INTERFACE FOCUS

Spectral and spatial information from a novel dual-wavelength full-waveform terrestrial laser scanner data for forest ecology

Journal:	*Interface Focus*
Manuscript ID	RSFS-2017-0049.R2
Article Type:	Research
Date Submitted by the Author:	20-Dec-2017
Complete List of Authors:	Danson, Francis; University of Salford, School of Environment and Life Sciences Sasse, Fadal; University of Salford, School of Environment and Life Sciences Schofield, Lucy; York St John University, School of Humanities, Religion and Philosophy
Subject:	Biogeography < CROSS-DISCIPLINARY SCIENCES, Environmental science < CROSS-DISCIPLINARY SCIENCES
Keywords:	Terrestrial laser scanner, forest ecology, full-waveform, SALCA, dual-wavelength

Spectral and spatial information from a novel dual-wavelength full-waveform terrestrial laser scanner data for forest ecology

F. Mark Danson[1*], Fadal Sasse[1], and Lucy A. Schofield[2]

1. Ecosystems and Environment Research Centre, School of Environment and Life Sciences, University of Salford, Salford M5 4WT, UK

2. School of Humanities, Religion and Philosophy, York St. John University, York YO31 9EX, UK

* Corresponding author f.m.danson@salford.ac.uk

Abstract

The Salford Advanced Laser Canopy Analyser (SALCA) is an experimental terrestrial laser scanner designed and built specifically to measure the structural and biophysical properties of forest canopies. SALCA is a pulsed dual-wavelength instrument with co-aligned laser beams recording backscattered energy at 1063 and 1545 nm; it records full-waveform data by sampling the backscattered energy at 1 GHz giving a range resolution of 150 mm. The finest angular sampling resolution is 1 mrad and around 9 million waveforms are recorded over a hemisphere above the tripod-mounted scanner in around 110 minutes. Starting in 2010, data pre-processing and calibration approaches, data analysis, and information extraction methods, were developed and a wide range of field experiments conducted. The overall objective is to exploit the spatial, spectral and temporal characteristics of the data to produce ecologically useful information on forest and woodland canopies including leaf area index, plant area volume density and leaf biomass, and to explore the potential for tree species identification and classification. This paper outlines the key challenges in instrument development, highlights the potential applications for providing new data for forest ecology, and describes new avenues for exploring information-rich data from the next generation of TLS instruments like SALCA.

Keywords: Terrestrial laser scanner, forest ecology, full-waveform, SALCA

1. Introduction

A new generation of multi-wavelength full-waveform recording terrestrial laser scanners (TLS) is now being developed for applications in forest ecology. To date most commercially available TLS instruments have been designed for very accurate range measurements, at high point density, for applications in building and infrastructure survey, archaeology and the mineral extraction industries. The new instruments aim to complement this range information with calibrated measurements of object reflectance and waveform shape, using two or more laser wavelengths. A key driver for these developments is the need to classify objects in TLS point clouds according to their reflectance properties and, in the case of forest ecology, to separate returns from foliage and those from the woody material of the tree trunks and branches. Potential solutions tested in this paper are to make measurements using two or more laser wavelengths where there is a contrast between the reflectance of foliage and wood in a forest stand, or to analyse the local-scale three dimensional (3d) characteristics of the point cloud.

A further driver is that most commercial instruments record only the X,Y,Z position of each laser return, and some uncalibrated measure of the intensity of laser backscatter and conversion of these data to calibrated reflectance may be difficult in the absence of detailed information about sensor characteristics. For quantitative use of TLS intensities in forest applications, calibrated data are required in order to compare object reflectance, and this requires the full waveform of backscattered energy to be recorded. It is also the case that the detailed instrument characteristics, data recording methods and pre-processing algorithms are unknown and often commercially protected; for a research instrument the specification and design are known, and the instrument developer can provide access to data processing algorithms in an open way. Finally, forests present a particular geometric arrangement of material that requires a hemispherical or even spherical field of view, in order to observe the stems, canopy, and forest understorey. Many commercial instruments operate using a panoramic viewing configuration (for example the Riegl VZ-400) so that two or more scans are required to measure a complete hemisphere above the instrument. An instrument designed for forest canopy measurements is therefore required to measure a complete hemisphere above the instrument and some part of the understorey or forest floor.

These drivers together have stimulated the development of a new generation of TLS for forestry applications starting with the first full-waveform TLS, Echidna [1]. This instrument was designed to provide calibrated full-waveform data in order to measure forest plant area index (PAI) and plant area volume densities (PAVD) [2] with attempts to classify the points as wood or foliage using the width of the returns. Subsequently, two dual-wavelength instruments were developed independently, based on the heritage of Echidna, the dual-wavelength Echidna lidar (DWEL) [3] and the Salford Advanced Laser Canopy Analyser (SALCA) [4]. The development and applications of DWEL are described in Li et al. [5], and in this paper the

development and applications of SALCA are described.

2. Previous research

Separation of foliage and woody material using light detection and ranging (lidar) data and other indirect approaches, remains a key goal of improving measurements of forest structure and function. One method of achieving this separation with TLS data is to apply a manually determined threshold based on the intensity of the reflected signal. This technique relies on there being a measurable difference in reflectance between the woody material and the foliage at the wavelength of the TLS instrument, for example in the region of absorption by leaf water near 1550 nm [6], or in green wavelengths [7]. This method has been tested for broadleaved [6] and coniferous [8] species on an individual tree scale, sometimes requiring a different threshold for each tree [9]. Béland et al. [6] compared leaf-off and leaf-on scans from two positions either side of a tree using a first return Optech ILRIS-3D TLS at 1545 nm wavelength. They first normalised the TLS return intensities to 20 m range using linear fitting, ignoring near and far range as the tree was located approximately 20 m from the scanner, and then identified an intensity threshold by analysing normalised intensity histograms for leaf-off and leaf-on scans. To compensate for parts of branches not visible to the TLS in the leaf-on scans (occluded by leaves) the bin values for the leaf-off histograms were multiplied by a correction factor (between 0 and 1). Once separated into leaves and wood, the data were then voxelised to complete the analysis. Béland et al. [10] extended this approach by using a multiple return Riegl VZ-400 TLS at 1550 nm wavelength to scan ten oak trees (mean height 9.5 m). Histograms of normalised intensities were analysed to distinguish between multiple and single returns from leaf-off and leaf-on scans, and to categorise the returns as foliage, wood, or noise. These studies highlighted the potential of using TLS intensity measurements for foliage-wood separation but also showed that it is important to account for the impact of 'partial hits', where the laser beam is not completely filled by the object of interest.

Methods have also been developed for foliage-wood separation by classifying point clouds based on their three dimensional (3d) geometrical properties. Brodu & Lague [11] developed a procedure for examining point clouds at a given location and scale, identifying lines, planes, and volumes to classify point clouds. Combining this information from multiple scales allowed classification of objects in a given scene, with apparently high levels of accuracy. This method has been used along with apparent reflectance data to provide a classification of foliage and wood in forest stands at Harvard Forest, U.S., from the DWEL instrument [12]. Similar research by Tao et al. [13] used geometrical features for classification, but in their approach, tree point clouds were sliced into horizontal layers and within each segment, circle (and circle-like) features were identified as wood, and line segments as leaves. This achieved misclassification errors of 10.7% on a 'virtual' tree, and 13.4-16.9% on real individual trees. Additionally, the incorporation of other attributes, such as pulse shape, have been considered for improving classification. For instance, Yang et al.

[14] classified trunk/branch or foliage by thresholding the relative width of the return waveforms, and Fieber et al. [15] classified an orchard according to pulse width and a backscattering coefficient using airborne full-waveform data. With the recent and rapidly evolving development of tree structure models, new methods are now on the horizon for modelling foliage and needle distributions based on tree structure models derived from TLS data [16].

