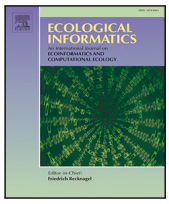
Ecological Informatics 91 (2025) 103398

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Enhancing forest inventory via a videogrammetry approach for robust 3D reconstruction: A study using Insta 360 Pro 2

Hristina Hristova [a,*], Clemens Blattert [a], Bogdan Candrea [c], Mihai Nita [b,c], Sergiu Florea [b,c], Sunni Kanta Prasad Kushwaha [a], Janine Schweier [a]

[a] Sustainable Forestry, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, Birmensdorf, 8903, Switzerland
[b] Faculty of Silviculture and Forest Engineering, Transilvania University of Brasov, Sirul Beethoven 1, Brasov, Ro-500123, Romania
[c] Forest Design SRL, Nicovalei 33, Brasov, Ro-500473, Romania

ARTICLE INFO

Dataset link: https://doi.org/10.5281/zenodo.16258209

Keywords:
3D reconstruction
Digital sensor
Forestry
Insta 360 Pro 2
Remote sensing
Videogrammetry

ABSTRACT

In this study, we explore the application of videogrammetry for 3D reconstruction in complex forest environments, aiming to enhance forest inventory measurement methods. Traditional techniques are often labor-intensive and lack scalability in dense or challenging terrain. We assess the efficacy of videogrammetry for generating 3D point clouds in complex forest environments, focusing on an Insta 360 Pro 2 setup with six fish-eye cameras. Harnessing this lightweight and user-friendly technology, we aim to elevate the process of data collection while delivering realistic visual representations of forest areas. Our approach enables the estimation of key forest characteristics, such as tree distribution and Diameter at Breast Height (DBH). The average errors for tree position and DBH measurements range from 5.2 cm to 18.8 cm and from 0.9 cm to 1.9 cm, respectively. The reconstructed 3D tree information is structurally similar to data obtained with Terrestrial Laser Scanning (TLS), with normally distributed Multiscale Model-to-Model Cloud Comparison (M3C2) errors with a mean of 0 cm and a standard deviation of 15 cm to 22 cm. Our method reduces the need for manual data collection, thus supporting effective forest management and planning.

1. Introduction

Assessing a forest's role in mitigating climate change, optimizing protection against natural hazards, and supporting sustainable biodiversity management require extensive knowledge about the forest on a large scale (Ferretti et al., 2024; Breidenbach et al., 2020). However, traditional inventory methods, which involve manually measuring trees, are labor-intensive, do not refer to large forest areas, and might be less accurate than data provided through modern equipment (Mattamala et al., 2024; Apostol et al., 2007). To enable scalable forest inventories, we need a reliable, practical, and replicable methodology to estimate key forest features with minimal manual effort and optimized costs (Malladi et al., 2024; Hyyppä et al., 2020; Liang et al., 2016a).

Recent advancements in remote sensing technology, like Light Detection and Ranging (LiDAR), have made it possible to accurately assess fundamental single tree and forest stand features from Terrestrial Laser Scanning (TLS) point clouds, like DBH and tree position (Hristova et al., 2024; Kükenbrink et al., 2022; Mokroš et al., 2018), tree species (Puliti et al., 2024), biomass (Su et al., 2025; Bornand et al., 2023), and carbon sequestration (Fasihi et al., 2024). TLS is considered the gold standard for generating point clouds in forestry (Calders et al., 2020; Liang et al., 2016b). However, the heavy weight of the equipment may make it impractical for certain forest conditions, such as dense understory and steep slopes. Additionally, using TLS requires expert knowledge and is costly compared with other ground-based survey methods (Fassnacht et al., 2024; Griebel et al., 2015). A more cost-effective and lighter alternative to TLS is Mobile Laser Scanning (MLS), which addresses some of the challenges of using TLS. However, similar to TLS, MLS requires expert knowledge on data acquisition and processing, and it may produce noisier point clouds than TLS (Cabo et al., 2018).

In recent years, there has been a surge in the popularity of cameras and other passive digital sensors for point cloud generation and 3D reconstruction (Pepe et al., 2022; Murtiyoso and Grussenmeyer, 2021; Sun and Zhang, 2019). Mass-produced cameras, such as those found in mobile phones, as well as low-resolution spherical cameras, are favored for their availability, ease of use, lightweight design, and low costs compared with TLS and MLS (Zhu et al., 2021; Bienert et al., 2018). Additionally, videogrammetry, a method in which 3D models are created from digital camera videos, has emerged as an efficient

* Corresponding author.
E-mail address: hristina.hristova@wsl.ch (H. Hristova).

https://doi.org/10.1016/j.ecoinf.2025.103398
Received 6 May 2025; Received in revised form 13 August 2025; Accepted 16 August 2025
Available online 8 September 2025

approach for point cloud generation (Pepe et al., 2022; Alsadik et al., 2015; Kwiatek and Tokarczyk, 2014). With this approach, it is possible to combine 3D structures from point clouds with Red Green Blue (RGB) information from photographs, providing additional information for data analysis. While low-cost videogrammetry has primarily been used in the context of cultural heritage, its potential application to the forest environment remains largely unexplored.

In this study, we address the effectiveness of videogrammetry in various forest environments characterized by high levels of structure and heterogeneity. These conditions may present challenges in accurately identifying and aligning key points in images. Our goal is to enable the creation of a lifelike 3D representation of the forest for immersive virtual environment experiences. These representations may also enable the estimation of forest stand features, such as the occurrence, density, and spatial distribution of trees, as well as the dimensions of tree diameters and trunks. These features provide crucial information for managing and planning forest ecosystems. They serve as the basis for assessing biodiversity (e.g., by considering variations in structure, deadwood, and habitat trees) and for understanding essential ecosystem services like timber production (e.g., based on diameter, stem length, and shape), carbon sequestration (e.g., by assessing changes in diameter and yield between two data collections), and recreation (e.g., based on tree density per hectare) (Knoke et al., 2021; Oettel and Lapin, 2021; Blattert et al., 2017). Furthermore, this extensive data, including tree coordinates, can guide single-tree-based harvesting interventions. This approach applies to all silvicultural practices, but is particularly relevant for continuous cover forestry, where felling and tree extraction are based on selective harvesting (Mason et al., 2022; Larsen et al., 2022). Identifying trees designated for removal is crucial to prevent damage to the soil and the remaining trees. Providing digital information in advance about the trees targeted for removal is likely to streamline the overall workflow. A central focus of this study is the acquisition of structural forest information while minimizing data collection efforts and also gaining timely knowledge of the forest environment to support management decisions and optimize the impact of ecosystem management.

We propose a videogrammetry approach to reconstruct 3D point clouds in forest environments, leveraging an Insta 360 Pro 2 acquisition setup. In a recent study, Bruno et al. (2024) assessed the performance of this spherical camera in indoor spaces. In contrast to indoor environments, forest areas are open and highly unstructured, marked by significant occlusion. The current camera setup has primarily been evaluated in narrow spaces in the presence of Ground Control Points (GCP), where stereo photogrammetry principles support its strong performance (Hasegawa et al., 2000). However, reconstructions in forest environments require covering much greater distances and demand larger baseline distances between camera lenses when using static images (Dai et al., 2014). To tackle these challenges with a single-body camera, we utilized videos captured while walking through a forest plot. The complexity of forest scenes, combined with the absence of GCP, further complicates reconstruction efforts. To our knowledge, this specific scenario has not yet been investigated using the Insta 360 Pro 2 setup, and the camera's potential applications in forestry have not been explored.

Compared with more cumbersome alternatives (Hristova et al., 2024; Forsman et al., 2016), the Insta 360 Pro 2 is a single-camera setup that simplifies and enhances data collection in various forest conditions. A stereo approach for forest acquisition recently introduced by Hristova et al. (2024) employs two Ricoh Theta Z1 cameras placed at a certain distance from each other. Despite its potential efficiency, such a stereo system with a distance between the cameras of more than 60 cm could be less mobile and more challenging to use in dense forest conditions (Hristova et al., 2024). In contrast, the Insta 360 Pro 2 offers a single-camera solution that is user-friendly, enabling easy handling, transportation, and use.

The key novelties and contributions of this work are fivefold:

- The proposed Insta 360 Pro 2 method is applicable to various forest study areas differing in forest stand characteristics and tree species.
- We demonstrate that the 3D reconstruction of forest plots based on multi-lens videos is feasible.
- Our approach enables the generation of scaled reconstructions directly from the captured videos, eliminating the need for GCP and visually coded targets.
- Our method facilitates the successful estimation of tree positions and DBH from the generated point clouds with relatively small errors.
- Using the compact Insta 360 Pro 2 equipment helps to minimize the amount of data acquisition time needed in the field, enhancing the overall efficiency of forest inventories and reducing manual work.

