Sub-canopy terrain modelling for archaeological prospecting in forested areas through multiple-echo discrete-pulse laser ranging: a case study from Chopwell Wood, Tyne & Wear

LANDY, JACK

Sub-canopy terrain modelling for archaeological prospecting in forested areas through multiple-echo discrete-pulse laser ranging: a case study from Chopwell Wood, Tyne & Wear

Thesis submitted for Master of Science

Department of Geography, Durham University

Jack Landy, December 2010

Abstract

Airborne Light Detection and Ranging (LiDAR) technology is assessed for its effectiveness as a tool for measuring terrain under forest canopy. To evaluate the capability of multiple-return discrete-pulse airborne laser ranging for detecting and resolving sub-canopy archaeological features, LiDAR data were collected from a helicopter over a forest near Gateshead in July 2009. Coal mining and timber felling have characterised Chopwell Wood, a mixed coniferous and deciduous woodland of 360 hectares, since the Industrial Revolution. The state-of-the-art Optech ALTM 3100EA LiDAR system operated at 70,000 pulses per second and raw data were acquired over the study area at a point density of over 30 points per square metre.

Reference terrain elevation data were acquired on-site to 'train' the progressive densification filtering algorithm of Axelsson (1999; 2000) to identify laser reflections from the terrain surface. A number of sites, offering a variety of tree species, variable terrain roughness & gradient and understorey vegetation cover of varying density, were identified in the wood to assess the accuracy of filtered LiDAR terrain data. Results showed that the laser scanner over-estimated the elevation of reference terrain data by 13±17 cm under deciduous canopy and 23±18 cm under coniferous canopy. Terrain point density was calculated as 4.1 and 2.4 points per square metre under deciduous and coniferous forest, respectively. Classified terrain points were modelled with the kriging interpolation technique and topographic archaeological features, such as coal tubways (transportation routes) and areas of subsidence over relic mine shafts, were identified in digital terrain models (DTMs) using advanced exaggeration and artificial illumination techniques.

Airborne LiDAR is capable of recording high quality terrain data even under the most dense forest canopy, but the accuracy and density of terrain data are controlled by a combination of tree species, forest management practices and understorey vegetation.

Contents

List of Figures

List of Tables

List of Equations

Declaration

The material submitted in this thesis has not previously been submitted by the candidate for a degree in this or any other university.

Statement of Copyright

Acknowledgements

I would like to acknowledge One North East, the Friends of Chopwell Wood and the Forestry Commission for financial support which enabled me to carry out this study. In particular, I am grateful to David Woodhouse of the North East branch of the Forestry Commission for his guidance on forest mensuration practices, many hours of help in the field and assistance interpreting laser/canopy interactions. Also, Liz Searle of the Friends of Chopwell Wood for providing comprehensive information on the history and context of the forest's archaeological features and Paul Fountain of the Friends of Chopwell Wood for supplying valuable historic aerial photography. Special thanks go to Mervyn Brown for his advice on complex field methods throughout an extensive field season and to Dr Samantha Waugh for her assistance with field data registration procedures. Finally, I would like to express my appreciation of the support and supervision offered by Dr Danny Donoghue throughout this project and the breakthroughs that only came about following lengthy discussion with my postgraduate colleagues in the Geography Department at Durham University.

Chapter 1. Introduction and Literature Review 1: Airborne LiDAR

1.1 Introduction

Collaboration between academic research institutions and Regional Development Agencies (RDAs) in the United Kingdom offers an effective means of concentrating research and development on projects beneficial to the local community. One North East has provided funding to conduct research at a forest near Gateshead with a view to forming a working partnership between the University of Durham, the North East branch of the Forestry Commission and local community groups, including the Friends of Chopwell Wood. Additional complimentary funding has been provided by each of these groups. Through knowledge transfer between agencies, the University is able to offer the wider community access to an innovative new surveying technique, airborne laser scanning, while utilising the experience of the Forestry Commission in forest practices and theory and the knowledge of the Friends of Chopwell Wood on local history. In addition to contributing resources, each group benefits from the research; for example the Forestry Commission obtains information on the viability of introducing remote sensing technology to forest operations and the Friends of Chopwell Wood acquire otherwise inaccessible information of their local landscape. Cooperation between interested parties, with a motivation for regional development, enhances the quality of research while maintaining a relevant focus on local community applications.

1.1.1 Sub-forest canopy archaeology

Topographic feature detection within forested areas remains one of the few obstacles to modern archaeological prospecting. Areas covered by woodland are potentially concealing undetected archaeological features and consequently local history relating to these features remains unknown. Aerial reconnaissance and photography have been employed by

archaeologists for over a century to compliment terrestrial surveys, in addition to contemporary multispectral satellite remote sensing and image analysis (De Laet et al., 2007; Fowler, 2002; Neubauer, 2004). However these techniques can, in most cases, only recover detailed information in open areas without complex topography and/or vegetation (Barnes, 2003; Bewley et al., 2005).

Aerial photogrammetry offers the only method previously utilised for remote measurement of the terrain surface beneath forest canopy for archaeological applications (e.g. Bewley, 2003; Kucukkaya, 2004). Aerial photographs of the same location, but taken from multiple viewing angles, are manually or automatically joined with terrestrial tie-points (taken within gaps in the canopy) and elevation is triangulated (Baltsavias, 1999b). Accuracy is lost when few tie-points can be identified. Given the passive nature of the technique this is often due to shadow on the forest floor (Kraus and Pfeifer, 1998; Pfeifer and Mandlburger, 2009), but is also due to a lack of full automation and versatility (Nardinocchi et al., 2003).

In an attempt to address these issues, some research has been conducted into an alternative remote sensing method for deriving accurate and detailed sub-canopy terrain data. Innovative 'Light Detection and Ranging' (LiDAR) technology is capable of measuring over 100,000 points per second across the Earth's surface (Mallet and Bretar, 2009), through the transmission of a laser beam between a moving aerial platform (typically a plane or helicopter) and the Earth. The three-dimensional location of a single point is measured from the laser pulse time-of-flight and the GPS co-ordinates of the aerial platform. Given the high measurement rate, points at the Earth's surface are dense in space and thus regular penetration through even the thickest forest canopy is expected, advocating laser scanning (LiDAR) as a tool amenable to sub-canopy prospection.

1.1.2 Airborne laser scanning in archaeology

Pfeifer and Mandlburger (2009) noted that through the 2000s, the consensus between leading national agencies, whose primary function concerned the acquisition of terrain elevation data, gradually changed from generating photogrammetric models to LiDAR alternatives of greater point density and precision. Anticipating further commercial development in this field, I describe in this thesis the derivation of a high spatial measurement resolution sub-forest canopy terrain surface from airborne LiDAR data with a view to testing the limits of this tool for archaeological prospection.

Airborne laser scanning presents a non-invasive technique for acquiring topographic data in forested areas, eliminating the destructive facets of terrestrial surveying (Barber, 2007) and difficulties in feature interpretation at ground-level (Doneus et al., 2008). Barber (2007) recommended airborne scanning as the prime archaeological tool for topographic surveys over areas greater than 0.01 km^2, where a point measurement density of less than 1 point per m^2 is required (figure 1.1). It is appropriate where the scale of a survey is too large for terrestrial-based scanning or photogrammetry applications and/or when the requisite detail is prohibitive to ground-based GPS reconnaissance. Compared with the man-power required for both terrestrial methods, airborne LiDAR offers a cost-effective alternative.

Yet it is the capacity to digitally remove the forest canopy that makes LiDAR so valuable to archaeology. Early work by Barnes (2003) and Holden at el. (2002) hinted at the potential of LiDAR for recovering topographic earthworks, but only as a complement to more established remote sensing techniques and in sparsely vegetated areas. It was not until English Heritage commissioned the UK Environment Agency to assess the technique at the Stonehenge World Heritage Site between 2001 and 2005 that the 'virtual deforestation' function was properly explored (Bewley et al., 2005; Devereux et al., 2005; Haugerud and Harding, 2001). The assertion by Devereux et al. (2005) that Great Britain has approximately 11 percent woodland cover, 'with the ability to hide even major features from the lens of a camera', acted as a

stimulus for basic experimentation with vegetation-removal algorithms and testing against known sub-canopy field boundaries, fort ramparts, tracks etc. at Welshbury Hill, Gloucestershire (Crow, 2003; Crow et al., 2007). Further investigation into method and application followed, with notable contributions from Crutchley (2006; 2008), Challis (et al., 2006; et al., 2008) and Doneus (and Briese, 2006a; 2006b; et al., 2008).

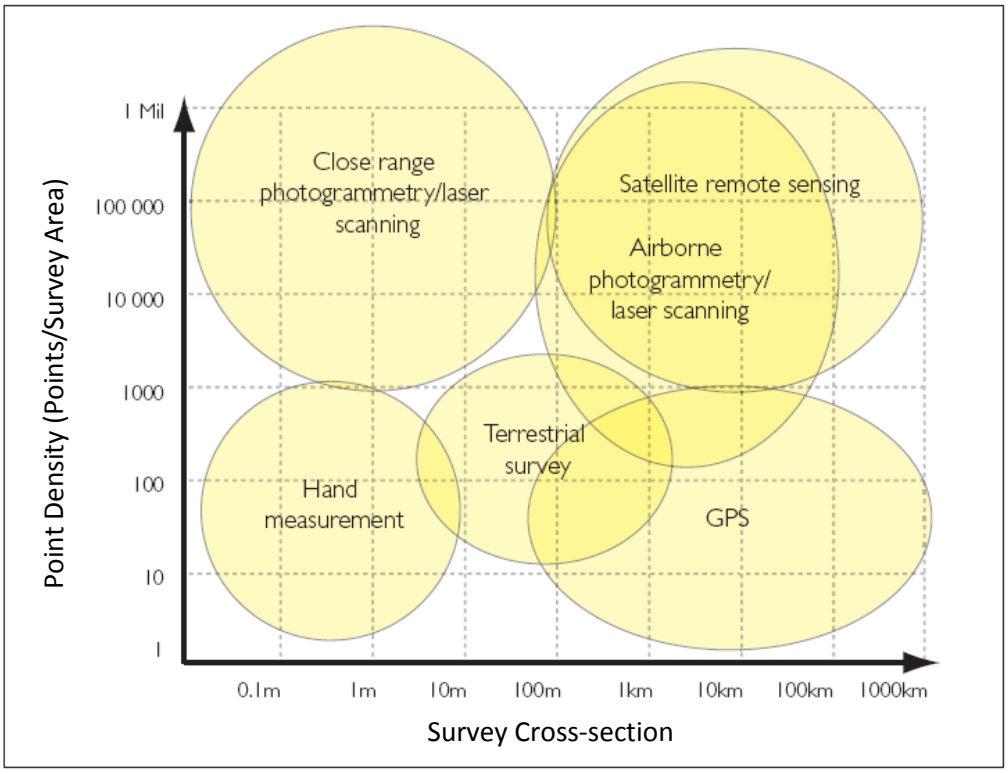


Figure 1.1 Appropriate archaeological survey techniques for recording three-dimensional data characterised by exterior and interior scale of survey. Adapted from Barber (2007).

The success of airborne laser scanning as a serviceable tool has, to some extent, limited research and development, with recent and prospective studies typically application oriented. Many practitioners have suggested that inadequate research has been directed towards establishing functional, specialised, archaeological vegetation-removal procedures (e.g. Devereux et al., 2005; Doneus, 2006a) and that auxiliary information on the role of canopy and understorey vegetation structure on terrain derivates is necessary for accurate archaeological feature identification (Pfeifer et al., 2004; Rosette et al., 2010). Investigation into the role of

both canopy and understorey vegetation could facilitate enhanced understanding of the capability of airborne LiDAR for conducting archaeological prospection in forested area. This study will examine the procedure for modelling sub-forest canopy topography with airborne LiDAR technology, the quality of modelled terrain and whether quality varies with environmental factors, e.g. tree species, and the capacity to detect archaeological features within topography.

1.2 Technical Laser Scanning

1.2.1 Defining a ranging system

A basic laser ranging system measures the time for a pulse of light to travel between the sensor and an illuminated spot on the ground. The distance (or 'range') *R* is subsequently calculated as:

$$R = \frac{1}{2} t_L c \tag{1.1}$$

where t_L is the travel time and *c* is the speed of light through air (Baltsavias, 1999a). Ranging systems typically operate under one of two principal concepts: discrete-pulse or full-waveform, both referring to the procedure by which reflected light is recorded in the sensor. Discrete-pulse systems can differentiate between one and five returns per pulse, i.e. they record the first and last 'echoes' from reflective surfaces within the laser footprint and up to 3 echoes in between. Full-waveform systems record the phase difference between the received and transmitted signal, subsequently permitting digitisation of the full waveform of surface interactions (Wehr and Lohr, 1999).

Both concepts are illustrated in figure 1.2, with the true reflected signal (power) recorded as five discrete echoes and also fully digitised (see the 'Ranging Unit' box). Noticeable in figure 1.2, and something that will be discussed in greater length below, is the typical identification of a mixture of canopy, scrub and terrain in the reflected signal (over forested areas). While

modern high-energy discrete-pulse or full-waveform LiDAR systems are capable of recording up to 5 or effectively unlimited echoes, respectively, per pulse, research to date into laser scanning for deriving terrain beneath forest, including that applied to archaeology, has commonly used discrete-pulse systems with the capacity to resolve only the first and last echo.

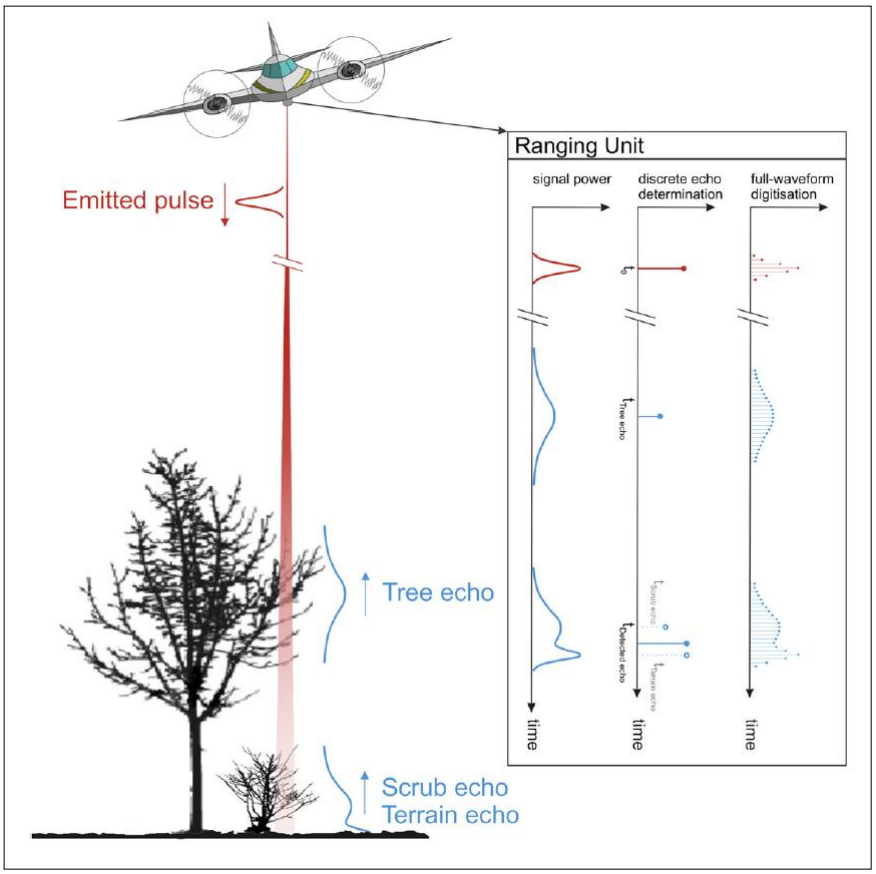


Fig 1.2 Laser ranging principle and differences between conventional discrete-pulse and full-waveform sensors. Source: Doneus et al. (2008).

Laser pulses are emitted regularly from the sensor by one of a number of scanning instruments, including oscillating mirror, rotating polygon, Palmer and fibre scanners. For more information the reader is referred to Wehr and Lohr (1999). An oscillating mirror apparatus is utilised in this study, with the laser deflected by an oscillating mirror to produce a zigzag line of points termed a bidirectional scan (Wehr and Lohr, 1999). State-of-the-art LiDAR systems can emit up to 100,000 pulses every second by this method, with each echo recorded upon return

as co-ordinates *x , y, z* and *intensity*, where intensity is the power of the returning signal. A group of echoes (or returns) in the same spatial assemblage is termed a 'point cloud', which is well defined by Barber (2007) as 'a collection of co-ordinates in a common co-ordinate system that portrays to the viewer an understanding of the spatial distribution of a subject'.

In addition to the scanning and recording apparatus, the full LiDAR ranging unit consists of a Position and Orientation System (POS) and computerised control system. While the detector measures the range, through travel time, and receiving angle of a laser return, the POS locates the position and incidence angle of the scanner aperture in space to an accuracy at least as high as the laser range accuracy (Wehr and Lohr, 1999). Location of the airborne platform (airplane or helicopter) is calculated with a differential Global Positioning System (dGPS) (communicating with a base station on the ground) and the vector between the GPS and scanner aperture and detector (which is constantly changing with aircraft attitude) with an Inertial Measurement Unit (IMU). Scanning and positioning apparatus are synchronised by the computer control system.

1.2.2 Technical sensing

It is necessary to clarify the theory associated with the laser and scanning system that is relevant to terrain derivation. A comprehensive formula-oriented review of LiDAR technology is presented by Baltsavias (1999a) and concepts offered in this section are either obtained directly, indirectly or based on theory from this review.

Maximum laser range is proportional to the square root of the laser power and the square root of the surface reflectivity, i.e. the magnitude of received energy following scattering and absorption at the land surface. Variation in maximum laser range with target reflectivity is given in figure 1.3, with variation in range described as a correction factor based on target reflectivity of 80 % providing a factor of 1. As an example of materials relevant to this study, deciduous trees, coniferous trees and typical sub-canopy terrain are approximately 60, 30 and

20 % reflective, respectively (modified from the web pages of the firm Riegl in Wehr and Lohr, 1999). This suggests that maximum range is roughly twice as large over the most reflective surfaces than the least reflective ones, but since this study is concerned with identifying terrain, maximum range here is limited to the range for ~ 20 % surface reflectivity. Range resolution is directly proportional to the time interval of the detector clock and specifies a threshold for differentiating between laser returns, consequently regulating the number of echoes recorded per laser pulse.

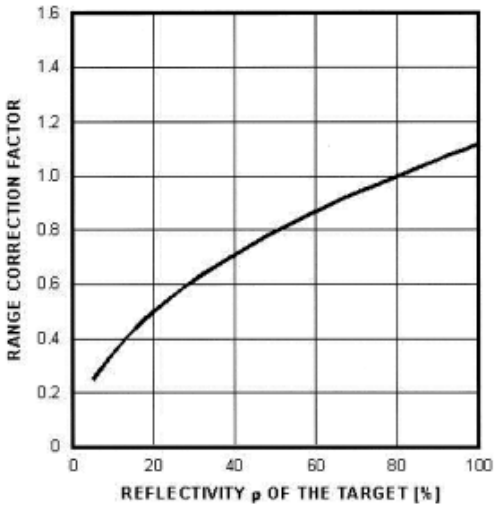


Figure 1.3 Correction factor for maximum laser range, depending on target reflectivity (900nm wavelength laser with diffuse targets). Source: Wehr and Lohr (1999).

Ranging precision σ_R is a product of the ratio between photoelectric signal and noise:

$$\sigma_R \sim \frac{1}{\sqrt{S/N}} \tag{1.2}$$

This describes how clear the echo reflected from an obvious land surface is, in amongst background noise. 'Noise' originates both naturally and artificially from numerous sources including complex pulse reflection, e.g. from scrub and terrain, and imperfect laser pulse emission (Huising and Gomes Pereira, 1998). The sensor in the ranging unit is calibrated pre-survey to reject candidate surface 'returns' that exhibit low signal to noise – where a significant deviation from background radiation cannot be detected. Supplementary both to

this relation and the range resolution for understanding limits on detection is the horizontal diameter of the minimum detectable object on the ground Λ, given by:

$$\Lambda \sim (A_L)^2 \rho \tag{1.3}$$

where A_L is the laser footprint diameter and ρ is the reflectivity at maximum range (adapted from Baltsavias, 1999a). Λ depends primarily on the structure of the target, which controls the quantity of energy reflected, but also on factors affecting the size of the laser footprint, like terrain gradient for example (see section 3.5.2.3). This relation is vital for archaeological feature detection owing to the frequent subtlety of relevant objects, e.g. earthworks, within topography.

The quality of a model generated from LiDAR-derived terrain measurements is enhanced with an increasing number of data points. By assuming (1) equal distance between returns on the scan-line and (2) flat terrain, across-track point spacing, dx_{across} is given by:

$$dx_{across} = \frac{SW}{N} = h\frac{\theta_{max}}{N-1} \tag{1.4}$$

where *SW* is the swath width, *N* is the number of points per scan-line, *h* is the flying height and θ_{max} is the maximum laser pointing angle with respect to the nadir. Along-track point spacing is directly proportional to the speed of the aircraft and the pulse rate. Each flight track produces a strip of points and the strips are then geo-referenced to each other using analogous pairs of tie-points and finally registered to a single point cloud (Kilian et al., 1996). The point cloud is validated against field-acquired reference data by calibrating an area of points in the survey area to an asphalt or concrete surface, for example a car park, measured with surveying equipment on-site, and removing any systematic bias in height from the point cloud. Kraus and Pfeifer (1998) emphasized the necessity of high quality calibration, with disruption to their final terrain model highlighted in their figure 9 (in Kraus and Pfeifer, 1998).

1.2.3 Sources of error associated with the LiDAR point cloud

Pfeifer et al. (2004) outlined the various factors affecting the 'journey' of a single laser pulse, with error in the raw point cloud principally generated during: geo-referencing; determination of the angle of emission; emission; travel through the atmosphere; interaction (i.e. spectral and diffuse reflection and absorption) with the ground, structures and vegetation; travel back; and finally signal detection and time measurement.

These sources of error can be grouped by error relating to position, to optical transfer and to detection, excluding error relating to interface with the land surface. Positional error is either random, originating from poor communication between the aircraft and GPS base station or insufficient satellite reception at the base station, or systematic, originating from poor quality control/tie points and subsequent mis-calibration of the whole point cloud (Baltsavias, 1999a; Kraus and Pfeifer, 1998). In-flight optical error, typically random, is primarily caused by attenuation and scattering off dust particles (smoke, smog) or precipitation, with ideal surveying conditions therefore very dependent on the weather (Baltsavias, 1999a; 1999b). Error in laser return detection has already been touched on and is strongly related to the sensor range resolution, which controls the minimum detectable height difference between echoes and is a critical restriction in LiDAR-based archaeology (Kraus, 2004).

Experiments with the earliest terrain models often identified systematic overestimation of the true terrain surface. Early suggestions for the cause of this bias were geoid undulations and subsequent mis-registration of the point cloud or GPS phase ambiguities (Kraus and Pfeifer, 1998), but although quite possible, neither accounted for the degree of bias. The height shift has since been explained by low vegetation, with the systematic overestimation σ illustrated in figure 1.4. Pfeifer et al. (2004) described characteristic offset over long grass of 8 cm and over low vegetation of 15 cm, while Huising and Gomes Pereira (1998) suggested decimetre-to-metre-scale systematic overestimation over grass and scrub, on flat or hilly terrain. A typical laser scanner is able to discriminate between successive pulses if they are greater than the

pulse duration of 5 ns apart, corresponding to a minimum detectable height difference on the ground of approximately 15 cm (Kraus, 2004; Doneus and Briese, 2006b). Even if discrete echoes are received both from the vegetation and terrain, if they are within the threshold minimum detectable height difference, the 'terrain' signal will be recorded as a mixture of ground and understorey/bush.

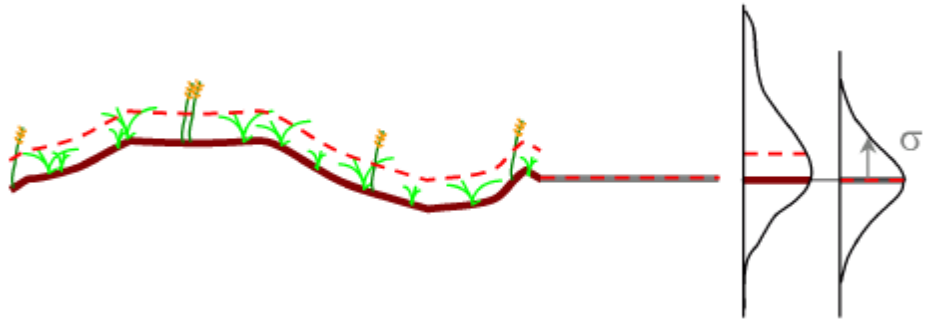


Figure 1.4 Systematic upward shift *σ* in laser measurements over low vegetation. The dashed line illustrates the expectancy of LiDAR-derived height values. Source: Pfeifer et al. (2004).

Various correction methods for different species/levels of understorey have been tested, for example using (1) reference data and linear regression of bias (i.e. empirical) (Hopkinson et al., 2004), (2) texture analysis by identification of grey-level co-occurrences between pairs of points (Pfeifer et al., 2004) and (3) using echo width in full waveform data (see section 2.3) (Doneus et al., 2008). However, these techniques typically either require intensive and comprehensive field data collection or offer only limited success. Consequently with falling restrictions on detectable height difference in rapidly evolving modern LiDAR systems, recent research has tended to ignore this problem or find ways of removing erroneous areas following data processing (e.g. Hesse, 2010).

The final significant source of error also originates from interaction of the laser pulse with the land surface. So-called 'time-walk' is caused by horizontal and especially vertical spreading of the laser pulse on sloped terrain (figure 1.5), leading to deterioration of the signal and increasing rise time *dt*, the latter breeding ranging error *dR*. Rise time describes the time required for pulse energy returning to the sensor to rise from 10 % to 90 % of maximum power

(Baltsavias, 1999a). Error in resolved terrain elevation increases when a measurement is taken across a wider area (figure 1.5 A) and this is illustrated in the case of the laser footprint in figure 1.5 B. Ranging error is proportional to the width of the footprint, which itself is a product of the height *h* and scanning angle *γ* (Baltsavias, 1999a).

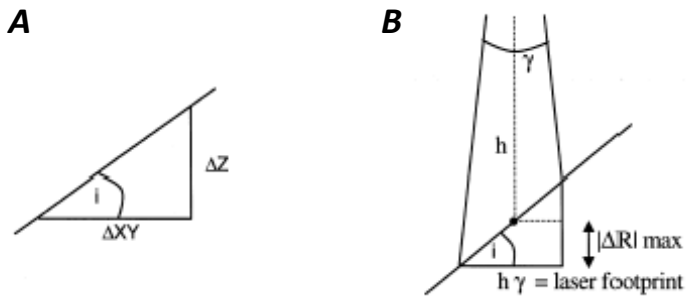


Figure 1.5 (A) Relation of planimetric and height errors over sloped terrain and **(B)** the influence of sloped terrain on measured range. Source: Baltsavias (1999a).

1.3 Sub-forest Canopy Three-dimensional Terrain Modelling

Following the acquisition of raw LiDAR data, it is necessary to process the point cloud in order to obtain digital models of the topography. Flood (2001) estimated that the data processing stage, as opposed to the acquisition and registration stages, typically requires approximately 60 to 80 % of total survey time. With agencies who supply sub-forest canopy terrain models limited in the level of accuracy they are able to provide to a client by the quality of data processing, there is significant commercial potential in designing proficient processing software. As a result, extensive research has been afforded to this issue and a choice must be made between numerous brands of processing and modelling software.

1.3.1 Defining the processing chain

It is necessary to outline the sequence of procedures involved in processing raw LiDAR data prior to examining the research conducted into each procedure. The registration and calibration stage has already been described in section 1.2.2 and is characteristically part of the data acquisition routine.

Essential to generating high quality terrain models is the process of *filtering* terrain points from the original point cloud. Points are classified into either terrain or non-terrain categories using computerised algorithms (see section 1.3.3.1). Terrain and non-terrain classifications refer to the attributes of the surface reflecting the laser beam, with examples of non-terrain surfaces including buildings and vegetation. While *classification* describes the process of 'labelling' points as terrain, building, vegetation etc. (Sithole, 2005), *filtering* explicitly refers to the removal of unwanted non-terrain points from the point cloud, leaving only returns from the terrain surface (Pfeifer and Mandlburger, 2009).

Filtered point cloud data, i.e. terrain points, are subsequently used to generate a digital model of the terrain and this may feasibly be seen as the final stage of processing. However, with a view to ascertaining the quality of a completed model, accuracy assessment is occasionally conducted in a further processing step. This assessment typically involves quantitative evaluation of the terrain model against field-acquired reference terrain data.

The full data processing chain comprises: (1) classifying the raw LiDAR point cloud and filtering only terrain points; (2) modelling terrain with the filtered data and (3) assessing the accuracy of the model with respect to reference data.

1.3.2 Processing and modelling terminology

With the development of LiDAR as an established technology, processing and modelling nomenclature has evolved to deal with the complexity of the subject. At the point cloud

classification stage, *optimum filtering parameters* refer to the parameters selected for a filtering algorithm which offer most successful classification of terrain points. *Type I (omission) errors* refer to the false rejection of terrain points from the terrain category and *Type II (commission) errors* to the mistaken acceptance of non-terrain points in the terrain category, e.g. rejection of returns from a hilltop and acceptance of returns from low bushes in the terrain category, respectively (Pfeifer and Mandlburger, 2009). For many algorithms, *training data* are used to derive optimum filtering parameters, and these are typically *in situ* field-acquired elevation reference data which *train* the filter to a particular location or environment (see section 1.3.3.3) (Zhang et al., 2003).

At the modelling stage, terrain points are typically either used to generate a raster surface or a triangulated irregular network (TIN). Raster surfaces comprise regularly-gridded cellular arrays with cell elevation modelled by a continuous function that maps from the 2D planimetric position of LiDAR points to terrain elevation, $z = f(x,y)$ (El-Sheimy et al., 2005). This function utilises the so-called *2.5D principle*, where each cell position (x,y) can only be expressed by one height value (Tóvári and Pfeifer, 2005). Where a cell does not correspond with LiDAR data in plan, e.g. in gaps between points, elevation values are interpolated by a pre-described mathematical function (see section 1.3.5). Raster surfaces generated with LiDAR data are termed digital elevation models (DEMs). Where the model describes a surface over the highest points in the LiDAR cloud, i.e. the top of tree canopies and buildings in addition to terrain in open areas, it is termed, interchangeably, a digital surface or canopy model (DSM/DCM) (Pfeifer and Mandlburger, 2009). More relevant to this study however is the digital terrain model (DTM) which describes only the terrain or 'bare-earth' (i.e. ground without even scrub vegetation) (El-Sheimy et al., 2005). The triangulated irregular network (TIN) represents a vector- (rather than raster-) based method for modelling LiDAR points, with points (or nodes) connected in a tetrahedral lattice structure (Li et al., 2005). The relative merits of each are evaluated in sections 1.3.5 and 3.5.1.

1.3.3 LiDAR point cloud classification

It is pertinent, with so many varieties of terrain filter algorithm, to outline the concept behind each general group of algorithms prior to reviewing their relative efficacy. While additional algorithms for classifying further groups of LiDAR points, i.e. vegetation and buildings, are available and numerous, they are not directly relevant to this study and, as such, will not be discussed. So the following are purely algorithms published for filtering terrain points from raw LiDAR point cloud data. All filters will characteristically assess the correlation of a 'candidate' terrain point to an initial surface or its nearest neighbour points through statistical analyses, thereby grouping contiguous points, likely reflected by the same target.

1.3.3.1 Semi-chronological review of terrain filtering algorithms

Morphological – block minimum (Lindenberger, 1993)

In this filter, the lowest points in a 'neighbourhood' are assumed to belong to the bare earth. An initial terrain model is refined through repeating the filtering procedure, a two-step process, with increasingly smaller moving windows. Points within the filtering window (or neighbourhood) are only evaluated against each other and not the full point cloud. At each stage, a coarse model is first calculated using the lowest point in each window and all additional points exceeding a given threshold height from this surface are then filtered out.

The concept is illustrated in figure 1.6, with the circles representing raw LiDAR points, the dotted line representing the true terrain surface and the vertical lines representing the regular boundaries of the moving window. Within each 'block' (neighbourhood), minimum point elevation is given by the flat horizontal line with the grey region above this line representing the threshold for accepting further points. Raw LiDAR points are classified as terrain (white) or non-terrain (grey). It is apparent that while areas of 'terrain'-classified points correspond with the true terrain surface, at four example scenarios (a-d in figure 1.6) LiDAR points are

incorrectly classified. At (a), the window size is smaller than a building so reflections from the roof are classified as terrain; at (b), over steep terrain the threshold is too low causing Type I errors; at (c), points over a bridge are commissioned as terrain (although this is the case with many filters and is often required by the data user, i.e. a bridge or ramp should be classified as terrain) and at (d), a sub-ground outlier (see section 1.3.4) causes Type I errors at the overlying terrain (Sithole and Vosselman, 2004; Sithole, 2005).

Lindenberger's (1993) algorithm was developed by Kilian et al. (1996), who assigned weights to points within a less severe height threshold based on the elevation difference between a LiDAR point and the lowest point in the neighbourhood and proportional to the size of moving window, with the ultimate surface a function of points with highest final weight. However, many of the same problems with the original algorithm remain.

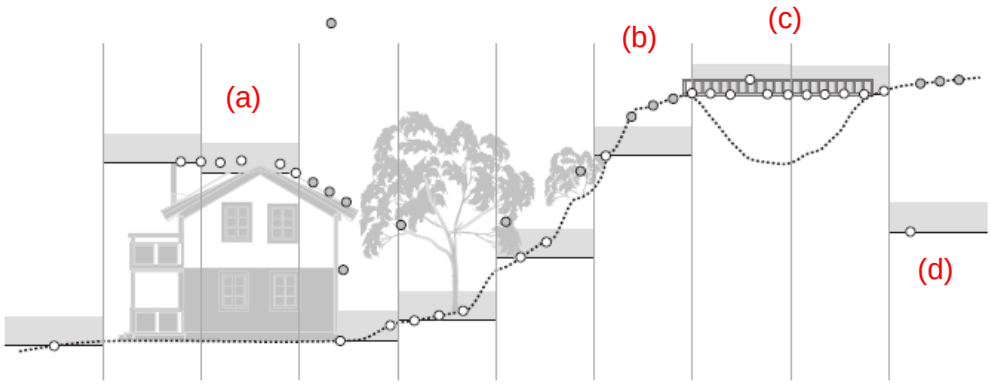


Figure 1.6 Visual demonstration of the block minimum procedure. Raw LiDAR points are classified as terrain (white circles) and non-terrain (grey circles). Letters are referred to in the text. Source: Sithole (2005).

Hierarchical surface regularisation (Briese et al., 2002; Kraus and Pfeifer, 1998)

In this filter, the bare earth is assumed to constitute a patchwork of piecewise, continuous overlapping surfaces. A rough first approximation of the terrain surface is calculated by performing robust kriging interpolation on initial terrain points and the vertical distances (residuals) of all additional LiDAR points to this surface are then calculated. Classification is

performed in a hierarchic method with a coarse to fine approach, where only points within a predefined threshold from the coarser level are selected as initial terrain points at the finer level. At each level, weights are allocated to LiDAR points based on their residual and position above or below the terrain surface (figure 1.7). Following weight allocation, the surface is recomputed and the weights recalculated; this iterative, hierarchical process continuing until the surface changes insignificantly with further iterations.

Early development of the algorithm was conducted by Kraus and Pfeifer (1998) and Kraus and Rieger (1999), while Kraus and Pfeifer (2001), Pfeifer and Stadler (2001) and Pfeifer et al. (2001) improved the hierarchical approach. It was observed that the filter tended to smoothen out discontinuities in the true terrain surface, so Briese et al. (2002) enhanced the regularisation technique to incorporate breaklines in the final model. Breaklines are linear features describing local terrain discontinuities (Briese et al., 2009) and can be explicitly modelled as part of a digital terrain product (e.g. Briese, 2004; et al., 2009; Brugelmann, 2004; Brzank et al., 2005). More recently, Goepfert et al. (2008) attempted to include LiDAR intensity data in the surface regularisation algorithm by modelling point-to-point height shifts against the quantity of energy returning to the airborne sensor, but with only limited success.

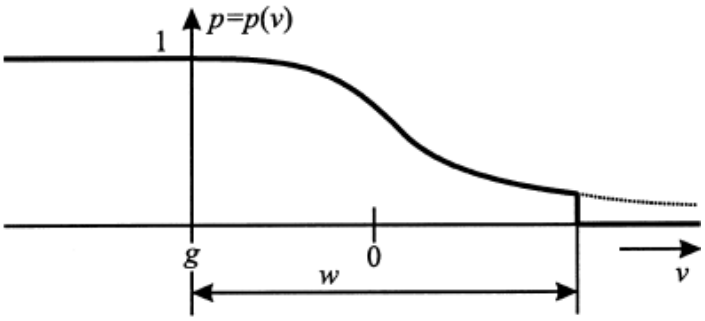


Figure 1.7 Weight function of the surface regularisation filtering algorithm. The x-axis describes the residual height of a point to the interpolated surface *(v)* while the y-axis gives the weight *(p)* from 0 to 1. Above the surface (residual > 0) the weight gets progressively lower and a LiDAR point has less effect on the terrain surface. Below the surface (residual < 0) the weight gets progressively higher and a point has more effect on the terrain surface, up to a threshold *(g)* whereupon the point has maximum effect (p = 1). If a candidate point is at *w* above *g*, the point is classified as non-terrain and ignored in future iterations. Source: Kraus and Pfeifer (1998).

Progressive TIN densification (Axelsson, 1999; 2000)

In this filter, the bare earth is assumed locally and globally flat. A first approximation of the terrain is obtained by block minimum, with the lowest point per x m^2 classified as terrain and triangulated to a TIN. The size of the neighbourhood must be larger than the size of the largest building in the LiDAR point cloud to prevent initial classification of terrain points on the tops of buildings. Further points are classified as terrain and added iteratively to the TIN if they fall within pre-defined description length and angle criteria, i.e. a point must not be over a threshold distance or angle from the terrain surface. In this way, the modelled surface almost 'fluctuates' up beneath the true terrain with newly classified terrain points repositioning the model closer to the real surface level.

The concept is illustrated in figure 1.8, with symbols equivalent to those in figure 1.6 and the bold line representing the modelled terrain surface. At (a), the description length and angle criteria are illustrated; at (b), by the third iteration step a reflection from low vegetation is still included in the terrain (Type II error) and at (c), a sub-ground outlier is classified as terrain at the initial block minimum stage causing Type I errors at the LiDAR points above (Sithole and

Vosselman, 2004; Sithole, 2005). However, realistically, prior to processing LiDAR data, outliers are typically removed using alternative (but very simple) algorithms and the effects of (c) are eliminated.

Later variants on this filter used the lower part of the convex hull of the point cloud as the initial reference surface and tested candidate points by the vertical distance to this surface (Von Hansen and Vögtle, 1999; Vögtle and Steinle, 2003). Sohn and Dowman (2002) also developed this method using a segmentation-based downward 'divide-and-conquer' optimality criterion, in addition to the upward densification process, to help to reduce the Type II errors illustrated in figure 1.8 (b), thus enhancing filter performance.

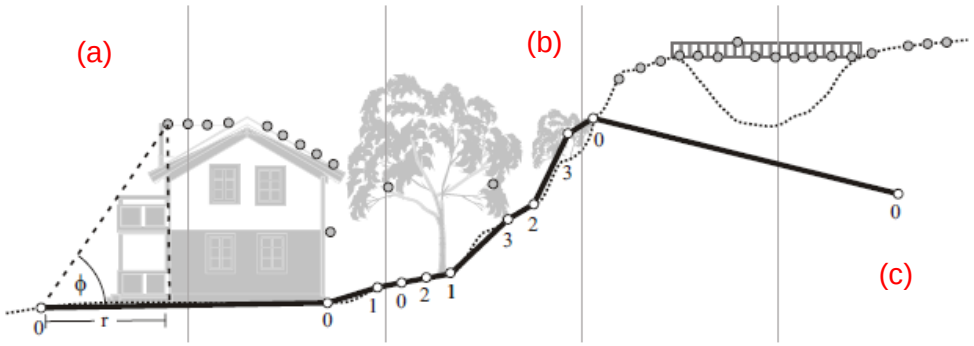


Figure 1.8 Visual demonstration of the progressive densification procedure. Raw LiDAR points are classified as terrain (white circles) and non-terrain (grey circles), with the number next to a point referring to classification at a particular iteration. Letters are referred to in the text. Source: Sithole (2005).

Morphological – adaptive slope-based (Sithole and Vosselman, 2001; Vosselman, 2000)

In this filter, it is assumed that gradients in the bare earth are locally bounded, i.e. terrain is not continuous and local surfaces do not smoothly link up with each other. Fundamental to this filter is the concept of the 'structure element' which is a function describing admissible height differences between candidate LiDAR terrain points as a function of the horizontal distance between them (figure 1.9 B). The structure element is shaped like an inverted funnel or cone (figure 1.9 A) and classifies points one-by-one with respect to their neighbours. The

function for the structure element must be determined from either assumptions made on the maximum terrain slope in the area of interest and associated computation of the probable threshold admissible height difference (probabilistic curve in figure 1.9 B) or field-acquired terrain reference data and evaluation of the maximum admissible height difference (maximum curve in figure 1.9 B) (Pfeifer and Mandlburger, 2009). Reference data consist of correctly classified terrain and non-terrain points in an environment representative of the full area of interest (see section 1.3.3.3) (Vosselman, 2000).

Morphological filters of this type are liable to fail when their relatively crude assumption fails, for example over very steep slopes and hilltops. Developments have been made to the early algorithms in an attempt to *adapt* the structure element to a local estimate of the terrain using contextual information. A coarse slope map and crude DTM obtained from block minimum filtering (Sithole and Vosselman, 2001) and local plane fitting (Roggero, 2001) have been utilised and changing the size of the structure element has been tested (Zhang et al., 2003), while more recent workers have inclined the element with the terrain and used repetitive interpolation to refine the model (Kobler et al., 2007). Meng et al. (2009) added a further step to standard morphological filtering procedure by removing LiDAR points above a threshold slope (with respect to local profiles through points), but crucially with a threshold that changes based on the original scanning direction of the airborne sensor.

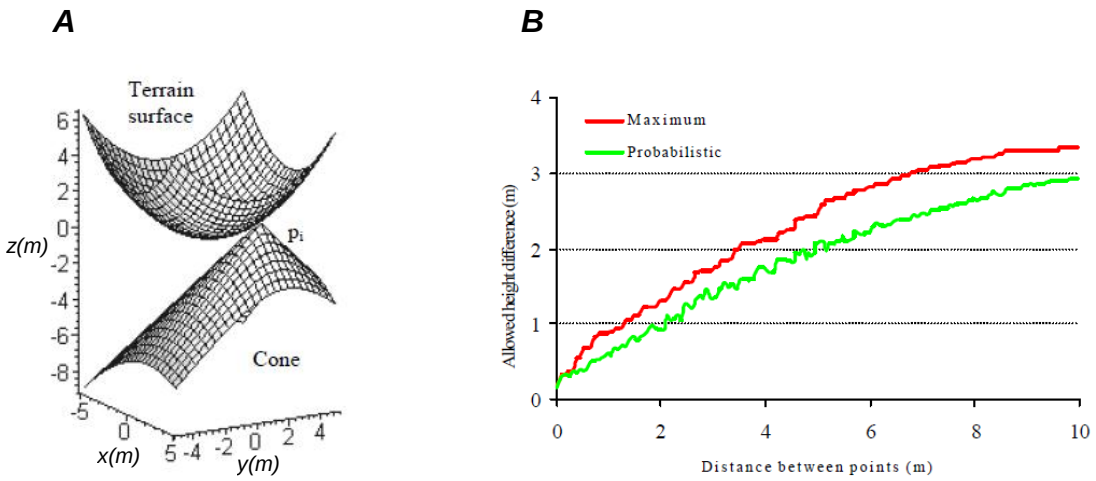


Figure 1.9 (A) Conical structure element moving between candidate LiDAR terrain points on the upper surface. Source: Sithole and Vosselman (2001). **(B)** Two filter functions for assessing admissible height differences between LiDAR terrain points. Threshold acceptable height difference changes exponentially with distance between points for both functions. Source: Vosselman (2000).

TIN thinning and de-spiking (Haugerud and Harding, 2001)

In this filter, it is assumed that curvature in the bare earth is locally bounded. A TIN is generated between all LiDAR points in the cloud and points that exhibit sharp upwards convexity are preferentially removed. Given that the geometry of the surface changes as points are removed, the procedure is iterated and the surface at each step is interpolated by linear prediction to a new TIN. This process helps to maintain topographic discontinuity but typically over 10 iterations are necessary to achieve high geomorphologic quality in the final terrain model, with associated heavy demand on computing time.

Active shape models (Elmqvist, 2001; et al., 2001; 2002)

In this filter, the bare earth is assumed to constitute a patchwork of piecewise, continuous overlapping surfaces. A 'membrane' is floated up from beneath the LiDAR point cloud and allowed to cling to low points. The membrane consists of a surface of active shape models whose material properties (i.e. the forces on the membrane) are defined by an internal energy

function. The energy function is minimised when the active contours find edges in the terrain and the material properties of the membrane (for example: rigidity, elasticity) control the structure of the modelled terrain, i.e. the degree of topographic discontinuity. Any LiDAR point within a threshold height of the membrane is classified as terrain following minimisation of the energy function.

Edge-based clustering (Brovelli et al., 2002; Filin, 2002)

In this filter, it is assumed that points inside closed edges belong to objects (non-terrain), i.e. terrain is unbounded. The full LiDAR point cloud is first gridded and gaps are modelled using spline interpolation. Residual differences in elevation between LiDAR points and the grid are calculated and standard image analysis edge-detection techniques are employed to locate boundaries between residuals exhibiting sharp contrast above and below the gridded surface (figure 1.10). A region-growing algorithm encloses the convex hull of a candidate object – a group of points within a closed boundary – and the object is removed (filtered out of the point cloud) if the mean height of the set of points is equal to or greater than the mean boundary height.

The original concept introduced by Brovelli et al. (2002) was developed by Filin (2002), who clustered points into four classes prior to filtering based on three measures: (1) the position of a candidate point relative to its neighbours, (2) morphological height differences between clusters and (3) distance parameters of points on the tangent plane to a candidate point. ThuyVu and Tokunaga (2002) took the technique a stage further by using wavelet analysis and k-means clustering to classify LiDAR points and filter non-terrain.

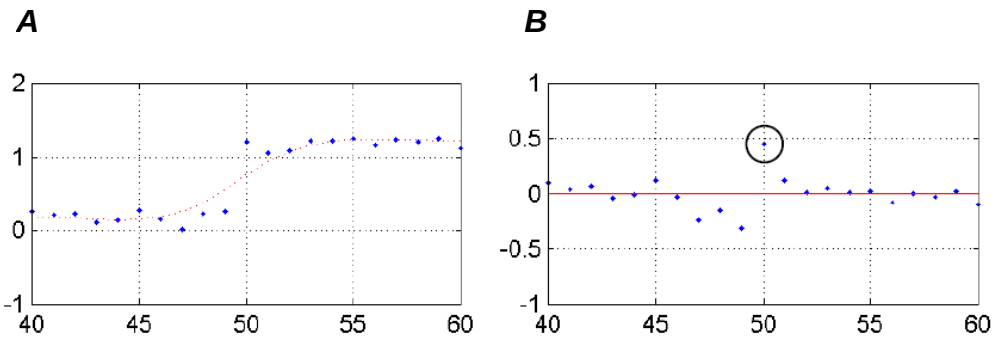


Figure 1.10 (A) LiDAR points (blue dots) along a transect over the preliminary grid (red dotted line) in an edge-based clustering algorithm. **(B)** Elevation residuals between the points and grid. Circled is a candidate point with high residual distance to an adjacent point and is hence suitable for edge detection. Source: Brovelli et al. (2002).

Advanced block minimum (Wack and Wimmer, 2002)

In this filter, the lowest points in a neighbourhood are assumed to belong to the bare earth. Significant improvement to the standard block minimum algorithm is achieved through incorporating the basic morphological concept into a hierarchical framework. An initial DTM of coarse spatial resolution (for example 9 m) is refined to models of higher resolution, e.g. 3 m, 1 m, 0.3 m etc., at each iteration. A 'Laplacian of Gauss' operator is employed at each level to identify and remove LiDAR points exhibiting elevation residuals above a pre-defined threshold to the block minima. Points not removed, for example at the 3 m level, are weighted based on their distance to the DTM at the previous level (the 9 m model) and the 3 m DTM is then recomputed, and so on.

Segmentation-based (Jacobsen and Lohmann, 2003; Nardinocchi et al., 2003)

In this filter, it is assumed that the bare earth is a collection of connected components, occupying a large area. Raw LiDAR data are first gridded, with cells interpolated using the nearest neighbour technique. Cells are segmented (grouped) if an area subscribes to a pre-defined geometric and topologic pattern, e.g. of that expected over tree canopy, building or terrain, via a region-growing algorithm. Following segmentation, regions are connected

through node and edge theory and non-terrain *regions* (rather than individual LiDAR points, as in most other filters) are filtered if their gradient orientations to adjacent regions are above a maximum threshold.

Tóvári and Pfeifer (2005) added a segmentation-based algorithm to standard surface regularisation, while an innovative method was proposed by Sithole (2005) where LiDAR points are classified as terrain if a human could 'reach' them by walking (by means of GIS-based network analysis).

1.3.3.2 Critical review of terrain filtering algorithms

To be able to perform effective qualitative or quantitative evaluation of available terrain filtering algorithms, assessment must be conducted in a controlled environment, i.e. by testing each filter on the same (correctly prior-classified) LiDAR data set(s). However, to date, few experimental comparison studies have been carried out because filtering algorithms tend to be proprietary and their details are not published. Two relevant studies are available: the first, a comparison of eight filters over twelve LiDAR data sets by an ISPRS Working Group in 2004 intended to identify directions for future research (Sithole and Vosselman, 2004) and the second, a comparison of nine filters over fifteen data sets (provided by ISPRS) by Texan Universities in 2010 intended to highlight the limits of existing algorithms (Meng et al., 2010).

Sithole and Vosselman (2004) recognised that the following scenarios tended to cause problems for filters: outliers, complex objects (e.g. buildings or urban landscapes), slopes, vegetation and discontinuous terrain; so appropriate LiDAR data sets were acquired and qualitative assessment was conducted (table 1.1). Most relevant to this study is the ability of a filter to preserve complex terrain and remove vegetation (highlighted sections in table 1.1), given that the environment of interest is forest rather than urban and outliers can typically be removed with alternative algorithms prior to classification.

The progressive densification and advanced morphological algorithms were best at preserving discontinuities in terrain, with the surface regularisation filter widely reported to reduce geomorphic quality of modelled terrain (Kraus and Pfeifer, 1998; Elmqvist, 2002), while segmentation-based and basic morphological filters experience difficulties with preservation (table 1.1) (Brovelli, et al., 2002). The quantitative ISPRS assessment described the basic morphological filters (Roggero, 2001; Sithole and Vosselman, 2001) as producing the highest number of Type I errors over discontinuous terrain. Sithole and Vosselman (2004) presented results for their quantitative assessment over one site in detail (unfortunately for this work, an urban site), but the progressive densification filter performed particularly well (only 4 % of points were incorrectly classified) while the basic morphological and segmentation-based filters yielded abundant Type I errors (over 29 % of terrain points were incorrectly classified as non-terrain).

Algorithm	Active Shape	Densification & Segmentation	Local Plane Fitting	Edge-based Clustering	Advanced Morphological	Progressive Densification	Adaptive Slope-based	Surface Regularisation
Outliers								
High points	***	***	***	***	***	***	***	***
High points influence	***	***	***	***	***	***	***	***
Low points	***	**	**	***	***	**	**	***
Low points influence	***	***	***	***	***	*	*	***
Object complexity								
Large objects	***	***	***	***	***	***	***	***
Small objects	**	**	***	**	**	***	**	***
Complex objects	**	**	**	**	**	**	**	**
Low objects	*	*	***	***	***	**	**	**
Disconnected terrain	**	**	**	**	**	**	**	**
Detached objects								
Building on slopes	***	**	**	**	**	***	**	***
Bridges	***/R	***/R	***/R	***/R	***/R	**	***/R	***/R
Ramps	*	*	*	*	*	**	*	*
Vegetation								
Vegetation	***	***	***	***	***	***	***	***
Vegetation on slopes	***	***	**	**	**	**	**	***
Low vegetation	***	**	**	**	***	**	**	***
Discontinuity								
Preservation	*	*	*	*	**	**	*	**
Sharp ridges	*	*	*	*	**	*	*	*

***: good; **: fair; *: poor; R: removed.

Table 1.1 Qualitative comparison of filtering algorithms by the ISPRS Working Group. Good (***) indicates > 90 % of LiDAR points were classified correctly, fair (**) indicates < 90 % but > 50 % of points were classified correctly and poor (*) indicates < 50 % of points were classified correctly. R denotes successful removal of a non-terrain object. Source: Sithole and Vosselman (2004).

Table 1.1 illustrates that all filtering algorithms were competent at filtering vegetation, even low bushes/scrub and vegetation on slopes. However, when assessing filters against each other, this issue was where segmentation-based filters typically struggled, given lack of clear region/class structure over vegetation (Filin, 2002; Tóvári and Pfeifer, 2005; Pfeifer and Mandlburger, 2009). Only the active shape, advanced morphological and surface regularisation algorithms were very successful at removing both vegetation on slopes and understorey/low vegetation (table 1.1).

Meng et al. (2010) conducted comprehensive quantitative assessment for the same range of filter algorithms evaluated by the ISPRS Working Group, but also included a multi-directional morphological algorithm (MGF algorithm) developed the previous year (Meng et al., 2009). The Kappa Index of Agreement (Jensen, 2005) was employed to assess the accuracy of filtered LiDAR data, with 100 % indicating perfect agreement between filtered data and prior-classified (reference) LiDAR data and 0 indicating agreement arising by chance, with negative values caused by low matching data sets. Axelsson's (1999; 2000) progressive densification algorithm yielded the highest Kappa average across the full range of fifteen test sites (with half urban and half rural sites) (figure 1.11). It was clear from the tests completed that the progressive densification terrain filter dominated over sites exhibiting rough or discontinuous terrain and was most successful at three of four sites with steep slopes and dense forest.

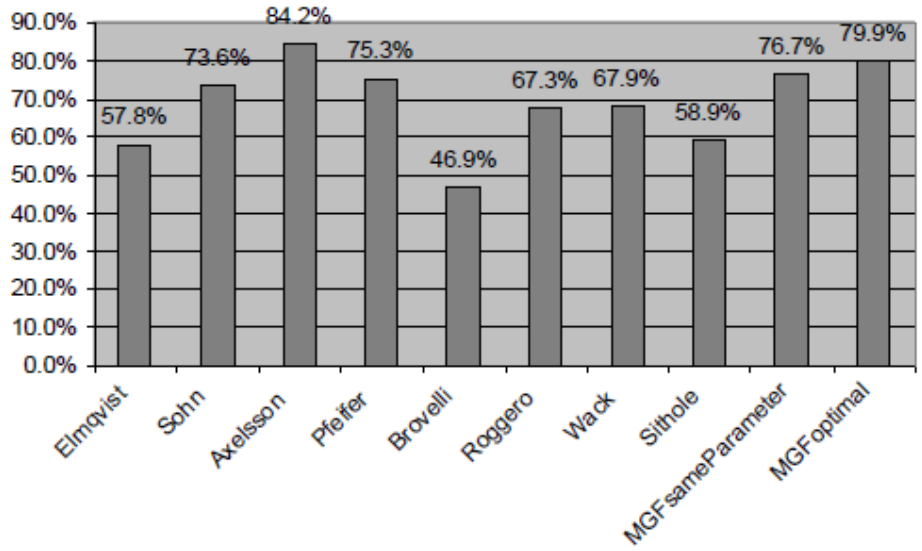


Figure 1.11 Kappa averages for nine filters based on 15 LiDAR data sets covering a range of environment types. Source: Meng et al. (2009).

Each filter tended to demonstrate particular aptitude over a specific environment or scenario, so there is no reason to advocate the disqualification of any algorithm. However, the experimental reviews conducted by Sithole and Vosselman (2004) and Meng et al. (2010) identified the progressive densification algorithm as the preeminent filter over a majority of environments, including dense forest and scenarios comprising sloped and discontinuous terrain – those most relevant to this study. Pfeifer and Mandlburger (2009) also promoted the use of filters with surface-based structures (e.g. Axelsson, 1999; 2000; Elmqvist, 2002) for generating terrain models of high geomorphic quality under forest canopy. The MGF and alternative progressive densification algorithms of Meng et al. (2009) and Sohn and Dowman (2002), respectively, also impressed when filtering both tree canopy and low scrub vegetation.

1.3.3.3 Filter adaptation to survey location

If an algorithm is not entirely autonomous, i.e. it requires user input prior to filtering LiDAR data, typically in the form of parameter configuration, then effective removal of non-terrain points is conditional on the identification of acceptable input. Few algorithms are 100 %

autonomous, but an example of one is the modified surface regularisation filter developed by Doneus and Briese (2006a) for full-waveform LiDAR data (see section 2.3) which thresholded the width of a returning laser 'echo' to better detect terrain points. However, if, for example, full-waveform ranging capabilities are unavailable or a basic, unmodified algorithm is used, as is standard in the majority of surveys, additional data must be employed to 'train' parameters to a specific site (Pfeifer and Mandlburger, 2009).

Training data are typically field-acquired reference points covering a small region of the full survey area, classified manually into terrain and non-terrain categories, which are then used to constrain filter parameters. LiDAR points covering the same area as the reference data are iteratively filtered with varying parameters until they match the reference points classified as terrain (e.g. Axelsson, 1999; Vosselman, 2000; Zhang et al., 2003). Numerous variations to this procedure exist; for example, Sithole (2001) used a coarse slope map obtained from preliminary LiDAR terrain filtering (i.e. less accurate but validated with less field data and therefore much quicker) to manipulate the rotation of a morphological structure element and Schickler and Thorpe (2001) created a classification map of vegetation types, water bodies, urban areas etc. and associated a parameter set with each class to train the surface regularisation filter to location.

Generally, a filtering algorithm is capable of classifying LiDAR data to a reasonable level of accuracy without exploiting training data; for example, in Meng et al.'s (2010) comparison study, default parameters supplied by the software providers were able to provide typically > 50 % accurate classifications across a range of environments (LiDAR data sets). Yet to obtain filtering accuracy closer to (and sometimes above) 90 %, algorithms must be trained to location with, characteristically field-acquired, reference data (Schickler and Thorpe, 2001).

1.3.4 Removing outlying points

It has been noted that several LiDAR terrain filtering algorithms have difficulties eliminating the influences of points far removed from the bulk point cloud (see table 1.1). These 'outliers' are either way above ground, typically caused by laser reflections from birds, or way below, caused by multi-path error: an energy pulse bouncing off multiple surfaces leading to erroneously long travel-time and mis-location of co-ordinates (Kobler et al., 2007), or by artefacts generated within the altimeter from high-energy laser returns (Haugerud and Harding, 2001). However, almost all outliers can be removed prior to the core filtering operation using maximum height difference algorithms, which are commonly available in most LiDAR filtering packages.

1.3.5 Modelling terrain with LiDAR data

Following classification, it is necessary to model filtered LiDAR terrain points as a digital terrain surface in order to facilitate computer visualisation of topography. As noted in section 1.3.2, the data structure of a model is typically either raster-based (DTM) or triangulated (TIN). To generate a raster surface, regularly-gridded cellular arrays are assigned elevation values by LiDAR terrain points when they intersect cells (Pfeifer and Mandlburger, 2009). However, where cells do not intersect any LiDAR points, i.e. at gaps between terrain returns, a mathematical function must be used to *interpolate* (model) between points.

The relative influences of different interpolation methods on DTM accuracy were investigated by Desmet (1997) and Smith et al. (2005) who calculated that spline and advanced kriging techniques offered the highest accuracy with respect to field-acquired reference terrain data. Nearest-neighbour interpolation generated the highest overall error and introduced artefacts to DTMs, primarily in sudden changes in elevation and regular 'stepping' (an oscillation around the true terrain surface on slopes). Since the options for the choice of interpolation procedure

are often limited to the type/brand of software package used and the motivation, i.e. prerequisite macro- or micro-scale modelling for construction or fluvial purposes, for example, this issue will be discussed in greater detail in the methodology in Chapter 3.

