International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: https://www.tandfonline.com/loi/tres20

Mapping woodland species composition and structure using airborne spectral and LiDAR data

R. A. Hill & A. G. Thomson

International Journal of Remote Sensing
Vol. 26, No. 17, 10 September 2005, 3763–3779

Mapping woodland species composition and structure using airborne spectral and LiDAR data

R. A. HILL and A. G. THOMSON
Centre for Ecology and Hydrology (CEH), Monks Wood, Abbots Ripton, Huntingdon, Cambs, PE28 2LS, UK

(*Received 24 June 2004; in final form 5 January 2005*)

Tree and shrub species composition and vegetation structure are key components influencing the quality of woodland or forest habitat for a wide range of organisms. This paper investigates the unique thematic classes that can be derived using integrated airborne LiDAR and spectral data. The study area consists of a heterogeneous, semi-natural broadleaf woodland on an ancient site and homogeneous broadleaf and conifer woodland on an adjoining plantation. A parcel-based unsupervised classification approach was employed, using the first two Principal Components from 12 selected wavebands of HyMap data and a Digital Canopy Height Model extracted from LiDAR data. The resultant 52 data clusters were amalgamated into 10 distinct thematic classes that contain information on species composition and vegetation structure. The thematic classes are relevant to the National Vegetation Classification (NVC) scheme for woodlands and scrub of Great Britain. Furthermore, in distinguishing structural subdivisions within the species-based NVC classes, the thematic classification provides greater information for quantifying woodland habitat. The classes show degeneration from and regeneration to mature woodland communities and thus reflect the underlying processes of vegetation succession and woodland management. This thematic classification is ecologically relevant and is a forward development in woodland maps created from remote sensing data.

1. Introduction

Forests and woodlands are important habitats exploited by a wide range of organisms. Tree and shrub species composition and vegetation structure are key components influencing the quality of forest and woodland habitat (Southwood *et al*. 1982, Fuller and Henderson 1992, Milazzo *et al*. 2003). Additional factors include spatial characteristics such as size, pattern, fragmentation (Saunders *et al*. 1991, Verbeylen *et al*. 2003), disturbance and management regimes (Donald *et al*. 1998, Rooney 2001, Chen and Tso 2004), intra- and inter-specific competition (Matthysen 1990, Mahoney and Virgl 2003) and weather conditions (Kindvall 1995, Pasinelli 2001, Menendez and Gutierrez 2004). Of those factors, it may be possible to derive vegetation structure, species composition and spatial metrics by remote sensing methods.

The airborne technique of Light Detection And Ranging (LiDAR) can supply detailed information on tree canopy structure (Lefsky *et al*. 2002, Lim *et al*. 2003b). Forest stand inventory measures that may be derived from 'small footprint' airborne

*Corresponding author. Email: rhill@ceh.ac.uk

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2005 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160500114706

LiDAR data include mean and dominant tree height, mean stem diameter, stem number, basal area and timber volume (Means *et al.* 2000, Næsset 2002, 2004). Canopy related measures include per cent canopy openness, leaf area index, ellipsoidal crown closure (Lim *et al.* 2003a) and crown diameter for individual trees (Persson *et al.* 2002, Popescu *et al.* 2003). The use of time series airborne LiDAR data has been demonstrated for detecting tree harvesting and forest growth (Yu *et al.* 2004) and changes in spatial patterns of forest stand structure following disturbance events (Boutet and Weishampel 2003). LiDAR data of forests and woodland have been shown to have applications beyond forestry. Canopy height data have been demonstrated to be of direct relevance for modelling habitat quality for woodland birds in the UK (Hinsley *et al.* 2002, Hill *et al.* 2004) and Delmarva Fox squirrels in the state of Delaware, USA (Nelson *et al.* 2003). Using small footprint airborne LiDAR data, Holmgren and Persson (2004) discriminated Scots pine and Norway spruce in mixed pine–spruce stands in Scandinavian boreal forest. However, the identification of a range of species or of individual tree species in more heterogeneous forest has yet to be demonstrated in small footprint LiDAR data.

High spatial resolution airborne multi-spectral data can be used to derive thematic maps of forest species composition. For example, Treitz and Howarth (2000) examined the discrimination of six forest ecosystem classes in northern Ontario, Canada, using data acquired by a Compact Airborne Spectrographic Imager (CASI). Franklin *et al.* (2001) used CASI data for the classification of 30 forest stands in New Brunswick, Canada, which differed in species composition, crown closure and stem density. They reported an increase in classification accuracy for the 30 stand types from 54% to 75% by the inclusion with spectral response of image texture. This was due to the influence of shadowing in stands of different forest canopy structure, which affected image texture. The resulting output thus detected species composition and structure, but the structure was not measured directly. Zhang *et al.* (2004) also used CASI data to discriminate pure and mixed wood stands in seven forest classes (aspen, aspen–pine, pine–spruce–aspen, pine, pine–spruce, spruce, and spruce–pine) in Alberta, Canada. They reached a similar conclusion to Treitz and Howarth (2000) and Franklin *et al.* (2000, 2001), that the combination of spectral and textural data from CASI imagery increased classification accuracy compared with the use of spectral response only, as textural measures related to structural features of the forest. Airborne spectral data can thus be used to derive surrogates of forest stand structure, but cannot supply direct measurements of canopy height and surface roughness, or estimates of tree stem diameter and timber volume, as can LiDAR data. Nonetheless, high spatial resolution airborne multi-spectral data have been demonstrated to estimate forest canopy characteristics of crown diameter, crown closure, and stand density (Smith *et al.* 1991, St-Onge and Cavayas 1995) and biophysical properties such as Leaf Area Index, leaf and woody biomass, and net primary productivity (Baulies and Pons 1995, Wulder *et al.* 1996, Peddle *et al.* 2001).

It has long been recognized that forest classification using multispectral data can be improved by the incorporation of ancillary environmental data (White *et al.* 1995, Wulder *et al.* 2004). Furthermore, Gillespie *et al.* (2004) recognized the potential for quantifying forest structure and species composition using combined spectral and LiDAR data. To date, the integration of spectral and LiDAR data has been demonstrated for deriving estimates of tree canopy height, crown size and density (Scarth *et al.* 2001, Hudak *et al.* 2002, McCombs *et al.* 2003), mapping individual tree crowns (Leckie *et al*. 2003), and more general aspects of land cover mapping (Hill *et al*. 2002b), landscape visualization (Hill and Veitch 2002) and parameterizing organism-habitat models (Mason *et al*. 2003). However, the integration of spectral and LiDAR data has yet to be demonstrated for woodland mapping.

