Final report

International Benchmarking of Terrestrial
Image-based Point Clouds for Forestry

ISPRS Scientific initiative 2019

Principal Investigators

Martin Mokroš[1], Markus Hollaus[2] and Yunsheng Wang[3]

Co-investigators

Peter Surový[1], Livia Piermattei[2], Xinlian Liang[3], Milan Koreň[4], Julián Tomaštík[4] and Lin Cao[5]

[1]Czech University of Life Sciences Prague, Czech Republic; [2]TU Wien, Austria; [3]Finnish Geospatial Research Institute, Finland; [4]Technical University in Zvolen, Slovakia; [5]Nanjing Forestry University, China

Participants

Gábor Brolly[1], Carlos Cabo[2,3,4], Bartłomiej Kraszewski[5], Grzegorz Krok[5], Martin Krůček[6], Karel Kuželka[7], Nizar Polat[8], Atticus Stovall[9], Di Wang[10], Jinhu Wang[11]

[1]University of Sopron, Hungary; [2]Swansea University, United Kingdom; [3]University of Oviedo, Spain; [4]CETEMAS Research Institute, Spain; [5]Forest Research Institute, Poland; [6]The Silva Tarouca Research Institute for Landscape and Ornamental Gardening, Czech Republic; [7]Czech University of Life Sciences Prague, Czech Republic, [8]Harran University, Turkey; [9]NASA Goddard Space Flight Center, USA; [10]Aalto University, Finland; [11]Chinese Academy of Sciences, China

1. Project goals

The project aims to evaluate the performance of terrestrial image-based point clouds in plot-level forest inventory through an international benchmarking with comparison to ground truth data measured by conventional methods. We focused on whether the image-based point clouds can be an alternative solution to the more expensive terrestrial laser scanning (TLS) derived point clouds. Variety of algorithms from different research groups were used to explore the influence of algorithms on the accuracy of the estimation of the diameter at breast height (DBH).

Goals:

- Is it possible to use image-based point clouds for individual tree mapping and stem modelling in various types of forest stands?
- Is it possible to achieve similar accuracy from image-based point clouds as from TLS point clouds regarding DBH and tree position estimation within the research plots?
- Is there any significant difference between the applied algorithms for the DBH estimation on image- and TLS-based point clouds?
- What are the influences of the different algorithms on the accuracy of tree mapping and modelling?

2. Datasets

Altogether, we established ten plots in the five countries Austria, China, Czech Republic, Finland, and Slovakia. One plot for each type of forest stand. The plots vary in size, tree species composition, tree density, and topography. In Table 1 an overview of the different forest characteristics is given.

Table 1. The characteristics of the research plots

Study site	Shape	Size (m) Diameter / square length	Dominant tree species	Stem Density [stems/ha]	Plot description	Plot No.
Austria 1	Circular	40	*Picea abies*	533	Even-aged, well managed spruce forest	plot 1
Austria 2	Circular	40	*Fagus sylvatica*	390	Uneven-aged, managed deciduous (beech) forest, multi- layer structures	plot 2
China 1	Circular	30	*Taxodium distichum*	410	Pure-even-aged, plantation of bald cypress	plot 3
China 2	Circular	30	*Liriodendron chinensis*	609	Pure-even-aged, plantation of Chinese tulip poplar	plot 4
Czechia 1	Square	50	*Fagus sylvatica*	280	Even-aged, well managed beech forest	plot 5
Czechia 2	Square	50	*Picea abies*	272	Even-aged, well managed spruce forest	plot 6
Finland 1	Square	32	*Pinus sylvestris*	479	Unmanaged, even-aged, Scots pine forest	plot 7
Finland 2	Square	32	*Pinus sylvestris, Betula sp.*	869	Unmanaged Mixed pine and birch forest, with multi- layer structures	plot 8
Slovakia 1	Circular	15	*Quercus petraea*	651	Even-aged, well managed Oak forest	plot 9
Slovakia 2	Circular	20	*Abies alba*	875	Even-aged, silver fir managed forest	plot 10

Reference Data
DBH and tree position were measured in-situ using conventional field measurement instruments. The tree position was measured by Total Station, and the DBH was measured by a diameter tape.

Point Cloud Data Sets
Both Image- and TLS- based point clouds were acquired for the test plots (Figure 1).

- Image-based point clouds
  Images were acquired using a stop-and-go mode. With this setting, the operator is capturing images only in a stable position. The paths were different depending on plot conditions. Plots situated in Austria, China, Czech Republic, Slovakia were collected by a camera held on a tripod, and the path of data collection was around and inside the plots and two diagonal lines. Plots in Finland were collected by a hand-held camera from a path surrounding the plots. Agisoft Metashape was used to align the images and generate the scaled dense point clouds.

