

## **Response to the reviewers (#ESSD-2023-143)**

Thanks for the positive comments from the Reviewers. The reviewers' requests are repeated below, in italics, and with our responses written below each suggestion. We have responded in full to each request.

### **Reviewer #4 (Remarks to the Author):**

We sincerely appreciate the reviewer's encouraging words and constructive comments. All issues have been adequately addressed both below and in the revised version of the manuscript.

Major questions:

#### **[Reviewer #4 Comment 1] Data validation**

*The text says, "We collected other validation datasets to assess the tree cover products during 2016-2021" (Line 111). Does it mean the authors went to the validation regions to check the forest states in person? Or, you are using other datasets or interpretation methods to do validation?*

*In the following paragraph, the authors says, "we randomly generated 1,515 points to ensure the representativeness of collected visual samples". It seems there were field checking. However, in the following sentence "these points were labeled these points as forests or non-forests by four human interpreters using Planet Explorer of QGIS.". Please clarify it.*

**[Response]** Thanks a lot for pointing this out.

Here, we enlisted four human interpreters and employed Planet Explorer within QGIS by interpretation methods to generate a robust annual validation dataset from 2016 to 2021 except for 2019 for investigating the accuracy of the generated time series tree cover map product.

We have revised the text in Section 2.2, i.e., "(Yang et al., 2023). However, despite the

advancements in the Land Cover Land Use Change (LCLUC) community, a notable gap remains the absence of publicly available high-resolution (e.g.,  $\leq 10$  m) tree cover/non-tree cover labels. The existing coarse-resolution labels for tree cover/non-tree cover can introduce considerable uncertainties when evaluating high-resolution tree cover maps. As a result, our ability to delve deeper into the accuracy of time-series tree cover map datasets was hindered.

Following the methodology established by Yang et al. (2023), we undertook a rigorous process to generate a robust validation dataset for our study. Firstly, we randomly generated 1,515 points to ensure a representative sample of collected visual data, as illustrated in Fig. 1. Next, to classify these points as trees or non-trees, we enlisted four human interpreters and employed Planet Explorer within QGIS. Our approach involved visually identifying tree cover/non-tree cover pixels in the true color composite of Planet-NICFI imagery where the points were located. To ensure accuracy, we superimposed the 10 m tree height data, previously developed by Lang et al. (2022), onto the Planet-NICFI imagery. This step ensured that the labels adhered to the specified tree height criteria (i.e.,  $\geq 5$  m). Subsequently, we thoroughly evaluated and refined the labels using Google Earth. To make time series tree cover/non-tree cover labels, we maintained the geographic location of the 1,515 points and changed the year of the Planet-NICFI imagery. The resulting labels encompassed data from the years 2016, 2017, 2018, 2020, and 2021. Comprehensive information about the validation dataset can be found in Table 1.” (P6L116-P7L146 in the track version of the revised manuscript).

**[Reviewer #4 Comment 2]** *In the section “Statistical accuracy assessment”, the authors mentioned “the user’s accuracy, producer’s accuracy, and overall accuracy”. Please explain what they are. These terms may be well known in the authors’ discipline. However, as a reader of this paper, I don’t know what they are, and many readers (and potential users) may have the same issue.*

**[Response]** Thanks a lot for pointing this out.

We employ widely used accuracy assessment metrics in the LCLUC community to investigate the accuracy of our data. Specifically, the generated tree cover map products are compared pixel by pixel with the labels. Then, a confusion matrix can be obtained, including true tree cover (TP), true non-tree cover (TN), false tree cover (FP), and false non-tree cover (FN). These four values were used to calculate the accuracy assessment metrics of the draft (Table R1).

**Table R1** Product evaluation metrics and corresponding equations.

Metric	Equation
User's accuracy (UA)	$\frac{TP}{TP + FP}$
Producer's accuracy (PA)	$\frac{TP}{TP + FN}$
F1-score	$\frac{2 \times UA \times PA}{UA + PA}$
Overall accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$

We have also added the text in Section 2.3.4 of the revised manuscript, which is “product. The generated tree cover maps product is compared pixel by pixel with the tree cover/non-tree cover labels. We then obtained a confusion matrix, including true tree cover (TP), true non-tree cover (TN), false tree cover (FP), and false non-tree cover (FN). These four values are used ..... based on Eqs. (1)-(4), respectively.

$$\text{User's accuracy (UA)} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Producer's accuracy (PA)} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Overall accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{F1 score} = \frac{2 \times UA \times PA}{UA + PA} \quad (4)$$

” (P10L226-P11L242 in the track version of the revised manuscript).

**[Reviewer #4 Comment 3]** Section “6 Conclusions”

*For a data paper, we have read the abstract, and understand how the data were generated/collected, the scope and uncertainty of the data, and know how to get them.*

*Do we really need a “conclusion” section?*

**[Response]** The points are well taken!

In our study, we make use of published data description papers that contain distinct components, such as "abstract" and "conclusion." To ensure clarity and differentiation between these sections, we adopt specific approaches for each.

In the abstract, we introduce the background and research question as well as how the data were generated/collected, the scope and uncertainty of the data, and how to get them, etc.

