Automatic estimation of tree stem attributes using terrestrial laser scanning in central Indian dry deciduous forests

R. Suraj Reddy[1,*], Rakesh[1], C. S. Jha[1] and K. S. Rajan[2]

[1]Forestry and Ecology Group, National Remote Sensing Centre, Hyderabad 500 037, India
[2]Lab of Spatial Informatics, Indian Institute of Information Technology, Hyderabad 500 032, India

Forest inventories are critical for effective management of forest resources. Recently, the use of terrestrial laser scanning (TLS) to automatically extract forest inventory parameters at tree level (e.g. tree location, diameter at breast height (DBH) and height) has gained significant importance. TLS using both single-scan and multi-scan techniques, not only helps in detailed and accurate measurements of tree objects but also helps increase the measurement frequency. In the current study, we develop an automated solution to extract forest inventory parameters at individual tree level from TLS data by using random sample consensus (RANSAC)-based circle fitting algorithm. The method was evaluated on both single- and multi-scan data by characterizing four circular plots of radius 20 m in dry deciduous forests of Betul, Madhya Pradesh (India). Over all the plots, tree detection rates of 75% and 97% were obtained using single- and multi-scan TLS data respectively. Tree detection rates were significantly affected by increase in distance from the scanner, in single-scan approach when compared to multi-scan approach. Field based DBH measurements correlated well using both single (R^2 = 0.96) and multiple scans (R^2 = 0.99). The DBH estimates from multi-scan TLS data resulted in low root-mean-square error (RMSE) of 2.2 cm compared to that of 4.1 cm using single-scan. Further, tree heights were extracted from TLS data and validated with selectively measured trees on field (R^2 = 0.98; N = 65). The RMSE of tree height was estimated to be 1.65 m. The current results show the potential use of TLS in automatically deriving forest inventory parameters with reliable accuracy at individual tree level.

Keywords: DBH, forest inventory parameters, multiscan, single-scan, terrestrial laser scanner.

FOREST inventory parameters, such as tree counts, location, species, DBH and height, are vital for obtaining above-ground biomass and assessment of carbon sequestration potential[1]. However, estimation of forest inventory parameters using conventional field methods is a daunting task. Simple measurements such as DBH and height are labour-intensive and time-consuming. The complexities and time only increase when other geometrical properties of trees including branch and crown parameters are to be measured. Thus, ways towards faster inventories with reliable accuracies are much appreciated.

TLS is a fast, affordable and accurate method used in the extraction of forest inventory parameters (e.g. DBH, tree heights and crown shape parameters) at the individual tree level by capturing three-dimensional (3D) point cloud of the forest area[2–5]. Also, dense point cloud data from TLS is used to determine stand/plot level attributes, such as leaf area index[6], non-destructive tree volume estimates[7], which aid in the calibration of satellite-based remote sensing estimations[8]. Both single and multiple scans have been used to extract forest inventory parameters[2,5,9,10].

The most common method used to extract inventory parameters using TLS point cloud commonly focused on extraction of a slice of the point cloud and identified tree stems by clustering methods and finding circles[2,5,9–12]. Maas *et al.*[2] used a mathematical morphological-based clustering approach to find trees and estimate DBH using single-scan and multiple-scan approaches in 5 circular plots of 15 m (212–410 stems/ha). They reported an accuracy of 97% with an RMSE of 1.8 cm in DBH measurement. An iterative close point algorithm, to cluster trees from ground vegetation, in combination with Hough transform to find circles was used by Huang *et al.*[10], to detect trees in a sparse forest plot with 33 trees/ha, using multiple-scan TLS data. They detected all trees (100%) with a DBH variation of 3.4 cm. Olofsson *et al.*[5] recovered tree locations from single-scan TLS data with an accuracy of 73% and DBH with an RMSE ranging from 2 to 9.6 cm over 16 circular plots of 20 m (358–1042 stems/ha) using both Hough transform and RANSAC based circle fitting. Studies have also noted that tree detection rates in both single- and multiple-scans are majorly impacted by forest structure[13]. These highly varying results from different forest types of varying tree densities, plot sizes and scanning scenarios suggest the need for further research on this topic.

