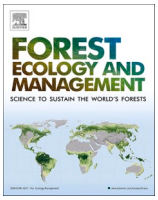
Forest Ecology and Management 532 (2023) 120795

Contents lists available at ScienceDirect

Forest Ecology and Management

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



Improving the accuracy of timber volume and basal area prediction in heterogeneously structured and mixed forests by automated co-registration of forest inventory plots and remote sensing data

Simon Janssen [a,*], Hans Pretzsch [b], Anton Bürgi [a], Laura Ramstein [a], Leo Gallus Bont [a]

[a] Sustainable Forestry Group, Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Zuercherstrasse 111, 8903 Birmensdorf, Switzerland
[b] Chair of Forest Growth and Yield Science, School of Life Sciences Weihenstephan, Technical University of Munich, Hans-Carl-Von-Carlowitz-Platz 2, 85354 Freising, Germany

ARTICLE INFO

Keywords:
Canopy height model
Single-tree detection
LiDAR
Orthophoto
Forest inventory
Regression analysis

ABSTRACT

Accurate georeferencing is essential if forest indicators, such as timber volume, are to be modelled and predicted area-wide by establishing a linkage between local inventories and remote sensing data. Nevertheless, due to inaccuracies in global navigation satellite system (GNSS) measurements under a closed canopy, determining the exact position of sample-plots in a forest inventory is a major challenge. In this study different methods were evaluated, each of which is designed to improve the co-registration between field measurements and remote sensing data in a forest inventory. The methods were evaluated in two areas in Switzerland (Bremgarten and Zurich), all of which have heterogeneously structured and mixed forests. A simple algorithm that searches for the best match between tree-tops, detected in remote sensing data, and an inventory tree-top point-cloud led to unsatisfactory results. The failure of the algorithm was primarily related to the lack of accurate single-tree-identification (STI) methods in airborne laser scanning (ALS) data (10 points m^{-2}) on deciduous and mixed forest stands. These inaccuracies hampered a successful co-registration of the two data sources. To omit the single-tree-identification (STI) step, methods that rely on the comparison between an artificial canopy height model (CHM) calculated from the inventory data and the CHM generated from ALS data were tested. Correlating the two CHMs made it possible to identify plausible positions within the search range. The quality of all co-registration methods was assessed by the leave-one-out cross-validated root-mean-squared error (RMSE) of the timber volume estimate of the subsequently calibrated regression model. The best results were achieved with a method that modelled inventory tree-crowns as spheres and that applied a correlation metric named SQDIFF_-NORMED. Co-registration made it possible to increase the model accuracy of timber volume estimates in Bremgarten by 31 RMSE-%, i.e. from 134.4 m^3 ha^{-1} (without co-registration) to 92.6 m^3 ha^{-1} with co-registered positions. Additional testing of the identified superior method with a larger inventory dataset of the canton of Zurich, Switzerland confirmed these results. There, the RMSE of the basal-area estimate was improved by 10 RMSE-%, from 13.56 m^2 ha^{-1} to 12.18 m^2 ha^{-1}. For these CHM-based methods, integrating the information from a deciduous–evergreen (DecEv) raster improved the positional accuracy but not the overall predictive power of the regression models.

1. Introduction

Forests provide various ecosystem services. They are important for biodiversity conservation and climate change mitigation, but also produce timber for various industrial purposes and function as a location for retreat and recreation. Provision of these various services requires efficient forest management. Effective forest management in turn requires precise knowledge about the current state of the forest. As a census of the whole forest area is usually impossible, inventories with a sampling scheme are usually applied to estimate, for example, the timber volume (VOL) and the basal area (BA). To increase the accuracy of these estimates two-phase inventories are common (Mandallaz, 2007; Naesset & Jonmeister, 2002). After inventory data are collected at the terrestrial sampling locations (i.e. in sample-plots), a relationship between the target variable (e.g. VOL) in the sample-plots and predictors derived from area-wide remote sensing data is established by means of

* Corresponding author.

https://doi.org/10.1016/j.foreco.2023.120795
Received 17 August 2022; Received in revised form 7 January 2023; Accepted 7 January 2023

regression analysis (Bont et al., 2020; Hauglin et al., 2014; Hollaus et al., 2009; Immitzer et al., 2012; Waser, 2012; Waser et al., 2011). To form accurate and reliable regression models, it is essential that the remote sensing data and the sampling locations are adequately geo-referenced (co-registered). Inaccurate co-registration between terrestrial inventory data and remote sensing data leads to erroneous regression models and therefore less accurate predictions.

The location of the sample-plots is only approximately known, even if positions are measured with a global navigation satellite system (GNSS). With non-differential GNSS devices, positional accuracies of 10–20 m are attainable under a closed canopy (Lamprecht et al., 2017; Monnet & Mermin, 2014; Wing et al., 2005). For example Zimbelman et al. (2018) reported an RMSE ranging from 1.81 m to 16.69 m, with a mean of 6.61 m, which was depending on the stand characteristics.

By using a differential GNSS (dGNSS) device it is possible to account for atmospheric diffraction effects, but precise measurements require a logging time of approximately 20 min (Hauglin et al., 2014; Valbuena et al., 2010). Without canopy cover, positional accuracies of ≤ 0.5 m are realisable. Under closed canopy conditions the satellite-constellation-dependent refraction of the GNSS signal leads to positional errors $>$ 0.5 m. For stands with a BA < 30 m^2 ha^{-1}, Naesset & Jonmeister (2002) measured dGNSS positional errors between 0.49 and 3.60 m. The "true" reference positions were found by means of transverses performed from a reference point in the open area with a positional accuracy of 0.1–0.2 m. For BA values > 45 m^2 ha^{-1} the authors found positional errors to be between 2.15 and 5.60 m. Kaartinen et al. (2015) showed that combinations of dGNSS sensors combined with an inertial measurement unit (IMU) further improved accuracy with a positional accuracy (RMSE) ranging between 1.3 m (young spruce) to 2.7 m (mature spruce).

Even though GNSS and dGNSS devices are now used in most forest inventories to measure the sampling location exactly, positional errors of $>$ 10 m are common due to erroneous instrument handling or a weak reference signal. The accuracy of statistical regression models that are derived from that position decreases with increasing positional error (Dorigo et al., 2010; Fadili et al., 2019; Hernández-Stefanoni et al., 2018). Therefore, predictions made with these models also have a greater error. Post-processing techniques to correct positional measurement errors by comparing field data with remote sensing data could improve the quality of area-wide predictions of forest parameters and multi-phase forest inventories. Several authors have studied such co-registration techniques, mostly under advantageous forest conditions with a high proportion of conifers (conical crown shape) and relatively open forest stands. Frequently, use of an accurate single-tree-identification (STI) method from remote sensing data is required for the success of these techniques (Hauglin et al., 2014; Olofsson et al., 2008).

Different STI methods have already been intensively researched. Today, most STI approaches rely directly on airborne laser scanning (ALS) point-clouds or datasets derived from an ALS point-cloud, such as canopy height models (CHM; Kaartinen et al., 2012; Vauhkonen et al., 2012). Some STI methods estimate tree positions (Eysn et al., 2015; Kaartinen et al., 2012; Koch et al., 2006; Menk et al., 2017; Vauhkonen et al., 2012), whereas others attempt to identify entire tree-crowns by means of different segmentation algorithms (Eysn et al., 2015; Kaartinen et al., 2012; Koukoulas and Blackburn, 2005; Leckie et al., 2003; Vauhkonen et al., 2012). The most sophisticated but also data intensive algorithms identify complete trees (Lu et al., 2014; Parkan, 2019; Parkan & Tuia, 2015; Reitberger et al., 2009). When co-registration is attempted by means of a previously conducted STI algorithm, the presumed sample-plot centre is shifted within a certain search range based on the identified single-tree distribution compared with the distribution based on the terrestrial record (Hauglin et al., 2014). However, little research has been done on the influence of STI quality on the accuracy of co-registration.

For clarity and readability, the terms deciduous and evergreen were used in this paper. Only, when the relevant distinguishing feature is the crown shape, the terms coniferous and broadleaf were applied. The larch is treated as deciduous and coniferous (non-broadleaf). All other observed conifers were treated as evergreen.

1.1. Co-registration via single-tree identification

If co-registration is carried out via a previous STI, different data sources can be linked by comparing and matching single-tree distribution patterns or single-tree characteristics. In contrast to an area- or pixel-based approach, this method can integrate structural characteristics of single trees into a subsequent modelling process.

Hauglin et al. (2014) use terrestrial laser scanner (TLS) point-clouds of each sample-plot to co-register it with the ALS point-cloud in a conifer-dominated study area. As the initial position, measurements with a non-differential GNSS device were used. By identifying local maxima (LM) in an ALS-derived CHM raster, tree positions were determined, as well as tree-top positions extracted from the TLS data. For co-registration, a match score was calculated based on pair-wise point distances within a search range.

Lamprecht et al. (2017) developed an algorithm that links trees from field measurements to trees that were detected with remote sensing data. GNSS measurements of the sample-plot centres were used as initial positions for co-registration. The authors argued that most co-registration methods are limited by the small number of sample-plots for algorithm calibration, and that sufficient training data for a reliable model can be provided only by modelling forest stands. The method proposed by Lamprecht et al. (2017) used many synthetic forest stands to achieve an optimal parametrization of the co-registration algorithm. From these forest stands artificial point-clouds were generated and an LM-based STI was applied. These identified tree-tops were subsequently matched to the tree-tops of the simulated stand (3D-point-matching). Samples were considered correctly co-registered if at least 50% of the individual trees have been correctly assigned. Olofsson et al. (2008) presented a method which linked terrestrial inventory trees and trees detected in remote sensing data. As a starting position for the co-registration algorithm, GNSS measurements of the sample-plot centres were used. In an ALS-based CHM, an LM-based STI was performed. Terrestrially measured, as well as remotely detected, trees were modelled as a Gaussian surface with an amplitude proportional to the tree height to calculate an artificial CHM (aCHM). Within a search range, this aCHM was compared with the ALS-based CHM by calculating a normalized correlation coefficient for each overlaying position. At the end, the overlaying position with the best coefficient was chosen. The Swedish study area was dominated by coniferous trees. It should be emphasized that the actual co-registration process was done by CHM image-matching, and only the preparatory work was based on single-tree detection.

1.2. Co-Registration with CHM image-matching

In contrast to the co-registration methods described in section 1.1, the CHM-based methods described here do not require an STI. This can be an advantage in broadleaf-dominated forests (irregular crown shape), where STI by LM detection is error prone (Monnet & Mermin, 2014).

Dorigo et al. (2010) chose an approach that strongly resembles the one of Olofsson et al. (2008) but avoids STI. The crown of each terrestrially sampled tree was modelled to generate an artificial CHM (aCHM) for each sample-plot. In a search range of 30 m, the ALS-CHM and the aCHM were compared and the most similar position was identified, giving more weight to pixels that presumably belonged to a crown-tip. As a reference and ground truth, all sample-plots were manually co-registered. Most broadleaf-dominated sample-plots had to be excluded from the analysis because manual co-registration was impossible.

Monnet and Mermin (2014) presented another approach, which performed a co-registration by comparing CHM values to the size of terrestrially recorded trees. Based on tree-size-related parameters (DBH

and tree height), a tree-size raster was calculated for each sample-plot. In this raster only cells with a tree trunk in it were kept as non-zero. This artificial inventory-plot raster was moved over the ALS-based CHM within a search range of 20 m, and the correlation (same equation as in Olofsson et al. (2008)) was calculated at each position.

Previous work primarily focused on co-registration in stands dominated by conifers with a single and pronounced tree-top maximum and therefore accurate STI results. In Central Europe, broadleaf and mixed forests dominate and are predicted to become even more important with advancing climate change (Seidl et al., 2017). Therefore, there is a need to investigate these existing methods under more diverse forest conditions and to develop methods that are better suited to highly structured and mixed forest stands. Recent co-registration studies have emphasized the use of single-tree point-matching methods, which search for the best match between two corresponding tree-top point-clouds. It is likely that co-registration methods that require an STI based on ALS data are not yet suitable for mixed and diversely structured forest stands, as STI is highly unreliable under such conditions. Raster-based methods not only omit STI, but also use area-wide information. Therefore, positional information, for example about gaps in the forest area, is maintained. Such non-STI-reliant methods can therefore be expected to become increasingly important in Central Europe as forest composition shifts under climate change.

The objective of this study was to test the effectiveness of different (STI and non-STI-reliant) co-registration methods using a variety of forest stand types. Emphasis was placed on the generalisability of the methods under different forest conditions. As acquiring reliable reference data, e.g. by manual co-registration, is a major challenge and a source of error, the quality of the co-registration was primarily assessed based on the quality of OLS regression models that predicted VOL and BA from ALS-derived predictors. OLS regression models were applied because the frequently used two-phase estimators introduced by Mandallaz (2013) require the internal models to be fitted with an OLS technique.

1.3. Research questions

The following research questions were addressed in this study:

1. Is there a gain in the model performance of OLS regression models, in predicting VOL or BA from remote-sensing-derived predictors, when terrestrial samples are co-registered with remote sensing data prior to model calibration?
2. Which co-registration methods are best suited in study areas dominated by mixed and heterogeneously structured stands?
3. Is co-registration further improved when deciduous–evergreen information from remote sensing data is integrated into the co-registration process?

2. Materials and methods

2.1. Study areas

The main comparison of different co-registration methods was performed with forest inventory data from Bremgarten (Canton of Aargau, Switzerland) (Fig. 1, left). The best-performing methods were then tested with a considerably larger dataset from a cantonal inventory in Zurich (Switzerland) (Fig. 1, right). Additional study area parameters can be found in Table 1.

The response variables were chosen according to the requirements of

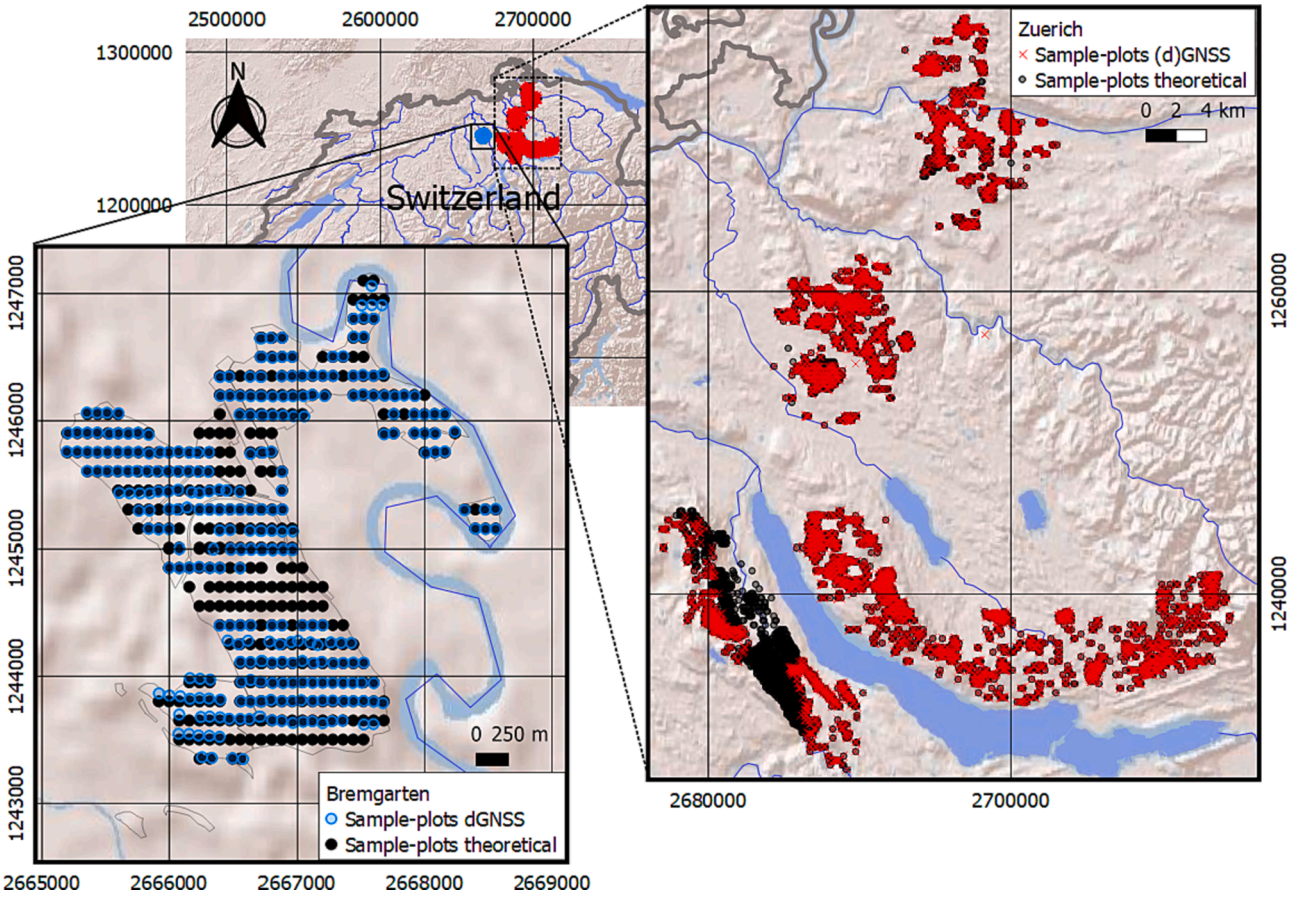


Fig. 1. Investigated case study areas in Switzerland and the locations of the sample-plots in Bremgarten and Zurich (coordinate system: EPSG 2056, CH 1903+ / LV 95).

