijk et al., 2017) and the NP dataset reconstructed by historical records Wei et al. (2019) and their relative deviation (RD).

land use or cover known on the basis of relevant expertise. Accordingly, we propose two rationality assessment methods for historical global land cover data. The first is to assess the extent of rationality through a comparison of the temporal and spatial conformity of the land cover data from the global dataset with the historical facts relating to regional development. The second is to assess the rationality upon which the expertise is based through an examination of how the data, assumptions, and methods of global land cover datasets match with the related spatial and temporal rules of nature and society, such as features of the natural environment or agricultural characteristics. The rationality assessments lead to outcomes that are determined to be either rational or irrational.

4.1.1 The rationality assessment based on regional historical facts

Historical facts relating to regional agricultural development are used as evidence for a rationality assessment. A variety of historical or archeological records could be used directly or indirectly to indicate the characteristics of regional agricultural development. Examples of direct records include historical descriptions of processes of regional land reclamation and agriculture development and archeological relics of crops and farming tools. Examples of indirect records include descriptions of land policies, advancing technologies, population and migration characteristics, major social events, settlement relics and sediments relating to the region's agricultural history.

Evidence of key historical processes within the regional agriculture can be extracted from systematically collected historical and natural records relating to the history of regional agriculture, for example, major events, critical turning points, and spatial patterns associated with each time section. The rationality of the data of the global land cover dataset is assessed through a comparison of the extent of their spatio-temporal similarities with relevant historical facts relating to agricultural development in the region.

4.1.2 The rationality assessment based on expertise

The rationality assessment based on expertise is an indirect assessment method. In some regions or during some periods for which fewer quantitative or qualitative historical land use or land cover records are available, this method can be employed for indirectly assessing the rationality of the land cover data from global datasets. This method entails the use, as evidence, of known principles or facts regarding natural or social factors and their relations to land use/cover within a region. The rationality of the land cover data from global datasets can be assessed in terms of their spatial and temporal dimensions by examining whether the land cover data matches the relevant principles or facts in the data sources used for reconstruction and in the methods, and assumptions relating to data conversion and spatial allocation. The rationality of the data can also be assessed by examining whether the spatiotemporal distribution of the land cover data are in agreement with that of the natural and social facts. The following are examples of approaches for conducting the assessments, as mentioned above.

(1) For a type of land cover within a certain area or period, the rationality can be assessed in terms of the tolerable natural conditions of that type resulting from the environmental differences and changes over time and space.

(2) The rationality of the methods and assumptions used in historical land use/cover reconstruction or allocation can be assessed according to whether they conform to the relevant natural or social principles, whether they can be supported by modern observations or historical records, and whether they exhibit spatial or temporal generality.

(3) The rationality of the data from the global land cover datasets can be assessed according to whether the spatiotemporal distribution of the data matches that of the factors affecting land cover, such as natural features relating to temperature, precipitation, topography and rivers, and social factors associated with the stages of economic and technological development. For example, the HYDE 3.2 cropland fraction of Germany during the period AD1700–1950 revealed a stepwise change every half century (Figure 2). This change would be considered irrational from the perspective of the general characteristics of regional development. Moreover, the evidently uneven pattern in the distribution of cropland fractions extracted from HYDE 3.2 cropland fractions in the northern and southern parts of the North China Plain during the period AD1700–1930 does not tally with the relatively even natural environment in the North China Plain (Figure 3).

4.2 Case study 3: A rationality assessment of the distribution of the HYDE 3.2 cropland cover in Northeast China over the last 1000 years

This case study was conducted to assess the rationality of the HYDE 3.2 cropland cover data (Goldewijk et al., 2017) in Northeast China (37°–55°N, 120°–136°E) over the last 1000 years. The records used as evidence were those relating to changes in the spatial distribution and amount of settlement relics from the Liao (AD916–1125) to Ming (AD1368–1644) dynasties, and those of towns used as the administrative headquarters during the Qing Dynasty (AD1644–1911).

In ancient Northeast China, the establishment of permanent settlements was closely associated with settled farming activities. The amount and spatial distribution of the settlements could reflect changes in agricultural development and in agricultural areas (Jia et al., 2018). During the Liao (AD916–1125) and Jin (AD1115–1234) dynasties, settlements were numerous. The northern boundaries of the settlements during these two dynasties extended as far north as 47°N, which is very close to the contemporary northern agricultural boundary. However, over the Yuan (AD1279–1368) and Ming (AD1368–1644) dynasties, the number of ancient settlements decreased significantly, and the northern boundaries of the settlements shifted southward by 3–4 latitudinal degrees compared with their positions during the Liao and Jin dynasties (Figure 4).