2. Materials and methods

2.1. Equipment

We captured videos using a spherical camera setup. Our setup includes an Insta 360 Pro 2 camera, which is equipped with six fish-eye lenses aligned horizontally, each providing a 200-degree Field of View (FoV). Each lens is capable of capturing 4K videos. When the content from each lens is stitched together, an 8K video is produced. This setup captures the entire 360-degree FoV of the forest in a single video. The Insta 360 Pro 2 camera was mounted on a Monopole backpack, as shown in Fig. 1(b). This mounting method served three purposes: (1) it kept the camera at a consistent height during data collection, (2) it reduced motion blur in captured videos because it acted as a stabilizer, and (3) it made data acquisition faster and more efficient.

2.2. Study area

We conducted the study in 2024 in the Teliu forest in Romania, managed by a Common Ownership Association comprising approximately 650 forest owners. This association oversees a total area of around 1200 hectares, subdivided into various parcels according to the management plan. The stands feature a wide range of diameters (DBH ranging from 2 cm to 75 cm) and include a variety of tree species, such as beech, spruce, pine, oak, larch, hornbeam, and sessile oak. The forest area is sampled using 85 plots, and TLS scans have been performed across these plots.

We visited six of these forest plots and recorded videos using our Insta 360 Pro 2 equipment. The selected plots displayed a variety of tree species, ages, and management practices and reflected the overall diversity of the area. Plot 1 featured mature beech trees, while Plot 2 included a combination of mature beech and conifer trees. Plot 3 contained a mix of sessile oak and beech trees, and Plot 4 featured a mixture of young spruce, beech, and hornbeam trees. The dominant tree species in Plot 5 was beech, and in Plot 6 it was pine. Snippets of the forest plots and equipment are shown in Fig. 1.

Table 1 illustrates the variability in forest stand characteristics across the six plots in our study, focusing on tree species composition, DBH, and tree height distributions. Plots 1 and 2 contained the largest trees in terms of both DBH and height, while Plot 4 had the highest number of young trees. Additionally, Plot 4 was the densest among all the plots, followed by Plots 5 and 6. Furthermore, Plots 5 and 6 exhibited a diverse composition of tree species. In all plots, except Plots 1 and 2, there was an understory layer present, which increased the likelihood of occlusion.

(a) (b) (c) (d)

Fig. 1. Impressions of the data collection and the study area in the forest region of Teliu, Romania (image source: Mihai Nita/Forest Design SRL). Images (a), (c), and (d) represent the different forest conditions observed during data collection: a mature stand, a young stand, and a mixed stand, respectively. Image (b) illustrates our camera setup.

Table 1
Tree species composition, Diameter at Breast Height (DBH) and height distributions (minimum, mean, maximum, and standard deviation values), and number of trees per hectare for each plot in our study.

		DBH (cm)				Height (m)				
	Composition	Min	Mean	Max	St. Dev.	Min	Mean	Max	St. Dev.	Number of trees per hectare
Plot 1	100 Fagus sylvatica	20.5	42.4	60.3	10.3	31.3	35.3	38.7	1.7	310
Plot 2	58 Fagus sylvatica 42 Picea abies	5.8	36.4	74.2	14.7	6.8	35.4	42.9	5.7	550
Plot 3	56 Fagus sylvatica 42 Quercus petraea	1.8	17.3	35.9	8.8	2.4	17	23	5.7	890
Plot 4	66 Picea abies 32 Fagus sylvatica 1 Pinus sylvestris 1 Alnus incana	2.1	11.8	35.7	6.5	2.6	13.8	22.8	5.7	4,530
Plot 5	57 Carpinus betulus 15 Fagus sylvatica 11 Pinus sylvestris 9 Quercus petraea 6 Pinus nigra 1 Acer campestre 1 Prunus avium	2.2	15.9	65.7	16.1	3.3	13.6	29.4	6.6	1,060
Plot 6	87 Carpinus betulus 4 Quercus petraea 4 Picea abies 2 Fagus sylvatica 2 Acer campestre 1 Prunus avium	2.4	14.1	32.1	4.7	3.7	16.2	19.7	3.5	1,170

2.3. Reference mobile TLS acquisition

For this research, we systematically collected high-resolution terrestrial LiDAR data from across the study area in 2024, using the GeoSLAM Horizon scanner. The GeoSLAM Horizon is a hand-held mobile TLS that emits laser pulses and measures the time taken for these pulses to return after striking objects in the environment. Its rotating mirror provides a 360-degree FoV by emitting laser pulses in all directions and capturing the reflected signals with integrated sensors. As the scanner moves through the environment, it continuously gathers data and generates a real-time 3D map using Simultaneous Localization and Mapping (SLAM) technology. The GeoSlam Horizon scanner is designed for outdoor scanning at distances up to 100 m with an accuracy of 1–3 cm. The resulting mobile TLS dataset achieved a point density of approximately 1500 points per square meter, providing detailed spatial information about the scanned environment.

2.4. Methodology

2.4.1. Multi-rig setup with constraints

Our approach involved creating point clouds from the data captured by the six fish-eye lenses of the Insta 360 Pro 2 camera (Fig. 2). Similar to Bruno et al. (2024), we designated one camera as the master and the other five as slaves. In our approach, we constrained the multi-rig system in the following way. First, we set the (x, y, z) baseline distances between the master and slave cameras, as per estimations in Bruno et al. (2024). The baseline distances were computed via a multi-rig camera calibration. Second, we considered the rotational angle between the master camera and each slave camera. Each camera is positioned at a 60-degree rotation around the z-axis of its neighboring camera. Both the (x, y, z) distances and the rotational angles were specified in our point cloud generation pipeline.

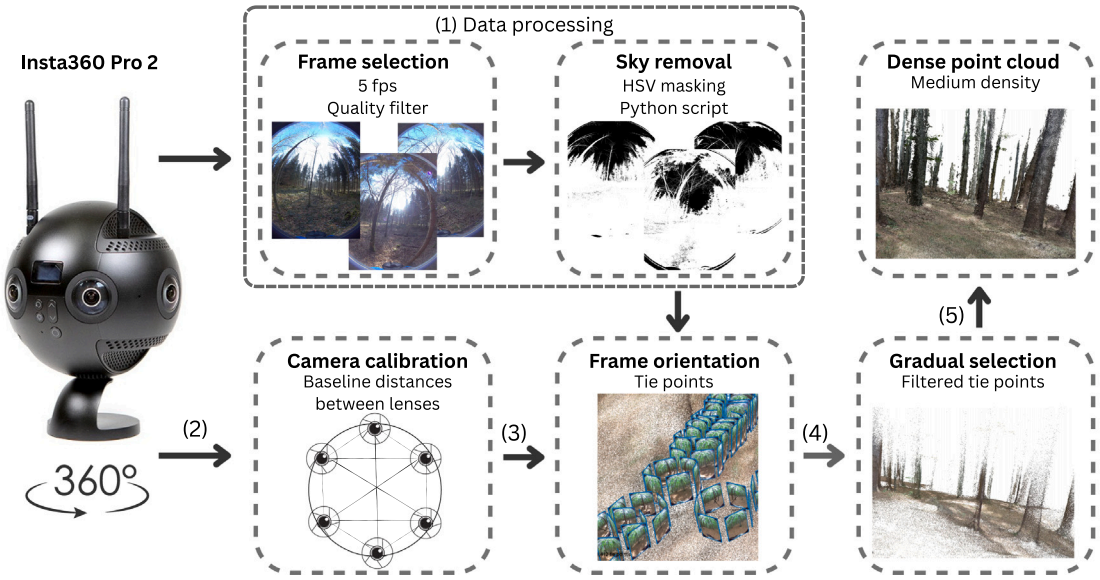


Fig. 2. Methodology pipeline. (1) We process the captured videos as a set of 5 frames per second (fps) synchronized video frames, for which we mask out the sky. (2) We calibrate the multi-rig system to find the proper distances between the cameras. (3) Given the multi-rig constraints, we proceed with the orientation of the video frames, before (4) filtering the computed tie points using gradual selection. Finally, (5) we generate the dense point cloud.

2.4.2. Video frame selection

We generated point clouds using content from the six fish-eye cameras by extracting a number of video frames. Each video frame represented a camera position at a specific time. The number of video frames we used was determined by the frame rate at which we extracted the frames. The frame rate affects (1) the distance between each pair of consecutive camera positions and (2) the level of detail in the resulting point cloud. A higher frame rate leads to more detail, allowing more features to be matched between images and ensuring successful camera orientation. However, a higher frame rate also means smaller distances between consecutive camera positions, which can result in larger errors in depth estimation, according to photogrammetric principles (Hasegawa et al., 2000).

We conducted tests on point cloud generation using three different frame rates: 1, 5, and 10 fps. At 1 fps, the level of detail in the point cloud was insufficient and required improvement, as illustrated in Fig. 3. The optimal frame rate for capturing detail and ensuring accurate camera orientation can vary based on factors such as forest type, density, number of trees, and the complexity and occlusion of the environment. Our findings indicated that utilizing a frame rate of only 1 fps could lead to frame misalignment during point cloud generation.

To address this issue, we opted for a frame rate of 5 fps, which considerably enhanced the quality of the point cloud. However, it is essential to note that increasing the frame rate also elevated the computation time required for generating the point cloud. For example, using a frame rate of 10 fps increased computation time by 2 to 3 times compared with 5 fps. Additionally, we observed that 10 fps often led to incorrect reconstructions, where a tree stem appeared multiple times within a single point cloud, as depicted in Fig. 4. This was caused by increased errors in depth estimation as a result of decreased distances between two consecutive video frames when 10 fps was used. Through empirical evaluation, we determined that 5 fps offers results with a good balance between computational efficiency and point cloud quality.

