Assessing the accuracy of forest height estimation with long pulse waveform lidar through Monte-Carlo ray tracing

Steven Hancock[1,2] , Philip Lewis[1], Mathias Disney[1] Mike Foster[3], Jan-Peter Muller[2]

1, Department of Geography, University College London, Gower street, London, WC1E 6BT. United Kingdom and NERC Centre for Terrestrial Carbon Dynamics.
shancock@geog.ucl.ac.uk
2, Mullard Space Science Laboratory, University College London
3, Lidar technologies ltd, Hovemere, United kingdom.

Abstract

The practicality of using an atmospheric differential absorption lidar (DIAL) such as ESA's A-scope for measuring vegetation is explored. Monte-Carlo ray tracing is used to simulate full waveform lidar responses over explicitly represented 3D forest models with both short and long temporal pulses. Deconvolution and Gaussian decomposition are used to estimate tree top and ground positions over a range of forest ages and stand densities. The errors of the height estimates are precisely quantified by comparison with the 3D model height. It is shown that (at least with a 12.5cm range resolution) an instrument optimised for atmospheric CO_2 measurement can successfully measure forest height over reasonably flat ground.

Keywords: *Forestry, lidar, vegetation, simulation, 3D modelling.*

1. Background

Carbon flux models are essential for understanding the complex processes involved in the Earth's climate (Woodward et al, 2004). These models need variables, such as biomass and leaf area index (LAI) at a range of scales and locations (Williams et al, 2005). Many areas are inaccessible and it would be prohibitively expensive to cover the world with airborne sensors. Space-borne remote sensing may be the solution.

One of ESA's six proposed Earth explorer missions, due for launch in 2012, is a space-borne full waveform lidar; the A-scope satellite (ESA, 2007). It will be optimised for measuring atmospheric CO_2 by differential absorption lidar (DIAL) with two laser wavelengths, one which causes resonance in the CO_2 molecule, one that does not. These will be close to either 1.65μm or 2.06μm. This paper investigates the ability of such an instrument to measure forest parameters.

2. Simulation system

Studies on estimating forest parameters from waveform lidar are promising; however positional uncertainty of remote measurements and the difficulty of field measurements make validation of real data difficult (Hyde et al, 2005). Computer simulations allow validation as the true parameters of the virtual forest are known. A Monte-Carlo ray tracer based upon the RAT library developed from "frat" (Lewis, 1999) was used to simulate a waveform lidar.

Explicit geometric forest models, in which every needle is described were used for the

simulations. Their creation is described in Disney et al. (2006). In a change to the method of Disney et al. (2006) needles were allowed to transmit light; trusting in the accuracy of the Prospect model (Jacquemond and Baret, 1990) in the absence of reliable transmittance data.

Using explicit 3D models is computationally expensive but avoids the assumption that canopies behave as turbid media; an assumption that ignores the heterogeneity of real trees. It is not clear how such an assumption would affect derived results, especially when derivation uses the same assumptions used to create the forest models (Widlowski et al, 2005).

Simulations were run with a range resolution of 12.5cm, a wavelength of 2.06μm, a 30m ground footprint and with and without a temporal laser pulse (100ns is proposed for A-scope). The laser pulse shape is applied to each return before binning so that quantisation noise is not ignored.

3. Realistic noise

A real direct detection instrument will suffer from noise from photon statistics, background light and detector noise. Photon statistic noise, n_s is modelled as Gaussian with a sigma of the square root of the number of photons measured in that bin. Background power, P_b, is given by the following equation (values used in this investigation are shown in brackets);

$$P_b = \rho \times E_\lambda \times \frac{TFOV^2}{4} \times A_r \times \cos(\vartheta_s) \times T_{atm} \times \Delta b \quad (1)$$

Where ρ is surface albedo (calculated from waveform), E_λ is solar energy in $Wm^{-2}sr^{-1}nm^{-1}$ (0.67), TFOV is the field of view in radians (0.0002rads), A_r is the receiver telescope area in m^{-2} ($0.79m^{-2}$), ϑ_s is the solar incidence angle (30^o), T_{atm} is the atmospheric transmission (0.8) and Δb is the bandwidth in nm (10nm). This is combined with detector noise and converted to detected photon count to get background and detector noise $n_{b,d}$ with the following;

$$n_{b,d} = QE \times P_b \times \left(\frac{\lambda}{h \times c}\right) + \frac{1}{2 \times F} \times \left(\frac{QE \times \lambda}{h \times c}\right)^2 \times NEP^2 \times t \quad (2)$$

