

# Overview and Benchmark on Multi-Modal Lidar Point Cloud Registration for Forest Applications

Mariana Batista Campos<sup>1</sup>, Leticia Ferrari Castanheiro<sup>2</sup>, Dipal Shah<sup>1</sup>, Yunsheng Wang<sup>1</sup>, Antero Kukko<sup>1</sup>, Eetu Puttonen<sup>1</sup>

<sup>1</sup>Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute, National Land Survey of Finland, 02150 Espoo, Finland; (mariana.campos, yunsheng.wang, dipal.shah, antero.kukko, eetu.puttonen)@nls.fi

<sup>2</sup>Department of Cartography, São Paulo State University (UNESP), 19060-900, Presidente Prudente, São Paulo, Brazil; leticia.ferrari@unesp.br

**Keywords:** feature-based methods, cross-platform LiDAR, stem matching, coarse registration, ALS-TLS

## Abstract

Light Detection and Ranging (LIDAR) is widely acknowledged as a robust tool for monitoring forest structure, dynamics, and changes. To achieve a high-complete forest structural model, LiDAR data acquisition from both aerial (above-canopy) and terrestrial (below-canopy) platforms is commonplace. Consequently, in such multi-modal LiDAR cases, robust data registration is required for accurate forest analysis, such as biomass and canopy growth. Yet, multi-modal LiDAR registration remains a significant challenge due to differences in observation perspectives, deficient data overlap, and often inhomogeneity in point distributions and densities. The challenge increases in complex forest environments due to the abundance of unstable features (e.g., leaves) and occlusions. Thus, the dynamic nature of forest scenes needs to be considered when applying registration methods on forest point clouds. In this paper, we overview the latest advancements in registering forest point clouds from multi-modal data acquisitions, aiming to discuss the strengths and weaknesses of the most used LiDAR registration methods for forest applications. To support our investigations, we benchmark two multi-modal registration methods especially designed for forest mapping against traditional global and feature-based approaches. Experiment assessments were conducted using two point clouds acquired from a permanent laser scanning and airborne laser scanning systems at a boreal forest plot.

## 1. Introduction

Point cloud registration is widely known as a method of applying a rigid transformation to align two or more point clouds acquired from different positions, platforms or times (Vosselman and Maas, 2010), particularly when direct georeferencing is either not possible or unsuccessful. As an essential step in many LiDAR applications, the photogrammetry and computer vision communities have a long tradition of developing methods for point cloud registration that do not require artificial reference targets (e.g., spheres) placed on the mapped area. These include LiDAR data registration between point clouds acquired from multiple perspectives, multi-temporal instances, and/or varied modalities or platforms, such as airborne laser scanning (ALS) combined with terrestrial laser scanning (TLS) or ground-based mobile laser scanning (MLS).

Over the last five years, significant efforts have been made to review and benchmark LiDAR point cloud registration methods (Cheng et al., 2018; Dong et al., 2020; Zhang et al., 2020; Huang et al., 2021; Si et al., 2022; Monji-Azad et al., 2023; Huang et al., 2023), particularly with the rapid development in learning approaches. For instance, Cheng et al. (2018) reviewed the main feature-based coarse and fine registration methods applied for LiDAR point clouds. These authors emphasised the need to assess the sensitivity, robustness, and accuracy of these methods across diverse and complex data environments. In this direction, Dong et al. (2020) provided a TLS benchmark dataset from 11 different environments, including complex scenarios such as forests. These authors highlighted the improved performance of learning-based methods for solving TLS point cloud registration problems, especially in small-scale indoor point cloud registration (Dong et al., 2020). However, as future direction, they also point out the need for further developments regarding outdoor and irregular LiDAR point clouds. Still according to Dong et al. (2020), the dynamics of the scene and changes over time need to be considered when applying registration methods, which significantly intensifies the challenge of aligning two point cloud datasets. Zhang et al. (2020) and Monji-Azad et al. (2023) conducted reviews of learning-based 3D point cloud registration

methods, highlighting that feature extraction and matching remain challenging in scenes lacking stable features and prior information, such as in GNSS-denied environments. Thus, a common remark among these reviews is that developing generalized registration methods for complex environments (e.g. dynamic objects, unstructured, GNSS-denied) still poses a significant challenge (Si et al., 2022).

Most LiDAR point cloud registration methods were not developed or assessed in complex environments with unstable features. For instance, forest and agricultural areas pose greater challenges for point cloud registration due to the intricacy and unstable nature of their features (Castanheiro et al., 2023) and the lack of external information. Most forest point clouds have multiple pulse echoes returned from leaves and underground vegetation. As a result, LiDAR point clouds obtained from forests often exhibit increased noise level. External information, such as sensor pose from GNSS positioning solutions, can significantly assist in feature selection and outlier removal. However, under canopy, GNSS signal is typically weak or even absent over substantial periods of time, hampering implementation of automated tools. For instance, TLS and MLS forest point clouds are typically acquired in a local coordinate system to avoid disparities resulting from poor GNSS circumstances. Transforming such data into a georeferenced system requires additional effort.

The challenge becomes more pronounced when conducting multi-modal point cloud registration, as the scene is scanned from different perspectives and often with inhomogeneous point distributions (Lin et al., 2022). Wide-baseline observations often suffer from deficient overlap between observations and self-similar structures, such as tree stems. These factors can lead to matching ambiguities during registration. Therefore, current multi-modal point cloud registration usually necessitates substantial interactive effort.

Here, we aim to provide a better understanding of the strengths and weaknesses of the most used LiDAR registration methods for forest point cloud registration to decrease uncertainties in modelling tree structure. Laser scanning technology is well

acknowledged as a robust tool for monitoring forests. LiDAR datasets enable 3D data collection and representation of tree structures due to the penetrability of the laser beam through the canopy. A high-complexity forest structural model requires LiDAR data acquisitions from aerial (above-canopy) and terrestrial (below-canopy) platforms. Such integration of multiple data acquisition geometries can be beneficial for complex forest analyses, such as aboveground biomass estimation and canopy space occupation. Therefore, we focus particularly on the registration of forest point clouds at plot level acquired with multi-modal means (e.g., ALS-TLS and ALS-MLS).

We review the latest advancements in registering multi-modal forest point clouds (Hauglin et al., 2014; Polewski et al., 2016; Polewski et al., 2019; Guan et al., 2020; Dai et al., 2019; Liu et al., 2021; Hyypä et al., 2021; Shao et al., 2022; Dai et al., 2022; Pohjavirta et al., 2022; Huang et al., 2023; Chen et al., 2024; Ghorbani et al., 2024), with an aim of investigate the future directions of global, feature-based or deep-learning registration methods for forest applications. To support our discussion, we benchmark multi-modal registration methods, especially design for forest mapping (Hyypä et al., 2021; Pohjavirta et al., 2022) against traditional global (iterative closest point - ICP, normal distribution transform - NDT, coherent point drift - CPD) and feature-based (signature of histogram of orientations - SHOT and fast point feature histogram - FPFH) registration approaches, using two point clouds acquired with permanent laser scanning (PLS) and ALS-helicopter systems.

