# VALIDATION OF AIRBORNE LIDAR INTENSITY VALUES FROM A FORESTED LANDSCAPE USING HYMAP DATA: PRELIMINARY ANALYSES

D.S. Boyd[a], R.A. Hill[b]

[a] School of Geography, University of Nottingham, University Park, Nottingham, NG7 2RD, UK - Doreen.Boyd@nottingham.ac.uk

[b] School of Conservation Sciences, Bournemouth University, Talbot Campus, Poole, BH12 5BB, UK - rhill@bournemouth.ac.uk

**ISPRS WG III/3, III/4, V/3, VIII/11**

**KEY WORDS:** Lidar intensity, interpolation, coniferous, deciduous, validation, HyMap, canopy height model

**ABSTRACT:**

There is a growing body of literature that points to the value of using the intensity measures of the backscattered laser light in addition to the pulse range measurements for studying a range of environments, including forests. However, there is a lack of literature that has validated the lidar intensity values captured in a campaign, and therefore limited understanding in the full utility of these data. This paper presents preliminary analyses of lidar intensity values captured over an area of woodland in the UK in comparison with concurrently acquired HyMap data, which measures the passive reflected radiation at the same wavelengths. The study concludes that lidar intensity values are broadly representative of the NIR radiation reflected from the forested landscape and therefore could be utilised. However, there is a real need for calibration of intensity data, particularly if flight lines are to be merged. Furthermore, if lidar intensity values are to be interpolated into a raster and used in a similar way to conventional image analysis, the selected interpolation technique significantly affects the resultant lidar values.

## 1. INTRODUCTION

### 1.1 Background

Airborne lidar systems are able to record the intensity of the backscattered laser light, with intensity measured either as the maximum of the returned pulse or signal integration over the returned pulse width. This is in addition to the pulse range measurements (Wehr and Lohr, 1999). Intensity data thus provide a record of the backscattered intensity of reflection for each laser pulse, supplying information about the reflecting surface or object at sampled points across the landscape. This ability to capture backscattered reflectance from returning pulses has proved useful for the identification of broad land cover types (e.g., Brennan et al., 2006) and as ancillary data for post-processing (e.g., Liu et al., 2007). This intensity information within lidar echos is a function of the wavelength of the source energy (often within the near infrared spectral region (NIR: 0.7 – 1.5μm) for terrestrial applications), path length and the composition and orientation of the surface or object which the pulse has hit. For any data capture project, the system specific factors are known (but may be unavailable), whilst those that are site specific are typically unknown. However, tabulated values of reflectances of materials are available through endeavours of spectroscopy (e.g., Clark et al., 2003 - http://pubs.usgs.gov/of/2003/ofr-03-395/ofr-03-395.html) and suggest that there is scope in using lidar intensity for applications common place in remote sensing.

The potential in the exploitation of lidar intensity has recently being realised and been demonstrated in a number of application areas. These include the identification and mapping of the age of lava flows from active volcanoes (e.g., Mazzarini et al., 2007); glacial features (e.g., Arnold et al., 2006; Kaasalainen et al., 2006); features of archeological interest such as palaeochannels (e.g., Carey et al.., 2007); and vegetation types (e.g., Farid et al., 2006). Within forestry, lidar intensity has been used to estimate forest volume and biomass in a temperate forest of coniferous, deciduous, and mixed stands (van Aardt et al., 2006); to filter lidar-height to estimate the basal area of northern hardwood forests (Lim et al., 2003). and as a predictor in tree species classification (Holmgren and Persson, 2004). Hudak et al. (2006) concluded that lidar intensity was more useful than the EO-1 Advanced Land Imager multispectral data acquired concurrently for predicting basal area and tree density of coniferous forests. All these studies illustrate that lidar intensity values are being utilized in ways beyond perhaps originally intended. The emergence of full-waveform laser scanners may well increase this trend.