1.3.6 Accuracy assessment

With a view to evaluating the quality of processed LiDAR terrain data, qualitative and quantitative accuracy assessment can be performed. The former will simply involve visual inspection of a surface for obvious data artefacts while the latter typically requires cross-tabulation with classified ground truth data (Huising and Gomes Pereira, 1998). A wide-scale random sampling approach has been tested for quantitative assessment, where filtered terrain points are selected at random over the full study area and their co-ordinates located in the field (e.g. Bowen and Waltermire, 2002; Norheim et al., 2002); but this has been found to overestimate DTM accuracy because poorly classified LiDAR points tend to be clustered and thus locally influential, for example where a dense bush causes Type II filtering errors, and random sampling characteristically misses the majority of these (relatively isolated) scenarios (Meng et al., 2010). The cross-tabulation approach is stratified-random and is, as such, biased towards these difficult filtering scenarios. *In-situ* field reference points are classified, with terrain points categorised by land-cover, e.g. those under grass, scrub/shrub, deciduous and coniferous forest (Su, 2004; Su and Bork, 2006) and evaluated against either nearby LiDAR terrain points (i.e. point-pairs within a set radius) (Hodgson and Bresnahan, 2004) or the interpolated terrain model (Reutebuch et al., 2003).

Measures of LiDAR terrain quality are ascertained by evaluating differences in elevation (deviations) between reference points and LiDAR points or terrain models. The actual statistics used vary between accuracy assessments in different studies, with little cohesion in approach, but all offer one, or a number of, statistical measures that describe: (1) bias – the over- or under-estimation of LiDAR-derived elevations with respect to reference elevation (e.g.

Norheim et al., 2002) and/or (2) precision – the variability of LiDAR-derived elevations with respect to reference elevation (e.g. Hodgson and Bresnahan, 2004; Hyyppä et al., 2005).

1.3.7 Seasonal impact on airborne LiDAR survey

It has been recognised that the time of year is influential on the value of airborne LiDAR survey over forest, as winter tree conditions, i.e. deciduous leaf-off, provide better canopy penetration rates than summer conditions and thus a greater (relative) number of returns from the terrain rather than canopy (Amable et al., 2004; Raber et al., 2002b). Devereux et al. (2005), Gallagher and Josephs (2008) and Harmon et al. (2006) conducted archaeological LiDAR surveys during winter for this reason and, although the former noted that this does not necessarily have any advantage for surveys over coniferous forest, Crutchley (2008) suggested that low understorey vegetation in winter, under both coniferous and deciduous stands, is also conducive to higher quality terrain derivation.

Chapter 2. Literature Review 2: Archaeological Prospection with LiDAR Data

Whilst archaeological features under forests are difficult to detect in the field and frequently inaccessible, the sheltered environment is ideal for long-term feature preservation without the threats typical of open areas, primarily from agriculture, but additionally from weather-related erosion and military damage (Barnes, 2003; Doneus and Briese, 2006a; Gallagher and Josephs, 2008). Consequently, the ability to 'remove' forest canopy and conduct archaeological prospection, for which airborne LiDAR is advocated as a non-invasive tool (see section 1.3), could, quite feasibly, make possible the detection of better-preserved features than is typical in open areas.

Key contributors to this young discipline (of archaeology-based airborne laser scanning) include: Bernard Devereux (Devereux et al., 2005; et al., 2008) for initial experimentation into the viability of using airborne LiDAR; Simon Crutchley (Crutchley, 2006; 2008) for his work on LiDAR processing and analysing methodology; Keith Challis (Challis et al., 2006; et al., 2008) for his work on classifying archaeological features from LiDAR data; and Michael Doneus (Doneus and Briese 2006a; b; et al., 2008) for adapting the standard processing chain to concentrate on detecting archaeological features. The following section will outline the techniques employed to analyse LiDAR data for archaeological purposes and the critical work undertaken to improve its performance.

2.1 Limitations of Ground Survey and Photogrammetry for Sub-canopy Archaeology

Shortcomings to previous sub-forest canopy feature prospection methods were mentioned early in this work as a premise for the initial experimentation with, and later development of, airborne LiDAR as a possible tool for detecting archaeological features beneath forests. Historically, prospection and monitoring have been conducted on foot on a site-by-site basis,

with the more recent addition of aerial photography support. However the cost of visiting and reporting on vast numbers of sites, especially under forests, is prohibitive and consequently, feature validation and condition surveys, supplementary to original field surveys, are characteristically not undertaken on a regular basis (Barnes, 2003). Aerial photogrammetry has also been employed to measure sub-canopy elevation; but as an intrinsically passive technique, in order to detect subtle features, lighting conditions and sun elevation through the surveying period must be at near optimum levels, and some low earthworks covered in dense vegetation are almost impossible to locate on the ground, let alone from the air, in low spatial resolution photogrammetry-derived elevation models (Doneus and Briese, 2006b; Gallagher and Josephs, 2008).

While airborne laser scanning suffers various measurement errors (see section 1.2.3), it has several crucial advantages over former techniques. Through LiDAR surveying, a client can obtain very high density elevation measurements at an acquisition rate of greater than 10 km^2 per day, and processing for terrain models is relatively automated and accurate; so for these respective reasons it is a more efficient technique than either ground survey or photogrammetry (Baltsavias, 1999b; Barber, 2007; Crutchley, 2008). For archaeological purposes, the coherence of LiDAR terrain models and images allows the human eye to detect small changes in topography that cannot be identified in the characteristically coarse terrain models derived from photogrammetry (Barber, 2007; Gallagher and Josephs, 2008).

2.2 Standard Methodology for Conducting Archaeological Feature Prospection with LiDAR

Following the decision to conduct an airborne LiDAR survey over a particular site, with the aim of detecting and resolving archaeological features within modelled topography, it is necessary to make several logistical and data processing-directed considerations. A survey is commissioned with a view to recording LiDAR points at a specific point density or spacing, i.e. *n* points per m^2 or one point every *x* m, respectively, and flight parameters, for example flying

height and laser pointing angle (equation 1.4 in section 1.2.2), are set accordingly. Crutchley (2008) recommended point spacing of at least 1 m to detect cultural features like barrows, enclosures and mining pits, most of the time; while Barber (2007) suggested that 50 cm spacing will permit detection of features of this size 95 % of the time (table 2.1). Table 2.1 gives appropriate point spacing for the 66 or 95 % probability of detecting features (standard cultural heritage) of various sizes, based on mathematical analysis of points on a grid (Barber, 2007). The obstacles to this 'inverted' approach of determining flight parameters from prior knowledge of the characteristic archaeological feature to be measured are: (1) if little or nothing is known about the features to be detected; (2) if there is forest cover and thus no idea of the laser penetration rate and (3) if there is forest cover, understorey or low vegetation cover etc., with terrain filtering required and thus no idea of the accuracy of processed LiDAR terrain data. These problems indicate both a direction for further research and the common, present lack of an informed procedure for selecting flight parameters (i.e. choosing suitable LiDAR point densities).

Feature size (mm)	Example feature	Point spacing required to give 66 % probability that a feature will be visible (mm)	Point spacing required to give 95 % probability that a feature will be visible (mm)
10,000	Large earth work	3500	500
1,000	Small earth work/ditch	350	50
100	Large stone masonry	35	5
10	Large tool marks	3.5	0.5
1	Weathered masonry	0.35	0.05

Table 2.1 Appropriate point spacing (sampling resolution) for various sizes of cultural heritage, based on standard mathematics to determine appropriate minimum sampling intervals over a regular grid of data. Adapted from: Barber (2007).

Raw LiDAR data are filtered and terrain points are modelled as a raster (DTM) or TIN. While triangulation maintains all the accuracy of the input data and rasterisation tends to smooth

(often important) topographic discontinuities, TINs will enhance artefacts caused by Type II errors introduced at the filtering stage where a smoothed raster surface will reveal hidden features (Crutchley, 2008). Investigators have typically favoured rasters over TINs for interpreting archaeological features in LiDAR data (e.g. Challis et al., 2008; Devereux et al., 2005; Harmon et al., 2006).

Numerous post-processing techniques, of varying sophistication, have been utilised to analyse LiDAR data. Typically, the more advanced techniques have greater processing (time/cost) requirements. The most basic is vertical exaggeration of the terrain using 3-D visualisation software packages, but this can equally emphasise background noise and data artefacts (Type II filtering errors) (Crutchley, 2006; Harmon et al., 2006). So the common first step of visual analysis is to generate hillshade images from terrain models, in either greyscale (figure 2.1) or colour. The 'scene' (area of interest) is artificially illuminated from a specific light angle and azimuth, with shadows cast behind the digital topography and, ideally, features accentuated against the general terrain surface. This technique is imitating passive aerial photography (with a static solar light source), but resolves characteristic photographic problems of excessive shadow and feature alignment parallel to the light source, with a single LiDAR scene usually lit from a minimum of four angles (Crutchley, 2008). Preferably, a scene should be lit and interpreted from every viewing angle and azimuth to maximise feature detection, but the processing requirements become unwieldy.

Hence, various developments have been made to the basic hillshade technique. The first is to give each hillshade model, of the same scene but lit from a different viewing angle (a 'series'), a separate colour scale and combine them into a single colour composite (Crutchley, 2008). This creates an image with shadows from four angles, but this is at the significant risk of masking subtler features like slight banks or ditches (Devereux et al., 2005; Crutchley, 2008) which require a specific (single) illumination direction to be detected. Alternatively, statistical analysis of the variation in a series can be employed to create a single interpretation image. Through principle components analysis (PCA), an image is generated which possesses the

property of equal illumination from all angles, thus yielding a far lower likelihood of feature masking (Devereux et al., 2008). This technique will be covered in greater detail in section 2.3.

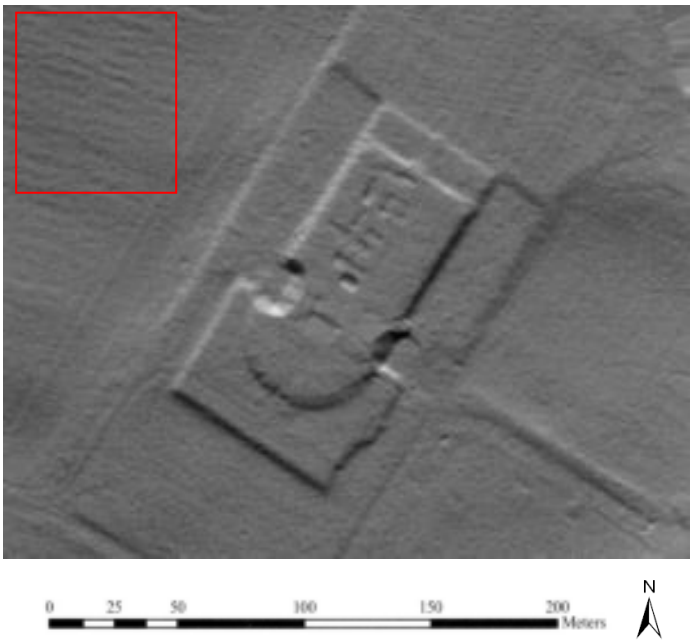


Figure 2.1 LiDAR hillshade image of an eighteenth-century plantation terrace in Maryland, U.S., illuminated from the north, with excavation trenches of 20-60 cm depth visible in the highlighted area (red box). Source: Harmon et al. (2006).

Techniques for the direct manipulation of digital terrain data have also progressed from basic vertical exaggeration, with slope and difference models generated to enhance archaeological features against natural, background topography (Doneus and Briese, 2006b; Hesse, 2010; Sittler, 2004). The slope model exhibits a graded colour scheme based on slope rather than elevation while the difference model is formed by subtraction of a full resolution terrain model from a resampled model of lower resolution. Again these techniques will be covered in greater detail in section 2.3.

Crutchley (2006) emphasised the importance of exploiting contextual data both in the interpretation and validation of LiDAR data. He suggested checking OS and historical maps to rule in or out candidate archaeological features identified in a terrain model and prevent the erroneous classification of, for example, modern field boundaries and man-made forest

clearance piles ('brash' from tree felling operations) as cultural heritage. Barber (2007) advocated the compilation of an integrated GIS with mapping, photography, points of interest and land cover information, in addition to LiDAR terrain information. Hyperspectral data and LiDAR have also been acquired for the same area and collated for archaeological applications (Coren et al., 2005; Pietrapertosa et al., 2008).

2.3 Examples of Critical Work Using LiDAR in Archaeology

Ridge and furrow extraction (Sittler, 2004; Sittler and Schellberg, 2006)

Fossilised ridge and furrow micro-topography (typically of approximately 30 - 60 cm depth), formed over years of ploughing and hoeing, was measured under mixed-species forest canopy at a site in Germany converted to forest from arable land in the 18th century. LiDAR data were acquired in broadleaf-on conditions, but variation in canopy density and species did not demonstrate significant control over the quality of terrain filtering. LiDAR point spacing of one point every 1.5 m was sufficient to resolve topography because the vertical accuracy of the data was less than the height of the majority of features. Following computation of a difference model between the 2-D planimetric surface and true undulating topography, individual furrows were extracted through manual delineation (figure 2.2). Although this method appeared to succeed, lack of validation against ground truth data indicates that its genuine potential is yet to be fully tested.

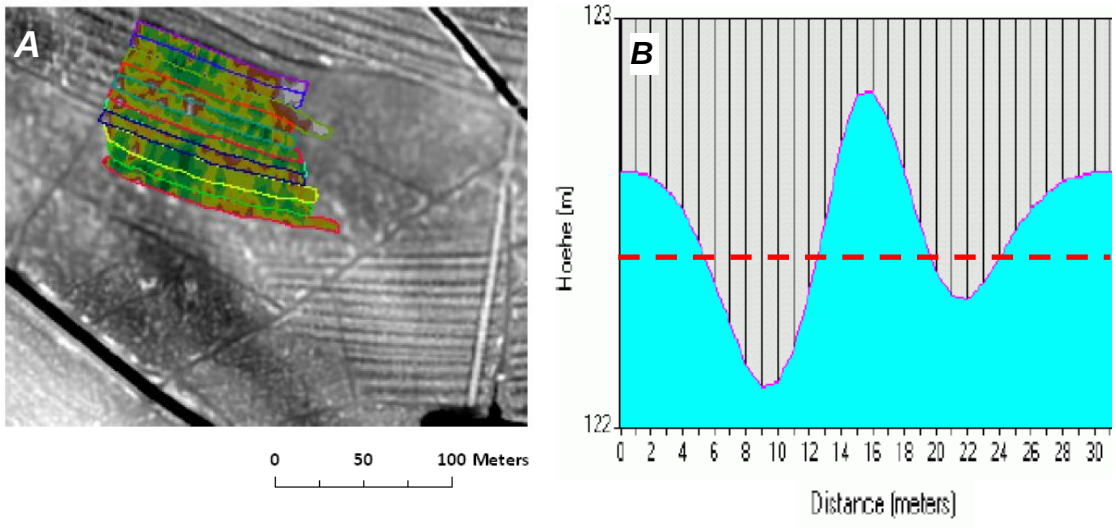


Figure 2.2 (A) Extraction of individual furrows from a difference model (greyscale image) as polygons (coloured shapes) by manual delineation in GIS software. The difference model was generated by subtracting the LiDAR terrain model from a 2-D planimetric surface. **(B)** Vertically exaggerated transect across two furrows with the position of the 2-D planimetric surface highlighted by the dashed red line. Source: Sittler (2004).

Advanced visualisation of sub-forest canopy features (Devereux et al., 2005; et al., 2008)

Broadleaf-off airborne LiDAR survey was employed to attempt to map a prehistoric hillfort at Welshbury Hill, Gloucestershire, characterised by obvious bank and ditch earthworks, subtle hut and charcoal platforms and a possible Bronze Age field system under dense, mixed-species forest, beneath which ground survey had been described as an 'impossible' task (SAM 31186, 2004 in Devereux et al., 2005). DTMs were generated from LiDAR terrain data, filtered with an algorithm that classified last echoes as terrain if they were above a threshold distance from their equivalent first echoes (i.e. of the same laser pulse) (this is a modified version of the algorithm developed by Raber et al., 2002a). A preliminary experiment was conducted to test the quality of filtered LiDAR terrain (and archaeological feature detection potential) against 76 vegetation zones (categorised by tree species and canopy form), the reference data set compiled by field survey (Crow, 2003). Leaf-off deciduous trees were conducive to regular laser penetration and thus high terrain point density, but even in mature, thinned conifer plantation, feature detection was described as 'surprisingly successful' (Devereux et al., 2005).

Very subtle features, like charcoal platforms, were frequently detected, but in some areas man-made forest clearance piles could not be filtered out and were evident as 'fishbone' patterns in the modelled terrain; these are likely, in some places, to mask underlying features.

Devereux's group extended their method in 2008 by pioneering innovative principal components analysis (PCA) for visualising the same Welshbury Hill LiDAR terrain models. This method takes a set of inter-correlated variables (i.e. a series of hillshade images illuminated from different angles) and transforms them to a new set of uncorrelated variables using the eigenvalues and eigenvectors of their correlation matrix (Devereux et al., 2008). The first three components typically contain over 95 % of the variability in the original data set, so a false colour composite of these components was generated by the group (figure 2.3 B), which negated the effects of over-illumination and over-shadow apparent in their earlier work. The third component (figure 2.3 C) was additionally found to exhibit almost equal loading on all sixteen of the original illumination directions (there were 16 hillshade images in the series), highlighting finer detail in the landscape and aiding in, especially, the detection of subtle platform earthworks (Devereux et al., 2008).

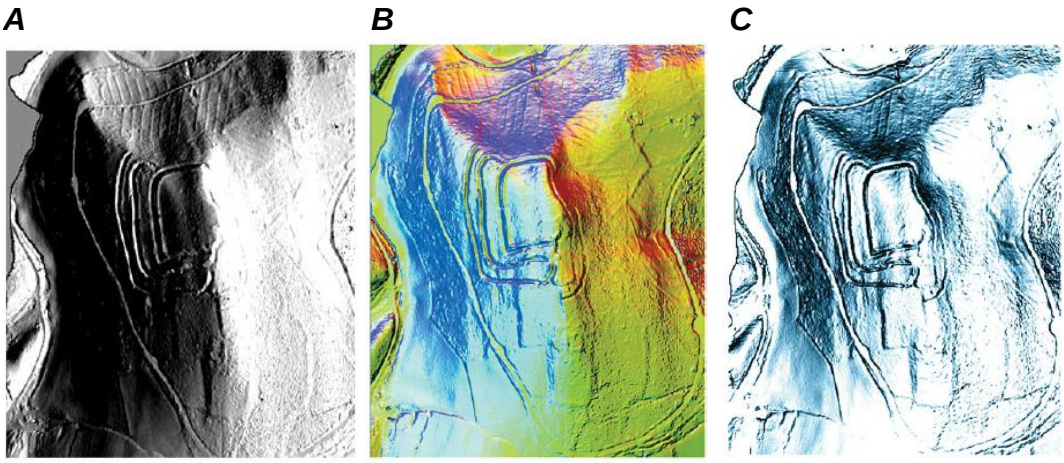


Figure 2.3 (A) Hillshade model of Welshbury Hillfort illuminated from the east (90°) with over-illumination and shadowing clearly visible. **(B)** Colour composite image of the first three principal components. **(C)** Third principal component with bank and ditch features noticeably emphasised. Spatial scale is unavailable in the original work for these images, but the scene is hundreds of metres across. Source: Devereux et al. (2008).

Feature edge detection from hillshade models (Harmon, 2006)

Terraced gardens and shaped acreage were identified in LiDAR terrain data acquired over a former plantation landscape in Maryland, U.S., now dominated by low vegetation (scrub) cover (figure 2.1). DTMs of both 1 and 2 m spatial resolution were generated by interpolating LiDAR last echoes (i.e. without advanced filtering) and following experimentation, 2 m resolution could not provide sufficient detail to pick up the edge-type terrace features. However, breakline modelling and image-based edge-detection techniques were utilised to successfully identify terraces in the 1 m resolution model (figure 2.1) and following feature mapping, the authors performed viewshed analysis to interpret the relationship of features to each other in the landscape.

Geoarchaeological prospection and feature cataloguing (Challis et al., 2006; et al., 2008)

Initially Challis' working group focused on assessing the capability of airborne laser scanning for detecting archaeological geomorphology, i.e. relict river channels (Challis, 2006; Challis et al., 2006). Although this work had some application for topographic feature detection, Challis' group later concentrated on archaeology and specifically on compiling feature inventories from LiDAR data (Challis et al., 2008). The group investigated the potential for LiDAR to enhance historic records using a data set acquired over a relatively open (unforested) corridor of the River Dove on the Derbyshire/Staffordshire border.

Fossilised ridge and furrow topography, historic field boundaries and settlement remains were identified by their surface expression in LiDAR hillshade models, digitised in GIS software and finally categorised. Mapped features were compared to the Historic Environment Records (HER) and 84.4 % detected by LiDAR were not, at that time, on-record, reflecting a tendency of the HER to exclude large-scale landscape features like ridge and furrow; but also some ditch/platform earthworks had been missed. Laser scanning was recognised by Challis et al. (2008) as missing the majority of crop and soil marks (typically identified in oblique aerial photography) and missing some of the 'smoother' earthworks.

Sub-forest canopy prospection with full-waveform LiDAR (Doneus and Briese, 2006a; 2006b; Doneus et al., 2008)

In their earliest work, Doneus and Briese (2006a) recognised the influence of low vegetation and understorey on LiDAR filtering accuracy, especially on the mistaken acceptance of points over vegetation in the terrain class (Type II errors), and the subsequent capabilities of airborne laser scanning for feature detection. As intimated in section 1.2.3, successive echoes from the same laser pulse must be separated by distances larger than the range discrimination of the LiDAR sensor (approximately 15 cm on the ground) to be recorded as distinct returns. The group attempted to identify earthwork features at an Iron Age Hillfort located under dense multi-story, but monoculture, deciduous forest in Austria, using the surface regularisation (robust linear prediction) filter (Pfeifer et al., 2001) to remove vegetation in leaf-off LiDAR data. However, by employing standard filtering procedure, the processed DTM included many anomalous features of diameter 5-6 m and height 0.2-1 m (figure 2.4 A), identified on the ground as man-made forest clearance piles (brash) (Doneus and Briese, 2006a).

The full-waveform LiDAR is capable of quantifying the entire power waveform of each returning laser pulse (see full-waveform digitisation function in figure 1.2) and, as such, the power (or 'width') of each echo – i.e. the range distribution of individual scatterers contributing to an echo (Baltsavias, 1999a). A small echo width is recorded from flat, clear surface elements (e.g. terrain), while a large echo combines scatter from various elements (e.g. low vegetation and terrain) (Doneus and Briese, 2006a). Thus, Doneus et al. (2008) performed exploratory analysis to determine a threshold echo width of 1.7 ns, above which LiDAR returns were classified as non-terrain, and added an auxiliary step to the standard surface regularisation filtering procedure. The superior algorithm was tested on the same Austrian Hillfort LiDAR data set, with the majority of clearance piles and noise 'purged' from modelled terrain (figure 2.4 A and B) and many additional barrows of 0.2 to 2 m height detected to those mapped in a 1960s field survey (figure 2.4 C and D) (Doneus et al., 2008;

Ulbrich, 1962). Over some barrows, shallow (10 to 20 cm) depressions could even be observed where graves had been robbed.

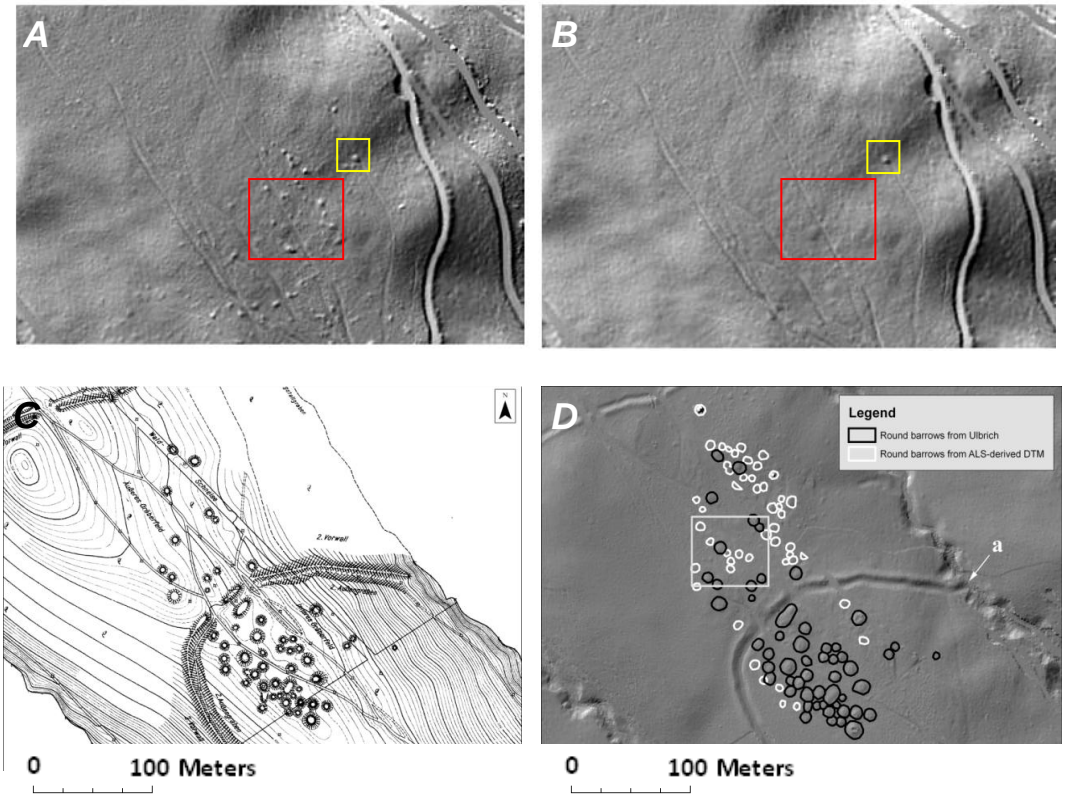


Figure 2.4 (A) Hillshade model of part of an Iron Age Hillfort (image approximately 0.5 km across) after standard filtering, with circular features highlighted in the boxes. **(B)** The same scene after supplementary removal of LiDAR returns with high echo width, with brash piles in the red box purged from the terrain model and a genuine barrow highlighted in the yellow box. **(C)** Terrestrial survey map of separate scene (Ulbrich, 1962). **(D)** Additional barrows detected with LiDAR (white circles) to those mapped by terrestrial survey (black circles) in the same scene as C. Sources: Doneus and Briese (2006a); Doneus et al. (2008).

Detecting and interpreting industrial archaeology (Gallagher and Josephs, 2008)

With the aim of detecting features of cultural heritage under boreal, mixed coniferous and deciduous forest in Michigan, U.S., airborne LiDAR data were acquired (in broadleaf-off conditions) and filtered by classifying only last echo returns as terrain. Exhaustive copper mining was carried out during the 18th and 19th centuries, with the area therefore likely

containing extensive, previously undetected, relict industrial infrastructure. Hillshade models generated from the processed LiDAR terrain data were examined and candidate features identified by four visual criteria: particular shape, pattern, texture and shadow. Features were then categorised into, for example, mining pit, prospecting pit, spoil heap etc. and validated by pedestrian reconnaissance with handheld GPS. By this procedure, Gallagher and Josephs (2008) managed to locate numerous, previously unrecorded features, but suggested that areas of their DTMs were frequently masked by either heavy shrub vegetation or data artefacts over sloped terrain.

Local relief modelling (Hesse, 2010)

Recent work has been undertaken to improve existing techniques for manipulating digital terrain data, i.e. to enhance the basic difference model. The Local Relief Model (LRM) emphasises the visibility of medium-scale, shallow topographic features in a hillshade model, irrespective of illumination angle, by creating a DTM purged of variable, small-scale landforms from the original LiDAR data. An LRM reflects less bias towards small-scale features relative to the landscape at large than is typical in a simple difference model (Hesse, 2010). When testing at a site in Germany, the technique was found to be particularly successful at exposing the most subtle topographic features, like ridge and furrow for example, but suffered from distortion on steep slopes and required site-specific parameter selection.

2.4 Common Obstacles to Conducting Viable Prospection with LiDAR Data

Airborne laser scanning is limited in its aptitude for remote archaeological prospection, with the limitations mentioned in the previous section now collated. Commission errors (mistakenly accepted returns from vegetation in the terrain category during filtering) appear to affect the bulk of research, with several studies intentionally designing procedures to remove artificial features in terrain models (Doneus and Briese, 2006a; Hesse, 2010). Gallagher and Josephs

(2008) also described some omission errors causing problems for feature detection, especially over steep terrain, where interpolation caused discontinuity smoothing. Problems when artificially illuminating hillshade models are frequently referred to and include: (1) obscuring features in shadow, (2) over-illumination with an intense light source and (3) feature alignment parallel to the light source (Crutchley, 2008); although advanced PCA goes some way to solving these issues (Devereux et al., 2008). Ease of feature detection is observed as being proportional to the size and shape of feature, with larger, negative, deeper and isolated features more easily identified (Gallagher and Josephs, 2008).

Some features are semi-erased by years of agricultural disturbance (e.g. ploughing), although this occurs much less frequently under woodland (Crutchley, 2006; Challis et al., 2008). Bewley et al. (2005) suggested that almost fully-erased features can still be detected with LiDAR, with relict field boundaries in particular surviving in modelled terrain but offering no visible surface expression in the field.

2.5 Research Aim and Objectives

To this date, insufficient research has concentrated on the use of high-powered multiple-return laser ranging for sub-canopy archaeological prospection. Studies are too broad, focused on feature extraction over a wide area of typically open terrain, or motivated by application alone, with inadequate evaluation of method. However, in light of the research identified in this chapter, it is plausible that airborne laser ranging may offer a more effective alternative to existing methods of sub-canopy prospection, like photogrammetry or ground survey. It is necessary to consider a suitable framework, comprising a series of research objectives, to address the following aim:

To evaluate the capability of multiple-return discrete-pulse airborne LiDAR for detecting and resolving archaeological features under mixed coniferous and deciduous forest.

Research objectives:

1. To acquire high density airborne LiDAR data for a range of forest species characteristic of British woodland.

2. To obtain *in situ* training data from which LiDAR filtering parameters can be constrained, with a view to removing non-terrain returns and retaining archaeological features.

3. To classify the airborne laser data into terrain and non-terrain classes and derive three-dimensional models of the sub-canopy terrain surface.

4. To evaluate the accuracy of filtered LiDAR terrain data against *in situ* reference data of the true terrain surface under a variety of forest species.

5. To detect features of archaeological interest in the terrain models through image analysis and artificial exaggeration and manipulation.

6. To validate candidate features by integrating the terrain models with existing aerial photography, historic mapping, ground survey information and other contextual data.

7. To evaluate the viability of using airborne laser scanning for sub-canopy archaeological prospection and in what situation(s) it is merited.

Chapter 3. Methodology

3.1 Study Area

Chopwell Wood is a mixed coniferous and deciduous woodland of 360 hectares (3.4 km^2) located on the northern slopes of the Derwent Valley, 16 km southwest of Gateshead in the northeast of England (figure 3.1). It is classified as an Ancient Woodland site by the UK Forestry Commission (2009), a designation reserved for only 2 % of the country's forested area, and is one of the rarest habitats in the UK. The wood is characterised by rough topography and sharp elevation changes, with the ground level quickly rising from 40 to 230 m a.s.l. from south to north up the valley side. Vegetation is diverse, with over 250 plant species recorded in the wood and, although the tree canopy is dominated by species of beech, larch, pine and Douglas fir, cover is variable and mixed stands are common (Forestry Commission, 2009). A full list of the tree species in Chopwell Wood is given in Appendix 1.

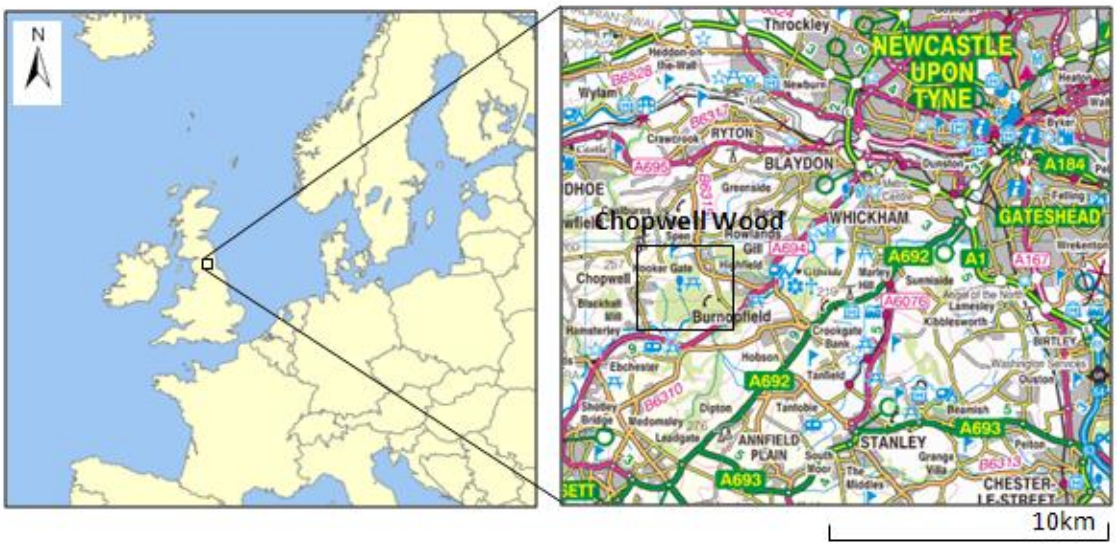


Figure 3.1 Location of Chopwell Wood. Source: EDINA (2009).

Early tribal clearance from 300 BC through to extensive Roman deforestation from 50 AD has reduced a forest covering the full north side of the Derwent Valley to a few small wooded

areas, of which the early Chopwell Wood was one. The forest was routinely felled through the second millennium AD, reaching a minimum one hundred acres of scrub and trees in the late 1600s which prompted replanting for the first time in the 1800s. Nine hundred acres of larch, oak and elm were replanted at this time and an even larger area was planted with Scots pine later on in the nineteenth century, although much of this new cover was felled during the two World Wars (figure 3.4) (Searle, 2000). With the Forestry Commission taking control of UK forests in 1923, full scale replanting of pine, fir, spruce and larch continued at Chopwell and many trees planted at this time are still standing in the forest today.

Although an ancient woodland of great value, both to environmental groups and to the local community for leisure activities, there is a commercial side to contemporary forest management. At Chopwell Wood the Forestry Commission primarily deals with thinning forest crops (removing, for example, one in every five trees), while also occasionally harvesting full areas. This has a number of implications for the work proposed in this study. When considering laser pulse interactions with human-planted, and thus regularly-spaced, trees, it is difficult to extrapolate results to natural forested area. Moreover, results from thinned forest will differ to those from non-thinned areas and harvesting operations also leave quite heavy remains on the forest floor – this compact branch and leaf material (brash) has been observed to severely limit laser pulse penetration to the terrain surface (Doneus et al., 2008). However, while these points are valid, Chopwell Wood's diversity and management is characteristic of forests across the UK and, provided that tests sites within the wood are as representative of the forest as possible, results will potentially be significant at the national level.

Chopwell Wood is host to an abundance of industrial archaeology, including relicts from the coal mining period and associated transportation infrastructure (figure 3.2). Coal mining began in approximately 1600 and by 1724 there were seven small collieries (bell-pits or drift-mines) working in and around the wood. Many drift- and shaft-mines were dug after 1850, with small (likely horse-drawn) tubways erected to move the coal out of the wood. A large mineral railway connection was constructed through the wood in 1896 (figure 3.3) to transport 300

tons of coal per day from Chopwell and mines in the surrounding area. Roads were also laid parallel to the track as firebreaks and a stone overbridge built in 1894. The railway operated until the last of the wood's collieries closed in 1960 and the track was lifted in 1961, but its impression on the sub-forest canopy topography is still clear (Searle, 2000).

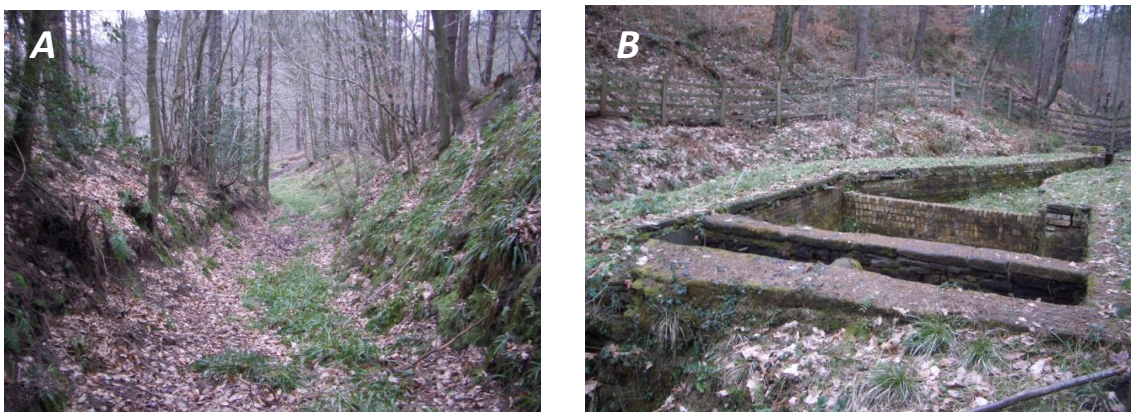


Figure 3.2 (A) Relic coal tubway and **(B)** disused sluice gate under Chopwell Wood. Locations are given in figure 3.3.

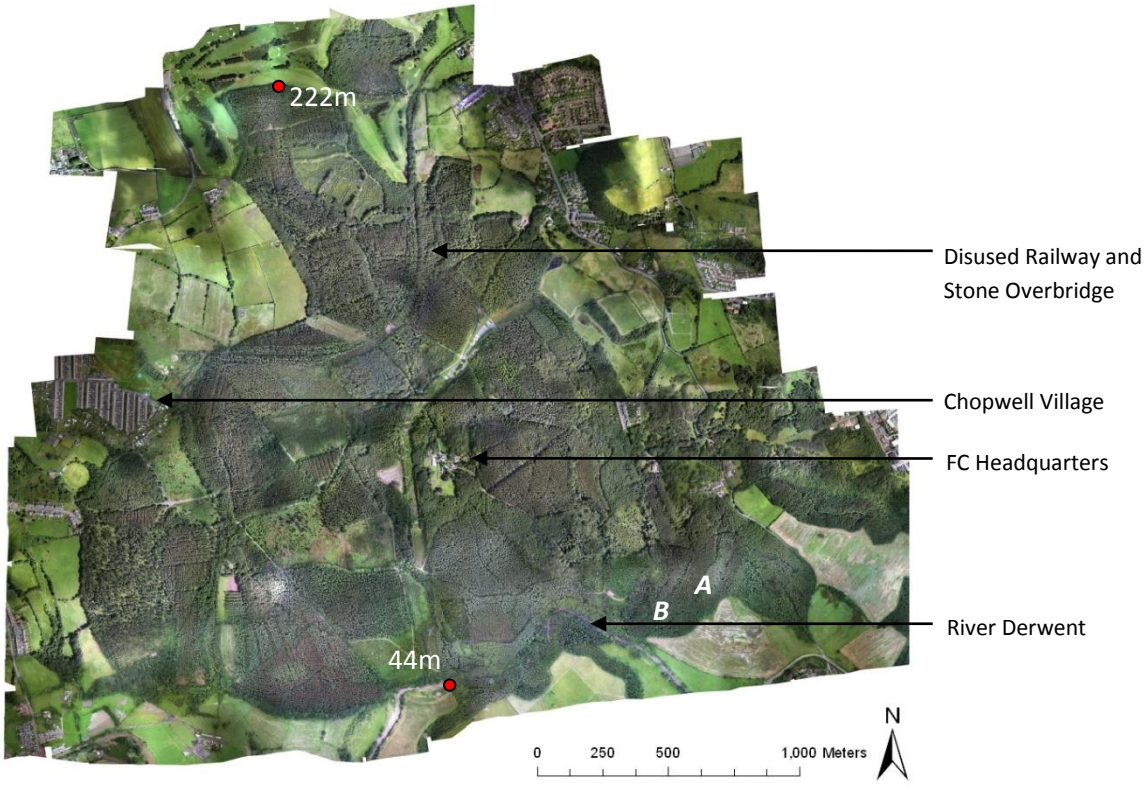


Figure 3.3 Aerial photograph of Chopwell Wood with some important features highlighted and spot height minima and maxima identified. Photo positions for figure 3.2 A and B are given.

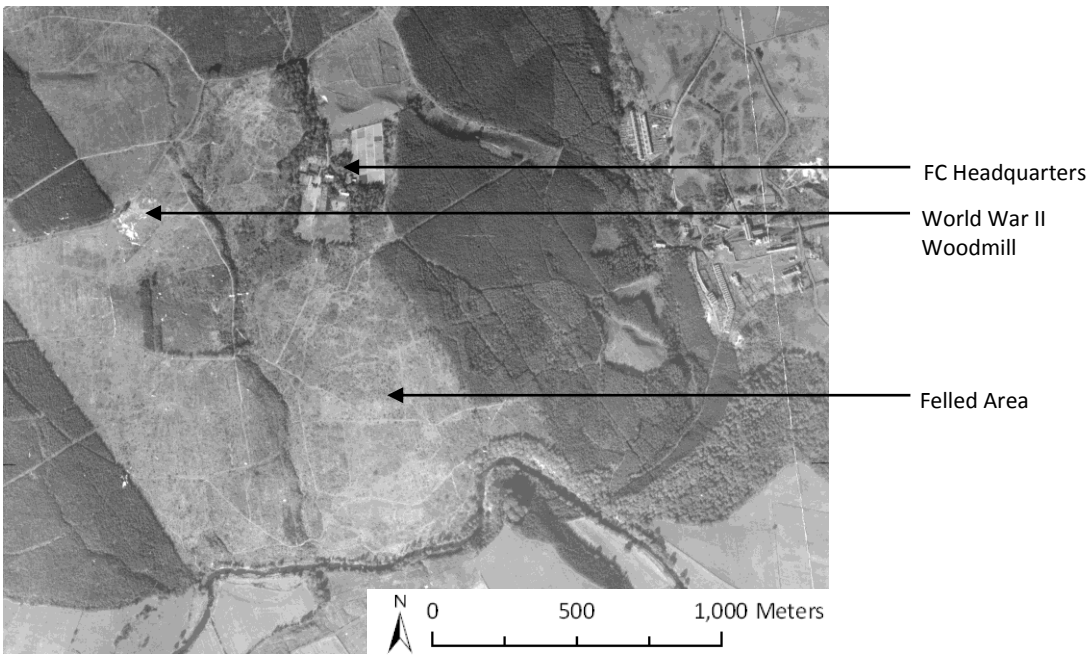


Figure 3.4 Oblique aerial photograph of south Chopwell Wood. Date: 4/10/1947. Source: Declassified Air Force Photographs from World War 2 UK aerial survey.

It is very possible that there is, as of yet, undetected industrial archaeology hidden beneath the forest canopy, whether relevant to mining activity, timber production or the railway connection. This, in addition to its unique ecological diversity, makes Chopwell Wood an ideal location to test the use of airborne LiDAR for remote modelling of sub-canopy terrain and detection of archaeological features.

This chapter is organised in the sequence of the data processing chain: (1) fieldwork preparation & data sources, (2) data acquisition, (3) data pre-processing & LiDAR point cloud classification, (4) LiDAR terrain modelling and accuracy & bias assessment, and (5) LiDAR post-processing and data analysis.

3.2 Fieldwork Preparation and Data Sources

3.2.1 Preliminary site assessment

To fully test the capability of airborne laser scanning for measuring sub-canopy terrain, experimentation was conducted under the harshest possible scientific conditions. As mentioned in the previous section, Chopwell Wood offers diversity in vegetation density and species, variable topography and variety in size and subtlety of archaeological features.

Prior to data acquisition, it was necessary to identify a number of sites within the forest representative of Chopwell's variety for *in situ* field data collection. These sites had to cover a range of tree canopy structure, topography and archaeological feature, wide enough to ensure that all key factors expected to affect the airborne LiDAR could be tested. Information collected at these sites would provide both training data for constraining optimum LiDAR filtering parameters and reference data against which the accuracy of the airborne LiDAR points could be evaluated.

During several early field visits, some of the twenty or so candidate sites (identified beforehand using mapping, aerial photography and GIS data) could be rejected under logistical criteria. These included sites in young Sitka spruce, where high canopy density was initially thought to provide an ideal challenge to sub-canopy airborne ranging; while this may be so, the density in fact proved too high even for comprehensive field data collection. Other sites were rejected in areas too steep to conduct fieldwork and where stands had been so recently thinned or fully harvested that they were not represented in Forestry Commission GIS data.

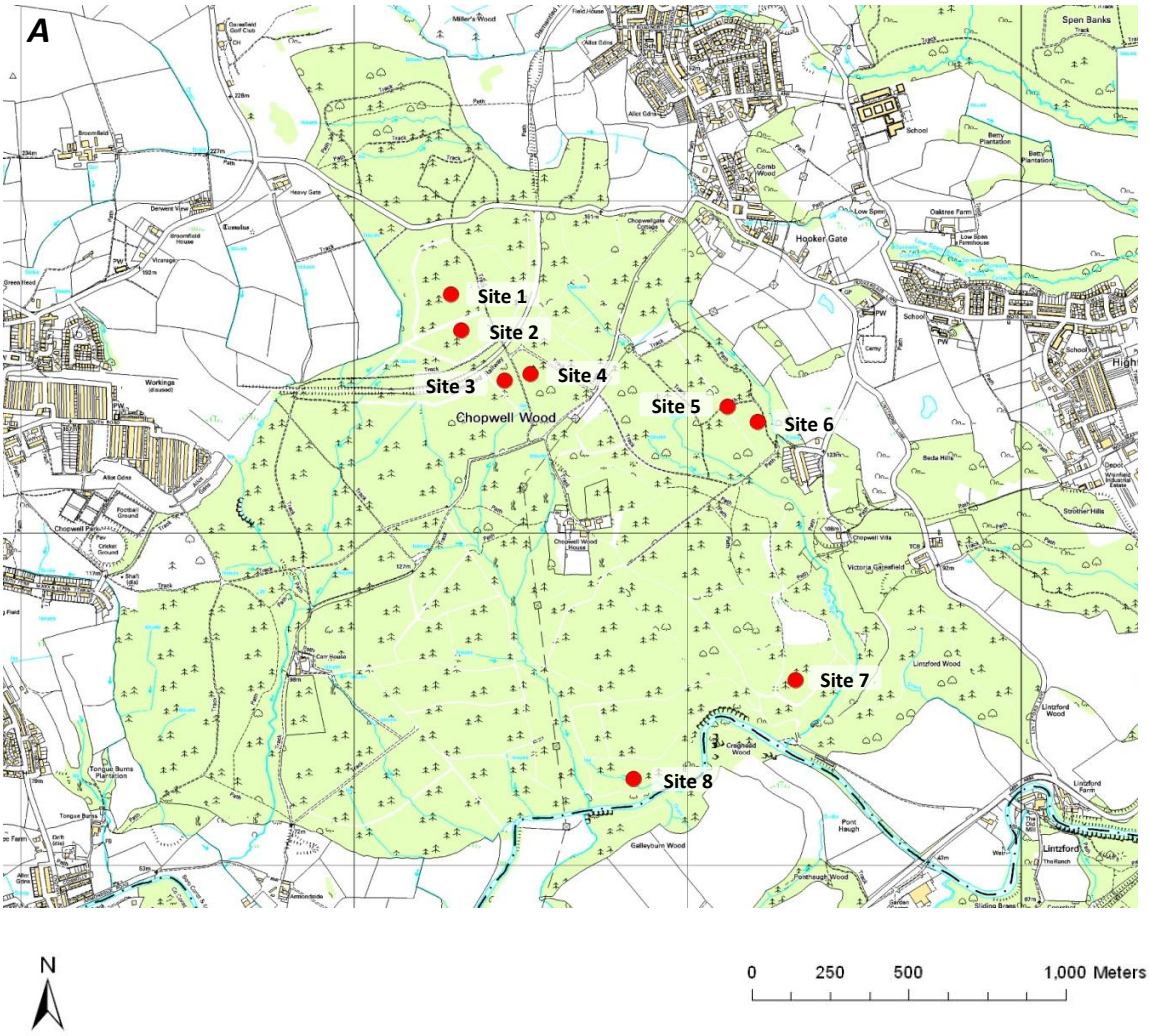


Figure 3.5 (A) Locations of sites for *in situ* field data collection on an Ordnance Survey map (1:10,000 scale). Source: EDINA (2009). **(B)** Site 4 in beech (during winter) with clear channelling (man-made, for drainage) and medium-rough topography. **(C)** Site 2 in pine (during winter) with some brash left from recent thinning operations.

Site No.	Easting	Northing	Primary Tree Species	Planting Year	FC Terrain Ratings		Known Archaeology	Understory Vegetation Rating
					Roughness	Gradient		
1	413319	558609	Corsican Pine	1947	2	1	None	5
2	413289	558718	Corsican Pine	1947	2	1	None	5
3	413527	558479	Beech	1934	2	2	None	3
4	413450	558459	Beech	1934	2	2	None	4
5	414119	558380	Beech	1923	2	1	Avenue (at possible location of pre-1800 house)	4
6	414209	558334	Beech	1923	2	1	Former reservoir	2
7	414322	557558	Douglas Fir	1987	2	3	None	1
8	414351	557617	Douglas Fir	1991	3	5	Stonework, low walls and building rubble	1

Table 3.1 Site details for *in situ* field data collection. FC (Forestry Commission) terrain ratings are subjective human observations of local terrain roughness and gradient on a scale from lowest roughness/gradient at 1 to highest at 5, obtained from the Forestry Commission GIS Database (1998). Understory vegetation ratings are subjective human observations of combined understory cover and thickness (density) on a scale from lowest density at 1 to highest at 5, made by the author and co-workers. Easting and northing are given in the OSGB36 coordinate system.

Eight candidate sites were identified during these early field visits, covering a cross-section of tree species, ages, terrain roughness, slope and size and shape of sub-canopy archaeological feature (figure 3.5 and table 3.1). Winter site photographs in figure 3.5 illustrate typical tree spacing and especially the structure of the terrain (often hidden by low vegetation in summer) at beech (B) and pine (C) stands. The number of sites was considered sufficient to provide a comprehensive assessment of the research aim, given both the range in site attributes and repeat measurement i.e. two sites in each area, four deciduous and four conifer sites (table 3.1). Data collected at all field sites were utilised in the accuracy assessment, while site 8 also served as a training data set for LiDAR classification and, as such, acquisition at this site was more intensive and comprehensive.

3.2.2 Data sources

This section describes the types of data collected whereas the data acquisition section describes the exact details of sample sizes, plot areas etc.

An Optech ALTM 3100EA airborne laser ranging system was mounted on a helicopter platform and used to collect point cloud data over the study area. The system comprised a pulsed infra-red laser and a sensor capable of recording up to four returns in the time domain of each pulse, a differential GPS in communication with a ground-based station and an IMU. The laser can deliver up to 70,000 pulses per second and can typically differentiate objects on the ground larger than approximately 30 cm in diameter at an elevation of 1 km (table 3.2).

	LiDAR System	
	Optech ALTM 3100EA	Trimble GS200
Platform	Helicopter	Terrestrial
Laser Type	Pulsed	Pulsed
Beam Deflection	Oscillating mirror	Oscillating mirror
Wavelength (nm)	1064	516
Flying Height (km)	≤3.5	-
Pulse Rate (kHz)	≤70	≤5
Footprint Size	0.3/0.8 m @ 1 km	3 mm @ 50 m
Accuracy (cm)	Hor: 18@1km, Ver: 5-35	Hor:0.3@100m, Ver: ~0.1
Range (m)	4000	200
Field of View (deg)	Hor:50, Ver: 50	Hor: 360, Ver: 60

Table 3.2 Technical specifications for the airborne and terrestrial LiDAR systems used in this study (Mallet and Bretar, 2009; Optech ALTM Gemini Datasheet, 2004; Trimble GS Series Datasheet, 2005).

Elevation data were collected in the field by three different means: terrestrial laser scanner (TLS), total station (EDM: Electronic Distance Measurement) and differential GPS (figure 3.6). Technical details for the Trimble GS200 terrestrial laser scanner are given in table 3.2. Like its

airborne counterpart, the TLS uses a pulsed time-of-flight system to record points at 5000 pulses per second and can operate up to a range of 200 m (far beyond anything required in this study). The scanner uses a visible (green) laser, as opposed to the infra-red laser in the airborne system, but this makes little relevant difference to how it collects data. The Leica TCRP1205 total station records the locations of specific points relative to a known point or 'base station' by transmitting and receiving an infra-red laser beam to and from a specular target, such as a prism. Within a distance of 800 m between base and target, never exceeded in this study, points are located to an accuracy of 1 mm (Leica TCRP1205 Datasheet, 2009) and can be identified beneath the forest canopy by ID number and attributes attached to this identification by the user. While the TLS measures numerous points over a surface; known objects on the ground i.e. archaeology can be measured *and* identified with the total station. However, in order to utilise either technique, one or several control points must be referenced at a site to a national or international co-ordinate system using a Global Positioning System (GPS). The Leica GX1230+, utilised in this study, offers static and real-time RTK (base and rover) applications and with the addition of up to 6 m aerial extension, providing stronger communication between the GPS unit and satellites under forest canopy, points can be surveyed rapidly and typically with an accuracy of less than 10 cm. Raw GPS points are post-processed using RINEX data from the Ordnance Survey, including real-time models of atmospheric conditions, to obtain a more accurate position.

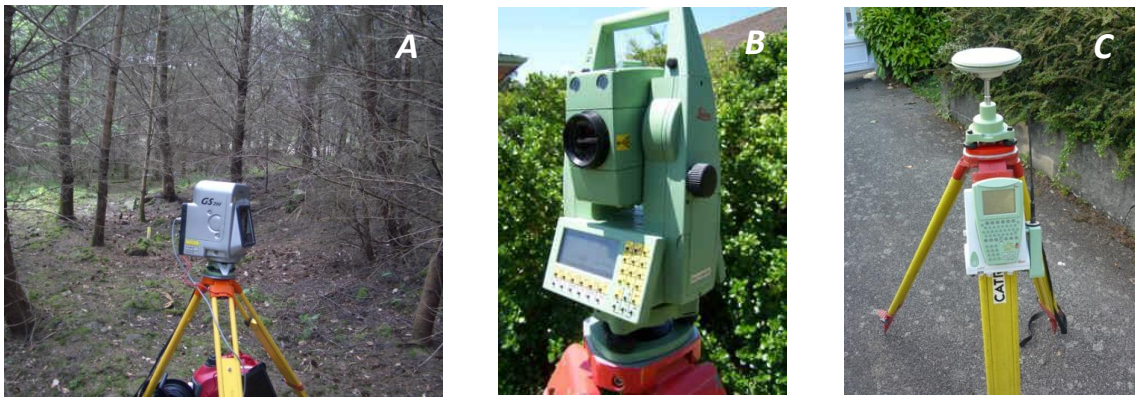


Figure 3.6 (A) Terrestrial laser scanner. **(B)** Total station. **(C)** Differential Global Positioning System.

The result of having to 'survey in' the TLS and total station data with the differential GPS is that the accuracy of a GPS control point is translated to its associated group of TLS and/or total station points, with error in the position or height of the GPS point systematically translated to all associated points. So a control point with three-dimensional accuracy of ± 10 cm assigns equivalent accuracy to all associated total station points, despite being surveyed themselves at an accuracy of 1 mm. Since the airborne LiDAR data (the primary data set) employed in this study is registered to an external reference frame (a national or international coordinate system), the terrestrial LiDAR and total station data has to be registered to the same reference frame otherwise direct comparison between data sets cannot be made. Therefore, although absolute uncertainty in field-acquired elevation data may be only 1 mm, the translation of points to an external reference frame introduces additional relative uncertainty.

Individual trees were surveyed according to the Forestry Commission Field Guide to Timber Measurement (Mackie and Matthews, 2008). Diameters were measured at *breast height* (DBH), which is the point on the tree 1.3 m above ground level or, on sloping ground, the 1.3 m level on the upslope of the tree. They were measured with girthing tape in rounded-down centimetre classes. Commonly *stand basal area* is calculated from the DBH by totalling the cross-sectional areas (at breast height) of all trees in a stand, followed by *stem density*, which is the stand basal area divided by the total plot area in hectares (Mackie and Matthews, 2008). The *stem height* of each tree was also measured with a Vertex-II hypsometer – the user standing at a distance from the tree of at least 1.5 times its height so as to judge the highest point of the trunk more precisely. In the same way, the *canopy base height* was measured as the lowest *live* branch in each tree canopy.

Various quantitative relationships have been established between characteristic tree dimensions, like those detailed above, and variables more difficult to measure directly, like, for example, tree volume (Zianis, 2005). One of these 'allometric relationships' can be exploited to estimate canopy density, through assessment of tree species and DBH against Forestry Commission form growth tables (Mackie and Matthews, 2008). Stem density and *canopy depth*

(the difference between stem and canopy base height) are, assuming robust allometric correlation, proportional to canopy density and, as such, are related to the degree of laser penetration through forest canopy.

Understorey vegetation height and density were measured across each site at randomly distributed points inside the plot area. The 'volume' of influential material covering the terrain surface was then calculated from density (in plan) and mean vegetation height.

Mapping, aerial photography and GIS data were required to inform initial site choices and also to provide context to the airborne LiDAR data, in addition to supporting archaeological feature validation following data analysis. High spatial resolution (10 cm) aerial photography was acquired simultaneously from the helicopter platform alongside the airborne LiDAR data with a Canon EOS 1d Mk3 digital camera. Ordnance Survey mapping dating from 1849 to the present day was obtained from the EDINA Digimap geo-data portal (EDINA, 2009), while oblique aerial photographs of the wood from 1947 and 1956, declassified from World War II Air Force surveys of the UK, were donated by a member of the local Chopwell Wood community (Fountain, 2010). GIS data were provided by the Forestry Commission, whose 1998 assessment includes information on tree species, planting year, sub-canopy topographic roughness and gradient and, given its status as ancient woodland, Chopwell Wood has changed little since.

3.3 Data Acquisition

3.3.1 Airborne LiDAR data

Airborne LiDAR data were acquired on the 18th and 19th July, 2009 by Network Mapping Ltd under favourable flying conditions. Data were obtained during the summer 'leaf-on' period for deciduous tree species when the seasonal variability in laser penetration through the forest canopy was at its minimum and understory vegetation density was also at its annual peak. This was a consequence of both restrictions on data availability, due to survey costs and the availability of the Network Mapping helicopter, and a desire to investigate the viability of sub-canopy archaeological prospection under the most rigorous conditions.

The data were pre-processed in-house by the service providers under the following routine. GPS and IMU data were processed through Applanix POSPac 5.0 under the loosely coupled GNSS mode using information from two ground-based stations. Optech's Dashmap was used to process the trajectory and range information into a georeferenced point cloud in the UTM Zone 30N WGS84 coordinate system. This was later converted to the OSGB36 co-ordinate system so that airborne LiDAR data matched the full data set. Finally Terrasolid's Terrascan software was used to match 17 flight strips with up to 50% overlap and rectify the point cloud. The mean first return point density of the pre-processed (raw) data, for all returns – i.e. from terrain, vegetation, buildings and anomalies – was 31 points per square meter.

3.3.2 *In situ* elevation training data for filtering

Typically commercial filtering algorithms for airborne LiDAR data require some user input in setting classification parameters (see section 1.3.3.3). Although there does not appear to be consensus on a single, outstanding method for deriving optimum filtering parameters, research frequently indicates that the most robust technique involves tuning a terrain filter to

known ground points, irrespective of the type of algorithm (Axelsson, 2000; Wack and Wimmer, 2002; Zhang et al., 2003). A training data set of pre-classified reference points must therefore be collected on which the filter algorithm can be tested and parameters constrained. The characteristics of the training data heavily influence the characteristics of the LiDAR points classified as terrain and, as such, the severity of the filter; i.e. in this study it is necessary that the filter must be tuned not only to reference points on the ground, but also to archaeological features on or within the topography, typical of those observed under Chopwell Wood. While the filter must be rigorous enough to limit Type II errors, it must also be sensitive to features on or in the terrain and not mistakenly filter them out.

