Is waveform worth it? A comparison of LiDAR approaches for vegetation and landscape characterisation.

Karen Anderson[1*], Steven Hancock[1], Mathias Disney[2] and Kevin J. Gaston[1]

* Corresponding author: Karen.Anderson@exeter.ac.uk

[1] Environment and Sustainability Institute, University of Exeter, Penryn Campus, TR10 9FE, UK

[2] University College London, Department of Geography, Gower Street, London, WC1E 6BT and National Centre for Earth Observation (NCEO)

Keywords: LiDAR, vegetation, waveform, structure, validation, urban

Running title: Waveform LiDAR considerations in ecology

Abstract

Light Detection And Ranging (LiDAR) systems are frequently used in ecological studies to measure vegetation canopy structure. Waveform LiDAR systems offer new capabilities for vegetation modelling by measuring the time-varying signal of the laser pulse as it illuminates different elements of the canopy, providing an opportunity to describe the 3D structure of vegetation canopies more fully. This paper provides a comparison between waveform airborne laser scanning (ALS) data and discrete return ALS data using terrestrial laser scanning (TLS) data as an independent validation. With reference to two urban landscape typologies we demonstrate that discrete return ALS data provided more biased and less consistent measurements of woodland canopy height (in a 100% tree covered plot, height underestimation bias = 0.82 m; SD = 1.78m) than waveform ALS data (height overestimation bias = -0.65 m; SD = 1.45 m). The same biases were found in suburban data (in a plot consisting of 100% hard targets e.g. roads and pavements), but discrete return ALS were more consistent here than waveform data (SD = 0.57 m compared to waveform SD = 0.76 m). Discrete return ALS data performed poorly in describing the canopy understorey compared to waveform data. Results also highlighted errors in discrete return ALS intensity, which were not present with waveform data. Waveform ALS data therefore offer an improved method for measuring the three dimensional structure of vegetation systems but carry a higher data processing cost. New toolkits for analysing waveform data will expedite future analysis and allow ecologists to exploit the information content of waveform LiDAR.

1. Introduction

The spatial and volumetric structure of vegetation in ecosystems is a key driver of function (Shugart et al., 2010) and Light Detection And Ranging (LiDAR) instruments provide critical data for describing and modelling vegetation structure (Vierling et al., 2008). LiDAR instruments can be operated from the ground (e.g. Terrestrial Laser Scanning; TLS) from airborne platforms (e.g. Airborne Laser Scanning; ALS) or from satellites (e.g. freely available data from ICESat (Harding and Carabajal, 2005)), and come in two forms – discrete return and full waveform systems (Lefsky et al., 2002, Vierling et al., 2008). The difference between these is the way in which data are recorded. Discrete return systems (most commonly used) measure the time taken for a laser pulse to travel to an object and are used to determine height. In products derived from ALS data there are usually two datasets: a digital surface model (DSM) provides an estimate of the top-of-canopy height whilst the digital terrain model (DTM) shows topographic variability in the neighbouring ground surface. Such data can be used to describe canopy patterns (Anderson et al., 2010, Luscombe et al., 2014), model hydrological flowpaths (Jones et al., 2014), monitor wildlife habitat (Hyde et al., 2006), or produce carbon inventories at patch (Calders et al., 2015) or landscape (Asner et al., 2011) scales. Waveform ALS data (Figure 1), however, have the potential to provide much richer spatial information about canopy characteristics in three dimensions. This is because these systems record the range to multiple targets within the canopy (Danson et al., 2014). By measuring the time-varying signal of the laser pulse as it illuminates different elements of the canopy, these systems can be used to model the spatial character and arrangement of structures that drive canopy biophysical processes such as canopy architecture and size and woody biomass (Armston et al., 2013, Mallet and Bretar, 2009) and can provide useful data for studies requiring tree species discrimination (Alonzo et al., 2014).

It is only since around 2010 that waveform systems have begun to be heavily explored in ecological contexts (with limited earlier examples by Anderson et al. (2006), and Hyde et al. (2005), for example). This is probably because of the high data volumes requiring high computing power, and the complexity of analysing the return signal (e.g. rather than a few 'hits' (typically, up to five) from a discrete return system, waveform systems give a near-continuous pulse; Figure 1). Waveform data represent a significant signal processing task - tracing the photon from the sensor to the ground and understanding what the interactions represent is a potential barrier to their application in ecology and beyond. Extracting 3D canopy information from the waveform is challenging because the pulse can be perturbed on its path through the canopy – e.g. the electromagnetic radiation in the pulse can be redirected within the canopy and is known to suffer 'multiple scattering' between different elements (e.g. leaves and woody biomass). This leads to highly complex signals requiring de-noising and correction using signal processing approaches, followed by product validation. Despite this challenge there are a variety of new waveform signal processing approaches emerging, particularly for vegetation applications, with most studies following one of three methods:

1) Decomposition into points and attributes using function fitting (Hofton et al., 2000, Wagner et al., 2008);

2) Decomposition into points using deconvolution (Jiaying et al., 2011, Roncat et al., 2011, Hancock et al., 2008);

3) Extracting metrics such as height of median energy (Drake et al., 2002).

The points or metrics from the resulting models can then be used to infer plot-level characteristics or calculate canopy height (Boudreau et al., 2008), fit geometric primitives to crowns (Lindberg et al., 2012); or fill voxels to enable construction of 3-dimensional models from a regular grid of cubes (e.g. as in Minecraft) where canopy structure can be optimally modelled (Hosoi et al., 2013).

Waveform laser scanning technology is now at a tipping point, evidenced by NASA's forthcoming 'Global Ecosystem Dynamics Investigation LiDAR' space mission, due for launch in 2018 (GEDI (NASA, 2014, Krainak et al., 2012)). It is hoped that the enhanced capability of the waveform system on GEDI will provide superior global estimates of vegetation carbon stocks.