Multi-wavelength TLS measurements also have potential for measuring the biophysical properties of vegetation canopies, for example leaf chlorophyll and nitrogen content [17,18]. Data from the dual-wavelength SALCA instrument has been shown to accurately measure leaf water content using three different species and a normalised difference index of the near- and shortwave infrared wavelengths [19], and the index has been shown to be insensitive to lidar incidence angle [20]. These results have been further supported in more recent research using multi-wavelength TLS [21,22].

3. Dual wavelength measurement principle

The received power recorded by a TLS (P_r) is often described using the radar equation that defines the relationship with emitted energy, range, instrument optics, and object backscatter (equation 1) [23]:

$$P_r = \frac{P_t D_r^2}{4\pi R^4 \beta_t^2} \sigma \quad (1)$$

where P_t is the transmitted power, D_r aperture diameter of receiver optics, R range to the target, β_t^2 beam divergence, and σ the backscatter cross section computed as equation 2 [23]:

$$\sigma = \frac{4\pi}{\Omega} \rho A_s \quad (2)$$

where Ω is a solid angle defining the backscattering cone, ρ is the object reflectance and A_s is the area of the scattering element.

If the instrument-related measurement characteristics are the same for two different laser wavelengths, then the received power is dependent only on range, object reflectance characteristics (spectral and directional) and the area of scatterer. For single wavelength TLS the range-dependent variation in received power can be defined and normalised, but the object reflectance and area of scatterer are always confounded. This in turn means that it is not possible to determine the reflectance of an object without knowing the area it occupies in the laser beam, or *vice versa*. For single wavelengths the measurement of backscatter may be referred to as 'apparent reflectance' (AR), equivalent to the reflectance of a diffuse target filling the beam that returns the same intensity as the measured target. For example, an apparent

reflectance of 20% could be from an object with a reflectance of 40% occupying 50% of the beam, or from an object with a reflectance of 80% occupying 25% of the beam. For a two-wavelength measurement, where the backscattering characteristics of the object are similar, the ratio of the apparent reflectance values for a given object is independent of the area of the scattering element (As) since the area is cancelled out. This means that the ratio is dependent only on the ratio of the reflectance of the object in the two wavelengths, assuming that the beams are co-aligned and the footprints are the same size at any given range. The two-wavelength measurement principle is therefore that objects that have a different reflectance ratios at two laser wavelengths may be classified based on this spectral ratio, independent of the area of those objects in the laser beam. Conversely if the object reflectance is known, then the area intercepted by the beam is given by the ratio of the apparent reflectance to actual object reflectance.

The spectral properties of foliage and woody material in a forest are determined by their structural and biophysical properties with typical spectral signatures shown in Figure 1. The green leaf spectrum is driven by the absorption properties of chlorophyll in the visible wavelength region, water in leaves in the shortwave infrared region, and scattering caused by leaf internal structure in the near-infrared. The bark spectrum is also determined by pigments and scattering, but since the water content is generally lower than the leaves, the SWIR reflectance is expected to be higher than that for leaves. At the wavelengths used by SALCA, foliage and wood are expected to have a similar reflectance at 1063 nm but contrasting reflectance at 1545 nm. This contrast has led to the concept of using spectral ratios with two-wavelength TLS to aid the classification of point clouds and in this research a Normalised Difference Index (NDI) was tested (equation 3):

$$NDI = \frac{(\rho 1063 - \rho 1545)}{(\rho 1063 + \rho 1545)} \quad (3)$$

where ρ is reflectance at a given wavelength. This index has maximum range of -1 to +1 with positive values indicating that the 1545 nm reflectance is lower than the 1063 nm reflectance and *vice versa.* For example in figure 1, the NDI for the leaf is 0.23 and for the wood 0.03. In 'leaf-off' conditions in a forest stand, with no leaves in the canopy, the NDI is expected to be controlled solely by the wood reflectance, and the NDI values to be clustered around zero. In leaf-on conditions the NDI for the woody targets should be around zero and for the leaf hits a positive value around 0.2 – 0.3 for example.

4. SALCA data characteristics and data pre-processing

The Salford Advanced Laser Canopy Analyser (SALCA) was developed as an experimental system in response to the limitations of existing commercial laser scanners outlined earlier. SALCA is dual-wavelength TLS recording the full waveform of backscattered energy with a range resolution of 150 mm. The two laser

wavelengths (1063 and 1545 nm) are emitted sequentially and measurements are made to a maximum range of 105m. The instrument scans in hemispherical mode measuring from 6 degrees below the horizontal with a minimum zenith and azimuth step angle of 1 mrad (0.06 degrees). The development of the instrument and its characteristics are described in more detail in Danson et al. [4] and the approach used for reflectance calibration of SALCA data using a trained neural network is described in Schofield et al. [24].

The backscattered energy detected by SALCA is recorded as an 8-bit digital number stored initially over a range of -127 to +127. The 1 Ghz sampling rate of 1 ns leads to a digitized sample equivalent to a measurement every 150 mm in range, with 700 measurements for the 1545 wavelength, followed by a maximum of 700 measurements for the 1063 nm wavelength. Part of the outgoing 1545 nm pulse is normally visible in the data due to internal reflections within the scanner optics, and this is used to define zero range for this wavelength and, with the corresponding part of the outgoing 1063 nm pulse, to match the waveforms for each wavelength. Instrument noise has been found to be highly stable and independent of wavelength so that a simple threshold can be applied to define returns above this threshold.

An example of a single two-wavelength waveform is shown in Figure 2. The outgoing 1545 nm pulse is just visible at zero range on the far left of the figure, with a series of 1545 nm wavelength returns between 12 and 16 m range. The stable noise signal is then visible in the 1545 nm data up to 147 m where the outgoing pulse for the 1063 nm wavelength is then seen at zero range for this wavelength followed by corresponding 1063 nm returns between 12 and 16 m. After dividing the waveform into two separate wavelengths, the data for both wavelengths is matched in range using the outgoing pulse signals, and an intensity threshold is applied to remove instrument noise. The intensity values (DN) for each return are then summed to provide an 'intensity sum' for each measured return that is detected above the noise threshold (figure 3). The range to each return is calculated using the 'centre of gravity' which uses both range and intensity information. These methods have been shown to be the optimum data extraction approaches for SALCA as they produce a higher signal-to-noise-ratio compared with other approaches and facilitate sub-bin range resolution measurement [25].

The geometric information required to determine the three-dimensional coordinate of each return is derived from a look-up-table relating the scan zenith and azimuth index to a 'true' scan zenith and azimuth angle accounting the specific scan geometry of the instrument, and this is used with the range estimate, to determine its Cartesian X,Y,Z coordinate. Additional post-processing steps analyse the width of each return in 'bins', and the return number (first, second etc.). Simple algorithms are also used to remove false returns caused by 'ringing', and to correct very high intensity sums caused by multiple returns over consecutive range bins that cannot be separated with our current return detection approach.

The final pre-processing step is to convert the intensity sums for each return into apparent reflectance. This is achieved using field-measured data on a painted multi-reflectance calibration panel that was measured at multiple ranges with SALCA under field conditions over a period of twelve months. Part of the calibration process involves accounting for variations in laser output as a function of its temperature. There is a well-known drop in laser output power as the temperature of the laser module increases [26] and although there are small cooling fans located in the scan head of SALCA, they are not able to maintain constant laser temperature. To address this issue the laser case temperatures are measured every 5 minutes during a scan using thermistors attached to each laser module. These data were used along with the range and calibration panel intensity from the field measurements to train a neural network that is used to convert the recorded laser intensities to apparent reflectance, taking into account wavelength, range, laser temperature and instrument characteristics [24]. Finally, to take advantage of the dual-wavelength capability, and compute the NDI, the corresponding returns need to be paired. That is, each return is coupled with the equivalent return in the other wavelength and the NDI calculated. The NDI can only computed when there is a pair of matching returns in each wavelength; other returns with no match for either wavelength are discarded. In the current algorithm a return match is only accepted when the ranges of the matching pair of returns are within a range of less than 0.12 m, that is less than one range bin.