On the other hand, the "conclusion" section offers a deeper analysis of the implications of our data. Here, we go beyond the introductory aspects mentioned in the abstract and delve into the broader significance and potential applications of our findings. By emphasizing the implications of our data, we aim to provide a comprehensive understanding of its relevance and impact in the context of the research question and beyond.

Minor questions and edits:

**[Reviewer #1 Specific Comment 1]** *Line 28: Please clarify what “annual samples” are.*

**[Response]** Thanks. We have changed annual samples to annual tree cover/non-tree cover samples.

**[Reviewer #1 Specific Comment 2]** *Lines 29~30 “with an overall accuracy of  $0.867\pm 0.017$  and a mean F1 score of 0.921, respectively.” Please explain what “overall accuracy and F1 score” are. My opinion, either explain it clearly or don’t mention it. In the previous sentence, authors have said “high accuracy”. Since I don’t know “overall accuracy and F1 score”, it is still “high accuracy” to me.*

**[Response]** Thanks. We reported high accuracy of our tree cover map product measured by overall accuracy and F1 score.

**[Reviewer #1 Specific Comment 4]** Lines 31~34 “Compared to existing maps ...”. These sentences can be removed. Add more details about your data published with this paper.

**[Response]** Thanks. It is Okay to keep it.

**[Reviewer #1 Specific Comment 4]** Lines 37 “The annual Planet-NICFI V1.0 tree cover map products from 2016 to 2021 at 4.77 m resolution”. I would replace it with “Our data”. It is not necessary to repeat the same information in the abstract.

**[Response]** Thanks. Have revised it.

**[Reviewer #1 Specific Comment 5]** Lines 112~115. Please rephrase this section so we can understand why “except 2019”.

**[Response]** Thanks. We have added the descriptions “as it has been provided by Yang et al. (2023).” (P6L115 in the track version of the revised manuscript).

**[Reviewer #1 Specific Comment 6]** Line 118: remove “these points”. repeated.

**[Response]** Thanks. We have removed it.

**[Reviewer #1 Specific Comment 7]** Line 119 “four human interpreters”: Do you mean you asked four colleagues (Homo sapiens) to do a test of identification? Are they acknowledged?

**[Response]** No, we enlisted four human interpreters and have acknowledged them in the revised manuscript.

**[Reviewer #1 Specific Comment 8]** Line 139: remove “For example,”.

**[Response]** Thanks. We have removed it.

**[Reviewer #1 Specific Comment 9]** Line 176. Please explain what are “user’s accuracy, producer’s accuracy, and overall accuracy”.

**[Response]** Thanks. We have added four equations to explain them, which are expressed as

$$\text{User's accuracy (UA)} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Producer's accuracy (PA)} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Overall accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{F1 score} = \frac{2 \times UA \times PA}{UA + PA} \quad (4)$$

We have added them to Section 2.3.4 and also added the text to describe them in the revised manuscript. Please see the response of [Reviewer #4 Comment 2].

**[Reviewer #1 Specific Comment 10]** *Lines 186~191: I cannot get what the first approach is.*

**[Response]** Thanks a lot. The first approach is four metrics (user's accuracy, producer's accuracy, and overall accuracy). They are the most commonly used metrics for evaluating data generated.

**[Reviewer #1 Specific Comment 11]** *Line 188 "based on a study by Tsendbazar": Please explain what it is.*

**[Response]** Thanks. We have added the words "**the methods developed**" (P11L253 in the track version of the revised manuscript).

**[Reviewer #1 Specific Comment 12]** *Line 195 "The results for 2019 were provided by .." can be move to the method section*

**[Response]** Thanks a lot. We have changed "The results for 2019 were provided by ..." to "The tree cover accuracy results for 2019 were provided by...".

**[Reviewer #1 Specific Comment 13]** *Line 261 "minimum tree height ...": is this about the definition of forest?*

[Response] Yes, we have added the words “...for further generating forest data” (P19L362-363 in the track version of the revised manuscript).

[Reviewer #1 Specific Comment 14] Lines 268~269. I think this sentence is about algorithm “random forest”. Please rephrase this sentence.

[Response] Thanks. We have added word “based” after random forests.

[Reviewer #1 Specific Comment 15] Line 270; “U-net”?

[Response] Thanks. U-Net is a convolutional neural network that was developed for biomedical image segmentation at the Computer Science Department of the University of Freiburg.

We have also added a reference ((Falk et al., 2019)) after the “U-net”.

**Reference:**

*Falk, T., Mai, D., Bensch, R., et al. U-Net: deep learning for cell counting, detection, and morphometry. Nature methods, 16(1), pp.67-70, 2019.*

[Reviewer #1 Specific Comment 16] Line 301 Section “Acknowledgements” who are those four “human interpreters” mentioned in the main text?

[Response] Thanks. We added them to “acknowledgments”, i.e., “We thank Sen Jiang, Haowen Duan, Hao Li, and Fangdong Fu for making tree cover/non-tree cover label data that are used to assess the time series tree cover map products.” ” (P21L408-410 in the track version of the revised manuscript).