The actual distribution of the agricultural area during the Qing dynasty is not reasonably represented by the amount and spatial distribution of the settlement relics alone because most of the settlements in Northeast China that were established during this period are still in use (Zeng et al., 2011). However, the areas cultivated during the Qing Dynasty could be identified from the location and the build-up time of the newly established settlements. Among these settlements, the towns as administrative headquarters were set up to rule the Han people who migrated into Northeast China and engaged in agricultural activities. These towns in different administrative levels, namely Fu (district level), Zhou (district or county level) and Xian (county level), may be indicative of the change in the amount and distribution of agricultural areas over time and space during the Qing Dynasty in the three provinces (Liaoning, Jilin, and Heilongjiang) of Northeast China. The development of these towns reveals significant differences in land reclamation before and after AD1860. Before AD1860, the number of towns slowly increased in the three provinces. In AD1700, the towns were only distributed in Liaoning Province. Up to AD1860, only a few newly established towns were distributed in Jilin Province and in the southern part of Heilongjiang Province. The newly established towns expanded relatively fast from AD1860 to AD1895 and even rapidly during AD1895–1911. More than 75% of the towns in Northeast China that came up during the Qing Dynasty were established during AD1860–1911. Specifically, 75.5% and 87.8% of the towns in Jilin and Heilongjiang provinces, respectively, were established during AD1901–1911 (Fang et al., 2005) (Figure 4).

