ForestScan Manuscript: essd-2025-67

Reviewer 1 — Hannah Weiser

We would like to reiterate our sincere gratitude to both reviewers for their thoughtful and constructive comments throughout the review process. Their feedback has been instrumental in improving the clarity, quality, and overall presentation of our manuscript.

Below, we provide detailed responses to the final comments. We have carefully considered all suggestions and implemented revisions where appropriate. In instances where we have retained the original wording, we have provided clear justification for doing so. We trust that our responses satisfactorily address the few remaining concerns.

Given that the manuscript has now been accepted subject to minor corrections, we would greatly appreciate it if the review process could be concluded and the paper advanced to publication at the earliest opportunity. Timely publication is particularly important as the CEDA archive has contacted us again requesting a copy of the accepted paper to add to our ForestScan Dataset Collection, which has become the most accessed dataset in their archive. To date, the ForestScan Collection has been accessed by more than 800 users across 77 countries, with over 385,000 individual accesses.

We thank you once again for your attention and look forward to finalising this process promptly.

GENERAL COMMENTS

3) In line with the comment by Reviewer 2, could you add a paragraph on leaf-wood segmentation quality? This may include a quantitative quality assessment (e.g., classification scores on a subset using manually labelled ground truth), or at least performance metrics from previous studies using the same TLS2trees method, so that users get an idea of the potential errors and/or limitations of the approach.	The below paragraph has been added to step 3 in the TLS data processing section: A comparison of the leaf-wood separation between *TLS2trees* and manual labelling showed a Jaccard index of between 54 - 87% across varying tropical sites (Wilkes et al., 2023). A number of TLS leaf-wood separation approaches have been developed, using deep learning, or geometric approaches. Unsurprisingly, they all tend to perform worse for taller trees, higher in the canopy (Arrizza et al., 2024). In *TLS2trees*, the impact of misclassifying (or missing) leaves, is to truncate smaller branches (Wilkes et al., 2023), reducing the contribution to volume (and hence biomass). This tends to have less impact on tall tropical trees, than on smaller more dense crowns of deciduous woodland (Calders et al., 2022).

SPECIFIC COMMENTS

1.1) Individual scan registration into plot-level	Done, the section below has been revised to:

point cloud (TLS) Please make the different processes for a) the VZ-400 (coarse registration via reflective targets) and b) VZ-400i (Auto Registration without targets) clearer. From the text, it currently seems like AR2 is used for both scanners, but I assume this is not the case?	**1. Individual scan registration into plot-level point cloud** This process was carried out using retro-reflective targets positioned between scan locations to facilitate coarse registration for data collected with the RIEGL VZ-400 or in a near-automated manner using the RIEGL VZ-400i's GNSS RTK positioning capabilities in conjunction with the enhanced RIEGL RiSCAN Pro software (versions 2.14–2.17). The integrated Auto Registration 2 (AR2) function employs GNSS RTK data to update the scanner's position and orientation, including in tilt mode, thereby enabling real-time automated coarse registration during scanning **without the use of retro-reflective targets**. Major registration errors are easily detected, typically occurring during pre-processing in RiSCAN Pro when individual scans fail to register (i.e., no coherent solution is found) or are incorrectly positioned, which is visually apparent. In cases where coarse registration/auto-registration fails, unregistered scans can be identified, adjusted, and refined using Multi Station Adjustment 2 (MSA2), **which is also used for final precise registration of data initially coarse-registered using retro-reflective targets**. **The registered plot point cloud is provided in the project's local coordinate system.** Following this workflow, the co-registration of all TLS point clouds achieves sub-centimetre accuracy, as confirmed through post-registration inspection. Wind and occlusion are key sources of uncertainty for the scan registration process, highlighting the necessity of scanning under low or zero wind conditions and capturing both tilt and upright scans at each location.
1.2) Recommendations for aligning and matching datasets - remove "in each case" (seems redundant). - L1364f.: Please fix this sentence, it currently does not make sense. - You could also investigate flight strip differences using overlapping regions (without ground control points), have you considered doing this?	- Done: removed "in each case" - Done: for clarity, the paragraph has been revised to: UAV-LS and ALS datasets are geo-referenced, with positional accuracy determined by IMU and GNSS measurements. These measurements can introduce errors that manifest as height biases between individual flight lines. Although no such discrepancies were observed in our data, a definitive assessment would require a rigorous comparison with ground control points -a step we have not undertaken. These datasets have not been explicitly aligned or matched to one another. Alignment is possible but requires manual identification of control points within each dataset, as noted above, should be undertaken only if necessary for the intended application of the data. - No, as mentioned in the paragraph above, the datasets are shared as they are and users can undertake alignment if their intended use requires it.
3-7) General: - The text in labels and legends is still too small; especially in Figure 3 - Is there a reason why the coordinate grid is not	General: - The text size in Figure 3 has been increased - Yes, to fit the different elements mapped at each site requested by both our reviewers while making all elements clearly visible. Figure 1:

