remote sensing

MDPI

Article

Quantifying Seasonal Changes in Forest Three-Dimensional Reconstruction Using UAV-SfM Point Clouds: Voxel Entity Density (VED) as an Indicator of Reconstruction Dynamics

Kazuho Ohnishi [1,*], Hidetake Hirayama [2], Ayako Sekiyama [2] and Sawahiko Shimada [2]

[1] Graduate School of Agro-Environmental Science, Tokyo University of Agriculture, Sakuragaoka 1-1-1, Tokyo 156-8502, Japan

[2] Faculty of Regional Environment Science, Tokyo University of Agriculture, Sakuragaoka 1-1-1, Tokyo 156-8502, Japan; hh207501@nodai.ac.jp (H.H.); a3sekiya@nodai.ac.jp (A.S.); shima123@nodai.ac.jp (S.S.)

* Correspondence: 12624002@nodai.ac.jp

Highlights

What are the main findings?

- Voxel Entity Density (VED) was used to quantify the local density of voxel centers reconstructed from multitemporal UAV-SfM point clouds.
- VED revealed contrasting seasonal patterns between selected upper-canopy foliage and stem regions, while the understory shrub exhibited a distinct pattern.

What are the implications of the main findings?

- VED enables quantitative comparison of seasonal changes in forest reconstruction within a common three-dimensional voxel space.
- Seasonal VED analysis may help characterize forest observation conditions for stems and understory vegetation affected by canopy occlusion.

Abstract

In three-dimensional forest reconstruction using structure from motion (SfM), branches, stems, and understory vegetation obscured by upper-canopy foliage after leaf-out are difficult to capture in images. In deciduous forests, this occlusion changes seasonally and may therefore alter point cloud reconstruction. We integrated UAV-SfM point clouds from 26 dates over approximately 14 months into a common three-dimensional voxel space. We compared Voxel Entity Density (VED), the number of reconstructed neighboring voxel centers per unit volume, among selected upper-canopy foliage, stem, and understory shrub regions. VED for the selected upper-canopy foliage decreased from autumn to winter and increased from spring onward. Mean VED in the selected stem regions increased from autumn to winter and subsequently declined, while the understory shrub showed a different temporal pattern.

Keywords: UAV photogrammetry; SfM point clouds; voxel; VED; deciduous forest; seasonal change

Academic Editors: Antonio Luis Montealegre Gracia, Alberto García-Martín and Raúl Hoffrén

Received: 6 October 2026
Revised: 16 November 2026
Accepted: 19 November 2026
Published: 24 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

In forest remote sensing using unmanned aerial vehicles (UAVs), RGB and multispectral images are widely used to assess vegetation conditions and classify tree species. In particular, spectral information in the visible, red-edge, and near-infrared regions, to-

gether with vegetation indices such as the Normalized Difference Vegetation Index (NDVI), has been used to assess vegetation conditions and capture differences in spectral characteristics among species. Numerous studies have also reported tree species classification at the individual-tree level using the high spatial resolution of UAV imagery, and the use of multitemporal imagery for species classification has been investigated in recent years [1–4]. However, two-dimensional images, including orthophotos, cannot directly observe branches, stems, or understory vegetation obscured by upper-canopy crowns, limiting their ability to represent the vertical structure of forests.

Point cloud measurements using light detection and ranging (LiDAR) are widely used to characterize three-dimensional forest structure. In addition to tree height and vertical structure, methods have been developed to estimate Leaf Area Density (LAD) and Plant Area Density (PAD) by dividing three-dimensional space into voxels, thereby quantifying the area density of foliage and woody material [5–8]. In deciduous forests, changes in laser penetration and point cloud distribution between leaf-on and leaf-off conditions have also been used to distinguish upper-canopy crowns from understory vegetation and evaluate seasonal differences in leaf area and forest structure [9,10]. Thus, three-dimensional evaluation of the area density of foliage and woody material has already been studied, primarily using LiDAR.

In contrast, photogrammetry using structure from motion/multi-view stereo (SfM/MVS) is increasingly used in forest measurements because it can generate dense three-dimensional point clouds from images at relatively low cost [11]. Three-dimensional features, such as tree height and crown structure, can be obtained from SfM point clouds and combined with spectral information for tree species classification [4]. Unlike LiDAR, however, SfM reconstruction requires the target to be visible in multiple images. In multilayered forests, occlusion by upper vegetation therefore limits the visibility of the ground surface, branches, stems, and understory vegetation, making their reconstruction as point clouds difficult. Studies have reported that SfM point clouds cannot fully reconstruct the vertical structure beneath dense canopies compared with LiDAR point clouds, and that reconstruction completeness varies with image overlap, spatial resolution, and vegetation occlusion [11–13]. This property constrains the application of SfM to forest measurements.

Nevertheless, the extent of this occlusion is not constant throughout the year in deciduous forests. If branches, stems, and understory vegetation that are hidden by upper-canopy crowns during leaf-out become visible in images following leaf fall, the three-dimensional space reconstructed by SfM may also change over time. In addition to using multiseason UAV data for tree species classification [4], previous studies have combined stem positions extracted from leaf-off UAV-SfM point clouds with leaf-on imagery and canopy height information to delineate individual tree crowns [14]. Such differences in the information obtained across seasons have been used to characterize forest structure in a complementary manner. Beyond this complementary use, retaining differences in reconstruction at each time within a common three-dimensional space and evaluating temporal changes in their local density may allow seasonal observation conditions for upper-canopy foliage, branches and stems, and understory vegetation to be quantified.

In this study, we therefore integrated UAV-SfM point clouds acquired at multiple times into a common three-dimensional voxel space and used Voxel Entity Density (VED) to quantify the local density of reconstructed voxel centers. We conducted UAV time-series observations over approximately 14 months in a deciduous forest and compared seasonal VED changes in selected stem, upper-canopy foliage, and understory regions. The objective was to examine whether VED can quantify seasonal changes in three-dimensional reconstruction at these locations.

2. Materials and Methods

2.1. Study Area

The study area was a forested zone between Oyama Plaza and Picnic Plaza in Kanagawa Prefectural Nanasawa Forest Park, Japan (Figure 1).



Figure 1. Location of study site and survey area.

The centroid of the park boundary data (WGS 84) is located at 35.4480° N, 139.3015° E and represents the park as a whole. Elevation across the park ranges from approximately 88 to 188 m, with an elevation difference of approximately 100 m [15]. These elevations do not refer specifically to the analysis plot.

The study plot is a secondary forest dominated by *Quercus acutissima* and *Quercus serrata*, rather than an entirely natural forest. *Acer pictum* subsp. *dissectum* f. *dissectum*, *Magnolia kobus*, and *Stachyurus praecox* are also present.

