Use of Mobile Laser Scanning (MLS) to Monitor Vegetation Recovery on Linear Disturbances

Caren E. Jones 1, Angeline Van Dongen 1, Jolan Aubry 2, Stefan G. Schreiber 3 and Dani Degenhardt 1,*

1 Canadian Forest Service, Natural Resources Canada, Northern Forestry Centre, Edmonton, AB T6H 3S5, Canada
2 Wild + Pine, Edmonton, AB T9E 0V4, Canada
3 EnviroStats Solutions Inc., Edmonton, AB T6H 3R9, Canada
* Correspondence: dani.degenhardt@nrcan-rncan.gc.ca; Tel.: +1-587-340-6391

Abstract: Seismic lines are narrow, linear corridors cleared through forests for oil and gas exploration. Their inconsistent recovery has led to Alberta’s forests being highly fragmented, resulting in the need for seismic line restoration programs and subsequent monitoring. Light detection and ranging (LiDAR) is becoming an increasingly popular technology for the fast and accurate measurement of forests. Mobile LiDAR scanners (MLS) are emerging as an alternative to traditional aerial LiDAR due to their increased point cloud density. To determine whether MLS could be effective for collecting vegetation data on seismic lines, we sampled 17 seismic lines using the Emesent Hovermap™ in leaf-on and leaf-off conditions. Processing the LiDAR data was conducted with GreenValley International’s LiDAR 360 software, and data derived from the point clouds were compared to physically measured field data. Overall, the tree detection algorithm was unsuccessful at accurately segmenting the point clouds. Complex vegetation environments on seismic lines, including small conifers with obscured stems or extremely dense and tall shrubs with overlapping canopies, posed a challenge for the software’s capacity to differentiate trees. As a result, tree densities and diameters were overestimated, while tree heights were underestimated. Exploration of alternative algorithms and software is needed if measuring vegetation data on seismic lines with MLS is to be implemented.

Keywords: remote sensing; boreal forest; mobile LiDAR; seismic line; linear disturbance; vegetation recovery

1. Introduction

As a result of natural resource exploration and extraction, Alberta’s boreal forest is extensively fragmented, posing significant challenges for the re-establishment of native species, canopy cover, and ecosystem functions in these areas. The most prominent anthropogenic disturbance type in this landscape are seismic lines—linear access routes through the forest created during oil and gas exploration in order to facilitate equipment access used to identify geologic formations of interest [1,2]. The province of Alberta is estimated to have over 1.8 million km of seismic lines, with densities in some regions reaching 10–40 km/km² [3–5]. Currently there are no regulatory requirements to reclaim legacy or new seismic lines in Alberta. However, their restoration has been identified as a priority in recovering the threatened woodland caribou populations, as these access routes facilitate the mobility and encounter rates of predators (e.g., wolves and bears) with caribou in areas where historically they would not overlap [6–8]. The restoration of seismic lines is not without its challenges, particularly those optimizing the methods used in the prioritization of restoration efforts and the associated costs of intensive versus extensive vegetation monitoring.
An integral first step for any reforestation program is determining the current state of vegetation recovery, thereby enabling the identification and prioritization of areas of concern [9]. Conventional physical sampling of vegetation inherently provides the highest accuracy in quantifying current conditions; however, the cost and time required to execute them can be prohibitive and sometimes impossible due to access and safety concerns. Consequently, remote sensing technologies continue to be explored as alternatives that not only provide an acceptable compromise between vegetation-data quality and survey time and effort but are also able to characterize complex vegetation metrics that cannot be quantified by conventional field measurement techniques [10].

The use of remote sensing to measure forest vegetation is becoming increasingly common as various sensors become more widely available and more easily deployed [11–13]. In particular, light detection and ranging (LiDAR) has emerged prominently in forest assessments due to its ability to operate in multi-layered vegetation environments, often accurately providing georeferenced data along multiple axes, producing 3-dimensional point clouds of scanned vegetation structure over large areas [14,15]. From a forest inventory perspective, LiDAR has proven to be an effective tool for accurately assessing a variety of stand forest features, including measuring individual tree characteristics [16–18], estimating stand density [19], and determining wood fiber attributes [20,21]. In the broader field of forest management, LiDAR has been used to monitor fire fuel loads [22–24], insect-related stand damage and mortality [25–27], and restoration progress [28,29].