- Terrestrial laser scanning point clouds
  Plots situated in Austria, Czech Republic and Slovakia were scanned by Riegl VZ-2000 scanner, in Finland by Leica HDS6100 scanner, in China by Riegl VZ-400i. The positions of the scanner were around plots and also inside them. The number of positions was based on the plot conditions. For each plot the point clouds from all scan positions were co-registered and merged into one point cloud. This method is known as a multi-scan method.
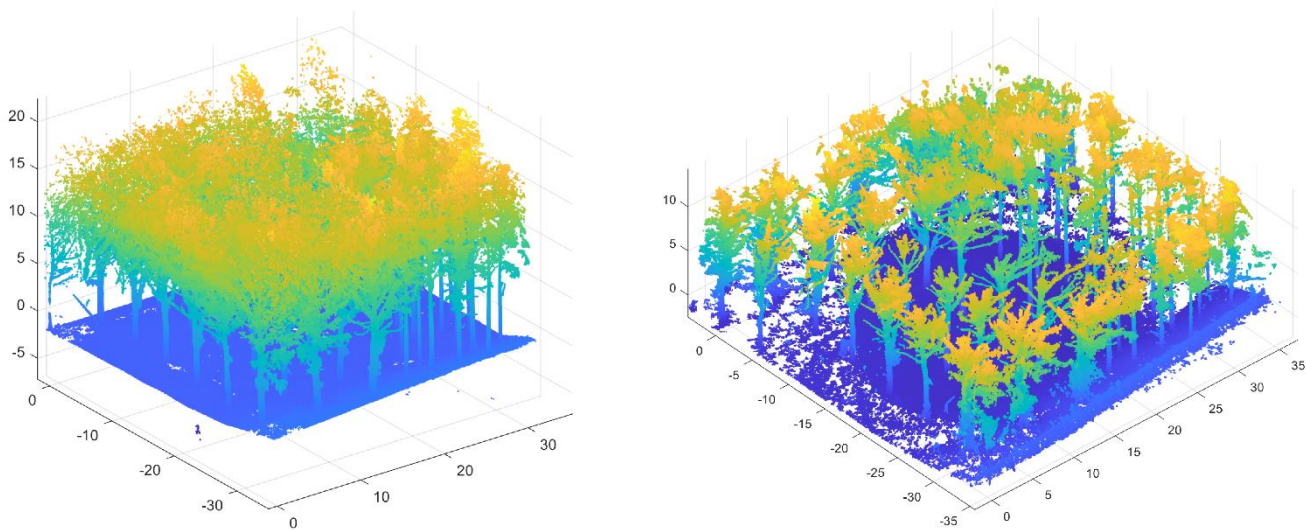

Figure 1. Example of point clouds plot number 3. On the left TLS-based point cloud on the right image-based point cloud.

3. Participants and their Algorithms

The project assembled 14 participants from ten countries worldwide. Altogether fifteen different algorithms were applied by the 14 research groups. A list of involved institutions for the algorithms are given in Table 2. An abbreviation is given to each of the algorithms, and the abbreviations are used later in the figures as the name of the algorithms. The algorithms are listed in an alphabetical order of the abbreviations in table 2.

Table 2. List of involved researchers responsible for the algorithms used

Institution	Abbreviation	Country
The Silva Tarouca Research Institute for Landscape and Ornamental Gardening	3DForest	Czechia
Aalto University	Aalto	Finland
Chinese Academy of Sciences	CAS	China
Czech University of Life Sciences Prague	CULS	Czechia
Finnish Geospatial Research Institute	FGI	Finland
Forest Research Institute	FR1	Poland
Forest Research Institute	FR2	Poland
Harran University	Harran	Turkey
NASA Goddard Space Flight Center	NASA	USA
Nanjing Forestry University	NFUa	China
Nanjing Forestry University	NFUm	China
Technical University of Vienna	TUWien	Austria
Technical University in Zvolen	TUZVO	Slovakia
University of Oviedo	S-O-C	Spain
University of Sopron	UniSopron	Hungary

The 15 algorithms used varied with the level of automation in the approaches for stem detection, DBH estimation, and pre-processing. The main characteristics of each algorithm are summarized in Table 3. The algorithms present a great methodological variety in all three approaches. Except one manual algorithm provided by NFU, all other 14 algorithms are fully automatic. Results from the manual algorithm reveals differences between the automatic approaches and the human visual examinations.

Table 3. List of algorithms with description of main attributes

Institution (abbreviation)	Level of automation	Pre-processing methods	Stem detection method	DBH estimation method
3DForest	Automatic	Extraction of terrain points, noise filtering on SfM data	Voxel clustering based on principal component analysis	Hough Transformation
Aalto	Automatic	Uniform downsampling, terrain removal	Recursive segmentation	RANSAC - Cylinder
CAS	Automatic	Off-ground points filtering, Voxelization and Outliers removal	Geometric feature + Density-based clustering	RANSAC-based circle fitting
CULS	Automatic	noise filtering	Vertical continuity of voxel densities	RLTS circle
FGI	Automatic	Point equivalent sampling	Point distribution analyses, stem modeling	Robust least square fitting
FR1	Automatic	Point cloud subsampling, SOR filtration, Low points removal	Geometric feature based segmentation	RANSAC - Circle
FR2	Automatic	SOR filtration, Point cloud cutting, Normal filtration	Point cloud clustering, RANSAC Cylinder Fitting	RANSAC - Cylinder
Harran	Automatic	Data reduction, SOR filtration, Verticality calculation	Geometrically segmentation, RANSAC Cylinder fitting	RANSAC Cylinder fitting (then diameter thresholding)
NASA	Automatic	Uniform downsampling, terrain removal, SOR filter	Geometrically segmentation, RANSAC Cylinder fitting	Iterative LS Circle Fitting
NFUa	Automatic	Point cloud normalization, SOR filtration	Density-based spatial clustering, Cone Fitting	DBSCAN - Cylinder
NFUm	Manual	Point cloud normalization, SOR filtration	Manual delineation	Cylinder fitting (based on the manually selected stem points)