5. Exploring SALCA data characteristics in a forest stand

To illustrate the wide range of information available in SALCA dual-wavelength point clouds, figure 4 shows the key variables measured in a forest stand. The data were recorded on 19th June 2014 in an oak-dominated stand of deciduous trees in full leaf-on conditions, at Delamere Forest, Cheshire, UK (53.2397° N, 2.6857° W), using the finest sampling resolution (0.06 degrees) in both zenith and azimuth direction. The recorded data sets consist of 9.7 million laser shots for each wavelength. The figures show the processed point clouds for the 1545 nm wavelength plotted in an equal angle projection with the vertical axis representing zenith angles of 90 to 0 degrees and the horizontal axis azimuth angles from 0 (on the left) to 180 degrees (on the right), that is, half a hemisphere above the scanner. The visual appearance of the data was very similar for the 1063 nm wavelength except for the spectral information (intensity and apparent reflectance) which is analysed in detail in the next section. Overall for the 1545 nm wavelength there were 10,827,134 detected returns recorded, with 62 % of these single returns, and 48 % multiple (two or more) returns; the equivalent data for the 1063 nm wavelength were 10,159,879, 71% and 29% (table 1).

The Intensity image is a combination of the range effect and variation in object reflectance. Intensity decreases with range, as expected, but at a given range the stems appear brighter than the foliage in the 1545 nm wavelength data. The Range image shows objects up to 60 m from the scanner since the data are

trimmed to this range as the reflectance calibration is unreliable (due to lack of measurements) beyond this range.

The Returns image shows that single returns come mainly from full hits on the main tree stems and branches but some are also visible at longer ranges for leafy area where the backscattered energy is so low that only one return can be detected. Single returns may also include some single partial hits from objects at all ranges. Two or more returns are recorded in the areas with leafy material, including areas of understorey vegetation. The Return Width image is derived from a simple measure of the pulse width (in bins of 150 mm). The widest pulses (red) are from the leafy areas at near to middle range indicating a volume of well distributed scattering elements over scales of 0.5 m or greater. The mid-width pulses (black) come from the stems at near to medium range and correlate with the single returns. The narrowest returns are from far range where the return intensities are much lower; here the waveforms have the same width as the outgoing pulse but the lower signal strength means that the tails of the waveforms fall below the noise threshold so that the data are recorded over fewer bins. At far range tree trunks may only be recorded as a single intensity measurement (one bin).The Apparent Reflectance provides a first indication that there is a spectral difference between the foliage and woody material; the stems and branches have a higher reflectance than the foliage and this difference is visible at all ranges. Finally the NDI image shows the spectral contrast between the 1063 and 1545 nm apparent reflectance data, with some evidence that the main stems and branches have a lower value than the leafy area of the canopy.

This visual interpretation of the 1545 nm point clouds serves to illustrate the rich information content of SALCA data. There is range dependence for intensity, returns and width which means that it will be difficult to use such metrics to extract information about the structure of a forest stand; recent analysis has therefore focussed on assessing the potential of the apparent reflectance and NDI metric for foliage-wood separation.

6. Analysis of the spectral properties of SALCA dual-wavelength point clouds at stand level

To test the potential application of the SALCA dual-wavelength data for foliage-wood separation, the spectral information for the oak stand described above was analysed. In this case a 'leaf-off' scan recorded on 1st April 2014 was compared with the leaf-on scan acquired on 19th June 2014, and described in the previous section. The AR values are displayed as histograms in figure 5, for both leaf-off and leaf-on conditions, and are a function of target reflectance and the amount of material occupying the beam footprint, with partial hits accounting for high frequencies at lower values of apparent reflectance.

The 1063 nm leaf-off histogram (figure 5a) shows the wide range in apparent reflectance recorded by SALCA in this stand in the absence of green foliage. The high AR values represent full hits on objects with high reflectance. Lower AR values represent a combination of full hits on objects with lower reflectance and

partial hits for both single and multiple returns. The 1063 nm leaf-on histogram shows a higher overall frequency of hits and is a combination of both full and partial hits on foliage and wood. Although the reflectance of both components is expected to be similar, there is some evidence that the leaf material has a higher AR than the woody material with larger frequencies in the 0.3 - 0.4 region.

For the 1545 nm leaf-off histogram (figure 5b) the shape is similar to that for 1063 nm supporting the hypothesis that the reflectance properties of the woody material are similar in both wavelengths. The leaf-on 1545 nm histogram is a combination of full foliage and wood hits, single full and partial hits on foliage and wood, and partial hits from foliage and wood multiple returns. The main right-hand peak represents full hits on leaves, the right-hand tail the full hits on wood, and the left-hand peak lower AR values and partial hits on foliage and wood. There is evidence here that there is significant overlap in the spectral reflectance of leaves and wood in this stand. The normalised histograms (figure 5c and figure 5d) show the spectral separation more clearly, with two distinct peaks in the normalised frequencies for both wavelengths. Both wavelengths show a lower normalised frequency around AR values of 0.4 indicating fewer hits on wood with high reflectance values.

This visual analysis of these data illustrates the problem with using a single wavelength to separate foliage and woody material hits in a TLS point cloud. Classification using a single wavelength would produce misclassifications due to the mixing of full and partial hits. The analysis also shows that there appears to be a large range in woody reflectance in this forest stand, and a relatively narrow range in foliage reflectance. This stand includes some trees of silver birch and sweet chestnut which have higher bark reflectance than the dominant oak and this may explain this observation.

The NDI histogram for the oak-dominated stand (figure 6a) is a normalised ratio of the AR in the two wavelengths, independent of the area intercepted, allowing the data to be interpreted solely in terms of the NDI for every point in the TLS data. The leaf-off histogram is approximately normally distributed with a mean close to zero as expected. The range of NDI is very large however suggesting that the woody material may have a large positive contrast in reflectance between the two wavelengths, or a large negative contrast. The leaf-on histogram shows the higher frequency of hits expected, and a shift in the peak of the histogram to slightly higher values, although not as high as expected from the earlier discussion. The normalised frequency histogram (figure 6b) shows the same slight shift to higher values in the leaf-on data, but also shows a narrower range of NDI values compared to the leaf-off data. If the foliage and woody material had distinctly different reflectance distributions, the NDI histogram would be expected to show two distinct peaks, one near zero NDI (wood) and one in the 0.2-0.3 range (foliage). It is clear that no such pattern exists so that again simple thresholding of the NDI to separate leaves and wood in TLS point clouds in this stand is unlikely to be successful.

Similar analyses were performed on three further stands to compare leaf-off and leaf-on NDI distributions. In all three cases there was a contrast in the NDI distribution between the two dates, but no clear distinction between the NDI for foliage and woody material. These results raise questions about the accuracy of the calibration of SALCA across different measurement periods and conditions, and about the nature of the spectral variability and temporal changes in reflectance properties of woody material in forest stands.

7. Comparison of spatial and spectral classification approaches for single trees

It is apparent from the previous section that classification of forest TLS point clouds to separate foliage and wood is unlikely to be successful using dual-wavelength information alone. However it may be possible to integrate spatial information on the local-scale three-dimensional arrangements data in the point cloud data to complement the spectral data. The approach of Brodu and Lague [11] implemented in the open source software CAractérisation de NUages de POints (CANUPO) was tested here and compared with the NDI spectral classification explored earlier. The targets selected were three isolated trees of the same species where it was expected that the woody material of the main stem and branches would have similar spectral properties at the SALCA wavelengths, and that the foliage would have contrasting spectral properties. The three trees were in different states of health providing an opportunity to compare targets with different proportions of foliage and wood.