In order to use TLS for forest inventory, algorithms have to be developed and tested across different forest types. In the current paper, we develop an automatic algorithm to identify tree stems and extract inventory information (tree positions, DBH and tree heights) from both single-scan and multiple-scan TLS data. The method uses stem probability score to cluster and identify tree stems and RANSAC-based circle fitting to estimate location, DBH and height. The algorithm is evaluated over 4 plots (different densities) in central Indian dry deciduous forests (Betul, Madhya Pradesh).

Four circular plots of 20 m radius were established in the teak mixed dry deciduous forests of Betul (Madhya Pradesh, India) around 21°51′46.84″N and 77°25′33.67″E. The plots were mostly flat with a slope variation of about

*For correspondence. (e-mail: surajreddy.rs@gmail.com)

Table 1. Plot level statistics of field measurements

Plot ID	Stem count	Density (stems/ha)	Diameter at breast height		
			Minimum (cm)	Maximum (cm)	Mean (cm)
S1	51	406	9.55	41.00	18.84
S2	60	477	9.55	41.00	18.71
S3	65	517	9.45	38.04	16.41
S4	52	414	9.61	42.53	19.27

5–10 degrees. The dominant tree species found in the plots are *Tectona grandis*, *Lagerostromia parviflora*, *Miliusa tomentosa* and *Diospyros melanoxylon*. In each plot, for all the trees greater than 30 cm in girth, inventory parameters such as tree location and DBH were measured. Tree locations were accurately measured up to 1 m using metre tape and DBH was measured using standard metre tape. Tree heights were measured for selected trees (~15 trees in each plot) using Nikon forestry pro instrument. The stem density across plots varied from 400 to 500 stems/ha. The summary of plot level statistics is given in Table 1.

A Reigl VZ-1000 was used to collect 5 scans for each plot, one at the centre of the plot and the remaining four at the plot boundaries during May 2016. Scans were merged using artificially set up targets. The data was captured in a high speed mode (pulse repetition rate of 300 kHz) with a vertical and horizontal resolution of 0.05 degree. The scanner acquired each scan within a time span of 2 min, in the field of view of 360° in horizontal direction and 100° (set at 30°–130° for the present study) in vertical direction.

In the current method, tree locations and associated parameters were obtained automatically from both single-scan and multiple-scan TLS data in four major steps: (i) Estimation of ground model (digital elevation model – DEM) to normalize the point cloud; (ii) Identification of tree boles by clustering and filtering the point cloud; (3) Circle fitting at 1.3 m height using RANSAC (random sample and consensus) based least square fitting; (4) Validation of estimated circles to detect trees. All TLS point clouds were pre-processed to remove all points with erroneous reflectance values, with range more than 20 m (plot size). All methods mentioned in the current paper for automated analysis of the data were developed by the authors using MATLAB, Python and R programing languages.

Estimation of ground model (DEM) is the primary step in TLS processing to normalize the point cloud and estimate breast height level for all trees. In the current study, DEM is based on the generation of minimum elevation (min$Z$) raster with a grid size of 50 cm. For all grids with point density >200 points/m^2, the lowest elevation value was recorded. For the missing grids, min$Z$ value was interpolated using Delaunay triangulation. Finally, DEM was created by smoothing min$Z$ raster using a 2D median filter. Using DEM, ground points were separated as points with less than 50 cm height and vegetation points were used for further processing.

The vegetation points were then sliced at 1.3 m height (breast level) with a 0.5 m buffer (i.e. points between 0.8 m and 1.8 m) and used to detect tree positions. The assumption that the tree trunks are continuous in vertical direction, was used to filter trees boles from low vegetation and branches.