square in Figures 1 and 3 (grid lines in 0.01° vs. 0.005° intervals)? Figure 1: - Maybe adapt the plot numbers/IDs to match the tables later and to better differentiate the 15 experimental 4 ha plots and the 40 ha biodiversity plot from the 10 GuyaFlux plots (see also my comment 20). - UAV-LS coverage is really hard to see. Maybe for ALS and UAV-LS coverage, find another way of visualization (e.g., outlined boxes; solid line and dotted?) This would also allow to better see the background map. - GuyaFlux tower plots do not look "solid green", but rather blue with thick white outline. Please adapt the caption. - Figure 2: Fix in caption "is marked with a yellow square" → "is marked with a white square" - Figure 3: The black text over dark background is very difficult to read, consider changing to white or adding background. The white labels with Plot IDs are way too small to be readable.	- The plot IDs have been further clarified in the revised Tables 2 – 4 which now provide clarification between ForestScan plot ID's (which include 2-3 letters for the country/site/local plot IDs + local plot number) and Census plot and Subplot IDs as now clarified in each table's caption: "We provide both the ForestScan plot IDs and their corresponding census plot and subplot IDs used by the census internet-based data repositories". - Done - Done, the colour has been changed and the caption updated Figure 2: done Figure 3: done, I have made all text and numbers as large as possible, unfortunately, as the reviewers requested for the ALS coverage to be included in this figure, the scale of the plots are very small compared to the ALS coverage.
10) Table 1 - Typo "REIGL VZ-400i" → RIEGL - Suggestion: Start with general scanner characteristics (not changeable), i.e., wavelength, Ranging accuracy/precision, max range, beam divergence, beam diameter, returns per pulse. Then continue	All done, pls note we prefer to keep the Max Pulse Repetition Rate [kHz] as is, these settings are also mentioned in Section 4.

