Quality Assessment of Terrestrial Laser Scanner Ecosystem Observations Using Pulse Trajectories

Ian Paynter, Daniel Genest, Edward Saenz, Francesco Peri, Zhan Li, Alan Strahler, *Member, IEEE*, and Crystal Schaaf

***Abstract*— Considering the trajectories of pulses from terrestrial laser scanners (TLS) can provide refined models of occlusion and improve the assessment of observation quality in forests and other ecosystems. By considering the space traversed by light detection and ranging (lidar) pulses, we can separate empty regions of an ecosystem sample from unobserved regions of an ecosystem sample. We apply this method of TLS observation quality assessment, and analyze Compact Biomass Lidar 2 (CBL2) TLS observations of a single tree and of a deciduous forest stand. We show the contribution of information from each TLS scan to be inconsistent and the combination of multiple scans to have diminishing returns for new information, without guaranteeing complete coverage of a sample. We quantitatively investigate the effects of imposing information quality requirements on TLS sampling, for example, requiring minimum numbers of observations in each region or requiring regions to be observed from a minimum number of independent scans. We show empirically that rigid, predefined TLS sampling schemes, even with hypothetically dense coverage, cannot guarantee successful samples in geometrically complex systems such as forests. Through these methods, we lay the groundwork for on-the-fly assessment of observation quality according to several modeling-relevant metrics which enhance TLS ecosystem assessment. We also establish the value of flexible deployment options for TLS instruments, including the ability to deploy at a variety of heights.**

***Index Terms*— Forestry, laser beams, laser measurements.**

Manuscript received March 23, 2017; revised August 23, 2017 and March 3, 2018; accepted March 25, 2018. Date of publication June 15, 2018; date of current version October 25, 2018. This work was supported in part by the Oracle Graduate Fellowship, in part by the NASA Harriet Jenkins Graduate Fellowship, and in part by NASA under Grant NNX14AK12G and Grant NNX14AI73G. *(Corresponding author: Ian Paynter.)*

I. Paynter is with the Universities Space Research Association, Greenbelt, MD 20771 USA, and also with the School for the Environment, University of Massachusetts Boston, Boston, MA 02125-3393 USA (e-mail: ianpaynter1@gmail.com).

D. Genest, F. Peri, and C. Schaaf are with the School for the Environment, University of Massachusetts Boston, Boston, MA 02125-3393 USA (e-mail: daniel.genest001@umb.edu; francesco.peri@umb.edu; crystal.schaaf@umb.edu).

E. Saenz is with Spectral Evolution Inc., Lawrence, MA 01840 USA, and also with the School for the Environment, University of Massachusetts Boston, Boston, MA 02125-3393 USA (e-mail: edward.saenz@umb.edu).

Z. Li is with the Pacific Forestry Centre, Canadian Forest Service, BC V8Z1M5, Canada, and also with the School for the Environment, University of Massachusetts Boston, Boston, MA 02125-3393 USA (e-mail: zhan.li@umb.edu).

A. Strahler is with the Department of Earth and Environment, Boston University, Boston, MA 02215 USA (e-mail: alan@bu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2018.2836947

I. Introduction

In THIS paper, we compute the trajectories of light detection and ranging (lidar) pulses from terrestrial laser scanner (TLS) observations of an individual tree and a forest stand to assess the extent and quality of observations resulting from various sampling strategies. TLS make large numbers of discrete lidar observations of their surroundings, which are generally used to infer the position and shape of objects [1]. The ability of TLS to provide detailed representations and reconstructions of their surroundings has led to the derivation of estimates of ecosystem properties [1]–[3], as well as validation for estimates of ecosystem properties derived from satellite and airborne remote sensing instruments [4]–[6]. However, lidar is limited to line-of-sight observation. Since the complex structure of ecosystems (the number, geometry, arrangement, and spectral properties of objects) guarantees that some regions of an ecosystem will be occluded from any given position, some regions of an ecosystem will inevitably be omitted from a sample [7], [8]. Forests, where TLS have gained considerable traction as observation tools, are particularly challenging [8] as trees are complex objects of highly variable morphology, which occur in highly variable arrangements. Forests also have stand-scale structural tendencies that can greatly increase occlusion, such as vertical stratification, understory development, and crown intermingling.

Current attempts to mitigate the uncertainty from occlusion in TLS data typically involve the use of rigid sampling schemes that hypothetically produce dense coverage. The remaining uncertainty in the TLS observations is then characterized by comparing the TLS-derived estimation of an ecosystem property to a validation measurement obtained by some other means. For example, a tree volume calculated from a quantitative structure model [9], constructed from TLS point-cloud data, can be compared to the measured volume of that tree after it has been destructively sampled [1]. However, these methods only allow us to understand the uncertainty arising from occlusion after it is propagated to error in models, by which point it is confounded by other sources of uncertainty [10]. Therefore, it is desirable to assess occlusion and other metrics of TLS observation quality, further upstream in the analytical process, ideally at the level of the individual lidar samples of an environment.

Metrics that consider the presence and density of lidar returns are already employed to infer TLS observation extent,

since the presence of a lidar return at a specific point in space indicates that a pulse was scattered from an object in that location, and therefore that the point in space was observed [1], [11], [12]. However, existing work has shown that a more direct model of occlusion can be formed by considering the complete optical path of each lidar pulse in TLS scans, as this process can distinguish between unobserved regions of a sample (not checked for an object during sampling) versus empty regions (checked for an object during sampling, and none was found).

The consideration of lidar pulse trajectories has been utilized for a variety of purposes in several previous studies. Certain methods for retrieving ecosystem properties from lidar observations include the path length and angle of pulses as a parameter implicitly describing occlusion [13], [14]. Béland *et al.* [15], [16] assessed the occlusion in voxels according to the proportion of pulses that traversed those voxels. In addition, Béland *et al.*'s [15], [16] studies linked the quantification of occlusion to accuracy in downstream modeling results. Hancock *et al.* [11], [17] utilized metrics derived from pulse pathways as a surrogate for uncertainty, feeding this back into upstream determination of sampling parameters such as plot size and scan numbers. Most recently, Kükenbrink *et al.* [18] evaluated regions of occlusion in forest canopies, and estimated uncertainties in canopy volume retrievals by modeling the paths of pulses from an airborne lidar scanner.

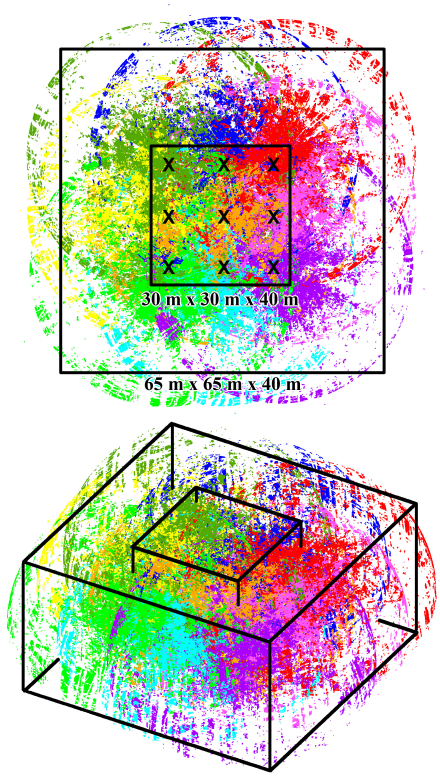


Fig. 1. Point clouds from the nine scans (one color per scan, black x denotes scan location) of the CBL2 of the Harvard Forest hardwood plot (MA, USA). Includes points corresponding to the maximum range of pulses with no returns. The full stand and subset used in this paper are delineated by the black boxes in (top) plan and (bottom) oblique views.