The sub-vegetation point cloud was rasterized with a horizontal grid size of 1 cm and all the points in each grid were used to understand height distribution in order to find points with high trunk probability. In each grid, all points were further divided into height (Z) sub-intervals of 4 cm (25 bins) and total filled bins with points (Z_t) and the largest continuously filled bins (Z_c) were calculated. High probable trunk points were then identified as points with high vertical continuity and point distribution ($Z_t \geq 5$ and $Z_c \geq 3$). Undergrowth and branches were then removed by employing less vertical connectivity thresholds ($Z_t \leq 2$ and $Z_c < 2$). The thresholds were set on the basis that a tree trunk should at least have 20% of the bins filled and continuous in at least half of them. Finally, tree boles were identified using high probable trunk points as seeds and clustering the filtered point cloud by using Euclidean distance criteria (5 cm threshold). Figure 1 depicts the screenshots of original sub-vegetation point cloud which is filtered and clustered using seed points to identify tree boles.

Trees are often modelled by circles or cylinders. Each cluster (possible tree bole), is used to fit circles to estimate DBH and tree location. In each cluster, all points in 20 cm disk centred at 1.3 height are used to fit circles using RANSAC method[14,15].

RANSAC is an iterative algorithm to estimate model parameters with minimal influence from outliers in the set of observed data. RANSAC algorithm was further modified to have valid circle estimates by using certain conditions as follows. The circles with estimated centre between the scan position and the first trunk point were considered invalid. Since no data point can penetrate the tree trunk, the estimated circles with more than 2% of the cluster points inside the trunk were considered invalid. Randomly estimated circles with radius more than 50 cm

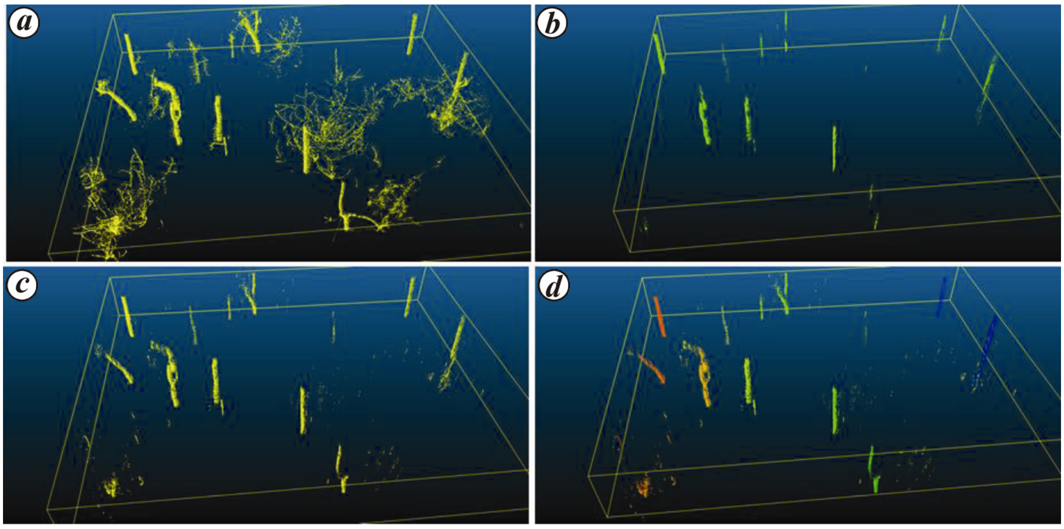

Figure 1. *a*, Represents subset of the original point cloud. *b*, Filtered point cloud. *c*, Seed points as identified. *d*, Clustered point cloud using seed points and filtered point cloud.