S-O-C	Automatic	Multiple voxeling-devoxeling, group isolation filtering	Distance clustering + consistency checking in height range	Iterative geometric circle fitting + consistence checkings in sections and with neighbor sections
TUWien	Automatic	DTM derivation; point cloud normalization	Geometric and density-based feature for horizontal cross section	Robust cone- and cylinder fitting
TUZVO	Automatic	DEM derivation, tree identification, tree base estimation, cross-section, noise removal	Spatial clustering of horizontal cross-section	Circle fitting
UniSopron	Automatic	Selection of terrain points by directional filtering, TIN interpolation to create DTM	Directional filtering in voxel space	Linear regression of stem circles fitted at multiple heights

4. Evaluation Procedure

Evaluation of results was carried out using an automatic evaluation approach.

The first criteria was the accuracy of tree detection, including the correctness, the completeness, and the mean accuracy (1-3).

$$Correctness = \frac{Number\ of\ Match\ case}{Number\ of\ detected\ trees} \tag{1}$$

$$Completeness = \frac{Number\ of\ Match\ case}{Number\ of\ reference\ trees} \tag{2}$$

$$Mean\ Accuracy = 2 * \frac{Number\ of\ Match\ case}{Number\ of\ detected\ trees\ +\ Number\ of\ reference\ trees} \tag{3}$$

For each reference tree, the existence of detected trees was searched within a circular area with one meter radius. If one and only one detected tree was found in the area, a case of match was confirmed. If more than one detected tree was found in the area, the one that stood at the nearest to the reference tree was considered as the matched tree, and a match case was confirmed. An omission error was counted if no detected tree was found in the area. At the end, if there were detected trees that remained unmatched with any of the reference trees, they were regarded as commission errors. Furthermore, for all matched trees the error was calculated as difference between estimated and reference diameter. If the error exceeded 20% of the reference diameter it was removed.

The second criteria was the accuracy of the estimated DBH of the detected trees. The accuracy was evaluated using the Root Mean Square Error (RMSE) and the Bias (4-5) between the estimated and the reference DBHs of the match cases. The RMSE and Bias was calculated in plotwise, namely, for each sample plot, the RMSE and the Bias were calculated to evaluate the overall accuracy of the estimated DBH. Gross Errors that were higher than 20% of reference DBH were not included in the evaluation.

$$Bias = \sum_{i=1}^{n} \frac{(e_i - r_i)}{n} \tag{4}$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(e_i - r_i)^2}{n}} \tag{5}$$

where e_i is estimated tree diameter and r_i is reference tree diameter.

Altogether, 300 datasets were evaluated. Results were collected from 15 algorithms, and each algorithm produced 20 results, including ten results for ten image-based point clouds and ten results for TLS-based point clouds of the ten test plots. All the results were evaluated using identical automated evaluation procedures and parameter settings.

