Response to referee comments

General reply. Thank you for your insight comments which have improved our work greatly. The object of this study is to develop a longer (compared to the work of Li et al. (2013)) historical gridded cropland dataset over the period of 1690–1999 at the resolution of 1 km. Although in the ear of remote sensing, land cover data could reach a resolution of 30 m (such as GlobeLand30), our dataset has the highest resolution over the historical period of 1690–1970 at least. We have deleted "high-resolution" in this revised version in accordance with your advice. Besides traditional allocation methods, remote sensing based CORINE Land Cover (CLC, https://land.copernicus.eu/pan-european/corine-land-cover) data is used as a reference in the cropland area allocation. We used CLC instead of CCI-LC map because CLC has agricultural land cover in eleven classes whereas CCI-LC map has only five classes. The reasonableness of our dataset is further explained using multiple independent methods. In this revision, modifications were carefully made.

My brief comment on each given comment is as follows:
1. My previous comment of major limitation about data validation and calibration is not explained satisfactory. In the revision, author had compared past statistical or other study dataset to compare their datasets at the grid level but that comparison do not add any value because the dataset developed by author has used baseline of the statistical data to allocate each value into grids. Since the baseline data used for mapping has itself used for validation (in this case areal comparisons at administrative level), which cannot be necessarily qualify validation. Thus, major limitation of this study still stands.

Reply: Thank you for your comments. We agree that the validation in this study is insufficient. However, validation is always the most difficult part because the reconstruction of historical datasets as they have no direct fitted observations. The actual past land cover data (referred to as the "true value") that serves as the credibility assessment baseline is not directly accessible and needs to be reconstructed in most cases. However, 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 (Fang et al., 2020). Therefore, to demonstrate the data production reasonableness, cross-comparison with other independent datasets is the most common approach (Yu et al., 2018; Zhao et al., 2020). The reasonableness of our cropland data before 1980 was investigated using other regional historical works (Anderberg, 1991; Groth et al., 1998; Magnusson, 2000; Lindstad, 2002; Jansson, 2011; FSS, https://ec.europa.eu/) in the previous revision.

In this revision, the reasonableness of our 1999 cropland data was validated using satellite-based land cover datasets (CLC and GlobeLand30 map, http://www.globallandcover.com/) from 2000. More explanations about the validation of our 1999 cropland dataset are listed in the responses to question 2.

We must explain that the comparison with global datasets is not intended to validate our datasets. One purpose of this study is to produce a historical cropland dataset based on cropland area at the **parish/municipality/county** levels, and to use our dataset to assess and improve the global cropland dataset, which is based on cropland area allocation at the **national** level. Reconstructed regional land cover data derived from historical records are regarded as the baseline in most existing studies of credibility assessments of historical global land cover data (Fang et al., 2020). In the first version of our manuscript, we only compared our dataset with the most widely used global dataset HYDE 3.2. Subsequently, following referee #1's advice, we compared our dataset with additional global datasets in the previous revision.

2. This study has mentioned that they have used CCI-LC maps of 2000, but revised paper has maps of 1690 to 1999, thus this study does not use any satellite data and has reorganized and spatially allocated the historical dataset where no validation is available. To provide suggestion on validation, the developed dataset may not needed to validate for all the past years but if author can validate dataset using the satellite data maps from satellite such as **Landsat** where map can be developed at **30m resolution** from 1980 to 1999 or at least from 1985 with some limitations of data and compare these maps with the allocated maps developed by this study: those results may be some kind of comparison and provide the base to validate the results of this study for some years at least.
Reply: Thank you for your comments. As you suggested, we have tried our best to collect all the public available datasets we can access.

In the first version of our manuscript, we used CCI-LC map in 2000 as reference for cropland area allocation. According to your comments, we found that CORINE Land Cover (CLC) had more detailed information about the cropland area in Scandinavia. Thus, CLC data were applied to our cropland area allocation model, because CLC had agricultural land cover in **eleven** classes whereas CCI-LC map had only **five** classes (Table 1).