In this paper we address the pragmatic research question of what benefits waveform ALS data can offer ecologists over more easily obtainable discrete return ALS products, using urban systems as an exemplar. Quantitative description of the pattern and 3D structure of urban vegetation demands fine-scale spatially-distributed information describing canopy architecture (Yan et al., 2015). This is because the pattern and extent of green infrastructure (e.g. street trees, parks, domestic yards and gardens) is a key determinant of the provision of ecosystem services in cities and towns, including nutrient cycling, temperature and flood risk regulation, reduction in atmospheric pollution, aesthetics, and multiple dimensions of human health (Gaston et al., 2013). Most examples of remote sensing approaches for mapping urban greenspace rely on either optical classification of aerial photographs, or height-based classification of discrete return ALS to determine the spatial distribution of basic classes such as trees, bushes and grass (Yan et al., 2015, Chen et al., 2014). Whilst these data are appropriate to the particular scale range of the texture of urban vegetation variance, and allow the small patch sizes of urban greenspace to be mapped (e.g. in yards and gardens) they neglect to characterise the important vertical distribution of vegetation and photosynthetic material through the depth of the canopy and its spatial form. Furthermore they cannot account for important habitat features such as the understorey which are important in driving urban ecological connectivity. This work sought to establish the impact of those omissions in describing urban vegetation complexity.

Here, we compare a simply processed waveform ALS product with discrete return ALS data from the perspective of ecologists working in urban environments. We validate the findings using a ground-based TLS survey, quantify differences in each approach and evaluate the relative processing costs of each. Finally, we discuss the wider implications for using waveform ALS data for vegetation monitoring in other ecological settings.

2. Materials and Methods

2.1 ALS survey data

An ALS survey was carried out over the town of Luton, UK on 5 and 6 September 2012 (Figure 2) when the urban vegetation was in full leaf-on stage. The survey utilised the UK Natural Environment Research Council (NERC) Airborne Research and Survey Facility (ARSF) Dornier 228 aircraft platform and the Leica ALS50-II ALS system with a WDM65 full waveform digitiser, measuring at 1064 nm. Geo-registration of the scans was achieved using differential global positioning system (GPS) data from the aircraft and at a linked GPS ground-station. All ALS data were collected by a single instrument with separate discrete return and waveform output streams. The footprint density of ALS data (waveform and discrete return data) were collected with a density of between one point per 25 cm^2 and one point per 4 m^2 – this variability is normal and is dependent on scan angle and overlap between flight lines. The discrete return ALS data had up to four returns per pulse. Raw ALS data were processed into a geolocated point cloud with associated waveforms using Leica ALSPP software (version 2.75). More detailed documentation about the data processing can be found online (NERC ARSF, 2014a, NERC ARSF, 2014b)

Two data products from the ALS survey were compared: a discrete return ALS point cloud describing *x,y,z* spot heights and intensity; and a waveform ALS dataset which required pre-processing before it could be used.

2.2 Field site description

Data from two field validation sites (both within an area of Luton, UK, called Little Bramingham Woods) are presented in this manuscript (Figure 2). The first site was in an area of dense and varied tree cover with a clear understorey (referred to as the ‘woodland’ site) and the second was from a residential area (referred to as the ‘suburban’ site). A very simple 2 m resolution land cover map (LCM) was generated for these sites using data from an airborne hyperspectral survey (with the AISA Eagle 12 bit pushbroom scanner) carried out at the same time as the ALS survey. The LCM was generated by applying an unsupervised classification algorithm to discrete return ALS data and a Normalised Difference Vegetation Index (NDVI) product. The NDVI was calculated using equation 1 where ρ_{vis} was the mean visible reflectance in channels from 500 nm to 680 nm and ρ_{nir} was the mean infrared reflectance between 761 nm and 961 nm.

$$NDVI = \frac{\rho_{nir} - \rho vis}{\rho_{nir} + \rho vis} \qquad (1)$$

A 70 cm threshold for discriminating tall from short vegetation and an NDVI threshold of 0.2 for discriminating vegetated from non-vegetated areas was used. In the woodland area, the LCM showed that the majority of the site was covered by tall vegetation. In the suburban area, as was expected, there was a mix of tall and short vegetation and vegetated and non-vegetated areas. For both woodland and suburban sites the discrete return and waveform ALS data were extracted for a 20 m by 20 m square at the centre of each TLS ground-validation site for comparison. These comparison areas were chosen because they were proximal to sampling sites where complementary ecological data were being collected – specifically bird feeders where population counts were being collected and where flows of biodiversity through urban systems were being measured. These sites were also evaluated in the waveform LiDAR datasets prior to collection of the TLS validation data, and were found to be representative areas with a variety of waveform shapes and widths.

2.3 Method for processing waveform ALS data

The ALS50-II system recorded the intensity of reflected light as an eight-bit value every 1 nanosecond. The first step in signal processing the waveform data was to remove background electronic noise – which is known to be very stable in the Leica ALS50-II (Hancock et al., 2015). Here we used a simple method to extract canopy signals from the waveform ALS data. The first peak in the waveform above the noise threshold was traced back to the mean noise level (DN=12, derived from a histogram) to provide a consistent estimate of the canopy maxima. The histograms of signal intensity from Hancock et al. (2015) were then used to set the simple noise threshold at DN=16 (see Hancock et al. (2015) Figure 5b) to remove all background noise, and the result was a product showing point height information that could be used to compare datasets quantitatively. Further processing - for example using function fitting, deconvolution or pulse width subtraction may have further improved the retrieval of the ‘true’ canopy top (Hofton et al., 2000). These more complex signal processing methods were not the focus of this paper and will be discussed in a subsequent manuscript which develops a validated voxel-based approach for 3-D canopy description in urban settings.

2.4 Validation data from TLS survey

To validate the information content of the two ALS products, a waveform TLS system was deployed (Riegl VZ-400, operating at 1545 nm (near infra-red)) to measure vegetation structure (from the ground up) on 5 and 7 August 2014. The TLS instrument had a reported 5 mm accuracy and 3 mm repeatability which was far greater than the ALS data. Previous work by Calders et al. (2015) has shown how this approach provides a good validation where accurate tree heights could be obtained, and demonstrating that attenuation was not significant. The dates of field sampling with TLS were chosen to ensure that the vegetation was in a similar state to the time of the ALS survey. Validation sites were chosen to cover a range of observed habitat structures, and a variety of ALS waveform shapes and urban typologies. As a result the TLS scan methodology had to be adapted for each site so as to capture the variability in canopy structure appropriately. The plot sizes also varied, with small (5 m) plots sometimes requiring three scan positions to capture variability in the dense vegetation whilst sparsely vegetated plots measuring tens of metres in size only required two scan positions due to reduced occlusion. Each site was scanned from two or three different positions so as to infill shadowed areas, and multiple scans were co-registered using reflector targets. TLS point clouds were then manually translated to align the roofs of buildings with the geolocated ALS data to within 10 cm vertically and < 30 cm horizontally.