At the nearby Ebina meteorological station, the 1991–2020 climate normals indicate a mean annual temperature of 15.6 °C and annual precipitation of 1760.9 mm [16]. These are not measurements taken within the analysis plot.

2.2. UAV Image Acquisition and Point Cloud Generation

The analytical workflow used in this study is shown in Figure 2.

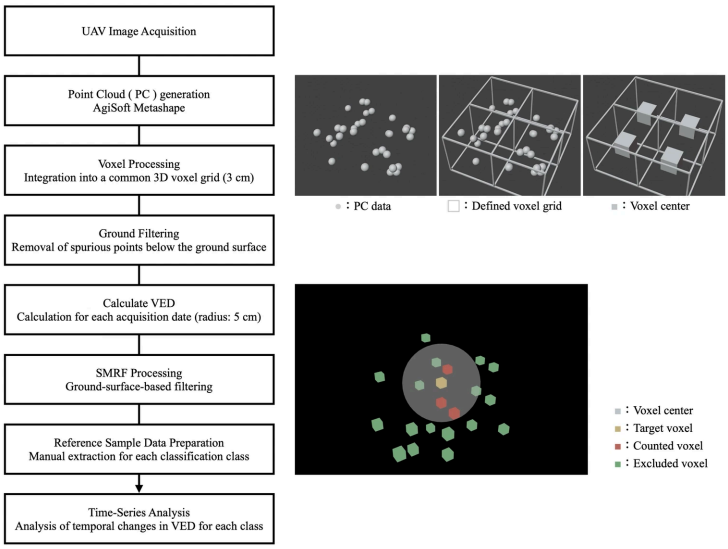

Figure 2. Analytical workflow.

Aerial images were acquired using a DJI Mavic 3 Multispectral (SZ DJI Technology Co., Ltd., Shenzhen, China). This UAV is equipped with an RGB camera featuring a 4/3 CMOS sensor with 20 MP effective resolution [17]. Separate SfM/MVS point clouds from

green, red, red-edge, and near-infrared imagery supplied attributes for preprocessing (Section 2.3).

An oblique collection mode was used to maintain a height of 50 m above the ground surface based on a previously generated Digital Surface Model (DSM). In addition to nadir imaging, oblique images were acquired in four directions—east, west, south, and north—with a camera angle of 45° from nadir. Forward and side overlap were set to 80% and 70%, respectively, for nadir imaging, and to 70% and 60%, respectively, for oblique imaging. Real-time kinematic (RTK) positioning was used to estimate camera positions at exposure time. Across 36,244 valid exposure records from the 26 acquisition dates, the mean GNSS position uncertainty reported in the flight metadata was 1.58 cm horizontally (root-sum-square of the north and east standard deviations) and 2.67 cm vertically (vertical standard deviation).

Point clouds derived from images acquired on 26 dates (6 September 2024 to 14 November 2025) were used for VED calculation and time-series analysis. Acquisition conditions and reconstructed point counts for each date are listed in Table 1.

Table 1. Dates and details of UAV aerial data acquisition.

Acquisition Date	DOY	Acquisition Time	Solar Elevation (°)	Weather	Air Temperature (°C)	Wind Speed ($m\ s^{-1}$)	Number of Points
6 September 2024	250	10:07:50–11:12:12	57.8	Clear	31.6	1.1	1,639,618,598
18 September 2024	262	10:06:46–11:22:54	54.4	Clear	33.4	1.5	1,639,345,668
3 October 2024	277	10:27:36–11:50:19	50.3	Cloudy	24.4	1.8	1,741,526,553
15 October 2024	289	10:10:36–11:29:15	45.2	Clear	25.1	1.1	1,651,644,959
31 October 2024	305	10:09:13–11:20:10	39.5	Clear	19.2	1.6	1,678,045,239
26 November 2024	331	10:03:53–10:52:50	31.7	Clear	—	—	1,422,876,539
5 December 2024	340	10:03:57–11:20:07	30.9	Clear	14.7	1.1	1,438,216,688
23 December 2024	358	10:07:28–11:18:17	29.5	Clear	8.6	1.3	1,037,638,905
9 January 2025	9	10:05:01–11:12:49	29.9	Clear	10.6	4.7	874,812,057
23 January 2025	23	10:04:49–11:14:55	32.1	Clear	10.5	1.6	907,919,421
5 February 2025	36	10:07:13–11:18:59	35.4	Clear	7.2	1.8	945,367,455
21 February 2025	52	10:03:35–11:21:45	40.4	Clear	8.1	1.7	937,158,737
7 March 2025	66	09:59:58–11:11:01	44.8	Clear	10.1	5.1	835,443,590
21 March 2025	80	10:05:38–11:15:24	50.7	Clear	12.6	3.4	849,093,196
8 April 2025	98	10:10:46–11:24:22	58.4	Cloudy	17.8	2.5	1,005,889,567
17 April 2025	107	10:11:10–11:25:43	61.8	Clear	20.9	2.0	1,324,979,702
9 May 2025	129	10:10:16–11:26:15	68.4	Cloudy	20.9	3.1	1,737,576,817
23 May 2025	143	10:16:59–11:26:52	71.7	Clear	22.1	3.6	1,668,614,862
5 June 2025	156	10:12:59–11:23:11	72.6	Clear	26.6	4.2	1,260,785,907
20 June 2025	171	10:13:52–11:23:41	73.1	Cloudy	28.5	3.0	1,862,337,223
3 July 2025	184	10:03:41–11:13:07	70.7	Cloudy	30.6	3.7	1,779,578,922
20 August 2025	232	10:08:28–11:17:40	63.2	Clear	33.5	2.6	1,695,736,853
26 September 2025	269	10:16:09–11:28:25	52.5	Clear	31.1	1.8	1,439,681,568
10 October 2025	283	10:10:48–11:20:31	47.1	Clear	22.9	1.4	1,583,699,151
31 October 2025	304	10:07:15–11:17:20	39.7	Cloudy	17.7	2.8	1,762,598,455
14 November 2025	318	10:04:23–11:20:53	35.6	Clear	17.7	1.0	1,559,620,241

Note: DOY denotes the number of days from 1 January of each year, which is assigned a value of

1, and differs from the continuous DOY used in Section 2.5. Air temperature and wind speed are the means of 10 min observations during the image acquisition period at the nearest Japan Meteorological Agency AMeDAS station, Ebina (station number 46091; 7.7 km from the study site), rather than measurements taken at the site. Relevant data were unavailable for 26 November 2024; these values are shown as missing (—) and were not estimated or filled.

