LiDAR mapping of canopy gaps in continuous cover forests; a comparison of canopy height model and point cloud based techniques

Rachel Gaulton & Tim J. Malthus

The School of GeoSciences, The University of Edinburgh, Drummond Street, Edinburgh, EH8 9XP, R.Gaulton@sms.ed.ac.uk, tjm@geo.ed.ac.uk.

Abstract

In continuous cover forest systems, canopy gaps are created by management activities with an aim of encouraging natural regeneration and of increasing structural heterogeneity. Canopy gaps are difficult to map from the ground and LiDAR may provide a means to accurately assess gap distribution, allowing more effective monitoring. This paper presents a new approach to gap delineation, based on identifying gaps directly from the point cloud and avoiding the need for interpolation of returns to a canopy height model (with associated errors). Areas of canopy are identified through local maxima identification, filtering and clustering of the point cloud, and gaps are then identified and delineated in a GIS environment.

When compared to field mapped gap outlines, the algorithm has an overall accuracy of 88% for data with a high LiDAR point density and accuracy up to 77% for lower density data. The method provides an increase in producer's accuracy, of on average 8%, over a method based on the use of a canopy height model. Results indicate that LiDAR data can be used to accurately delineate gaps in managed forests, potentially allowing more accurate and spatially explicit modelling of understorey light conditions.

Keywords: spatial structure, airborne laser scanning, gap delineation, clustering, forestry.

1. Introduction

The distribution, shape and extent of gaps in forest canopies can influence a range of ecological factors including the understorey light regime and vegetation, microclimate and soil moisture (Page and Cameron 2006) as well as influencing other considerations such as the aesthetic appeal of the stand. Continuous cover forestry (CCF) is a management approach that maintains a forest cover over time, by the selective felling of single or small patches of trees, and aims to produce a more diverse forest structure (Mason and Kerr 2004). Canopy gaps within CCF forests result from both management activities and natural disturbances such as windthrow and allow the occurrence of natural regeneration in the understorey, reducing the need for re-planting. Monitoring such stands presents new challenges, as detailed and spatially explicit information is needed on a more frequent time-scale than for traditional clear-cut systems.

Airborne laser scanning provides high resolution information on vertical and horizontal aspects of forest structure and may allow the study and monitoring of fine-scale spatial structural heterogeneity of stands in the process of conversion to CCF, through the identification and delineation of canopy gaps. By examining the spatial distribution and characteristics of canopy gaps, indices of spatial structure can be developed with direct relevance to important ecological parameters, whilst the gap distribution itself is useful in predicting understorey conditions.

This paper examines the use of discrete return, small footprint LiDAR data for the delineation of canopy gaps (mainly resulting from management) in Sitka spruce (*Picea sitchensis*) plantations. A new approach, delineating gaps directly from the LiDAR point cloud, is developed and compared to previous approaches based on the use of a canopy height model (CHM). The accuracy of delineations is assessed against independent field measurements.

2. Background

A canopy gap can be viewed as a lowering of the surrounding forest canopy below a certain height threshold. Information on the height of the canopy surface can therefore aid the identification of gaps. A number of studies have examined the potential of surface models derived from digital photogrammetric techniques for identifying canopy openings (Nuske *et al*. 2007; Betts *et al*. 2005; Nakashizuka *et al*. 1995). Small-footprint discrete return LiDAR sensors have been widely used to derive locations, heights and crown sizes of individual trees (Falkowski *et al*. 2006; Koukoulas and Blackburn 2005; Suárez *et al.* 2005). However, very little explicit consideration has been given to delineation of canopy gaps from LiDAR data.