5. Results

5.1 The accuracy of tree detection

The accuracy of the tree detection was evaluated with the completeness and the correctness. As shown in Figure 2, it is obvious that almost all the algorithms provided higher completeness of tree detection when using the TLS-based point cloud. A significantly lower performance of tree detection completeness occurred in plot 7 and 8, where the plots had a great amount of undergrowth such as regenerations and bushes.
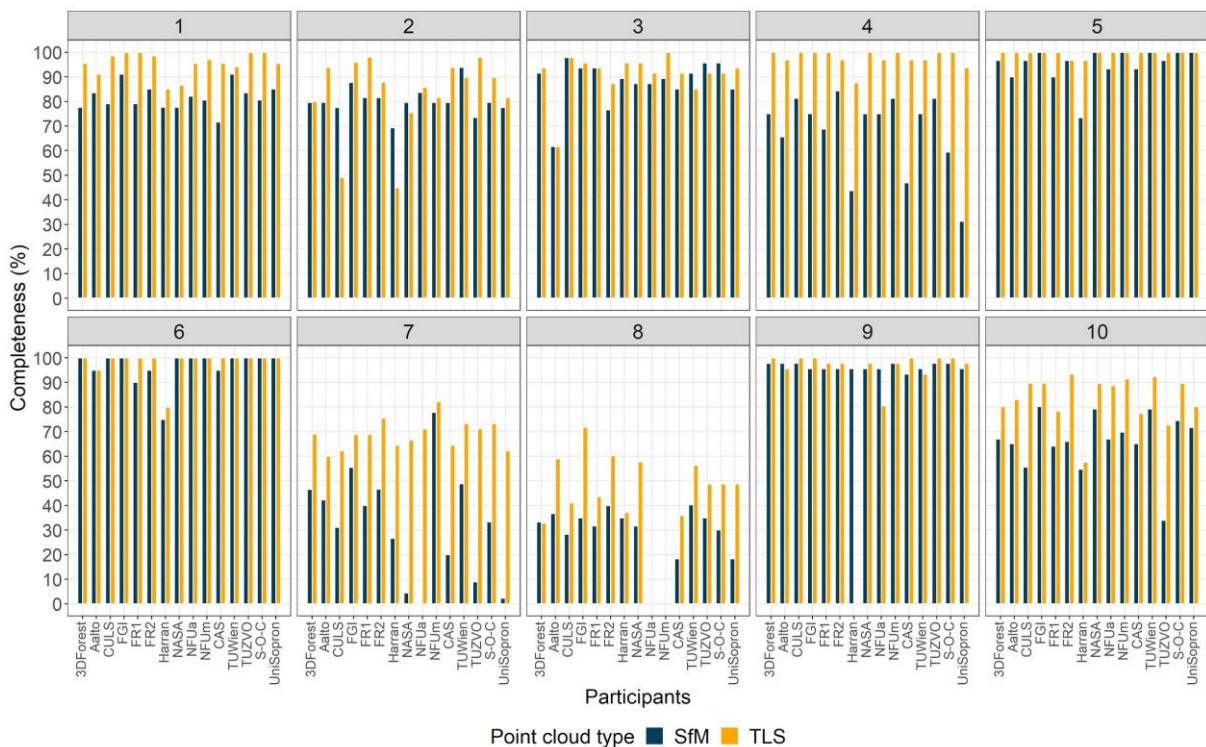

Figure 2. Results of completeness (%) for each algorithm within all sample plots (plot number is marked on top of each subfigure). For each algorithm (abbreviations of the algorithms are listed along the x axis in each sub figure), the completeness of stem detection results using SfM- and TLS-based point clouds are reported in parallel.

The median of the completeness among all the algorithms in each plot are presented in Figure 3. Except in plot 6 where both SfM- and TLS- based results have a median completeness of 100%, the

median of completeness was higher when TLS-based point clouds were used in all other plots. The median of the completeness ranged from 33% to 100% and from 49% to 100% for image-based and TLS-based point clouds, respectively.

The results suggested that the existence of the undergrowth in the sample plots (e.g., plot 7 and 8) has a significant influence on the stem detection algorithms. The results also suggested that, although the overall completeness of tree detection from SfM-based point clouds inclined to be lower than that from TLS-based point clouds, in most of the cases, the differences are not as high as commonly expected that the TLS is superior to the image-based point cloud. In five out of ten plots, the difference between the SfM- and TLS-based completeness is less than 10%. This result indicates that, in easy forest stand conditions, such as the managed or the matured pure forest stands, the SfM-based point clouds can provide a similar completeness for stem detection as the TLS-based point clouds.
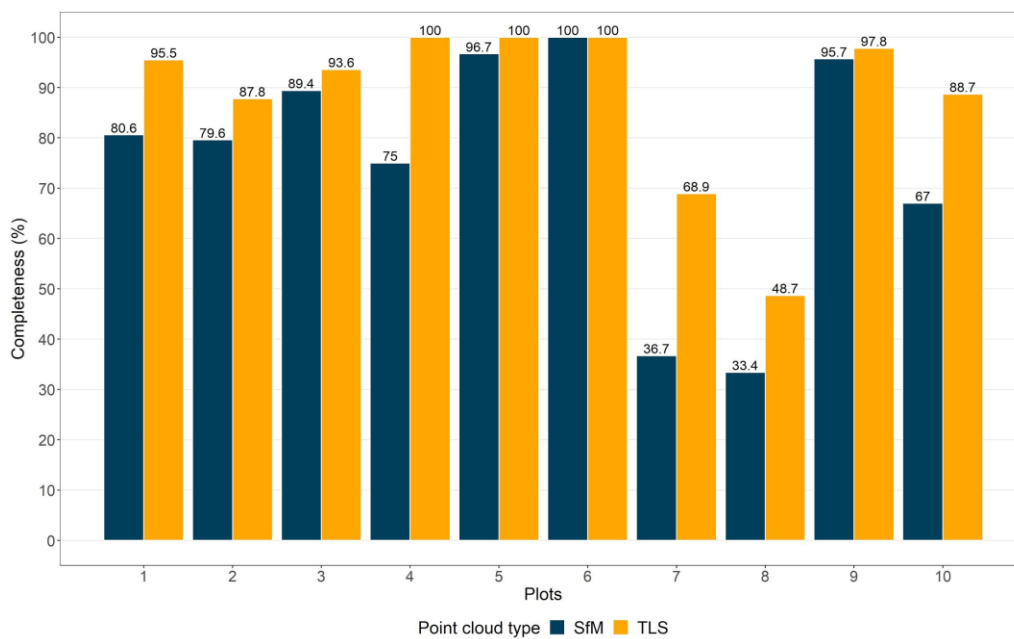