This paper investigates the unique thematic classes that can be derived using integrated airborne LiDAR and spectral data for mapping heterogeneous, semi-natural broadleaf woodland on an ancient site and homogeneous broadleaf and conifer woodland on an adjoining plantation. The focus is on the ecological relevance of the derived thematic classes, which is judged in relation to the National Vegetation Classification (NVC) scheme for woodlands and scrub of Great Britain (Rodwell 1991).

2. Study area

The study area comprises two ancient woodland sites in Cambridgeshire, UK. Monks Wood, a National Nature Reserve, is a broadleaf woodland covering an area of 157 hectares. Bevill's Wood, which is separated from Monks Wood by a minor road, is a 36-hectare site that was almost entirely clear-felled and replanted in the 1950s–1960s. The geology of the study area is dominated by grey calcareous Oxford clay, chalky boulder clay and calcareous clay with loamy and sandy drift (Steele and Welch 1973). These deposits give rise to base-rich soils: gleyic brown calcareous earths and surface water gleys. The canopy species composition of these two woodland sites is dominated by the influence of their drainage conditions, base-rich soils and management histories.

Monks Wood is extremely heterogeneous in terms of the woody species making up the canopy and understorey, their relative proportions in any area, canopy closure and density, tree height and stem density. The canopy tree species of Monks Wood are ash (*Fraxinus excelsior*), pedunculate oak (*Quercus robur*), field maple (*Acer campestre*), silver birch (*Betula pendula*), downy birch (*Betula pubescens*), aspen (*Populus tremula*) and small-leaved elm (*Ulmus carpinifolia*). Ash is the most common and widespread species, occurring mostly as coppice stems but regenerating naturally wherever the canopy is opened (Steele and Welch 1973, Massey and Welch 1993). Oak occurs less frequently because of intense felling during the First World War. Field maple, silver and downy birch are found scattered throughout Monks Wood; maple most commonly as coppice stems, birch regenerating from seeds in canopy gaps. Aspen and elm form occasional clusters on the wetter soils. Elm declined significantly in the 1970s due to an outbreak of Dutch elm disease. The former elm stands were left to regenerate naturally and today tend to be rather scrubby in nature. The dominant shrub species making up the understorey and fringes of Monks Wood are hawthorn (*Crataegus monogyna*), common hazel (*Corylus avellana*), blackthorn (*Prunus spinosa*), dogwood (*Cornus sanguinea*) and common privet (*Ligustrum vulgare*). Hazel, along with ash, was coppiced until 1995. Hazel now occurs mixed with hawthorn and blackthorn throughout Monks Wood (Massey and Welch 1993). Also to be found are elder (*Sambucus nigra*), buckthorn (*Rhamnus catharticus*), grey willow (*Salix cinerea*), goat willow (*Salix caprea*), and bramble (*Rubus fruticosus*).

Bevill's Wood has stands dominated by beech (*Fagus sylvatica*), Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*) that were planted in the 1950s–1960s. These patches of woodland have a relatively homogeneous structure and tend to lack an understorey. There are, however, stands of pine and spruce that have areas of ash and scattered beech intermingled. The edges of stands inside Bevill's Wood are ringed with ash or willow trees. The outer fringe of Bevill's Wood is similar in character to Monks Wood, with ash, oak, field maple, hazel, hawthorn and blackthorn, reflecting the character of the wood prior to the 1950s felling. Also similar to Monks Wood are the open areas of herbaceous vegetation with scattered shrubs.

In relation to the National Vegetation Classification (NVC) scheme for woodlands and scrub (Rodwell 1991), the composition of Bevill's Wood appears as complicated as that of Monks Wood. The woodland and scrub types recognized in the NVC scheme are listed (using English names) in table 1. This is a hierarchical classification scheme in which woodland and scrub vegetation is divided into 10 general types and 25 communities (labelled W1 to W25). The vegetation communities are named according to the two or three most frequent species they contain across Britain as a whole. It is recognized that within the range of environmental conditions that each community occupies across Britain, a variety of

Table 1. The National Vegetation Classification (NVC) scheme for woodlands and scrub of Great Britain.

Sallow and willow carr	
W1	grey willow–marsh bedstraw
W2	grey willow–downy birch–common reed
W3	bay willow–bottle sedge
Wet birch and alder wood	
W4	downy birch–purple moor grass
W5	common alder–tussock sedge
W6	common alder–stinging nettle
W7	common alder–ash–yellow pimpernel
Ash, hazel, maple, rowan and allied basic woodland	
W8	ash–field maple–dog's mercury
W9	ash–rowan–dog's mercury
Mesophilous mixed broadleaf woodland (often oak dominated)	
W10	pedunculate oak–bracken–bramble
W11	sessile oak–downy birch–wood sorrel
Calcareous beechwood and allied yew woodland	
W12	European beech–dog's mercury
W13	English yew
Acid beech wood	
W14	European beech–bramble
W15	European beech–wavy hair-grass
Upland oak and birch wood	
W16	oak–birch–wavy hair-grass
W17	sessile oak–downy birch–fork moss
Native pine and juniper wood	
W18	Scots pine–feather moss
W19	Common juniper–wood sorrel
Subarctic willow scrub	
W20	downy willow–great wood-rush
Thorny scrub thickets and areas of recolonization	
W21	hawthorn–common ivy
W22	blackthorn–bramble
W23	common gorse–bramble
W24	bramble–velvet-grass
W25	bracken–bramble

sub-communities exist with distinct preferential species (Rodwell 1991). Based on the NVC scheme, Monks Wood is almost entirely W8 *ash–field maple–dog's mercury*, with patches of W21 *hawthorn–common ivy*, W22 *blackthorn–bramble*, and W24 *bramble–velvet-grass* in areas of recolonization and woodland fringes. Bevill's Wood contains W12 *European beech–dog's mercury*, W18 *Scots pine–feather moss* and remnants of W8 *ash–field maple–dog's mercury*. There is far more variation in the W8 *ash–field maple–dog's mercury* community in Monks Wood than in Bevill's Wood.

3. Remote sensing data

Hyper-spectral data of Monks and Bevill's Woods were acquired in June 2000 using the HyMap sensor (see http://www.hyvista.com). This sensor recorded reflected radiation in 126 wavebands, for pixels with a 4-m spatial resolution. The wavebands covered the spectral range from visible to shortwave infrared (0.437–2.486 μm): bands 1–5 visible blue, bands 6–11 visible green, bands 12–17 visible red, bands 18–22 red-edge, bands 23–62 near infrared, bands 63–94 shortwave infrared (sensor 1), bands 95–126 shortwave infrared (sensor 2). The HyMap data were supplied by the National Remote Sensing Centre, UK (now Infoterra Ltd, UK) as a 126-waveband raster image with DN values converted to radiance.