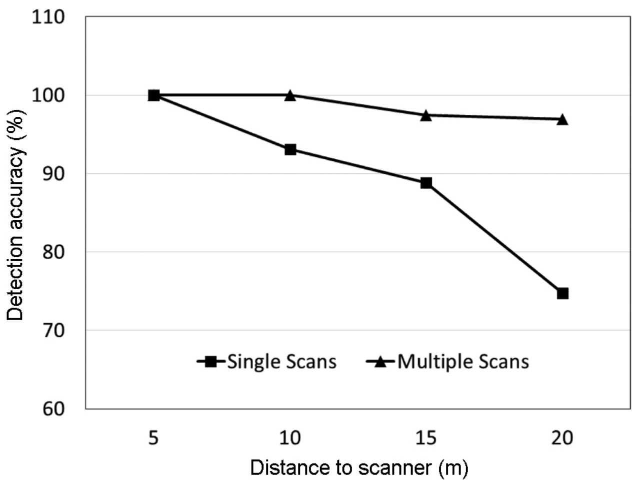

Figure 2. Proportion detected trees as a function of distance to the scanner using both single and multiple scans.

and radius less than 4 cm were also considered invalid (minimum threshold of 4 cm was based on the field data). Finally, using modified RANSAC, maximum inliers (iterations = 100) were selected and a circle was fitted by using least square method to estimate centre and radius.

Further validation included examination of vertical consistency to finalize tree locations and DBH estimates. Since trunks are vertically distributed, the circle centre and radius have to maintain uniformity in vertical direction. For each cluster with valid radius and centre at 1.3 m, two more circles were fitted at heights 1.2 m and 1.4 m for verification of centre and radius. The centre location error of 10 cm and a radius error of 20% were allowed to validate each circle. Finally, such validated circles were used to provide the tree locations and DBH estimates.

Tree heights were estimated using vegetation points (after normalization) and tree locations. After fitting circles for each cluster, a cylindrical subset of point cloud was extracted with centre as circle centre and radius as 5 times circle radius. The height (Z) values were binned into intervals of 0.5 m. The continuous bins were separated in order to remove the canopy layers from surrounding trees. Finally, the maximum height value of those continuous bins was marked as tree height. Since the scans were done during dry season with least hindrance due to leaf, occlusions to the top of the tree were minimal and helped in accurate estimation of tree height.

The tree detection rates across four plots of 20 m radius with stem density between 400 and 500 stems/ha have varied from 73% to 82.5% using single-scan data and from 95% to 98% using multiple scan-data. In the range of 10 m, over all plots, an average accuracy of 91% and 100% was observed using single- and multiple-scans respectively.

The detection accuracies decreased with increase in distance to the scanner. The detection rates decreased up to 30% from 5 m to 20 m range in case of single-scan approach whereas it was only 5% using multiple-scan approach (Figure 2). This was largely due to the increasing occlusion of trees leading to less visibility of the plot to the scanner. However, increase in plot visibility using multiple scan positions led to higher detection accuracy as expected. It was also observed that the trees with larger diameter were identified with more accuracy in both single- and multiple-scans. The high detection rates were also possible due to the dry season which allowed minimal hindrance from undergrowth. The plot-wise accuracy estimates of detection accuracy are provided in Table 2 and Figure 3. The tree location bias for majority of trees (90% using single-scan and 97% using multiple-scan) was

Table 2. Plot wise detection accuracy with respect to distance to the scanner (R) using both single- and multiple-scans

	Detection accuracy (%)					
	$R \leq 10$ m		$R \leq 15$ m		$R \leq 20$ m	
Plot ID	Single-scan	Multiple-scan	Single-scan	Multiple-scan	Single-scan	Multiple-scan
S1	100% (12/12)	100% (12/12)	96.5% (28/29)	96.5% (28/29)	82.5% (42/51)	98.0% (50/51)
S2	91.6% (11/12)	100% (12/12)	85.2% (29/34)	97.1% (33/34)	70% (42/60)	98.3% (59/60)
S3	94.1% (16/17)	100% (17/17)	84.8% (33/33)	100% (33/33)	73.4% (47/65)	95.3% (61/64)
S4	86.7% (12/15)	100% (15/15)	88.9% (24/27)	96.2% (26/27)	73.1% (38/52)	96.2% (50/52)