2.5 Quantitative comparison

To quantitatively compare the consistency of the height estimate error in the datasets, the mean difference between the ALS and TLS derived ranges to the tallest object, and the standard deviation (SD) of those differences were calculated for a 5 m x 5 m area around the plot centres of the 20 m x 20 m extracts. In the woodland area this 5 m x 5 m measurement area was covered with dense trees. The LCM classification indicated that the woodland plot comprised 100% tall vegetation. In the surburban zone, the 5 m x 5 m measurement area was a road surface with neighbouring pavement and lamp posts with no green elements. The LCM classification indicated that this plot comprised 75% short non-vegetation (e.g. roads, footpaths, gravel driveways or cars), and 25% tall non-vegetation (e.g. buildings or lamp posts). These comparison plots therefore represent endmembers of urban structural diversity and so offer the most effective insight into the relative merits of waveform versus discrete return ALS products.

The ALS waveform-derived canopy top was calculated using the method described in Hancock et al. (2011) using a mean noise level of 12 and a noise threshold of 16. Calders et al. (2015) have demonstrated that TLS-derived estimates of canopy height are very reliable (see figure 6 in (Calders et al., 2015)) and our comparisons therefore rely on TLS being able to provide a robust validation of true canopy height. Biases between TLS measuring the leaf-underside versus the ALS measuring the leaf-topside are treated as negligible here.

3. Results

3.1 Validation of airborne discrete return and waveform ALS data with TLS

Figure 3 shows the results of comparing waveform and discrete return ALS data with TLS data. Over hard surfaces with little spatial complexity in height and structure, such as roads and buildings in the suburban area (Figure 3(a) and (b)), the discrete return data provided a height model that indicated basic trends, whilst the waveform data showed pulse blurring caused by the 3.55 nanosecond system pulse (Hancock et al., 2015). Conversely, the waveform pulses (coloured green) in Figure 3(b) travelled through urban greenspace components like bushes and shrubs and so provided potentially useful within-canopy structural information, whilst the discrete return points failed to capture the detail of the canopy profile.

In the woodland setting the ALS waveform system recorded returns from throughout the canopy and could be used to provide useful information on the canopy understorey (e.g. presence/absence, density and structure). In some settings there was penetration of the ALS waveform all the way to the ground, allowing the urban habitat to be described much more accurately than with discrete return data (Figure 3(c) and (d)). In some places, however, there were data shadows – e.g. beneath the centre of a large tree (Figure 3(d)). This same figure shows that in a few places the discrete return ALS heights of the tree tops appear to be under-estimated relative to the height derived from TLS. A few further issues are evident with the waveform data – in figure 3(b) and (d) some of the waveform returns appear below the TLS-derived ground surface. These errors are caused by the combination of multiple scattering of photons in the canopy and automatic instrument settings applied at the point of data collection. These erroneous points can be corrected using signal processing approaches (see section 1), but these are computationally complex and require extensive testing and validation.

3.1.1 Quantitative comparison

Applying the method explained in 2.3 and 2.4, statistics were generated that showed that discrete return ALS data consistently overestimated the range (and so underestimated height), with a bias of 0.82 m (SD = 1.78 m) in the 5 m x 5 m woodland test area. Conversely the waveform ALS data consistently underestimated range (and so overestimated height), but with a smaller bias, and provided a more consistent estimate of height (i.e. smaller SD) than the discrete return data (bias = -0.65 m; SD = 1.45 m). In the 5 m x 5 m suburban test area the biases showed similar patterns (discrete return bias = 0.78 m; waveform bias = -0.29 m) but the discrete return data had a lower SD (0.57 m) compared to the waveform data (0.76 m), indicating that more consistent results were achieved with discrete return data where vegetation was not present. This analysis adds weight to the suggestion that the discrete return algorithms are optimised for hard surfaces (such as roads), where they outperform simply processed waveform data, and that waveform data provide more accurate results over vegetation. It should be noted that the waveform ALS product could be processed to generate a product which performed as well as the discrete return data over hard surfaces, but the computational costs of doing so would be high.

3.2 ALS intensity measures

Further issues with discrete return ALS products are apparent when evaluating discrete return ALS intensity values over vegetated surfaces. Figure 4 demonstrates this by comparing the intensity measured from the discrete return ALS product with the reflected energy from the waveform data (the integral of the waveform intensity with time) over a mixed urban landscape in Luton. Areas of high intensity appear brighter than those with lower intensity. At 1064 nm healthy green vegetation would be expected to reflect radiation strongly and yet some of the vegetated areas in Figure 4(a) show low intensity (indicated by dark areas) which is an artefact of the diffuse return containing a large amount of energy but having a low, broad peak (Hancock et al., 2015). Therefore, there are often non-physical effects caused by signal distortion, and these could lead to large errors in interpretation of discrete return ALS data if used for automated land cover determination. This is frequently overlooked - for example studies by Antonarkakis et al. (2008) and Donoghue et al. (2007) both utilised discrete return ALS intensity as an additional measure to derive a supervised classification of vegetation types. The discrete return intensity is a function of vegetation structure (e.g. foliage profile), albedo (e.g. phenology) and the processing algorithm applied, so will confound classification accuracy if one or more of those variables is changed. Waveform ALS data are much less prone to such limitations, being able to record a much more accurate measure of reflected radiation and shape of the signal response of the target, allowing the same discrimination using the physically based shape rather than an artefact (Figure 4(b)).

3.3 Computational requirements

When deciding which ALS product to use one must consider data volumes and computational requirements underpinning information extraction. Data volume and processing costs are currently much higher with waveform data than with discrete return data. For example, the waveform files used here (LAS1.3 format (ASPRS, 2015)) were 6 to 10 times larger than the discrete return (LAS1.0 format) files. For example, 1 strip of discrete return ALS data would occupy 700Mb of disk space, whilst the same spatial extent of waveform ALS data would occupy 4.2Gb. Much of this additional data volume is occupied by wavebins that contain no usable signal but which must be retained for post-processing. Once the background noise is removed, file sizes can be reduced by roughly an order of magnitude by simple run length encoding. The signal processing needed to extract target properties is computationally expensive: applying the method described in Hancock *et al.* (2008) took 25 processor days on a computer with a 3Ghz CPU, although this could be parallelised on a cluster workstation to expedite processing time. In comparison, the discrete return point cloud is processed by the instrument during collection and typically is ready for use in geographical information systems or other image processing software on delivery (although some users will subsequently choose to apply additional topographic normalisation techniques or post-process the data using other tools).

