Policy of the Countryside Survey partners on access to survey square locations (http://www.countrysidesurvey.org.uk/square-access-policy)

CS square locations are considered confidential to preserve the representativeness of sampling sites and the goodwill of landowners. These are both essential elements to the future of the survey to ensure the scientific integrity of the sampling strategy, the protection of the environment, and help to ensure future permission from landowners to survey their land. Future surveys would be compromised if either of these elements were to be jeopardised and our capacity to reliably inform environmental policy would be diminished.

We believe our position on confidentiality to be in the public interest. Any requests for this information will be dealt with by CEH, on behalf of NERC and the CS partners, under the terms of the UK Freedom of Information Act 2000 (The Act) or the associated Environmental Information Regulations 2004 (The Regulations).

Spatial data from within survey squares (e.g. maps of habitats, linear features) could be used relatively easily to identify the location of a square. It is therefore considered to carry equivalence to the location data and will only be released under the same exceptional circumstances and terms as the six-digit grid references location data.

The numbers of individual survey squares and the figures containing images of those squares have therefore been redacted in the publicly available version of this report.

Ian McCulloch
CEH Archives and Records Coordinator
17th June 2011

CENTRE FOR ECOLOGY AND HYDROLOGY
(NATURAL ENVIRONMENT RESEARCH COUNCIL)

Project. C00324 (T02083j5)

Countryside Survey 2000 - Part I

Module 8: Airborne Scanner Applications

Classification of airborne CASI and LIDAR data of selected CS2000 sample squares

Fifth Interim Report

CSCL/Int5

R.A. Hill[1,] R.M. Fuller[1] & N. Veitch[2]

[1]Centre for Ecology and Hydrology, Monks Wood,
Abbots Ripton, Huntingdon, Cambridgeshire, PE28 2LS

[2] The Environment Agency, National Centre for Environmental Data and Surveillance,
Rivers House, Lower Bristol Road, Bath, Avon, BA2 9ES

Corresponding author:
R.A. Hill
Section for Earth Observation
Centre for Ecology and Hydrology
Monks Wood
Abbots Ripton
Huntingdon
Cambs PE28 2LS

Telephone: 01487 772 485
Fax: 01487 773 467
Email: rhill@ceh.ac.uk

6 October 2000

CONTENTS

1. EXECUTIVE SUMMARY 3

2. INTRODUCTION 5

3. DATA PROCESSING FLOWLINE FOR THE TRIAL SQUARES 5
   3.1 LIDAR data pre-processing 5
   3.2 CASI image pre-processing 5
   3.3 Image classification 6
   3.4 Knowledge-based correction 6
   3.5 Validation of classification output 6

4. APPLICATION OF PROCESSING FLOWLINE TO THE CHECK SQUARES 6

5. RESULTS OF CLASSIFICATION 8
   5.1 Arable, Pastoral and Marginal check squares 8
   5.2 Upland check square 9

6. ADITIONAL ANALYSES FOR ARABLE TRIAL SQUARE 10

7. PROGRESS MEASURED AGAINST SCHEDULE 11

8. CONCLUSIONS 11

9. REFERENCES 12

Figures 1-4 14-17

1. EXECUTIVE SUMMARY

- This project is evaluating the use of a Compact Airborne Spectrographic Imager (CASI) and LIght Direction And Ranging instrument (LIDAR) in Countryside Survey 2000, for environmental monitoring at an extent and scale which is intermediate to the field and satellite surveys.
- Pairs of example survey squares are being studied in each of the Arable, Pastoral, Marginal and Upland Landscapes of GB - as defined in Countryside Survey 1990. Each pair has been divided into a *trial* and a *check* square, to allow the development, refinement, and validation of methods and their subsequent 'blind testing'.
- After difficulties with 1998 imagery, replacement CASI and LIDAR data-sets were flown by the Environment Agency (EA) during summer 1999.
- CEH developed a data processing flowline using the four 1 x 1 km trial squares. This involved: LIDAR data pre-processing; CASI image pre-processing; image classification; knowledge-based correction (KBC); validation of classification output by comparison with field survey data.
- The applicability of the methods and data-sets derived using the trial squares was tested in 'blind trials' on the four check squares.
- Many of the stages involved in both the CASI and LIDAR pre-processing (e.g. creating a Digital Surface Model from the LIDAR point sample information, and CASI flightline normalisation) involved running a series of command lines and software applications, and so were readily transferable to new data sets.
- The pre-processing stages of cleaning the LIDAR-derived height data, and CASI flightline registration and mosaicking, however, were manual processes requiring considerable user input. There was no correction algorithm that could be applied to all image files for these processes.
- An additional pre-processing stage to correct topographically induced variations in illumination was investigated for the Marginal check square, which contained a notable level of shadowing due to variations in aspect and relief. Software that was developed for operational use in the Land Cover Map (LCM) was used to correct differential illumination across the landscape.
- Classifications used the segment-based approach of LCM. Input thresholds for image segmentation were directly transferable from the trial squares to the check squares.
- The classification of trial and check squares as pairs required their spectral normalisation, which was achieved by an averaging process for the Arable and Pastoral sites (creating a set of four normalised squares) and for the Marginal sites (creating a normalised image pair). However, the land-cover of the central 1 km square of the two Upland sites was too distinct to enable spectral normalisation in these examples.
- For the Arable and Pastoral sites, the check squares were classified together, using the combined training data from the trial squares, but with the addition of three land-cover types not present in the trial squares.
- For the Marginal site, the check square contained fewer land-cover types than the trial square, and so the training data for those land-cover types not present were excluded from the classification.
- The Upland check square was trained independently using examples of known land-cover types.
- The classification procedure was performed in the same way as for the trial squares, with the Maximum Likelihood algorithm applied to each parcel, using mean statistics from a shrunken polygon to select the most likely class in statistical terms.