Table 1
Properties of the study areas Bremgarten and Zurich (DBH = diameter at breast height, 1.3 m above ground).

	Bremgarten	Zurich
Number of terrestrial sample-plots (GNSS in brackets)	363 (276)	4,687 (3,446)
Number of sample-plots after cleaning (section 2.5) (GNSS in brackets)	329 (244)	4,687 (3,446) – no cleaning
Theoretical sample-plot arrangement grid	80 m × 150 m	80 m × 300 m
GNSS measurement quality	dGNSS measurements	dGNSS or GNSS measurements
Recording method	400 m^2 circle, min. DBH threshold of 12 cm	300 m^2 circle, min. DBH threshold of 12 cm
Date of terrestrial sampling	Autumn to winter 2011/2012	2017/2018
Date of the LiDAR flight	3 datasets: 9.11.2011, 18.3.2014 – 4.4.2014 and 19.6.2014 – 25.7.2014	8.3.2014 – 15.4.2014
Point density of the LiDAR raw data	≥ 8 points m^{-2}, ≥ 16 points m^{-2} and ≥ 14 points m^{-2}	≥ 8 points m^{-2}, mean: 15 points m^{-2}
Date of orthoimage flight	Summer 2016	Summer 2018
Orthoimage resolution	25 cm	10 cm

the data provider. In Bremgarten the local forest owners primary focus is the merchantable VOL. For the Canton of Zurich, the BA is a valuable monitoring dataset.

2.1.1. Bremgarten

The Bremgarten study area is located in the Central Plateau of Switzerland (Fig. 1). A detailed description can be found in Bont et al. (2020). The natural forest types are deciduous-dominated mixed forests. The actual forest stands are dominated by evergreen trees (conifers without larch; 52 % of both VOL and BA) and have an average VOL of 274 m^3 ha^{-1} and BA of 20.9 m^2 ha^{-1}. As a result of storm Lothar in December 1999, there is a surplus of young forest stands compared with a normal age class structure (Fig. 2, left and Fig. 3), with a higher proportion of deciduous volume than in older forest stands (Bont, 2019). As a consequence, at least in terms of area, deciduous trees dominate the study area (75 %; calculated according to section 2.9). In winter 2011/2012, an inventory with a total of 363 permanent sample-plots was carried out, following the protocol of Schmid-Haas et al. (1993). The tree species beech (*Fagus sylvatica,* 17 % VOL), oak (*Quercus* sp., 9.1 %), maple (*Acer* sp., 5.8 %), ash (*Fraxinus excelsior,* 5.2 %), basswood (*Tilia* sp., 0.9 %), cherry (*Prunus avium,* 0.5 %), elm (*Ulmus* sp., 0.1 %), Norway spruce (*Picea abies,* 41.6 %), silver fir (*Abies alba,* 6.2 %), Scots pine (*Pinus sylvestris,* 1.4 %) and larch (*Larix* sp., 2.9 %) were distinguished. Other species were summarized as other deciduous trees (0.1 %) and other evergreen trees (3.0 %). The same design had already been applied in inventories in 1971, 1976, 1986 and 1996. The sampling design featured circular sample-plots with a radius of 11.3 m (area = 400 m^2). A sample was taken in a W-E direction every 80 m and in a S-N direction every 150 m. During the inventory of 2011/2012, the position of sample-plots on which trees were present (276 sample-plots in total) was measured with a dGNSS receiver. In addition to the terrestrial measurements, 4-band (RGB and near infrared [NIR]) orthoimages from 2016 with a resolution of 25 cm were available, along with data from three ALS flights. One ALS flight was made in November 2011 (leaf-off) with a point density of 8 points m^{-2} and two flights were made in 2014 (leaf-on and leaf-off) with 14 points m^{-2} (Abteilung Wald, 2014; BFS Swissphoto AG, 2011).

2.1.2. Zurich

The canton of Zurich maintains a forest inventory with permanent sample-plots surveyed according to Schmid-Haas et al. (1993). The inventory design features circular sample-plots with a radius of 9.8 m (300 m^2) on a 80 m (W-E) × 300 m (S-N) grid (Baudirektion Kanton Zürich, 2017). The DBH threshold for recording is 12 cm. The tree species beech (*Fagus sylvatica,* 25.5 % VOL), oak (*Quercus* sp., 3.6 %), maple (*Acer* sp., 5.9 %), ash (*Fraxinus excelsior,* 8.8 %), basswood (*Tilia* sp., 0.4 %), cherry (*Prunus avium,* 0.6 %), elm (*Ulmus* sp., 0.4 %), Norway spruce (*Picea abies,* 32.8 %), silver fir (*Abies alba,* 10.1 %), Scots pine (*Pinus sylvestris,* 6.5 %) and larch (*Larix* sp., 1.8 %) were distinguished. Other species were summarized as other deciduous trees (2.4 %) and other evergreen trees (0.5 %). In this study, records from the years 2017 and 2018 were used (4,687 sample-plots; Fig. 1). The corresponding forest stands have an average VOL of 414 m^3 ha^{-1} and BA of 30.9 m^2 ha^{-1}. For 3,446 of the total 4,687 sampling-plots, dGNSS or GNSS measurements were available. ALS data with a point density of 8 points m^{-2} were acquired in spring 2014 (mostly leaf-off) (BSF Swissphoto AG, 2014). Four-band (RGB and NIR) orthoimages with a resolution of 10 cm were acquired in summer 2018.

2.2. Allometry data

For the derivation of tree-species-specific regression models, which were used to estimate tree height and crown size based on DBH, data from the Chair of Forest Growth and Yield Science at the Technical University of Munich (TUM), as well as the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) were used. The WSL data originated from the Experimental Forest Management (EFM) plots (Forrester et al., 2019), while the TUM data were gathered by Professor Hans Pretzsch and his group for various research projects related to the growth pattern in pure and mixed forest stands (Dieler & Pretzsch, 2013; Naudts et al., 2015; Pretzsch, 2014; Pretzsch et al., 2002, 2015; Pretzsch & Dieler, 2012). In total 213,006 single-tree measurements were

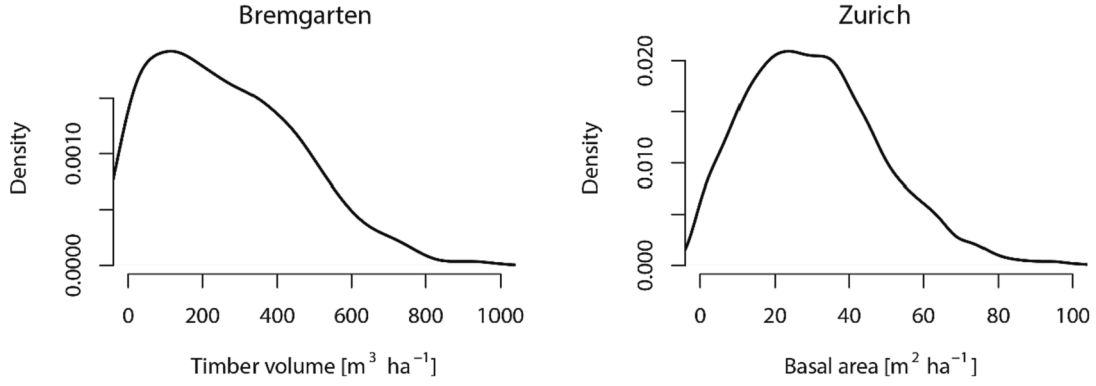


Fig. 2. Densities of the measured timber volume (Bremgarten, left) and basal area (Zurich, right) distributions of the sample-plots. The density scale differs between panels.

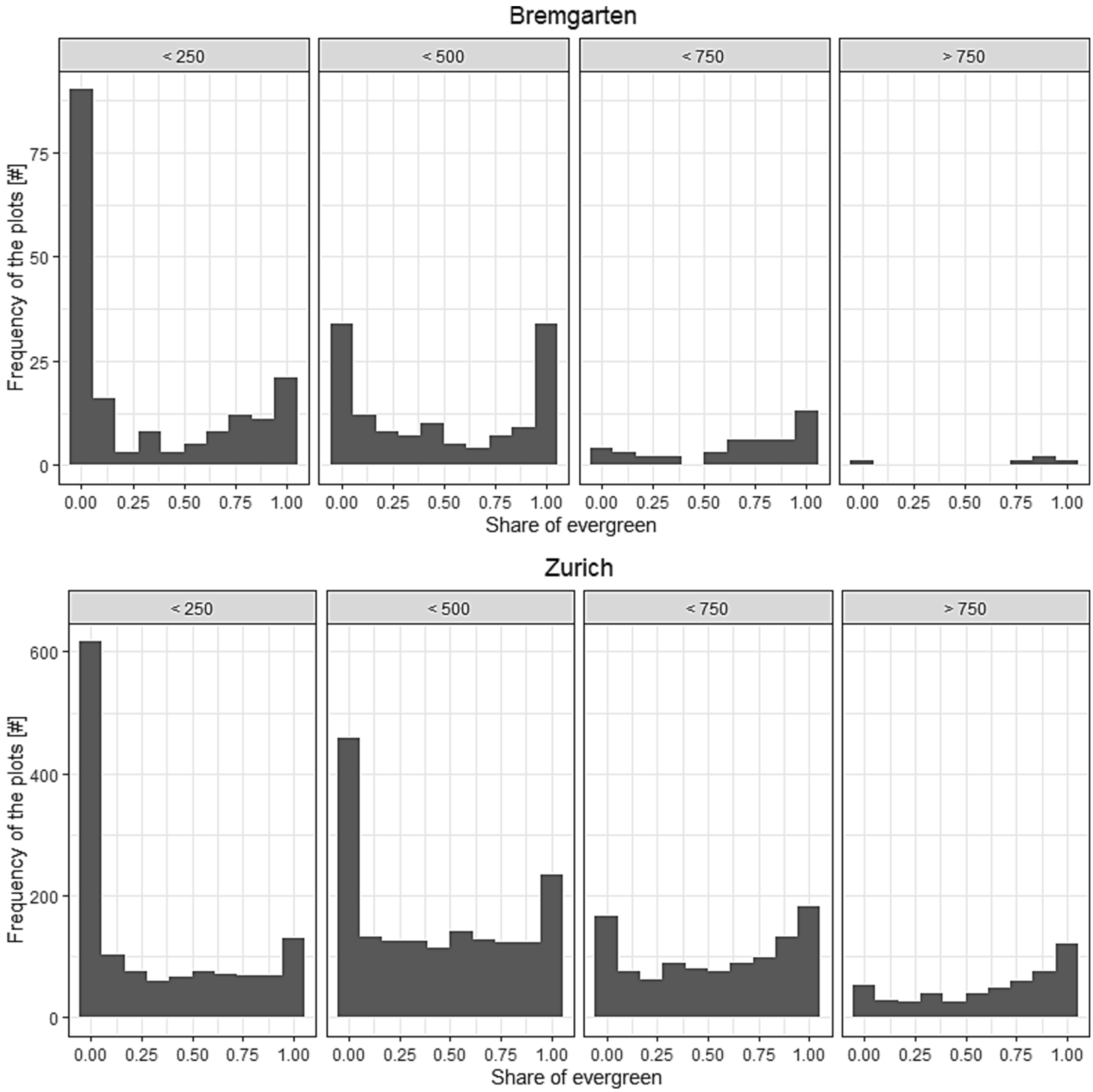


Fig. 3. Histograms of the number of sample-plots aggregated by timber volume class and ordered by the share of evergreen tree species (conifers without larch) in Bremgarten and Zurich. Timber volume classes are < 250 m^3 ha^{-1}, 250–500 m^3 ha^{-1}, 500–750 m^3 ha^{-1}, and > 750 m^3 ha^{-1}.

analysed. Additional Information can be found in the Appendix.

2.3. Methods overview

An overview of the workflow is given in Fig. 4. In Bremgarten all of the methods were evaluated, whereas in Zurich only the ones that performed best in Bremgarten were tested. In total three different sources of initial coordinates for co-registration were tested: (1) theoretical positions, (2) dGNSS positions and (3) dGNSS measurements with an artificial random error of up to 10 m to simulate non-differential GNSS measurements. The workflow for Bremgarten is explained below, as the application to the Zurich dataset was merely a simplification of the procedure.

In a first step (section 2.4; Fig. 4a) the ALS data were validated, classified, normalized and clipped to the extent of the search window of each sample-plot. Starting from a known sample-plot position, two different co-registration approaches were tested.

In the first approach (Fig. 4b), the tree-tops of the inventory trees were modelled via tree-species-specific allometric relationships to generate an artificial CHM (aCHM) (section 2.7). Subsequently, (Fig. 4c) within a quadratic search-window with a side length of 100 m for theoretical sample-plot positions and 50 m for (d) GNSS initial positions, the best match with the pit-free CHM (Khosravipour et al., 2014) was identified (section 2.11.1). The choice of the search window resulted from deviations from the theoretical sample-plot positions of up to 50 m in Bremgarten.

In the second approach (Fig. 4d), the height of each inventory tree was estimated from the DBH, and the tree-tops were saved as a three-dimensional point-cloud (section 2.6). In parallel, (Fig. 4e) various STI methods were applied to extract single-tree positions from the CHM or

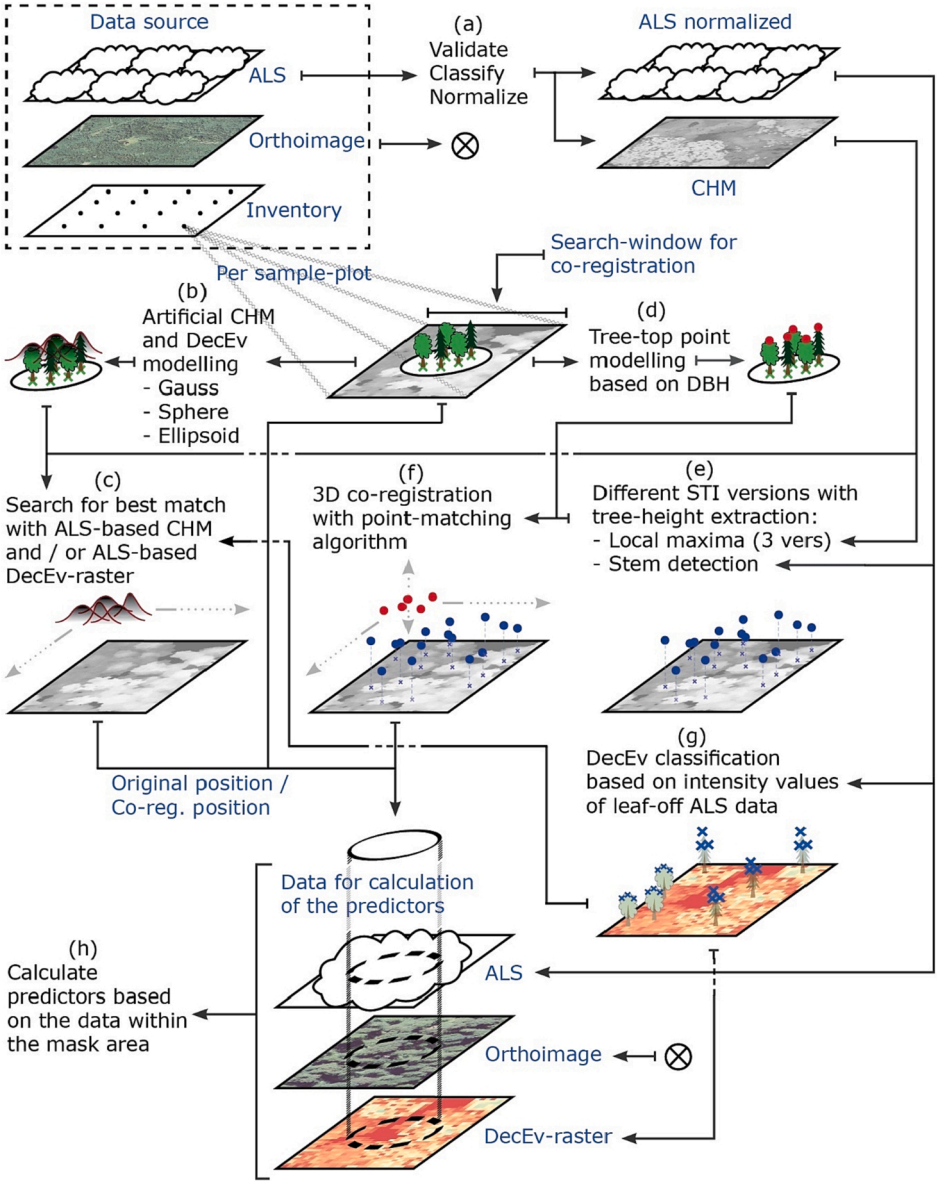


Fig. 4. Illustration of the data-processing steps applied in this study.

the ALS point-cloud (section 2.10). The two tree-top point-clouds were aligned with a simple 3D-point-matching algorithm (section 2.11.2; Fig. 4f). To limit the computational burden, a slightly smaller search window of 60 m for theoretical sample-plot positions and 30 m for GNSS / dGNSS surveyed positions was chosen.

In a next step (section 2.9; Fig. 4g), a deciduous–evergreen (DecEv) map was produced based on the leaf-off ALS data. This procedure made use of the fact that first return intensity values of deciduous trees in a leafless condition are smaller than those of evergreen trees.

After this DecEv extraction, predictors were calculated per sample-plot (Fig. 4h). The co-registered and non-co-registered sample-plot polygons were used as masks, so that only the values within this mask were taken into account in the calculation of predictors from ALS data, orthoimages and DecEv rasters (section 2.13).

