Simulated full-waveform LiDAR compared to Riegl VZ-400 terrestrial laser scans

Angela M. Kim[a], Richard C. Olsen[a], and Martin Béland[b]

[a]Naval Postgraduate School, Remote Sensing Center and Physics Department, , 833 Dyer Road, Monterey, CA, USA;
[b]Laval University, Department of Geomatics Sciences, , 2325 rue de l'Université, Québec, Québec, Canada

ABSTRACT

A 3-D Monte Carlo ray-tracing simulation of LiDAR propagation models the reflection, transmission and absorption interactions of laser energy with materials in a simulated scene. In this presentation, a model scene consisting of a single Victorian Boxwood (*Pittosporum undulatum*) tree is generated by the high-fidelity tree voxel model VoxLAD using high-spatial resolution point cloud data from a Riegl VZ-400 terrestrial laser scanner. The VoxLAD model uses terrestrial LiDAR scanner data to determine Leaf Area Density (LAD) measurements for small volume voxels (20 cm sides) of a single tree canopy. VoxLAD is also used in a non-traditional fashion in this case to generate a voxel model of wood density. Information from the VoxLAD model is used within the LiDAR simulation to determine the probability of LiDAR energy interacting with materials at a given voxel location. The LiDAR simulation is defined to replicate the scanning arrangement of the Riegl VZ-400; the resulting simulated full-waveform LiDAR signals compare favorably to those obtained with the Riegl VZ-400 terrestrial laser scanner.

Keywords: LiDAR, Monte Carlo, ray tracing, VoxLAD, simulation, small footprint, full-waveform

1. INTRODUCTION

LiDAR is an active remote sensing technique in which pulses of laser energy are transmitted from a sensor, and the precise amount of time required for laser energy to be reflected from the scene back to the sensor is recorded. The time-of-flight of the reflected pulses (and exact location in space of the sensor and pointing direction of the laser) is used to build up a 3D point cloud representation of the scene. A full-waveform LiDAR system records the amount of laser energy being returned to the sensor over time for each transmitted pulse. Processing the digitized waveform signal enables extraction of individual returns corresponding to peaks in the waveform data as in a traditional LiDAR point cloud, and additional information based on the shape of the returned signal.

Differences between the emitted and returned laser pulse shapes give some indication of the material the pulse interacted with on its path from and back to the sensor. Attributes derived from the returned pulse shape have been shown to be useful for improving point cloud classification;[1] however, these attributes are not widely available, or used in cases where the information is available. A simulation of full-waveform LiDAR will aid in the understanding and development of algorithms for processing full-waveform signals for extracting information about a scene. In this work, we present results of an ongoing project to simulate full-waveform LiDAR signals.

The goal of this work is to validate the LiDAR simulation by comparing real waveforms as recorded by the Riegl VZ-400 to simulated signals generated by the LiDAR simulation. To achieve this, a Victorian Boxwood tree was scanned with a Riegl VZ-400. The point cloud data from the Riegl scans were used as input to the VoxLAD model, which returns a 3-D voxel representation of the tree, where each voxel contains information about the density of material at each location. The 3-D voxel representation of the tree was combined with a voxel ground surface, and used within the LiDAR simulation as the scene model. The LiDAR simulation replicates the real-world collection geometry of the Riegl VZ-400 data collect. Real waveforms from the Riegl VZ-400 acquisition and simulated waveforms are compared in Section 8. Several complicating factors need to be addressed in order to truly validate the simulation results, but the preliminary results are promising.

Further author information:
A.M.K.: E-mail: amkim@nps.edu, Telephone: 1 401 647 3536

Laser Radar Technology and Applications XXI, edited by Monte D. Turner,
Gary W. Kamerman, Proc. of SPIE Vol. 9832, 98320T · © 2016 SPIE
CCC code: 0277-786X/16/$18 · doi: 10.1117/12.2223929

2. BACKGROUND

Previous results and discussions of the development of the LiDAR simulation are given in Kim 2015.[2] Recent advances include the ability to simulate the scene in 3-D, including the option to locate the sensor to the side of the scene as in a terrestrial LiDAR scanning arrangement. We make use of this new utility here to simulate the scanning arrangement of a Riegl VZ-400 TLS.

The VoxLAD model, developed by M. Béland, uses terrestrial laser scanning data to develop a 3-D voxel representation of the leaf area density (LAD) of a tree canopy.[3–5] The VoxLAD model emulates the contact frequency approach of using a physical probe to measure the amount of material in a given volume by counting the number of leaves contacted by a thin probe inserted into a canopy. In this case, the laser pulse from the Riegl VZ-400 terrestrial laser scanner is the "probe", and the number of returns from each voxel location is related to the density of material at each given location. The VoxLAD model accounts for the occlusion of interior voxels by materials at the surface of the canopy, and returns an estimate of material density for each voxel location.

3. APPROACH

A Riegl VZ-400 TLS was used to acquire a high-density point cloud dataset of a Victorian Boxwood (*Pittosporum undulatum*) tree. Details about the Victorian Boxwood tree are given in Section 4. The Riegl VZ-400 is capable of optionally recording waveforms, and in this case, waveforms were digitized for a portion of the data collection. See Section 5 for a description of the Riegl VZ-400 scanning process. The high-density Riegl VZ-400 point cloud dataset was then classified (by means of an apparent reflectance threshold) into leaf, wood, and noise classes. The classified Riegl VZ-400 point cloud was used within the VoxLAD model to create a 3-D voxel representation of the Victorian Boxwood tree with a measure of leaf and bark density for each voxel (Section 6). A flat ground surface was combined with the voxel model of the tree to complete the simulated scene. Finally, parameters for the LiDAR simulation were defined to replicate the real-world scanning arrangement used when scanning the Victorian Boxwood canopy with the Riegl VZ-400 TLS (Section 7). The resulting simulated waveforms are compared to the real waveforms as recorded by the Riegl VZ-400 in Section 8.

4. VICTORIAN BOXWOOD

The Victorian Boxwood used in this work is approximately 8.6 m tall. Images of the tree are shown in Figure 1, and close-up views of the leaves in Figure 2. Data were collected on 14 October 2015, and the leaf canopy is somewhat sparse. The leaves are oblong shaped, mostly flat, and somewhat shiny.

Figure 1: Two views of the Victorian Boxwood *(Pittosporum undulatum)*.

Figure 2: Victorian Boxwood *(Pittosporum undulatum)* leaves.

5. RIEGL VZ-400 SCANS

The Riegl VZ-400 is a terrestrial laser scanning system that records waveform signals by echo digitization. A laser pulse is emitted from the sensor, and the reflected optical signal is converted to an electronic signal, amplified and sampled at a rate of 2 GHz. The digitized waveform signals enable extraction of the point cloud data, and also attributes including amplitude, pulse shape deviation, and FWHM of the returned pulse.[6]

The Riegl VZ-400 has the ability to convert recorded pulse intensity values to a measure of relative reflectivity that is independent of the range of the object from the sensor. The relative reflectivity measurement relies on assumptions of flat targets being perpendicular to the laser beam, the whole laser beam hitting the target, and the target being the first target the laser beam interacts with.[6] Although these assumptions are obviously not satisfied for many of the returns from a tree canopy, the relative reflectivity measure is still useful for distinguishing materials. We make use of the relative reflectivity measurement to classify the resulting Riegl VZ-400 point cloud data into leaf, wood, and noise classes.