- The KBC rules were applied directly to the Arable, Pastoral and Marginal check squares, with only minor alterations required. This involved: slight changes to the height thresholds used for correcting woodland / hedge classifications; the addition of extra rules to address problems of water mis-classification in the Arable and Pastoral check squares; and the removal of rules relating to classes not present in the Marginal check square.
- Validation of the classification outputs was performed by comparison with the digitised CS2000 Broad Habitat data. As with the trial squares, a degree of editing was necessary to make the two data-sets more comparable (e.g. giving the field survey hedges a width rather than treating them as vector lines with no thickness).
- Correspondence between the classified CASI-LIDAR imagery (after full knowledge-based correction) and the field survey data (after edits) was 68.7%, 69.1%, 77.4% and 80.1% for the Marginal, Upland, Arable and Pastoral sites respectively.
- For those check squares classified using training data from the trial squares, the correspondence with field survey data was on average 10 percentage points below that achieved for the trial squares (after knowledge-based correction).
- Residual discrepancies with the field survey data related to confusion between: improved grassland neutral grassland and young green crops; arable harvested and hay fields, deciduous and coniferous woodland; arable bare and built surfaces; and between the land-cover classified in the airborne imagery and the land-use mapped in the field survey.
- Independent classification of the two Upland squares gave correspondences with field survey data of approximately 70%. Discrepancies in this Landscape type related in part to image mis-registration and mis-classification, but also to a greater degree of land-cover detail available in the airborne imagery than was portrayed in the generalised parcel boundaries of the field survey.
- The results of the blind trials must be seen in the context that: (i) the field survey had a repeatability level of 88% in identifying primary land cover codes, and so should not be viewed as absolute ground-truth data; (ii) the trials made use of CASI and LIDAR data acquired before integration had become an operational process in airborne data collection; (iii) the sample data-set of one trial square per Landscape type was inadequate to derive a comprehensive set of training statistics for the classification of the check squares.
- Additional analyses are being performed using the Arable trial square investigating the robustness of integrated CASI-LIDAR data: for inter-annual studies; for extending the study area beyond the central 1 km square; and for the detailed sub-division of woody vegetation classes.
- A two-stage process has been devised to generate a hierarchical classification of woody vegetation, operating firstly per-pixel based on geometry and structure, and secondly per-parcel based on context. This defines: areas of woodland / scrub; isolated trees / clumps; hedgerows and treelines (with the location of trees identified).
- Progress is currently on schedule: the data processing flowline has been devised using the trial squares and blind testing on the check squares is complete. Remaining work is on the Arable trial square and the comparison of all eight sites with the Land Cover Map UK. The Final Report should be completed on time, and work is due to begin on the first scientific paper from this work.
- Results are thus very promising. Issues relating to the lengthy image pre-processing stages and the transferability of training statistics between sites should be solved by current developments in EA and NERC image acquisition and processing systems.
- Airborne scanning could therefore make a significant contribution to future Countryside Surveys, as an operational system supplying data to compliment the field survey (in terms of spatial detail, extent of coverage, topographic and height data).

• 2. INTRODUCTION

Airborne remote sensing can provide data with a spatial scale and extent that is intermediate between field survey and satellite data. This project aims to evaluate the use of airborne scanner applications in the context of the 1 km field squares and satellite census of Countryside Survey 2000. Data acquired by the Compact Airborne Spectrographic Imager (CASI) and LIght Detection and Ranging instrument (LIDAR) may enable a greater understanding of the links between ground-based sample survey and the satellite-derived census. This may allow elements of the field survey to be replaced, and may also allow extension of the 1 km study sites to record extra details of their wider contexts.

The focus of this work is on identifying the extent and spatial patterns of land cover, linear landscape features and widespread Broad Habitats. Pairs of example survey squares are being studied in each of the Arable, Pastoral, Marginal and Upland Landscapes of GB - as defined in Countryside Survey 1990 (Barr *et al*. 1993). Each pair has been divided into a *trial* and a *check* square, to allow the development, refinement, and validation of methods and their subsequent 'blind testing'. Analysis has centred on 1 km squares using integrated CASI and LIDAR data acquired in summer 1999. Additionally, for one site in the Arable Landscape, a comparison is being made with CASI imagery from summer 1998; and the 1 km square is being studied in the context of the surrounding 3 x 3 km area.

This report is the fifth of a series of Interim Reports on the CASI-LIDAR Module of CS2000 (see Fuller *et al.* 1998, Hill *et al*. 1998, 1999, 2000a) and covers work between June and September 2000.

3. DATA PROCESSING FLOWLINE FOR THE TRIAL SQUARES

The development of a processing flowline for operational image analysis has constituted a major part of this project. Methods development focussed on the four trial squares, and is described in detail in Hill *et al*. (2000). The image analysis flowline involves the following procedures:

3.1 **LIDAR data pre-processing**:

(a) creating a Digital Surface Model (DSM) from the 'first pulse' LIDAR point sample information;
(b) creating a Digital Terrain Model (DTM) by the removal of surface height features such as buildings, trees, etc., and interpolation across the resultant data gaps;
(c) deriving height data for surface features by subtracting the DTM from the DSM elevation values;
(d) manual cleaning of height data due to residual edge-of-flightline data errors.

3.2 **CASI image pre-processing**:

(a) image normalisation (based on averaging) to reduce the effect of atmospheric 'noise' across and between flightlines;
(b) geometric correction involving the registration of individual CASI flightlines to corresponding LIDAR DSM data and the mosaicking and trimming of adjacent registered flightlines to create 1 km square data-sets;
(c) normalisation of spectral response between 1 km image mosaics, for squares with 'roll-over' training data;

(d) image segmentation by a process of edge-detection and region-growing, to derive the parcel boundaries to be used in the classification procedure.