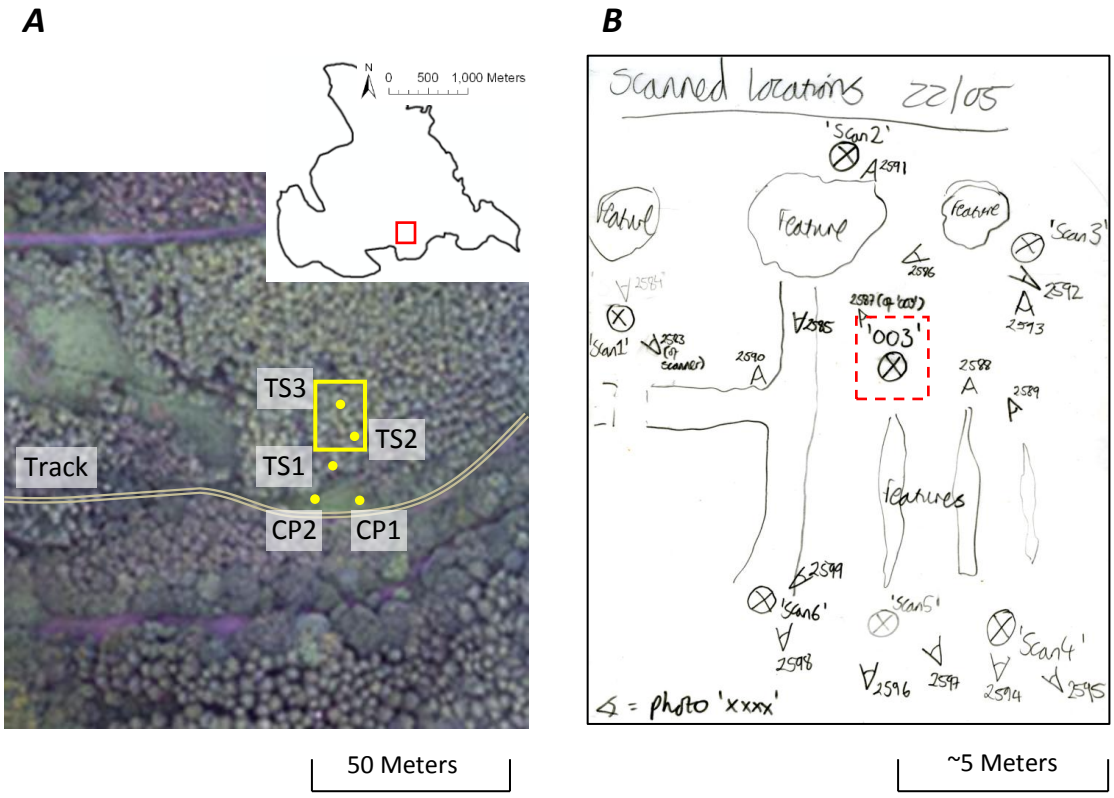


Figure 3.7 (A) Aerial photograph, the extent of which is marked by the red box on the inset map of Chopwell Wood, showing the location of Site 8 (yellow box). The locations of GPS control points, CP# and Total Station traverse points, TS# are given as yellow dots. **(B)** Field Map of Site 8 (22/05/10), the extent of which is marked by the yellow box on figure 3.7 A. illustrating the distribution of the seven TLS scan positions (X), including the central scan position which is highlighted by the dashed red box and is also the position of the centre of the site (TS3 on figure 3.7 A). Archaeological features are labelled and photograph IDs and locations are given by < symbols.

Site 8 was chosen to collect the training data for a number of reasons. Primarily it was apparent during the early field visits that the crowded Douglas fir at this site, only twenty years old, and the complex and steep topography (table 3.1) would likely limit both the number of laser pulses penetrating the canopy and the accuracy of the ones that did. The area had also been highlighted by members of a local community group as a site of archaeological interest, given the presence of a number of low walls, man-made channels and structural rubble beneath the canopy (figures 3.7 and 3.8). It was thought that this could have been the site of a forest house pre-dating the 19th century; and the size and shape of the features, in addition to the complexity of the vegetation and topography, offered an ideally harsh environment to fine tune the filtering algorithms.

Preliminary tests with the GPS confirmed original assessments of the dense tree canopy, with even hour-long surveys over a single position, using the full aerial extension, yielding poor horizontal and vertical accuracy (over 1 m and 3 m, respectively). So two GPS control points were set up in an open area approximately twenty meters from the site centre and the total station was then used to traverse into the forest (via points TS1 and TS2 in figure 3.7 A) to the site centre at point TS3 (003 in figure 3.7 B). Terrain elevation at the base of each tree was surveyed with the total station (for a regular distribution of ground points) along with points at the edges of and over the archaeological features. Full rotation terrestrial laser scans (360° FOV at 20 m range and 3 cm point resolution at 20 m) were taken from seven distributed locations over the ~0.4 ha site (points 003 and Scan 1-6 in figure 3.7 B) and the scan locations were recorded with the total station.

Figure 3.8 (A) Photograph Id2590 – Man-made channels. **(B)** Photograph Id2586 – Fir trees over what is thought to be building rubble. See Ids in figure 3.7 B for photograph locations.

3.3.3 *In situ* data for accuracy assessment

Further to acquiring training data to process the airborne LiDAR, field data were also collected for accuracy assessment (see section 1.3.6). To facilitate meaningful comparison between LiDAR and reference elevations, the field data were collected in comparable conditions to the airborne laser data: during spring and summer, so that deciduous trees were in-foliage and understory vegetation was at its most dense (e.g. figure 3.9 A). Hence, all *in situ* data were acquired during a field season lasting from approximately the start of March to mid-July and involving over 20 days on-site. Data were collected inside eight circular sites (figure 3.5 and table 3.1), including site 8 despite its additional purpose, of radius 20 m (approximately 0.12 ha). Typically the Forestry Commission use plot sizes of 0.01 to 0.05 ha as a representative sample for timber measurement, depending on the tree spacing in a particular stand, so a plot size of > 0.1 ha would offer a large sample whatever the tree species or spacing (Mackie and Matthews, 2008). Plot centres were located with GPS (utilising aerial extension capabilities) so that full rotation terrestrial laser scans (360° FOV at 20 m range and 3 cm point resolution) could be taken of the entire site.

Individual tree diameters and stem and canopy heights were surveyed using girthing tape and a Vertex-II hypsometer. However, tree attribute acquisition was not straightforward due to the occurrence of, for example, multiple trees at the same location – a product of split trunks below the 1.3 m DBH measurement point or numerous germinations. Ensuring that data were consistently acquired with a view to how it would be expected to affect the airborne LiDAR, a number of rules were established for these incidences. The height of a 'multi-tree' was taken as the highest stem, as the most elevated section of the canopy has principal control over penetration (Amable et al., 2004), while multi-tree DBH was taken as the sum of all stem diameters, as each stem contributes to the density of the total canopy cluster. By considering the probable effect of stem density on laser penetration rate, a multi-tree was counted as: 1 + 0.5*[number of stems] in order to describe a larger/denser canopy but limiting its influence given that, for example, a multi-tree with 3 stems was assumed to provide typically greater obstruction to the laser than a single stem, but less than three distributed trees.

Understory vegetation 'volume' (density multiplied by mean height) was recorded at each of the eight sites. Density was measured as the percentage cover (in plan view) to the nearest 5 %, so, for example, close to total cover under the pine (figure 3.9 A) and entirely exposed terrain under the Douglas fir (figure 3.9 B). Mean height was calculated as the average understory height to the nearest 10 cm, with a greater number of measurements taken at sites with higher density vegetation, i.e. up to one every 5 m^2 at sites like the pine (figure 3.9 A).

Figure 3.9 (A) Dense low vegetation under the pine at Site 2 (with the plot marker hidden). **(B)** No low vegetation under the Douglas fir at Site 7.

3.3.4 Terrestrial survey for feature validation

Following processing and analysis of the airborne LiDAR data, candidate archaeological features were validated by a number of methods. Doneus et al. (2008), among others, argued that 'low micro topographic earthworks, especially when covered with dense vegetation... can be difficult to locate on the ground, even by an experienced surveyor', so handheld GPS was used to pinpoint the general location of subtle features for visual assessment. Qualitative assessment was conducted primarily by taking photographs of features from below the forest canopy to provide additional information on stonework, colour of soil, surrounding features etc. to the topographic information provided by the LiDAR data. Where a candidate archaeological feature was identified in the airborne LiDAR data but greater topographic detail was required for validation, the feature, and often its surrounding area, was surveyed with the Leica GPS 1200.

3.4 Data Pre-processing and LiDAR Point Cloud Classification

3.4.1 Elevation data registration

Field-acquired GPS (and associated total station) data were post-processed with active RINEX data from the Newcastle Station, approximately 14 km from Chopwell Wood, using Leica's Geo Office software. Processed points were registered to the OSGB36 co-ordinate system. The terrestrial LiDAR data were georeferenced with corresponding GPS control points in Trimble's Pointscape but, having no control on scan rotation, each point cloud had to then be registered in the z-axis. Scans were matched to the local terrain slope (derived from preliminary classification of the airborne LiDAR) through fine-tune rotation in Archaeoptics' Demon software. The airborne LiDAR data were registered to the OSGB36 coordinate system in-house by the service providers.

3.4.2 Quality control for airborne LiDAR data

The airborne LiDAR data were calibrated in-house by the service providers, using 85 ground truthing points at a Network Mapping calibration site as control. A mean offset (systematic bias) of 0.06 m between the laser and ground was removed from the point cloud. High and low outliers were removed with Terrascan's 'isolated points' and 'low points' filtering algorithms – simple procedures for identifying anomalies based on unreasonable distances between points.

3.4.3 LiDAR point cloud classification

Of the numerous algorithms presented for filtering LiDAR data (see section 1.3.3.1), the progressive densification method of Axelsson (1999; 2000) offers most to this study. It fares well in both Sithole and Vosselman's (2004) ISPRS filter comparison and Meng et al.'s (2010)

review, and is especially impressive at filtering complex objects, discontinuous terrain and vegetation (table 1.1). It is additionally employed as the 'ground' filter in the powerful Terrascan processing software, which was freely available to the author. An approximation of the terrain surface is first obtained through block minimum analysis, with the user defining the block diameter by setting the maximum building size – the largest possible horizontal distance between ground points (figure 3.10 A). Points are triangulated and the TIN then 'fluctuates' up through the point cloud with further ground points added if they fall within pre-defined description length and angle criteria (figure 3.10 A 'iteration distance' and 'iteration angle', figure 3.10 B and see also figure 1.8). The maximum terrain angle can also be defined to prevent unrealistic commission of ground points over larger areas (figure 3.10 A). An example of a transect through classified ground points (and other categories) is given in figure 3.10 C.

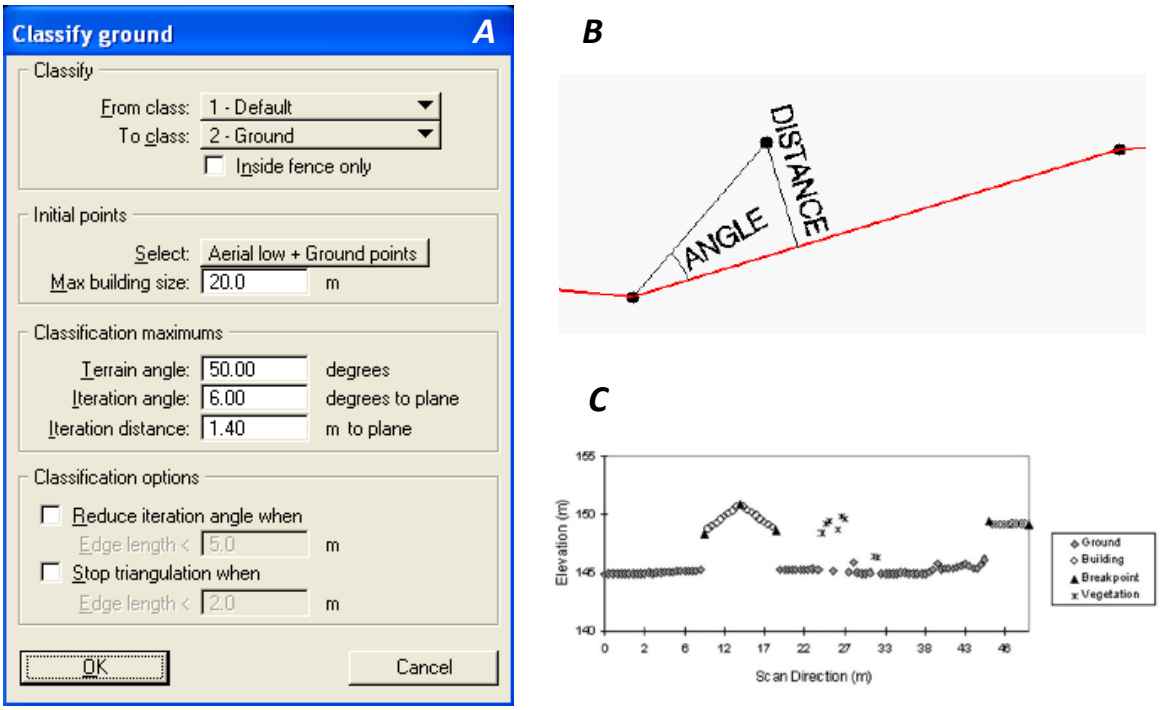


Figure 3.10 (A) User-defined parameters in the Terrascan 'ground' classification algorithm. **(B)** Iteration distance and iteration angle between candidate points. Source: Terrascan Manual (2010). **(C)** A transect through LiDAR points classified by the progressive densification method. Source: Axelsson (2000).

Terrain filtering parameters were constrained using the *in situ* elevation training data collected at site 8. The GPS and total station ground control were used to classify the terrestrial LiDAR data, which subsequently provided control for classifying the airborne LiDAR (figure 3.11). This two stage process was necessary because airborne laser penetration through the fir canopy was infrequent and ALS (airborne LiDAR) terrain points were randomly distributed. Very few ALS and ground control points related spatially and thus assessing correlation in elevation between ALS and training data was relatively meaningless. Employing the terrestrial laser scan points, relatively concentrated in space, as intermediaries between ground control and airborne data provided more reliable filtering.

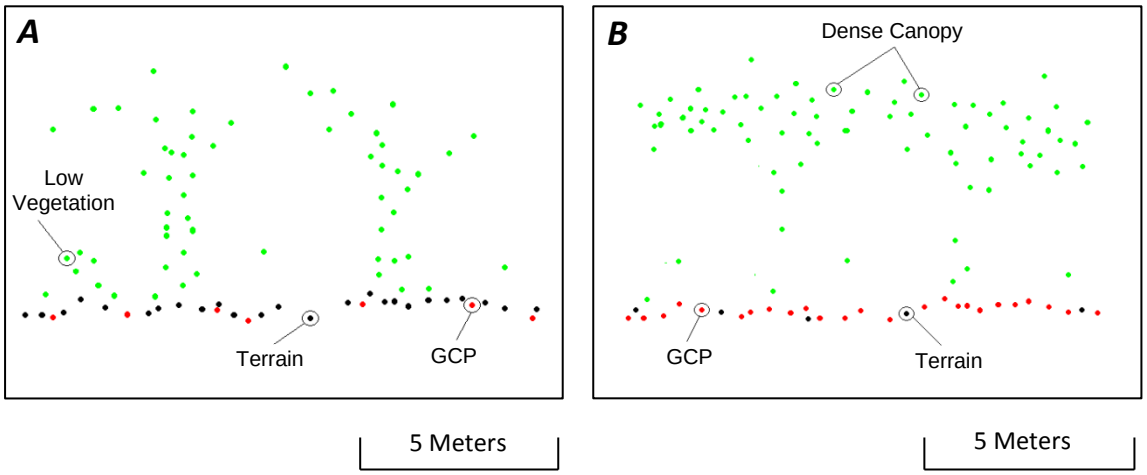


Figure 3.11 Cross-sections through hypothetical LiDAR data. Point cloud classification: green = vegetation, black = terrain and red = ground control point (GCP). **(A)** Terrestrial LiDAR classified into vegetation and terrain using total station GCP's. **(B)** Airborne LiDAR classified into vegetation and terrain using the terrain returns from the terrestrial LiDAR classification as GCP's.

Terrascan offers an effective means of testing the correlation between two point clouds covering the same area. The 'control report' procedure selects points between the two datasets closest in the x-y plane and runs, on these point-pairs, simple least-squares regression in elevation, with output including: (1) the mean offset between all tested points, (2) the standard deviation of this offset and (3) the root mean square error of this offset. Mean offset describes the systematic over- or under-estimation (bias) of LiDAR terrain with respect to

reference data while standard deviation describes the variability (precision) of LiDAR terrain with respect to reference and RMSE describes the variability of the height measurement between the LiDAR and reference data. Collectively, these three statistics describe LiDAR accuracy. The control report procedure can be managed by changing the maximum permitted distance and slope between a pair of tested points.

Owing to the prerequisites on filter sensitivity mentioned in section 3.3.2, the filtered terrestrial LiDAR training data were compared, using the control report function, to both the full sample of total station ground control points and to only those identifying archaeological features on the terrain surface. In this way, the ability of filtered LiDAR terrain to pick out specific features, in addition to the full ground surface, could be analysed. Optimum filtering parameters for the terrestrial LiDAR data were obtained by iterating round the routine: (1) filtering with angle x and distance y, (2) performing a control report against all ground control points, (3) performing a control report against the archaeological control points, (4) output mean offset, standard deviation and RMSE. Low mean offset and standard deviation, combined with a reasonable retention of points, identified the strongest set of parameters.

Airborne LiDAR data for the area covering site 8 were filtered by the same procedure as that described above, using the points retained as terrain in the classification of the terrestrial LiDAR data as control. Only the single control report was produced at each cycle, however, due to a lack of archaeological feature identification in the terrestrial laser points; but with the optimum parameters at the first step purposely sensitive to archaeological features in the topography, testing at the second against these filtered points was, as such, equally sensitive. Accordingly, optimum filtering parameters for the airborne data were derived and these were used to classify the point cloud over the full study area into terrain and non-terrain categories. Figure 3.12 is a flow diagram outlining the various steps of the point cloud classification procedure.

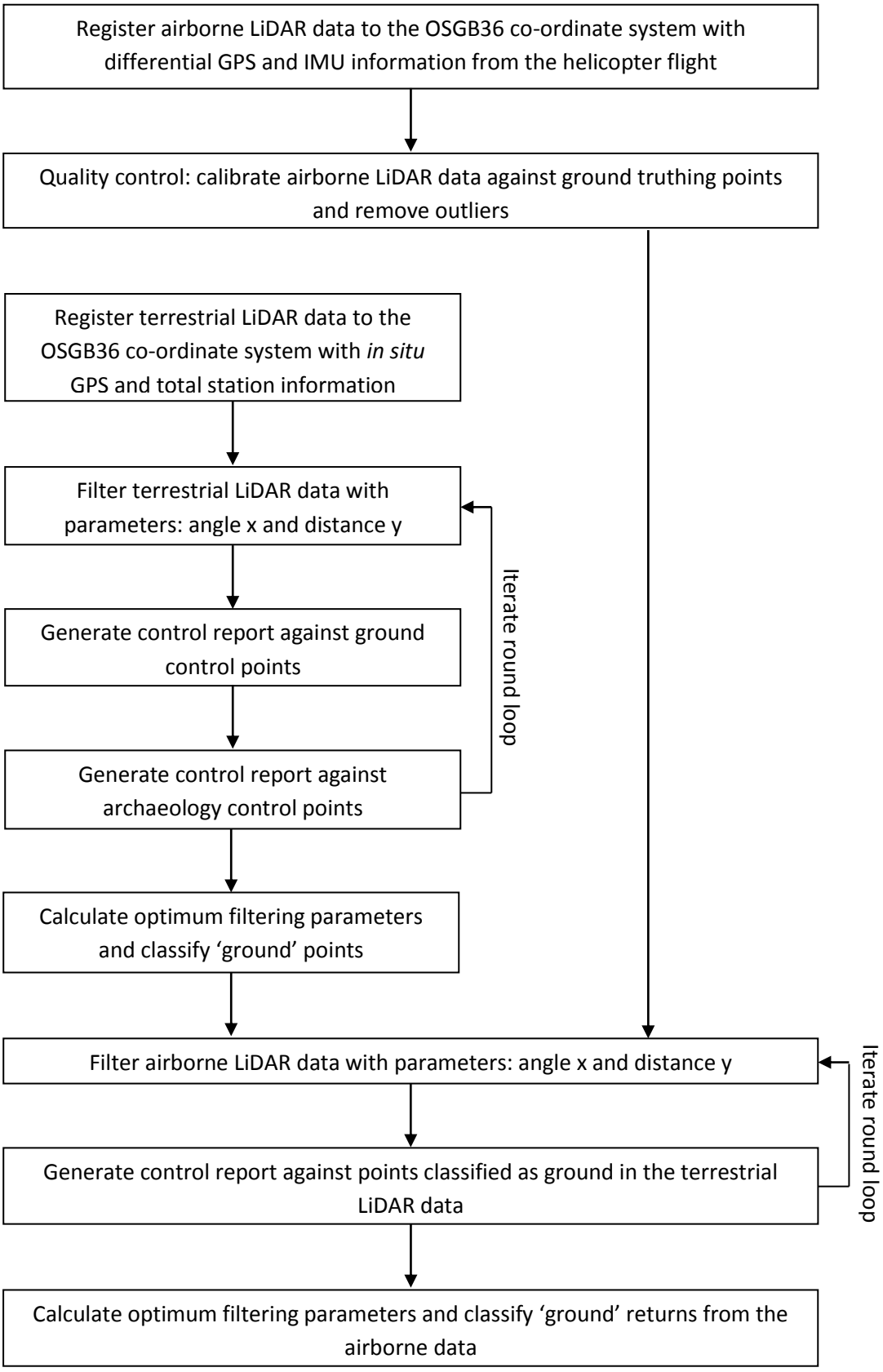


Figure 3.12 Flow diagram for the airborne LiDAR point cloud classification procedure

3.5 LiDAR Terrain Modelling and Accuracy & Terrain Point Density Assessment

3.5.1 Evaluation of terrain model derivation procedure

After the pre-processing stage, filtered airborne LiDAR data were modelled for computer visualisation. The structure of a terrain model can, to some extent, be modified depending on the technique employed to interpolate between LiDAR points (see section 1.3.5) and the ability to visualise and identify topographic, archaeological features may then depend on this choice.

It was mentioned in the previous chapter that options for the choice of modelling procedure are often restricted by computer software availability and that the optimum technique can change depending on the modelling scenario. Though primarily utilised for pre-processing LiDAR data, the Terrascan software additionally offers a number of output models for filtered terrain data. The first is a triangular irregular network (TIN), which, as stated in section 2.2, maintains all the accuracy of the input data; but a lack of interpolation, or smoothing, between nodes can hide objects in the terrain that would otherwise be clear (Li et al., 2005; Crutchley, 2008). Also provided are raster-based rectangular arrays of grid cells, with elevation values determined from lowest, highest or average point elevation within each cell and basic linear interpolation performed at cells without point coverage. However, isolated points cause biased elevation estimates, and clustering and isolation are typical in real sub-canopy LiDAR terrain data. Interpolation techniques of greater sophistication were available and are detailed below, but the principal advantage of using Terrascan was its processing power and subsequent rapid modelling procedure.

Various multivariate interpolation techniques for vector data are available in commercial GIS and image processing software, including ERDAS Imagine, ITT Visual Information Solutions' ENVI and ESRI's ArcGIS. The outstanding techniques are: inverse distance weighting (IDW), based on calculating a value at an unknown location as the weighted sum of a user-defined

sample of known points; kriging, which calculates the value at a point by linear least squares estimation; and spline interpolation, which fits a surface defined piecewise by polynomial functions (Li et al., 2005). Where terrain points are relatively sparse or widely spaced, for example under some areas of forest canopy, kriging has been recommended as the best predictor of the three (Lloyd and Atkinson, 2002; Pfeifer and Mandlburger, 2009), consequently providing better computer visualisation.

In order to establish the most accurate modelling procedure, the three preeminent interpolation techniques given above, along with standard linear interpolation, were tested by comparing modelled airborne LiDAR surfaces with reference terrain data. Airborne LiDAR surfaces for each interpolation technique were generated through the same filtered terrain points at Site 8 and evaluated against the terrestrial LiDAR terrain data. The terrestrial LiDAR data was chosen as reference terrain data as it provided as close a match to the true terrain surface as possible. Mean offset and standard deviation between surface and reference were calculated and these statistics were used to determine the most effective point modelling procedure.

3.5.2 Accuracy, bias and terrain point density assessment

Comprehensive accuracy and terrain point density assessment covered four areas: (1) standard assessment involving cross-tabulation of filtered LiDAR data against classified ground truth data (see section 1.3.6) for eight study sites; (2) within-site LiDAR terrain point clustering assessment; (3) assessment of the influence of slope on LiDAR terrain accuracy and (4) full study area assessment. The latter extends the results obtained at the eight test sites to the entire wood and the second and third are supplementary analyses, conducted to investigate issues with LiDAR terrain accuracy identified during preliminary experimentation with the Chopwell Wood data set, but barely reported in previous research.

3.5.2.1 Standard accuracy, bias and terrain point density assessment

A comparison between the airborne LiDAR and *in situ* reference elevation data across all eight study sites helps to ascertain the relative success of airborne laser scanning as a viable technique for archaeological prospection over a variety of land covers.

Airborne LiDAR terrain data, filtered with the optimum parameter set (established in section 3.4.3), were tested for accuracy in elevation at each site by evaluation point-to-point against terrestrial LiDAR terrain reference data. After removing vegetation (using tested filtering parameters), terrestrial LiDAR data were assumed to represent the true terrain surface. The 'control report' function in Terrascan was again utilised to calculate mean offset, standard deviation and RMSE between closest pairs of points in the x-y plane (figure 3.13 A). A surface-to-surface assessment was also performed at each site after interpolation through respective airborne LiDAR and terrestrial LiDAR (reference) points with the chosen modelling procedure (see section 3.5.1). As processed LiDAR data are displayed as a digital terrain model prior to visual analysis, it was crucial to ascertain the difference between terrain point and surface accuracies. This involved calculating mean absolute offset (positive plus negative offset) in elevation between airborne LiDAR and reference surfaces (figure 3.13 B) using Arc GIS software. Absolute offset describes the mean difference between surfaces, irrespective of direction, rather than mean offset (often termed 'signed' offset, e.g. Hodgson et al., 2003) which describes net over- or under-estimation (calculated in the point-to-point assessment). Both mean offset and mean absolute offset are measures of bias (systematic error), while standard deviation is a measure of precision (random error) and RMSE measures the variability of the height measurement between LiDAR and reference. Published accuracy assessments offer one or more of these four statistical measures, so, in order to comprehensively evaluate the results from this study, all four measures were calculated.

Filtered LiDAR terrain was evaluated quantitatively against cross-tabulated terrain reference data for the tree species beech, pine and fir. With a view to clarifying broader relationships

between the quality of airborne LiDAR sub-canopy terrain measurement and various attributes of the forest environment, the three LiDAR elevation accuracy statistics and spatial point density were compared with tree canopy variables, understory vegetation and roughness/gradient of the topography through linear regression analysis.

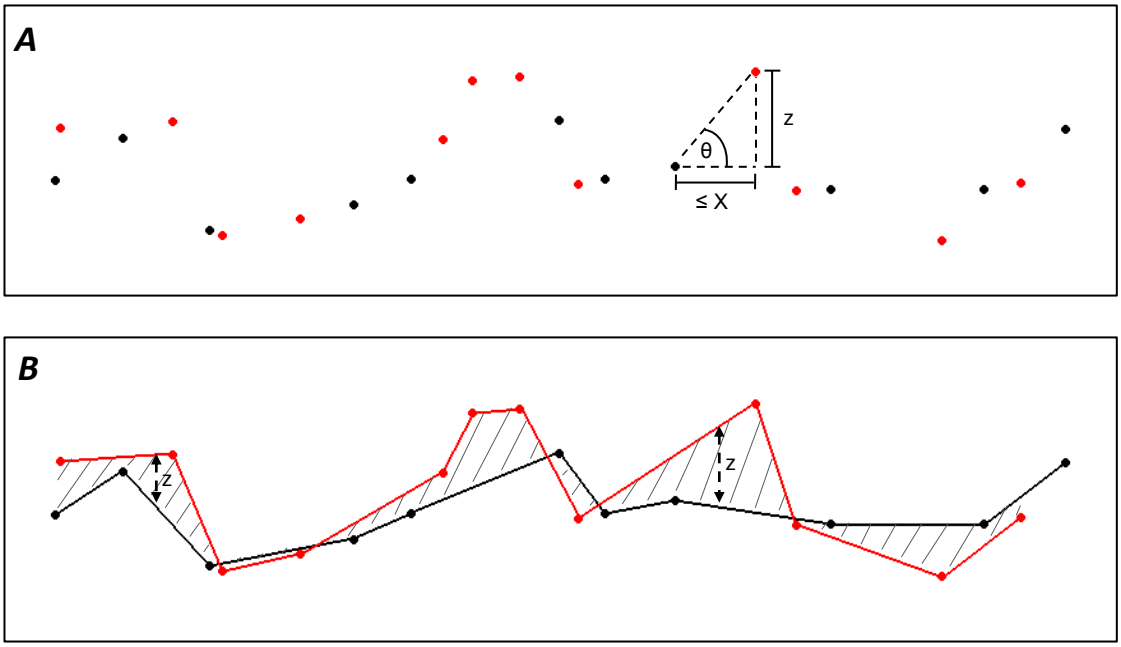


Fig 3.13 (A) Point-to-point accuracy assessment, where X is the maximum tolerated horizontal distance between tested pairs of points and z is the offset in elevation between the points. **(B)** Surface-to-surface assessment, where z is the offset in elevation between the surfaces, with absolute offset represented by the hashed grey area.

3.5.2.2 Terrain point clustering assessment

It was noticed during early field visits, and later discussed at the preliminary stage of method planning, that several of the factors highlighted in previous research as influential to the passage and reflection of a laser pulse through the forest, i.e. canopy structure and understory vegetation, were characteristically regular and clustered. Stand-scale canopy alternates between dense and dispersed areas of foliage depending on the positions of individual trees, which are typically regularly spaced in managed forest, and understory vegetation appears

most dense at gaps between tree crowns where light penetrates to the forest floor. There is, of course, a significant random component to this pattern, but statistical autocorrelation and clustering analyses were undertaken to evaluate to what degree this affected the positions of terrain points in the filtered airborne LiDAR data.

Moran's I three-dimensional autocorrelation evaluates whether a pattern between points is clustered, dispersed or random based on XYZ co-ordinates. A null hypothesis was setup prior to undertaking the analysis that stated there was complete spatial randomness between points, i.e. that there was *no* pattern between LiDAR returns in space. The index is given by Griffith (1987) as:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i \, z_j}{\sum_{i=1}^{n} z_i^2} \tag{3.1}$$

where z_i is the deviation in elevation of feature i from its mean $(z_i - \bar{Z}), w_{i,j}$ is the spatial weight between feature *i* and *j*, *n* is equal to the total number of features and S_0 is the aggregate of all the spatial weights:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \tag{3.2}$$

An index above 1 indicates increased clustering, near 0 a random distribution and below 1 increased dispersion between points. To test the statistical significance of autocorrelation index results, p-value and Z-score values were also calculated. A p-value describes the probability that the observed spatial pattern between points was created by some random process and the Z-score is a measure of standard deviations. A subjective judgement must be made regarding the accepted degree of risk that the null hypothesis is falsely rejected, i.e. accepting a spatial pattern between points to a confidence level of 95, 99% etc.; but the higher confidence levels are associated with small p-values and very high or very low (negative) Z-scores (relating to the tails of the normal distribution through, in this case, LiDAR point elevations). Where autocorrelation analysis yields a small p-value and either a very high or a

very low (negative) Z score, it is extremely unlikely that the observed pattern is some version of the theoretical spatially random pattern represented by the null hypothesis.

The Getis-Ord General G clustering analysis evaluates the degree of concentration between points in two-dimensional space. The index is given by Getis and Ord (1992) as:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j} x_i \, x_j}{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i \, x_j}, \forall j \neq i \qquad (3.3)$$

where x_i and x_j are spatial attribute values for features *i* and *j* and $w_{i,j}$ is the spatial weight between them. The index value presents less information than the autocorrelation equivalent, but if the confidence (p-value) is small and the absolute value of the Z score is large enough that it falls outside of the desired confidence level, there is spatial clustering in the x-y plane.

Both tests were conducted between airborne LiDAR returns within each of the eight test sites and compared to the following field data: (1) understory vegetation density, (2) the stem density of primary tree species, (3) mean canopy depth and (4) tree count. In this way, the original suggestions that regular vegetation structure might lead to spots of high laser penetration, and thus clustered terrain returns, were properly evaluated.

3.5.2.3 Influence of terrain slope on accuracy

It was necessary to attempt to establish whether terrain slope influenced the accuracy of airborne LiDAR terrain measurements. This was investigated by comparing the accuracy of a LiDAR point when the laser was assumed to illuminate a flat footprint on the ground and a footprint whose size was proportional to the terrain slope.

In a modern sensor, the diameter of the laser footprint on the ground is typically > 20 cm (Mallet and Bretar, 2009) and depends primarily on the height of the aircraft and the angle of the scanner at any point in time. While the LiDAR sensor utilised in this study could locate up to four surfaces within the footprint; it could not discriminate between surfaces separated by less than 15 cm elevation (see section 1.2.3) and therefore could not resolve topography at a

scale below the diameter of the laser footprint. Where topography within the footprint was complex, the LiDAR sensor would record only the energy peak of the returning pulse (Wehr and Lohr, 1999). It was necessary to establish whether within-footprint topography recorded by the sensor changed with terrain slope.

The footprint assessment involved testing filtered airborne LiDAR terrain points against terrestrial LiDAR reference terrain points across the eight field test sites. Buffer zones were first created around each ALS return based on the *average* size of the laser footprint during helicopter flight. Obviously the size of the footprint would have varied throughout LiDAR data acquisition, but the information required to calculate the size of each footprint individually could not feasibly be collected. Instantaneous laser footprint diameter, A_{Linst} is given by:

$$A_{L_{inst}} = \frac{h}{\cos^2(\theta_{inst})}\gamma \tag{3.4}$$

where h is the flying height, γ is the laser beam divergence and ϑ_{inst} is the instantaneous scan angle, itself given by:

$$\theta_{inst} \in [0, \pm\frac{\theta}{2}] \tag{3.5}$$

where ϑ is the maximum scan angle (Wehr and Lohr, 1999). The maximum and minimum laser footprint size during flight was calculated using the range in aircraft altitude and maximum scan angle, obtained from the Network Mapping flight reports, and the beam divergence for the laser in the Optech ALTM 3100EA ranging system, taken from Mallet and Bretar (2009). Given the regularity of the scanning procedure and low deviation in aircraft altitude during data acquisition, average footprint size was calculated as the midrange size. Without the capacity to identify exact footprint shapes at ground level, the buffer zone around each LiDAR return ('footprint centre') was assumed circular.

However, equation 3.4 is valid only under the assumption that the terrain surface is consistently flat, which is not always the case, especially at Chopwell Wood given the range in elevation across the study area. To be able to assess the influence of terrain slope on LiDAR point accuracy, a more complex version of equation 3.4, which includes terrain inclination as

an independent variable, was used to calculate a second version of each laser footprint at each field test site.

Across track instantaneous laser footprint diameter, A_{Linst} is given by:

$$A_{L_{inst}} = \frac{2h \sin\left(\frac{\gamma}{2}\right)\left(\cos\ (\theta_{inst} + \ i) + \sin(\theta_{inst} + \ i)\tan(\theta_{inst} + \ i) + \frac{\gamma}{2}\right)}{\cos\left(\theta_{inst} - \frac{\gamma}{2}\right)} \quad (3.6)$$

where *i* is the inclination angle of the local planar terrain surface. At each LiDAR point, the footprint was calculated with: (1) a constant diameter based on average flying height and scan angle and (2) a diameter based also on the inclination of the local slope. Local terrain inclination was previously calculated by generating a low resolution slope map from the filtered airborne LiDAR data. The significant range in gradient across the eight test sites (see table 3.1) helped to determine whether terrain slope was a controlling factor on laser accuracy.

The accuracy of each airborne LiDAR terrain point with respect to the detail of the true topography was calculated by evaluating the mean offset and standard deviation in elevation between all terrestrial LiDAR reference points within the boundary of the laser footprint and the elevation of the airborne LiDAR return. Mean offset and standard deviation were calculated for both constant diameter footprints and footprints with diameter proportional to the terrain slope, thereby emphasising the role of slope, if any, on LiDAR measurement accuracy.

3.5.2.4 Full study area assessment

With the intention of evaluating the viability of airborne laser scanning for sub-forest canopy terrain derivation, accuracy and terrain point density assessments were conducted across the entire study area. This involved assessing both LiDAR terrain point density and systematic &

random error under each of Chopwell Wood's nineteen dominant tree species (see Appendix 1 for a list of species). The optimum filtering parameters, determined in section 3.4.3, were employed to classify the complete airborne LiDAR dataset into 'terrain' and 'non-terrain' categories, and subsequently terrain points were separated spatially according to forest boundaries. These boundaries were obtained from the Forestry Commission GIS (1998) and located single (or mixed) species 'compartments' (stands) within the forest.

After isolating terrain returns under discrete areas of canopy, LiDAR terrain point density was calculated by compartment. Despite the airborne LiDAR system acquiring raw data at regular across track point spacing (see section 1.2.2), *along* track spacing, proportional to the speed of the aircraft, varied and accordingly, terrain point density had to be normalised by the point density of the original, raw data. Mean raw point density (points per m^2) was first calculated for the full study area. The mean penetration rate (terrain point density divided by raw point density) for each compartment was then determined and normalised terrain point density for each compartment was calculated as penetration rate multiplied by the mean raw point density the full study area. Finally, mean terrain point density for each species was obtained by calculating the average across all compartments of the same species, weighting the contribution of a compartment to the mean based on its area. Thus, mean, normalised LiDAR terrain point density was determined for Corsican pine, Norway spruce, Japanese larch etc. (see Appendix 1). Mean terrain point density was also determined for five classes of similar species: (1) pine family minus fir, (2) fir, (3) cypress family, (4) beech, sycamore and ash families and (5) birch family (see Appendix 2 for details on classes), and for deciduous and coniferous forest.

The quality of a digital terrain model is proportional to both the number of points contained within it and the accuracy of those points (Pfeifer and Mandlburger, 2009). Accuracy was estimated over the full study area by extrapolating information obtained at the eight test sites, i.e. at beech, pine and fir sites, in section 3.5.2.1, to all forest compartments. A universal assumption was made that these deciduous, pine and fir sites were symptomatic of equivalent

stands across the wood. A measure of 'viability' for archaeological prospection (i.e. for measuring sub-canopy terrain) at each forest compartment was determined by weighting terrain point density according to high/medium/low point accuracy (so that high point density with high point accuracy would generate the highest viability and vice versa). This measure provided an indication of areas of Chopwell Wood amenable to resurvey and of comparable areas at other locations that would be suitable for survey with airborne LiDAR.

3.6 Visualisation and Analysis

3.6.1 Vertical exaggeration for feature detection

The first step in post-processing the digital terrain data involved exaggerating topography to highlight archaeological features against their surroundings. Exaggerating the DTM vertically in areas of interest was followed by generation of difference models against the local topography and slope models, based on the methods of Sittler (2004) and Crutchley (2008).

Local relief models (LRMs), as suggested by Hesse (2010), were also generated where required by extracting low-relief features, covering a range of sizes typical of industrial and pre-industrial archaeology (e.g. charcoal platforms, mining and quarrying pits and heaps etc.), from the DTM and eliminating as far as possible the small-scale (<1 m diameter) landscape forms from the data (see section 2.3). An LRM reflects less biased elevation information for small-scale features relative to the landscape at large than is typical in a simple difference model (Hesse, 2010). Deriving a finished LRM is more complex than any of the other elevation manipulation techniques and comprises several processing steps: (1) application of a low pass filter (with a 3 x 3 kernel) over the DTM, (2) extraction of zero-metre contours from a difference map between the original and filtered DTM and (3) interpolation of these contours to a new DTM 'purged' of the small-scale features. However, the finished product can significantly enhance some of the most subtle topographic features.

3.5.2 Artificial illumination for feature detection

Further to modifying a digital terrain model directly, the manner in which the model is viewed can also be varied by changing its environment. Visualisation depends on the position of an artificial light source over the model and subsequent topographic shading effects (see sections 2.2 and 2.3). The light source is fully adjustable, both in terms of intensity and position; the latter varied by angle around the azimuth and altitude (hemispheric angle). This is the basis of the 'hillshade' model (figure 3.14) and was the second step in feature detection over the processed Chopwell terrain data after vertical exaggeration.

Areas proving difficult to interpret following hillshade analysis from single light sources, due to excessive shadowing or feature alignment parallel to the source, were post-processed as false colour composite images of several illumination angles and azimuths, according to the method of Devereux et al. (2005) (see section 2.3). In areas of extreme complexity, composite images were generated by principal components analysis (PCA). As stated in section 2.3, the first three components of a PCA typically contain over 95% of the variability in the original dataset and Devereux et al. (2008) suggested that the third component of a PCA between 16 images demonstrates almost equal loading in all directions, aiding detection of the most subtle features. Devereux et al. (2008) also found that the third component displayed this same trait across a variety of sites. So PCA between 16 images was utilised in areas of complex topography, where feature detection using standard techniques was impractical, with the aim of identifying subtle, previously concealed archaeological features.

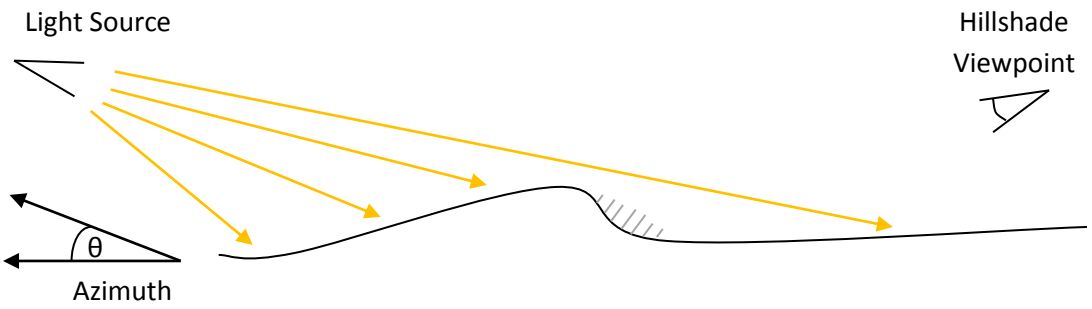


Figure 3.14 Hillshade technique. The light source over a DTM shades the backside of topographic features (grey hashed area).

Chapter 4. Results and Analysis 1: Airborne Laser Scanning Over a Forest Environment

Results are separated into two chapters. Chapter Four covers objectives 3 and 4: constraining optimum LiDAR filtering parameters, selecting the most favourable modelling procedure and accuracy assessment. Chapter Five covers objectives 5 and 6: archaeological feature detection and validation.

4.1 Airborne LiDAR Data

Following quality control and outlier removal, the raw airborne LiDAR data were investigated for unresolved issues. The size of the data set (> 20 million points) was identified early on as problematic, given that the Terrascan software is capable of loading roughly only half of these and processing (i.e. running filtering routines on) even fewer (Terrascan Manual, 2010). So the data were separated into numerous smaller point cloud tiles of 250 x 250 m area. The Forestry Commission GIS 'subcompartment' database was imported into Terrascan to allow point clouds and individual forest compartments to be analysed when, for example, looking at one of the eight field test sites.

Although outlier removal had already been performed, it was clear that some anomalous returns remained in the data set. All sub-ground outliers appeared to have been removed correctly, but a few, way above the ground surface, potentially presented unavoidable obstacles to the filtering procedure and had to be identified before continuing. Some of these outliers were clearly recognisable as high-voltage power lines (figure 4.1) or returns from low-flying birds, but were sufficiently isolated from each other that they were not expected to cause problems. Outliers could only influence the filtering procedure if they were clustered to such an extent that the initial block minimum algorithm selected one for the starting TIN (Sithole, 2005).

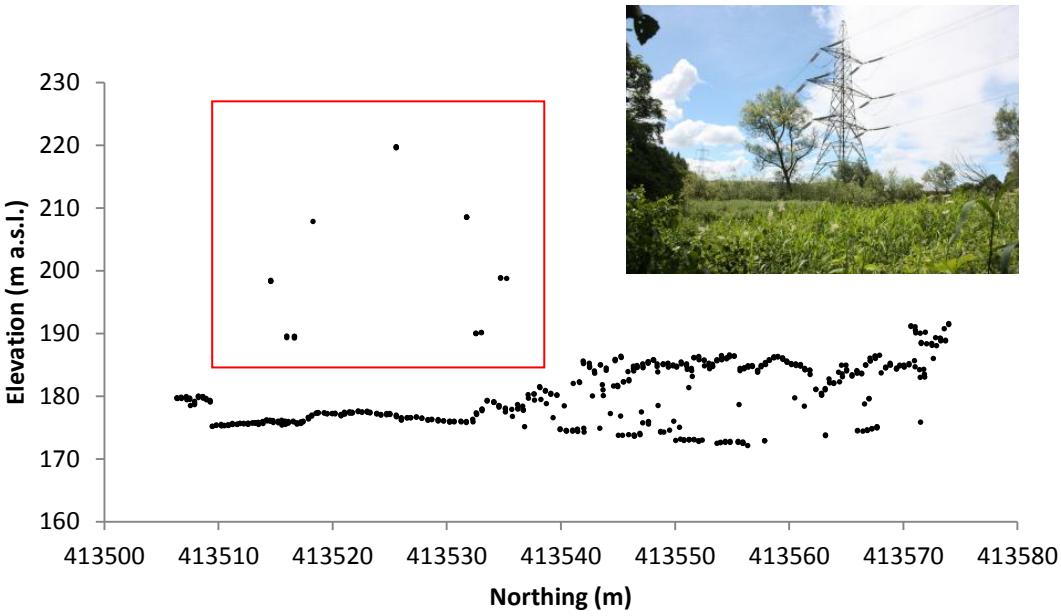


Figure 4.1 Transect through raw airborne LiDAR points at Chopwell Wood, with 'outlier' returns from a high voltage power-line highlighted. Inset is a photograph at the same location.

4.2 Constraining Optimum LiDAR Terrain Filtering Parameters

Site 8 provided an ideal location to evaluate the effect of varying terrain filtering parameters over a range of topography and different types of archaeological feature under dense overlying conifer plantation, typical of many European forests. The locations of the ground control points (GCPs) recorded with the total station and GPS equipment are illustrated in figure 4.2 B. GCPs were evenly distributed and split between terrain and man-made structures (note the corresponding feature outlines in figure 4.2 A and B).

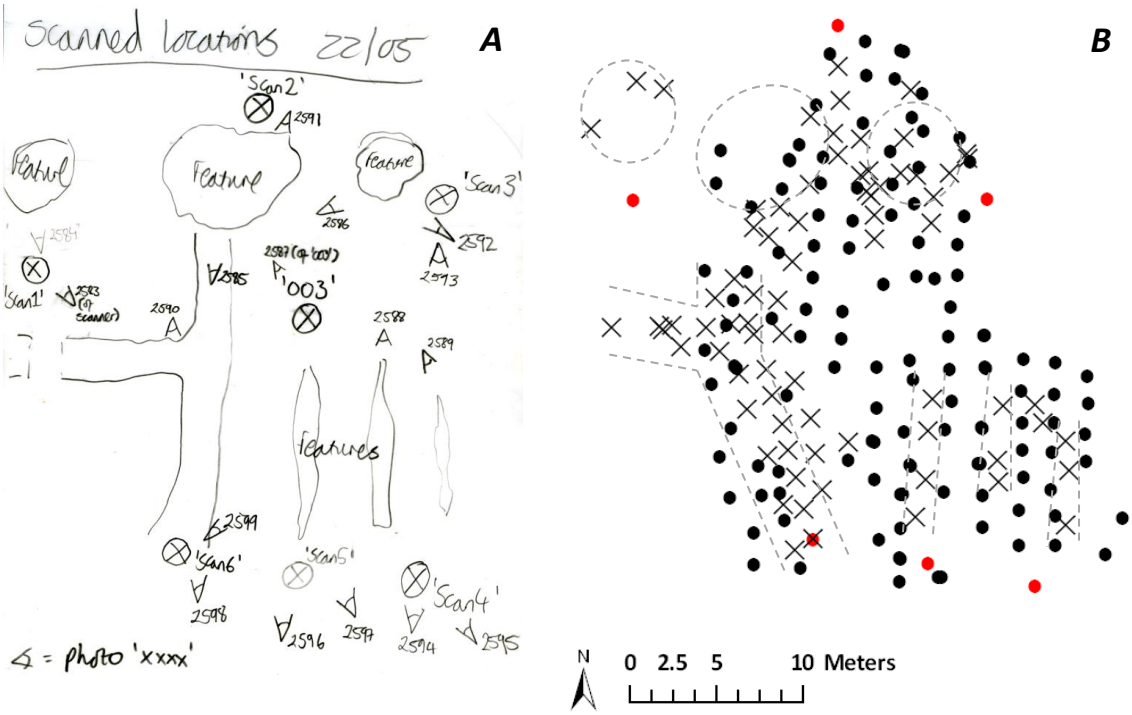


Figure 4.2 (A) Field map of Site 8 showing scan locations and archaeological features. **(B)** *In situ* reference data (GCPs) at Site 8, where: black dot = control point over terrain, red dot = terrestrial LiDAR scan location, cross = control point over archaeological feature and dashed grey line = outline of archaeological feature (corresponding with figure 4.2 A).

Parameter	Description and Derivation	TLS Filter	ALS Filter
Maximum Terrain Angle	The maximum tolerated terrain angle during filtering - designed to prevent artificial terrain gradient. Determined from a rough terrain model through GCPs (i.e. true terrain surface).	50 degrees	50 degrees
Maximum Building Size	The maximum distance between points used as initial ground points in the filter. Determined from the maximum distance between pre-filtered points.	5 m	12 m
Maximum Triangle Size	The largest accepted horizontal distance between filtered point and control point pairs during parameter testing. Determined from pre-measured micro-topography.	2 m	2 m
Z Tolerance	The largest accepted vertical distance between filtered point and control point pairs during parameter testing, so that small triangles do not exceed the max slope. Determined from pre-measured micro-topography.	0.15 m	0.15 m
Maximum Slope	The maximum slope between filtered point and control point pairs during parameter testing. Determined from a rough terrain model through GCPs .	45 degrees	45 degrees

Table 4.1 Parameters kept constant in the progressive densification filtering and testing procedures.

Two of the parameters required for the progressive densification filtering procedure, the iteration angle and distance, were systematically changed (figure 4.3) while maximum terrain angle, building size and triangle size (table 4.1) were kept constant. The iteration angle was examined over a range of 2 to 5° and the iteration distance over a range of 0.01 to 0.1 m. Following registration of the terrestrial LiDAR (TLS) at Site 8, optimum iteration angle and distance for filtering terrain from TLS data were selected by comparing filtered points against GCPs (table 4.2). Comparisons were based on three measures of accuracy: (1) mean offset, which describes the systematic over- or under-estimation (bias) of LiDAR terrain with respect to reference data; (2) standard deviation, which describes the variability (precision) of LiDAR terrain with respect to reference and (3) RMSE, which describes the variability of the bias. The

raw TLS training data numbered 194,103 points and the final column of table 4.2 gives the percentage of these points accepted as terrain after each iteration. The choice of optimum parameters reflects the need to balance between obtaining filtered points as close as possible to the GCPs and accepting an adequate number of points (i.e. not severely reducing resolution for a small gain in accuracy). The optimum filtering parameters are an iteration angle of 3.2 degrees and a distance of 0.05 metres (see dashed grey lines in figure 4.3).

Iteration ID	Classification Thresholds		Accuracy Statistics for all GCPs (m)			Accuracy Statistics for Feature Points Only (m)			Accepted Points (%)
	Angle (deg)	Dist (m)	Mean Offset	Std Dev	RMSE	Mean Offset	Std Dev	RMSE	
1	2	0.05	0.032	0.170	0.172	-0.095	0.186	0.206	3.57
2	4	0.05	0.044	0.178	0.182	-0.095	0.196	0.215	7.30
3	6	0.05	0.061	0.207	0.215	-0.045	0.285	0.285	10.42
4	4	0.02	0.065	0.194	0.202	-0.121	0.261	0.276	3.92
5	4	0.04	0.048	0.180	0.185	-0.095	0.197	0.215	6.97
6	4	0.06	0.035	0.205	0.207	-0.105	0.244	0.262	7.72
Optimum	3.2	0.05	0.047	0.170	0.176	-0.094	0.189	0.208	5.79

Table 4.2 A subset of the terrestrial LiDAR filtering optimisation with iterations at varying angle and distance parameters.

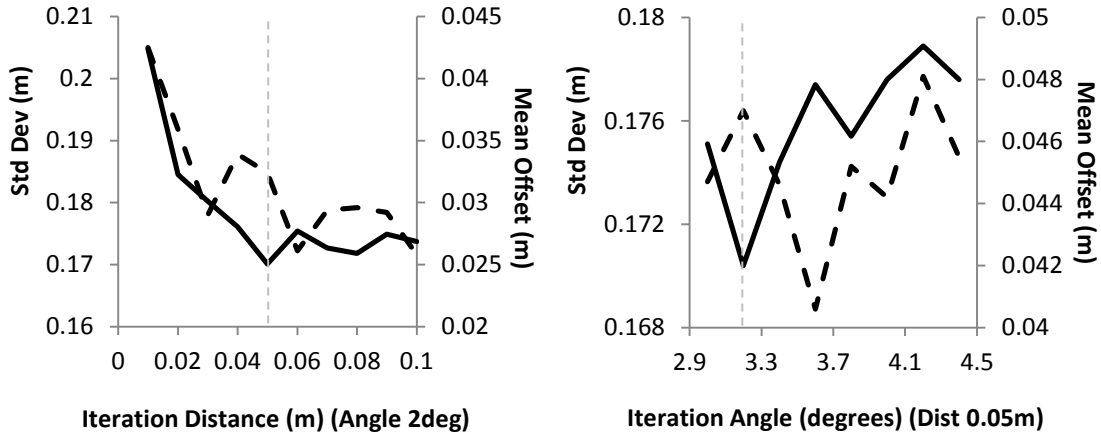
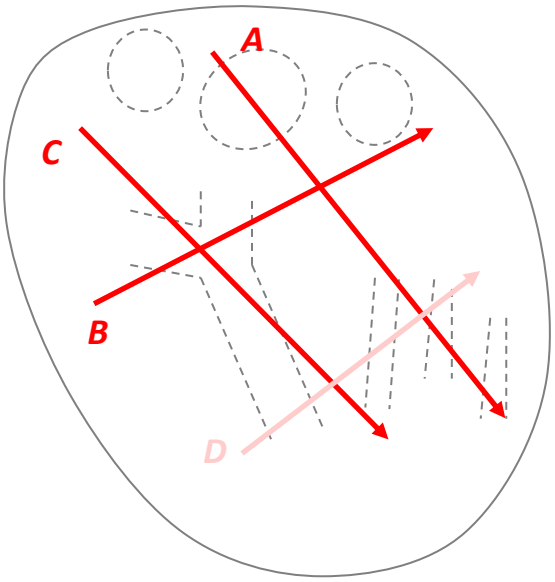


Figure 4.3 Change in accuracy statistics between terrestrial LiDAR terrain points and the full GCP set with varying filter parameters. Full line = standard deviation and dashed = mean offset.

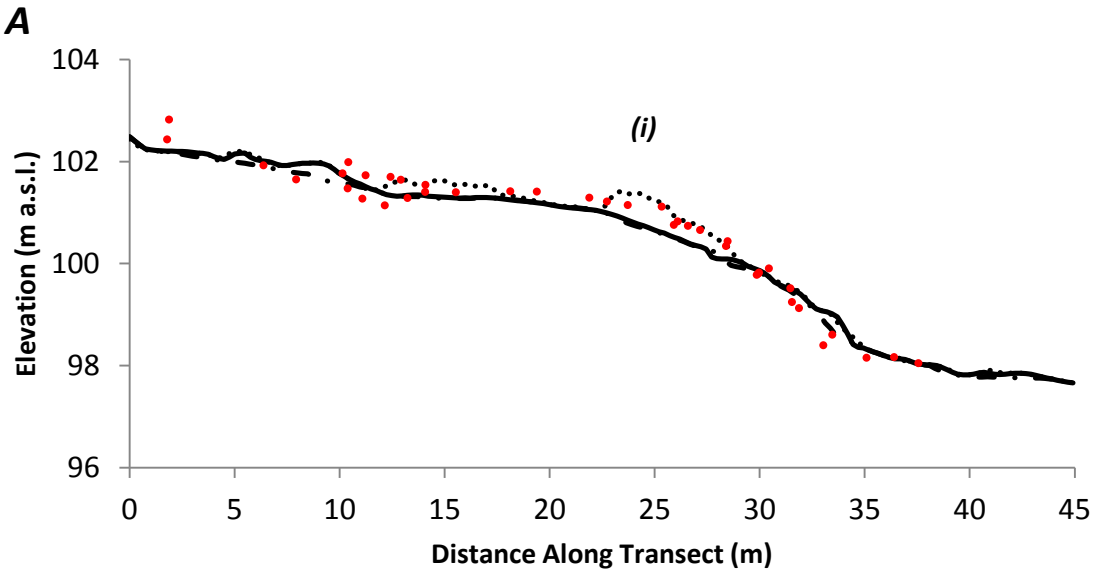


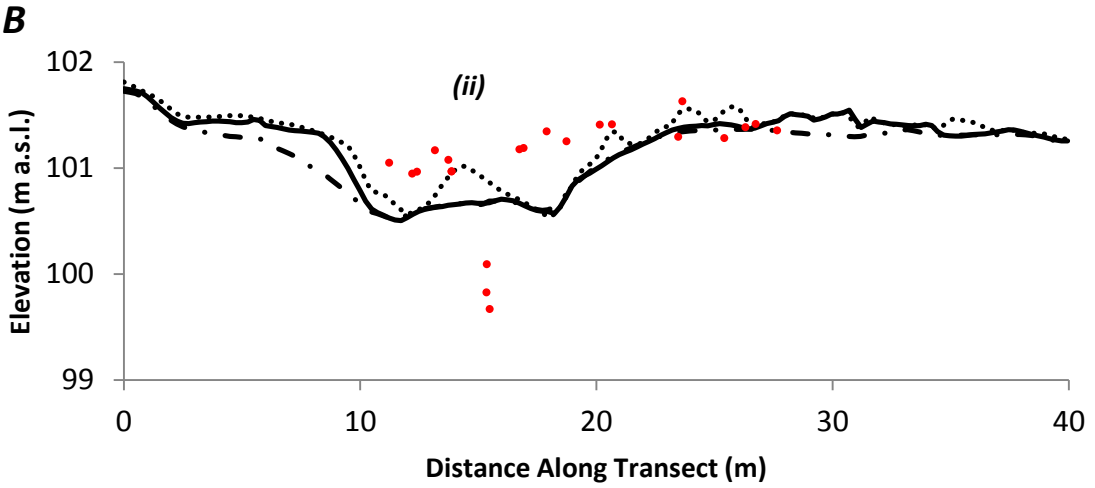
Graph Legend

- GCPs
- TLS Optimum Params: 3.2deg, 0.05m
- TLS Loose Params: 6deg, 0.5m
- TLS Tight Params: 2deg, 0.01m

Left: Outline Map of Site 8 with 4 Transects

N 0 2.5 5 10 Meters





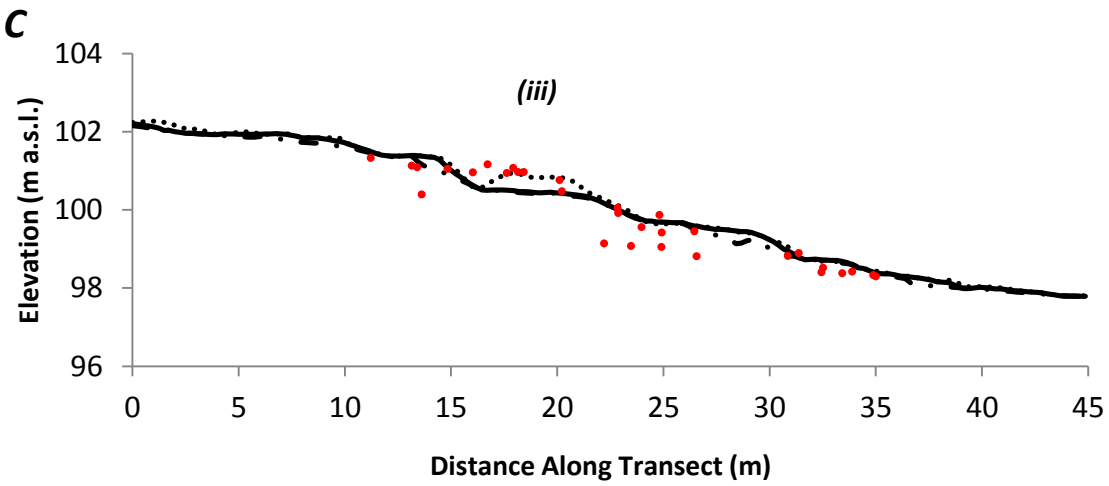


Figure 4.4 Transects A, B and C through the training data at Site 8. Red point = GCP within 1 m either side of the transect, full line = interpolated surface through TLS points filtered with optimum parameters, dotted line = surface after loose filtering and alternate line = surface after tight filtering.