Conventionally, LiDAR sensors are mounted onto piloted aircraft (airborne laser scanning (ALS)), and while this method can efficiently collect data from large spatial scales, it is associated with low data resolution (particularly of small vegetation) and high acquisition costs [30–32]. Advancements in LiDAR sensors have enabled them to be mounted onto and deployed via remotely piloted aircraft systems (RPAS, i.e., drones) and terrestrial LiDAR scanning (TLS) units, which have gained popularity due to their relatively low acquisition cost, higher point densities, and effectiveness at detecting mid-canopy structural components compared to ALS [33–35]. The use of RPASs may have regulatory limitations and be restricted to individuals with appropriate licenses, potentially significant barriers to the implementation of these systems [32]. While TLS has similar data acquisition flexibility as RPAS and lacks the regulatory barriers, stationary TLS is limited to short functional ranges and spatial areas and, unless deployed in multiple locations (significantly increasing operation time), the risk of occlusion is high [31,33,36]. To overcome these limitations, backpack and handheld mobile laser scanning (MLS) units have become of particular interest for forest management assessments, as they can quickly collect robust point clouds with minimal occlusion [36–38]. MLS systems often utilize automatic co-registration algorithms (e.g., simultaneous localization and mapping [SLAM]) to orient point clouds and do not require navigation satellite systems to accurately function [39,40].

As a result of their improved data quality and reduced field implementation time, remote sensing technologies are being used more frequently to explore a variety of environmental and site conditions on linear disturbances, including vegetation height, seedling densities, coarse woody debris volume, and terrain characteristics [41–45]. LiDAR scans utilized in linear disturbance restoration studies in the boreal forest have primarily been ALS or RPAS, with little testing of MLS in this field. Due to the re-use of seismic lines and the slow rate of recovering vegetation, seismic lines in Alberta’s boreal forest rarely have dense vegetation above 5 m [43], meaning the low-canopy accuracy of MLS could be well suited for monitoring vegetation recovery on these disturbances accurately and efficiently [35,38–40]. Additionally, software and algorithms necessary for processing of LiDAR point clouds are becoming readily available and user accessible, further incentivising the use of MLS.

In this study, we assessed the effectiveness of using a handheld Emesent Hovermap™ MLS system with GreenValley International’s LiDAR 360 processing software for measuring vegetation on recovering seismic lines by comparing data derived from MLS point clouds to physically measured data on seismic lines in Northwestern Alberta. Furthermore, we
aimed to better understand the ideal conditions for the use of MLS on seismic lines by testing this technology in different seasons and across a range of tree densities. Specifically, the objectives of this study were:

1. To test the accuracy of using MLS to measure the height, DBH, and density of trees on seismic lines.
2. To determine whether the accuracy of data from MLS systems is impacted by leaf-on or leaf-off conditions.

2. Materials and Methods

2.1. Site Selection

The study area is located approximately 70 km northwest of Manning, Alberta, Canada within the lower boreal highlands subregion of the boreal forest (Figure A1). This subregion is characterized by extensive wetlands and diverse mixedwood upland stands, with widespread forestry and oil and gas operations, including seismic lines, throughout [46,47]. The stands selected for this study were characterized as upland forests developed on fine-textured soils, where early successional species, including *Populus tremuloides* Michx. (aspen), *Betula papyrifera* Marshall (paper birch), and *Pinus contorta* Douglas ex Loudon (lodgepole pine), are replaced by *Picea glauca* (L.) Mill (balsam fir) over time [47]. In addition to the unmanaged stands, areas under intensive forest management (cutblocks) were also included in the study. Seismic lines in cutblocks were measured as a representation of lines that have received active restoration (e.g., site preparation and planting), which has not otherwise been performed in this area. Based on available historical imagery [48], seismic lines in the study area were cleared in the mid-1980s. Cutblocks were harvested between 2004 and 2011 and subsequently planted with either *P. contorta* or *P. glauca*. As a result, seismic line recovery ages in this study were between 15 and 40 years, and all seismic lines were approximately 5 m wide.

2.2. Plot Layout & Ground Reference Data Collection

To ensure that a breadth of vegetation densities were sampled and tested (Figure 1), seismic line segments were assigned to one of three predicted recovery categories using aerial imagery based on the methods described in Van Dongen et al. [49]: recovered (R), fractionally recovered (FR), and not recovered (NR). In total, 17 seismic line areas were measured and scanned: nine in cutblocks and eight in unmanaged stands, distributed between four R, seven FR, and six NR lines (Table A1).

![Figure 1. Examples of sampled seismic line areas; the left image shows an area with dense woody vegetation on a cutblock, and the right image shows a sparsely vegetated area in an untreated forest.](image-url)
In July of either 2020 or 2021, a 55 m long transect was established on each seismic line consisting of 12 offset circular plots, each with an area of 10 m² (Figure A2). Within each circular plot, the species, height, and diameter at breast height (DBH) of all trees ≥ 1.3 m tall were measured using telescopic sectional poles and calipers. Counts of all shrubs that were taller than 1.3 m were also recorded, and the tallest individual for each species was measured within each circular plot. Values collected from these plots are referred to as the “ground” condition data.