3.3 **Image classification**:

(a) creation of a vector data-base within Laser-Scan IGIS software by raster-to-vector conversion;
(b) training the classification on known examples of individual land-cover types;
(c) amalgamating the training data into spectral sub-classes;
(d) applying the maximum likelihood algorithm to classify each parcel, based on mean spectral statistics from a shrunken area in all 12 CASI wavebands.

3.4 **Knowledge-Based Corrections (KBC**):

(a) the phase-1 KBC procedure operates per-parcel using rules based on a combination of context, LIDAR height data, class probabilities, and (in the case of sub-classes of Broad Habitats 3 and 17) CS 1990 reporting codes;
(b) the phase-2 KBC procedure involves similar correction rules, but operates first on a per-pixel basis, and subsequently after aggregating all contiguous parcels of the same Broad Habitat type.

3.5 **Validation of classification output by comparison with field survey data**:

(a) editing the digital field survey data to adjust for differences in the timing of data collection, the alignment of the two data sets, distinctions between land-use mapped in the field survey and land-cover mapped in the airborne imagery, and the representation of linear features;
(b) calculating correspondence (at a 1 m spatial resolution and at the Broad Habitat level) between the classified and corrected airborne data and the field survey data.

Comparison of the classified airborne imagery and the relevant portions of the Land Cover Map 2000 will involve the same method as the validation exercise.

4. APPLICATION OF PROCESSING FLOWLINE TO THE CHECK SQUARES

Having established a data processing flowline using the trial squares, the test of its operational capabilities was in its application to the check squares. Of interest was the wider applicability of the techniques and of variables such as segmentation thresholds, classification training data, and KBC rules.

The LIDAR pre-processing and CASI flightline normalisation involved running a series of software applications so the methods developed for the trial squares could be applied directly and objectively. Cleaning the LIDAR height data to remove the slight edge-of-flightline errors and registering the CASI flightlines to corresponding LIDAR DSM data to remove geometric errors, however, were manual processes requiring considerable operator interaction. The techniques were readily applicable, but the cleaning and registering were unique to each flightline and so there was no ‘correction algorithm’ that could be transferred from one image file to another. Recent developments in the Environment Agency image acquisition system should eliminate data quality problems (and therefore remove this highly interactive pre-processing phase) from any repeat exercise.

An additional stage in the processing flowline was investigated during the check square analysis. Topographic variation influences the spectral response of ground features recorded in CASI imagery. This is because undulating terrain is illuminated differentially according to whether facets of terrain are horizontal, face the sun, or face away from the sun (potentially shaded from direct solar illumination). Only in the case of the Marginal check square was the nature of relief in the central 1 km square considered significant enough to warrant attempted topographic-illumination correction. This was carried out using software developed by Cambridge University Geography Department and used operationally in creating the Land Cover Map 2000 (Fuller *et al*. 1998). Differential illumination across the landscape, and its consequent effects on the radiation recorded by the CASI sensor, were modelled using a smoothed version of the LIDAR-derived DTM and compensated for in the topographic correction software.

The image segmentation process was performed in two phases, the first phase used default input variables, but the second required two input thresholds to guide the processes of segment generation and region merging. Optimal values for these thresholds were derived for the trial squares, and these were found to be directly applicable to the check squares.

Since one of the objectives of this project was to investigate the transfer of training data from trial to check squares, it was necessary to normalise the spectral data of each check square to match its corresponding trial square. This was achieved by shifting the mean radiance value in each waveband of the check squares to match those of the trial squares. An inherent assumption in this procedure was that the type and proportions of land cover were similar between different sites. This restricted the transferability of spectral characteristics. Thus, for the Arable and Pastoral sites (which have a mixture of grassland and agricultural land-cover types), the check squares were normalised to the trial squares, creating a set of four 1 km^2 images in which the combined classification training data for the two trial squares could be applied to the check squares. The land cover of the Marginal sites was too distinct to allow their normalisation with the Arable and Pastoral sites, but did allow the normalisation of the check square to the trial square enabling the roll-over of the classification training data. For the Upland sites, however, the land-cover of the check square was very different to that of the trial square, and so these images were not normalised.

After the vectorisation of the segmented images and import into Laser-Scan IGIS software, the check squares were ready for parcel-based classification. For the Arable and Pastoral sites, the check squares were classified together using the combined training data from the trial squares. The training data were applied with only minor modifications, inserting three additional land-cover types not present in the trial squares (arable field beans, arable lucerne, and calcareous grassland. For the Marginal site, the check square contained fewer land-cover types than the trial square, and so the training data for those land-cover types not present in the check square were not included in the classification. The two Upland sites had such different land-cover that no attempt was made to transfer training data from the trial square to the check square. Instead the check square was trained independently using examples of known land-cover type. The classification procedure was performed in the same way as for the trial squares, with the Maximum Likelihood algorithm applied to each parcel, using mean statistics from a shrunken polygon to select the most likely class in statistical terms.

The KBC rules were also applied directly to the check squares, with only minor alterations required. This involved: slight changes to the height thresholds used for correcting woodland / hedge classification; the addition of extra rules to address problems of water mis-classification in the Arable and Pastoral check squares; and the removal of rules relating to classes not present in the Marginal check square. No KBC rules were applied to the Upland site.

Validation of the classification output was by comparison with the digitised CS2000 Broad Habitat data. As with the trial squares, a degree of editing was necessary to make the two datasets more comparable. Firstly, 1999 field reconnaissance data were used to update the distribution of arable fields in the 1998 field survey data. Secondly, all boundaries identified in the field survey GIS linework as hedges or walls were ‘burnt into’ the 1 m rasterised grid data as features with a real width. The inserted hedges were assigned to BH 1 (Broadleaved, mixed and yew woodland), and the dry stone walls to Broad Habitat 3 (Boundary and linear features). Individual trees, identified in the field survey, were also inserted into the 1 m rasterized grid data. Thirdly, the CS2000 field survey data were edited to improve the registration with the classified airborne data.