II. Methods

TLS data utilized in this paper were collected with the Compact Biomass Lidar 2 (CBL2) instrument (University of Massachusetts Boston, Boston, MA, USA), based on the SICK LMS-151 lidar unit [10]. The CBL2 is a 905-nm wavelength, time-of-flight, discrete scanning lidar with a 0.25° angular resolution (separation between pulses) and a 360° horizontal (azimuthal) and a 270° vertical (zenithal) angular scanning range. The pulse rate of 50 kHz provides a very rapid 33-s scan with a maximum range of 45 m. Beam divergence is 0.86° (15 mrad), and the beamwidth at the optical center is 8 mm.

It should be noted that the CBL2 has a coarser angular resolution (angular separation between emitted pulses), shorter maximum range, and wider beam divergence than some other contemporary instruments [1], [19], [20]. For reference, the Riegl VZ-400, used in many relevant studies [1], [12], [21], has an angular resolution of 0.0024°, a maximum range of 600 m, and a beam divergence of 0.02° (0.35 mrad). However, the rapid scanning and portability of the CBL2 make it a useful tool for investigative studies such as these.

Two sets of CBL2 scans were acquired, one covering a temperate forest stand and the other capturing a single tree to act as demonstrations of this observation quality strategy. All scan positions of both sets were established with surveying methods and equipment. The forest stand is a long-term hardwood research site in Harvard Forest, MA, USA, and scans were acquired in August 2014 during leaf-ON conditions. Scans were taken on a grid consisting of nine vertices, with internal edges of 12.5 m in length (Fig. 1). Field measurements for five 20-m diameter plots within the forest stand had a mean stem density of 0.76 stems per square meter (standard deviation 0.08) and a mean diameter at breast height (DBH) of 20.56 cm (standard deviation 13.84). A subsample of 60 trees had a mean height of 20.33 m (standard deviation 6.28 m), a mean crown base height of 9.21 m (standard deviation 3.63), and a mean crown diameter of 8.80 m (standard deviation 3.05 m). The most common tree species in the plots were yellow and black birch, red oak, and red maple.

The single tree was a dead black locust tree, devoid of leaves but otherwise intact, located in managed parkland in Dorchester, MA, USA. A dead tree was used so that scans could be taken in leaf-OFF conditions to better understand the relationship of scanning configuration to tree structure, while utilizing the longer daylight hours of the summer to make all observations in a single day. Scans of the single tree were acquired in June 2015. Scans were taken at four cardinal points from a distance of 5 m and at two heights (1.4 and 3 m), totaling eight scans (Fig. 2).

The lidar data were processed into returns with local polar and Cartesian co-ordinates. The positions of the returns from each scan were then transformed according to the scan positions in the field, which had been established with traditional surveying techniques. Any residual disagreement between the scans evident in visualizations was mostly from small rotational discrepancies, likely resulting from slight differences in the placement of the instrument in the field. These were

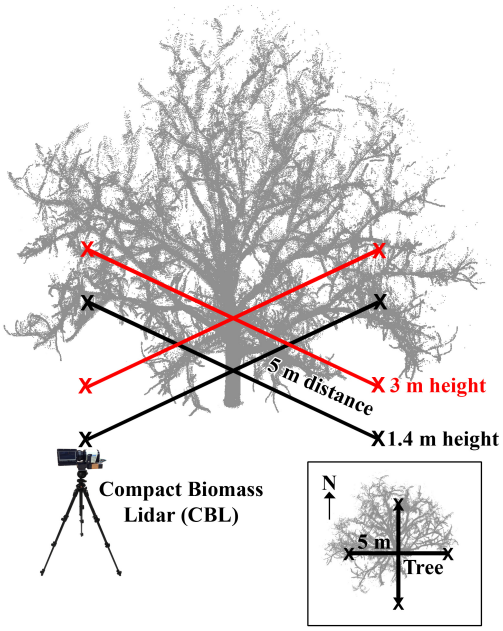


Fig. 2. Sampling scheme for the single, leafless black locust tree located in managed parkland in Dorchester, MA, USA, scanned in June 2015. Four scans were taken in cardinal directions at each height (black: 1.4 m; red: 3 m).

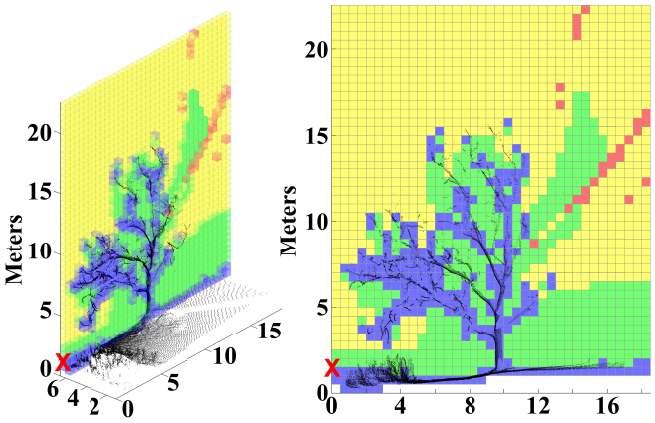


Fig. 3. Demonstration of types of pulse/voxel interactions for a single vertical "slice" of voxels from a single scan location (red X). Black points are the CBL2 point cloud. Blue voxels contained a return. Green voxels were traversed by a pulse which eventually generated a return. Yellow voxels were traversed by a pulse which did not generate a return. Red voxels were not observed. Note that voxels often contain more than one kind of interaction (Fig. 4).

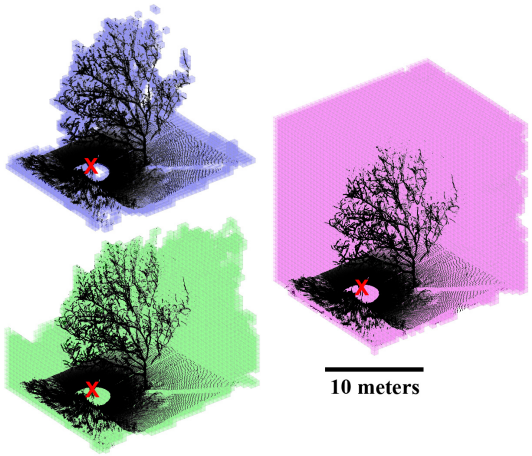


Fig. 4. Returns (black points) from CBL2 scan (red X) of the single tree. (Top) Blue voxels contain returns. (Middle) Green voxels were traversed by pulses which eventually generated a return. (Bottom) Pink voxels were traversed by pulses which did not generate a return. The addition of the pulse trajectory information to characterize space which is observed, but empty, enables refined estimates of unobserved space (Fig. 5).