All iterations in table 4.2 yield terrain points with mean offset approximately 5 cm higher than the true terrain (full GCP set). This is expected, since laser pulses will frequently intercept low vegetation and dead 'brash' material, leading to points immediately above the true terrain surface, and if the filtering routine accepts a sufficient number of these points as terrain, the LiDAR terrain overestimates true elevation. The archaeological features at Site 8, for example the building rubble and breaks of slope at the channel sides (see figure 3.8), are characteristically 'positive' with respect to the local topography, i.e. convex upwards and angular, rather than concave. The filtered LiDAR terrain is anticipated to underestimate the height of these features because the filtering tends to smooth out positive, angular discontinuities. A rectilinear pile of building rubble is 'smoothed' by 10 to 20 cm in the filtered LiDAR (at (i) in transect A in figure 4.4), but alternatively the depth of the man-made channel is underestimated by 50 cm (at (ii) in transect B). Although it could be argued that filtering with a 'looser' parameter set, i.e. using parameters that allow more points to be classified as terrain, can help to fully retain some of these features (as at (i) and (iii) in figure 4.4), other features are just as easily obscured by non-terrain points, e.g. those from low vegetation (as at (ii) in

figure 4.4) and missed as Type II filtering errors. The issue with using a 'tighter' parameter set (figure 4.4), i.e. using parameters that allow fewer points to be classified as terrain, is that the terrain point density is reduced.

However, 5 cm systematic error (mean offset) between terrestrial LiDAR and true terrain is low and the random error (standard deviation) of only 17 cm (table 4.2), in addition to visual analysis of the transects in figure 4.4, suggests that the terrestrial LiDAR terrain points filtered with optimum parameters describe the true terrain surface very closely. After filtering, over 11,200 LiDAR terrain points remain and are used as control points for calculating the optimum airborne LiDAR filtering parameters. This number of control points corresponds to a density on the ground of > 3 points per m^2, which provides a sufficiently detailed representation of the topography to validate the airborne LiDAR filtering routine.

While the procedure for classifying the airborne LiDAR was identical to that utilised in classifying the terrestrial LiDAR, one of the control parameters had to be altered. The *maximum building size* parameter limits the size of the moving widow in the initial block minimum routine (table 4.1) and while 5 m was satisfactory when filtering the terrestrial LiDAR, given the low distances between points, it had to be raised to 12 m when filtering the airborne LiDAR. This was obviously not due to buildings within Site 8 (the largest building in a point cloud characteristically defines the largest possible distance between terrain points), but to the lower resolution of the airborne LiDAR and the possibility that adjacent pulses may not penetrate the canopy within 5 metres. The value of 12 m was ascertained through preliminary investigation of all points penetrating the high canopy at Site 8. Additionally, ground control points over different surfaces, i.e. archaeological features or terrain, could not be differentiated in the airborne LiDAR optimisation like they were in the terrestrial LiDAR optimisation, so airborne LiDAR filtering parameters could only be validated against the full GCP set.

Following registration of the airborne LiDAR at Site 8, the optimum parameters for filtering terrain from ALS data were acquired by comparing filtered points against the 11,200 (TLS)

control points (reference data). Accuracy statistics for a number of example iterations and the optimum parameters are given in table 4.3, with some subjective judgement again required to ensure a balance between tight validation of ALS points to the control and accepting an adequate number of terrain points. The iteration angle was examined over a range of 7 to 13° and the iteration distance over a range of 0.1 to 0.5 m. The changes in mean offset, standard deviation and RMSE with varying iteration angle and distance parameters are illustrated in figure 4.5. Mean offset and standard deviation values are low for the optimum parameters of an 11.2 degree angle and 0.26 metre distance (see dashed grey lines in figure 4.5), with a relatively large number of terrain points accepted.

Iteration ID	Classification Thresholds		Accuracy Statistics for all GCPs (m)			Proportion of Points Accepted (%)
	Angle (deg)	Distance (m)	Mean Offset	Std Dev	RMSE	
1	9	0.5	0.144	0.204	0.249	2.61
2	11	0.5	0.149	0.197	0.247	2.98
3	13	0.5	0.157	0.206	0.259	3.33
4	11	0.1	0.126	0.265	0.293	1.47
5	11	0.2	0.139	0.204	0.246	2.47
6	11	0.3	0.146	0.195	0.243	2.87
Optimum	11.2	0.26	0.144	0.195	0.242	2.63

Table 4.3 A subset of the airborne LiDAR filtering optimisation with iterations at varying angle and distance parameters.

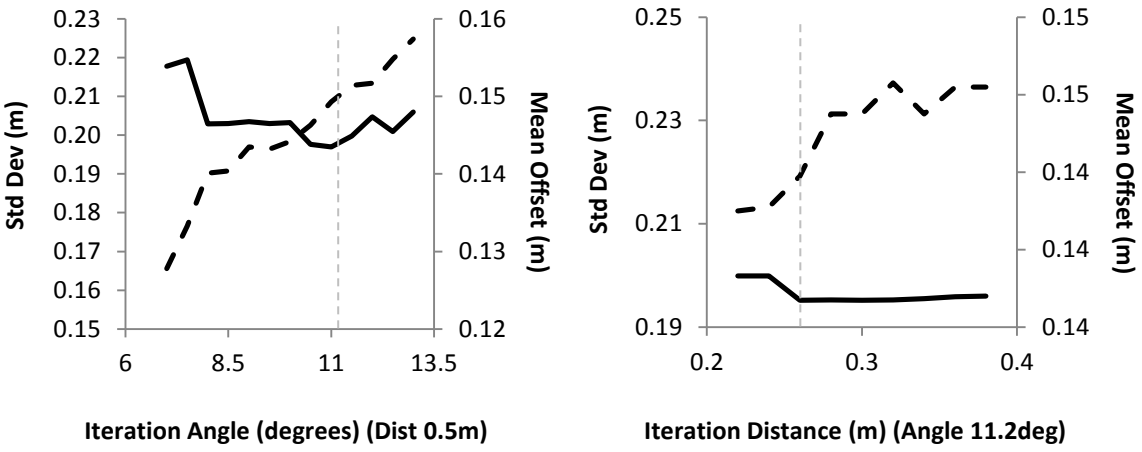
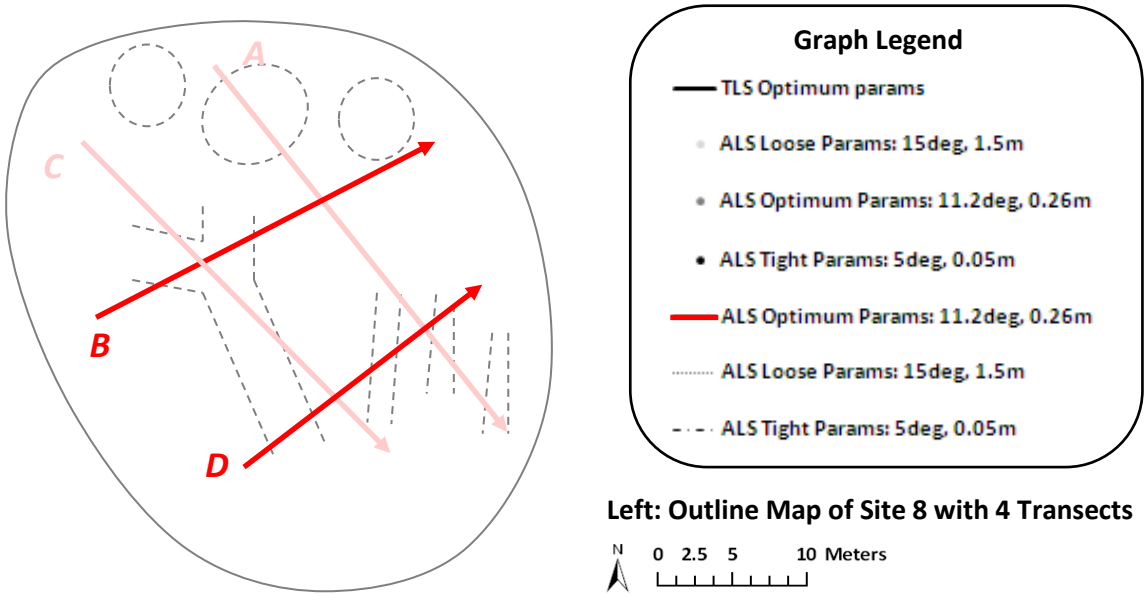
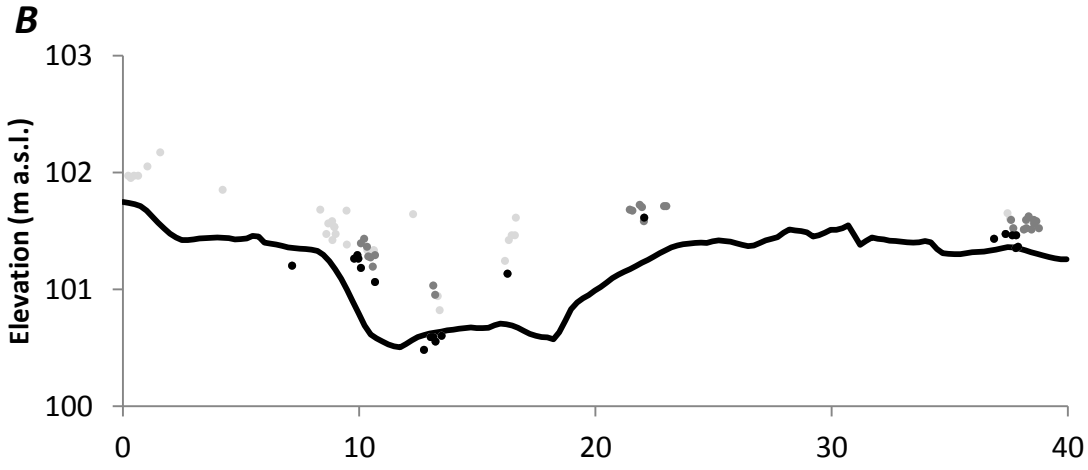


Figure 4.5 Change in accuracy statistics between airborne LiDAR terrain points and the full GCP set with varying filter parameters. Full line = standard deviation and dashed = mean offset.





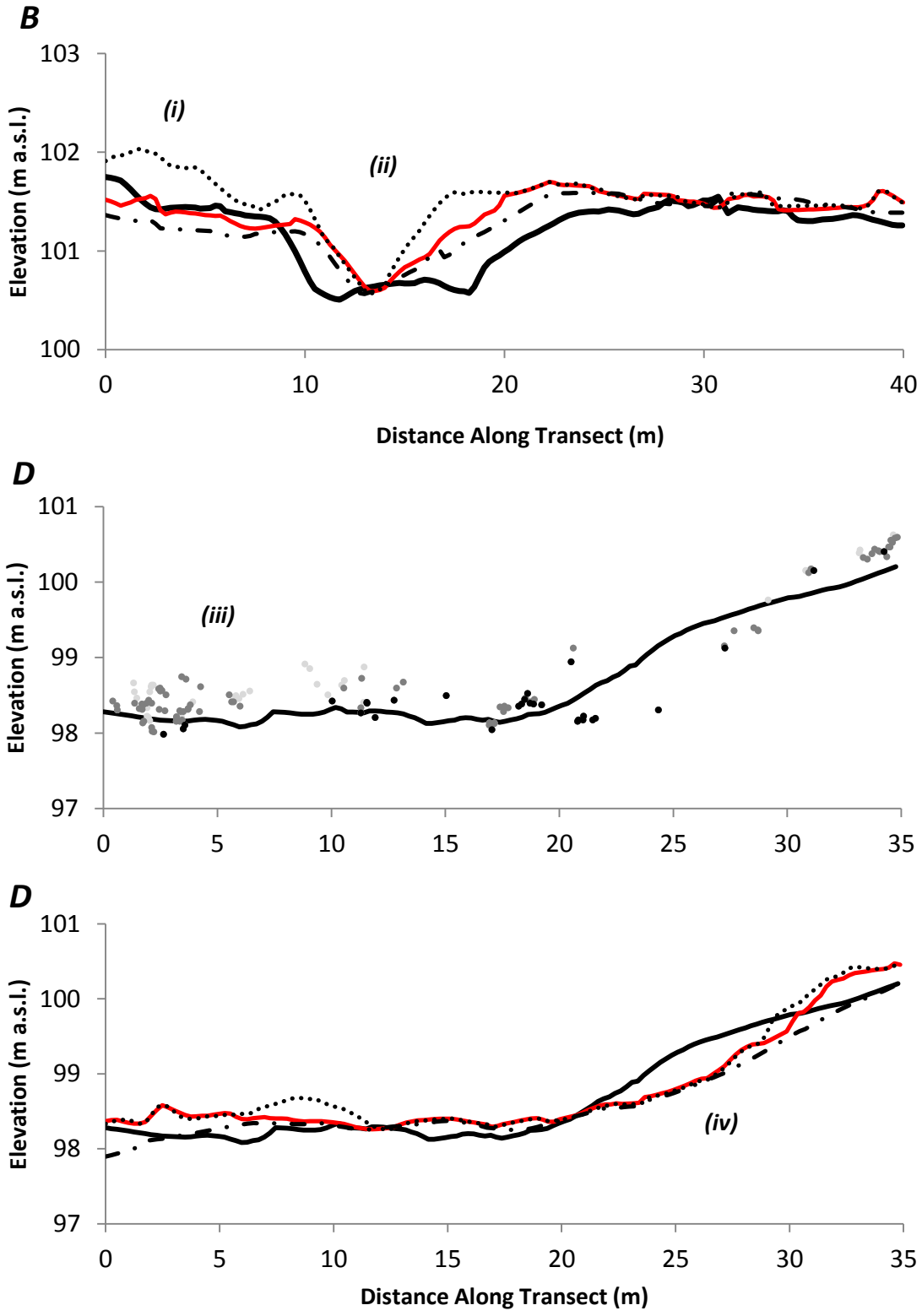


Figure 4.6 Transects B and D through the training data at Site 8. Bold line = interpolated surface through TLS points filtered with optimum terrestrial parameters (filtering stage 1), light grey point = ALS point from loose airborne filtering (filtering stage 2) within 1 m either side of the transect, grey point = ALS point from optimum filtering, black point = ALS points from tight filtering, red line = interpolated surface through ALS points filtered with optimum airborne parameters, dotted line = interpolated surface after loose filtering and alternate line = interpolated surface after tight filtering.

The airborne LiDAR systematically overestimates the true terrain (reference data) by 14 cm, with random error (standard deviation) of 20 cm (table 4.3). Systematic error is larger than the 5 cm observed at the first filtering stage between terrestrial LiDAR and GCPs, and can be explained by: (1) low vegetation having greater influence on the airborne LiDAR, owing to a higher distance between ALS terrain points and subsequently greater likelihood of Type II errors during filtering, and the larger airborne laser footprint occasionally measuring points as a mixture of terrain and scrub reflection (Baltsavias, 1999b; Doneus et al., 2008); and (2) the Optech ALTM 3100EA laser having inferior vertical accuracy at long-range compared to the terrestrial Trimble GS200: 5 to 35 cm rather than 0.1 cm (Mallet and Bretar, 2009; Optech ALTM Gemini Datasheet, 2004; Trimble GS Series Datasheet, 2005). Essentially, the 14 cm bias (and ±20 cm precision) could be explained by either, but is likely a combination of the two and cannot easily be unravelled. A lower proportion of terrain points accepted from the ALS than the TLS, 2.6 % as opposed to 5.8 %, is also expected, as a larger fraction of the airborne point cloud will include returns from the forest canopy due to scanning from a position above the wood (airborne) rather than from the ground (terrestrial). What is important is that a *relatively* large number of points are accepted (table 4.3).

The transects in figure 4.6 highlight some of the areas where airborne LiDAR points are, in general, higher than the true terrain surface, or don't absolutely locate the surface. At (iii) in transect D, many ALS points lie 10 to 20 cm above the true terrain and the optimum ALS surface (red line) runs through these points. The variability in elevation of points at this position suggests local low vegetation, and thick enough shrub/scrub that the filter with optimum parameters accepts these points as terrain, i.e. the algorithm cannot identify a clear number or distribution of points on the true terrain surface below. The full width of the man-made channel is also not successfully resolved at (ii) in transect B (figure 4.6). However, full channel depth is measured correctly and as a result, the channel can clearly be detected within the topography. It is significant that this feature can be identified, given that the site is located

in a particularly difficult environment for laser canopy penetration, and this offers promise for detecting similar features elsewhere.

Filtering LiDAR data with less severe (i.e. loose) parameters has already been mentioned as a possible method for retaining topographic features, but with the caveat of increasing Type II errors, which lead to unrealistic deviations (artefacts) in modelled terrain. These deviations are observed frequently at Site 8 when the airborne LiDAR is filtered with loose parameters and one instance is demonstrated at (i) in transect B (figure 4.6), with LiDAR terrain overestimating the reference data by up to 50 cm. At the same location, the LiDAR terrain filtered using optimum parameters is very close to the true terrain surface.

The final noticeable concern with the filtering procedure is the tendency of the progressive densification algorithm to occasionally remove convex slopes from the topography, such as the 8 m length slope at (iv) in transect D (figure 4.6). This lack of slope preservation has been observed in previous research (Huising and Gomez Pereira, 1998), but it also apparent that almost all algorithms suffer from the same problem (Sithole and Vosselman, 2004) and little research has been directed towards identifying its cause.

In summary, a two-stage filter parameter identification procedure has been conducted at Test Site 8 in Chopwell Wood, with (1) filtered terrestrial LiDAR terrain points validated against GPS and total station ground control (reference data) and (2) filtered airborne LiDAR terrain points validated against terrestrial LiDAR terrain points. Optimum parameters for filtering terrain with the progressive densification algorithm of Axelsson (1999; 2000) over complex forest canopy and sub-canopy topographic features are established as an iteration angle of 11.2 degrees and an iteration distance of 0.26 metres.

4.3 Evaluation of Interpolation Techniques for Modelling LiDAR Terrain Data

Prior to testing the accuracy of the filtered airborne LiDAR terrain data, candidate methods for deriving digital terrain models (DTMs) through terrain points were evaluated at Site 8. The following techniques were advocated in previous research: (1) raster-array cell-minimum with linear interpolation, (2) inverse distance weighted (IDW) interpolation, (3) kriging interpolation and (4) spline interpolation. All methods calculate the elevation of cells in a regular, rectangular grid and in this experiment, a cell size of 0.25 m was selected based on the point density of the filtered airborne LiDAR terrain data at Site 8 and pre-requisites for detecting topographic features of sub-metre dimensions. It was decided that the Triangular Integrated Network (TIN) technique would not be tested, given that a lack of smoothing between LiDAR points has regularly been observed in previous research to conceal otherwise obvious features in the terrain (see section 3.5.1) (Crutchley, 2008).

Each interpolation procedure required some user input in setting operational parameters (constants) and these were determined through preliminary observation of the LiDAR data, recommendations from secondary research and/or inductive reasoning. All functions searched for the closest 12 LiDAR points to a particular cell and the value of the interpolation was based on elevation as a weighted mean of these returns. This was preferred to a search radius based on a threshold distance, to reduce bias at cells where few accepted returns could be identified – a problem limited when interpolation is consistently a function of a pre-determined number of points. The kriging algorithm used a linear variogram with a sill (Oliver, 1990) and the spline, a stiff-interpolant tension method (Franke, 1982), to prohibit the generation of artificial features in the topography (artefacts). Both methods forced the surface closer to terrain points, thereby providing a realistic representation of the terrain, more faithful to the original data.

The reference data employed to test and validate the interpolation techniques were a TIN between terrestrial LiDAR points (i.e. as true a representation of the actual terrain surface as

possible) and the GPS/total station GCPs over archaeological features at Site 8 (table 4.4). An interpolated surface with lower offset and standard deviation to reference data was considered higher quality, i.e. of greater value to visualising terrain and detecting archaeological features. It is apparent that the surface interpolated through linear cell-minimum (method (1)) offers the lowest quality terrain model. Surfaces generated by methods (2) – (4) are compared visually with the reference surface along a 30 m transect through the site in figure 4.7. Table 4.4 indicates that kriging offers the closest representation of the true terrain surface, with standard deviation of only 32 and 43 cm to the terrestrial LiDAR TIN and archaeological feature GCPs, respectively. Figure 4.7 illustrates that the kriging technique slightly overestimates the concavity of the slope and the size of the slope crest along the transect; the tension spline technique appears to suffer from the same problem and the inverse distance weighted interpolation causes 'stepping', due to a problem inherent to the linear distance weighting concept.

		Surface to Terrestrial LiDAR TIN		Surface to Archaeological Features	
		Mean Offset (m)	Standard Deviation (m)	Mean Offset (m)	Standard Deviation (m)
1	**Cell-Min Linear Interpolation**	0.385	0.352	-0.105	0.519
2	**Inverse Distance Weighted**	0.353	0.372	0.069	0.450
3	**Krige with Linear Variogram**	0.333	0.324	0.098	0.426
4	**Tension Spline**	0.361	0.346	0.082	0.471

Table 4.4 Mean offset and the standard deviation of offset between airborne LiDAR digital terrain surfaces, interpolated through the best-fitting iterations of the 4 different techniques and (1) a TIN of terrestrial LiDAR terrain points and (2) GCPs over archaeological features.

Based on this evaluation and on recommendations from published research, the kriging interpolation technique with linear variogram is selected as the most favourable method for modelling digital terrain from the airborne LiDAR terrain data.

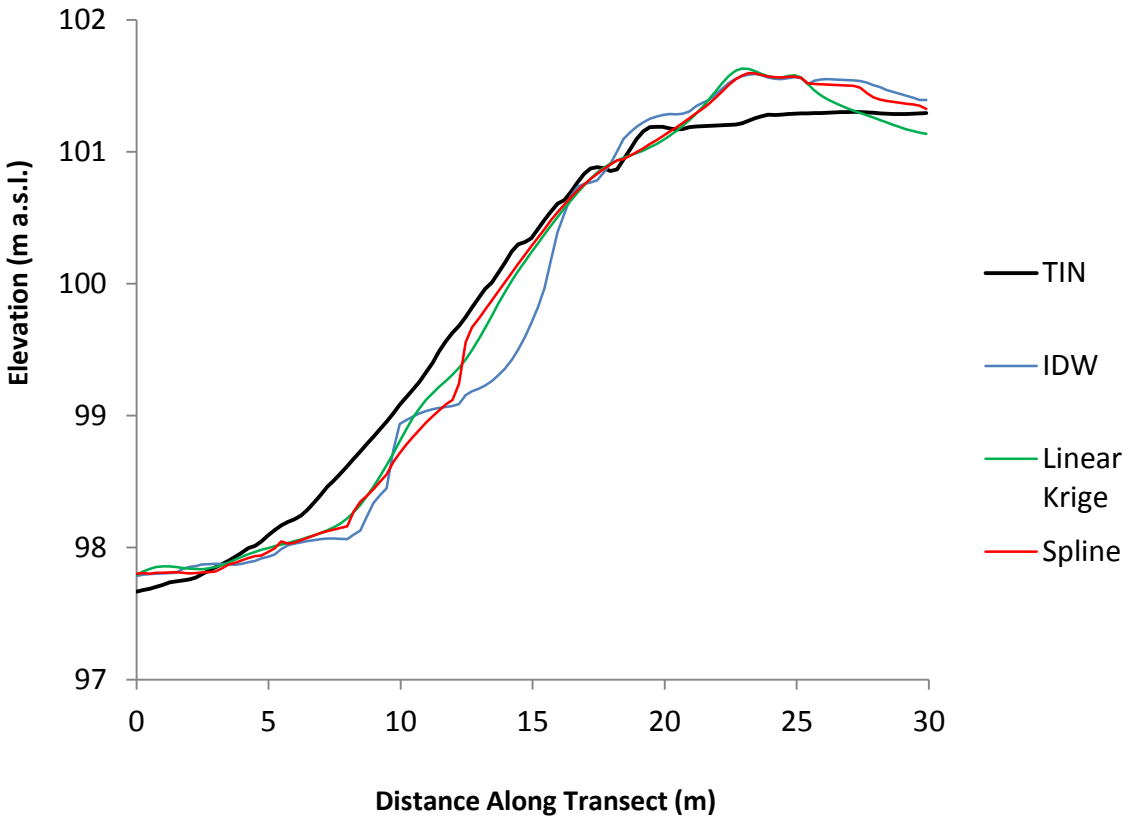


Figure 4.7 Vertically exaggerated transect of terrain surfaces interpolated through (1) inverse distance weighting, (2) kriging with linear variogram and (3) tension spline (see legend). The bold, black line illustrates the position of a TIN through reference terrain data. The transect covers 30 m from southeast to northwest through Site 8.

4.4 Accuracy and Terrain Point Density Assessment

4.4.1 Standard accuracy and terrain point density assessment

Following the identification of optimum filtering parameters for processing airborne LiDAR data with the progressive densification algorithm at Chopwell Wood, accuracy and terrain point density were assessed by comparing filtered LiDAR with cross-tabulated terrain reference data. Reference data were acquired in the field at eight, 0.12 ha test sites which comprise two sites in Corsican pine, two in Douglas fir and four in beech (see table 3.1 and figure 4.5 in section 3.2.1). Terrestrial LiDAR were registered and terrain was filtered using the optimum TLS parameters determined in section 4.2, with the terrain point density of reference data at least 5 points per m^2. It is unprecedented in research published to date to have such a high quality reference data set for assessing sub-forest canopy LiDAR terrain data.

Prior to quantitative assessment, transects through the test sites were studied to compare airborne LiDAR and reference data qualitatively. Summer photographs and transects through Sites 1, 3, 6 and 7 are illustrated in figures 4.8 to 4.11. These sites offered variety of tree species, terrain slope and roughness, and density of understorey vegetation. Photograph viewpoints and transect locations are illustrated on the circular (0.12 ha) site outline map, with each transect displaying a profile through the reference data (full line) and all airborne LiDAR points within a ±1 m (lateral distance) buffer of the transect (black points). Polynomial trends are also presented through LiDAR points at each transect to highlight over- or under-estimation with respect to the true terrain surface.

The low vegetation at Site 1, a mixture of bracken and fern, is considerably denser and consistently higher than at any other site (Photo ID1 and 2, figure 4.8), which appears to have a strong impact on the nature of the airborne LiDAR terrain points (transect 1A and 1B, figure 4.8). Elevation differences between adjacent points are erratic and points are typically above the true terrain – in places over 20 cm higher ((i) and (iv) in figure 4.8) (note the different

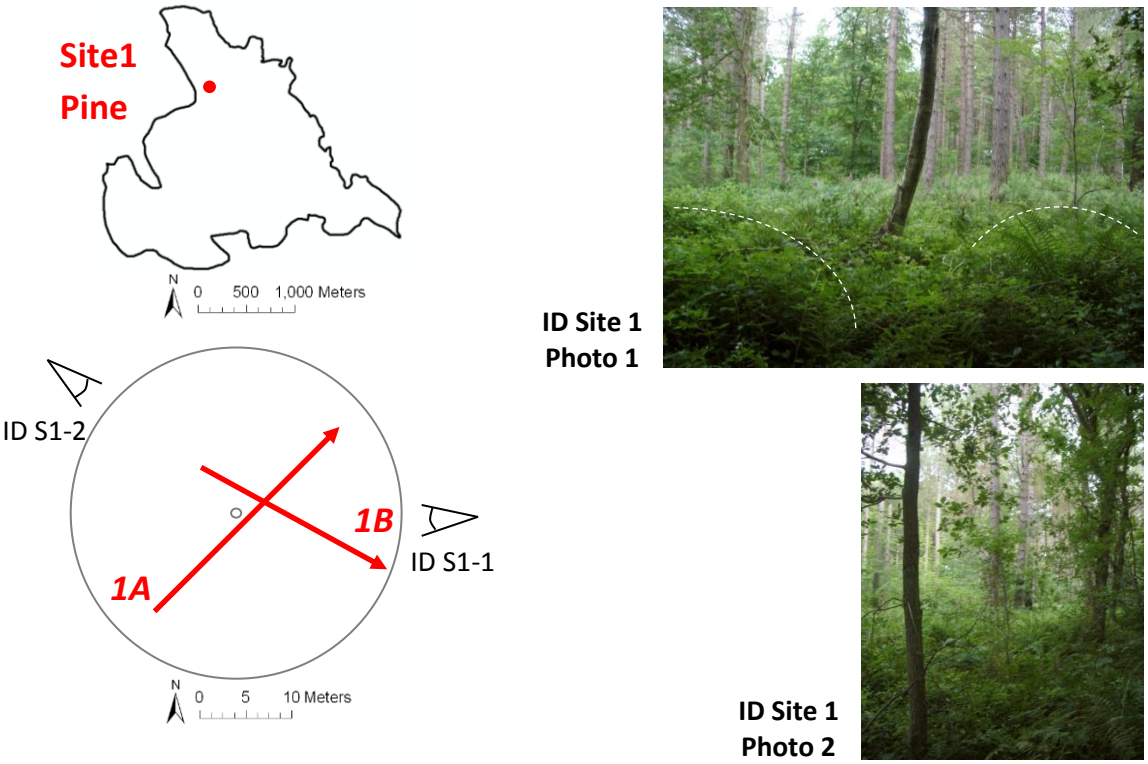
scales of the y-axes). The domed and shaped structure of the LiDAR 'terrain' at (iii) (figure 4.8), along with the variability in point elevations and approximate 50 cm overestimation compared to the reference data, suggests that laser returns over a particularly large and compact (single, isolated) plant (similar to those highlighted by the dashed white lines in photograph ID1 in figure 4.8), were mistakenly accepted in the terrain class during filtering. It is possible that the tree canopy is responsible for these effects; but while LiDAR terrain points at Site 2 (also pine) exhibit comparable, erratic overestimation relative to reference elevation, this particular scenario cannot be identified at any other site, including the other coniferous (fir) sites.

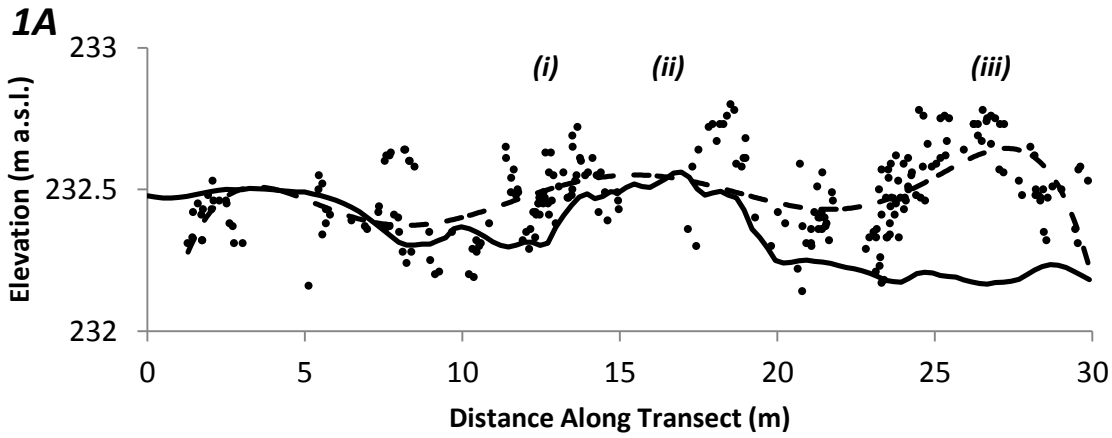
Summer tree canopy cover appeared denser in the beech at Site 3 than the pine at Sites 1 and 2, with significantly less dense understorey vegetation likely a result of lower sunlight penetration (photos in figure 4.9). A complex LiDAR scrub/terrain mixture is not observed, but surface vegetation, noticeable in Photo ID2 (figure 4.9), causes the laser scanner to typically overestimate the elevation of the true terrain. This bias is never greater than 30 cm or so and the LiDAR generally appears to correspond well with the reference data, although the LiDAR points do not follow a 30 to 40 cm deep natural drainage channel in the topography at (i) in figure 4.9.

Airborne LiDAR terrain points also appear to correspond well with reference data along transects 6A and 6B (figure 4.10) at another beech stand in a different area of the forest, with points characteristically < 10 cm from the true terrain elevation. In spite of this, some systematic under-estimation by the LiDAR of the true terrain surface is evident at (i) in transect 6A (figure 4.10). Possible explanations for this include: (1) systematic error in the LiDAR sensor, (2) a cluster of laser scattering (multi-path) errors caused by the structure of broadleaves (Ni-meister et al., 2001) or (3) mis-registration of the reference data.

Alternation between 10 to 20 cm over- and under-estimation compared to the reference data is observed under the Douglas fir at Site 7 (transect 7B, figure 4.11). This site is located on a steep, convex slope and a phenomenon associated with laser ranging over steep topography termed 'time-walk' (see section 1.2.3) has previously been recognised to disrupt terrain

filtering algorithms in this way (Baltsavias, 1999a; Huising and Gomes Pereira, 1998). However, the three scenarios mentioned in the previous paragraph are also possible explanations for the minor LiDAR deviations with respect to the true terrain elevation. There is little understorey vegetation at this site (photos in figure 4.11), so time-walk, error in the sensor and/or mis-registration of the reference data are all plausible. In any case, deviation from the true terrain surface appears very low at all sites excluding Sites 1 and 2, under the thinned pine.





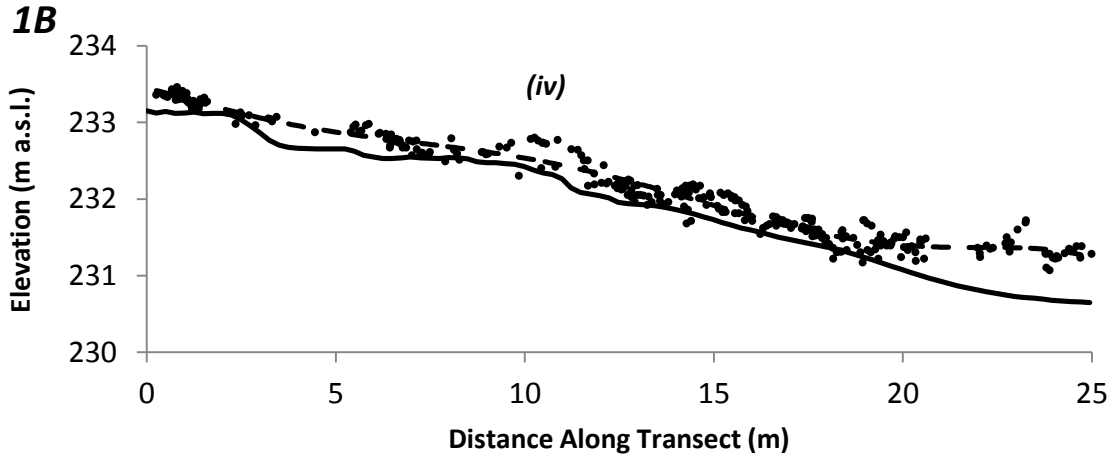


Figure 4.8 Site 1. Clockwise from top left: site location and outline map of site, site photographs, transects 1A and 1B through the site where: full line = surface through reference data, black point = airborne LiDAR and dashed line = trend through airborne LiDAR.

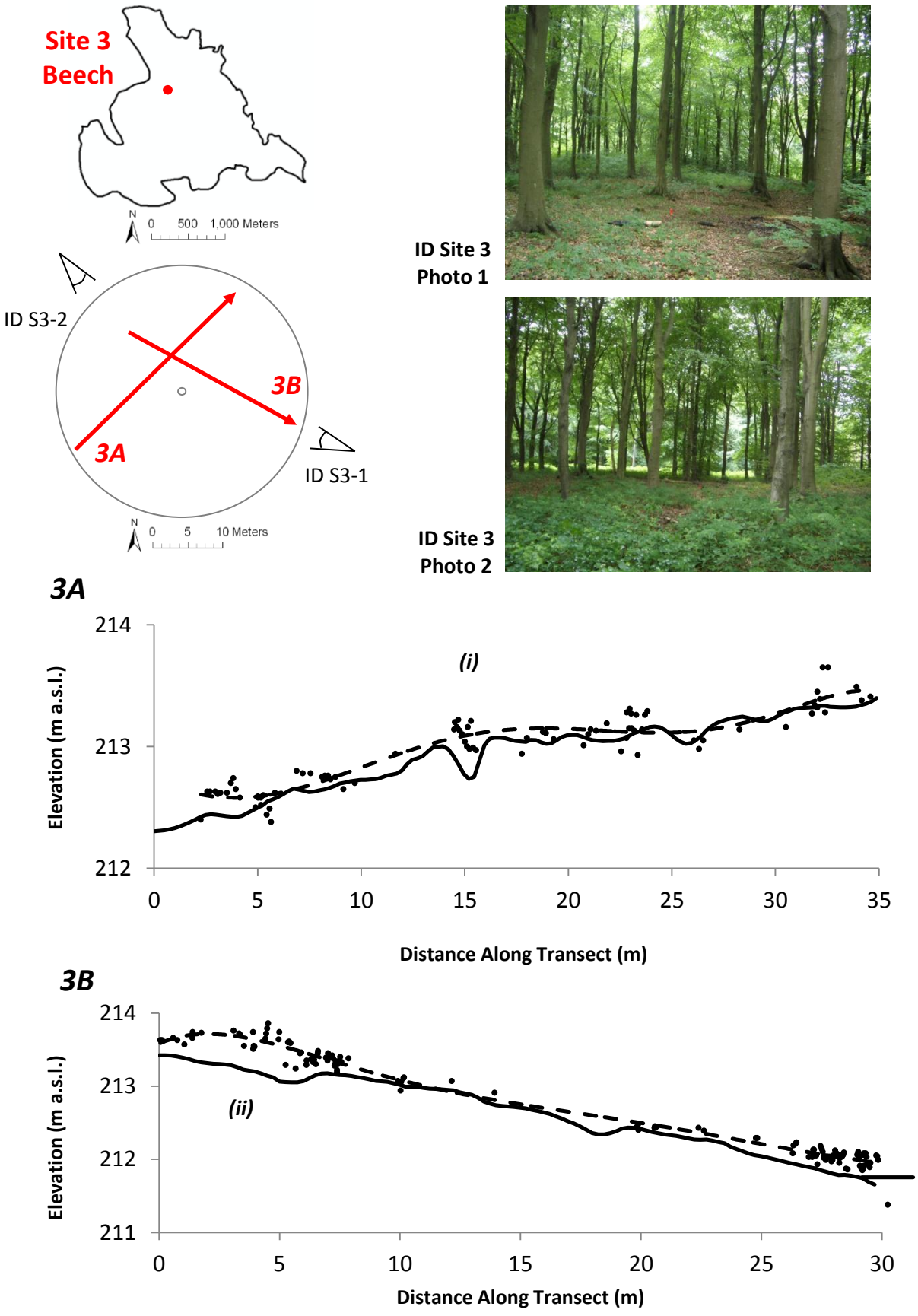


Figure 4.9 Site 3. Clockwise from top left: site location and outline map of site, site photographs, transects 3A and 3B through the site where: full line = surface through reference data, black point = airborne LiDAR and dashed line = trend through airborne LiDAR.

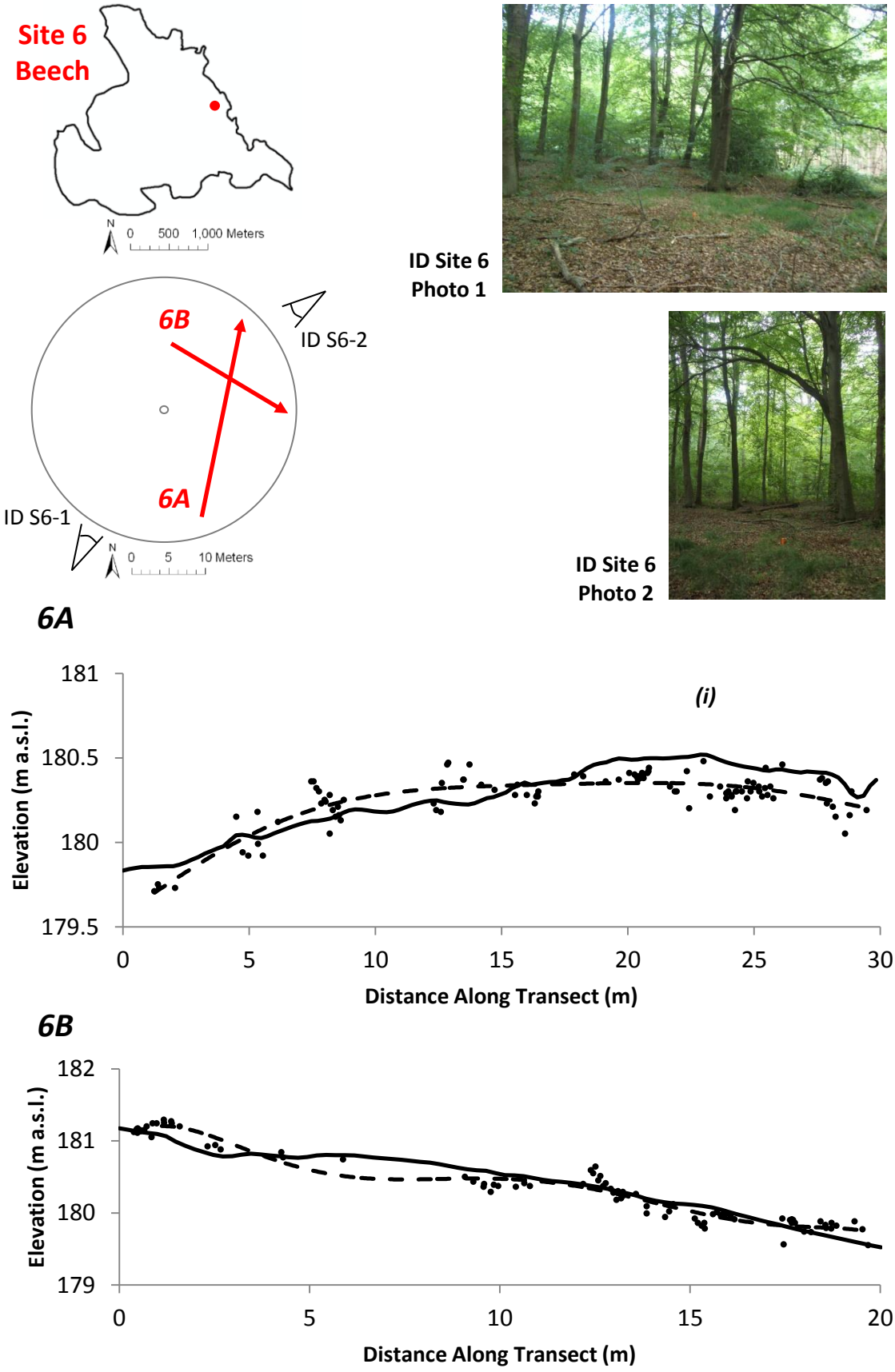


Figure 4.10 Site 6. Clockwise from top left: site location and outline map of site, site photographs, transects 6A and 6B through the site where: full line = surface through reference data, black point = airborne LiDAR and dashed line = trend through airborne LiDAR.

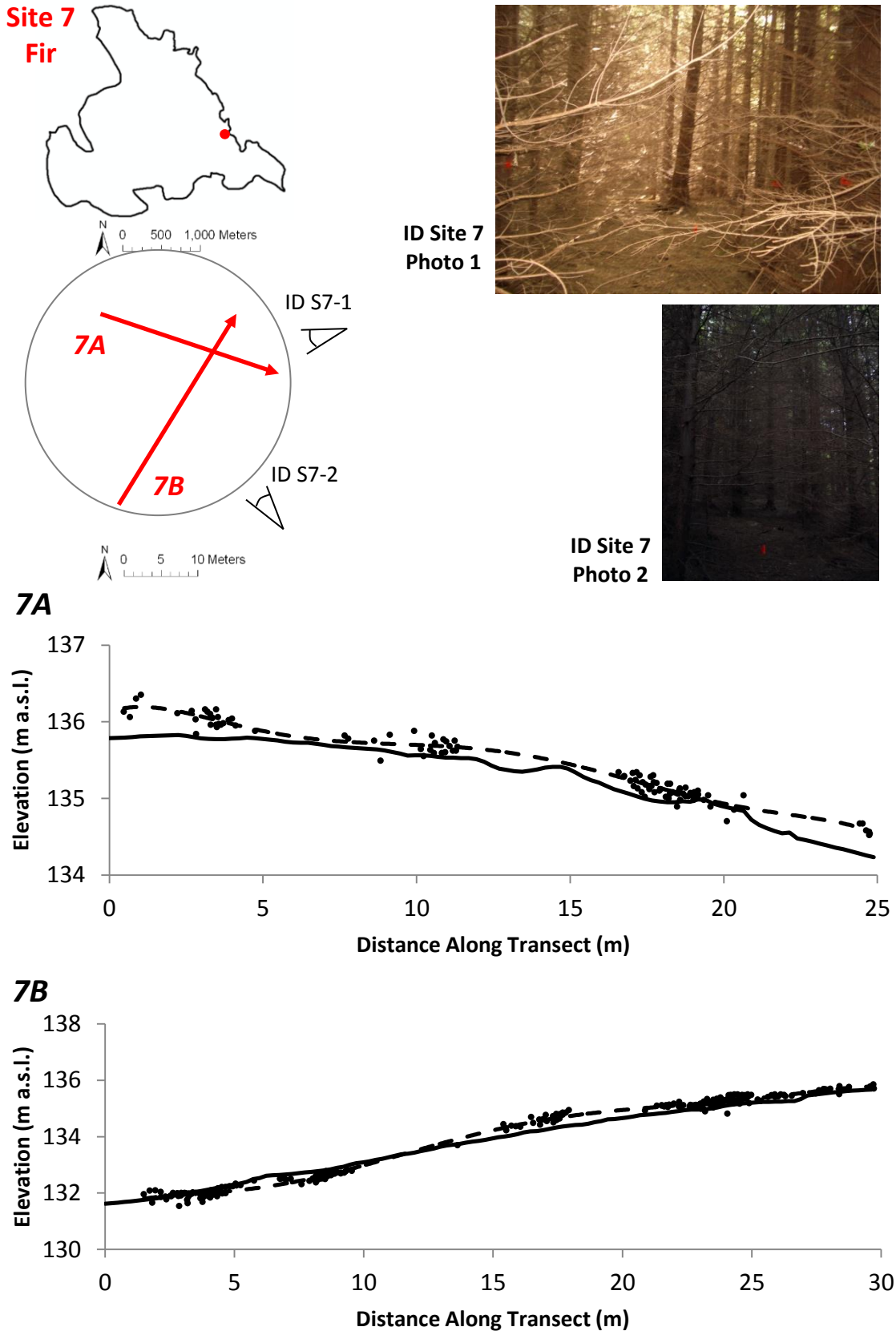


Figure 4.11 Site 7. Clockwise from top left: site location and outline map of site, site photographs, transects 7A and 7B through the site where: full line = surface through reference data, black point = airborne LiDAR and dashed line = trend through airborne LiDAR.

Quantitative assessment of airborne LiDAR terrain point accuracy is undertaken by evaluating airborne LiDAR terrain points and modelled surfaces against respective reference terrain points and surfaces. Reference data comprise filtered terrestrial LiDAR point clouds, acquired at each of the eight test sites at Chopwell Wood. The surface-to-surface assessment involves comparison of 0.25 m spatial resolution LiDAR and reference terrain models, both interpolated by the kriging technique (see section 4.3). This resolution is based on the minimum airborne LiDAR terrain point density, which is 3 points per m^2 under the Douglas fir at Site 8. Point-to-point and surface-to-surface assessments provide four significant statistical measures: (1) mean offset and (2) mean absolute offset – measures of bias, (3) standard deviation of offset – a measure of precision and (4) root mean square error (RMSE) – a measure of the variability of the height estimate (see section 3.5.2.1 for more information). Measures (1), (3) and (4) describe systematic error in the LiDAR terrain data with respect to the true terrain elevation, while (2) describes random error.

The results of the site-by-site statistical assessment are given in table 4.5. Mean offset and absolute offset are highest at Site 2 (54 and 46 cm, respectively) and lowest at Site 6 (-3 and 11 cm, respectively), while standard deviation is highest at Site 5 (27 cm) and lowest at Sites 3 (12 cm) and 6 (10 cm). Maximum absolute offset, which is a crude indication of the largest non-terrain artefact (i.e. a product of filtering and/or sensor errors) that could incorrectly be interpreted as an archaeological feature, is recorded as 141 cm at Site 7. LiDAR terrain point density is highest at Sites 1 and 2 (43 and 39 points per m^2, respectively) and lowest at Sites 6 and 8 (7 and 3 points per m^2). Sites 1 and 2, in the thinned Corsican pine, offer approximately 40 'terrain' returns per m^2, indicating canopy laser penetration of up to almost 100 %.

Site	Airborne LiDAR Accuracy Statistics (cm)				No. of LiDAR Points in Site	LiDAR Point Density (points per m^2)
	Mean Offset (1)	Standard Deviation (2)	RMSE (3)	Mean Absolute Offset (4)		
1	21	20	29	27	5440	43
2	54	21	57	46	4948	39
3	9	12	15	12	1294	10
4	39	18	43	30	1020	8
5	6	27	28	41	1338	11
6	-3	10	11	13	901	7
7	3	16	16	25	2411	19
8	15	15	21	21	401	3

Table 4.5 Statistical assessment between airborne LiDAR and reference terrain data at the eight test sites. Individual statistical measures are referred to in the text. Positive standard offset indicates that LiDAR over-estimating the elevation relative to the reference terrain data. Measures (1) to (3) are determined through point-to-point assessment, while (4) is determined though surface-to-surface assessment.

Species	Airborne LiDAR Accuracy Statistics (cm)				No. LiDAR/ Reference Point Pairs Compared	LiDAR Point Density (points per m^2)
	Standard Offset	Standard Deviation	RMSE	Mean Absolute Offset		
Pine	38	21	43	37	10,388	41
Douglas fir	9	16	19	23	2,812	11
Beech	13	17	24	24	4,553	9
Coniferous	23	18	31	30	13,200	26
Deciduous	13	17	24	24	4,553	9

Table 4.6 Statistical assessment between airborne LiDAR and reference terrain data, cross-tabulated by canopy species. Refer to the caption of table 4.5 for more information on statistical measures.

LiDAR accuracy statistics for the eight test sites are cross-tabulated by canopy species in table 4.6. The pine yields particularly poor results, for reasons already mentioned. However, for Douglas fir, the LiDAR over-estimates true terrain elevation by only 9 cm, four times less than the pine, and is 5 cm more precise. As suggested in the qualitative assessment, this can be at least partly attributed to a lack of understorey vegetation cover (photographs in figure 4.11). Under the beech canopy, LiDAR terrain accuracy is similar to the fir canopy, and when comparing deciduous against coniferous sites, the deciduous (beech) sub-canopy terrain offers almost exactly the same precision as the coniferous (pine and fir) terrain, but over-estimates the true terrain by a significantly lower quantity (13 cm as opposed to 23 cm). The pine is considerably less amenable to summer LiDAR terrain survey than either the fir or beech, but this cross-section of tree species and other site attributes, e.g. understorey vegetation, is representative of the entire wood and, as such, deciduous forest appears, in general, more amenable to LiDAR terrain survey than coniferous forest.

With a view to assessing potential cause and effect relationships between LiDAR accuracy statistics, LiDAR terrain point density and various site attributes, the results from table 4.5 are evaluated against information on the tree canopy, understorey vegetation and topography. All site attribute data were acquired in the field, apart from the data for topographic roughness and gradient and stand planting year, which were obtained from the Forestry Commission UK GIS (1998). The entire data set, referenced by site, is presented in table 4.7. Multiple regression analysis is performed among variables to ascertain statistically significant relationships and a number of the most interesting relationships are investigated further. The Pearson's Product Moment Correlation Coefficient (PMCC) *r* is employed to evaluate the strength of linear least squares relationships between variables, and provides values between -1 and 1 depending on the direction (positive equals positive direction) and strength (one equals perfect correlation, zero equals no correlation) of a relationship. The correlation matrix between variables is presented in table 4.8.

Site		1	2	3	4	5	6	7	8	Mean
LiDAR Accuracy Statistics (cm)	Std Offset	21	54	9	39	6	-3	3	15	18
	Std Dev	20	21	12	18	27	10	16	15	18
	RMSE	29	57	15	43	28	11	16	21	27
	Abslt Offset	27	46	12	30	41	13	25	21	27
LiDAR Point Density (pm^{-2})		43	39	10	8	11	7	19	3	18
Primary Tree Species		CP	CP	BE	BE	BE	BE	DF	DF	-
Dominant Secondary Tree Species		SY	-	-	-	HO/ OK	MA/ EL	BI	BE/ BI	-
Planting Year		1947	1947	1934	1934	1923	1923	1987	1991	-
Tree Survey Stats (0.12ha plot)	Tree Count	45	36	61	36	35	55	175	252	87
	Primary vs. Secondary (%)	58	78	95	100	80	45	98	85	80
	Conifer vs. Deciduous (%)	58	78	0	0	3	5	98	85	41
	Mean Stem Height (m)	17.5	20.0	22.0	19.0	14.0	16.5	18.5	11.5	17.5
	Mean Canopy Depth (m)	9.5	10.0	16.5	13.0	10.5	11.0	7.0	5.5	10.5
	Mean DBH (m)	0.31	0.35	0.32	0.30	0.31	0.27	0.24	0.14	0.28
	Mean DBH of Primary Species (m)	0.44	0.41	0.32	0.30	0.34	0.40	0.24	0.14	0.32
	Stem Density (basal area per ha)	0.38	0.33	0.41	0.23	0.29	0.32	0.73	0.36	0.38
	Stem Density of Primary Species (basal area per ha)	0.35	0.32	0.40	0.23	0.27	0.27	0.73	0.30	0.36
PMCC Stem Height vs. DBH		0.95	0.89	0.59	0.84	0.93	0.87	0.79	0.53	0.80
Forestry Commission Terrain Ratings (1-5)	Roughness	2	2	2	2	2	2	2	3	2
	Gradient	1	1	2	2	1	1	3	5	2
Understorey Vegetation	Density (%)	95	95	35	60	40	25	0	5	45
	Mean Height (m)	1.2	1.1	1.0	1.0	0.9	0.7	0	0.3	0.8
	Volume (m^3)	1400	1200	400	800	400	200	0	20	400

Table 4.7 Airborne LiDAR and site attribute information for the eight test sites. Tree species codes are: CP = Corsican pine, BE = Beech and DF = Douglas fir.

		LiDAR				Tree						Topo		Low Veg
		Off	Std	RMS	PD	Age	Cou	SH	CD	DBH	SD	Rou	Gra	Vol
LiDAR	Offset	-	-	-	-	-	-	-	-	-	-	-	-	-
	Std Dev	0.36	-	-	-	-	-	-	-	-	-	-	-	-
	RMSE	0.96	0.58	-	-	-	-	-	-	-	-	-	-	-
	Point Density	0.50	0.39	0.52	-	-	-	-	-	-	-	-	-	-
Tree	Age	0.03	0.15	0.14	-0.02	-	-	-	-	-	-	-	-	-
	Count	-0.30	-0.31	-0.41	-0.35	-	-	-	-	-	-	-	-	-
	Stem Height	0.32	-0.22	0.23	0.36	0.31	-	-	-	-	-	-	-	-
	Canopy Depth	0.07	-0.17	0.03	-0.12	0.76	-	-	-	-	-	-	-	-
	DBH	0.40 0.22	0.36 0.21	0.48 0.31	0.52 0.64	0.74 0.72	-	-	-	-	-	-	-	-
	Stem Density	-0.39 -0.32	-0.23 -0.17	-0.43 -0.35	0.13 0.16	-0.64 -0.59	-	-	-	-	-	-	-	-
Topo	Roughness	-0.06	-0.18	-0.17	-0.38	-	-	-	-	-	-	-	-	-
	Gradient	-0.16	-0.32	-0.30	-0.48	-	-	-	-	-	-	-	-	-
Summer Understorey Volume		0.71	0.44	-0.18	0.81	0.36	-0.65	0.40	0.24	0.69	-0.39	-	-	-

Table 4.8 Correlation matrix of PMCC (r) values between LiDAR accuracy statistics, LiDAR point density and test site attributes. Highlighted in red are correlation coefficients above 0.4, indicating a strong relationship. When there clearly cannot be a relationship between two variables, e.g. tree age and topographic roughness, correlation is not performed. Two coefficients are given for DBH and stem density: (1) for the entire plot, on the top row, and (2) only primary species, on the bottom row. Abbreviations are: Off = mean offset, Std = standard deviation, RMS = root mean square error, PD = point density, Cou = tree count, SH = stem height, CD = canopy depth, DBH = diameter at breast height, SD = stem density, Rou = terrain roughness and Gra = terrain gradient.

Caution must be used when attempting to interpret these relationships, because: (1) the independent and dependent variables are not identified and, as such, cause and effect is only recognised through logical reasoning; (2) a strong relationship may arise indirectly if many variables are dependent, i.e. one affects another which in turn affects another so the first and last correlate, and (3) each least squares regression considers only eight values (eight test sites). The latter argument is reasonably weak, given that each value is a product of a data set rigorously sampled over a large forest plot. For example, the mean DBH value at Site 8 is calculated from 252 individual trees (table 4.7).

It has been suggested that tree structure is highly influential over the rate of laser pulse penetration through forest canopy and accordingly the LiDAR terrain point density (Hodgson et al., 2005; Raber et al., 2007). In this study, tree structural properties are estimated by assuming allometric association between forest variables, i.e. that stem diameter, stem height and canopy depth are proxies for canopy density and diameter (Zianis, 2005). Some criteria for the existence of such a relationship between stem height and DBH are given in table 4.7 at with an average PMCC *r* of 0.8 across all plots.

Least squares regression analyses between LiDAR terrain point density and four forest structure variables (stem density, stem density of primary species only, mean stem height and mean canopy depth) are illustrated in figure 4.12. Point density is positively related to tree stem diameter (DBH) (table 4.8) with a high *r* (0.52 and 0.64, for full and only primary species, respectively). Correlations between stem density, stem height and point density are positive but weak ($r < 0.4$) (table 4.8; figure 4.12). Increasing canopy cover and density are assumed to reduce laser penetration rate, and these positive relationships are inconsistent with that assumption. The following explanations are offered: (1) stem diameter, height and density are not appropriate proxies for stand-scale canopy structure; (2) allometric assumptions are not valid for all tree species; (3) an insufficient number of plots are analysed to recognize the true relationships and/or (4) the relationships are non-linear. Whilst canopy depth exhibits negative correlation with terrain point density (figure 4.12 B), in line with the original theory on laser

penetration rate, the relationship is weak (*r* of -0.12). However, tree count and terrain point density also exhibit negative correlation, and although an *r* value of -0.35 is relatively weak in comparison to, for example, the correlation between terrain point density and DBH, it is strong enough to suggest that a relationship could exist. It is assumed that a higher tree count (with individual trees closer to each other) is likely to reduce laser penetration and consequently lower terrain point density, but, based on the weak correlation obtained in this study, more intensive investigation, for example using a dataset with a much larger number of forest plots, may be required to determine a strong relationship between these variables.