To provide results that would be directly comparable with the trial squares, validation was carried out comparing the product of full KBC against field survey data *with* and *without* the hedges/trees inserted and registration.

5. RESULTS OF CLASSIFICATION

5.1 Arable, Pastoral and Marginal squares

The classification of the check squares was notably worse than the trial squares. This was not surprising, given the basic nature of the normalisation procedure applied to the check squares to allow the application of trial square training statistics. The mis-classification was mostly a higher occurrence of the same errors recorded in the trial squares. Exceptions to this were shadow and water (which produced a wider range of classification errors), and the mis-classification of dry stone walls in the Marginal check square. The need for post-classification knowledge-based correction was therefore greater for the check squares than for the trial squares. Application of the KBC rules developed for the trial squares (with only slight modifications) resulted in correspondence values between the classified airborne imagery and the raw field survey data of 70-78% (Table 1).

The correspondence between the classified airborne imagery (after 2-phase KBC) and the field survey data (with edits) was lower for the check squares than for the trial squares by an average of 10 percentage points (Table 1, Figure 1). The discrepancy between the check and trial squares was least for the Pastoral site (correspondence of 80.1% and 86.7% respectively), increased for the Arable site (77.4% and 89.0% respectively), and was greatest for the Marginal site (68.7% and 80.9% respectively). The residual errors in check square classification include confusion between: improved grassland and young green crops such as rape or lucurne; improved and neutral grassland; arable harvested and hay fields; deciduous and coniferous woodland; arable bare and built surfaces. In addition, subtle differences occurred between the land-cover classified in airborne imagery and land-use identified in the field survey data. This was especially the case in Broad Habitat 17 (Built up areas and gardens) which need not necessarily have a built surface land-cover. There were also less explicable classification errors such as a field of improved grassland classifying as built surface in the Pastoral check square. Although not accounting for the above example, it should be noted that the field survey has been shown to have a repeatability level of 88% in identifying primary land cover codes, which are used objectively to generate Broad Habitat data. Thus, the field survey data should not be viewed as absolute ground-truth data with which to compare the accuracy of airborne image classification.

Landscape type	Classified airborne imagery *vs* unedited field survey data	Classified airborne imagery *vs* edited field survey data
Arable trial square	86.6%	89.0%
Arable check square	76.3%	77.4%
Pastoral trial square	83.5%	86.7%
Pastoral check square	78.0%	80.1%
Marginal trial square	79.7%	80.9%
Marginal check square	69.9%	68.7%
Upland trial square	71.0%	-
Upland check square	69.1%	-

Table 1. Correspondence between CASI-LIDAR image classifications (after Phase-2 KBC) and CS 2000 field survey data.

(Note 1: edits to the field survey data involve the insertion of boundary features with a given width, and slight registration shifts – both sets of field survey data include 1999 land-cover up-dates).
(Note 2: the correspondence figures treat Broad Habitats 13 (standing open water and canals) and 14 (rivers and streams) as one water class.

Editing the field survey data to add in the hedgerow boundary features increased the correspondence with the fully corrected (i.e. KBC phase-2) imagery for the Arable and Pastoral check squares by one and two percentage points respectively. However, for the Marginal check square, the addition of the boundary dry stone walls into the field survey data decreased the correspondence by one percentage point. This highlights the ability of integrated CASI and LIDAR data to identify hedges and treelines but not dry stone walls, which in the present data-set are beyond the accuracy of the integration process (Figure 1).

5.2 Upland check square

Classification of the Upland check site was not performed using trial square training data, and no KBC rules were applied. Correspondence with the field survey Broad Habitat data was very similar for the trial and check squares (71.0% and 69.1% respectively), thereby demonstrating an obvious consistency in Upland habitat mapping from airborne imagery. The classification errors in the check square relate, in part, to residual errors in geo-registration resulting in discrepancies in location of habitat boundaries in the field survey and airborne image data. However, the airborne data reveal a much finer spatial mosaic of habitats than the field survey polygon boundaries portray (Figure 2). Hence, whilst part of the discrepancy between the field survey and airborne data relates to image mis-registration and mis-classification, part also relates to a degree of detail greater than the field surveyors could be expected to achieve in what can be extremely difficult environments to map. It should be remembered that the field survey trials in 1997 showed that ground-survey could not repeatably record boundaries in the heterogeneous and continuously variable vegetation of semi-natural uplands and it was agreed to adopt 1990 boundaries (themselves subject to the same errors) unless changes were clearly evident. It was felt that proportional cover was determined reasonably accurately, but not the exact distributions. It seems likely that the differences between field and airborne surveys may relate largely to this problem and that the airborne survey might be the better record of spatial patterns.

6. ADDITIONAL ANALYSES FOR THE ARABLE TRIAL SQUARE

Given the time constraints placed on this project by the early problems with data quality and the lengthy geometric correction process required for the CASI imagery, it was decided to explore more detailed issues in one site only (Hill *et al*. 2000b). Thus, for the Arable trial square, additional topics for investigation include:

- independent training, classification and validation of 1998 and 1999 CASI imagery, enabling a comparison of land-cover statistics between corresponding years;
- roll-over of 1998 training data to 1999 CASI imagery (after normalisation) to examine the inter-annual transferability of training statistics;
- classification of a 3 x 3 km area for 1998 CASI data to allow the analysis of the central 1 km square in the context of the surrounding countryside;
- investigation of the additional uses of LIDAR data in image classification, particularly for woody vegetation types, and use of a more object-oriented approach.