Three-dimensional point clouds were generated from the acquired images through SfM/MVS processing in Agisoft Metashape 2.2.1 (build 20641) [18]. For image matching, downscale was set to 1, generic preselection and reference preselection were enabled, and the key-point limit and tie-point limit were set to 80,000 and 40,000, respectively. After camera alignment, depth maps were generated with downscale set to 1 and used to construct dense point clouds. The maximum number of neighboring images used for depth map filtering was set to 100. The coordinate reference system of the Metashape project was WGS 84 (EPSG:4326), and the generated point clouds were exported in LAS format using WGS 84/UTM zone 54N (EPSG:32654). Each point cloud was clipped to the plot polygon to create an analysis point cloud for each acquisition date. Before clipping, the plot polygon was transformed to the coordinate reference system recorded in the LAS header. Orthophotos of the study site in each season are shown in Figure 3.



Figure 3. Orthophotos for four seasons (summer, autumn, winter, and spring).

In each plot-clipped RGB LAS file, we computed three-dimensional nearest-neighbor distances within 10 blocks of 50,000 consecutive records spaced evenly by file index, counting each XYZ coordinate once per block. The date-wise medians before voxelization ranged from 8.5 to 12.0 mm across 26 dates.

2.3. Point Cloud Processing

Point clouds from each acquisition date were fused on a common three-dimensional voxel grid. Voxels with an edge length of v = 0.03 m were defined over the bounding extent of all input point clouds. The voxel indices of a point (x,y,z) were calculated as

$$i_x = \left\lfloor \frac{x - x_{\min}}{v} \right\rfloor, \qquad i_y = \left\lfloor \frac{y - y_{\min}}{v} \right\rfloor, \qquad i_z = \left\lfloor \frac{z - z_{\min}}{v} \right\rfloor. \tag{1}$$

Here, $x_{\min}$, $y_{\min}$, and $z_{\min}$ are the minimum values of the respective coordinates across all input point clouds. The union of voxels containing at least one point was constructed, and each voxel center was used as the three-dimensional coordinate of a fused point. When multiple points from the same acquisition date fell within the same voxel, the median of each RGB component was used as the representative value. RGB values for each date were stored as additional LAS attributes tagged with the acquisition date.

We combined RGB and multispectral attributes on the common grid and retained

voxels with finite attributes from at least two date-specific groups (RGB, green, red and red edge combined, or near infrared). This reduced the count from 54,811,675 to 38,132,013. We transferred multispectral attributes to the nearest RGB-supported center in three dimensions on the same date and retained centers with all three RGB components positive on at least one date, yielding 28,261,980 voxels.

For visual interpretation during subsequent reference sample data preparation, standard LAS RGB attributes were created from the date-specific RGB values. These values were checked in descending order of acquisition month and, within the same month, in descending order of acquisition date. The first valid RGB values found for each point were assigned as its display color.

Finally, the Simple Morphological Filter (SMRF) [19] was applied to the fused point cloud to separate ground and non-ground points. Ground filters, including SMRF, have also been applied to point clouds derived from UAV photogrammetry [20]. The SMRF parameters were slope =0.3, window =16.0 m, elevation threshold =0.5 m, and scalar =1.5. Points assigned classification 2 by SMRF were treated as ground points, and all other points as non-ground points.

2.4. Reference Sample Data Preparation

The author matched trees identified during field surveys to the fused point cloud using their locations. The non-ground point cloud extracted through SMRF processing was visually interpreted in CloudCompare. Clearly identifiable foliage and main-stem regions were selected and labeled, avoiding areas where neighboring structures were difficult to distinguish. The Stem class comprises main stems.

The nine labeled regions comprised two main-stem regions and foliage regions for one *Acer pictum* subsp. *dissectum* f. *dissectum*, one *Stachyurus praecox*, one *Magnolia kobus*, two *Quercus serrata*, and two *Quercus acutissima* trees.

The locations of the reference sample regions are shown in Figure 4, and details of the reference sample trees are provided in Table 2.

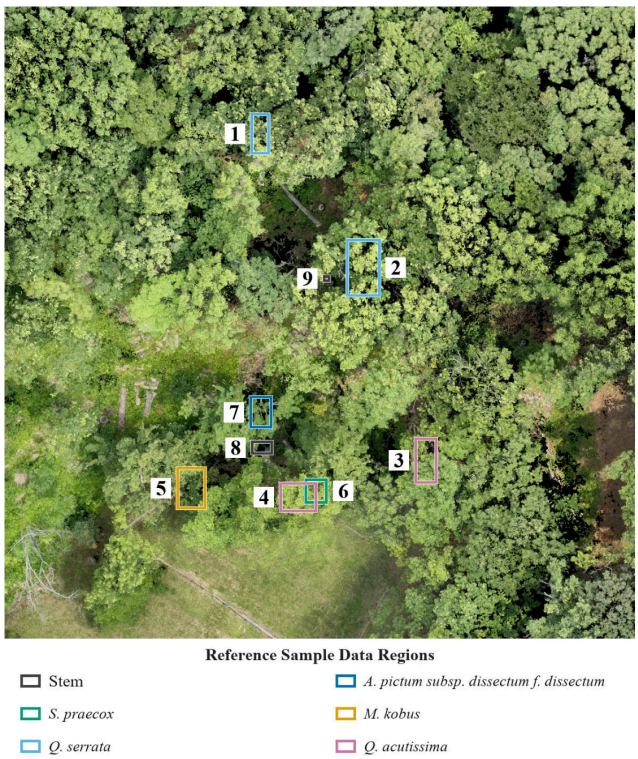

Figure 4. Locations of reference sample regions.

Table 2. Reference regions and associated trees.

Region ID	Species/Component	Tree Height (m)	Diameter at Breast Height (DBH) (cm)	Notes
1	*Quercus serrata*	17.8	42.0	
2	*Quercus serrata*	16.1	39.8	
3	*Quercus acutissima*	17.7	24.4	
4	*Quercus acutissima*	14.2	35.0	
5	*Magnolia kobus*	8.0	23.6	
6	*Stachyurus praecox*	2.8	8.6	
7	*Acer pictum* subsp. *dissectum* f. *dissectum*	11.0	28.0	Same tree as ID 8
8	Stem			Same tree as ID 7
9	Stem	20.7	55	

An axis-aligned bounding box (AABB) was defined from the minimum and maximum x, y, and z coordinates recorded in the header of each labeled LAS file. Points in the fused non-ground point cloud located inside or on the boundary of this AABB were extracted using a three-dimensional containment test. No margin was added around the AABB (0 m). When multiple labeled LAS files belonged to the same class, the union of their AABBs was used as the reference sample region for that class. The labeled regions were selected to limit inclusion of adjacent vegetation of other classes, including near their borders. This was a visual selection criterion, and residual class mixing is possible where components are close together.