Figure 3. Median of completeness (%) of participants for each plot divided by point cloud type (image-based (SfM) and TLS based point clouds)

The correctness of the stem detection algorithms did not have a clear difference between TLS and SfM-based point cloud within all plots, as in the completeness. The difference is clearly higher within plot 4, 7 and 8, where the correctness is higher when TLS-based point clouds were used. In overall the correctness varies for SfM- and TLS-based point clouds from 8% to 100% and from 40% to 100%, respectively.
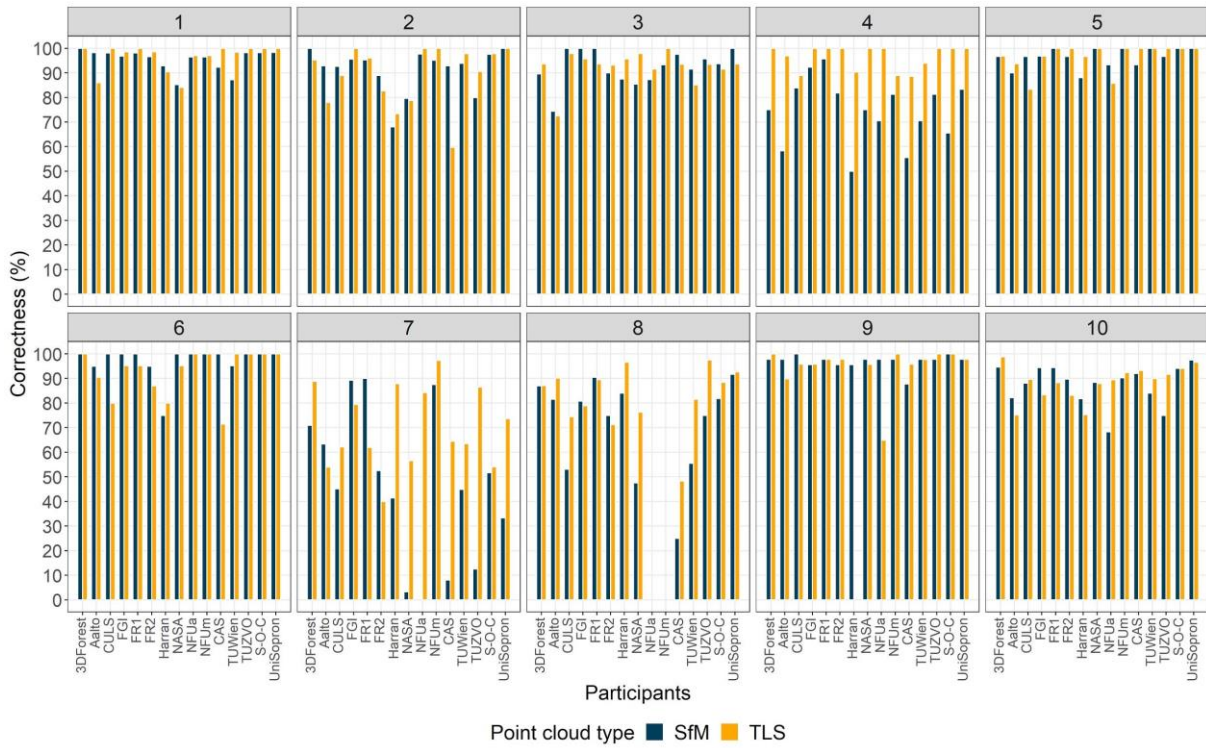
Figure 4. Results of correctness (%) for each algorithm within all sample plots (plot number is marked on top of each subfigure). For each algorithm (abbreviations of the algorithms are listed along the x axis in each sub figure), the completeness of stem detection results using SfM- and TLS-based point clouds are reported in parallel.

