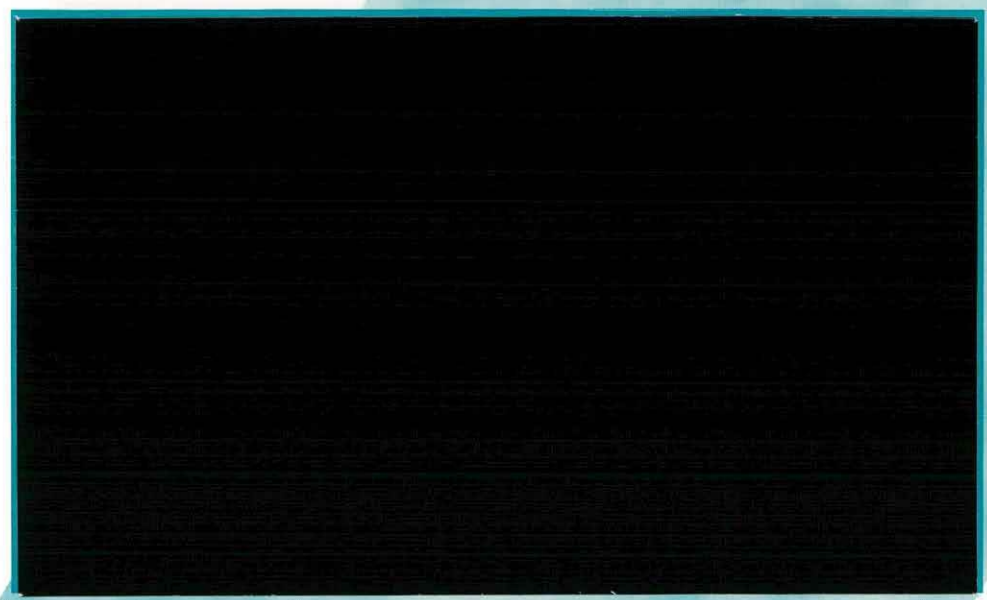
Institute of Terrestrial Ecology



Centre for Ecology & Hydrology

Natural Environment Research Council

ITE has six Research Stations throughout Britain, which allows the efficient use of resources for regional studies and provides an understanding of local ecological and land use characteristics. The Institute's administrative headquarters is at Monks Wood.

ITE sites

Monks Wood
(Admin HQ)
Abbots Ripton
HUNTINGDON PE17 2LS
Telephone 01487 773381-8
Fax 01487 773467
Email MONKSWOOD@ITE.AC.UK

Merlewood Research Station
GRANGE-OVER-SANDS
Cumbria LA11 6JU
Telephone 015395 32264
Fax 015395 34705
Email MERLEWOOD@ITE.AC.UK

Edinburgh Research Station
Bush Estate
PENICUIK
Midlothian EH26 0QB
Telephone 0131 445 4343
Fax 0131 445 3943
Email BUSH@ITE.AC.UK

Furzebrook Research Station
WAREHAM
Dorset BH20 5AS
Telephone 01929 551518-9, 551491
Fax 01929 551087
Email FURZEBROOK@ITE.AC.UK

Banchory Research Station
Hill of Brathens
Glassel, BANCHORY
Kincardineshire AB31 4BY
Telephone 01330 823434
Fax 01330 823303
Email BANCHORY@ITE.AC.UK

Bangor Research Unit
University of Wales, Bangor
Deiniol Road
BANGOR, Gwynedd LL57 2UP
Telephone 01248 370045
Fax 01248 355365
Email BANGOR@ITE.AC.UK

Details about the Institute are available on the Internet via the World Wide Web (http:/www.nmw.ac.uk/ite)

INSTITUTE OF TERRESTRIAL ECOLOGY
(NATURAL ENVIRONMENT RESEARCH COUNCIL)