2.3. LiDAR Data Collection and Processing

A handheld Hovermap™ scanner by Emesent Pty Ltd. (Figure 2) was used for LiDAR data acquisition; the scanner has a weight of 1.8 kg, 480 GB of internal storage, a LiDAR range up to 100 m, position accuracy of 30 mm, a 360° × 360° angular field of view, and either 300,000 or 600,000 points per second data acquisition speed [50]. While holding the Hovermap™ sensor, the perimeter of each seismic line area was walked (approximately 1 m in from the seismic line edge with the adjacent forest) in a closed loop to minimize internal drift and improve navigation [35,51]; on seismic lines with dense vegetation, a zig-zag pattern was walked when needed around breaks in line-of-sight to minimize occlusion. Prior to each scan, a marker was placed at the beginning and end of the scan area to assist with delineation within the point cloud. Data was acquired in leaf-on (July 2021) and leaf-off (October 2021) conditions, although due to weather and access issues, not all lines were scanned during both campaigns. In total, 14 seismic lines were scanned during the leaf-on campaign and 15 during leaf-off (Table A1).

Figure 2. Side and top view of the Emesent Hovermap™ scanning system and carrying case.

Processing of the LiDAR data was conducted using GreenValley International’s LiDAR 360, a graphical user interface (GUI) software with built-in functions to manipulate and process LiDAR point cloud data [52]. Specifically, the software’s TLS Forest module can directly use terrestrial LiDAR data (including MLS) to estimate the number of trees, segment individual trees, and measure a variety of parameters including DBH, height, and crown volume. The software uses a point cloud segmentation method to identify trees based on the comparative shortest-path algorithm (CSP) [52,53]. A benefit of this tree detection method is that it was developed with the challenges of irregular and overlapping canopies in mind [53]. The CSP algorithm consists of three primary steps: point cloud normalization, trunk detection and DBH calculation, and crown segmentation. The normalization step calculates the elevation (height) for each point within the cloud based on its relation to a defined digital elevation or ground layer, which in LiDAR 360 can be calculated through its Terrain module. The trunk detection and DBH calculation step focuses on the automatic detection and measurement of individual stems using a density-based spatial clustering of applications with noise (DBSCAN) approach, where only the minimum cluster size (number of points) needs to be user defined. This is followed by DBH measurements derived from radius calculations from weighted trunk centres of the clusters [54]. The final step, crown
segmentation, aims to distinguish individual tree crowns from one another based on the metabolic ecology theory. The application of this theory for crown segmentation assumes, as a generalization of the method, that any given crown point likely belongs to the trunk with which it has the shortest transportation (i.e., xylem, phloem) distance. To account for irregular tree and crown sizes, a DBH scaling factor is applied to determine a relative transportation distance. Therefore, each point not previously assigned to a stem (in step 2) is assigned a DBH-scaled transporting distance value, and these are used to calculate transportation paths within the point cloud and facilitate clustering. If transportation paths cannot be identified, the point is identified as noise and not included in a tree cluster.

To process the collected data, once a LiDAR file was loaded into LiDAR 360, outliers (defined as any points not within a distance of 3 standard deviations of the average distance between points within the cloud to at least three nearby points) were removed (Figure 3A). The point cloud was then clipped to a 3 × 55 m area based on locating the targets within the point cloud (Figure 3B); although the seismic lines were roughly 5 m wide, they were clipped to 3 m in the point cloud to ensure exclusion of off-line trees as the seismic lines were rarely truly straight (Figure 3C) [55]. Ground point classification was then executed in order to identify the ground layer within the point cloud. This was done by first determining the maximum terrain angle of the scanned area, which was then adjusted by 10° to account for variability and applied with the software’s Classify Ground Points function (Figure 3D).

The point cloud was then reclassified based on the ground points and run through the software’s Seed Point Editor to enable segmentation (Figure 3E), the parameters for which were set to a cluster tolerance of 0.01 m, a minimum cluster size of 5 points, eligible DBH values of interest were constrained to 0.01–1.0 m, the minimum height of eligible vegetation was set to 1.3 m (to only include vegetation that was deemed a tree or tall shrub in the field measurements), and a maximum angle of 30° was used to allow for the inclusion of leaning trees. Using these seed points, point cloud segmentation and batch extraction of DBH was conducted to produce an output table with vegetation characteristics (Figure 3F). While DBH is ideally measured at 1.3 m above the ground, the DBH measurement window was set between 0.3–1.6 m to account for instances where a clear trunk was not visible at 1.3 m (e.g., young spruce).