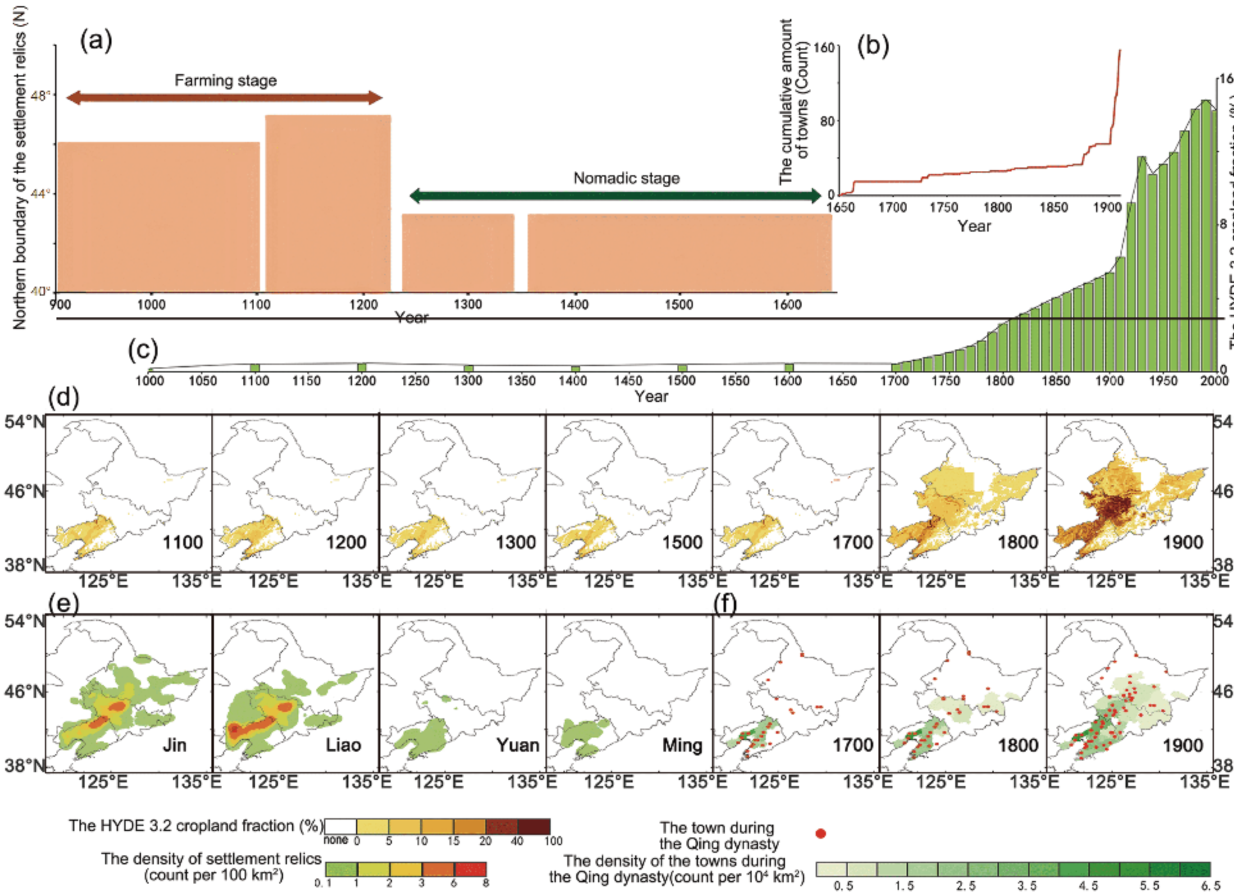

Figure 4 Comparison of the temporal and spatial distribution of HYDE 3.2 cropland fractions over the last 1000 years and settlement relics (the Liao dynasty to the Ming dynasty) and towns (administrative headquarters) during the Qing dynasty in Northeast China. (a) The northern boundary of the settlement relics in Northeast China (37°–55°N, 120°–136°E) from the Liao to the Ming Dynasty (Jia et al., 2018); (b) the cumulative number of towns (administrative headquarters) in the three provinces of Northeast China during the Qing Dynasty (Fang et al., 2005); (c) the HYDE 3.2 cropland fraction in the three provinces of Northeast China during the last 1000 years (Goldewijk et al., 2017); (d) the spatial distribution of the HYDE 3.2 cropland fractions in the three provinces of Northeast China during the last 1000 years (Goldewijk et al., 2017); (e) the spatial distribution of settlement relics in Northeast China from the Liao to the Ming Dynasty (Jia et al., 2018); (f) and the spatial distribution of the towns (administrative headquarters) in Northeast China during the Qing Dynasty (Fang et al., 2005).

Spatial and temporal changes in the amount and distribution of HYDE 3.2 cropland fractions in Northeast China over the last 1000 years evidently differ from the facts indicated by the number and distribution of the settlement relics. From a temporal perspective, the HYDE 3.2 cropland fraction in Northeast China has increased continuously over the last 1000 years. However, this increase neither reflects the peak in land reclamation that occurred in the Liao and Jin dynasties, nor the sharp decrease in the agricultural area during the transition from the Jin to the Yuan dynasties. The HYDE 3.2 cropland fraction during the Qing Dynasty showed an almost linear increase that is inconsistent with the significant differences in the land reclamation rates indicated by the increase of towns before and after AD1860. From a spatial perspective, during the Liao and Jin dynasties, the HYDE 3.2 cropland was only distributed in Liaoning Province. This is obviously inconsistent with the historical facts indicated by the settlement relics that the large expanses of cropland existed in Jilin and Heilongjiang provinces during that period. Moreover, in AD1800 and in AD1900, the northern boundaries of the HYDE 3.2 cropland were also significantly further north compared with the actual distribution of agricultural areas, as indicated by the towns during the Qing Dynasty (Figure 4).

The above results reveal that changes in the HYDE 3.2 cropland in Northeast China over the last 1000 years are irrational in terms of its spatial and temporal distribution. One of the reasons is that Northeast China's unique history of regional development is not factored into the HYDE 3.2. The HYDE 3.2 did not take account of the flourishing agriculture during the Liao and Jin dynasties attributed to the farming people migrating in and the ethnic minority groups adapting the farming techniques. HYDE 3.2 did not also consider the dramatic transformation in land use patterns from farming to pasturing following this area governed by Mongolian nomads during the Yuan dynasty (Jia et al., 2018). Moreover, the HYDE 3.2 did not capture the rapid expansion of cropland area and its gradual northward expansion as a result of extensive migration into Northeast China after AD1860 following the repeal of the "farming-prohibition" policy in Northeast China (Fang et al., 2005; Ye et al., 2009).

4.3 Case study 4: A rationality assessment of the HYDE 3.2 cropland cover in the coastal plain adjoining the Bohai Sea over the last 7000 years

In the HYDE 3.2, historical cropland was allocated within the scope of modern cropland distribution and the part of build-up areas suitable for reclamation. A higher degree of suitability corresponds to a higher cropland fraction (Goldewijk et al., 2011, 2017). This allocation method is only reasonable if the natural environment has remained unchanged over time. Once the environment changed, the order of the grids in suitability for reclamation should be reranked. Consequently, the rationality of the assumption underlying cropland allocation and its outcomes would be questioned. Based on reconstructed historical changes in the Bohai Sea coastline and in the Yellow River channel, this case study entailed an assessment of the rationality of HYDE 3.2 cropland in the region of the contemporary coastal plain adjoining the Bohai Sea over the last 7000 years by examining the validity of the premise of cropland allocation.

The coastal plain has been changing over time resulting from the advances and retreats of coastline, which is affected by both sea-level changes and river delta sedimentation. Thus, the amount of land that was suitable for reclamation in the past coastal plain would differ from the contemporary one. Over the last 7000 years, the western coastal plain adjoining the Bohai Sea has continuously expanded about 31200 km^2 as a result of the phased retreats of the coastline (Xue, 2009; Shang et al., 2018). In the absence of available land, cropland cannot be developed. Thus, previously submerged land in the western coastal plain adjoining the Bohai Sea could not have been cropland until these lands emerged, even if they are now suitable for cropland. The spatial distribution of HYDE 3.2 cropland cover in this area was compared with coastline locations in 5000BC and AD1200 (Figure 5). This comparison revealed that the HYDE 3.2 cropland data is unreasonable because changes in the coastline within the study area were not accounted for. Some of the cropland during both 5000BC and AD1200 were still submerged and therefore unavailable for cultivation.