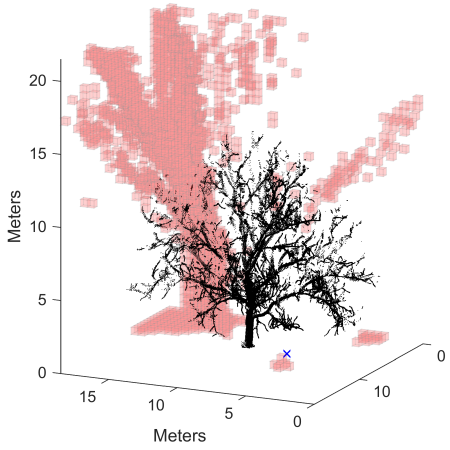


Fig. 5. Returns (black points) from CBL2 scan (blue X) of the single tree. Red voxels are refined model of unobserved regions after consideration of pulse trajectories (Fig. 4). The major region of occlusion is directly behind the tree, relative to the scan position. A small region of occlusion under the scan position results from the 270° vertical sampling range of the CBL2. An additional small region of unobserved voxels to the bottom right of the plot corresponds to occlusion from tall grass.

removed by fine-scale, iterative, supervised rotational transformations of each scan.

The trajectory of each pulse was computed. Pulses with no return were assumed to have a maximum range of 40 m. In pulses with returns, only portions of pulses up to the first return were utilized, since object detection becomes more inconsistent after the first scattering event. To spatially summarize the data, interactions precise to 1-mm intervals were derived between the trajectory of each pulse and a 3-D grid of 0.5-m voxels (discrete cubic regions existing on a continuous Cartesian coordinate grid).

A record was made of each time a scan contributed a pulse to a voxel. These pulse/voxel interactions were categorized into three types: pulses that terminated in a return within the voxel; pulses that terminated in a return in a subsequent voxel; and pulses that passed through the voxel but did not ultimately terminate in a return (Figs. 3 and 4). The remaining voxels with no record of any pulse/voxel interactions were deemed unobserved due to occlusion or the angular separation between pulses (Fig. 5).

For the stand-scale study, voxel data were then analyzed both in a $65 \times 65 \times 40$ m^3 volume encompassing the maximum ranges of all the scans and in a $30 \times 30 \times 40$ m^3 subset bounded by the centers of the scans (Fig. 1). For the single tree study, voxel data were assessed in a $16.5 \times 18 \times 22$ m^3

TABLE I
SUMMARY OF RESULTS AND ATTRIBUTES

Attribute	Full Forest Stand	Forest Stand Subset	Single Tree
dimensions (meters)	65 x 65 x 40	30 x 30 x 40	16.5 x 18 x 22
number of scans	9	9	8
total voxels	1,273,776	321,164	55,260
voxels observed	711,065	207,225	55,247
voxels unobserved	562,711	113,939	13
total pulse/voxel (p/v) interactions	153,043,018	57,010,990	162,665,643
lidar returns	6,675,611	2,721,943	2,542,224
p/v interactions (with returns)	132,311,056	50,025,949	65,621,481
p/v interactions (no returns)	14,056,351	4,263,098	94,501,938
min. voxels observed by a scan	133,084	39,077	52,311
max. voxels observed by a scan	204,141	90,199	53,635
coef. variation of voxels observed	0.1598	0.2614	0.0094
min. observations of voxels by a scan	12,757,340	2,290,928	18,506,491
max. observations of voxels by a scan	20,140,457	16,245,052	21,932,168
coef. variation of observations of voxels	0.1373	0.6142	0.06

volume, fully encompassing the tree (Fig. 2). In both studies, unobserved voxels occurring below the ground due to ground slope were excluded from analyses.

The relationship between the sampling schemes and the quality of the observations of the plots was investigated in two ways. First, to examine the influence of the total number of scans on the amount of information obtained, all sizes of scan sets from a single scan to all scans in each study with all possible combinations of scans for each set were considered. Second, to examine the variation in information gained for each additional scan, all possible permutations of the order in which scans could be combined were considered (362 880 for the forest stand, and 40 320 for the single tree). In other words, the contribution of each scan was assessed as if it was the first scan performed or as if it was the second scan performed, given each other possible first scan, and so on. In this way, the second process described information gain from a particular scan, given a particular preceding combination of scans, which is distinct from the total information contained within the resulting combination of scans.

For each permutation, the number of new voxels which were observed for the first time by each consecutive scan was calculated, along with the amount of observations made of the newly observed voxels. To investigate the potential impact of prerequisite information quality standards on sampling cost, the information gain was also recalculated to simulate the effects of imposing a range of potential threshold values.

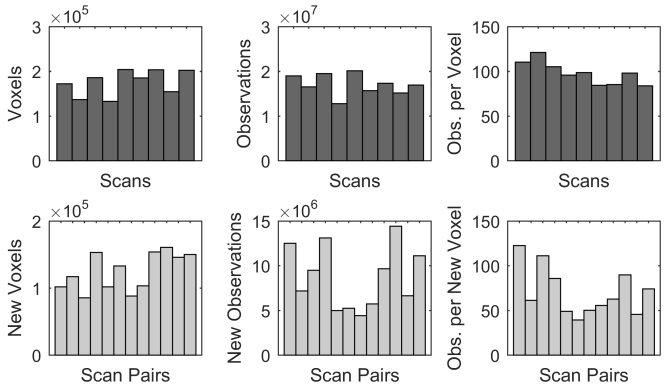


Fig. 6. (Top row) For each CBL2 scan and (Bottom row) for each pair of adjacent CBL2 scans in the full forest stand, (Left) total voxels observed, (Middle) total observations of voxels, and (Right) observations per voxel. Note the inconsistent information gain between scans and between pairs of scans. These data suggest that single TLS scans offer variable amounts of information, and spatial coverage is neither complete nor consistent, even when multiple overlapping scans are combined.

Potential threshold values for minimum observations per voxel (10, 25, 50, 75, and 100); and minimum numbers of scans observing a voxel (2, 3, and 4) were considered.

III. RESULTS

The forest stand and single tree studies considered a total of 17 scans and 13.2 million total pulses. The numbers of observed and unobserved voxels and occurrence of each type of pulse/voxel interaction for each study can be found in Table I. In the full stand-scale data set, which included a large area beyond the boundaries of the scanning grid (Fig. 1), 44.2% of voxels were unobserved. The percentage of unobserved voxels dropped to 35.5% for the subset of the volume within the boundaries of the scanning grid. In the full forest stand data sets (Fig. 6), individual scans observed varying numbers of voxels [0.16 coefficient of variation (CV) (see Table I)], and contributed varying numbers of total observations of voxels (0.14 CV) (Fig. 6). The variation in contribution per scan was considerably higher for the forest subset (0.26 CV voxels observed, 0.62 CV total observations), but considerably lower for the single tree (0.009 CV voxels observed, 0.06 CV total observations).

For the single tree study, when all eight scans were combined, only 0.02% of voxels were unobserved (Fig. 7). Due to the relative openness of the single tree scene, there was a much higher occurrence of pulses intersecting voxels, without ever producing a return, in the single tree data set (11.81 million events per scan) than either the full forest stand (1.56 million events per scan) or the forest stand subset (0.47 million events per scan). Once even a few scans were combined, this leafless tree and its immediate surroundings were almost completely observed (Fig. 7). Once all eight scans were combined, only a single, localized region of unobserved voxels remained, which was associated with occlusion from a patch of tall grass (visible in Fig. 5).