Whilst considering the various costs of extracting information from waveform ALS data, it is also important to highlight the recent development of new software tools for expeditious analysis of such data. Not all of these tools are mature but they offer a means by which most users could extract useful information from both discrete return and waveform-capable LiDAR systems (from both ALS and TLS systems). Such tools (we list only free-to-use (FTU) or open source (O/S) options) are briefly summarised in table 1.

4. Summary and conclusions

The results shown here suggest that discrete return ALS data are optimised for use in measurement of simple hard targets (i.e. roads), and that the methods and assumptions used to generate discrete return ALS products do not permit accurate description of the three dimensional structural complexity of vegetated areas. Using two urban landscape typologies we have shown that if discrete return data were used alone, measurements of the vegetation system would be biased in terms of canopy height (underestimation), inaccurate in terms of intensity (likely resulting in physical misclassifications of greenspace) and missing vital data on the characteristics of the canopy understorey. Inaccuracies arising from the use of discrete return ALS data in measuring tree canopy height have been reported previously, for example by Zimble et al. (2003) who showed bias in deriving canopy height models from discrete return ALS (in this example, the underestimation was caused by the points missing tree tops, hitting the shoulders of tree crowns and thus, underestimating canopy height). The bias in canopy height in the discrete return ALS data reported in our study is most likely caused by the signal processing algorithms used to generate the discrete return products and has also previously been reported also by Gaveau and Hill (2003). This is a different, and additional effect to that described by Zimble et al. (2003). Such biases in discrete return ALS data could be addressed on a site-by-site basis using an empirical calibration against ground data, although using the waveform allows this bias to be removed in a more consistent way (Hancock et al., 2011).

By adopting a waveform ALS approach, there are benefits and costs for the ecologist. The major benefits are a more complete three dimensional description of the vegetation canopy. With waveform data, we show how ecologists can obtain improved canopy height models, which are critical for improving understanding of spatial carbon assessment and biomass, for example (Lefsky et al., 2005, Hilker et al., 2010). We also show the potential of the waveform approach for improved detection and description of understorey characteristics which are important if spatial models of biodiversity, resource availability (Decocq et al., 2004), and variables such as propagule abundance and connectivity (Jules and Shahani, 2003) are to be determined. To date, there have only been a limited number of studies that have investigated canopy understorey characteristics with LiDAR systems, and none currently exist which use waveform ALS for this purpose. For example, Hill and Broughton (2009) used leaf-off and leaf-on discrete return ALS data to map the spatial characteristics of suppressed trees and shrubs growing beneath an overstorey canopy, and Ashcroft et al. (2014) have demonstrated the capability of TLS to capture three-dimensional vegetation structure, including understorey. With waveform data we have shown that there exists an unexplored capability to model canopy understorey in leaf-on stage, over large areal extents: an exciting scientific opportunity. The costs are a high data storage and processing demand (see section 3.3) and in this thread there is certainly a great need for more work to improve and optimize the processing of waveform data to account for multiple scattering effects and for accounting for the waveform pulse shape. It is also worth noting that currently there are many LiDAR systems (both ALS and TLS systems) that are waveform-capable but the waveforms are often discarded during the automated process of generating discrete return data (e.g. Riegl LMS-Q1560 (Disney et al., 2010)).

In answering the question posed in the title of the paper, we therefore conclude that there is a hidden and rich resource in data from waveform ALS systems that would provide added value for spatial ecologists investigating vegetation systems and dynamics across a range of ecological systems. The ‘costs’ of processing waveform data should not be overlooked, but a growing suite of processing tools (table 1) will reduce the processing costs and the technical requirements for users of waveform data to have signal processing expertise. As waveform data become more readily available (e.g. through new global missions such as NASA’s GEDI (NASA, 2014, Krainak et al., 2012)) and tools become available to make those data easier to process, we suggest that these will provide a rich source of accurate, three dimensional spatial information for describing vegetation canopies. This will improve scientific understanding of the functional relationships between vegetation structure and related, important ecological and environmental parameters in a wide range of settings.

5. Acknowledgements

This work was funded under the NERC Biodiversity and Ecosystem Services Sustainability (BESS) thematic programme for the “Fragments Functions and Flows in Urban Ecosystems” project (Reference: NE/J015237/1; http://bess-urban.group.shef.ac.uk/). The waveform ALS data were acquired by the NERC Airborne Research and Survey Facility (ARSF) and the team from the ARSF Data Analysis Node at Plymouth Marine Laboratory is acknowledged for undertaking initial ALS processing. The Riegl VZ-400 TLS used in the validation work is owned by University College London.

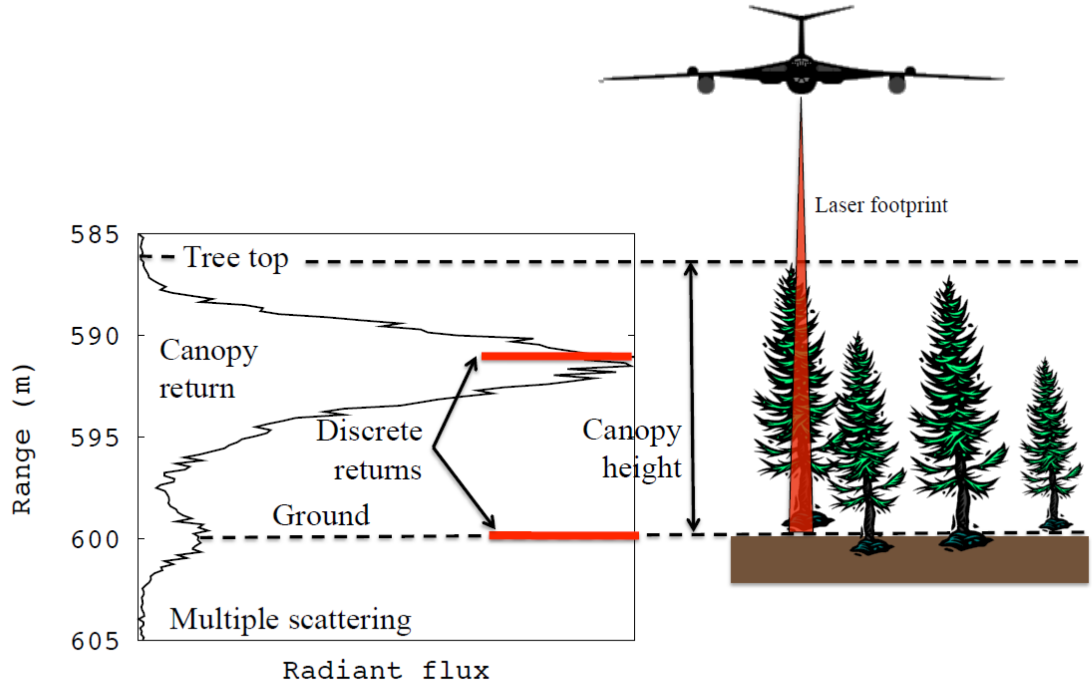

Figure 1: Stylised representation of a waveform ALS system over a tree canopy, showing a typical waveform pulse return (left of figure). In contrast, a discrete return system would not provide details of the pulse, but would instead report a series of 'hits' from various components of the landscape being monitored, typically from near to the top of the tree and from somewhere close to the ground surface (sometimes with further returns from points in between). Simulated discrete returns are shown on the plot in the left of the figure.