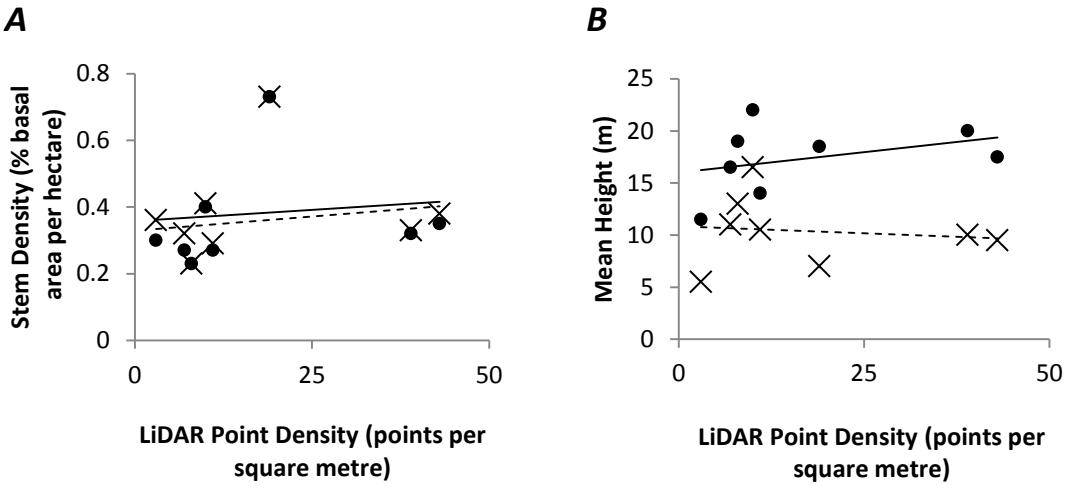


Figure 4.12 (A) Tree stem density against LiDAR terrain point density, where black point (full trend line) = primary species only, cross (hashed trend line) = all trees in site. **(B)** Mean height against LiDAR terrain point density, where black point (full trend line) = tree top height and cross (hashed trend line) = tree canopy height.

The influence of specific tree species on point density will be covered in greater detail in the full study area assessment (section 4.4.4). However, it is clear from the results in table 4.5 that the laser penetrated the pine canopy far more easily than either the beech or fir. It was initially apparent, both from studying Sites 1 and 2 in the field and the tree count at each of these sites (table 4.7), that the Corsican pine stands had been significantly thinned. This clearly contributes to the laser penetration rate, but it has also been suggested that needle foliage is

less obstructive to the laser than broad-shaped and horizontally oriented leaves of deciduous trees, which tend to absorb or heavily attenuate the pulse energy (Su and Bork, 2006). It was observed at Site 8 that secondary species amongst the fir (typically beech and birch), although constituting only 15 % of stems, provided significant canopy cover, and this, to some extent, explains the considerably lower terrain point density at Site 8 than Site 7, of 3 and 19 points per m^2, respectively.

Figure 4.14 A illustrates the relationship between the LiDAR accuracy and understorey vegetation. Understorey vegetation is given as the volume of influential material covering the sub-canopy terrain surface, e.g. percentage cover * 0.12 ha plot (converted to square metres) * mean height. Based on preliminary investigation, including the qualitative assessment (e.g. the photographs in figure 4.8), and published research (e.g. Doneus et al., 2008) it is assumed that a thicker and greater coverage of low vegetation will: (1) cause the airborne laser scanner to over-estimate the elevation of the terrain, due to Type II filtering errors, and (2) reduce the precision of LiDAR terrain measurements, due primarily to deterioration of the laser pulse following interaction with vegetation.

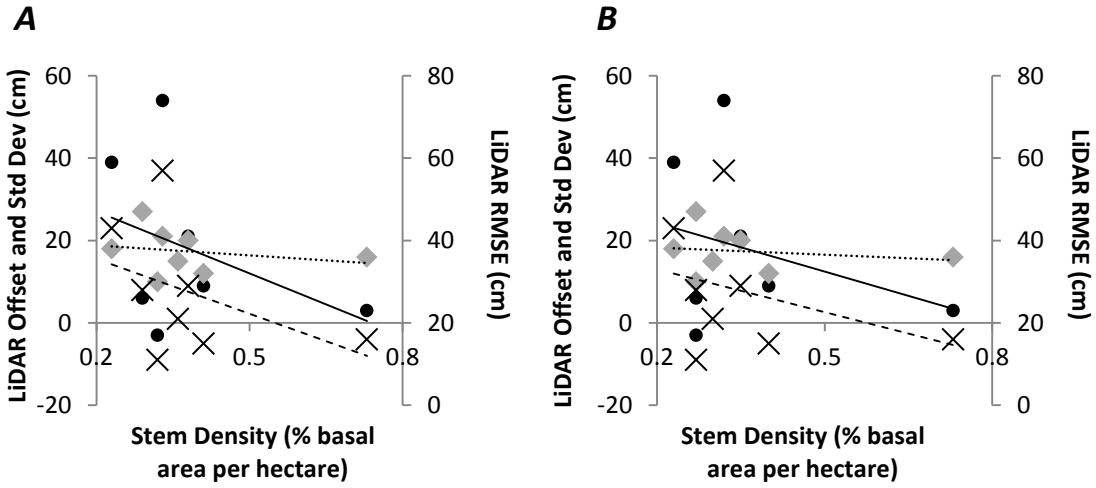


Figure 4.13 Tree stem density against LiDAR terrain accuracy statistics, where black points (solid trend line) = mean offset, grey diamonds (dotted grey trend line) = standard deviation and crosses (dashed trend line) = RMSE. **(A)** Stem density of all species in stand. **(B)** Stem density of primary species only.

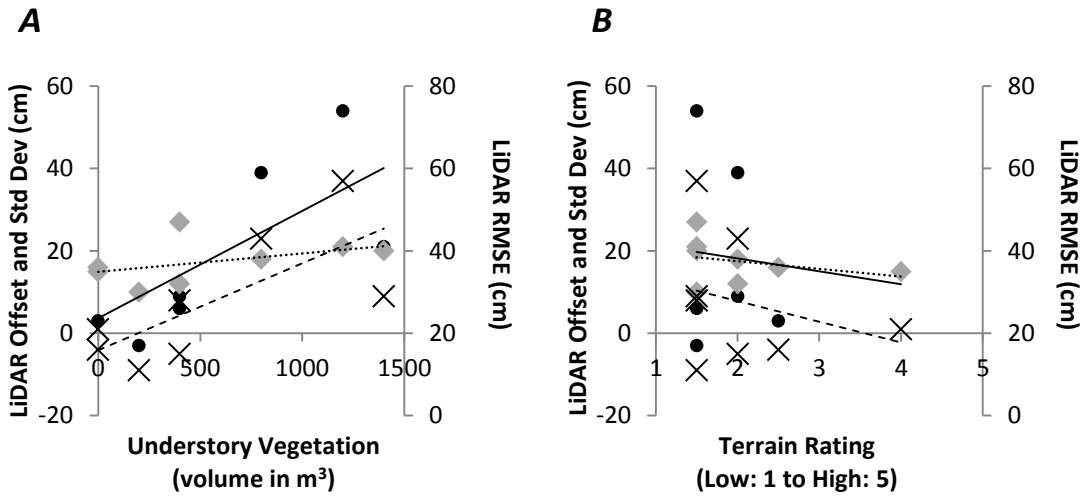


Figure 4.14 LiDAR terrain accuracy statistics, where black points (solid trend line) = mean offset, grey diamonds (dotted grey trend line) = standard deviation and crosses (dashed trend line) = RMSE; against **(A)** understorey vegetation volume and **(B)** combined (average of) Forestry Commission roughness and gradient terrain ratings (recorded in 1998).

Both assumptions are validated by strong, positive correlations between understorey volume and mean offset (*r* of 0.71), RMSE (0.72) and standard deviation (0.44) (table 4.8 and figure 4.14 A). The relative strength of these relationships, in comparison to those associated with the tree canopy, suggests that understorey vegetation plays a crucial role in controlling the precision and especially over-estimation of LiDAR terrain data. In identifying a deterministic relationship such as this within the complex matrix of variables influencing the travel of a laser pulse (see section 1.2.3), considerable caution must be taken. Previous concerns expressed over the quantity of data contributing to the least squares analysis, for example that the number of test sites is insufficient, still apply; but it is clear that understorey vegetation is more influential on sub-forest canopy LiDAR terrain measurement than previously thought.

Besides tree canopy and understorey vegetation, terrain roughness and gradient are believed to influence LiDAR terrain point accuracy. Steep slopes have, in previous studies, caused systematic error associated with 'time-walk' (see section 1.2.3) (Baltsavias, 1999a) and convex slopes have caused problems for both the progressive densification filtering algorithm and terrain modelling procedures (Pfeifer and Mandlburger, 2009), while Huising and Gomes

Pereira (1998) have observed that slope crests can be detached during filtering. The relationships between Forestry Commission terrain ratings and LiDAR bias and accuracy are illustrated in figure 4.14 B. The rating in figure 4.14 B, from 1 to 5, is an average of roughness and gradient and it is expected, based on the research cited above, that a low rating will correlate with higher quality LiDAR terrain measurements. It appears that, contrary to expectation, terrain of greater roughness and gradient provides LiDAR points of higher precision and lower bias (figure 4.14 B). However, the relationships between terrain gradient and especially roughness and LiDAR mean offset, RMSE and standard deviation are weak (table 4.8), typically with $r < 0.4$. Weak statistical correlation implies that variation in topography has little or no control over the quality of sub-forest canopy LiDAR terrain measurements, but, given that the Forestry Commission GIS terrain ratings are a subjective and crude (stand-wide) measure, the relationship between terrain slope, roughness and LiDAR terrain point accuracy will be investigated in greater detail in section 4.4.3.

In summary, the most significant outcome of this section of the accuracy assessment is that, under thinned coniferous forest, which might previously have been considered relatively amenable to summer (leaf-on) LiDAR terrain survey and despite regular laser penetration and subsequently high terrain point density, dense understorey vegetation caused numerous Type II filtering errors. This in turn led to reduced terrain accuracy and significant over-estimation of the true terrain elevation.

4.4.2 Within-site terrain point clustering assessment

Based on field observation, stand-scale canopy and understorey often appeared regular and clustered, suggesting that selective laser penetration around non-terrain objects could potentially lead to patterns in terrain points. This is particularly significant for sub-canopy archaeological feature prospection capabilities, because a feature could be hidden in a gap between intermittent 'patches' of dense terrain points, for example if it is located directly beneath a tree crown. A digital terrain model (DTM) specified as a product of 5 measurement points per m^2 may actually describe a surface with patches of 20 points per m^2 but also considerable gaps without points. To evaluate the influence of forest canopy and understorey vegetation on the distribution of LiDAR terrain returns, two statistical clustering analyses were conducted.

The two statistical tests are Moran's I three-dimensional autocorrelation and the Getis-Ord General G clustering analysis, with the results for each across the eight Chopwell Wood test sites given in table 4.9 and cross-tabulated against understorey vegetation density, the stem density of primary tree species, mean canopy depth and the tree count. Moran's I autocorrelation evaluates whether points are clustered, dispersed or random based on XYZ co-ordinates. Points are evaluated against neighbours; so, to rule out the possibility of mistakenly identifying patterns in random data, a fixed sphere of 2 m was established for point-to-point comparison. If the confidence (p-value) is small and the absolute value of the Z score is large enough that it falls outside of the desired confidence level, a pattern between points is likely at that confidence level. An index above 1 indicates increased clustering, near 0 a random distribution and below 1 increased dispersion between points (i.e. points are increasingly isolated).

Site		1	2	3	4	5	6	7	8
Moran's I Three-dimensional Auto-correlation	Index (p-value)	0.78	0.83	0.98	0.68	0.53	0.69	1.12	0.74
	Z Score (Std Deviations)	371	346	151	99	91	66	274	48
	Significance Level	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	% Likelihood Random Chance	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Getis-Ord General G Clustering	Index (p-value)	0.01	0.01	0.03	0.04	0.03	0.02	0.02	0.05
	Z Score (Std Deviations)	-2.24	15.4	-9.67	-3.53	-0.54	-7.58	-3.65	-4.60
	Clustering	Mod	Very	Very	Very	None	Very	Very	Very
	Significance Level	0.05	0.01	0.01	0.01	-	0.01	0.01	0.01
	% Likelihood Random Chance	< 5	< 1	< 1	< 1	-	< 1	< 1	< 1
Tree Species		CP	CP	BE	BE	BE	BE	DF	DF
Understorey Vegetation Density (%)		95	95	35	60	40	25	0	5
Stem Density of Primary Species (basal area per ha)		0.35	0.32	0.40	0.23	0.27	0.27	0.73	0.30
Mean Canopy Depth (m)		9.5	10.0	16.5	13.0	10.5	11.0	7.0	5.5
Tree Count		45	36	61	36	35	55	175	252

Table 4.9 Site-by-site Moran's I autocorrelation and Getis-Ord general G clustering analyses, cross-referenced against field-acquired site attribute information. Tree species codes: CP = Corsican pine, BE = Beech and DF = Douglas fir.

All sites offer less than 1 % likelihood that the distribution of LiDAR terrain returns in three-dimensions are the result of random chance (table 4.9). The points at Site 7 are most clustered with Moran's I Index, M_I = 1.12, while those at Site 5 are least clustered, while still exhibiting a statistically significant pattern, with M_I = 0.53. It was originally hypothesised that clustering would increase with higher percentage understorey vegetation cover, which itself grows in patches, until a peak value, after which the cover becomes uniform and clustering declines. This would yield a negative second-order polynomial relationship between understorey cover and Moran's I, but no clear correlation is visible (figure 4.15 A). Likewise, it was hypothesised that clustering would increase with greater stem density, canopy depth and a higher number of trees, given that gaps in the canopy should theoretically diminish with a rise in these variables. Canopy depth and tree count offer little trend, each yielding a Pearson's correlation coefficient, *r* of less than 0.4. However, the Moran's I Index – stem density relationship exhibits a clear positive linear correlation (*r* of 0.87; figure 4.15 B), which suggests that total canopy cover, i.e. the size of canopy in addition to the number of trees, has primary control over the pattern of LiDAR terrain points. Over 80 % of the raw LiDAR data over Chopwell Wood were reflections from the tree canopy, so the distribution of canopy penetrating laser pulses was understandably organised by the canopy density (an allometric proxy of stem density).

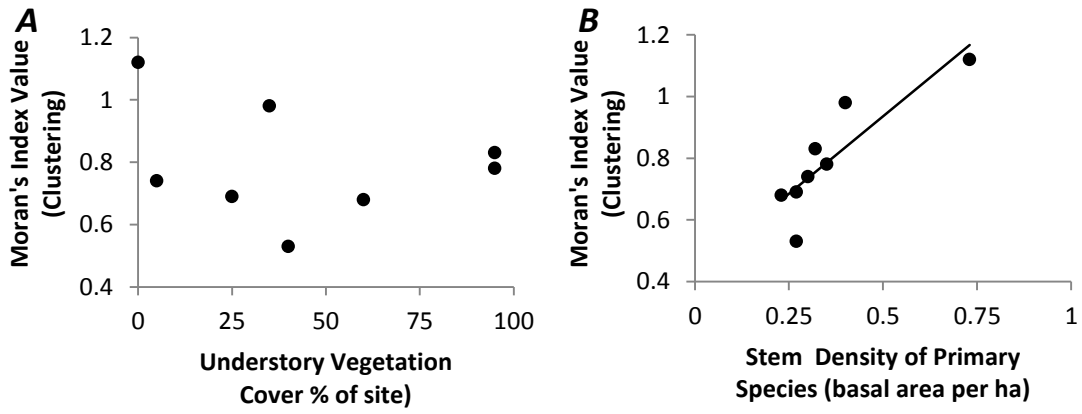


Figure 4.15 (A) Moran's I index against understorey vegetation cover. **(B)** Moran's I index against stem density of primary species, with a linear least squares regression trend line.

Getis-Ord General G clustering analysis evaluates to what extent points are concentrated in two dimensions. A fixed radius of 2 m was again established for point-to-point comparison, to rule out the possibility of mistakenly identifying patterns in random data. If the confidence (p-value) is small and the absolute value of the Z score (number of standard deviations from the mean) is large enough that it falls outside of the desired confidence level, a pattern between points is likely to that confidence level. The Index value presents little clear information but a higher 'Z Score' indicates greater spatial clustering.

Sites 2 to 4 and 6 to 8 offer less than 1 % likelihood that the distributions of LiDAR returns in two dimensions are the result of random chance (table 4.9). Site 1 demonstrates only moderate clustering, with a less than 5 % likelihood of random chance, and Site 5 demonstrates no clustering. While Site 3 has the highest Z score of +15.4, suggesting that terrain points are relatively clustered in two dimensions, the Z score of -0.54 at Site 5 is too few (< 1) standard deviations from the mean to indicate clustering. Again, with a view to establishing potential deterministic relationships between forest canopy, understorey cover, and two-dimensional clustering, least squares linear regression analysis was conducted between variables. Only the understorey vegetation density (percentage cover over the site) yielded strong positive correlation with the Z Score (*r* of 0.61), with the coefficients for stem density, canopy depth and tree count all < 0.4 and negative.

This suggests that, although canopy cover organises the distribution of canopy penetrating laser pulses and accordingly the distribution of LiDAR terrain points in three-dimensions, understorey vegetation controls the planimetric distribution of these points. In other words, the Getis-Ord clustering analysis ignores the elevation of terrain points and thus also misclassified points; so the understorey vegetation, understood to have caused Type II errors at the filtering stage of LiDAR data processing (see section 4.4.1), is strongly related to the distribution of terrain points in two dimensions, but not in three. Full, three-dimensional clusters of points are associated with gaps or low foliage density areas of the forest canopy.

4.4.3 Influence of terrain slope on accuracy

In order to establish whether terrain slope influenced the accuracy of airborne LiDAR terrain measurements, LiDAR terrain returns were assessed against field-acquired terrain reference data at each of the eight field test sites.

The average flying height during the Chopwell Wood helicopter LiDAR survey was approximately 490 m a.s.l. By evaluating equation 3.4 (section 3.5.2.3) with this altitude and with an average laser beam divergence of 0.5 (γ is unitless) and average scan angle of 10°, the average diameter of the laser footprint on the ground was determined as 0.29 m, equating to an area of 0.07 m^2. This equation assumes that the terrain surface is consistently flat, so equation 3.6 (section 3.5.2.3) was employed to calculate the size of the laser footprint at individual terrain returns, depending on the local slope. Local slope was obtained from a coarse, 1 m spatial resolution slope model through terrestrial LiDAR terrain reference data.

Site	Species	Footprint Diameter Constant 0.29 m Accuracy against Reference Points (cm)		Footprint Diameter α to Slope Accuracy against Reference Points (cm)	
		Offset	Std Dev	Offset	Std Dev
1	CP	13	21	22	21
2	CP	46	25	55	24
3	BE	0	13	9	12
4	BE	35	14	45	12
5	BE	12	28	12	27
6	BE	-2	12	-2	12
7	DF	8	18	8	17
8	DF	11	12	11	12
Average		20	18	20	17

Table 4.10 Statistical comparison of LiDAR terrain point accuracy when terrain slope is, or is not, taken into account. Positive mean offset describes the *over*-estimation in elevation of an airborne LiDAR point with respect to a reference point. See Appendix 1 for tree species codes.

Reference data comprised terrestrial LiDAR terrain points at spatial resolution of at least 5 points per m^2. All reference points within the horizontal bounds of each airborne LiDAR terrain return were evaluated, and statistics for mean elevation offset (i.e. between a LiDAR point and its associated reference points) and standard deviation of offset were generated at each site (table 4.10). These measures of accuracy were generated both for a constant laser footprint size and for a footprint size proportional to the local slope, so that the influence of terrain inclination on laser scanning accuracy could be assessed.

The largest differences in mean offset and standard deviation between accuracy calculated with a constant footprint diameter and accuracy calculated with a diameter proportional to the local slope are 1.4 and 1.8 cm, respectively, at Site 4 (figure 4.16). Relevant differences were expected at sites with steep terrain, i.e. Sites 7 and 8, but average differences in mean offset and standard deviation over all sites of only 0.4 and 0.5 cm, respectively, suggest that slope does not greatly influence laser terrain measurement accuracy.

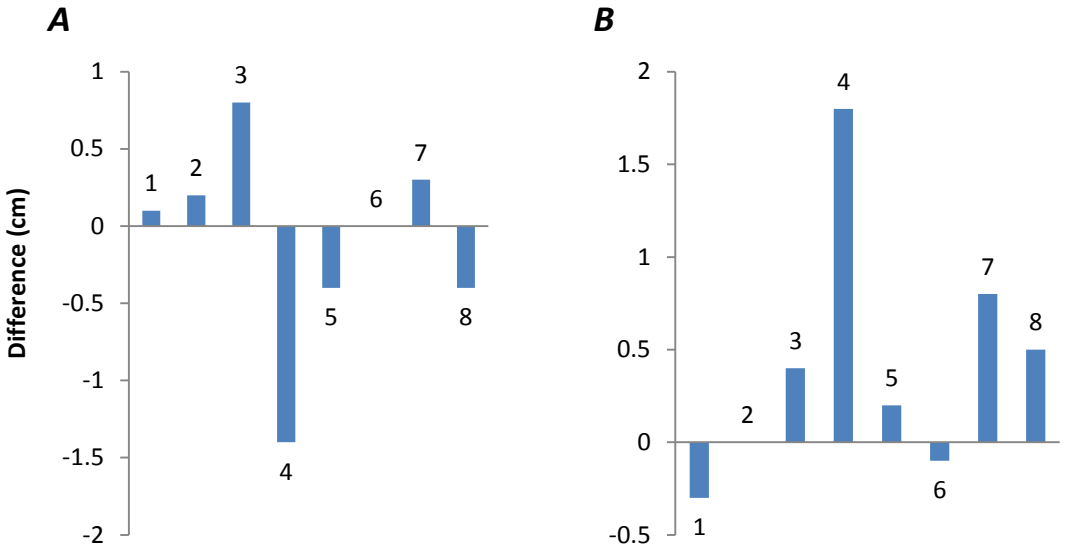


Figure 4.16 Difference between footprint analysis with constant diameter and diameter proportional to slope, where **(A)** refers to mean offset and **(B)** refers to standard deviation of offset. Positive difference describes greater accuracy using footprint diameters proportional to slope than using constant diameter footprints and vice versa. Data labels refer to the test site numbers, i.e. from 1 to 8.

4.4.4 Full study area accuracy and terrain point density assessment

Since airborne LiDAR data were acquired for the whole of Chopwell Wood, several parts of the standard accuracy and terrain point density assessment (section 4.4.1) could be extrapolated to the full study area. Without impractically extensive *in situ* field data collection, accuracy and assessment like that performed in section 4.4.1 could not be carried out over the entire forest. However, the Forestry Commission UK GIS (1998) offered spatial demarcation of individual tree stands, with each stand characterised by a single, or group of, tree species; and, consequently, terrain point density could be evaluated against tree species across the wood.

With the Forestry Commission GIS constructed primarily for commercial operations, several important facets of the data structure must be considered prior to use in research. One peculiarity is the procedure for splitting regions inside 'compartments' (stands) with multiple (but not mixed) tree species. Establishing the precise location of small stands is not essential to the Forestry Commission, so for the sake of visualising the data, a secondary species is represented by randomly distributed blocks (note the artificial square cells in figure 4.17). The areal fraction of a compartment covered by these blocks is accurate, but the locations of individual smaller stands are not and the random distribution is purely for visualising data. A second significant aspect of the GIS is the date of data acquisition. For almost all forest compartments this was 1998, which means that thinning and clearing operations, for example, contracted after this date, are not recorded. Some stands planted during the 2000s have been added to the initial data set and felled areas noted, but the most recent operations (2009-2010) have not and so field observations and communication with Forestry Commission rangers provide up to date information.

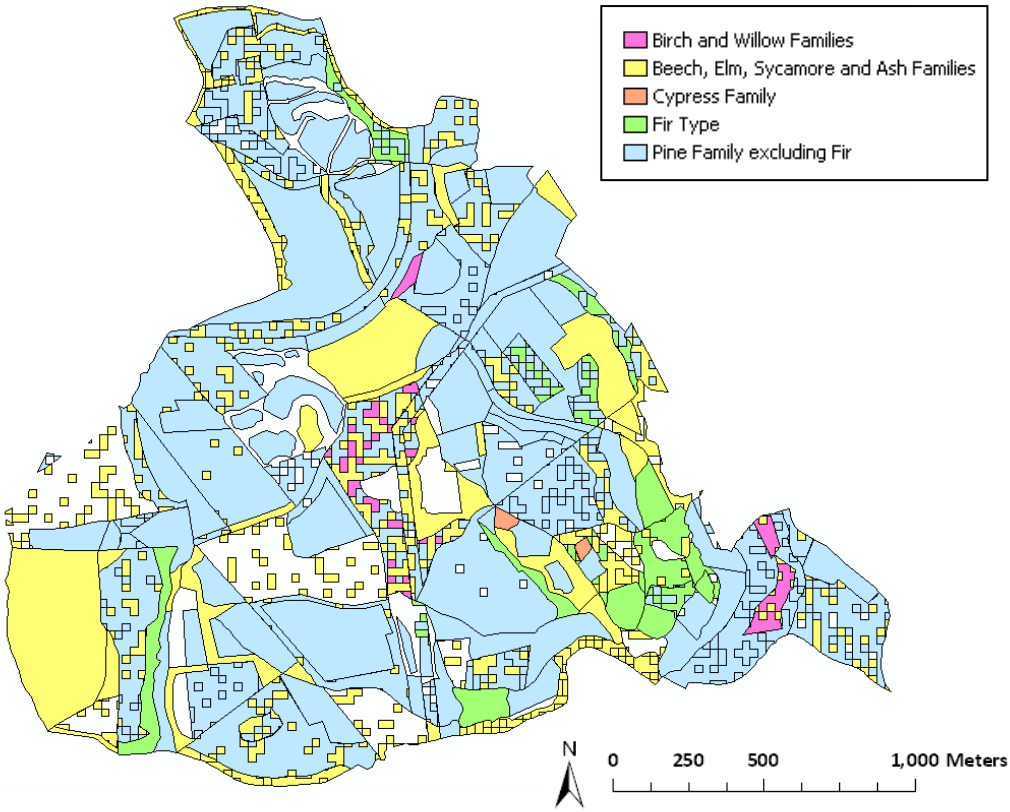


Figure 4.17 Spatial distribution of tree cover at Chopwell Wood. Tree species are grouped into five species classes. For details of tree species classes see Appendix 2.

Raw LiDAR data for the entire wood were classified with the progressive densification filtering algorithm, using the optimum parameters determined in section 4.2. Terrain point density (i.e. terrain points per m^2) was normalised by the average point density of the raw LiDAR data (31 points per m^2) (see section 3.5.2.4) and then calculated for each forest compartment. Mean terrain point density for each tree species was calculated as the average point density across all compartments of the same species, ensuring that the contribution of a compartment to the mean was weighted by its area. Terrain point density and canopy penetration rate for each species in Chopwell Wood are given in table 4.11, along with the area covered by that species. With the aim of identifying patterns across similar tree species, point density and penetration rate are also given for five species classes, and for coniferous and deciduous forest (table 4.11), and are represented graphically in figures 4.18 and 4.19.

Table 4.11 illustrates canopy penetration rates (and equivalent terrain point densities) in descending order, i.e. from most to least favourable penetration. It must be kept in mind,

throughout this assessment, that penetration rates are only valid for summer, broadleaf-on canopy and dense understorey conditions.

Mean normalised terrain point density is highest under Norway spruce at 6.3 points per m^2 and lowest under Western hemlock at 0.9 points per m^2. The latter was anticipated, prior even to commissioning the LiDAR survey, because the hemlock canopy lets almost no sunlight filter through to the forest floor even on the brightest summers day (D. Woodhouse, personal communication 2010) and it was expected to have the same 'impenetrable' effect on the laser scanner. However, the high point density under the Norway spruce was not expected, because published research has not generally advocated coniferous trees with needle foliage as being especially amenable to laser penetration (e.g. Hodgson et al., 2003; Hodgson and Bresnahan, 2004). It was only following discussion with Forestry Commission rangers, who knew the forest exceptionally well, that it was deduced that large stands of Norway spruce at Chopwell Wood had been heavily thinned.

	Tree Species	Mean Penetration Rate (%)	Mean Terrain Point Density (points per m^2)	Area (km^2)
	Norway Spruce	20.3	6.3	0.41
	Mixed Deciduous	17.5	5.4	0.53
	Oak	15.9	4.9	0.38
	Birch	14.5	4.5	0.04
	Lodgepole Pine	13.6	4.2	0.02
	Sycamore	9.7	3.0	0.01
	Scots Pine	8.5	2.6	0.39
	Douglas Fir	8.0	2.5	0.41
	Lawsons Cypress	7.8	2.4	0.02
	Japanese Larch	7.7	2.4	0.47
	Corsican Pine	7.4	2.3	0.51
	Noble Fir	7.3	2.3	0.01
	Hybrid Larch	7.2	2.2	0.01
	Beech	6.6	2.1	0.39
	Grand Fir	6.3	1.9	0.03
	Sitka Spruce	6.2	1.9	0.25
	European Larch	3.6	1.1	0.02
	Mixed Coniferous	3.3	1.0	0.03
	Western Hemlock	3.0	0.9	0.04
5 Classes	Description	Mean Penetration Rate (%)	Mean Terrain Point Density (points per m^2)	Area (km^2)
1	Pine Family minus Fir	7.9	2.4	2.12
2	Fir	7.6	2.3	0.18
3	Cypress Family	7.8	2.4	0.01
4	Beech, Elm, Sycamore and Ash Families	13.2	4.1	0.77
5	Birch and Willow Families	14.5	4.5	0.04
		Mean Penetration Rate (%)	Mean Terrain Point Density (points per m^2)	Area (km^2)
	Coniferous	7.9	2.4	2.31
	Deciduous	13.2	4.1	0.81

Table 4.11 Airborne LiDAR mean terrain point density and mean canopy penetration rate assessed by (1) tree species, (2) five species family classes (see Appendix 2) and (3) coniferous and deciduous forest.

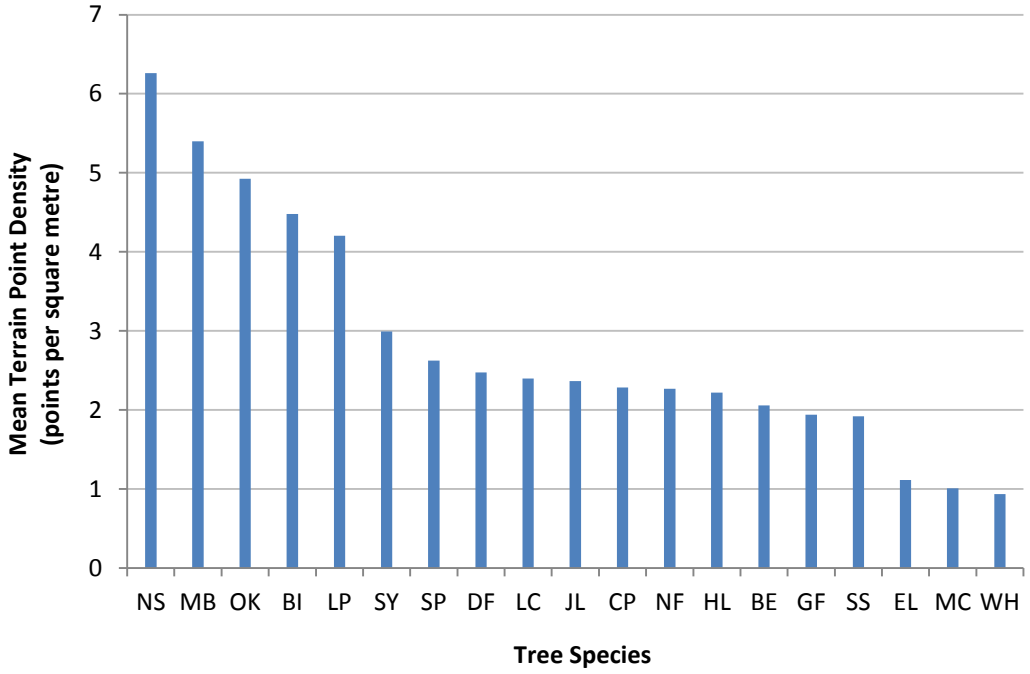


Figure 4.18 Mean airborne LiDAR terrain point density under specific tree species. See Appendix 1 for full tree species codes.

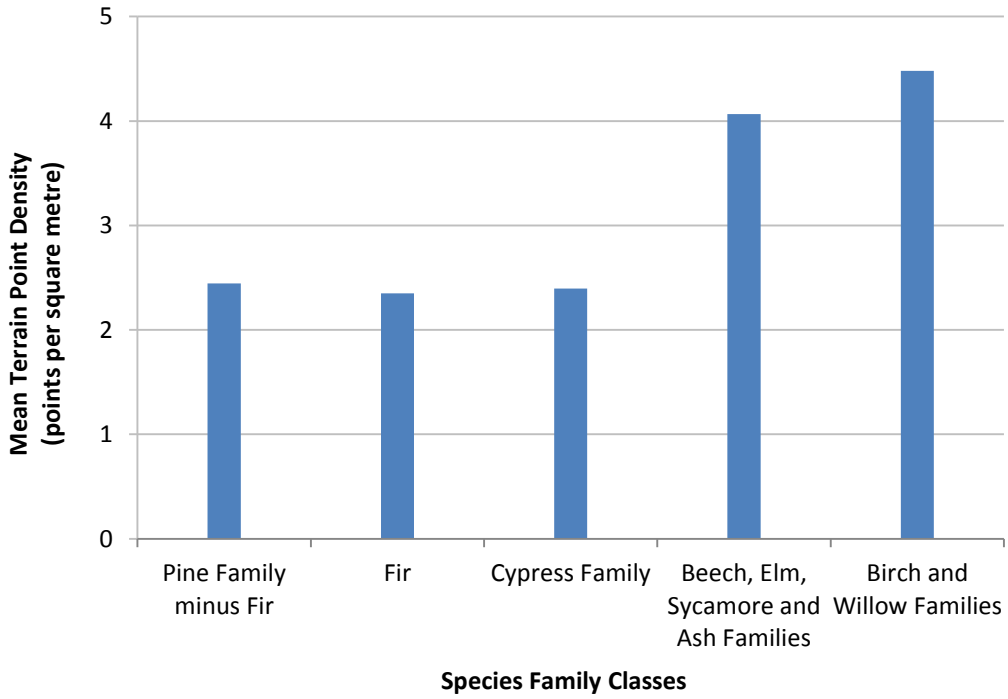


Figure 4.19 Mean airborne LiDAR terrain point density under five tree species family classes. See Appendix 2 for details on tree species classes.

It is clear from the Norway spruce that forest management practices can be highly influential on laser penetration rates. In general, un-thinned coniferous species of high density stocking,

for example the Western hemlock, Sitka spruce and Grand fir, offer lowest terrain point density (< 2 points per m^2), while thinned coniferous and broad-leaved (foliage-on) deciduous species, for example Corsican pine and beech, respectively, offer moderate point densities (2 to 3 points per m^2) and typically deciduous species with low stocking densities, like oak and birch, offer highest point densities (> 4 points per m^2).

It is also clear that the age of a stand has some control on the canopy penetration rate. Three stands of different age are presented in figure 4.20: (A) a juvenile Sitka spruce stand planted in 2002, (B) a Douglas fir stand planted in 1991 and (C) a Douglas fir stand planted in 1984 which has experienced rapid growth. All three stands are coniferous but demonstrate a trend of increasing mean terrain point density with age: the juvenile Sitka of 1.5 points per m^2, the 1991 fir of 2.4 and the 1984 fir of 3.1. The large canopy depth to stem height ratio evident in the Sitka (figure 4.19 A) is not conducive to either full-canopy laser penetration or terrain filtering, hence the low terrain point density. Alternatively, the obvious canopy in the older Douglas fir (figure 4.19 C) and lack of understorey are very beneficial to the filtering procedure. It is beyond the scope of this study to attempt to model or account for the effects of varying age and management strategy within individual species on LiDAR terrain modelling quality, but it is theoretically possible if comprehensive data for stand planting year and species growth models are readily available.

Figure 4.20 (A) 7 year old Sitka spruce. **(B)** 18 year old Douglas fir. **(C)** 25 year old Douglas fir. Ages are given as they were at the time of airborne LiDAR data acquisition.

Individual tree species are classified into five family groups in table 4.11. It is instantly noticeable that LiDAR terrain point density is very similar among the three coniferous species groups and also the two deciduous species groups, suggesting that generalisations in pre-survey point density forecasts are valid over particular areas of coniferous or deciduous forest. Mean terrain point density is 2.4 points per m^2 under coniferous forest and 4.1 points per m^2 under deciduous forest (table 4.11). The canopy penetration rates, of 7.9 % for coniferous and 13.2 % for deciduous forest, are perhaps more significant for future research or commercial operations, as these figures could be exploited, pre-survey, to map sub-canopy terrain point density forecasts based on the expected raw point density of a particular LiDAR system at a particular average flying height. Equally, this procedure could be employed inversely to determine flight parameters (including altitude) from pre-requisite sub-coniferous and sub-deciduous canopy terrain point densities.

A measure of sub-canopy terrain surveying viability is presented with a view to informing archaeological feature prospection strategy. 'Viability' was calculated by weighting normalised terrain point density at each forest compartment (stand) by the expected accuracy of terrain points based on the results of this study. The results of the standard accuracy assessment (see section 4.4.1) suggested that LiDAR terrain returns under beech forest are typically more accurate than those under fir, and that terrain returns under fir are typically more accurate than those under pine (table 4.6). So terrain point density was given a weighting of 2, 1 or 0.5 for respective deciduous species, fir species and non-fir coniferous species (pine, larch, spruce and hemlock). It must be taken into account that generalisations over deciduous, fir and pine forest are based on stocking densities and average stand ages typical of Chopwell Wood, which is itself characteristic of woodland across the UK. The viability measure is relative and is, as such, not quantified.

The spatial distribution of tree species family classes is illustrated in figure 4.17 as reference for the viability map in figure 4.21. Maximum viability is predictably located under oak and birch stands, where laser penetration through the canopy is high and terrain measurement

accuracy is also high. Minimum viability is located under Corsican pine, Japanese larch and Western hemlock stands in the west, south and east of the wood, respectively, where canopy is almost impenetrable and/or understorey vegetation causes significant terrain measurement bias and imprecision.

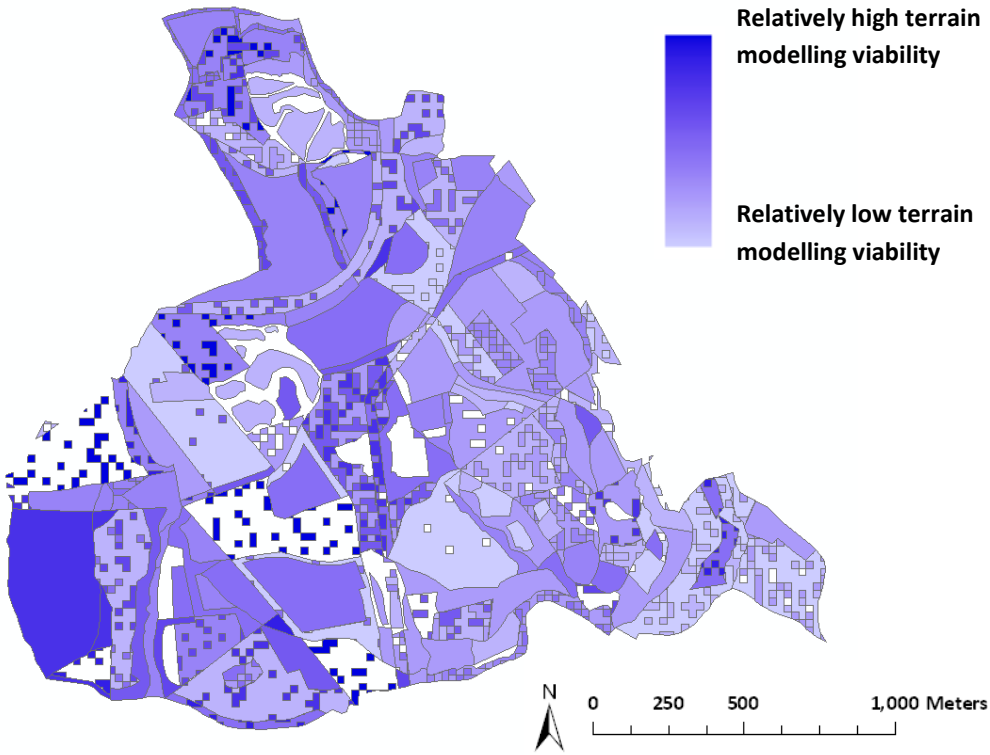


Figure 4.21 Relative viability of airborne LiDAR sub-canopy terrain survey, based on post-processed terrain point density and accuracy with respect to the true terrain elevation. White regions are open areas within the boundary of the wood.

The viability map is purely a tool for visualising prospective differences in terrain model quality between areas of the Chopwell Wood LiDAR data set. It might be argued that the ranges in both terrain point density and measurement precision, of approximately 1 to 6 points per m^2 and 15 to 20 cm standard deviation, respectively, for various tree species, are not sufficiently wide to have a significant effect on the detail or quality of a terrain model. This argument would be supported by the fact that the lower ends of each range, i.e. 1 point per m^2 and 20 cm precision, are typical of, and frequently better than, many published archaeological feature prospection studies (e.g. Devereux et al., 2005; Sittler, 2004). While archaeological feature

detection may be possible across all areas of Chopwell Wood, in order to resolve the most subtle topographic features, for example shallow mounds and ditches with height and width less than one metre, higher point density and accuracy may be required. This subject will be explored further in the following chapters.

Chapter 5. Results and Analysis 2: Terrain Modelling and Archaeological Prospection

5.1 Digital Terrain Modelling

Filtered airborne LiDAR terrain data are modelled into a surface structure to facilitate sub-canopy archaeological prospection. After evaluating interpolation methods (see section 4.3), the raster-based kriging technique was selected for modelling gaps between LiDAR points. Interpolated digital terrain models (DTMs) are generated and used first for hillshade analysis, thereby emphasizing topographic features with artificial shadows. A 0.25 m spatial resolution hillshade terrain model of the whole of Chopwell Wood is illustrated in figure 5.1, but, from this relatively remote viewpoint, only large river tributaries and topography at a spatial scale greater than approximately 100 m can be identified. Several white areas are noticeable in the model where terrain points are too isolated to accurately model intervening topography.

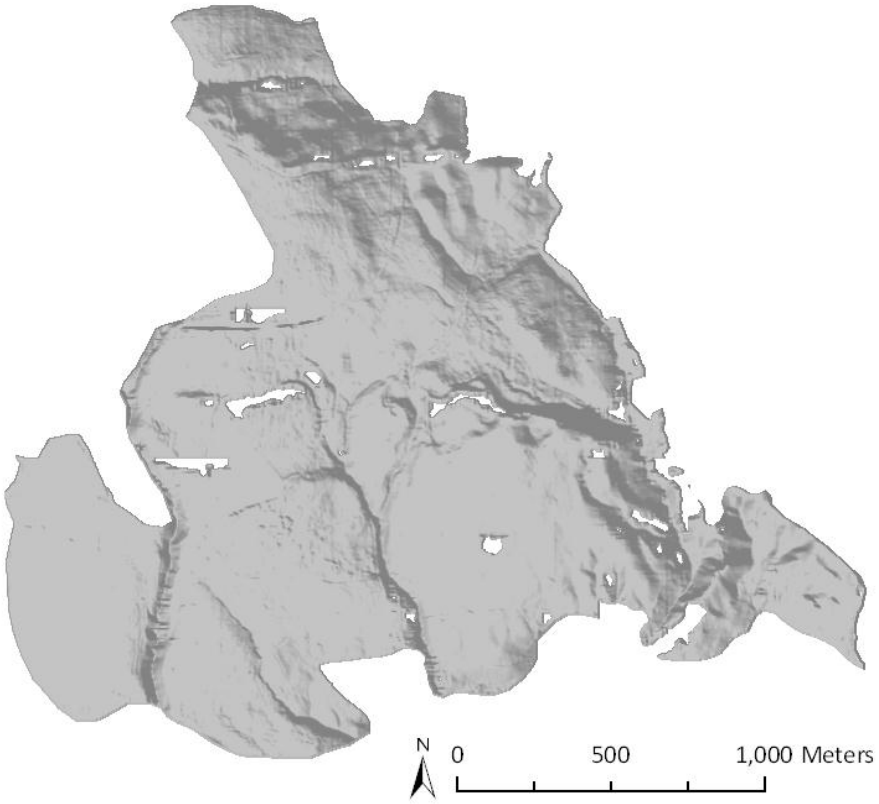


Figure 5.1 Hillshade terrain model of Chopwell Wood with 0.25 m spatial cell resolution.

With the aim of demonstrating variation between terrain models under different areas of the forest, 0.25 m spatial resolution hillshade models at three of the field test sites are presented in figures 5.2 to 5.4, covering each of the tree species: pine, beech and fir. While LiDAR terrain points are dense under the pine at Site 2 (39 points per m^2) (figure 5.2), some quasi-circular gaps are evident between point clusters (note the planimetric point distribution), suggesting that the large understorey plants, evident in figure 5.2 A, prevent all laser pulses from reaching the terrain. In reality, the true terrain at Site 2 is not characteristic of the 'bumpy' topography apparent in figure 5.2 C and emphasized in figure 5.2 D, which also suggests that some LiDAR returns from scrub, or mixed scrub and terrain reflections, have been mistakenly classified as terrain during filtering. Each of the 'bumps' evident across the profile in figure 5.2 D is therefore likely an individual or amalgamated group of low plants, rather than rough topography. However, the linear feature highlighted in figure 5.2 C, identified in the field as a natural drainage channel of approximately 50 cm depth, offers some promise for detecting topographic features under pine canopy in spite of the 'noise' associated with understorey vegetation.

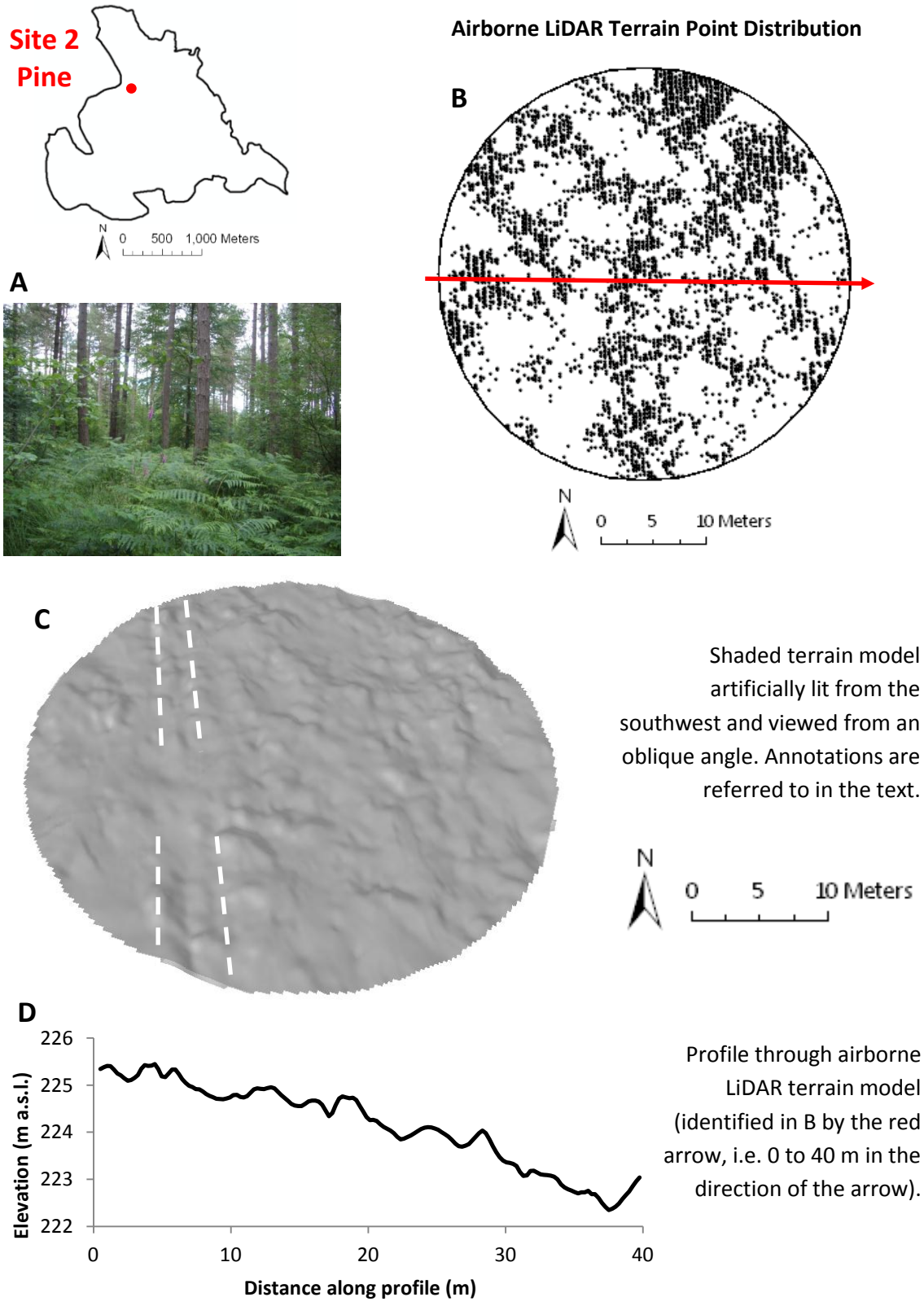


Figure 5.2 Site 2. **(A)** Site photograph. **(B)** Planimetric distribution of airborne LiDAR terrain points, with the location of the profile in D identified by the red arrow. **(C)** A shaded terrain model through airborne LiDAR terrain points. **(D)** A profile through the terrain model.

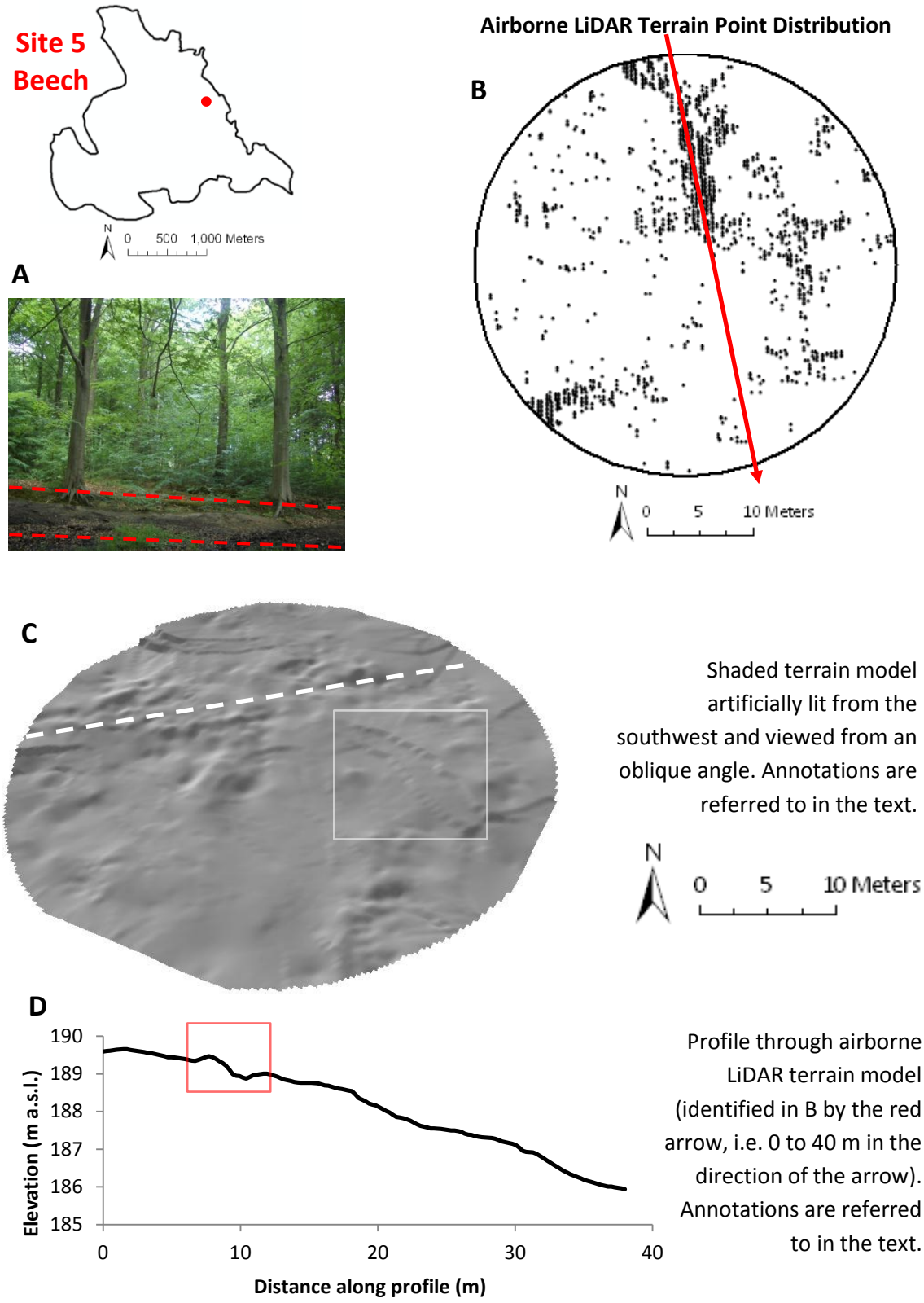


Figure 5.3 Site 5. **(A)** Site photograph. **(B)** Planimetric distribution of airborne LiDAR terrain points, with the location of the profile in D identified by the red arrow. **(C)** A shaded terrain model through airborne LiDAR terrain points. **(D)** A profile through the terrain model.

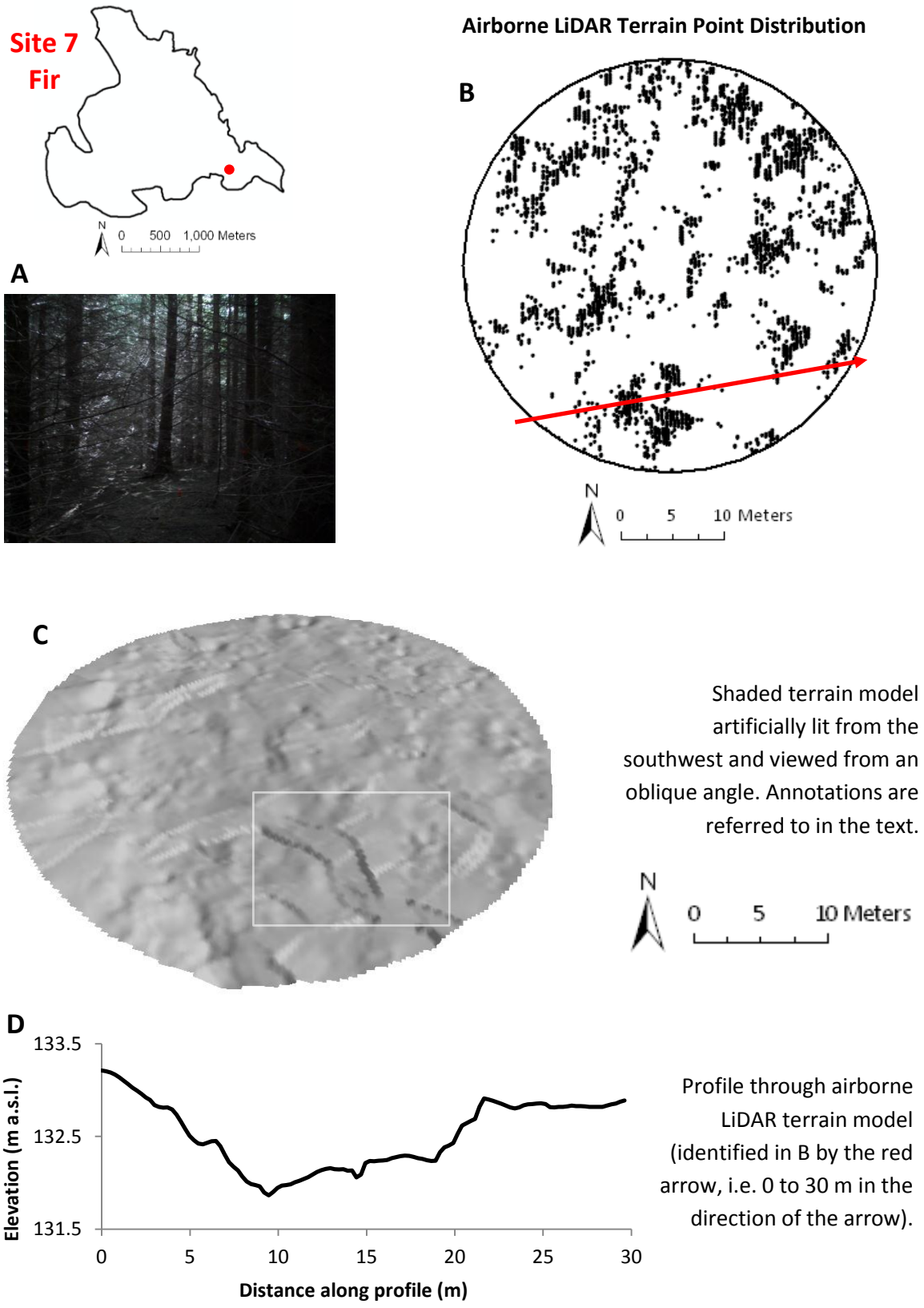


Figure 5.4 Site 7. **(A)** Site photograph. **(B)** Planimetric distribution of airborne LiDAR terrain points, with the location of the profile in D identified by the red arrow. **(C)** A shaded terrain model through airborne LiDAR terrain points. **(D)** A profile through the terrain model.

Although laser pulse penetration through the beech canopy is less frequent than the pine, resulting in a lower terrain point density at Site 5 than Site 2 (11 points per m^2, figure 5.3), understorey vegetation is less dense (figure 5.3 A) and the terrain model (figure 5.3 C) is consequently much smoother, exhibiting little of the 'noise' so evident under the pine (figures 5.2 C and D). However, an artefact of the interpolation procedure, `stepping`, mentioned previously in section 4.3 and typically associated with the inverse distance weighting technique rather than kriging, is evident in the terrain model in figure 5.3 C and highlighted by a white box. This artefact can cause problems when it occurs over archaeological features in terrain models and affects feature detection. A linear step-type feature is highlighted by the dashed white line in figure 5.3 C and identified in the field as one edge of a clear avenue of approximately 3 to 4 m width. Both edges of the avenue can be identified in the field, but only the northern edge has an obvious topographic imprint, highlighted by the red lines in figure 5.3 A and the red box in figure 5.3 D. This feature will be discussed in greater detail in section 5.2.3 but it was mentioned in the early field visits that a pre-1800 AD forest house may have stood at this site and, if so, the feature may have been an avenue to this house (L. Searle, personal communication 2010). Clustering of terrain points is evident, especially in the northern part of Site 5 (figure 5.3 B), caused by greater canopy laser penetration over the avenue feature where there are few stems and only sparse foliage from adjacent beech trees.

Canopy penetration through the fir at Site 7 is more frequent than the beech and less than the pine (figure 5.4 B), but a complete absence of understorey vegetation (figure 5.4 A) yields a smooth terrain model (figure 5.4 C). Site 7 is located at the lip of a steep slope and a ditch-type feature of approximately 10 m width and 1 m depth is visible in the profile across the southern part of the terrain model (figure 5.4 D). This feature is too large to be a drainage channel but, following validation in the field, is discovered to be the source of natural surface flow down the local slope.

While the topographic features identified at the test sites are primarily not archaeological, they give a preliminary indication of the minimum size and most favourable shapes for

detection. Linear, channel-type features of < 1 m width and depth are resolved even under the thick scrub vegetation in the pine, due to their consistency over 4 m plus distances, even if sections are missed. Other linear features, for example the edge of the 'avenue' at Site 5, are less obvious but can still be located if point density is high enough. There appears to be an element of chance involved in sub-canopy archaeological prospection with airborne LiDAR terrain data, given that LiDAR points tend to cluster in space (see section 4.4.2). It is plausible that one of two adjacent archaeological features of equivalent size, under the same forest species, could be resolved in detail while the other is missed altogether, depending on their relative positions with respect to tree crowns and understorey. Non-linear features appear to be more difficult to identify in shaded terrain models as they often resemble natural topography.

5.2 Archaeological Prospection and Validation

Rigourous LiDAR terrain filtering and comprehensive accuracy and point density assessments were performed with the research aim consistently kept in mind: to test the capability of airborne laser ranging for detecting and resolving archaeological features in sub-canopy topography. The following section will utilise the information obtained in Chapter 4 and additional visualisation and post-processing analyses to evaluate this aim. A number of archaeological features in the forest were either familiar to a local community group or located during preliminary field visits by the author, prior to carrying out this investigation, although information on context or history is not necessarily established. These 'known' features offer a means of testing the capability of airborne LiDAR and are detailed in table 5.1. Four groups of features are selected for full analysis, labelled in table 5.1 by a number in brackets, with their locations given in figure 5.5.