In the forest, the uneven terrain and occasional poor lighting can sometimes cause image blur. To prevent any incorrect image feature matching caused by blur, we filtered out low-quality video frames. As demonstrated by Sun and Zhang (2019), removing blurry images improves the point cloud quality. After selecting the video frame rate, we used the Image Quality Estimation feature in the Agisoft Metashape software (AgiSoft, 2018) to assess image quality, and we filtered out frames with a quality rating below 0.75. This resulted in the removal of 5% to approximately 40% of the video frames, depending on the plot.

2.4.3. Masking

To enhance the quality of the point cloud and minimize noise, we developed a Python script that filtered out the sky in each video frame by generating image masks. We created the masks by isolating all pixels in the video frames with Hue Saturation Values (HSVs) falling within a specific range corresponding to the color blue. Furthermore, we eliminated pixels with the highest luminance (channel V in HSV), ranging from 200 to 255. This removed the brightest regions in the frames, belonging to the sky. The mask was generated for each fish-eye video, specifically for the upper half, as the sky was only visible in that portion. During our data acquisition, the weather was clear and sunny. However, the mask computation can be adjusted for different weather conditions by altering the color range during pixel filtering. Although the masking algorithm is simple, a visual inspection of the generated masks for all videos in our study revealed the efficiency of the filtering procedure.

2.4.4. Camera orientation

The first step in generating a point cloud from video frames using photogrammetry involves orienting the video frames according to the different camera positions. An important part of this process is camera calibration. We generated initial estimates of the camera parameters through an auto-calibration step in Agisoft Metashape. The camera parameters k_4, p_1, and p_2 were equal to 0, as per findings by Bruno et al. (2024).

Following the initialization of camera calibration, we proceeded with the camera orientation step using the “medium” settings in Agisoft Metashape, downsampling the original video resolution by a factor of 16. The calibration parameters remained adjustable during the camera orientation process, enabling the bundle adjustment algorithm to derive improved calibration estimates. For the feature extraction, we used 20,000 key points and 4000 tie points. We used the “generic” preselection option for computational efficiency, and we applied adaptive camera settings.

(a) 1 fps (b) 5 fps

Fig. 3. Quality comparison between a point cloud generated using our videogrammetric approach with (a) 1 frame per second (fps) and (b) 5 fps. The snippets shown in (a) and (b) were captured from the same viewpoint. Except for the difference in the frame rates, the two point clouds were generated with the same settings. A larger number of video frames (5 fps) results in a higher level of detail in the point cloud.

(a) 5 fps (b) 10 fps

Fig. 4. Quality comparison between a point cloud generated with (a) 5 frames per second (fps) and (b) 10 fps. The snippets shown in (a) and (b) were captured from the same viewpoint. Except for the difference in the frame rates, the two point clouds were generated with the same settings. Using 10 fps resulted in an incorrect stem reconstruction, making the same tree appear several times in the point cloud.

After completing the frame orientation, we used the gradual selection tool in Agisoft Metashape to filter out tie points. First, we removed tie points with a high reconstruction uncertainty, i.e., above 100. This helped us to filter out noise in the point cloud caused by distant objects without losing too many points. Then, we removed tie points with a reconstruction error above 0.8 and optimized the camera orientation by rerunning the bundle adjustment algorithm. We repeated this process three times to address any poor localization of the point projections. Finally, we filtered out all tie points with a projection accuracy greater than 10 and ran the bundle adjustment algorithm to optimize camera locations. The three threshold values used in the gradual selection tool were carefully chosen based on (1) standard practices to filter out noise without removing too many tie points and (2) empirical testing.

2.4.5. Point cloud generation

Upon successful orientation, we generated a dense point cloud with "medium" quality and "moderate" filtering parameters. The quality parameter "medium" denotes multi-view dense matching on sub-sampled pixels of the original image, representing 1/16 of the original pixel resolution. The filtering parameters pertain to the specific depth map filtering strategies employed by the software. Snippets of the point clouds generated using our pipeline[1] are visualized in Fig. 5. All the Insta 360 point clouds were computed on a MacBook with a dedicated Radeon Pro 580 GPU with 16 GB of memory.

2.5. Point cloud alignment

We used TLS-derived point clouds as reference data for our evaluation. To align the videogrammetric point clouds with the TLS point clouds, we initially performed a rough manual alignment of the two types of point clouds. This was followed by an automatic fine registration using the Iterative Closest Point (ICP) algorithm of CloudCompare (CloudCompare, 2023).

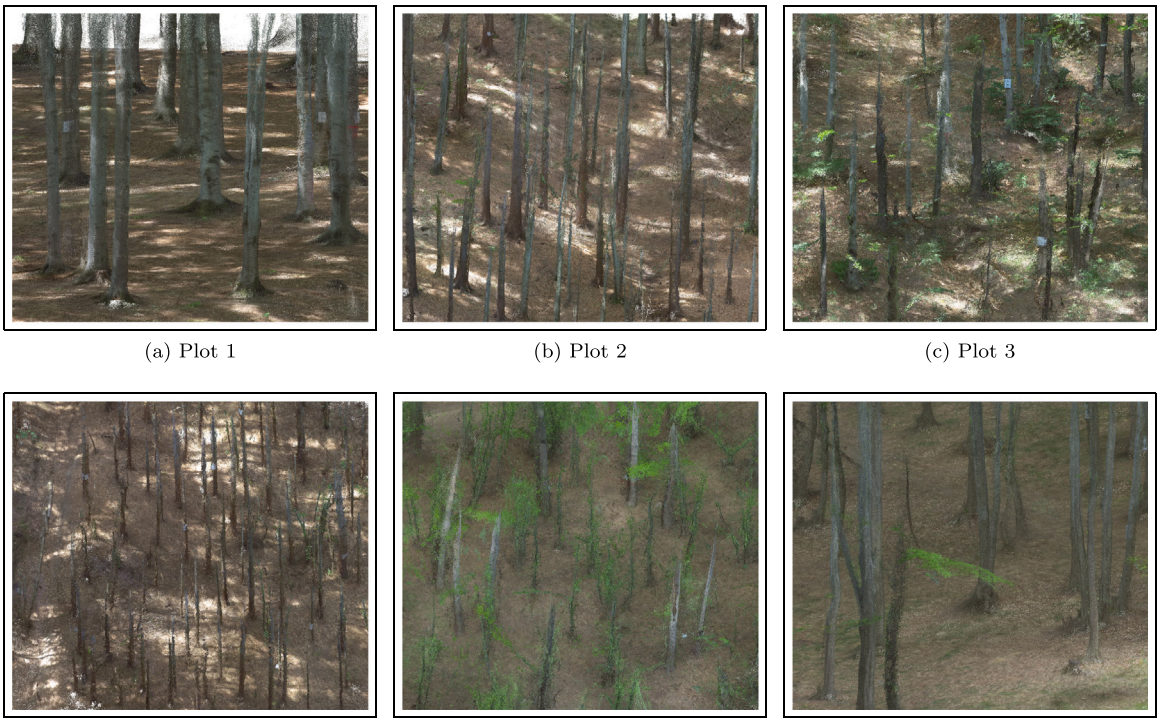
We excluded trees located at the very edge of the reconstructed forest area to eliminate those that were highly incomplete and not included in the acquisition path, i.e., we intentionally did not capture data for these trees. These edge trees were only partially reconstructed, with less than 20% of the tree stem generated, and were therefore considered unreliable for evaluation.

Within each plot, there were trees that were not completely reconstructed due to the acquisition path and occlusion. Unlike the edge trees, all trees completely within the plots were included in our evaluation, regardless of their level of reconstruction.

2.5.1. Point cloud scale

Our method generates point clouds with a specific scale. The scale is solely a result of the multi-rig constraint in our pipeline, and it does not depend on any external factors. We first analyzed how well this

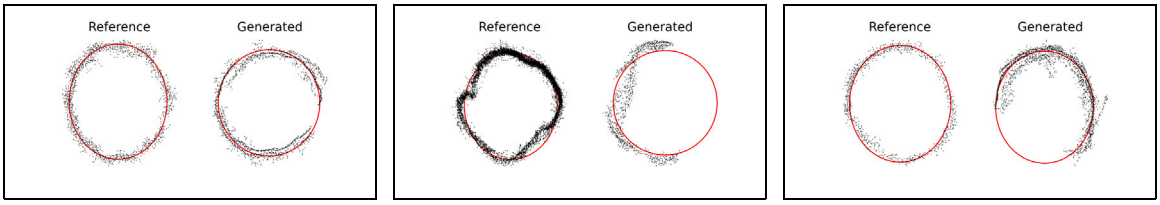
[1] The generated Insta 360 Pro 2 point clouds are available here: https://doi.org/10.5281/zenodo.16258209.



(a) Plot 1 (b) Plot 2 (c) Plot 3

(d) Plot 4 (e) Plot 5 (f) Plot 6

Fig. 5. Snippets of the point clouds generated with our method for each forest plot in our study.