## 2. Multi-modal Forest Point Cloud Registration: Review

Overall, the majority of registration methods utilize a coarse-to-fine strategy. Coarse registration algorithms, such as feature-based or global methods, aim to estimate the three orientations and three translation parameters (6 degrees of freedom - DoF) between one point cloud and another (Cheng et al., 2018). This is accomplished, for instance, by identifying common features (e.g., points, lines, planes) or point-to-point distance between nearest neighbours (e.g. ICP method) in the point clouds, which are then used to estimate the 6 DoF parameter transformation (Mikhail, 1976). In this review, we focus specifically on describing extracted features, matching approaches, and registration accuracy achieved by particular point cloud registration methods designed specifically for forest applications.

Multi-modal point cloud feature detection and matching, especially for forest applications, need to be robust against rotation, scale, point cloud density, partial occlusions and unstable features (outliers). As previously mentioned, due to the abundance of unstable features in complex forest environments, automatic point-level registration based on 3D keypoints often fails (Castanheiro et al., 2023). As a potential solution to address this challenge, previous works suggested the removal of dynamic objects (e.g. leaves), focusing on aligning two or more LiDAR datasets based on elementary forest structures, such as tree stem location (Liang and Hyypä, 2013; Hyypä, 2021; Ghorbani et al., 2024), ground points and canopy height (Liu et al., 2017) and shape (Dai et al., 2019; Shao et al., 2022).

Stems location is the most explored feature for forest point cloud registration. Tree stems as features for feature-based matching and forest point cloud registration were initially introduced to register LiDAR point clouds obtained from the same platform but at different locations and times (Liang and Hyypä 2013). Approaches utilizing stem locations as features for multi-modal feature-based matching and registration were proposed by Hauglin et al. (2014), Polewski et al. (2016), Polewski et al.

(2019), Guan et al.(2020), Hyypä et al. (2021) and Ghorbani et al. (2024).

Hauglin et al. (2014) proposed a TLS-ALS point cloud co-registration by using individual tree positions (planimetric coordinates) and tree relative size (e.g. using diameter at the breast height - DBH and height) detected on both ALS and TLS data. In this study, both ALS and TLS datasets were georeferenced, or rather, external GNSS information was expected as an initial input. Considering a stem position detected in a TLS point cloud, a circular search area is defined to find the best match in a corresponding ALS-derived tree map. The search space are defined according to the expected error in the initial estimated position. The best match is obtained considering the minimum distance between stems and the closest match with estimated relative tree size. Tree size proved important as TLS detects small trees frequently occluded in ALS point clouds. Hence, the smallest TLS-detected trees can be discarded, as finding a match among ALS-detected trees for these small trees is unlikely. After TLS-ALS stem detection and matching, a least-squares method to estimate the optimal translation and rotation matrix between two equally sized 2D stem datasets was applied. The method was evaluated in a boreal forest, achieving co-registered positioning accuracy between 0.5 and 1 m. These authors highlight the need for robust methods to identify and handle erroneous stem matches.

Aiming to improve stem matching, Polewski et al. (2016) introduced a similarity stem descriptor designed based on the planimetric and vertical distances between the target stem and other stem centres. This approach requires the ALS point cloud in a georeferenced coordinate system and the terrestrial point cloud in a local coordinate system, both with preserved object scale, as inputs. Considering that planimetric and vertical distances between trees are used as stem descriptors, TLS and ALS point clouds required same Z orientation. Tree stem orientations (principal axis) in TLS point clouds were computed to ensure alignment with the ALS reference plane. The similarities between all pairs of descriptors from the terrestrial-aerial point clouds are computed. Subsequently standard graph maximum matching technique is then employed to determine corresponding stem pairs. Finally, the matched stem positions are used to estimate the rigid transformation parameters that map the terrestrial point cloud to the ALS georeference. This method was evaluated by registering terrestrial photogrammetric and ALS point clouds acquired in the silvicultural stands of Douglas-fir and vine maple. An average 2D position accuracy of 66 cm was achieved, in which, according to the authors, ALS tree center estimation was the main source of error affecting the registration results. The registration average accuracy obtained by Polewski et al. (2016) aligns with the results reported by Hauglin et al. (2014). However, this method represents an advancement from Hauglin et al. (2014) as it does not require initial external information for the terrestrial point cloud or additional tree attributes, such as tree height or DBH). Previous registration methods between terrestrial and aerial datasets, which also rely on georeferenced tree locations, for instance, obtained from GNSS positioning, were presented by Lindberg et al. (2012) and Paris et al. (2017). Recent works have also explored combining tree location and parameters, such as DBH, as features for MLS-ALS stem matches, as demonstrated by Olofsson and Holmgren (2022).

Polewski et al. (2019) enhanced the approach originally presented by Polewski et al. (2016) by introducing a scale term into the 2D registration transformation and employing a weighted bipartite graph for stem descriptor matching. This method was evaluated by registering a terrestrial MLS point cloud acquired

with a backpack system with an unmanned aerial vehicle (UAV) point cloud. The dataset was collected in planted temperate forest plots (Jiangsu, China), featuring both coniferous and broadleaf trees. These authors discuss that predominant broadleaf plots can be more challenging than coniferous ones. For instance, more matched trees were obtained in coniferous plots, resulting in higher registration position accuracy (27–36 cm) compared to that achieved in broadleaf plots (54–67 cm).