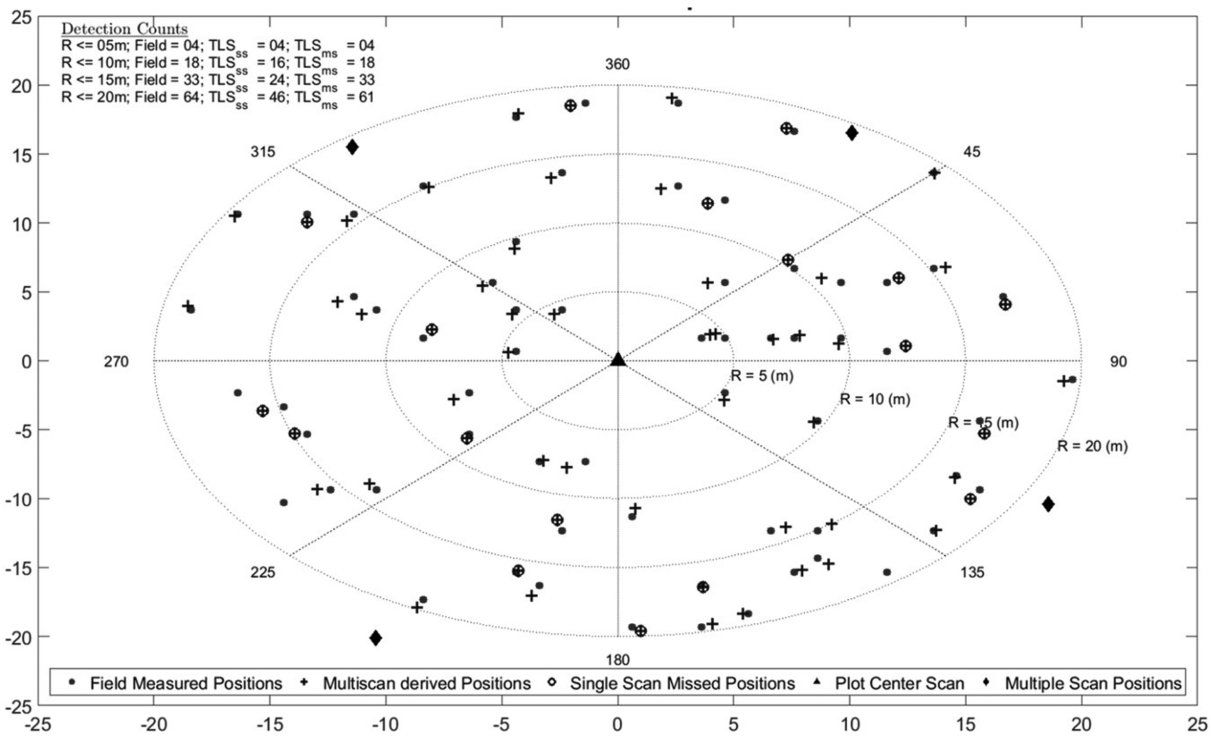

Figure 3. Tree location map of plot S3 as derived using TLS.

found to be less than 1 m (Figure 4). The result was anticipated since the tree locations measured on ground using metre tape were marked with accuracy up to 1 m. Further, the deviations in bias could also be due to the deviations in tree start positions and stem position at 1.3 m.

Out of 227 trees (across 4 plots), 169 trees were detected using single-scan approach and 220 trees were detected using multiple-scan approach. Estimated DBH values for all detected trees ranged from 10 to 45 cm. The estimated DBH values were well in agreement with the field measured values, with an R^2 of 0.96 and 0.99 for single- and multiple-scans respectively. The RMSE in DBH measurements was found to be 4.1 cm in case of single and 2.2 cm using multiple scans.

When plots are described using single-scan, DBH was estimated with a relative RMSE of about 22%. In the range of 10 m, RMSE in estimated DBH using single-scan was similar to that of multiple-scan. The higher error in DBH estimation using single-scan was caused by trees with large distances from the scanner. Another possible reason could be due to large trunks, wherein only a limited portion was seen using single direction. It was observed that error in DBH estimation increased with increase in trunk diameter using single-scan. The relative RMSE increased by 5% when trunks with diameters greater than 30 cm in comparison with trunks with diameter less than 15 cm were measured using single-scan (Table 3). This is due to the fact that large diameter trees appear to form different circles when viewed from different directions and estimation of circles using single-scan always underestimates/overestimates the actual value.