The assumption that mapping canopy gaps using LiDAR is straightforward (Nuske *et al.* 2007) has not been tested. Koukoulas and Blackburn (2004) mapped canopy gaps from LiDAR data of deciduous woodland, with the aim of describing gap spatial structure and within-gap vegetation types. A CHM was generated from the raw data and a fixed height threshold (four metres) was used to define a gap. Morphological shrinking functions were then applied to isolate individual gaps. However, by applying a 'hard' height threshold to delineate gaps, the algorithms used do not lend themselves to applications in uneven-aged stands likely to result from CCF. A recent study by Zhang (2007) examined the use of LiDAR for delineating gaps in mangrove forests. The fixed height threshold method of Koukoulas and Blackburn (2004) was compared to a mathematical morphology based method, using opening and closing operations, in which the threshold height was determined as a ratio of whole gap height. However, neither study undertook a formal accuracy assessment of the gap delineations or any comparison to field data. In Koukoulas and Blackburn (2004) field measurements were not felt to be comparable and delineations from aerial photographs too inaccurate. Zhang (2007) compared LiDAR delineations to a slope grid generated from the canopy model but not to field data. Further work is therefore needed to assess the potential of LiDAR data for accurate gap delineation, including comparison to appropriate field data.

Acquisition parameters such as scanning angle and point density could have considerable impact on the accuracy of gap delineations. A number of studies have considered the influence of flight altitude (Morsdorf *et al*. 2008; Goodwin *et al.* 2006; Næsset 2004) and scan angle (Morsdorf *et al*. 2008; Ahokas *et al*. 2005) on the retrieval of biophysical properties, including individual tree height and crown width. Maximum scan angle (from nadir) is likely to influence the range of angles at which LiDAR pulses interact with the canopy and ground, the spacing of pulses (across track) and the distance that laser pulses travel through the forest canopy. As large scan angles can result in areas of shadow where the pulses do not penetrate adjacent tree crowns, this seems likely to have a significant influence on the delineation of canopy gaps. Point density is also likely to influence accuracy, as at low point density the canopy drip-line of trees (defined below) may be frequently missed. Although the influence of point density and scanning angle on the accuracy of gap delineations is not explicitly investigated in this paper, at lower densities or in areas of shadow, the averaging involved in interpolation of point data to CHMs is likely to compound these problems and reduce the precision of delineations. The process of producing a CHM also leads to the loss of data from different levels in the canopy, preserving just lidar returns from the surface, leading to a further reduction in the information available for gap delineation. This paper therefore develops an alternative approach, avoiding the use of a CHM and instead delineating gaps directly from the point cloud.

3. Data collection and pre-processing

3.1 Study sites

The three study sites are located in Sitka spruce (*Picea sitchensis*) plantations in the UK. Aberfoyle, Scotland is the site of a CCF trial area. Two one hectare plots were used for field measurements, one in a stand in the early stages of transformation (AbP1) and the other in a traditionally managed plantation that had been lightly thinned (AbP2). Clocaenog Forest in North Wales is also a CCF trial area, managed by the University of Wales, Bangor. One plot (CLG1) is a thinned control plot, whilst the other (CLG4) has been managed as a uniform shelterwood since 2004. The final study site is Glasfynydd Forest in South Wales. This is the site of a Forest Research thinning experiment. Plots were established in 50 year old stands and thinned (in 2002 and 2005) to 69% (plot GLT) and 60% (plot GHT) of their original basal area. The plots exhibit a range of canopy gap sizes and distributions due to their management history.

3.2 LiDAR and field data

Airborne campaigns were carried out over all three sites in summer 2006. Discrete return LiDAR data (first and last returns) for Glasfynydd and Clocaenog were collected by the NERC Airborne Research and Survey Facility at an average point density of 1.2 returns per m^2 with a maximum scan half-angle of 20 degrees. Data for Aberfoyle was acquired by the Environment Agency at a much higher point density (11.4 returns per m^2) and a maximum scan half-angle of 10 degrees.

A gap can be viewed as a 'hole' in the forest canopy caused by the loss or removal of one or multiple trees, thereby excluding small gaps within tree crowns or between neighbouring trees. To allow the comparison of field and LiDAR data, a more complete definition is needed. For the purposes of this study, a gap boundary is defined by a line at ground level (the drip-line) located vertically beneath the inner most point reached by the foliage of a tree crown at any level, at that point on the gap perimeter, as suggested by Brokaw (1982). A gap must have a minimum area of $5m^2$ and must extend down to at least ten metres from ground level (closure occurs when regeneration reaches an average height of ten metres).