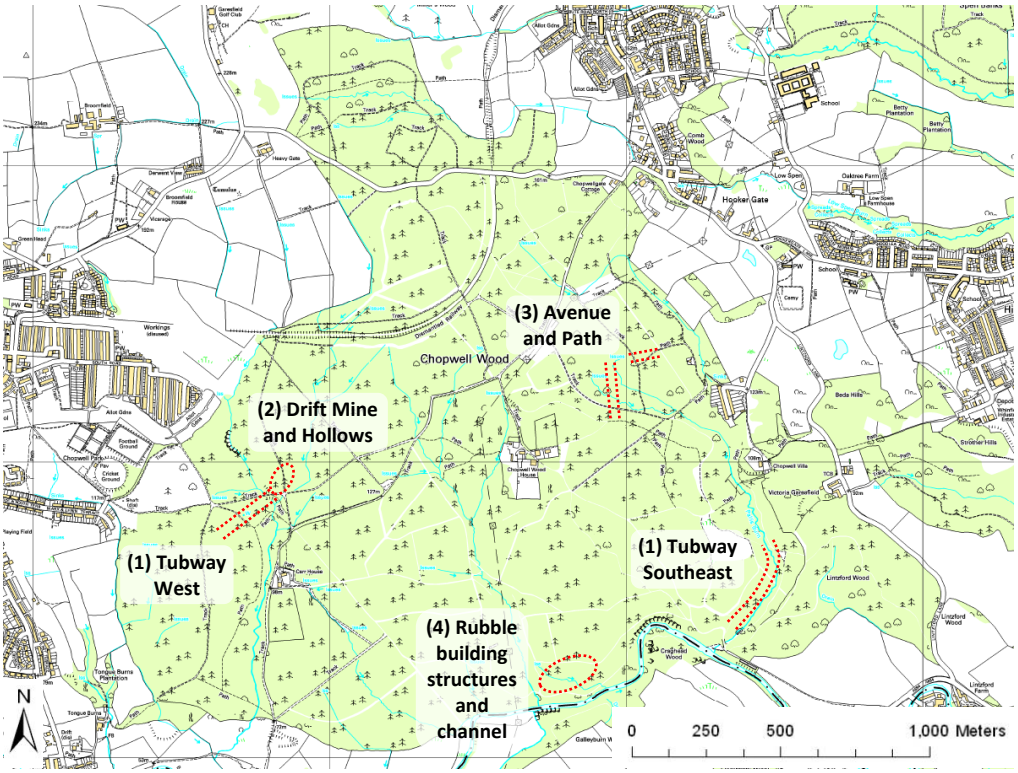


Figure 5.5 Locations of the archaeological features concentrated on in the prospection evaluation. Base is a 1:10,000 scale OS map. Source: EDINA (2009).

Feature (Group)	Easting (m)	Northing (m)	Canopy Species	Topographic Signature	Width, Depth (m)	Shape	History
Avenue	414153	558392	BE	Obvious edge	3-4, 0.1-0.5	Linear, wide	Avenue to Dene House?
Pathway	414041	558236	BE, CP	Brick material in path	1, 0.1-0.3	Linear	Path to Dene House?
Reservoir	414209	558334	BE	None	?, ?	Circular	
Building Rubble (4)	413834	557270	DF (MB)	5+ piles of building rubble	5-6, <1	Circular, linear	House/ building?
Man-made Channel (4)	413829	557256	DF (MB)	Channel, small bridge	2, 2	Linear, forked	Part of building?
Disused Railway	413337	558504	JL, BE, OK	Large channel/cutting	10, 3-4	Linear, curved	Coal transport
Stone Overbridge	413546	558974	DF, CP, JL, BE	Large intact bridge	10, 8	Linear	Access over trainline
Tumulus	412695	558904	Few Stems	Mound	5, <2	Elongate	Ancient burial mound
Enclosure	414838	557362	Open	Artifically shaped topo	>10, 2-3	Circular	Ancient
Tubway West (1)	412725	557850	Open	Large ditch	5, 2-3	Linear	Coal from drift mine
Drift Mine (2)	412899	557932	CP	Large slope	8, 5-6	Concave Slope	After 1893, pre-1939
3+ Hollows (2)	412899	557932	CP	Line of small pits at mine	<2, <1	Circular	
Site of Sawmill	413229	557916	CP	Shaped platform	>10, ?	Square	World War II
Drover's Road	412778	557339	MB, JL	Livestock passage	>5, 0.5	Linear	
2 Pits	414116	557569	DF	Small pits and spoil heap	1-2, <1	Circular	
Tubway SE (1)	414443	557441	SP, CP, MB	Large ditch	5, 2-3	Linear	Coal transport
Dam/Sluice	414371	557422	SP	Obvious brick structure	10, 2-3	Square	Pre-1854

Table 5.1 'Known' features of archaeological interest at Chopwell Wood. Easting and northing coordinates refer to the OSGB36 system. Tree species codes are given in Appendix 1.

Of the features not selected for evaluation, the large disused railway cutting and stone bridge crossing the railway are very obvious in the filtered LiDAR terrain model and can be identified even at the scale of figure 5.1. The tumulus and enclosure, although interesting features, are just outside the boundary of Chopwell Wood and are therefore not covered by plantation forest, so the benefit in testing at these sites is reduced. The topographic signature of the Drover's Road has been disrupted by its contemporary use as a forest track and little field validation has been conducted at either this or the site of the World War II sawmill, again reducing the benefit in testing the LiDAR at these sites. All LiDAR terrain models presented in this chapter have spatial cell resolution of 0.25 m.

It must be acknowledged at this stage that prospection is conducted with the primary aim of resolving and identifying archaeological features in the sub-canopy topography. Interpretation and understanding of features and their context is secondary and in this study is limited to constraining age/period and validating function. It must be understood that the author is not an archaeologist and that any interpretation offered either for currently poorly understood features or for newly identified features is not necessarily correct and would need professional evaluation at a later date.

5.2.1 Coal tubways

Two trough-type features, both of approximately 2 m depth and 5 m width, can clearly be identified at the south east corner and western edge of Chopwell Wood (figure 5.5). These are known 'tubways', constructed to transport coal from mine shafts within the forest. Photographs of the two tubways are presented in figure 5.6, both exhibiting comparable shape and dimensions. Given the (relatively large) size of the features, locating them in the airborne LiDAR terrain data is not difficult; but what these features offer is the possibility of testing to what level detail, especially of the trough rims and base, can be resolved. While the west tubway is located in open terrain, free from canopy cover, the southeast tubway is completely

hidden under discrete regions of Corsican and Scots pine, with some intermittent mixed broadleaf species. The latter is also situated on the side of a steep slope (up to 40°), so reduced laser return accuracy is expected.

Figure 5.6 (A) Winter photograph of the western coal tubway cross-section from the position of transect B in figure 5.12. Note the tubs and tracks are replicas placed at this site for heritage value. **(B)** Summer photograph up the southeast tubway from the position of transect B in figure 5.7, with pine forest cover.

A shaded terrain model of the southeast tubway, artificially illuminated from the northwest at altitude 45 degrees, is shown in figure 5.7. Various features are evident in the model, most notably a significant tributary of the Derwent River as well as the main river channel, an open field (observe the even surface) and a forest track. The tubway is highlighted by the dashed red lines and could be identified by the prominent shadow over one side of the trough after shading. Three transects (on equivalent vertical scales) are illustrated in figure 5.8 and referred to in figure 5.7. The high gradient of the area is clear at transect B and, while along-profile channel dimensions are not regular, the tubway can be identified at all transects, with the signature in the topography strongest at transect A (width > 5 m, depth > 2 m). One of the least obvious sections of the tubway, at transect B (depth < 1 m), is still resolved in the LiDAR, and even topography of < 50 cm height is evident in the model.

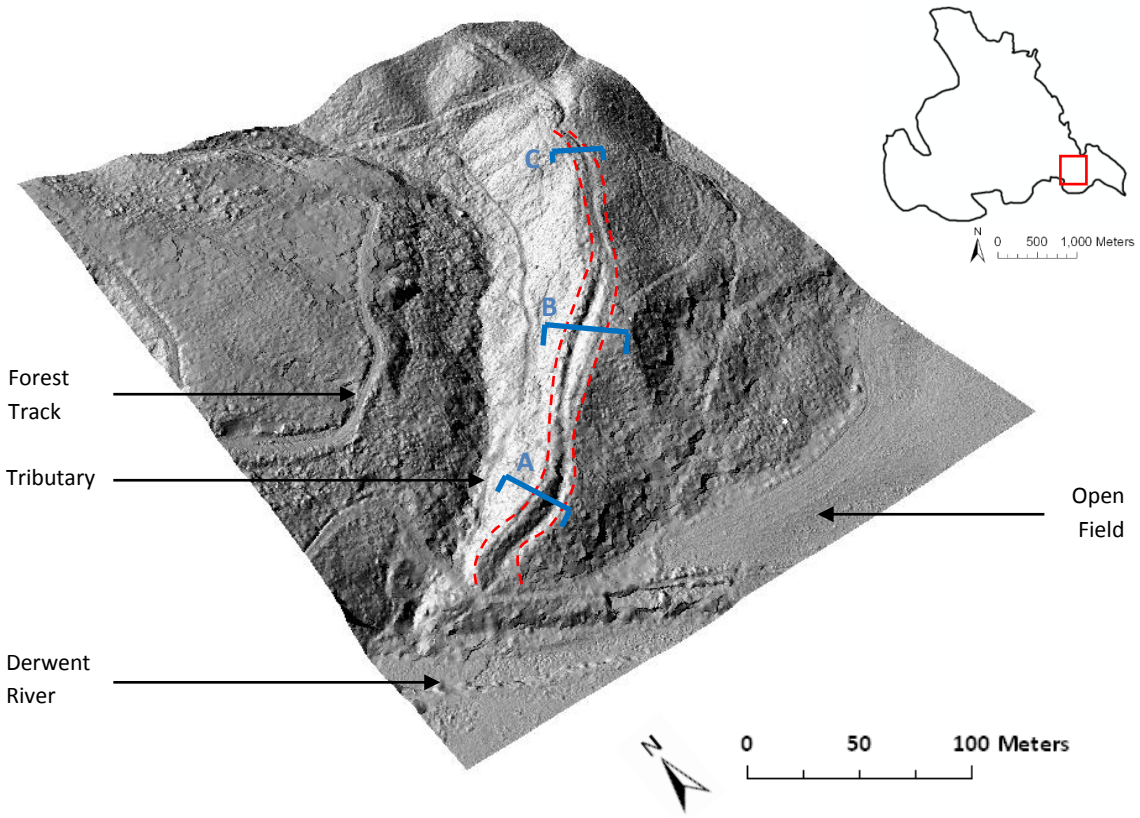


Figure 5.7 Shaded terrain model of the southeast coal tubway with illumination from the NW at 45°. Note that illumination parameters are consistently given as: 'azimuth', 'altitude' (in degrees). The edges of the tubway are highlighted by the two dashed red lines and lines A to C indicate the locations of transects in figure 5.8.

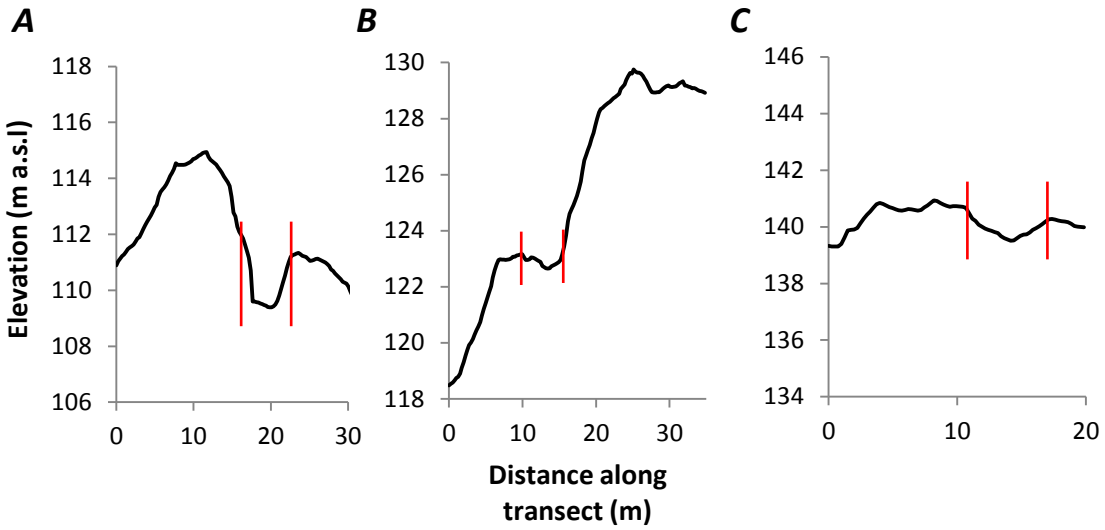


Figure 5.8 Vertically exaggerated transects A, B and C across the southeast coal tubway (referenced in figure 5.7) with the cross-sectional extent of the tubway highlighted. Distance along transect is always given in a west to east direction.

In order to validate both the accuracy and detail of this LiDAR-derived feature, control points along the base of the tubway profile, acquired through differential GPS survey in the field, are compared with a transect through the LiDAR DTM (figure 5.9). GCPs are located at approximate 5 m intervals along the profile and thus horizontal characteristics of the topography can only be validated at this spatial resolution. Additionally, the horizontal accuracy of the GPS was compromised to some extent by the forest canopy and rough terrain, with some difficulty detecting satellites; so mean control point horizontal and vertical accuracy is ± 1.08 and ± 3.02 m, respectively. Hence, while LiDAR elevation overestimates the elevation of the GPS data by 0.24 m (± 0.50 m), the GCPs are up to 3 m out and, based on the vertical accuracy of the LiDAR under pine during accuracy assessment (see section 4.4.1), the terrain model is up to 0.21 m out. So although it could be expected that the LiDAR might over- rather than under-estimate the elevation of control points, since returns from low vegetation (figure 5.6 B) may have been mistaken for terrain at the LiDAR filtering stage, the estimated 0.24 m difference in elevation between the two data sets is still smaller than the total LiDAR plus reference (control point) data error of ± 3.23 m.

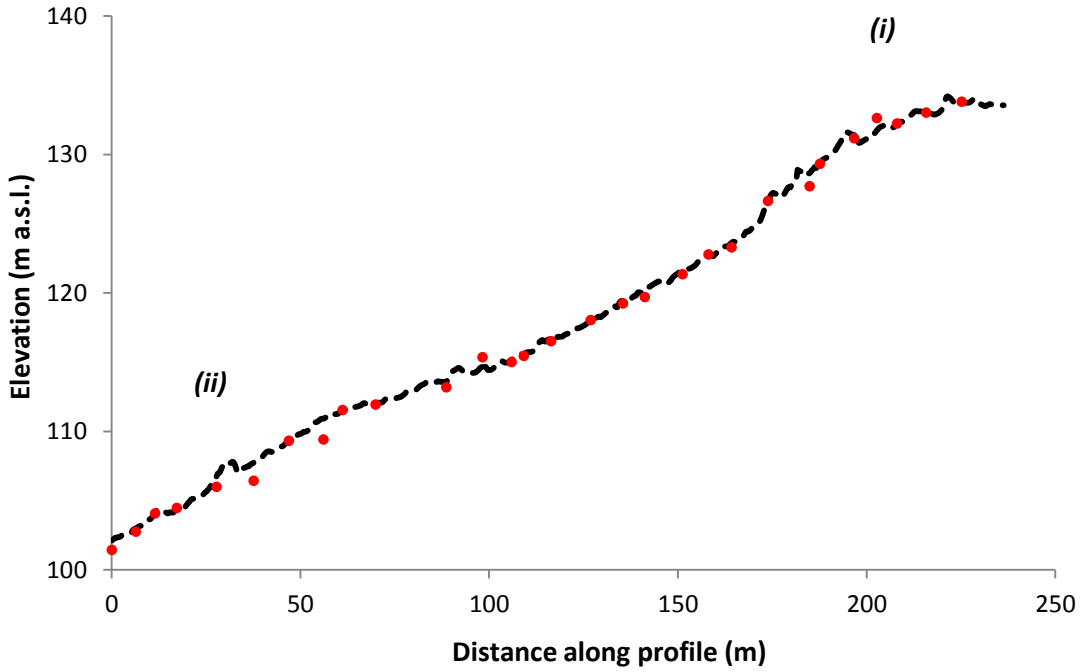


Figure 5.9 Profile through the LiDAR terrain model up the southeast coal tubway, along the base of the trough (black line) with GPS control point locations and elevations (red points).

Yet, despite expected error, the airborne LiDAR terrain data models the true ground surface closely and incidences where the LiDAR resolves even minor deviations in the terrain, for example at (i) in figure 5.9, are more common than irregularities in the modelled surface, for example at (ii). It is evident from the LiDAR terrain model that the rims of the tubway are also extracted correctly and, where no laser return falls directly over the edge, kriging interpolation aids in the derivation of breaks in slope between points either side of the edge.

Aerial photographs of Chopwell Wood, provided by a member of a local community group (P. Fountain, 2010), are available for 4th October 1947 and 13th March 1956. By fortunate coincidence the area of forest over the southeast tubway was felled for timber in 1956, immediately prior to image acquisition (with the current pine stands replanted later that same year), and the canopy covered and uncovered photographs are shown in figure 5.10. Figure 5.10 B offers, following image registration, high accuracy feature positioning; but further analysis using these images is beyond the scope of this study and they are presented as an example of data that could be utilised, in combination with LiDAR, for advanced archaeological interpretation.

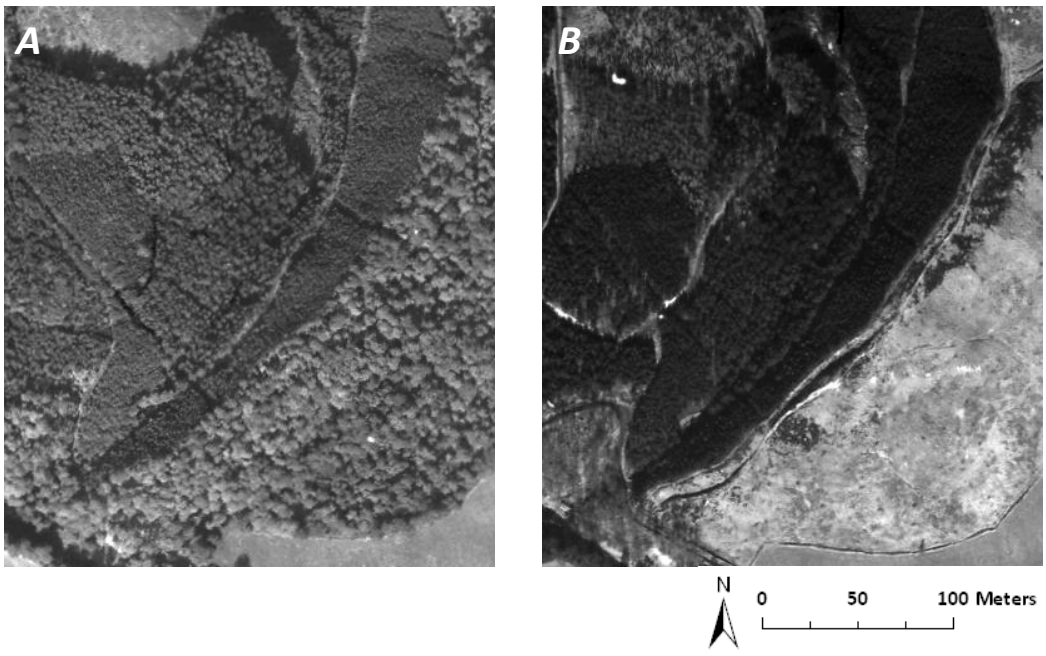


Figure 5.10 (A) Aerial photograph of the southeast coal tubway below tree cover. Date: 4/10/1947. **(B)** Aerial photograph of same area immediately after deforestation. Date: 13/03/1956. Source: Royal Air Force Photographs from World War 2 UK aerial survey.

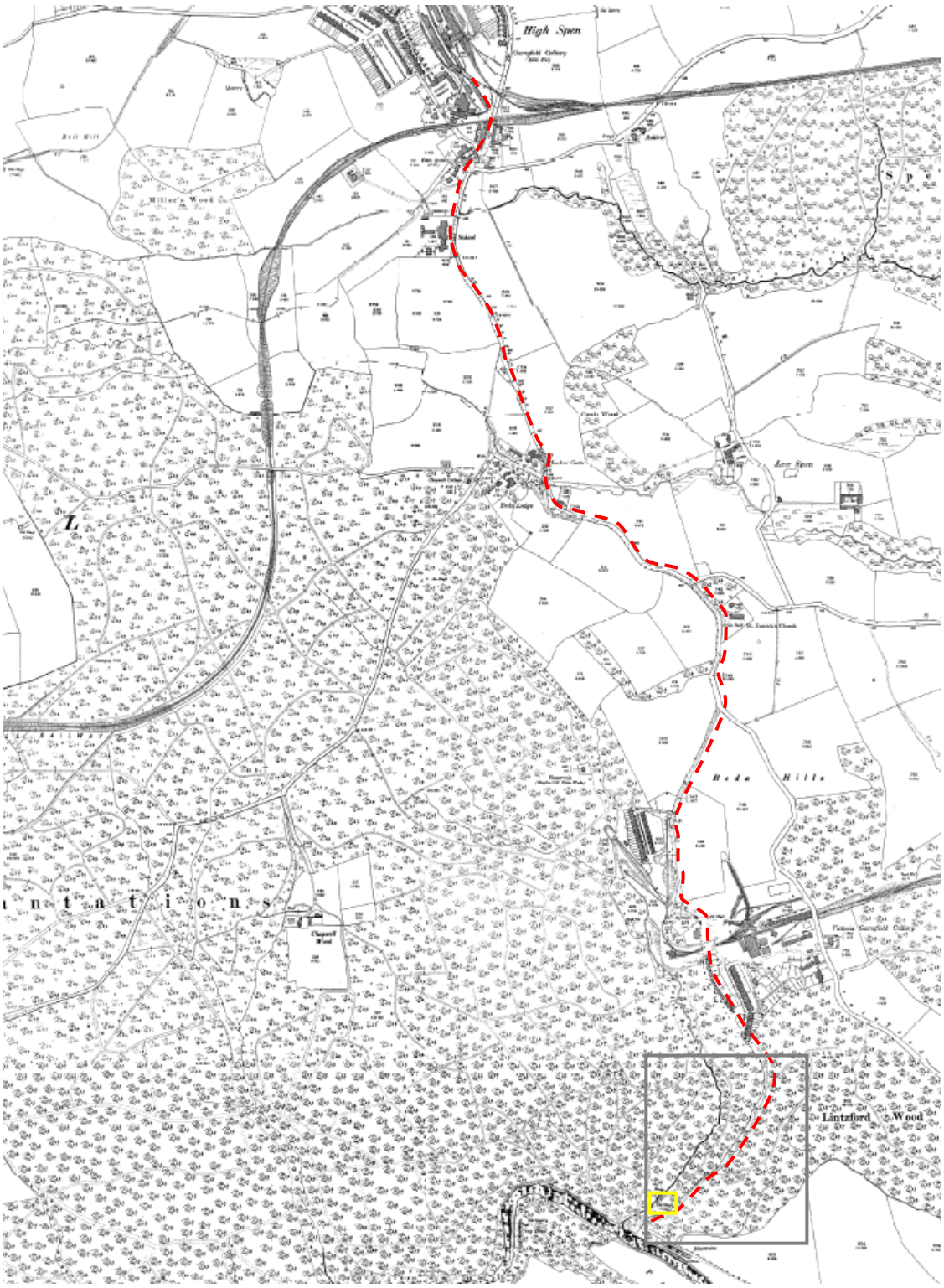


Figure 5.11 Historic 1:2500 County Series map, registered as in operation between 1893-1915, with the investigated section of the southeast coal tubway inside the grey box and the continuation of the tubway to Garesfield Colliery in High Spen highlighted by the dashed red line. A coal shaft can be identified on this map inside the yellow box, but may be difficult to pick out. Source: EDINA (2009).

Basic clarification of the history and context of this tubway can be made through interpretation of historic Ordnance Survey 6" Series maps (EDINA, 2009). A coal shaft and early tubway configuration can be identified in County Series mapping from 1854, but the edition operational from 1893-1915 illustrates the continuation of the tubway from the section recognised in the LiDAR data (grey box in figure 5.11), along the eastern edge of the wood up to Garesfield Colliery at High Spen (dashed red line in figure 5.11). Apart from the section identified, this tubway is known to have passed *underground* along this route (Searle, personal communication). At Garesfield colliery it intersects the railway line which up until 1961 transported coal through and from the wood. The 1924 edition of the same map is the first to show the tubway as inoperative and the coal shaft is labelled as 'old shaft'. The southeast tubway can, as such, be constrained as active for an unknown period between 1854 and 1924, with full operation in place by 1893. To validate this interpretation, intermittent sections of the tubway were located in the field along the full length highlighted in figure 5.11.

The tubway at the western edge of the forest is highlighted by dashed red lines on a shaded terrain model in figure 5.12, artificially illuminated from the southeast at an altitude of 45°. A large tributary to the Derwent River is identified in the model, intersecting a forest road and the tubway at approximately the same location. Prominent shadow over one side of the trough strongly emphasizes the feature and the similarity in detail between this model and that of figure 5.7 suggests that LiDAR terrain model quality at the southeast tubway is not severely degraded after filtering the pine canopy. Two transects across the feature are given in figure 5.13 and referred to in figure 5.12. The trough has a maximum depth of 2 to 3 m at transect B and while the topographic signature is diminished further west (transect A), over 100 metres of tubway length can be identified (figure 5.12).

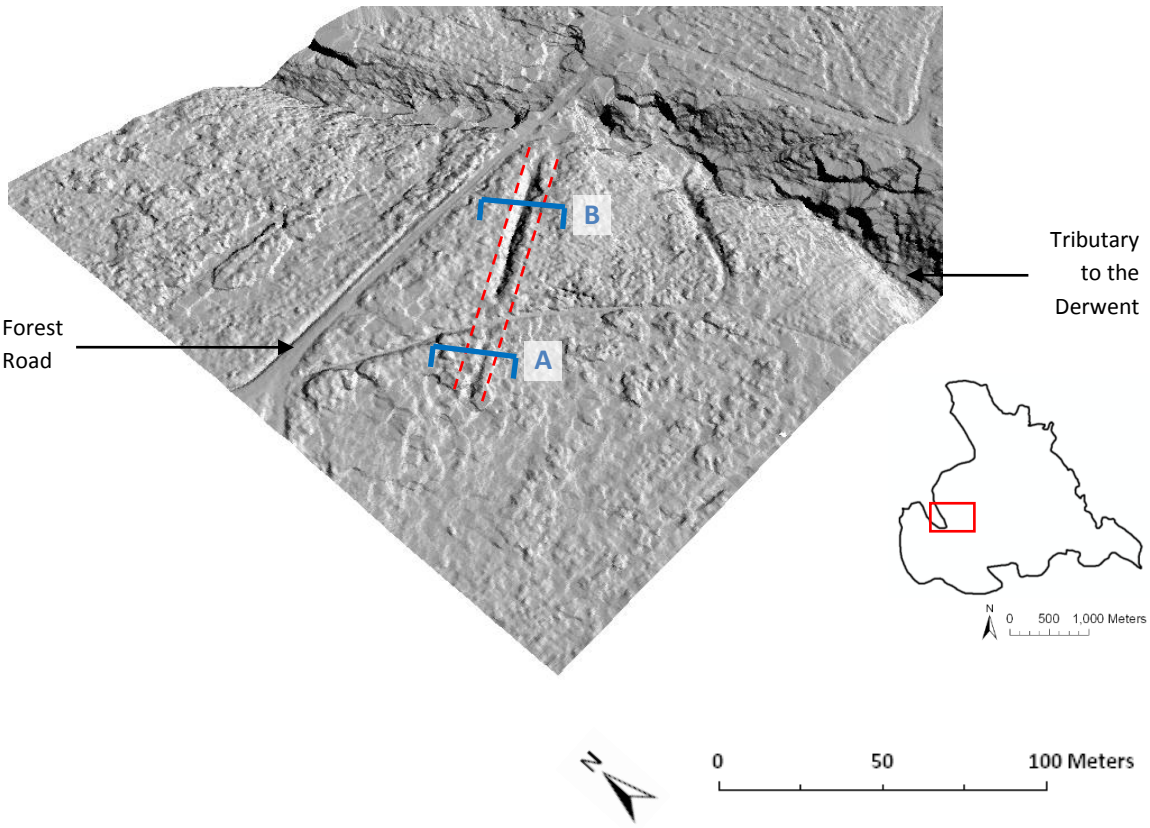


Figure 5.12 Shaded terrain model of the western coal tubway with illumination from the SE at 45°. The edges of the tubway are highlighted by the two dashed red lines and lines A and B indicate the locations of transects in figure 5.13.

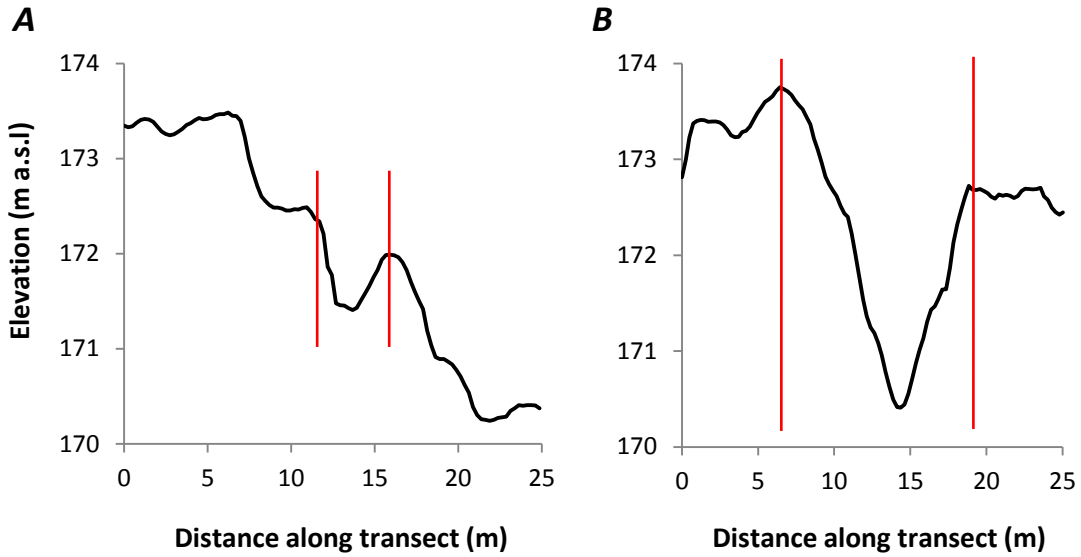


Figure 5.13 Vertically exaggerated transects A and B across the western coal tubway (referenced in figure 5.12) with the cross-sectional extent of the tubway highlighted. Distance along transect is always given in a west to east direction.

The tubway at the western edge of Chopwell Wood can be first identified in a County Series Ordnance Survey map from 1906 (EDINA, 2009) (figure 5.14). Given that it cannot be located on the previous edition, released in 1893, and that the coal drift mine at the eastern end (see section 5.2.2) is labelled as 'Old' in 1906 (figure 5.14), it can be reasonably assumed that this section of tubway was operational for an unknown period between 1893 and 1906. Although the contemporary topographic signature can only be distinguished between the red limits in figure 5.14, the original tubway extended both an additional 200 metres or so further west and 50 m east to the drift mine. The tubway was constructed to transport coal from the drift mine to the edge of the wood as this area was fully forested until very recently.

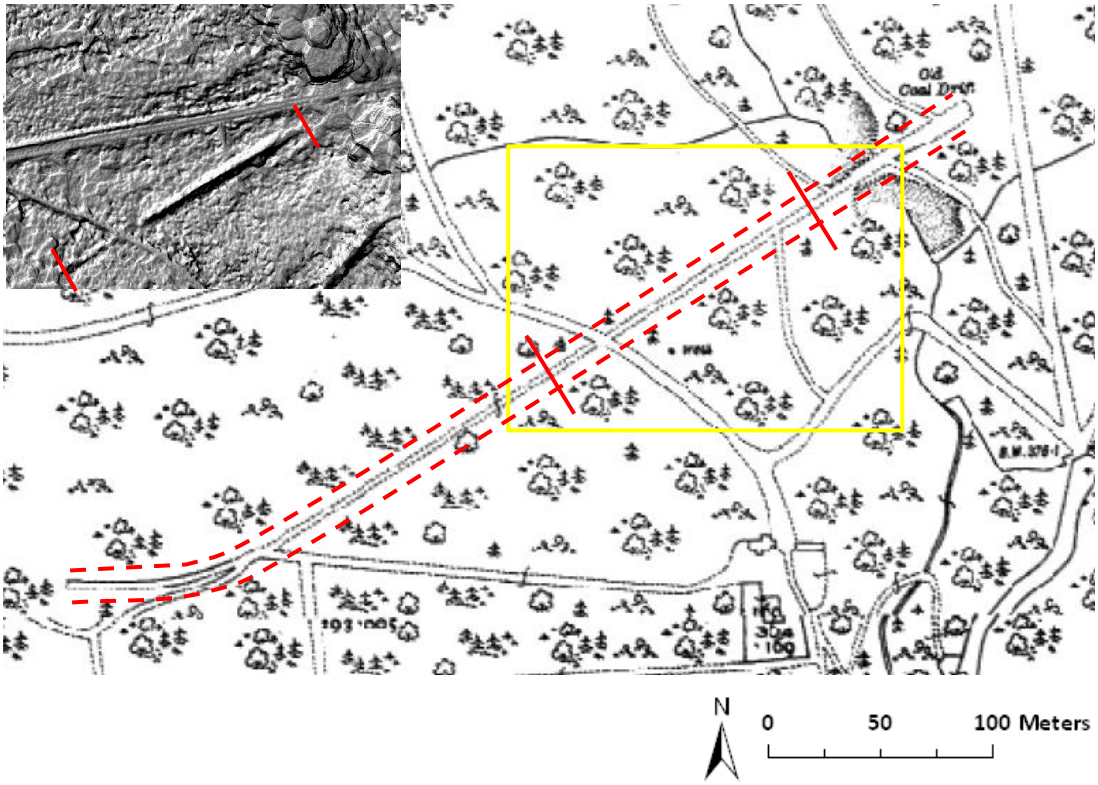


Figure 5.14 Historic 1:2500 County Series map, registered as in operation 1906-1936, with the full extent of the western coal tubway located between the two dashed red lines. The inset shaded LiDAR terrain model covers the area highlighted by the yellow box and the section of the tubway that can be clearly distinguished in the model is highlighted on both the historic map and terrain model by the full red lines. Note the coal drift mine `labelled on the map at the eastern end of the tubway (see section 5.2.2). Source of historic map: EDINA (2009).

5.2.2 Drift mine and hollows

Less well documented than the coal tubways is a drift mine at the eastern end of the tubway at the western edge of Chopwell Wood (figure 5.5), although it is established that the mine and tubway are associated (figure 5.14; table 5.1). A large, steep convex slope is observed in the contemporary topography at this location (figure 5.15 A) in addition to the brick foundations of industrial infrastructure, and a number of obscure hollows can also be identified in a line to the northeast of the slope (figure 5.15 B). While the convex slope is situated at the boundary between Japanese larch and Corsican pine stands, the hollows are covered exclusively by the pine, with moderately dense understorey cover over both (figure 5.15). Identifying the slope in the LiDAR terrain data may not prove a demanding task, given the size of the feature, and much like the tubway analysis, focus is placed on calculating how much detail can be resolved. However, the hollows are all < 2 m in width and < 1 m in depth, and the pine canopy and understorey provide significant obstruction to the laser, so it will be evaluated whether these features can even be extracted from the topography.

Figure 5.15 (A) Summer photograph of the convex slope at the drift mine under larch and pine forest. **(B)** Summer photograph of one of the hollows to the NE of the slope under pine forest.

A shaded terrain model, artificially illuminated from the southeast at an altitude of 45°, is presented in figure 5.16. Unrealistic edge effects just north of the feature are labelled (demarcated by the dotted yellow line in figure 5.16). These effects are a product of the interpolation procedure when there are few or no points in a large area, with the interpolated cell value at the edge of the true surface (along the dotted yellow line) spread to the boundary of the interpolation extent. No LiDAR terrain returns are located within this area due to a lack of pulse penetration and those that do penetrate do not then offer suitable terrain reflections. However this is not an issue as the area of interest (the extent of the features) is outside the modelled area. The uneven (bumpy) terrain surface across figure 5.16 is comparable to that observed at Test Sites 1 and 2 (see figure 5.2 in section 5.1) and is therefore explained as 'false' topography caused by understorey vegetation mistakenly classified as 'terrain' at the LiDAR filtering stage.

Transect A (figure 5.17) illustrates a profile through the features, with slope and hollows highlighted, while transect B constrains the width of the slope. It must first be noted that the lip of the slope is resolved successfully (figure 5.17), despite this situation being known to commonly cause problems to filtering algorithms, including the progressive densification algorithm (see section 1.3.3.2) (Sithole and Vosselman, 2004). This offers considerable promise for extracting similar discontinuities under thick conifer forest elsewhere. In contrast, the back wall of the convex slope is not modelled as accurately, with 'stepping' clear on transect A (figure 5.17). This is likely caused by either: (1) a lack of terrain returns over the back wall, with the subsequent stepping phenomenon a product of the interpolation process between distant returns of significantly different elevation (i.e. from the top and bottom of the slope), or (2) a laser footprint diameter too large to resolve the required gradient, causing 'time-walk' (see section 1.2.3). The hollows behind the convex slope are evident in a transect through the DTM (transect A in figure 5.17), but in order to detect similar features without prior knowledge (and therefore no recognition of where to position a transect, for example) at another location, more sophisticated visualisation would be necessary.

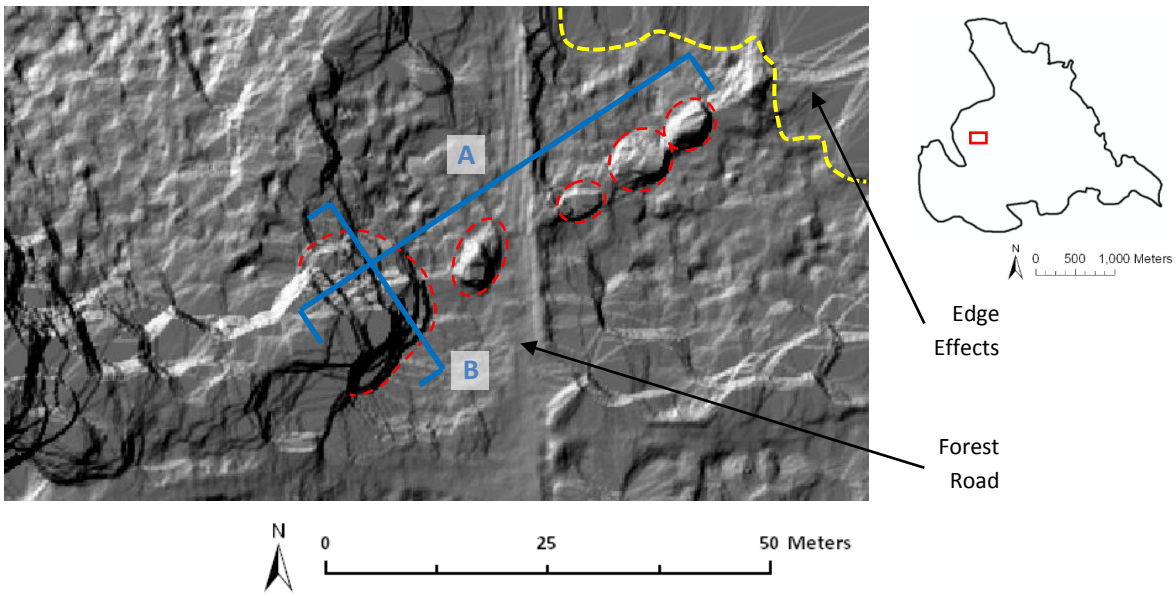


Figure 5.16 Shaded terrain model of the drift mine and hollows with illumination from the SE at 45°. The drift mine (westernmost feature) and hollows are highlighted by dashed red lines and lines A and B indicate the locations of transects in figure 5.17. The dashed yellow line marks edge effects in the terrain model.

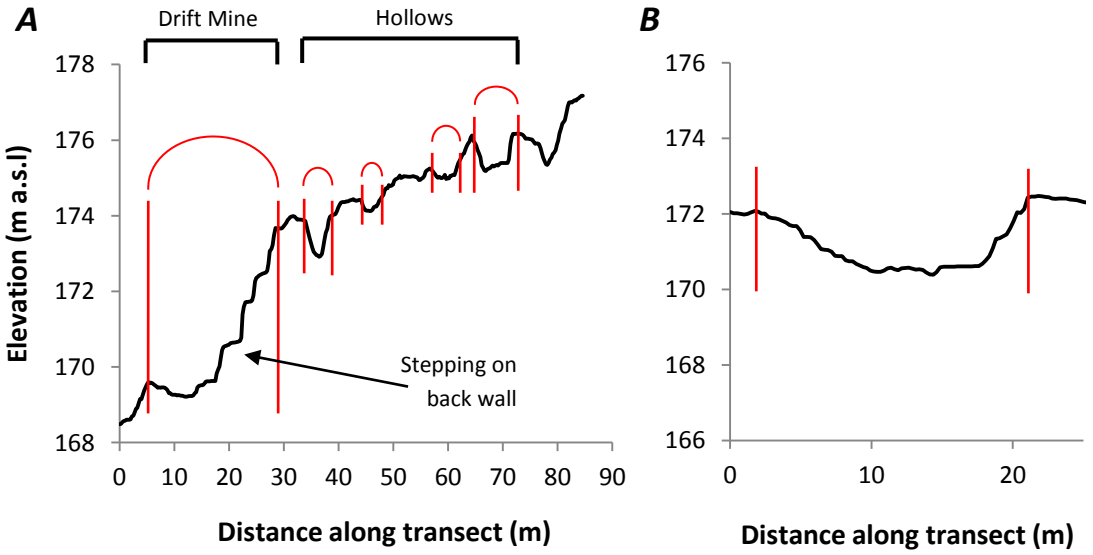


Figure 5.17 Vertically exaggerated transects A and B through and across the drift mine and hollows (referenced in figure 5.16), highlighted by pairs of dashed red lines. Distance along transect A is in a west to east direction and along transect B is in a north to south direction.

Given that it is difficult to detect and resolve the hollows (of < 2 m width and < 1 depth) under pine forest in standard shaded terrain models, further visualisation analysis is employed to attempt to emphasize the depressions against the local topography. With the aim of testing its efficacy for resolving particularly subtle archaeological features, principal components analysis (PCA), a visualisation technique for combining information from a number of artificially illuminated hillshade models (see section 3.5.2), is utilised over the drift mine and hollows. Shaded terrain models of the area from sixteen angles (at regular 22.5° intervals in azimuth and constant 45° altitude) are combined and transformed to a new set of sixteen uncorrelated variables. Histograms of illumination values for all sixteen original models are verified as normal distributions – a pre-requisite for PCA. Table 5.2 gives directional loadings for the first five components, with values > 0.3 (indicating high model intercorrelation) in bold, and the percentage of variance explained by each component. The first three components explain approximately 93 % of the variation between hillshade models (topography), with components 1 and 2 associated with northeast/southwest and northwest/southeast illumination directions, respectively (note high loadings in these directions). The importance of component 3 is advocated by the pioneers of this applied technique (Devereux et al., 2008) because it exhibits almost equal loading in all 16 directions (table 5.2), suggesting that its information content might be independent of illumination direction, with implications for detecting features without the problems of directional shading.

A single hill-shaded model from the southeast is presented in figure 5.18 A for comparison against a false colour composite image of the first three principal components figure 5.18 B and also component 3 independently figure 5.18 C. The 'stripe' effect across the top of the principal component images is an artefact of the intercorrelation procedure in the PCA algorithm, but it does not coincide with the features under consideration and is thus inconsequential. It is clear from the outset that component 3 offers high definition for feature detection at this scale, with the hollows particularly distinct and the lip of the convex slope obvious, but its interpretation is somewhat complicated by irregular shadowing and definition

of the (erroneous) 'stepping' on the back wall (figure 5.18 C). The true location of the lip is highlighted in figure 5.18 A. This technique offers promise for identifying features of less than 2 m width and 1 m depth in areas of forest, like the Corsican pine, where LiDAR-derived terrain data includes unfiltered returns from low vegetation.

Component		1	2	3	4	5
Illumination direction from North (0°)	**0**	-0.23	-0.21	**0.30**	0.21	**0.30**
	22.5	**-0.31**	-0.10	0.29	0.10	0.24
	45	**-0.35**	0.06	**0.30**	**-0.77**	0.20
	67.5	**-0.32**	0.18	0.25	-0.06	-0.25
	90	-0.26	0.28	0.20	0.12	-0.25
	112.5	-0.16	**0.36**	0.16	0.15	-0.18
	135	-0.03	**0.39**	0.13	0.14	-0.08
	157.5	0.10	**0.38**	0.13	0.11	0.04
	180	0.22	**0.32**	0.15	0.06	0.15
	202.5	**0.31**	0.22	0.19	0.01	0.23
	225	**0.35**	0.11	0.24	-0.05	0.25
	247.5	**0.34**	-0.01	0.29	-0.09	0.21
	270	0.29	**-0.45**	**0.31**	**-0.32**	**-0.59**
	292.5	0.18	-0.24	**0.31**	0.01	-0.21
	315	0.04	-0.29	**0.32**	**0.33**	-0.19
	337.5	-0.10	-0.28	**0.30**	0.21	0.17
	%	36.58	31.21	25.77	1.73	1.45
	Cum. %	36.58	67.79	93.56	95.28	96.73

Table 5.2 Loadings for the first five principal components of the sixteen illuminated models of the drift mine and hollows and the percentage of variance explained by each.

The location of the drift mine is given in figure 5.14 in section 5.2.1 (labelled as `Old Coal Drift`) and the large convex slope likely represents the entry point, with the tubway formerly moving underground at this point. A historic Ordnance Survey map from 1924 suggests that the lateral extent of the slope has fallen substantially, with the extent of the feature in 1924 highlighted

in blue on figure 5.19 and the extent of the contemporary feature highlighted in red. The origin of the hollows is less apparent, but their trend exactly parallel to, and extending beyond, the coal tubway (see figure 5.14 in section 5.2.1), indicates that they might be sites of subsidence down into a possible drift tunnel below. Performing PCA in other locations, with a view to locating similar lines of hollows, may facilitate the identification of further drift mines currently lost beneath the forest canopy.

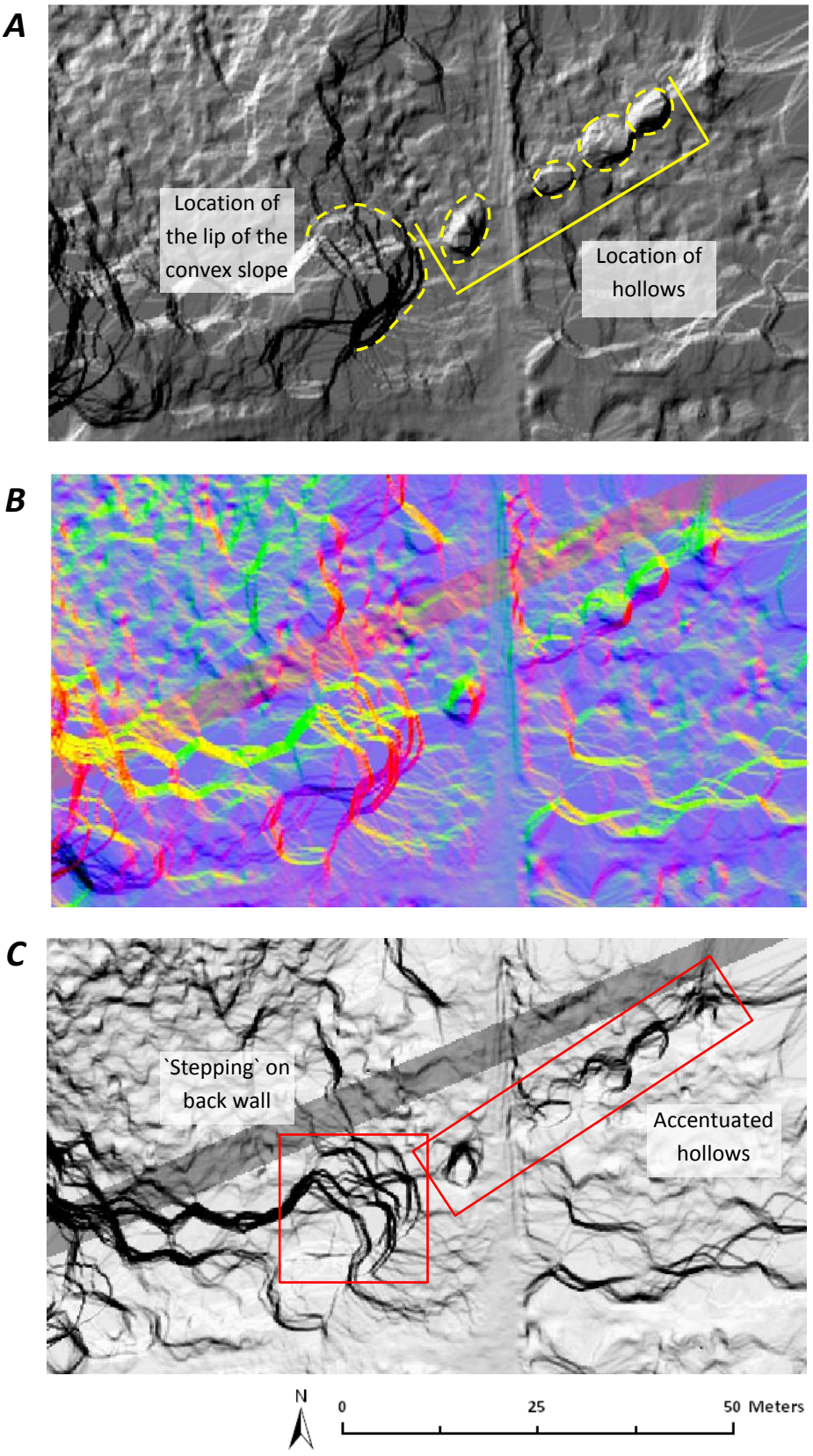


Figure 5.18 (A) Shaded terrain model of the drift mine and hollows (highlighted by the dashed yellow lines), illuminated from the SE at 45°. **(B)** False colour composite of the first three principal components of the sixteen band principal components image, where: component 1 is shaded red, 2 is shaded green and 3 is shaded blue. **(C)** Shaded relief image of component 3.

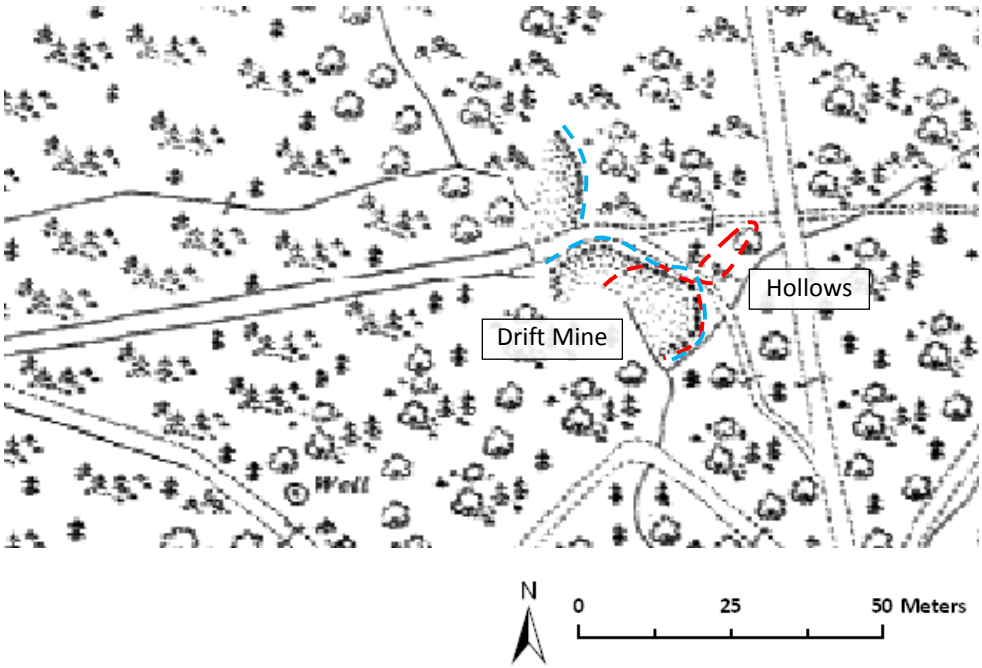


Figure 5.19 Historic 1:2500 County Series map, registered as in operation between 1924-1949, with the convex slope of the drift mine highlighted by the dashed blue line and the contemporary topographic signature of both the drift mine and hollows highlighted by the dashed red line. Note the coal tubway is not marked by this edition, although the new forest road has been laid (see figure 5.14). Source: EDINA (2009).

5.2.3 Avenue and pathway

It has already been suggested that the wide, linear feature in the topography at Test Site 5 may have been an avenue leading to a former house within the forest pre-dating the 19th century (table 5.1) (see section 5.1) (FoCW, 2009). A photograph down this 'avenue' is given in figure 5.20 A, with the edges (< 50 cm high ridges) especially prominent. Furthermore, a path constructed with what appears to be building material (i.e. brick, tiles etc.) is evident nearby (figure 5.20 B), and the manner in which the avenue and path appear to intersect indicates that there may have been several routes to this house. Both features are currently used as forest paths and are marked on the Ordnance Survey map as such (figure 5.5), but it is natural that even following demolition of the house, the same tracks have continued to provide access through this area of the forest.

An artificially illuminated shaded terrain model of the area surrounding the avenue is presented in figure 5.21. A stream at the eastern end of the feature is marked on the model and can be located in figure 5.5. A vertically exaggerated profile across the feature (figure 5.22) emphasizes the avenue platform raised from the surrounding terrain surface, and ridges along either side of approximately 10 cm height on the south side and 30 cm on the north side are identified. The ridges are almost impossible to detect in the standard terrain model (figure 5.21) so a vertically exaggerated shaded terrain model is generated of the same area with elevation exaggerated by ten times. Part of this model is presented in figure 5.23 and although it is still difficult to detect the 10 cm high southern edge of the avenue, the raised surface of the 30 cm high northern ridge is heavily accentuated against the adjacent topography. The canopy cover above this feature is beech, which offers reasonable laser penetration and relatively high terrain point accuracy (with respect to coniferous forest) (see table 4.6 in section 4.4.1), but there is noticeable clustering of LiDAR terrain returns under this type of canopy (see section 5.1). This is the most plausible explanation for the heavily undulating (rather than flat) modelled terrain surface over the avenue ridges, noticeable in figure 5.23.

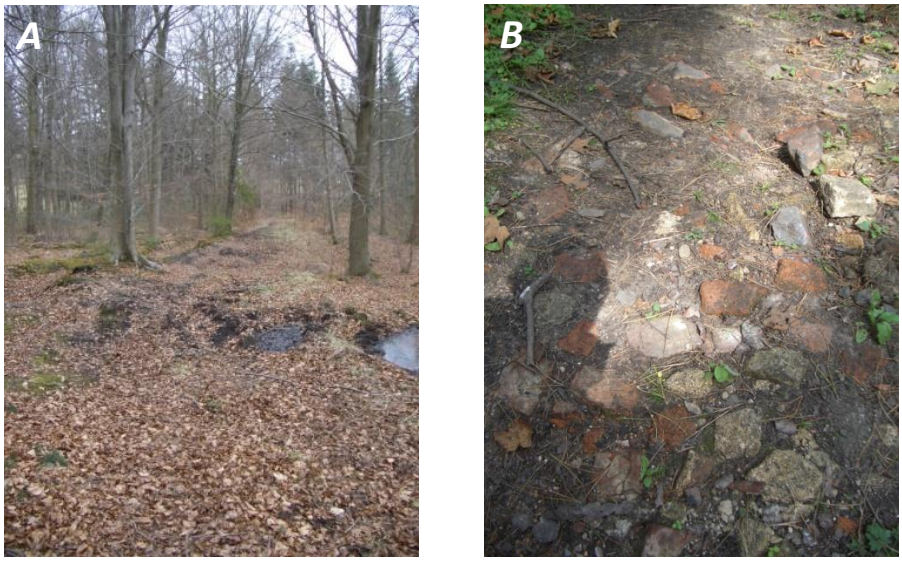


Figure 5.20 (A) Winter photograph from the west end of the avenue. **(B)** Close-up photograph of the pathway with brick and construction materials clear.

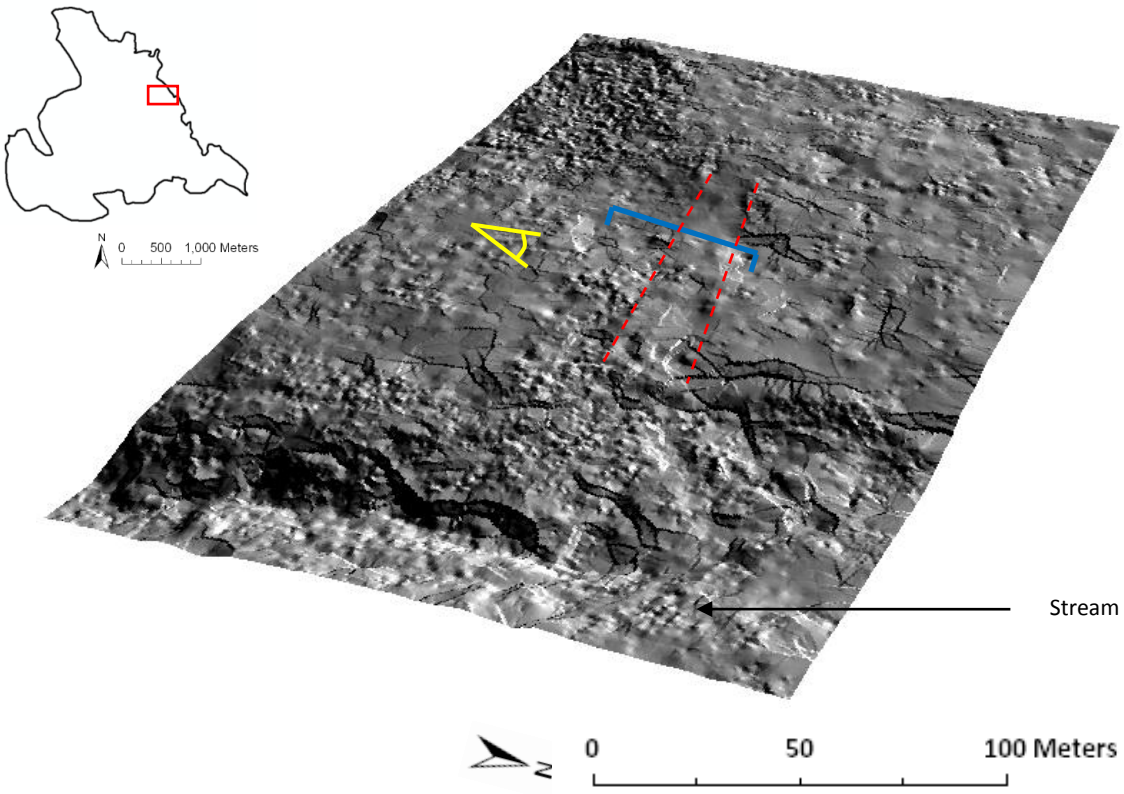


Figure 5.21 Shaded terrain model of the avenue with illumination from the NW at 45°. The lateral extent of the avenue is highlighted by the red lines. The blue line indicates the location of the transect in figure 5.22. The yellow symbol indicates the viewpoint of figure 5.23.