2.4. Data Analysis

Data analysis and visualization were carried out using the R language and environment for statistical computing (version 4.2.1, Vienna, Austria) [56]. Generalized linear mixed effects models were calculated using the glmmTMB() function of the glmmTMB package [57]. The fixed effects in these models were measurement technique/period, consisting of 3 levels: (1) individually measured shrubs and trees in circular plots along the seismic line which function as a control, (2) LiDAR measured vegetation along the seismic line in leaf-on condition and (3) LiDAR measured vegetation along the seismic line in leaf-off condition. The response variables were height, DBH, as well as stem count and stem density. The stem count and density for the individually measured trees in the circular plots (control) were counted in two versions, the raw and an adjusted version. The raw version includes only trees above 1.3 m, while the adjusted version considered all shrubs that were taller than 1.3 m as trees. This was used to account for the possibility that the LiDAR measurement identified something as a tree, which, in the circular plot control measurements, would have been considered a shrub. The random effect for all models was the seismic line in order to account for spatial dependency among the measurements. While useful for initial characterization, neither the site type (cutblock, forest) nor the classification classes (R, FR, NR) were used as factors in analyses. Height, DBH, and density were modeled using a Gamma distribution, which is a more appropriate choice when data are truncated at zero, positive, and typically right skewed [58]. Stem count on the other hand, was modeled using the negative binomial distribution, which is a good choice for count data that cannot be modeled using the Poisson distribution due to potential over dispersion [58]. All negative binomial models also included an offset to adjust the stem
count for the measured area. To assess the fit of each model, we used the `simulateResiduals()` function from the `DHARMa` package [59]. Wald chi-square tests were conducted to test the fixed effects for their significance via the `Anova()` function of the `car` package [60]. Estimated marginal means were calculated using the `emmeans()` function of the `emmeans` package based on the fitted models [61]. P-values for multiple mean comparisons were adjusted using the Tukey method.

![Visualizing of the data preprocessing in LiDAR 360.](image)

**Figure 3.** Visualizing of the data preprocessing in LiDAR 360. Specifically, (A) shows an entire point cloud of a sampled seismic line area following outlier removal; (B) includes both a photo and a point cloud clip of a target placed on the seismic line area; (C) is a point cloud clipped to the 3 × 55 m area of interest; (D) shows an isolation of the identified ground layer points; (E) is a profile view of the clipped point cloud normalized based on the identified ground points; (F) is a profile view of the same seismic line area as E following point cloud segmentation, with individual trees represented by unique colours relative to their neighbours and noise points in grey.

The accuracy of estimated values (derived from LiDAR point clouds) was assessed using three metrics: root-mean-square error (RMSE), relative RMSE (rRMSE), and relative bias (rBias). To calculate RMSE as well as the rBias, we used the `rmse()` and `pbias()` functions of the `hydroGOF` package [62] and the `rRMSE()` function from the `Fgmutils` package [63]. To measure the association/correlation between the LiDAR collected data and the field
measured data, we used Kendall’s $\tau$ (tau) (rather than Pearson’s correlation coefficient) as the data are right skewed and do not follow a normal distribution. The R function \texttt{cor.test()} from the \texttt{stats} package was used for these tests [56].

3. Results

Tree counts were significantly overestimated when derived from LiDAR data compared to the ground collected values, regardless of leaf condition ($p < 0.01$). This trend continued when count values were scaled-up to stems-per-hectare (sph) density values, with the density being significantly overestimated ($p < 0.01$). When the ground values were adjusted to include all tall shrubs (> 1.3 m), the LiDAR-derived count and density values were still significantly greater than the ground values ($p < 0.01$; Tables 1 and 2). No count or density values were significantly different between the leaf-on and leaf-off sample periods.

Table 1. Estimated means (and standard errors) of the generalized linear model output for the measured count, density, height, and DBH response variables measured. Lowercase letters represent significant differences ($\alpha = 0.05$) following pairwise comparisons. Adjusted values were calculated by adding the counted shrubs ($\geq$ 1.3 m) to the tree values in the ground measurement.

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Sampling Period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ground</td>
</tr>
<tr>
<td>Count</td>
<td>75 (13) b</td>
</tr>
<tr>
<td>Density (stems/hectare)</td>
<td>5262 (893) b</td>
</tr>
<tr>
<td>Adjusted Count</td>
<td>219 (34) b</td>
</tr>
<tr>
<td>Adjusted Density (stems/hectare)</td>
<td>15,417 (2378) b</td>
</tr>
<tr>
<td>Height (m)</td>
<td>4.3 (0.1) a</td>
</tr>
<tr>
<td>DBH (cm)</td>
<td>2.9 (0.1) c</td>
</tr>
</tbody>
</table>

Table 2. Root-mean-square error (RMSE), relative RMSE (rRMSE), relative bias (rBias), Kendall’s $\tau$ correlation values, and $p$ values from pairwise comparisons for the various density measurements derived from LiDAR point clouds collected in leaf-on and leaf-off campaigns compared to ground data. Adjusted values were calculated by adding the counted shrubs ($\geq$ 1.3 m) to the tree values in the ground measurement.