Table 1 Classes of agricultural areas from CLC-LC maps and CLC maps

CCI-LC	Cropland, rainfed (Herbaceous cover)	Cropland, rainfed (Tree or shrub cover)	Cropland, irrigated or post-flooding	Mosaic cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%)	Mosaic natural vegetation (tree, shrub, herbaceous cover) (>50%) / cropland (<50%)	
CLC	Non-irrigated arable land	Permanently irrigated land	Rice fields	Vineyards	Fruit trees and berry plantations	Olive groves
	Pastures	Annual crops associated with permanent crops	Complex cultivation patterns	Land principally occupied by agriculture, with significant areas of natural vegetation	Agro-forestry areas	

The CLC inventory was initiated in 1985 and updates have been produced in 2000, 2006, 2012, and 2018. However, CLC maps are only available after 2000 for Scandinavia. Thus, we used 2000 CLC map in this study. CLC2000 is produced by many countries in Europe by visual interpretation of high-resolution satellite imagery from **Landsat-7**

ETM (https://land.copernicus.eu/pan-european/corine-land-cover).

The CLC map of 2000 plays an essential role in our allocation methods; however, this was not explained clearly in the previous version. In this revision, modifications were made in the data and methods sections. **Please check the noted Sections (2.3 Satellite-based data, 3.2 Cropland area allocation into grid cells).**

According to your advice, to validate our cropland dataset from 1980 to 1999, we find that there are two global land cover maps at 30m resolution, Global Food Security Analysis-Support Data at 30 Meters (GFSAD30, https://croplands.org/app/map?lat=0&lng=0&zoom=2) Project and GlobeLand30 maps. However, GFSAD30 provides a land cover map at 30-m resolution only in 2015. GlobeLand30 has maps at 30-m resolution in 2000, 2010, and 2020. **Although there are differences between satellite-based maps of multiple time points in modern times, compared to the history over the past 300 years, 1985–2000 can be regarded as one time point. Thus, we chose the GlobeLand30 map from 2000 to validate our 1999 cropland dataset.**

The images for land cover development classification and update of **GlobeLand30** were mainly 30-m multispectral images, including **TM5 ETM+**, and **OLI multispectral images from Landsat** (USA) and **HJ-1** (China Environment and Disaster Reduction Satellite). GlobeLand30 includes ten land-cover classes in total; namely **cultivated land**, forest, grassland, shrubland, wetland, water bodies, tundra, artificial surface, bare land, and perennial snow and ice. **Cultivated land** refers to the land used for cultivating crops. Paddy fields, irrigated upland, rainfed upland, vegetable land, cultivated pasture, greenhouse land, land mainly planted with crops and rarely with fruit trees or other trees, tea plantations, coffee plantations, and other economic croplands are included in this category (http://www.globallandcover.com/).

However, cropland in our datasets only includes arable land (areas under temporary crops, temporary meadows and pastures, land temporarily fallow) and areas under permanent crops. We compared the **cropland area** at the parish and county levels in 1999 in this study with **cultivated land area** from the 2000 GlobeLand30 map to validate our statistics. Then, we aggregated the 30-m resolution GlobeLand30 to 1 km and compared the result with our 1999 gridded cropland dataset to validate our allocation method.

Comparing the cropland area at the parish and county levels in 1999 in this study with the cultivated land from the 2000 GlobeLand30 map shows that the **cultivated land area** from GlobeLand30 is 1.4 times that of our **cropland area** (Figure 1).