### 1.2 Factors determining lidar intensity

There are a number of factors that determine the lidar intensity values captured by a system and can be conveniently grouped into system variables and target variables, with the effect of exhibiting co-dependency. The system variables include target-emitter distance, beam divergence (there is a loss of intensity with the diverging beam), the laser footprint size, angle of incidence, atmospheric attenuation and signal processing. Target variables include target reflectivity, surface roughness and bidirectional properties and the size of the target (Wagner et al., 2006). Within forests this is mainly a function of leaf area, leaf inclination, species type, and tree density. Another factor to consider, if the data are converted from a point cloud into an interpolated 2-dimensional surface, is the post-processing procedure. Interpolation technique and selected output cell size will influence the nature of the resulting surfaces. All of these factors need to be considered and understood if lidar intensity values are to be used optimally.

Limitations in the effective use of lidar intensity values are the lack of calibration techniques (Kaasalainen et al., 2005) and the lack of validation of the lidar intensity values obtained over a particular environment. Much progress has been made to calibrate intensity both under laboratory and field conditions (e.g., Coren and Sterzai, 2006; Ahokos et al., 2006). Validation of lidar intensity by means of comparison with a similar product derived by more "conventional" means should lead to a better understanding of the parameters within which lidar intensity values can be employed. The challenge for validating lidar intensity data is the lack of reference data at appropriate spatial, spectral and temporal resolution to compare with lidar intensity values. In this study, HyMap data have been acquired concurrently with small-footprint lidar data over a woodland area in the UK, thus enabling an exploration of the lidar intensity across a landscape.

## 2. STUDY AREA

The study area focuses on two woodland sites, Monks Wood and Bevill's Wood, and their immediate agricultural vicinity, in Cambridgeshire, UK (52° 24' N, 0° 14' W). Monks Wood, covers 157 hectares and is a National Nature Reserve comprising broadleaf forest. Monks Wood is divided up into 30 compartments for management purposes. It is a complex woodland environment and extremely heterogeneous in terms of the woody species comprising the canopy and understorey, their relative proportions in any area, canopy closure and density, tree height and stem density (Hill and Thomson, 2005). The dominant tree species are ash *Fraxinus excelsior* L., oak *Quercus robur* L., field maple *Acer campestre* L., elm *Ulmus carpinifolia* Gleditsch. and aspen *Populus tremula* L., while the dominant shrub species are hawthorn *Crataegus monogyna* Jacq., hazel *Corylus avellana* L., blackthorn *Prunus spinosa* L., dogwood *Cornus sanguinea* L., and wild privet *Ligustrum vulgare* L. The majority of overstorey trees are 70–80 years old. The soils are gleyed brown calcareous and surface water gley resting on impervious clay. To the south of Monks Wood, separated by a minor road, is Bevill's Wood, a 36-hectare site that was almost entirely clear-felled and replanted in the 1950s–1960s. Bevill's Wood has stands dominated by beech *Fagus sylvatica* L., Scots pine *Pinus sylvestris* L. and Norway spruce *Picea abies* L.. These patches of woodland have a relatively homogeneous structure and tend to lack an understorey. There are, however, stands of pine and spruce that have areas of ash and scattered beech intermingled. The edges of stands inside Bevill's Wood are ringed with ash or willow trees. Both Monks Wood and Bevill's Wood have an outer fringe comprising ash, oak, field maple, hazel, hawthorn and blackthorn, and throughout both woods are open areas of herbaceous vegetation with scattered shrubs.

The woods are divided up into compartments. Within the compartments are stands and in July 2000, five contrasting stands in Monks Wood were surveyed. These covered the species composition and structure present within Monks Wood, providing a representative sample of the composition of the broadleaved woodland area. For further information on the stands refer to Table 1 in Patenaude et al., 2003 and Hill 2007.

