

Comments from Referee 2:

We have thoroughly revised our manuscript according to the comments and suggestions provided by the reviewer. We would like to thank her/him for its review, which allowed us to improve our manuscript. All the new and modified contents are in blue in the revised version of the manuscript.

In remainder of the document, lines in **bold** echo your comments for ease of reading, lines in **red** provide direct answers to your comments, followed in case by proposed modifications to our paper (with new elements in **green**).

We sincerely hope that these corrections will match your expectations.

This study has presented the ground data collected from nine selected sites in tropical countries within JECAM initiative. The area of site ranges from 250 to 11700 according to the location. This data is collected for either one or two years except Burkina Faso (7 years), Madagascar (4 years) and Brazil (3 years). As it is ground data collection manually – there is not accuracy assessment except set up of standardized procedures to collect data. Overall, this work has potential to add value in the problem area of availability of ground data for agricultural monitoring and can be considered for publication with major edits.

I have major and minor comments as follows:

Major comments

1. Why is the dataset called harmonized although it only depicts to be ground data?

→ In fact it is the ground data sets collected on each of the nine sites that are harmonized with each other, from data collection to final database format ready to use. We prefer the term “harmonized” to “standardized”, because part of the work was done *a posteriori*.

2. What is the clear definition of classes used to label? Author may add these definitions for more clarity and avoid confusion? For example, what is mean by croplands in this study? (do croplands include agroforestry, rangelands and horticultural crops too)

→ In the “2.2 Data Collection” section the attribute “keywords” is described (see extract below) and aims to provide harmonized and generic keywords describing each class. These keywords are based on existing standards in terms of land use definitions (FAO, 2020 - see link below) that can be consulted to have a clear definition of each class or term used. When the JECAM nomenclature showed disagreements with the FAO land use definitions (eg. In the FAO definitions, the “fallow” class is considered as cropland, while it is not in JECAM nomenclature), priority was given to the FAO definitions. The full list of the land use classes and their associated keywords are now provided in Appendix B.

Extract of “2.2 Data Collection” Section:

“An attribute referred to as “Keywords” was also created in order to associate various generic terms (land cover, crop group, crop type, cropping practice, etc. (Appendix B)) to each polygon. This attribute has two objectives: (i) facilitating keyword search for the user, (ii) allowing the user to create his own nomenclature (hierarchic or not) with different levels of detail so that the nomenclature can be dedicated to the user's needs. These terms are based on the FAO land use definitions (FAO, 2020) and JECAM hierarchic nomenclature (Defourny et al., 2014), which were adapted to take into account the diversity of the farming systems in the surveyed sites.”

Link of FAO land use definitions (see “References”):

FAO: Land use, irrigation and agricultural USE, IRRIGATION AND AGRICULTURAL PRACTICES - DEFINITIONS:

http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Definitions/Land_Use_Definitions_FAOSTAT.xlsx, last access: 9 September 2020.

3. For every site and region – summer crop and winter crop seasons are separate – if author is using this term – I would recommend to use it carefully as the seasons although name is same but months are different location-wise. In short , provide definition of winter crop and summer crop – and in the label – add unique name if possible.

→ We agree with the reviewer, and it is why in the database we have the attributes Start of Season (SOS) and End of Season (EOS) that gives the validity period for the crop type recorded for each polygon (see Table 1), and why the terms “winter crop” and “summer crops” do not appear in the keywords list (see Appendix B).

However, we maintained the term “summer” and “winter crops” in the description of the Brazilian sites (Section 2.1) because the months concerned are specified, but in Table 1, to avoid confusion, we replaced “For each field in the Tocantins site, the operator recorded the crop type of the 2 seasons (summer / winter) by observing the crop residues on the field or by interviewing the farmers” by *“For each field at the Tocantins site, the operator was able to record the crop type for the two cropping seasons by observing the crop residues in the field or by interviewing the farmers”*.

4. Figure 1 is totally misleading. The selected site cannot possibly represent the entire tropical farming systems. I would suggest to remove it as it is not adding any value. Main important is to highlight the sites and author can do that by showing site zoomed regions rather than misleading with unnecessary presentation.

→ We regret that the representation in Figure 1 may have given the impression that we consider our sites to be representative of the entire tropical farming systems, while we just intended to illustrate the diversity of our sites in terms of tropical agrosystems. But as Figure 1 is misleading, we have simplified it to provide only information on the location of the sites.