CANUPO was designed for automatic classification of point clouds, based on characterising the local-scale geometric properties of the scene around each point at multiple scales. The 'local dimensionality' characterises the distribution of points within a sphere of variable diameter, or scale of interest, around a given point as 1d, 2d or 3d. For example, at a scale of 0.5 m a tree branch appears 1d, with the points forming a linear feature, the main trunk appears 2d, with the points lying a plane, and leaves on a branch appear 3d with points distributed throughout the sphere. CANUPO analyses features in the point cloud at a range of scales and automatically finds the best combination of scales to separate two or more classes in the point cloud. The classification algorithm uses principal components to analyse the 3d arrangement of points within spheres of variable size (scale) centred on points representing a given class within the point cloud. If a single eigenvalue accounts for the total spatial variance of the points in a sphere they are distributed as 1d. If two eigenvalues are required to account for the variance they are distributed as a 2d plane, and when all three eigenvalues have the same magnitude the cloud is 3d.

The proportions of eigenvalues thus describe how the points appear at a given scale and a given location in the point cloud. The process is repeated for a second class and linear discriminant analysis is then used to define a plane of maximum separability between the two classes. In practical terms the implementation of CANUPO requires the user to interactively identify a set, or sets, of points belonging to each class. These

are then used to train the classifier before applying the classification to the full point cloud. The user also defines the range of scales applied as part of the training process based on subjective assessment of the typical size of objects in the scene.

Full resolution SALCA point clouds of three isolated oak trees in different condition were acquired at a site near Silverdale (54.1766° N, 2.8151° W) in Lancashire, UK, in July and August 2015 (figure 7). The trees are referred to as being in 'Good', 'Moderate' and 'Poor' condition, based on a visual assessment of their health. The Good tree was 17 m high and had a crown in full leaf with no visible signs of damage or stress. The Moderate tree was 16 m high, had a hollow tree stem at the base, and lower leaf cover overall. The Poor tree was 14 m high, had damaged bark in the lower part of the main stem, and a small number of leaves on only a few branches. Four scan positions (north, south, east and west) located 20 m from the main stem were used for each tree although only the scans recorded from a position south of the tree are used here. On the two days of data collection the weather was sunny and the sky clear, and the wind speed was very low with no significant movement of the tree canopies observed. Photographs of the three trees taken from the scanner viewpoint provided a basic means of visually validating the classification outputs.

After pre-processing, the three point clouds were classified into 'wood' and 'foliage', using two separate approaches based on spatial (CANUPO) and spectral (NDI) methods, and using the 1545 nm data. The CANUPO classification process requires interactive training to define polygons that characterise each of the required classes. For the wood class this normally mean polygons that capture the main stem and areas of smaller branches devoid of leaves, and for the foliage class smaller polygons approximately 0.5 – 1.0 m in diameter, representing leafy parts of the respective point clouds. The NDI was computed for each point cloud as described earlier, and a threshold applied to classify the points as foliage or wood. The threshold was defined by iterative visual interpretation and comparison with the photographs of the trees. Three different thresholds were required -0.1, -0.1 and 0.2 for the Good, Moderate and Poor trees respectively.

In order to compare the classifications for the three trees, a point-by-point classification matrix was produced with four classes: spatial classifier wood (Sw), spatial classifier foliage (Sf), NDI classifier wood (Nw) and NDI classifier foliage (Nf). The matrices allowed the frequency of points where the spatial classifier and NDI classifier agreed (overall classification agreement), and the points where they disagreed, and the spatial distribution of that disagreement. In making this comparison no assumptions were made about which points were 'correct' or which were 'incorrect' in terms of the foliage-wood classification; the analysis was designed to help understand the different types of information in the SALCA data.

The results (figure 7) show that both classification approaches produced visually plausible separation of foliage and wood. For the Good tree the spatial classifier showed a scattered distribution of foliage in the crown and erroneous classification as foliage of the small branches with no foliage. The appearance of the NDI classifier was similar although the smaller branches are more clearly defined in general. The dense

clumps of leaves were clear in the Moderate condition tree in both classifiers although again there was clearer discrimination of the smaller branches in the canopy in the NDI classifier. Points classified by the NDI as leaves appear at the bottom of the main stem of the Moderate tree suggesting misclassification in this area. The Poor condition tree had very few leaves but both classifiers appeared to pick out the main foliage clumps. The spatial classifier misclassified the finer branches as leaves, and the NDI classifier misclassified some points on the lower part of the stem.

The two classifications for the three trees showed variation in total agreement of 62, 58, and 85% for the Good, Moderate and Poor tree respectively. Visualisation of the cross-comparison (figure 8) showed the areas where the classifiers agreed: dark green where both classify the points as wood, and red where both classify the points as leaves. There was generally good agreement for the main stem and some agreement for the leaves in the canopies, but for the Good and Moderate conditions trees, there were large numbers of points of disagreement in the canopy. It is apparent that the spectral classifier, based on thresholding the NDI, produces finer detail in the output data most likely because it treats each point separately, whereas the spatial classifier uses a probabilistic decision rule based on groups of points.

These results show the contrasting information content of the TLS point clouds. The NDI classifier is applied on a point-by-point basis and uses no other information; the spatial classifier uses only the X,Y,Z coordinates of all returns and classifies points based on a local-scale analysis of the 3d point distribution. The variation in the results for a single species where the spectral properties of foliage and wood are assumed to be similar shows that there is different information in in the two approaches. Future research should examine and exploit the full range of spatial, spectral and waveform-related information in TLS point clouds for 3d mapping of forest structure.

8. Discussion and conclusions

Novel TLS instruments like SALCA have the potential to provide new approaches for assessing foliage distributions and dynamics in forest ecosystems. The dual-wavelength measurements add additional information beyond that from the traditional point-based analyses used with most laser scanner data. Access to the waveform data allows other metrics like waveform width and return number to be derived, although this type of information has yet to be exploited in forestry applications. This research has shown that a simple threshold-based classification of foliage and woody material in mixed-species forest stands is unlikely to be accurate because of the large variation in the spectral properties of the material in the scene including the main stems, branches, leaves and finer woody material. Comparison of leaf-off and leaf-on measurements depends on reliable calibration of the instrument, and an assumption that there is no time-related change in the reflectance of the woody material. Both these aspects require further field-based investigations along with quantitative analysis of the spectral data, including analysis of the errors in reflectance calibration and NDI.

Spatial classifiers like CANUPO provide a powerful means of point cloud classification. However they rely on the identification of homogeneous distributions of points within a sphere, so that areas within a tree canopy that consist of a mix of leaves (2d) and fine branches (1d), at scales of around 0.5 m are likely to be misclassified. It is also likely that in a mixed species stand with different leaf sizes and leaf arrangements, a simple binary foliage/wood classifier will be unsuccessful. A species-specific training approach for spatial classification may be required in this case. Further work is necessary to investigate how spatial classification is affected by the wide variation in range measurements in TLS scans where, in a high resolution scan, leaves at close range may appear as a plane, but at far range may appear as a 3d cloud.

Future multi-wavelength full-waveform TLS will provide a wide range of new information for forest ecology. There are significant challenges associated with data pre-processing, calibration and information extraction that can now be explored with instruments like SALCA. These instruments provide information-rich point clouds that have yet to be fully exploited for forest applications. Classification methods that combine both spectral and spatial information are likely to provide a way forward, but there is still a clear need for basic field data collection on the spectral and spatial characteristics of foliage and wood in a range of forest environments to inform these research developments.

Data, code and materials

Data and code used in this research are available from the authors on request.

Competing interests

We have no competing interests

Authors' contributions

All authors were involved in field data collection, data processing, data analysis, and drafting the manuscript. All authors gave final approval for publication.

Acknowledgements

We would like to acknowledge Jon Murray from Lancaster University for arranging access to the Silverdale field site and for assistance with data collection, and the Forestry Commission for access to Delamere Forest. Dr Rachel Gaulton, Dr Alberto Ramirez and Dr Steven Hancock contributed significantly to SALCA algorithm developments.

Funding

This work was supported research grants to FMD from the UK Natural Environment Research Council (NE/H002685/1 and NE/I01702X/1). FMD was supported by a Royal Society Leverhulme Trust Senior

Research Fellowship (2013–14). FS is undertaking a PhD supported by the Libyan Government. LS was supported by a UK Natural Environment Research Council PhD studentship (2012-2016) at the University of Salford.