## 3. REMOTELY SENSED DATA

### 3.1 Airborne Lidar data

An Airborne Laser Terrain Mapper (Optech ALTM 1210—see http://www.optech.on.ca) was flown over the study site in a east-west direction in June 2000. Laser pulses were emitted by the ALTM with a NIR wavelength of 1.047 μm. By scanning in sweeps perpendicular to the flight-line, the forward motion of the aircraft generated a saw-toothed pattern of point sample elevation and intensity recordings. A small scan angle range of ±10° was selected to minimize the influence of varying incidence angle on the penetration into the canopy of each laser pulse (Leckie, 1990) and thus the effect of incidence angle on intensity (Ahokas et al., 2006). The parallel flight lines had overlapping swaths of data acquisition, resulting in an irregular distribution of points. On average, one point was recorded every 4.83$m^2$ across the study site. Both first and last return range and intensity data were recorded for each laser pulse, which generated a circular footprint on the ground surface with a diameter of approximately 0.25m at nadir. Based on the instrument specifications supplied by the manufacturer and the flying altitude, the lidar data had a horizontal and vertical accuracy of approximately 0.6m and 0.15m respectively. The Lidar data acquired by the ALTM were supplied by the Environment Agency of England and Wales as an ASCII file of x-, y- and z- British National Grid co-ordinates for the first and last significant return of each laser pulse and the associated intensity values. The intensity values themselves are unitless as no method was applied to calibrate them. The individual flight lines of point sample data were supplied merged together into a single point cloud.

### 3.2 Hymap image data

The HyMap sensor records reflected radiation in 126 wavebands, for pixels with a 4-m spatial resolution (see http://www.hyvista.com). The Hymap provides a signal to noise ratio (>500:1) and image quality that is setting the industry standard and thus provides a reliable validation dataset for use in this study. Moreover, the sensor operates close to backscatter providing similarity to the laser scanner which operates practically at exact backscattering (Kaasalainen et al., 2005). This sensor was flown over the study site at a time coincident with the lidar and the acquired data were supplied as a 126-waveband raster image with DN values converted to radiance. The HyMap data were geo-registered to British National Grid co-ordinates with a 4-m spatial resolution using the aircraft telemetry from the time of data acquisition and a plug-in routine for ENVI software supplied by the HyVista Corporation. A subsequent comparison with the lidar data showed geometric accuracy to be within 1 pixel (i.e. 4-m) in the x- and y- directions. Here the reflected radiation in waveband 42 (band centre 1.0475 μm, width 0.0188 μm) were used as the validation data for the lidar intensity values.

## 4. DATA PROCESSING AND ANALYSIS

### 4.1 Data processing

Since comparisons were being made with the HyMap Band 42 data, only the first return intensity values were processed; over a wooded landscape it is most likely that the first returning pulse is from the top part of the canopy (leaves or branches), similar to that of the passive NIR radiation reflected from a canopy recorded by the HyMap sensor (Gaveau and Hill, 2003). The point-sample intensity data were interpolated into three sets of raster images, at 4-m and 1-m spatial resolution using three interpolation routines; Delaunay Triangulation (DT), Inverse Distance Weighting (IDW) and Ordinary Kriging (OK) for direct comparison with the HyMap data. Previous studies have used the lidar range data for Monks Wood to produce a digital terrain model (DTM) for the site which was then used to extract canopy height from the first-return lidar data as a grid-based digital canopy height model (DCHM). Both the DTM and DCHM have a 1-m spatial resolution (Patenaude et al., 2004). These were also available for use in the analyses.

### 4.2. Preliminary assessment of intensity rasters

A qualitative visual assessment of all the intensity images was undertaken. A grainy texture is evident in the intensity raster and this speckle is similar to that seen in radar images and a function of echo fading. Despite this, similar landscape features were visible in both the HyMap (Figure 1a) and lidar intensity data (Figure 1b), and this was most evident in the krigged intensity data. Particularly evident are the different crop types and management, the rides between compartments, clearings in the woodland, as well as areas of shrub. Also strongly evident in the interpolated lidar intensity data are the differences between individual flight lines. This demonstrates that lidar intensity data could be useful for visualisation purposes and developing an understanding of the area of interest, but that for more detailed analyses some form of calibration within each flight line is required prior to interpolation to a raster.