The spatial distribution of the different types of pulse/voxel interactions are intuitively linked to the scan positions and the

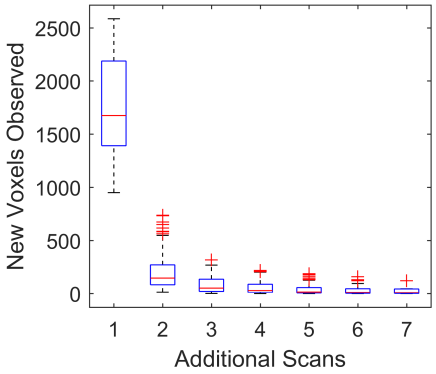


Fig. 7. Box plot of the number of new voxels that contained returns observed by all permutations of sequentially added CBL2 scans of the single tree. These data suggest that a great deal of important information is consistently added by a second scan, but additional new portions of the tree structure were also observed by third and fourth scans. New regions were observed even when an eighth scan was added. It could be argued that 3–4 scans consistently capture most of the tree structure, in this case.

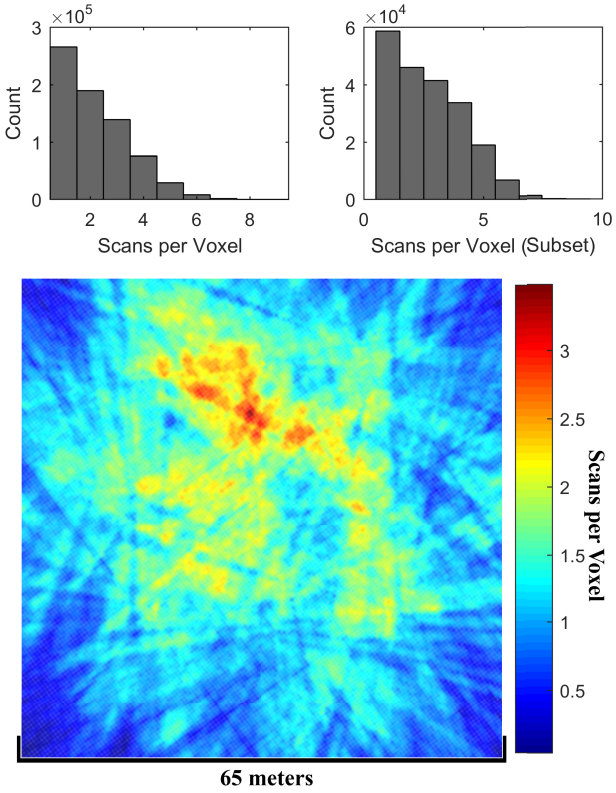


Fig. 8. Number of scans which independently observed voxels (Top left) for the full forest stand and the (Top right) stand subset. (Bottom) Within the full stand, the spatial distribution of the mean number of scans observing each voxel creates distinct patterns attributable to occlusion by tree and demonstrates many rarely and exclusively observed regions.

ecosystem and object structure (Figs. 8–10). The unobserved regions were visibly occluded by particular objects (such as trunks, branches, and vegetation), with objects closer to the scanner creating larger regions of occlusion. Unsurprisingly, the forest stand data set contained a lot of unobserved voxels in regions that were behind trees, relative to scan positions.

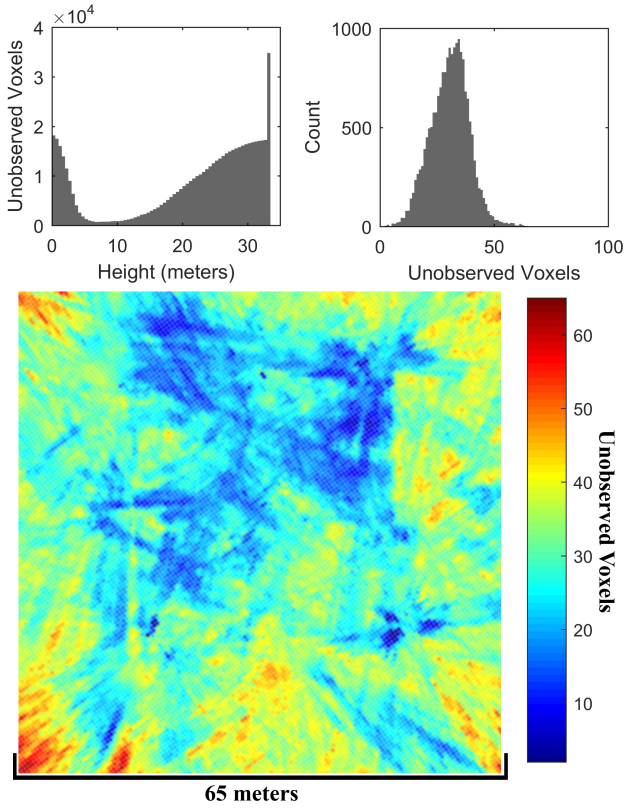


Fig. 9. (Top left) Number of unobserved voxels for the full forest stand tended to increase with height, as canopy structure creates occlusion, and range increases the angular separation between CBL2 pulses. Once the height exceeded the range at non-zenith angles, almost all voxels were unobserved. (Bottom) There was also substantial spatial variation in the amount of occlusion. (Top right) Left-hand skew of the distribution suggests that there were small regions of occlusion throughout much of the extent of the region, with the long right tail suggesting that a few regions were severely occluded. (Bottom) These highly occluded regions tended to be toward the edges of the study site, but there was also inconsistent light to moderate occlusion everywhere.

IV. Discussion

The variation in the contribution of scans from different positions within the same ecosystem samples illustrates the important point that the observation extent, and therefore information gained, from any TLS scan cannot be guaranteed. Furthermore, the contribution of scans can be expected to vary substantially, especially in more geometrically complex samples, as demonstrated by the higher variability in the forest stands versus the single tree. Extrapolating, a rigid sampling scheme (for example, grid-based scan positions, as used here for the forest stand) cannot be expected to produce consistent and complete observations of an ecosystem.

In this paper, we also see empirical evidence that observation quality is inconsistent even when a large numbers of scans and a high density of TLS observations are acquired. The 12 possible pairs of adjacent scans (12.5 m apart) for the forest stand were compared for the information added by the second scan in each pair. There was substantial variation in the number of newly observed voxels and observations within new voxels by the additional scans (Figs. 8 and 9). This variability suggests that simply increasing observation density without

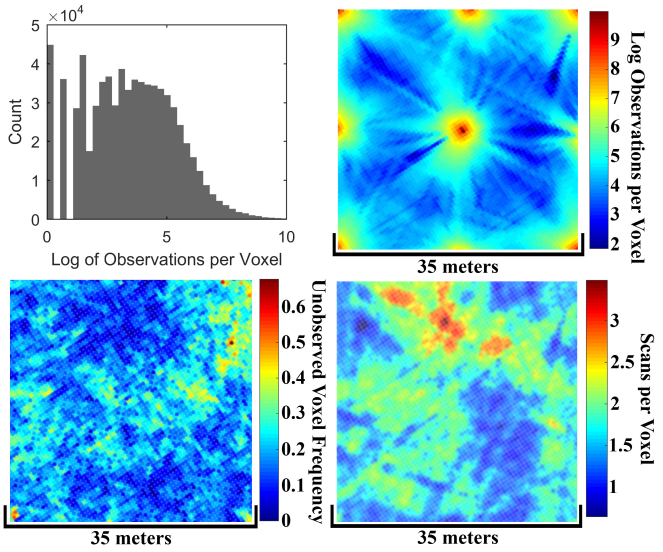


Fig. 10. (Top left) Density of observations varies greatly in the forest stand subset. (Top right) Voxels which are observed many hundreds of thousands of times occur at the scan locations. Generally, observation density declines with distance from a scan, though patterns resulting from occlusion by trees are obvious both in (Bottom left) reduced observation density and (Bottom right) reduced numbers of scans observing voxels.

assessing the specific occlusion in an ecosystem sample is not likely to overcome the strong and variable occlusion patterns of forest ecosystem. In the subset of the forest stand surrounded by eight scan positions, and containing a ninth, there was almost complete hypothetical coverage of the volume (40-m range scans, separated by a maximum of 12.5 m, evaluated over a 40-m-height range). Regardless, even when all the nine scans were actually combined, 1492.2 m^3 remained unobserved (Fig. 10) through occlusion, though this represents only 3% of the cubic volume of the sample area.