A detailed ground survey of all plots was carried out in September 2005 for Aberfoyle and summer 2006 for the Welsh sites. Full data sets including tree positions and dimensions were available for Clocaenog and Glasfynydd and these were up-dated by re-measuring a subset of trees. A 50m by 50m subplot was established in each plot and the boundaries of all canopy gaps (fitting the study definition) were surveyed. As it was not possible to survey every point along the canopy drip-line of gap boundaries, the resulting field-mapped boundaries were to some degree generalised. A Total Station survey, from a GPS baseline, was used to record points located directly below the drip-line wherever significant changes in the orientation of the gap boundary occurred. Several points were recorded per tree crown. Gap boundaries were mapped for all plots except CLG1.

3.3 LiDAR pre-processing

Pre-processing of LiDAR data was carried out using FUSION (McGaughey 2007). A digital terrain model was generated from the last return data by filtering to leave bare earth points, using an iterative process adapted from Kraus and Pfeifer (1998). A gridded DTM was generated from these points with a cell size of 2m. The accuracy of the DTM was assessed against height measurements obtained during the GPS and total station survey, resulting in an average RMSE of 1.1m. This error falls within the vertical accuracy levels of the GPS positions so the absolute accuracy of the DTM could not be assessed further. To determine the degree to which the DTM represented variation in the ground surface (a factor perhaps more important to this study than absolute height), the relationships between total station surveyed points and the heights extracted for the same locations from the DTMs were examined, giving a strong linear relationship (R^2 of 0.93-0.99, $P<0.001$) and suggesting that the DTM accurately represented variation in the ground surface within the plots. The corresponding value of the DTM was then subtracted from the first return point data to convert from elevation to height above ground.

A canopy height model (cell size of one metre for the Welsh sites and 0.5 metres for Aberfoyle) was generated from the first returns, also within FUSION. The highest first return (from the point cloud, converted to height above ground) was assigned to each grid cell, with missing values interpolated by averaging of neighbouring values. The use of a TIN to generate the CHM was also investigated but produced poorer results in this study. Heights derived from the final CHM were compared against field measured tree heights (from 2007) for plot CLG1, resulting in a correlation co-efficient of 0.95 ($P<0.001$) and a mean error of -1.55m.

4. Methodology

Two methods were used to delineate gaps from the first return LiDAR data. In the first, a relative height threshold of 66% of local tree height was applied to the CHM. This threshold was found to correspond to the canopy drip-line through manual measurements of the point clouds of 45 individual trees located in open areas or on gap edges. A surface representing the top of the canopy was first generated by applying a moving window (with a radius of 5.5 metres) and taking the maximum value in the window to represent the maximum local tree height for the centre pixel. This filter size was selected as large enough not to cause the resulting canopy top raster to fall into small gaps between trees whilst preserving the spatial variation in tree heights within the stand. Pixels in the original CHM were then classified as gaps if they had heights lower than 66% of the local height of the canopy top raster. All pixels in the CHM with a height of less than 10 m were also included in the gap class to account for those in the centre of large gaps.

The alternative method delineates gaps directly from the LiDAR point cloud without interpolation to a CHM. As many gaps contain large areas of shadow with few LiDAR returns, the algorithm focuses on the identification and delineation of areas of canopy, with gaps being found subsequently by default. The algorithm is composed of a number of stages including the identification of local maxima (as points higher than their neighbours), filtering to remove returns from below the canopy drip-line (less than 66% of local tree height), clustering of lower canopy returns into separate clusters around each identified maxima (limited by a radius corresponding to the maximum tree crown size) and merger and delineation of clustered points to retrieve gap delineations. The initial stages (identifying local maxima and clustering of returns) are similar to the approach taken by Tiede *et al.* (2005) to assign point data to individual tree crowns, but the clustering method differs by first filtering the point cloud to leave only returns from above the canopy drip-line. Figure 1 shows the full point-based processing scheme.