with the user-defined settings, i.e., pulse repetition rate (300 kHz), angular resolution, FOV (please also specify the vertical FOV), and scan time per scan. - Max Pulse Repetition Rate [kHz] → change to "Pulse Repetition Rate" and list only the one you used, i.e., 300 kHz - Inconsistency in row "Angular resolution" - Please change the caption to mention both scanner characteristics and (user-defined) scanner settings	
13) If extra attributes are always the same, please list all of them specifically, instead of the current vague way "such as […], etc.". Consider omitting "XYZ" coordinates. I would consider only data stored on top of the point's spatial location as "attributes". If the attributes differ between the point clouds, please explain how and why.	Thank you for your comment. We refer to *attributes* rather than *extra attributes*, which is why we include XYZ coordinates in the list (we have now explicitly added the word "coordinates" for clarity). We believe retaining the XYZ coordinates is helpful, particularly for beginner users of TLS data. Regarding other attributes, these are not always consistent across point clouds as they are generated by different processing steps. For this reason, we have opted to keep the description general rather than listing all possible variations.
17) No, this comment referred to the point clouds themselves. If they are in a local coordinate system, please state so.	Done, pls see response to 1.1
20) The plot IDs from the tables vs. in the text are confusing. Is FG6c2 the same as "P6".	FG6c2 and P6 are partly the same as explained in the figure's caption (see below). In Paracou, there are 15 experimental 4 ha plots with 4 subplots each containing four 1 ha subplots numbered 1 – 4, the three 1 ha ForestScan FBRSM are subplots FG5c1, FG6c2 and FG8c4 correspond to subplots 1, 2 and 4 in plots 5, 6 and 8, respectively. This is explained in detail in the revised caption for Figure 1 (see below), and in subsections TLS: FBRMS-01: Paracou, French Guiana and UAV-LS: FBRMS-01: Paracou, French Guiana **Figure 1:** Multi-scale map depicting the location and spatial distribution of research plots at Paracou Research Station, French

<table>
<tr><td></td><td>Guiana. (a) Location of French Guiana (green) within South America. (b) Location of Paracou Research Station (green) within French Guiana. (c) Detailed site map showing the spatial distribution of research plots with treatment-specific colours, UAV-LS coverage (yellow solid outline), and ALS coverage (yellow dashed outline). The map displays 15 experimental 4 ha plots, each containing four 1 ha subplots numbered 1 - 4 (60 subplots in total; plots 1 - 12: silvicultural treatments; plots 13 - 15: Biodiversity monitoring), one large 40 ha Biodiversity plot (plot 16; red), and 10 GuyaFlux plots (yellow). Treatment categories include: Biodiversity monitoring plots (plots 13, 14, 15, 16; red), T0 Control (plots 1, 6, 11; green), T1 Selective logging (plots 2, 7, 9; dark blue), T2 Selective logging + thinning by timber stand improvement (TSI; plots 3, 5, 10; cyan), and T3 Selective logging + TSI + fuelwood harvesting/FW (plots 4, 8, 12; pink). The three FBRMS-01 subplots -FG5c1 (subplot 1 of plot 5), FG6c2 (subplot 2 of plot 6), and FG8c4 (subplot 4 of plot 8)- are shown in solid orange and were surveyed using terrestrial laser scanning (TLS) with corresponding tree census data. The GuyaFlux tower location is indicated by a black triangle with radiating transmission waves, and the Base Camp location is marked with a white square. Scale bar: 800 m. Map data: Natural Earth 10 m cultural vectors. Satellite imagery basemap: Imagery ©2024 Google. Map projection: WGS84 (EPSG:4326).

The plot IDs have been further clarified in the revised Tables 2 – 4 which now provide clarification between ForestScan plot ID’s (which include 2-3 letters for the country/site/local plot IDs + local plot number) and Census plot and Subplot IDs as now clarified in each table’s caption: “We provide both the ForestScan plot IDs and their corresponding census plot and subplot IDs used by the census internet-based data repositories”</td></tr>
<tr><td>21) Also in Table 8, please refer to Figure 1 (see above) so the reader understands how to match the plot numbers.</td><td>Done, we have added reference to Table 2 as this references is more appropriate (see response 20)</td></tr>
<tr><td>24) L1240: How was this "geometric accuracy" quantified? Does this refer to georeferencing error (quantified with additional check points??) or flight strip differences or something else? Please define.</td><td rowspan="2">By “geometric accuracy,” we refer to the overall positional accuracy of the LiDAR-derived point cloud after all trajectory corrections and ground control adjustments were applied. This is not limited to georeferencing error or flight strip differences alone but encompasses the cumulative accuracy achieved through the following steps:
1. Reconstruction of flight trajectories using GNSS/IMU measurements and differential corrections in Applanix POSPac.
2. Integration of corrected flight paths and laser data in RiPROCESS.
3. Refinement of relative position and orientation using small buildings as reference features.</td></tr>
<tr><td>26) But also in general: Please make it clearer in which coordinate reference system each</td></tr>
</table>