Project. T02083j5/C00878

Countryside Survey 2000
Module 7

LAND COVER MAP 2000

FOURTH INTERIM REPORT

incorporating the

Eighth Quarterly Progress Report

CSLCM/Int4/Prog8

R.M. Fuller, G.M. Smith, J.M. Sanderson,
R.A. Hill, M. Hall, A.G. Thomson & N.J. Brown
Institute of Terrestrial Ecology
Monks Wood
Abbots Ripton
Huntingdon
Cambs PE17 2LS

Corresponding author:
R.M. Fuller
Section for Earth Observation
Institute of Terrestrial Ecology
Monks Wood
Abbots Ripton
Huntingdon
Cambs PE17 2LS

Telephone: 01487 772519
Fax: 01487 773467
Email: rf@ceh.ac.uk

6 March 2000

MAIN REPORT

EIGHTH QUARTERLY PROGRESS REPORT AND EXECUTIVE SUMMARY TO THE FOURTH INTERIM REPORT

- This Executive Summary to the Fourth Interim Report on Land Cover Map 2000 (LCM2000) also serves as the Eighth Quarterly Progress Report. The Report covers work done to 25 February 1999.
- LCM2000 is making a census survey of the widespread Broad Habitats of the United Kingdom using satellite imagery and automated image processing techniques to map target classes with a classification accuracy of 90%.
- Image purchases now cover, 90% of the UK as summer-winter data, with 96% covered by a single date of imagery. Due to poor weather, no summer scene was recorded within the ideal target period of June / July 1998. Only 28% of winter imagery was in the target period, namely October 1997 to March 1998, excluding the 3-4 weeks either side of the winter solstice. Image purchases were either compromises on the seasonal definitions, or they were based on 1999 coverage, missing the field survey date except in parts of Scotland. Image coverage of northern Scotland is still waiting for new acquisitions in winter 1999-2000.
- To compensate for part-clouded scenes, the estimated number of images required has increased well above the 43 scenes used for mapping Britain in 1990, to 69 scenes for the UK in LCM2000 (64 of which are in GB).
- The mosaic of summer-winter composite sections used to map the UK is much more complex than expected. LCMGB 1990 mapped Britain in 32 sections; the LCM2000 processing job will be twice that. The workload substantially exceeds the 'worst case scenario' anticipated in contingency plans: these allowed, as a maximum, 54 mosaic sections for full UK coverage. In practice, at least 61 will be required.
- By increasing staff-inputs, adding extra processors and software, and streamlining procedures, much of the extra workload has been accommodated. However, completion by the time of the launch is in question.
- Completing England and Wales would require 26 additional sections - more than is possible, even allowing that nearly half the pieces are very small infill sections; and, because much time would be spent on these small sections, UK coverage would by then only be 67%.
- If we were to work with the biggest summer-winter sections first, coverage could, by the time of the launch, extend to 80% of the UK, but with no one country complete. At that point, the remaining 20% of the UK would require 33 mosaic sections - still 54% of the total number. However, these would be processed at the rate of one per week, extending the project conclusion to the end of May 2001. However, as the last sections to be mapped will be very small indeed, coverage should, by March 2001, be 99%.
- It is considered that high quality outputs are of paramount importance. The LCM2000 GIS will see routine use for probably a decade and, for change detection purposes, perhaps for a long time thereafter. A delay of a few months in completion, while unfortunate for the purposes of the launch, will have relatively little impact on operational uses. While corners could be cut to achieve a timely completion, it is suggested that this would not achieve the quality demanded by the Consortium members.
- If LCM2000 production overshoots the launch, there would be ample demonstrable output: it would be possible to drop the 80% coverage onto LCMGB and show the 'national' picture: this approach would be no different to OS mapping, which commonly shows partial revision. Local and regional detail from 1998-99 would be widely available,

both as maps and data. Validation data would, however, be incomplete and we would need to talk in general terms about accuracy. National statistics would need to come from field survey estimates: but this would overcome the risk of any conflict. The timescale was discussed with the Advisory group on 2 March 2000; a proposed revision to the milestones (Annex C to the contract) is given in Annex I.