(a) (b) (c)

Fig. 6. Close-ups of cluster cross-sections at 1.3 m height along the tree stem in the generated point clouds and the corresponding reference point clouds. The two types of cross-sections are visualized using the same scale. The red circle illustrates the approximated diameter. Example (a) illustrates a nearly complete tree stem, while example (b) showcases a more natural shape that deviates from a perfect circle. Examples (b) and (c) both depict partially incomplete diameters, with example (c) containing more noise.

scale represents the real-world distances within the videogrammetric point clouds. To achieve this, we distributed 25 pairs of visually coded targets around each forest plot during data collection. These targets were automatically detected in the video frames and resulting point clouds. The distance between each pair of visually coded targets was known to be 13.3 cm. Such an approach was used by Murtiyoso et al. (2022) and Kükenbrink et al. (2022) to scale photogrammetric point clouds in the forest. In our study, we used it exclusively for evaluation purposes.

Once a target pair was detected, we created a scale bar reflecting the distance between the targets, which we used as control points. We then computed the error by comparing the real-world distance between each pair of targets with the estimated distance derived from the point cloud.

2.6. Tree position estimation

We estimated tree positions from each videogrammetric point cloud and compared them with the corresponding reference tree positions from the TLS point clouds. To accomplish this, we first normalized both types of point clouds using the ground information obtained from the TLS point cloud to remove the influence of the terrain. Then, we applied the algorithm presented by Hristova et al. (2024) to the normalized point clouds. This algorithm efficiently clusters the tree stems. To make it more robust to noise, we added two more steps by computing the SPHERICITY feature of the point clouds using CloudCompare (CloudCompare, 2023) before applying the Connected Components (CC) algorithm for clustering. We removed all points with a SPHERICITY value less than 0.08, which helped us to filter out excess noise around the tree stems. After the stem clustering, we generated a cross-section at 1.3 m height for each cluster. We consider the center of each cluster as the tree position. Examples of cluster cross-sections at 1.3 m height for both the videogrammetric (Insta 360 Pro 2) and the corresponding TLS point clouds are shown in Fig. 6. These examples highlight different degrees of stem completion and variation in shape.

2.7. DBH estimation

Next, we evaluated the quality of the videogrammetric point cloud by comparing the estimated DBH from the point cloud with a ground-truth DBH estimated from the TLS point cloud. We extracted the DBH from both the generated point cloud and the TLS-derived point cloud corresponding to a tree stem, applying the modified Random Sample Consensus (RANSAC) method proposed in Hristova et al. (2024).

Table 2
Acquisition and processing times for our approach, in hh:mm format, the number of frames and tie points, and the scale error in mm. Point cloud generation was completed with 5 frames per second (fps).

	Number of frames	Number of tie points	Acquisition	Frame orientation	Point cloud generation	Scale
Plot 1	1886	5 278 259	00:08	10:11	16:20	4.5
Plot 2	1165	2 914 922	00:07	04:00	06:59	4
Plot 3	1846	1 105 081	00:06	08:46	05:04	4
Plot 4	651	1 076 615	00:04	01:32	01:58	5
Plot 5	1816	2 286 486	00:06	08:30	06:18	5
Plot 6	1527	2 104 681	00:06	06:30	07:33	5

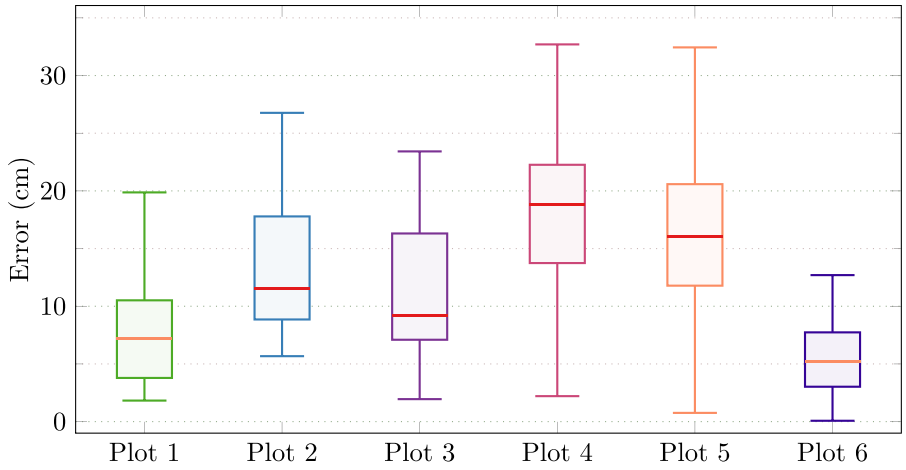


Fig. 7. Distribution of tree position errors estimated from the videogrammetric point cloud and the respective Terrestrial Laser Scanning (TLS) point cloud for each forest plot in our study. The average position error per forest plot is indicated as a horizontal line within each box plot.

2.8. Multiscale model-to-model cloud comparison

We used the M3C2 metric (Lague et al., 2013) to measure the structural distance between point clouds by identifying local distance variations at reference core points extracted from the TLS data. For each core point, we calculated the M3C2 metric based on its distance to neighboring points in the videogrammetric point clouds. We excluded core points that had no neighboring points from our analysis. We calculated the M3C2 distance using projection and normal scale diameters of 0.5954 and 0.1488, respectively. We selected these parameters based on empirical observations, starting with an initial guess from the M3C2 algorithm. This was followed by a refinement stage where we aimed to achieve a balance between estimating broader-scale features and capturing finer details of the tree point cloud.

3. Results

We assessed the performance of our approach by analyzing the scale and the geometry of the videogrammetric point clouds, as well as the potential for extracting realistic tree position and DBH information.

3.1. Computational time and scale

Table 2 summarizes the acquisition and processing times for each forest plot in our study, along with the number of frames retained after low-quality frame removal and the number of tie points generated after applying gradual filtering to the sparse point cloud. The acquisition time, which reflects the duration spent in the field, ranged from 4 to 8 min. Computation time is split into two parts: frame orientation and point cloud generation. The shortest computation time, with the fewest frames and tie points, was achieved for Plot 4. Conversely, Plot 1 exhibited the longest computation time, as it had the largest number of frames and tie points. The total computation time per plot varied between 3 h and 30 min and 26 h and 30 min.

The last column in Table 2 showcases the scale error. The scale error represents the degree to which the distance between a pair of visually coded targets, as computed on the point cloud, deviates from the actual distance between these targets. The error was calculated using control scale bars in Agisoft Metashape software, and it was consistent over all six plots.

3.2. Tree position estimation

Fig. 7 shows the horizontal tree position errors between the videogrammetric point cloud and the reference TLS data. We estimated the tree positions as presented in Section 2.6. The average tree position errors ranged from 5.2 cm to 18.8 cm, with a percentile range spanning from 12.6 cm to 31.7 cm. Fig. 8 shows snippets of the cross-sections of the tree stems for the generated point cloud and the TLS point cloud.

3.3. DBH estimation

Fig. 9 shows the distribution of errors between the DBH calculated from the videogrammetric point cloud and the DBH obtained from the TLS point cloud. The errors are presented both in absolute terms and relative to the reference DBH. The average absolute DBH error ranged from 0.99 cm to 1.89 cm, while the average relative DBH error ranged from 3.15% to 9.33% (reported in Table 3). Moreover, the percentage of tree stems with a relative DBH error of less than 10% ranged from 55.55% to 91.89%, as shown in Table 3. The lowest percentage of trees exceeding a DBH error of 2 cm was recorded for Plot 4 (4.76%), whereas the highest was recorded for Plot 1 (37.8%).