Hyypä et al. (2021) proposed a 2D coarse registration method for forest point clouds using translation- and rotation-invariant local descriptors computed based on tree locations. The feature descriptors for each of the trees describe the relative locations of the neighbouring objects. The descriptors are constructed as follows: first, by detecting the closest neighbouring tree within the same point cloud. Subsequently, the tree's neighbourhood (xy plane around the tree) within a certain radius (e.g. 10 m) is divided into four quadrants. The descriptor comprises the closest neighbouring tree in each quadrant and the angle between them. Utilizing the distances and angles between the closest trees in each quadrant, the rotation- and translation-invariant feature descriptor is generated. Note that in this method, it is also expected that the Z axes of both point clouds (e.g. TLS and ALS) are approximately in the same direction. Feature descriptors from both terrestrial and aerial point clouds are compared using the Euclidean distance in the feature space as the similarity criterion. Once matched, an optimal rigid-body 2D transformation between the two point clouds is estimated. The authors aim to enhance the approach presented by Polewski et al. (2019) in terms of computational time (quadratic time complexity). Therefore, the proposed feature descriptor vector relies only on the immediate local neighbourhood of each tree, rather than the entire plot as designed by Polewski et al. (2019). The method was evaluated using simulated and real datasets acquired in a boreal forest scenario. Tree stems and crown tops were utilized as objects to construct the descriptors, and the matching was assessed between stems (TLSstem-to-ALSstems) and between stems and tops (TLSstem-to-ALStop). It is important to highlight that despite the method being evaluated using tree locations as objects, it is not limited to only those features. Additional tree attributes can be incorporated into the proposed feature descriptor. An overall root means square error (RMSE) of 28.6 cm was achieved for the matching tree pairs using the stem-to-top method, while the corresponding RMSE for the stem-to-stem method was below 10 cm post-registration. The simulated dataset study demonstrates that the method reliably performs in the presence of moderate tree location errors but exhibits sensitivity to tree omission. Results deteriorate when more than 10% of the trees are missing. The authors also emphasize the need for additional tests to evaluate the effectiveness of the proposed registration method in more complex forests. These may include scenarios involving a large number of young trees, often occluded in ALS data acquisition, or a significant proportion of broadleaved trees that may intensify the challenge of tree matching, as reported by Polewski et al. (2019).

The mentioned previous works assume equal Z axes orientation in both coordinate systems, focusing on calculating in-plane rotation angle and 2D translation based on feature descriptors generated from tree position information and its surrounds. Therefore, oblique TLS point clouds (e.g., Campos et al., 2021) or terrestrial photogrammetric point clouds require prior Z orientation estimation based, for example, on the stem direction and a digital terrain model (DTM). Additionally, these works also assume that individual tree locations were identified using previously developed stem detection algorithms. Thus a key assumption is that both captured scenes contain sufficiently

common trees, and that the densities of both point clouds are high enough to reliably extract tree stems or crown tops.

Guan et al. (2020) and Ghorbani et al. (2024) proposed a TLS-ALS registration based on 3D tree location (X, Y, Z), in which initial Z orientation estimation is not needed (3D rigid transformation). Guan et al. (2020) introduced a triangulated irregular network (TIN) matching approach, in which the framework's input is the pre-extracting individual tree positions. Subsequently, a TIN is generated for each tree, considering its position and the locations of the neighbouring trees. The neighbouring tree locations, found using the k-nearest neighbour search method, are then input for a Delaunay triangulation. As performed by Polewski et al. (2019) and Hyypä et al. (2021), the goal was to identify a spatial pattern of tree distributions for each tree. The TIN matching is conducted using a voting strategy, which counts the number of similar triangles between two TINs and iteratively finds the best match. The individual TIN pattern of each tree is highly sensitive to tree neighbourhood omissions, which can result in insufficient number of matches or false positives. Thus, robust tree detection is needed. To minimize the false positives, matched tree pairs are further filtered and optimized by the RANSAC algorithm. The selected matched tree locations are used to estimate 3D rigid body transformation parameters. Further improvements are achieved by applying the ICP algorithm in a fine-registration step. The method is assessed in three plots acquired at a coniferous dominant planted forest achieving an average MLS (backpack system)-UAV point cloud registration accuracy of 30 cm in planimetry and 20 cm in altimetry

Similarly, Ghorbani et al. (2024) introduced a method for registering TLS and ALS forest point clouds, also relying on 3D individual tree location correspondence and a 3D body rigid transformation. However, the authors aim to advance by reducing the reliance on the accuracy and completeness of individual tree locations during point cloud registration, addressing a limitation highlighted by Polewski et al. (2016), Guan et al. (2020) and Hyypä et al. (2021). In this regard, a filtering approach is introduced to remove positions unlikely to have corresponding matches between TLS and ALS datasets (e.g. suppressed trees). Locations of small tree detected in the TLS data are removed from the dataset according to DBH information, specially targeting larger trees in the TLS data. The filtered tree locations from both TLS and ALS datasets serve as input for the proposed algorithm. As an initial step, a search space in the ALS data is defined by using a local neighbourhood triangulation (TIN), similar to that proposed by Guan et al. (2020). The matching between TLS-ALS stem locations is performed by correlating a set of three distances between stems in the ALS data with the distance constraints observed in the TLS data. The matching process is iterative, in which the accuracy of the registration performed with the selected set of matches is used as criteria. The method was assessed using six plots collected in Vienna Austria. A registration accuracy of 55 cm was obtained in the best scenario, not exceeding 1 m in all other plots. Comparative analysis with Hyypä et al. (2021), showed improved registration accuracy at the expense of increased computational time and additional tree information (DBH).

As UAV technology advances, high-density above-canopy point clouds with visible tree stems are increasingly common. Those methods based on stem location have proved to be a consistent solution for a combination of ground (TLS or MLS) and aerial (UAV, ALS) data acquisition on temperate and boreal forests, especially coniferous dominant ones. However, none of these methods have been evaluated in complex forest scenarios where stems are frequently omitted. As highlighted by Castanheiro et

al. (2023), tree stems may not always be visible in dense forest areas or specific agricultural environments (e.g. orange fields). Consequently, there is a need for further advancements in feature-based and learning methods to address this limitation. Recent studies propose also to explore the use of additional features, such as ground points (Pohjavirta et al., 2022), canopy attributes (Liu et al., 2021; Shao et al. 2022; Zhou et al., 2023) or 3D keypoints (Dai et al., 2022; Zhang et al., 2021 and Chen et al., 2024)

Pohjavirta et al. (2022) integrated three types of features: stem positions, stem points, and ground points, in aiming to improve wide-baseline registration using a two-step registration approach. The objective of the study was to address the challenge of limited overlap between point clouds, which is a common problem, for instance, in the registration of terrestrial and aerial point clouds. The extracted feature points from tree stem and ground are used as input in a planar registration (2D). The shift in the Z direction was further determined by aligning target and reference DTMs. Subsequently, tree stem and ground point detection is redone and matches are refined for a fine-segmentation step using ICP algorithm. The method was assessed in boreal forest plots located at Evo, Finland. The proposed method achieved a 3D registration accuracy between TLS-UAV ranging between 7.2 and 13.6 cm, according to the plot complexity.

Exploring canopy attributes, Liu et al. (2021) propose a TLS-UAV point-cloud registration method utilizing crown centre position and height derived from a canopy height model (CHM). Point correspondences are established by considering the distance between two tree tops detected in the CHM and their height similarities. Rotation and translation between TLS-UAV point-clouds are then calculated through singular value decomposition. The point cloud registration is further refined using the ICP algorithm. This method was applied to alpine forest land, achieving a reported average accuracy of 43 cm.