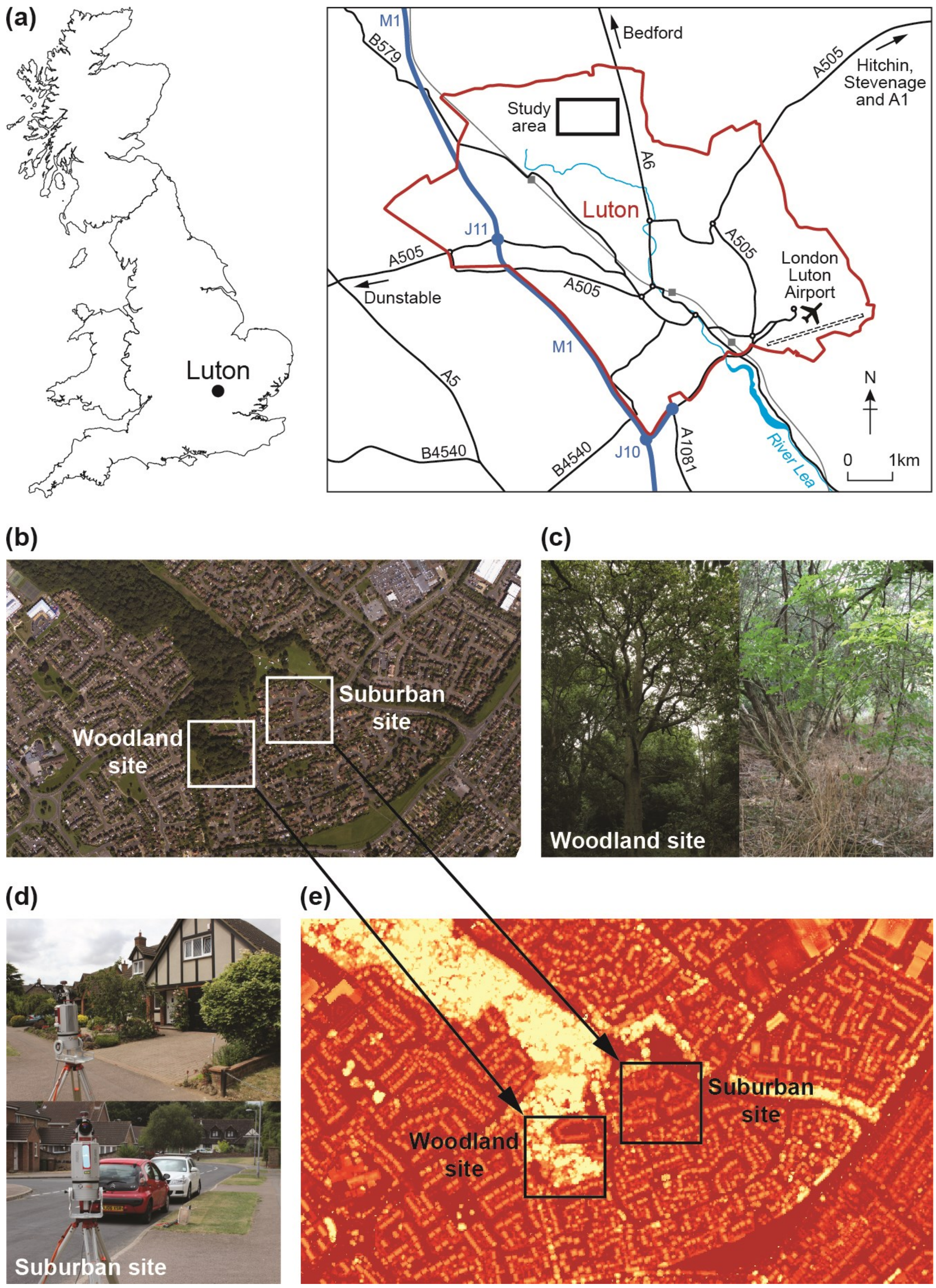

Figure 2: (a) a map of Luton with position in the UK shown inset, (b) air photo with two urban endmember typologies shown, (c) Photographs showing typical vegetation structure at the woodland site and (d) at the suburban site, (e) ALS discrete return dataset showing a basic vegetation height model of the focus area in Luton, UK.