Three high-density scans of the Victorian Boxwood tree were acquired on 14 October 2015. The Riegl VZ-400 was mounted on a tripod, and the three scans were evenly spaced at approximately 120° intervals around the tree canopy, and at a distance of approximately 20 meters (Figure 3). The angular resolution of the scanner (the minimum possible angular distance between two consecutive laser measurements[6]) was set at 0.03 degrees, resulting in a distance between points of approximately 1 cm at a distance of 20 m. The beam divergence of the Riegl VZ-400 is 0.3 mrad,[6] resulting in a spot size of approximately 0.6 cm at a distance of 20 m. The Riegl VZ-400 operates at a wavelength of 1550 nm, and each emitted pulse is approximately 3 ns long.[6] Figure 3 shows an overhead view of the Riegl VZ-400 point cloud with the locations of the Riegl VZ-400 for each of 3 scans as indicated, and the resulting point cloud of the Victorian Boxwood tree trunk and canopy.

Due to the extra time required to scan and record waveform data, waveforms were only acquired at the "Scan 2" location (Figure 3). A lower resolution scan pattern was used to collect waveforms, with the point spacing between samples at 20 m being approximately 2.5 cm.

Riegl provides dynamic link libraries (DLLs) to facilitate accessing the raw waveform data as acquired by the sensor. Waveforms are identified within the raw waveform files (.wfm format) by means of a timestamp, and each waveform has an associated origin and direction vector defining the laser pointing direction at the time of pulse emission.[7] This information enables exact replication of the scanning arrangement within the LiDAR simulation.

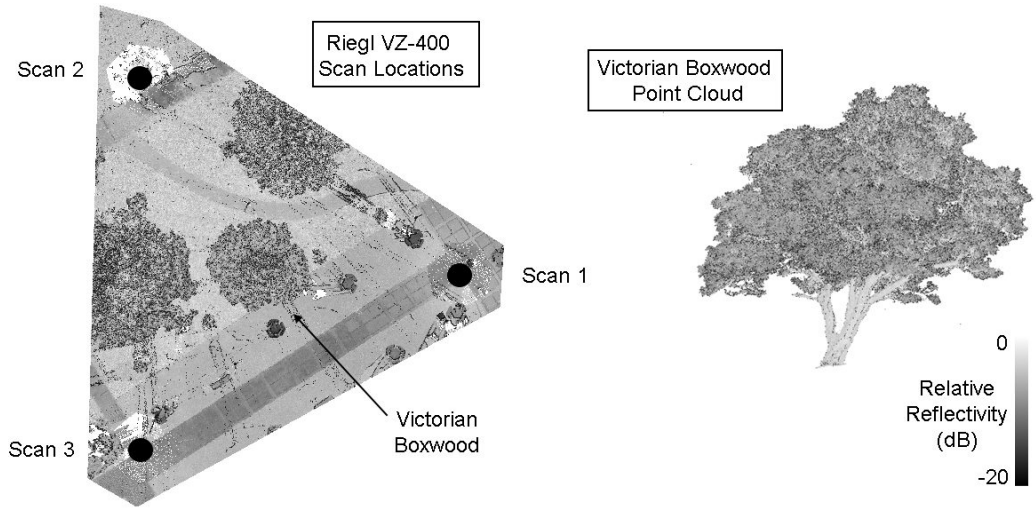

Figure 3: (Left) Overhead view of Riegl VZ-400 point cloud showing scan locations used to acquire a high density point cloud of the Victorian Boxwood canopy. Each of the scans were acquired from a distance of approximately 20 m from the tree trunk. (Right) Point cloud of the Victorian Boxwood tree, with points colored in gray scale according to relative reflectivity.

6. VOXLAD MODEL FOR DERIVING LEAF AREA DENSITY FROM TERRESTRIAL LIDAR DATA

The high-density Riegl VZ-400 TLS point cloud data of the Victorian Boxwood were used as input to the VoxLAD model. For details about the operation of the VoxLAD model, the reader is referred to Béland 2011, Béland 2014 and Béland 2014.[3–5]

The VoxLAD model takes scans from a terrestrial laser scanner as input, where each of the points has been classified as a wood, leaf, or noise return. The output of the VoxLAD model includes estimates of Leaf Area Density (LAD) for individual 3D voxels covering the spatial extent of the input point clouds. While the VoxLAD model is designed to measure LAD, it was used in a non-traditional fashion in this case to also create the voxels defining the probability of the existence of wood material. This was accomplished simply by switching the classification of wood and leaf in the point clouds input to the VoxLAD model. The principle of treating the laser pulses as equivalent to a physical probe, and the additional modifications to account for occlusion of the interior voxels, is assumed to be valid for wood materials to the extent required for this work. For the wood materials, the returned values from the VoxLAD model were treated as probability values indicating the likelihood of the presence of wood material at each voxel.

The VoxLAD model requires a number of inputs and parameters to be defined, including classification of the input point clouds into leaf, wood, and noise classes; definition of the voxel size; and definition of correction factors based on the chosen voxel size. In this case, the classification of the point cloud data into leaf, wood, and noise classes was accomplished by means of a threshold applied to the apparent reflectivity measurements obtained by the Riegl VZ-400. The histograms of apparent reflectance values for each of the three scans are shown in Figure 4, along with the corresponding threshold values chosen to separate the point cloud into noise, wood, and leaf classes. Further manual "cleaning" was used to classify points that were obviously misclassified using the simple threshold method; for example, some points along the edge of the tree trunk were classified as leaf returns. These are most likely partial returns from the edges of the trunk, where only a portion of the laser pulse hit the trunk, and so the amplitude of the returned pulse is less than would be expected from a return where the entire laser pulse beam was reflected from wood material. These points were manually reclassified as noise. The VoxLAD model also requires the user to define the voxel size and the related multiplicative correction factor; a discussion of the choice of optimal voxel size and correction factors is given in Béland 2014.[4] The optimal choice of voxel size depends upon the average size of the leaf, the clumping structure of the leaves and

branches, and the likelihood of occlusion due to the scanning geometry used to collect the point cloud data. The choice of voxel size must also satisfy the assumption within the VoxLAD model that leaves are randomly distributed within the voxel volume. In this case, a voxel side-length of 20 cm was chosen, and the returned LAD values were multiplied by a correction factor of 0.95. More work is required to verify and validate the choices of voxel size and correction factors for this particular Victorian Boxwood tree. The resulting voxel models of leaf and woody material are given in Figure 5.

