

Comment on **essd-2022-298**

Anonymous Referee #2

Referee comment on "AI4Boundaries: an open AI-ready dataset to map field boundaries with Sentinel-2 and aerial photography" by Raphaël d'Andrimont et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-298-RC2>, 2022

The preprint "AI4Boundaries: an open AI-ready dataset to map field boundaries with Sentinel-2 and aerial photography" introduced two data sets for extracting field boundaries (10-m Sentinel-2 monthly composites and 1-m aerial orthophoto), with the labels extracted from the GSAA data. In general, the datasets proposed by the authors could be promising in the algorithm development for parcel boundary detection.

However, I think this preprint needs more attention to get ready for publication. I have several specific comments:

- The GSAA data quality: There is no information about this data, such as how parcel boundaries are derived, their accuracy, etc. The authors may add more information in the introduction section. Further, Figure 6 also shows some missing fields in vector label (B). Although the authors mentioned some limitations of label data, the ability and accuracy should have been addressed in the beginning to show the potential of this data. Note that this is the main data used for training and testing all ML/DL models to map field boundaries; thus, it has to be highly accurate and confident.
- The ability of the 10 m Sentinel 2 monthly composite dataset for mapping field boundaries in European countries: A table of field-size statistics is needed. In some areas, field size may be small for large-scale analysis using 10 m data. Do the authors think 10 m data is able to capture small-size fields like in Figure 7? I think it is a critical concern here.
- The median monthly composite approach: This seems to lack agreement. What are the current data preparation methods used for field boundary mapping? Do the authors think we may add more noises to the data when we do linear interpolation and smoothing before calculating the median composite? Although the authors provide some insights in the limitations and perspective section, I would suggest adding references that used this composition method for field boundary detection and other applications (e.g., crop mapping) to show the effectiveness of the method.
- It is worth expanding the algorithm development for field boundaries detection in Europe and other countries like the US, their advantages, disadvantages, data,

accuracy, etc. This is important to lead the use of Sentinel-2 composite data for large-scale analysis and median composition method. Is there a difference if we use only "clear" for median compositing and then perform interpolation later to fill gaps?

- For aerial orthophoto and labels: we need a deeper understanding of matching them with satellite Sentinel-2 images instead of only using reprojection. We know that the accuracy of geolocations of aerial orthophoto could be less accurate than well-operated satellites. So, it is also a vital concern to provide accurate labels for training and testing (Sections 2.2.2 and 2.3).
- Besides, the manuscript lacks citations to support its statements. For example, a statement in lines 25-26 needs citations, the GSAA needs a link to the source, etc. Therefore, I encourage the authors to go through the paper and add citations, links to sources, and evidence to support the statements.
- Many abbreviations without pre-definition include EO, BDAP, NDVI, WMTS, WMS, EU, ESA SNAP, etc. It may be easy for some users, but it could be hard to understand for others from different backgrounds, such as computer science.
- Section 2.4: Is it a random split for all samples? Is it better to randomly split for each country with a similar percentage for training, validation, and testing?

Line comments:

Lines 40-48: Worth to add more information about the GSAA data as my first main concern.

Lines 49-52: Contents look quite similar in lines 32-35

Line 69: Add more information. Why do we need to use stratified random sampling rather than other sampling methods?

Lines 77-80: It is hard for me to follow. "A sample of 10,000 units..." means only one sample with 10,000 units of parcel? Why do we have 170 per stratum?

Line 82: Explain more about the distribution of samples for each country. It seems like we have denser samples in Austria and Netherlands. (Figure 3 needs a background map of the countries' boundaries).

Lines 84-85: Not clear

Lines 89-90: This sentence is not clear as well. "The number of pixels for orthophoto data set had to be increased" from ... to ...?

Lines 94-98: We already have ready-to-use surface reflectance (BOA-Level 2A). Why do we use TOA?

Figure 4: Can we separate Sentinel-2 RGB and aerial orthophoto (don't overlap them) and scale them up to clearly see the fields? It is good to compare two data sets to understand the field boundary characteristics.

Lines 103-104: The SCL is 20 m spatial resolution. Did the authors resample the SCL band to 10 m and then match it with 10 m bands?

Figure 5: This figure is important, but subpanels are too small to compare the Sentinel-2 field boundaries with their labels. The coordinates are too small, missing where it is (which country?). Missing D explanation in the caption. Scale sizes should match with each other (D-H do match with A, B, C).

Figure 6: Caption - "Example of the aerial ortophoto 10-m dataset..." is "Example of the aerial **orthophoto 1-m** dataset..."? The figure needs a side map showing the chip's locations.

Line 141: B is the vector label. Is it "(example in Figure 6 **C, D, E, F**)"?

Figure 7: Similar comments as Figure 5. I may have misunderstood the paper concept, but I feel it is crucial to compare 10 m Sentinel-2 chips, 1 m orthophotos, and labels to show their correspondence. This is because (i) Sentinel-2 monthly composite is proposed for large-scale field boundaries mapping, and (ii) ML/DL algorithms need labels for training and testing, but currently, I did not see the evidence between Sentinel-2 chips and labels.

Line 147: Add where the data are. (Data availability section)

Table A1: Some links do not work.