References

1. Strahler, A.H., D.L.B. Jupp, C.E. Woodcock, C.B. Schaaf, T. Yao, F. Zhao, X. Yang, J. Lovell, D. Culvenor, G. Newnham, W. Ni-Miester, and W. Boykin-Morris (2008). Retrieval of Forest Structural Parameters using a Ground-Based Lidar Instrument (Echidna®), Canadian Journal of Remote Sensing, 34: S426-S440.
   doi: 10.5589/m08-046
2. Jupp, D.L.B., Culvenor, D.S., Lovell, J.L., Newnham, G.J., Strahler, A.H., Woodcock, C.E. (2009) Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna'. Tree Physiology, 29, 171-181
   doi: 10.1093/treephys/tpn022
3. Douglas, E. S., Martel, J., Howe, G., Hewawasam, K., Marshall, R. A., Schaaf, C. L., Cook, T. A., Newnham, G. J., Strahler, A., and Chakrabarti, S., 2015, Finding leaves in the forest: The Dual-Wavelength Echidna Lidar, IEEE Geoscience and Remote Sensing Letters, 12, 776-780.
   doi: 10.1109/LGRS.2014.2361812
4. Danson, F.M., Gaulton, R., Armitage, R.P., Disney, M.I., Gunawan, O., Lewis, P., Pearson, G., Ramirez, A.F., 2014. Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure. Agricultural and Forest Meteorology, 198, 7-14.
   doi: 10.1016/j.agrformet.2014.07.007
5. Li, Z., Schaefer, M. Strahler, A.H., Schaaf, C. and Jupp, D. (2018) On the utilisation of novel spectral laser scanning for three-dimensional classification of vegetation elements (this issue)
6. Béland, M., Widlowski, J., Fournier, R. A., Côté, J. and Verstraete, M. M. (2011). Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements. Agricultural and Forest Meteorology, 151, 1252–1266.
   doi: 10.1016/j.agrformet.2011.05.004
7. Clawges, R., Vierling, L., Calhoon, M., & Toomey, M. (2007). Use of a ground-based scanning lidar for estimation of biophysical properties of western larch (Larix occidentalis). International Journal of Remote Sensing, 28, 4331-4344.
   doi: 10.1080/01431160701243460

8. Seielstad, C., Stonesifer, C., Rowell, E. and Queen, L. (2011). Deriving fuel mass by size class in douglas-fir (Pseudotsuga menziesii) using terrestrial laser scanning. Remote Sensing, 3, 1691-1709. doi: 10.3390/rs3081691

9. Yanez, L., Homolova, L., Malenovky, Z. and Schaepman, M. E. (2008). Geometrical and structural parameterization of forest canopy radiatove transfer by LIDAR measurements. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, B7, p45-50. ISPRS, Beijing, China, 3-11 July 2008.

10. Béland, M., Baldocchi, D. D., Widlowski, J., Fournier, R. A. and Verstraete, M. M. (2014). 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. doi: https://doi.org/10.1016/j.agrformet.2013.09.005

11. Brodu, N. and Lague, D. (2012). 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. ISPRS Journal of Photogrammetry and Remote Sensing, 68, 121-134. doi: 10.1016/j.isprsjprs.2012.01.006

12. Li, Z., Strahler, A., Schaaf, C., Jupp, D., Howe, G., Hewawasam, K., Chakrabarti, S., Cook, T., Paynter, I., Saenz, E., Yang, X. and Yao, T. (2015). Improving canopy vertical structure measurements with dual-wavelength laser scanning. Paper presented at AGU, San Francisco, 14-18 December 2015.

13. Tao, S., Guo, Q., Xu, S., Su, Y., Li, Y. and Wu, F. (2015). A geometric method for wood-leaf separation using terrestrial and simulated lidar data. Photogrammetric Engineering and Remote Sensing, 91, 767-776. doi: 10.14358/PERS.81.10.767

14. Yang, X.Y., A.H. Strahler, C.B. Schaaf, D.L.B. Jupp, T. Yao, F. Zhao, Z.S. Wang, D.S. Culvenor, G.J. Newnham, J.L. Lovell, R. Dubayah, C.E. Woodcock, and W. Ni-Meister. (2013). Three- dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform Lidar instrument (Echidna (R)), Remote Sensing of Environment, 135, 36–51. doi: 10.1016/j.rse.2013.03.020

15. Fieber, K. D., Davenport, I. J., Ferryman, J. M., Gurney, R. J., Walker, J. P. and Hacker, J. M. (2013). Analysis of full-waveform LiDAR data for classification of an orange orchard scene. ISPRS Journal of Photogrammetry and Remote Sensing, 82, 63-82. doi: 10.1016/j.isprsjprs.2013.05.002

16. Akerblom, M., Raumonen, P., Casella, E., Disney, M., Danson, F.M., Gaulton, R., Schofield, L.A, Kaasalainen, M. (2018) Non-intersecting leaf insertion algorithm for tree structure models (this issue).

17. Eitel, J.U.H., Vierling, L.A., Long, D.S. (2010) Simultaneous measurements of plant structure and chlorophyll content in broadleaf saplings with a terrestrial laser scanner. Remote Sensing of

Environment, 114, 2229–2237.
doi:10.1016/j.rse.2010.04.025.

18. Hakala, T., Nevalainen, O., Kaasalainen, S., Mäkipää, R. (2015) Technical note: multispectral lidar times series of a pine canopy chlorophyll content. Biogeosciences, 12, 1629-1634.
doi: 10.5194/bg-12-1629-2015
19. Gaulton, R., Danson, F. M., Ramirez, F. A. and Gunawan, O. (2013). The potential of dual-wavelength laser scanning for estimating vegetation moisture content. Remote Sensing of Environment, 132, 32-39.
doi: 10.1016/j.rse.2013.01.001
20. Hancock, S., Gaulton, R. and Danson, F.M. (2017). Angular reflectance of leaves with a dual-wavelength terrestrial lidar and its implications for leaf-bark separation and leaf moisture estimation. IEEE Transactions on Geoscience and Remote Sensing, 55, 3084-3090 .
doi: 10.1109/TGRS.2017.2652140
21. Zhu, X., Wang, T., Skidmore, A.K., Darvishzadeh, R., Olaf Niemann, K., Liu, J. (2017). Canopy leaf water content estimated using terrestrial LiDAR, Agricultural and Forest Meteorology, 232, 152-162.
doi: 10.1016/j.agrformet.2016.08.016
22. Junttila, S., Vastaranta, M. Liang, X., Kaartinen, H., Kukko, A., Kssalainen, S., Holopainen, M., Hyyppä, H, Hyyppä, J. (2017) Measuring leaf water content with dual-wavelength intensity data from terrestrial laser scanners. Remote Sensing, 9, 8.
doi: 10.3390/rs9010008
23. Wagner, W., Ullrich, A, Ducic, V., Melzer, T, Studnicka, A. (2006) Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. Journal of Photogrammetry and Remote Sensing, 60, 100-112.
doi: 10.1016/j.isprsjprs.2005.12.001
24. Schofield, L. A., Danson, F. M., Entwistle, N. S., Gaulton, R., & Hancock, S. (2016). Radiometric calibration of a dual-wavelength terrestrial laser scanner using neural networks. Remote Sensing Letters, 7, 299-308.
doi: 10.1080/2150704X.2015.1134843
25. Hancock, S., Armston, J., Li, Z., Gaulton, R., Lewis, P., Disney, M., Danson, F.M., Strahler, A.H. Schaaf, C.B., Anderson, K., Gaston, K.J. (2015). Waveform lidar over vegetation: An evaluation of inversion methods for estimating return energy. Remote Sensing of Environment, 164: 208-224.
doi: 10.1016/j.rse.2015.04.013
26. Welford, D., A. Mooradian. (1982). Output power and temperature dependence of the linewidth of single-frequency CW (Gaai) as Diode Lasers. Applied Physics Letters 40 (10): 865–867.
doi: 10.1063/1.92945