The independent classification of the central 1 km square in 1998 and 1999 CASI data was reported in Hill *et al*. (2000b). Correspondence of the classified images (after phase-2 KBC) with field survey data (edited for land-cover change, the insertion of hedges and scattered trees, and mis-registration) was 86% and 89% for the 1998 and 1999 CASI data respectively. The processes of registering and mosaicking the 1998 CASI flightlines to build a 3 x 3 km square have been completed, which will enable the more routine processes of segmentation, vectorisation, classification (using the central 1 km square training statistics) and knowledge-based correction to be performed. Once the cross-over classification of the 1999 CASI data with the 1998 image training data has been carried out, it will be possible to make comparisons of land-cover statistics for the various data-sets generated for the Arable trial square.

Work has also begun investigating the potential use of LIDAR data for sub-dividing the ‘woody vegetation’ classified in the 1998 data for the Arable trial square. A two-stage process has been developed, operating firstly at the pixel level by a buffering process, and secondly by a more object-oriented approach at the parcel level (after vectorisation), which splits the ‘woody vegetation’ class into: areas of woodland or scrub; isolated trees or clumps; linear woody features (*i.e.* hedgerows and treelines). In addition, the location of trees within linear woody features has also been identified as a further level of sub-division (Figure 3).

The per-pixel approach involved shrinking a mask of the woody vegetation class, to a point that removed all scattered trees and linear features, and subsequently re-growing the mask remnants outwards guided by a height threshold. This identified patches of woodland and scrub from scattered trees, hedges and treelines. Within these latter woody vegetation types, trees could be discriminated from hedges by a greater width and height. The next phase of image analysis was therefore to create a ‘width image’ from the remaining woody vegetation mask (by calculating a focal sum in a moving pixel window) and apply a threshold to both the ‘width image’ and height image. This identified the approximate centroids of trees. Buffering these tree centroids outwards within the area covered by the woody vegetation class allowed the separation of trees. By subsquently vectorising these data, it was possible to separate trees in hedgerows or treelines from scattered individuals or clumps by the analysis of surrounding land-cover classes.

For the Arable trial square it has therefore become possible to calculate the proportion of woody vegetation in patches of woodland and scrub, or as linear features, or as scattered trees.

It is also possible to identify the proportion of linear woody features that is composed of trees. Further investigation will examine whether the use of per-parcel height statistics (e.g. mean, maximum, variance) can be used to further sub-divide the woody vegetation classes; for example, separating woodland from scrub, or possibly deciduous from coniferous woodland.

7. PROGRESS MEASURED AGAINST SCHEDULE

Work is currently on schedule and should be complete by March 2001 (see updated GANNT). Methods development and refinement using the trial squares and 'blind testing' on the check squares is complete. Creation of a 3 x 3 km integrated data set for the Arable trial squares is complete, as is the classification and validation of 1998 1km CASI data for this site. A more detailed and part object-oriented analysis sub-dividing woody vegetation types has demonstrated positive results that will form the focus of a short paper for the *International Journal of Remote Sensing* to be submitted by the end of 2000.

Remaining work includes: the classification of the 3 x 3 km data set for the Arable trial square using training statistics already derived for the central 1 km square; comparison of the classified airborne imagery with Land Cover Map UK for all eight sites; the analysis of land-cover statistics for all eight sites and the various test classifications for the Arable trial square; writing the Final Report, from which the principal project publication will be extracted.

8. CONCLUSIONS

This project has a very strong element of research and development, containing several novel aspects that have posed a unique set of challenges:

- The integration of multi-spectral CASI imagery with LIDAR elevation data to derive information on landscape features and structure.
- The use of remotely sensed data products gathered as part of an operational airborne remote sensing programme, rather than under research specifications.
- The use of automated image segmentation procedures (e.g. per-parcel mapping, object-orientation) for analysis of airborne images.

The results of image classification (both for the trial and check squares) are influenced by the use of CASI and LIDAR data acquired before integration had become operational, and by a project design which was restricted to a limited number of study sites due to a high R & D content. Hence, the issue of CASI image registration influenced the accuracy of data integration, whilst the need for spectral normalisation within and between flightlines restricted the transferability of classification training data. In addition, the sample size of one trial square for each Landscape type could not be expected to supply a comprehensive set of training data for application in blind trials. A more realistic operational procedure would be to build a 'spectral library' of training data from several sites, adding new (or distinct) examples of land-cover where encountered but not re-training previously occurring examples. This would avoid issues such as the spectral effects of phenological, soil-background, or species differences between sites. The results of image classification must therefore be viewed in the above context, but nonetheless, demonstrate very strong potential for airborne scanner applications to supplement field survey data in future Countryside Surveys.

Methods development (in terms of both image acquisition and processing) has constituted a major part of this project for both the EA and CEH. The requirements of data supply for the operational application of integrated CASI and LIDAR data in Countryside Survey have become clear, and this has fed back into Environment Agency image acquisition systems. Recent developments by the EA have included the simultaneous acquisition of CASI and dual-pulse LIDAR data and an integrated processing system. This is currently generating geo-registered CASI and LIDAR data for test sites. By recording both 'first' and 'last' pulse LIDAR response, both Digital Surface and Digital Terrain models can be created directly from the LIDAR point sample data. This would remove the lengthy processes outlined in this project of image registration for CASI data and surface feature removal in the LIDAR data. In addition, the NERC Airborne Remote Sensing Facility is currently developing a CASI image acquisition and processing system (for completion by March 2002) that will involve geometric and radiometric correction (removing geometric distortions and compensating for illumination conditions, viewing geometry and atmospheric attenuation). This would enable the immediate integration of CASI imagery with LIDAR data and remove problems of spectral normalisation across flightlines and study sites. The transfer of spectral training data between Countryside Survey squares would then become a realistic proposition within an operational process.