- Despite problems which will extend the size of the processing job and delay its completion, other issues have gone to plan.
- Procedural developments continue but now at a reduced level. These are being made to improve performance and address any new problems which arise.
- The widespread Broad Habitats are now fixed in type and definition.
- Sample field reconnaissance data cover *c.* 85% of the UK, with north Scotland and Northern Ireland to be covered in May-June 2000.
- The classification procedure is finalised. This uses the same maximum likelihood algorithm as does per-pixel classification. However, it applies the procedure to the polygon, using mean statistics for the raster data within the polygon to select the most likely class in statistical terms. CLEVER-Mapping in IGIS records the probabilities for the top five subclass options. Per-pixel classifications are being made as a 'standalone' product and also to record the natural heterogeneity associated with CLEVER-Mapped polygons.
- A combination of class probabilities, internal context and external data is being used in knowledge-based correction.
- Initial 'validation' has used field reconnaissance data, compared with resulting classifications, to score correspondence per-class. Results have been weighted according to class coverage per scene. Overall correspondences range between 87-93% at Broad Habitat level, prior to the final stages of knowledge-based correction.
- Independent validation will use squares from the CS2000 field survey. The first examples of these have just been delivered and methods development and testing has commenced.
- In conclusion: the significant production issues have all been fully addressed; the pre-processing steps in operation are more elaborate and more effective than had been expected at the start of the programme; all the issues concerning segmentation, per-segment classification and edge-matching of segmented classifications have been addressed and made operational; widespread Broad Habitats have been shown generally to be mappable and with the target 90% accuracy, measured per-segment; furthermore, the detail provided by subclasses and variants builds much more flexibility into the dataset than does the Broad Habitat listing, better serving the diverse needs of the Consortium partners and other potential end-users.
- Only the time schedule is under threat. To ensure maximum throughput, there are 7 staff working on LCM2000, generally 4-5 at any one time; a fourth SUN workstation has been dedicated to LCM2000, 2 mostly on pre-processing and two on classification. It is expected that 99% of the UK will be mapped by the project's intended end-date on March 2001, with just small sections to be completed over the following two months.

1. INTRODUCTION

1.1 Context

Land Cover Map 2000 (LCM2000), a part of Countryside Survey 2000 (CS2000), will provide a census of the countryside of the United Kingdom. LCM2000 outputs will be in the form of digital maps and databases, plus a range of derived products held in a geographical information system (GIS). A Consortium of eleven Departments and Agencies are co-funding LCM2000 (see previous reports).

1.2 Background

LCM2000 updates and upgrades the Land Cover Map of Great Britain (LCMGB), made in 1990-92 (Fuller *et al.*, 1994). Refinements include:

- Improved accuracy of classification;
- Added thematic detail;
- Compatibility with other systems of environmental survey and evaluation;
- Closer integration between field and satellite data.

1.3 Aims

The aims for LCM2000 are:

- To undertake a census survey of the land cover / widespread Broad Habitats of the UK at the turn of the Millennium;
- To apply the most appropriate satellite imagery and automated image processing techniques to achieve a classification accuracy of 90% for target classes;
- To produce and make available, under licence, a range of geographically referenced data outputs on land cover characteristics, tailored to the needs of Consortium members;
- To calibrate and validate satellite-derived classifications against ground reference data, publish results of the correspondence analyses, and provide a guide to their interpretation.

This Report is the fourth six-monthly *Interim Report* on LCM2000 (see Fuller *et al.* 1998a, 1999a and 1999c) and includes, in its Executive Summary, the eighth *Quarterly Progress Report.* It covers work done up to the 25 February 1999. It is part of a series of reports on LCM2000, which are listed in the references.