In addition to the completely unobserved regions of the ecosystem samples, there were also regions which were observed quite rarely. Even in the densely scanned forest subset, a large number of voxels (59 041 voxels totaling 7380 m^3) were observed exclusively by single scan, and few voxels were observed by more than five scans. The fact that substantial regions in these ecosystem samples were rarely observed further supports the notion that individual scans will make inconsistent contributions to overall sampling extent and quality. Therefore, even densely gridded sampling schemes may not provide complete coverage of a sample.

This notion is further supported by the single tree data set, where additional scans across all possible permutations regularly observed new voxels even up to the total of eight scans (Fig. 7). Although the addition of newly observed voxels drops off rapidly as more scans are added (Fig. 7), the persistence of occlusion in a single tree data set is still remarkable because the data set only considers a compact volume containing a single object. This persistent occlusion implies that certain regions of ecosystem samples will be observable from a relatively small subset of view angles, and therefore quite specific scan positions.

Examining the number of scans which observed voxels, and the density of unobserved voxels (Figs. 8 and 9), over the horizontal extent of the forest stand, reveals that occlusion does have some predictable and intuitive tendencies. For example, the frequency of unobserved voxels increases toward the boundaries of the sample (Fig. 9). This tendency would be expected as these regions are both farthest from any one scan and are viewed from the smallest range of potential view angles between scans.

While the frequency of unobserved voxels in the forest stand data set may vary unpredictably in horizontal space, it does seem to be a function of height (Fig. 9). Frequency of unobserved voxels rises steadily with height, until it spikes as the height exceeds the maximum range of the instrument. Of course, TLS observation extent should be expected to decrease with height, since there is both a greater angular separation between pulses with range, and greater potential for occlusion from the convoluted structure of forest canopies. Understory vegetation likely explains the small rise in frequency of unobserved voxels at low heights (Fig. 9).

The total number of observations of each voxel was also found to be highly inconsistent. This inconsistency is most clearly represented in the forest stand subset (Fig. 10), where we can see that observation density decreases with range from the nearest scan position, as would be expected. Voxels close to the scanner subtend a larger solid angle from the scanner and therefore receive more pulses. In addition, voxels directly above a scanner are sampled with every rotation of the scan mirror, since scans converge at the zenith. Extremely highly observed voxels are, therefore, intrinsically associated with scan positions. Prominent regions of low observation density in the horizontal distribution (Fig. 10) are again consistent with occlusion from trees.

During the sampling of the single tree, we utilized an extended tripod to obtain scans from an increased height (3 m). We compared these data to the scans acquired from the same horizontal position at the instrument's normal deployment height of 1.4 m. There were no voxels which were exclusively observed by either the four high scans or the four low scans. Also, there was no observed benefit of height, in terms of number of new voxels observed, when an additional scan was added to an initial low scan. This result runs counter to the intuitive notion that a higher scan position would result in more observations in the upper regions of the tree, and furthermore the strong relationship of observation quality metrics to height in the full stand also supports this expectation. However, the lack of occlusion in the single tree study, as well as the relatively low height of the single tree (17 m), likely serves to mask the intuitive benefit of deploying the TLS at greater heights.

Examining the number of observations of the tree structure over height, we see that high scans offered comparable numbers of observations of lower portions of the tree structure, but substantially higher numbers of observation for higher portions of the tree structure (Fig. 11). In other words, the higher scan positions offered more total observations of the tree per scan, with a large improvement in the observation density in the upper portions of the tree, at only a slight cost in the density of the observation in the lower portions.

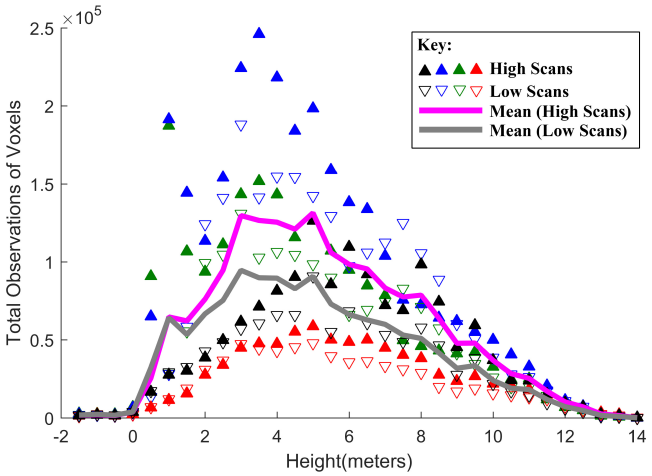


Fig. 11. For the single tree study, the number and height distribution of returns varied substantially between the different scan locations. High scans consistently made more observations of higher regions of tree structure than low scans taken at the same positions. However, neither high nor low scan positions make many observations of the uppermost portions of the tree structure, due to the integration of occlusion from canopy structural features.

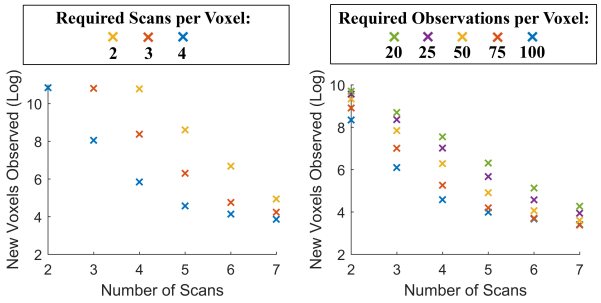


Fig. 12. Establishing information requirements for thresholds of (Left) observations per voxel and (Right) scans per voxel reduced the number of voxels considered successfully observed for the forest stand subset. This reduced success persisted even with the addition of several scans, exacerbated by imposing more stringent requirements.

Due to the emergence and refinement of technologies to provide real time, highly accurate estimates of the position and orientation of instruments, such as real-time kernel global positioning systems, and compact inertial movement units, as well as the continuing increase in mobile computing power and the continued improvements of TLS quality and efficiency, we foresee the eventual capability to assess TLS observations quality during the sampling of an ecosystem. The ability to track the position and orientation of an instrument permits the derivation of transformation matrices to register one scan with another, which can implemented with mobile computers of ever-increasing power and portability. However, operational responsive sampling requires a relevant metric of observation quality to track within the data and an appropriate threshold value of that metric to target in order to consider the observations complete. We suggest that appropriate metrics and threshold values could be established as follows.

