

Earth Syst. Sci. Data Discuss., referee comment RC4  
<https://doi.org/10.5194/essd-2020-382-RC4>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on **essd-2020-382**

Anonymous Referee #4

---

Referee comment on "A historical reconstruction of cropland in China from 1900 to 2016" by Zhen Yu et al., Earth Syst. Sci. Data Discuss.,  
<https://doi.org/10.5194/essd-2020-382-RC4>, 2021

---

Reconstructing of long-term statistical areas as well as the spatial distribution of historical cropland are essential to track the dynamics of the agriculture development and to analysis the driving factors from natural and anthropological aspects. The manuscript collected and integrated different data source to reconstruct time series cropland layers since 1900. The topic is of interests to researchers and has potential to be published on ESSD. However, some major issues must be addressed before considered in ESSD.

1. When working on the spatializing of cropland distribution, the first type of gridded images was used assuming higher weighting factors with a higher value given to a higher resolution image. However, did the authors test the consistency or inconsistency of different existing land cover dataset? At many cases, the consistency at many regions is low which introduced high uncertainty at the cropland allocation steps.

2. When spatializing the cropland data from statistical area to grid level, why first generated 100m binary cropland & noncropland map and then resampled to 5km grided percentage map? Is it more straightforward to generate the possibility of cropland (0-100) at 5km grid level?

3. After 1980s when Landsat imageries became available, why not to use multi-annual actual satellite data as a validation source to prove the effectiveness of the spatializing approach?

Some minor comments:

1. CNLUCC products are illustrated in figure 1 at different time period but in the description

of the figure 1, only CAS1990 is described. Are those two different sources? In Figure 2, both CNLUCC and CAS1990 are not included as a comparison source. Is the ignorance of those dataset by purpose or by mistake?

2.The authors compared the provincial cropland areas and coverage percentages using data derived from HYDE, Yang et al. (2015), and our study with officially released NLRB data (Fig. 3). Do the scatter points include only one year or all years of NLRB data?