ClassID values of 0, 1, 2, 3, 4, and 5 were assigned to Stem, *Acer pictum* subsp. *dissectum* f. *dissectum*, *Stachyurus praecox*, *Magnolia kobus*, *Quercus serrata*, and *Quercus acutissima*, respectively. Extracted points inherited the date-specific attributes retained in the fused point cloud, and ClassID and Class_name were added.

Points were then randomly sampled from each class so that all classes contained 25,000 evaluation locations. The same locations were used on all dates. These locations were spatial subsamples of the selected regions, with one or two trees per foliage class (Table 2).

For classes with two reference regions, we also calculated VED time series separately for each region using all its labeled locations on the same dates.

2.5. VED Calculation and Time-Series Analysis

Voxel Entity Density (VED) was calculated for the fused point cloud for each acquisition date d.

For each acquisition date, a KD-tree was constructed from the three-dimensional coordinates of voxel centers with a finite value in the RGB red channel on that date. For each evaluation location p, we counted the date-specific centers $n_d(p)$ within a sphere of radius r = 0.05 m. The center itself was excluded from the count if it was reconstructed on that date. The neighbor search uses all reconstructed RGB-supported voxel centers on that date, regardless of class or reference-region boundary. Class labels identify the evaluation locations used to summarize VED. Equation (2) defines VED:

$$\mathrm{VED}_d(p) = \frac{n_d(p)}{\frac{4}{3}\pi r^3} \quad (2)$$

The volume of a sphere with a 0.05 m radius is 0.000523598776 m^3, and VED is expressed in voxel centers m^{-3}. The 3 cm voxel edge exceeded the measured date-wise median nearest-neighbor distances (Section 2.2). The 5 cm radius includes the centers of six face-adjacent and twelve edge-adjacent voxels. Section 4.4 compares other settings.

For each acquisition date and class, the arithmetic mean was calculated across the same 25,000 reference locations. VED was zero when no reconstructed neighbor was found within the radius after excluding the center. All values were finite and included in the means.

$$\overline{\mathrm{VED}}_{c,d} = \frac{1}{n_{c,d}} \sum_{i=1}^{n_{c,d}} \mathrm{VED}_{i,c,d} \tag{3}$$

Here, $\overline{VED}_{c,d}$ is the mean VED for class c on acquisition date d, $VED_{i,c,d}$ is the VED at the ith location in that class on that date, and $n_{c,d}$ is the number of locations in the mean (25,000 for each class and date).

Whittaker smoothing was applied to the time series of mean VED for each class [21] to visualize broad seasonal trajectories. Whittaker smoothing considers both the fit to the observations and a penalty on the roughness of the smoothed series. The smoothed series $\mathbf{z}$ was obtained by minimizing the following objective function:

$$\| \mathbf{y} - \mathbf{z} \|^2 + \lambda \| \mathbf{D}_2 \mathbf{z} \|^2 \tag{4}$$

Here, $\mathbf{y}$ is the observed mean VED series, $\mathbf{z}$ is the smoothed series, λ is the smoothing parameter, and $\mathbf{D}_2$ is the second-order difference matrix. For the initial plots, we provisionally used $\lambda = 80$. We then compared 11 candidate values by holding out one acquisition date at a time. Section 3.3 reports the results.

Because the acquisition intervals were not constant, we used second-order differences that reflect the actual intervals between acquisition dates rather than treating the acquisition sequence as equally spaced. For irregularly spaced data, second-order divided differences that account for the time intervals can be constructed by incorporating adjacent observation intervals into the difference operator [21,22].

Acquisition dates were converted to continuous DOY, with 1 January 2024 assigned a value of 1. Thus, 31 December 2024 was assigned 366 and 1 January 2025 was assigned 367, providing a time coordinate that preserves the actual intervals in days across years. The 25 intervals between the 26 acquisition dates ranged from 9 to 48 days, with a median of 14 days. To express the representative acquisition interval as one time unit, continuous DOY was therefore normalized by the median interval of 14 days, and the time coordinate x_i for each acquisition date was defined as follows:

$$x_i = \frac{t_i - t_1}{\widetilde{\Delta t}} \tag{5}$$

Here, t_i is the continuous DOY of acquisition date i, t_1 is the continuous DOY of the first acquisition date, and $\widetilde{\Delta t}$ is the median interval between adjacent acquisitions, which was 14 days in this study.

The adjacent normalized time intervals were defined as

$$h_{\mathrm{left}} = x_i - x_{i-1}, \qquad h_{\mathrm{right}} = x_{i+1} - x_i, \tag{6}$$

and the nonzero coefficients in each interior row of the second-order difference matrix $\mathbf{D}_2$ for irregularly spaced observations were set as follows [22]:

$$\left[\frac{2}{h_{\text{left}}\left(h_{\text{left}}+h_{\text{right}}\right)},\ -\frac{2}{h_{\text{left}}h_{\text{right}}},\ \frac{2}{h_{\text{right}}\left(h_{\text{left}}+h_{\text{right}}\right)}\right] \tag{7}$$

The smoothed series was obtained by solving a system of linear equations equivalent to

$$\mathbf{z}=\left(\mathbf{I}+\lambda\mathbf{D}_2^{\mathsf{T}}\mathbf{D}_2\right)^{-1}\mathbf{y}. \tag{8}$$

Each of the 26 acquired-date means had equal weight. All were available for every class, and no values were imputed on unsurveyed dates. To check sensitivity to the smoothing parameter, we held out one acquisition date at a time, retained its time coordinate with zero fitting weight, and predicted the excluded class means from the remaining dates. We compared positive λ values of 0.01, 0.1, 1, 5, 10, 20, 40, 80, 160, 320, and 640. The root mean squared prediction error (RMSE) was pooled across the 156 date-by-class predictions.

As an exploratory analysis of trajectory shape, we compared the four *Quercus* foliage trajectories with the two stem trajectories using silhouette coefficients [23]. Each region's time series was centered and scaled to unit standard deviation across the 26 dates. Distances between standardized trajectories were Euclidean.

3. Results

3.1. Point Cloud Integration

The appearance of point clouds generated in each season is shown in Figure 5. The filtered 0.03 m grid from 26 acquisitions contained 28,261,980 voxel centers. Within the 5 cm radius, the daily mean number of reconstructed neighbors per center ranged from 1.14 to 2.44 across the fused cloud, and the maximum count was 18.