Following filtering and clustering, the locations and cluster memberships of the points are processed in a GIS environment to produce vector gap outlines. Any points that have not been assigned to a cluster are removed at this stage. First, a buffer (of fixed radius) is applied to the points, and those points with the corresponding cluster memberships (associated with the same maxima) are merged into a single polygon. Any 'holes' fully enclosed within a cluster polygon are removed leaving polygons representing tree crowns. The areas of test plots representing gaps (i.e. all areas not included in the canopy polygons) are then retrieved. The gap polygons are then dilated to reconstruct full gap extent. The point density of the data determines the optimum distances for buffers to merge points, with the radius of buffers approximately equal to the distance between adjacent scan lines in the along-track direction (the direction in which the largest point spacing was present). A small buffer remains surrounding the clustered canopy points (i.e. the gaps are not dilated by the full buffer distance), as such points are unlikely to be located precisely on the canopy edge. All specified parameters for each plot are presented in Table 1 and vary according to the LiDAR point density. The clustering radius (R) can be estimated as approximately double the average along-track spacing of LiDAR returns. The search radius, S, used for maxima location can be estimated from field data (or in the absence of such data, from optical or lidar intensity images) so as to be slightly larger than the average crown diameter.

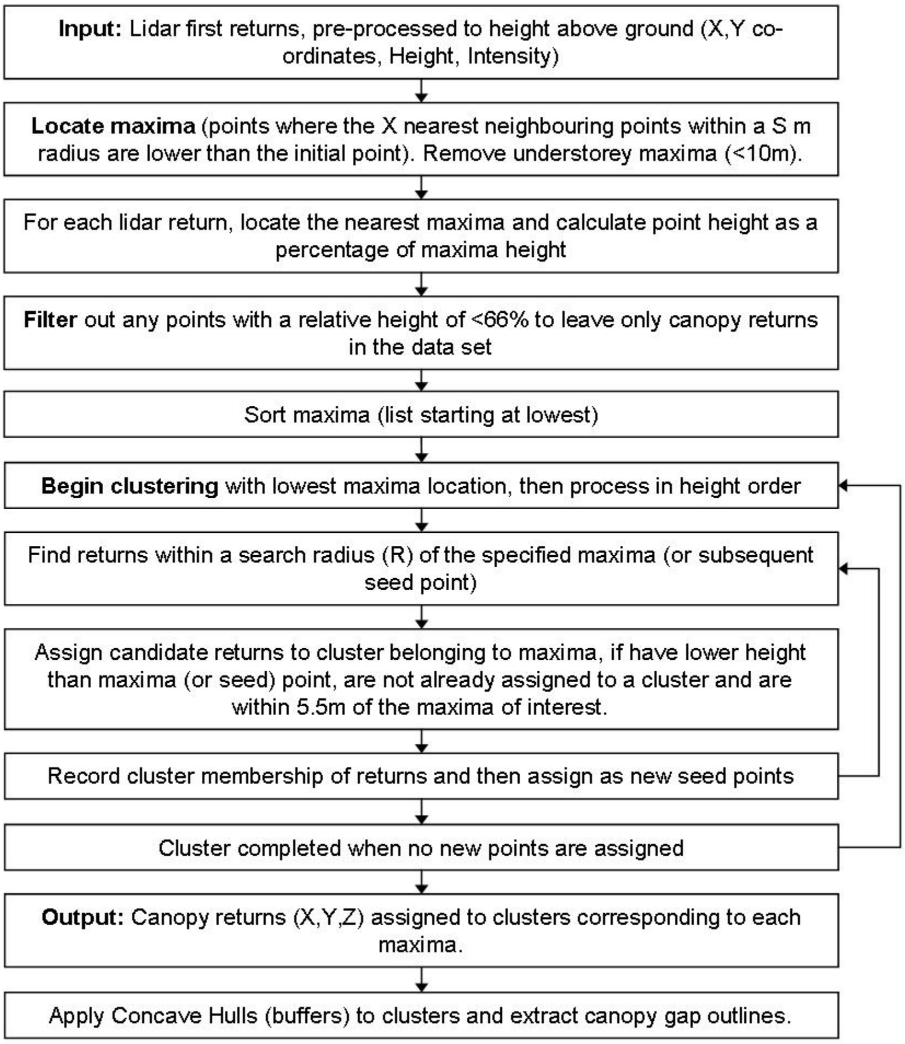

Figure 1: Point-based processing scheme for delineation of gaps from LiDAR data.

Table 1: Final parameters selected for point-based processing of each plot.