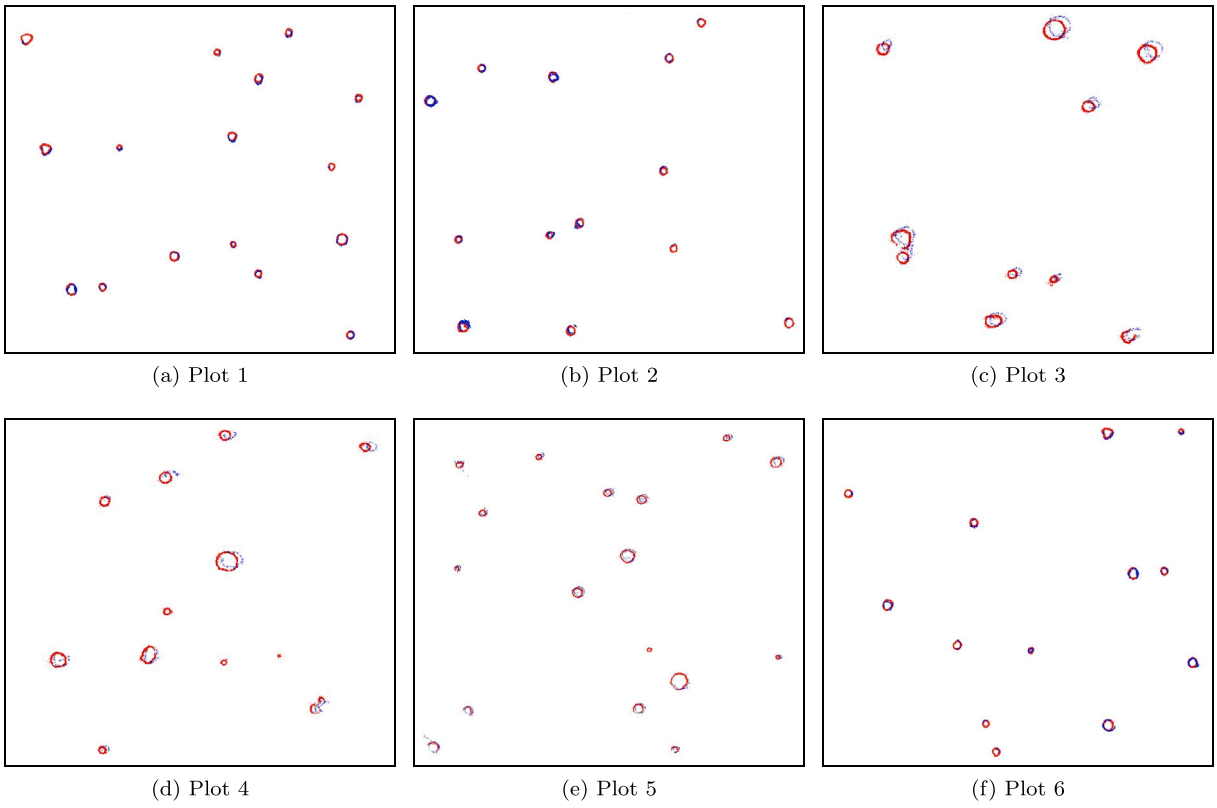


Fig. 8. Snippets from the six forest plots of cross-sections of tree stems obtained from videogrammetric point clouds (in red) and Terrestrial Laser Scanning (TLS) point clouds (in blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

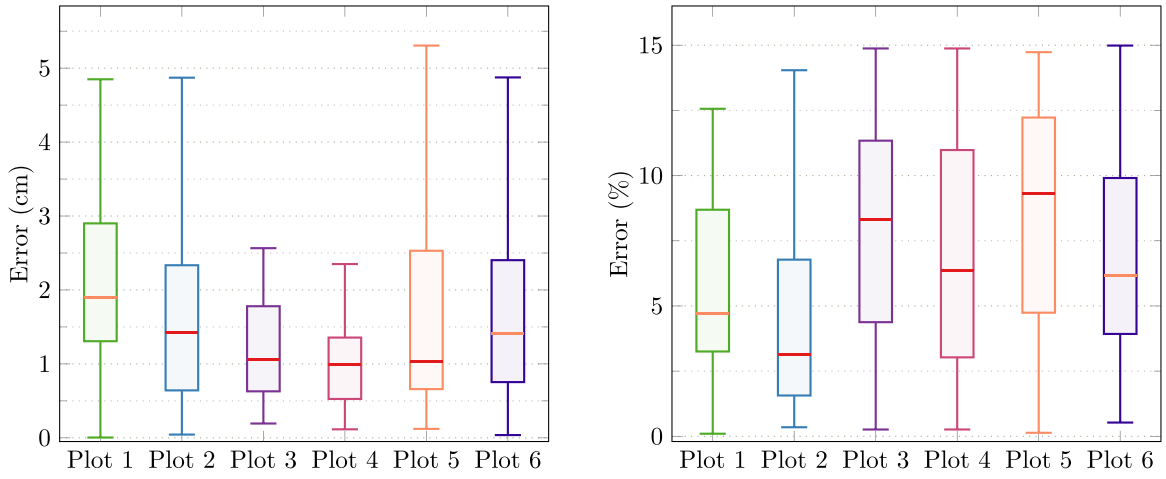


Fig. 9. Distributions of (a) absolute Diameter at Breast Height (DBH) errors in (cm), and (b) relative DBH errors for each forest plot in our study area. The average DBH error per forest plot is indicated as a horizontal line within each box plot.

3.4. Multiscale model-to-model cloud comparison

Figs. 10 and 11 display a subset of M3C2 point clouds computed for single trees from our videogrammetric (Insta 360 Pro 2) point clouds. Well-reconstructed points are shown in blue, representing the majority of the points, as illustrated in the histograms (b) and (d) in Fig. 10. Points differing substantially from their counterparts in the TLS point cloud are highlighted in orange. White indicates points from the TLS point cloud that are not reconstructed in the Insta 360 Pro 2 point clouds. The M3C2 errors are normally distributed, with a mean of approximately 0 cm and a standard deviation of 15 cm to 22 cm.

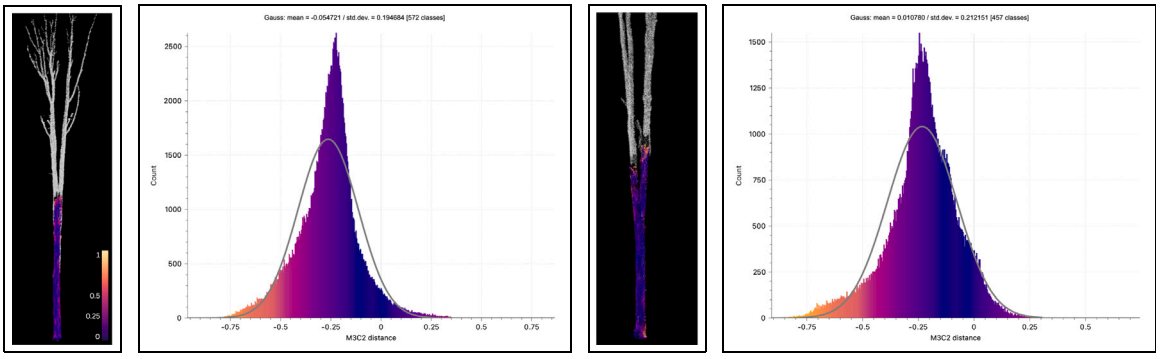
4. Discussion

4.1. Reconstruction quality

The generated point clouds can be used to visualize the forest from a ground-level perspective while minimizing visual noise (Fig. 5). As indicated by the M3C2 metric in Fig. 10, the reconstruction of most tree stems is close to the structure from the TLS point cloud. Larger errors are observed for tree branches and leaves. Moreover, the point clouds sometimes show incomplete tree stems due to insufficient data captured for the tree stem, often caused by occlusions or acquisition path issues.

Table 3
Average error values for tree position, absolute Diameter at Breast Height (DBH), and relative DBH; percentage of trees with a DBH error of less than 10% and percentage of trees with an absolute DBH error greater than 2 cm.

	Average value			Percentage of trees	
	Tree position	DBH (absolute)	DBH (relative)	DBH (less than 10%)	DBH (greater than 2 cm)
Plot 1	7.19	1.89	4.70%	91.89%	37.80%
Plot 2	11.54	1.42	3.15%	83.08%	35.38%
Plot 3	9.23	1.05	8.32%	55.55%	29.63%
Plot 4	18.79	0.99	6.36%	83.33%	4.76%
Plot 5	16.08	1.03	9.33%	61.76%	23.53%
Plot 6	5.25	1.40	6.16%	73.33%	5.33%



(a) M3C2 (b) M3C2 histogram (c) M3C2 (d) M3C2 histogram

Fig. 10. Examples of the Multiscale Model-to-Model Cloud Comparison (M3C2) metric computed for two single trees, visualized in false colors. The color range used in the M3C2 point clouds in (a) and (b) is depicted in the M3C2 histograms in (b) and (d), and shown as a color bar in example (a). Blue indicates small M3C2 distances, whereas orange corresponds to large distances, i.e., considerable reconstruction errors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Additionally, the point clouds do not provide much information about the upper layers of the canopy. The generated videogrammetric point clouds include points up to approximately 15 m in height. Although we used a spherical camera, we captured data from a terrestrial viewpoint at a considerable distance from the canopy. With our multi-lens approach, we cannot guarantee that the estimated scale will be accurate for large distances from the camera sensor, as is the case for the upper parts of the canopy.

Furthermore, the low resolution around the projected spherical poles of a spherical video (the upper and lower quarters of the projected spherical video frames) combined with the complexity of the canopy may hinder effective feature matching on leaves and small branches. Moreover, the significant distance between the canopy and the camera sensor poses a challenge in reconstructing reliable depth information for these canopy regions. This difficulty in accurately reconstructing the upper layers of the canopy is a common issue in terrestrial videogrammetric methods used in forestry, as also demonstrated in previous studies (Murtiyoso et al., 2022; Kükenbrink et al., 2022; Hristova et al., 2024). To complete the point cloud, approaches combining terrestrial and Unmanned Aerial Vehicle (UAV) point clouds can be considered (Kushwaha et al., 2023). This topic lies beyond the scope of this paper and is identified as a direction for future work.

4.2. Point cloud scale and orientation

By employing our multi-rig approach, we achieved a good scale for the generated point cloud using only a single-body camera without the requirement of visually coded targets. We analyzed the scale of the generated point clouds based on the distances measured between the visually coded targets, which serve as a representation of the real-world scale (final column in Table 2). Our analysis revealed that the scale of the generated point clouds deviated by no more than 5 mm from the actual real-world scale. Using the Insta 360 Pro 2 system with two spatial constraints thus eliminates the need for visually coded targets, simplifying data acquisition for non-expert users and making the process more time-efficient.

