Reviewer#3

Summary

Mapping the spatial distribution and temporal dynamics of crop area and crop types is essential for crop management. However, long-term and relatively high-spatial resolution crop area data with detailed crop type information is lacking. Here Ye et al., combined multi-sources of crop area and type information and generated the 1-km maps of crop area and types in the conterminous US during 1850-2021. The manuscript is generally well organized, and easy to follow. I only have few specific concerns listed below.

Response: We appreciate the reviewer's positive feedback and have revised our manuscript. Please see our response below.

Specific comments

1) Line 8, 'climate, air', there could have some overlaps between these two terms.

Response: We appreciate this suggestion. We have changed "climate, air" to climate in lines 7-8.

Lines 7-8: Agricultural activities have been recognized as an important driver of land cover/land use change (LCLUC) and have significantly impacted the ecosystem feedback to climate by altering land surface properties.

2) Line 11, each time when I mention 'high resolution', I am always very cautious. In remote sensing, high resolution could be meter to submeter. So 'relatively high resolution' could be more rigorous.

Response: We thank for this suggestion. We have revised "high resolution" to "relatively high resolution" line 11.

Lines 11: there remains a dearth of a relatively high-resolution dataset with crop type details over a long period

3) Line 85, any justification about the linear assumption?

Response: We appreciate this valuable suggestion. After conducting a comprehensive data search, we discovered that the crop rotation information was recorded by the "Tailored Reports: Crop Production Practices" (USDA- Agricultural Resource Management Survey). We further extracted the state-level crop rotation information from 1996 to 2010, and found the rotation rate in each state kept relatively stable during ARMS-available years. Consequently, we adopted a new assumption that the rotation percentage between corn and soybean remained stable in each state when the crop rotation information was unavailable from 1940 to 2009. Furthermore, we updated the crop type maps under the new corn-soybean rotation assumption. More details about the rotation process are in lines 96-97, 230-239 and Table S3.

Lines 96-97: The rotation percentage between corn and soybean remained constant when the rotation information was unavailable from 1940 to 2009.

Lines 230-239: Considering the dominant crop rotation type in US, soybean and corn rotation, we simulated the corn-soybean rotation by randomly switching a certain area between corn and soybean according to the rotation rate. The crop rotation information from 1996 to 2010 at state level was documented by the "Tailored Reports: Crop Production Practices" of USDA's Agricultural Resource Management Survey (ARMS) (https://data.ers.usda.gov/reports.aspx?ID=17883). The rotation rate was calculated as the ratio of the sum of corn-soybean and soybean-corn acreage to the total area of corn and soybean. We found that the rotation rate in each state kept relatively stable in the ARMS-available years, and assumed that the rotation rate in the missing years is the same as the mean rate from available years (Table S3), which is further applied to corresponding counties. Because soybean was rarely planted in the Corn Belt before 1940 (Yu et al., 2018), we only considered the corn-soybean rotation during the period 1940-2009 in 17 states (Table S3) (Padgitt et al., 1990).

Table S1. The mean corn-soybean ration ratio calculated from the available years.

State	**Rotation Rate (%)**	**Standard Deviation**	**State**	**Rotation Rate (%)**	**Standard Deviation**
IL	85	5	MO	58	8
IN	81	10	NB	59	12
IA	86	5	NC	30	10
KS	35	9	OH	64	12
KY	58	15	PA	27	10
LA	16	10	SD	69	7
MI	49	11	TN	35	15
MN	74	6	WI	43	13
MS	14	9			

4) Line 87, references are needed following 'diversity index'

Response: Thanks for this suggestion. We have added a reference.

L Jost: Entropy and diversity, Opinion, 2, 363–375, 2006.

5) Line 109-110: please also briefly mention how you conducted the filtering operation in the main text.

Response: We are grateful for this advice. We added a brief description in lines 125-129.