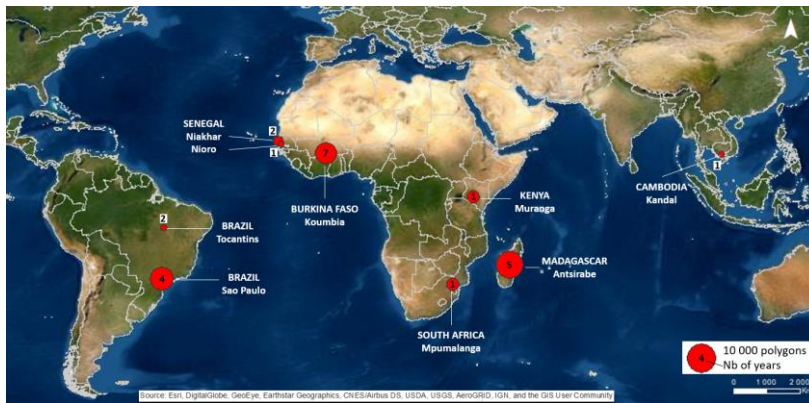


Figure 1. Location map of the study sites, and the associated number of collection years and sampled plots (symbolized by the size of the red circles).

5. Study site – the explanation of study sites in section 2.1 is repetitive of what’s there is in table 1. I would recommend rewriting to avoid repetition.

→ Thank you for bringing this to our attention. There is probably a misunderstanding about Table 1 (now Table 2 in the new version of the manuscript) due to its positioning prior to the database description. Indeed, the table mainly provides statistics about the final database, such as the total number of records per-site (polygons), their average size, etc. That is why we named it “Database synthesis”.

To avoid confusion, Table 1 was moved to section 3 Data Records, just after the database description and corresponds now to Table 2; the column “cropping pattern” and the size of the study site were removed because of information redundancy with the text in section 2.1.

6. Table 1 name is synthesis of database – which is not quite correct – I would recommend just naming it as “study area description” or related.

→ Please, cf. to the answer to Comment 5. We hope that in the revised version of the paper the description of the sites and the description of the database will be clearer.

7. Table 1 can have additional columns such as season, temperature, major crops, average precipitation etc.

→ As explained in the answer to Comment 5, Table 1 was not intended to report geographical description of the sites, which is indeed provided in the text of Section 2.1. Most of the requested information is already in the aforementioned text. Sorry again for the misunderstanding.

8. Data collection protocol need to provide with more details as it is important step in this data. I would recommend to explain it with sample examples of data points and showing the standardized format along with flowchart if possible.

→ Data collection protocol is indeed a very important step of our work. In order to help the understanding, we added a workflow of the data acquisition in the Section 2.2, showing: the field data

entry form filled for one plot, GPS waypoints and then boundaries digitized for each GPS point, displayed on a satellite image in false color.