Figure 1a. HyMap sensor image of the study site (displayed in band 42).

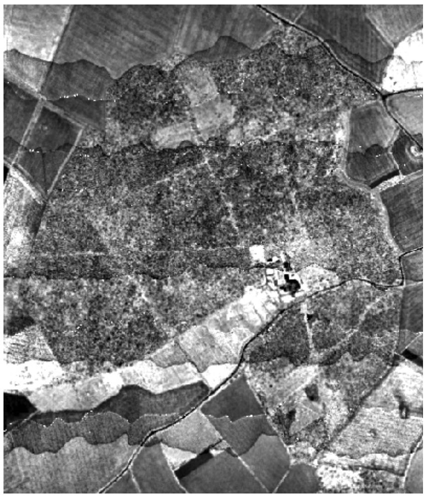

Figure 1b. First return lidar intensity image derived through ordinary kriging of the study site.

Landscape features such as deciduous and coniferous forest stands, shrubs, grassland and crops were examined for their intensity characteristics and compared with corresponding Hymap reflectances. Generally, intensity values are as expected (e.g., bare soil has low intensity and shrub a high intensity). These plots illustrate large variance in lidar intensity from each landscape feature in relation to the HyMap values. Additionally, lidar intensity from coniferous forests are high, such that there is no radiometric separability between this feature and deciduous forest and shrubs (Figure 2). This was was not the case for Hymap reflectances illustrating that the mature forests have a complex returning echo causing a similar backscatter despite structural and physiological differences between them. These results illustrate the complexity of factors that influence the lidar intensity data, further work is required to fully understand the data prior to its optimal use.

### 4.3 Stand analysis of intensity values

Per pixel analysis was conducted that compared the intensity and HyMap data at 4-m spatial resolution for all three interpolation methods for the five sampled stands in Monks Wood. The focus on these five stands should provide a range of NIR values from both the HyMap and lidar sensors for a forest of this type. Table 1 documents the regression equations computed for each stand and for each interpolation method, while Figure 3 illustrates the plots obtained for stand 5, as an illustrative example. Both the table and plots of Figure 3 show that correlation coefficients are small (although in the case of stand 4 significant at $p<0.01$) and that a large degree of scatter exists in the relationships between HyMap and lidar intensity data. This scatter may be a function of pixel mis-alignment where the geometric correction of the HyMap data was not absolute and the lack of calibration applied to the intensity data. Moreover, the intensity values for stands 1 to 3 were from areas where flightlines over lap and thus interpolated values were calculated from a double set of intensity data.

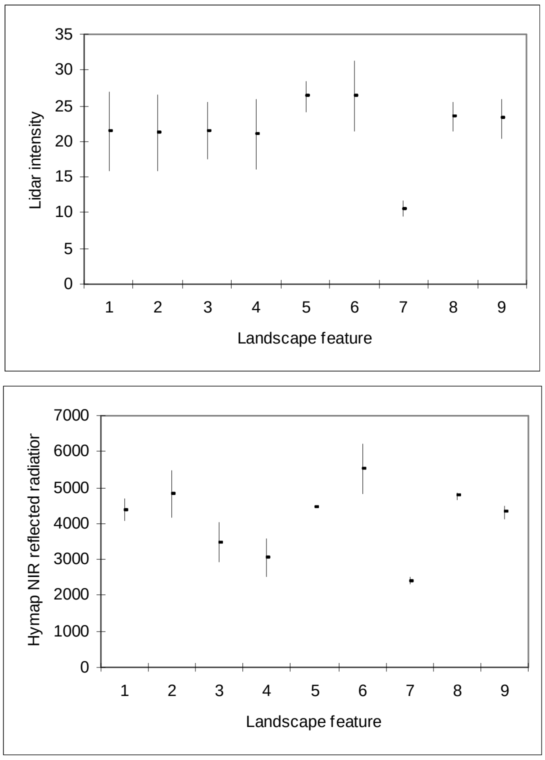


Figure 2. Characteristics (mean and ±1 standard deviation) of lidar intensity and HyMap reflectances for landscape features (1) stand 4 (Ash dominated); (2) stand 5 (Elm dominated); (3) coniferous forest; (4) coniferous forest; (5) beech; (6) shrub; (7) bare soil; (8) crop and (9) grass.