Plot	Number of neighbours (X)	Search radius (S)	Cluster radius (R)	Buffer distance	Distance gaps dilated
AbP1	100	3 m	1 m	0.5 m	0.25 m
AbP2	100	3 m	1 m	0.5 m	0.25 m
GLT	20	3 m	2 m	1 m	0.5 m
GHT	20	3 m	2 m	1 m	0.5 m
CLG1	20	3 m	2 m	1 m	0.5 m
CLG4	20	3 m	3 m	1.5 m	1 m

Before assessing accuracy against field data, any gaps identified as having an area of less than 5 m^2 were removed. In the case of the CHM-based method of gap delineation, accuracy was assessed by a pixel based comparison (confusion matrix) of the classified raster with a raster of the field mapped gap distributions for sub-plots of AberP1, CLG4, GHT and GLT. AberP2 was excluded from the assessment due to difficulties in accurately co-registering the field and LiDAR data. For the point based method, a confusion matrix was calculated directly from the vector delineations. The total area of canopy gap identified in each sub-plot using both methods was also compared with the overall area delineated in the field. A separate assessment of the accuracy of the maxima identification stage for the point-based method was also carried out for plots where tree positions were available (either field mapped or obtained visually from the CHM). A maxima was considered to correspond to a field mapped tree if it lay within a 2m radius of the field position (the average crown radius for the stands).

5. Results

Table 2 shows the correspondence between identified maxima (from the point-based method) and field mapped tree positions. Good levels of accuracy were achieved for all plots, with the majority of trees identified. Some commission errors occurred, but this is unlikely to have a significant effect on the performance of the algorithm as a whole.

Table 2: Accuracy of maxima identified from LiDAR first return point data compared to field measured tree positions (except for AbP1, where 'field' positions were manually identified from the CHM).

Plot	Producer's accuracy (%)	User's accuracy (%)
CLG1	85.7	75.2
CLG4	88.0	79.0
GLT	73.2	68.0
GHT	86.2	69.7
AbP1	94.8	83.8
Mean	85.6	75.1

Figure 2 shows an example of the results of the clustering stage of the algorithm for two plots, one with a high point density (AbP1) and the other with much lower point density (CLG4). In general the points form compact clusters around field mapped tree positions, although there are some instances in which clusters also include parts of neighbouring crowns.

The final gap delineations using this method for the same two plots are shown in Figure 3. The majority of gap areas are correctly identified in both cases but greater errors can be seen for the plot with lower density LiDAR data. This was confirmed by the results of the formal accuracy assessment, as presented in Table 3. The corresponding results for the CHM-based method are also included. Overall accuracy was slightly higher for all plots using the point-based method (an increase of 3.7% on average), whilst the producer's accuracy of the gap class was considerably improved in most cases (average increase of 8.3%). Table 4 compares the total derived gap area in each sub-plot to that mapped in the field. It can be seen that both the CHM and point-based methods generally under-estimate gap area. The point-based approach usually retrieves a greater gap area than the CHM-based method and the results are also more consistent between plots.