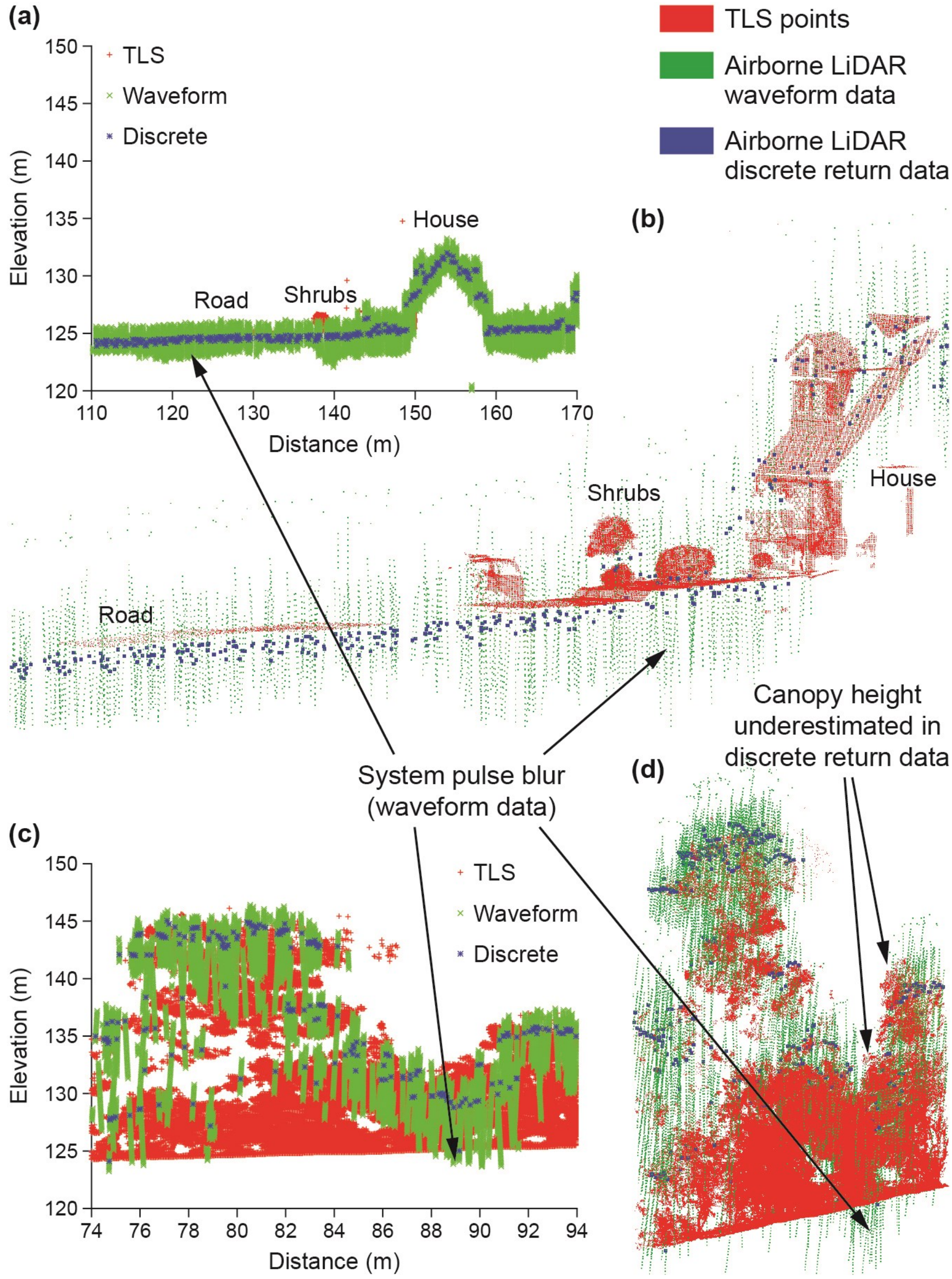

Figure 3: Comparison of TLS, and waveform and discrete return ALS data for two urban typologies. (a) and (b) show sections through the 'suburban' scanning site whilst (c) and (d) show sections through the 'woodland' scanning site. The simple plots (a) and (c) show a cross section through a 2 m deep area, whilst the more complex plots (b) and (d) show a cross section through a 20 m deep area to give a broader perspective to the comparison. The results highlight where waveform ALS intensity carries information on within-canopy structures whilst also demonstrating how discrete return ALS performs best over hard surfaces such as roads.

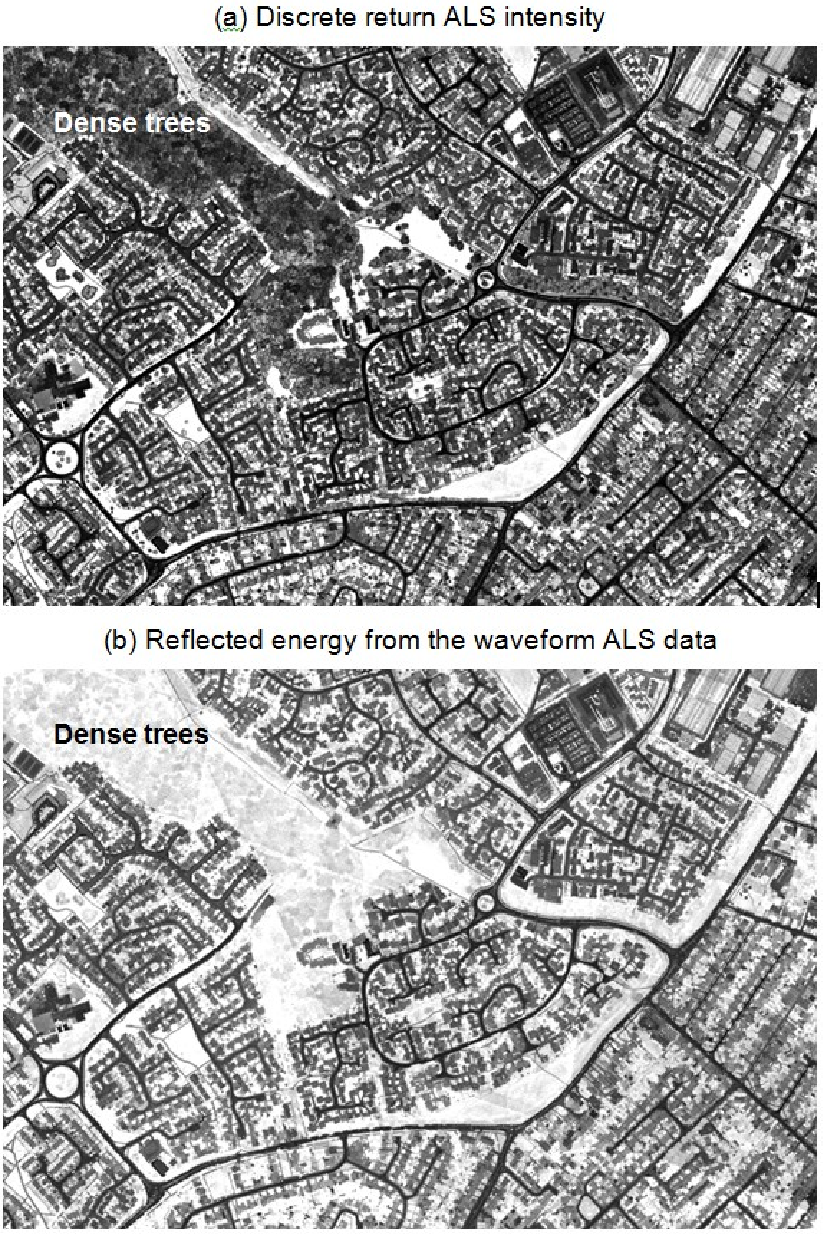

Figure 4: The impact of using discrete return intensity vs. waveform ALS in the near infra-red (1064 nm) is shown for a mixed zone in the focal area of Luton. In (a) the intensity of the discrete return ALS data are shown, whilst (b) shows the difference when waveform ALS intensity is used. The major differences in intensity appear in zones with dense vegetation. These data show that relying on discrete return intensity would lead to bias – the area of dense trees appear as having low intensity (low reflectance at 1064 nm) when they should have high reflectance (the two are related). This bias is not present in waveform intensity which shows both the mown grass and the dense trees as having high intensity which is correct given the known strong vegetation reflectance response in this region of the spectrum.