List of tables and figures

Table 1. Return frequencies for leaf-off and leaf-on scans for Oak plot at Delamere Forest

Condition	Wavelength (nm)	Total number of returns	Single returns	Multiple returns			
				1	2	3+	Max return no
Leaf-off 1st April 2014	1063	6088572	5242396 (86%)	846176	846176	50229	21
	1545	7189120	5056238 (70%)	1027504	1027504	77874	17
Leaf-on 19th June 2014	1063	10159879	7198529 (71%)	1436443	1436443	88464	7
	1545	10827134	6770167 (62%)	1943573	1943573	169821	7

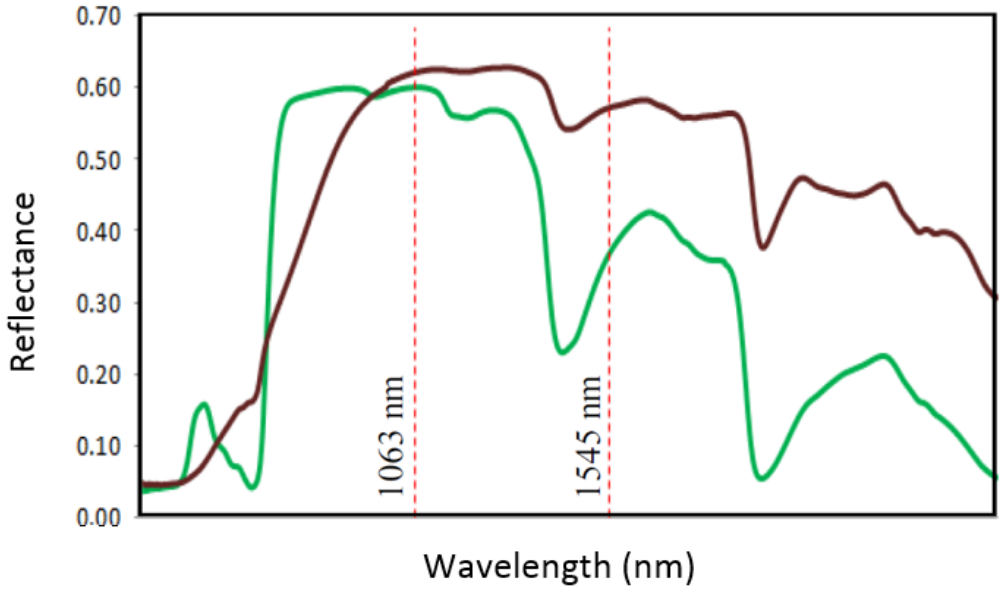


Figure 1. Spectral reflectance of an oak leaf (green) and oak bark (brown) showing the wavelengths measured by SALCA.

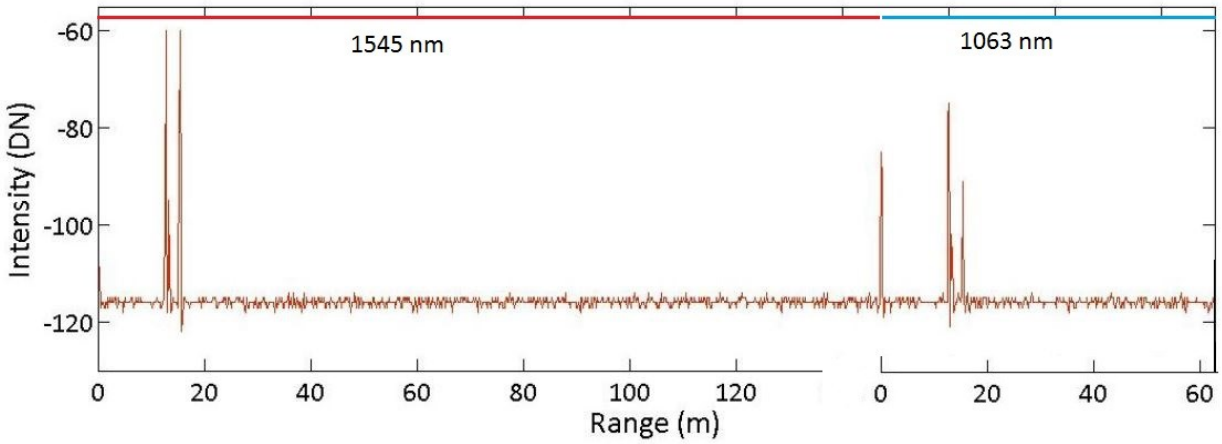


Figure 2. A single dual-wavelength waveform from SALCA. In the figure time is expressed as the relative two-way distance (meters) travelled by a pulse from the detection of the outgoing 1545 nm and 1063 pulses at zero range.

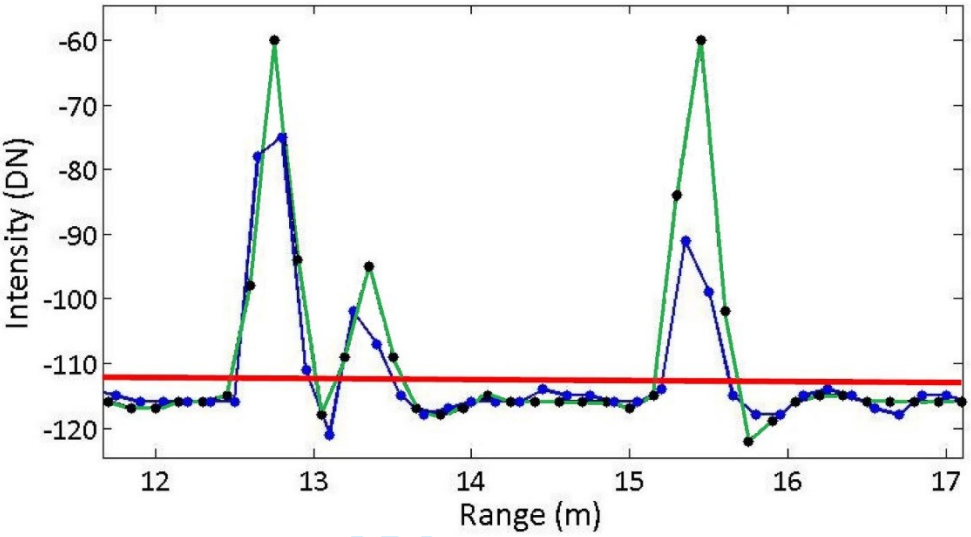
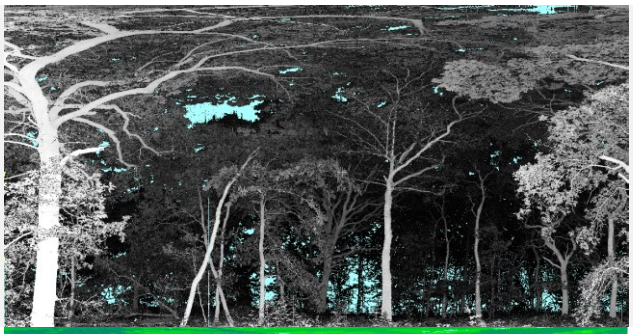
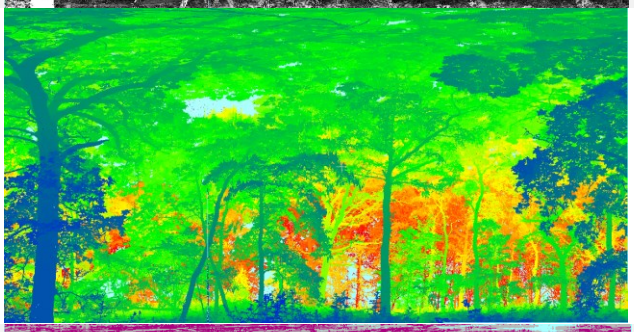


Figure 3. Dual-wavelength data from Figure 2 combined. Green line is 1545 nm data and blue line is 1063 nm data. Points show discrete samples of backscattered energy. Red line is noise threshold applied.