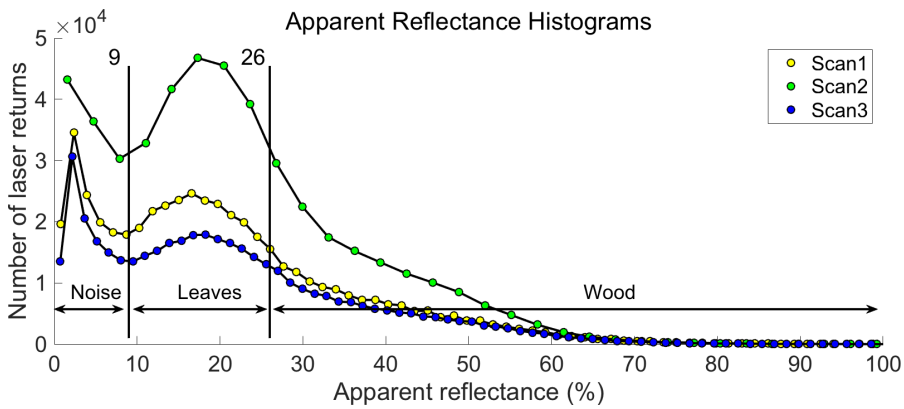

Figure 4: Apparent reflectance histograms for each of the three Riegl VZ-400 TLS scans of the Victorian Boxwood. The thresholds chosen to classify the points as noise, leaves, or wood are indicated by the vertical lines.

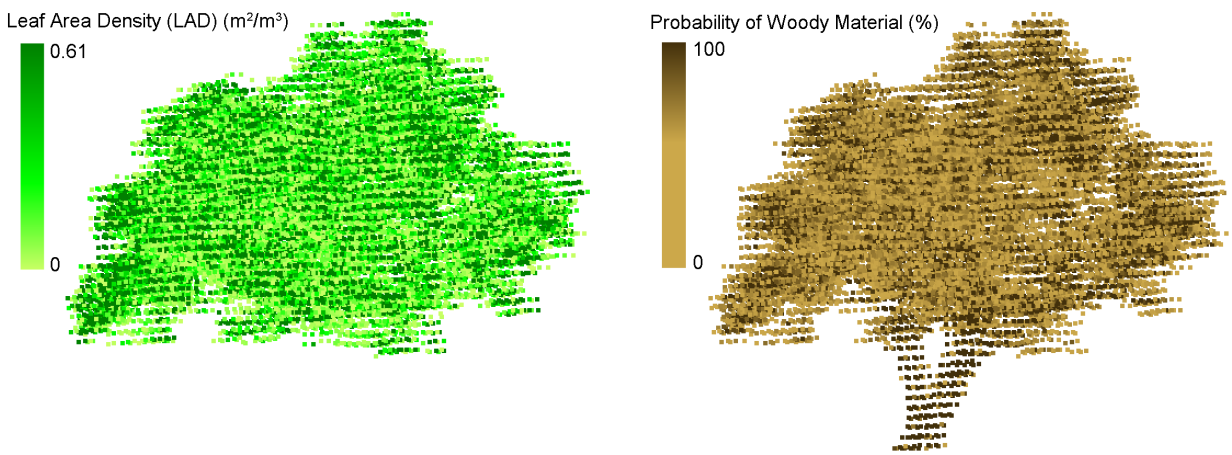

Figure 5: Voxel model of tree with voxels colored according to (L) Leaf Area Density and (R) probability of existence of woody material. These voxel models are the output of the VoxLAD model, based on terrestrial laser scans of a Victorian Boxwood tree. The output of the VoxLAD model for the woody material was converted from density values to a range from 0-100%. Note: For clarity, the voxels are displayed here as points; in reality, the voxel map completely covers the spatial extent of the tree with no gaps.

7. LIDAR SIMULATION

The LiDAR simulation includes the model scene composed of the leaf, wood, and ground voxel maps; functions to define the pulse of laser energy and the interaction of energy with each of the materials in the scene; and parameters to define the geometry of the sensing arrangement. The voxel-based estimates of material density from the VoxLAD model are used within the LiDAR simulation to define the model scene. Creation of the leaf and

wood voxel maps using the high-density Riegl VZ-400 point cloud data as input to the VoxLAD model is described in Section 6. A flat ground surface was added to the scene, with the ground modeled as having Lambertian scattering properties. The leaf and wood materials are also modeled as Lambertian scattering surfaces. The reflectance properties of the materials in the scene were defined as follows: the leaves were defined as having a reflectivity, transmittance, and absorbance of 75%, 20%, and 5% respectively; the bark was likewise defined with 33% reflectivity, 67% absorbance, and no transmittance; the ground was modeled as having reflectivity of 85%, absorbance of 15% and no transmittance. These properties can be modified as needed to simulate the reflectance behaviors of materials at any given wavelength.

The simulated LiDAR sensor transmits a Gaussian-shaped 3 ns long pulse, and a beam spread of 0.3 mrad. Two simulations were run – in the first case (Figure 6), the simulation replicates a situation where the laser was directed at the tree trunk, while in the second case (Figure 7), a more complex situation is simulated with the laser being directed into the tree canopy. In both cases, the location and arrangement of the simulated scanner replicates the scanning geometry used for the Riegl VZ-400 collect of the Victorian Boxwood tree. Figures 6 and 7 show the simulated scanner and tree voxel model within the Riegl VZ-400 point cloud, demonstrating the alignment between the simulated and real-world data.

The LiDAR simulation operates by probabilistically determining the likelihood of interactions of laser energy at each voxel, and the distribution of energy should an interaction occur. When an interaction with leaves is simulated, the surface normal of the leaf is determined according to a function of the location of the leaf within the tree canopy. The azimuthal component of the leaf surface normal is chosen so that the leaves face away from the center of the canopy. The zenith component (defining the degree to which the leaf is horizontal or vertical) is chosen according to a leaf angle distribution function, where the height of the leaf above ground level determines the most likely orientation of the leaf. This same characterization of leaf angle distribution is used within the VoxLAD model. In this case, a uniform distribution was used. The probability of energy interacting with leaf materials in the simulated scene is a function of both the LAD and the surface normal of the leaf – a leaf that is "edge-on" to the scanner will be much less likely to be interacted with than a leaf which is normal to the incident pulse. The parameters defining the leaf surface normal also have a significant effect on the values of LAD returned by the VoxLAD model; further work is needed to determine the most appropriate distribution to be used in this case. See Section 9 for further discussion.

The resulting simulated LiDAR waveforms for the trunk and canopy simulations are shown in Figures 9 and 11; comparisons to real waveforms from the Riegl VZ-400 TLS are made in Section 8.

