PREDICTING FOREST HEIGHT FROM IKONOS, LANDSAT AND LiDAR IMAGERY

P.J. Watt[1], D.N.M. Donoghue, K.B. McManus, R.W. Dunford

Department of Geography, University of Durham, South Road, Durham, DH1 3LE, England

KEY WORDS: Forestry, LiDAR, IKONOS, Landsat, Tree height estimation

ABSTRACT:

This paper compares and contrasts predictions of forest height in Sitka spruce (*Picea sitchensis*) plantations based on medium-resolution Landsat ETM+, high-resolution IKONOS satellite imagery and airborne **Li**ght **D**etection **A**nd **R**anging (LiDAR) data. The relationship between field-measured height and LiDAR height is linear and highly significant (R^2 0.98) and so LiDAR height measurements were used to improve the height predictions derived from Landsat and IKONOS data. The results showed that despite the difference in spatial resolution and radiometry between Landsat ETM+ and IKONOS data, the strength of the relationship between field height and predicted height using the green spectral band was very similar, with R^2 values of 0.84 and 0.85 respectively. The inclusion of additional observations taken from the LiDAR data improved the strength of the relationship slightly for the Landsat ETM+ data (R^2 = 0.87), but did not change the relationship for the IKONOS data (R^2 = 0.84). Comparison of the height models derived from the satellite and LiDAR data shows that the optical models provide accurate predictions up to the point of forest canopy closure (10 m) in densely stocked plantations (>2000 stems ha^{-1}), beyond this point only the LiDAR model is able to provide a reliable estimate of forest height.

1. AIMS

This study compares forest height predictions from LiDAR, IKONOS and Landsat ETM+ data for managed Sitka spruce stands. Regression analysis is used to evaluate the quality of predictions from each of these sensors against measured tree heights. It is time consuming and expensive variable to obtain measurements of tree height for large areas of forestry. Therefore, we compare two different empirical models to predict height from multi-spectral IKONOS and Landsat ETM+ satellite image data. The first approach uses only tree height measured in the field as the dependent variable; the second approach uses height data derived from LiDAR to complement the field measurements.

2. STUDY AREA

The forest stands used in this study are located in Kielder Forest District, northern England Sitka spruce is the dominant crop and, in the Atlantic maritime climate of the UK, it is a fast growing tree that is tolerant of acid and waterlogged soils. The topography consists of low undulating hills with an altitude range of 30 to 600 metres. Planting occurs on land with a mean altitude of 270 metres and a mean slope angle of 6°. The initial density of the plantation usually exceeds 2,500 trees per hectare (a spacing of 2 m by 2 m between trees and rows). These plantations are almost never thinned so closure of the forest canopy normally occurs between 15 and 20 years after planting.

3. FIELD DATA

Mensuration data was collected between January and May 2003 and comprised twenty eight 0.02 ha circular ground survey plots. Sample measurements included; tree height recorded using digital hypsometer and tree diameter at breast height (dbh) using a diameter tape (Table 1). In all plots the position of each tree was derived using either differential GPS or laser total station initialized on an established survey point. Canopy closure status was determined by recording the type and proportion of understorey vegetation present in each sample plot using a 1-metre square quadrat split into 4 equal quadrants. Plots consisting of greater than 50% dead vegetation on forest floor were classified as closed canopy and conversely those with less than 50% understorey vegetation, open canopy.

	Mean	Standard Devation	Minimum	Maximum
Age (years)	33	18.8	8	59
Density (trees ha^{-1})	2,732	2,614	1,150	12,300
Basal area (m^2 ha^{-1})	47	17.4	4.5	69.4
Height (m)	11.1	6.7	1.5	22.3
Diameter (cm)	17.1	5.1	4.3	23.8

Table 1. Summary of plot data

4. IMAGE DATA

4.1 LiDAR data

The LiDAR data was acquired over the Kielder forest site on 26 March 2003, by the UK Environment Agency using an Optech ALTM 2033 system. The ALTM 2033 is a discrete return system, that operates at 1047 nm (near infrared), capturing two returns (first and last) for each laser pulse. The 6 km^2 study area was covered by four parallel flight lines orientated in the east-west direction. The system collects data by scanning perpendicular to the direction of flight resulting in a zig zag pattern of irregularly spaced data points. On average laser measurements were made at a density of two returns per m^2 from a flight altitude of 950 m. At this altitude, the footprint diameter of laser on the ground is approximately 0.25 m at nadir. The *xyz* position and intensity of each pulse were supplied geo-referenced to British National Grid. The height of the *z* position was supplied as elevation above the Ordnance Survey of Great Britain 1936 Datum.