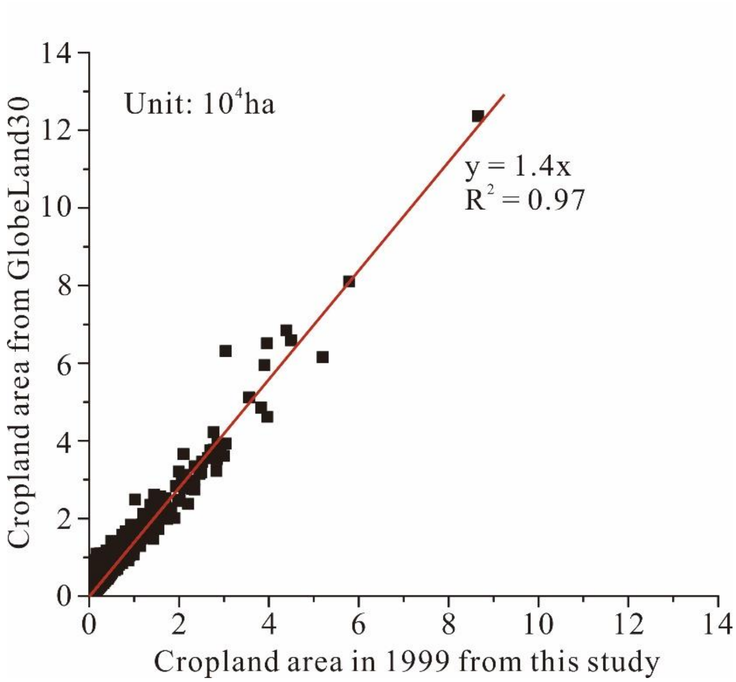
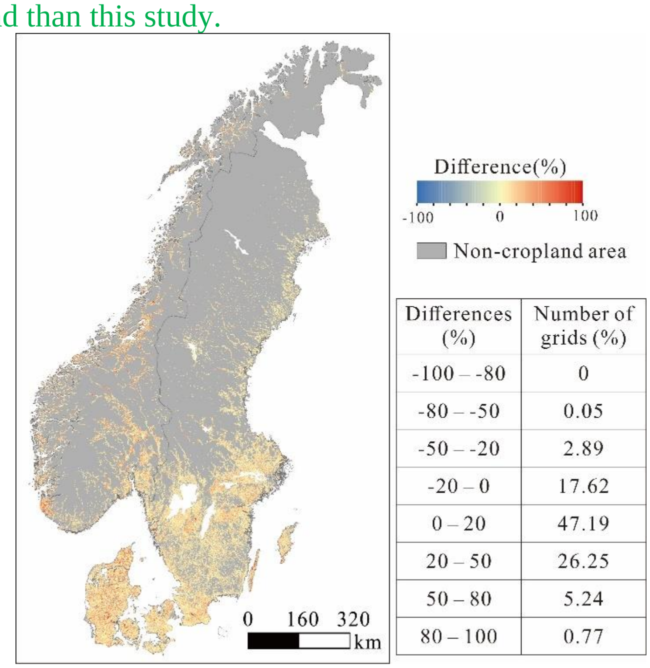


Figure 1 Comparison of the area at the parish and county levels between GlobeLand30 and this study

Comparison of the gridded cropland area at 1-km resolution between GlobeLand30 and this study shows grids with differences between -20% and 20% account for 64.81% of the total number of grids with the cropland area > 0 (Figure 2). Because cultivated pasture, greenhouse land, gardens and so on are included in cultivated land from GlobeLand30, whereas only arable land and permanent crops comprise cropland in this study, among all grids with cultivated land > 0, 79.44% of grids from GlobeLand30 have more cultivated land than this study.



Differences (%)	Number of grids (%)
-100 – -80	0
-80 – -50	0.05
-50 – -20	2.89
-20 – 0	17.62
0 – 20	47.19
20 – 50	26.25
50 – 80	5.24
80 – 100	0.77

Figure 2 Comparison of the gridded cropland area between GlobeLand30 and this study

Because CLC2000 provides more agricultural (cultivated) land classes, we used CLC2000 to further validate our cropland dataset. In Scandinavia, CLC2000 divides agricultural areas into four categories: **“Arable land”** (including one class, “Non-irrigated arable land”), **“Permanent crops”** (including one class, “Fruit trees and berry plantations”), **“Heterogeneous agricultural areas”** (including two classes, “Complex cultivation patterns” and “Land principally occupied by agriculture, with significant areas of natural vegetation”) and **“Pastures”** . Comparison of the gridded cultivated land area and agricultural areas at 1-km resolution between GlobeLand30 and CLC2000 shows the consistency of these two datasets (Figure 3).