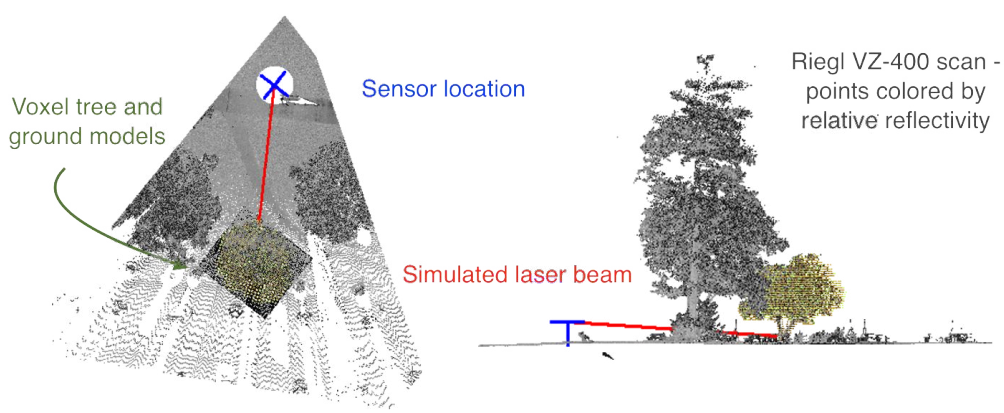

Figure 6: LiDAR simulation arrangement with laser directed at the tree trunk. (Left) Top view and (Right) side view of the voxel model of tree (in green and brown) within the Riegl VZ-400 point cloud (with points colored in gray-scale according to relative reflectivity). The location of the simulated sensor is shown in blue, and the simulated laser beam is shown in red. The location of the sensor and direction of the laser beam replicate the real-world scanning arrangement of the Riegl VZ-400 data collect.

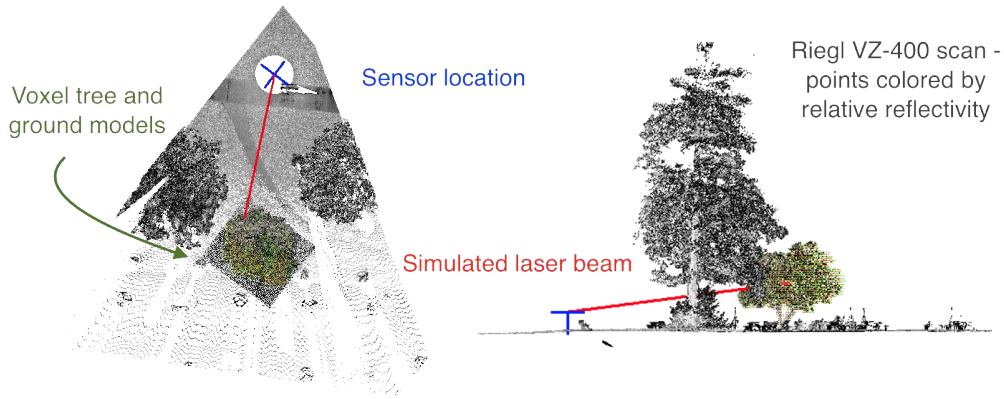

Figure 7: LiDAR simulation arrangement with laser directed at the tree canopy. (Left) Top view and (Right) side view of the voxel model of tree (in green and brown) within the Riegl VZ-400 point cloud (with points colored in gray-scale according to relative reflectivity). The location of the simulated sensor is shown in blue, and the simulated laser beam is shown in red. The location of the sensor and direction of the laser beam replicate the real-world scanning arrangement of the Riegl VZ-400 data collect.

8. COMPARISON TO RIEGL VZ-400 WAVEFORMS

Each of the LiDAR simulations replicates the scanning arrangement for a single specific Riegl VZ-400 waveform. The spatial resolution of the LiDAR simulation, however, is only as high as the voxel model of the scene; in this case, voxels have side lengths of 20 cm. To compare the simulated waveforms to the real waveforms collected by the Riegl VZ-400, we first collected a subset of the Riegl waveforms from an area approximately equal to the voxel size of the simulation. Using the laser pointing direction stored in the raw waveform files from the Riegl VZ-400, a series of points were projected to a uniform distance of 20 m. A subset of these projected points was defined representing a 20 x 20 cm^2 region, centered around the waveform that was used to define the LiDAR simulation parameters. The timestamps of the subset points were then used to extract the associated Riegl waveforms. The Riegl VZ-400 waveforms were collected at a lower density than the point cloud data, with points in the waveform collection being spaced approximately 2.5 cm in each direction (at a distance of 20 m from the scanner), so a total of 64 waveforms were extracted to represent each 20 x 20 cm^2 voxel area. The 64 raw waveforms from the Riegl VZ-400 scanner and the average waveform for the trunk simulation are shown in Figure 8, and in Figure 10 for the canopy simulation.

8.1 Comparisons of simulated and real waveforms from tree trunk

Riegl VZ-400 waveforms from the 20 x 20 cm^2 area of the tree trunk are shown in Figure 8, and the corresponding simulated waveform is shown in Figure 9. A significant amount of variability is apparent in the waveforms from the Riegl VZ-400 TLS, even from the relatively smooth, flat tree trunk surface. A similar illustration of the variability in LiDAR waveforms is shown in Parish 2011.[8] In that study, data were collected with an Optech ALTM Gemini system, and results show waveforms from flat targets having a significant amount of variability, even when collected in a controlled laboratory environment. In the Parish 2011 study, some of the variability in returned waveforms was attributed to not allowing sufficient time for the system to stabilize before data collection; however, variability is still present even in the case where sufficient stabilization time was allowed. This variability in LiDAR waveforms is something that is not currently captured within the LiDAR simulation, and is an example of one of the complicating factors in creating a realistic LiDAR simulation.

When comparing the simulated and real waveforms of the tree trunk, it is observed that the location of the peak maximum in both the simulated and real waveforms occurs at a range of approximately 20.5 m. The most notable difference between the real and simulated waveform is the width of the peak, and the trailing edge present in the real waveform. It is hypothesized that these differences may be attributable to the sensor impulse response function. Relatively little effort has been applied to examining the affect of the instrument response function within the simulation as of yet. In this case, a 3 ns Gaussian impulse response function is assumed, along with a 1.8 GHz sampling rate. The waveforms from the Riegl VZ-400 exhibit a trailing edge response that

decays slowly, indicating that a Gaussian impulse response function may not be the most appropriate choice. Also, noise effects have not been included in the LiDAR simulation as of yet. The width of the peak is also influenced by multiple scattering. Figure 12 shows the trace of simulated energy for the tree trunk waveform simulation; even in the case of the relatively smooth and flat tree trunk surface, multiple scattering occurs, as evidenced by the vertical offset between the multiple reflection paths of energy returning to the sensor. A larger number of simulation iterations may be needed to fully capture the amount of multiple scattering occurring in the scene.