Our method generates videogrammetric point clouds that are initially not oriented. The orientation can be adjusted during post-processing, either through georeferencing or by rotating the point clouds in CloudCompare to align with the ground. Georeferencing can be achieved by placing visual markers on the forest floor, usually three in total, and recording their coordinates with a Global Navigation Satellite System (GNSS) receiver. These markers can then be automatically identified in the point cloud. A transformation can be applied to the unoriented point cloud based on the coordinates of the detected markers, thereby georeferencing the point cloud. The point cloud orientation is essential when applying algorithms for tree position and DBH estimation (Hristova et al., 2024). In our evaluation, all generated point clouds were georeferenced, as we aligned them with the reference TLS point clouds.

4.3. DBH and tree position estimation

The average DBH error was less than 2 cm, and its relative consistency across the study plots suggests that our method is reliable for reconstructing tree DBH (see Fig. 9). The expected error can reach up to approximately 7% relative to the DBH estimated from the TLS point cloud. The distributions of DBH errors exhibited a small standard deviation, ranging from 1 cm to 2 cm, further confirming the robustness of our approach. In contrast, Hristova et al. (2024) report an average DBH error of 20% for their approach with two spherical cameras positioned 60 cm apart. Their results are based on a study in a temperate forest. Our study and the study by Hristova et al. (2024) cannot be compared directly since the analyses were done on different forest plots. However, our method estimated the tree positions with an average position error ranging from 5.2 cm to 18.8 cm, which is about two to six times lower than the best average position error obtained by Hristova et al. (2024).

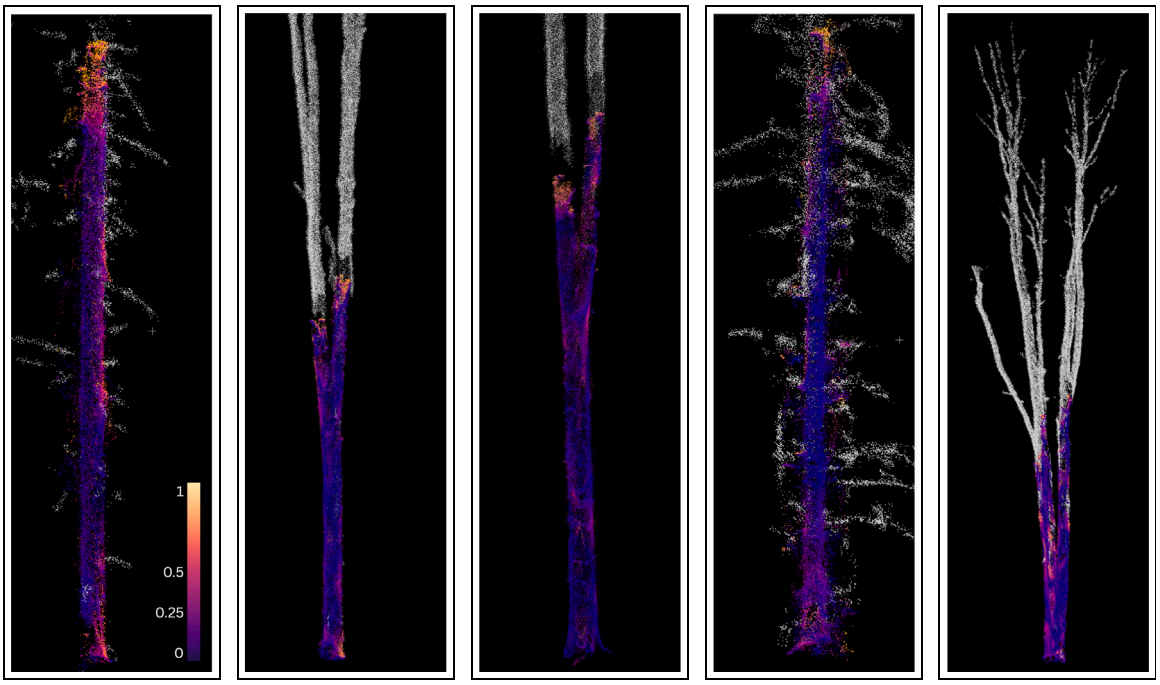


Fig. 11. Snippets of the Multiscale Model-to-Model Cloud Comparison (M3C2) point clouds, visualized in false colors. Purple and blue indicate small errors, whereas orange depicts large errors, i.e., substantial reconstruction distances. The color bar shown in example (a) illustrates the color range for all examples in this figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Our analysis indicates that Plot 4 has the highest average tree position error, measuring 18.79 cm, followed by Plot 5 with an average positional error of 16.08 cm. Plot 4 also had the greatest tree density and the steepest average slope among all the plots, along with the largest number of young trees. The combination of these three factors made Plot 4 particularly challenging, likely contributing to the increased positional errors. Similarly, Plot 5 was characterized by a high average slope and was partially obscured by an underlayer that is less than 10 m in height.

In our study, the smallest average absolute DBH error and the lowest deviation in relative DBH error were found for Plot 3 and Plot 4. However, both of these plots exhibited relatively large percentage DBH errors, which were similar to those observed for Plot 5 and Plot 6. This can be attributed to the presence of many young trees with small DBH values in Plot 3 and Plot 4 (Fig. 9). Yet, Plots 3 and 4 differed in the percentage of trees with a relative DBH value of less than 10% and an absolute DBH value greater than 2 cm. Plot 4 had the fewest trees whose DBH measurement differed by more than 2 cm from the DBH estimated using TLS. This is likely due to Plot 4 containing the highest number of young trees compared to all the other plots in our study (refer to Table 1), resulting in an overall decrease in the relative DBH errors.

Some of the absolute DBH errors exceeding 2 cm (reported in Table 3) were attributed to stem incompleteness and noise, as shown in Fig. 12(a). Stem incompleteness was influenced by factors such as the acquisition path and occlusion. For instance, Plot 3 had the lowest percentage of trees with a relative DBH error below 10%, and approximately 30% of the trees in this plot exhibited an absolute DBH error greater than 2 cm. Additionally, Plot 3 was one of the areas with the highest level of occlusion, which increased the risk of stem incompleteness during 3D reconstruction. Our analysis revealed that substantial DBH errors can also arise from incorrect clustering of data points. For instance, all the points presented in Fig. 12(b) belong to a single cluster despite representing two separate tree stems. This led to an overestimation of the DBH.

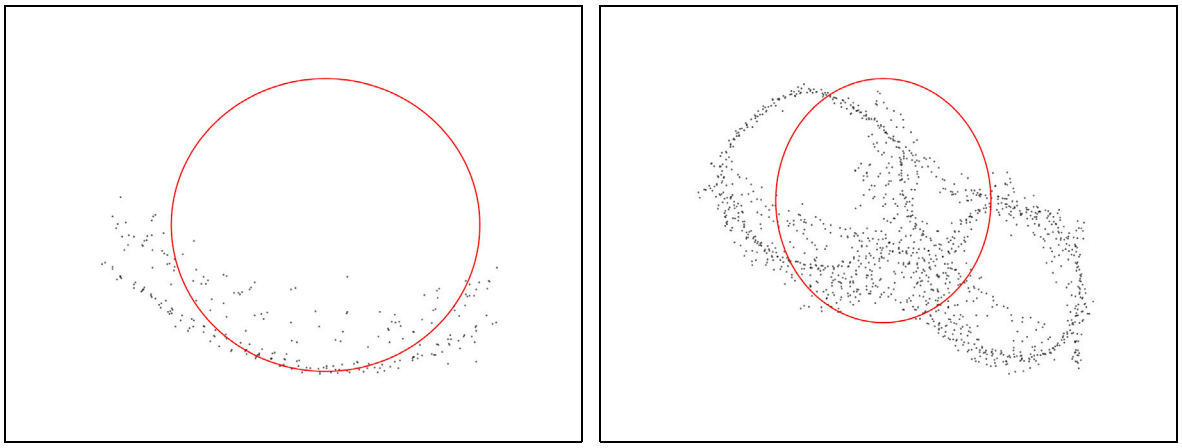
We discovered that occasional incorrect clustering was caused by residual noise in the data. Although we aimed to minimize noise near tree stems by incorporating a filtering step during the DBH estimation – based on the SPHERICITY feature of the point cloud – some noise may still have persisted, depending on the density of the generated point cloud. Overall, the proposed approach is effective for applications that include estimating DBH distribution and assessing forest stand maturity.

4.4. Multi-rig videogrammetry with Insta 360 Pro 2

In a recent study, the performance of the Insta 360 Pro 2 spherical camera was evaluated in narrow spaces, particularly through the work of Bruno et al. (2024). Our study builds upon their findings by utilizing the same camera and following their recommendations for a constrained multi-rig system. However, while Bruno et al. (2024) focused on narrow environments, our research targets outdoor forest settings characterized by varying illumination and textural differences. We aimed to analyze the effectiveness of the multi-rig setup in highly structured and complex open conditions that may involve occlusion, as well as to assess our method's ability to accurately reconstruct real-world distances and scales in both open and complex environments.