| | |
|---|---|
| | Stand 1 (N = 1747) |
| DT | y = 0.0013x + 15.736; $r^2$ = 0.01 |
| IDW | y = 0.001x + 16.169; $r^2$ = 0.01 |
| OK | y = 0.0013x + 15.064; $r^2$ = 0.03 |
| | Stand 2 (N = 2088) |
| DT | y = 0.0001x + 17.845; $r^2$ = 0.0001 |
| IDW | y = 0.0004x + 16.963; $r^2$ = 0.002 |
| OK | y = 0.0007x + 15.54; $r^2$ = 0.012 |
| | Stand 3 (N = 2010 ) |
| DT | y = 0.0005x + 16.424; $r^2$ = 0.003 |
| IDW | y = 0.003x + 17.729; $r^2$ = 0.002 |
| OK | y = 0.0004x + 17.245; $r^2$ = 0.005 |
| | Stand 4 (N = 3131 ) |
| DT | y = 0.0018x + 13.259; $r^2$ = 0.01 |
| IDW | y = 0.0017x + 13.808; $r^2$ = 0.014 |
| OK | y = 0.0017x + 14.003; $r^2$ = 0.032 |
| | Stand 5 (N = 430) |
| DT | y = 0.0024x + 9.7862; $r^2$ = 0.09 |
| IDW | y = 0.002x + 11.483; $r^2$ = 0.08 |
| OK | y = 0.0021x + 10.832; $r^2$ = 0.19 |

Table 1. Regression equations computed between HyMap data and lidar intensity data derived using each interpolation method for each of the five stands.

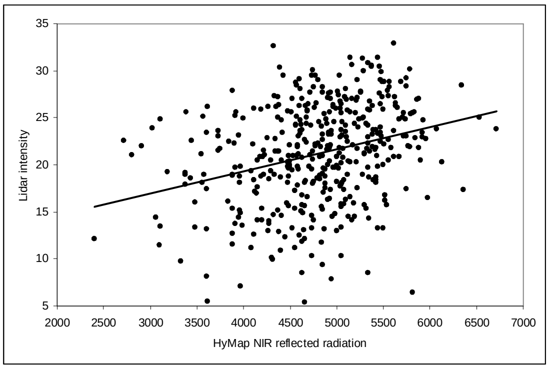


Figure 3a. Plot of Hymap NIR data against lidar intensity data derived using DT for stand 5

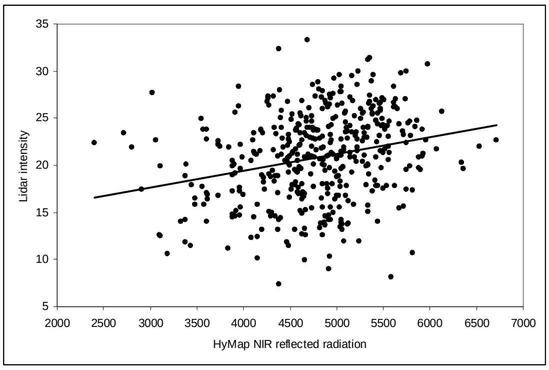


Figure 3b. Plot of Hymap NIR data against lidar intensity data derived using IDW for stand 5.