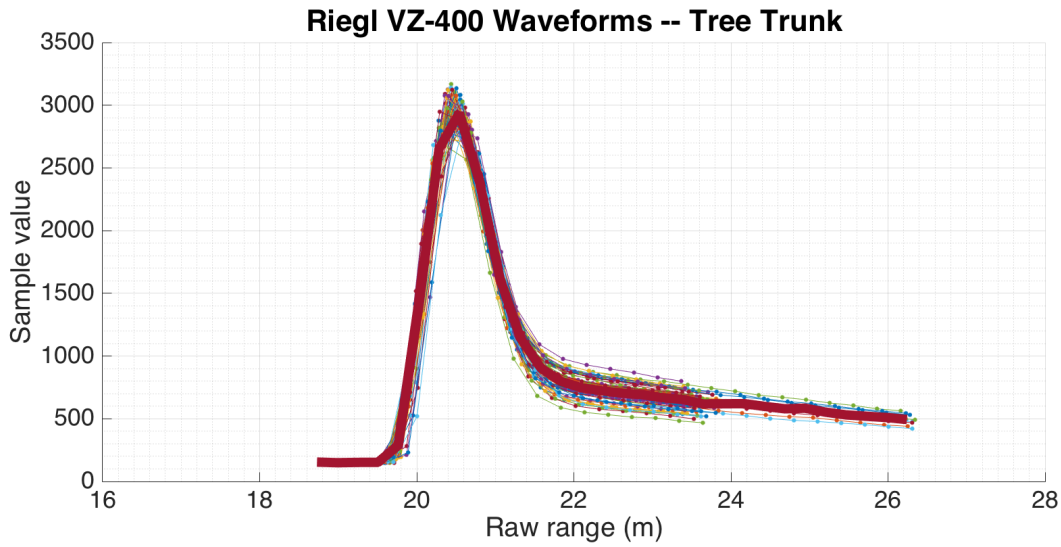

Figure 8: A total of 64 Riegl waveforms representing a subset approximately equal in size to the voxel resolution of the LiDAR simulation. The waveforms are plotted individually, and the average waveform is shown as the bold red line.

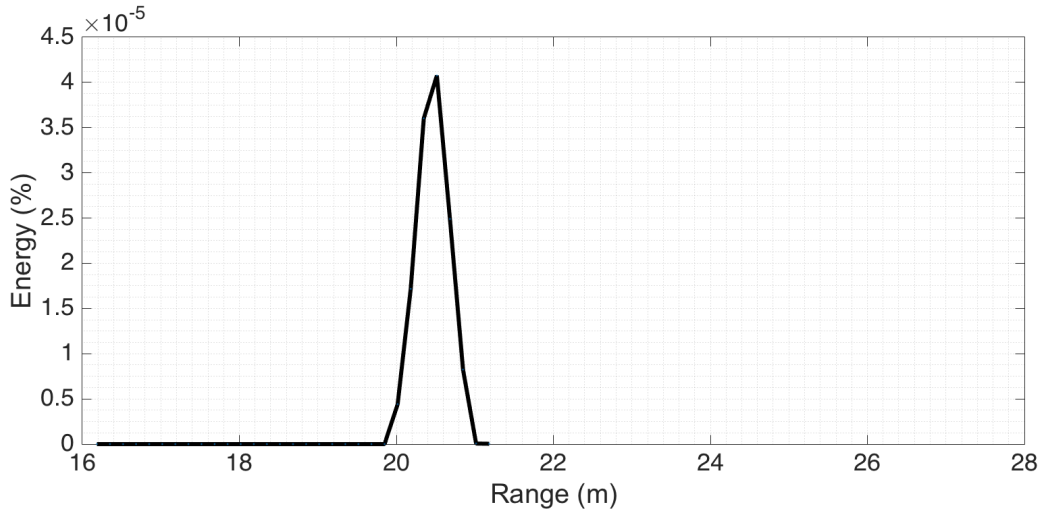

Figure 9: The simulated waveform based on replication of the Riegl VZ-400 scanning arrangement used to collect the waveforms shown in Figure 8.

8.2 Comparisons of simulated and real waveforms from tree canopy

For the Riegl VZ-400 waveforms from within the tree canopy (Figure 10), there is significant amount of variability in the 20 x 20 cm^2 region. The simulated waveform for the canopy simulation is shown in Figure 11. Several

differences between the simulated and real waveforms of the tree canopy are readily apparent. The most notable are the ranges to the initial and final peaks, the width of the peaks, and the relative amplitudes of the first versus the second and third peaks.

In the simulated tree canopy waveform, the initial peak occurs at a range of approximately 17 m, whereas in the real Riegl VZ-400 data, the first peak (on average) occurs closer to a range of 18 m. It is hypothesized that the difference in the initial peak location is at least partially due to the voxel nature of the simulation. An examination of the trace of the simulated energy that returned to the sensor (Figure 13) reveals that some of the returned energy is being reflected from the leading edge of the canopy (at a range of approximately 17 m from the sensor). In the case of the Riegl VZ-400, with a beam diameter of only 0.6 cm at 20 m, it is very likely that the laser beam penetrates into the canopy, rather than returning from the leaves or wood at the very edge of the canopy. In the LiDAR simulation, although the voxels along the canopy edge have appropriately low values of LAD or wood density, there is some probability that the simulated laser pulse will interact with these voxels. The definition of the leaf surface normals also has a significant effect on both the LAD value returned from VoxLAD, and the determination of the likelihood of interaction with leaf materials at each voxel. More work is needed to determine the most appropriate definitions of leaf surface normal directions.

The final peak in the simulated tree canopy waveform occurs at a range of approximately 24.5 m, and at a range of 24 m in the real data. Investigation of the trace of simulated energy that returned to the sensor (Figure 13) indicates that the "delay" in the simulated waveform may be due to multiple scattering in the simulation. It is unclear from this particular situation whether the degree of multiple scattering occurring in the simulation is correct or not. Further simulations and comparisons to real waveforms may help to illuminate this. Several parameters within the LiDAR simulation that affect the degree of multiple scattering can be adjusted, including multiplication factors for the material density maps, and reflectance properties of the materials. (Note: The model can also be restricted to artificially disallow multiple scattering completely for the purpose of examining the effects of multiple scattering on the waveform.)

The replication of the exact scanning geometry of the Riegl VZ-400 scans within the simulated scene enables viewing the projection of the transmitted pulse within the point cloud data. It should be noted that this is also possible within the Riegl's RiScan Pro software;[9] however, the outputs from the LiDAR simulation are easy-to-use ASCII text files, enabling use of open source or commercial software packages for point cloud viewing. Viewing the projection of the transmitted laser pulse within the point cloud data provides several insights. The furthest edge of the canopy, along the path of the transmitted laser pulse, occurs at a range of approximately 24.4 m from the sensor. It isn't clear from examination of these particular Riegl waveforms whether the laser is returning energy from the very edge of the canopy, or if the peak located at 24 m and the spread in that peak is due to multiple scattering within the canopy. Again, it is expected that further comparisons of simulated and real waveforms will help to illuminate this.

There are obvious differences in widths of the simulated and real waveform peaks; in the Riegl VZ-400 waveforms, the peaks are wider than the simulated waveform. As discussed in Section 8.2, the difference in widths may be partially due to an improper simulation of the sensor impulse response function. Another possible explanation of the peak widths is multiple scattering within the canopy, which is influenced by the reflectivity properties of the canopy materials, and the shapes of the surfaces interacted with. A more extensive investigation of these factors is needed.