In addition, the HYDE 3.2 cropland fractions in 5000BC and AD1200 were significantly higher in areas located along the contemporary Yellow River channel (Figure 5). This is clearly an outcome of the allocation of more cropland to the grids situated near the river in the HYDE 3.2 premised on the assumption that land near the river would be preferentially reclaimed. Regardless of whether or not the river preference was rational, the higher cropland fraction for the area near the contemporary Yellow River during 5000BC and AD1200 is irrational because the Yellow River did not flow into the contemporary channel until AD1855. Around 5000BC (the early Holocene), the Yellow River flowed on the north side

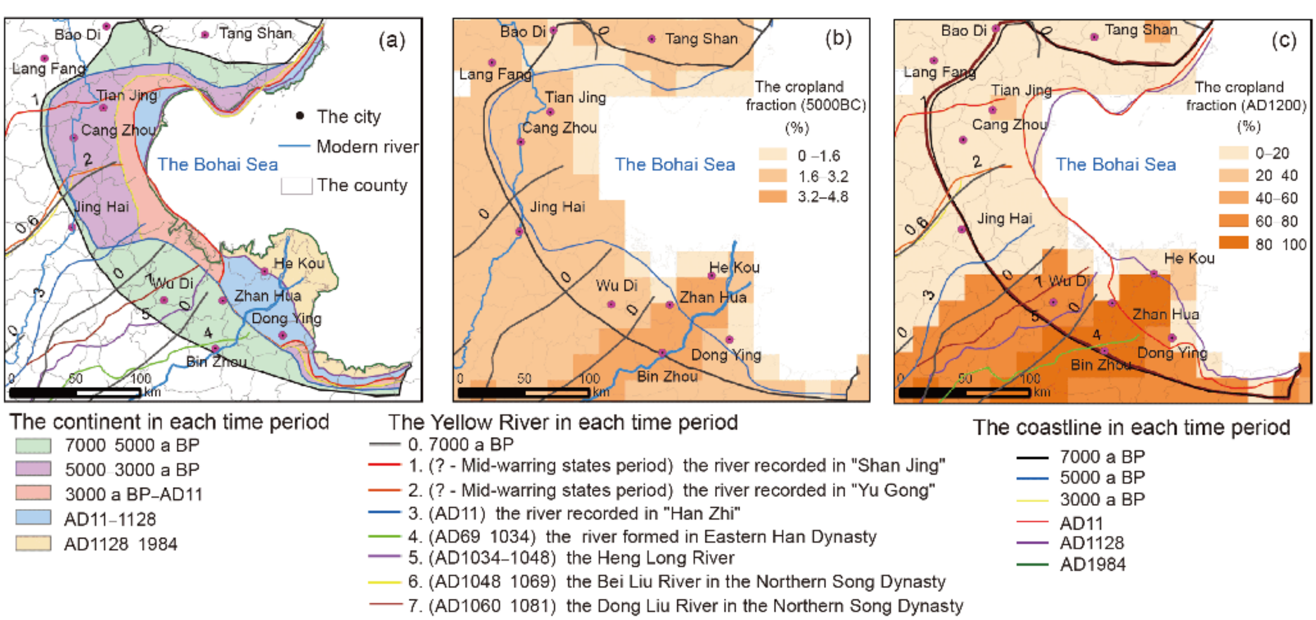

Figure 5 Comparison of the historical coastlines and the Yellow River channels with the distribution of HYDE 3.2 cropland in the western coastal plain adjoining the Bohai Sea. (a) The Bohai coastlines (Xue, 2009) and Yellow River channels (Niu et al., 1994; Liu, 2012) during different historical time sections; (b) the HYDE 3.2 cropland fraction in 5000BC (Goldewijk et al., 2017); (c) the HYDE 3.2 cropland fraction in 1200 AD (Goldewijk et al., 2017).

of the contemporary Yellow River and entered the Bohai Sea (Wu and He, 1991). In AD1200, it flowed southward and entered the Yellow Sea, because it captured the Huaihe River's channel after AD1128 (Niu, 1994).

5. Likelihood assessment based on the level of consistency of the land cover data from multiple datasets

5.1 The likelihood assessment method

The likelihood assessment applies the consistency of land cover data obtained from more than one dataset as an indicator of credibility. This method is employed when it is difficult to determine the credibility of more than one existing dataset. This method is based on the assumption that the methods and data of each land cover dataset demonstrate some degree of rationality but uncertainties and limitations as well. It means that a considerable proportion of the land cover data in each dataset may be consistent with or approximate the actual land cover ("true value"). Simultaneously, a certain proportion of the data are inconsistent with the actual land cover because of the methodological weakness or uncertainties relating to the data source. Considering each global land cover dataset as an expert determination regarding the actual land cover, the likelihood of the credibility of the land cover data for a given spatial or temporal unit is assessed by measuring the degree of consistency of the data on the unit derived from multiple datasets. According to the degree of consistency of the data, ranging from high to low, the likelihood level is classified as high, medium, or low.

Temporal or spatial overlay analysis can be used to measure the consistency of land cover data from multiple datasets. The consistency is described by the grid value, which refers to the number of datasets or the proportion of the total datasets that have a particular land cover type or fraction range of a land cover type within the given spatial and temporal unit. A higher grid value corresponds to a higher degree of consistency and thus to a higher likelihood of credibility.