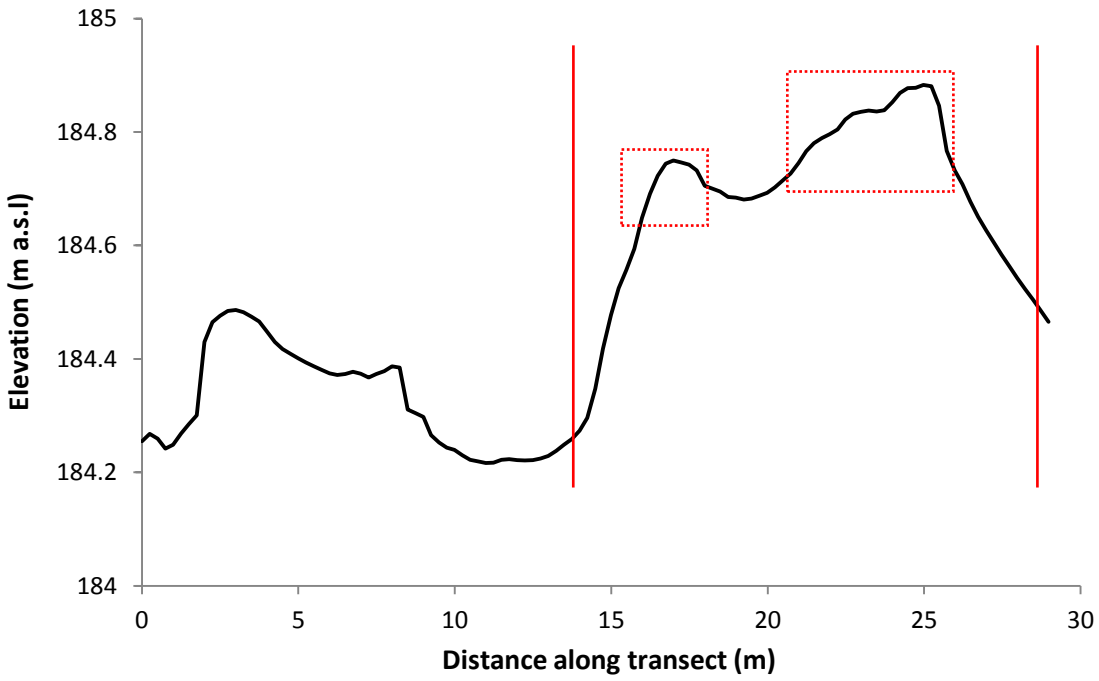


Figure 5.22 Vertically exaggerated transect across the avenue (blue line in figure 5.21) with the width of the major platform marked by the two red lines and the two ridges highlighted by the dotted red boxes. Distance along transect is in a south to north direction.

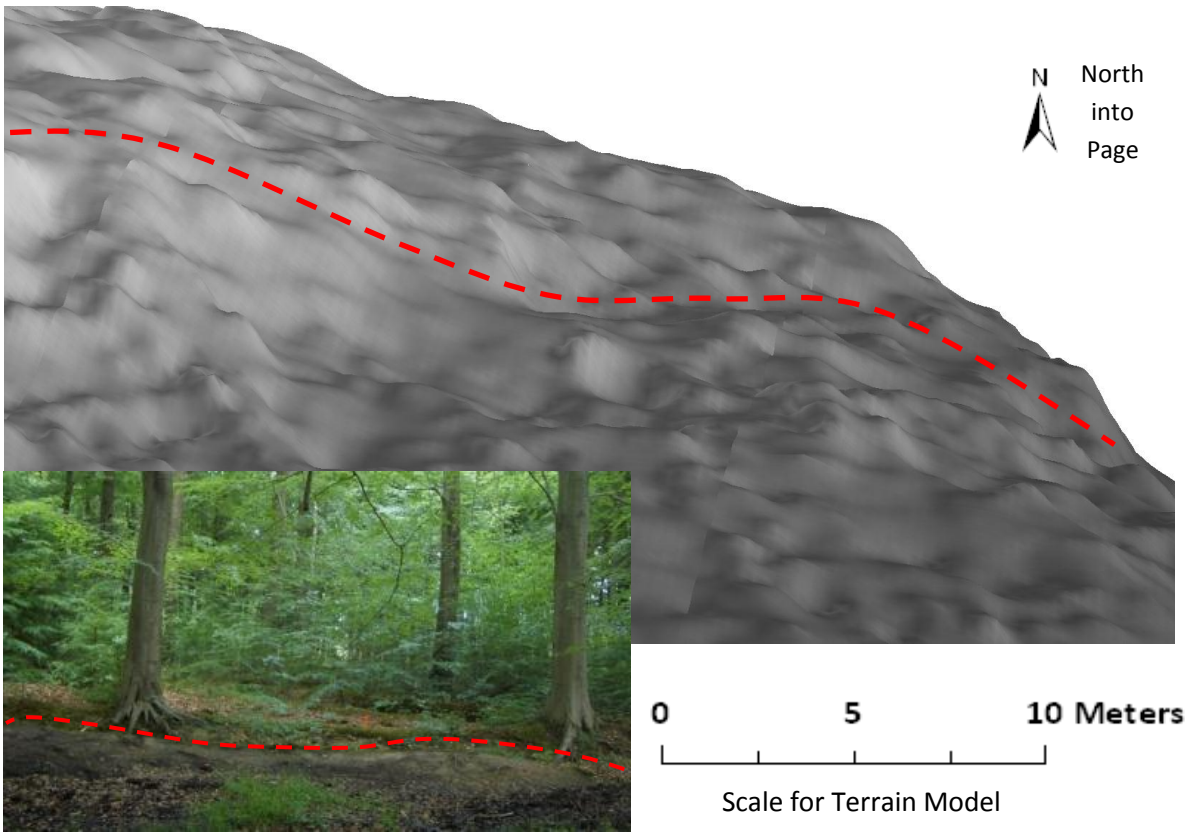


Figure 5.23 Shaded terrain model of the avenue, vertically exaggerated by 5 times and illuminated from the northwest. The same section of the northern (prominent) ridge of the avenue is highlighted by the dashed red lines in both the terrain model and inset photograph. The location of the viewpoint for both the model and photograph is given by the yellow symbol in figure 5.21.

The pathway is a much longer archaeological feature than the avenue and is covered by both beech and Corsican pine stands. The subtlety of the feature, measured at less than 30 cm depth in the field, and widespread 'false' topography in the terrain model (figure 5.24), characteristic of LiDAR data under pine and caused by poor filtering over dense low vegetation cover (noted in sections 4.4.1 and 5.1), mean that only a single, small section of the pathway can be identified in the standard hillshade model (which is illuminated from the northwest at 45° altitude; figure 5.24). Exaggerating the model vertically would only accentuate these disruptive features further and likewise principal components analysis would do the same.

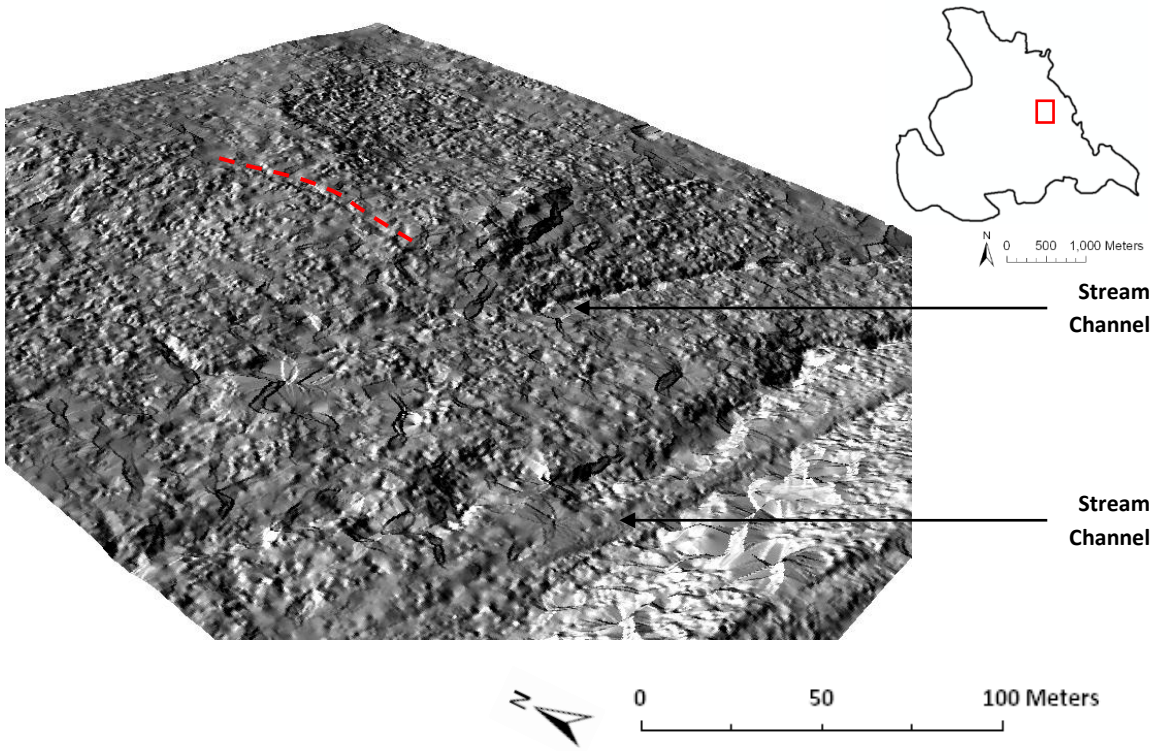


Figure 5.24 Shaded terrain model of the area surrounding the pathway with illumination from the NW at 45°. The small section of the pathway that can be detected is highlighted by the dashed red line.

The Local Relief method (Hesse, 2010) enhances the visibility of medium-scale (> 1 m diameter), shallow topographic features, irrespective of the illumination angle, by creating a DTM 'purged' of variable landforms from the original LiDAR terrain data. This technique was designed for archaeological prospection and appears to be ideal for extracting more information from the terrain model at the pathway. Three of the steps are presented in figure 5.25, with B illustrating a difference map obtained after performing low-pass filtering (with a 3 x 3 kernel) over the original DTM and calculating the difference between the two terrain models. Contours at the zero-metre level are extracted from the difference map in figure 5.25 C and a purged terrain model is interpolated from these contours (D) (see Hesse, 2010 for details). Comparison between the two shaded surface models (figure 5.25 A and D), both computed with identical illumination parameters, demonstrates the uniform removal of small, disruptive features from the topography and consequently a longer section of the pathway can be identified more easily in the purged model .

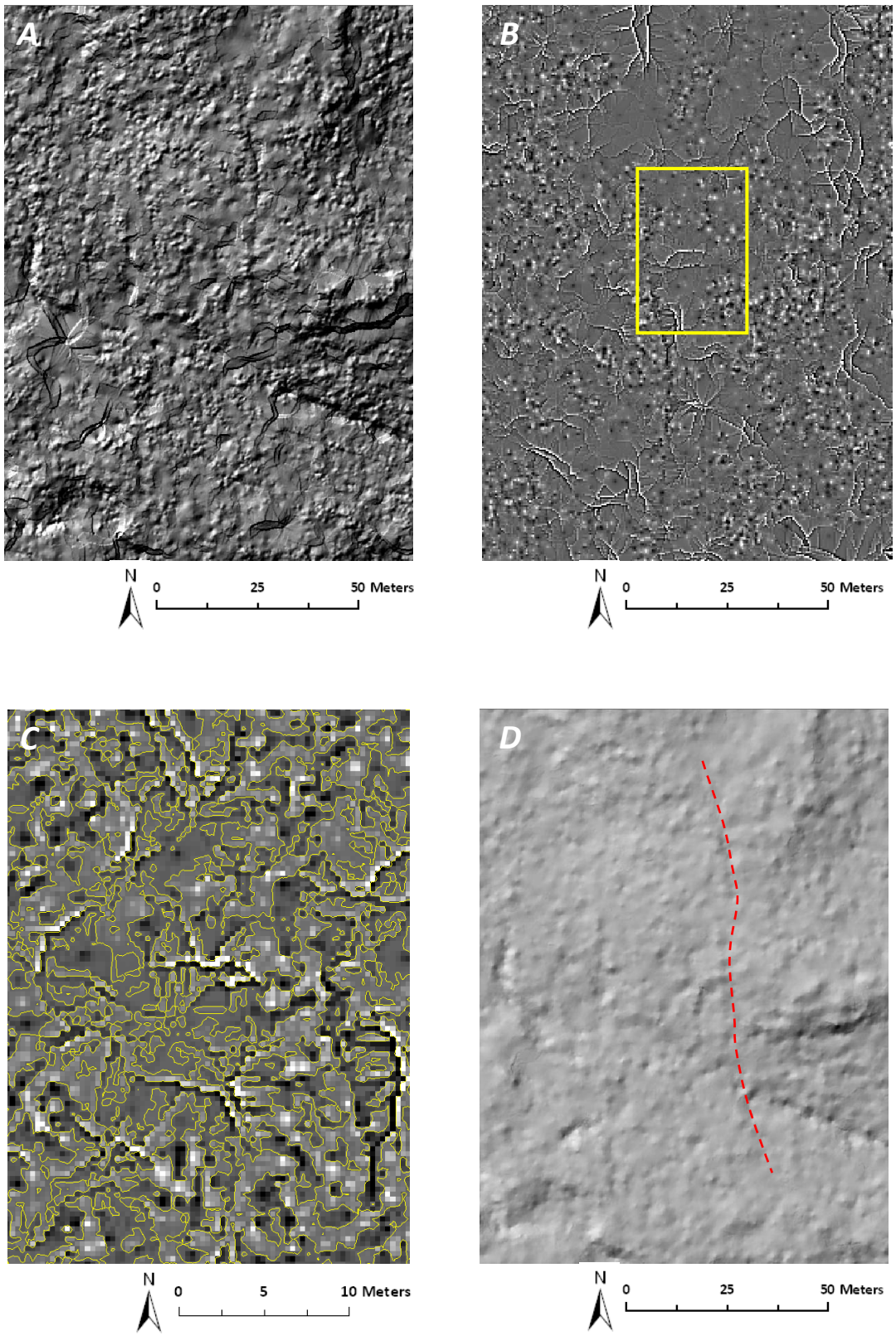


Figure 5.25 (A) Shaded terrain model of the pathway with illumination from the NW at 45°. **(B)** Difference map of A representing local relief variations. **(C)** Extraction of zero metre contour lines from the difference map (inside the yellow box in B). **(D)** Enhanced local relief model of the same area as A, with the same illumination parameters, following 'purging' of small-scale features, with the more obvious pathway highlighted by the dashed red line.

Assuming that both the avenue and pathway are associated with the former position of this mysterious 'Dene House', it can be hypothesised that the building could have been located close to the intersection of the two linear features. Figure 5.26 illustrates a local relief model generated from purged avenue and pathway DTMs, and artificially illuminated from the northwest at an altitude of 45°. Despite the features not physically intersecting, it is assumed that neither route can deviate by > 20° from their expected course, based on their contemporary position/direction and the hypothesis that the two features are linked. Hence, if a house was located here prior to 1800, it is most likely that it was situated within the hashed red zone marked on figure 5.26.

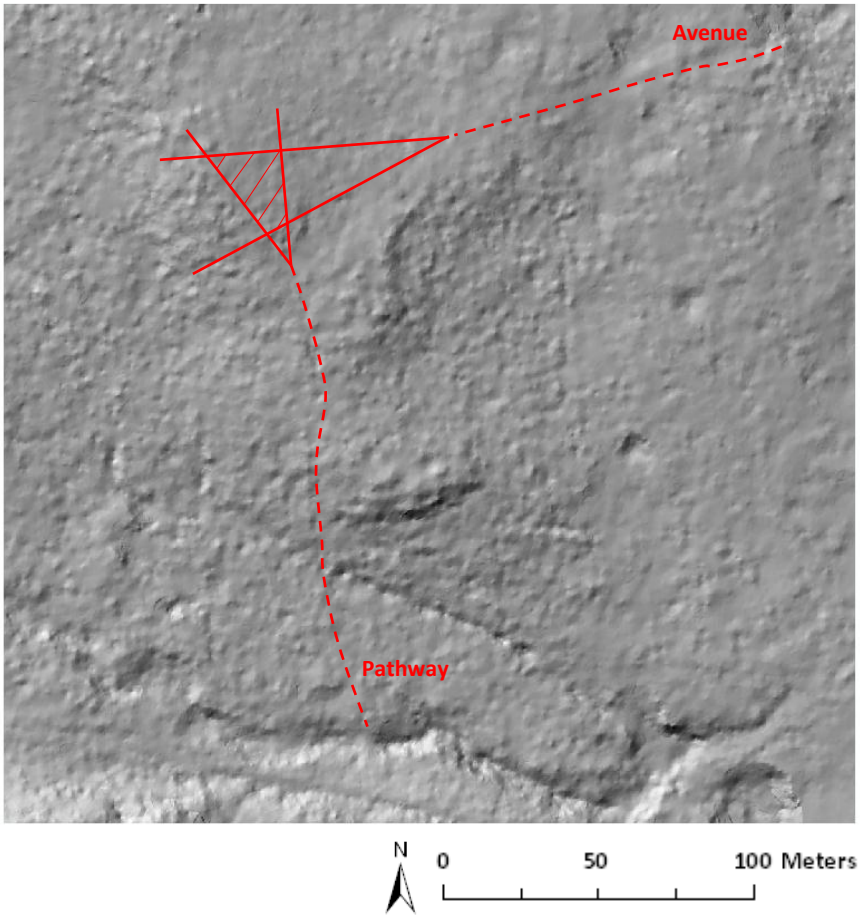


Figure 5.26 Local relief model (LRM) of the pathway and avenue with illumination from the NW at 45°. A building (Dene House?) at the intersection between these routes was likely situated inside the zone marked by the hashed red lines.

5.2.4 Rubble building structures and man-made channel

The limit of airborne LiDAR as a technique for archaeological prospection is demonstrated by a relative inability to detect the piles of building material and man-made channel beneath the young Douglas fir (and partial mixed deciduous) canopy at field Test Site 8 (figure 5.5). These features are thought to be the remains of a building that stood at this location at some point prior to 1850 (L. Searle, personal communication 2009). It was noted in section 4.4.4 that LiDAR terrain point density was higher under the older, taller, 25 year old fir canopy and negligible understorey vegetation at Site 7 than under the 19 year old canopy at Site 8 (19 and 3 points per m^2, respectively). Consequently the interpolated terrain model is based on fewer points at Site 8 and is less detailed and less representative of the true terrain than at Site 7 (see figure 5.4 in section 5.1). For this reason, the probability of detecting, let alone resolving, archaeological features such as the piles of building material and man-made channel at Site 8, of approximately 2 to 3 m width and 1 to 2 m depth (figure 5.27), is much lower.

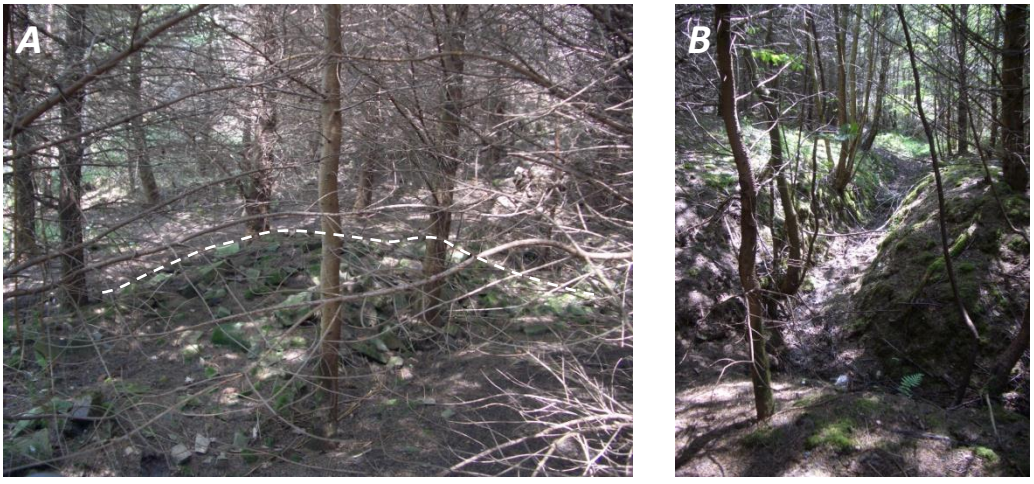


Figure 5.27 (A) Winter photograph of one of the building rubble piles at test site 8 with the shape of the pile highlighted. **(B)** Winter photograph from the northern end of the man-made channel at test site 8 (note the fork). Both features lie under Douglas fir canopy.

A shaded terrain model of the area surrounding test site 8 is presented in figure 5.28, artificially illuminated from the northeast at an angle of 45°. Control points along the edges of the archaeological features, acquired in the field with the total station (and previously utilised during filter parameterisation in section 4.2), are superimposed onto the image and emphasize that neither the piles of building rubble or man-made channel are visible in the standard shaded terrain model (figure 5.28). In addition to sparse LiDAR terrain point density limiting feature definition, the whole area of interest is situated on the side of a significant slope, with several of the piles of rubble located on an even steeper backslope (figure 5.28). This has important implications for both the accuracy of returns (the reason of which has previously been alluded to) and the process by which terrain is interpolated between relatively remote returns on a slope. 'Stepping' has already been noticed in a terrain model over the convex slope of the drift mine (see section 5.2.2) and can also be observed systematically over the shaded surface in figure 5.28, including over a majority of the archaeological features. Where features are smaller than the typical size of the largest artefact, i.e. < 2 m depth, they cannot realistically be resolved or identified at all.

One feature that can be detected is the lower part of the man-made channel (figure 5.27 B). A profile acquired along the base of the channel (which covers the western rather than north pointing fork; figure 5.28) is compared with field-derived total station control points in figure 5.29 and suggests that the airborne LiDAR struggles to extract the full channel depth. It could be argued, as with previous scenarios, that overestimation of the true channel base elevation by 81 cm (with a standard deviation of 28 cm) is caused by dense low vegetation; but, following examination in summer (figure 5.27 A), this is found not to be the case. It is more likely that it is beyond the capability of the airborne LiDAR to resolve subtle topographic earthworks, like the rounded piles of rubble, under young (19-year old) fir and likely also spruce and hemlock species (see table 4.11 in section 4.4.4). In this scenario, the airborne LiDAR data can, at best, be used to infrequently detect features like the channel.

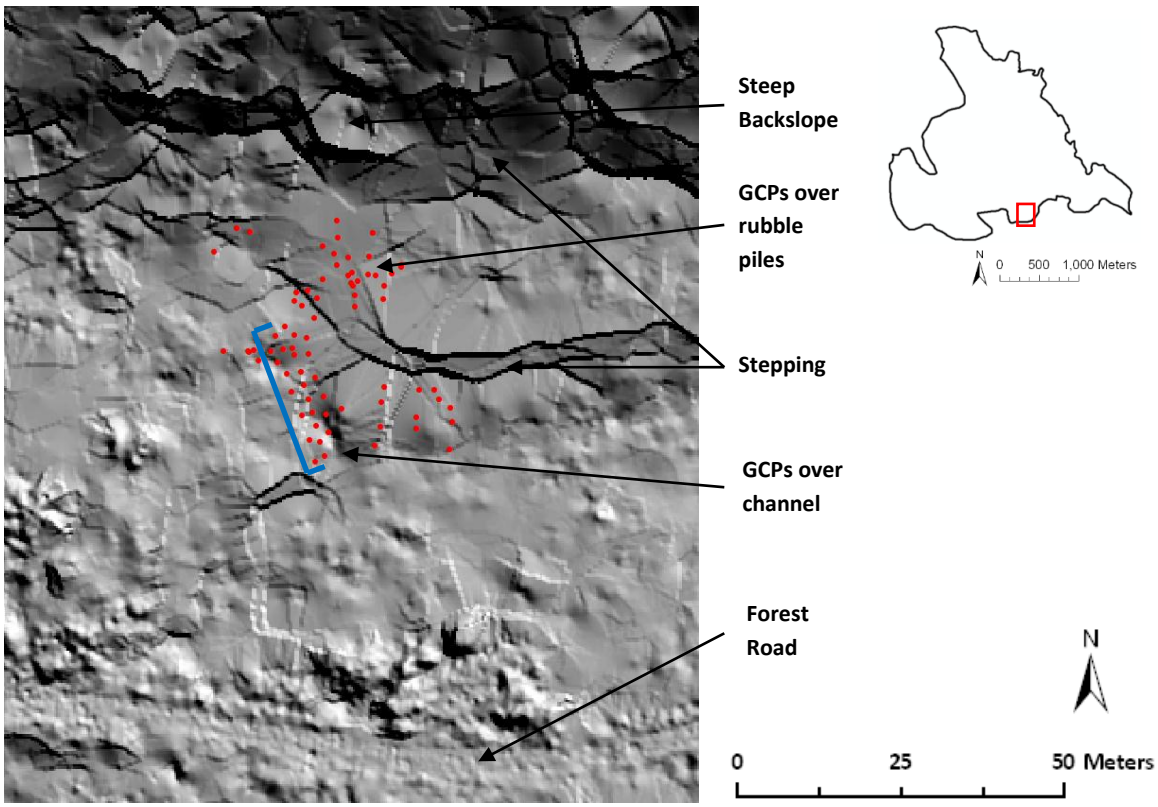


Figure 5.28 Shaded terrain model of the area surrounding the rubble structures and man-made channel, with illumination from the NE at 45°. The blue line indicates the location of a profile along the floor of the channel shown in figure 5.17. Ground control points (GCPs) over the features are illustrated as red dots.

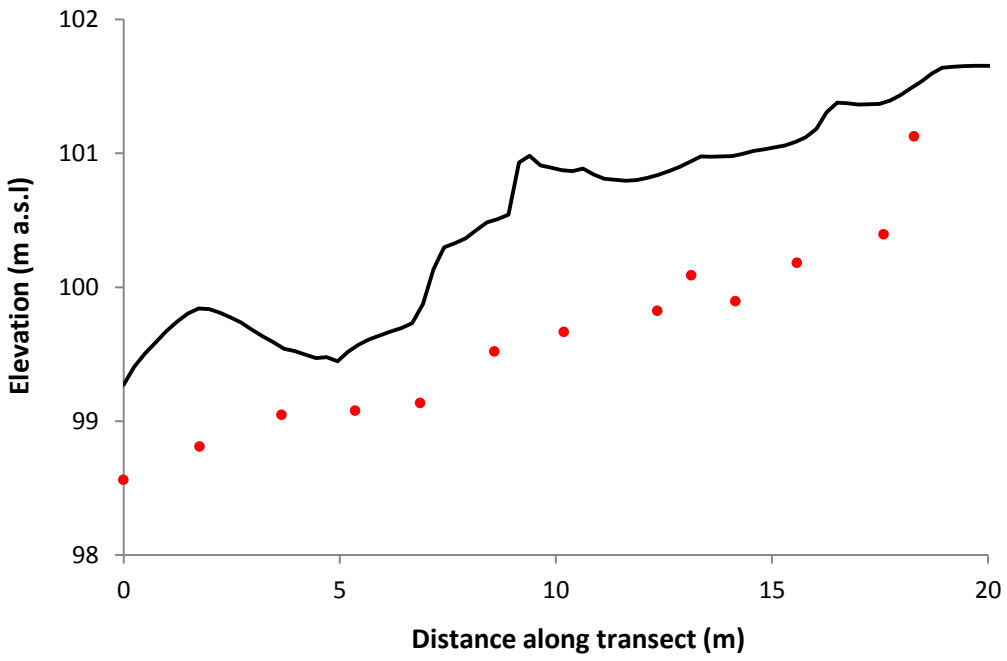


Figure 5.29 LiDAR profile along the floor of the channel (black line), from the southernmost end up along the western fork, compared with ground control points (red points). The extent of the profile is given by the blue line in figure 5.28.

In summary, this section (5.2) has emphasised both the capabilities and limitations of airborne laser scanning as a technique for sub-forest canopy archaeological feature prospection. It has answered the question of whether minor differences in LiDAR point density and accuracy have a significant effect on the detail of a terrain model and subsequently on the user's ability to detect archaeological features. Whilst features of less than one metre height and width can be detected under beech and pine canopy at the avenue, pathway and drift mine, features of comparable subtlety, i.e. piles of building rubble and most of a man-made channel, cannot be detected under dense fir canopy.

The viability of terrain modelling and associated archaeological feature prospection using LiDAR will be discussed in greater detail in the following chapter.

Chapter 6. Discussion

6.1 Evaluation of Laser Scanning as a Tool for Sub-canopy Terrain Modelling

6.1.1 Reasons for variations in LiDAR terrain point accuracy and density under different forest species

Variations in the accuracy and density of LiDAR terrain points under different types of forest cover were demonstrated in sections 4.4.1 and 4.4.4. In this section, the possible and probable reasons for these variations, and the implications they have for predicting terrain modelling (and consequently also archaeological surveying) viability, are discussed.

It has been suggested, in numerous studies (e.g. Clark et al., 2004; Hodgson and Bresnahan, 2004), that canopy cover is highly influential over the accuracy of LiDAR terrain measurements. Canopy of greater cover and density is, based on previous research, expected to increase both the systematic and random error exhibited by terrain points. By this logic, mean standard offset, standard deviation and RMSE ought to rise with increasing stem height, canopy depth, stem diameter (DBH) and stem density. However, it is clear from the correlation analysis in section 4.4.1 (table 4.8) that only mean DBH demonstrated a consistent positive correlation with LiDAR bias and precision. The relationships between stem density and bias and precision were poor and, surprisingly, all negative (figure 4.13 A), with little change if only trees of primary species were examined (figure 4.13 B). The following explanations are offered: (1) stem height, canopy depth and stem density are not appropriate proxies for stand-scale canopy structure; (2) allometric assumptions are not valid for all tree species; (3) an insufficient number of plots are analysed to recognize the true relationships; (4) the relationships are non-linear and/or (5) that the filtering procedure was so effective that the quality of accepted terrain points will not necessarily decline with increasing canopy cover.

It is clear from the qualitative assessment at the start of section 4.4.1 that the final argument is invalid, given that Type II filtering errors were numerous at Sites 1 and 2 (figure 4.8). The

fourth, and to some extent, third arguments can be discounted for stem density because the trends between variables were strong enough that their negative characteristics were unmistakable (figure 4.13). Although weak, the positive relationships between stem height and canopy depth and LiDAR bias suggest that stem height and canopy depth describe stand-scale canopy structure more closely than stem density. Tree diameter (at breast height) appeared to predict LiDAR bias and precision well (*r* of 0.40, 0.36 and 0.48 for mean offset, standard deviation and RMSE, respectively; table 4.8).

But despite the relatively strong relationships between tree DBH and LiDAR bias and precision, the accuracy of LiDAR-derived terrain elevation appears to be controlled primarily by the understorey vegetation. The weak, and often negative, relationships between tree canopy structural variables, like canopy depth and stem density for example, terrain roughness and gradient, and LiDAR bias and precision obtained in the correlation analysis (table 4.8) were either coincidental or were, more likely, a by-product of the dominance of understorey vegetation as a controlling factor. For instance, Test Site 1 was 95 % covered by (on average) 1.2 m high fern and bracken, and LiDAR at this site demonstrated mean offset of 21 cm and standard deviation of 20 cm compared to the true terrain. In contrast, Test Site 7 had negligible understorey and demonstrated only 3 and 16 cm offset and standard deviation, respectively. This is a pronounced difference and, as such, it matters little that the terrain was much steeper and stem density was higher at Site 7 than Site 1, because the understorey vegetation was the principal controlling factor over the quality of LiDAR-derived terrain measurements. LiDAR terrain point density at the test sites also appeared to be controlled by the relative rate of laser penetration through understorey vegetation, and to some extent through the tree canopy, with strong least squares relationships determined between point density and understorey volume (*r* of 0.81; table 4.8 in section 4.4.1), and tree diameter (*r* of 0.52 and 0.64, for the entire plot and for only primary species, respectively).

Both the accuracy and density of LiDAR terrain points were higher under deciduous than coniferous forest, even in deciduous leaf-on conditions (table 4.6 in section 4.4.1 and table 4.11 in section 4.4.4). This suggests that the quality of terrain modelled with LiDAR data should, in general, be greater under deciduous canopy cover. However, either terrain point accuracy or density can have a dominant influence on the quality of a terrain model, something that is emphasised in areas of coniferous forest in Chopwell Wood that have been subject to thinning operations.

It has been assumed until now that thinned forest, with greater laser canopy penetration, is ideal for sub-canopy LiDAR survey (e.g. Devereux et al., 2005). Following this assumption, the thinned Corsican pine cover in Chopwell Wood ought to be highly conducive to sub-canopy survey, as should other thinned coniferous species like the Norway spruce (which offers a mean penetration rate of over 20%; table 4.11). But, given that thick summer understorey was observed throughout the Corsican pine test sites in Chopwell Wood, there is a convincing argument that thinned and, as such, relatively low density coniferous forest (compared to that observed in the fir – compare the photos in figures 4.8 and 4.11), allows the sub-canopy vegetation to thrive. This indirectly increases Type II filtering errors and reduces the accuracy of LiDAR-derived terrain elevation. The assumption that thinned forest is ideal for sub-canopy survey is not necessarily the case with summer-acquired leaf-on LiDAR data.

Significant point bias and low accuracy are as degrading to a digital terrain model as low point density (Pfeifer and Mandlburger, 2009), if not more so, and so have considerable influence on the interpretation of topographic features. For example, a feature with relief < 20 cm in the topography at (i) in transect 1A in figure 4.8 (section 4.1.1) cannot be resolved with the LiDAR data. Relatively extreme over-estimation of true terrain elevation and low precision at these sites indicates that, in spite of the high spatial measurement density, thinned Corsican pine in summer is not especially amenable to LiDAR-based sub-canopy terrain survey. In contrast, little summer understorey vegetation was observed under the thinned Norway spruce (during investigation in the field). This suggests that the accuracy of LiDAR terrain points under the

Norway spruce could be expected to be similar to the accuracy of points recorded under the Douglas fir at Test Sites 7 and 8 in Chopwell Wood, which also contained little or no understorey. High accuracy measurements of the terrain combined with a greater than 20% canopy penetration rate make thinned Norway spruce particularly amenable to sub-canopy terrain survey. The contrasting examples from the Corsican pine and Norway spruce in Chopwell Wood indicate that the role of forest management practices, for example thinning operations, on LiDAR-based terrain surveying viability during peak summer understorey conditions is not clear cut. However, it is clear that low terrain point accuracy can result in poor quality modelled terrain in spite of high point density, and vice versa, with either point accuracy or density acting as a limiting factor for conducting sub-canopy survey.

6.1.2 Evaluation against published LiDAR terrain accuracy assessments

Airborne LiDAR-derived terrain model accuracy has not been comprehensively investigated in any single published study to date, which is surprising given that vegetation-filtered terrain models are commonly utilised in numerous commercial and scientific applications. In fact, far more attention has been paid to assessing the relative success of vegetation-removal algorithms against each other, than to identifying the limit of airborne LiDAR as a technique for sub-canopy terrain modelling. That said, a number of groups have conducted tests on LiDAR products over the last decade to determine accuracy against field-acquired reference data of true terrain elevation. This section will review all relevant published studies and compare results to those obtained from this study.

Eight studies contributing significant data are identified, but the lack of investigation into LiDAR accuracy assessment has led to significant inconsistency in approach. Typically, reference point elevation is evaluated against either proximate LiDAR terrain point elevation or the elevation of an interpolated surface through LiDAR terrain points (see section 1.3.6). Where testing point-to-point examines the true difference between LiDAR and reference

(Hodgson and Bresnahan, 2004), testing surface to reference emphasizes the critical role of the final modelled surface, which is usually the required product, and is, as such, the more *applicable* assessment (Hyyppä et al., 2005). Given that only eight suitable studies can be identified, tests performed under similar conditions are uncommon. Published studies include data collected under a variety of conditions, such as leaf-on or –off, different point densities, over different tree species and mixtures, so all relevant studies are included. Likewise, tests have been conducted in areas with different forest management strategies, so care must be taken when comparing, for example, 'equivalent' results from coniferous or deciduous forest.

Results for the eight studies are given in table 6.1 and offer varying quantities and types of data. Another consequence of the lack of cohesion between investigations is the range of statistical approaches, with each study presenting results for at least one of the following: (1) mean ('signed') offset or (2) mean absolute offset, which describes the *bias* of LiDAR against reference terrain elevation; (3) standard deviation, which describes the *precision* of LiDAR against reference; and (4) root mean square error (RMSE), which describes the *bias* precision. In this work, results are offered for all four measures, under both deciduous and coniferous forest (see table 4.6 in section 4.4.1). Comparable tests taken from published studies are grouped by the categories: deciduous forest, conifer forest and shrub/brush. Mixed conifer and deciduous forest is included in the deciduous category given the dominant role of broadleaf trees on laser canopy penetration (Su, 2004; Su and Bork, 2006).

Studies 2 and 8 and this work describe very similar 10 to 20 cm over-estimation by LiDAR of the true terrain surface elevation under deciduous forest (table 6.1). Both Hodgson et al. (2003) (study 2) (study 8), who conducted their research in mixed deciduous forest, and Su and Bork (2006), who conducted theirs in aspen forest, chose to use leaf-on conditions – comparable to the data collected from Chopwell Wood. So, if it assumed that these data are describing a genuine pattern, minor overestimation of the terrain surface can be expected under leaf-on deciduous canopy. The overestimation is explained in each paper by multi-story vegetation affecting the LiDAR filtering procedure (Hodgson et al., 2003; see also Pfeifer and

Mandlburger, 2009; Pfeifer et al., 2004) and the broad-shaped and horizontally oriented leaves of deciduous trees (and typical understorey plants) causing attenuation of the laser beam (Su, 2004; Su and Bork, 2006). Ni-Meister et al. (2001) suggested that a corollary of the latter process is that LiDAR last returns may originate from the forest canopy or understorey vegetation rather than the true ground, particularly with a small footprint system. The results of Norheim et al. (2002) (study 1) and Hodgson et al. (2005) (study 6) indicated that airborne LiDAR *under-estimated* reference terrain data by 36 and 16 cm, respectively, and while Norheim et al. recognized that the source of their bias was likely a product of poor control point co-ordinate transformations (an issue they mentioned they were looking into), Hodgson et al. suggested that their under-estimation was relatively insignificant and was just as plausible a result as slight over-estimation, given the lack of tree foliage and understorey vegetation during *leaf-off* survey.

The results presented in this study have remarkably similar approximate 20 cm absolute offset and 25 cm RMSE values under deciduous forest to studies 5 and 6 (table 6.1). Both were leaf-off investigations of forest in North and South Carolina in the U.S. conducted by Prof. Michael Hodgson's group. As previously mentioned, these landscapes offer forest conditions not dissimilar to the deciduous stands at Chopwell Wood, so analogous results for LiDAR terrain bias are promising and suggest that the overestimation that can be expected when commissioning an airborne LiDAR survey for measuring sub-deciduous canopy terrain is approximately 10 to 25 cm (Hodgson and Bresnahan, 2004; Hodgson et al., 2005).

Norheim et al. (2002) (study 1), Clark et al. (2004) (study 4) and Su and Bork (2006) (study 8) offered values for LiDAR terrain precision of 74, 166 and 52 cm, respectively (table 6.1), that are far greater than the 17 cm presented in this study. Clark et al. (2004) explained the low precision and high mean offset and RMSE (101 and 195 cm) for their accuracy assessment at a tropical rainforest in Costa Rica by erroneous inclusion of laser returns from dense understorey vegetation and dead wood on the forest floor as terrain during classification (Type II filtering errors). Only Hyyppä et al. (2005) offered precision values as low as the 17 cm calculated for

Chopwell Wood – 9 and 8 cm for leaf-on and –off surveys, respectively. While it is promising that the precision estimate for terrain under deciduous forest at Chopwell Wood is low, it does not correlate well with published data and it is therefore inconclusive whether a precision of approximately 20 cm is typical of global deciduous forest or relates more to managed broadleaf stands like those at Chopwell.

Although Reutebuch et al. (2003) (study 3) offered LiDAR overestimation of the true terrain surface under coniferous forest (of 31 cm), three other studies (1, 2 and 6) offered underestimation by approximately 10 to 25 cm (table 6.1). Reutebuch et al. (2003) obtained similar over-estimation to the results presented in this work, 18 cm compared to 23 cm, in 70+ year-old lightly thinned coniferous stands in Washington State, U.S.A. Slightly lower bias under the predominantly Douglas fir cover in Washington than the combined pine and fir coniferous species investigated in this study is expected, because the understorey vegetation under pine in summer is observed in this work to increase bias (see section 4.4.1). Mean offset under the Douglas fir species in this study is only 9 cm. The LiDAR underestimation of true terrain elevation observed by Norheim et al. (2002) (study 1) and Hodgson et al. (2005) (study 6) could again be reasonably explained by poor control point co-ordinate transformation and winter survey (with little understorey vegetation), respectively (see above); but bias of a similar magnitude was presented by Hodgson et al. (2003) for terrain estimates under pine during summer. Assuming that understorey vegetation was as dense under the pine forests of North Carolina as it was at Chopwell Wood in the summer of 2009, this underestimation is alarming. Moreover, the repeat determination of this negative bias in three separate studies (1, 2 and 6), reduces confidence in the 23 cm overestimation under coniferous forest acquired at Chopwell, although differences in method between the three published studies and this one are considerable.

Study	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	Here
Data Acquisition Date	Sep 99	Jun 00	Mar-May 99	Oct 97	Mar 00	Jan, Feb 01	Sep 98-May 03	Oct 00	Jul 09
Leaf-On or -Off	ON	ON	ON	ON	OFF	OFF	Both	ON	ON
GCP Technique	GPS & TS	RTK	GPS & TS	GPS & TS	RTK	RTK	RTK & TS	GPS & TS	GPS, TS & TLS
GCP Sampling Method	random	trans-ects	random	random	strat-random	trans-ects	random	strat-random	strat-random
LiDAR Full Point Density (per m^2)	0.3	0.1	4.2	9	0.25	0.1	-	0.8	31
Deciduous Type	Mixed	Dec		Dec	Dec	Dec	Mixed	Dec	Dec
GCP Count	444	281		~1000	82	282	~250	74	~3,800
Mean Offset	-36	21		101	-	-16	-	20	13
Absolute Offset	-	86			20	20	-	-	24
Std Dev	74	-		166	-	-	8, 9	52	17
RMSE	-	122		195	26	27	-	-	24
Conifer Type	Con	Pine	Con		Con	Pine	Con		Con
GCP Count	81	112	120		119	113	~250		~3,800
Mean Offset	-16	-11	18		-	-24	-		23
Absolute Offset	-	29	-		13	25	-		30
Std Dev	27	-	18		-	-	13, 11		18
RMSE	-	46	-		17	28	-		31
Shrub/Brush	Scrub	Shrub			Brush	Shrub		Shrub	
GCP Count	591	177			98	178		79	
Signed Offset	-16	112			-	-16		7	
Absolute Offset	-	122			19	26		-	
Std Dev	99	-			-	-		46	
RMSE	-	153			23	36		-	

Table 6.1 Accuracy assessment results for all relevant published studies preceding this research. Where, in order of publishing date, (1) Norheim et al. (2002) for mixed forest in Washington; (2) Hodgson et al. (2003) for stream corridors in North Carolina; (3) Reutebuch et al. (2003) for conifer forest in Washington; (4) Clark et al. (2004) for tropical rainforest in Costa Rica; (5) Hodgson and Bresnahan (2004) for mixed forest in South Carolina; (6) Hodgson et al. (2005) for stream corridors in North Carolina; (7) Hyyppä et al. 2005 for mixed forest in Finland and (8) Su and Bork (2006) for deciduous forest in Canada. Accuracy statistics are given in centimetres and to the nearest centimetre. 'GPS' is static acquisition of GPS control points, 'RTK' is acquisition of GPS control points with a rover and base-station, 'TS' is acquisition of control points with a total station, and TLS with a laser scanner. 'Strat-random' is a stratified-random sampling technique for sampling control points.

Results for the root mean square error (RMSE) of LiDAR terrain measurements under coniferous forest are reasonably similar in all published studies (table 6.1). RMSE varies by approximately only 10 cm between studies 2, 5 and 6 and this work, and these consistent values for bias precision, in addition to the 23 cm value for bias calculated in this study, suggest that terrain overestimation of 25 to 45 cm can be expected when commissioning a survey with modern LiDAR technology over coniferous forest (Hodgson and Bresnahan, 2004; Hodgson et al., 2003; Hodgson et al., 2005). However, this is only valid under the assumption that the positive bias obtained in this study is typical of terrain measurements acquired with a state-of-the-art LiDAR system under coniferous forest and that the negative estimates of bias obtained in studies 1, 2 and 6 are not applicable to modern laser scanning, given that all three studies were conducted approximately a decade ago using now out-dated LiDAR technology. RMSE is higher under pine-only forest (average 35 cm) than both the mixed coniferous forest (17 cm) (table 6.1) and Douglas fir stands evaluated in this study (18 cm) so terrain overestimation appears to be consistently higher under pine canopy than other coniferous tree species.

The precision of LiDAR terrain measurement under coniferous forest is very similar between this study (18 cm) and studies 1 and 7, (27 and 12 cm, respectively), and is identical to that found by Reutebuch et al. (2003) for coniferous forest in Washington (18 cm). Standard deviation values are consistently low across all studies and suggest that precision of 10 to 25 cm can be expected of a LiDAR terrain survey under coniferous canopy. It should be noted that both Norheim et al. (2002) and Hyyppä et al. (2005) obtained lower values of precision under coniferous forest than deciduous forest, which contradicts the results presented in this study.

Assessment results for shrub and brush cover are given in table 6.1, but as these types of environment were not (and could not be) investigated at Chopwell, they are provided only as reference. It is inconclusive whether LiDAR-derived terrain measurements under shrub over- or under-estimate the true surface elevation, but precision is relatively low (> 45 cm) (Norheim et al., 2002; Su and Bork, 2006).

Based on comparisons between published results and the bias and precision assessment carried out in this study, airborne LiDAR-derived terrain measurement accuracy of approximately 10-25 ± 20 cm is expected under deciduous forest and 20-45 ± 20 cm is expected under coniferous forest.

6.1.3 Evaluation of point density on the quality of a terrain model

The quality of a LiDAR-derived digital terrain model is controlled by the density and accuracy of terrain points contained within it. Topographic feature detection capability is heavily influenced by the quality of a terrain model, so the effectiveness of a model depends on terrain point density, bias and precision. The effect of high bias and low precision on a terrain model is relatively clear – the elevation at a particular location will not describe the true terrain elevation – but the effect of changing point density is less obvious and will be investigated in the following section.

Raber (2003) and Raber et al. (2007) explored the effect of changing point density on terrain model accuracy across a range of land-cover classes, including deciduous forest, mixed forest and shrub. These studies involved testing the mean absolute error and RMSE between increasingly decimated terrain models of the same area, i.e. models generated with 1 of every 2, 1 of every 4 etc. points in the original 0.6 points per m^2 dataset, against field reference control points (Raber et al., 2007). Random error patterns were observed in the increasingly decimated terrain models across all land-cover classes, that is to say areas of decreasing accuracy were *not* spatially related to reductions in point density. In addition, several artefacts, including 'crystal forests' and 'bomb craters', were identified with increasing frequency in the terrain models of lower point density. The former refers to large triangular facets caused by single returns highly deviated from the general surface trend, and the latter to sub-ground outliers (Raber, 2003). Intensive outlier removal was conducted on the raw Chopwell Wood LiDAR data, so bomb craters were not a problem; but crystal forest artefacts have been

observed in the filtered terrain model and were especially evident on the backslope of Test Site 8 (see figure 5.28 in section 5.2.4). The presence of these artefacts indicates areas of the terrain model with insufficient detail to allow topographic feature recognition.

Contrary to Raber's work, Chow and Hodgson (2009) found that the absolute error from reference terrain elevation increased monotonically with declining LiDAR point density. However, the relationship they presented as 'significant' was based on using the original DTM as reference, rather than field-acquired ground control, and therefore seems self-evident. But this could equally be interpreted as advantageous in the context of understanding different point densities in the Chopwell Wood data, as this *relative* fall in accuracy reflects reduced DTM quality at low point density, with subsequent difficulty in detecting topographic features. Hence for different point densities under coniferous or deciduous forest cover, from the same original LiDAR data, higher or lower accuracy can represent the difference between detecting and missing a feature.

Hodgson and Bresnahan's (2004) study appeared to be one of the only ones to offer terrain point densities classified by forest cover, comparable to those presented in this work (see section 4.4.4). These are understandably hard to interpret, because raw mean LiDAR point density in Hodgson and Bresnahan's study was 0.25 points per m^2 in comparison to the 31 points per m^2 at Chopwell Wood, but the laser penetration rate though deciduous and coniferous canopy was given as 48 and 36 %, respectively (Hodgson and Bresnahan, 2004), in comparison to the 13 and 8 % calculated in this study. Importantly, the rate is higher under deciduous forest in both investigations, but a considerably higher proportion of laser pulses penetrated the canopy and understorey in Hodgson and Bresnahan's study. This is likely explained by: (1) Hodgson and Bresnahan (2004) acquiring LiDAR data in *leaf-off* conditions, which heavily influences both penetration through deciduous canopy and though understorey vegetation under either deciduous or coniferous forest and/or (2) a universal difference between the attributes of deciduous and coniferous canopy in the temperate Chopwell Wood and boreal South Carolina forests that is not well understood from the information presented

in Hodgson and Bresnahan (2004). However, both studies suggest that the detail of a terrain model is typically higher under deciduous than coniferous forest, owing to a higher rate of laser penetration and subsequently greater terrain point density under broadleaf species.

The effect of changing point density on the accuracy of a terrain model has not been established, with opposing views as to whether reducing point density is spatially related to loss in accuracy (Chow and Hodgson, 2009, Raber et al., 2007). But it is clear from the results of the clustering assessment in this study (section 4.4.2) that sub-canopy terrain models under both deciduous and coniferous forest exhibit intermittent areas of high and low detail, depending on gaps or low foliage density regions of the canopy. While deciduous species offer relatively higher terrain model detail, the *quality* of a model, whether under deciduous or coniferous forest, is particularly influenced by the level of terrain point clustering.

The implications of varying terrain point density, accuracy and bias under deciduous vs. coniferous canopy, and of terrain point clustering, on archaeological feature prospection will be explored further in the next section.

6.2 Archaeological Feature Prospection with LiDAR-derived Terrain Models

6.2.1 Minimum detectable feature size and shape

To establish the viability of LiDAR-based survey for detecting archaeological features under a particular type of forest cover, the size and shape of the most challenging, but still detectable, feature must be determined. The primary control over this is the quantity (density) of laser reflections from the terrain surface, but a threshold minimum detectable size can initially be ascertained from the accuracy of the final LiDAR terrain in height and position.

Assuming an area with unlimited terrain returns, the work at Chopwell Wood and results from other published studies suggest, based on mean offset, precision and RMSE between LiDAR and reference terrain data, that only a feature with an elevation signature (i.e. elevation relative to the surrounding area, on flat or sloped ground) of approximately >20 and >30 cm under deciduous and coniferous forest, respectively (table 6.1 in section 6.1.2), can be resolved. A lower feature is just as likely to be an artefact of measurement error or the modelling procedure representative of archaeological imprint in the topography. But realistically, even expert prospection in the field will also routinely miss such features, especially in shadowed areas or under shrub or brush (Doneus and Briese, 2006b). It may additionally be true that a feature smaller than these 'limits' may be detectable if the feature is covered by densely-packed LiDAR points, but the chance of such circumstances occurring is low given typical distance between terrain points of approximately 0.5 and 0.7 m for deciduous and coniferous forest, respectively (calculated from table 4.11 in section 4.4.4). Even over an area of maximum terrain point clustering, i.e. below a gap in the tree cover and exhibiting no understorey vegetation, the smallest possible point spacing of approximately 0.18 m (based on the raw point density of 31 points per m^2) is too wide for a feature with an elevation signature smaller than 20 cm to be identified.

The effect of LiDAR point coverage on the difference between detecting and the failure to detect particularly subtle archaeological features is illustrated in figure 6.1. These schematic diagrams highlight the difference between detecting (figure 6.1 A and B) and failure to detect (C and D) two small and subtle archaeological features typical of areas such as Chopwell Wood: a 30 x 100 cm mound, for example a barrow or small pile of rubble, and a 30 x 30 cm relict drainage gulley or trench. These feature sizes were selected based on the sizes of particularly difficult to detect features at Chopwell Wood and on those described in previous studies using high-resolution airborne laser scanners (e.g. Doneus et al., 2008; Gallagher and Josephs, 2008). It must be kept in mind that these diagrams only show cross-sections through features and that archaeological feature detection over terrain models is a three-dimensional process, but when LiDAR terrain points are spaced at approximately 0.5-0.7 m (the mean spacings under coniferous and deciduous forest at Chopwell Wood) as in figure 6.1 C and D, the topography of neither feature is adequately resolved. The shape of interpolated topography in figure 6.1 C and D would change if different points were removed from A and B (i.e. if the original data were equally decimated, but different points had, for example, penetrated the tree canopy as those shown), but any combination of points at 0.5-0.7 m spacing would not aid in the detection of either archaeological feature.

The resolving power of a LiDAR system is further degraded as a result of error in position, with the horizontal accuracy of the LiDAR system employed in this study approximately ±20 cm at 1000 m flying altitude (Optech ALTM Gemini Datasheet, 2004). Error in height can mean the difference between detecting a feature or not, but the magnitude of potential positional error was not high enough to regularly disguise archaeology on the ground at Chopwell Wood, given that even the smallest features analysed (e.g. hollows, edges etc.) were always >30 cm and typically >100 cm in diameter. These feature sizes are characteristic of industrial infrastructure investigated at other locations (e.g. Gallagher and Josephs) and the horizontal accuracy of a modern airborne laser scanning system is therefore generally not a limiting factor for conducting prospection over industrial archaeology.

Terrain point density was measured at 4.1 and 2.4 points per square metre under deciduous and coniferous forest, respectively, at Chopwell Wood, and in terrain modelled from these point densities, many archaeological features were identified under the forest canopy. The smallest features still resolved in detail were the hollows behind the western drift mine (see section 5.2.2) and the narrow pathway in the east of the wood (see section 5.2.3). Numerous additional features of similar size, approximately 50 cm height and < 100 cm width, were identified and validated during extensive prospection. However, a number of features, including the 50 to 100 cm high piles of building rubble under Douglas fir in the south of the wood, could not be resolved (see section 5.2.4).

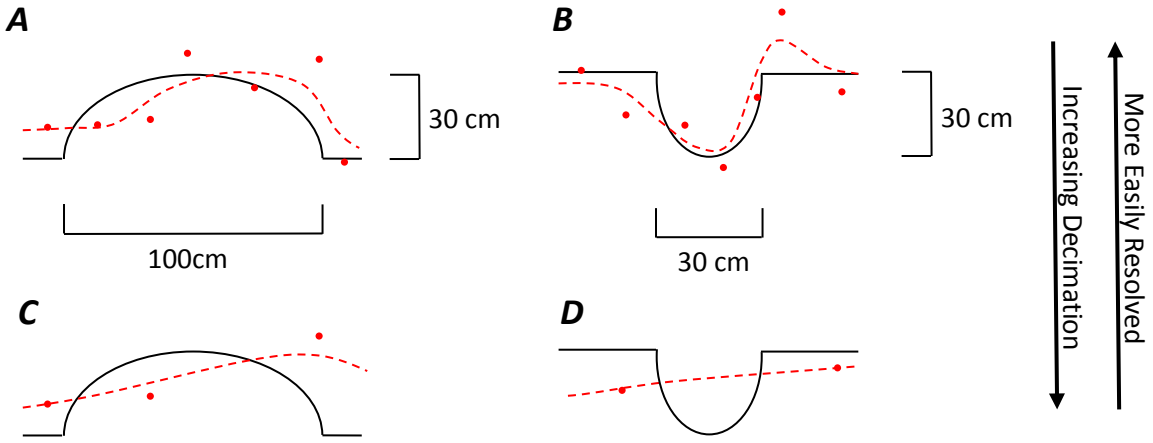


Figure 6.1 Schematic diagrams of interpolated surfaces (dashed red lines) through airborne LiDAR terrain returns (red points) against true ground elevation (black lines). **(A)** 30 x 100 cm mound, for example a barrow or small pile of building rubble, with low LiDAR point spacing. **(B)** 30 x 30 cm gulley/trench, for example relic drainage or similar to feature edge, with low point spacing. **(C)** Mound with high (wide) point spacing, typical of that observed under Chopwell Wood using the laser scanner employed in this study. **(D)** Gulley with high point spacing.

Previous studies have attempted to estimate threshold minimum LiDAR point densities required to identify various archaeological features. Challis et al. (2008) suggested that terrain modelled with LiDAR points at a density of 0.25 per m^2 tended to overlook barrows and earthworks at the River Dove in Derbyshire and plantation terrace edges of < 50 cm height in

Maryland, U.S. were not well resolved at the same point density by Harmon et al. (2006). In contrast, Crutchley (2008) suggested that terrain point density of only 1 point per m^2 was adequate to record the majority of approximately 1 m diameter barrows, enclosures and mining pits at a number of English Heritage sites (Bewley et al., 2005; Crutchley, 2006). Gallagher and Josephs (2008) used LiDAR terrain models generated from 0.7 terrain points per m^2 under mixed forest in Michigan, U.S., to locate various linear and circular topographic depressions, later validated in the field as copper prospecting trenches and pits. They identified a minimum depth for detecting negative features, i.e. those below the surface (trenches, ditches etc.), as 1.5 m and indicated that differentiating between adjacent, similar features was challenging. They also noted inaccurate LiDAR elevations on the steeply-sloping surfaces of several larger relic mines, equivalent to the problems observed at the western Drift Mine at Chopwell Wood (see section 5.2.2).

Barber (2007) identified LiDAR terrain point density requirements for various sizes of cultural heritage based on the probability of visualising a feature – the two sizes relevant to this study being 1 and 10 m diameter earthworks (highlighted in red in table 6.2). For a 'large earthwork', point densities of 0.08 and 4 points per m^2 provided 66 and 95% probability, respectively, that the feature will be visible; for a 'small' earthwork or ditch, the corresponding figures are 8 and 400 points per m^2. These calculations take no account of realistic detection scenarios and are based on strict mathematical concepts, so caution must be taken in their interpretation, but they suggest that terrain modelled using 8 points per m^2 LiDAR will regularly resolve the small earthwork/ditch type features characteristic of those at Chopwell Wood, e.g. the piles of building rubble or drift mine hollows. In this study, mean point density under deciduous forest is 4.1 points per m^2, a density of the same order of magnitude to the 66% detection probability, so it can loosely be assumed that just fewer than 66% of approximately 1 m sized features are detected with the LiDAR under deciduous tree species at Chopwell Wood. Lower point density under coniferous forest of 2.4 points per m^2 indicates that less of these smallest features are identified. It must be noted that forest management

strategies were identified as being highly influential on laser penetration rates, and associated terrain point densities, through forest canopy (see section 4.4.4), so the separation of deciduous vs. coniferous canopy would not necessarily adequately explain feature detection potential at a particular stand.

Feature size (m)	Example feature	Point density required to give 66% probability that a feature will be visible (points per m^2)	Point density required to give 95% probability that a feature will be visible (points per m^2)
10	Large earth work	0.08	4
1	Small earth work/ditch	8	400
0.1	Large stone masonry	800	40,000
0.01	Large tool marks	80,000	4,000,000
0.001	Weathered masonry	8,000,000	400,000,000

Table 6.2 Appropriate point density requirements for resolving various sizes of cultural heritage, based on standard mathematics to determine appropriate minimum sampling intervals over a regular grid of data. The two feature sizes most relevant to this study are highlighted in red. This table is based on data from table 2.1 in section 2.2. Adapted from: Barber (2007).

In an attempt to identify the minimum detectable feature size under coniferous and deciduous forest, one final issue must be addressed. It was discovered in section 4.4.2 that laser reflections from the sub-canopy terrain surface under all tree species at Chopwell were typically significantly clustered, both in two and three dimensions, resulting from selective absorption by individual canopies and/or understorey bushes/patches of vegetation. This has important implications for feature prospection, given that archaeology may be 'hit' or 'missed' depending on planimetric alignment with respect to canopy and understorey vegetation. In comparison to the bulk of published research into LiDAR-based archaeological feature prospection, point density in this study is relatively high – high enough even in sparse areas to

distinguish larger (> 1 x 1 m) features; but where features are as small and subtle as those in figure 6.1, comparative clustering can be the difference between detection (for example A and B) and not (for example C and D). This phenomenon has rarely been discussed in previous research, is difficult to predict and requires further research.