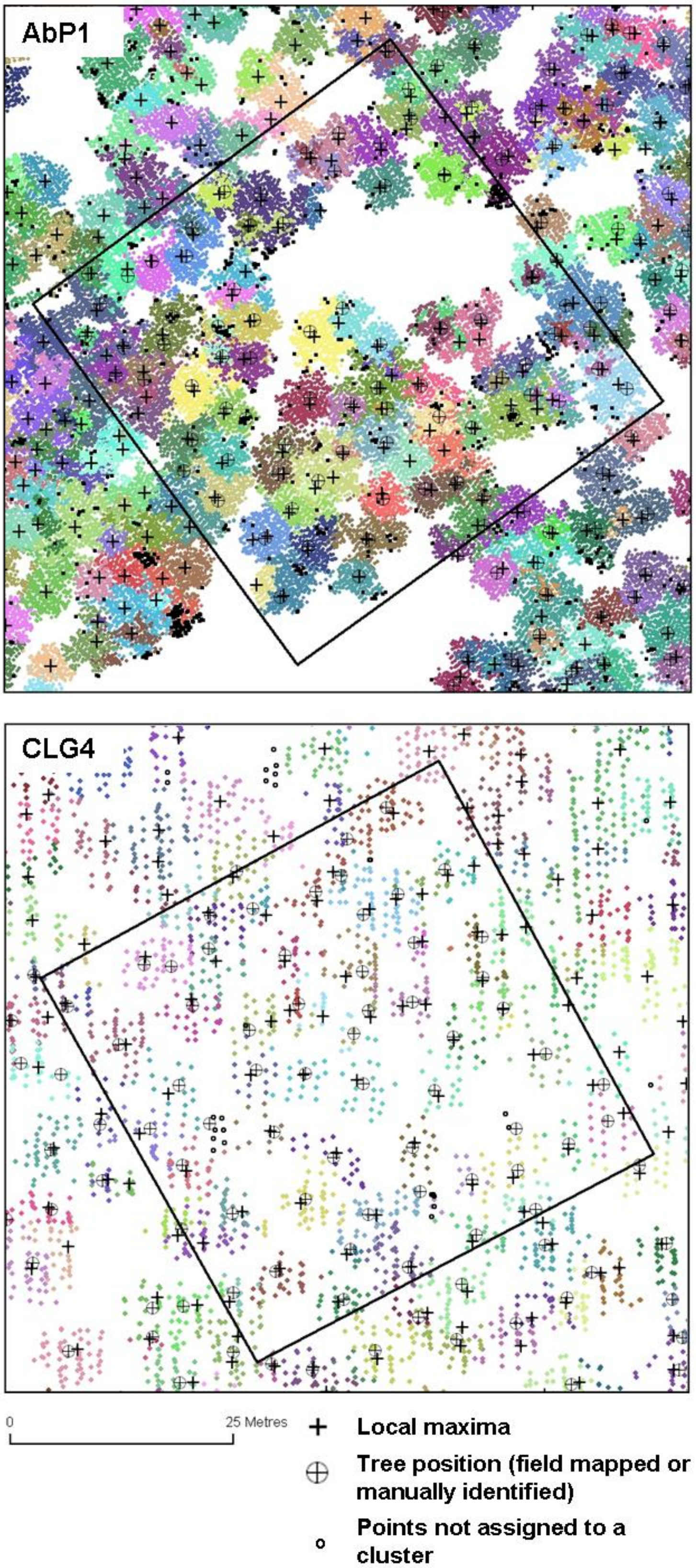

Figure 2: Examples of the results of the clustering stage of the point-based algorithm for plots CLG4 and AbP1. Locations of local maxima used as seeds and the locations of trees (field mapped for CLG4, visually identified for AbP1) are also shown. Colours are randomly assigned to clusters.