A method was applied which randomly removed one out of two strongly correlated predictors and subsequently performed a predictor-selection procedure. Finally, an OLS regression model (section 2.14) was used to predict VOL (Bremgarten) or BA (Zurich) from the previously derived predictors. To get an accurate measure of model quality, the leave-one-out cross-validated (LOOCV) root-mean-squared error (RMSE) and adjusted R^2 were calculated for the entire study area. Since the implemented predictor variable selection was a stochastic process, this complete procedure was repeated three times. The quality of the models was ultimately assessed based on the mean RMSE and adjusted R^2 values of these three repetitions.

The presented method enabled the comparison between different point-based and raster-based co-registration methods. Additionally, raster-based (CHM) methods in previous research were limited to the comparison of a selection of pixels close to a tree-top. Performing the co-registration by comparing a complete aCHM and ALS-based CHM was new to the present study. Similarly, the integration of a DecEv raster into the co-registration procedure had not previously been attempted.

2.4. ALS data pre-processing

The entire processing of the ALS data (top-right in Fig. 4) was carried out using the LAStools v. 4.4.2019 software (Rapidlasso GmbH, 2019). After cleaning the ALS data, a digital terrain Model (DTM), a digital surface model (DSM) and a normalized DSM (nDSM or CHM) with a resolution of 50 cm was calculated. Processing details can be found in

the Appendix.

2.5. Cleaning of sample-plots

For Bremgarten data from three ALS flights (2011 leaf-on; 2014 leaf-on and leaf-off) exist. Data from 2011 are in good temporal synchronicity with the terrestrial measurements but have a lower point density. To use the data from 2014 without introducing great uncertainty regarding major changes on some sample-plots (e.g. harvesting), both leaf-on datasets were processed according to section 2.4. Subsequently, for each sample-plot an absolute difference raster was calculated by subtracting the two CHMs. Visual inspection of the difference rasters and comparison with aerial photographs led to the decision to exclude all sample-plots where the mean deviation of the two CHMs was >4.5 m (Fig. 5). Furthermore, five sample-plots were excluded because ALS data coverage did not include the whole search window of 100 m × 100 m.

Since only one ALS dataset was available for the Zurich study area, no sample-plots were excluded.

2.6. Derivation of tree-species-specific allometric relationships

To model the tree shape of all inventory trees, tree-species-specific allometric relationships were derived, which establish a relationship between DBH and tree height and between DBH and crown size. A model was calibrated for each of the separately recorded tree species in Bremgarten (section 2.1.1). The models were calibrated based on the synthesis of the data from WSL and TUM (section 2.2). For each of the tree species, five different non-linear models and one linear model with a quadratic interaction were tested. A general formulation of these models is given in Table 2 (Peng et al., 2001). Each model was calibrated separately for each individual tree species. Subsequently, the model with the lowest Akaike Information Criterion (AIC) value was chosen. For model calibration, the software R 3.6.1 (R Core Team, 2019) with the user interface RStudio 1.2 was used, including the package *minpack.lm* (for nonlinear functions) and the function 'lm' (linear regression) within the package *stats*. As initial parameter values for the numeric minimization of the non-linear regression models, the mean value of the respective parameters in Peng et al. (2001) were used. The resulting best model formulations, as well as the corresponding AIC values, are given in Tables 5, 6 and 7 in the Appendix.

2.7. Artificial CHM (aCHM) computation

For co-registration by searching for the best match between two CHMs, an aCHM raster with a size of 30 m × 30 m and a resolution of 50 cm was created for each sample-plot, based on the terrestrial data. Each tree was modelled, considering the tree species and the respective tree height h and crown radius r (section 2.6) (Fig. 6). The following crown forms were tested:

Table 2
Regression model equations tested to model the species-specific relationship between DBH and tree height and between DBH and crown size, where $\boldsymbol{h}$ stands for tree height and $\boldsymbol{r}$ for crown radius. The parameters marked $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ were calculated during model calibration. The parameter $\boldsymbol{i}$ represents the intercept and was set to 1.3 m for tree height modelling and 0 m for crown size modelling.

Model			Source
Linear model:	$h \text{ or } r = a + b^{*}DBH + c^{*}DBH^{2}$	(1)	this paper
Chapman-Richards:	$h \text{ or } r = i + a(1 - e^{-b^{*}DBH})^{c}$	(2)	Chapman (1961)
Weibull:	$h \text{ or } r = i + a(1 - e^{-b^{*}DBH^{c}})$	(3)	Yang et al. (1978)
Exponential:	$h \text{ or } r = i + a^{*}e^{\frac{b}{(DBH + c)}}$	(4)	Ratkowsky & Giles (1990)
Logistic modified:	$h \text{ or } r = i + \frac{a}{(1 + b^{-1}{}^{*}DBH^{-c})}$	(5)	Ratkowsky & Reedy (1986)
Korf:	$h \text{ or } r = i + a^{*}e^{(-b^{*}DBH^{-c})}$	(6)	Stage (1963)

- Gaussian bell-curve (bivariate normal distribution) with an amplitude of h at the tree-top position (x_0, y_0) and standard deviation r:

$$f(x,y) = h^{*}\exp\left(-\left(\frac{(x-x_0)^2}{2r^2} + \frac{(y-y_0)^2}{2r^2}\right)\right) \quad (1)$$

- Sphere with radius r, with the highest point of the sphere located at a height of h above the terrain at the tree-top position (x_0, y_0):

$$f(x,y) = h - r + \sqrt{\left(1 - \frac{(x-x_0)^2}{r^2} - \frac{(y-y_0)^2}{r^2}\right)^{*}r^2} \quad (2)$$

- Ellipsoid with radius r in the two horizontal dimensions and radius $3^{*}r$ in the vertical dimension, with the highest point of the ellipsoid at a height of h above the terrain at the tree-top position (x_0, y_0):

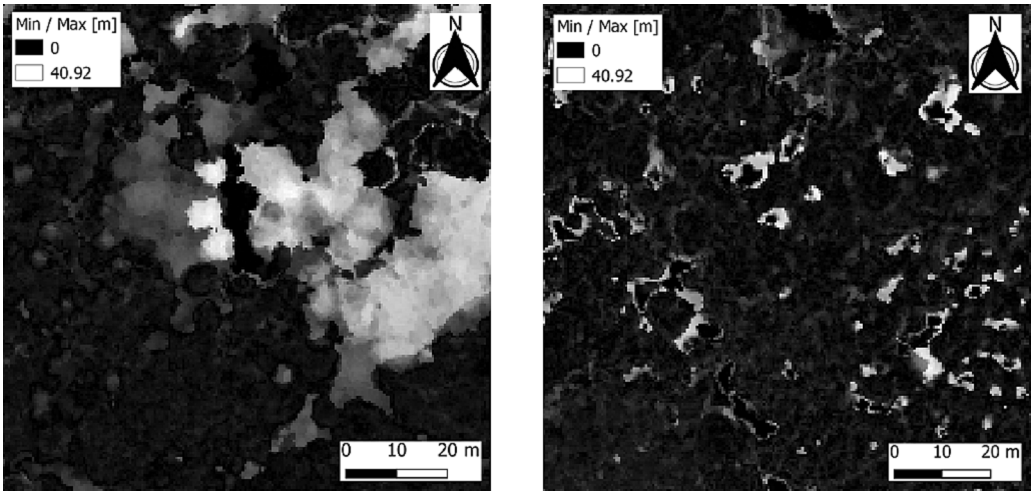


Fig. 5. CHM difference rasters between 2011 and 2014. The sample-plot related to the raster on the left was excluded due to major differences between the years, presumably caused by a harvesting operation. The raster on the right was considered sufficiently similar.

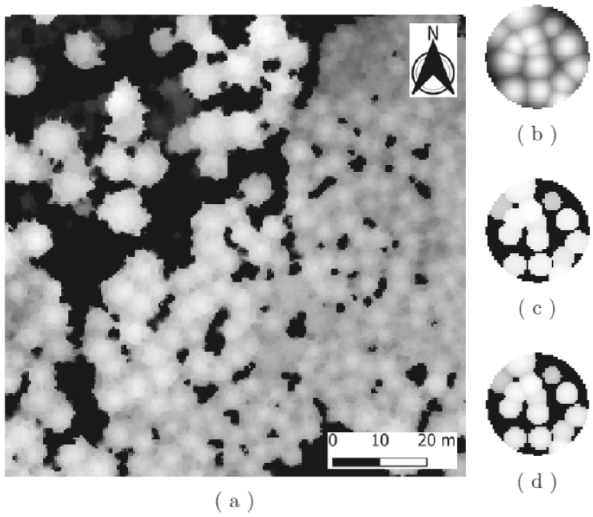


Fig. 6. (a) ALS-derived CHM (100 × 100 m) of a sample-plot in Bremgarten (No. 234'667), (b) – (d) aCHM (diameter 22.57 m), with (b) gaussian aCHM model, (c) sphere model and (d) ellipsoid model.

$$f(x,y) = h - 3^{*}r + \sqrt{\left(1 - \frac{(x-x_0)^2}{r^2} - \frac{(y-y_0)^2}{r^2}\right)^{*}(3^{*}r)^2} \quad (3)$$

2.8. Artificial deciduous–evergreen (DecEv) raster computation

Incorporating deciduous–evergreen (DecEv) information could improve co-registration quality. To generate the artificial DecEv raster based on the inventory data, a check was performed for each cell in the aCHM to evaluate whether the underlying maximum value was based on an evergreen or deciduous tree. Cells with an aCHM value below 1.5 m were set to a noData code in the DecEv raster. During the co-registration process this raster was compared with an ALS-derived DecEv raster (section 2.11.1).

2.9. Deciduous–evergreen (DecEv) raster from leaf-off ALS

The DecEv raster was computed by using the mean return intensity of leaf-off ALS points located high up in the canopy, as proposed in the `Digital Forestry Toolbox` by Parkan (2018). The DecEv Raster had a spatial resolution of 1 m.

For the ALS tile of each individual sample-plot, only medium and high vegetation points (≥50 cm) that were < 1 m below the calculated CHM surface were used to calculate a mean intensity raster with a resolution of 1 m (Fig. 7a). Segments were calculated using the `Digital Forestry Toolbox`'s `slic` algorithm (Fig. 7b). Subsequently, an intensity histogram was calculated and the intensity limits (x) and (y) were set manually (Fig. 7c) to calculate the classified raster (Fig. 7d). Segments with an intensity less than (x) were assigned to a non-vegetation class, those between (x) and (y) were classified as deciduous, and those above (y) were classified as evergreen. The limits (x) and (y) were set to 7 and 18 in Bremgarten and to 17 and 70 in Zurich.

2.10. Single-tree-identification (STI) methods

Various single-tree-identification (STI) methods were implemented in a separate co-registration process. The local maxima (LM) based methods from Menk et al. (2017) [method `LM1`], Kaartinen et al. (2012) [method `LM2`] and Parkan (2018) [method `LM3`] were applied. In addition, stem detection from leaf-off ALS data, designed for deciduous stands, was applied, as proposed by Parkan (2018) in the `Digital Forestry Toolbox` [method `Stems`]. Processing details can be found in the Appendix.

2.11. Co-registration

2.11.1. aCHM-based Co-registration

With theoretical sample-plot locations as initial positions, a search window (default size: 100 m × 100 m, 50 m × 50 m for GNSS measured positions) was used. The aCHM (size: 30 m × 30 m) was moved across the ALS-based CHM in 0.5 m steps within the search range. Thus, the centre position moved by a total of 70 m [± 35 m] for theoretical initial positions, and 20 m [± 10 m] with GNSS measurements. At each position different measures of similarity between the two CHMs were calculated, as listed in Table 3 (except the Simple method). As the co-registered position of the sample-plot, the position with the highest correlation (Olofsson, CCORR_NORMED) or the smallest deviation (SQDIFF_NORMED) was returned (see equations in Table 3).

As shown in section 3.2.1, SQDIFF_NORMED led to promising results. Therefore, subsequent incorporation of forest type information, to improve co-registration, was only tested with SQDIFF_NORMED. For this purpose, the correlation between the DecEv map, based on leaf-off ALS data (section 2.9), and the artificial tree-type map, based on field data (section 2.8), was calculated with the Simple method. Both resulting correlation matrices (forest type, SQDIFF_NORMED CHM) were normalized, weighted and added together, with weighting factors 5.0, 3.0, 2.0, 1.0 and 0.5 for the latter, and the co-registered position of the sample-plot was assigned to the location where the value was smallest.

2.11.2. Point-matching-based co-registration

The methodology used to co-register two point-clouds is illustrated in Fig. 4f. The central idea is to identify the best match between two tree-top point-clouds, where the reference (blue dots in Fig. 4f, remotely sensed trees [RSTs]) is obtained using one of the STI methods described in section 2.10. The tree height is taken from the ALS CHM. The red dots represent the tops of the terrestrial measured inventory trees (ITs), which were calculated with the species-specific relationship between DBH and tree height.

The IT point-cloud was shifted in 1-m steps across the AST point-cloud, with a moving window size of 60 m [± 30 m] for theoretical initial positions and 30 m [± 15 m] with GNSS measurements. At each position the following point-matching algorithm was used:

- Step 1: For each IT, find the RST whose tree-top is closest to that of the IT (3D) and store the respective distance.
- Step 2: Sum the distances of all ITs.
- Step 3: Compare the distance sum with the smallest distance sum measured so far (previous positions in the moving window) and save the position as the new best co-registration position if it is smaller than the smallest distance sum so far.

2.12. Response variable

The response variable in Bremgarten was the local density of the merchantable VOL [m^3 ha^{-1}] (volume with a diameter ≥ 7 cm). It is based on field measurements on which a tariff function was applied, as described in Bont et al. (2020). In Zurich the BA was used as the response variable, which was directly derived from field measurements.

2.13. Predictors from remote-sensing data

Predictors formed the basis for model building and were derived from various sources, such as the normalized leaf-on and leaf-off ALS data (Bremgarten; Zurich only leaf-off), four-band orthophotos and the leaf-off ALS-derived DecEv raster. Based on the orthophotos, the Normalized Difference Vegetation Index (NDVI; equation (4)), the Shannon entropy, and the Grey Level Co-occurrence Matrix (GLCM)

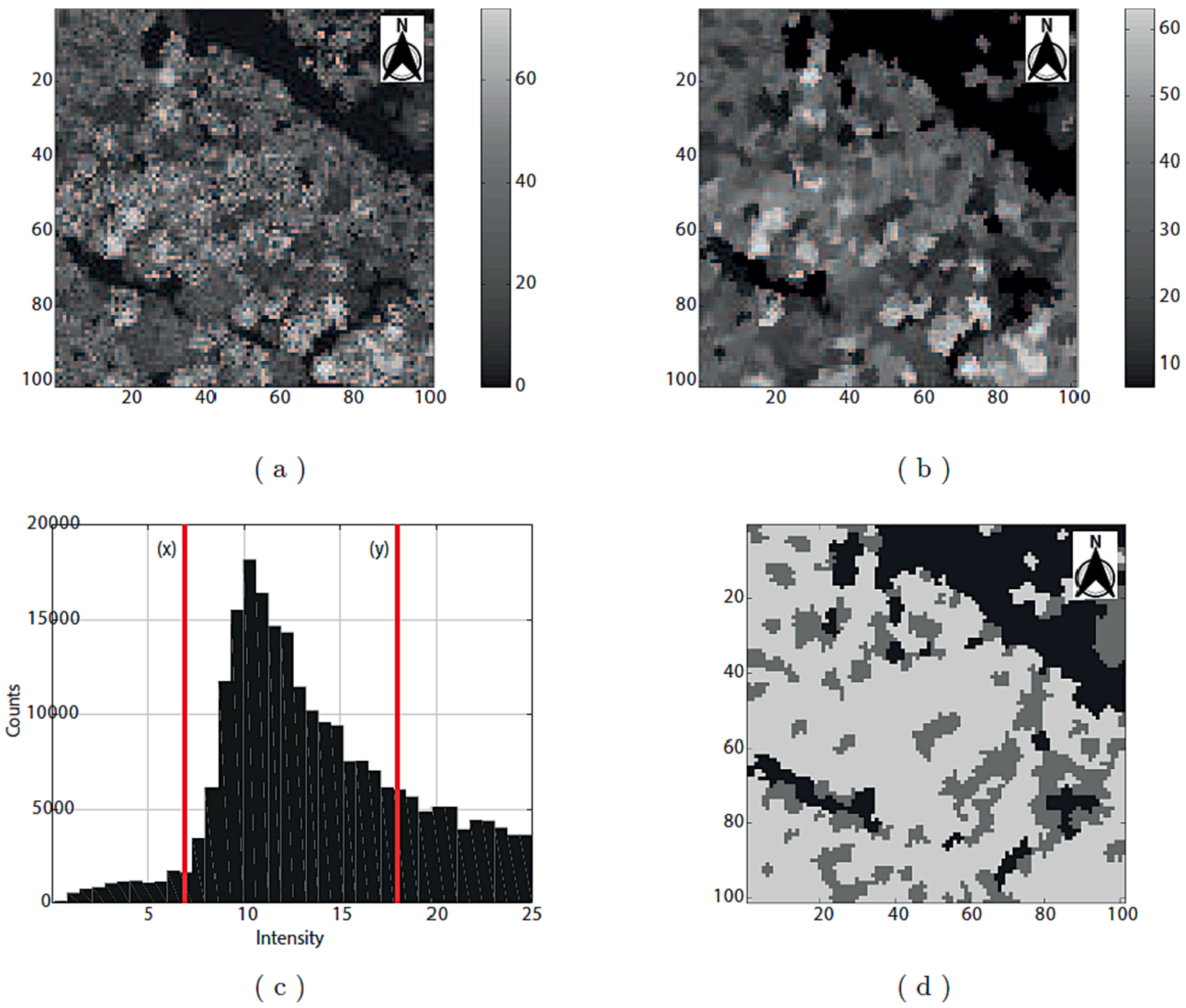


Fig. 7. Illustration of the deciduous–evergreen classification. Panels (a), (b) and (d) show the 100 × 100 m tile around the theoretical sample-plot centre of plot No. 259'680 in Bremgarten. (a) shows the raw intensity raster. (b) shows the segmented version of (a), and (d) shows the same raster classified as *non-vegetation* (light grey), *deciduous* (dark grey) and *evergreen* (black). (c) is the histogram used to set the intensity limits (x) and (y) in Bremgarten.