[1] Corresponding author

4.2 LiDAR height estimation

An initial screening of the LiDAR was conducted to eliminate any erroneous elevation values present in the data. The LiDAR first and second return point data were separated into ground and above-ground (canopy) returns using the Terrascan software (Soininen, 1999). The classification was further simplified by re-sampling the point data into a regular grid format with 4 m spacing. The variable *maximum canopy height* was derived by selecting only the highest returns in each 4 m cell and subtracting the canopy surface from the ground surface. The accuracy of LiDAR ground surface was verified using a Total Station and traditional survey methods to measure height profiles in open and closed canopy plantations. The RMS error for all profiles was 0.34 cm. In all cases the LiDAR derived ground height gave a small overestimation when compared to the measured ground surface, a result that is observed in most studies of this kind (Means et al. 2000). This is illustrated in Figure 1, which shows the most topographically variable profile in a very densely planted forest stand. In this example, the profile begins on a road and runs downwards through a sample plot ending on the other side of a deep drainage channel. Despite the density of the overhead canopy, approximately 35% of the LiDAR hits still penetrated to the forest floor.

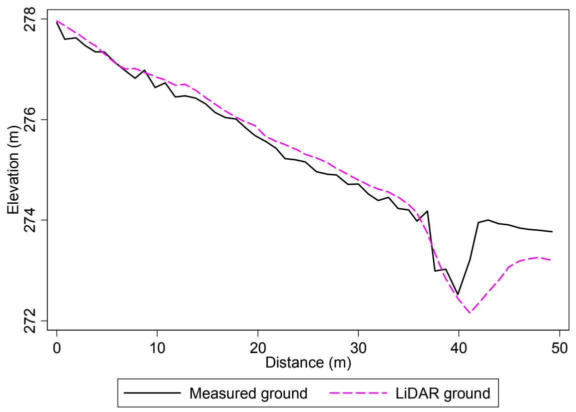

Figure 1. Profiles of measured ground elevation and LiDAR derived ground elevation across the same forest stand

4.3 Satellite data

The two satellite images used in this study were acquired in cloud-free conditions six months apart, with IKONOS multi-spectral (Geo product) data acquired on 3 March 2002 and the Landsat 7 ETM+ (level 1G format) data on the 2 September 2002. Both sensors have almost identical spectral band passes in the blue-green (**I**KONOS 0.45-0.52 µm : **L**andsat 0.45-0.52 µm) green (I 0.51-0.60 µm : L 0.52-0.60 µm), red (I 0.63-0.70 µm : L 0.63-0.69 µm), near infrared (I 0.76-0.85 µm : L 0.76-0.90 µm). A Landsat image covers an area of 185 km x 170 km at a 30 m spatial resolution in all bands except for the thermal band (60 m). An IKONOS image covers a nominal area of 16 km x 16 km at nadir at a spatial resolution of 4 m in all multi-spectral bands.

The IKONOS and Landsat images were clipped to the spatial extent of the study area and geo-rectified using ground control points derived from forest compartment boundaries and the LiDAR data. The RMS error was less than a pixel for both images. The Landsat ETM+ image was re-sampled using a nearest neighbour algorithm to 4 metres to allow overlay and comparison with the IKONOS and LiDAR data.