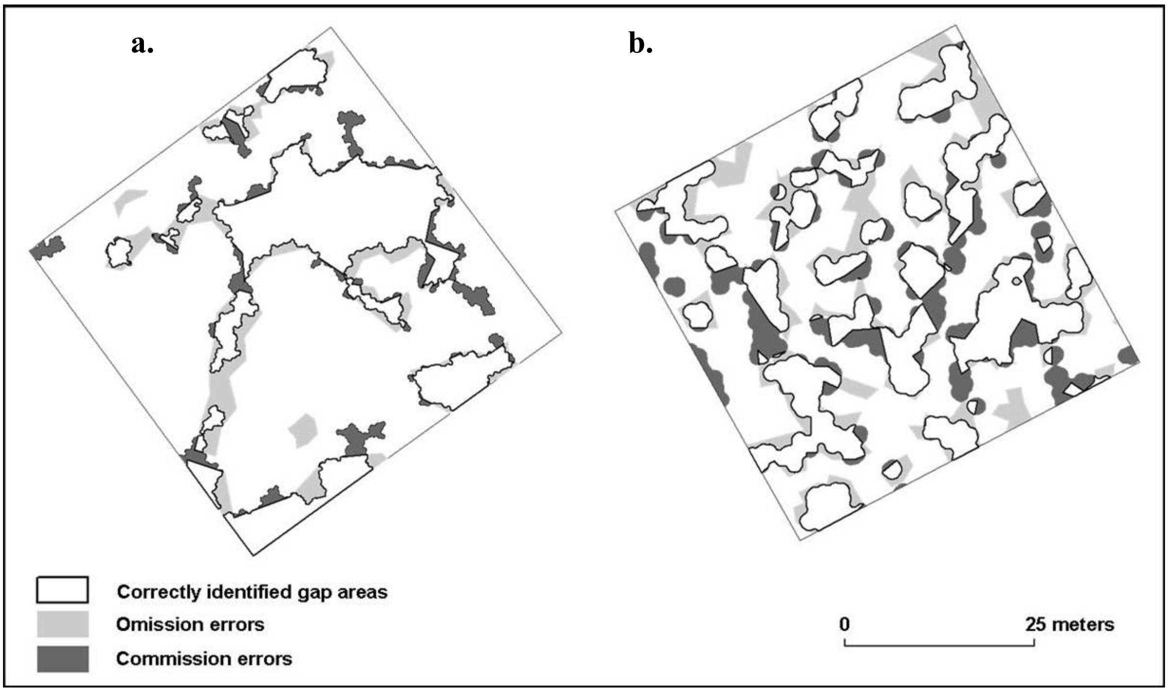

Figure 3: Results of the point-based gap delineation algorithm for plots AbP1 (a.) and CLG4 (b.).

Table 3: Confusion matrix results for a comparison of field mapped canopy gaps to point and CHM –based LiDAR delineations. The producer's and user's accuracies are those of the 'gap' class.

Plot	Method	Overall accuracy (%)	Producer's accuracy (%)	User's accuracy (%)	Kappa co-efficient
AbP1	Point-based	87.8	74.9	82.7	0.70
	CHM-based	85.2	79.7	73.1	0.66
CLG4	Point-based	77.2	69.3	72.5	0.52
	CHM-based	70.7	62.6	62.8	0.39
GHT	Point-based	72.8	61.4	70.9	0.43
	CHM-based	71.6	47.6	76.8	0.39
GLT	Point-based	74.8	46.7	81.0	0.43
	CHM-based	70.6	29.3	89.2	0.31
Mean	Point-based	78.2	63.1	76.8	0.52
	CHM-based	74.5	54.8	75.5	0.44

Table 4: Comparison of the total gap area identified within sub-plots by field mapping and LiDAR delineation using the point and CHM –based methods.

Plot	**Method**	**Field mapped gap area (m^2)**	**LiDAR derived gap area (m^2)**	**Error (%)**
AbP1	Point-based	750	679	-9.5
	CHM-based		817	9.0
CLG4	Point-based	1016	972	-4.3
	CHM-based		997	-1.9
GHT	Point-based	1080	935	-13.5
	CHM-based		671	-37.9
GLT	Point-based	957	552	-42.3
	CHM-based		316	-66.9
RMSE	Point-based			22.8
	CHM-based			38.7

6. Discussion and Conclusions

Both methods provide accurate gap delineations when applied to LiDAR data collected with a high point density. However, when lower point density data is used, the method based on an interpolated CHM can result in significant errors, probably due to the low number of returns located in canopy gaps and resulting interpolation errors. In these circumstances, the accuracy of gap retrieval can be improved by the use of methods based on the LiDAR point cloud, although the point-based method does require the careful selection of appropriate parameters for each data set. These can be estimated from basic field data (or the examination of optical imagery) and from the along-track point spacing of the LiDAR data, as described in the methodology.

It is not possible to compare the accuracy of the developed algorithm to those used in other studies of LiDAR gap delineation (Zhang 2007; Koukoulas and Blackburn 2004) as these studies did not attempt to assess the accuracy of resulting delineations. However, the results compare reasonably with those of Nuske *et al*. (2007) who used colour, texture and height information from aerial photographs to delineate gaps in Beech stands with a recall of 57-79% and a precision of 68-77% when compared to manual delineations. Whilst field and LiDAR delineations were felt to be generally comparable in this study, allowing the assessment of accuracy, the approach did have limitations. Errors in the GPS baselines used for the field survey led to error in registration of the field and LiDAR data for some plots. As only a limited number of points on the gap boundaries could be surveyed, the field boundaries are unavoidably generalised and some smaller gaps may have been missed altogether. These factors could account for a significant proportion of the remaining error in the delineations, suggesting 'true' accuracy may be higher than that reported.

LiDAR data can allow the accurate delineation of canopy gaps in stands in the process of transformation to CCF systems, although the influence of scan angle on results is yet to be determined. The resulting gap distributions may be used to develop indices of spatial structure that allow the monitoring of such stands but further work is needed to assess the usefulness of such indices and to integrate the spatially explicit gap information into models of understorey light levels.