Table 3
The table shows the different correlation metrics, which were used for co-registration with an aCHM. The correlation (1) was introduced as a simple measure for binary deciduous-evergreen comparison. The correlation measure (2) corresponds to the one used by Olofsson et al. (2008). The other metrics were derived from the `MatchTemplate` function of the `open-cv` python package. $\boldsymbol{p}$ denotes the $\boldsymbol{m}$ pixel of the artificial CHM within the circular moving window. $\boldsymbol{c}$ stands for the corresponding pixels in the leaf-on ALS CHM.$\overline{\boldsymbol{p}}$ and $\overline{\boldsymbol{c}}$ correspond to the respective mean values in $\boldsymbol{m}$ and $\boldsymbol{cor}_{\boldsymbol{dx,dy}}$ is the correlation (2, 3), respectively the deviation (1, 4) at the position $\boldsymbol{dx}$, $\boldsymbol{dy}$.

Name	Formula	
Simple:	$cor_{dx,dy} = \frac{\sum_{i\in m}\sqrt{(p_i - c_i)^2}}{m}$	(1)
Olofsson:	$cor_{dx,dy} = \frac{\sum_{i\in m}(p_i - \overline{p})(c_i - \overline{c})}{\sqrt{\sum_{i\in m}(p_i - \overline{p})^2 * \sum_{i\in m}(c_i - \overline{c})^2}}$	(2)
CCORR_NORMED:	$cor_{dx,dy} = \frac{\sum_{i\in m} p_i * c_i}{\sqrt{\sum_{i\in m}(p_i)^2 * \sum_{i\in m}(c_i)^2}}$	(3)
SQDIFF_NORMED:	$cor_{dx,dy} = \frac{\sum_{i\in m}(p_i - c_i)^2}{\sqrt{\sum_{i\in m}(p_i)^2 * \sum_{i\in m}(c_i)^2}}$	(4)

were calculated.

$$NDVI = \frac{NIR - red}{NIR + red} \quad (4)$$

The Shannon entropy was computed according to Singh & Singh (2008), the GLCM according to Soh & Tsatsoulis (1999). Details can be found in the Appendix. Finally, the deciduous, evergreen and non-vegetation proportions within the masks were calculated from the DecEv raster. All predictors used are listed in Table 4.

2.14. Statistical methods

The aim of co-registration is to achieve a more precise link between different data sources by searching for plausible sample-plot positions within a search window. Predictors are calculated from the remote sensing data at these corrected positions. VOL and BA were modelled as a response variable.

For the analysis, ordinary least squares (OLS) regression models were calibrated. The root-mean-squared error (RMSE; equation (7)) and the adjusted R^2 (equation (6)) after leave-one-out cross-validation (LOOCV) were used to denote the accuracy improvement after co-registration, compared with un-co-registered positions. The adjusted R^2, unlike the

Table 4
This table shows the various predictors which were calculated based on the remote sensing data (ALS and orthophotos). They were subsequently used for model building to derive a model for the response variable (VOL or BA).

	Description	Abbreviation	Unit
Percentile-Values	Specifies the height percentiles of the normalized ALS points within the mask. Example: 95th percentile: p95	pXX; XX = {05, 10, 20, 30, 40, 50, 60, 70, 80, 90}	[m]
	Specifies the intensity percentiles of the ALS points within the mask. Example: 95th percentile: int_p95	int_pXX; XX = {05, 10, 20, 30, 40, 50, 60, 70, 80, 90}	[-]
Quantile-Values	Specifies the quantile of the NDVI-values within the mask. Example: 0.95 quantile: q0.95	qXX; XX = {0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9}	[-]
Minimum	Lowest point in the normalized ALS data within the mask.	min	[m]
	Minimal intensity value of the ALS points within the mask.	int_min	[-]
	Minimal NDVI value of the pixels within the mask.	NDVI_min	[-]
Maximum	Highest point in the normalized ALS data within the mask.	max	[m]
	Maximal intensity value of the ALS points within the mask.	int_max	[-]
	Maximal NDVI value of the pixels within the mask.	NDVI_max	[-]
Mean	Average height of the normalized ALS points within the mask.	avg	[m]
	Mean of the intensity values of the ALS points within the mask.	int_avg	[-]
	Mean of the NDVI values of the pixels within the mask.	NDVI_mean	[-]
Standard deviation	Standard deviation of the height of the normalized ALS points within the mask.	std	[m]
	Standard deviation of the intensity values of the ALS points within the mask.	int_std	[-]
	Standard deviation of the NDVI values of the pixels within the mask.	NDVI_stdev	[-]
Asymmetry (skewness)	Pearson's moment coefficient of skewness of the height of the normalized ALS points within the mask.	ske	[-]
	Pearson's moment coefficient of skewness of the intensity values of the ALS points within the mask.	int_ske	[-]
Curvature (kurtosis)	Kurtosis of the height of the normalized ALS points within the mask.	kur	[-]
	Kurtosis of the intensity values of all the ALS points within the mask.	int_kur	[-]
Shannon entropy	Shannon entropy defined as $H = -\sum_k p_k * \log_2(\max(p_k, 1))$, where p_k is the probability that a pixel has the value k (Singh & Singh, 2008).	S_entropy	[-]
GLCM	GLCM entropy for each band as $H = -\sum_i \sum_j p_{i,j} * \log_2\left(\max\left(p_{i,j}, 1\right)\right)$	XX_glcm_entropy	[-]
	GLCM dissimilarity for each band as $D = \sum_i \sum_j p_{i,j} * \lvert i-j \rvert$	XX_glcm_dissimilarity	[-]
	GLCM homogeneity for each band as $H = \sum_i \sum_j \frac{p_{i,j}}{1+(i-j)^2}$	XX_glcm_homogeneity	[-]
	GLCM ASM for each band as $ASM = (\text{energy})^2$ as $A = \sum_i \sum_j p_{i,j}^2$	XX_glcm_asm	[-]
	GLCM correlation for each band as $K = \sum_i \sum_j p_{i,j} * \frac{(i-\mu_i) * \left(j-\mu_j\right)}{\sqrt{\left(\sigma_i^2\right) * \left(\sigma_j^2\right)}}$ with the mean values $\mu_i = \sum_i \sum_j i * p_{i,j}$ and $\mu_j = \sum_i \sum_j j * p_{i,j}$ and the standard deviations $\sigma_i = \sum_i \sum_j (i-\mu_i)^2 * p_{i,j}$ and $\sigma_j = \sum_i \sum_j \left(j-\mu_j\right)^2 * p_{i,j}$	XX_glcm_correlation	[-]
	With RGB and NIR bands B1, B2, B3, B4 for all GLCM metrics	XX = {B1, B2, B3, B4}	
DecEv	Area share of deciduous trees within the mask	share_dec	[%]
	Area share of evergreen trees within the mask	share_ev	[%]

R^2, also reflects the complexity of the models. To keep the procedure easily adaptable to changes in available predictor input data, expert-based models according to Bont et al. (2020), which do not rely on automatic predictor selection, were not tested.

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (5)$$

$$\text{adj.}R^2 = 1 - \left(1 - R^2\right) * \left(\frac{n-1}{n-(k+1)}\right) \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (7)$$

where y_i are the individual observations, $\hat{y}_i$ are the estimated values of these observations, $\bar{y}$ is the mean value of all observations, n is the number of observations, and k is the number of predictors used.

2.14.1. Removal of highly correlated features

The previously described predictors (section 2.13) were standardized (mean 0, standard deviation 1). Subsequently, a subset of predictors that had a correlation of <0.85 was selected in an iterative automated

process. Details can be found in the Appendix.

2.14.2. Predictor selection procedure

The remaining set of predictors was subject to a feature selection algorithm that relied on the coefficients of the calibrated parametric OLS model. This Recursive Feature Elimination (RFE) algorithm was (n/10)-fold (n-observations) cross-validated (CV) to obtain the optimal number of predictors (Guyon et al., 2002).

2.14.3. Model calibration

The calibration of the OLS model, as well as all previous steps, was performed three times to compensate for fluctuations resulting from the random removal of highly correlated features. Each of these three runs was leave-one-out cross-validated. The accuracy metrics (RMSE, adjusted R^2) were calculated based on the cross-validated results of each run. The mean value of the accuracy metrics of the three runs was ultimately used to assess model accuracy.

3. Results

3.1. Co-registration

The quantitative evaluation of the co-registration, presented in the following section 3.1.1, was kept relatively short. On the one hand, in homogeneous stands of deciduous forest it is impossible to manually co-register sample-plots by visually inspecting CHM or orthoimage rasters. On the other hand, the assessment of the co-registration quality, with certain restrictions (section 4), was possible by measuring the quality of the subsequent VOL or BA modelling. Furthermore, for the accuracy of the subsequent VOL model calibration, the identification of *one* plausible position was considered more important than *the exact* recording position. Qualitative evaluations are presented in sections 3.1.2 and 3.1.3.

3.1.1. Quantitative consideration of co-registration methods

The quantitative evaluation of the co-registration was based on a comparison between the co-registered positions and the location of the relatively exact dGNSS sample-plot positions. The results of this distance evaluation are shown in Fig. 8. For non-co-registered theoretical sample-positions, distances of up to 74 m from the dGNSS position were measured.

The first two panels in Fig. 8 show that, without the integration of DecEv information, the sample-plot positions were on average further away from the dGNSS positions than before co-registration. As reference lines the following distances were observed: mean GNSS position with artificial error: 8.21 m; mean theoretical position: 13.31 m; median

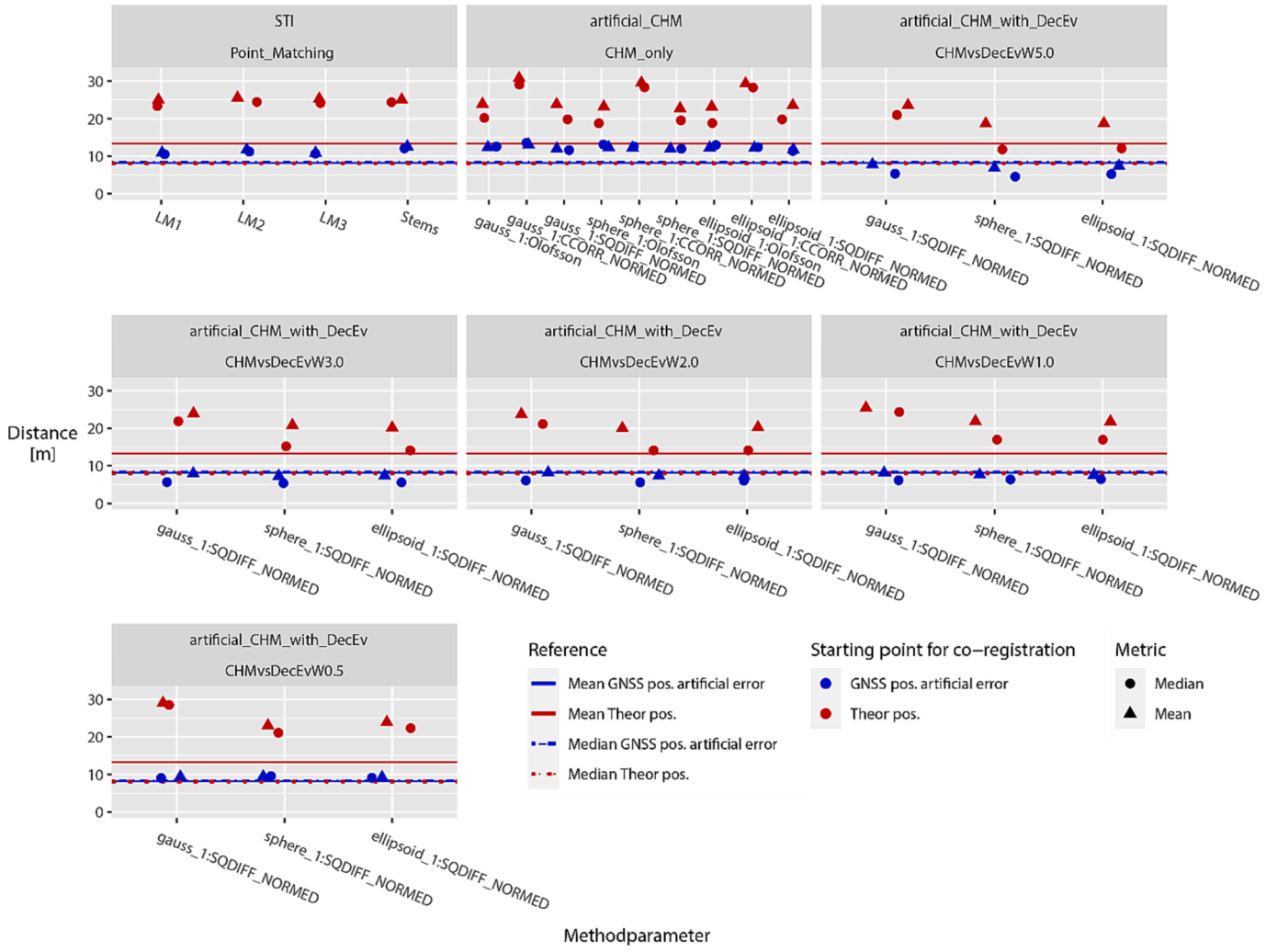


Fig. 8. The mean (and median) distance between the co-registered sample positions and the relatively accurate dGNSS positions. Large distances denote a poor co-registration. The co-registration results which used the theoretical sample positions as starting points are shown in red. The blue data points are based on randomly distorted dGNSS positions (shifted up to 10 m) as initialization of the co-registration algorithm. The non-co-registered reference is shown as a line (red and blue line) in all panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

GNSS position with artificial error: 8.49 m; median theoretical position: 8.10 m. And the following distances after co-registration (mean and median of the first two panels): mean GNSS position with artificial error: 12.02 m; mean theoretical position: 25.47 m; median GNSS position with artificial error: 12.06 m; median theoretical position: 23.00 m. The co-registration method combinations, which used the theoretical sample-plot positions as starting points, generally showed a larger deviation from the dGNSS positions than those which were initialized with the artificially distorted dGNSS positions. This observation is presumably related to the larger search window with the theoretical positions. The largest deviations are associated with the aCHM-based CCORR_NORMED co-registration methods. The finding that the largest distances occurred with one of the aCHM co-registration methods in combination with theoretical sample-plot positions is consistent with the fact that this is where the largest search window was used (100 m CHM section, which means a 70 m shift [± 35 m] of the sample positions; see section 2.11.1). Despite the smaller search window (computing power limitation; 60 m [± 30 m] for theoretical positions), the point-matching algorithm led to mean deviations of up to 25 m. When DecEv information was integrated into the co-registration process, the distances (mean and median) to the dGNSS positions were smaller, especially when the CHM correlation was weighted to be higher than the DecEv correlation with weighting factors 5.0, 3.0 and 2.0 (Fig. 8, subfig. 3, 4 and 5). The mean of the results for weighting factors of 5.0, 3.0 and 2.0 is 7.51 m (8.21 m without co-registration). The corresponding median is 5.50 m (8.49 m without co-registration). Giving higher weights to the DecEv correlation (factors 1.0 and 0.5; Fig. 8, subfig. 6 and 7) increased the distances to the dGNSS positions, especially for theoretical positions as the co-registration starting point. For all weighting factors, the results with distinct tree-crowns (ellipsoid or sphere) were superior to the ones with a Gaussian crown surface. The combination with DecEv co-registration was only tested for the CHM co-registration method SQDIFF_NORMED.

These quantitative considerations are followed by a qualitative assessment of the plausibility of the co-registered position found for the example of Bremgarten sample-plots No. 234'664 and No. 253'683 in section 3.1.2 and No. 259'680 in section 3.1.3.