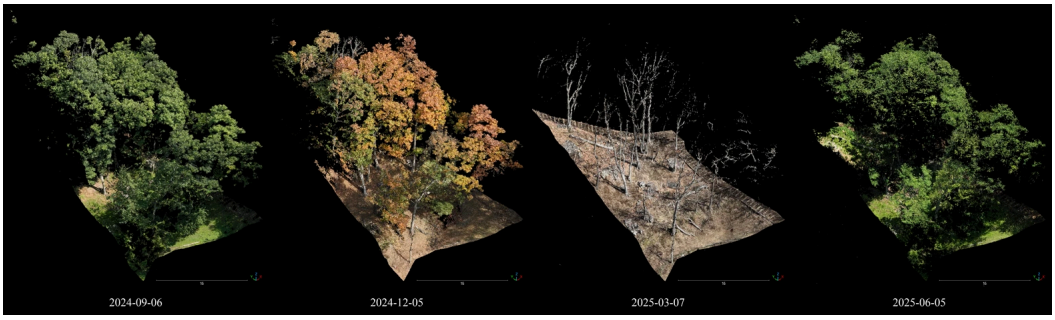

Figure 5. Appearance of point clouds generated in each season.

SMRF ground filtering classified 7,334,044 points as ground points and 20,927,936 points as non-ground points.

3.2. Reference Sample Data

Extraction of the reference sample data yielded 990,879 labeled voxel centers before subsampling. The number in each class is shown in Table 3.

Table 3. Number of labeled voxel centers by class.

Class ID	Class	Number of Points
0	Stem	25,985
1	Leaves of *Acer pictum* subsp. *dissectum* f. *dissectum*	49,736
2	Leaves of *Stachyurus praecox*	57,343
3	Leaves of *Magnolia kobus*	132,007
4	Leaves of *Quercus serrata*	383,550
5	Leaves of *Quercus acutissima*	342,258
Total		990,879

3.3. VED Time-Series Analysis

Figure 6 shows the date-wise means and standard deviations across voxel subsamples. Figure 7 shows the same date-wise means with and without smoothing. In the reference regions for *Acer pictum* subsp. *dissectum* f. *dissectum*, *Magnolia kobus*, *Quercus serrata*, and *Quercus acutissima*, mean VED decreased from autumn 2024 to winter 2025, was low around February–March, and increased from spring into summer. Mean VED in the Stem regions rose from autumn to winter, reached high values around February, and declined thereafter. The selected understory *Stachyurus praecox* region showed a different trajectory. The class-date means ranged from 0 to 8357 voxel centers m^{-3}.

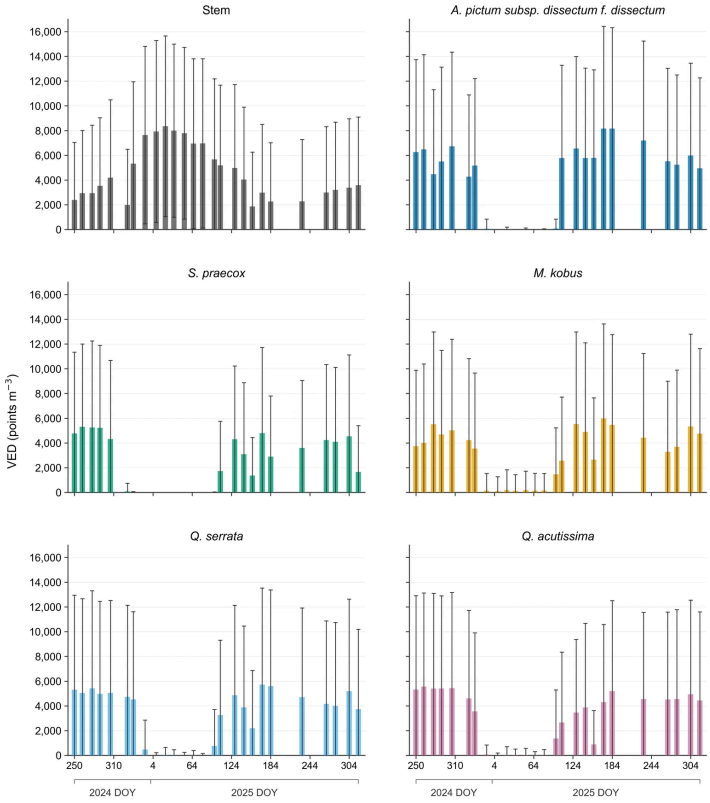

Figure 6. Date-wise VED from 6 September 2024 to 14 November 2025. Bar heights represent means, including zeros, and error bars represent standard deviations across the 25,000 sampled locations per class. A bar with zero height at an acquisition date represents a mean of zero. The spacing between bars follows the actual intervals between surveys.

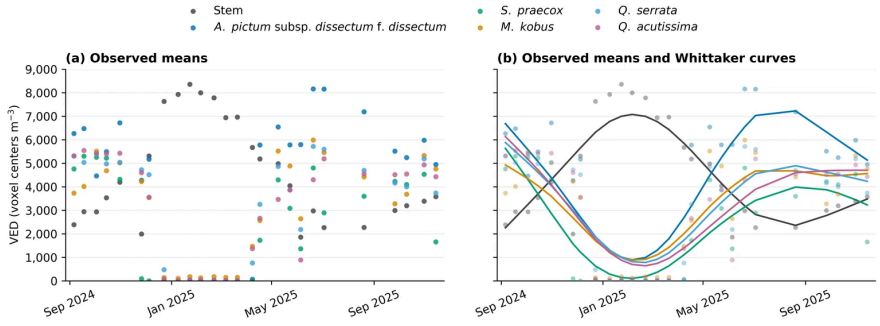

Figure 7. Date-wise mean VED for each class. Panel (**a**) shows the observed means. Panel (**b**) overlays Whittaker curves (λ = 80) fitted using the actual intervals between acquisitions. Each class-date mean includes zero values and represents 25,000 sampled locations.

The lowest holdout RMSE was 1097 voxel centers m^{-3} at λ = 1, compared with 1321 at λ = 80. Pearson correlations between the two fits were 0.916–0.957 across the six classes. At the acquisition dates, 18 of 156 fitted values at λ = 1 were negative, and all fits at λ = 80 were non-negative. Figure 7b uses λ = 80 to visualize broad seasonal trajectories.

Figure 8 shows the 26-date series for the four sampled *Quercus* trees (regions 1–4) and the two stem regions (8 and 9). Mean VED in all four *Quercus* foliage regions declined in winter and increased from spring onward. Pairwise correlations across the four unsmoothed foliage series ranged from 0.626 to 0.970. These patterns describe the sampled trees.