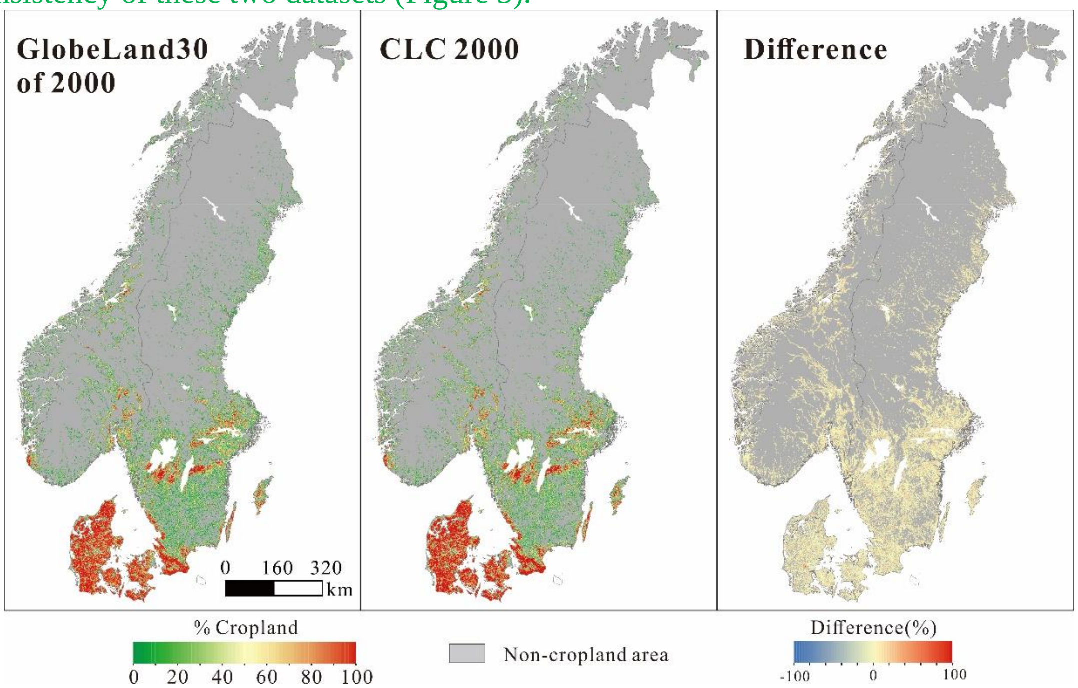


Figure 3 Comparison of agricultural land between GlobeLand30 and CLC2000

We compared the total area of “**Arable land,**” “**Permanent crops,**” and “**Complex cultivation patterns**” from CLC2000 with our cropland area for each parish and county (Figure 4). Figure 4 shows that the area from CLC2000 is 1.17 times that of our cropland area. Removing some marginal roads and natural lands included in “**Arable land**,” “**Permanent crops,**” and “**Complex cultivation patterns**” from CLC2000, our cropland areas at the parish and county levels are close to those from CLC2000.