Bruno et al. (2024) reported a Root-Mean-Square Error (RMSE) of 1.6 cm when employing a constrained multi-rig setup with six Ground Control Points GCPs. While the use of GCPs generally enhances accuracy, in our approach we intentionally omit them to streamline data acquisition. Furthermore, our method is video-based, which significantly accelerates the collection process compared with that of Bruno et al. (2024).

We employ a single-camera setup in our approach to capture data in forested areas, allowing for adaptability across various types of forests. In contrast to the dual-camera approach using stereo Ricoh Theta Z1 cameras, demonstrated by Hristova et al. (2024), the Insta 360 Pro 2 features automatic lens synchronization. This capability reduces alignment errors between frames and thus indirectly enhances



(a) Incomplete stem (b) Incorrect stem clustering

Fig. 12. Examples of clusters corresponding to tree stems with substantial Diameter at Breast Height (DBH) errors (larger than 2 cm). Example (a) illustrates an incomplete DBH, whereas example (b) demonstrates incorrect clustering caused by noise. The points belonging to the cross-section at 1.3 m height along a tree stem are shown in black, and the estimated DBH is shown in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the quality of the resulting point clouds. As illustrated in our findings (Fig. 5), we have determined that using videos from a single-body camera while navigating through forest terrain does not compromise the quality of reconstruction or the accuracy of the scale.

Our setup is significantly more affordable and user-friendly than TLS and MLS devices, while representing an upgrade from consumer-grade cameras like those found in smartphones. One key benefit of our camera equipment over smartphone cameras is its use of multiple lenses, which can capture the forest scene from various angles simultaneously. This capability can significantly speed up data collection. Additionally, these multiple lenses help ensure accurate real-world scaling of the 3D reconstructions of the forest scene (see Section 4.2). In contrast, smartphone cameras rely on visually coded targets to scale the generated point clouds (Murtiyoso et al., 2022). The need for these visually coded targets can increase data acquisition time and requires careful placement throughout the forest area (Hristova et al., 2024). In our approach, the scale is inherently integrated into the pipeline, eliminating these complications.

4.5. Implications for forest management planning

The advances in digital camera technology enable the reconstruction of rich, spatially explicit data on individual trees, including DBH, spatial position, and, in the near future, species identification (Puliti et al., 2024). This enables reliable 3D reconstructions of forest stands and detailed assessments of tree social status (e.g., spatial distribution, competition). Such fine-scale information opens new possibilities for forest management planning, allowing interventions to be downscaled from the traditional stand level (O'Hara and Nagel, 2013) to the level of individual trees. Thus, it can inform tree-level harvesting decisions by identifying which trees should be retained or removed – and, in the near future, by estimating tree volumes of removals.

Beyond operational planning, data collected through our method can also serve as initialization input for forest growth simulations (Neudam et al., 2023). By using 3D reconstructions, we can accurately estimate the characteristics of each tree within a defined forest area, which eliminates the need for traditional inventory methods. This advancement helps to enhance the accuracy of forest growth simulations as initialization data are provided for small areas; the growth, competition, and mortality of trees can thus be simulated per patch and therefore more accurately. When exploring forest dynamics under various intervention scenarios and climate change conditions, a more solid foundation for adaptive management can be provided (Thrippleton et al., 2021).

Finally, policies such as the EU Forest Strategy (EC, 2021) and the EU Biodiversity Strategy (EC, 2020) can be supported by this approach. These policies increasingly advocate for climate-adaptive and single-tree-based continuous-cover forestry practices to support multifunctional forest goals and biodiversity conservation (Mason et al., 2022; Brang et al., 2014). Implementing such policy objectives requires detailed, high-resolution information on current forest conditions. In this context, the data obtained with our spherical camera approach offers valuable insights. For instance, it enables the linkage of forest structural attributes to indicators of key ecosystem services and biodiversity (Oettel and Lapin, 2021; Blattert et al., 2017).

5. Conclusion

Our study demonstrates the potential of the proposed videogrammetry pipeline using the Insta 360 Pro 2 setup for effectively reconstructing 3D representations of complex forest environments. Notably, the scale error is less than 5 mm, indicating that the videogrammetric reconstructions closely match real-world dimensions. Our findings show that the proposed approach not only streamlines data collection efforts but also obtains good estimates of critical forest metrics, such as tree position and DBH. The average values of tree position and DBH estimates remain below 20 cm and 2 cm, respectively, and are comparable to estimates derived from TLS for the same tree features. By eliminating the need for visually coded targets, we have simplified the acquisition workflow, making it more user-friendly and accessible for forest management applications. Creating reliable 3D reconstructions efficiently supports forest ecosystem management and planning, providing essential insights into stand structure and composition. Overall, our research highlights the robustness and practicality of the Insta 360 Pro 2 method, paving the way for its broader adoption in forestry studies.

CRediT authorship contribution statement

Hristina Hristova: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Clemens Blattert:** Writing – original draft, Methodology, Conceptualization. **Bogdan Candrea:** Writing – original draft, Resources, Data curation. **Mihai Nita:** Writing – original draft, Resources, Data curation, Conceptualization. **Sergiu Florea:** Writing – original draft,

Resources, Data curation. **Sunni Kanta Prasad Kushwaha:** Writing – original draft, Conceptualization. **Janine Schweier:** Writing – original draft, Supervision, Project administration, Funding acquisition, Conceptualization.

Funding

This work received funding from the European Research Council (ERC) under the European Union Horizon Europe (HORIZON) Research and Innovation program with grant agreement No. 101135517 (Small4Good), and from the Swiss State Secretariat for Education, Research, and Innovation (SERI).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hristina Hristova reports financial support was provided by European Union Horizon Europe (HORIZON) Research and Innovation program. Hristina Hristova reports financial support was provided by Swiss State Secretariat for Education, Research, and Innovation (SERI). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ecoinf.2025.103398.

Data availability

The generated videogrammetric point clouds using the proposed pipeline are available for download here: https://doi.org/10.5281/zenodo.16258209. The folder also contains walkthrough demos of the point clouds, as well as video comparisons with TLS-derived point clouds.

References

AgiSoft, 2018. AgiSoft metashape professional (version 1.4.5) (software),. Available Online: http://www.agisoft.com.

Alsadik, B., Gerke, M., Vosselman, G., 2015. Efficient use of video for 3D modelling of cultural heritage objects. ISPRS Ann. Photogramm. Remote. Sens. Spatial Inf. Sci. II-3/W4, 1–8.

Apostol, B., Chivulescu, S., Ciceu, A., Petrila, M., Pascu, I.-S., Apostol, E.N., Leca, S., Lorent, A., Tanase, M., Badea, O., 2007. Data collection methods for forest inventory: a comparison between an integrated conventional equipment and terrestrial laser scanning. Ann. For. Res. 50, 189–202.

Bienert, A., Georgi, L., Kunz, M., Maas, H.-G., Von Oheimb, G., 2018. Comparison and combination of mobile and terrestrial laser scanning for natural forest inventories. Forests 9 (7), 395.

Blattert, C., Lemm, R., Thees, O., Lexer, M.J., Hanewinkel, M., 2017. Management of ecosystem services in mountain forests: Review of indicators and value functions for model based multi-criteria decision analysis. Ecol. Indic. 79, 391–409.

Bornand, A., Rehush, N., Morsdorf, F., Thürig, E., Abegg, M., 2023. Individual tree volume estimation with terrestrial laser scanning: Evaluating reconstructive and allometric approaches. Agricult. Forest. Meterol. 341, 109654. http://dx.doi.org/10.1016/j.agrformet.2023.109654.

Brang, P., Spathelf, P., Larsen, J.B., Bauhus, J., Boncčìna, A., Chauvin, C., Drössler, L., García-Güemes, C., Heiri, C., Kerr, G., et al., 2014. Suitability of close-to-nature silviculture for adapting temperate European forests to climate change. For.: An Int. J. For. Res. 87 (4), 492–503.

Breidenbach, J., Granhus, A., Hylen, G., Eriksen, R., Astrup, R., 2020. A century of national forest inventory in Norway–informing past, present, and future decisions. Forest ecosyst 7 (1): 46.

Bruno, N., Perfetti, L., Fassi, F., Roncella, R., et al., 2024. Photogrammetric survey of narrow spaces in cultural heritage: Comparison of two multi-camera approaches. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 48, 87–94.

Cabo, C., Del Pozo, S., Rodríguez-Gonzálvez, P., Ordóñez, C., González-Aguilera, D., 2018. Comparing terrestrial laser scanning (TLS) and wearable laser scanning (WLS) for individual tree modeling at plot level. Remote. Sens. 10 (4), 540.