Acknowledgements

This work was funded by a NERC CASE studentship with the Forestry Commission (NER/S/A/2004/12732). We would like to thank the NERC ARSF and the Forestry Commission for supplying LiDAR data and the NERC FSF and GEF for providing ground instrumentation. We are very grateful to Sophie Hale (Forest Research) and the Tyfiant Coed Project at the University of Wales, Bangor for access to field data and to all those who assisted during the fieldwork, especially V. Barron and P. Wieland. Particular thanks are due to J.C. Suárez and I.H. Woodhouse for advice and support.

References

Ahokas, E., Yu, X., Oksanen, J., Hyyppä, J., Kaartinen, H. and Hyyppä, H., 2005. Optimization of the scanning angle for countrywide laser scanning. In *Proceeedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, G. Vosselman and C. Brenner (Eds). XXXVI, Part 3/W19, 115-119.

Betts, H.D., Brown, L.J. and Stewart, G.H., 2005. Forest canopy gap detection and characterisation by the use of high resolution Digital Elevation Models. *New Zealand Journal of Ecology*, 29, 95-103.

Brokaw, N.V.L., 1982. The definition of a treefall gap and its effect on measures of forest dynamics. *Biotropica*, 14, 158-160.

Falkowski, M.J., Smith, A.M.S., Hudak, A.T., Gessler, P.E., Vierling, L.A. and Crookston, N.L., 2006. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. *Canadian Journal of Remote Sensing*, 32, 153-161.

Goodwin, N.R., Coops, N.C. and Culvenor, D.S., 2006. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sensing of Environment*, 103, 140-152.

Koukoulas , S. and Blackburn, G.A., 2004. Quantifying the spatial properties of forest canopy gaps using LiDAR imagery and GIS. *International Journal of Remote Sensing*, 25, 3049-3071.

Koukoulas, S. and Blackburn, G.A., 2005. Mapping individual tree location, height and species in broadleaved deciduous forest using airborne LiDAR and multi-spectral remotely sensed data. *International Journal of Remote Sensing*, 26, 431-455.

Kraus, K. and Pfeifer, N. 1998. Determination of terrain models in wooded areas with airborne laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53, 193-203.

Mason, B. and Kerr, G., 2004. Transforming even-aged conifer stands to continuous cover management. *Forestry Commission Information Note 40.* Forestry Commission, Edinburgh.

McGaughey, R., 2007. *FUSION / LDV: Software for LiDAR Data Analysis and Visualisation*. USDA Forest Service, Pacific Northwest Research Station.

Morsdorf, F., Frey, O., Meier, E., Itten, K. and Allgöwer, B., 2008. Assessment of the influence of flying height and scan angle on biophysical vegetation products derived from airborne laser scanning. *International Journal of Remote Sensing*, 29, 1387-1406.

Næsset, E., 2004. Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. *Remote Sensing of Environment*, 91, 243-255.

Nakashizuka, T., Katsuki, T. and Tanaka, H., 1995. Forest canopy structure analyzed using aerial photographs. *Ecological Research*, 10, 13-18.

Nuske, R., Ronneberger, O., Burkhardt, H. and Saborowski, J., 2007. Self-learning canopy gap mapping for aerial images using photogrammetric height, colour, and texture information. In: *Proceedings of ForestSat 2007*. Montpellier, France.

Page, L.M. and Cameron, A.D., 2006. Regeneration dynamics of Sitka spruce in artificially created forest gaps. *Forest Ecology and Management*, 221, 260-266.

Suárez, J., Ontiveros, C., Smith, S. and Snape, S., 2005. Use of airborne LiDAR and aerial photography in the estimation of individual tree heights in forestry. *Computers and Geosciences*, 31, 253-262.

Tiede, D., Hochleitner, G. and Blaschke, T., 2005. A full GIS-based workflow for tree identification and tree crown delineation using laser scanning. In: U. Stilla, F. Rottensteiner and S. Hinz (Eds.). *The Proceeedings of the ISPRS Workshop CMRT 2005*. Vienna, Austria: 9-14.

Zhang, K., 2007. Identification of gaps in mangrove forests with airborne LIDAR. *Remote Sensing of Environment*, doi:10.1016/j.rse.2007.10.003.