There are also differences in the relative amplitudes of the peaks. In the Riegl VZ-400 waveform data, the amplitude values are raw sample values (Figure 10); in the simulated waveform, the amplitude is given as a percentage of the initial transmitted energy (Figure 11). In the simulated waveform, the initial peak is significantly taller than the following two peaks; in the Riegl VZ-400 waveform, the amplitudes of the peaks diminish more gradually with range. It is hypothesized that the amplitude of the peaks in the simulated waveforms is impacted by the number of simulated points. Further work is needed to determine appropriate stopping criteria for the LiDAR simulation.

A final difference to discuss is the presence of noise in the real waveform data (minimum sample values>0). Noise effects have not been included in the LiDAR simulation up to this point.

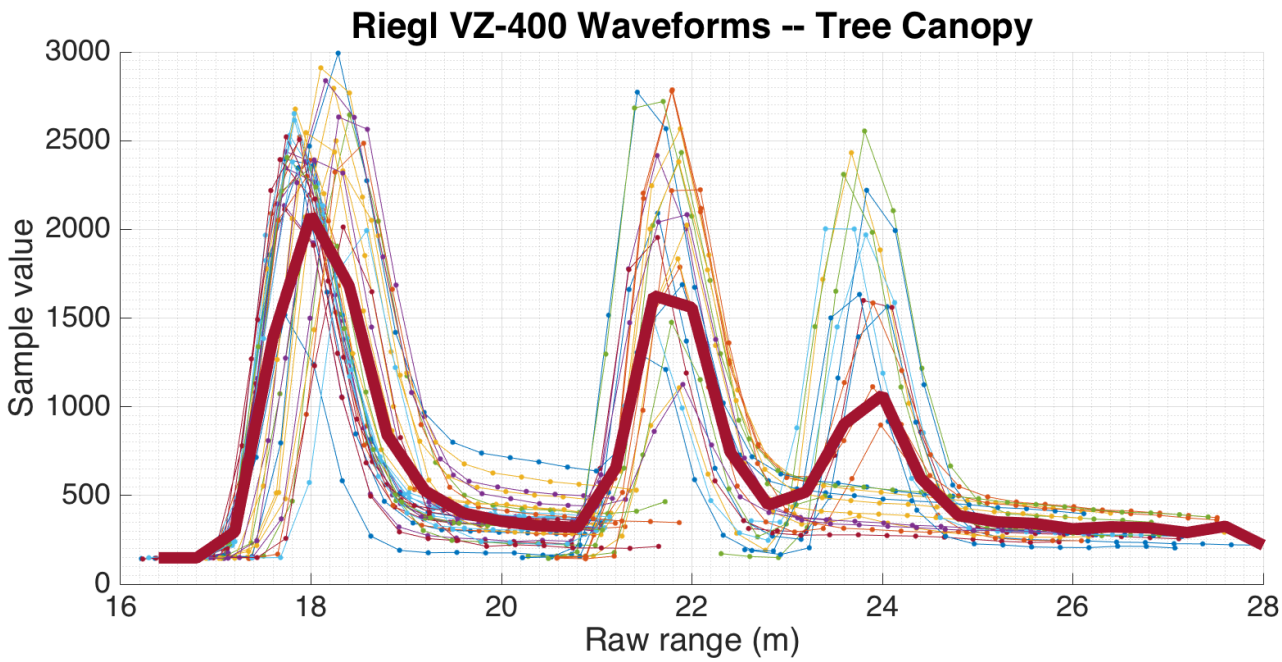

Figure 10: A total of 64 Riegl waveforms representing a subset approximately equal in size to the voxel resolution of the LiDAR simulation. The waveforms are plotted individually, and the average waveform is shown as the bold red line.

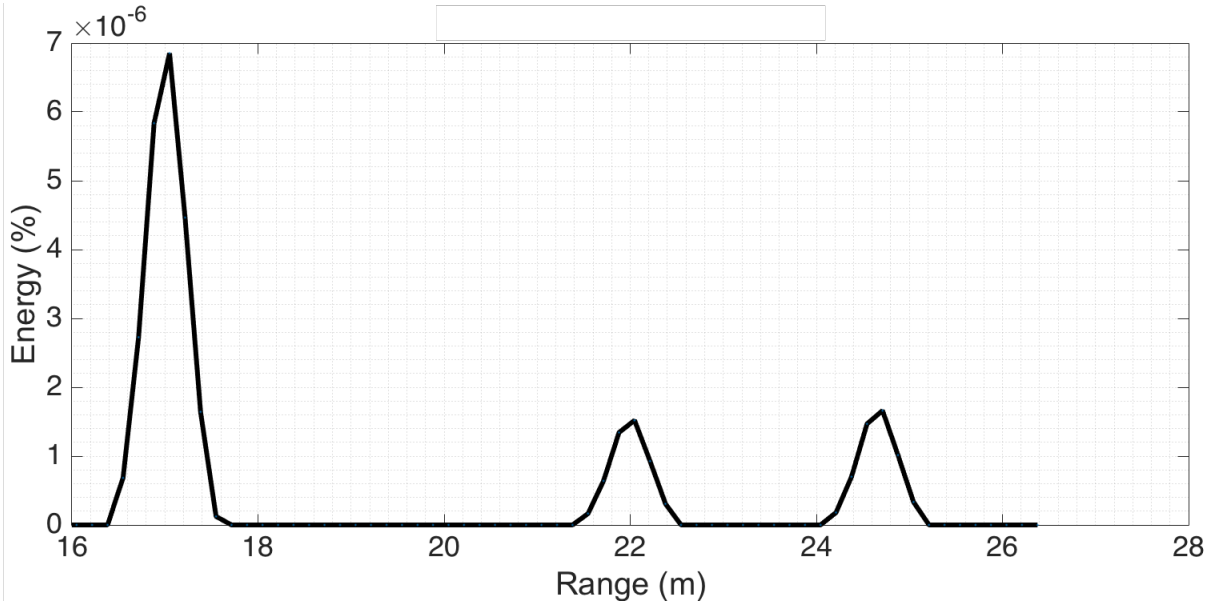

Figure 11: The simulated waveform based on replication of the Riegl VZ-400 scanning arrangement used to collect the waveforms shown in Figure 10.