Table 1: Summarising free-to-use (FTU) and open-source (O/S) tools for processing and visualizing waveform LiDAR data

Software	FTU or O/S	Function	Coding expertise required	References
LAStools	FTU	Handling and visualising discrete return LiDAR	Low	http://www.cs.unc.edu/~isenburg/lastools/ (Podobnikar and Vrecko, 2012).
Pulsewaves	FTU	Waveform LiDAR analysis	Low	http://rapidlasso.com/category/pulsewaves/
SPDLib	O/S	Processing LiDAR data including waveform formats	High, requires C++ coding	http://www.spdlib.org/doku.php (Bunting et al., 2013)
PyLAS	O/S	Converts LiDAR formats into GIS layers	Medium, requires Python coding	https://code.google.com/p/pylas/
LibLAS	O/S	Converts LiDAR formats and links with GDAL functionality	Medium, requires Python coding	http://www.liblas.org/
Cloudcompare	O/S	Visualising 3D LiDAR point clouds	Medium, requires data in specific formats	http://www.danielgm.net/cc/

6. References

ALONZO, M., BOOKHAGEN, B. & ROBERTS, D. A. 2014. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sensing of Environment,* 148**,** 70-83.

ANDERSON, J., MARTIN, M. E., SMITH, M. L., DUBAYAH, R. O., HOFTON, M. A., HYDE, P., PETERSON, B. E., BLAIR, J. B. & KNOX, R. G. 2006. The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sensing of Environment,* 105**,** 248-261.

ANDERSON, K., BENNIE, J. & WETHERELT, A. 2010. Laser scanning of fine scale pattern along a hydrological gradient in a peatland ecosystem. *Landscape Ecology,* 25**,** 477-492.

ANTONARAKIS, A. S., RICHARDS, K. S. & BRASINGTON, J. 2008. Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment,* 112**,** 2988-2998.

ARMSTON, J., DISNEY, M., LEWIS, P., SCARTH, P., PHINN, S., LUCAS, R., BUNTING, P. & GOODWIN, N. 2013. Direct retrieval of canopy gap probability using airborne waveform lidar. *Remote Sensing of Environment,* 134**,** 24-38.

ASHCROFT, M. B., GOLLAN, J. R. & RAMP, D. 2014. Creating vegetation density profiles for a diverse range of ecological habitats using terrestrial laser scanning. . *Methods in Ecology and Evolution,* 5.

ASNER, G. P., HUGHES, R. F., MASCARO, J., UOWOLO, A. L., KNAPP, D. E., JACOBSON, J., KENNEDY-BOWDOIN, T. & CLARK, J. K. 2011. High-resolution carbon mapping on the million-hectare Island of Hawaii. *Frontiers in Ecology and the Environment,* 9**,** 434-439.

ASPRS. 2015. *LASer (LAS) File Format Exchange Activities - what is the LAS format? [http://www.asprs.org/Committee-General/LASer-LAS-File-Format-Exchange-Activities.html] Date accessed: 3 March 2015* [Online].

BOUDREAU, J., NELSON, R. F., MARGOLIS, H. A., BEAUDOIN, A., GUINDON, L. & KIMES, D. S. 2008. Regional aboveground forest biomass using airborne and spaceborne LiDAR in Quebec. *Stochastic Environmental Risk Assessment,* 23**,** 387-397.

BUNTING, P., ARMSTON, J., LUCAS, R. M. & CLEWLEY, D. 2013. Sorted pulse data (SPD) library. Part I: A generic file format for LiDAR data from pulsed laser systems in terrestrial environments. *Computers & Geosciences,* 56**,** 197-206.

CALDERS, K., NEWNHAM, G., BURT, A., MURPHY, S., RAUMONEN, P., HEROLD, M., CULVENOR, D., AVITABILE, V., DISNEY, M., ARMSTON, J. & KAASALAINEN, M. 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods in Ecology and Evolution,* 6**,** 198-208.

CHEN, Z., XU, B. & DEVEREUX, B. 2014. Urban landscape pattern analysis based on 3D landscape models. *Applied Geography,* 55**,** 82-91.

DANSON, F. M., GAULTON, R., ARMITAGE, R. P., DISNEY, M., 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.

DECOCQ, G., AUBERT, M., DUPONT, F., ALARD, D., SAGUEZ, R., WATTEZ-FRANGER, A., DE FOUCAULT, B., DELELIS-DUSOLLIER, A. & BARDAT, J. 2004. Plant diversity in a managed temperate deciduous forest: understorey response to two silvicultural systems. *Journal of Applied Ecology,* 41**,** 1065-1079.

DISNEY, M., KALOGIROU, V., LEWIS, P., PRIETO-BLANCO, A., HANCOCK, S. & PFEIFER, M. 2010. Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. *Remote Sensing of Environment,* 114**,** 1546-1560.

DONOGHUE, D. N. M., WATT, P. J., COX, N. J. & WILSON, J. 2007. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment,* 110**,** 509-522.

DRAKE, J. B., DUBAYAH, R. O., CLARK, D. B., KNOX, R. G., BLAIR, J. B., HOFTON, M. A., CHAZDON, R. L., WEISHAMPEL, J. F. & PRINCE, S. D. 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment,* 79**,** 305-319.

GASTON, K. J., ÁVILA-JIMÉNEZ, M. L. & EDMONDSON, J. L. 2013. REVIEW: Managing urban ecosystems for goods and services. *Journal of Applied Ecology,* 50**,** 830-840.

GAVEAU, D. L. A. & HILL, R. A. 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing,* 29**,** 650-657.

HANCOCK, S., ARMSTON, J., LI, Z., GAULTON, R., LEWIS, P., DISNEY, M., MARK DANSON, F., STRAHLER, A., SCHAAF, C., 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.