An Airborne Laser Terrain Mapper (Optech ALTM 1210—see http://www.optech.on.ca) was flown over Monks and Bevill's Woods also in June 2000. Laser pulses were emitted by the ALTM with a wavelength of 1.047 μm (near infrared). The parallel flight lines had overlapping swaths of data acquisition, resulting in an irregular distribution of points. On average, one point was recorded every 4.83 m^2 across the study site. Both first and last return elevation data were recorded for each laser pulse, which generated a circular footprint on the ground surface with a diameter of approximately 0.25 m at nadir. Based on the instrument specifications supplied by the manufacturer and the flying altitude, the LiDAR data had a horizontal and vertical accuracy of approximately 0.6 m and 0.15 m respectively.

4. Remote sensing data pre-processing

The HyMap data were geo-registered to British National Grid co-ordinates with a 4-m spatial resolution using the aircraft telemetry from the time of data acquisition and a plug-in routine for ENVI software supplied by the HyVista Corporation. A subsequent comparison with the LiDAR data showed the accuracy of this geo-registration to be within 1 pixel (i.e. 4 m) in the *x*- and *y*- directions. To reduce the data volume for image analysis, 12 of the available 126 HyMap wavebands were extracted. These were selected on the basis of spectral profiles for woodland and other vegetated surfaces in Monks and Bevill's Woods. The 12 wavebands sampled the most significant points in the spectral curves, as explained in table 2.

The LiDAR data acquired by the ALTM were supplied by the Environment Agency of England and Wales as an ASCII file of *x*-, *y*- and *z*- co-ordinates for the first and last significant return of each laser pulse. The *x*- and *y*- location of each scanned point was supplied in British National Grid (BNG) co-ordinates, whilst the *z*- elevation was supplied in metres above the Ordnance Survey of Great Britain 1936 Datum. The point-sample elevation data were interpolated into 1-m spatial resolution raster Digital Surface Models (DSMs). Separate DSMs were created for the first and last return elevation measurements. The last return DSM had the

Table 2. The 12 wavebands selected from the HyMap hyper-spectral data of Monks and Bevill's Woods based on their spectral profiles. (NIR=near infrared, SWIR=shortwave infrared)

Waveband no. in 12 band image	Waveband no. in 126 band image	Wavelength (μm) band centre	Comments
1	6	0.5074	Band with minimum blue-green response (chlorophyll absorption)
2	9	0.5542	Band with maximum green response
3	17	0.6769	Band with minimum red response (chlorophyll absorption)
4	19	0.7078	Band at mid-point of red edge
5	21	0.7381	Band at top of red-edge
6	42	1.0475	Band with maximum response on NIR plateau
7	62	1.3368	Band with minimum response at base of NIR plateau
8	66	1.4485	Sensor 1 band with minimum SWIR response (water absorption)
9	85	1.6975	Sensor 1 band with maximum SWIR response
10	96	1.9707	Sensor 2 band with low SWIR response (water absorption)
11	109	2.2066	Sensor 2 band with maximum SWIR response
12	123	2.4399	Sensor 2 band with low SWIR response (water absorption)

higher proportion of ground hits and so was used to model the terrain underlying Monks and Bevill's Woods. This was achieved by a process of adaptive morphological filtering, extracting ground hits as local elevation minima. A Digital Terrain Model (DTM) was then rendered by applying a thin-plate spline interpolation to the extracted ground hits (Hill *et al.* 2003). The DTM had a 1-m pixel size and supplied terrain elevation with a root mean square error (RMSE) of ± 0.51 m (Hill *et al.* 2002a). By the per-pixel subtraction of the DTM from the first return DSM a Digital Canopy Height Model (DCHM) was created with canopy height in metres above the ground for Monks and Bevill's Woods. The DCHM under-estimated canopy height in Monks Wood by an average of 1.02 m for shrubs and 2.12 m for trees (RMSE ± 1.12 m and ± 2.45 m respectively). This was due mostly to penetration of laser pulses into the canopy before reflecting a first significant return (Gaveau and Hill 2003). Taking height as a surrogate for canopy density (which is principally what determined laser pulse penetration and therefore canopy height under-estimation), the DCHM was calibrated using a linear relationship between height and height under-estimation (Patenaude *et al.* 2004).

5. Remote sensing data integration and classification

A parcel-based approach was selected for the mapping of woodland species and structure using the spectral information from HyMap (figure 1(a)) and the canopy height information from LiDAR (figure 1(b)). This reduced the difficulties associated with (i) local heterogeneity of species composition and structure in