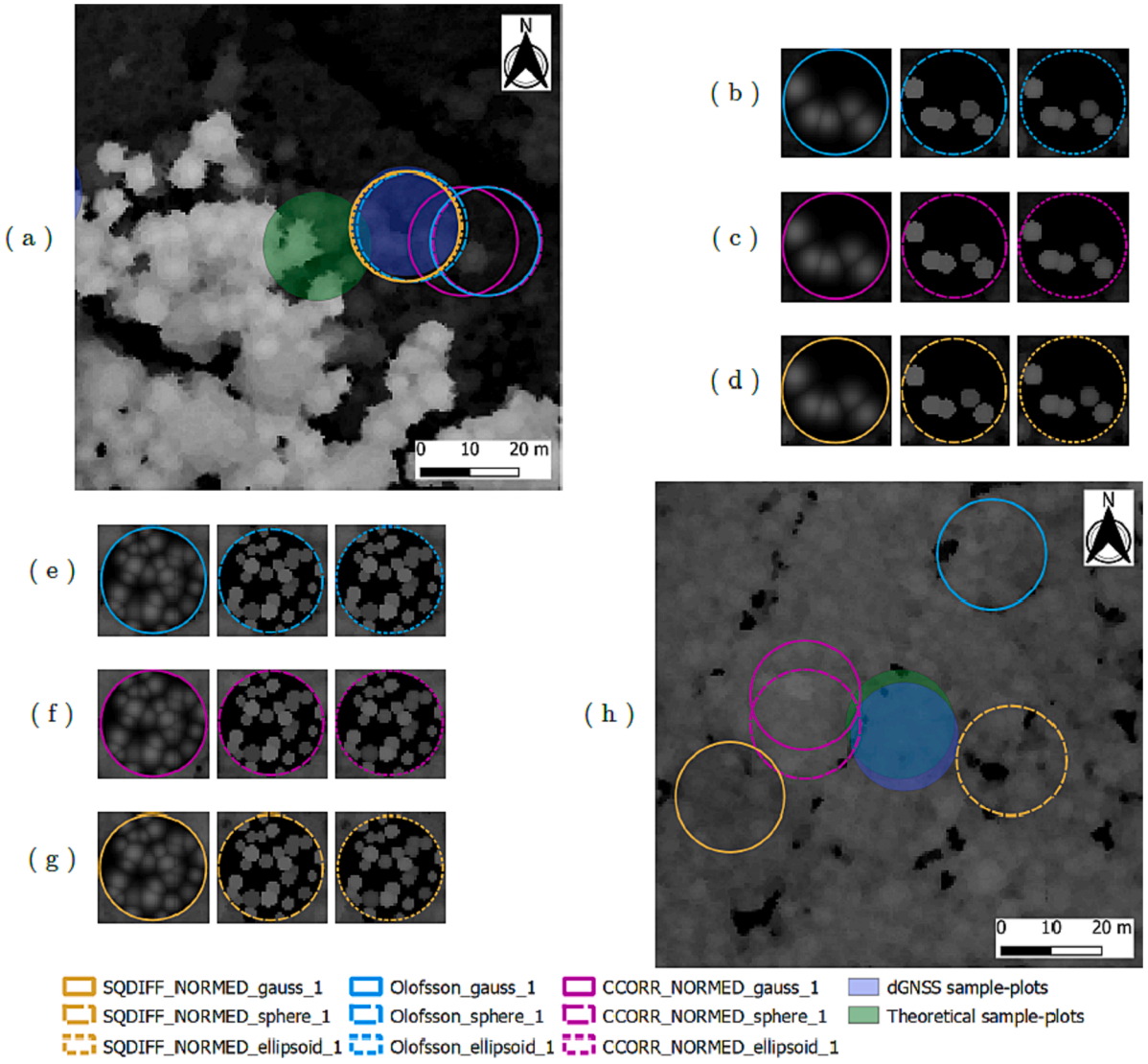


Fig. 9. The leaf-on ALS-based CHM (100*x*100**m**) of the sample-plots No. 234'664 (a) and No. 253'683 (h) in Bremgarten, as well as the corresponding aCHMs, which were derived from the inventory data (b–d, e–g). The positions identified by the co-registration methods, with theoretical positions as initial starting points, are illustrated in colour, overlying the ALS-based CHM. The aCHMs presented above one another in (b)–(d) and (e)–(g) are identical and are only shown to facilitate the visual comparison with the ALS-CHM. The diameter of an aCHM raster and the sample-plot circles was 22.57 m. For some correlation metrics positions were identical, which made the circles drawn underneath invisible.

3.1.2. Examples of aCHM-based co-registration

Fig. 9 illustrates two typical cases of co-registration via aCHMs generated based on inventory data (Bremgarten sample-plots No. 234'664 and No. 253'683). The co-registration used the theoretical sample positions (green) as starting points.

The first graphs (Fig. 9a–d) show sample-plot No. 234'664, with a convincing co-registration outcome. All variants of the SQDIFF_NORMED method were identical with the dGNSS measured position (blue). Visual comparison of the CHM suggests that co-registration led to a plausible result for the other methods as well. All co-registration methods achieved a shift of the sample-plot away from the erroneous position on the edge of an old stand towards a young stand, where the dGNSS located the sample-plot.

Sample-plot No. 253'683 (Fig. 9e–h) serves as an example of a sample-plot where the co-registration was difficult, yet it is rather irrelevant for the derivation of better VOL or BA regression models. This sample-plot lies within a homogeneous stand dominated by oak trees. In the CHM, there were no clear stand boundaries or deviations from the prevailing pattern. As a consequence, the co-registration results differed strongly.

The sphere- and ellipsoid-based SQDIFF_NORMED methods identified a position in the CHM where the gap on the left edge of the aCHM was well represented. When 3D Gaussian-bells were used, such gaps were smoothed and therefore were given less weight. In contrast, Olofsson's correlation method seemed to give more weight to the gap at the left edge of the sample-plot. For the CCORR_NORMED method, it is rather ambiguous why the respective positions were chosen.

3.1.3. Examples of point-matching-based co-registration

Fig. 10 illustrates an example of the co-registration with the 3D point-matching method for sample-plot No. 259'680 in Bremgarten. As the co-registration starting point the theoretical sample position was used. All development stages, from thickening trees to trees at a harvestable age, and both broadleaf and conifer wood are present. In addition, the theoretical sample-plot position is distant from the dGNSS position to the extent that the latter is no longer completely within the search window. These are conditions which make a correct co-registration important but also difficult. Since the algorithm took tree height information into account, the position was shifted southwards where the smaller trees, as well as the dGNSS position, were located. The algorithm identified positions that seem plausible in all four cases. However, if one compares the inventory tree positions with the detected individual trees, none of the single-tree detection methods identified a tree pattern at the dGNSS position with a strong similarity to the inventory tree distribution. It presumably would not have been found even if it was completely within the search window. The chosen implementation of the point-matching algorithm made it favourable for positions in the search window where many detected single trees occurred.

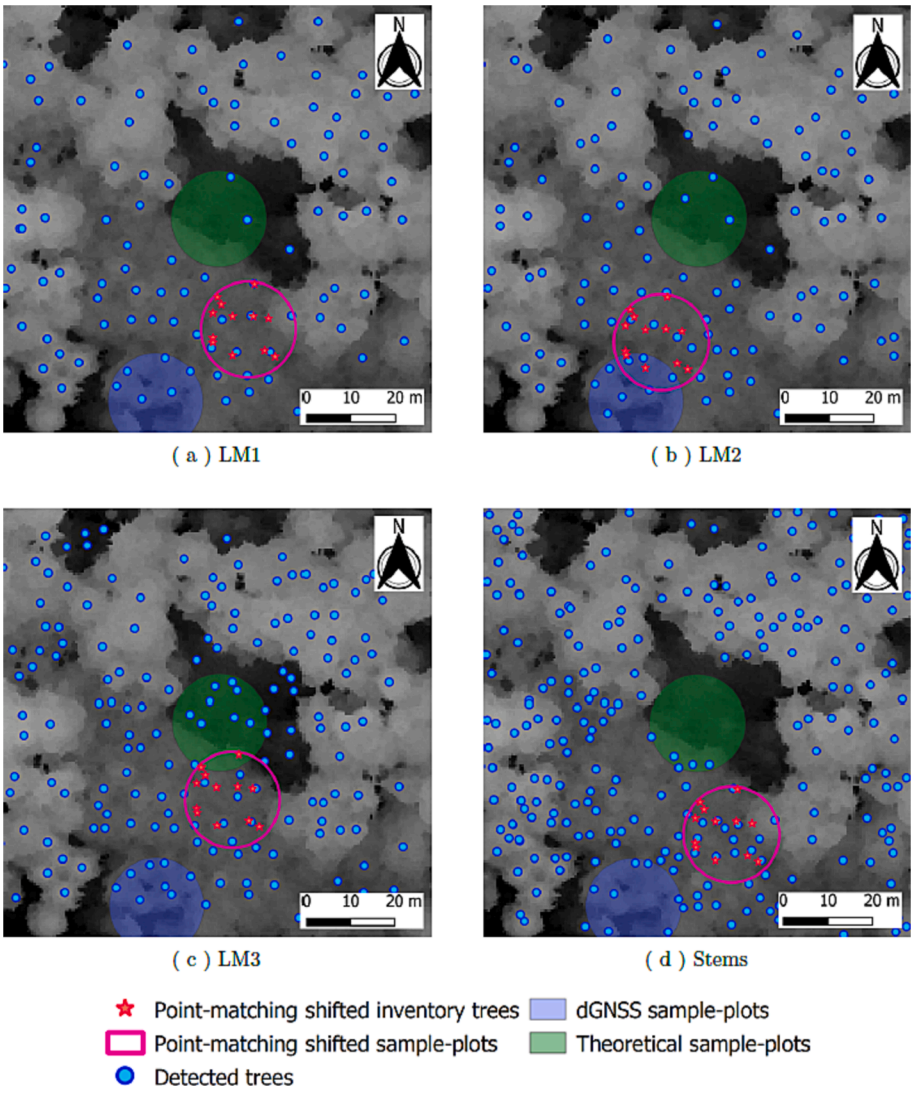


Fig. 10. The leaf-on ALS-CHM (100*x*100 m) of sample-plot No. 259'680 in Bremgarten. Panels (a)–(d) show different STI results in blue, calculated according to section 2.10. Furthermore, the result of the 3D-tree-top co-registration algorithm (section 2.11.2), which matches the red inventory trees to the blue STI-detected trees, is shown in pink. The dGNSS sample-plot position is shown in blue and protrudes slightly beyond the search window with a side length of 100 m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Under such conditions, the likelihood of finding a closely matching detected tree for each inventory tree was high.

3.2. Timber volume (VOL) or basal area (BA) prediction

The co-registration was assessed by comparing the regression model quality before and after co-registration for VOL and BA in Bremgarten and Zurich, respectively. As shown in the Appendix (Fig. 18), the correlation between BA and VOL is almost linear. Therefore, the results (RMSE) can be compared without further effort.

3.2.1. Bremgarten

Fig. 11 gives an overview of the results (RMSE of VOL modelling) of all the OLS regression models with previous predictor variable selection (reduced predictor set) in Bremgarten. Each data point is the average RMSE of three separate leave-one-out cross-validated OLS calibrations, including the described probabilistic predictor variable selection procedure. Applying a co-registration method led to an improved RMSE, provided that the corresponding points were below the respective horizontal reference line without co-registration. The respective adjusted R^2 values are given in Fig. 16 in the Appendix.

The first two panels of Fig. 11 show the RMSE when co-registration was performed without taking DecEv information into account. The other five panels show the result with a different weighting between the match of the CHM pattern versus the deciduous–evergreen pattern. These were therefore an extension of the aCHM matching techniques.

No matter which STI method was chosen (`LM1`, `LM2`, `LM3` or `Stems`), applying a point-matching co-registration method did not improve the results. The second panel in the first line, however, showed promising results. These methods relied on aCHM modelling. The underlying inventory tree-crown model seemed to be less relevant than the correlation metrics used. When the correlation metric of Olofsson et al. (2008) (Methods *_1:Olofsson) was used, the RMSE deteriorated compared with the dGNSS positions. In comparison with the theoretical sample-plot positions, the RMSE remained unchanged. If the CCORR_NORMED correlation was used for CHM co-registration, the results for dGNSS initial positions were similar to the ones with the Olofsson correlation methods. Compared with the theoretical sample positions, the RMSE deteriorated significantly. A noticeable improvement in the RMSE was achieved only when the SQDIFF_NORMED correlation metric was used. Compared with the non-co-registered theoretical sample-plot positions, with an RMSE of 134.4 m^3 ha^{-1}, the co-registration with the best method combination sphere_1:SQDIFF_NORMED led to a 31% lower RMSE of 92.6 m^3 ha^{-1}. Only the co-registration with inventory tree-tops modelled as 3D Gaussian-bells, starting from the already very accurate dGNSS positions, did not improve the RMSE.

Adding DecEv information to the SQDIFF_NORMED correlation method did not substantially improve the respective model accuracy in

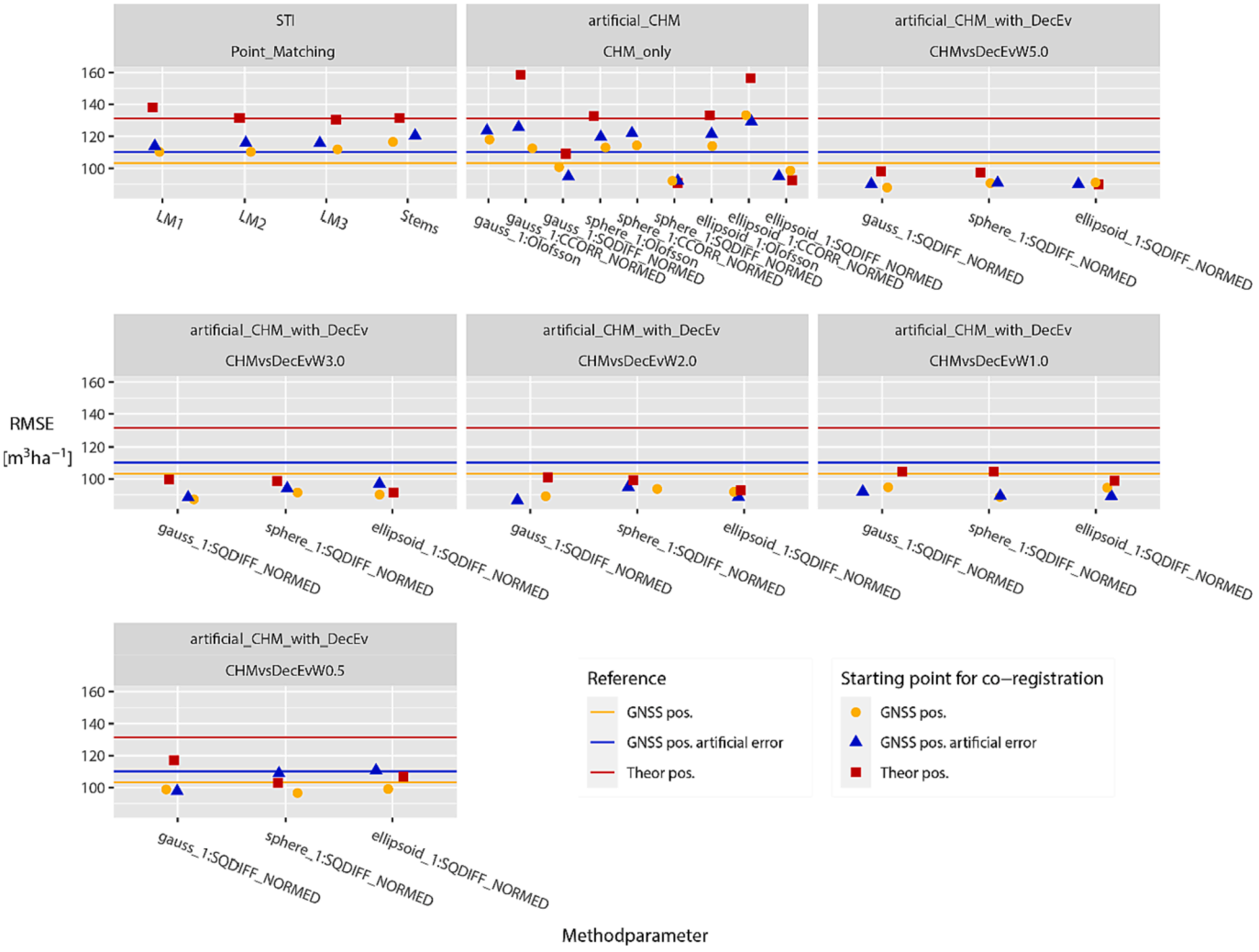


Fig. 11. Comparison of co-registration methods (horizontal axis) based on the leave-one-out cross-validated RMSE (vertical axis) of the timber volume in Bremgarten. The values shown are the average values of three OLS regression model calibration runs, including predictor variable selection and random removal of strongly correlated predictors. The colour indicates the different co-registration starting positions. Horizontal lines show the reference values without co-registration.

terms of RMSE (Fig. 11) or adjusted R^2 (Fig. 16 in the Appendix). However, Fig. 8 shows that adding this DecEv information decreased the mean and median distance between dGNSS positions and co-registered positions, if the CHM correlation was weighted to be higher than the DecEv correlation. This indicates that the co-registration quality improved with regards to the correct identification of sampling positions, but the effect was smaller regarding the subsequently calibrated VOL regression models. However, the integration of DecEv correlation information did not decrease the regression model quality if the weight given to the CHM correlation was higher than that given to the DecEv resemblance. Only for weights equal to or less than one did the quality decrease significantly.