The results of image classification thus clearly show the potential of airborne scanning in future Countryside Surveys, especially in light of recent developments in image acquisition. The airborne imagery provides complete and simultaneous coverage of the entire CS square (and potentially the surrounding countryside). The level of detail present within the airborne imagery is greater than field surveyors could be expected to map, and this if of particular significance in the continuously variable vegetation of semi-natural Upland Landscapes where field survey boundaries are known to be approximations. In addition, the airborne data provides detailed height data for surface features and landscape topographic information; which together with the land-cover classification could provide valuable input for hydrological models. However, it should be remembered that the airborne imagery provides land-cover rather than land-use information, and will always have difficulties in separating those land-cover types that require botanical examination for identification, such as grassland types. As such, airborne imagery should be seen as providing information to compliment the work of field survey, and to ease some of the burden of field surveyors, but not to replace them.

9. REFERENCES

CASI-LIDAR Reports:

Fuller, R.M., Hill, R.A. and Veitch, N. (1998). Countryside Survey 2000 – Part I. Module 8: Airborne Scanner Applications: Classification of airborne CASI and LIDAR data of selected CS2000 sample squares. *First Interim Report, CSCL/Int.1. Unpublished report to the DETR.*

Hill, R.A., Fuller, R.M., and Veitch, N. (1998). Countryside Survey 2000 – Part I. Module 8: Airborne Scanner Applications: Classification of airborne CASI and LIDAR data of selected CS2000 sample squares. *Second Interim Report, CSCL/Int.2. Unpublished report to the DETR.*

Hill, R.A., Fuller, R.M., & Veitch, N. (1999). Countryside Survey 2000 - Part I, Module 8: Airborne Scanner Applications. Classification of airborne CASI and LIDAR data of selected

CS2000 sample squares. *Third Interim Report, CSCL/Int 3*. Unpublished ITE report to the DETR.

Hill, R.A., Fuller, R.M., & Veitch, N. (2000a). Countryside Survey 2000 - Part I, Module 8: Airborne Scanner Applications. Classification of airborne CASI and LIDAR data of selected CS2000 sample squares. *Fourth Interim Report, CSCL/Int 4*. Unpublished ITE report to the DETR.

Hill, R.A., Fuller, R.M., & Veitch N., (2000b). *CS2000 Project Specification Module 8: Airborne Scanner Applications - classification of airborne CASI and LIDAR data of selected CS2000 sample squares* - a revised joint proposal to DETR.

Other references:

Barr C.J., Bunce R.G.H., Clarke R.T., Fuller R.M., Furse M.T., Gillespie M.K., Groom G.B., Hallam C.J., Hornung M., & Howard D.C. (1993) *Countryside Survey 1990: Main Report.* London: Department of the Environment.

Fuller, R.M., Smith, G.M , Hill, R.A., Thomson, A.G. & Gerard, F.F., (1998). C*ountryside Survey 2000 – Part III. Module 7. Land Cover Map 2000. First Interim Report incorporating the Second Quarterly Progress Report. CSLC/Interim 1* Unpublished ITE Report to the LCM2000 Consortium.