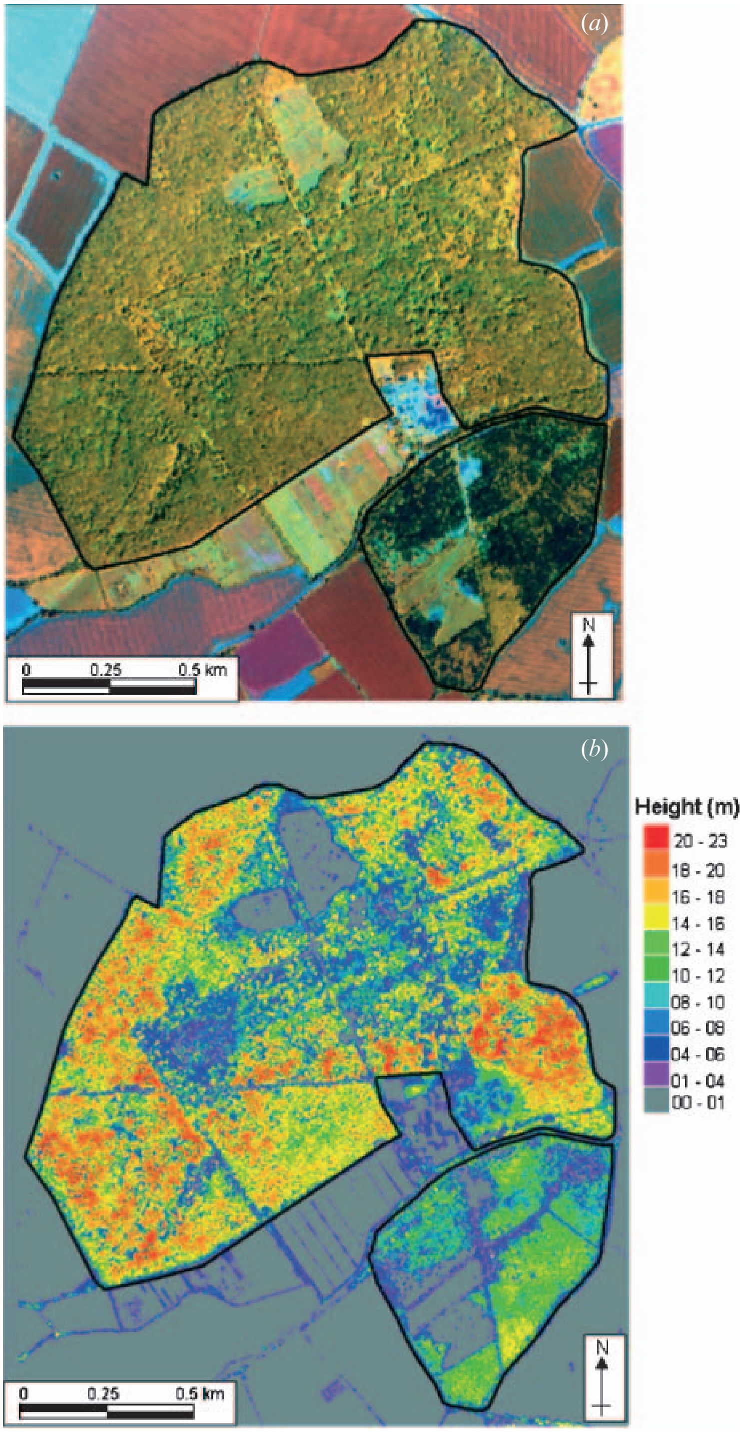

Figure 1. Remote sensing data of the study area. (a) False colour composite of HyMap data (RGB=bands 21, 19, 17: near infrared, red-edge, visible red); (b) Digital Canopy Height Model coloured in height intervals. The boundaries of Monks and Bevill's Woods are shown; Monks Wood lies to the north of Bevill's Wood. The difference between broadleaf and conifer tree species in spectral response and canopy structural heterogeneity is evident.

Monks Wood; (ii) the DCHM and HyMap data having different pixel sizes and showing spatial variability at different scales; and (iii) pixel mis-alignment where the geometric correction of the HyMap data was not absolute. The benefits of parcel-based over pixel-based mapping for forest environments using spectral data have been demonstrated by Hill (1999).

A segmentation algorithm was used to identify spatial groups of pixels with similar characteristics (Devereux *et al.* 2004). As this algorithm could only operate on three channels of data, a Principal Components Analysis (PCA) was run on the 12 selected HyMap wavebands. PCA is designed to capture the variance in a dataset by identifying a set of variables that encapsulate the maximum amount of variation in the dataset and are orthogonal to each other. This reduces the dimensionality of the data, summarizing the most important (i.e. defining) parts whilst simultaneously filtering out noise. Six Principal Components were derived from the 12 HyMap wavebands. PCs 1 and 2 contained 56% and 36% respectively of the variation in the 12 wavebands (i.e. 92% of the total). Correspondence analysis demonstrated that PC1 related to image 'brightness' (albedo) and PC2 to 'greenness' (vegetation biomass). PCs 1 and 2 were used along with the DCHM as input to the segmentation procedure (figure 2(a)). PCs 1 and 2 were, of course, uncorrelated with each other and were only weakly correlated with the DCHM: correlation co-efficient between PC1 and DCHM was -0.12, and PC2 and DCHM was $+0.06$. The segmentation process involved (i) edge-detection using a Sobel filter to identify spatial disjunctions in the input data; (ii) region-growing from seed points selected to avoid the identified edges; and (iii) region-merging. The region-growing and -merging algorithms were guided by setting segmentation thresholds for the three channels of data and by establishing the number of standard deviations expected to contain the majority of the population of a segment. The resulting parcels did not relate to individual compartments of Monks Wood or Bevill's Wood, but to areas of similar species composition and structure, as indicated by 'brightness' and 'greenness' in the spectral data and canopy height in the DCHM (figure 2(b)). Parcel size ranged from 32 m^2 to 62 048 m^2, with an average size of 480 m^2. The total of 4141 parcels was a reduction in the database from 123 114 HyMap pixels and 492 456 DCHM pixels for Monks and Bevill's Woods.

Mapping the study area was achieved by performing an unsupervised classification of the segmented PC1, PC2 and DCHM data. Each parcel contained the average value of PCs 1 and 2 and of canopy height from its constituent pixels. The unsupervised classification method (ISODATA) identified a number of clusters in the three-dimensional feature space of the input data (Tou and Gonzalez 1974). This was an iterative procedure calculated to ensure maximum statistical separability of the data clusters. Because the three input channels of data were uncorrelated or only weakly correlated, the ISODATA algorithm was able to identify 52 separate clusters. However, 25 of these data clusters were represented by three or fewer parcels in the database. A field survey was carried out in Monks and Bevill's Woods to describe the species composition and structure of parcels assigned to each of the 52 data clusters. In total, 118 parcels were examined in the field, identifying between one and 10 parcels for each of the 52 data clusters. In addition, the average and maximum canopy height for each of the 118 parcels was extracted from the DCHM. This information on species composition, structure and canopy height characteristics was combined with a statistical separability measure (Transformed Divergence Analysis) to combine the 52 data clusters identified from the remotely sensed data

