

Comments

Title: ChinaSoyArea10m: a dataset of soybean planting areas with a spatial resolution of 10 m across China from 2017 to 2021

Author(s): Qinghang Mei et al.

MS No.: essd-2023-467

MS type: Data description paper

Mei et al's work mapped the soybean planting areas across China with a high spatial resolution of 10 meters, spanning from 2017 to 2021, provided important information for sustainable soybean production and management, as well as agricultural system modeling and optimization. In this work, authors summarized five methods of mapping crops by remote sensing. The advantages and uncertainties of each method were compared, and a highly effective for accurately mapping crops over a larger region method named combining unsupervised classification and post-classification methods applied in this paper. They accomplished this by Sentinel-2 remote sensing images from the GEE platform with cropland layer and detailed phenology observations. They validated the results with the census data at both county- and prefecture-level, and with the two existing datasets (CDL and GLAD maize-soybean map).

Overall, I find this work to be valuable. However, I have some concerns regarding the robustness from the sparse number of AMSs in SW Zonal IV and uncertainty in quality of satellite imagery. I hope the authors will consider these points and provide further clarification in their responses and/or revisions. Please find my major comments and minor for clarification below.

Major comments:

1. The text mentions the need for 10-day time series composite images per month, but in certain areas, the average monthly count of clear observations is insufficient to meet this requirement. Can the existing time series composite methods be optimized to accommodate the inadequacy of observational data?
2. The observations per month of satellite imagery in SW Zonal IV are less, and the AMSs in this zonal also only have two sites. Whether it is possible to increase the observational data or phenological data from remote sensing to test the robust.
3. To determine the potential cropping areas, authors filtered the pixels exhibiting an EVI maximum value during the growing season greater than 0.4 to remove fallow land. For spatial variation across four zonal, the constant threshold would bring some uncertainty. I expect to see more evidence for selecting 0.4 or a sensitivity analysis of threshold can also be implemented.

Minor comments:

Line 58: "same areas" means the north China?

Line 180, Figure2: The label on the left in Figure2 (i.e. 'Data processing' and 'Accuracy assessment') are set to rotate 180° to match reading habits.

Line 180, Figure2: In step2, part (2) of the dashed box is confusing. What the color

represents? If I understand correctly, they represent different layers of indexes. It is recommended to put the abbreviation to the right of the color layers.