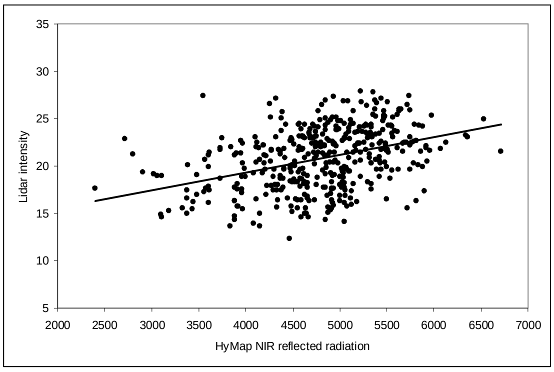


Figure 3c. Plot of Hymap NIR data against lidar intensity data derived using OK for stand 5.

Further examination of the per pixel stand data focused on the differences in the plots between HyMap and lidar intensity data apparent as a function of interpolation method. Figure 4 illustrates the histograms of intensity values derived using each interpolation method for stand 5 and the three plots between the three pairs of interpolation methods. It is evident that there are differences in the lidar intensity values, with Fisher's Z test calculations showing that each relationship is significantly different from the other ($p<0.01$). This demonstrates the significance of the selected interpolation technique if lidar intensity values are to be interpolated into a raster and used in a similar way to conventional image analysis. The full implications of this require further investigation and will be a function of point support characteristics. Similar findings can be found in the literature pertaining to the derivation of DSMs from lidar range values (e.g., Lloyd and Atkinson, 2002).

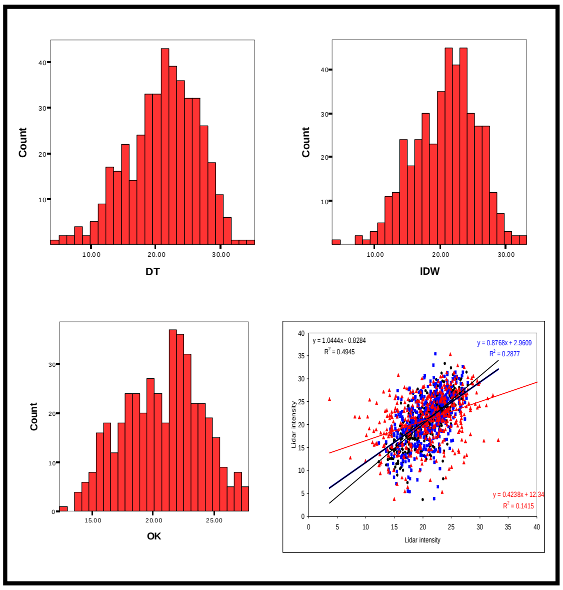

Figure 4. Illustrating the differences in lidar intensity values derived from three interpolation methods (stand 5 data): Histograms and plots of lidar intensity from OK against lidar intensity from IDW (●); lidar intensity from OK against lidar intensity from DT (■) and lidar intensity from IDW against lidar intensity from DT (▲).

### 4.4. Per parcel analysis of intensity values

To overcome the uncertainty in the per pixel analysis, further analysis focused on using parcels of pixels sampled from compartments within Monks Wood and Bevill's Wood.. Here plots were drawn for the OK derived lidar intensity values against HyMap data for 28 broadleaved parcels and 11 coniferous parcels (Figure 5a and b respectively).

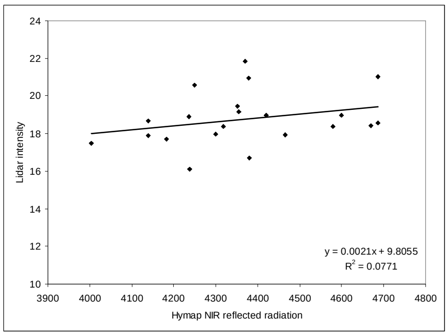


Figure 5a. Plot of HyMap NIR against lidar intensity values for broadleaved forest.