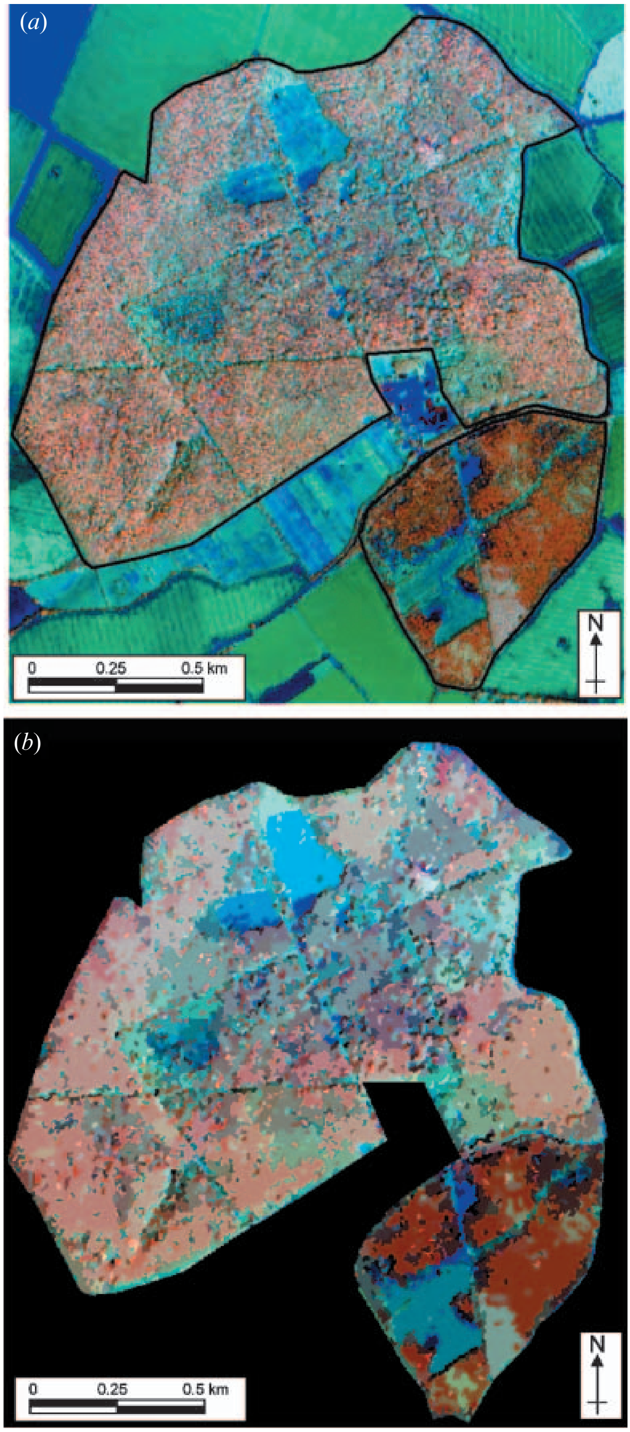

Figure 2. (a) The three channels of per-pixel data used as input for image segmentation; (b) the derived parcels used in classification of Monks and Bevill's Wood. RGB=DCHM, PC2, PC1. For clarity, the parcel boundaries are not shown.

into 10 distinct thematic classes. The 'validity' of these classes was determined by investigating the class assignment of a further 86 field surveyed parcels and by comparing the thematic descriptions with the National Vegetation Classification (NVC) of woodland and scrub (Rodwell 1991).

6. Results and discussion

The outcome of remote sensing data integration and analysis is shown in figure 3, and the 10 classes are explained in table 3. There is overlap between some of the classes in terms of top canopy species composition or canopy height ranges, but the 10 classes are all unique when both the canopy height and species composition are considered. This highlights the benefit of integrating spectral and LiDAR data, since these thematic distinctions could not have been derived using HyMap or ALTM data alone. It should be noted that this is not a classification of structurally distinct woodland patches that happen to have a certain array of species composition, or vice versa, but rather is a thematic combination of species composition and structure that has been extracted from remotely sensed data by the amalgamation of objectively-derived classes.

Reducing the input from the hyper-spectral data from 126 original wavebands to a selection of 12 representative wavebands from which the first two Principal Components were used in classification had the advantage of data volume reduction and ensuring that the three channels of data used in classification were uncorrelated or only weakly correlated. This facilitated the extraction of data clusters by unsupervised classification. The process of amalgamating the 52 data clusters into 10 thematic classes involved an element of interpretation, but followed the method used by Thomson *et al*. (2003) for coastal habitat mapping. The benefits of deriving a parcel structure by image segmentation were two-fold. First, the procedure reduced pixel-scale (i.e. very local) variation in the spectral and canopy height data, enabling the more general patterns of vegetation within the woods to be discerned. Second, the parcel structure provided a spatial framework within which variables such as average and maximum canopy height could be calculated.

The shrub layer characteristics presented in table 3 were not mapped directly, except where the shrub layer was the top canopy. These were identified in the field as being associated with the mapped top canopy classes. Of the 10 thematic classes, three occur only in Bevill's Wood (Classes 3, 4 and 7), one only in Monks Wood (Class 2) and the remaining six classes occur in both woods. It is noticeable that the class definitions are much more specific for the three classes of plantation woodland found in Bevill's Wood than for the semi-natural woodland found primarily in Monks Wood. This reflects the heterogeneity of Monks Wood, where the jumble of species composition, tree ages, and canopy structure makes characterization difficult. In Bevill's Wood, by contrast, it is a relatively simple task to identify and characterize the more homogeneous patches of vegetation.

The 86 parcels that were used to examine the validity of our interpretation of the classification output were distributed across all 10 thematic classes. Of the 86 parcels, only five were apparently assigned to an incorrect class. All five parcels had valid species compositions for their assigned thematic class, but had canopy structure outside the height limits specified in the class interpretations. Thus, the 118 parcels used in the interpretation of classification output did not cover fully the structural variation within all of the thematic classes identified across Monks and Bevill's Woods.

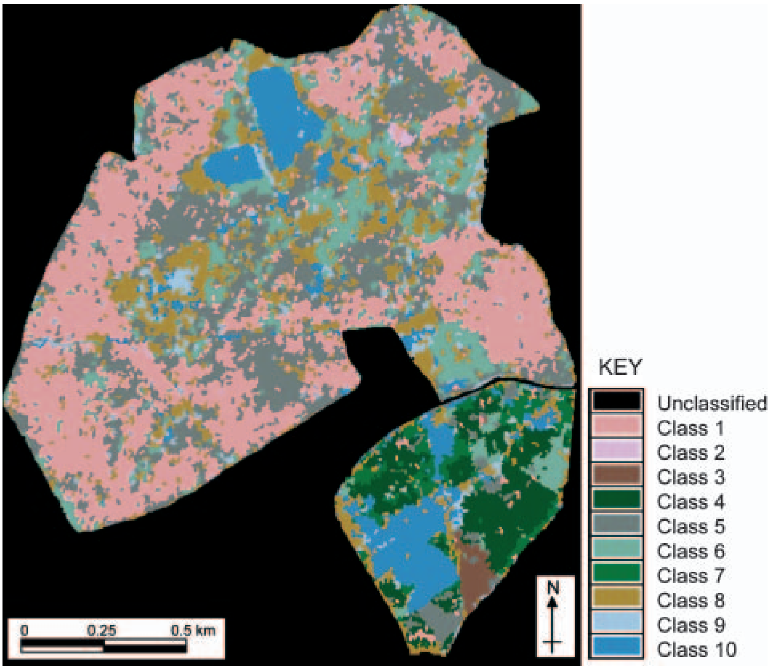

Figure 3. Classified vegetation map of Monks Wood and Bevill's Wood derived from the analysis of integrated airborne remote sensing data (for description of Classes 1 to 10 see table 3).