Tree crown and canopy gap shapes are used for registering TLS and ALS point clouds by Shao et al. (2022) and Zhou et al. (2023), respectively. Shao et al. (2022) propose a method for registering TLS and ALS forest point clouds by detecting keypoints in canopy shape edges for subsequent image matching. Initially, tree canopies are projected onto the 2D plane, followed by an image-processing pipeline that includes canopy alignment (with Z parallel to the XY plane), canopy binary image generation, edge detection, keypoint extraction using an adaptation of Harris approach (Harris and Stephens, 1988), and keypoint matching. The method was evaluated in a subtropical forest, achieving a coarse alignment accuracy of 20 cm. A final TLS-ALS registration accuracy of 15 cm was obtained after applying a fine-segmentation using the ICP algorithm. Similarly, Zhou et al. (2023) evaluated a TLS-ALS point clouds registration method utilizing canopy gap shapes. First, canopy gap boundaries are extracted from CHM, followed by obtaining feature points from the canopy gap vectors using the weighted effective area algorithm. Coarse registration transformation parameters were obtained using the CPD algorithm. The ICP algorithm was applied for fine registration. Meter level accuracy was obtained on coarse registration, which was subsequently significantly improved by ICP algorithm (~ 15 cm). These methods are stem-independent and computationally efficient, as they reduce 3D point clouds to 2D space. However, TLS-ALS different perspective and image processing threshold, such as binarization can affect the shape of tree crown and canopy gaps.

As most registration methods rely on initial point cloud segmentation into stems and canopy, the question that remains is whether those features are the only viable option? Traditional feature-based matching, such as FPFH and BSC (Binary Shape

Context), were explored by Zhang et al. (2021), Chen et al. (2024) and Dai et al. (2022).

As an innovative combination of stem detection and classical 3D feature-based matching, Dai et al. (2022) proposed a TLS-UAV coarse registration method by semantically guiding keypoint detection based on previously classified point clouds of wood and leaves. In this approach, points classified as wood material in both TLS and UAV point clouds are utilized for keypoint detection and feature-based matching using BSC (Dong et al., 2017). Subsequently, outliers from the initial correspondence are eliminated using RANSAC. The reported accuracy of 29 cm was achieved for UAV-TLS datasets collected in a coniferous dominant forest with understory vegetation in Guangxi, China.

Zhang et al. (2021) employed FPFH method for the initial alignment of TLS plots and a UAV point cloud, followed by fine-registration using the ICP algorithm and a graph-based global adjustment method. The initial Coarse registration using FPFH provided registration accuracy at the meter level (< 2 m). However, subsequent fine-registration and global adjustment steps resulted in a reported relative accuracy between TLS-UAV clouds around 5 cm.

Chen et al. (2024) introduced a novel approach that combines hierarchical clustering and the FPFH algorithm. Initially, point clouds are normalized for height. Multi-layer tree maps for both ULS and TLS data were established by segmenting the point cloud into height segments and employing hierarchical clustering (DBSCAN). Clusters were defined based on two predetermined parameters: radius and minimum cluster size. Subsequently, FPFH features were extracted for each cluster obtained at different layers (height). Feature matching and transformation estimation were conducted using nearest-neighbour search and least-squares method. A transformation matrix for each cluster was estimated. The matrix with the best matching score is selected for coarse registration. ICP algorithm is subsequently applied for fine segmentation. Similar to Shao et al. (2022), the method was assessed in a subtropical humid forest ecoregion in southern China, achieving a RMSE in the registration accuracy of 15 cm.

As final remarks, the coarse registration accuracy between aerial and terrestrial point clouds typically falls within the range of 10 to 50 cm. Regarding feature-based methods, coarse registration approaches relying on stem locations as features have demonstrated robust performance in boreal forests (RMSE < 15 cm). However, they are typically sensitive to the accuracy of tree position estimation and particularly susceptible to errors due to tree omissions. Automatically detecting young tree stems or partial-occluded stems due to understory vegetation is still a challenge. Alternatively, approaches that integrate established 3D keypoint detection and feature descriptors (e.g., Harris, FPFH, and BSC) within a specified search space, such as canopy edges, hierarchical clusters, or wood-leaf classifications, achieved in overall coarse registration accuracies exceeding 20 cm during the coarse registration step. However, those methods were also evaluated in more complex forest scenarios, proving to be a promising alternative for areas where stem information is not available. Regarding global-based methods, CPD and ICP emerge as the most applied approaches, particularly for fine-registration. The reported results indicate the current robustness of global implementations, especially in ALS-TLS forest datasets with high overlap, highlighting the possibility of further exploring the use of global based registration methods in forest scenarios. Regarding learning methods, no specific approach designed for the registration of aerial and terrestrial forest point clouds was identified by the authors during this overview. According to Monji-Azad et al. (2023), deep learning-based

registration methods still struggle to achieve acceptable results in real datasets, especially in forest environments. However, deep-learning methods have the potential to bring new perspectives and future advancements in point cloud registration. For example, they could explore more complex characteristics of forest components, like stems or leaves. In Section 3, we benchmarked feature-based and global methods, aiming to enhance the discussion towards these different directions.

### 3. Multi-modal Forest Point Cloud Registration: Benchmark

#### 3.1 Datasets

The benchmarking was performed using PLS and ALS point clouds obtained from a coniferous-dominant boreal forest situated at the Hyytiälä forestry field station in southern Finland (61°51'N, 24°17'E). The terrestrial LiDAR data was acquired by a PLS station, named Lidar Phenology station (LiPhe) (Campos et al., 2021). LiPhe was specifically designed for continuous monitoring of a fixed forest scene by using a time-of-flight Riegl VZ-2000i scanner (RIEGL Measurement Systems, Horn, Austria) installed at 30-meter height in an observation tower (Figure 1.a). The forest scene was scanned by a 1550 nm laser wavelength with a fine angular resolution of 0.006 degrees and a scan frequency of 1200 kHz. More technical details about LiPhe can be found at Campos et al. (2021). LiPhe point clouds were collected in a local coordinate system, with the scanner position as the origin. Since most of the methods assume equal Z axes orientation in both TLS and ALS coordinate systems. These point clouds pose a particular challenge due to their oblique point of view and non-uniform density across the scanned area. PLS point clouds were previously rectified to the ground in the same ALS reference plane. A 3D passive rotation ( $R_{\omega\phi\kappa}$ ) was performed to normalize the full point cloud to the ground, in which a right-hand system was defined with origin at the scanner and Z to up. Rotation parameters were  $\omega=0^\circ$ ;  $\phi=-60^\circ$  and  $\kappa=-90^\circ$  and translation parameters were (0, 0, 0) for X, Y, Z respectively.