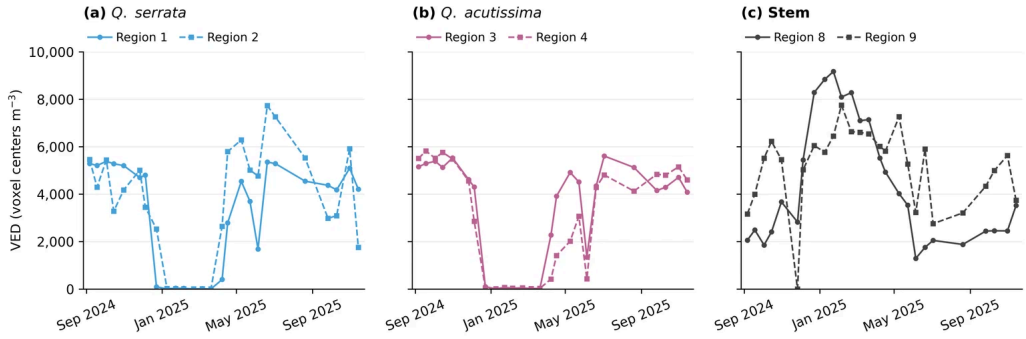

Figure 8. Individual reference-region mean VED across the 26 acquisition dates for the two *Quercus serrata* trees, two *Quercus acutissima* trees, and two stem regions. Numbers are the reference region IDs in Figure 4 and Table 2. Lines connect successive date-wise means.

The mean silhouette coefficient was 0.71 for the four *Quercus* foliage trajectories and 0.50 for the two stem trajectories. Relative to the stem group, the four foliage trajectories were more similar to one another. This comparison concerns the sampled region types and does not establish discrimination among oak species or among genera.

3.4. Acquisition and Processing Characteristics

Figure 9 shows the location of PlotWA within the reconstructed footprint on 3 July 2025. Across the 26 RGB surveys, there were 1109–1529 images per date, a ground sampling distance of 1.49–1.69 cm $pixel^{-1}$, and 2.87–5.10 million tie points (Table 4). Table 1 gives reconstructed point counts before clipping.

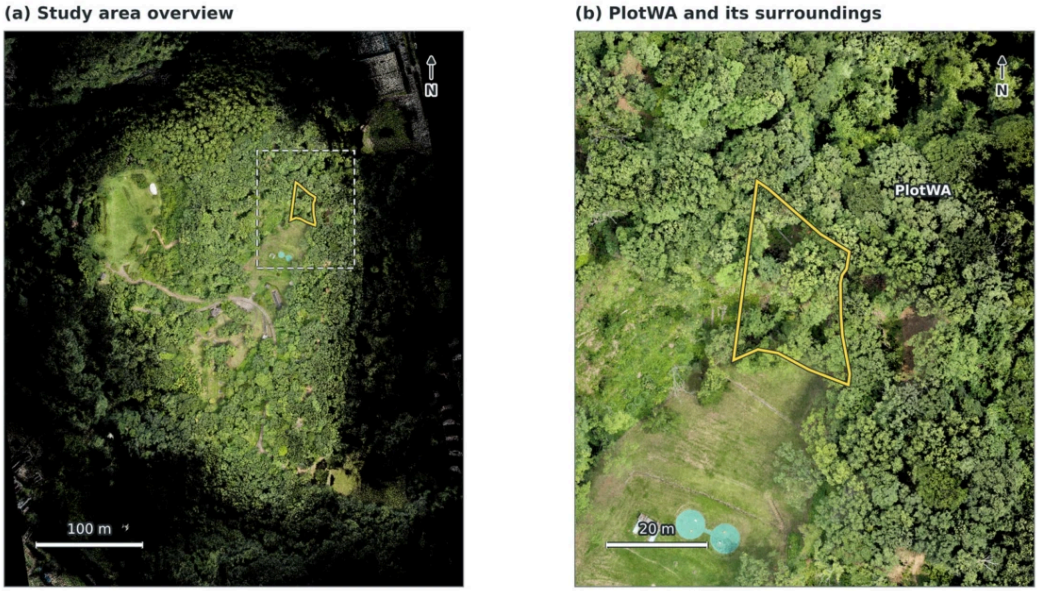

Figure 9. A plan view of the RGB point cloud acquired on 3 July 2025 and rasterized on a 2 cm grid. (**a**) Study area overview. (**b**) An enlarged view of PlotWA and its surroundings. The yellow outline marks PlotWA. The dashed rectangle in (**a**) indicates the extent of (**b**). Black indicates cells with no point data.

Table 4. RGB acquisition and SfM processing diagnostics by date.

Date	Images	GSD (cm Pixel^{-1})	Reported Flight Height (m)	Tie Points	Mean Tie-Point Projections
6 September 2024	1395	1.65	43.2	3,997,577	3.29
18 September 2024	1463	1.69	46.3	4,145,275	3.28
3 October 2024	1529	1.62	43.8	5,102,849	3.31
15 October 2024	1514	1.65	47.6	4,382,474	3.36
31 October 2024	1390	1.61	44.6	4,070,762	3.33
26 November 2024	1109	1.49	42.3	3,065,896	3.46
5 December 2024	1453	1.59	49.8	3,743,142	3.53
23 December 2024	1389	1.59	53.0	3,277,849	3.66
9 January 2025	1384	1.63	52.2	3,327,118	3.65
23 January 2025	1388	1.60	51.2	3,445,900	3.76
5 February 2025	1385	1.61	51.3	3,457,060	3.74
21 February 2025	1387	1.60	50.9	3,410,089	3.84
7 March 2025	1385	1.61	51.5	3,274,506	3.96
21 March 2025	1386	1.61	51.2	3,326,202	3.97
8 April 2025	1388	1.60	53.8	2,871,242	4.06
17 April 2025	1482	1.61	51.2	3,221,819	3.86
9 May 2025	1385	1.56	43.9	4,292,048	3.46
23 May 2025	1386	1.58	44.8	4,030,267	3.35
5 June 2025	1385	1.62	48.8	3,179,122	3.48
20 June 2025	1387	1.52	43.5	4,362,155	3.40
3 July 2025	1387	1.54	43.9	4,123,951	3.36
20 August 2025	1390	1.56	44.9	3,837,073	3.38
26 September 2025	1331	1.69	46.9	3,537,752	3.36

Date	Images	GSD (cm Pixel^{-1})	Reported Flight Height (m)	Tie Points	Mean Tie-Point Projections
10 October 2025	1388	1.62	45.4	3,708,209	3.39
31 October 2025	1390	1.59	43.6	4,163,822	3.51
14 November 2025	1396	1.59	45.8	3,645,905	3.37

Note. Values are from the Metashape RGB processing report for each date. GSD is ground sampling distance. Mean tie-point projections are the mean number of images in which a tie point was observed.

4. Discussion

4.1. Seasonal Changes in Three-Dimensional Reconstruction Captured by VED

Within the selected upper-canopy foliage regions, mean VED decreased from autumn to winter and rose from spring onward. The selected stem regions showed the reverse autumn-to-winter change. The understory *Stachyurus praecox* region followed a different pattern. These are changes in local SfM reconstruction at the selected reference locations.