Where QE is detector quantum efficiency (0.5), λ is wavelength in m (2.06μm), h is Planck's constant in m^2kgs^{-1}, c is the speed of light in ms^{-1}, F is the excess noise factor of the detector (2) and NEP is the noise equivalent power on the detector after amplification in W $Hz^{-0.5}$ (assumed negligible). This is then multiplied by a random number between 0 and 1 and combined with the photon statistics, n_s to get total noise by;

$$noise = \sqrt{(n_{stats} + n_{b,d}) \times F} \quad (3)$$

Different levels of noise were simulated by assuming that the signal (noiseless waveform) included a certain number of photons. Noise effects were added and the resultant waveform scaled from photon count to reflectance for analysis. Different random number seeds were used to fully investigate the effect of noise on inversions. As the noise is added based upon signal photons the wavelength is irrelevant. 2.06μm will need a more powerful laser to get the same photon count from a forest than at 1.064μm. Figures 1 to 5 used 1.064μm (they are for an optimised canopy lidar), though 2.06μm should behave in a similar fashion for the same signal level.

4. Derivation of parameters

Two of the most important biophysical parameters for ecological models are biomass and leaf area index (LAI). These cannot be directly measured by lidar or any current remote instrument but can be related to tree height and canopy coverage through empirical relationships. More complex metrics combining height and canopy coverage with height (foliage profile) can be used to improve the accuracy of estimates (Lefsky et al, 1999). With any method tree height and canopy are the measurables needed to derive any parameters.

For tree height to be measured the position of the tree top and ground must be distinguishable from the waveform. If the topography is negligible over the laser footprint tree height can be found directly. Topography complicates the matter. It may be possible to use multi-spectral lidar to extract ground position from topographically blurred waveforms.

The tree top is the signal start above background noise in the absence of a pulse length. Taking it as the point at which the signal rises above the noise threshold will always lead to an underestimate. This contributes to the "well known underestimate of tree height by lidar" (Morsdorf et al. 2008). Data assimilation schemes such as the Kalman filter rely on unbiased observations (Williams et al, 2005). Tracking back through the waveform from the noise threshold to the mean noise level should provide an unbiased estimate. Figure 1 shows a histogram of the signal start position error with and without tracking back from the noise threshold. A negative error means a premature signal trigger; this was common in both methods.

Figure 2 shows the mean and modal signal start position errors against signal photon count. The means are biased by some premature triggerings caused by noise. It is hoped that these can be removed by looking at their distance from the ground and rejecting unrealistic tree heights. No attempt was made to calculate the ground position in this experiment due to the calculation's computational expense. The modal error does not display this bias and shows that both methods giving similar outputs for large photon numbers (small noise) and the tracking method's superiority at low signal levels (high noise).

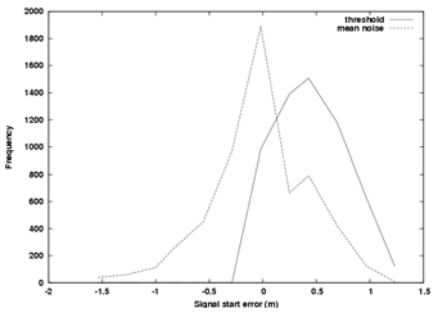

Figure 1. Signal start error histogram for 7,000 signal photons. The negative tail has been clipped for clarity.

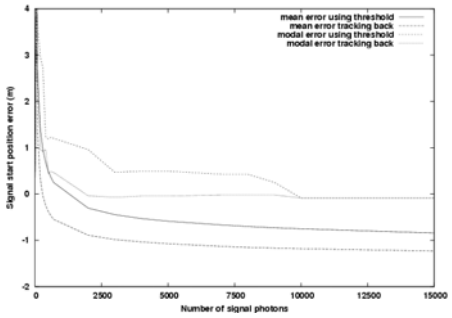

Figure 2. Mean and modal tree top error against number of signal photons for the two methods.

The ground position is much harder to extract. The traditional method is to decompose the waveform into a set of Gaussians by non-linear regression (Hofton et al, 2000). It must then be decided which Gaussian corresponds to the ground. An appropriate threshold (either amplitude or energy contained in the Gaussian) must be chosen to avoid any Gaussians caused by multiple scattering, noise or the canopy. This threshold is dependent upon canopy cover and wavelength. In denser canopies the ground return will be weaker, requiring a lower threshold. In sparser canopies more subterranean multiple scattering may be recorded requiring a higher threshold.

An implementation of the Levenberg-Marquardt method was used to minimise the root mean square difference between the fitted Gaussians and original signal (Press et al, 1994). This method is unstable, the error being affected by waveform shape, canopy cover and noise. Figure 3 shows one of the more successful attempts.