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>LiDAR Period</th>
<th>RMSE (Unit of Variable)</th>
<th>rRMSE (%)</th>
<th>rBias (%)</th>
<th>Kendall’s $\tau$</th>
<th>$p$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Count</strong></td>
<td>Leaf-On</td>
<td>179</td>
<td>60.0</td>
<td>-53.2</td>
<td>0.21</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Leaf-Off</td>
<td>125</td>
<td>39.6</td>
<td>-45.7</td>
<td>0.44 *</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Density (stems/hectare)</strong></td>
<td>Leaf-On</td>
<td>5754</td>
<td>23.2</td>
<td>-14.4</td>
<td>0.41 *</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Leaf-Off</td>
<td>5097</td>
<td>19.3</td>
<td>13.1</td>
<td>0.44 *</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Adjusted Count</strong></td>
<td>Leaf-On</td>
<td>321</td>
<td>39.6</td>
<td>-54.2</td>
<td>0.30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Leaf-Off</td>
<td>484</td>
<td>56.7</td>
<td>-63.2</td>
<td>0.45 *</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Adjusted Density (stems/hectare)</strong></td>
<td>Leaf-On</td>
<td>20,344</td>
<td>30.0</td>
<td>-46.6</td>
<td>0.32</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Leaf-Off</td>
<td>25,408</td>
<td>35.7</td>
<td>-49.5</td>
<td>0.45 *</td>
<td>&lt;0.002</td>
</tr>
</tbody>
</table>

* indicates Kendall’s $\tau$ values that were significant at $\alpha = 0.05$.

Relative to the ground data, the RMSEs for the leaf-on count, density, adjusted count, and adjusted density were 179 trees, 5754 sph, 321 trees, and 20,344 sph, respectively; rRMSE values were 60%, 23%, 40%, and 30%, respectively; and rBias values were all negative, indicating an overestimation of densities (Table 2). The correlation coefficients (Kendall’s $\tau$) between the leaf-on count, density, adjusted count, and adjusted density values to the ground data were 0.21, 0.41 ($p < 0.05$), 0.30, and 0.32, respectively, although only the density relationship was significant (Table 2). The RMSEs for the leaf-off count,
density, adjusted count, and adjusted density were 125 trees, 5097 sph, 484 trees, and 25,408 sph, respectively; rRMSE values were 40%, 19%, 57%, and 36%, respectively; and rBias values were also all negative, indicating an overestimation of densities (Table 2). The correlation coefficients between the leaf-off count, density, adjusted count, and adjusted density values to the ground data were 0.44 \( (p < 0.05) \), 0.44 \( (p < 0.05) \), 0.45 \( (p < 0.05) \), and 0.45 \( (p < 0.05) \), respectively; all relationships were significant (Table 2).

There was a significant effect of sample period on both tree height \( (p < 0.001) \) and tree DBH \( (p < 0.001) \). Field-measured tree heights ranged from 1.3 to 13.7 m, leaf-on from 1.3 to 17.2 m, and leaf-off from 1.3 to 15.6 m, although the majority of LiDAR-derived heights were below 5 m (Figure 4). As a result, the field-measured tree heights were significantly greater than those derived from the leaf-on \( (p < 0.001) \) and leaf-off \( (p < 0.001) \) sampling periods; heights derived from both LiDAR datasets were not significantly different between sampling periods (Table 1). In contrast to the height values, the DBH values derived from the leaf-on \( (p < 0.001) \) and leaf-off \( (p < 0.001) \) scans were significantly greater than those measured in the field (Table 1). DBH values measured in the field ranged from 0.1 to 15.5 cm, leaf-on from 1.2 to 77.7 cm, and leaf-off from 1.1 to 85.9 cm (Figure 4). Additionally, data derived from the leaf-on scans were significantly greater than those from the leaf-off \( (p < 0.001) \).

![Boxplots showing the distribution of tree height and DBH values from the three measurement periods (ground validated, leaf-off LiDAR, leaf-on LiDAR). The total number of trees included for each period is shown along the x-axis.](image)

**Figure 4.**

### 4. Discussion

#### 4.1. Vegetation Density

The overestimated density values derived from the LiDAR point clouds compared to the field validated ground data represents a significant commission difference when using this combination of MLS and LiDAR 360 to identify and count trees in these environments. The presence of a commission difference is in stark contrast to other studies in which omission differences were observed, often as a result of insufficient point returns on narrow trees that resulted in whole stems being absent from the point cloud [36,64]. The accuracy of the Hovermap™ system is stated to be 30 mm by the manufacturer [50], which is similar to the accuracy of other tested systems [36,64], suggesting that the commission difference is unlikely to be hardware based and is likely a result of model failure during data processing.

Specifically, multi-stemmed individual trees and shrubs were identified by the software during point cloud classification as multiple individual trees rather than as a single tree. Over-segmentation has been observed in highly clustered environments when DBSCAN methods are utilized [65], and this is supported by the adjusted ground count and density data (those that included all shrubs \( \geq 1.3 \) m tall in the plots) still being significantly
lower than the counts and densities produced from the point cloud. A single *Alnus viridis* shrub can have dozens of stems, and although they all originate from a single base, these stems could be spaced apart enough at 1.3 m to be considered individual trees by the CSP algorithm (as per the parameters set for Point Cloud Segmentation in this study). For example, in Figure 5, a single *A. viridis* individual has been identified as at least five individual trees. In the ground dataset, this one individual would not have been included in the tree count and density values and would only have been represented by a single shrub in the adjusted values (which includes both shrubs and trees). As *A. viridis* formed dense thickets on some of the lines measured (Figure 1), having each shrub stem identified as an individual tree likely accounts for the greater density values derived from the point clouds than those measured in the field. The development of algorithms to differentiate multi-stemmed individuals as a single tree, as well as to identify and remove tall, multi-stemmed shrubs within the point clouds, is needed to improve accuracy in complex environments like those found in this study.