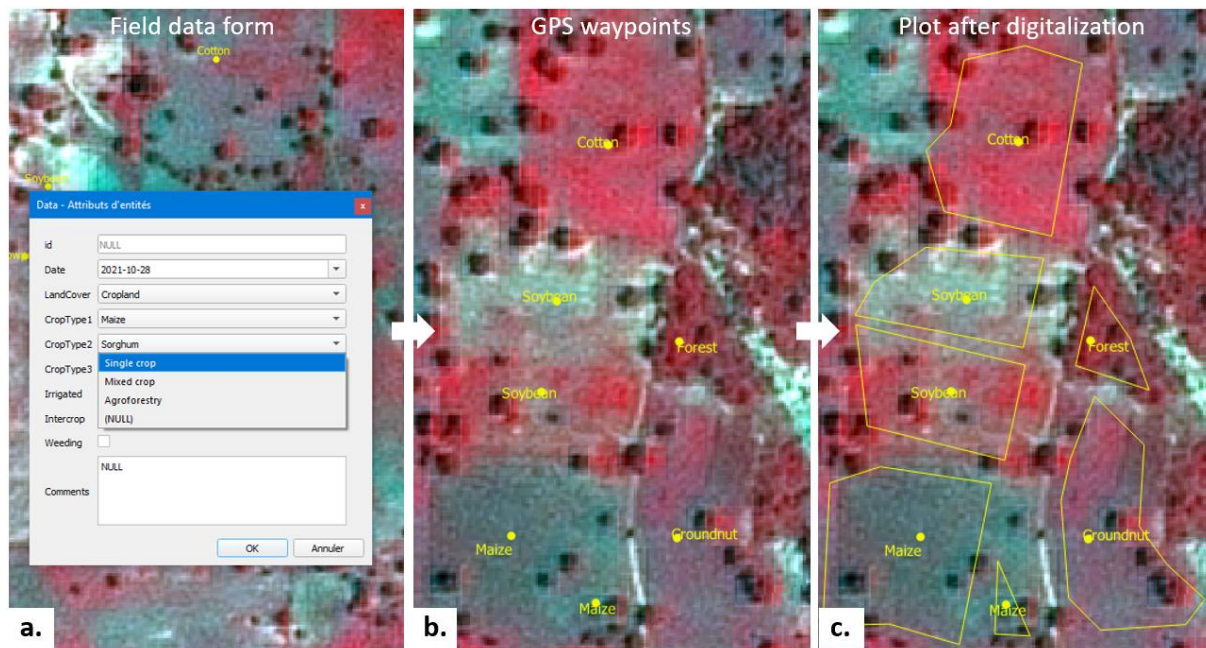


Figure 3. Workflow of the data acquisition: (a) field data entry form used on the GPS tablet, (b) GPS waypoints acquired in the field and (c) corresponding plots after digitalization of the boundaries, displayed on a satellite image in false color (Red: Near Infrared band, Green: Red band, Blue: Green band).

9. Post processing of data may add many additional errors to the raw data point collection with the provided steps by author. I would suggest to provide more details and explanation about how the manual error were avoided? In short, provide the framework in methodological format. (although – the step is performed by same personnel who did survey – it is not valid explanation of expertise or scientific explanation)

➔ Digitization of each field or non-crop entity boundaries was performed on very high spatial resolution images that were acquired specifically for this purpose just before the field campaigns. This step was done quickly after the field campaigns and in the same order as the GPS waypoints records so that the operator can well remember each one. This digitization was performed by remote sensing experts who also participated in the field surveys, and had very good skills in visual interpretation of satellite images. Furthermore, photographs were available for each waypoint to ensure the consistency between the ground and satellite information. Finally, no doubts on entities boundaries were accepted, and the entities having fuzzy boundaries or too much heterogeneity were discarded.

10. In section data records – one of the column is data source – and there are three data sources – As it is important information – I would recommend author to provide a number on how many samples are “0” , “1” and “2” as data is mainly labelled as in situ

➔ To illustrate the relative importance of the data source, the following figure was added in the document :

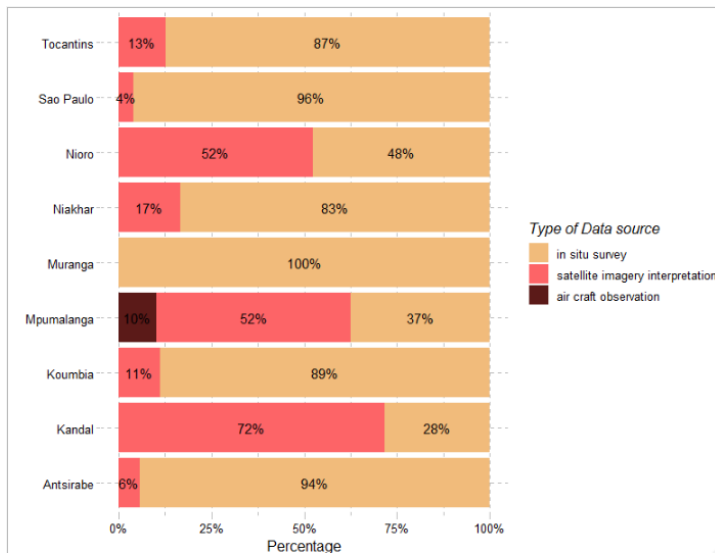


Figure 6. Distribution of the data sources, given in percentage of the total number of polygons per site.

11. Detailed explanation on data source or data collection is needed regarding minimum size unit (MMU), labeling strategies, mix land use class and other details related to land use.

→ In section 2.2, we refer to a minimum homogeneous field size which can be considered for waypoints recording of 0.04 ha. We modified the text to explicitly mention the notion of Minimum Sampling Unit.

“Waypoints were only recorded for homogenous fields/entities of at least 20 x 20 m² (against a *minimum sampling unit* of 0.25 ha with a minimum width of 30 m in JECAM guidelines).”

Regarding the other information acquisition (such as labeling strategies, mix land use class, irrigation,..), we used a data entry form to facilitate the data entry : scrollable lists, checked box,... Some examples of lists are now displayed on Appendix A.

12. Total crop samples are ~20,257 and non-crop are related very low – how did you decide this number? Need further explanation on classes, their sampling size, sampling and labeling strategy.