Of greater importance than whether patches of vegetation of 'unknown' character adhere to the thematic class description to which they are assigned, is whether the thematic classes themselves, and their spatial manifestation (as determined by image segmentation), are meaningful. In the absence of detailed, up-to-date vegetation maps of Monks and Bevill's Woods it is difficult to address the latter point. Monks Wood shows both continuous variation and patchiness in vegetation characteristics. This provokes a discourse on the merits of pixel-based versus parcel-based classification (e.g. Pedley and Curran 1991, Lobo *et al.* 1996) and of crisp versus fuzzy classification (e.g. Foody 1992, Fisher 1997) in a semi-natural environment. It is not the intention of this paper to pursue this discourse. Of greater concern here is the ecological significance of the derived thematic classes. These correspond well with the NVC scheme (table 3), although it should be noted that they provide no information on ground flora. Class 10 does not correspond to an NVC woodland and scrub community type since this is primarily non-woody vegetation. Classes 8 and 9 contain examples of W22 *blackthorn–bramble* and W21 *hawthorn–common ivy*. Classes 1, 2, 5 and 6 are all examples of W8 *ash–field maple–dog's mercury*. For the classes exclusive to Bevill's Wood, Class 3 is an example of W12 *European beech–dog's mercury*, and Classes 4 and 7 are examples of W18 *Scots pine–feather moss*.

The thematic classes derived from remotely sensed data are ecologically relevant in relation to the NVC scheme (which is species-based). Furthermore, they contain additional information on the character of woodland since, in many cases, the NVC classes are sub-divided according to canopy structure. What is being picked up here, particularly in Monks Wood, is degeneration from and regeneration to mature

Table 3. The 10 vegetation classes mapped across Monks and Bevill's Woods using integrated airborne spectral and LiDAR data.

Class number	Structure and species composition of 'top canopy'	Shrub layer characteristics	NVC community
1	Mature closed canopy (mean 14–16 m tall, up to 20–23 m). Dominant species: ash and/or oak, with field maple and in some areas scattered birch.	Sporadic under-storey: predominant species is hawthorn. (Also possible: blackthorn, hazel, young oak/ash/maple, willow, crab apple, privet, elder).	W8
2	Mature closed canopy (mean c. 15–18 m tall, up to 23 m). Dominant species: elm with oak.	Sporadic under-storey: predominant species is hawthorn.	W8
3	Closed canopy (mean c. 12 m, up to 17 m). Dominant species: beech.	Sporadic: some hawthorn	W12
4	Closed canopy (c. 10–13 m, up to 16 m). Dominant species: Norway spruce or Scots pine.	None	W18
5	Closed canopy (mean c. 10–14 m tall, up to 17–21 m). Dominant species: ash and/or oak, with birch and/or field maple. Compared with classes 1–3, there is often a higher density of younger trees with thinner trunks (especially ash), and a shorter canopy which is less dense resulting in a notable shrub layer.	Obvious under-storey of variable species composition, including any of: hawthorn, blackthorn, hazel, young birch or ash, aspen, maple, wild service, willow, elder, dogwood.	W8
6	Closed canopy of irregular nature resulting from selective thinning. A mixture of mature oak, ash and field maple trees (mean 10–15 m, up to 19 m tall) and a shorter canopy (5–10 m tall) composed of either (i) birch and/or aspen trees, (ii) a high density of thin 'pole' trunks (predominantly ash), or (iii) shrubs—see shrub layer.	Obvious under-storey of variable species composition, including any of: hawthorn, blackthorn, hazel, aspen, wild service, willow, elder, dogwood.	W8
7	Open conifer canopy composed of Norway spruce or Scots pine trees (up to 15 m tall) with gaps.	None	W18
8	Scrub canopy (3–10 m tall: see shrub layer), with occasional 10–15 m tall trees (typically oak, ash or birch), and possibly with occasional gaps. NB frequently found in areas that were previously elm dominated.	A combination of birch, aspen, or hazel coppice or a mixture of blackthorn, hawthorn, hazel, elder, willow, young oak/ash/maple, bramble.	W21, W22
9	Open scrub or coppice (2–5 m tall) with gaps.	Hazel, birch, willow, bramble	W21, W22
10	Herbaceous vegetation (grasses, sedges, soft rushes) or artificial surfaces. May contain some scattered scrub (bramble).	–	W24

communities. These processes are taking place in an environment of base-rich calcareous soils, which determines the exact nature of the NVC sub-communities being mapped. The influence of the two most significant events to impact Monks Wood, extensive felling in the 1910s and the impacts of Dutch elm disease in the 1970s, are identified clearly in the unique species-structure thematic classes presented here. The successional pathway (Clements 1916) at Monks Wood progresses from herbaceous vegetation to open scrubby canopies of birch–aspen on wetter soils, birch–ash or birch–oak on drier soils, and developing into taller and more closed canopies in which birch gives way to some combination of ash–oak–field maple (elm remaining in limited patches) with an understorey typically dominated by hazel, hawthorn and blackthorn. As the top canopy becomes more closed and mature the shrubby understorey becomes increasingly sporadic. In NVC terms, this is a progression from W21 or W22 maturing towards varieties of W8, depending on environmental conditions. The thematic classes in Monks Wood derived from the remotely sensed data contain much of this information in the narrative descriptions written following field reconnaissance. The species-structure classes arguably show five seral phases (Classes 10, 9, 8, 6 and 5) and two mature woodland communities (Classes 1 and 2) for the NVC woodland community W8 *ash–field maple–dog's mercury*. The thematic classification of Bevill's Wood also shows seral phases of NVC community W8, as well as 'open' and 'closed' canopy variants of NVC community W18 and closed canopy W12. To an extent, this shows the felling, planting and management history as represented by species composition and structure.

7. Conclusions

The integration of spectral information from HyMap and canopy height data from LiDAR into a parcel-based unsupervised classification has been demonstrated to produce an ecologically meaningful thematic product for a complex woodland environment. The resultant thematic classes contain information on species composition and structure, which reflect the underlying processes of vegetation succession and woodland management.

The characterization of the structure and species composition of Monks and Bevill's Woods presented in this paper is just one way of mapping these habitats. The derivation of classes was not entirely objective, as the process of amalgamating the output of unsupervised classification involved interpretation. The thematic classes include narrative description of structure and species composition acquired from field reconnaissance. This procedure would be repeatable at another woodland site, but with dissimilar environmental conditions the interpretation process would be different. The resultant classes would be dependent on the species–structure relationship of an individual woodland, and could only be labelled after field survey. This paper thus demonstrates the potential of data integration, which has produced a unique thematic classification, rather than an operational technique.

The derived map is of direct relevance to the study of organisms (e.g. phytophagous insects) that live in the tree/shrub canopy and those that depend on such organisms as a food source (e.g. woodland birds). This thematic classification represents a step forward in terms of woodland maps created from remote sensing data, as the integration of spectral and LiDAR data has enabled more detail to be identified than has been achieved in the past when only one data type has been used. Furthermore, in distinguishing structural subdivisions within the species-based NVC classes this thematic classification provides greater information for quantifying woodland habitat.