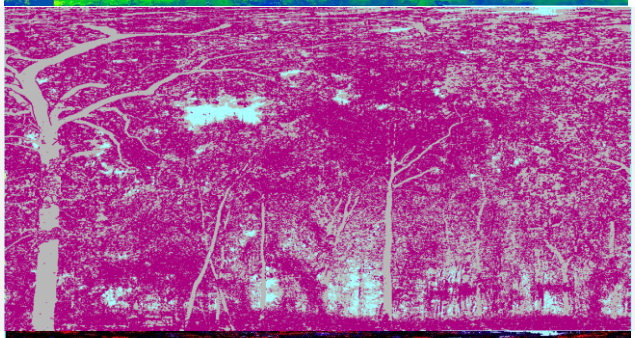
Intensity
Scale range 0 – 500 DN

Pale blue area are canopy gaps and areas beyond 60 m range



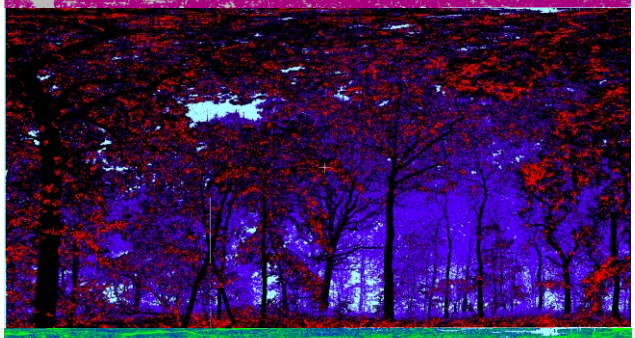
Range
Scale range 1 – 60 m

Data recorded up to a maximum range of 60 m



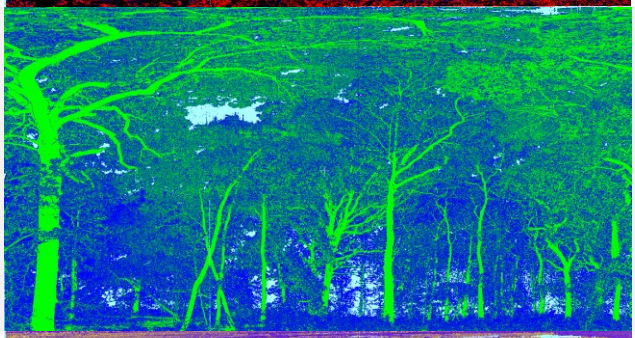
Single returns vs multiple returns

Single: grey – panel, trunks, branches
2+ returns: purple - leafy areas and partial hits



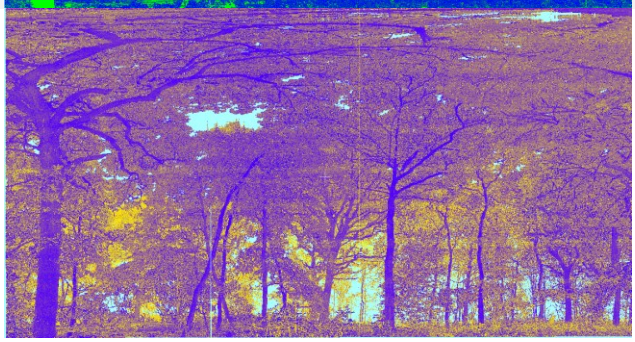
Return width in bins

1-2: blue – mostly leafy areas
3-4: black – mostly wood and some close-range leaves
5-12: red – leafy areas (size of this class enlarged for clarity)



Apparent reflectance
Scale range 0 – 1.0

Main stems and branches have similar apparent reflectance



Normalised difference index (NDI)
Scale range -1 to +1

Low NDI for stems, moderate NDI for leafy areas

Figure 4. Visualisation of variables from SALCA for Oak plot at Delamere Forest, UK, recorded 19th June 2014. The top five images were generated using data from the 1545 nm wavelength.

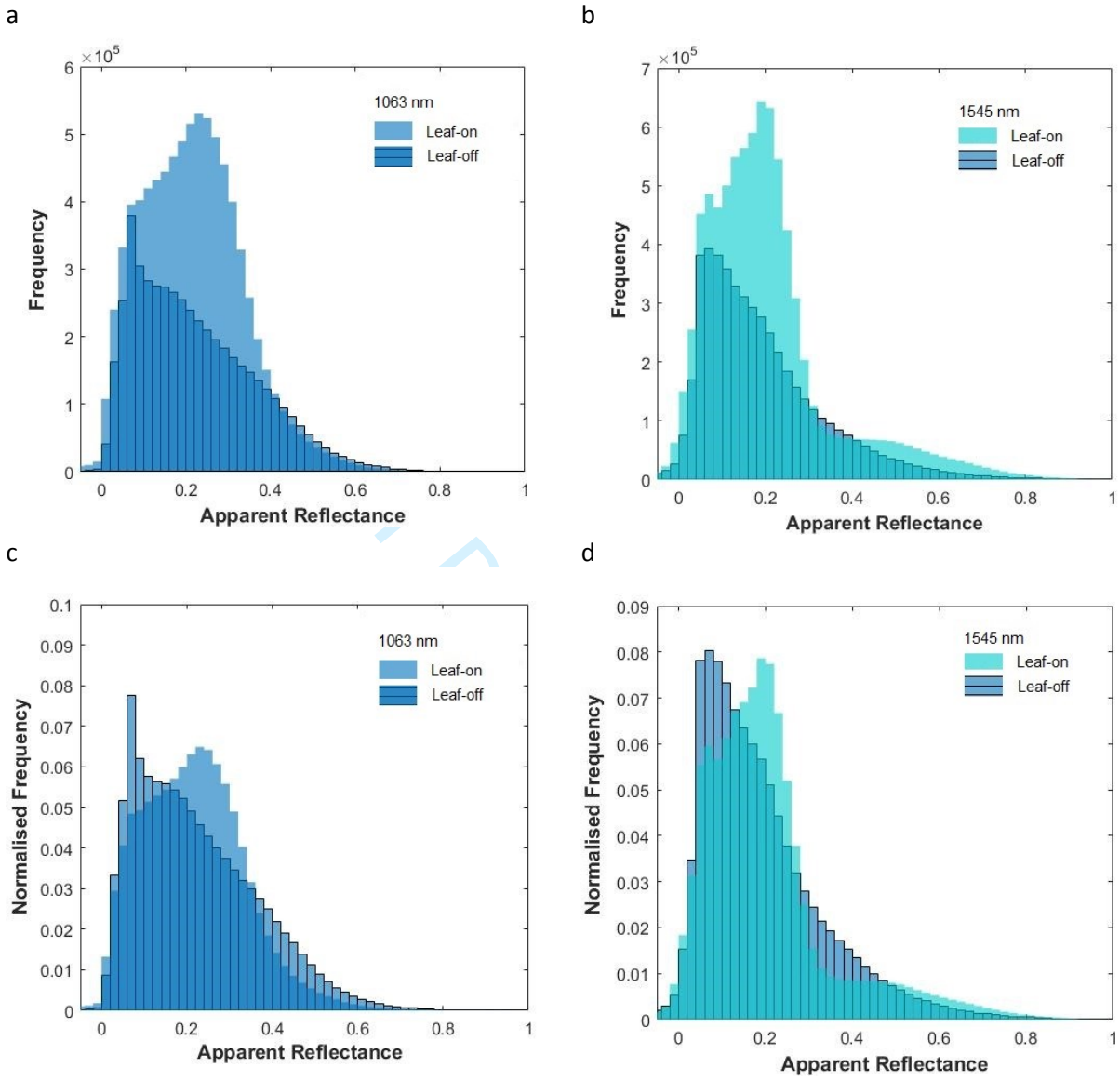
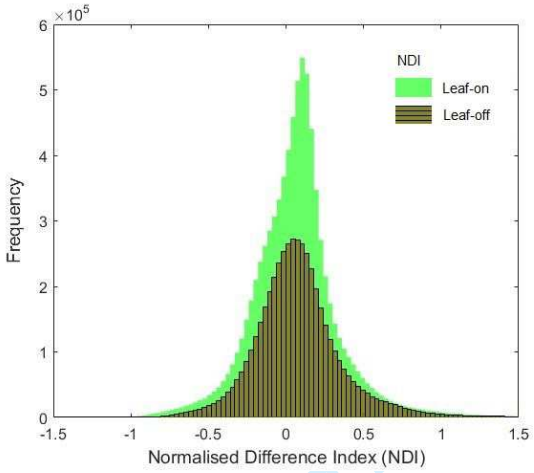