5.2 Case study 5: The likelihood assessment based on the consistency of cropland cover data in China derived from 10 modern global land cover datasets

Cropland grids (1 km×1 km) were extracted from 10 widely used modern global land cover datasets to assess the consistency of cropland cover data in China. The cropland grids were spatially overlaid using the Map Algebra tool in ArcGIS to calculate the grid value within a 0–10 range. A "0" value indicates that the grid was not considered to contain cropland within any dataset, and a "10" value indicates that the grid was considered to contain cropland within all of the datasets. A higher grid value corresponded to a higher likelihood of the distribution of cropland within the grid. However, there is no commonly accepted standard for determining how consistency can be used to measure credibility. In this case, the cropland grids were divided into three levels of credibility according to the grid values. Values of 1–3 represented a low degree of consistency and therefore a low likelihood of credibility; values of 4–5 represented a moderate degree of consistency and therefore a moderate likelihood of credibility; and values of 6–10 represented a high degree of consistency and therefore a high likelihood of credibility (Figure 6).

Of the total number of cropland grids in China, 42% and 43%, respectively, demonstrated high and low degrees of consistency and 15% demonstrated a medium degree of consistency. This means that only 42% of all of the cropland grids evidenced a high likelihood of credibility. At the provincial scale, more than 50% of cropland grids evidencing a high degree of consistency were located in 20 provinces. These provinces were situated in eastern and central China, mainly in the North China Plain, the Chengdu Plain, and the Northeast China Plain. The proportions of cropland grids that demonstrated both high and low degrees of consistency were less than 50% in six provinces, namely Shaanxi, Gansu, Ningxia, Zhejiang, Fujian, and Yunnan. These provinces are mainly located in the agriculture–pasture transitional zone of northern China and in the Southern China Hills. The proportion of cropland grids demonstrating a low degree of consistency was more than 50% in six other provinces, namely Xizang, Qinghai, Xinjiang, Sichuan, Taiwan, and Inner Mongolia, which are mainly located in the agriculture-pasture transitional zone of North China, the Southern China Hills, the Qinghai-Tibet Plateau, and the mountains and piedmonts in Northwest China.

The cropland grids demonstrating a high degree of consistency were concentrated in China's major agricultural areas. These areas generally have favorable agro-climatic conditions and flat topographies suitable for agriculture. The cropland grids in these areas are easily identified in remote sensing images because they distribute continuously with high level of reclamation intensity, are planted steady crop types year-by-year without large fallow areas. By contrast, the cropland grids with low degrees of consistency are mostly located in areas with moderate to low levels of reclamation intensity that have relatively harsh agro-climatic conditions and complex topographies. These grids mainly belong to the mixed cropland grids in the Boolean datasets. The low degree of consistency of the cropland grids may be attributed to the uncertainty entailed in the identification of these mixed cropland grids and the different cropland classification criteria used for each dataset.