Acknowledgements

We are grateful to the Environment Agency of England and Wales for funding and acquiring the ALTM 1210 data used in this study. The HyMap data were acquired during the SAR and Hyper-spectral Airborne Campaign (SHAC 2000) and are courtesy of the Natural Environment Research Council (NERC) and the British National Space Centre (BNSC). HyMap data processing was carried out by Laine Skinner, University of Wales, Swansea. We are grateful to Richard Wadsworth (CEH Monks Wood) for advice on LiDAR processing methods, and to David Gaveau and Jennifer Grant (formerly of CEH Monks Wood) for assistance with LiDAR data pre-processing. We are also grateful for helpful comments on the manuscript by Geoff Smith, Andrew Wilson, and Shelley Hinsley (of CEH Monks Wood) and by the anonymous reviewers.

References

BAULIES, X. and PONS, X., 1995, Approach to forestry inventory and mapping by means of multispectral airborne data. *International Journal of Remote Sensing*, **16**, pp. 61–80.

BOUTET, J.C. and WEISHAMPEL, J.F., 2003, Spatial pattern analysis of pre- and post-hurricane forest canopy structure in North Carolina, USA. *Landscape Ecology*, **18**, pp. 553–559.

CHEN, K.C. and TSO, I.M., 2004, Spider diversity on Orchid Island, Taiwan: A comparison between habitats receiving different degrees of human disturbance. *Zoological Studies*, **43**, pp. 598–611.

CLEMENTS, F.E., 1916, *Plant Succession: An Analysis of the Development of Vegetation* ((Washington, DC: Carnegie Institute).

DEVEREUX, B.J., AMABLE, G. and COSTA POSADA, C., 2004, An efficient image segmentation algorithm for landscape analysis. *International Journal of Applied Earth Observation and Geoinformation*, **6**, pp. 47–61.

DONALD, P.F., FULLER, R.J., EVANS, A.D. and GOUGH, S., 1998, Effects of forest management and grazing on breeding bird communities in plantations of broadleaved and coniferous trees in western England. *Biological Conservation*, **85**, pp. 183–197.

FISHER, P., 1997, The pixel: a snare and a delusion. *International Journal of Remote Sensing*, **18**, pp. 679–685.

FOODY, G.M., 1992, A fuzzy sets approach to the representation of vegetation continua from remotely sensed data: an example from lowland heath. *Photogrammetric Engineering and Remote Sensing*, **58**, pp. 221–225.

FRANKLIN, S.E., HALL, R.J., MOSKAL, L.M., MAUDIE, A.J. and LAVIGNE, M.B., 2000, Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing*, **21**, pp. 61–79.

FRANKLIN, S.E., MAUDIE, A.J. and LAVIGNE, M.B., 2001, Using spatial co-occurrence texture to increase forest structure and species classification accuracy. *Photogrammetric Engineering and Remote Sensing*, **67**, pp. 849–855.

FULLER, R.J. and HENDERSON, A.C.B., 1992, Distribution of breeding songbirds in Bradfield Woods, Suffolk, in relation to vegetation and coppice management. *Bird Study*, **39**, pp. 73–88.

GAVEAU, D.L.A. and HILL, R.A., 2003, Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, **29**, pp. 650–657.

GILLESPIE, T.W., BROCK, J. and WRIGHT, C.W., 2004, Prospects for quantifying structure, floristic composition and species richness of tropical forests. *International Journal of Remote Sensing*, **25**, pp. 707–715.

HILL, R.A., 1999, Image segmentation for humid tropical forest classification in Landsat TM data. *International Journal of Remote Sensing*, **20**, pp. 1039–1044.

HILL, R.A. and VEITCH, N., 2002, Landscape visualisation: rendering a virtual reality simulation from airborne laser altimetry and multi-spectral scanning data. *International Journal of Remote Sensing*, **23**, pp. 3307–3309.

HILL, R.A., GAVEAU, D.L.A. and SPENDLOVE, M., 2002a, Accuracy issues in the assessment of deciduous woodland canopy height by airborne laser scanning: a case study. In *Operational tools in forestry using remote sensing techniques. ForestSAT Symposium 2002* ((Edinburgh: Heriot Watt University).

HILL, R.A., SMITH, G.M., FULLER, R.M. and VEITCH, N., 2002b, Landscape modelling using integrated airborne multi-spectral and elevation data. *International Journal of Remote Sensing*, **23**, pp. 2327–2334.

HILL, R.A., HINSLEY, S.A., BELLAMY, P.E. and BALZTER, H., 2003, Ecological applications of airborne laser scanner data: modelling woodland bird habitats. In *Proceedings of the Scandlaser Scientific Workshop on Airborne Laser Scanning of Forests*, J. Hyypä, E. Næsset, H. Olsson, T. Granqvist Pahlén and H. Reese (Eds). 2–3 September 2003, Umeå (Umeå: Swedish University of Agricultural Sciences), pp. 78–87.

HILL, R.A., HINSLEY, S.A., GAVEAU, D.L.A. and BELLAMY, P.E., 2004, Predicting habitat quality for Great Tits (Parus major) with airborne laser scanning data. *International Journal of Remote Sensing*, **25**, pp. 4851–4855.

HINSLEY, S.A., HILL, R.A., GAVEAU, D.L.A. and BELLAMY, P.E., 2002, Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology*, **16**, pp. 851–857.

HOLMGREN, J. and PERSSON, Å., 2004, Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, **90**, pp. 415–423.

HUDAK, A.T., LEFSKY, M.A., COHEN, W.B. and BERTERRETCHE, M., 2002, Integration of lidar and ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, **82**, pp. 397–416.

KINDVALL, O., 1995, The impact of extreme weather on habitat preference and survival in a metapopulation of the bush-cricket *Metrioptera bicolor* in Sweden. *Biological Conservation*, **73**, pp. 51–58.

LECKIE, D., GOURGEON, F., HILL, D., QUINN, R., ARMSTRONG, L. and SHREENAN, R., 2003, Combined high-density lidar and multispectral imagery for individual tree crown analysis. *Canadian Journal of Remote Sensing*, **29**, pp. 633–649.

LEFSKY, M.A., COHEN, W.B., PARKER, G.G. and HARDING, D.J., 2002, Lidar remote sensing for ecosystem studies. *BioScience*, **52**, pp. 19–30.

LIM, K., TREITZ, P., BALDWIN, K., MORRISON, I. and GREEN, J., 2003a, Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, **29**, pp. 658–678.

LIM, K., TREITZ, P., WULDER, M., ST-ONGE, B. and FLOOD, M., 2003b, LiDAR remote sensing of forest structure. *Progress in Physical Geography*, **27**, pp. 88–106.