Figure 5. Frequency histograms for all returns in Oak plot at Delamere Forest, U.K.

a



b

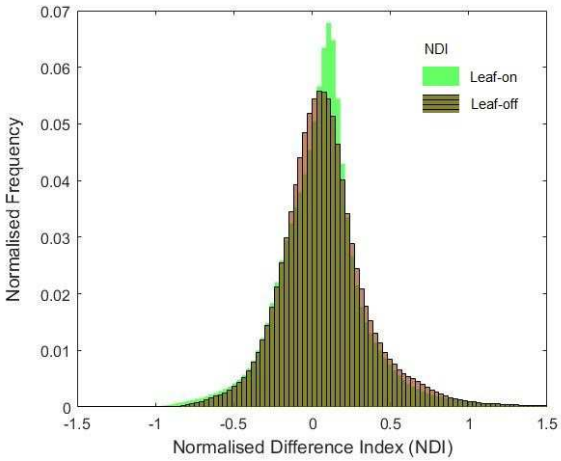


Figure 6. NDI frequency histograms for Oak plot at Delamere Forest, U.K.

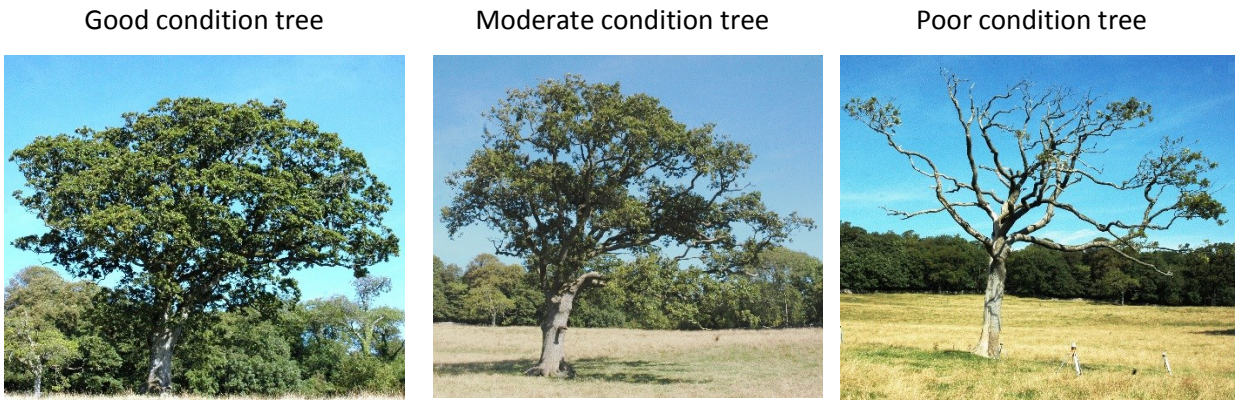


Figure 7. Comparison of spatial and spectral classifiers for three oak trees at Silverdale, UK

Good tree: classification agreement = 62%

Moderate tree: classification agreement = 58%

Poor tree: classification agreement = 82%

Figure 8. Four class classification of three oak trees at Silverdale, UK: Sw-Nw (dark green), Sf-Nf (red), Sw-Nf (yellow) Sf-Nw (light green).

Spectral and spatial information from a novel dual-wavelength full-waveform terrestrial laser scanner for forest ecology

RSFS-2017-0049.R1

Responses to Referee 2

Changes are highlighted in yellow in the main document

1. About figure 5 and figure 6, another reviewer and I both raised the concern over the normalization of the frequency for better comparison of distributions. I concur with the authors' argument on the need to demonstrate the different number of returns/points between leaf-on and leaf-off scans. However, these frequency-based histograms hamper the comparison of distributions of leaves versus woody materials which is a key subject of this paper. I still strongly suggest the authors to add normalized histogram, i.e. distribution density histograms in figure 5 and figure 6 (rather than change the frequency-based).

>>*Normalised histograms have now been added and a short additional interpretation added to the text*

2. Figure 2 now becomes a bit more confusing. Reading the texts in line 10-28 on page 6, it looks like in figure 2, the long tail of 1545nm waveform is plot there only to illustrate the noise in the waveforms. The pulse at the "0" under the 1063nm blue line is the outgoing pulse at 1063 nm? Then where is the outgoing pulse at 1545nm? Can you also put the outgoing pulse at 1545nm there? It's even better to place text boxes explaining what each pulse is for an easier read. And add some simplified version of main text to the figure caption to facilitate the interpretation of the figure "The 1 Ghz sampling rate of 1 ns leads to a digitized sample equivalent to a measurement every 150 mm in range, with 700 measurements for the 1545 wavelength, followed by a maximum of 700 measurements for the 1063 nm wavelength"

>>*Additional text has been added to further clarify these points.*

3. Page 7, line 60 to Page 8 line 2, the latter half of the first sentence in this paragraph, i.e. "some are also visible at longer ranges for leafy area where the backscattered energy is so low that only one return can be detected.", it seems to connect better with the next sentence "Single returns also include some single partial hits." because that is the reason why single returns could also be from single partial hits. Prob. consider restructure the words in this part.

>>*Two sentences have been modified to further clarify this section*

4. Page 8, line 7 – 11, about the meaning of different pulse widths. You have better explanation in the response (referee 2, number 5), esp. about why the narrowest ones actually indicate returns from leaves due to fewer bins when leaf-return signals drop under the noise thresholds. In the main text, you miss this explanation. Please consider adding it there.

>>*Additional text clarifies this point*

5. Page 8, line 11 – 14, prob. better to highlight the more pronounced contrast between leaves and wood in the AR image than the intensity image to show why you use AR rather than intensity directly.

>>We feel that the text does highlight the contrast in the AR images. No action taken

6. Page 9, line 19 – 36, these texts on the interpretation of NDI image sound pessimistic and depreciate the value of the SALCA instrument and its concept. Though the leaf on NDI histogram is not as dual-peaked as expected, the skew towards the right does suggest the impact of leafy presence on the NDI distribution and hence its contribution to the discrimination between leaves and woody materials.

>>We feel that we have reported the results in a clear way and note that the results may be specific to one type of forest stand. We do not see the text as pessimistic. No action taken

7. Page 11, line 11 -12, "The NDI was computed for each point cloud as described earlier, and a single AR threshold applied to classify the points as foliage or wood.", this reads like you did one classification using NDI and the other classification using 1545-nm AR? But I don't think that's the case.

>>"AR" deleted – this was a typographical error

8. Figure 8, it would be better to have a matrix of agreement of leaf and woody points, and also disagreement of the two, rather than just a simple overall agreement.

>>We provide a visualisation of the misclassifications in the figure (ie. these are the data from the matrix of agreement) and do not want to add complexity to the paper by adding a further table here. No action taken

9. Section "4. SALCA data characteristics and data pre-processing" can be reduced and simplified as the previous papers on SALCA have introduced more details on the instrument and preprocessing.

>>This section complements and adds new material to the previous papers and without it the reader may find the rest of the paper difficult to follow. No action taken

10. Page 2, line 52, "in order to measure forest plant area index (LAI)", shouldn't be PAI? Or better spell out directly leaf area index or explain the difference.

>>Corrected.