

Response to the reviewers

We thank the reviewer for taking the time to go over our work. In the following, we address their concerns point by point.

Where required, we also provide excerpts of the manuscript (MS) with removed text marked in red and crossed over, new text in blue and underlined.

Reviewer 2

Reviewer Point P 2.1 — * In 20: the link to the data description paper is broken

Reply: We have checked this link, and found it working and pointing to the right repository.

Reviewer Point P 2.2 — * In 230: how were the splits performed? how were features selected for in the random forest classifier to avoid overfitting?

Reply: We have now reported the average overall accuracy resulting from a 5-fold cross-validation to test the robustness of the reported statistics. We also now include a figure showing the feature importance for the classifier, and a brief discussion in Section 4:

(Section 3.5)

5-fold crossvalidation was also used to evaluate the robustness of the results shown in Table 2, which resulted in an overall accuracy of 0.72 (standard deviation 0.12). The importance of each considered features is shown in the Gini index plot shown in Fig. 15, which ranks different features by classifier importance.

Section 4.3

The maize mask that was developed in this paper demonstrates that the data can be used as an input to a classifier. However, the limited number of samples for 2021 (where the main aim of the field campaign was biophysical parameter collection) result in a crop mask that is probably only reliable around the collected data points. Also, since the surveyed fields were selected as late sown, this may also bias the field selection. Fig. 15 indicates that the classifier is mostly being driven by observations around the first half of June (DoYs 150-165), suggesting that early crop development may be more informative for crop discrimination than late crop development.

Reviewer Point P 2.3 — *In 445: the link to the code for classification seems broken or unavailable

Reply: Fixed, thanks!

Reviewer Point P 2.4 — * In 455: In situ biophysical parameter time series csv file does not seem available at the zenodo page provided.

Reply: The biophysical parameters are stored in file `Ghana_ground_data_v5.csv`. This file has been part of the data set, but as a way of clarifying where data are, we have added the filenames to Section 7.

Reviewer Point P 2.5 — The dataset lacks a metadata which could be useful to describe the data for users who may want to explore the data further.

Reply: We hope that this publication acts as a thorough description of the data set and a list of possible uses, and the dataset will be submitted to MLHub (<https://mlhub.earth/>), to give it wider visibility.