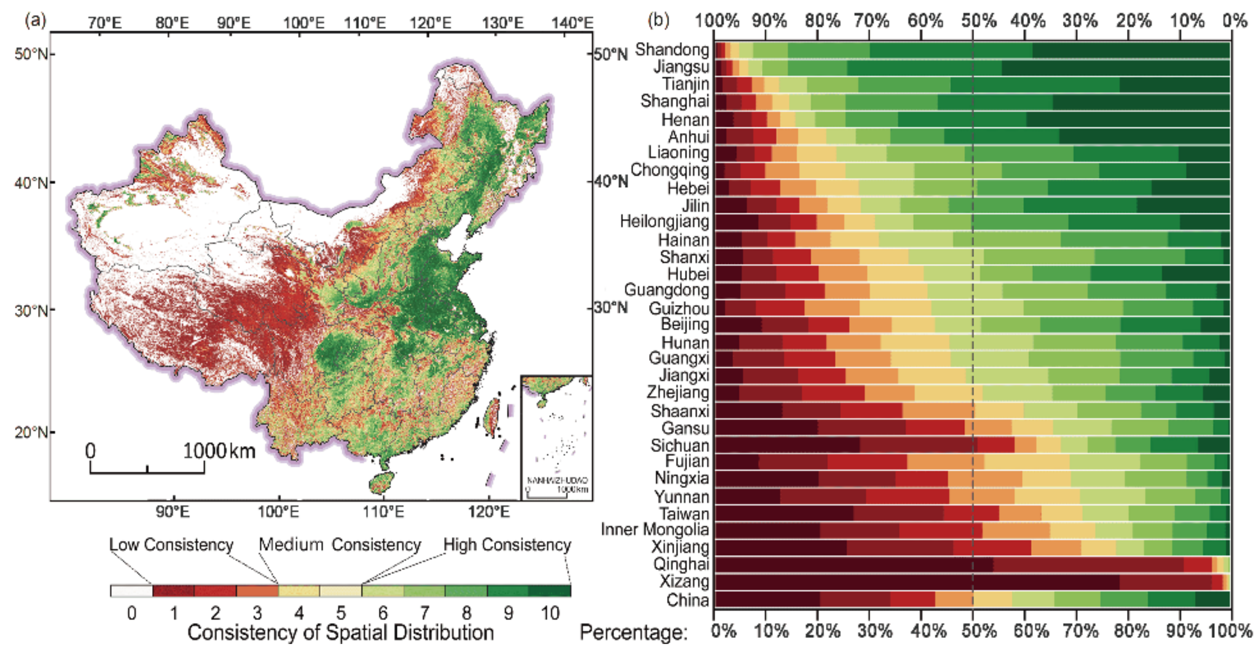

Figure 6 Consistency of cropland distribution in China described by the spatial overlay of 10 modern global cropland datasets. (a) The spatial distribution of grid values; and (b) the percentages of different grid values of 32 provincial units, excluding Hong Kong and Macao. Of the 10 modern global land cover datasets, seven were Boolean datasets whose cropland grids referred to cropland types as well as to mixed types of cropland and natural vegetation: IGBP-DISCover (Loveland et al., 2000), GLC-UMD (Hansen et al., 2000), GLC-MODIS (Friedl et al., 2002), GLC2000 (Bartholomé and Belward, 2005), GLCNMO (Tateishi et al., 2001), ESA-CCI_LC (Bontemps et al., 2013), and GlobeLand30 (Chen et al., 2015). The remaining three were fusion datasets whose cropland grids has a cropland fraction greater than or equal to 1%: GLC-Consensus (Tuanmu and Jetz, 2014), GLC-Share (Latham et al., 2014), and Hybrid Cropland (Fritz et al., 2015).

6. Conclusion

LUCC has influenced, and is influencing, many aspects of regional and global environments. More credible historical global land use/cover datasets, entailing a comprehensive understanding of land use and land cover histories are required for quantitative assessments of the impacts of LUCC on global change. Since the 1990s, several historical global land use/cover datasets have been developed. However, existing historical global land use/cover datasets demonstrate inconsistencies relating to temporal and spatial changes in the amount and distribution of land cover, and the use of historical record-based reconstructed data has shown inaccuracies in global data for many regions worldwide.

Assessing the credibility of existing historical global land cover datasets is a precondition for improving data quality. To address the challenge of limited availability of historical land cover data or evidence, we have proposed a methodological framework for conducting credibility assessments of these data. We outlined four different credibility assessment methods that respond to differences in the data to be assessed and in available evidence and presented five case studies of credibility assessments of cropland cover.

(1) Credibility assessments of historical global land cover data entail an examination of temporal and spatial changes in the amount and distribution of land cover data. However, it is difficult to assess the credibility of the data because actual land cover ("true value"), which is regarded as the baseline in credibility assessments, also needs to be reconstructed. Apart from a few accuracy assessments that have entailed the use of quantitative reconstructed data as approximations of the "true value," credibility has mostly to be assessed from the perspectives of rationality and consistency.

(2) Accuracy assessment is a quantitative assessment based on quantitatively reconstructed regional land cover data. The assumption underlying this method is that reconstructed quantitative regional land cover data that is based on historical or natural records a fine approximation the actual historical land cover. The accuracy of the data from the global land cover dataset is quantitatively assessed based on the extent of deviation from the reconstructed regional data in the study area.

(3) Rationality assessment is a qualitative assessment. Being rational is the underlying premise of the credible land cover data. There are two assessment methods according to the type of available evidence, the regional historical facts-based rationality assessment and the expertise-based rationality assessment. Data credibility is assessed by comparing similarities in data derived from the global land cover dataset and historical facts of regional development, or by examining the extent of the match between the data, assumptions, and methods of global land cover datasets with established

expertise.

(4) In likelihood assessment, the credibility of the land cover data for given spatial or temporal units is inferred according to the degree of consistency of land cover data extracted from multiple datasets. Each global land cover dataset is considered an expert determination of the actual land cover. Land cover credibility is assessed by measuring the degree of consistency of the data derived from multiple datasets for given spatial or temporal units. A higher degree of consistency of the data corresponds to a higher likelihood of credibility.

The methodological framework of credibility assessment that we have proposed can be used to conduct credibility assessments of cropland cover as well as of other land cover types, such as forest and grassland. Future studies should explore how the criteria or thresholds of credibility can be objectively and rationally established for accuracy assessments and how the scope of credibility assessments can be expanded from regional case studies to global scale assessments.