Three-dimensional reconstruction using SfM depends on target visibility, and occlusion by upper vegetation has been reported to limit point cloud generation within crowns and in the understory [11,13]. Dietenberger et al. [14] used leaf-off UAV-SfM point clouds to detect stems and showed that leaf-on and leaf-off data provide complementary information on stems and crowns. The observed autumn-to-winter increase in stem VED is consistent with increased visibility after leaf fall. The contributions of canopy occlusion and acquisition conditions remain unresolved.

Zero VED indicates that no neighboring voxel center was reconstructed within the search radius. This can reflect vegetation absence, occlusion, or poor reconstruction. Inferring vegetation or phenological change therefore requires independent observations.

4.2. Implications for Observing Understory Vegetation

Because three-dimensional reconstruction using SfM requires targets to be visible in multiple images, occlusion by upper-canopy crowns limits point cloud generation for the ground surface, crown interiors, and understory vegetation [11–13]. Dietenberger et al. [14] combined leaf-on and leaf-off observations to extract stem positions and crown extents.

The selected *Stachyurus praecox* region showed a different temporal pattern from the selected upper-canopy foliage regions. Both changes in the shrub itself and changes in upper-canopy visibility could contribute to this pattern. The VED series alone does not resolve those contributions.

A potential application is to compare acquisition dates for surveying understory vegetation or stems. Independent observations would be needed to determine how closely VED changes track leaf-out, leaf fall, or canopy cover.

4.3. Relationship to Point Density, Voxel Occupancy, and Plant Area Density

Methods for quantifying the three-dimensional density of foliage and vegetation, including Leaf Area Density (LAD) and Plant Area Density (PAD), have already been proposed for forest measurements using LiDAR [5–8]. LAD represents leaf area per unit spatial volume, whereas PAD represents the area density of vegetation, including foliage and woody material; both are expressed in m^2 m^{-3}. These are physical quantities based on area and are distinct from vegetation volume occupancy and biomass density.

PAD estimation using terrestrial laser scanning (TLS) uses the number of laser

beams that actually reach each voxel, the number intercepted by vegetation within the voxel, and the distances traveled by the beams. Pimont et al. [7] evaluated bias corrections and uncertainty for estimators based on the Beer–Lambert law and free path lengths, accounting for the number of observed beams and the size of plant elements. Nguyen et al. [8] implemented these estimators in L-Vox and examined the effects of voxel size and scan configuration in virtual and real forests. These methods estimate area density using laser sampling of space and assumptions about plant element orientation, rather than directly treating the number of return points as density.

Local point density divides the number of raw points in the search sphere by its volume V [24]. VED counts each reconstructed voxel center once. Binary voxel occupancy records whether a voxel contains a point [25]. Wu et al. [26] calculated a local volume fraction within a radius of each voxel center. If n of K neighboring voxels are occupied, the local occupancy fraction is n/K and VED is n/V. Given the same occupancy criterion, grid, neighborhood, and treatment of the center, VED equals K/V times the local occupancy fraction.

LAD and PAD express plant surface area per unit volume. VED applies the established concept of voxel occupancy to compare seasonal changes in local SfM reconstruction across forest components at fixed locations on a common grid.

4.4. Limitations of VED and Future Work

Because VED uses SfM reconstruction, it is influenced by image acquisition and point cloud generation conditions in addition to vegetation structure. The reconstruction quality of SfM point clouds has been reported to depend on image overlap, spatial resolution, viewing angle, and illumination [11,13]. For example, differences in flight height, viewing direction, image overlap, illumination, or SfM processing settings may alter the density of the generated point cloud even when the vegetation state is unchanged. In Dietenberger et al. [14], viewing directions, image numbers, and point cloud quality settings also differed between leaf-on and leaf-off acquisitions; differences in point cloud distribution between those periods therefore cannot be attributed solely to foliage conditions. In the present study, temporal VED comparisons likewise require acquisition and processing conditions to be kept as consistent as possible, while considering the effects of differences in these conditions on reconstruction.

Tables 1 and 4 document survey-level acquisition and processing conditions. Local reconstruction can vary with illumination, wind, image overlap, and canopy visibility. Their effects on VED could not be separated in this study.

Misalignment between time-series point clouds can cause the same plant element to be assigned to different voxels, affecting VED. The GNSS uncertainties in Section 2.2 describe camera positions at acquisition. Inter-date point cloud misalignment and its effect on VED remain unquantified. Future work could assess these effects using stable surfaces or independent control data.

VED values also depend on voxel size and the search radius used to calculate local density. A sensitivity check using 41 acquisition dates from the same plot compared voxel sizes of 3, 5, and 7 cm and search radii of 5, 8.5, and 12 cm. For the seven nonzero alternatives, Pearson correlations with the unsmoothed 3 cm/5 cm class means ranged from 0.968 to 1.000 across the six classes. Absolute values varied. The 7 cm grid with a 5 cm search radius yielded zero VED because the radius was smaller than the grid spacing.

Nearby voxel subsamples may share reconstructed neighbors, so their variability reflects within-region heterogeneity. With one or two trees per foliage class, broader species-level conclusions require more biological replicates. Future work could test

whether VED and spectral data jointly aid tree species classification.

Future work could separately report the frequency of dates with reconstructed neighbors and the mean VED on those dates.

5. Conclusions

We integrated multitemporal UAV-SfM point clouds into a common three-dimensional voxel space and compared VED over approximately 14 months. VED for selected upper-canopy foliage decreased from autumn to winter and increased from spring onward, whereas the stem regions showed the opposite autumn-to-winter change. The understory shrub showed a different temporal pattern.

These results indicate that VED can quantify changes in three-dimensional SfM reconstruction at fixed locations. Independent observations could test how closely these changes track seasonal visibility.

VED may be influenced by acquisition conditions, SfM processing settings, point cloud registration accuracy, voxel size, and search radius. Future work should examine these effects across more trees, forest types, and observation conditions.

Author Contributions: Conceptualization, K.O., H.H., A.S. and S.S.; methodology, K.O., H.H. and S.S.; investigation, K.O.; writing—original draft preparation, K.O.; writing—review and editing, S.S.; supervision, S.S.; funding acquisition, K.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by internal funding from Tokyo University of Agriculture through its graduate student support program in 2026.

Data Availability Statement: The authors plan to make the VED analysis code available on GitHub. The source point cloud dataset includes multispectral information reserved for ongoing research and is currently restricted.