HANCOCK, S., DISNEY, M., MULLER, J.-P., LEWIS, P. & FOSTER, M. 2011. A threshold insensitive method for locating the forest canopy top with waveform lidar. *Remote Sensing of Environment,* 115**,** 3286-3297.

HANCOCK, S., LEWIS, P., DISNEY, M., FOSTER, M. & MUELLER, J. P. 2008. Assessing the accuracy of forest height estimation with long pulse waveform lidar through Monte-Carlo ray tracing. *Proc. SilviLaser 2008, Sept. 17–19, Edinburgh, UK (2008): 199-206.,* [available online: http://geography.swan.ac.uk/silvilaser/papers/oral_papers/Waveform%20LiDAR/Hancock.pdf].

HARDING, D. J. & CARABAJAL, C. C. 2005. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophysical research letters,* 32.

HILKER, T., VAN LEEUWEN, M., COOPS, N. C., WULDER, M. A., NEWNHAM, G. J., JUPP, D. L. B. & CULVENOR, D. S. 2010. Comparing canopy metrics derived from terrestrial and airborne laser scanning in a Douglas-fir dominated forest stand. *Trees-Structure and Function,* 24**,** 819-832.

HILL, R. & BROUGHTON, R. K. 2009. Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: A case study in lowland Britain. *ISPRS Journal of Photogrammetry and Remote Sensing,* 64.

HOFTON, M. A., MINSTER, J. B. & BLAIR, J. B. 2000. Decomposition of laser altimeter waveforms. *Ieee Transactions on Geoscience and Remote Sensing,* 38**,** 1989-1996.

HOSOI, F., NAKAI, Y. & OMASA, K. 2013. 3-D voxel-based solid modeling of a broad-leaved tree for accurate volume estimation using portable scanning lidar. *Isprs Journal of Photogrammetry and Remote Sensing,* 82**,** 41-48.

HYDE, P., DUBAYAH, R., PETERSON, B., BLAIR, J. B., HOFTON, M., HUNSAKER, C., KNOX, R. & WALKER, W. 2005. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment,* 96**,** 427-437.

HYDE, P., DUBAYAH, R., WALKER, W., BLAIR, J. B., HOFTON, M. & HUNSAKER, C. 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM plus , Quickbird) synergy. *Remote Sensing of Environment,* 102**,** 63-73.

JIAYING, W., VAN AARDT, J. A. N. & ASNER, G. P. 2011. A Comparison of Signal Deconvolution Algorithms Based on Small-Footprint LiDAR Waveform Simulation. *Geoscience and Remote Sensing, IEEE Transactions on,* 49**,** 2402-2414.

JONES, D. K., BAKER, M. E., MILLER, A. J., JARNAGIN, S. T. & HOGAN, D. M. 2014. Tracking geomorphic signatures of watershed suburbanization with multitemporal LiDAR. *Geomorphology,* 219**,** 42-52.

JULES, E. S. & SHAHANI, P. 2003. A broader ecological context to habitat fragmentation: Why matrix habitat is more important than we thought. *Journal of Vegetation Science,* 14**,** 459-464.

KRAINAK, M. A., ABSHIRE, J. B., CAMP, J., CHEN, J. R., COYLE, B., LI, S. X., NUMATA, K., RIRIS, H., STEPHEN, M. A. & STYSLEY, P. 2012. Laser transceivers for future NASA missions. *SPIE Defense, Security, and Sensing***,** 83810Y-83810Y-11.

LEFSKY, M. A., COHEN, W. B., PARKER, G. G. & HARDING, D. J. 2002. Lidar remote sensing for ecosystem studies. *Bioscience,* 52**,** 19-30.

LEFSKY, M. A., HARDING, D. J., KELLER, M., COHEN, W. B., CARABAJAL, C. C., ESPIRITO-SANTO, F. D., HUNTER, M. O. & DE OLIVEIRA, R. 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters,* 32.

LINDBERG, E., OLOFSSON, K., HOLMGREN, J. & H., O. 2012. Estimation of 3D vegetation structure from waveform and discrete return airborne laser scanning data. *Remote Sensing of Environment,* 118**,** 151-161.

LUSCOMBE, D. J., ANDERSON, K., GATIS, N., WETHERELT, A., GRAND-CLEMENT, E. & BRAZIER, R. E. 2014. What does airborne LiDAR really measure in upland ecosystems? *Ecohydrology*.

MALLET, C. & BRETAR, F. 2009. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing,* 64**,** 1-16.

NASA. 2014. *"New NASA Probe Will Study Earth's Forests in 3-D" (September 8 2014) [http://www.nasa.gov/content/goddard/new-nasa-probe-will-study-earth-s-forests-in-3-d/] Date accessed: 7 April 2015* [Online].

NERC ARSF. 2014a. *Processing report: RG12/10, flight day 249/2012, Luton [http://arsf-dan.nerc.ac.uk/trac/ticket/463] Date accessed: 7 April 2015* [Online].

NERC ARSF. 2014b. *Processing report: RG12/10, flight day 250/2012, Luton Bedford [https://arsf-dan.nerc.ac.uk/trac/ticket/457] Date accessed: 7 April 2015* [Online].

PODOBNIKAR, T. & VRECKO, A. 2012. Digital Elevation Model from the Best Results of Different Filtering of a LiDAR Point Cloud. *Transactions in Gis,* 16**,** 603-617.

RONCAT, A., BERGAUER, G. & PFEIFER, N. 2011. B-spline deconvolution for differential target cross-section determination in full-waveform laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66**,** 418-428.

SHUGART, H. H., SAATCHI, S. & HALL, F. G. 2010. Importance of structure and its measurement in quantifying function of forest ecosystems. *Journal of Geophysical Research-Biogeosciences,* 115.

VIERLING, K. T., VIERLING, L. A., GOULD, W. A., MARTINUZZI, S. & CLAWGES, R. M. 2008. Lidar: shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment,* 6**,** 90-98.

WAGNER, W., HOLLAUS, M., BRIESE, C. & DUCIC, V. 2008. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. *International Journal of Remote Sensing,* 29**,** 1433-1452.

YAN, W. Y., SHAKER, A. & EL-ASHMAWY, N. 2015. Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment,* 158**,** 295-310.

ZIMBLE, D. A., EVANS, D. L., CARLSON, G. C., PARKER, R. C., GRADO, S. C. & GERARD, P. D. 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment,* 87**,** 171-182.