→ This is an important point. Overallly speaking, the number of crop vs. non-crop samples is not fixed beforehand in each field campaign, but is a direct consequence of the opportunistic sampling approach described in Section 2.2. In other words, these numbers reflect the ratio of crop vs. non-crop surfaces observed along the covered tracks. However, literature on remote sensing based crop mapping shows that the problem of “extracting cropland” in a landscape is a relatively easier task with respect to detecting different agricultural land uses (e.g. crop types). This also means that the need for annotated surfaces over crop classes is definitely more important that for non-crop ones. Of course, this does not mean that the proposed database can be considered exhaustive for non-crop classes, but this is not its purpose. Concerning sampling strategies, there is no difference between crop and non-crop surfaces.

13. What is the overarching goal and novelty in this dataset? I understand it is very important

dataset – but author need to add its novelty and goal of research in introduction for more clarity.

→ Yes, you are right. We need to emphasize the dataset quality and the final goal. We added a part at the end of the introduction section in that respect:

The experiment has been operating since 2013, and some in situ datasets produced at the field scale have been used in different benchmarking mapping studies (Waldner et al., 2016; Inglada et al., 2015). However, only a part of the collected ground data was used in these studies and the databases are not publicly shared.

To make agricultural land use data publicly available to the remote sensing community, for classification algorithm benchmarking or LULC product validation for example, an important work of harmonization of in situ JECAM and JECAM-like agricultural land use datasets was undertaken for nine sites located in the tropical belt. The acquisition protocol was adapted from Defourny et al. (2014) to take into account the characteristics of tropical agriculture (e.g. small field size, accessibility). At each site, information on crop type and cropping practices was collected locally, at the field level, with a detailed nomenclature. The acquisition period was between 2013 and 2020, and the number of monitoring years per site was between 1 and 7.

In this paper, we describe in detail the study sites, the data collection protocol and the structure of the final database. We then discuss how the harmonization of the dataset and the diversity of the studied agrosystems, including small-holder farming, make our dataset unique and valuable for applications in the emerging/developing countries in the tropics.

Minor comments:

14. Line 40 and 41 – can be split to two sentences to avoid complexity

→ Indeed, the sentence is very long and not very pleasant to read. We splitted it in 2 sentences as recommended.

These datasets can be used to produce and validate agricultural land use maps in the tropics. They can also be used to assess the performances and robustness of classification methods of cropland and crop types/practices in a large range of tropical farming systems.

15. Validation through study cases – is confusing section – is it application of dataset or validation?

→ Thank you for this comment. The sub-section “4.3. Validation through study cases” was removed from the “4. Technical validation” section, and converted into a new section titled to “5. Dataset application study cases”.

16. Overall, writing needs to improve for spelling and grammar – I would recommend professional English proof-reading – I had real difficult time in reading this paper.

→ We are sorry to hear that you had difficult time reading the paper. The revised version has been reviewed by a professional English native speaker (cf. the join certificate).

17. Title of article contains “JECAM” – which needs to be expanded?

→ In order not to make the title of the article too long, the word JECAM has been removed from the title. However, the word “JECAM” appears in the expanded form in the abstract.

18. Abstract is misleading in many aspects such as – data time from 2013-2020 (which is not true as most of the sites has data from 2 years only). I would recommend author to be careful and precise facts in the abstract for more clarity and description of work.

→ Thank you for this comment. In order to be more precise, we changed the text from *“In this paper, we present nine datasets collected in a standardized manner between 2013 and 2020 in seven tropical and subtropical countries within the framework of the international JECAM (Joint Experiment for Crop Assessment and Monitoring) initiative”*,

to

“In this paper, we present a database made of 24 datasets collected in a standardized manner over nine sites within the framework of the international JECAM (Joint Experiment for Crop Assessment and Monitoring) initiative; the sites were spread over seven countries of the tropical belt, and the number of data collection years depended on the site (from 1 to 7 years between 2013 and 2020)”.

19. For small field sizes – what was the strategy to collect data – how would it be homogeneous to the data collection strategy/

→ No specific strategy was applied for small field sizes. On the protocol, the minimum sampling size was 20*20m (0.04 ha) and attention focused primarily on the homogeneity of the fields, both on the ground and on satellite images used for boundary delineation.

20. Overall, I would recommend author to improve readability of the article.

→ We hope that the English proof-reading and the revisions following the reviewer's comments have improved the readability of the article.

We want to thank you again for your thoughtful reading.