The ALS point cloud was acquired using the FGI HeliALS-TW system. The HeliALS-TW system consists of a RIEGL miniVUX-1UAV scanner (RIEGL Laser Measurement Systems GmbH, Austria) and an inertial navigation system integrated onto a helicopter platform (Figure 1.b). The scanner also operated at a wavelength of 1550 nm. The flight altitude was set at 100 m above the ground, with a flight speed of 50 km/h. The estimated ALS point cloud density is 500 pts/m<sup>2</sup>, at least 200 times less than LiPhe point clouds. The ALS point cloud is georeferenced at ETRS89/ TM35FIN. PLS and ALS point clouds were both acquired on April 6th, 2020, around 10 A.M.

To estimate the registration accuracy of the benchmarked methods, the PLS and ALS point clouds were initially registered and georeferenced using control features (e.g. points and lines), such as spheres and building corners. Eight control features were manually identified in the local point cloud and measured in situ using RTK-GNSS positioning. The transformation parameters were computed in the least square adjustment. A Helmert 3D transformation was performed to georeference the LiPhe point cloud to ETRS89/ TM35FIN. The estimated transformation parameters and distances between LiPhe point clouds (reference vs. after applying each benchmarked method estimated transformation) were compared to evaluate the performance of the benchmarked methods. Figure 1 shows the top view of the PLS (local coordinate system) and ALS (ETRS89/ TM35FIN) point clouds in panels (c) and (d) respectively. The visualizations are colorized based on normalized LiDAR reflectance, ranging from 0 to 2, expressed in decibels (DB).

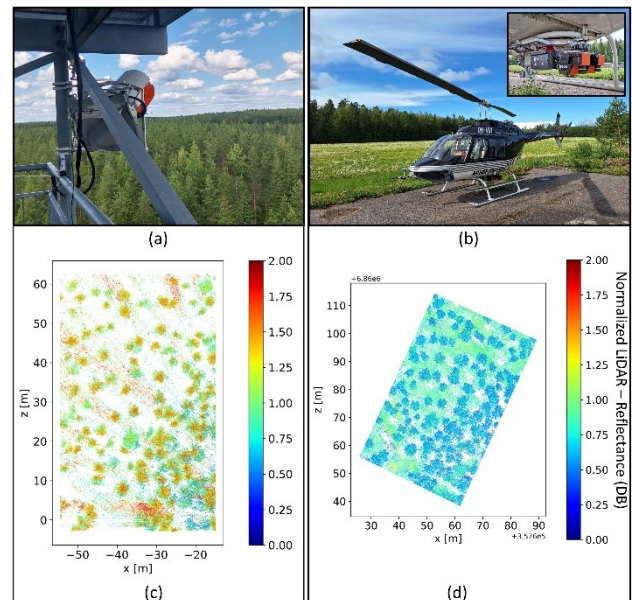


Figure 1. (a) Permanent laser scanner system at Lidar Phenology station (LiPhe), (b) FGI HeliALS-TW system, (c) top view (x,y) of LiPhe point cloud (sample) in local coordinates and colorized normalized LiDAR reflectance [0,2], (d) top view (x,y) of HeliALS-TW point cloud (sample) in georeferenced coordinates (ETRS89/ TM35FIN) and colorized normalized LiDAR reflectance [0,2].

#### 3.2 Methods

Many proposed multi-modal registration methods for forest point clouds rely on segmenting forest features such as stems and canopy. To represent these registration approaches, we selected for this benchmark the methods developed by Hyypä et al. (2021) and Pohjavirta et al. (2022) with reported registration accuracy superior to 20 cm. Additionally, common algorithms such as CPD, FPFH, and ICP were also explored in previous related works. Therefore, we benchmarked the multi-modal feature-based registration method, especially designed for forest (Hyypä et al., 2021; Pohjavirta et al., 2022), against traditional global (ICP, NDT, CPD) and feature-based (SHOT and FPFH).

ICP, NDT and CPD are the most used and adapted approaches in the literature for point cloud registration. Classified as global registration methods, these techniques do not rely on feature descriptors. ICP method (Best and McKay, 1992) is based on a point-to-point distance between nearest neighbours, in which the rotation and translation were parameterised in terms of the unit quaternion. This function minimizes the sum of squares of Euclidean distances between a set of points, leading to the estimation of transformation parameters between point clouds based on convergence criteria. Several variants of ICP, including plane-to-plane, have been proposed in previous works, which is also explored in this paper. NDT (Biber et al., 2003) consists of converting the point clouds into a 3D grid, represented as a continuously differentiable probability distribution function. Point cloud registration is achieved by optimizing the probability distributions of two point cloud datasets using the Hessian Matrix method. Achieving optimal results with ICP and NDT original algorithms typically requires an initial approximate registration of point clouds. Consequently, they are commonly utilized in the fine-registration step. CPD (Myronenko and Song, 2010) consider point cloud registration as a probability density estimation task. In this framework, one point cloud represents the centroids of a Gaussian Mixture Model (GMM), while the other represents the data points. Correspondences are determined by maximizing the GMM posterior probability for each data point. Consequently, CPD ensures that GMM centroids move

collectively as a group, maintaining the topological structure of the point cloud.

SHOT (Salti et al., 2014) and FPFH (Rusu et al., 2009) are 3D feature-based methods. We chose FPFH for its reported performance in real-time systems with a small number of points, while SHOT proved to be more effective for larger datasets, which is often the case of PLS data. FPFH is an optimized version of point feature histograms (PFH) to reduce the computation times (Rusu et al., 2008). FPFH computes a feature vector for each point in the point cloud based on the geometric properties of its local neighborhood, which are expressed into histograms. The histograms computed for each point are concatenated into a single feature vector for feature-based matching. Besides that, there are methods based on signature, which describe the 3D surface neighborhood of a given point by defining an invariant local reference frame. SHOT combines both signatures and histogram features. First, a local reference frame is established for a keypoint and its neighborhood. A spherical grid centered at this point is divided along the radial, azimuth and elevation axes. Subsequently, a locally weighted histogram is computed in each grid according to the normal at the keypoint and the angles between the normal at the neighboring points. In this work, we utilized uniform sampling from the Point Cloud Library (PCL) to downsample the point clouds and extract keypoints. Keypoint features were then obtained using FPFH and SHOT.

