

Comment on **essd-2022-297**

Anonymous Referee #1

Referee comment on "GlobalWheatYield4km: a global wheat yield dataset at 4-km resolution during 1982–2020 based on deep learning approaches" by Yuchuan Luo et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-297-RC1>, 2022

This paper produced a global wheat yield dataset named GlobalWheatYield4km using deep learning method (LSTM). More efforts are needed to improve the manuscripts.

- Add a table in Section 2.2 to list all the datasets used as inputs and outputs, including the name, spatial and temporal resolutions, time period covered, the purposes of those datasets used in the paper.
- Add a figure in Section 2.3 to describe the detailed workflow of how the dataset was produced from inputs to outputs.
- Why the time period of the produced yield dataset is 1982-2020 instead of 1981-2021 since the AVHRR data used in this manuscript is from 1981 to 2021?
- In this paper, only the deep learning approach LSTM was used (RF is a traditional machine learning approach instead of deep learning approach), thus change "deep learning approaches" to "deep learning approach" in title.
- Line 15, the sentence is incomplete. "to map spatial distribution of xxx".
- Line 19, how 82% calculated was not described in the main content, and at which spatial scale, for which time period?
- Line 19-20, the comparison of GlobalWheatYield4km and SPAM at which spatial and temporal scale?
- Line 40, the training of statistical models needs a lot of data, and heavily depend on calibration data. Why you say that less dependence on calibration data?
- Line 73-74, the sentence is incomplete.
- Line 88, the GFSAD1KCM only provides a combined cropland mask, how you get the spatial distributions of wheat from this dataset?
- Line 145, how you deal with the gaps in NDVI dataset?
- Line 149, since the cropland mask is in 4km, almost all the pixels are mixing pixels, how you deal with that?
- Line 185, the nRMSE is 13.1 and 16.2 for LSTM and RF, are some typos here? 13.1% and 16.2%? Same for Line 263 and Figure 3.
- Line 196, is this the out of sample performance, at which spatial scale, for which time period?
- Line 217, for the comparison of GlobalWheatYield4km and SPAM in 2000, 2005, 2010, it may be not fair, due to the crop yield of SPAM is the nominal value for several years (e.g., 2010 is for 2009-2011). Please check the papers about SPAM to make sure the

comparison is correct.

- Please provide the maps of GlobalWheatYield4km and SPAM performance. Since the SPAM used less data than GlobalWheatYield4km in China, maybe the major improvement of GlobalWheatYield4km compared with SPAM is in this region.
- Please provide the uncertainty of GlobalWheatYield4km in the released dataset also the result part.