Potential candidates for information quality metrics would first be related to validated observations of the target ecological variable, after the lidar observations have passed through the relevant processing and analysis. Suitable data sets for these preliminary analyses exist across a wide variety of ecosystems and instrument specifications for several key forest ecology variables such as DBH, tree height, and stem density. Candidate observation quality metrics could be as simple as a requirement that each region of the forest space, subdivided in some way, such as by voxels, is observed. Alternatively, observation quality metrics could be as specific as requiring a certain angular distribution of pulses to observe each region, such as the requirement for voxels to be observed by multiple discrete scans demonstrated in this paper. This particular metric would be appropriate for retrieving any ecological parameter that may have view-angle dependencies, such as gap probability.

Once a candidate metric has been established, the relationship of that metric to the downstream error in validated observations would be assessed. Major inflection changes or pronounced elbow points in this relationship would represent observation quality thresholds corresponding to a rapid change in the accuracy of downstream modeling. These influential values would make appropriate threshold values for the chosen observation quality metric. Different sampling scheme approaches would then be tested, evaluating their reliability in achieving the required threshold of quality. It should be noted that many TLS applications in forestry have achieved, or are in the process of achieving, the necessary steps to understand the most important attributes of data quality for the quality of their downstream modeling, and have established relevant target values for their attributes. In particular, lidar retrievals of DBH, tree height, and leaf area index are well-characterized for certain aspects of sampling-related uncertainty [11], [12], [15], [16], [22], [23]. Implementing responsive sampling schemes could be a useful next step for these applications.

To demonstrate the potential influence that information requirements would have on the necessary sampling effort, we analyzed the success of the TLS observations in this paper according to two intuitive types of information requirements. The first was a required number of observations per voxel, predicated on the assumption that the more observations of a particular region of space are obtained, the lower the errors in modeling its contents. The second information requirement tested was a specified minimum number of scans independently observing a voxel, based on the idea that acquiring additional view angles of an object would more completely observe its structure. We also investigated the effect of varying the stringency of these information requirements from 10 to 100 observations and 2–4 scans per voxel, respectively.

We found that increasing the required numbers of observations per voxel greatly decreased the number of "successfully" observed voxels in all data sets (Fig. 12). Considering all the possible permutations of order for the scans to be added to each data set, increasing the required numbers of observations per voxel also greatly reduced the rate at which voxels were observed (Fig. 12), considerably increasing the scanning time in the field. A similar reduction in observation totals and rate was evident with an increase in the number of scans required to observe each voxel (Fig. 12). While a reduction in observation

total and rate was inevitable in both cases, it was notable for its rapid fall off, particularly when requiring increasing numbers of observations per voxel.

Evaluating this study, it should be restated that the CBL2 employed in this paper has a coarser angular resolution (0.25°) than many contemporary instruments, a shorter maximum range (45 m), and a wider beam divergence (0.86°) [1], [19], [20]. An instrument such as the Riegl VZ-400 (angular resolution of 0.0024°, range of 600 m, and beam divergence of 0.02°) [1], [12], [21] would have far superior coverage and provide a much higher frequency of observations in voxels. In addition, the finer beam divergence would allow pulses to penetrate and see through and beyond smaller gaps, also providing more observations, and better observation coverage. Also noteworthy is the potential of mobile, continuous-scanning TLS such as the Zebedee, which, while quite coarse resolution and quite limited range may utilize its mobility to considerably mitigate occlusion issues. In fact, scanners with enhanced deployment flexibility, such as the CBL2 and Zebedee, may be suitable to use in tandem with instruments with higher overall performance to complete the coverage of observations. Overall, therefore, the potential influence of instrument-dependent factors on observation quality should always be considered in the planning of an acquisition campaign [10], [11], [14], [19], [20], [22], [23].

It should be noted that there are also several potential sources of error that could result in pulse/voxel interactions being placed in the incorrect voxel. Errors in scan registration, the atypically wide beam divergence of the CBL2, and a lack of detection of strongly absorbing or strongly specular surfaces, or surfaces orientated other than orthogonally to the pulse, may all contribute to false observations of false pulse/voxel interactions. The careful registration of scans, and the filtering the data to use only first returns, and portions of pulses up to first returns were employed to mitigate these concerns for this paper. However, care should also be taken when assessing particular characteristics of individual forests, for example for a forest following a fire, as burned bark may be strongly absorptive.

In addition, while voxelization is a convenient method to assess the spatial attributes of pulse trajectory information, it must be understood that voxelization spatially averages the data. This spatial averaging may offer some protection against individual, high magnitude errors in the position of lidar returns, but at the cost of potential information from the specific spatial properties of any given pulse. Voxelization is just one potential method of partitioning the space to assess observation quality, but it was utilized in this paper as it offers intuitive quantification of the distribution of lidar observations; it is a common method employed in TLS ecological research, and it may also offer the computational efficiency to permit the assessment of occluded areas and information quality on-the-fly during sampling in the future.

The voxel size of 0.5 m used in this paper is coarse compared to contemporary voxelization studies aimed at estimating vegetation content [12], [13], [16]. While this voxel size is appropriate to the specific aim of this paper, namely, to establish the irregular and unpredictable patterns of occlusion that occur in trees and forests, and highlight potential techniques to mitigate the resulting uncertainty in observations, such coarse voxels may not be appropriate for eventual operational use. The appropriate voxel size for a particular study should depend on the overall size and specific morphology of the trees, as well as the spacing of the trees within the environment. Large tree trunks, for example, could entirely encompass smaller voxels, preventing their observation and biasing quality assessments. Appropriate voxel size may also depend on the importance to a specific study of capturing the particular structural features of a forest. For example, in a forest with very high stem density, the smaller gaps between the closely spaced stems could be bridged by larger voxels. While pulses encountering the stems would technically have made observations of the region described by the voxel, they may not be representing the important aspects of the structure within it, namely, the individual stems and the gaps between them.

Particular types of forest may also require specific parameterization of quality assessment methods. Forests predominantly consisting of broadleaf trees, for example, may benefit from a smaller voxel size to precisely describe the observations of the complex canopy structure and gaps, while regions of more compact and uniform coniferous trees might be better assessed with a larger voxel size. For ecosystems such as tropical forests, which contain a wide variety of tree morphologies, as well as complex understory, and distinct vertical stratification, hybridized models involving different parameterizations, and even different quality assessment metrics, may be appropriate. Also worthy of consideration when selecting an appropriate voxel size will be the resulting computational efficiency, and the frequency of misplacement of pulse/voxel interactions, influenced by both voxel size and individual scanner properties such as beam divergence.

V. Conclusion

Utilizing the information contained implicitly in the trajectories of TLS pulses has been previously shown to add valuable context to TLS observations, and in this paper, trajectories are offered as a more refined model of observation quality for forest ecosystem samples, particularly by distinguishing between unobserved regions versus empty regions. Through this paper, we have empirically demonstrated that the extent and quality of TLS observations of an ecosystem sample can vary considerably. It is, therefore, often difficult to capture a complete ecosystem sample even when dense TLS observations are available, particularly in geometrically complex ecosystems such as natural forests. Even on the scale of a single, isolated object like the single tree in this paper, complex occlusion patterns still challenge the ability of TLS to observe every region of a sample. The most important conclusion from these findings is that rigid, predetermined TLS sampling schemes (number and arrangement of scans) can never guarantee a particular observation extent or quality in ecosystems with irregular geometry.