It was found that inclusion of scans from different directions in a multi-scan approach better described the

Table 3. DBH class wise detection accuracy and relative RMSE error in DBH using both single- and multiple-scans

	Detected trees (%)		DBH–RMSE Rel. (%)	
DBH class	Single-scan (%)	Multiple-scan (%)	Single-scan (%)	Multiple-scan (%)
DBH ≤ 15 cm	67.01	95.88	18.25	14.21
15 < DBH ≤ 30 cm	81.65	96.33	19.49	10.92
DBH > 30 cm	68.18	100.00	23.42	10.04

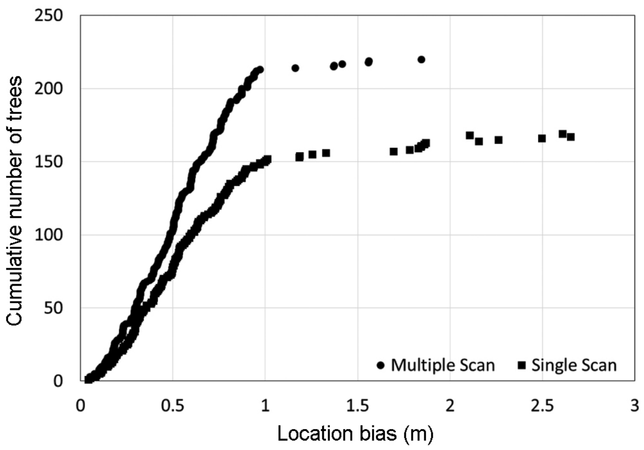

Figure 4. Cumulative number of trees with location bias in tree positions.

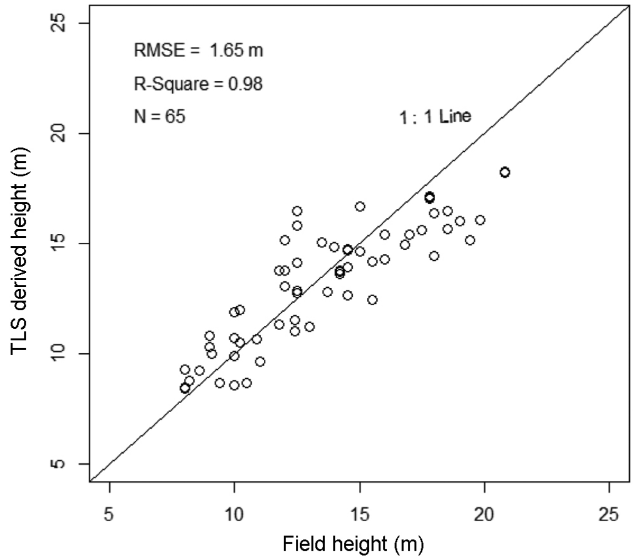

Figure 6. Comparison between field measured tree height and TLS derived tree height over selected trees across all plots.

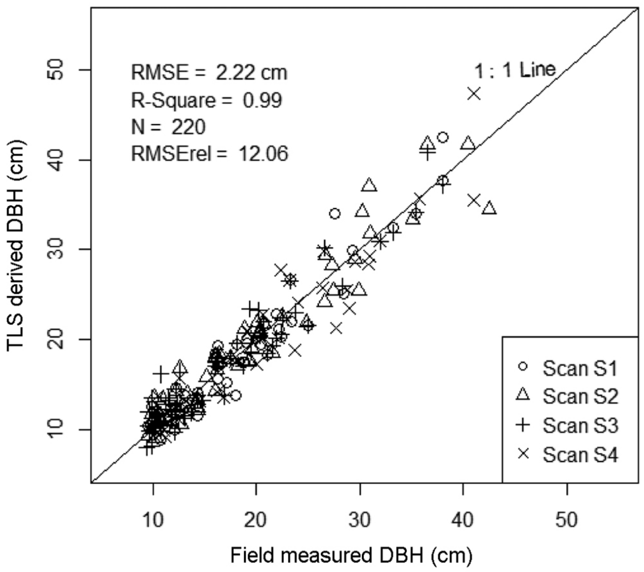

Figure 5. Scatter plot for TLS derived DBH values versus field measured DBH values using multiple scans for four plots.