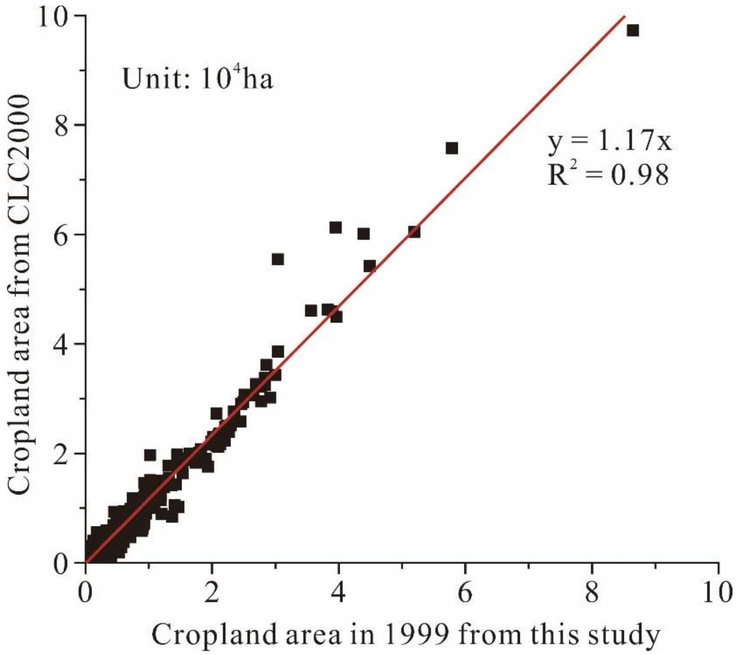
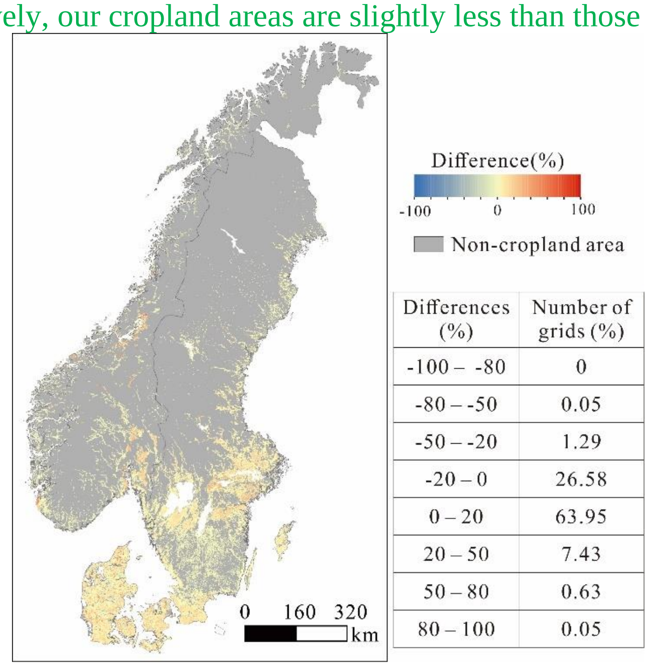


Figure 4 Comparison of the cropland area at the parish and county levels between CLC2000 and this study

Comparison of the gridded cropland area at 1-km resolution between CLC2000 and this study shows grids with differences between -20% and 20% account for 90.54% of the total number of grids with the cropland area > 0 (Figure 5). Our gridded cropland area is close to that in CLC2000 overall, which indicates the reliability of our cropland dataset. However, around Stockholm and Trondheim, in southeastern Norway and northern Denmark, respectively, our cropland areas are slightly less than those from CLC2000.



Differences (%)	Number of grids (%)
-100 – -80	0
-80 – -50	0.05
-50 – -20	1.29
-20 – 0	26.58
0 – 20	63.95
20 – 50	7.43
50 – 80	0.63
80 – 100	0.05

Figure 5 Comparison of the gridded cropland area between CLC2000 and this study

Please check the Section 5.1 Validation of the dataset developed in this study.

3. As observed in this study, cropland data is collected from several studies and different governments : thus **do not hold a single cropland definition** and need further explanation. This study did not provide clear explanation on it. How did author combine all this datasets when cropland definitions of different dataset were different. What was the basis, how did it affect the fusion?

Reply: Thank you for your comments. We regret not providing a clear explanation of our cropland definition. Table 2 is used here to give a clear explanation of the cropland definitions for all datasets. The category "Cropland" defined by the FAO (http://www.fao.org/) was used in this study. Thus, cropland in this study includes areas under temporary crops (A), temporary meadows and pastures (B), land temporarily fallow (C), and areas under permanent crops (D).

Statistics of all countries in Scandinavia recorded land use areas of all crops, temporary meadows and pastures and fallow land. Therefore, we selected the land areas classified as cropland defined by the FAO and calculated their total area. In Norway, statistics only provide the area of A. Based on the census from the Farm Structure Survey (FSS, https://ec.europa.eu/) in 2000 and 2010, the total size of fallow land and land under permanent crops accounted for approximately 0.9% of all cropland area. Thus, we used the total areas under temporary crops (A) as cropland area before 1810 in Norway.

