

Report 1:**Comment:**

The authors highlight the importance of forest monitoring for climate change analysis, referencing shifts in species composition and biomass. While the dataset provides valuable spatial coverage, such analyses typically require time-series data in addition to spatial distribution. For instance, shifts on species composition usually need long time-series (~ 25 years). This aspect should be better explained in the Introduction. Please clarify how the present dataset contributes to climate-related research.

And:

Regarding the following sentence: “Shifts in species composition require the delineation of individual trees and their crowns with high spatial resolution remote sensing, also providing large area coverage.” The sentence attempts to link climate-related forest changes to the data processing requirements. However, stating that delineating individual trees is the main requirement for determining species composition is a bit strange. Typically, species classification approaches are central to such analyses. I recommend reformulating this sentence to emphasize that climate impact studies require tree-level information, which must first be obtained through close-range data acquisition and processing, including tree identification, segmentation and species identification. Consider improving the logical connection from ecological objectives to the data processing methodology.

Response:

We thank the reviewer for this observation. The following new sentences were integrated to improve flow and logical connection between ecological objective and methodology, as suggested by the reviewer.

Assessing climate driven shifts in forest species composition requires detailed insights at tree level. Very high resolution remote sensing such as using drones and planes as platforms enables the detection, segmentation, and subsequent species identification of individual tree crowns over larger areas. This tree level detail allows the assessment of satellite derived forest cover products and therefore an increase in detail in timeseries analysis.

Comment:

Most of the references cited in the first paragraph are over six years old. I recommend updating the references to include recent state-of-the-art studies, particularly to support the ecological relevance and importance of the work.

Response:

We acknowledge, that previously provided references were outdated and that more recent developments in 2D and 3D based segmentation techniques apply deep learning (DL) for the task. The new references reflect this now.

Approaches for tree segmentation using 2D true-color images, infrared photography, or 3D-derived canopy height models (CHM) from Structure-from-Motion (SfM) previously relied on area-based techniques. Today, there is a strong shift toward 2D deep learning (DL) approaches for delineating crowns and extracting individual trees (Jaafar et al., 2018; Tong and Zhang, 2025; Aldaeri et al., 2026).

AND

Point cloud voxelisation and deep learning are more recent approaches for individual tree segmentation from 3D point cloud data (Jaafar et al., 2018; Brunner and Houtmeyers, 2022; Liu et al., 2023).

Comment:

I suggest removing the explanation of TOF and waveform from the Introduction, as it disrupts the narrative flow.

Response:

The sentence has been removed, as it indeed did not contribute and could be discarded.

Comment:

Figures 1 and 6 should be improved by increasing the font size to ensure they are suitable for publication

Response:

Both figures are now 10% larger

Comment:

Is the validation Section 2.8? Now it marks 2.1

Response:

It indeed is, the structure has been adapted since the Preprint.

Comment:

Figure 11-13 need to be better explained in the text.

Response:

Sentences have been added for each figure to incorporate them more and contribute to their explanatory value:

Shape differs between species, but also between individuals and age (height). This is clearly visible when looking at the relation of a species pointedness and height (Fig. 11), and the overlap of density functions of structural features for each species (Fig. 12, Fig. 13). Therefore, structural features tend to be less distinctive at the species level because many species share similar structural forms, especially within mixed stands or complex forests.

Comment

Table 6 could be merged adding one more column (RF1, RF2, RF3, RF4)

Response:

It has been merged into a short and wide lookup table

Comment:

Line 405-408 – This sentence is super nice, some similar insight could be include in the introduction when mentioning the related datasets

Response:

We appreciate the comment and adapted the last paragraph of the Introduction to incorporate the possible application of BorFIT for validation of large scale landcover products:

Here, we present the novel reference dataset "Boreal Forest Individual Tree" (BorFIT) (Kruse et al., 2025a), designed for artificial intelligence applications in individual tree segmentation. BorFIT comprises 384 LiDAR point clouds collected across 146 locations along bio-climatic gradients in Alaska, Canada, and Eastern Siberia, featuring up to 200 manually segmented trees per point cloud. Tree species were assigned using a Random Forest classifier trained on both structural (LiDAR) and spectral (RGB) metrics. By providing detailed tree-level training data across the northern boreal zone, BorFIT facilitates the assessment and validation of large-scale landcover and forest structure maps upscaled via spaceborne LiDAR (e.g., GEDI, ICESat-2), Synthetic Aperture Radar (SAR) or optical satellite data. Ultimately, dataset resources like BorFIT are essential for effective boreal forest monitoring, offering critical baseline information to evaluate climate change impacts on species distribution patterns.

Report 2:

Comment:

Comments 8 and 10 from reviewer 2 were not addressed in the text; I'd suggest revising the text to provide more detail about how sites were positioned, why you used this approach for selecting reference plot locations -- the stratification, 20m plot size, and number of plots selected? Were there formal criteria used or was this more ad-hoc? Documenting these decisions in the manuscript is necessary for assessing the data-generating process.

Response:

Thanks for pointing this out. We adapted the Section "Study Region" as follows:

Transects were established to enable long-term monitoring of ecosystem dynamics and responses to climate change and disturbances, such as wildfires. Sampling locations along these transects were selected prior to fieldwork based on topography, forest types, and proximity to highways for logistical accessibility. Site selection was guided by satellite-derived change detection and forest density metrics, alongside optical indicators of fire history, such as fire scars and time-since-fire products derived from datasets like the Hansen global forest cover change product (Hansen et al 2013).

And in the section "Reference plot generation":

This size was chosen because it allows for the segmentation of up to 200 trees per point cloud without cutting off branches. An area of 20 m x 20 m was found to be a good balance between file size, easy handling in CloudCompare and number of complete trees.

[...]

The number of subgroups was chosen to strike a balance between structural granularity and practical sample size, ensuring that the full vertical canopy profile of boreal forests was adequately captured without over fragmenting the tree height distribution.

[...]

To balance data processing efficiency with structural representation, tree selection prioritized canopy extremes and median height classes (groups 1, 5, and 10). Groups 3 and 8 were processed secondarily for gap-filling afterwards.