dataset is provided. I still do not seem to be able to find it. Or is it always local coordinates?	4. Final adjustment using ground control points (checkerboard targets). The resulting geometric accuracy of **1.8 cm** was quantified based on the residuals between the LiDAR point cloud and the surveyed ground control points after all corrections. All elevation data are expressed as ellipsoidal heights within the UTM 32S coordinate system. We added this line to the text for clarification: Geometric accuracy refers to the absolute positional accuracy of the final point cloud after these corrections, quantified by the residuals between LiDAR points and surveyed ground control points.: The CRS is: WGS84 (ESPG:4326), this clarification has been added to the text.
29) If the section is about "Recommendations for aligning and matching datasets", then I do not understand the value of the second section in 3.3 (L1427-1433). Here, you basically mention aspects to consider prior to data acquisition but in the paper, you describe a dataset that has already been acquired.. So in my opinion, this is misplaced here and can be left out, unless it can be addressed via post-processing.	Thank you for your comment and for raising this point. We appreciate the observation; however, we believe that the section in question provides important context and practical guidance that complements the preceding sections on aligning TLS with census data and TLS with UAV-LS data. While the manuscript primarily describes datasets that have already been acquired, the discussion of acquisition parameters and their influence on alignment is highly relevant for readers seeking to understand the limitations and challenges inherent in multi-scale LiDAR integration. These considerations -i.e. point density, scan angle distribution, footprint size, and pulse power- cannot always be fully corrected during post-processing. They often determine the feasibility and accuracy of subsequent alignment steps. Including this information therefore serves two purposes: 1. It clarifies why alignment between TLS, UAV-LS, and ALS datasets is complex and why manual intervention is frequently required. 2. It provides valuable recommendations for future campaigns, ensuring that readers who intend to use or replicate these datasets are aware of factors that influence comparability and bias. For these reasons, we believe the section is appropriately placed and adds meaningful value to the manuscript by bridging acquisition considerations with alignment strategies.
30) My last sentence was about the TLS point clouds/QSMs, not about the scan positions. If I see correctly, they are provided in a local coordinate system. And it would be interesting for users how to transform these into a global	This is all already explained in detail in subsection 3.2 in section 3. Recommendations for aligning and matching datasets: **3.2 Aligning TLS to UAV-LS data (and other spatial data)** Through its accurate global registration via PPK processing, UAV-LS can be regarded as a high-quality geometric reference for registration. For the purpose of comparison with accurate ALS data or satellite observations, a registration of TLS to the UAV-LS point cloud is highly recommended. The integration of GNSS directly into TLS data collection now ensures that registered plot-level point clouds are aligned within a global coordinate system. This

coordinate reference system. This would also be needed if users would want to align the ALS/ULS data to the TLS data, right? If TLS data is already provided in a global coordinate reference system, please state so.	significantly facilitates the co-registration of TLS and UAV-LS point clouds, given that GNSS accuracy is typically within 1 metre. Historically, placing all LiDAR point clouds within accurate global coordinate systems necessitated dedicated survey measurements of plot corners or TLS locations via GNSS, a process often hindered by signal attenuation in dense forests. Consequently, GNSS surveying of plot corner locations is not a standard component of forest census protocols, although it should be considered essential for plots intended for EO calibration and validation purposes. The reduced cost of RTK GNSS equipment and its subsequent routine integration into TLS workflows have made this more feasible, despite the challenges in obtaining fixed positions, and maintaining radio link with a base positioned on a well-known point under deep forest canopy cover. While this may not benefit ALS directly, UAV-LS is likely to serve as a valuable intermediary between TLS (and census data) and ALS. The requirement for global GNSS positioning also extends to other spatial datasets.
Other) - L 781ff.: Can you add the folder or file names (or patterns) for each entry, so that users can find the data more easily? You may additionally hint to the "ForestScan_example_directory_structure.pdf" in the TLS data directories.	All this information is already included in this PDF document, we have added a reference to this document in the text as below: These TLS ForestScan FBRMS 1 ha plot datasets are freely available via the Centre for Environmental Data Analysis (CEDA) with URLs and DOIs provided in section 5, and are accompanied by the **ForestScan_example_directory_structure.pdf** document for guidance on dataset organisation.