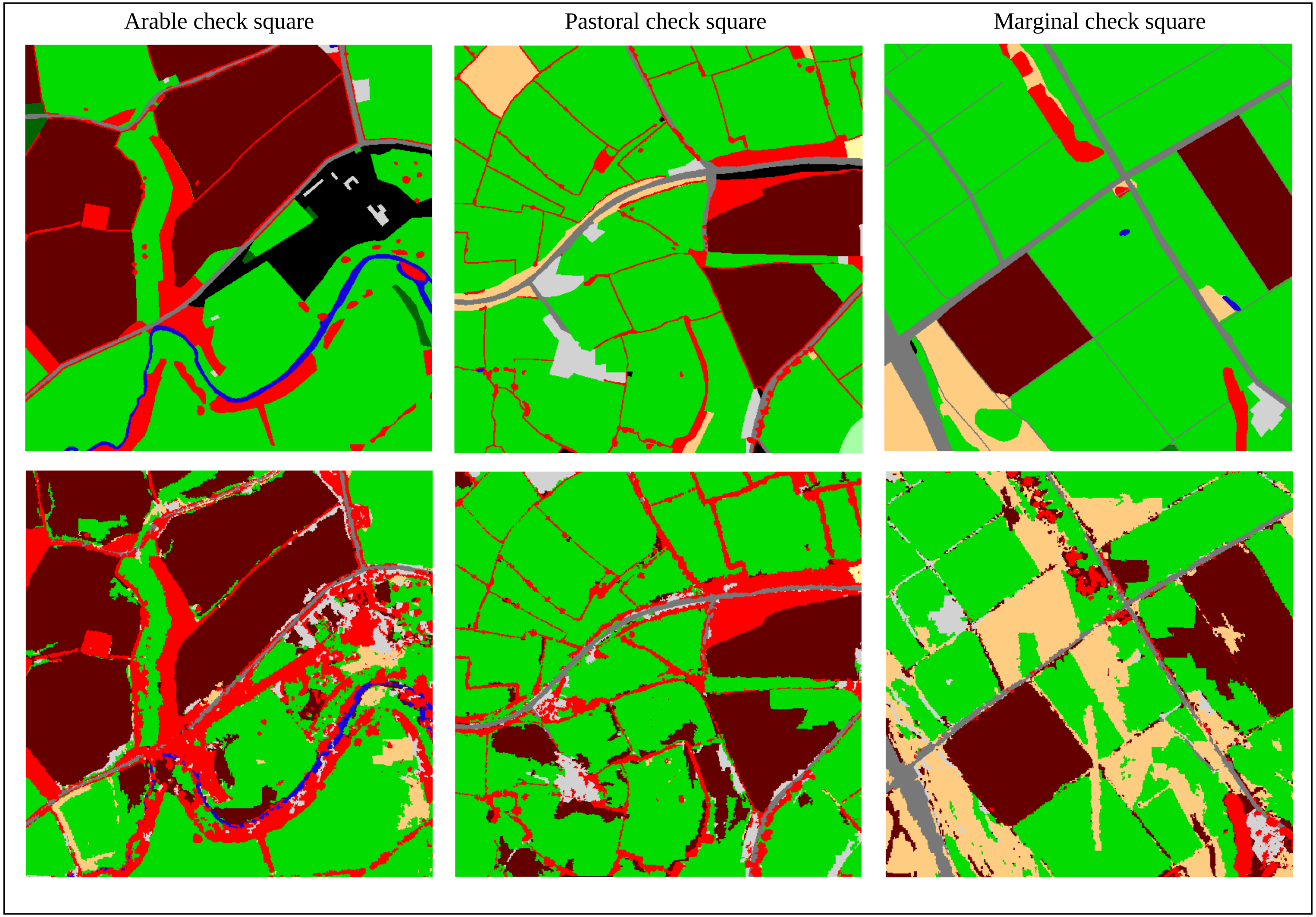
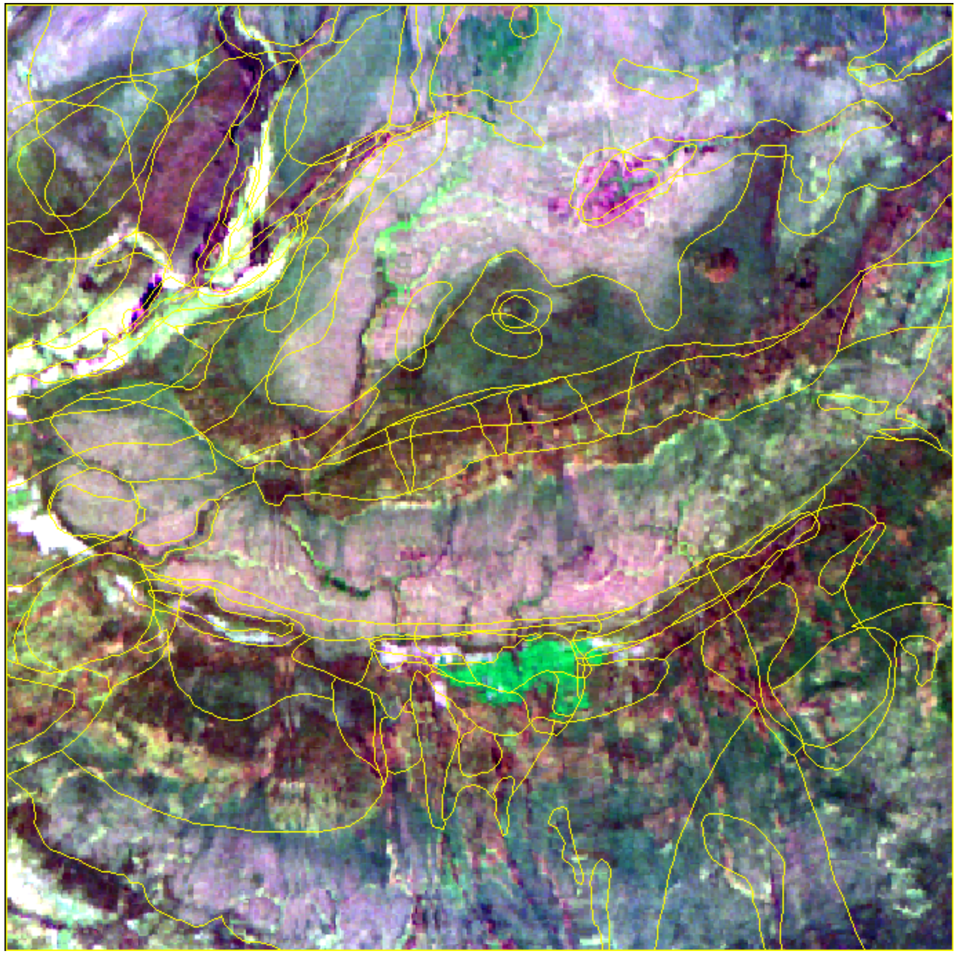
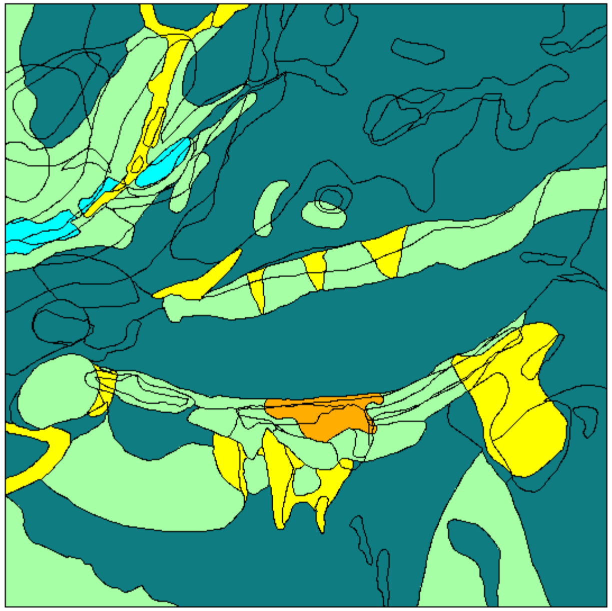