Lines 125-129: LCMAP adopts Anderson Level I-based legend, grouping cropland and pasture into one category (Xian et al., 2022). In contrast, NLCD uses a Level II-based legend where cropland and pasture are separately tracked (Xian et al., 2022) (Table S4). To generate a reliable cropland distribution, the long-term non-crop trajectory derived from NLCD was used to exclude all grids identified as cropland the LCMAP map (more details are presented in Supplementary Methods: (1) Preprocesses for LCMAP).

6) Table 1, how did you resample 30 m data to 1km, how did you interpolate the 5arc-min data to 1km?

Response: Thanks for this suggestion. We added the description of resampling map in lines 136-141

Lines 136-141: All gridded datasets were resampled to 1km. We employed a 1km*1km window to aggregate the total cropland area from the 30m*30m map and assigned the area to the corresponding 1km*1km grid. To resample the CDL crop type map from 30m to 1km, the crop type in each 1km by 1km pixel was assigned to the dominant crop type with the largest fraction of land area within the 1km*1km window. Conversely, the cropland percentage in each 5 arc-min grid is interpolated to 1km*1km grid cells with an assumption that cropland percentage is evenly distributed within the 5 arc-min by 5 arc-min window.

7) Line 110: why not use the data from Zumkehr and Cambell (2013). I assume differences exist between these two long-term datasets.

Response: In reconstructing historical cropland maps, HDYE adopted a weighing map involving social and multiple environmental factors (soil suitability, temperature, and topography) to spatialize cropland area (Goldewijk, 2001; Goldewijk et al., 2011; Goldewijk et al., 2017), while Zumkehr and Cambell (2013) adopted a land-use model of Romankutty and Foley (1999) to calculate historical cropland grid area based on a satellite-derived map. Compared with ZCMAP, HYDE considers the impacts of various environmental factors on crop distribution. Consequently, we adopted HYDE products as potential cropland distributions to retrieve historical cropland maps. Please see lines 129-134.

Lines 129-134: For the period of 1850-1984, although both ZCMAP and HYDE provide the cropland distribution, HYDE considers the impacts of various environmental factors (soil suitability, temperature, and topography) on crop distribution compared with ZCMAP (Goldewijk, 2001; Goldewijk et al., 2011; Goldewijk et al., 2017; Zumkehr and Campbell, 2013). Consequently, HYDE (available every 10 years) was initially used to identify the cropland distribution by calculating the fraction of cropland to the physical area for each grid.

8) Line 119-121: please briefly explain what do the differences between NASS-CPAS and USDA-COA mean in the main text.

Response: We appreciate this suggestion. We have explained the difference between NASS-CPAS and USDA-COA in lines 148-153.

Lines 148-153: NASS-CPAS reports the annual total planting area of major crops for each state from 1909 to 2021. However, some minor crop types, such as vegetables and fruits, are excluded. USDA-COA provides the total areas of crop harvest, failure, and fallow for each state from 1925 to 2017 with 4~5-year intervals. We computed the difference between these two datasets for available years and linearly interpolated unavailable years during 1909-2021. The difference was assumed to be the planting area of those minor crops.

9) Line 122: ‘interpolate’->’extrapolate’

Response: We appreciate the reviewer’s suggestion. We have revised it in lines 154-156.

Lines 154-156: We used the interannual variations of arable land of each state extracted from HYDE to extrapolate the total planting area from 1908 to 1850 (Equation 1).

10) Fig. 2, Please explain what does each point mean in the caption.

Response: We thank for this suggestion. We moved Figure 2 to Figure S2, and added the explanation in the title of Figure S2.

The title of Figure S2: Comparison of crop-specific cropland area between reconstructed maps and raw inventory data at county level in 1920, 1960, 2000, and 2020 (Kha is a thousand hectares). The point in subfigures represents the paired cropland area from the reconstructed map and raw inventory data for a certain county and year. The color bar in each subfigure indicates the probability density of paired point calculated by the gaussian kernel

11) Fig. 2, In addition to this kind of scatter plot, how about plotting the time series of crop-specific cropland area between reconstructed maps and raw inventory data to show that the reconstructed maps captured the interannual variations/trend of the raw data.