### 3.3 Benchmark results

Figure 3 provides a visual assessment of the transformed PLS point cloud generated by the benchmarked methods, colorized based on the distance to the PLS point cloud obtained using control points (named as PLS reference). Panel (a) shows the overlay of the PLS reference (blue) and ALS point clouds (yellow). Panel (b) to (g) displays the PLS point cloud results produced by Hyypä et al. (2021), Pohjavirta et al. (2022), ICP (plane to plane), NDT, CPD, SHOT and FPFH, respectively. In each panel, the coordinate component (E, N, or h) with the highest estimated RMSE is indicated in the upper-left corner. Additionally, the proportion of the RMSE corresponding to the E (green), N (orange), and h (blue) coordinates components is displayed in the bottom-right corner of each panel. The estimated planimetric and altimetric RMSE for each benchmarked method are presented in Table 2, in which the errors were obtained from the discrepancies between the estimated point cloud coordinates and PLS reference coordinates.

Table 1 presents the average point-to-point distance between the PLS reference and the PLS point cloud after coarse registration by the benchmarked methods. The corresponding standard deviation between those distances are also shown. This comparison is feasible because both point clouds remain consistent the same, with identical numbers of points and order, with the only distinction being the transformation method applied and the resulting coordinates. Subsequently, we check the point-to-point distance from the registered PLS to ALS point clouds. Figure 2 shows in detail the distribution of point-to-point PLS-ALS distances in a histogram for the methods that achieved an average point-to-point distance smaller than 15 cm, regarding PLS reference. The results obtained by Hyypä et al. (2021), ICP (plane to plane), and NDT meet this criterion, as indicated in Table 1.

Method	Distance between PC (m)	
	Average	Std
Hyypä et al. (2021)	0.175	0.058
Pohjavirta et al. (2022)	0.207	0.103
ICP (plane to plane)	0.159	0.093
NDT	0.157	0.084
CPD	0.756	0.165
SHOT	0.346	0.133
FPFH	0.766	0.060

Table 1. Average point-to-point distance between the LiPhe reference point cloud and LiPhe point clouds registered by the benchmarked methods and standard deviation (Std) between the estimated distances.

Method	RMSE (m)	
	Planimetry	Altimetry
Hyypä et al. (2021)	0.119	0.142
Pohjavirta et al. (2022)	0.192	0.130
ICP (plane to plane)	0.176	0.055
ICP (point to point)	8.907	0.608
NDT	0.169	0.057
CPD	0.289	0.718
SHOT	0.289	0.232
FPFH	0.507	0.578

Table 2. RMSE of the differences between the estimated point cloud coordinates obtained from the benchmark methods and the PLS reference, expressed in terms of planimetry and altimetry.

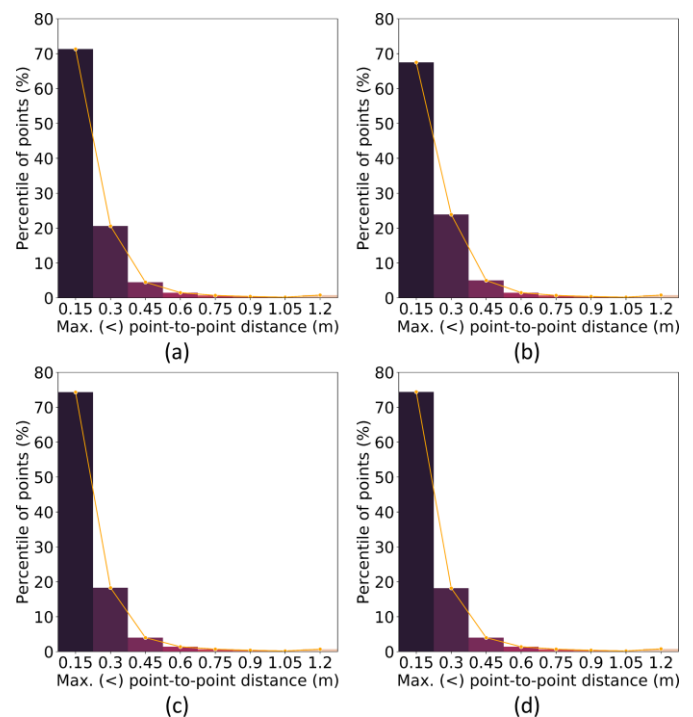


Figure 2. Point-to-point PLS-ALS distance distribution: (a) PLS reference, (b) Hyypä et al. (2021), (c) ICP (plane to plane) and (d) NDT.