An initial examination of the plots in Figure 5, revealed insignificant relationships between HyMap data and lidar intensity values. However, on further examination of Figure 5 there is evidence of an effect of different flight lines. Within a flight line plots exhibit strong relationships between HyMap data and lidar intensity values. This is illustrated using the broadleaved compartment data. Within a flight line strong (significant at 0.01 level; two tailed) relationships exist between HyMap data and lidar intensity values (Figure 6). This suggests a real need for calibration of intensity values of different flight lines of an area of interest, if they are to be used to produce one raster for subsequent analysis. In particular there is a need to correct for observation angle (Ahokas, 2006).

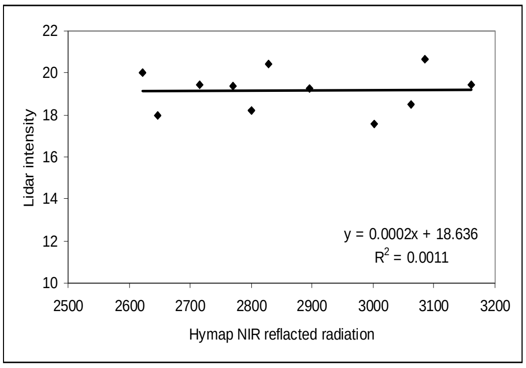


Figure 5b. Plot of HyMap NIR against lidar intensity values for coniferous forest.

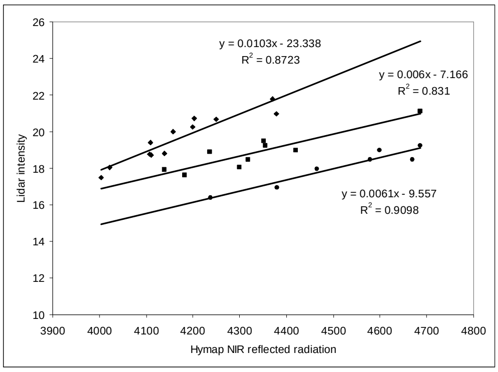


Figure 6. Plot of HyMap data against lidar intensity values for 28 compartments of broadleaved forest for each flight line.

## 5. DISCUSSION AND CONCLUSION

This paper reports on analyses into the validity and utility of lidar intensity values for a woodland environment. A number of factors influencing lidar intensity values have been explored, including interpolation methods to derive a two dimensional surface and the effect of merging flight lines on the resulting lidar intensity values to be used as an image in direct comparison with a "conventional" remotely sensed image (HyMap data). The results show that lidar intensity values correspond strongly with the HyMap data, however, there is a real need for calibration of the intensity values on an individual flight line basis so they can be used readily. The limiting factor here are a lack of calibration techniques that can be applied to a lidar dataset of this nature, in particular one that has been provided with flight lines merged. Furthermore, the lessons learned in using the lidar range values via interpolation to a grid should be heeded when using the intensity values. There is potential in using lidar intensity values, but there is still much to do to explain some of the lidar intensity values obtained. One specific query relates to the intensity from coniferous forest. Once a full understanding of the lidar intensity values obtained is achieved, it may be that future lidar campaigns need to consider the specification for mapping classes of interest via intensity in addition to the specification for terrain mapping (Reutebuch et al., 2005).

Two main conclusions can be drawn from this study: (i) lidar intensity values are broadly representative of the NIR radiation reflected from the landscape, though there are some features, such as coniferous forest, that require further analysis to understand their backscatter (ii) there is a real need for calibration of intensity data, particularly if flight lines are to be merged and interpolated.

## 6. ACKNOWLEDGEMENTS

This work was carried out whilst the RAH was based at the Centre for Ecology & Hydrology (CEH), at Monks Wood. We are grateful to the Environment Agency of England and Wales for funding and acquiring the ALTM 1210 data used in this study. The HyMap data were acquired during the SAR and Hyper-spectral Airborne Campaign (SHAC 2000) and are courtesy of the Natural Environment Research Council (NERC) and the British National Space Centre (BNSC). HyMap data pre-processing was carried out by Laine Skinner, University of Wales Swansea.