![Figure 5](image_url)

**Figure 5.** A tall (> 1.3 m), multi-stemmed *Alnus viridis* individual; the left is a photo showing its base and stems on a seismic line measured this study, and the right is a clip of a similarly sized *A. viridis* plant from LiDAR 360 following point cloud segmentation where each colour indicates an individual “tree” stem identified by the CSP algorithm.

### 4.2. Tree Height

A limitation of early MLS systems in monitoring forest structure was their relatively short scan ranges [33,66]. While this could be compensated for in the horizontal plane, it limited the types of trees that could be sampled to those shorter than the scan range, as taller trees could not be entirely measured. The scanner used in this study, the Emesent Hovermap™, is reported to have a LiDAR range up to 100 m [50], well above the maximum height measured in the ground-validated plots for this study (14 m). Despite its increased scan range, tree height was severely underestimated when derived from the MLS point clouds.

While the data points for the treetops were present in the point cloud, tree segmentation in LiDAR 360 was unable to appropriately distinguish individual trees once canopies or branches began overlapping, with many trees only being identified up to the base of their crowns. As a result, the measured heights of the trees from the point clouds were much shorter than they should have been had their canopy been distinguished and included. This is demonstrated in Figure 6, in which stems assigned to a tree following point cloud segmentation are shown in colour, whereas aspects of the point cloud that were not assigned to a tree remain grey. In areas where vegetation was sparse, entire trees were successfully identified (Figure 6A), unlike in areas with dense vegetation (Figure 6B). Although similar CSP algorithms have successfully delineated trees with overlapping canopies in LiDAR
point clouds [53,65,67], the crown conditions present on most of the seismic lines found in this study appear to have been too complex to be accurately separated by LiDAR 360’s CSP-based tree detection algorithm. The algorithm was able to identify individual trees in the canopy (Figure 6C), albeit very rarely.

Figure 6. Examples of point clouds shown in the LiDAR 360 software following point cloud segmentation; grey represents points that were unassigned to a tree, whereas coloured areas denote points assigned to an individual tree in a poorly vegetated (A) and densely vegetated (B) seismic line; (C) shows a tree where a portion of the canopy was successfully identified following segmentation.

The complex nature of natural stands makes single tree detection in point clouds challenging; utilizing a simpler canopy detection method that does not explicitly attempt to distinguish overlapping canopies may be an alternative solution for tree segmentation in these types of environments. An example of such a method that has been successful in mature mixedwood forests is the distance judgment clustering approach [68,69]. This method assumes that the distance between trees is greater at their apex compared to the ground and uses those distances to cluster and segment trees. This technique could improve accuracy in low-density or conifer-dominated seismic lines (Figure 6A) where tree crowns are clearly distinguishable, but in high-density areas with uniform crown heights (Figure 6B) or deciduous stands, this method may struggle to identify tree “tops”. As this method is based on height relationship assumptions and not clearly distinguishable branch relationships and pathways, there is a risk that an assigned canopy may either not be truly associated with its assigned trunk or may not be the trunk’s true height.

Another commonly used tree segmentation approach is the cubical voxel method [70–72]. This is a type of connected component labeling that utilizes octree segmentation to detect connected areas within a point cloud. This is a hierarchal approach, where a central node voxel is identified and repetitively subdivided, with empty voxels being removed from the process until a minimum threshold (e.g., pixel size, levels of divisions) is achieved. As this method does take into consideration potential branching patterns of tree canopies, it may better relate canopies to their stem. However, in environments such as those found in this study, this segmentation process may also be unable to detect patterns in the cluttered and overlapping canopies.

The trunk growth (TG) algorithm segmentation method also takes branching patterns of canopies into consideration [65]. This method has four primary steps: trunk extraction, fragment merging, trunk growth, and localization of branch and leaf points. While it has a similar workflow to the CSP algorithm, it does not utilize DBSCAN for trunk detection or a DBH adjustment to determine branch and stem relationships. A study comparing the two segmentation methods in a vertically complex natural forest observed similar detection results, although the TG overall was more accurate compared to the CSP operation [65]. While the TG method is of interest for the processing of LiDAR data from complex environments, it is relatively new and has yet to be adopted as a pre-programmed segmentation method in LiDAR processing software, limiting its ability to be widely tested.
and implemented in analyses. Further exploration of tree segmentation methods is critical if LiDAR is to be used for stand characterization in densely vegetated seismic lines.