5. RESULTS AND DISCUSSION

5.1 Estimation methods

For each 0.02 ha plot field data mean height was calculated with the corresponding LiDAR grid and satellite pixel values extracted from the IKONOS and Landsat ETM+ data (with the exception of the ETM+ thermal band). Ten to twelve pixels were extracted per sample plot with the mean of these pixels taken to represent the canopy height or radiance value of the plot. A second data set was constructed using LiDAR data as a substitute for field data. A regular 100 m grid was generated and overlaid with the image data sets. A circular buffer of 7.98 m was generated around each sample to correspond with the 0.02 ha ground sample plot. Using the forestry GIS layers the sample points were stratified into non-forest, forest and those points that were located within 30 m of forest compartments and forest rides (gaps). Only points located clearly within a forest stand were retained. This process resulted in 410 additional samples over the study area.

The relationship between LiDAR derived height and measured height is linear and so a conventional least squares linear regression model is appropriate. For IKONOS and Landsat ETM+ data, various single and multiple band regression models were tested but single band models were preferred over the multi-band models for two reasons. First, the amount of variation explained by the addition of other bands did not improve the fit of the models to the ground survey data as summarised by the R^2 value and RMSE values. Secondly, a simple model based on a single band yields a simple model that can be understood in a physical sense and can easily be transferred to other locations. Inspection of the scatter plots for IKONOS and Landsat ETM+ and data suggests that the relationship between reflectance and height is non-linear. A number of non-linear regression models were applied to the data, and a model of the type selected below best describing the relationship:

$$y = ax^b$$

Where y is mean sample plot height, x is either the IKONOS or Landsat ETM+ band digital number (DN) value and a and b are empirically derived constants. The same regression approach was used generate a second model which were based on the inclusion of 410 LiDAR measured sample plots. Height estimation images for each band of the IKONOS and Landsat ETM+ were generated by applying the regression equation to the DN value. A map of the height residuals wa created by subtracting the LiDAR height from the image bands for the two models.

5.2 Discussion

The regression models between the ground reference data, satellite spectral data and airborne LiDAR data are summarised in Table 2. The LiDAR height model is very strongly related (R^2=0.98) to mean height within the sample plot. Figure 2a shows the LiDAR height plotted against mean sample plot height values. It is interesting to note that the largest amount of variability in the relationship, as shown by the residual plot (Figure 2b), occurs for heights of 11-16 m, just after the crown canopy has closed but is still very dense. At this stage of growth, fewer laser pulses reach the ground and this probably accounts for the small amount of variance observed in the otherwise very strong relationship. Most previous studies that compare LiDAR and measured height have concentrated on mature crops (e.g. Hudak et al. 2002, Means et al. 2000). In this study we show that LiDAR predicts height over the full range of age classes studied (8-59 years).

Sensor	Band	Height model 1 (n=28)		Height model 2 (n=438)	
		R^2	*RMSE*	R^2	*RMSE*
LiDAR		0.98	0.83		
LANDSAT ETM+					
Band 1	Blue	0.65	0.42	0.71	0.70
Band 2	Green	0.84	0.28	0.87	0.48
Band 3	Red	0.86	0.26	0.84	0.52
Band 4	NIR	0.60	0.45	0.46	0.98
Band 5	SWIR	0.80	0.33	0.85	0.50
Band 7	SWIR	0.80	0.32	0.80	0.60
IKONOS					
Band 1	Blue	0.75	0.35	0.81	0.56
Band 2	Green	0.85	0.27	0.84	0.53
Band 3	Red	0.82	0.30	0.86	0.49
Band 4	NIR	0.45	0.53	0.30	1.09

Table 2. Summary of regression models used to estimate forest height from LiDAR and satellite data

Mean height is most strongly related to IKONOS and Landsat ETM+ data in the green and red wavelength bands. Table 2 provides a summary of regression of imagery and field plot data (Model 1) and imagery and field plot data supplemented by LiDAR height data (Model 2). The inclusion of LiDAR height information is helpful because it extends the range of heights over which the statistical models are fitted, and, the additional number of samples gives us more confidence that the empirical statistical models give a sensible prediction of forest height (Wulder and Seeman 2003). As a quantitative measure of fit, the R^2 values show that inclusion of LiDAR data improves the Landsat height predictions slightly, but makes little difference to the IKONOS predictions (Figures 2c and 2d).