LOBO, A., CHIC, O. and CASTERAD, A., 1996, Classification of Mediterranean crops with multi-sensor data: per-pixel versus per-object statistics and image segmentation. *International Journal of Remote Sensing*, **17**, pp. 2385–2400.

MAHONEY, S.P. and VIRGL, J.A., 2003, Habitat selection and demography of a nonmigratory woodland caribou population in Newfoundland. *Canadian Journal of Zoology*, **81**, pp. 321–334.

MASON, D.C., ANDERSON, G.Q.A., BRADBURY, R.B., COBBY, D.M., DAVENPORT, I.J., VANDEPOLL, M. and WILSON, J.D., 2003, Measurement of habitat predictor variables for organism/habitat models using remote sensing and image segmentation. *International Journal of Remote Sensing*, **24**, pp. 2515–2532.

MASSEY, M.E. and WELCH, R.C., 1993, *Monks Wood National Nature Reserve: The experience of 40 years 1953–1993* ((Peterborough: English Nature).

Matthysen, E., 1990, Behavioural and ecological correlates of territory quality in the Eurasian Nuthatch (*Sitta europaea*). *Auk*, **107**, pp. 86–95.

McCombs, J.W., Roberts, S.D. and Evans, D.L., 2003, Influence of fusing lidar and multispectral imagery on remotely sensed estimates of stand density and mean tree height in a managed loblolly pine plantation. *Forest Science*, **49**, pp. 547–466.

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

Menendez, R. and Gutierrez, D., 2004, Shifts in habitat associations of dung beetles in northern Spain: climate change implications. *Ecoscience*, **11**, pp. 329–337.

Milazzo, A., Falleta, W. and Sara, M., 2003, Habitat selection of Fat dormouse (*Glis glis italicus*) in deciduous woodlands of Sicily. *Acta Zoologica Academiae Scientiarum Hungaricae*, **49**, pp. 117–124.

Næsset, E., 2002, Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, **80**, pp. 88–99.

Næsset, E., 2004, Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, **19**, pp. 164–179.

Nelson, R., Valenti, M.A., Short, A. and Keller, C., 2003, A multiple resource inventory of Delaware using airborne laser data. *BioScience*, **53**, pp. 981–992.

Pasinelli, G., 2001, Breeding performance of the Middle Spotted Woodpecker *Dendrocopos medius* in relation to weather and territory quality. *Ardea*, **89**, pp. 353–361.

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

Peddle, D.R., Brunke, S.P. and Hall, F.G., 2001, A comparison of spectral mixture analysis and ten vegetation indices for estimating boreal forest biophysical information from airborne data. *Canadian Journal of Remote Sensing*, **27**, pp. 627–635.

Pedley, M.I. and Curran, P.J., 1991, Per-field classification: an example using SPOT HRV imagery. *International Journal of Remote Sensing*, **12**, pp. 2181–2192.

Persson, Å., Holmgren, J. and Söderman, U., 2002, Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, **68**, pp. 925–932.

Popescu, S.C., Wynne, R.H. and Nelson, R.F., 2003, Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, **29**, pp. 564–577.

Rodwell, J.S. (ed), 1991, *British Plant Communities Volume 1 Woodlands and scrub* ((Cambridge: Cambridge University Press).

Rooney, T.P., 2001, Deer impacts on forest ecosystems: a North American perspective. *Forestry*, **74**, pp. 201–208.

Saunders, D.A., Hobbs, R.J. and Margules, C.R., 1991, Biological consequences of ecosystem fragmentation: a review. *Conservation Biology*, **5**, pp. 18–32.

Scarth, P., Phinn, S.R. and McAlpine, C., 2001, Integrating high and moderate spatial resolution image data to estimate forest age structure. *Canadian Journal of Remote Sensing*, **27**, pp. 129–142.

Smith, N., Borstad, G., Hill, D. and Kerr, R., 1991, Using high-resolution airborne spectral data to estimate forest leaf area and stand structure. *Canadian Journal of Forest Research*, **21**, pp. 1127–1132.

Southwood, T.R.E., Moran, V.C. and Kennedy, C.E.J., 1982, The richness, abundance and biomass of the arthropod communities on trees. *Journal of Animal Ecology*, **51**, pp. 635–649.

Steele, R.C. and Welch, R.C. (Eds), 1973, *Monks Wood, A Nature Reserve Record* ((Peterborough: The Nature Conservancy).

St-Onge, B. and Cavayas, F., 1995, Estimating forest stand structure from high resolution imagery using the directional variogram. *International Journal of Remote Sensing*, **16**, pp. 1999–2021.

Thomson, A.G., Fuller, R.M., Yates, M.G., Brown, S.L., Cox, R. and Wadsworth, R.A., 2003, The use of airborne remote sensing for extensive mapping of intertidal sediments and saltmarshes in eastern England. *International Journal of Remote Sensing*, **24**, pp. 2717–2737.

Tou, J.T. and Gonzalez, R.C., 1974, *Pattern Recognition Principles* ((Reading, Massachusetts: Addison-Wesley).

Treitz, P.M. and Howarth, P.J., 2000, Integrating spectral, spatial, and terrain variables for forest ecosystem classification. *Photogrammetric Engineering and Remote Sensing*, **66**, pp. 305–317.

Verbeylen, G., De Bruyn, L. and Matthysen, E., 2003, Patch occupancy, population density and dynamics in a fragmented red squirrel *Sciurus vulgaris* population, *Ecography*, **26**, pp. 118–128.

White, J.D., Kroh, G.C. and Pinder, J.E., 1995, Forest mapping at Lassen-Volcanic-National-Park, California, using Landsat TM data and a geographical information-system. *Photogrammetric Engineering and Remote Sensing*, **61**, pp. 299–305.

Wulder, M.A., Franklin, S.E. and Lavigne, M.B., 1996, High spatial resolution optical image texture for improved estimation of forest stand leaf area index. *Canadian Journal of Remote Sensing*, **22**, pp. 441–449.

Wulder, M.A., Franklin, S.E., White, J.C., Cranny, M.M. and Dechka, J.A., 2004, Inclusion of topographic variables in an unsupervised classification of satellite imagery. *Canadian Journal of Remote Sensing*, **30**, pp. 137–149.

Yu, X.W., Hyyppä, J., Kaartinen, H. and Maltamo, M., 2004, Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment*, **90**, pp. 451–462.

Zhang, C., Franklin, S.E. and Wulder, M.A., 2004, Geostatistical and texture analysis of airborne-acquired images used in forest classification. *International Journal of Remote Sensing*, **25**, pp. 859–865.