4.3. Tree DBH

One of the benefits of MLS relative to RPAS LiDAR is the increased point cloud density in the lower canopy which facilitates DBH measurements [39,69]. A number of studies have observed acceptable ranges of DBH accuracy when sampling trees in both planted and natural stands; however, the general consensus is that accuracy is highest for trees with DBHs between 10 and 20 cm [40,66,73]. The minimum measured DBH in the field for this study was 0.1 cm, and the average was 2.9 cm, well below the recommended DBH limits. The accuracy of the Hovermap™ hardware itself is only 3 cm [50], far larger than many of the tree diameters sampled in this study. While recent work with MLS has achieved more accurate tree detection and even DBH measurements at 4–6 cm [39,74,75], these require ideal environmental conditions and complex preprocessing to reduce noise in the point clouds.

The DBH values from the LiDAR data were highly skewed and much exceeded the maximum DBH value measured in the field. The stems of young conifer trees were often obscured by branches in the point cloud (Figure 7A), and the related DBH estimation included those branches, resulting in a large overestimation. In a study where trees with DBHs as small as 6 cm were accurately measured from an MLS point cloud, there was no branching or obstruction in the forest below 2 m to facilitate clear lines-of-sight between the sensor and tree [75]. While this may be achievable in highly managed environments, this is not a feasible condition to expect in most natural forests.

![Figure 7. Close-ups of individual trees identified within LiDAR 360 following point cloud segmentation. Specifically, (A) illustrates a section where two coniferous trees were identified as a single tree (orange), a single coniferous tree (dark blue) and a single deciduous tree (light blue) were identified appropriately, and a tall shrub (cyan) was identified as a tree; while (B) shows an area with multiple deciduous trees identified.](image)

While the challenge with overestimation of DBH as a result of conifer branches is a concern on recovering seismic lines, many of the sites measured in this study were dominated by deciduous trees and shrubs. In these sites, stems were clearly visible and point cloud segmentation to identify individual “trees” (including tall shrubs) was feasible within the DBH measurement zone (0.3–1.6 m) (Figure 7B). Despite these unobstructed conditions, DBH measurements were still highly overestimated for these trees, likely a result of the combination of high stem densities and relatively small trees. For small trees that are detected, the proportion of valid reference points relative to the surrounding noise points is lower compared to large trees (i.e., > 10 cm DBH), resulting in overestimated DBHs and larger error values [39]. An integral part of any MLS hardware is the automatic
co-registration of points, which in the Hovermap™ is through SLAM algorithms. In areas with high stem densities, the co-registration accuracy is reduced as a result of increased slight offsets, creating “fuzzier” point clouds than in lower density areas [40,75].

Accurate detection of stems > 6 cm in diameter has occurred in studies with stand densities around 2000 sph [75,76], which is less than half the tree density of the sites sampled for this study (not including shrubs). In Alberta, vegetation recovery standards for forested uplands is a minimum stand density of 2000 sph and vegetation height of 3 m [55,77]. As such, some seismic lines included in this study were far denser than the normal conditions, particularly in the cutblocks as a result of the reforestation work (i.e., site preparation and planting). Even though the lines sampled in this study had densities above the minimum recovery standard, it is important to note that accurate measurements of vegetation from LiDAR begins to deteriorate at 2000 sph, when all other conditions are ideal. As observed in this study, seismic line recovery is highly variable, and lines rarely have ideal sampling conditions. Some seismic lines can have high densities of trees and shrubs, significantly limiting the applicability of current MLS hardware and software to accurately assess these environments.

4.4. Leaf Condition

Until appropriate tree detection algorithms can be identified, disseminating the differences between sampling periods and subsequent conclusions regarding the ideal conditions for MLS systems cannot be determined. Anecdotally, the majority of density values derived from the leaf-on period were not significantly correlated to the ground data, whereas all of the leaf-off data were correlated. Although the density values were overestimated regardless of leaf condition, the leaf-off data more accurately captured the density patterns present. In complex canopies where the CSP algorithm is unable to segment trees above the stem, it stands to reason that this issue would be compounded by the presence of leaves, as they would further complicate the point cloud environment. The presence of leaves has been found to increase point cloud “fuzziness” from MLS systems as a result of increased co-registration errors associated with moving vegetation [75,78].

4.5. Field Challenges

There were a number of challenges associated with implementing the MLS scans and data processing in the environments selected for this study. High stem densities on some seismic lines, particularly in the cutblocks where A. viridis could exceed 23,000 sph, proved challenging. Unlike TLS, which can be placed in a safe stationary location where the hardware will not be damaged by nearby vegetation, there is a much higher risk of damaging the hardware while walking through dense vegetation. The Hovermap™ is light, easy to maneuver, and can compensate for the movement of the operator; however, in densely vegetated areas, preventing the equipment from coming into contact with nearby vegetation was nearly impossible (Figure 1). In areas where dense vegetation proves to be a risk to the hardware and TLS occlusion risk is high, ALS or RPAS may be the preferred LiDAR scan options.