In regards to correctness, The difference between point cloud types is bigger within the plot with more complex structure. The TLS is achieving better results then the terrestrial photogrammetry in these plots. The median of the correctness varies for SfM- and TLS-based point clouds from 48.4% to 100% and from 64.4% to 100%, respectively.
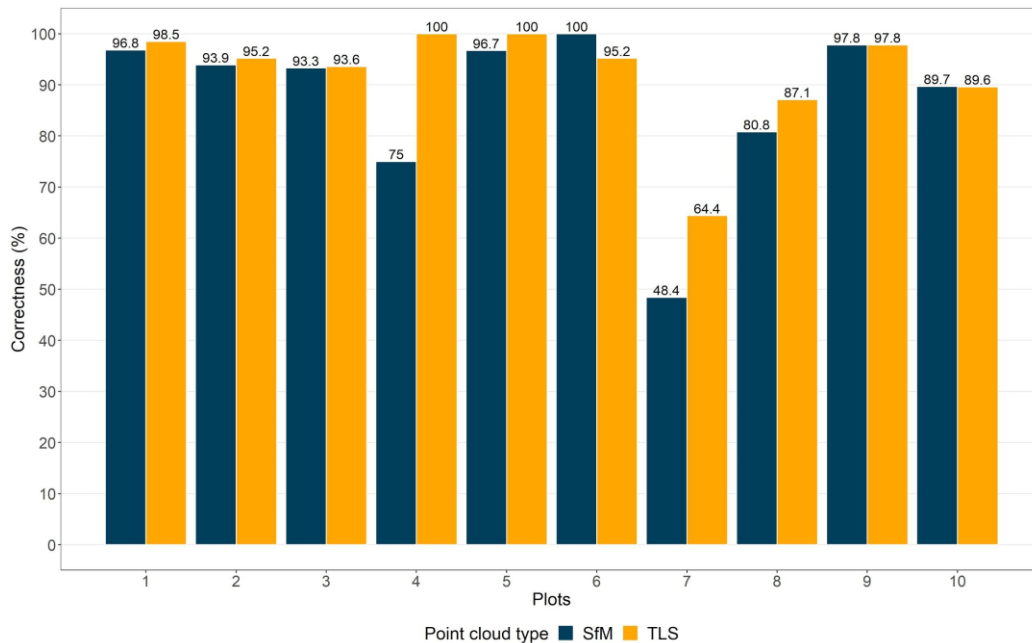

Figure 5. Median of correctness (%) of participants for each plot divided by point cloud type (image-based (SfM) and TLS based point clouds)

5.2 The accuracy of DBH estimates

The Biases and RMSE values of the DBH estimates achieved by each algorithm are reported in Figure 6. As shown in the Figure, the results agreed with the assumption that TLS-based point clouds in

general provide more accurate point cloud data, which consequently provide more accurate DBH estimates, regardless of the algorithms applied and regardless of the stand conditions in the sample plots. Another important finding in the results is that the algorithms performed similarly within the sample plots for SfM and TLS point clouds. The RMSE of the DBH estimates from the majority of the algorithms is less than 2 cm.
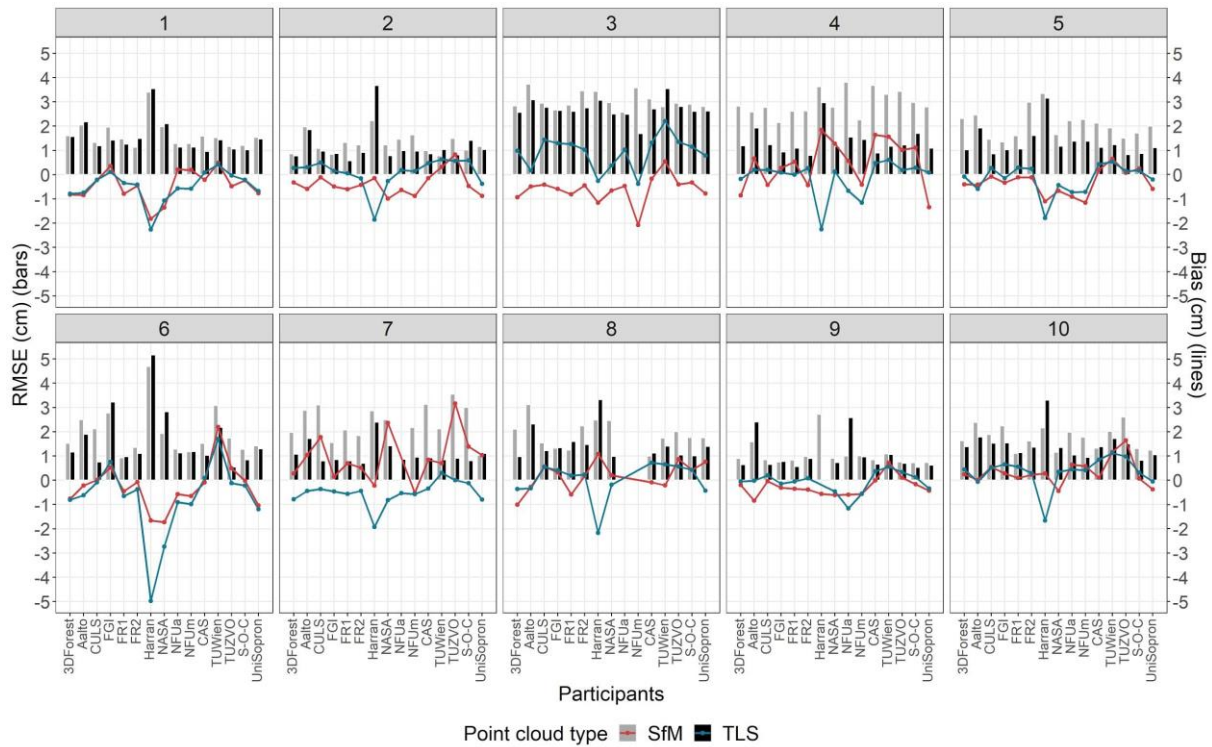

Figure 6. Root mean square error (bars) and bias (lines) for each algorithm within all research plots divided to image-based and TLS-based point cloud.