Acknowledgements *This work was supported by the National Key Research and Development Program of China on Global Change (Grant No. 2017YFA0603304) and the National Natural Science Foundation of China (Grant No. 41807433).*

References

Bartholomé E, Belward A S. 2005. GLC2000: A new approach to global land cover mapping from Earth observation data. Int J Remote Sens, 26: 1959–1977

Bontemps S, Defourny P, Radoux J, Van Bogaert E, Lamarche C, Achard F, Mayaux P, Boettcher M, Brockmann C, Kirches G, Zülkhe M, Kalogirou V, Seifert F M, Arino O. 2013. Consistent global land cover maps for climate modelling communities: Current achievements of the ESA's land cover CCI. Edimburgh: Proceedings of the ESA Living Planet Symposium. 9–13

Bork H R, Bork H, Dalchow C, Faust B, Piorr H P, Schatz T. 1998. Landschaftsentwicklung in Mitteleuropa: Wirkung des Menschen auf Landschaften. Gotham: Klett. 328

Chen J, Chen J, Liao A, Cao X, Chen L, Chen X, He C, Han G, Peng S, Lu M, Zhang W, Tong X, Mills J. 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS J Photogramm Remote Sens, 103: 7–27

Chen J, Chen J. 2018. GlobeLand30: Operational global land cover mapping and big-data analysis. Sci China Earth Sci, 61: 1533–1534

Ellis E C, Kaplan J O, Fuller D Q, Vavrus S, Klein Goldewijk K, Verburg P H. 2013. Used planet: A global history. Proc Natl Acad Sci USA, 110: 7978–7985

Fang X Q, Ye Y, Ge Q S, Zheng J Y. 2005. Land exploitation in the Northeast China during the Qing Dynasty inferred from the development of town system (in Chinese). Sci Geogr Sin, 25: 129–134

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

Friedl M A, McIver D K, Hodges J C F, Zhang X Y, Muchoney D, Strahler A H, Woodcock C E, Gopal S, Schneider A, Cooper A, Baccini A, Gao F, Schaaf C. 2002. Global land cover mapping from MODIS: Algorithms and early results. Remote Sens Environ, 83: 287–302

Fritz S, See L, McCallum I, You L, Bun A, Moltchanova E, Duerauer M, Albrecht F, Schill C, Perger C, Havlik P, Mosnier A, Thornton P, Wood-Sichra U, Herrero M, Becker-Reshef I, Justice C, Hansen M, Gong P, Abdel Aziz S, Cipriani A, Cumani R, Cecchi G, Conchedda G, Ferreira S, Gomez A, Haffani M, Kayitakire F, Malanding J, Mueller R, Newby T, Nonguierma A, Olusegun A, Ortner S, Rajak D R, Rocha J, Schepaschenko D, Schepaschenko M, Terekhov A, Tiangwa A, Vancutsem C, Vintrou E, Wenbin W, van der Velde M, Dunwoody A, Kraxner F, Obersteiner M. 2015. Mapping global cropland and field size. Glob Change Biol, 21: 1980–1992

Gaillard M J, LandCover6k Interim Steering Group members. 2015. LandCover6k: Global anthropogenic land-cover change and its role in past climate. PAGES Program News, 23: 38–39

Gaillard M J, Morrison K, Madella M, Whitehouse N. 2018. Past land-use and land-cover change: The challenge of quantification at the subcontinental to global scales. Past Land Use and Land Cover, 26: 3

Ge Q S, Dai J H, He F N, Pan Y, Wang M M. 2008. Land use changes and their relations with carbon cycles over the past 300 a in China. Sci China Ser D-Earth Sci, 51: 871–884

Goldewijk K K. 2001. Estimating global land use change over the past 300 years: The HYDE database. Glob Biogeochem Cycle, 15: 417–433

Goldewijk K K, Battjes J J. 1997. A Hundred Year (1890–1990) Database for Integrated Environmental Assessments (HYDE, version1.1). Bilthoven: National Institute of Public Health and the Environment

Goldewijk K K, Beusen A, Doelman J, Stehfest E. 2017. Anthropogenic land use estimates for the Holocene–HYDE 3.2. Earth Syst Sci Data, 9: 927–953

Goldewijk K K, Beusen A, Drecht G V, Martine D V. 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12000 years. Glob Ecol Biogeogr, 20: 73–86

Goldewijk K K, Drecht G V, Bouwman A F. 2007. Mapping contemporary global cropland and grassland distributions on a 5 × 5 minute resolution. J Land Use Sci, 2: 167–190

Hansen M C, Defries R S , Townshend J R G, Sohlberg R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int J Remote Sens, 21: 1331–1364

He F N, Li S C, Zhang X Z, Ge Q S, Dai J H. 2012. Comparisons of reconstructed cropland area from multiple datasets for the traditional cultivated region of China in the last 300 years (in Chinese). Sci Geogr Sin, 67: 1190–1200

He F N, Li S C, Zhang X Z, Ge Q S, Dai J H. 2013. Comparisons of cropland area from multiple datasets over the past 300 years in the traditional cultivated region of China. J Geogr Sci, 23: 978–990

He F, Li S, Yang F, Li M. 2018. Evaluating the accuracy of Chinese pasture data in global historical land use datasets. Sci China Earth Sci, 61: 1685–1696

Jia D, Li Y, Fang X. 2018. Complexity of factors influencing the spatiotemporal distribution of archaeological settlements in northeast China over the past millennium. Quat Res, 89: 413–424

Kaplan J O, Krumhardt K M, Ellis E C, Ruddiman W F, Lemmen C, Goldewijk K K. 2010. Holocene carbon emissions as a result of anthropogenic land cover change. Holocene, 21: 775–791