For datasets from previous studies, authors gave clear cropland definitions. **In Sweden**, "*åker*" in studies from SND (https://snd.gu.se/en/catalogue/study/SND0910) and Li et al. (2013) included land under temporary crops (A), land under temporary meadows and pastures (B), and temporarily fallow land (C). Because the census from FSS shows D accounts for only 0.1% of the total cropland area, we used the total area of A, B, and C as the cropland area. **In Norway**, Li et al. (2013) used the total area of A, B, C, D, and E (Permanent grassland and meadow) as the cropland area. We identified their data sources (NSD kommunedatabase, https://kdb.nsd.no/kdbbin/kdb_start.exe) and re-collected A, B, C, and D as cropland. **In Denmark**, both "ager" in the dataset of Dam and Jakobsen (2008) and "agerjord" in the dataset of Odgaard and Rømer (2009) indicate the total of A, B, and C. Because the census from FSS shows land area under permanent crops (D) accounted for approximately 0.4%–1% in Denmark, we used the total of A, B, and C as cropland.

We have explained the definitions of cropland from different sources more clearly. Please check the indicated sections (**2.1 Cropland data and 3.1.1 Cropland data collection and calibration**).

Tabel 2 Cropland definitions of the data sources

Data sources	Spatial coverage	Years	Reference	Cropland definitions (Categories included in recorded cropland)	Combination

Sockenvis jordbruksstatistik	Sweden	1690, 1750, 1810	SND	A, B, C (Åker)	Census from FSS shows the area of D accounted for about 0.1% of the total cropland area in Sweden. We use the total of A, B and C as cropland.
Statistiske studier over folkemængde og jordbrug i Norges	Norway	1665, 1723	Aschehoug, 1890	A	Census from FSS shows the total size of B, C, D accounted for about 0.9% of the total cropland area in Norway. We use A as cropland.
Historisk Tidsskrift	Norway	1809	Hovland, 1978	A	
Atlas over Denmark: Historisk-Geografisk Atlas	Denmark	1688	Dam and Jakobsen, 2008	A, B, C (Ager)	Census from FSS shows the area of D accounted for about 0.4%~1% of the total cropland area in Denmark. We use the total of A, B and C as cropland.
Danske landbrugs-landskaber gennem 2000 år	Denmark	1800, 1881, 1998	Odgaard and Rømer, 2009	A, B, C (Agerjord)	
Statistisk Aarbog 1912	Denmark	1907	Danmarks Statistik, 1912	A, B, C, D	We use the total of A, B, C and D as cropland
Statistiske Meddelelser 1936, 1950 and 1980	Denmark	1936, 1950, 1980	Danmarks Statistik, 1936, 1950 and 1980	A, B, C, D	
Cropland in Scandinavian Peninsula	Sweden, Norway	1875, 1910, 1930, 1950, 1980, 1999	Li et al., 2013	Norway: A, B, C, D, E	
				Sweden: A, B, C (Åker)	Census from FSS shows the area of D accounted for about 0.1% of the total cropland area in Sweden. We use the total of A, B and C as cropland.

Notes: A—Areas under temporary crops; B—Areas under temporary meadows and pastures; C—Land with temporary fallow; D—Areas under permanent crops; E—Permanent grassland and meadow

4. As I provided above solution about using those satellite dataset available in historical

years like Landsat is available from 1985 , which can be used to check spatially accuracy and allocation for precision of maps. I suggest authors to implement this method rather than not validating the results and providing blind spatial allocations with no base.

Reply: This comment is explained in the responses to question 2.

5. This work still needed high level of English correction and organization of writing. For example, in revised version, discussion section has lot of methodological details and results and very less discussion. This paper has lot of scope to work on organizing the sections and restructuring the paper while providing English corrections.

Reply: Thank you for your comments. We have moved the methodological details to **Method** Section and added datasets validation details in **Discussion** Section. We also restructured the paper. Our revised manuscript has been edited by Elsevier Language Editing Services.

6. This data is not high resolution maps: Author may call it spatial maps as previous dataset just have county level details but using high-resolution is not suitable. In remote sensing terms less than 10m pixel can be considered as high resolution according to definitions provided by several international research organizations such as USDA, UN,FAO.

Reply: Thank you for your comments. We have deleted “high-resolution” in the latest revised version in accordance with your comments.

7. Data reliability is still questionable as the validation and gap filling is not explained or analyzed properly. Although interpolation is the only way to gap fill data but the interpolated data need to be validate for further use.