Besides size, shape is also regarded as a possible control on archaeological feature detectability, whether owing to the pattern and density of LiDAR points over a feature or due to feature illumination parallel to the light source during post-processing (Devereux et al., 2005; 2008; Harmon et al., 2006). While the latter limitation can typically be avoided by employing principal components analysis (Devereux et al., 2008), the former has been recognized as a limitation for prospection. Gallagher and Josephs (2008) evaluated relative feature detectability on two criteria: whether mining features were (a) linear or circular and (b) positive or negative (in elevation, with respect to the terrain surface). They determined that (1) some linear features were difficult to pick up due to their frequent similarity to pathways and stream courses, (2) circular features in groups often appeared as a single feature in the terrain model and (3) negative features were much clearer than positive ones. At the prospection stage of this work (see section 5.2), it was often noticed that linear features were more easily detected than circular features. This was a consequence of circular features often demonstrating similar topographic signatures to 'false' topography in the modelled terrain – that is, the topography caused by mistakenly accepted LiDAR reflections from plants/bushes in the terrain class at the filtering stage of data processing.

With reference to the discussion above and secondary data acquired from published research, the following information is recognised as relevant to calculating a minimum detectable feature size. Consistent to both deciduous and coniferous forest are horizontal LiDAR system error of 20 cm and *in situ* reference (validation) data error of 7 cm, both random errors (see section 6.4.1). Vertical LiDAR point accuracies and horizontal point densities are approximately 20 cm and 4.1 points per m^2 and 30 cm and 2.4 points per m^2 for deciduous and coniferous forest species, respectively. Based on threshold point density requirements calculated for

detecting typical small and subtle industrial archaeological features both in this work and previous studies, and on the theoretical feature detection probabilities calculated by Barber (2007) (table 6.2), a threshold minimum coverage of four LiDAR terrain points over a feature is required for detection.

Minimum detectable feature size is calculated as approximately 50 x 100 cm (vertical x horizontal) under deciduous forest and 65 x 170 cm under coniferous forest. The vertical component of the minimum detectable feature size is a product of reference data error, vertical LiDAR point accuracy and the threshold minimum point coverage. The horizontal component of the minimum detectable feature size is a product of LiDAR system error, reference data error, horizontal LiDAR point densities and the threshold minimum point coverage. However, realistically, features are more easily detected and resolved where LiDAR terrain points are fortuitously clustered and linear features are characteristically more easily resolved than circular ones.

6.2.2 Evaluation against the work of Michael Doneus' group

It is necessary when evaluating an innovative piece of research to compare results with relevant published data. If one particular study or series of work is outstanding as the foremost piece of research to date, it seems prudent to conduct supplementary assessment. The work by Michael Doneus' group on full-waveform airborne laser scanning for the detection of Iron Age hillfort remains at Purbach, Austria represents the furthest anyone has taken the idea of LiDAR-based archaeology. The remains primarily comprise earthwork ditches and mounds (barrows) of a similar range of dimensions to, for example, the tubways, rubble piles etc. at Chopwell Wood, under a mixed oak and beech forest with a varying degree of understorey vegetation (Doneus et al., 2008). The contrasting methods of Doneus's group and this work, and the success of each with feature detection, are evaluated.

Despite both studies paying close attention to the same objective while designing LiDAR classification procedures – namely to base filtering rigidity on the end product (archaeology) so as to limit feature smoothing and removal (Doneus and Briese, 2006a) – the structure of each method remains very different. Where filtering in this work utilised the progressive densification algorithm of Axelsson (1999; 2000) and field-derived training data to constrain processing parameters, Doneus' group used robust linear prediction and surface regularisation (Briese et al., 2002; Kraus and Pfeifer, 1998; 2001; Pfeifer et al., 2001) coupled with a threshold procedure for refining terrain models by removing returns with large 'echo width' (Doneus and Briese, 2006a). Echo width, a measure of the range distribution of individual surface scatterers contributing to the echo, is a function available only to full-waveform sensors and highlights returns from multiple elements rather than the terrain (Doneus and Briese, 2006a). Doneus et al. (2008) identified a threshold echo width of 1.7 ns by exploratory analysis and removed all returns with echo width above this. The key distinction between these methods is that the procedure in this study utilised *in situ* data, obtained in on-site, while Doneus' data processing chain was entirely remote from the study area. Clearly the latter is advantageous in terms of processing ease and requirements on time, due to lengthy GPS, Total Station and TLS data acquisition and registration procedures, but it can only be considered more suitable if it is of greater benefit to archaeological feature detection.

It must be noted prior to comparing results that the Chopwell Wood and Purbach LiDAR surveys were conducted under very different conditions. Although full-waveform, the system used by Doneus' group in 2006 provided average, raw point density of approximately 8 points per m^2 (Doneus and Briese, 2006a) while the system used in this work provided over 30 points per m^2. However, the Purbach survey (over deciduous forest) was performed in *leaf-off* conditions and, although filtered terrain point density was not declared, the laser penetration rate was undoubtedly higher than the approximately 10% observed at Chopwell Wood during leaf-on conditions (equating to a mean terrain point density of 3 points per m^2) (see table 4.11). It can, as such, be reasonably assumed that filtered terrain point density was very close

between the two surveys, but without foliage, terrain returns from the Purbach survey can be expected to demonstrate higher accuracy than those from the Chopwell survey.

Doneus et al. (2008) suggested that barrows of down to 20 cm height could be detected in their terrain models following removal of forest clearance piles and understorey vegetation by echo width thresholding. The group added a considerable number of new features to those originally mapped by ground survey in the 1960s (see figure 2.4 D). Additionally, depressions between 7 and 20 cm in depth were identified on a number of barrows, interpreted as traces of looting (Doneus and Briese, 2006b). This capability in detection is an upshot of leaf-off survey, with terrain modelling elevation precision under deciduous forest only 17 cm in this work, thus limiting feature detection in leaf-on conditions to a scale of at least 20+ cm. This extended ability of the full-waveform system to model < 10 cm elevation topography over the discrete-pulse equivalent may be entirely attributable to canopy cover (i.e. leaf-off rather than -on) and subsequent variation in model accuracy; but alternatively it may, in part, be accounted for by superior performance of the echo-width filtering procedure. To fully understand the benefit, if any, of full-waveform LiDAR survey and echo-width filtering over discrete-pulse survey, both techniques would have to be tested under the same controlled environment, i.e. over the same study area and for either leaf-on or -off conditions.

6.2.3 Evaluation of terrain model post-processing analysis techniques

Various techniques have been advocated for visualising and analysing the LiDAR-derived terrain model, the majority of which have been utilised in this study (section 5.2). No technique has been rejected for poor performance or incompetence as all provide some level of additional support to the analysis, but this section will discuss the effectiveness of each and the appropriate scenarios where each should be employed, based on effectiveness vs. computation time.

Vertical terrain model exaggeration is employed as a first step of nearly every archaeological study and offers a rapid method of emphasizing artificial features in the topography (Crutchley, 2006; 2008). It was used in this work to accentuate the edges of the coal tubways (section 5.2.1) and avenue (section 5.2.3), much like Harmon et al. (2006) did to detect and map linear edge-based cultural features at Maryland plantation sites in the U.S., and is most effectively used to quickly identify features with obvious structure over smooth topography. Where features are hidden in discontinuous terrain, vertical exaggeration also emphasizes natural roughness and background noise (Crutchley, 2008). Difference and slope models were also used in this study at the preliminary stage of feature prospection, but typically suffered from over-illumination compared with PCA, and slope models were difficult to interpret owing to a lack of information on slope direction and therefore feature shape (i.e. ridge or gulley).

So following the recognition of these deficiencies, Hesse (2010) designed the Local Relief Model (LRM), an evolution of a technique originally conceived by Doneus and Briese (2006b), which reflects less biased elevation information of small-scale features relative to the landscape at large than is typical in a simple difference model. The LRM method was employed in this work to enhance the pathway feature (section 5.2.3) by producing a terrain model purged of small-scale features, for example bushes or plants mistakenly classified as terrain during LiDAR filtering, and allowed a greater length of the feature to be identified. While it can therefore be argued that LRMs should always be preferred to difference or slope models, generating a 'local relief' product is time-consuming and hence an apparently inferior option may prove cost-effective for large data sets or for detecting less subtle features. For example, Sittler (2004) aimed to delineate a wide area of obvious linear ridge and furrow topography at an ancient agricultural site in Germany and, despite having to filter a dense mixed forest canopy from the raw LiDAR data set, difference models appeared to highlight features perfectly adequately.

Image processing techniques are also employed to analyse terrain models for archaeological prospection, with colour composite images generated to visualise the same topography from

multiple shading angles simultaneously (Crutchley, 2008) and principle components analysis (PCA) used to create composite images possessing the property of 'equal illumination' from all directions (Devereux et al., 2008). Devereux et al. (2005) suggested that small features on sloped terrain can be hidden by over-illumination of a hillshade model and Crutchley (2008) indicated that linear features can be missed if they are oriented parallel to the light source, but PCA especially can minimise these problems. In this study, PCA was successfully exploited to identify the concave slope of a drift mine entrance and possible subsidence of the mine tunnel under thick pine forest (section 5.2.2). For this type of scenario, i.e. complex groups of features on rough topography or sloped terrain under dense vegetation cover, principal components analysis is ideal for maximising the probability of reliable feature detection and interpretation. However, 'the method adds to the already significant burden of LiDAR processing time' (Devereux et al., 2008) and must therefore be used selectively.

As stated in the opening paragraph to this section, all terrain model visualisation techniques provide some level of additional support to airborne archaeological survey, but each method must be chosen carefully based on a particular landscape or scenario. Where vertically exaggeration of a hillshade model is capable of highlighting relic field boundaries or walls, for example, it is not able to differentiate subtle topographic earthworks from complex terrain in a situation more suitable to advanced image processing, like PCA.

6.3 Evaluation of Methodology

While outlining a methodological framework for this study, field and laboratory procedures were justified against prevailing research and the study objectives. However, where alternative approaches and procedures were available, the suitability of the method must be evaluated with respect to the final results.

6.3.1 Point cloud classification

It has been mentioned that Michael Doneus' group designed an entirely remote procedure for filtering LiDAR data (section 6.2.2) (Doneus and Briese, 2006a) and that other attempts have been made to filter without training data (Pfeifer and Mandlburger, 2009). However, the LiDAR system utilised in this study lacked the capability to differentiate between similar objects on the ground, a capability that the full-waveform system has, and required additional data to tune the filter algorithm parameters. Overlooking logistical demands, the use of training data over full remote processing is not necessarily a disadvantage and has been employed successfully by Axelsson (2000) and Zhang et al. (2003). This success originally directed the choice and justification of 'tuning' the LiDAR filtering algorithm to known terrain points over alternative methods (see section 3.3.2).

Yet the progressive densification algorithm is often employed without locally-acquired training data for commercial applications (G. Stenson, personal communication 2010), i.e. by using universal filter parameters (default to the computer software), and in these situations it is argued that the demands on field data acquisition outweigh the benefit to accurate terrain derivation. This may well be the case for wider-scale applications, such as flood modelling or planning tree felling operations; but where the aim is to identify sub-metre scale topographic features, terrain filtering parameters must be close to perfect (see section 5.2.4 for a case in point). In this study, the bias and precision of filtered LiDAR terrain data against reference data

of the true terrain elevation varied by only a few 10s of cm over a reasonably wide range of parameter values when identifying optimum filter parameters (table 4.3), suggesting that little is gained from fine tuning the filter. However, localised deviations between surfaces filtered with slightly different parameters are evident in figure 4.6, especially at (i) and (ii), and are universal to all forest-canopy filtered terrain models in this study. As already stated, in order to identify subtle topographic features, terrain models must be proficiently filtered *universally*, i.e. in all areas. Obtaining low bias and high precision statistics with respect to reference data, when using default software parameters, are irrelevant if some areas containing archaeological features are filtered poorly.

Further to evaluating the use of training data for constraining filter parameters, it is also necessary to assess the competence of the progressive densification algorithm (Axelsson, 1999; 2000) against other available filters for archaeological applications. It was advocated in both Sithole and Vosselman's (2004) and Meng et al.'s (2010) filter review papers as being one of, if not the most, capable terrain filters in forested environments and over discontinuous surfaces. Under quantitative assessment, it was recognised as the most accurate overall algorithm, i.e. for all tested landscapes/environments, in both reviews.

There is no doubt that the algorithm is *relatively* strong at filtering trees and retaining rough topography, like that observed at Chopwell Wood, compared to other available algorithms, but it still has deficiencies over steep convex up-slopes (figure 5.17 A in section 5.2.2) and is frequently unable to remove dense understorey vegetation (e.g. transect 1A in figure 4.8 in section 4.4.1). However, like Doneus and Briese's (2006a) addition of echo width thresholding to the linear prediction and surface regularisation algorithm in their archaeologically-oriented study, the progressive densification algorithm could be updated in future work to counter these deficiencies, for example by using an additional segmentation-based, region-growing function (e.g. Nardinocchi et al., 2003) to help to discriminate between areas of terrain and localised groups of laser returns over low plants or bushes.

6.3.2 Accuracy assessment

Given consensus in the published research over the most suitable method for conducting accuracy assessment, i.e. of calculating 'closeness statistics' between the elevation of filtered LiDAR terrain and field-acquired reference data, there was little need or sense in attempting a new approach (Clark et al., 2004; Hodgson and Bresnahan, 2004; Norheim et al., 2002). Yet considerable diversity between studies in the type of statistic used indicates that it is prudent to evaluate where the accuracy and bias assessment carried out in this work lies within this wider research framework.

Values for mean ('signed') offset, absolute offset, standard deviation of offset and root mean square error were generated in this work, but previous studies tended to use only one or two of these and not the same ones (see table 6.1 in section 5.1.2). These tests measure bias and precision against reference data, so making robust comparisons between studies utilising different tests is extremely difficult.

Attempts have been made in previous research to isolate the relative contribution of various sources of error to the total error obtained during accuracy assessment (Hodgson and Bresnahan, 2004; Hodgson et al., 2005; Su and Bork, 2006). The sources of error segregated in published studies include: LiDAR system error, *in situ* control (reference) point error, error caused by sloped terrain and interpolation error. By removing these components, the true capability of LiDAR as a technique for modelling terrain can be determined (e.g. Hodgson and Bresnahan, 2004). The drawback of this development to the accuracy assessment is that it is rarely performed, principally owing to intensive requirements for supplementary data collection and processing. The evaluation of kriging interpolation in section 4.3 was not of the required standard to identify the true interpolation error component, as this processing step was exclusively conducted to assess a range of interpolation techniques against each other.

6.4 Sources of Error

Many sources have been cited as contributing error to the processed LiDAR terrain data and they are now outlined. Errors inherent within the study are divided into three main sections: those associated with the *in situ* field data, with the airborne LiDAR data and with the contextual data.

6.4.1 Field data

Reference data for validating and testing airborne LiDAR include those collected with differential GPS, total station (electronic distance measurement) and terrestrial laser scanning (TLS). The latter two systems offer typical three-dimensional accuracy of ±1 and < ±3 mm, respectively (Leica TCRP1205 Datasheet, 2009; Trimble GS Series Datasheet, 2005). Inherent to the TLS sensor are a number of component errors, e.g. those induced with low signal to noise ratio or at scan borders, and correspond to the same problems with the airborne system (section 6.4.2); but at typical terrestrial scanning ranges of <20 m (2% of the standard ALS range), their effects were insignificant. Dominant over the accuracy of the reference data was error relating to the differential GPS. As the control points measured with the total station and TLS were registered to a national co-ordinate system with GPS reference ('base station') points, any error in the base station was translated to its associated sub-group of control points. Additional error was generated by inaccurate registration of a terrestrial LiDAR point cloud, which was common given the complexity of the procedure (see section 3.4.1); but where data were checked following registration, the magnitude of this component was characteristically low compared to total error. Typical dGPS vertical and horizontal accuracy was ±1 and ±1 cm, respectively, for base station locations, with subsequent translation of ±1 cm base station accuracy to all TLS and total station reference points, and ±7 and ±2 cm for direct measurements under the forest canopy (utilising aerial extension capabilities). One

centimetre error in reference data was at least an order of magnitude smaller than typical three-dimensional error in the airborne LiDAR data (see section 6.4.2).

In addition to that introduced during field acquisition of elevation control points, error was introduced during individual tree attribute data acquisition procedures. Stem diameter (at breast height), stem height and canopy base height contributed to stand-scale canopy characteristics via the calculation of stand basal area, canopy density and canopy height (through allometric relationships between forest variables) (see section 3.2.2). Girthing tape is accurate to the nearest centimetre, but measures in rounded-down centimetre classes, so introduces systematic error (underestimation) at each stem. Given that basal area and stem density are proportional to the sum of squares of stem diameters, potential systematic error was $< 1*N^2$ cm where N was the number of stems at a sample site. However, stem density was calculated to a hundredth of a hectare (i.e. to the nearest 100 m^2; table 4.7) and even if DBH was consistently underestimated for every tree at the site with most stems, Site 8 (approximately 250 stems), stem density would be underestimated by only 6 m^2. Human error introduced when measuring stem height and canopy base height with the laser hypsometer was harder to quantify. The rangefinder offers accuracy of ±10 cm but measurements were taken to the nearest 50 cm which, following experimentation, was found to eliminate the typical standard deviation resulting from human perception of highest branch or lowest live branch in repeat measurements. This method was proposed and justified by Clark et al. (2004). Human error was also introduced in the estimation of understorey vegetation cover at field test sites. Vegetation height was measured to ±5 cm with a tape, but sampling points were randomly distributed which potentially introduced bias to final mean height calculation. In fact, sampling could not be described as totally random due to the selection of sampling points by field workers on-site rather than by a computerised (unbiased) system. Severe human error is associated with estimating vegetation density and for this reason it was recorded to the nearest 5% cover, with final estimates on 'volume of cover' subsequently calculated to the nearest 100 m^3 (table 4.7 in section 4.4.1). This volume statistic only needed to be a loose

estimate, i.e. no better than ±50 m^3 accuracy, due to vast differences between test sites (table 4.7) and the requirement to simply establish whether a relationship existed between understorey and LiDAR terrain accuracy, bias and point density (figure 4.14 A and table 4.8 in section 4.4.1), not determine a robust equation for a potential trend.

6.4.2 Airborne LiDAR data

It is worth noting prior to outlining the sources of error associated with LiDAR data that they are often very difficult to quantify. Attempts have been made to 'budget' error by source; for example, Hodgson and Bresnahan (2004) discriminated between errors caused by sloping terrain, by the modelling (interpolation) procedure and by the system itself. With the exception of an attempt to identify the influence of terrain slope on total error in LiDAR elevation (see section 4.4.3), the processing requirements to conduct such a comprehensive investigation were considered beyond the scope of this study. The following section will outline sources of error associated with the LiDAR system and registration and classification procedures before presenting estimates of quantity.

Sources of error relating to the LiDAR system and how it measures objects at the Earth's surface were identified in section 1.2.3 and, as stated, are related to (1) position, (2) optical transfer or (3) detection. These include: (1) poor communication between GPS base station and aircraft; insufficient satellite reception; Inertial Navigation System (INS) misalignment and gyro drift; and scan border phenomena; (2) attenuation and scattering of the laser beam off dust particles or precipitation; (3) poor target reflectivity; 'time-walk' over sloped terrain and low range resolution (Baltsavias, 1999a; 1999b; Huising and Gomes Pereira, 1998 ; Kraus, 2004).

Poor calibration of LiDAR point clouds to reference data, misregistration of clouds to each other and poor quality reference data introduce systematic positional error (Kraus and Pfeifer, 1998). Flight-strip adjustment procedures are based on the method first detailed by Kilian et

al. (1996) of transforming strips to a homogenous exterior coordinate system via registration of pairs of tie-points within point clouds. However, since registration was carried out by the service provider, the author was unable to obtain comprehensive information on calibration errors.

Error associated with LiDAR classification is divided into that relating to the filtering algorithm, to incorrect parameter determination and to commission and omission errors introduced into the final terrain data in spite of incorrect parameters. The relative merits of Axelsson's (1999; 2000) progressive densification algorithm and the 'training data technique' for constraining filter parameters were discussed in section 6.3.1, and although each appears favourable when compared with alternatives, both are imperfect and introduce error into filtered terrain models. In their quantitative filter comparison study, Sithole and Vosselman (2004) demonstrated that the progressive densification algorithm was the best performer over an (admittedly urban) test site with vegetation cover (figure 4 in Sithole and Vosselman, 2004), with only 3% of terrain returns mistakenly rejected (Type I errors) and 6% of non-terrain returns accepted (Type II errors). These percentage errors cannot, however, be extrapolated to all environments. While more complex training datasets have been employed to determine filter parameters, i.e. using different parameters based on regional gradient or environment type (Sithole, 2001), little information has been published on the degree of error reduction associated with more comprehensive reference data.

Although their analysis was based on outdated laser scanning technology (in addition to theoretical considerations), Huising and Gomes Pereira (1998) presented a thorough evaluation of LiDAR-based errors, with errors separated by environment type (table 6.3). Indeed, when these values are compared with updated measurements, the lower end of the estimates in Huising and Gomes Pereira (1998) can be considered close to modern error values (e.g. GPS base station errors from this study; vegetation errors from Pfeifer et al., 2004; slope errors from Hodgson and Bresnahan, 2004, etc.).

Detector bias and laser pulse delay relate to the sensor range resolution, which is typically 5 ns, corresponding to a minimum detectable height difference at the Earth's surface of approximately 15 cm (Kraus, 2004). Doneus and Briese (2006b) noted that this provided a vertical threshold for detecting archaeological features in discrete-pulse LiDAR systems of a minimum 15 cm height above adjacent topography. An estimate of horizontal accuracy was not offered by the service providers, but for the LiDAR system employed in this study it was approximately 20 cm at 1000 m flying altitude (Optech ALTM Gemini Datasheet, 2004). Additional system error, both systematic and random, is generated by the GPS and INS units, but is typically centimetre-scale (table 6.3) and can be limited through high quality surveying. At the data calibration stage, 6 cm systematic error in height was removed from the point cloud after evaluation against ground control (section 3.4.2).

Understorey vegetation (causing type II filtering errors) and 'time-walk' over sloped terrain have been argued as plausible explanations for the systematic overestimation by the LiDAR of true terrain elevation at Chopwell Wood (section 4.4.1). Huising and Gomes Pereira (1998) suggested that understorey and 'time-walk' typically contribute decimetre- and centimetre-scale error, respectively (table 6.3). Without error budgeting, the contribution of each factor to total terrain elevation error at Chopwell Wood cannot be offered. However, with bias between processed LiDAR terrain and reference elevation calculated as 13 and 23 cm for deciduous and coniferous forest, respectively (see table 4.6 in section 4.4.1), the contribution of either factor cannot be greater than decimetre-scale. Total random error (precision), a combination of system and terrain roughness/reflectivity errors (table 6.3), was 17 and 18 cm for deciduous and conifer forest, respectively (see table 4.6), which are again comparable to the lower end of the magnitudes presented by Huising and Gomes Pereira (1998) (among other studies, see section 6.1.2).

Error source	Terrain type					
	flat paved	flat barren	flat grass and scrubs	hilly paved	hilly barren	hilly grass and scrubs
Magnitude of systematic errors [a]						
Detector bias and gain	cm	cm	cm	cm	cm	cm
Laser pulse delay	cm	cm	cm	cm	cm	cm
INS misalignment	cm to dm	cm to dm	cm to dm	dm to m	dm to m	dm to m
INS gyro drift	cm	cm	cm	dm	dm	dm
GPS base-line error (for base lines < = 20 km)	cm	cm	cm	cm	cm	cm
GPS delays in troposphere	cm to dm	cm to dm	cm to dm	cm to dm	cm to dm	cm to dm
Terrain slope	0	0	0	cm	cm	cm
Vegetation	0	0	dm to m	0	0	dm to m
Positional integration	cm to dm	cm to dm	cm to dm	cm to dm	cm to dm	em to dm
Overall (cm)	5–20	5–20	20–200	5–20	5–20	20–200
Magnitude of random errors						
Pulse detection	dm	dm	dm	dm	dm	dm
Pointing jitter	0	0	0	dm to m	dm to m	dm to m
INS	cm	cm	cm	cm	cm	cm
GPS	cm	cm	cm	cm	cm	cm
Terrain roughness	0	cm	cm to dm	0	cm	cm to dm
Reflectivity	cm to dm	cm to dm	cm to dm	cm to dm	cm to dm	cm to dm
Overall (cm)	10–20	10–50	10–50	20–200	20–200	20–200

[a] 0 means negligible effect.

Table 6.3 Quantification of systematic and random errors based on the theoretical considerations of Lemmens and Fortuin (1997), with the type of terrain cover closest to Chopwell Wood highlighted by the red box. Source: Huising and Gomes Pereira (1998).

6.4.3 Contextual Data

Of lesser significance to the study than sources of error in the field or airborne LiDAR data, but still relevant, the context data (GIS, mapping and aerial photography) introduced some additional error. The accuracy of the Forestry Commission GIS is relevant to the quantitative section of this work as it was utilised to investigate LiDAR terrain point density across the full study area (section 4.4.4). However, while sources of error in the GIS include incorrect forest compartment boundary locations, incorrect compartment areas and mistaken tree species classification, they were not estimated by the Forestry Commission (1998) and therefore cannot be quantified. Error in feature position is an issue relevant also to the historic mapping, given the lack of GPS capabilities when these maps were created; but these data were only used for qualitative interpretation and validation of archaeological features. More important to feature interpretation (section 5.2) was human error introduced by misinterpretation of mapping and historic aerial photography.

Chapter 7. Conclusion and Future Developments

7.1 Conclusion

Airborne laser scanning offers a promising solution to one of the few obstacles remaining to modern, remote archaeological prospection – that of detecting features hidden under woodland. With the aim of evaluating this potential, airborne LiDAR data were acquired from a helicopter over Chopwell Wood, a mixed coniferous and deciduous forest in the northeast of England. The wood is characterised by rough topography and sharp elevation changes, and contains numerous archaeological features, most of which post-date the Industrial Revolution and include coal mining and transportation infrastructure. It therefore represented an ideal location at which to test the limit of laser scanning for airborne reconnaissance of archaeological features.

In order to conduct effective prospection, digital terrain models (DTMs) were generated from the LiDAR data through classification and modelling procedures. At the classification stage, the LiDAR data were filtered into terrain and non-terrain categories with the progressive densification algorithm, designed by Axelsson (1999; 2000). Optimum parameters for the algorithm were identified and validated with a training data set, consisting of *in situ* terrain reference points collected at a field sample site considered characteristic of the Wood at large. These parameters describe filtering angle and distance thresholds, and were calculated as 11.2 degrees and 26 cm, respectively. Following classification of the full LiDAR data with the optimum parameters, the kriging technique was employed to interpolate terrain points to digital terrain models.

Further *in situ* field data were collected from eight sample sites in the Wood to evaluate the accuracy of the filtered LiDAR terrain data. Sample sites covered a range of tree species, topographic roughness and gradients, and understorey vegetation densities, so as to test the accuracy of the filtered LiDAR terrain data in different environments. LiDAR terrain data and reference data were compared, yielding precision of 18 cm for deciduous forest and 17 cm for

coniferous forest. LiDAR terrain data were found to overestimate the true terrain surface by 13 and 23 cm for deciduous and coniferous forest, respectively. This overestimation was principally attributed to erroneous inclusion of laser returns from understorey vegetation in the terrain (rather than non-terrain) category at the filtering stage of data processing. These values compare favourably with previous research, but owing to a lack of consistent methodology between published studies in this field, it is difficult to cross-compare results among those that are published.

Both LiDAR terrain point accuracy and density are critical controlling factors on digital terrain model quality which in turn affects archaeological feature detection. Terrain point density was calculated as the percentage of laser returns penetrating the forest canopy, normalised by the mean raw LiDAR point density for the entire study area (31 points per square metre), because changing helicopter flying height and inconstant scan line overlap meant that the point density of the raw data was variable. Point density was calculated for the entire study area by tree species, with mean point density under deciduous forest of 4.1 points per square metre and under coniferous forest of 2.4 points per square metre. The penetration rates over deciduous and coniferous canopy were 13.2 and 7.9%, respectively. Forest management practices, especially thinning operations, were identified as being highly influential over the rate of canopy penetration at a particular stand.

The ability to detect and resolve archaeological features was evaluated by considering a number of case studies at Chopwell Wood. A relict coal tubway of 5 m depth and 2 to 3 m width under pine, a drift mine and tunnel subsidence with < 1 m depth hollows under pine and an avenue and path combination of < 50 cm elevation under beech were all identified and resolved in detail with the LiDAR terrain models – although typically only following post-processing with advanced artificial illumination and exaggeration techniques. However, at one site, < 1 m high piles of building rubble and a 2 m deep man-made channel from a (now demolished) 150+ year old house under Douglas fir were resolved either poorly or not at all. Considering these cases, in addition to the quantitative analyses of LiDAR terrain point

accuracy and density, threshold minimum detectable feature size was estimated at approximately 50 cm height and 100 cm diameter under deciduous forest and approximately 65 cm height and 170 cm diameter under coniferous forest. These values were based on LiDAR terrain point density and accuracy, LiDAR system accuracy and the accuracy of reference data (ground control). Features smaller than these thresholds, like the < 50 cm elevation avenue and pathway for example, could only be resolved when they demonstrated a particularly obvious shape, i.e. extreme linearity.

Modern airborne laser scanners, like the one employed in this study, are able to record high quality terrain measurements below all species of forest. Successive measurements of the terrain surface are typically less than one metre apart and areas of greater than 2.5 km^2 can be surveyed in a single day. In comparison to existing ground survey techniques, for example field survey and image photogrammetry, airborne LiDAR offers unrivalled speed of data acquisition (including data processing obligations) and is accordingly the most cost-effective option when surveys over areas greater than approximately 1 km^2 are required and field survey becomes impractical (Pfeifer and Mandlburger, 2009). LiDAR technology has the ability to revolutionise archaeology, with terrestrial scanners already commonly utilised to document recognised artefacts (Barber, 2007; Ortiz et al., 2006). Previous research has suggested that topographic earthworks can be identified under specific, characteristically monoculture, forests (Devereux et al., 2005; Doneus et al., 2008); but this study has tested the capability of airborne LiDAR to detect and resolve archaeological features under a variety of tree species and forest management situations. Furthermore the LiDAR data were acquired when deciduous trees were in-foliage (leaf-on) and understorey vegetation was at peak annual cover and density, in order to suitably evaluate the limits of the technique. Results have reinforced the early promise identified in pilot studies, with archaeological features from the Industrial Period, including relic mineral tubways and mine shafts, detected under a range of forest species, both coniferous and deciduous. Airborne LiDAR represents an ideal solution for the rapid visualisation of sub-canopy topography over a wide area, and with a view to detecting and

resolving archaeological features for either remote prospection or identification of interesting sites for future field survey, it has significant scientific benefits.

With regard to the original research aim – to evaluate the capability of multiple-return discrete-pulse airborne LiDAR for detecting and resolving archaeological features under mixed coniferous and deciduous forest – the following conclusions are presented:

1. In managed deciduous and coniferous forest, airborne LiDAR terrain data overestimated the elevation of the true terrain by 13±17 and 23±18 cm, respectively. This was during deciduous leaf-on conditions when understorey vegetation was at peak density. The principal explanation for this overestimation was the mistaken classification of laser reflections from understorey vegetation as terrain during LiDAR filtering.
2. The proportion of laser reflections from the terrain, i.e. penetrating the forest canopy, was greater under deciduous forest at 13.2% than under coniferous forest at 7.9%.
3. Based on the canopy penetration rates for airborne LiDAR over deciduous and coniferous forest and the accuracy of laser returns from the terrain, sub-canopy archaeological prospection through airborne laser scanning is concluded as being more viable under deciduous than coniferous forest, even in summer deciduous leaf-on conditions.
4. By combining canopy penetration rates and terrain point accuracies with the information obtained from conducting sub-canopy prospection over a range of archaeological feature sizes and shapes, the minimum feature sizes detectable using state-of-the-art discrete-pulse airborne LiDAR under deciduous and coniferous forest were estimated as 50x100 cm (vertical x horizontal) and 65x170 cm. Linear shaped features were more easily detected than round features.

7.2 Future Developments and Recommendations

With the aim of formulating an effective framework for sub-forest canopy terrain modelling and associated archaeological feature prospection strategies, research must supplement future developments in airborne LiDAR technology. At present, the processing chain for deriving digital terrain model (DTM) products is assembled through a combination of choices made by the LiDAR data service providers and/or the data users. Preferences vary over the LiDAR data acquisition procedure, LiDAR registration and calibration procedure, LiDAR classification procedure and choice of filtering algorithm, and method of acquisition and quantity of *in situ* reference training data for classification, and indeed whether reference data are even used. That is not to say the methodology presented in this study outlines a list of perfect options, but a clear framework for processing LiDAR terrain models, with a view to detecting archaeological features, must be developed through research, before the advantages of superior laser scanning technology can be exploited.

7.2.1 Using airborne LiDAR data

A considerable amount of scientific research has gone and is still going into developing increasingly effective LiDAR terrain filtering algorithms. The reviews of Sithole and Vosselman (2004) and Meng et al. (2010) tested the performances of most of the commercially available algorithms at filtering the same data sets and developers will build upon these results with future evolutions of their algorithms or new concepts. However, little can likely be done to improve the best algorithms, so perhaps more important is the type and detail of the reference training data set employed to constrain optimum filtering parameters. Different data sets (and consequently different parameter combinations) could be acquired for varying land cover type and/or tree species, for example above and below 45 degree terrain slope or coniferous and deciduous forest. Likewise, more comprehensive *in situ* reference data could

be collected for assessing the accuracy of LiDAR terrain under, for example, a greater range of tree species than the pine, Douglas fir and beech investigated in this study, or different levels of terrain roughness or slope.

One noticeable controlling factor on the quality of LiDAR derived terrain models was the degree of local terrain point clustering, resulting from variable laser penetration through tree crown/gaps and understorey shrub and bushes. It would be beneficial to be able to identify stands prior to LiDAR data acquisition likely to cause significant clustering, which could subsequently be flown twice or purposely included in multiple scan lines to increase the chances of penetration through tree crown or shrub. This would require canopy gap analysis or preliminary low-density laser scanning prior to full survey, in order to classify complex areas, with obvious caveats that greater surveying cost and processing time would be required.

7.2.2 Defining a methodological framework

Making improvements to the LiDAR data processing steps, like developing increasingly effective filtering algorithms and more comprehensive training data sets for example, will enhance the quality of processed terrain models. But, in order to rapidly improve the viability of airborne LiDAR as a technique for sub forest-canopy archaeological feature prospection, it is perhaps more necessary to define a methodological framework to guide future research. It is essential that research is directed towards identifying effective procedures at two key areas of the processing chain: terrain filtering and accuracy assessment. Filtering algorithms with a particular focus on retaining topographic features should be evaluated against each other by testing classification accuracy on the same LiDAR data sets, i.e. across a range of environments, forest species and over various sizes and shapes of archaeological feature, in much the same way that broad filtering reviews were conducted by Sithole and Vosselman (2004) and Meng et al. (2010). Likewise, methods for assessing LiDAR terrain point or model

accuracy against reference data should be evaluated, given that current procedures are exceptionally diverse (highlighted by table 6.1 in section 6.1.2).

Only by identifying an effective terrain classification procedure (including choice of filtering algorithm) and method of accuracy assessment can future research be organised and valid comparisons be made between studies conducted at different locations. By resolving inconsistencies in method, a research framework can be established to direct future study. After the entire LiDAR terrain filtering and modelling procedure is regulated, focus can be placed on evaluating the true potential of airborne laser scanning for sub-forest canopy archaeological feature prospection and developing strategies for future commercial and research endeavours.

7.2.3 Combining several remote sensing approaches

Various remote (and also *in situ*) techniques can be utilised collectively to improve archaeological survey strategies. Barnes (2003), Coren et al. (2005) and Rowlands and Sarris (2007) all supplemented LiDAR terrain model-based feature detection with additional data including multi- and hyper-spectral satellite imagery, oblique aerial photography and LiDAR intensity data. Typically each technique contributed to overall feature detection capability, but each tended to specialise in detection of a specific type of feature, e.g. buried or non-buried features, and only the LiDAR could provide information under forest. Only LiDAR co-ordinate (rather than intensity) data has, as of yet, been employed where it is necessary to prospect below forest canopy. The use of additional remote sensing technologies is subject to future research into, for example, LiDAR intensity data for establishing better terrain classifications or aerial photography for canopy gap analysis prior to laser scanning (see section 7.2.1). The viability of combining techniques is conditional on this research and whether it indicates that the costs of adding further apparatus to the airborne surveying platform or obtaining

secondary data, and additional processing requirements, enhance prospection capabilities sufficiently.

7.2.4 Further research

With LiDAR technology evolving so rapidly, for example the pulse rate of an airborne sensor can realistically double in a couple of years, it is crucial that research is sustained throughout sensor development. The preceding sections have identified a number of areas requiring further attention and a concerted effort towards standardising future research methodology would accelerate progress. The potential for airborne laser scanning to revolutionise archaeology over forested landscapes is evident in the results presented here and a second, NERC-funded survey to acquire airborne LiDAR data over Chopwell Wood during leaf-off conditions has been commissioned. But the continued exploitation of this technology by groups with archaeological expertise, rather than exclusively technical expertise, would help to improve and accelerate data pre-processing procedures and increase awareness of the technique in the archaeology community, thereby directing funding for further research.

References

Amable, G., Devereux, B., Cockerell, T., and Renshaw, G. (2004). Analysis of interaction patterns between vegetation canopies and small footprint, high-density, airborne LiDAR. *Proceedings of the 20th International Society for Photogrammetry and Remote Sensing Congress*, Istanbul, July 2004.

Axelsson, P. (1999). Processing of laser scanner data - algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, **54**(2-3)**,** 138-147.

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. *IAPRS*, **33**(B4/1), 110–117.

Baltsavias, E. (1999a*).* Airborne laser scanning: basic relations and formulas*. ISPRS Journal of Photogrammetry and Remote Sensing*, **54**(2)**,** 199-214.

Baltsavias, E. (1999b). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, **54**(2-3)**,** 83-94.

Barber, D. (2007). 3D Laser Scanning for Heritage, Advice and Guidance to Users on Laser Scanning in Archaeology and Architecture*. English Heritage Publishing.*

Barnes, I. (2003). Aerial remote-sensing techniques used in the management of archaeological monuments on the British Army's Salisbury Plain Training Area, Wiltshire, UK. *Archaeological Prospection*, **10**(2), 83-90.

Bewley, R. (2003). Aerial survey for archaeology. *The Photogrammetric Record***, 18**(104), 273-292.

Bewley, R., Crutchley, S., and Shell, C. (2005). New light on an ancient landscape: lidar survey in the Stonehenge World Heritage Site*. Antiquity Oxford,* **79**(305)**,** 636-647.

Bowen, Z. and Waltermire, R. (2002). Evaluation of light detection and ranging (LiDAR) for measuring river corridor topography. *Journal of the American Water Resources Association,* **38**(1), 33-41.

Briese, C. (2004). Three-dimensional modelling of breaklines from airborne laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, **35**(B3), 1097–1102.

Briese, C., Pfeifer, N., and Dorninger, P. (2002*).* Applications of the robust interpolation for DTM determination. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, **34**(3/A)**,** 55-61.

Briese, C., Mandlburger, G., Ressl, C., and Brockmann, H. (2009). Automatic breakline determination for the generation of a DTM along the River Main. In: Bretar, F., Pierrot-Deseilligny, M., Vosselman, G. (eds.) *Laser scanning 2009, IAPRS*, Vol. XXXVIII, Part 3/W8 – Paris, France, September 1-2nd 2009.

Brovelli, M., Cannata, M., and Longoni, U. (2002). Managing and processing LIDAR data within GRASS. *Proceedings of the Open source GIS – GRASS Users Conference 2002*, Trento, Italy, 11-13th of September 2002.

Brugelmann, R. (2000). Automatic breakline detection from airborne laser range data. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII, Part B3, 109–115.

Brzank, A., Lohmann, P., and Heipke, C. (2005). Automated extraction of pair wise structure lines using airborne laserscanner data in coastal areas*. International Archives of Photogrammetry, Remote Sensing and Spatial information Sciences*, **36**(3/W19)**,** 36–41.

Challis, K. (2006). Airborne laser altimetry in alluviated landscapes. *Archaeological Prospection*, **13**(2)**,** 103-127.

Challis, K., Howard, A.J., Smith, S.N., Gearey, B.R., Moscrop, D.,Carey, C.J. and Thompson, A. (2006). Using LiDAR intensity to predict the organic preservation of waterlogged deposits. In: Campana, S. and Forte, M. (eds.): *From Space to Place: Proceedings of the 2nd International Workshop on Remote Sensing in Archaeology,* CNR, Rome, Italy, Dec 4-7, 2006 (British Archaeological Reports International Series 1568), 93-98. Oxford: Archaeopress.

Challis, K., Kokalj, Z., Kincey, M., Moscrop, D., and Howard, A. (2008). Airborne lidar and historic environment records. *Antiquity*, **82**(318)**,** 1055-1064.

Chow, T. and Hodgson, M. (2009). Effects of lidar post-spacing and DEM resolution to mean slope estimation*. International Journal of Geographical Information Science,* **23**(10), 1277-1295.

Clark, M.L., Clark, D.B. and Roberts, D.A. (2004*).* Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Remote Sensing of Environment*, **91**, 68-89.

Coren, F., Visintini, D., Prearo, G., and Sterzai, P. (2005*).* Integrating LiDAR intensity measures and hyperspectral data for extracting of cultural heritage. *Workshop Italy-Canada for 3D Digital Imaging and Modelling: applications of heritage, industry, medicine and land*, 2005.

Crow, P., Benham, S., Devereux, B., and Amable, G. (2007*).* Woodland vegetation and its implications for archaeological survey using LiDAR. *Forestry***, 80**(3)**,** 241-252.

Crow, P. (2003). The potential applications of remote sensing as a method of prospecting for archaeological sites under woodland*. Unpublished Forest Research Report.*

Crutchley, S. (2006). Light detection and ranging (lidar) in the Witham Valley, Lincolnshire: an assessment of new remote sensing techniques. *Archaeological Prospection*, **13**(4), 251-257.

Crutchley, S. (2008). Shedding light on the past: using Lidar to understand ancient landscapes. In: Vassilopoulos, A., Bender, O., Evelpidou, N., and Krek, A. (eds.). *Geoinformation Technologies for Geocultural Landscapes: European Perspectives.* CRC.

De Laet, V., Paulissen, E., and Waelkens, M. (2007). Methods for the extraction of archaeological features from very high-resolution Ikonos-2 remote sensing imagery, Hisar (southwest Turkey). *Journal of Archaeological Science*, **34**(5), 830-841.

Desmet, P.J.J. (1997). Effects of interpolation errors on the analysis of DEMs. *Earth Surfaces Processes and Landforms*, **22**, 563-580.

Devereux, B., Amable, G., Crow, P., and Cliff, A. (2005). The potential of airborne lidar for detection of archaeological features under woodland canopies. *Antiquity Oxford,* **79**(305)**,** 648.

Devereux, B., Amable, G., and Crow, P. (2008). Visualisation of LiDAR terrain models for archaeological feature detection. *Antiquity Oxford,* **82**(316)**,** 470.

Doneus, M. and Briese, C. (2006a*).* Digital terrain modelling for archaeological interpretation within forested areas using full-waveform laserscanning. In: Ioannides, M., Arnold, D., Niccolucci, F., and Mania, K. (eds.) *The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage,* VAST 2006.

Doneus, M., and Briese, C. (2006b). Full-waveform airborne laser scanning as a tool for archaeological reconnaissance. *BAR International Series*, **1568,** 99-105.

Doneus, M., Briese, C., Fera, M., and Janner, M. (2008). Archaeological prospection of forested areas using full-waveform airborne laser scanning. *Journal of Archaeological Science*, **35**(4)**,** 882-893.

EDINA (2009). Digimap. Edina. online at: http://edina.ac.uk/digimap/ (accessed through Oct-09 to Mar-10).

Elmqvist, M. (2002). Ground surface estimation from airborne laser scanner data using active shape models*. International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences***, 34**(3/A)**,** 114-118.

Elmqvist, M., Jungert, E., Lantz, F., Persson, A., and Soderman, U. (2001). Terrain modelling and analysis using laser scanner data*. International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences***, 34**(3/W4)**,** 219-226.

Elmqvist, M. (2001). Ground Estimation of Laser Radar Data using Active Shape Models. Presented at the OEEPE workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models, Stockholm, Sweden, 2001.

El-Sheimy, N., Valeo, C. and Habib, A. (2005). Digital Terrain Modelling*. Artech House*, Boston, Massachustettes.

Filin, S. (2002). Surface clustering from airborne laser scanning data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, **34**(3/A)**,** 119-124.

Flood, M., (2001). LIDAR activities and research priorities in the commercial sector. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIV(Pt. 3/W4), 3-8.

Forestry Commission (2009). Wildlife at Chopwell Woodland Park. *Forestry Commission England*, online at: http://www.forestry.gov.uk/website/wildwoods.nsf/LUWebDocsByKey/EnglandTyneandWearNoForestChopwellWoodlandPark (last updated in 2009)

Forestry Commission UK GIS (1998). Chopwell Wood GIS database. *Northeast branch of the Forestry Commission*, accessed in 2009.

Fountain, P. (2010). 1947 and 1956 oblique aerial photographs of Chopwell Wood.Declassified Royal Air Force photographs from a World War II aerial survey of the UK, supplied by Paul Fountain of the Friends of Chopwell Wood.

Fowler, M. (2002). Satellite remote sensing and archaeology: a comparative study of satellite imagery of the environs of Figsbury Ring, Wiltshire. *Archaeological Prospection*, **9**(2)**,** 55-69.

Franke, R. (1982). Smooth interpolation of scattered data by local thin plate splines. *Computers and Mathematics with Applications*, **8**(4), 237–281.

Gallagher, J., and Josephs, R. (2008*).* Using LiDAR to detect cultural resources in a forested environment: an example from Isle Royale National Park, Michigan, USA. *Archaeological Prospection*, **15**(3), 187-206.

Getis, A. And Ord, J.K. (1992). The analysis of spatial association by use of distance statistics*. Geographical Analysis*, **24**, 189-206.

Goepfert, J., Soergel, U., and Brzank, A. (2008). Integration of intensity information and echo distribution in the filtering process of LIDAR data in vegetated areas. *SilviLaser*, September 17-19th, 2008 – Edinburgh, UK.

Griffith, D. (1987). Spatial Autocorrelation: A Primer. Resource Publications in Geography, Association of American Geographers.

Harmon, J., Leone, M., Prince, S., and Snyder, M. (2006). Lidar for archaeological landscape analysis: a case study of two eighteenth-century Maryland plantation sites. *American Antiquity*, **71**(4)**,** 649-670.

Haugerud, R. and Harding, D. (2001). Some algorithms for virtual deforestation (VDF) of LIDAR topographic survey data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, **34**(3/W4)**,** 211-218.

Hesse, R. (2010). LiDAR derived Local Relief Models–a new tool for archaeological prospection. *Archaeological Prospection*, **17**(2), 67-72.

Hodgson, M. and Bresnahan, P. (2004). Accuracy of airborne lidar-derived elevation: empirical assessment and error budget. *Photogrammetric engineering and remote sensing*, **70**(3), 331-340.

Hodgson, M., Jensen, J., Schmidt, L., Schill, S. and Davis, B. (2003). An evaluation of LIDAR-and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. Remote Sensing of Environment, **84**(2), 295-308.

Hodgson, M., Jensen, J., Raber, G., Tullis, J., Davis, B., Thomson, G. And Schuckman, K. (2005). An evaluation of lidar-derived elevation and terrain slope in leaf-off conditions. *Photogrammetric engineering and remote sensing*, **71**(7), 817-823.

Holden, N., Horne, P., Bewley, R.H. (2002). High-resolution Digital Airborne Mapping and Archaeology. In: Bewley, R., Raczkowski, W. (eds.) *Aerial Archaeology: Developing Future Practice.* NATO Series 1, **337**, IOS Press: Amsterdam, 173-180.

Hopkinson, C., Lim, K., Chasmer, L., Treitz, P., Creed, I., and Gynan, C. (2004). Wetland grass to plantation forest–estimating vegetation height from the standard deviation of lidar frequency distributions. Laser scanners for Forest and Landscape Assessment. *International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences*, **36**(8), 288-294.

Huising, E.J. and Gomes Pereira, L.M. (1998). Errors and accuracy estimates of laser altimetry data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, **53**(5), 245-261.

Hyyppä, J., Hyyppä, H., Litkey, P., Yu, X., Haggrén, H., Rönnholm, P., Pyysalo, U., Pitkänen, J. and Maltamo, M. (2004). Algorithms and methods of airborne laser-scanning for forest measurements. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **36**(8), 1682-1750.

Jacobsen, K. and Lohmann, P. (2003). Segmented filtering of laser scanner DSMs. *International Archives of Photogrammetry and Remote Sensing*, **34**(3/W13).

Kilian, J., Haala, N. and Englich, M. (1996). Capture and evaluation of airborne laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, **31,** 383-388.

Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Oštir, K. and Džeroski, S. (2007). Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. *Remote Sensing of Environment*, **108**(1), 9-23.

Kraus, K. (2004). Photogrammetrie. Band1. Geometrische Informationen aus Photographien und Laserscanneraufnahmen. *7., vollstandig bearbeitete und erweiterte Auflage*. Walter de Gruyter, Berlin, New York.

Kraus, K. and Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, **53**(4)**,** 193-203.

Kraus, K. and Rieger, W. (1999). Processing of Laser Scanning Data for Wooded Areas. In: Fritsch and Spille (eds.) Photogrammetric Week*,* **99**, 221-231.

Kucukkaya, A. (2004). Photogrammetry and remote sensing in archeology. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **88**(1-3), 83-88.

Leica TCRP1205 Datasheet (2009). Leica TPS1200+ Series Technical Data*. Leica Geosystems AG*, Switzerland, 2009.

Lemmens, M.J.P.M. and Fortuin, E.H.W. (1997). Error analysis of airborne laser altimetry (in Dutch). *Survey Department Rijkswaterstaat*, Delft, 70.

Li, Z., Zhu, Q. And Gold, C. (2005) Digital Terrain modeling: Principles and Methodology*,* CRC.

Lindenberger, J. (1993) Laser-Profilmenssungen zur topographischen Gelandeaufnahme. PhD Thesis, Institute for Photogrammetry, Stuttgart University. Deutshce Geodatische Kommission, Reihe C, **400**, Munich.

Lloyd, C. and Atkinson, P. (2002). Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing*, **23**(12), 2519-2524.

Mackie, E.D. and Matthews, R.W. (2008). Timber Measurement*. Forestry Commission*, Edinburgh.

Mallet, C. and Bretar, F. (2009). Full-waveform topographic lidar: State-of-the-art. *Journal of Photogrammetry and Remote Sensing*, **64**, 1-16.

Meng, X., Wang, L., Silvan-Cardenas, J.L., Currit, N. (2009). A multi-directional ground filtering algorithm for airborne LiDAR*. Journal of Photogrammetry and Remote Sensing*, **64**, 117-124.

Meng, X., Currit, N. And Zhao, K. (2010). Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sensing*, **2**, 833-860.

Nardinocchi, C., Forlani, G. and Zingaretti, P. (2003*).* Classification and filtering of laser data. *International Archives of Photogrammetry and Remote Sensing*, **34**(3/W13).

Neubauer, W. (2004). GIS in archaeology-the interface between prospection and excavation. *Archaeological Prospection*, **11**(3), 159-166.

Ni-Meister, W., Jupp, D.L.B. and Dubayah, R. (2001). Modeling LIDAR waveforms in heterogeneous and discrete canopies. *IEEE Transactions on Geoscience and Remote Sensing*, **39**(9), 1943–1958.

Norheim, R.A., Queija, V.R. and Haugerud, R.A. (2002). Comparison of LiDAR and INSAR DEMs with dense ground control. Online at: http://proceedings.esri.com/library/userconf/proc02/pap0442/p0442.htm (accessed in 08-10).

Oliver, M.A. (1990*).* Kriging: A method of interpolation for Geographical Information Systems. *International Journal of Geographic Information Systems*, **4**, 313-332.

Optech ALTM Gemini Datasheet (2004) ALTM Gemini Summary Specification Sheet*. Optech Incorporated*, Canada.

Ortiz, P., Sánchez, H., Pires, H. and Pérez, J. (2006). Experiences about fusioning 3D digitalization techniques for cultural heritage documentation. *ISPRS Commission V Symposium – Image engineering and vision metrology*, Dresden, Germany, 2006.

Pfeifer, N. and Mandlburger, G. (2009). LiDAR data filtering and DTM generation. In: Shan, J. and Toth C.K. (eds.) *Topographic Laser Ranging and Scanning: Principles and Processing*. CRC, 2009.

Pfeifer, N. and Stadler, P. (2001*).* Derivation of digital terrain models in the scop environment. *Proceedings of OEEPE workshop on airborne laser scanning and interferometric SAR for detailed Detailed Digital Elevation Models,* March 1-3rd, Stockholm, Sweden, 2001. Official Publication No. 40.

Pfeifer, N., Stadler, P. and Briese, C. (2001). Derivation of digital terrain models in the SCOP++ environment*. Proceedings of the OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Digital Elevation Models*, Stockholm, Sweden, 2001.

Pfeifer, N., Gorte, B., and Elberink, S. (2004). Influences of vegetation on laser altimetry–analysis and correction approaches. Proceedings of Natscan, Laser-Scanners for Forest and Landscape Assessment-Instruments, Processing Methods and Applications*. International Archives of Photogrammetry and Remote Sensing,* **36**(8)**,** 283-287.

Pietrapertosa, C., Vellico, M., Sterzai, P. and Coren, F. (2008). Remote sensing applied to the detection of archaeological buried structures in the Aquileia site. *GNGTS 2008,* Session 3.1, 368-372.

Raber, G. (2003). The effect of LiDAR posting density on DEM accuracy and flood extent delineation*.* NASA Affiliated Research Centre (ARC), University of South Carolina.

Raber, G.T., Jensen, J., Schill, S. and Schuckman, K. (2002a). Creation of digital terrain models using an adaptive lidar vegetation point removal process. *Photogrammetric engineering and remote sensing*, **68**(12)**,** 1307-1314.

Raber, G.T., Hodgson, M.E., Jensen J.R., Tullis, J.A., Thompson, G., Davis, B. and Schuckman, K. (2002b). Comparison of LIDAR data collected leaf-on vs. leaf-off for the creation of digital elevation models. *Proceedings of the ASPRS 2002 Annual Convention*, 19-26th April, Washington D.C., 2002.

Raber, G., Jensen, J., Hodgson, M., Tullis, J, Davis, B.A. and Berglund, J. (2007). Impact of LiDAR nominal post-spacing on DEM accuracy and flood zone delineation. *Photogrammetric engineering and remote sensing*, **73**(7), 793-804.

Reutebuch, S., McGaughey, R., Anderson, H-E. And Carson, W. (2003*).* Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing*, **29**(5), 527-535.

Roggero, M. (2001). Airborne Laser Scanning-Clustering in Raw Data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, **34**(3/W4)**,** 227-232.

Rosette, J.A.B., North, P.R.J., Suarez, J.C. and Los, S.O. (2010). Uncertainty within satellite estimates of vegetation and topography*. International Journal of Remote Sensing,* **31**(5), 1325-1342.

Rowlands, A., and Sarris, A. (2007). Detection of exposed and subsurface archaeological remains using multi-sensor remote sensing. *Journal of Archaeological Science*, **34**(5)**,** 795-803.

SAM 31186 (2004). Summary report for Area 5161. *Gloucestershire county council sites and monuments record 2004*. Accessed by Devereux, B., Amable, G., Crow, P., and Cliff, A. during Oct-04).

Schickler, W. and Thorpe, A. (2001). Surface estimation based on lidar. *Proceedings of the ASPRS conference*, April 23-27th 2001, St. Louis, Missouri.

Searle, L. (2000). Chopwell Wood: Past & Present. A booklet supplied by the Friends of Chopwell Wood.

Searle, L. (2009) Personal communication.

Sithole, G. (2001). Filtering of laser altimetry data using a slope adaptive filter. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXIV (Pt. 3/W4), 203– 210.

Sithole, G. (2005*).* Segmentation and Classification of Airborne Laser Scanner Data. Ph.D. thesis, TU Delft.

Sithole, G., and Vosselman, G. (2001). Filtering of laser altimetry data using a slope adaptive filter. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, **34**(3/W4)**,** 203-210.

Sithole, G., and Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds*. ISPRS Journal of Photogrammetry and Remote Sensing*, **59**(1-2)**,** 85-101.

Sittler, B. (2004). Revealing historical landscapes by using airborne laser scanning. *Laser scanners for forest and landscape assessment*, 3-6.

Sittler, B. and Schellberg, S., (2006). The potential of LIDAR in assessing elements of cultural heritage hidden under forest canopies or overgrown by vegetation: Possibilities and limits in detecting microrelief structures for archaeological surveys. In: Campana, S., Forte, M. (eds.) From Space to Place. *Proceedings of the 2nd International Conference on Remote Sensing in Archaeology*. BAR International Series, **1568**, 117.

Smith, S.L., Holland, D.A. and Longley, P.A. (2005). Quantifying interpolation errors in urban airborne laser scanning models. *Geographical Analysis*, **37**, 200-224.

Sohn, G. and Dowman, I. (2002). Terrain surface reconstruction by the use of tetrahedron model with the MDL criterion. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, **34**(3/A)**,** 336-344.

Stenson, G. (2010). Personal communication.

Su, Q. (2004). DEM modelling, vegetation characterization and mapping of Aspen Parkland Rangeland using LiDAR data. PhD Thesis, University of Alberta, Canada, 2004.

Su, J. and Bork, E. (2006*).* Influence of vegetation, slope, and lidar sampling angle on DEM accuracy. *Photogrammetric engineering and remote sensing*, **72**(11), 1265-1274.

Terrascan Manual (2010). TerrScan User's Guide. Terrasolid.

Vu, T. and Tokunaga, M. (2002). Designing of a wavelet-based processing system for airborne laser scanner segmentation. *Proceedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*.

Tóvári, D., and Pfeifer, N. (2005*).* Segmentation based robust interpolation – a new approach to laser data filtering. ISPRS WG III/3*,* **4,** 79-84.

Trimble GS Series Datasheet (2005). Trimble GS Series 3D Scanner Datasheet. *Trimble Navigation Limited*, Germany.

Ulbrich, K. (1962). Der Bzrgstall von Purbach im Nordburgenland. Burgenlandische Heimatblatter, **24**, 138-148.

Vögtle, T. and Steinle, E., (2003). On the quality of object classification and automated building modelling based on laserscanning data. *IAPRS*, Vol.XXXIV 3/W13, Dresden, Germany.

Von Hansen, N. and Vögtle, T. (1999). Extraktion der Gelandeoberflache aus flugzeuggetragenen Laserscanner-Aufnahmen. *Photogrammetrie, Fernerkundung, Geoinformation*, **4**, 229-236.

Vosselman, G. (2000). Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing*, **33**(B3/2; PART 3)**,** 935-942.

Wack, R., and Wimmer, A. (2002). Digital terrain models from airborne laserscanner data-a grid based approach. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences,* **34**(3/B)**,** 293-296.

Wehr, A., and Lohr, U. (1999). Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, **54**(2-3)**,** 68-82.

Woodhouse, D. (2010). Personal communication.

Zhang, K., Chen, S., Whitman, D., Shyu, M., Yan, J. and Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **41**(4)**,** 872-882.

Zianis, D. (2005). Aspects of tree allometry. In: Burk, A.R. (ed.) *New Research on Forest Ecosystems*. Nova Science, 113-144.

Appendices

Appendix 1

Tree species codes for all species at Chopwell Wood:

SP	Scots pine
CP	Corsican pine
LP	Lodgepole pine
EL	European larch
JL	Japanese larch
HL	Hybrid larch
NS	Norway spruce
SS	Sitka spruce
DF	Douglas fir
WH	Western hemlock
GF	Grand fir
NF	Noble fir
LC	Lawsons cypress
OK	Oak
BE	Beech
BI	Birch
SY	Sycamore
MC	Mixed coniferous
MB	Mixed broadleaf (deciduous)

Appendix 2

Tree species groups for all species at Chopwell Wood:

Group 1 – Pine Family minus firs:

Scots pine, Corsican pine, Lodgepole pine, European larch, Japanese larch, Hybrid larch, Norway spruce, Sitka Spruce, Western hemlock, mixed coniferous

Group 2 – Fir type:

Douglas fir, Grand fir, Noble fir

Group 3 – Cypress Family:

Lawsons Cypress

Group 4 – Beech, Elm, Sycamore and Ash families:

Oak, Beech, Sycamore, Mixed broadleaf

Group 5 – Birch family:

Birch