2. REFERENCE DATA FOR MAP PRODUCTION

2.1 Images

Earlier image acquisitions were described in previous reports (Fuller *et al.* 1998a, 1999a & 1999c). Purchases have built image coverage for England, Wales, southern Scotland and

Northern Ireland that is essentially complete, with *c.* 90% of the UK covered in both summer and winter (Figure 1a) and 96% covered on at least one date (Figure 1b). However, not a single summer scene was recorded within the ideal target summer period of June / July 1998. Only 28% of winter imagery was in the target period, October 1997 to March 1998 but excluding the 3-4 weeks either side of the winter solstice. The chosen images were either compromises on the seasonal definitions, or they were covered in 1999, mismatching the field survey date (except in Scotland where field survey stretched into 1999).

Aberdeenshire still records incomplete cover in both summer and winter. Shetland relies on very hazy 'winter' cover of 20 April 1999 and equally hazy 'summer' cover of 13 May 1999 - substantially stretching the target seasonal dates. North-west Scotland was covered by summer images but the winter images were recorded at 10.30 GMT on 22 December 1997, consequently with very poor light levels, especially on shaded north-west facing slopes. Central northern Scotland also relies on midwinter data for full multi-temporal coverage. With 4 suitable satellites now operating, compared with just 2 until summer 1999, we have opted to wait for new coverage of northern Scotland in winter 1999-2000. As a result, it was agreed with the Consortium (at the meeting of 3 November 1999) that ITE should leave Northern Scotland acquisitions until the spring of next year.

To compensate for part-clouded scenes, the estimated number of images required has increased, well above the 43 scenes used for GB in 1990, to 69 scenes for the UK in LCM2000 (64 of which are in GB). The increased costs were anticipated in the Specification and helped by an 80% reduction in the price of recent images with the advent of Landsat 7. With image acquisitions near complete, and most areas covered by purchased images which can be inspected in far greater detail than web-based 'quick-looks', it has been possible to take stock and review the likely consequences of poor image coverage in 1998-99.

The worst impact of the poor image quality has been that the mosaic of summer-winter composite sections is much more complex than expected (Figure 2). LCMGB 1990 used 32 sections to cover GB. Contingency plans allowed an **absolute maximum** of 54 scene-pairs to make full coverage. LCM2000 will in fact require approximately 61 sections (assuming cloud-free purchases for the remaining Scottish coverage). This difficulty reflects the extremely poor summer of 1998 and the mediocre conditions for winter and summer imagery since. The fact that most 1998 summer scenes were based on early to mid-May images, with known problems for the separation of arable crops, woodlands and grasslands, has also demanded much more painstaking attention to training, classification and knowledge based correction. Furthermore, other areas with summer and winter data which did not record the same growing season also offered significant and unwanted challenges.

It can be seen that the processing job is twice that of 1990; and because every section needs to be segmented and classified individually, even single-date patches need processing through much the same sequence as two-date composites. The result is that the workload substantially exceeds the worst expectations. By increasing staff-inputs, adding extra workstations and software, and by streamlining procedures, much of the extra workload has been accommodated. However, it has been made clear that completion by the time of the launch is in question (Fuller *et al.* 1999d). Given sight of the new images, it has been possible to estimate more accurately the true production timescale. In so doing, several observations and assumptions are made:

- There are 7 staff working on LCM2000, generally 4-5 at any one time;
- There will, from now on, be four workstations dedicated to LCM2000, two mostly on pre-processing and two on classification;
- Classification is the rate-determining step: it now takes about 4 weeks to classify a typical scene-pair, or 2 weeks for single-date infill;
- With two workstations dedicated to classification, a piece can be added to the mosaic every 2 weeks for large summer-winter scenes, and every week when working on infill.