To improve future TLS studies, we suggest the eventual implementation of two novel strategies. The first strategic

component is *a priori* definition of information requirements to consider an ecosystem sample successfully observed. The exact metrics and stringency of these information requirements would obviously be relative the specific purpose of the study. Information requirements for a study could generally be derived by relating the error in downstream estimations of ecosystem properties to the quality of the upstream TLS data. The data to derive these quantitative relationships could be acquired from a combination of rapidly deployed pilot studies (as represented herein), analogous previous studies, or by simulation.

The second strategic component to implement in future TLS studies is the measurement of chosen information quality metrics during sampling, in order to monitor the quality of the sample, and thus guide additional observations until the requirements for success are met. The computational power available through contemporary mobile devices is already sufficient to operationally employ methods such as those utilized in this paper, which consider the trajectory of lidar pulses to assess TLS observation quality.

We can also conclude that flexibility in the deployment of TLS instruments can be of considerable benefit to improving the extent and quality of observations. The ability to complete scans rapidly, acquire scans in quick succession, and deploy at multiple heights may be extremely beneficial to overcoming the challenges of occlusion in forests. Conceivably, in the future, more compact and rapid TLS instruments could be routinely deployed in tandem with highly capable instruments that offer natively higher information quality.

Overall, the approaches in this paper can be used to provide the immediate refinement of the quality assessment of TLS observations, and ultimately, these techniques are a step toward guaranteeing the quality of TLS observations at both the object and ecosystem scale.

ACKNOWLEDGMENT

The authors would like to thank Harvard Forest, Petersham, MA, USA, and Harvard University, Cambridge, MA, USA, for logistical assistance.

REFERENCES

[1] K. Calders *et al.*, "Nondestructive estimates of above-ground biomass using terrestrial laser scanning," *Methods Ecol. Evol.*, vol. 6, no. 2, pp. 198–208, Nov. 2015, doi: 10.1111/2041-210X.12301.

[2] G. Newnham, J. Armston, and J. S. Muir, "Evaluation of terrestrial laser scanners for measuring vegetation structure," CSIRO, Canberra, Australia, Tech. Rep. EP124571, Jun. 2012, doi: 10.4225/08/584af4ccc9a19.

[3] J.-F. Côté, J.-L. Widlowski, R. A. Fournier, and M. M. Verstraete, "The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial LiDAR," *Remote Sens. Environ.*, vol. 113, no. 5, pp. 1067–1081, May 2009, doi: 10.1016/j.rse.2009.01.017.

[4] R. O. Dubayah and J. B. Drake, "LiDAR remote sensing for forestry," *J. Forestry*, vol. 98, no. 6, pp. 44–46, Jun. 2000, doi: 10.1093/jof/98.6.44.

[5] J. L. Lovell, D. L. Jupp, D. S. Culvenor, and N. C. Coops, "Using airborne and ground-based ranging LiDAR to measure canopy structure in Australian forests," *Can. J. Remote Sens.*, vol. 29, no. 5, pp. 607–622, Jun. 2014, doi: 10.5589/m03-026.

[6] N. Pfeifer and C. Briese, "Geometrical aspects of airborne laser scanning and terrestrial laser scanning," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 36, no. 3, pp. 311–319, Sep. 2007.

[7] D. Van der Zande, W. Hoet, I. Jonckheere, J. van Aardt, and P. Coppin, "Influence of measurement set-up of ground-based LiDAR for derivation of tree structure," *Agricult. Forest Meteorol.*, vol. 141, nos. 2–4, pp. 147–160, Dec. 2006, doi: 10.1016/j.agrformet.2006.09.007.

[8] M. Dassot, T. Constant, and M. Fournier, "The use of terrestrial LiDAR technology in forest science: Application fields, benefits and challenges," *Ann. Forest Sci.*, vol. 68, no. 5, pp. 959–974, Aug. 2011, doi: 10.1007/s13595-011-0102-2.

[9] P. Raumonen *et al.*, "Fast automatic precision tree models from terrestrial laser scanner data," *Remote Sens.*, vol. 5, no. 2, pp. 491–520, Jan. 2013, doi: 10.3390/rs5020491.

[10] I. Paynter *et al.*, "Observing ecosystems with lightweight, rapid-scanning terrestrial LiDAR scanners," *Remote Sens. Ecol. Conservation*, vol. 2, no. 4, pp. 174–189, Sep. 2016, doi: 10.1002/rse2.26.

[11] S. Hancock *et al.*, "Characterising forest gap fraction with terrestrial LiDAR and photography: An examination of relative limitations," *Agricult. Forest Meteorol.*, vol. 189, pp. 105–114, Jun. 2014, doi: 10.1016/j.agrformet.2014.01.012.

[12] M. Béland, D. D. Baldocchi, J.-L. Widlowski, R. A. Fournier, and M. M. Verstraete, "On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR," *Agricult. Forest Meteorol.*, vol. 184, pp. 82–97, Jan. 2014, doi: 10.1016/j.agrformet.2013.09.005.

[13] F. Hosoi and K. Omasa, "Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning LiDAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3610–3618, Nov. 2006, doi: 10.1109/TGRS.2006.881743.

[14] F. Hosoi and K. Omasa, "Factors contributing to accuracy in the estimation of the woody canopy leaf area density profile using 3D portable LiDAR imaging," *J. Exp. Botany*, vol. 58, no. 12, pp. 3463–3473, Sep. 2007, doi: 10.1093/jxb/erm203.

[15] M. Béland, J.-L. Widlowski, R. A. Fournier, J.-F. Côté, and M. M. Verstraete, "Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements," *Agricult. Forest Meteorol.*, vol. 151, no. 9, pp. 1252–1266, Sep. 2011, doi: 10.1016/j.agrformet.2011.05.004.

[16] M. Béland, J.-L. Widlowski, and R. A. Fournier, "A model for deriving voxel-level tree leaf area density estimates from ground-based LiDAR," *Environ. Model. Softw.*, vol. 51, pp. 184–189, Jan. 2014, doi: 10.1016/j.envsoft.2013.09.034.

[17] S. Hancock, K. Anderson, M. Disney, and K. J. Gaston, "Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform LiDAR: Calibration and validation with voxelised terrestrial LiDAR," *Remote Sens. Environ.*, vol. 188, pp. 37–50, Jan. 2017, doi: 10.1016/j.rse.2016.10.041.

[18] D. Kükenbrink, F. D. Schneider, R. Leiterer, M. E. Schaepman, and F. Morsdorf, "Quantification of hidden canopy volume of airborne laser scanning data using a voxel traversal algorithm," *Remote Sens. Environ.*, vol. 194, pp. 424–436, Jun. 2017, doi: 10.1016/j.rse.2016.10.023.

[19] E. S. Douglas *et al.*, "Finding leaves in the forest: The dual-wavelength Echidna LiDAR," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 776–780, Apr. 2015, doi: 10.1109/LGRS.2014.2361812.

[20] F. M. Danson *et al.*, "Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure," *Agricult. Forest Meteorol.*, vol. 198, pp. 7–14, Nov. 2014, doi: 10.1016/j.agrformet.2014.07.007.

[21] A. Burt, M. I. Disney, P. Raumonen, J. Armston, K. Calders, and P. Lewis, "Rapid characterisation of forest structure from TLS and 3D modelling," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul 2013, pp. 3387–3390, doi: 10.1109/IGARSS.2013.6723555.