## 7. REFERENCES

Ahokas, E., Kaasalainen, S., Hyyppä, J., and Suomalainen, J., 2006. Calibration of the Optech ALTM 3100 laser scanner intensity data using brightness targets. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Marne-la-Vallee, France, Vol. 36, Part A1, CD-ROM.

Arnold, N.S., Rees, W.G., Devereux, B.J., & Amable, G.S., 2006, Evaluating the potential of high-resolution airborne LiDAR data in glaciology. *International Journal of Remote Sensing*, 27, 1233-1251.

Brennan, R., & Webster, T.L., 2006, Object-oriented land cover classifciation of lidar-derived surfaces. *Canadian Journal of Remote Sensing*, 32, pp.162-172.

Carey, C.J., Brown, T.G., Challis, K.C., Howard, A.J., & Cooper, L., 2007. Predictive modeling of multiperiod geoarchaeological resources at a river confluence: a case study from the Trent-Soar, UK. *Archaeological Prospection*, 11, pp. 37-53.

Coren, F., and Sterzai, P., 2006. Radiometric correction in laser scanning. *International Journal of Remote Sensning,* 27 (15-16), 3097-3104.

Farid, A., Rautenkranz, D., Goodrich, D.C., Marsh, S.E., & Sorooshian, S., 2006. Riprarian vegetation classification from airborne laser scanning dat with an emphasis on cottonwood trees. *Canadian Journal of Remote Sensing*, 32, pp. 15-18.

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, pp. 650-657.

Hill, R.A., & Thomson, A.G., 2005. Mapping woodland species composition and structure using airborne spectral and LiDAR data. *International Journal of Remote Sensing,* 17, pp. 3763-3779.

Holmgren, J., & Persson, Å., 2004. Identifying species of individual trees using airborne laser scanning. *Remote Sensing of Environment*, 90, pp. 415-423.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E., & Morgan, P., 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing*, 32, pp.126-138.

Kaasalainen, S., Ahokas, E., Hyyppä, J. & Suomalainen, J., 2005. study of surface brightness from backscattered laser intensity: Calibration of laser data. *IEEE GeoScience and Remote Sensing Letters*, 2, pp. 255-259.

Kaasalainen, S., Kaasalainen, M., Mielonen, T., Suomalainen, J., Peltoniemi, J.I. & Naranen, J., 2006. Optical properties of snow in backscatter. *Journal of Glaciology*, 52, pp. 574-584.

Lim, K., Treitz, P., Baldwin, K., Morrison, I., & Green, J., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, 29, pp.658-678.

Lloyd, C.D., & Atkinson, P.M., 2002. Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing*, 23, pp. 2519-2524.

Mazzarini, F., Pareschi, M.T., Favalli, M., Isola, I., Tarquini, S. and Boschi, E., 2007. Lava flow identification and ageing by means of lidar intensity : Mount Etna case. *Journal of Geophysical Research-Solid Earth*, 112, B2.

Patenaude, G.L., Briggs, B.D.J., Milne, R., Rowland, C.S., Dawson, T.P., & Pryor, S.N., 2003. The carbon pool in a British semi-natural woodland. *Forestry*, 76, pp. 109-119.

Patenaude, G.L., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.D.J., & Dawson, T.P., 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*, 93, pp. 368-380.

Reutebuch, S.E., Andersen, H.E., & McGaughey, R.J., 2005. Light Detection and ranging (LiDAR): An emerging tool for multiple resource inventory. *Journal of Forestry*, 103, pp. 286-292.

van Aardt, J.A.N., Wynne, R.H., & Oderwald, R.G., 2006. Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. *Forest Science*, 52, pp. 636-649.

Wehr, U., & Lohr, U., 1999. Airborne laser scanning – an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, pp. 68-82.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., & Studnicka, N., 2006,. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60, pp. 100-112.