Reply: Thank you for your comments. We have validated our cropland dataset using GlobeLand30 and CLC2000, please check the responses to question 2.

Missing data existed in Denmark in 1688, 1750, 1800, and 1881. In 1688, 1800 and 1881, cropland area data in 184 (2.3%), 56 (3.3%) and 2 (0.1%) “*ejerlavs*” were missing, respectively. We assumed that neighboring “*ejerlavs*” with similar terrain had the same cropland fraction and cropland growth rate. Then, the missing data were interpolated based on the cropland fractions of their neighboring "*ejerlavs*" in 1688 and the cropland area changes during 1688–1800 and 1800–1881. There was no cropland area record of Denmark in 1750. Therefore, we assumed that the cropland area change rate from 1690 to 1810 was constant and computed each parish’s cropland area in 1750 (linear interpolation). These interpolation methods were also used by Ramankutty and Foley (1999), Ye et al. (2015), Wei et al. (2016); He et al. (2017), Li et al. (2018), and Yu and Lu (2018); however, their interpolated data did not reduce the credibility of their datasets. As satellite-based data and survey data at the parish level were unavailable from 1688 to 1881, the interpolated data were impossible to validate using direct fitted observations.

Based on the study of Fang et al. (2020), three methods could be used to assess the credibility of historical land cover datasets, including accuracy assessment (quantitative assessment based on quantitatively reconstructed regional land cover data), rationality assessment (qualitative assessment, including the regional historical facts-based rationality assessment and the expertise-based rationality assessment) and likelihood assessment (the credibility of the land cover data for given spatial or temporal units is inferred according to the degree of consistency in land cover data extracted from multiple datasets). Because apart from the data sources we used, other quantitatively reconstructed regional land cover data in Denmark from 1688 to 1881 were unavailable, we employed a regional historical facts-based rationality assessment to analyze the reliability of our interpolated data in Denmark. The following Danish history suggests that linearly interpolated data are reasonable. The national tax system stipulated that each household in Denmark had 50 acres of cropland, no more and no less. The population of Denmark grew steadily from 1690 to 1881. During this period, wars did not cause sudden changes in cropland area in Denmark. The agricultural reform that began in 1789 changed the relationship between landlords and tenant farmers but did not cause a sudden change in cropland area (Jespersen, 2018).

We have added this explanation in Section **5.1 Validation of the dataset developed in this study**, please check the page 26, lines 10–22 and page 27, lines 1–10.

References:

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

He, F., Li, M., Li, S.: Reconstruction of Lu-level cropland areas in the Northern Song Dynasty (AD976-1078), Journal of Geographical Sciences, 27(5): 606-618, 2017.

Jespersen, K. J. V.: A history of Denmark (Macmillan essential histories), Red Globe Press, 2018.

Li, B., Jansson, U., Ye, Y., and Widgren, M.: The spatial and temporal change of cropland in the Scandinavian Peninsula during 1875–1999, Regional Environmental Change, 13, 1325-1336, https://doi.org/10.1007/s10113-013-0457-z, 2013.

Li, M., He, F., Li, S., Yang, F.: Reconstruction of the cropland cover changes in eastern China between the 10th century and 13th century using historical documents, Scientific Reports, 8: 13552, 2018.

Ramankutty, N., and Foley, J. A.: Estimating historical changes in global land cover: Cropland from 1700 to 1992, Global Biogeochemical Cycles, 13, 997-1027, 1999.

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

Ye, Y., Wei, X., Li, F., Fang, X.: Reconstruction of cropland cover changes in the Shandong Province over the past 300 years, Scientific Reports, 5: 13642, 2015.

Yu, Z., Lu, C.: Historical cropland expansion and abandonment in the continental U.S. during 1850 to 2016, Global Ecology and Biogeography, 27: 322-333, 2018.

Zhao, Z., Fang, X., Ye, Y., Zhang, C., and Zhang, D.: Reconstruction of cropland area in the European part of Tsarist Russia from 1696 to 1914 based on historical documents, Journal of Geographical Sciences, 30(8), 1307-1324, https://doi.org/10.1007/s11442-020-1783-y, 2020.