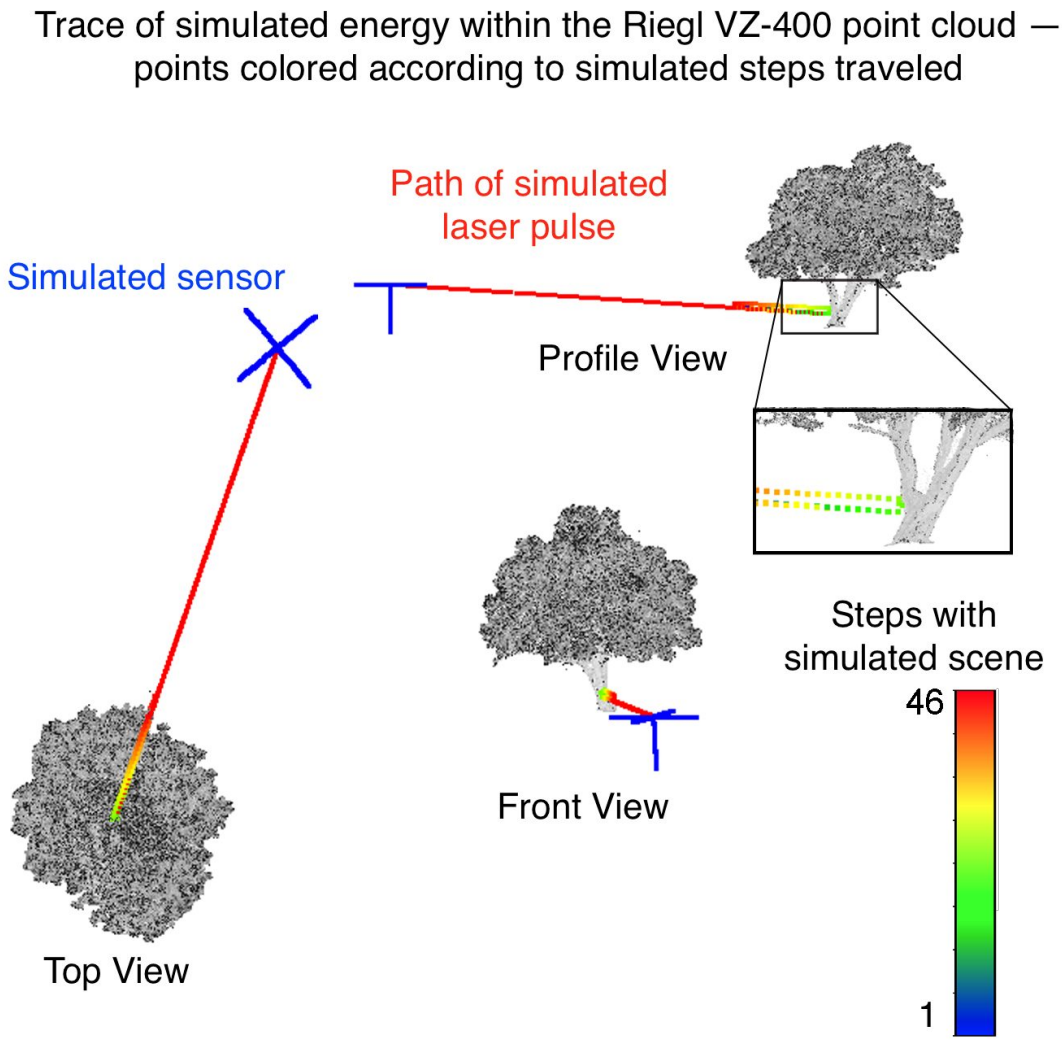

Figure 12: The trace of simulated energy that returned to the sensor and contributed to the simulated full-waveform signal for the tree trunk simulation. The simulation replicates the exact scanning geometry used in the Riegl VZ-400 scans, so the simulated data is shown overlaid on the Riegl point cloud data. The sensor location is indicated in blue and the transmitted pulse is shown in red. The trace of simulated energy is given in rainbow scale according to the number of pulses traveled within the simulated scene – the pulse is only traced within the simulated scene, so the path the simulated energy makes from scene back to the sensor is not shown. The inset box shows a zoomed in view of the pulse trace, illustrating multiple scattering as evidenced by the vertical offset between the multiple reflection paths of energy returning to the sensor.

9. DISCUSSION

There are obvious differences between the simulated and real waveforms as collected by the Riegl VZ-400, and issues to address before the simulation can be "validated"; however, the correspondence between the simulated and real waveforms is encouraging. The simulation of full-waveform LiDAR data has also already proved useful in viewing and examining the real Riegl VZ-400 full-waveform signals. In this section, we discuss some of the many factors that may be contributing to the differences between the simulated and real waveform signals.