Figures 2e and 2f show predicted height against LiDAR height for both Landsat and IKONOS data using the green wavelength data. It is apparent from both scatter plots that beyond a height value of approximately 10 m the relationships deteriorate significantly. This limitation of optical image data has been reported in previous studies when satellite observations have been compared with field measurements (Danson and Curran 1993, Puhr and Donoghue 2000, Donoghue et al. *in press,* Nilson and Peterson 1994). In this study the LiDAR observations allow us quantify the behaviour of the relationships in much more detail. The LiDAR data also allows us to study the spatial patterns in residual maps (not presented here).

There is no apparent difference in the accuracy of the height predictions based on green wavelength data between the IKONOS (4 m) multi-spectral and the Landsat ETM+ (30 m) in spite of the large difference in spatial resolution between the two sensors. IKONOS imagery although much more expensive than Landsat ETM+ and other medium resolution sensors such as SPOT HRVIR or IRS-1D, may be used for other forest applications such as detecting small patches of wind damage or ecological mapping in place of conventional aerial photography. Both IKONOS and Landsat ETM+ are able to survey an extensive forested area in a single image, the data is acquired instantaneously and the survey is easily repeated. In addition, each image requires very little computer processing and the predictive models are simple and easily understood. Predictions do not require a large ground survey effort, and, could be replaced by LiDAR data if available. An operational system would not have to rely on a single source of image data.

6. CONCLUSIONS

1. In very dense forest stands (>2000 stems ha^{-1}) there is sufficient LiDAR penetration to provide a reliable estimates of ground elevation and tree height.
2. In densely stocked Sitka spruce crops LiDAR is able to provide an estimate of height over the entire range of age classes studied (8-59 years).
3. Optical satellite data shows a good prediction of height from 0-10 m; above 10 m height predictions are very poor.
4. LiDAR data can be used in place of, or to supplement field measurements of tree height.
5. LiDAR data is still significantly more expensive than optical data. However, this paper shows that optical satellite image data can be used to map forest height up to canopy closure using optical satellite images that can cover large areas at very low unit cost.
6. LiDAR could be used in place of field measurements of tree height to derive prediction from optical satellite image data such as IKONOS and Landsat ETM+.

References:

Danson, F.M., and Curran, P.J., 1993. Factors affecting the remotely sensed response of coniferous forest plantations. *Remote Sensing of Environment*, 43, 55-65

Donoghue, D.N.M., Watt, P.J., Cox, N.J., Dunford, R.W., Wilson, J., Stables, S. and Smith, S. (*In press*) An evaluation of the use of satellite data for monitoring early development of young Sitka spruce plantation forest growth, *Forestry, 77(5)*

Hudak, A.T., Lefsky, M.A., Cohen, W.B. and Berterretche, M., 2002. Integration of LiDAR and Landsat ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, 82, 397-416

Means, J.E, Acker, S.A., Brandon J., Fitt, B.J., Renslow, M., Emerson, L. and Hendrix, C., 2000. Predicting forest stand characteristics with airborne scanning LiDAR. *Photogrammetric Engineering and Remote Sensing*, 66, 1367-1371

Nilson, T. and Peterson, U., 1994. Age dependence of forest reflectance: analysis of main driving factors. *Remote Sensing of Environment,* 48, 319-331

Puhr, C.B. and Donoghue, D.N.M., 2000. Remote sensing of upland conifer plantings using Landsat TM data: a case study from Galloway, south-west Scotland. *International Journal of Remote Sensing*, 21, 633-646

Soininen, A, 2002. Terrascan for Microstation, user's guide. Terrasolid Ltd., Helsinki, Finland. Available from www.terrasolid.fi

Wulder and Seemann, 2003. Forest inventory height update through the integration of LiDAR data with segmented Landsat imagery. *Canadian Journal of Remote Sensing*, 29, 536-543