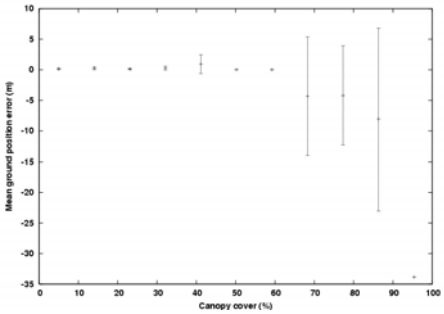

Figure 3. Mean ground position error for 4,000 photons against canopy cover. Bars show standard deviation

These methods perform reasonably well when the ground return contains significant energy and is distinguishable form the canopy return. In very dense canopies (>85% coverage) little signal reaches the ground and a proportion of the inversions will fail. The expected failure rate should be quantified to asses the method's global use as this canopy cover is not uncommon for evergreen broadleaf forests (Hofton et al, 2002). An iterative method to choose an appropriate threshold based upon an estimation of canopy coverage may be necessary.

Figure 4 shows the average energy contained in the nearest Gaussian to the ground, an indicator of how the threshold depends upon canopy cover. A failure is classed as a waveform without a Gaussian centred within (an arbitrary) 3m of the ground. The need for an iterative threshold selection is apparent.

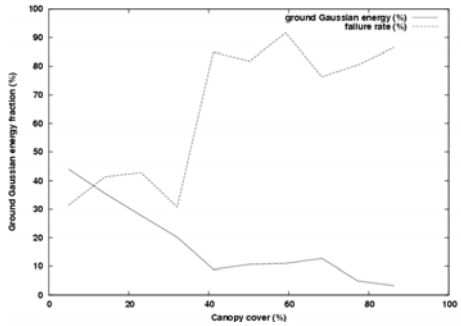

Figure 4. Fraction of waveform energy contained in nearest Gaussian to the ground.

These errors combine to give the tree height error. The signal start error is insensitive to canopy cover, possibly due to the shape of conifers (there is no more foliage at the tree top for dense than for sparse canopies). Figure 5 shows mean tree height error against canopy cover. An overestimate is suggested due to too low a threshold being used to select the ground Gaussian (0.75% of total energy).

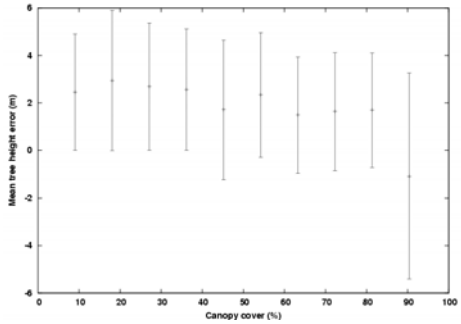

Figure 5. Mean tree height error against canopy cover. Bars show standard deviation.

5. Pulse length

All real lasers have a finite pulse length. This can range from short 6ns pulses such as ICESat's GLAS up to 100ns for ESA's proposed A-scope (for smaller linewidth). Any pulse length will blur the waveform, extending the signal start and merging ground and canopy returns. A Gaussian is a good approximation of the pulse shape. For 100ns pulses, corresponding to a Gaussian with a full width half maximum of 25m, this blurring is severe, obliterating any features (as shown in figure 6). If such an instrument is to be used for measuring vegetation some form of deconvolution is needed.

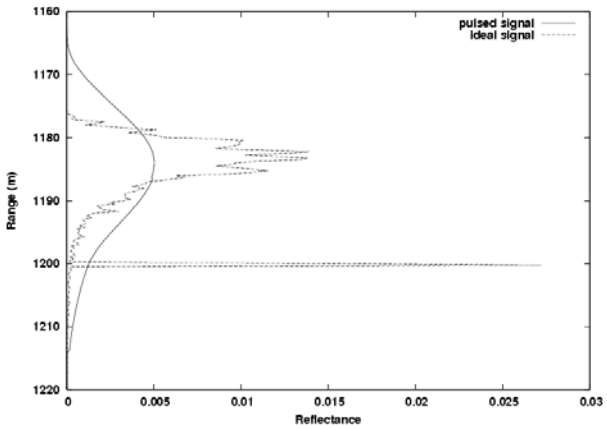

Figure 6. Simulations of an ideal and 100ns pulsed waveform over a Sitka spruce forest.

This can be done either by fitting functions with the known pulse width and shape to the waveform or a Fourier space deconvolution. As figure 6 shows, for long pulses there is little detail left to fit the function to. Any algorithm is more likely to fit a single larger amplitude Gaussian than the two Gaussians needed to de-blur.