The median of RMSE for each plot from all 15 algorithms is presented in Figure 7, which can be regarded as a clearer indication on the overall performance of the algorithms in estimating the DBH. The RMSE median of Image-based and TLS-based ranges from 0.90 cm to 2.95 cm, and from 0.72 cm to 2.64 cm, respectively.

Similar to the completeness of the stem detection, despite a stronger capacity from the TLS-based point cloud in DBH estimations, the difference between the RMSEs of TLS- and SfM- based DBH estimates is less than 10% with respect to TLS-based RMSE median. The exception is plot 4 and 7, where the differences in RMSE median is more than twice higher than that in other plots. Such results indicate the higher sensitivity of the SfM-based point clouds with respect to the stand conditions in forest environments.
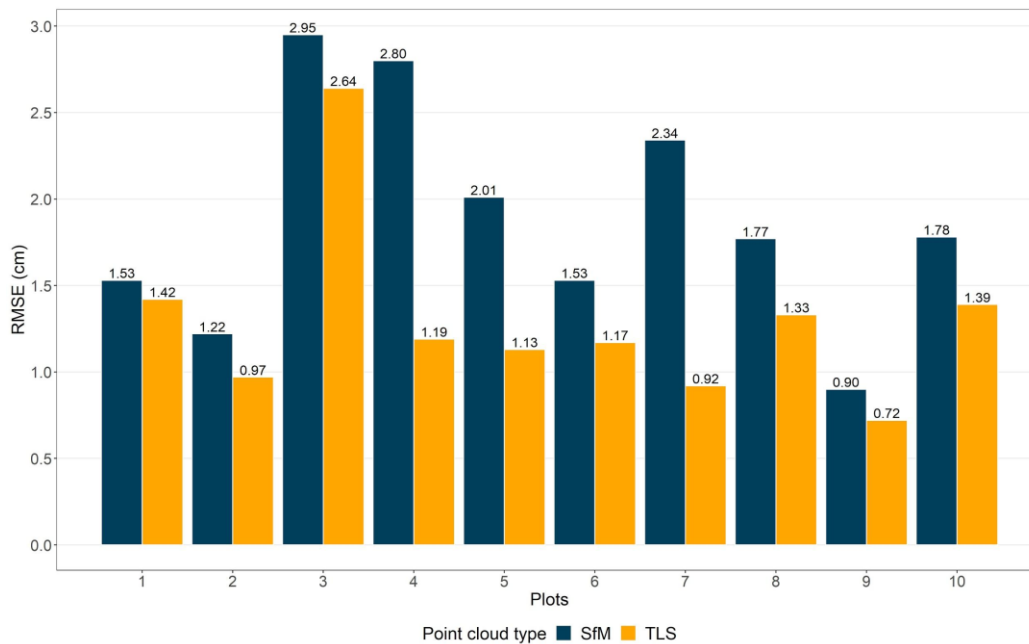
Figure 7. Median of RMSE of participants for each plot divided by point cloud type (image-based (SfM) and TLS based point clouds)

5.3 The impacts of the circle-based vs. cylinder-based methodologies

One of the distinguishing differences among the algorithms is the geometric model used for the stem modeling. The algorithms are divided into two groups based on the geometrical shape. Nine algorithms used a circle and six used a cylinder. Figure 8 and Figure 9 presented the median completeness and the median RMSE of DBH estimates of the algorithms in each group. As shown in Figure 8 and 9, the impact of the modeling approach has more significant impact on the accuracy of DBH estimates than the completeness of the stem detection. It can clearly be seen from Figure 8 that the differences of the stem detection completeness between different method groups are rather small for most of the plots, except the challenging ones where the cylinder approaches leads to higher completeness values.
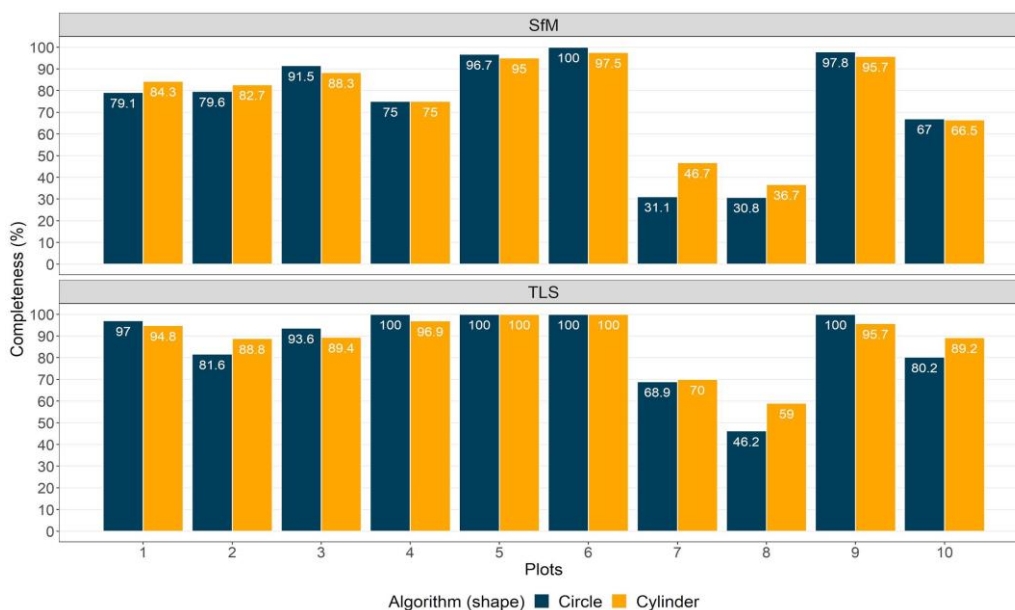