Fig. 12 illustrates the regression model accuracy with predictor variable selection and (b) with and (a) without co-registration (theoretical sample-plot positions) in Bremgarten. For the latter, the difference between the modelled and the measured values was rather high. These variances were in many cases the result of poor matching. Co-registration with the sphere_1:SQDIFF_NORMED method (Fig. 12b)

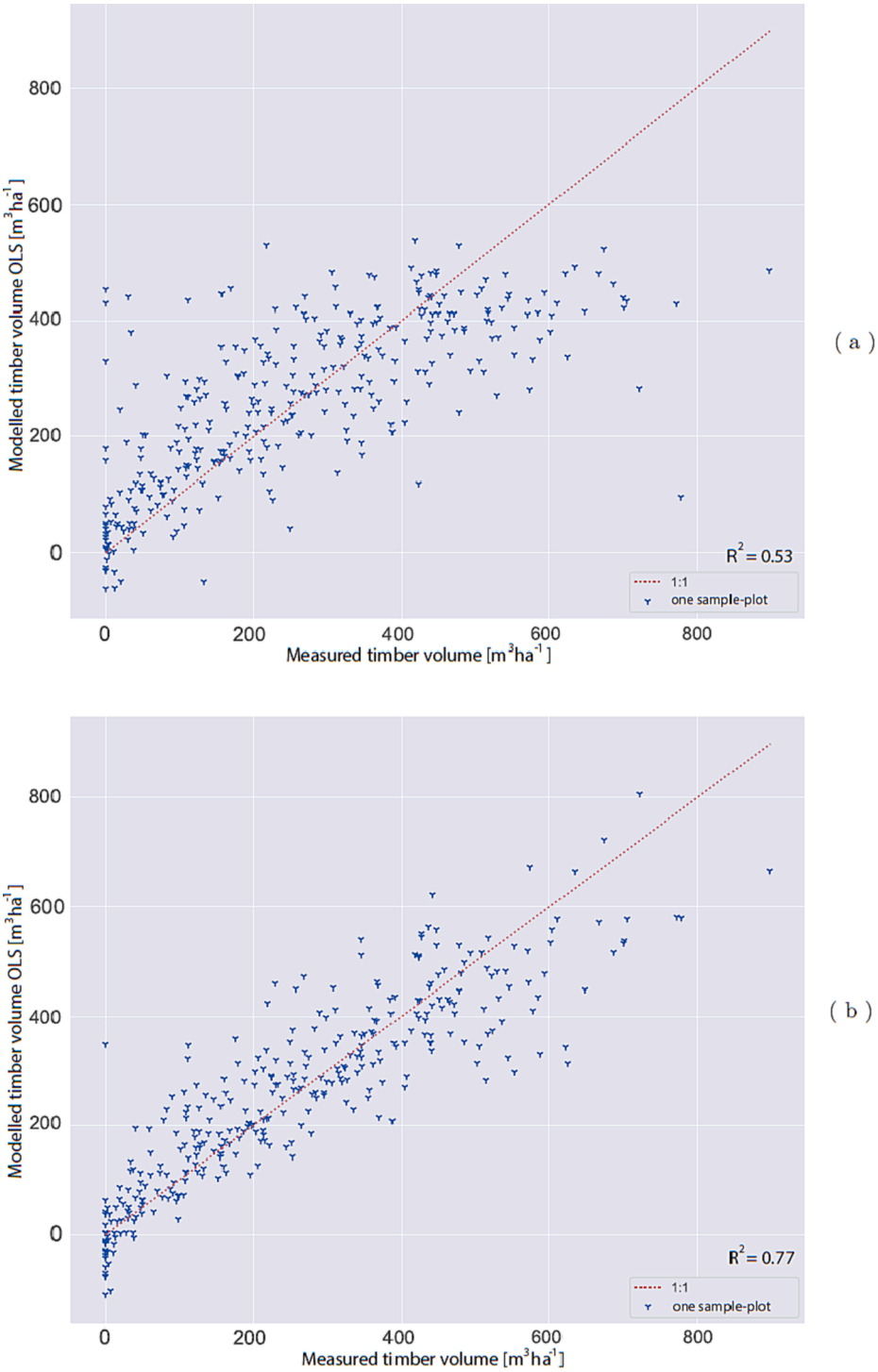


Fig. 12. The observed and the predicted timber volume values after predictor variable selection in Bremgarten. (a) displays the leave-one-out cross-validated values for the non-co-registered theoretical sample-plot positions. (b) shows these values after co-registration with the sphere_1:SQDIFF_NORMED method. All predictors were standardized before the calibration of the OLS model. Data points with a large model deviation are labelled with the respective sample-plot number.

achieved an improvement in R^2 from 0.53 to 0.77 (adj. R^2: 0.51 to 0.76; RMSE: 134.4 to 92.6 m^3 ha^{-1}).

The residuals of Fig. 12b are shown in Fig. 13a. Calibrating an OLS regression model relies on the assumptions that the residuals have a mean of zero, that they are uncorrelated, and that they have a constant variance (homoscedasticity) (Gauss–Markov-theorem; Hallin, 2014). Fig. 13a does not indicate a pronounced correlation between the residuals. As the measured values could not fall below zero, the variance for small VOL tended to be lower than for medium VOL. The residuals' average value was sufficiently close to zero (4.21 m^3 ha^{-1}). Only for high observed VOL, the quantile–quantile graph and the histogram (Fig. 13b) indicated that VOL might have been slightly underestimated. The deviations from the assumptions were, however, rather small.

Co-registration results show that the correct dGNSS position was not always identified, but rather one plausible position within the search window was found. This circumstance calls for additional validation. A systematically incorrect co-registration could result in a bias in the resulting regression models. To test for such a bias, an OLS sphere_1:SQDIFF_NORMED regression model was fitted with the co-registered positions and applied on the *non*-co-registered positions. Fig. 14 displays the results of this validation. The representation of the observed and the modelled VOL at the theoretical sample-plot positions, as well as the histogram and the quantile–quantile graph of the residuals, do not indicate any systematic error in the regression model. The quantile–quantile graph is symmetrical, and therefore the residuals do not correlate with the measured VOL. The histogram shows a slight overrepresentation of positive residuals.

3.2.2. Zurich

For Zurich, only BA data were available, and they were therefore used for modelling. Only the sphere_1:SQDIFF_NORMED co-registration method was tested. All other processing steps were identical to those in

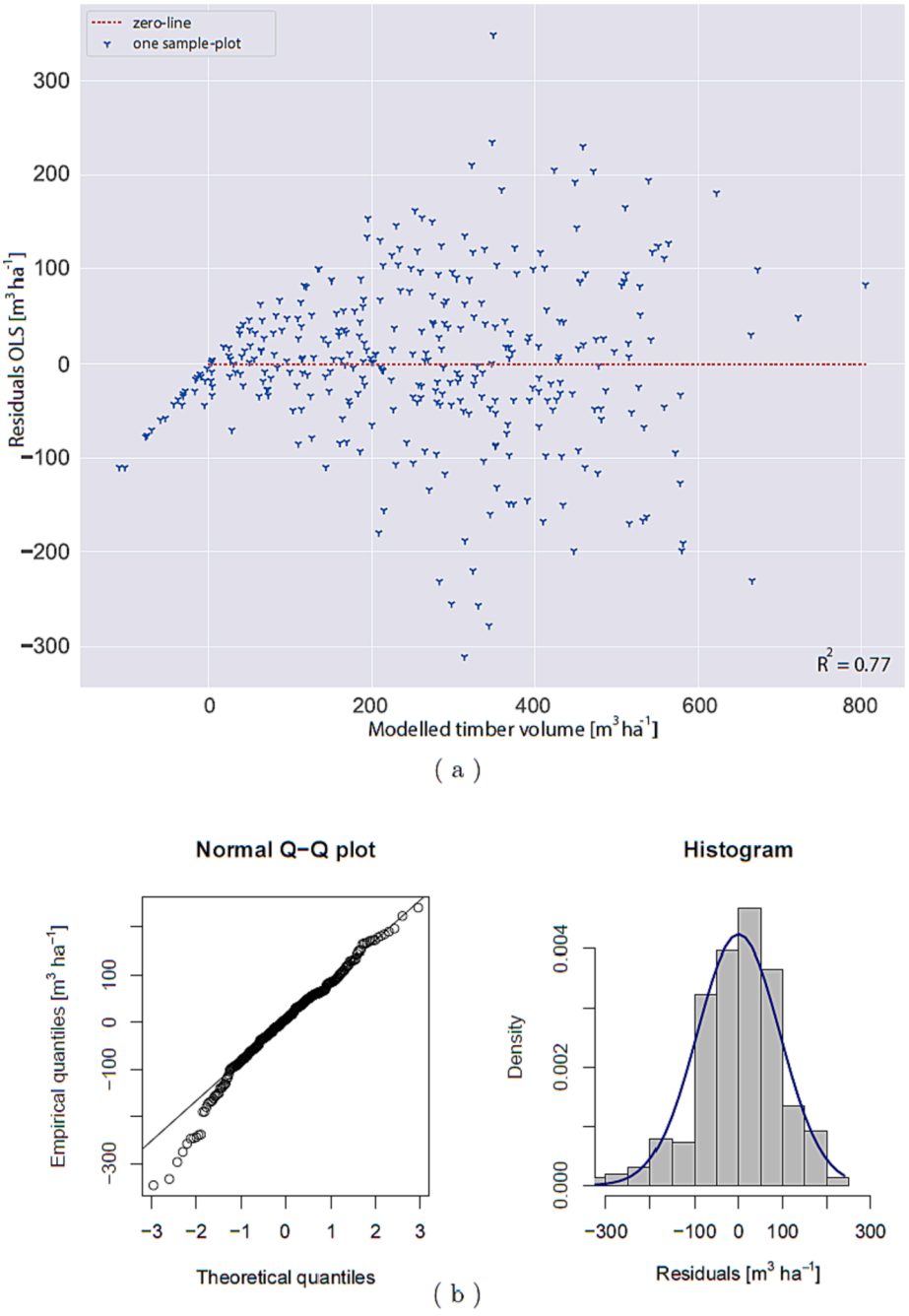


Fig. 13. (a) shows residuals and predicted timber volume values in Bremgarten (same data as in Fig. 12b). Data points with large residuals are labelled with the respective sample-plot number. The quantile–quantile graph and the histogram in (b) are based on the same residuals. The histogram is overlayed with a normal distribution with mean value and standard deviation of the empirical residuals. The sample-plot positions used to calculate the predictors were identified with the sphere_1:SQDIFF_NORMED co-registration method, starting from the theoretical sample-plot positions. All predictors were standardized before the calibration of the OLS model.

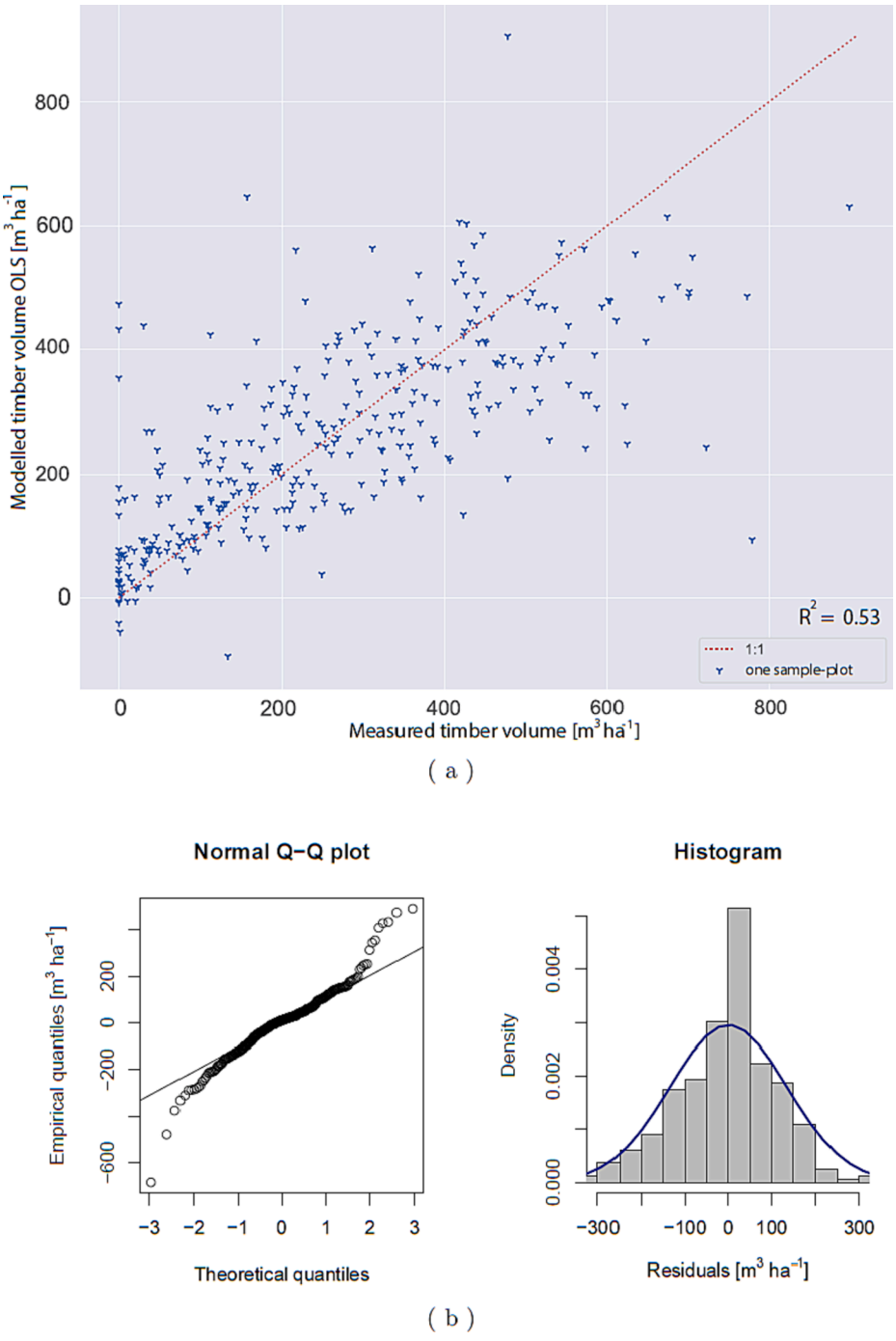


Fig. 14. The observed and the OLS predicted timber volume values after predictor variable selection in Bremgarten. The regression model was calibrated with the predictors calculated at the sphere_1: SQDIFF_NORMED co-registered sample-plot positions, and that model was used to predict the timber volume at the non-co-registered positions (a). All values are based on a leave-one-out cross-validation. The predictors were standardized before the calibration of the OLS regression model. The histogram is overlayed with a normal distribution with mean value and standard deviation of the empirical residuals. The graphics are suitable to check whether co-registration led to a systematic bias in the regression model.

Bremgarten. All regression model types were calibrated three times, and the mean value of the leave-one-out cross-validated RMSE and adjusted R^2 were used as quality measures. As can be seen in Fig. 15, co-registration improved the OLS regression model accuracy for all initial sample-plot positions, when the DecEv map was not considered for co-registration. For theoretical sample-plot positions, the RMSE decreased from 13.56 m^2 ha^{-1} without co-registration to 12.18 m^2 ha^{-1} with sphere_1:SQDIFF_NORMED co-registration (1.38 m^2 ha^{-1} or −10%). The integration of the DecEv map deteriorated the co-registration quality for all the tested weighting factors. The respective adjusted R^2 values are given in Fig. 17 in the Appendix.

4. Discussion

4.1. Co-registration

For the calibration of regression models, which establish a connection between terrestrial measurements and remote sensing data, an accurate spatial matching is necessary. In this section, the results of the various co-registration procedures are analysed.

To achieve a performant regression model, the identification of the *exact* recording position is only of secondary importance. It is, however, essential to identify *one* position for each sample-plot that is consistent with the inventory data. Whether the *exact* position of each sample-plot is found depends on the size of the search window and the associated assumed accuracy of the initial position for co-registration. The

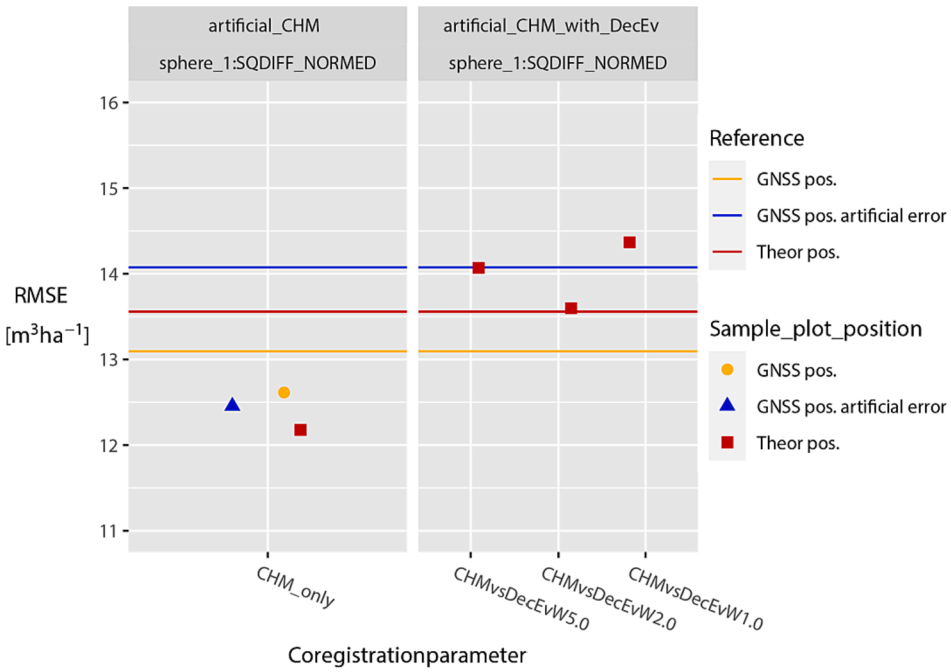


Fig. 15. Application of the co-registration method, which performed best in Bremgarten, to the Zurich dataset. Integration of the DecEv raster was only tested with the theoretical sample-plot positions. The figure shows the leave-one-out cross-validated RMSE (vertical axis) of the OLS modelled basal area (BA) in Zurich. The reference is given by the non-co-registered sample-plots (horizontal lines). All values are average values from three independent calibration runs, including predictor variable selection and random removal of strongly correlated predictors. The colour denotes the respective co-registration starting positions.

heterogeneity of the forest area in the search window plays a major role as well. Homogeneous stands or large search windows increase the probability that several plausible positions will appear within the search window. Positional errors of up to 74 m compared with dGNSS measurements in Bremgarten made large search windows necessary, which in turn meant that several plausible sample-plot positions within the window were likely. As a consequence, co-registration did not necessarily identify the true recording position.