Gold's iterative re-blurring deconvolution method (Jansson, 1997) was selected for its relative robustness to noise. This method is given by the equation;

$$\hat{o}^{(k+1)} = \hat{o}^{(k)} \frac{i}{s \otimes \hat{o}^{(k)}} \tag{4}$$

Where i is the original waveform, s is the deconvolution function (normally the laser pulse) and $\hat{o}^{(k)}$ is the kth estimate of the de-blurred waveform (initially taken as i).

Again simulations offer the advantage over reality of precise error analysis. Simulations were run with and without a pulse length. Deconvolved waveforms were compared to the ideal, pulseless waveforms.

6. Noise

Noise complicates the issue leading to wildly inaccurate products. We can be certain that there should not be any components of the waveform with a higher frequency than is contained in the laser pulse (Gurdev at al, 1993) and any such components can be taken as noise (which is high frequency). These can be removed by convolution with the laser pulse before deconvolution.

The following method was found to give the best results when deconvolving noised waveforms;

> Noise statistics were calculated from a known empty portion (all signal more than 70m above the maximum intensity return).
> Background noise was removed by subtracting a constant threshold, either the mean noise level plus three standard deviations or the maximum recorded noise level, whichever was greater.
> The waveform was smoothed with the laser pulse.
> The waveform was deconvolved with 6,000 iterations of Gold's method using the laser pulse convolved with the smoothing function as the deconvolution function.

Figure 7 shows that this gave an acceptable recreation of the ideal waveform for high noise levels (3,000 signal photons); an encouraging result. A waveform with clearly defined canopy and ground returns (without pulse length) was used for the initial investigation; it had a canopy coverage of 81% and a maximum tree height of 12.5m. This avoids the complications of trying to find low canopy or ground returns in the blurred waveforms.

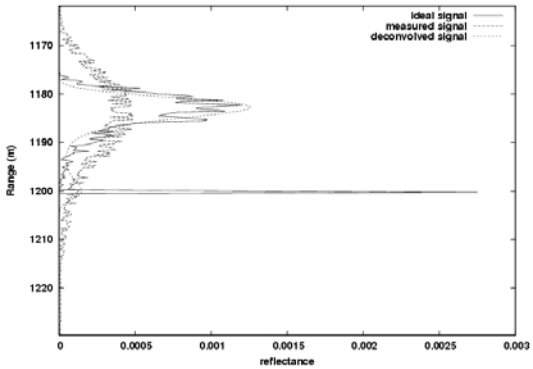

Figure 7, 3,000 signal photons, an acceptable recreation by deconvolution.

The above deconvolution method was applied to simulated A-scope waveforms for different noise levels. Figure 8 shows the mean accuracy of the inversion of height against signal photon count for a three different sets of noise added to a single waveform. The instability of the ground position estimate is apparent. A clear improvement of tree top position estimate with increasing signal to noise ratio is shown. More checks may highlight failures, improving certainty in the results. Care must be taken to separate the effects of noise, canopy cover and tree height.

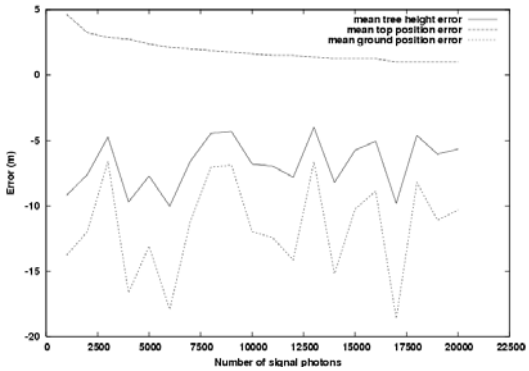

Figure 8. Tree height, tree top position and ground errors against signal photon count for a single waveform.

7. Conclusions

Tree height, top and ground position errors have been precisely quantified with simulated waveforms. A method for reducing tree top position bias has been tested and shown to perform well if at least 2,000 signal photons are measured (an easily achievable number). Some refining is needed to cope with the effect of varying canopy cover on ground position and premature triggering on tree top estimate. Above 3,000 photons errors are dominated by the algorithms. More robust algorithms may benefit from more signal photons (10,000 gives near perfect recreation of the ideal). This area needs more work before it can be considered operational.

This method relies on a clear separation between ground and canopy (after deconvolution). If the variation in ground height across the footprint is greater than the ground to foliage separation that will not be the case. This limits the areas such an instrument could be used. Smaller footprints aggregated together to ensure a tree top is recorded may be a solution; a high pulse repetition rate to allow a continuous track would be preferable. A second waveband with spectral contrast between ground and canopy will allow distinction in topographically mixed signals. Both of these would require extra equipment to be included which is unlikely in A-scope for such a secondary capability.