Figure 8. Median of completeness of tree detection divided by the geometrical shape used by algorithms for each plot divided by point cloud type (image-based (SfM) and TLS based point clouds).

Regarding the RMSE of DBH estimates, as shown in Figure 9, when image-based point clouds were used, the circle- based algorithms had an error ranging from 0.85 cm to 2.95 cm, whereas the error of cylinder-based algorithms ranged from 1.04 cm to 3.11 cm. When TLS-based point clouds were used, the RMSE of DBH estimates varied from 0.66 cm to 2.62 cm with the circle-based algorithms, and from 0.86 to 2.91 with the cylinder-based algorithms. The results suggest that the circle-based algorithms could be slightly more accurate on the estimation of DBH, however, more detailed analyses are needed to clarify the reasons behind the slightly higher median RMSE values of the DBH estimates from the cylinder-based algorithms.
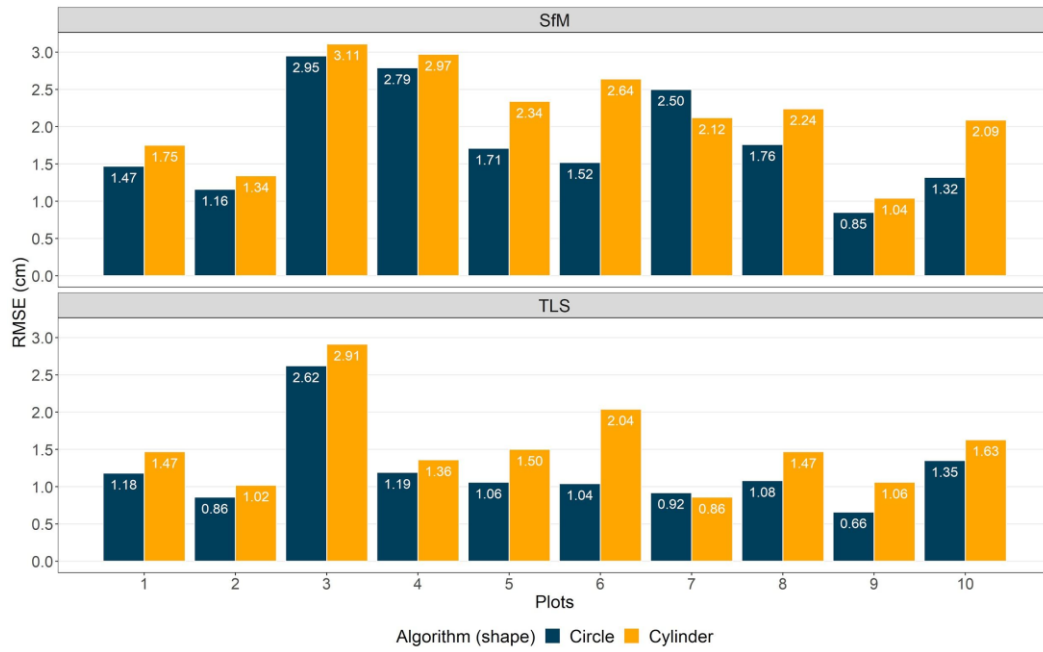

Figure 9. Median of RMSE (cm) of DBH estimation divided by the geometrical shape used by algorithms for each plot divided by point cloud type (image-based (SfM) and TLS based point clouds).

6. Published work related to the project

Mokroš, M., Hollaus, M., Wang, Y., & Liang, X. (2020, May). SFM-Forest-Benchmark project: The benchmarking of image-based point clouds for forest inventory. In *EGU General Assembly Conference Abstracts* (p. 5822).

Hollaus, M., Mokroš, M., & Wang, Y. (2019, April). International Benchmarking of terrestrial Image-based Point Clouds for Forestry. In *Geophysical Research Abstracts* (Vol. 21).

7. Project expenses

The whole budget from ISPRS Scientific Initiatives was 10,000 CHF. Budget was evenly divided among each partner.

Partner	Budget	Purpose of expenditure
Czech University of Life Sciences Prague	2,000	Personal cost
TU Wien	2,000	Personal cost
Finnish Geospatial Research Institute	2,000	Personal cost
Technical University in Zvolen	2,000	Personal cost
Nanjing Forestry University, China	2,000	Personal cost