4.1.1. Distance to dGNSS

As described above, it was expected that the co-registered positions would not necessarily be close to the dGNSS measured sample-plot positions. A comparison of the average distance to the dGNSS position, which used either the theoretical sample-plot positions or the artificially distorted dGNSS positions as starting points, confirmed this expectation. Due to the larger search window with theoretical positions, the average deviations after co-registration were greater than with distorted dGNSS initial positions. Without consideration of the DecEv map, none of the methods achieved an improvement with regard to the distance to the dGNSS measurement. However, when the error of the initial position was rather small (distorted dGNSS), the integration of the DecEv information into the CHM-based co-registration considerably reduced the median, but also the mean distance error, especially when the CHM correlation weight was higher than the DecEv weight. This indicates, that, within a small search window, the DecEv information made it possible to considerably narrow the possible CHM positions, down to the true one in many cases. However, the CHM information is still more important than the DecEv accordance. For large search windows, the positive effect of DecEv integration diminished, likely because multiple similar positions appeared within the search window.

4.1.2. Qualitative assessment of CHM co-registration

The quantitative observations were confirmed by the qualitative results, as described in sections 3.1.2 and 3.1.3. For the interpretation of the results of the aCHM co-registration, a basic understanding of the correlation equations is essential (Table 3). The method Olofsson includes an image-wise normalization step. As a result, this correlation equation (Table 3, formula 2) was unable to detect absolute height differences among the two CHMs. If a CHM was copied and all values increased or decreased by a few meters, the two CHMs would still correlate perfectly. The other two correlation equations (CCORR_NORMED and SQDIFF_NORMED) lack such a normalization with the average value of all pixels of an image. These were thus suitable to detect absolute height differences between the two CHMs.

As illustrated with examples in section 3.1.2, there were sample-plots for which all aCHM co-registration methods led to plausible results, as well as some for which the identified positions were less plausible and therefore some methods were less suited for regression modelling of the VOL. In particular, the qualitative examples confirmed that the *true* recording position in the search window was not always identified, but rather *one* plausible position was identified. This observation is consistent with the observed increase in mean distance to the dGNSS position after co-registration.

As the first example in Fig. 9 shows, the probability that the *true* sample-plot position was identified was higher in a heterogeneously structured forest. In contrast, the second example in Fig. 9 shows that the correct localization was almost impossible in a homogeneous forest. However, under such conditions, the positioning of the sample-plot is rather irrelevant for model fitting, as these metrics are similar within the entire search radius.

Visual inspection suggested that the modelled tree-crowns were in some cases rather small, at least in the spherical or ellipsoid case. An inadequate crown size can be problematic, if this leads to a systematic shift of the sample-plot towards positions with lower or higher average CHM values. Using such a co-registration to calibrate regression models would lead to a systematic over- or underestimation. The relevance of this issue can be checked by visually comparing the CHM and aCHM of some sample-plots that are easy to co-register. On average, the CHM and aCHM should have similar values at the correct position. If this is not the case, a crown-size scaling factor could be introduced, or another crown model could be implemented.

4.1.3. Qualitative assessment of point-matching co-registration

In section 3.1.3, point-matching co-registration examples were described. Visual inspection led to the conclusion that the SQDIFF_NORMED aCHM-based co-registration methods attain more plausible results than the application of a point-matching algorithm. The visual comparison of the inventory tree-top point-cloud with the detected tree-

tops at the dGNSS position in Fig. 10 indicated that, even with an optimal point matching algorithm, the dGNSS position would probably not have been identified, as the two point-clouds differed significantly. The basic problem is that all STI methods do not find all trees and that they locate the stem positions with too little precision. Since the tested 3D point-matching algorithm also included tree height information in the co-registration process, it could be expected that positions are identified where the estimated height of the forest stand matches the height of the inventory trees reasonably well. However, this information is already included in the faster aCHM co-registration. Whenever individual tree-crowns are hard to identify due to a relatively flat CHM with fuzzy LM, an aCHM method has the advantage that the influence of an individual LM pixel on the correlation result is smaller than if only distances between LM points and inventory tree-tops are compared.

Finally, the results showed that point-matching-based methods are not superior to aCHM methods, as long as no accurate STI methods are available or additional information on the single-tree level is used (e.g. crown margin points from ALS). If artificial point-clouds were simulated based on the inventory data and if they were compared with the observed ALS point-cloud by means of the point-matching algorithm, an STI, including all its uncertainties, could be omitted.

4.1.4. Relationship to other studies

Studies about co-registration of forest inventory and remote sensing data, which have been published in recent years, assume positional errors of<20 m and thus assume at least GNSS or dGNSS measurements are available. Only Hauglin et al. (2014) also tested search radii between 20 and 65 m. Using a point-matching algorithm, like the one in the present study, they achieved a co-registration error of<0.5 m for 78% of all sample-plots, if a search radius of 25 m was used. The search window used in the present study for dGNSS positions with artificial distortion corresponds approximately to the search radius of 25 m in Hauglin et al. (2014). The median distances to the approximately correct dGNSS position indicated that even the best co-registration method SQDIFF_NORMED, was less accurate than the one in Hauglin et al. (2014). Besides a slightly more sophisticated algorithm, the almost exclusive abundance of spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*) probably facilitated co-registration in the study of Hauglin et al. (2014).

Lamprecht et al. (2017) used real inventory samples to generate artificial forest stands and used an artificial ALS point-cloud for each of these forest stands. From these forest stands, inventory data were simulated and used as ground-truth in the subsequent modelling process. The position of the simulated inventory data was distorted with an artificial GNSS error. At these positions, an LM-STI was performed with the ALS point-cloud. Afterwards, probabilities of agreement between the trees of the simulated samples and the trees detected as LM were calculated and used for co-registration. The match probabilities were based on different distance metrics between the individual trees. This point-matching co-registration method is similar to the one applied in the present study, but it uses a more complex single-tree matching method and permits rotation of the inventory plot by superimposing the single-tree displacement vectors. With this method, the authors achieved a co-registration overall accuracy of 82.7%, where a co-registration was considered successful if >50% of all single trees were correctly assigned (single-tree assignment overall accuracy = 89.1%). Due to the lack of ground-truth reference data in the present study, no quantification of how often the point-matching algorithm correctly matched a pair of trees was done. The median and mean distances to the dGNSS positions, however, suggest that the frequency of correct allocation of both the sample-plots and the individual trees must be smaller than in the study of Lamprecht et al. (2017). The degree of simplification of the co-registration problem by simulating the ALS point-cloud, as done in Lamprecht et al. (2017), compared with using real ALS data remains to be clarified. The random addition of residuals to the ALS points generated based on the simulated crowns in Lamprecht et al. (2017) may lead to a more realistic ALS point distribution, but the crowns would remain approximately symmetrical. Therefore, unilateral, asymmetric crowns (especially broadleaf trees) would likely be underrepresented.

The co-registration process described by Olofsson et al. (2008) is similar to the one described as aCHM co-registration with the Olofsson correlation coefficient and gauss_1 crown shape described in this paper. However, Olofsson et al. (2008) used an LM-STI to model the detected single trees in the same way as the inventory trees. This STI step was omitted in the present study, and the aCHM was directly compared with the CHM. Olofsson et al. (2008) showed, with their simulations, that with a standard deviation of the STI positional error < 1 m and an STI omission and commission error of 20%, 70% of all inventory trees were correctly linked to the corresponding remotely sensed trees. The proportion of correctly co-registered sample-plots, even if not specifically quantified, is likely to have the same order of magnitude. The median and mean distances to the dGNSS positions presented in the present paper are rather high. Therefore, the achieved proportion of correctly assigned individual trees is significantly lower than in Olofsson et al. (2008) (Fig. 8). The results of Olofsson et al. (2008) are relativized by the fact that they assumed an STI positional error of < 1 m. This is likely the case for conifer stands, but rather unlikely for broadleaf stands.

Apart from slight adjustments regarding the weighting of individual pixels when comparing the CHM with the aCHM and the renunciation of an STI in the CHM, as well as adjustments to the sampling design (angle-counting sampling according to Bitterlich 1952), the co-registration method according to Dorigo et al. (2010) is identical to the method according to Olofsson et al. (2008) and to the aCHM-based co-registration applied in the present study. Despite a relatively large search radius of 30 m, Dorigo et al. (2010) successfully co-registered 67 out of 98 sample-plots with a positional error < 5 m. Such an accuracy was not achieved in the present study. However, Dorigo et al. (2010) excluded many broadleaf-dominated stands, as they could not be co-registered manually.

The co-registration method presented by Monnet and Mermin (2014) is similar to the ones by Olofsson et al. (2008) and Dorigo et al. (2010) and the one applied in the present study. Monnet and Mermin (2014) showed that the average positional accuracy could be improved compared with conventional GNSS measurements. However, the standard deviation of the distances to the correct positions increased, as incorrectly co-registered sample-plots were frequently positioned far from the correct locations. Compared with dGNSS measurements, the co-registration decreased the positional accuracy. The data in the present study (Fig. 8) show that co-registration, based on the artificial and uniformly distorted dGNSS positions, did not lead to any improvement in the average positional accuracy.

The above comparisons with other studies show that the co-registration methods implemented here, despite the similarity with the methods of other studies, found the correct position less often. This may result from a less sophisticated point-matching co-registration or a slightly too small crown size for the aCHM-based methods. On the other hand, the high proportion of broadleaf stands in Bremgarten probably increased the uncertainty of the co-registration results.

4.2. Timber volume (VOL) and basal area (BA) prediction

As shown in section 3.2.1, none of the point-matching methods could achieve a systematic improvement in accuracy of the subsequent VOL modelling with the Bremgarten data. From a theoretical point of view, it is not obvious why the point-matching algorithm did not achieve an improvement in regression accuracy, which is comparable to a good aCHM method. The distances given in section 3.1.1 show that the correct sample position was not always identified, and, as described in section 4.1.3, the STI accuracy was likely too low. However the use of 3D tree-top distances to quantify the co-registration quality ensures that the forest-stand height is approximately the same as that of the inventory trees at the identified positions. Extreme deviations, which occur, for

example, when sample-plots are accidentally located in a mature forest stand instead of a young growth stand, should be eliminated by the algorithm. Nevertheless, this factor was not enough to significantly improve the regression accuracy. A possible reason could be the previously described problem that the algorithm tends to place sample-plots at locations with many detected trees.

The co-registration with an aCHM and the SQDIFF_NORMED metric led to an improvement of the regression model accuracy of 31 RMSE-%, from 134.4 m^3 ha^{-1} to 92.6 m^3 ha^{-1} in Bremgarten. This result was confirmed by the modelling of the Zurich inventory data, at least if the DecEv raster was not used for co-registration. The fact that the `Olofsson` correlation metric gave poor results when tree position patterns occurred several times within the search window could be expected given that the correlation equation (Table 3) is not able to detect absolute CHM differences. This means that mature wood cannot be distinguished from a young growth stand if the CHM surface has a similar relative height pattern. The method will thus only lead to good results in open stands with gaps and small search radii. Such a limitation does not exist for the correlation equations CCORR_NORMED and SQDIFF_NORMED. However, they both use a standardization term in the denominator, which gives the sum of the CHM differences in the numerator more weight if one of the two CHMs has a low squared-sum of all pixel values. Using the difference-sum of the pixel values (SQDIFF_NORMED) was superior to using the product-sum of the pixel values (CCORR_NORMED). One possible reason is that the CCORR_NORMED correlation tended to move the sample-plots to positions where the CHM values were rather too high relative to the aCHM values. Such a trend was not observed with SQDIFF_NORMED. It would be worth investigating whether the inclusion of the normalization term in the correlation equations leads to better results, given that the observed height range is wider for large trees than for smaller trees. With the ALS-CHM the height error is independent of the absolute CHM values, which makes the chosen standardization less appropriate. It would be of interest to test an adapted SQDIFF_NORMED correlation equation, in which only the aCHM height is used for normalization:

$$cor_{dx,dy} = \frac{\sum_{i \in m}(p_i - c_i)^2}{\sum_{i \in m}(p_i)^2} \quad (8)$$

where p are the pixels of the aCHM and c the pixels of the CHM.

Besides testing other correlation equations for co-registration, a crown-size scaling factor or alternative crown shapes, for example according to Pretzsch (2001), could lead to better co-registration results and thus better regression models. This work has shown that the modelling of tree-crowns as Gaussian-bells is rather unsuitable. A possible explanation is that the information contained in the tree position is blurred when using a Gaussian-bell. A horizontally clearly defined geometry keeps more of the positional information. Instead of direct modelling of the crown shape as a Gaussian-bell, a subsequent smoothing of other crown models with a Gauss-filter could be tested. Furthermore, during the validation of most regression models, it was found that the VOL of samples in Bremgarten with an observed VOL of zero was systematically slightly overestimated. This is plausible because the DBH threshold for recording was 12 cm in Bremgarten. Therefore, an inventory value could be zero even though the actual VOL was slightly larger than zero. Excluding sample-plots with a VOL of zero from the model calibration might improve the model quality.

As previously described (section 4.1.1), the integration of the DecEv information into the co-registration process lowered the mean and median deviation between the co-registered positions and the dGNSS measured positions. When the weighting factor for the DecEv raster was smaller than the one for the CHM, the inclusion of the DecEv information had no influence on the regression model quality in Bremgarten. Unfortunately, for the Zurich dataset the inclusion of the DecEv raster in the co-registration process deteriorated the model quality in terms of RMSE. A possible explanation is that the ALS data were recorded in mid-April, when some deciduous trees might have already grown their foliage. Therefore, the quality of the leaf-off ALS data was probably lower in Zurich than in Bremgarten. The DecEv results in Bremgarten, especially in terms of positional error, make the integration of ALS-based DecEv rasters into the co-registration process advisable whenever ALS data that were clearly collected completely under leaf-off conditions are available.

4.3. Forest inventory and co-registration: Advantages and pitfalls

If a co-registration method accurately identifies the position for a large proportion of the sample-plots, it improves, in particular, the accuracy of regression models connecting field measurements with remote sensing data. If these prerequisites are met, co-registration is always recommended. However, cases in which co-registration identifies a *plausible but not true* sample-plot position must be discussed thoroughly.

Frequently, good performance of the co-registration methods described in literature has required the exclusion of sample-plots with a difficult co-registration (e.g. dense broadleaf stands) from the investigation, or a small proportion of broadleaf trees in the study area (section 4.1). One aim of this study was to show that an improvement in the modelling of forest indicators through co-registration is also possible if the inventory contains sampling areas that are difficult to co-register. For this reason, no sample-plots were excluded in advance and a search window was chosen, which should ensure that the actual recording position was within the search window for almost all sample-plots. Visual inspection of the identified tree-tops (point matching) or the CHM (aCHM co-registration) indicated that similar patterns occurred several times in the search window and that the identified tree-top pattern at the true position (dGNSS) did not always match the inventory data well. A manual co-registration would have been possible only with great uncertainty for many sample-plots. It was therefore expected that an automated co-registration procedure would not necessarily identify the correct position, but rather a suitable position with a similar tree distribution and tree height pattern in the search window of each sample-plot. As such, the true connection between the two data sources was no longer given, and the question of whether regression models based on this information can be used for further modelling in the framework of multi-phase forest inventories must be discussed in detail.

Since the SQDIFF_NORMED correlation method attained the best VOL and BA regression results, the following discussion will focus on this co-registration metric. In general, it would be a problem if the response variable in the co-registration process had a direct influence on the values of the response variables. This would lead to a selective choice of positions whose predictors would be optimally matched to the response variable, without necessarily representing the true conditions. Since the VOL correlates with the aCHM values and the values of aCHM should correlate with those of the CHM, an indirect link between the response variable and the predictors is given, but this correlation is ultimately necessary and desirable in order to achieve a meaningful calibration of a regression model. Because the response variable (VOL or BA) has no direct influence on the co-registration process, it cannot directly control the values of the predictors. If the co-registration process finds positions at which the calculated predictors perfectly explain the VOL or BA, despite leave-one-out cross-validation (LOOCV), the derived regression model is suitable, even if the true sampling positions were not found.

Caution is advised whenever the co-registration could lead to a systematic bias in the linkage of the two data sources. This is theoretically possible and best explained by means of an example. Assuming a co-registration method *systematically* identifies positions at which the forest stand height is too small (e.g. too small tree height or crown size in the aCHM), a calibrated regression model perfectly explains the putative VOL at this position. However, if the model is used to make predictions for areas with independent data, it overestimates the VOL *systematically*.

This systematic bias is possible with the modelling of all forest indicators if they are in any way related to the metric used for co-registration (CHM height in the aCHM- and point-matching case or stem number in the point-matching case). A major problem would be that the mean value of the quantity to be estimated would be distorted systematically and the confidence interval would remain small. This would lead to a false high precision. Less problematic is the situation when the identification of *plausible* but *not true* positions leads to a greater spread in the regression model. In this case, the residuals remain symmetrically distributed around the true mean value. Although this is undesirable, since the confidence interval of the estimate is enlarged, no false accuracy is given. As this comparison indicates, a systematic bias can be tested on the residuals, when the model calibrated at the co-registered positions is used for the prediction of the VOL at the initial positions. Fig. 14 shows an example of such a verification for the best-performing co-registration method sphere_1:SQDIFF_NORMED. No systematic bias was found.

If such a review of residuals were to show that a systematic bias exists, it is advisable to check whether the distortion can be eliminated by adjusting the parameters of the co-registration method. With the aCHM methods, a different crown model or a different crown-size scaling factor could be tested. For the point-matching methods, a correction of the CHM-extracted tree heights, for which it is known that they tend to underestimate the height of the trees, is possible (Hollaus et al., 2010). If a systematic error cannot be prevented by such means, a correction analogous to the correction via the residuals in the evaluation of sub-areas of a multi-phase inventory would be plausible (Mandallaz, 2007).