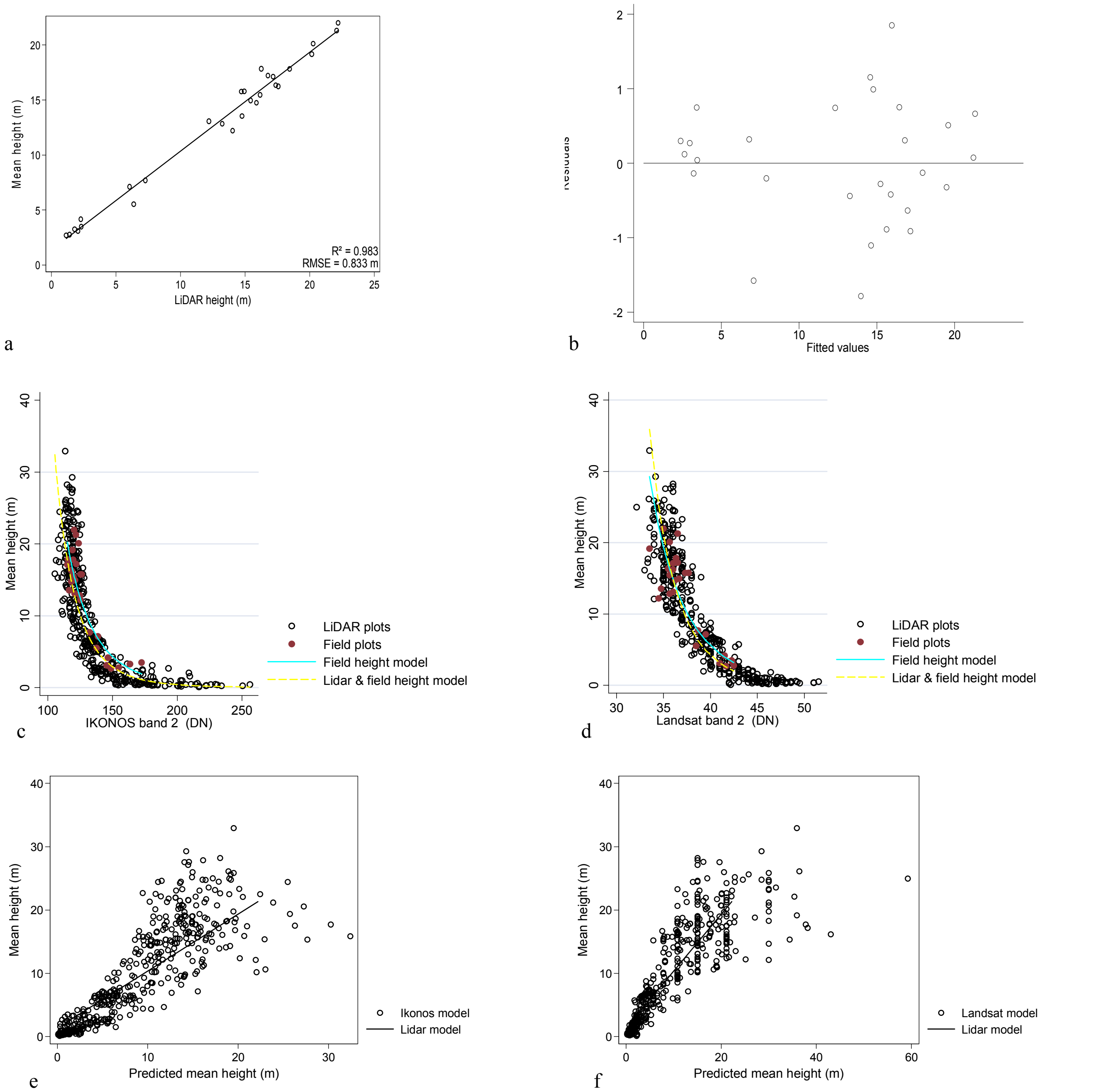

Figure 2. (a) LiDAR height against mean height, (b) LiDAR residual plot, (c) LiDAR height against Landsat Band 2, (d) LiDAR height against IKONOS Band 2, (e) LiDAR height against predicted height IKONOS band 2 (f) LiDAR height against predicted height Landsat Band 2