Calders, K., Adams, J., Armston, J., Bartholomeus, H., Bauwens, S., Bentley, L.P., Chave, J., Danson, F.M., Demol, M., Disney, M., Gaulton, R., Krishna Moorthy, S.M., Levick, S.R., Saarinen, N., Schaaf, C., Stovall, A., Terryn, L., Wilkes, P., Verbeeck, H., 2020. Terrestrial laser scanning in forest ecology: Expanding the horizon. Remote Sens. Environ. (ISSN: 0034-4257) 251, 112102. http://dx.doi.org/10.1016/j.rse.2020.112102, URL https://www.sciencedirect.com/science/article/pii/S0034425720304752.

CloudCompare, 2023. CloudCompare: 3D point cloud and mesh processing software.

Dai, F., Feng, Y., Hough, R., 2014. Photogrammetric error sources and impacts on modeling and surveying in construction engineering applications. Vis. Eng. 2 (1), 1–14. http://dx.doi.org/10.1186/2213-7459-2-2.

EC, 2020. Communication from the commission to the European parliament, the council, the economic and social committee and the committee of the regions. EU biodiversity strategy for 2030. Bringing nature back into our lives.. Eur. Comm.: Bruss. Belg..

EC, 2021. Communication from the commission to the European parliament, the council, the European economic and social committee and the committee for the regions. New EU forest strategy for 2030.. Eur. Comm.: Bruss. Belg..

Fasihi, M., Portelli, B., Cadez, L., Tomao, A., Falcon, A., Alberti, G., Serra, G., 2024. Assessing ensemble models for carbon sequestration and storage estimation in forests using remote sensing data. Ecol. Informatics 83, 102828.

Fassnacht, F.E., White, J.C., Wulder, M.A., Næsset, E., 2024. Remote sensing in forestry: current challenges, considerations and directions. For.: An Int. J. For. Res. 97 (1), 11–37.

Ferretti, M., Fischer, C., Gessler, A., Graham, C., Meusburger, K., Abegg, M., Bebi, P., Bergamini, A., Brockerhoff, E.G., Brunner, I., et al., 2024. Advancing forest inventorying and monitoring. Ann. For. Sci. 81 (1), 6.

Forsman, M., Börlin, N., Holmgren, J., 2016. Estimation of tree stem attributes using terrestrial photogrammetry with a camera rig. Forests 7 (3), 61.

Griebel, A., Bennett, L.T., Culvenor, D.S., Newnham, G.J., Arndt, S.K., 2015. Reliability and limitations of a novel terrestrial laser scanner for daily monitoring of forest canopy dynamics. Remote Sens. Environ. 166, 205–213.

Hasegawa, H., Matsuo, K., Koarai, M., Watanabe, N., Masaharu, H., Fukushima, Y., 2000. DEM accuracy and the base to height (b/h) ratio of stereo images. Int. Arch. Photogramm. Remote Sens. 33 (B4/1; PART 4), 356–359.

Hristova, H., Murtiyoso, A., Kükenbrink, D., Marty, M., Abegg, M., Fischer, C., Griess, V.C., Rehush, N., 2024. Viewing the forest in 3D: How spherical stereo videos enable low-cost reconstruction of forest plots. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens..

Hyyppä, E., Yu, X., Kaartinen, H., Hakala, T., Kukko, A., Vastaranta, M., Hyyppä, J., 2020. Comparison of backpack, handheld, under-canopy UAV, and above-canopy UAV laser scanning for field reference data collection in boreal forests. Remote. Sens. 12 (20), 3327.

Knoke, T., Kindu, M., Schneider, T., Gobakken, T., 2021. Inventory of forest attributes to support the integration of non-provisioning ecosystem services and biodiversity into forest planning—from collecting data to providing information. Curr. For. Rep. 7, 38–58.

Kükenbrink, D., Marty, M., Bösch, R., Ginzler, C., 2022. Benchmarking laser scanning and terrestrial photogrammetry to extract forest inventory parameters in a complex temperate forest. Int. J. Appl. Earth Obs. Geoinf. 113, 102999.

Kushwaha, S.K.P., Singh, A., Jain, K., Cabo, C., Mokros, M., 2023. Integrating airborne and terrestrial laser scanning for complete 3D model generation in dense forest. In: IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium. pp. 3137–3140. http://dx.doi.org/10.1109/IGARSS52108.2023.10283032.

Kwiatek, K., Tokarczyk, R., 2014. Photogrammetric applications of immersive video cameras. ISPRS Ann. Photogramm. Remote. Sens. Spatial Inf. Sci. 2 (5), 211.

Lague, D., Brodu, N., Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the rangitikei canyon (NZ). ISPRS J. Photogramm. Remote Sens. 82, 10–26.

Larsen, J.B., Angelstam, P., Bauhus, J., Carvalho, J.F., Diaci, J., Dobrowolska, D., Gazda, A., Gustafsson, L., Krumm, F., Knoke, T., et al., 2022. Closer-to-nature forest management. From science to policy 12.. vol. 12, EFI European Forest Institute.

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., Holopainen, M., Vastaranta, M., 2016b. Terrestrial laser scanning in forest inventories. ISPRS J. Photogramm. Remote Sens. (ISSN: 0924-2716) 115, 63–77. http://dx.doi.org/10.1016/j.isprsjprs.2016.01.006, URL https://www.sciencedirect.com/science/article/pii/S0924271616000204 Theme issue 'State-of-the-art in photogrammetry, remote sensing and spatial information science'.

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., et al., 2016a. Terrestrial laser scanning in forest inventories. ISPRS J. Photogramm. Remote Sens. 115, 63–77.

Malladi, M.V., Guadagnino, T., Lobefaro, L., Mattamala, M., Griess, H., Schweier, J., Chebrolu, N., Fallon, M., Behley, J., Stachniss, C., 2024. Tree instance segmentation and traits estimation for forestry environments exploiting lidar data collected by mobile robots. In: 2024 IEEE International Conference on Robotics and Automation. ICRA, IEEE, pp. 17933–17940.

Mason, W.L., Diaci, J., Carvalho, J., Valkonen, S., 2022. Continuous cover forestry in europe: usage and the knowledge gaps and challenges to wider adoption. For.: An Int. J. For. Res. 95 (1), 1–12.

Mattamala, M., Chebrolu, N., Casseau, B., Freißmuth, L., Frey, J., Tuna, T., Hutter, M., Fallon, M., 2024. Autonomous forest inventory with legged robots: system design and field deployment. arXiv preprint arXiv:2404.14157.

Mokroš, M., Liang, X., Surovỳ, P., Valent, P., Čerňava, J., Chudỳ, F., Tunák, D., Saloň, Š., Merganič, J., 2018. Evaluation of close-range photogrammetry image collection methods for estimating tree diameters. ISPRS Int. J. Geo-Information 7 (3), 93.

Murtiyoso, A., Grussenmeyer, P., 2021. Experiments using smartphone-based videogrammetry for low-cost cultural heritage documentation. In: 28th CIPA Symposium, 28 AoÛt-1er Septembre, UniversitÉ Tsinghua, Pekin, Chine. Vol. 46.

Murtiyoso, A., Hristova, H., Rehush, N., Griess, V., 2022. Low-cost mapping of forest under-storey vegetation using spherical photogrammetry. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 48, 185–190.

Neudam, L.C., Fuchs, J.M., Mjema, E., Johannmeier, A., Ammer, C., Annighöfer, P., Paul, C., Seidel, D., 2023. Simulation of silvicultural treatments based on real 3D forest data from mobile laser scanning point clouds. Trees, For. People 11, 100372.

Oettel, J., Lapin, K., 2021. Linking forest management and biodiversity indicators to strengthen sustainable forest management in europe. Ecol. Indic. 122, 107275.

O'Hara, K.L., Nagel, L.M., 2013. The stand: revisiting a central concept in forestry. J. For. 111 (5), 335–340.

Pepe, M., Alfio, V.S., Costantino, D., Herban, S., 2022. Rapid and accurate production of 3D point cloud via latest-generation sensors in the field of cultural heritage: A comparison between SLAM and spherical videogrammetry. Heritage 5 (3), 1910–1928.

Puliti, S., Lines, E.R., Müllerová, J., Frey, J., Schindler, Z., Straker, A., Allen, M.J., Winiwarter, L., Rehush, N., Hristova, H., et al., 2024. Benchmarking tree species classification from proximally-sensed laser scanning data: introducing the FOR-species20k dataset. arXiv preprint arXiv:2408.06507.

Su, Y., Wu, Z., Zheng, X., Qiu, Y., Ma, Z., Ren, Y., Bai, Y., 2025. Harmonizing remote sensing and ground data for forest aboveground biomass estimation. Ecol. Informatics 86, 103002.

Sun, Z., Zhang, Y., 2019. Accuracy evaluation of videogrammetry using a low-cost spherical camera for narrow architectural heritage: An observational study with variable baselines and blur filters. Sensors 19 (3), 496.

Thrippleton, T., Blattert, C., Bont, L.G., Mey, R., Zell, J., Thürig, E., Schweier, J., 2021. A multi-criteria decision support system for strategic planning at the swiss forest enterprise level: coping with climate change and shifting demands in ecosystem service provisioning. Front. For. Glob. Chang. 4, 693020.

Zhu, R., Guo, Z., Zhang, X., 2021. Forest 3D reconstruction and individual tree parameter extraction combining close-range photo enhancement and feature matching. Remote. Sens. 13 (9), 1633.