5. Conclusions

Various co-registration methods were evaluated in a study area in Bremgarten. The aim of these methods is to establish a more precise link between forest inventory and remote sensing data. These methods should make the calibration of more precise regression models possible, which could then be used in a multi-phase inventory. The point-matching algorithm, which searched for the correct position by comparing two tree-top point-clouds, did not lead to satisfactory results. The method failed primarily because of the insufficient positional accuracy of the tested STI methods. To address the problem of an inaccurate STI, it would be of interest to try integrating a weighting factor, which gives dominant trees a greater weight in the co-registration process. Furthermore, it would be of interest to develop a method that not only uses tree-top points but also compares whole ALS point-clouds. Analogous to the method of Lamprecht et al. (2017), synthetic ALS point-clouds could be generated based on the inventory data and subsequently compared with the point-cloud retrieved from the ALS flights.

In addition to the point-matching-based co-registration, methods that use an artificial CHM (aCHM) generated from the sample data were tested for georeferencing. As the comparison with the dGNSS measured sample positions showed, none of the co-registration methods succeeded in adequately identifying the *true* recording positions. Due to several similar positions within the search window, a certain manual (CHM or orthophoto) co-registration was impossible for many sample-plots.

Despite the partly uncertain co-registration, it could be shown from the positions identified by the SQDIFF_NORMED correlation metric that this method is better suited for deriving regression models that link forestry indicators and remote sensing data. The method did not necessarily identify the true sampling position, but it found one plausible position (out of many). The improved accuracy of the regression models led to more precise modelling of forestry indicators from remote sensing data. If the inventory trees were modelled as Gaussian-bells in the aCHM, co-registration led to worse regression models than if the tree-crowns were modelled as spheres or ellipsoids. In this respect, it would be interesting to develop a method that smooths the discretely limited crown shapes by an additional application of a Gaussian filter to achieve a more realistic aCHM surface. Additionally, the investigation of an alternative SQDIFF_NORMED correlation metric, where the normalization is performed using only the aCHM values, and not the CHM values, would be worthwhile (section 4.2: Equation (8)).

Including the DecEv information in the raster-based co-registration process improved the positional accuracy when an accurate DecEv raster source was available (Bremgarten) but deteriorated the BA regression model quality when the DecEv raster source was likely not accurate enough (Zurich). However, neither for Bremgarten nor for Zurich did the DecEv information improve the quality of the VOL or BA regression model.

The evaluated co-registration methods were primarily advantageous when no dGNSS measurements of the sample-plot positions were available as a reference. However, even compared with the dGNSS positions, a slight improvement of the regression model accuracy was possible. If predictors based on co-registered positions are used to calibrate the regression models, the confidence interval of the model estimate and thus the confidence interval of the estimated average value of the total inventory is improved. It is essential to check whether the identification of a *plausible*, but not the *true* recording position introduces a systematic bias in the regression models. If there is such a bias, it should be tested whether it can be dissolved by adapting the co-registration method parameters. If this is not possible, a correction via the empirical residuals, analogous to the evaluation of sub-areas in multiphase inventories, is indicated. A systematic bias was not found in the present study. As the *true* recording positions were not identified in some cases, it should be emphasized that none of the methods tested is suitable if a linkage (co-registration) at the single-tree level is intended.

Applying the best-performing co-registration method sphere_1: SQDIFF_NORMED to the Zurich forest inventory data showed that the methods also led to an improvement of the RMSE of the regression model for this study area, at least when the DecEv raster was not considered in the co-registration process. The DecEv raster quality was most likely not sufficient in Zurich. This shows that the results are transferrable to other areas of investigation, with other ALS data and orthophoto sources.

For further research, it would be of interest to refine the point-matching co-registration procedures. Since the method already fails due to insufficiently precise STI methods, stem-detection methods in deciduous stands and leaf-off ALS data with a point density > 50 points m^{-2} could show whether the method is able to produce good co-registration results, if the underlying STI is sufficiently accurate. Furthermore, point-matching procedures which not only co-register tree-top point-clouds but also use all ALS points could improve the co-registration process. On the one hand, the uncertainties of an STI could be avoided, while on the other hand more information from the ALS data could be used than if only LM were detected. Regarding the aCHM co-registration methods, possible improvements include other crown models or correlation metrics.

CRediT authorship contribution statement

Simon Janssen: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Visualization, Resources, Software, Writing - original draft, Writing - review & editing. **Hans Pretzsch:** Resources. **Anton Bürgi:** Conceptualization, Resources. **Laura Ramstein:** Conceptualization, Resources. **Leo Gallus Bont:** Conceptualization, Methodology, Validation, Supervision, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Anja Bader from the Canton of Zurich for supporting this project. We also thank Daniel Mandallaz from the Chair of Land Use Engineering (ETH Zurich) for his input concerning statistical analysis. We are especially grateful to the WSL, in particular Dr. Oliver Thees and Dr. Janine Schweier, for collaboration and support in this project.

We thank Melissa Dawes for English editing assistance.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foreco.2023.120795.

References

Abteilung Wald. (2014). *Datendokumentation, LiDAR-Rohdaten*.
Baudirektion Kanton Zürich. (2017). *Aufnahmeanleitung Stichprobeninventur*.
Bitterlich, W., 1952. Die Winkelzählprobe: Ein optisches Meßverfahren zur raschen Aufnahme besonders gearteter Probeflächen für die Bestimmung der Kreisflächen pro Hektar an stehenden Waldbeständen. Die Winkelzählprobe. *Forstwissenschaftliches Centralblatt* 71 (7-8), 215–225.
Bont, L., 2019. Combining canopy height and different auxiliaries of tree type map information for timber volume estimation for forest management inventories. ETH Zürich.
Bont, L., Hill, A., Waser, L., Bürgi, A., Ginzler, C., Blattert, C., 2020. Airborne-laser-scanning-derived auxiliary information discriminating between broadleaf and conifer trees improves the accuracy of models for predicting timber volume in mixed and heterogeneously structured forests. For. Ecol. Manage. 459, 117856 https://doi.org/https://doi.org/10.1016/j.foreco.2019.117856.
BSF Swissphoto AG. (2011). Technischer Abschlussbericht DTM und DOM Bremgarten.
BSF Swissphoto AG, 2014. Laserscanning Kanton Zürich 2014. Technischer Bericht.
Chapman, D., 1961. Statistical problems in dynamics of exploited fisheries populations. Proc. 4th Berkeley Symp. on Mathematics Statistics and Probability 4, 153–168.
Dieler, J., Pretzsch, H., 2013. Morphological plasticity of European beech (Fagus sylvatica L.) in pure and mixed-species stands. For. Ecol. Manage. 295, 97–108.
Dorigo, W., Hollaus, M., Wagner, W., Schadauer, K., 2010. An application-oriented automated approach for co-registration of forest inventory and airborne laser scanning data. Int. J. Remote Sens. 31 (5), 1133–1153. https://doi.org/10.1080/01431160903380581.
Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.-M., Dalponte, M., Kobal, M., Pellegrini, M., Lingua, E., Mongus, D., Pfeifer, N., 2015. A benchmark of lidar-based single tree detection methods using heterogeneous forest data from the alpine space. Forests 6 (12), 1721–1747.
Fadili, M., Renaud, J.-P., Bock, J., Vega, C., 2019. RegisTree: a registration algorithm to enhance forest inventory plot georeferencing. Ann. For. Sci. 76 (2), 1–13.
Forrester, D., Nitzsche, J., Schmid, H., 2019. The Experimental Forest Management project: An overview and methodology of the long-term growth and yield plot network.
Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46 (1), 389–422. https://doi.org/10.1023/A:1012487302797.
Hallin, M., 2014. Gauss–markov theorem in statistics. Statistics Reference Online, Wiley StatsRef.
Hauglin, M., Lien, V., Næsset, E., Gobakken, T., 2014. Geo-referencing forest field plots by co-registration of terrestrial and airborne laser scanning data. Int. J. Remote Sens. 35 (9), 3135–3149. https://doi.org/10.1080/01431161.2014.903440.
Hernández-Stefanoni, J.L., Reyes-Palomeque, G., Castillo-Santiago, M.Á., George-Chacón, S.P., Huechacona-Ruiz, A.H., Tun-Dzul, F., Rondon-Rivera, D., Dupuy, J.M., 2018. Effects of sample plot size and GPS location errors on aboveground biomass estimates from LiDAR in tropical dry forests. Remote Sens. (Basel) 10 (10), 1586.
Hollaus, M., Mandlburger, G., Pfeifer, N., Mücke, W., 2010. Land cover dependent derivation of digital surface models from airborne laser scanning data. IAPRS 38, 1–3.
Hollaus, M., Mücke, W., Höfle, B., Dorigo, W., Pfeifer, N., Wagner, W., Bauerhansl, C., Regner, B., 2009. Tree species classification based on full-waveform airborne laser scanning data. Proceedings of SILVILASER 54–62.
Immitzer, M., Atzberger, C., Koukal, T., 2012. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. Remote Sens. (Basel) 4 (9), 2661–2693. https://doi.org/10.3390/rs4092661.
Kaartinen, H., Hyyppä, J., Vastaranta, M., Kukko, A., Jaakkola, A., Yu, X., Pyörälä, J., Liang, X., Liu, J., Wang, Y., Kaijaluoto, R., Melkas, T., Holopainen, M., Hyyppä, H., 2015. Accuracy of kinematic positioning using global satellite navigation systems under forest canopies. Forests 6 (12), 3218–3236.
Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., Wu, J.-C., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sens. (Basel) 4 (4), 950–974.
Khosravipour, A., Skidmore, A., Isenburg, M., Wang, T., Hussin, Y., 2014. Generating pit-free canopy height models from airborne lidar. Photogramm. Eng. Remote Sens. 80 (9), 863–872.
Koch, B., Heyder, U., Weinacker, H., 2006. Detection of individual tree crowns in airborne lidar data. Photogrammetric Engineering & Remote Sensing 72 (4), 357–363 https://www.ingentaconnect.com/content/asprs/pers/2006/00000072/00000004/art00001.
Koukoulas, S., Blackburn, G., 2005. Mapping individual tree location, height and species in broadleaved deciduous forest using airborne LIDAR and multi-spectral remotely sensed data. Int. J. Remote Sens. 26 (3), 431–455. https://doi.org/10.1080/0143116042000298289.
Lamprecht, S., Hill, A., Stoffels, J., Udelhoven, T., 2017. A machine learning method for co-registration and individual tree matching of forest inventory and airborne laser scanning data. Remote Sensing 9 (5), 505.
Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., Shreenan, R., 2003. Combined high-density lidar and multispectral imagery for individual tree crown analysis. Can. J. Remote. Sens. 29 (5), 633–649. https://doi.org/10.5589/m03-024.
Lu, X., Guo, Q., Li, W., Flanagan, J., 2014. A bottom-up approach to segment individual deciduous trees using leaf-off lidar point cloud data. ISPRS J. Photogramm. Remote Sens. 94, 1–12 https://doi.org/https://doi.org/10.1016/j.isprsjprs.2014.03.014.
Mandallaz, D., 2007. Sampling techniques for forest inventories. Chapman and Hall/CRC.
Mandallaz, D., 2013. Design-based properties of some small-area estimators in forest inventory with two-phase sampling. Can. J. For. Res. 43 (5), 441–449. https://doi.org/10.1139/cjfr-2012-0381.
Menk, J., Dorren, L., Heinzel, J., Marty, M., Huber, M., 2017. Evaluation automatischer Einzelbaumerkennung aus luftgestützten Laserscanning-Daten. Schweizerische Zeitschrift Fur Forstwesen 168 (3), 151–159. https://doi.org/10.3188/szf.2017.0151.
Monnet, J.-M., Mermin, É., 2014. Cross-correlation of diameter measures for the co-registration of forest inventory plots with airborne laser scanning data. Forests 5 (9), 2307–2326.
Naesset, E., Jonmeister, T., 2002. Assessing point accuracy of DGPS under forest canopy before data acquisition, in the field and after postprocessing. Scand. J. For. Res. 17 (4), 351–358. https://doi.org/10.1080/02827580260138099.
Naudts, K., Ryder, J., McGrath, M.J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray, G., Bönisch, G., Campioli, M., Ghattas, J., De Groote, T., Haverd, V., Kattge, J., MacBean, N., Maignan, F., Merilä, P., Penuelas, J., Peylin, P., Pinty, B., Pretzsch, H., Schulze, E.D., Solyga, D., Vuichard, N., Yan, Y., Luyssaert, S., 2015. A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes. Geosci. Model Dev. 8 (7), 2035–2065.
Olofsson, K., Lindberg, E., Holmgren, J., 2008. A method for linking field-surveyed and aerial-detected single trees using cross correlation of position images and the optimization of weighted tree list graphs. Proceedings of SilviLaser, 2008, 8th.
Parkan, M., 2018. Digital-Forestry-Toolbox. https://mparkan.github.io/Digital-Forestry-Toolbox/.
Parkan, M., 2019. Combined use of airborne laser scanning and hyperspectral imaging for forest inventories. EPFL.
Parkan, M., Tuia, D., 2015. Individual tree segmentation in deciduous forests using geodesic voting. IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2015, 637–640. https://doi.org/10.1109/IGARSS.2015.7325844.
Peng, C., Zhang, L., Liu, J., 2001. Developing and validating nonlinear height–diameter models for major Tree species of Ontario's boreal forests. North. J. Appl. For. 18 (3), 87–94.
Pretzsch, H., 2001. Modellierung des Waldwachstums, Parey Buchverlag Berlin, 341. Germany.
Pretzsch, H., 2014. Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. For. Ecol. Manage. 327, 251–264.
Pretzsch, H., Biber, P., Ďurský, J., 2002. The single tree-based stand simulator SILVA: construction, application and evaluation. For. Ecol. Manage. 162 (1), 3–21.
Pretzsch, H., Biber, P., Uhl, E., Dahlhausen, J., Rötzer, T., Caldentey, J., Koike, T., van Con, T., Chavanne, A., Seifert, T., Toit, B.d., Farnden, C., Pauleit, S., 2015. Crown size and growing space requirement of common tree species in urban centres, parks, and forests. Urban For. Urban Green. 14 (3), 466–479.
Pretzsch, H., Dieler, J., 2012. Evidence of variant intra-and interspecific scaling of tree crown structure and relevance for allometric theory. Oecologia 169 (3), 637–649.
R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/.
Rapidlasso GmbH. (2019). *LASTools*. http://rapidlasso.com.
Ratkowsky, D., Giles, D., 1990. Handbook of nonlinear regression models.
Ratkowsky, D.A., Reedy, T.J., 1986. Choosing near-linear parameters in the four-parameter logistic model for radioligand and related assays. Biometrics 42 (3), 575.
Reitberger, J., Schnörr, C., Krzystek, P., Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS J. Photogramm. Remote Sens. 64 (6), 561–574 https://doi.org/https://doi.org/10.1016/j.isprsjprs.2009.04.002.
Schmid-Haas, P., Baumann, E., Werner, J., 1993. Kontrollstichproben: Aufnahmeinstruktion. Eidgenössische Forschungsanstalt für Wald Schnee und Landschaft.
Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D., Petr, M., Honkaniemi, J., Lexer, M.J., Trotsiuk, V., Mairota, P., Svoboda, M., Fabrika, M., Nagel, T.A., Reyer, C.P.O., 2017. Forest disturbances under climate change. Nat. Clim. Chang. 7 (6), 395–402.

Singh, B., Singh, A.P., 2008. Edge detection in gray level images based on the Shannon entropy. J. Comput. Sci. 4 (3), 186–191.
Soh, L., Tsatsoulis, C., 1999. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans. Geosci. Remote Sens. 37 (2), 780–795. https://doi.org/10.1109/36.752194.
Stage, A., 1963. A mathematical approach to polymorphic site index curves for grand fir. For. Sci. 9 (2), 167–180.
Valbuena, R., Mauro, F., Rodriguez-Solano, R., Manzanera, J.A., 2010. Accuracy and precision of GPS receivers under forest canopies in a mountainous environment. Spanish J. Agric. Res., 8(4 SE-Agricultural environment and ecology), 1047–1057. https://doi.org/10.5424/sjar/2010084-1242.
Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkanen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalen, P., Gobakken, T., Koch, B., Naesset, E., Tokola, T., Maltamo, M., 2012. Comparative testing of single-tree detection algorithms under different types of forest. Forestry 85 (1), 27–40.
Waser, L., 2012. Airborne remote sensing data for semi-automated extraction of tree area and classification of tree species. ETH Zurich.
Waser, L., Ginzler, C., Kuechler, M., Baltsavias, E., Hurni, L., 2011. Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. Remote Sens. Environ. 115 (1), 76–85 https://doi.org/https://doi.org/10.1016/j.rse.2010.08.006.
Wing, M., Eklund, A., Kellogg, L., 2005. Consumer-grade global positioning system (GPS) accuracy and reliability. J. For. 103 (4), 169–173. https://doi.org/10.1093/jof/103.4.169.
Yang, R., Kozak, A., Smith, H., 1978. The potential of Weibull-type functions as flexible growth curves. Can. J. For. Res. 8 (4), 424–431.
Zimbelman, E.G., Keefe, R.F., 2018. Real-time positioning in logging: Effects of forest stand characteristics, topography, and line-of-sight obstructions on GNSS-RF transponder accuracy and radio signal propagation. *PLoS One*, *13*(1), e0191017.