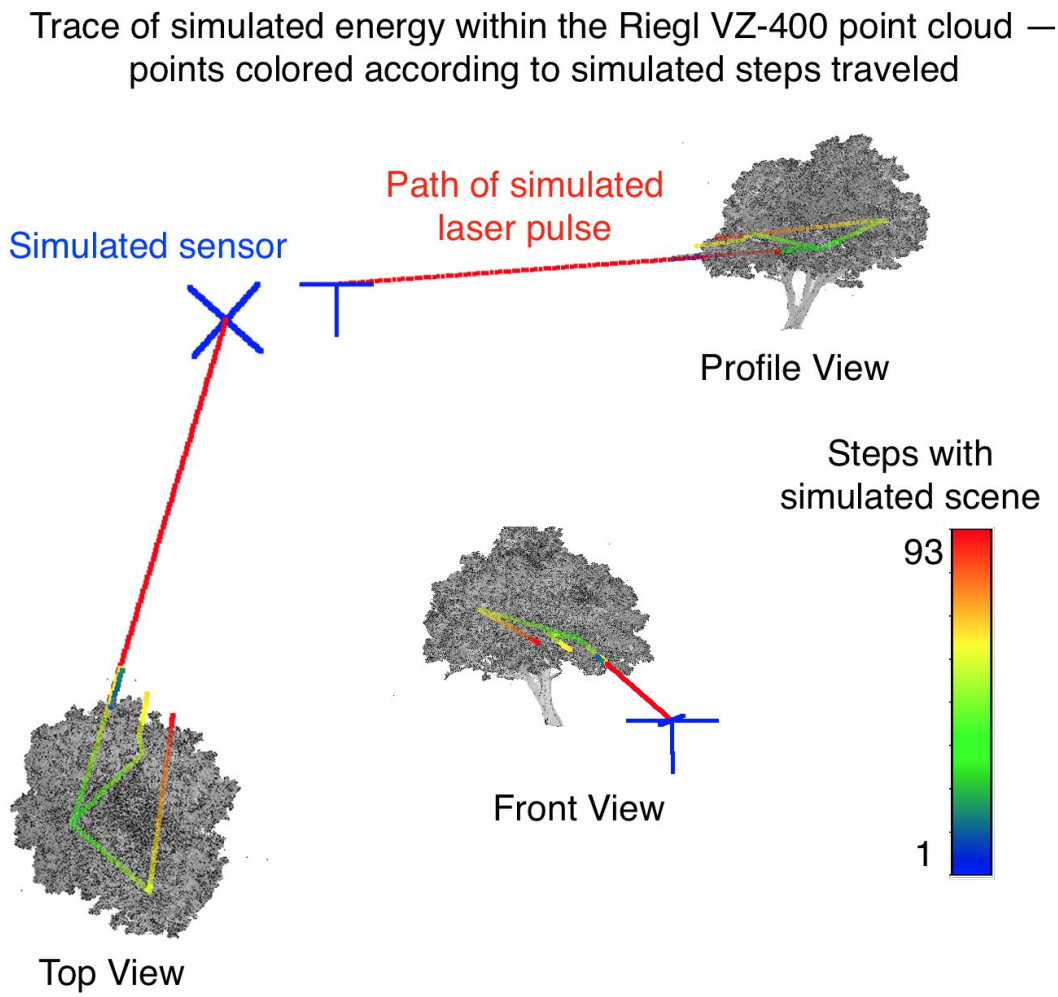

Figure 13: The trace of simulated energy that returned to the sensor and contributed to the simulated full-waveform signal for the complex tree canopy simulation. The simulation replicates the exact scanning geometry used in the Riegl VZ-400 scans, so the simulated data is shown overlaid on the Riegl point cloud data. The sensor location is indicated in blue and the transmitted pulse is shown in red. The trace of simulated energy is given in rainbow scale according to the number of pulses traveled within the simulated scene – the pulse is only traced within the simulated scene, so the path the simulated energy makes from scene back to the sensor is not shown.

The high spatial resolution of the Riegl VZ-400 data, as well as the extensive metadata available with the raw waveform data, enables creation of a simulation that replicates the real world data collect. In previous attempts to compare the simulated data to real-world data from an airborne full-waveform LiDAR system,[2] the simulation had a higher spatial resolution than the real-world data, whereas in this case, the Riegl has a much higher spatial resolution than the simulated scene. The high spatial resolution of the Riegl data introduces challenges in deciding how to compare simulated and real-world data; in this work, we averaged a series of waveforms representing an area approximately equal to the voxel size of the simulation. More work is needed to determine if this is a valid way of dealing with the differences in spatial resolution.

Further work is also required to ensure the correct settings are being used for the VoxLAD simulation for this particular type of tree and scanning arrangement. Settings within the VoxLAD simulation, including the voxel size, definition of leaf angle distribution, and classification of the input point cloud into leaves, wood, and noise classes all affect the returned values for LAD. The non-traditional use of the VoxLAD model to return a probability map for wood materials also needs to be investigated further.

Settings within the LiDAR simulation also need further investigation. The reflectance properties of the materials in the scene must be verified. In this case, the leaves were defined as having a reflectivity of 75%, transmittance of 20% and absorbance of 5% – these values are most likely not representative of the response of leaves at the laser wavelength of the Riegl VZ-400 (1550 nm). Spectral measurements of the leaves, wood and ground material in the scene are necessary to verify the correct settings. The definition of the leaf surface normals is another item that needs to be verified. The LiDAR simulation currently assumes the azimuthal component of the leaf surface normal is oriented away from the tree center; in reality, some randomness should be incorporated to make this a more realistic representation of the leaf arrangement. The zenith component of the leaf surface normal is defined according to a uniform distribution; a different distribution may be more appropriate. Finally, the relative importance of these factors needs to be analyzed in order to help guide the user's time and effort; the more realistic the simulation is, the better the results; however, trade-offs must be made between accuracy and time available to dedicate to defining the simulation parameters. This "sensitivity analysis" would also provide some indication of the information which can be reliably extracted from waveform signals.

Finally, some factors about the Riegl VZ-400 are unknown; in particular, specifics about the electronics of the system (such as the sensor instrument response function) and other factors of the physical system (such as the aperture size) have not been verified. It is clear from results shown in Parish 2011, that even with knowledge of these system factors, a significant amount of variability in LiDAR waveforms may be present in recorded signals.[8] The LiDAR simulation does not currently incorporate any of this variability due to system factors.

Further work is needed to address the uncertainties in the simulation, and discrepancies between the real and simulated data. A validated simulation of full-waveform LiDAR data will enable the analysis of full-waveform signatures, and the ability to design algorithms for extracting even greater amounts of information from LiDAR data.

10. CONCLUSION

A 3-D simulation of full-waveform LiDAR propagation, with a 3-D voxel tree created using the VoxLAD model, has been used to simulate a terrestrial laser scanner waveform signal. Real waveforms from the Riegl VZ-400 collection were compared to simulated waveforms. Results demonstrate that the scanning geometry of the real-world collect can be accurately replicated within the LiDAR simulation. Although the resulting simulated waveforms exhibit some obvious differences to the real waveforms, there is some correspondence between the two. In particular, the location of the returned waveform peak from the solid tree trunk surface corresponds well to the Riegl VZ-400 waveform.

Several complex factors may be contributing to the differences in the real-world Riegl VZ-400 waveforms and the simulated data, and more work is needed to determine which factors are causing the most significant differences, and ways these factors can be addressed in order to create the most realistic simulation possible. The ability to replicate the exact scanning arrangement of the Riegl VZ-400 scans has already proved useful in viewing the projection of the transmitted laser beam within the point cloud data.

11. ACKNOWLEDGEMENTS

The Naval Postgraduate School acknowledges the support of the US Army Corps of Engineers for this project.

REFERENCES

[1] Ullrich, A. and Pfennigbauer, M., "Echo digitization and waveform analysis in airborne and terrestrial laser scanning," *Proceedings of Photogrammetric Week* (2011).

[2] Kim, A. M., Olsen, R. C., and Béland, M., "Simulation of small-footprint full-waveform LiDAR propagation through a tree canopy in 3D," *Proceedings of SPIE Defense, Security, and Sensing* (2015).

[3] Béland, M., Widlowski, J.-L., Fournier, R. A., Côté, J.-F., and Verstraete, M. M., "Estimating leaf area distribution in savanna trees from terrestrial lidar measurements," *Agricultural and Forest Meteorology* **151**(9), 1252–1266 (2011).

[4] Béland, M., Baldocchi, D. D., Widlowski, J.-L., Fournier, R. A., and Verstraete, M. M., "On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial lidar," *Agricultural and Forest Meteorology* **184**, 82–97 (2014).

[5] Béland, M., Widlowski, J.-L., and Fournier, R. A., "A model for deriving voxel-level leaf area density estimates from ground-based lidar," *Environmental Modelling & Software* **51**, 184–189 (2014).

[6] RIEGL Laser Measurement Systems, *3D Terrestrial Laser Scanner Riegl VZ-400/Riegl VZ-100 General Description and Data Interfaces*. RIEGL Laser Measurement Systems GmbH, Austria (2012).

[7] RIEGL Laser Measurement Systems, "Waveform extraction library," tech. rep., RIELG Laser Measurement Systems GmbH (2014).

[8] Parish, C. E., Jeong, I., Nowak, R. D., and Smith, R. B., "Empirical comparison of full-waveform LiDAR algorithms," *Photogrammetric Engineering & Remote Sensing* **77**(8), 825–838 (2011).

[9] RIEGL Laser Measurement Systems, *RiScan Pro User's Manual Version 1.7.7*. RIEGL Laser Measurement Systems, Austria (February 2013).