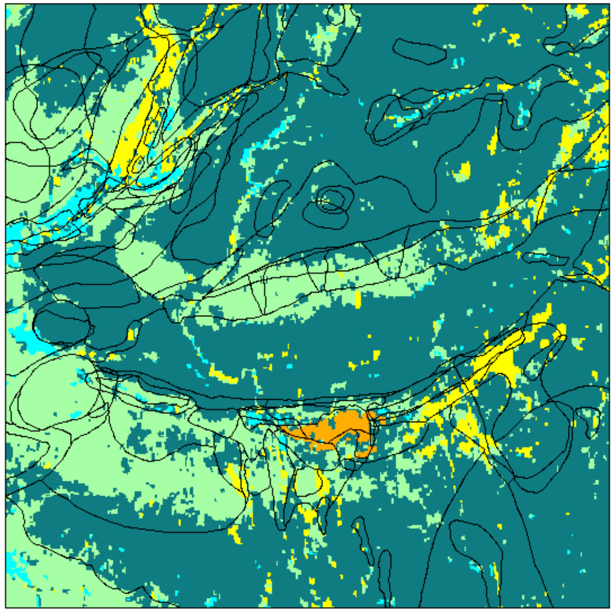
Figure 1. Comparison of edited field survey data (top) and the final product of classification and knowledge-based correction for the Arable, Pastoral and Marginal check squares (below). (NB black = no data)



CASI image (visible spectrum wavebands) with field survey boundaries overlaid

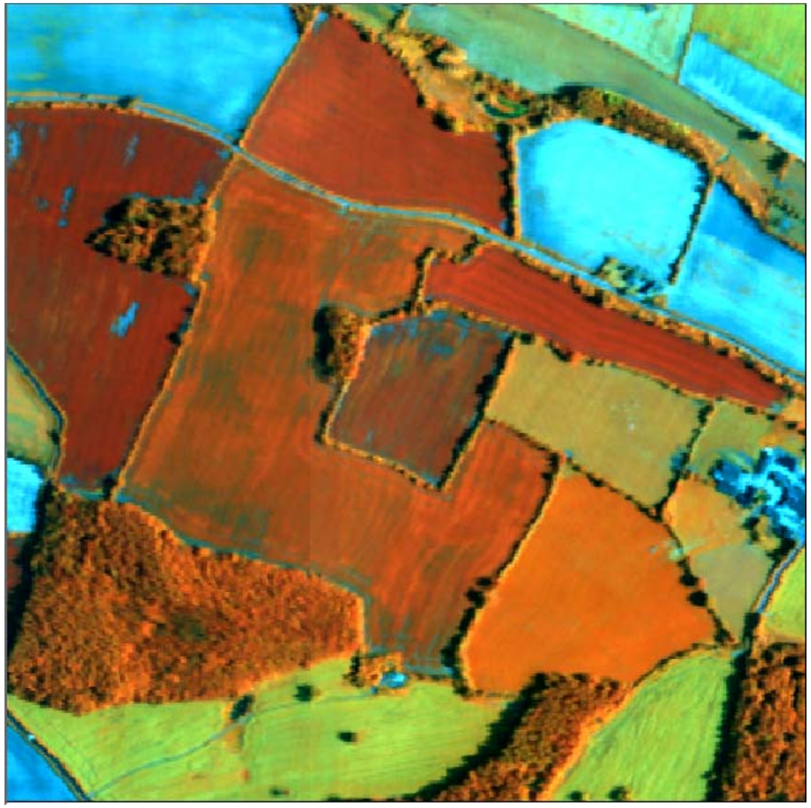


Field survey data

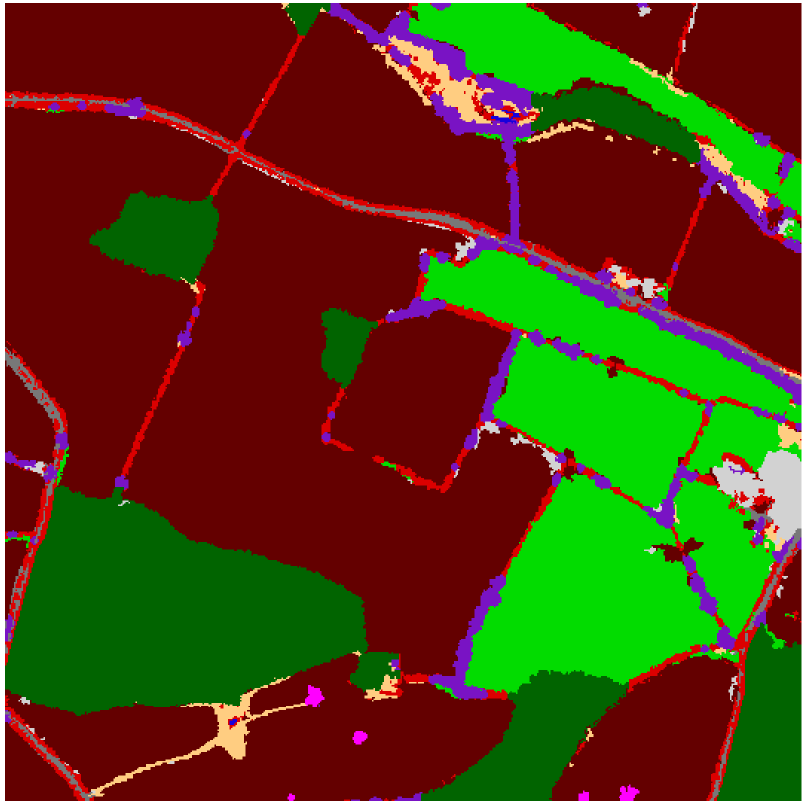


Classified CASI image showing field survey boundaries

Figure 2. The fine spatial mosaic of land cover shown in the CASI image compared with the field survey data for the Upland check square.



CASI image (false colour composite)



Classified image

Figure 3. Detailed classification of the Arable trial square in 1998 CASI data.

(Colour scheme for classified image is: brown = arable, mid-green = improved grassland, peach = neutral grassland, pale grey = built surface, dark grey = road, dark green= woodland, pink = scattered trees, purple = treelines & hedgerow trees, red = hedgerows).

Figure 4. Revised GANNT