[22] J.-F. Côté, R. A. Fournier, and R. Egli, "An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR," *Environ. Model. Softw.*, vol. 26, no. 6, pp. 761–777, Jun. 2011, doi: 10.1016/j.envsoft.2010.12.008.

[23] D. L. Jupp, D. S. Culvenor, J. L. Lovell, G. J. Newnham, A. H. Strahler, and C. E. Woodcock, "Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna'," *Tree Physiol.*, vol. 29, no. 2, pp. 171–181, Feb. 2009, doi: 10.1093/treephys/tpn022.

Ian Paynter received the B.Sc. degree in zoology from the University of Wales, Aberystwyth, U.K., in 2008, and the M.S. degree in marine science and the Ph.D. degree in environmental science from the University of Massachusetts Boston, Boston, MA, USA, in 2015 and 2016, respectively.

Since 2017, he has been a Research Scientist with the Universities Space Research Association, Greenbelt, MD, USA.

Daniel Genest received the B.S. degree in environmental science from the University of Massachusetts Boston, Boston, MA, USA, in 2015, where he is currently pursuing the Ph.D. degree. His Ph.D. dissertation involves salt marsh morphology and remote sensing techniques to observe and quantify such morphology.

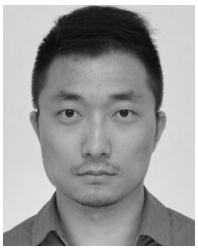
He was involved in the field of terrestrial laser scanning and contributed to four publications in the field of highly portable laser scanning for forested environments. He is currently a Lab Manager with the Coastal Processes and Ecosystems Laboratory (CaPE Lab), University of Massachusetts Boston. The CaPE Lab is a joint research effort between the Provincetown Center for Coastal Studies and the University of Massachusetts Boston, and features a group of natural scientists, social scientists, engineers, and students of all levels.

Edward Saenz received the B.S. and Ph.D. degrees in environmental science from the University of Massachusetts Boston, Boston, MA, USA, in 2012 and 2017, respectively.

He has been a Field Applications Scientist with Spectral Evolution Inc., Lawrence, MA, USA, since 2018. He utilizes various spectroradiometer instruments in addressing needs in the fields of vegetation, geological, and solar simulation. His research interests included applying remote sensing technologies to characterize and analyze a variety of temperate and tropical ecosystems, for which he received the NASA's Harriet G. Jenkins Graduate Fellowship.

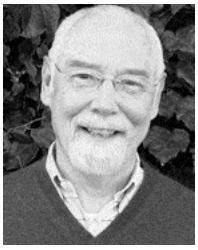
Francesco Peri received the Diploma degree in electrical engineering from the Leonardo Da Vinci Institute, Florence, Italy, and the B.S. degree in computer science and physics and the M.S. degree in computer science from the University of Massachusetts Boston, Boston, MA, USA, in 2000 and 2010, respectively.

He was a Senior Software Engineer with Softrax Company, from 2000 to 2002. He was a Research Associate and a Marine Technical Manager with the Earth Environment and Ocean Sciences Department, University of Massachusetts Boston, from 2002 to 2008. Since 2008, he has been the President and the CEO with Charybdis Group LLC, Boston, MA, USA, a Research Engineer with the University of Massachusetts Boston, and the Managing Director with Coastal Environmental Sensing Networks, Boston, MA, USA. He has authored A Real-Time Sea-Level Monitoring Network for Massachusetts Bay (ASLO Ocean Science Meeting, 2012), Boston Environmental Area Coastal Observation Network (ASLO Ocean Sciences Meeting, 2010), and Machine Reasoning About Anomalous Sensor Data (Journal of Ecological Informatics, ISEI6 special issue). He has engineered customized autonomous winch systems for tow-yow operations on research vessels, and holds patents for an e-coli real-time bacterial detector and a cryogenic laser ablation cell.



Zhan Li received the B.S. degree in cartography and geographic information systems from Nanjing University, Nanjing, China, in 2008, the M.S. degree in geographic information systems from the Chinese Academy of Sciences, Beijing, China, in 2011, and the Ph.D. degree in geography from Boston University, Boston, MA, USA, in 2015.

He is currently a Research Fellow with the University of Massachusetts Boston, Boston, MA, USA. His research interests include optical remote sensing of land surface properties such as forests, urban areas, and agricultural lands, and light detection and ranging remote sensing for forest structure measurements.



Alan Strahler (M'86) received the B.A. and Ph.D. degrees in geography from Johns Hopkins University, Baltimore, MD, USA, in 1964 and 1969, respectively.

He has held prior academic positions at Hunter College, City University of New York, New York, NY, USA, the University of California, Santa Barbara, CA, USA, and the University of Virginia, Charlottesville, VA, USA. He is currently an Emeritus Professor of remote sensing with the Department of Earth and Environment, Boston University, Boston, MA, USA. He was originally trained as a Biogeographer, and he has been actively involved in remote sensing research since 1978. He has been a Principal Investigator on numerous NASA and NSF grants and contracts. His research interests include modeling the bidirectional reflectance distribution function (BRDF) of discontinuous vegetation covers, retrieving physical parameters describing ground scenes through inversion of BRDF models using directional radiance measurements, and quantifying forest structure through terrestrial light detection and ranging scanning.

Dr. Strahler was a fellow of the American Association for the Advancement of Science in 2003. He was awarded the William T. Pecora Award for outstanding contributions to understanding the earth by means of remote sensing from the Department of Interior, NASA, in 2011, and the honorary degree Doctorem Scientarum Honoris Causa from the Université Catholique du Louvain, Belgium, in 2000.

Crystal Schaaf received the B.S. and M.S. degrees in meteorology from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1982, the M.S. degree in archaeology from Harvard University, Cambridge, MA, USA, in 1988, and the Ph.D. degree in geography from Boston University, Boston, MA, USA, in 1994.

She served as an Atmospheric Research Officer with the United States Air Force, Cloud Physics Branch, USAF Geophysics Lab, Dayton, OH, USA, from 1982 to 1986. She was a Research Meteorologist with Satellite Branch, Geophysics Directorate, USAF Phillips Lab, Kirtland AFB, Albuquerque, NM, USA, from 1986 to 1996. She was a Research Assistant Professor with the Center for Remote Sensing, Department of Earth and Environment, Boston University, from 1996 to 2000, and a Research Associate Professor from 2000 to 2008. She was a Science Team Member and a Principal Investigator on Moderate Resolution Imaging Spectroradiometer on board NASA's Aqua and Terra Earth Observing System Platforms from 1999 to 2003. She has been a Science Team Member and a Principal Investigator for Visible/Infrared Imager/Radiometer Suite on board the Suomi National Polarorbiting Parternship and future NOAA/NASA Joint Polar Satellite Systems since 2004, a Science Team Member for Landsat Data Continuity Mission since 2012, and a Collaborator for Global Ecosystem Dynamics Investigation since 2014. She has been a Professor with the School for the Environment, University of Massachusetts Boston, Boston, MA, USA, since 2011. She has authored over 100 articles.

Dr. Schaaf has received the NASA Group Achievement Award: Aqua Mission Team in 2003, the Goddard Space Flight Center Group Achievement Award: Earth Observing System Aqua Mission Team in 2003, and the Editors' Citations for Excellence in Refereeing (JGR-Atmospheres) in 2009.