5. Conclusions

Data acquisition using the MLS system was reasonably easy to implement on recovering seismic lines, where conditions allowed. However, analysis of the collected data in the LiDAR 360 software using the CSP detection algorithm was unsuccessful at accurately describing the density or characteristics of trees on the recovering seismic lines, regardless of leaf condition. Complex vegetation environments on seismic lines, including small conifers with obscured stems or extremely dense and tall shrubs with overlapping canopies posed a challenge for the software’s capacity to differentiate trees. Exploration of alternative tree detection algorithms and software, as well as functions to differentiate tall shrubs from young trees, are needed if measuring vegetation data on seismic lines with MLS is to
be widely implemented; until then, field-based measurements will likely continue to be the standard.

**Author Contributions:** C.E.J. collected field data, processed and analyzed LiDAR data, and lead the writing of the manuscript; A.V.D. collected field data and contributed to manuscript; J.A. collected and processed LiDAR data; S.G.S. analyzed data and contributed to manuscript; D.D. conceptualized the research, acquired funding, and contributed to manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was supported by the Cumulative Effects Program Funding from the Canadian Forest Service.

**Data Availability Statement:** Data underpinning the work are not available in the public domain. Please email the corresponding author for data access.

**Acknowledgments:** The forest in which this research was conducted is a part of the traditional lands of many Indigenous Peoples, including the Dene and Métis. Office work for this project was conducted in Edmonton, traditional lands of the Cree, Dene, Stoney, Saulteaux, Blackfoot, and Métis Peoples. We would like to thank Amanda Schoonmaker and their staff at NAIT’s Centre for Boreal Research for support in collecting the ground condition data, as well as Keith Wells and Steve Blanton from Manning Diversified Forest Products for providing harvest and reforestation data for the study area. We would also like to thank Candrone for providing access to the LiDAR 360 software.

**Conflicts of Interest:** The authors have no conflicts of interest to declare.

**Appendix A**

**Table A1.** Stand type, regeneration classification, and location of all seismic line areas measured, including whether MLS scans were completed in leaf-on (July) and/or leaf-off (October) conditions.

<table>
<thead>
<tr>
<th>Stand Type</th>
<th>Classification Direction</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Leaf-On Scan</th>
<th>Leaf-Off Scan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest FR</td>
<td>NE-SW</td>
<td>−118.287</td>
<td>57.226</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forest FR</td>
<td>E-W</td>
<td>−118.429</td>
<td>57.400</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Forest FR</td>
<td>E-W</td>
<td>−118.384</td>
<td>57.387</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Forest FR</td>
<td>NE-SW</td>
<td>−118.254</td>
<td>57.219</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Forest NR</td>
<td>NE-SW</td>
<td>−118.399</td>
<td>57.283</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Forest NR</td>
<td>NE-SW</td>
<td>−118.292</td>
<td>57.224</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Forest NR</td>
<td>E-W</td>
<td>−118.347</td>
<td>57.387</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Forest NR</td>
<td>NE-SW</td>
<td>−118.289</td>
<td>57.227</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Cutblock R</td>
<td>NE-SW</td>
<td>−118.290</td>
<td>57.226</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Cutblock R</td>
<td>E-W</td>
<td>−118.355</td>
<td>57.387</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Cutblock R</td>
<td>E-W</td>
<td>−118.372</td>
<td>57.385</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Cutblock R</td>
<td>N-S</td>
<td>−118.426</td>
<td>57.398</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cutblock FR</td>
<td>SE-NW</td>
<td>−118.393</td>
<td>57.275</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cutblock FR</td>
<td>NE-SW</td>
<td>−118.298</td>
<td>57.218</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cutblock FR</td>
<td>N-S</td>
<td>−118.379</td>
<td>57.383</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cutblock NR</td>
<td>E-W</td>
<td>−118.227</td>
<td>57.235</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cutblock NR</td>
<td>E-W</td>
<td>−118.221</td>
<td>57.235</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Figure A1. Location of all seismic line areas sampled. Stand type is indicated as either a blue diamond (unmanaged forest) or an orange circle (Cutblock).

Figure A2. Layout of the circular sampling plots used to measure woody vegetation on seismic lines. LiDAR data were collected by walking the perimeter of the 55 m long area.

References
5. Filicetti, A.; Cody, M.; Nielsen, S. Caribou Conservation: Restoring Trees on Seismic Lines in Alberta, Canada. Forests 2019, 10, 185. [CrossRef]


56. R Core Team. *R: A Language and Environment for Statistical Computing*, version 4.2.1; R Core Team: Vienna, Austria, 2022.


69. Gao, S.; Zhang, Z.; Cao, L. Individual Tree Structural Parameter Extraction and Volume Table Creation Based on Near-Field LiDAR Data: A Case Study in a Subtropical Planted Forest. *Sensors* 2021, 21, 8162. [CrossRef] [PubMed]


