

Earth Syst. Sci. Data Discuss., author comment AC3  
<https://doi.org/10.5194/essd-2021-281-AC3>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC3

Femke van Geffen et al.

---

Author comment on "SiDroForest: a comprehensive forest inventory of Siberian boreal forest investigations including drone-based point clouds, individually labeled trees, synthetically generated tree crowns, and Sentinel-2 labeled image patches" by Femke van Geffen et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-281-AC3>, 2022

---

Dear Reviewer, Thank you very much for your time and effort. The review we received is included below in regular text and the changes that are included in the manuscript and the responses are written below each point in bold and cursive starting with the word response. I, Femke van Geffen would like to, also on behalf of my team, thank the reviewers for their time and detailed comments on the manuscript and dataset. The comments greatly improve the work and are well appreciated. Kind regards, Femke van Geffen.

Review 3:

General comments

The author collected tree - and plot-level forest structure data based on unmanned aerial vehicle and field investigation in two vegetation transition zones of Siberia, Russia. The datasets, including field plot level individual tree and shrub records (tree height, crown diameter, and species) and UAV products (e.g., Canopy Height) can be used for calibration and verification of model output, experiments, or observations. They are useful and important for future carbon dynamic studies and help to inform forest management, especially for the area where historical records and monitoring tend to be scarce. However, the manuscript is poorly organized and difficult to read. I encourage publication after addressing the following issues.

Significance:

The data are useful, complete, and fill in the region field data gap for Siberia boreal forests.

Data quality:

The data are easy to understand and presented readily and accessible to be used in other studies.

Presentation quality:

The manuscript was not logically articulated and was poorly written.

I recommend reframing the Introduction section. It should be more concise. I suggest shortening the detailed SiDroForest dataset and collection method description, but introducing the necessary or implication of each dataset (such as tree-level or plot-level records or canopy Height) and discussing the importance and challenges to collect these data.

**Response: We have taken this suggestion into consideration. Please see new structured manuscript with the tracked changes that has been submitted.**

The results seem like a duplicate of the method. I suggest including some further information and analysis, not only what the data were included or collected in the four datasets. For example, the frequency or distribution of tree species in the area for 3.2 Dataset 2, and the dominant vegetation classes and their distribution for Dataset 4 can be described.

**Response: Thank you for your feedback. In fact, we detected some repetitions and optimized the data and methods and result section. We removed repetitive information and put highlights into the result chapter. Several of the results are visualized in the Appendix, e.g. the distribution of vegetation species in A2, and Percentage vegetation cover per plot in Yakutia for only large shrubs and trees (>1.3m) in A3. The dominant vegetation class of the Sentinel-2 patches are shown in Table 3. Please also see the track change file as an overview on our edits in chapter 2 'data and methods and chapter 3 'results'.**

Specific comments:

Line 68: change "are" to "is"

Lines 106-109: The sentence is difficult to read.

Lines 119-121: delete one of the "use" in the sentence

Line 124: add "which was" before "derived from"

Line 129: change "was" to "were"

Line 150: add "," before "and"

Response: Above specific comments were all taken into account and corrected for.

Lines 243-244 and Figure A4: Two 30-m-long tape or 15m?

**Response: We investigated a circular plot area and used two 30-m long tape measures to segment the circle with a 15 m radius into four quadrants.**

**L195-200 "In the field, two 30-m-long tape measures were laid out along the main cardinal directions, intersecting in the plot centre, marking the main axes of a circular area with a radius of 15 m. A minimum of ten individuals of each tree and shrub species present were selected per plot. For each individual tree we measured the stem diameter at breast height and at the base. The tree crown diameter, tree height, and vitality were estimated as described in Brieger et al. (2019). There were three deviations from the standard method of vegetation inventory. On plot EN1814 and EN1865, all trees were recorded, and plot EN18070 was recorded by a transect with three segments: edge, transition, and**

**centre.”**

Lines 243-248: why were a minimum of ten individuals selected per plot? Why were plots recorded differently? If only part of the trees and shrubs were recorded, the data may not be able to represent the real forest information.

**Response: All plots were recorded with the same method. All trees present on the plot area were recorded in the field by its height, species and vitality information. Additionally, a minimum of 10 individuals per species were selected for in-depth analyses and can be found in the Species polygons. Shrub cover per taxa was recorded in the field as well but only a minimum of three individuals selected for in-depth analyses. The UAV-derived Species polygon is supplied for classification tasks and marks a clear tree or shrub from a certain species. It includes a minimum of 10 recorded trees or shrubs per species, where possible the same as recorded in fieldwork.**

Line 252 please clarify the 11 vegetation classes here.

**Response:**

**A clarifying sentence is added to L201-204: "We post-fieldwork assigned 11 vegetation classes to the 64 plots (table A1). The class assignment was based on the previous classes determined by Shevtsova et al. (2020a) for Chukotka. For plots in Central Yakutia, we applied a similar method incorporating principal component analysis (PCA), tree density information from the UAV data, and recorded tree species information per plot (Fig. A2, A3 show the field data information)."**

Lines 256-259 and Figure A4: The vegetation plot looks smaller than 30m ×30m (the double grid) if the red line is 15m long.

**Response: Yes, the double grid in figure A4 is not the 30 m x 30 m vegetation plot, but the UAV double grid flight line set up around the 15 m radius vegetation plot, the blue grid covers around 50x50 m square overflight grid and the orange circular flight mission with a radius of ca. 25 m around the center was carried out with the camera pointing towards the center.**

**Figure A4: SiDroForest unmanned aerial vehicle (UAV) data acquisition and flight pattern consisting of a double grid (blue) and a circular mission (orange) around the vegetation plot. (Both images are attached with this message)**

Figure A1 The two 15 m long grid lines (red) divide the plot area into four quadrants of similar size (yellow). From Brieger et al. (2019).

Line 274: with very low vegetation?

**Response: The vegetation that is near to the ground i.e., low structure vegetation. Often this is small shrubs and small trees that are hard to segment because they are close together. We changed the text accordingly to 'low structure vegetation'**

Line 295: add "that was" between "area" and "not"

Lines 301-304: Please break the long sentence into several simple sentences.

Line 326: add "that were" before "corrected"

**Response: All three suggestions above were followed.**

Lines 338 -340: the sentence is difficult to read. Does that mean the tree crowns were captured by two methods: 1) watershed segmentation analysis and 2) successive automatic generation of a polygon around them?

**Response: Yes. I corrected the sentence to be clearer.**

Line 542: Link or Linking?

**Response: Link.**

Lines 553-557: what's the meaning of this paragraph? Do you compare your field-measured crown diameters and detected crown polygons? If so, what's the difference between your results and the results of Brieger et al.?

**Response: The paragraph introduces that the automatically detected tree crowns are likely better fitting than the field estimations for each tree. We present here the detected tree crowns and did not run an analysis such as in Brieger et al. 2019.**

Please also note the supplement to this comment:

<https://essd.copernicus.org/preprints/essd-2021-281/essd-2021-281-AC3-supplement.zip>