Response: We appreciate this valuable suggestion. To address this comment, we added two additional figures (Figure 2 and S3) to show the historical cropland residual and relative residual between maps and the inventory data at county level. The results show that the residual (the inventory-based crop area minus the rebuilt-map-based crop area) in planting area between the inventory data and maps is less than 0.2 Kha, and the relative residual (the ratio of residual to the inventory crop area) to the plant area is less than 10% for most counties (>75%) throughout the entire study period for nine major crop types, which indicate that the developed maps have strong capacity to capture the historical cropland area change. More details are presented in lines 243-248.

Lines 243-248: We further examined the time-series residual between the inventory data and maps (Figure 2 and S3). It is evident that the residuals (the inventory-based crop area minus the rebuilt-map-based crop area (Equation 7)) are generally smaller than 0.2 Kha for the majority counties (>75%) across all years for nine crop types. Relatively greater residuals are observed in spring wheat, durum wheat, and rice before 1875 (Figure 2d, g, and i), which might be attributed to the marginal area of these three crops during the early years. Similarly, the relative errors (the ratio of residual to the inventory crop area (Equation 8)) in most counties remain within ±2% for different crops, except for spring wheat, durum wheat, and rice before 1875 (Figure S3d, g, and i).

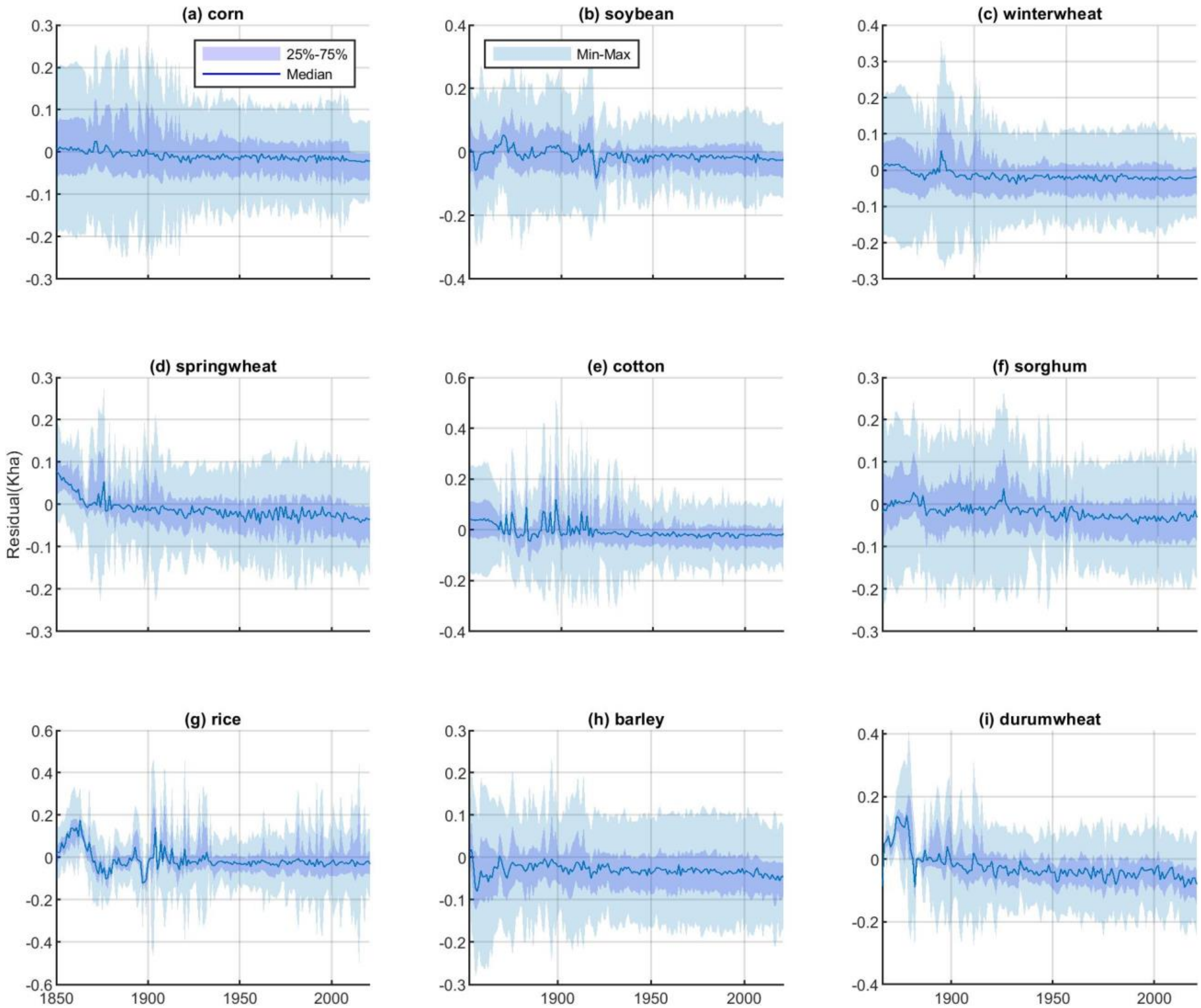

Figure 1. The distribution of residual (the inventory-based crop area minus the rebuilt-map-based crop area, defined by Equation 6) between the rebuilt inventory and maps from 1850 to 2021 (Kha is a thousand hectares). In each year, "Min-Max", "Median", and "25%-75%" reflects the extent of residual from all counties at levels of "minimum value to maximum value", "50th percentile", and "25th percentile to 75th percentile", respectively, which are corresponding to five percentiles in a box plot.

