



Comment on **essd-2021-219**

Anonymous Referee #1

Referee comment on "A 1 km global cropland dataset from 10000 BCE to 2100 CE" by Bowen Cao et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-219-RC1>, 2021

The contribution of the research is having produced the first 1km global cropland dataset with long time span by employing the newly developed spatially explicit allocation algorithms. The theoretical framework of the method is convincing, and the results are well presented. This spatiotemporally continuous dataset will provide a new opportunity for better understanding the past global changes caused by ALCC.

Although the new global cropland dataset has been created to such a high resolution (especially for historical periods), it is necessary to recognize the uncertainty of this dataset that was unmentioned in this manuscript. It is suggested to make a brief discussion further to address the issue that both the cropland area per capita estimated by HYDE and the historical population datasets adopted by this study are with huge uncertainty.

Firstly, the reviewer acknowledged that there is no more reliable cropland area data at the global scale than HYDE up to now, but it should be noticed that the historical cropland area data in HYDE have great uncertainty that has already been proved in many countries. One of the main reasons is due to the unreliable estimation of per capita cropland area. The continuous improvement of cropland area data in versions of HYDE mainly relies on the modification of the historical cropland area per capita curve according to the quantitative regional reconstruction results or other related indexes. Although the authors gave a clear definition of cropland (line 118), which was also the same as the definition adopted by HYDE during its historical cropland area estimation. The definition without considering the unignorable amount of fallow land or crop rotation in history would cause obvious underestimation about the cropland areas in some countries. Especially in countries that are far less intensively cultivated (completely different from the traditional agricultural period in China), like some countries in Europe. Some studies have also pointed out that the area of cultivated land in Europe in HYDE is obviously underestimated. Thus, this would cause a smaller extent of historical cropland distribution and lower fractions in the gridded allocation datasets.

Second, the historical population dataset of HYDE is actually derived from the national or subnational statistical/estimated population amount by downscaling method, which is basically the same as the allocation algorithm of historical cropland. The huge uncertainty also existed in this dataset caused by its allocation principles. Since both studies have

adopted the population factor in the allocation algorithm of historical cropland, the difference of gridded datasets between HYDE and this study is namely caused by the different usage of physiogeographic factors and their resolutions during the allocation. In future related researches, please cautiously use this unevaluated historical population dataset.

Additionally, the cropland results are displayed at a global scale, so the details cannot be seen clearly, and no administrative boundary was added on the regional maps (line 635-640, Fig4-6, it would be better to add some necessary labels and administrative boundaries on the map; it seems that the linear unit scale should not be added under the geographical coordinate system?).