TECHNICAL CORRECTIONS

8) Fig. 5: 8.5) The subplots a and b seem redundant (since only the labels differ). Can they be combined into one figure?	Thank you for your suggestion and for taking the time to provide detailed feedback. We appreciate your perspective on combining panels and enlarging the grid; however, after trying to combine them and careful consideration, we believe the current figure layout best serves its intended purpose. Separating panels (a) and (b) allows us to clearly illustrate both the sampling grid and the scan order without overcrowding the visual elements. Combining them would not significantly improve clarity and could potentially reduce readability. Regarding the upright and tilted scan indicators, while these may appear small at this scale, they are included for completeness and consistency with the caption and methodology. They are not essential as the scanning directions/positioning are clearly described in the figure caption. We have ensured that the caption provides sufficient explanation of the grid size, line positions, and scanning approach, which mitigates any potential ambiguity. For these reasons, we prefer to retain the figure in its current form.

14) I did not mean a colour bar, but a scale bar, so that the readers get an idea how tall the tree is. Regarding the colour, my comment holds that the discrete colours (green/brown) are difficult to distinguish (due to the shading) and I would still ask you to adapt the style.	Thank you for your comment and for sharing your perspective on improving the figure. We appreciate the suggestion; however, we believe the current figure effectively conveys the intended information. The height and canopy width of the tree are already clearly stated in the caption (~40 m tall and ~50 m wide), which provides readers with the necessary scale context. Adding a scale bar would therefore not enhance understanding. Regarding the colour scheme, changing this would require additional processing and would not materially enhance the scientific interpretation, as this colouring is primarily illustrative and does not affect interpretation of the data, as the classification (wood vs. leaf points) is clearly described in the caption. For these reasons, we prefer to retain the figure in its current form.
19) You clearly use more than 5 discrete colours in the DSM, so technically, a continuous colour scale (of course with discrete labels) would be correct here. Please fix.	This figure was removed from the manuscript in the previous revision round but we forgot to update our response.
22) It is a bit difficult to understand from the revised section that a reduction of available power leads to a lower range and this is why with higher PRR, the tree tops of tall trees would not be covered. Can you make this clearer?	These lines have been revised to: TLS data were collected using a pulse repetition rate (PRR) of 300 kHz on RIEGL VZ-400 and VZ-400i scanners, trading longer scan times for a fixed angular resolution to maximise coverage at the tops of tall trees. In the RIEGL configuration, PRR and emitted laser power are intrinsically linked: increasing the PRR reduces the available laser power, which in turn decreases the maximum range of the scanner. At very high PRR settings, this reduction in range means that the tops of tall trees may not be captured effectively. Therefore, selecting a lower PRR (300 kHz) ensures sufficient power and range to cover the full canopy height of forests, while maintaining the desired angular resolution.
25) There is one more occurrence of "section 5. Data access" in L1194.	Done
Other) - L833: "ope-source" → "open-source" - L836: Use "Open3D" (the official name and not the abbreviation) - L840: "deformations" and "warpRefMesh functions" is unclear. Please explain. - Harmonize notations for	- L833: done - L836: done - L840: done, this text was revised to: The surface may also be used as a mesh for visualising 3D deformations, which refer to changes or displacements in the geometry of the object compared to a reference state. This is achieved using the warpRefMesh function. - done, unit conations harmonised

units, e.g., m/ s^{-1} vs. pts/m^2 etc.	