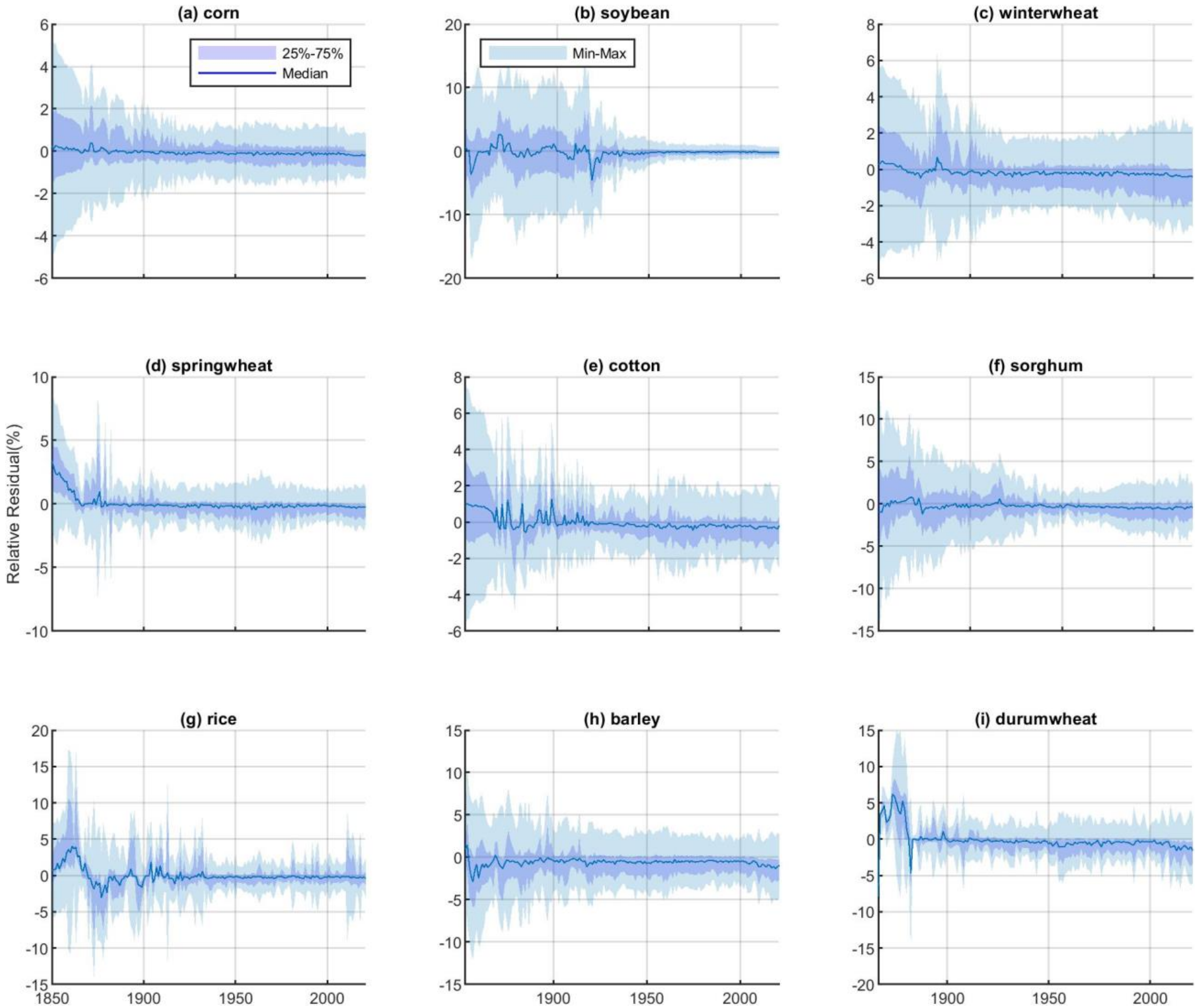

Figure S1. The distribution of relative residual (the ratio of the residual to the inventory crop area, defined by Equation 7) between the rebuilt inventory and rebuilt maps from 1850 to 2021. In each year, “Min-Max”, “Median”, and “25%-75%” reflects the extent of residual at levels of “minimum value to maximum value”, “50th percentile”, and “25th percentile to 75th percentile”, respectively, which are corresponding to five percentiles in a box plot. The counties with cropland areas less than 1kha are excluded to avoid the case with a relative residual greater than 100%.

References:

Goldewijk, K. K.: Estimating global land use change over the past 300 years: The HYDE database, Global Biogeochem. Cycles, 15, 417–433, https://doi.org/10.1029/1999GB001232, 2001.

Goldewijk, K. K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the Holocene - HYDE 3.2, Earth Syst. Sci. Data, 9, 927–953, https://doi.org/10.5194/essd-9-927-2017, 2017.

Klein Goldewijk, K., Beusen, A., van Drecht, G., and de Vos, M.: The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years, Glob. Ecol. Biogeogr., 20, 73–86, https://doi.org/https://doi.org/10.1111/j.1466-8238.2010.00587.x, 2011.

Ramankutty, N. and Foley, J. A.: Estimating historical changes in land cover North American croplands from 1850 to 1992, Glob. Ecol. Biogeogr., 8, 381–396, https://doi.org/10.1046/j.1365-2699.1999.00141.x, 1999.

Yu, Z., Lu, C., Cao, P., and Tian, H.: Long-term terrestrial carbon dynamics in the Midwestern United States during 1850-2015: Roles of land use and cover change and agricultural management, Glob. Chang. Biol., 24, 2673–2690, https://doi.org/10.1111/gcb.14074, 2018.

Zumkehr, A. and Campbell, J. E.: Historical U.S. cropland areas and the potential for bioenergy production on abandoned croplands, Environ. Sci. Technol., 47, 3840–3847, https://doi.org/10.1021/es3033132, 2013.