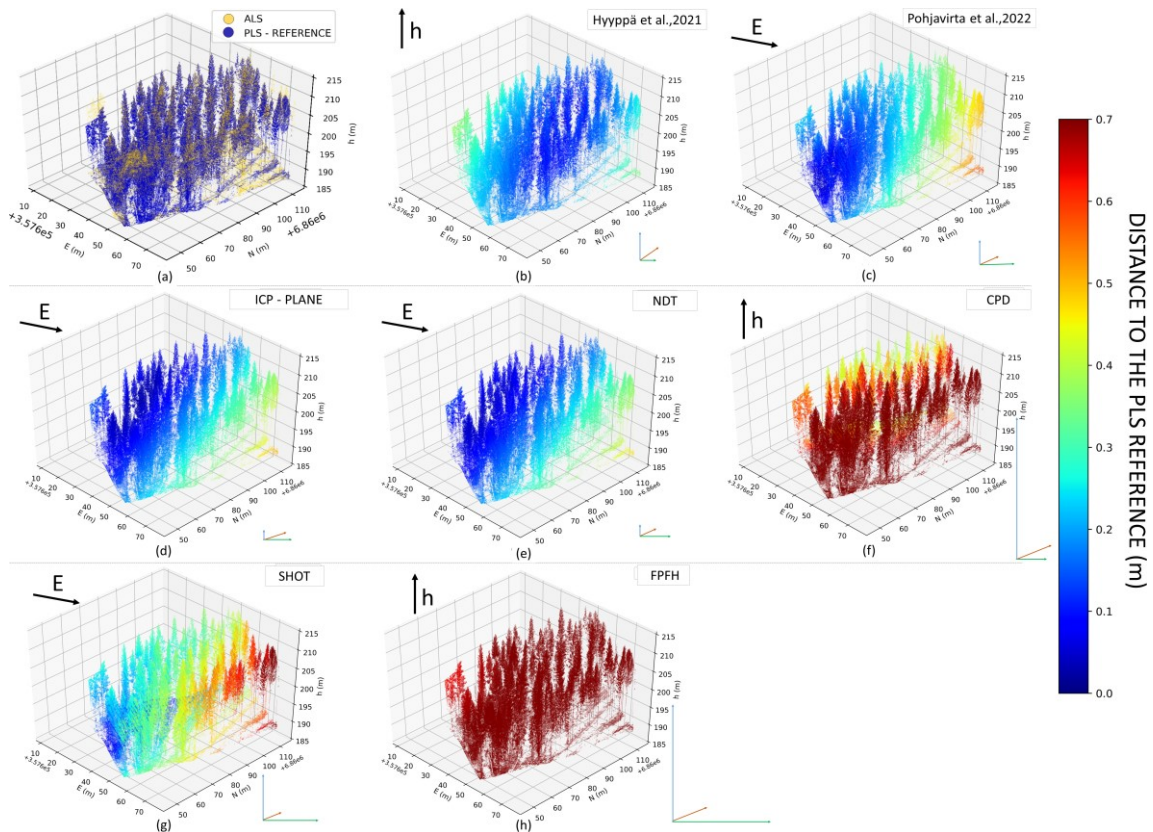


Figure 3. Benchmark results: (a) PLS-reference point cloud obtained by a 3D transformation estimated with control features (blue) against ALS point cloud (yellow). And, point-to-point distance between PLS-reference and the benchmarked methods: (b) Hyyppä et al. (2021), (c) Pohjavirta et al. (2022), (d) ICP (plane to plane), (e) NDT, (f) CPD, (g) SHOT and (h) PPFH.

#### 4. Conclusions

Here, we analysed the results concerning the coarse registration step. The majority of the methods achieved a coarse registration accuracy at cm-level, ranging from 12 to 50 cm in planimetry and from 5 to 57 cm in altimetry. These results may be further refined in a fine-segmentation step, which will not be considered in the assessments conducted here.

Overall, the benchmarked methods that were particularly designed for forest and the global methods, ICP (plane-to-plane) and NDT, exhibited similar performance in terms of RMSE, achieving results better than 20 cm in both planimetric and altimetric alignment. The achieved coarse registration accuracy closely aligns with the state-of-the-art methods reported in the review. Stem-based methods demonstrate stability across boreal and temperate forest datasets. For instance, Hyyppä et al. (2021) achieved an RMSE of 12 cm in planimetry against PLS reference, consistent with the accuracy reported by the author for boreal forest applications. Despite of its robustness, this method only provides 2D transformation. Altimetric alignment can accumulate errors due to the absence of 3D transformation estimation. An RMSE of 14 cm in altimetry was obtained. When comparing the transformed PLS point cloud directly to the ALS point cloud, over 67.5% of the points were within a 15 cm distance of the ALS point cloud (Figure 2.b). When comparing with the PLS reference (transformation obtained with control features), 71.3% of the points were within 15 cm distance of the ALS point cloud (Figure 2.a). Generally, distances greater than 15 cm may be also attributed to differing PLS-ALS perspectives, as some tree canopies were not fully visible to the LiDAR scanner. The main drawback of stem-based methods is their sensitivity to the required pre-processing steps, including point cloud

rectification, stem detection, and stem position estimation. Therefore their applicability in more complex forest environments can encounter challenges.

ICP and NDT are feature-independent, which can be advantageous in scenarios where feature extraction is challenging. Among the benchmarked global methods, we found that ICP (plane-to-plane) exhibited the most favourable performance in terms of usability and achieved RMSE. NDT achieved comparable planimetric and altimetric accuracy, however, it required initial values. We attribute ICP (plane-to-plane) performance to the high overlap between the ALS and PLS datasets and the constant developments of ICP variations. When compared to the ALS, the PLS point cloud registered with ICP (plane-to-plane) had 74.3% of its points within a distance of 15 cm from the ALS point clouds (Figure 2.c). These findings suggest that there is still potential for exploring variations of global registration methods as alternative solutions to complex forest environments. On the other hand, global methods demand high computational resources and they are sensitive to variations in point cloud overlaps, particularly due to the absence of one-to-one correspondence between LiDAR point sets from different platforms. For example, CPD and ICP (point-to-point) failed to produce satisfactory results with high planimetric (0.289 m and 8.907 m, respectively) and altimetric (0.7 m and 0.6 m, respectively) RMSE values. Zhou et al. (2023) achieved comparable outcomes by integrating CPD into a coarse-registration methodology, resulting in an average distance of 194.83 cm between ALS and TLS point clouds.

Less accurate results were obtained using traditional feature-based methods. As proposed by Dai et al. (2022) and Chen et al. (2024), keypoint matching in forest environments needs to be accompanied by strategies for reducing the search space and

filtering outliers. Consistent results for SHOT and FPFH were only achieved when reducing the search space and applying outlier filtering. Both SHOT and FPFH resulted in RMSE values larger than 20 cm in both planimetric and altimetric coordinates. SHOT outperformed FPFH in both the planimetry ( $28.9 < 50.7$  cm) and the altimetry ( $23.2 < 57.8$  cm) alignment. Chen et al. (2024) presented better results using FPFH with an RMSE of 18.2 cm after TLS-UAV coarse segmentation. These results are likely attributed to the feature detection steps, which provided insufficient and non-optimal correspondences between the point clouds, especially in the FPFH approach. More studies focusing on the multi-modal registration requirements and accuracy for advancing forest applications are still needed. Additionally, benchmark initiatives targeting more complex forests and the future of deep-learning methods are recommended.

### Acknowledgments

This study was funded by Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES), project: 88887.695922/2022-00; National Council for Scientific and Technological Development (CNPq); project: 141550/2020-1 and Research Council of Finland, projects: 316096/320075, 334060, 359204, 356137, 331708, 353264, 357902.

### References

Best, P.J., McKay, N.D., 1992: A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* 14, 239–256.

Biber, P., Straßer, W., 2003: The normal distributions transform: A new approach to laser scan matching. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2003*, 27–31.

Campos, M. B., Litkey, P., Wang, Y., Chen, Y., Hyyti, H., Hyyppä, J., Puttonen, E., 2021: A long-term terrestrial laser scanning measurement station to continuously monitor structural and phenological dynamics of boreal forest canopy. *Frontiers in Plant Science*, 11, 606752.

Castanheiro, L.F., Tommaselli, A.M.G., Garcia, T.A.C., Campos, M.B., Kukko, A., 2023: Point Cloud Registration Using Laser Data from an Orange Orchard. *The Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 48, 71-77.

Chen, J., Zhao, D., Zheng, Z., Xu, C., Pang, Y., Zeng, Y., 2024: A clustering-based automatic registration of UAV and terrestrial LiDAR forest point clouds. *Computers and Electronics in Agriculture*, 217, 108648.

Cheng, L., Chen, S., Liu, X., Xu, H., Wu, Y., Li, M., Chen, Y., 2018: Registration of laser scanning point clouds: A review. *Sensors*, 18(5), 1641.

Dai, W., Yang, B., Liang, X., Dong, Z., Huang, R., Wang, Y., Li, W., 2019: Automated fusion of forest airborne and terrestrial point clouds through canopy density analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 156, 94-107.