trunk and resulted in relative RMSE of 12% in DBH (Figure 5). Also, an error rate of 10% was observed in trees with DBH >15 cm (Table 3). For, smaller DBH trees (DBH ≤ 15 cm), the error was comparatively higher (14%) and was majorly due to the fact that smaller trees are likely to be partly occluded by bigger trees even though viewed from multiple scan positions, thus causing over/under estimation of estimated circle.

On the ground, heights of 65 trees across 4 plots were measured. The heights extracted from multi-scan TLS data for these 65 trees were found to be well in agreement ($R^2 = 0.98$) with the field measured height (Figure 6). The RMSE in height estimation was 1.65 m. The discrepancy in height estimation with TLS in comparison to ground was due to two facts. First, often, reliable tree height measurements from ground are difficult since it is tough to identify the top of a tree in a forest. Secondly, there is no way to assess whether the tree top points are captured in the point cloud since there are many occlusions from crown and branches of the tree and its neighbours.

In the current study, tree detection rates were found to be 75% and 97% using single- and multiple-scan approaches respectively. Detection rates were comparable to globally reported studies, if not higher. Most studies using single-scan approach to detect trees reported accuracies ranging from 22% to 73% depending on plot size, stand density and forest type[5,9,16,17]. Liang *et al.*[17] used single-scan TLS data to automatically detect trees over 9 circular plots of 10 m radius with average stem density of 1022 stems/ha, and reported an accuracy of 73%. In comparison with current results, over 10 m radius, detection rate over four plots using single-scan approach was 91% with an average stem density of 450 stems/ha. Olofsson *et al.*[5] characterized 16 circular plots of 20 m radius (average stem density 741 stems/ha) using single-scan TLS data and achieved an accuracy of 73%, which is slightly less when compared to the detection rate in the current study. Multi-scan approach resulted in tree detection rates between 62.1% and 100% depending on the forest structure and scanning setup[18]. Kankare *et al.*[13] estimated DBH from multi-scan TLS data in different forest types with an accuracy of 1.7 cm using 2 circular plots of 20 m (122–480 stems/ha).

Tree height estimation using TLS data had not been explored much. Results from previous studies showed that the tree height is typically underestimated up to several metres. An RMSE of 0.76 m in height estimation using multi-scan data over one plot (212 stems/ha) was reported by Huang *et al.*[10]. A relatively high RMSE of 4.55 m for 9 selected trees on four plots (212–410 stems/ha) was reported by Maas *et al.*[2]. In the current study, tree height was underestimated with an RMSE of 1.65 m. The main source of underestimation of tree height by TLS is that the top of the tree crown is occluded by itself or a neighbouring tree (i.e. wide crowns of trees do not allow the scanner to see the tree tops).

The current study presents an automatic solution to derive forest inventory parameters from both single- and multiple-scan TLS data. The experimental results suggests that 20 m circular plots are better described by using multiple scans from different directions in comparison with single scans from the plot centre. However, it was found that in the range of 10 m, single scans were also able to describe the plot with reliable accuracies. The results strongly support the potential use of TLS in forest inventory. Further studies towards accurate estimation of tree attributes would help understand the added advantage of using TLS and would likely challenge the efficiency of conventional methods in terms of accuracy. However, application of TLS in forest inventories is hampered by difficulties in automated point cloud processing and consistent estimation of forest inventory parameters. It is highly likely that in the near future, TLS would be used in individual tree measurements in field plots to support airborne or satellite driven regional and national field inventories.