A-scope has a proposed laser wavelength of either 2.06μm or 1.65μm, neither of which has a strong reflectance from vegetation (the two wavelengths used in DIAL are too close to be of any advantage for vegetation). In this investigation noise levels were calculated by assuming a certain number of signal photons therefore wavelength had little impact upon this investigation. The choice of laser wavelength is likely to limit the maximum number of measurable photons.

The possibility of using long pulse lidar for measuring forest parameters has been demonstrated, given sufficient range resolution. The effect of range resolution on inversion accuracy must be quantified as A-scope is unlikely to have such a fine resolution (12.5cm in this investigation). Few waveforms and inversions were available for this investigation due to the computational expense of Monte-Carlo ray tracing and deconvolution by Gold's method. More samples are needed to fully test the methods under a range of conditions. The finer the range resolution the more information the deconvolution has and the more accurate the result is likely to be. This effect needs exploring.

Acknowledgments

This work was funded by the NCEOI (Natural environment research council Centre for Earth Observation and Instrumentation) for project SA-014-DJ-2007 (Hyperspectral Imaging LiDAR) and the Hollow waveguide technology for CO_2 and canopy measurements (A-scope) project

through the Centre for Terrestrial Carbon Dynamics (CTCD) and the EPSRC studentship no. GR/9054658.

References

Disney, M., Lewis, P., Saich, P., 2006. 3-D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains. Remote Sensing of Environment, 100, pp. 114-132.

ESA, 2007, 17th October. http://www.esa.int/esaCP/SEMHQH9ATME_index_0.html

Gurdev LL, Dreischuh TN, Stoyanov DV, 1993. Deconvolution techniques for improving the resolution of long-pulse lidars. Journal of the optical society of America, 10, p2296-2306.

Hofton MA, Minster JB, Blair JB. 2000. Decomposition of laser altimeter waveforms. IEEE transactions on geoscience and remote sensing, 38, pp. 1989-1996.

Hofton MA, Rocchio LE, Blair JB, Dubayah R, 2002. Validation of vegetation canopy lidar sub-canopy topography measurements for a dense tropical forest. Journal of geodynamics, 34 pp. 491-502.

Hyde P, Dubayah R, Peterson B, Blair JB, Hofton M, Hunsacker 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, pp. 427-437.

Jacquemond and Baret, 1990. Prospect: A model of leaf optical properties spectra. Remote sensing of environment, 34, pp. 75-91.

Jansson PA, 1997. Deconvolution of images and spectra, second edition, Academic press, p 115.

Lefsky MA, Cohen WB, Acker SA, Parker GC, Spies TA, Harding D. 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir and western hemlock forests. Remote Sensing of Environment, 10, p339-361.

Lewis, P. 1999. Three-dimensional plant modelling for remote sensing simulation studies using the Botanical Plant Modelling System. Agronomie, 19(3-4), pp. 185-210.

Lovell, J.L., Jupp, D.L.B., Culvenor, D.S., Coops, N.C., 2003. "Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests." Canadian Journal of Remote Sensing 29(5), pp. 607-622.

Morsdorf F, Frey O, Meier E, Itten I, Allgöwer B, 2008. Assessment of the influence of flying altitude and scan angle on biophysical vegetation products derived from airborne laser scanning. International journal of remote sensing, 29, p 1387-1406.

Press W.H., Tuekolsky S.A., Vetterling W.T. Flanery B.P. 1994. Numerical Recipes in C. Cambridge University Press, second edition (with reprints).

Reitberger J, Krzystek P, Stilla U. 2008. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. International journal of remote sensing, 29, pp. 1407-1431.

Rosette JAB, North PRJ, Suarez JC. 2008. Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. International journal of remote sensing, 29, pp. 1475-1493.

Widlowski J-L, Pinty B, Lavergne T, Verstraete MM, Gobron N. 2005. Using 1-D models to interpret the reflectance anisotrpy of 3-D canopy targets: Issues and caveats. IEEE transactions on geoscience and remote sensing, 4(9), pp. 2008-2017.

Williams M, Schwarz PA, Law BE, Irvine J, Kurpuis MR. 2005. An improved analysis of forest carbon dynamics using data assimilation. Global change biology, 11, pp. 89-105.

Woodward FI, Lomas MR, 2004. Vegetation dynamics – simulating responses to climate change. Biological reviews, 79, pp. 643-670.