Kaplan J O, Krumhardt K M, Zimmermann N. 2009. The prehistoric and preindustrial deforestation of Europe. Quat Sci Rev, 28: 3016–3034

Latham J, Cumani R, Rosati I, Bloise M. 2014. Global land cover share (GLC-SHARE) database beta-release version 1.0-2014

Leite C C, Costa M H, Soares-Filho B S, de Barros Viana Hissa L. 2012. Historical land use change and associated carbon emissions in Brazil from 1940 to 1995. Glob Biogeochem Cycle, 26: 1–13

Li B B, Fang X Q, Ye Y, Zhang X Z. 2014. Carbon emissions induced by cropland expansion in Northeast China during the past 300 years. Sci China Earth Sci, 57: 2259–2268

Li B B, Fang X Q, Ye Y, Zhang X Z. 2010. Accuracy assessment of global historical cropland datasets based on regional reconstructed historical data—A case study in Northeast China. Sci China Earth Sci, 53: 1689–1699

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

Li Y, Ye Y, Zhang C, Li J, Fang X. 2019. A spatially explicit reconstruction of cropland based on expansion of polders in the Dongting Plain in China during 1750–1985. Reg Environ Change, doi: 10.1007/s10113-019-01575-8

Liu G W. 2012. On the geo-basis of river regulation in the lower reaches of the Yellow River. Sci China Earth Sci, 55: 530–544

Loveland T R, Reed B C, Brown J F, Ohlen D O, Zhu Z, Yang L, Merchant J W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int J Remote Sens, 21: 1303–1330

Niu Z X, eds. 1994. Map of Ancient Channels of the Lower Yellow River in Historical Period. Beijing: Surveying and Mapping Publishing House

Pielke R A, Pitman A, Niyogi D, Mahmood R, McAlpine C, Hossain F, Goldewijk K K, Nair U, Betts R, Fall S, Reichstein M, Kabat P, Noblet N D. 2011. Land use/land cover changes and climate: Modeling analysis and observational evidence. WIREs Clim Change, 2: 828–850

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: 1–16

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

Ruddiman W F. 2003. The anthropogenic greenhouse era began thousands of years ago. Clim Change, 61: 261–293

Shang Z W, Wang H, Li J F, Shi P Q. 2018. Land and se evolution since 11000 years in the coast of Bohai Bay (in Chinese). China Mining Magazine, 27(S2): 286–289

Tateishi R, Bayaer M A, Al-Bilbisi H, Tsendayush J, Shalaby A, Kasimu A, Hoan N T, Kobayashi T, Alsaaideh B, Rahman M M, Tsevengee E, Yamada Y, Kajikawa S. 2001. A new global land cover map, GLCNMO. Forest, 8: 1369–1372

Tuanmu M N, Jetz W. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Glob Ecol Biogeogr, 23: 1031–1045

Wei X Q, Ye Y, Zhang Q, Fang X Q. 2016. Reconstruction of cropland change over the past 300 years in the Jing-Jin-Ji area, China. Reg Environ Change, 16: 2097–2109

Wei X, Ye Y, Zhang Q, Li B, Wei Z. 2019. Reconstruction of cropland change in North China Plain Area over the past 300 years. Glob Planet Change, 176: 60–70

Wu C, He N H. 1991. Plaeochannel changes of the main rivers in the North China Plain over 20000 years. In: Wu C, ed. Collected Works on Plaeochannel in the North China Plain. Beijing: China Science and Technology Press. 132–148

Xue C T. 2009. Historical changes of coastlines on west and south coasts of Bohai Sea since 7000 a B.P. (in Chinese). Sci Geogr Sin, 2: 217–222

Yang F, He F N, Li M J, Li S C. 2019. Reliability assessment of global historical forest data in China (in Chinese). Sci Geog Sin, 74: 923–934

Yang X, Jin X, Xiang X, Fan Y, Liu J, Shan W, Zhou Y. 2019. Carbon emissions induced by farmland expansion in China during the past 300 years. Sci China Earth Sci, 62: 423–437

Ye Y, Fang X Q, Ren Y Y, Zhang X Z, Chen L. 2009. Cropland cover change in Northeast China during the past 300 years. Sci China Ser D-Earth Sci, 52: 1172–1182

Yuan C, Ye Y, Tang C, Fang X. 2017. Accuracy comparison of gridded historical cultivated land data in Jiangsu and Anhui Provinces. Chin Geogr Sci, 27: 273–285

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

Zhang X Z, Wang W C, Fang X Q, Ye Y, Zheng J Y. 2012. Agriculture development-induced surface albedo changes and climatic implications across northeastern China. Chin Geogr Sci, 22: 264–277

Zumkehr A, Campbell J E. 2013. Historical U.S. cropland areas and the potential for bioenergy production on abandoned croplands. Environ Sci Technol, 47: 3840–3847

(Responsible editor: Tianxiang YUE)