Acknowledgments: This study was also conducted with the support of Kanagawa Prefectural Nanasawa Forest Park. The authors would like to express their sincere gratitude to the park director, Takahashi, and to all staff members for their generous cooperation and assistance.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Fassnacht, F.E.; Latifi, H.; Stereńczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [Google Scholar] [CrossRef]
2. Guimarães, N.; Pádua, L.; Marques, P.; Silva, N.; Peres, E.; Sousa, J. Forestry Remote Sensing from Unmanned Aerial Vehicles: A Review Focusing on the Data, Processing and Potentialities. *Remote Sens.* **2020**, *12*, 1046. [Google Scholar] [CrossRef]
3. Yel, S.; Tunc Gormus, E. Exploiting hyperspectral and multispectral images in the detection of tree species: A review. *Front. Remote Sens.* **2023**, *4*, 1136289. [Google Scholar] [CrossRef]
4. Xu, Z.; Shen, X.; Cao, L.; Coops, N.C.; Goodbody, T.R.H.; Zhong, T.; Zhao, W.; Sun, Q.; Ba, S.; Zhang, Z.; et al. Tree species classification using UAS-based digital aerial photogrammetry point clouds and multispectral imageries in subtropical natural forests. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *92*, 102173. [Google Scholar] [CrossRef]
5. Hosoi, F.; Omasa, K. Estimation of Leaf Area Density Profiles of Japanese Maple and Camellia Woody Canopies Using Portable Scanning Lidars. *Eco-Engineering* **2011**, *23*, 105–109. [Google Scholar] [CrossRef]
6. Hosoi, F.; Nakai, Y.; Omasa, K. Voxel tree modeling for estimating leaf area density and woody material volume using 3-D LIDAR data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *II-5/W2*, 115–120. [Google Scholar] [CrossRef]
7. Pimont, F.; Allard, D.; Soma, M.; Dupuy, J.-L. Estimators and confidence intervals for plant area density at voxel scale with T-LiDAR. *Remote Sens. Environ.* **2018**, *215*, 343–370. [Google Scholar] [CrossRef]
8. Nguyen, V.-T.; Fournier, R.; Côté, J.-F.; Pimont, F. Estimation of vertical plant area density from single return terrestrial laser scanning point clouds acquired in forest environments. *Remote Sens. Environ.* **2022**, *279*, 113115. [Google Scholar] [CrossRef]

9. Hill, R.; Broughton, R. Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: A case study in lowland Britain. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 223–233. [Google Scholar] [CrossRef]
10. Davison, S.; Donoghue, D.; Galiatsatos, N. The effect of leaf-on and leaf-off forest canopy conditions on LiDAR derived estimations of forest structural diversity. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *92*, 102160. [Google Scholar] [CrossRef]
11. Iglhaut, J.; Cabo, C.; Puliti, S.; Piermattei, L.; O'Connor, J.; Rosette, J. Structure from Motion Photogrammetry in Forestry: A Review. *Curr. For. Rep.* **2019**, *5*, 155–168. [Google Scholar] [CrossRef]
12. Wallace, L.; Lucieer, A.; Malenovský, Z.; Turner, D.; Vopěnka, P. Assessment of Forest Structure Using Two UAV Techniques: A Comparison of Airborne Laser Scanning and Structure from Motion (SfM) Point Clouds. *Forests* **2016**, *7*, 62. [Google Scholar] [CrossRef]
13. Frey, J.; Kovach, K.; Stemmler, S.; Koch, B. UAV Photogrammetry of Forests as a Vulnerable Process. A Sensitivity Analysis for a Structure from Motion RGB-Image Pipeline. *Remote Sens.* **2018**, *10*, 912. [Google Scholar] [CrossRef]
14. Dietenberger, S.; Mueller, M.M.; Bachmann, F.; Nestler, M.; Ziemer, J.; Metz, F.; Heidenreich, M.G.; Koebsch, F.; Hese, S.; Dubois, C.; et al. Tree Stem Detection and Crown Delineation in a Structurally Diverse Deciduous Forest Combining Leaf-On and Leaf-Off UAV-SfM Data. *Remote Sens.* **2023**, *15*, 4366. [Google Scholar] [CrossRef]
15. Atsugi Civil Engineering Office, Kanagawa Prefecture. Nanasawa Forest Park Development and Management Plan. 2024. Available online: https://www.pref.kanagawa.jp/documents/100963/nanasawashinrinkouen-seibikanrikeikaku.pdf (accessed on 2 October 2026). (In Japanese) [Google Scholar]
16. Japan Meteorological Agency. Ebina (Kanagawa Prefecture), Climate Normals (Annual and Monthly Values), Main Elements (1991–2020). 2021. Available online: https://www.data.jma.go.jp/stats/etrn/view/nml_amd_ym.php?block_no=0388&prec_no=46 (accessed on 2 October 2026). (In Japanese) [Google Scholar]
17. DJI. DJI Mavic 3M Specifications. Available online: https://enterprise.dji.com/mavic-3-m/specs (accessed on 26 August 2026).
18. Westoby, M.; Brasington, J.; Glasser, N.; Hambrey, M.; Reynolds, J. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314. [Google Scholar] [CrossRef]
19. Pingel, T.; Clarke, K.; McBride, W. An improved simple morphological filter for the terrain classification of airborne LIDAR data. *ISPRS J. Photogramm. Remote Sens.* **2013**, *77*, 21–30. [Google Scholar] [CrossRef]
20. Klápště, P.; Fogl, M.; Barták, V.; Gdulová, K.; Urban, R.; Moudrý, V. Sensitivity analysis of parameters and contrasting performance of ground filtering algorithms with UAV photogrammetry-based and LiDAR point clouds. *Int. J. Digit. Earth* **2020**, *13*, 1672–1694. [Google Scholar] [CrossRef]
21. Eilers, P. A Perfect Smoother. *Anal. Chem.* **2003**, *75*, 3631–3636. [Google Scholar] [CrossRef]
22. Garcia, D. Robust smoothing of gridded data in one and higher dimensions with missing values. *Comput. Stat. Data Anal.* **2010**, *54*, 1167–1178. [Google Scholar] [CrossRef]
23. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [Google Scholar] [CrossRef]
24. CloudCompare. Density. Available online: https://www.cloudcompare.org/doc/wiki/index.php/Density (accessed on 16 November 2026).
25. Cimdins, R.; Yrttimaa, T.; Vastaranta, M.; Kankare, V. Assessing forest structural complexity: Insights from alternative laser scanning approaches. *Scand. J. For. Res.* **2026**, *41*, 40–53. [Google Scholar] [CrossRef]
26. Wu, J.; Aage, N.; Westermann, R.; Sigmund, O. Infill Optimization for Additive Manufacturing—Approaching Bone-Like Porous Structures. *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 1127–1140. [Google Scholar] [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.