1. Wang, B., Huang, J., Yang, X., Zhang, B. and Liu, M., Estimation of biomass, net primary production and net ecosystem production of China's forests based on the 1999–2003 National Forest Inventory. *Scand. J. For. Res.*, 2010, **25**, 544–553.
2. Maas, H.-G., Bienert, A., Scheller, S. and Keane, E., Automatic forest inventory parameter determination from terrestrial laser scanner data. *Int. J. Remote Sensing*, 2008, **29**, 1579–1593.
3. Bienert, A., Scheller, S., Keane, E., Mohan, F. and Nugent, C., Tree detection and diameter estimations by analysis of forest terrestrial laserscanner point clouds. In ISPRS Workshop on Laser Scanning, 2007, pp. 50–55.
4. Xu, W. *et al.*, Comparison of conventional measurement and LiDAR-based measurement for crown structures. *Comput. Electron. Agric.*, 2013, **98**, 242–251.
5. Olofsson, K., Holmgren, J. and Olsson, H., Tree stem and height measurements using terrestrial laser scanning and the RANSAC algorithm. *Remote Sensing*, 2014, **6**, 4323–4344.
6. Antonarakis, A. S., Richards, K. S., Brasington, J. and Muller, E., Determining leaf area index and leafy tree roughness using terrestrial laser scanning. *Water Resour. Res.*, 2010, **46**.
7. Hackenberg, J., Spiecker, H., Calders, K., Disney, M. and Raumonen, P., SimpleTree – an efficient open source tool to build tree models from TLS clouds. *Forests*, 2015, **6**, 4245–4294.
8. Zheng, G., Moskal, L. M. and Kim, S.-H., Retrieval of effective leaf area index in heterogeneous forests with terrestrial laser scanning. *IEEE Trans. Geosci. Remote Sensing*, 2013, **51**, 777–786.
9. Thies, M. and Spiecker, H., Evaluation and future prospects of terrestrial laser scanning for standardized forest inventories. *Forest*, 2004, **2**, 1.
10. Huang, H. *et al.*, Automated methods for measuring DBH and tree heights with a commercial scanning lidar. *Photogramm. Eng. Remote Sensing*, 2011, **77**, 219–227.
11. Simonse, M., Aschoff, T., Spiecker, H. and Thies, M., Automatic determination of forest inventory parameters using terrestrial laser scanning. In Proceedings of the Scand Laser Scientific Workshop on Airborne Laser Scanning of Forests, 2003, pp. 252–258.
12. Astrup, R., Ducey, M. J., Granhus, A., Ritter, T. and von Lüpke, N., Approaches for estimating stand-level volume using terrestrial laser scanning in a single-scan mode. *Can. J. For. Res.*, 2014, **44**, 666–676.
13. Kankare, V. *et al.*, Diameter distribution estimation with laser scanning based multisource single tree inventory. *ISPRS J. Photogramm. Remote Sensing*, 2015, **108**, 161–171.
14. Fischler, M. A. and Bolles, R. C., Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 1981, **24**, 381–395.
15. Chum, O., Two-view geometry estimation by random sample and consensus, Czech Technical University in Prague, 2005.
16. Strahler, A. H. *et al.*, Retrieval of forest structural parameters using a ground-based lidar instrument (Echidna{®}). *Can. J. Remote Sensing*, 2008, **34**, S426–S440.
17. Liang, X. *et al.*, Automatic stem mapping using single-scan terrestrial laser scanning. *IEEE Trans. Geosci. Remote Sensing*, 2012, **50**, 661–670.
18. Liang, X. *et al.*, Terrestrial laser scanning in forest inventories. *ISPRS J. Photogramm. Remote Sensing*, 2016, **115**, 63–77.

ACKNOWLEDGEMENT. The present study is funded by Indian Space Research Organisation's Geosphere Biosphere Program (IGBP).

Received 1 March 2017; revised accepted 1 August 2017

doi: 10.18520/cs/v114/i01/201-206