Dai, W., Kan, H., Tan, R., Yang, B., Guan, Q., Zhu, N., Xiao, W., Dong, Z. 2022: Multisource forest point cloud registration with semantic-guided keypoints and robust RANSAC mechanisms. *International Journal of Applied Earth Observation and Geoinformation*, 115, 103105.

Dong, Z., Liang, F., Yang, B., Xu, Y., Zang, Y., Li, J., Wang, Y., Dai, W., Fan, H., Hyyppä, J., Stilla, U., 2020: Registration of large-scale terrestrial laser scanner point clouds: A review and benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163, 327-342.

Ghorbani, F., Chen, Y.C., Hollaus, M., Pfeifer, N., 2024: A Robust and Automatic Algorithm for TLS-ALS Point Cloud Registration in Forest Environments based on Tree Locations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4015 – 4035.

Guan, H., Su, Y., Hu, T., Wang, R., Ma, Q., Yang, Q., ..., Guo, Q., 2019: A novel framework to automatically fuse multiplatform LiDAR data in forest environments based on tree locations. *IEEE Transactions on Geoscience and Remote Sensing*, 58(3), 2165-2177.

Harris, C., Stephens, M., 1988: A combined corner and edge detector. In: *Proceeding of Alvey vision conference*, 15(50), 10-5244.

Hauglin, M., Lien, V., Næsset, E., Gobakken, T., 2014: Geo-referencing forest field plots by co-registration of terrestrial and airborne laser scanning data. *International journal of remote sensing*, 35(9), 3135-3149.

Huang, X., Mei, G., Zhang, J., Abbas, R., 2021: A comprehensive survey on point cloud registration. *arXiv preprint arXiv*, 2103.02690.

Huang, X., Qu, W., Zuo, Y., Fang, Y., & Zhao, X., 2022: IMFNet: Interpretable multimodal fusion for point cloud registration. *IEEE Robotics and Automation Letters*, 7, 4, p. 12323-12330.

Huang, X., Mei, G., Zhang, J., 2023: Cross-source point cloud registration: Challenges, progress and prospects. *Neurocomputing*, 126383.

Hyyppä, E., Muhojoki, J., Yu, X., Kukko, A., Kaartinen, H., Hyyppä, J., 2021. Efficient coarse registration method using translation-and rotation-invariant local descriptors towards fully automated forest inventory. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 2, 100007.

Liang, X., Hyyppä, J., 2013: Automatic stem mapping by merging several terrestrial laser scans at the feature and decision levels. *Sensors*, 13(2), 1614-1634.

Lin, Y.C., Shao, J., Shin, S.Y., Saka, Z., Joseph, M., Manish, R., Fei, S., Habib, A., 2022: Comparative analysis of multi-modal, multi-resolution, multi-temporal LiDAR data for forest inventory. *Remote Sensing*, 14(3), 649.

Lindberg, E., Holmgren, J., Olofsson, K. and Olsson, H., 2012: Estimation of stem attributes using a combination of terrestrial and airborne laser scanning. *European Journal of Forest Research*, 131(6), 1917-1931

Liu, J., Liang, X., Hyyppä, J., Yu, X., Lehtomäki, M., Pyörälä, J., Zhu, L., Wang, Y., Chen, R., 2017: Automated matching of multiple terrestrial laser scans for stem mapping without the use of artificial references. *International journal of applied earth observation and geoinformation*, 56, 13-23.

Liu, Q., Wang, J., Ma, W., Zhang, J., Deng, Y., Shao, D., Xu, D., Liu, Y., 2021: Target-free ULS-TLS point-cloud registration for alpine forest lands. *Computers and Electronics in Agriculture*, 190, 106460.

Monji-Azad, S., Hesser, J., Löw, N., 2023: A review of non-rigid transformations and learning-based 3D point cloud registration methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 196, 58-72.

Myronenko, A., Song, X., 2010: Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12), 2262-2275.

Olofsson, K., Holmgren, J., 2022: Co-registration of single tree maps and data captured by a moving sensor using stem diameter weighted linking. *Silva Fennica*, 56(3).

Paris, C., Kelbe, D., van Aardt, J., Bruzzone, L., 2017: A novel automatic method for the fusion of ALS and TLS lidar data for robust assessment of tree crown structure. *IEEE Trans. Geosci. Remote Sens.* 55, 3679–3693.

Pohjavirta, O., Liang, X., Wang, Y., Kukko, A., Pyörälä, J., Hyyppä, E., Hyyppä, J., 2022: Automated registration of wide-baseline point clouds in forests using discrete overlap search. *Forest Ecosystems*, 9, 100080.

Polewski, P., Erickson, A., Yao, W., Coops, N., Krzystek, P., Stilla, U., 2016: Object-based coregistration of terrestrial photogrammetric and ALS point clouds in forested areas. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, 347-354.

Polewski, P., Yao, W., Cao, L., Gao, S., 2019: Marker-free coregistration of UAV and backpack LiDAR point clouds in forested areas. *ISPRS journal of photogrammetry and remote sensing*, 147, 307-318.

Rusu, R. B., Blodow, N., Beetz, M., 2009: Fast point feature histograms (FPFH) for 3D registration. In: *Proceedings of the IEEE international conference on robotics and automation*, 3212-3217.

Salti, S., Tombari, F., Di Stefano, L., 2014: SHOT: Unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding*, 125, 251-264.

Shao, J., Yao, W., Wan, P., Luo, L., Wang, P., Yang, L., Lyu, J., Zhang, W., 2022: Efficient co-registration of UAV and ground LiDAR forest point clouds based on canopy shapes. *International Journal of Applied Earth Observation and Geoinformation*, 114, 103067.

Si, H., Qiu, J., Li, Y., 2022: A review of point cloud registration algorithms for laser scanners: applications in large-scale aircraft measurement. *Applied Sciences*, 12(20), 10247.

Vosselman, G., Maas, H. G., 2010: Airborne and terrestrial laser scanning. CRC press, 311.

Zhang, Z., Dai, Y., Sun, J., 2020: Deep learning based point cloud registration: an overview. *Virtual Reality & Intelligent Hardware*, 2(3), 222-246.

Zhang, W., Shao, J., Jin, S., Luo, L., Ge, J., Peng, X., Zhou, G., 2022: Automated marker-free registration of multisource forest point clouds using a coarse-to-global adjustment strategy. *Forests*, 12(3), 269.

Zhou, R., Sun, H., Ma, K., Tang, J., Chen, S., Fu, L., Liu, Q., 2023: Improving estimation of tree parameters by fusing ALS and TLS point cloud data based on canopy gap shape feature points. *Drones*, 7(8), 524.