If image sections are processed in size-order, from largest to smallest and if pieces are added at the rate of one every 2 weeks then, by late September we might expect to add 15 further pieces prior to the launch which would not complete coverage of England and Wales as previously planned (Figure 3a). If we concentrate on completing England and Wales, it would require 26 sections (Figure 3b) - more than is possible, even allowing that nearly half the pieces are very small infill sections. And, because much time would be spent on these small sections, UK coverage by then would be only 67% (Figure 3a). If, however, we work with the biggest summer-winter sections first, coverage could, by the time of the launch, extend to 80% of the UK (Figure 3c) but with no one country complete. At that point, the remaining 20% of the UK would require 33 mosaic pieces - still 54% of the total number. However, these would be processed at the rate of one per week, taking the conclusion to the end of May 2001. As the last sections to be mapped will be very small indeed, coverage will be 99% by March 2001 (Figure 4), leaving less unclassified than the 2% which remained unclassified on LCMGB. A proposed revision to the milestones (Annex C to the contract) is given in Annex I to this report.

2.2 Broad habitat specifications

LCM2000 aims to map widespread examples of Broad Habitats, as defined under the Biodiversity Action Plan. A list of target cover classes was reported in detail (Fuller *et al.* 1998a). Amendments were agreed at subsequent Consortium Meetings, field-meetings with conservation agencies, plus reconnaissance surveys of summer 1999. A finalised classification has been provided by the Joint Nature Conservancy Council though full definitions are yet to be published. Further evolution of the classification accommodates new cover types and corrects misunderstandings of Broad Habitat definitions. The list of LCM2000 Target Classes appears in Table 1.

2.3 Ground reconnaissance data

Ground reference data collection (Fuller *et al.* 1999b) brings overall sample coverage up to *c.* 85% with the completion of Northern Ireland, north Scotland and the Scottish Islands timetabled for May and June 2000. Some 85 thematic subclasses have been recorded in field reconnaissance trips, each related to the Broad Habitats.

3. PRE-PROCESSING OF IMAGE DATA

The early investment of time in procedural developments is now benefiting productivity. Earlier reports outlined procedures for:

- The correction of atmospheric haze,
- Cloud and shadow masking within scenes,
- Geo-registration and resampling,
- Correction of differential illumination effects on undulating terrain,
- Band combinations for analyses.

Correction of atmospheric haze is working routinely and effectively. However, there are new scenes to enter the processing stream with far worse haze than scenes processed to date, where generally the haze has been almost invisible. These badly affected scenes, with streaks of haze, will prove a severe test for the procedure. If it fails, then dense haze will have to be treated as cloud and masked out, with appropriate patching.

Cloud-masking is also working well; however, on scenes where there are several layers of cloud at varying heights, the projection and masking of the position of the cloud-shadow has proved particularly difficult. Where necessary, manual intervention has been used to define reject areas.

The geo-registration process is testing the procedures further than before, as coverage moves into hilly terrain, often with co-registration of over-lapping summer and winter data recorded from adjacent flightpaths 90 km apart. Whilst difficulties have been encountered, solutions have always been achieved, if sometimes at the expense of additional, interactive, control-pointing. Typical root mean square errors on registration are <1 pixel (<25 m) and co-registration is achieved with generally no more than a pixel of visible misregistration between summer-winter paired scenes.

The correction of differential illumination in undulating terrain has worked well for summer data; terrain effects on summer scenes are barely perceptible after correction. Winter data, especially scenes made in mid-winter, can be corrected and visible bands may show little terrain effect. However, near infrared radiation is not scattered in the same way as visible light, and indirect illumination (from the blue sky) is minimal in infrared bands. Therefore compensation, which multiplies the recorded value by a correction factor, may be based on a zero response in the original dataset giving a zero value after multiplication. The consequence is that terrain facets may influence segmentation and that classification may need to use shaded and sunlit spectral variants (see Kershaw & Fuller 1992). The spectrally determined segments will be aggregated thematically, post-classification, where the cover is of the same type.

4. IMAGE SEGMENTATION

4.1 The segmentation software

Segmentation issues were discussed in the last Interim Report (Fuller *et al.* 1999c) and no developments have been made which need detailed reporting. All scenes classified to date

have continued to use one band (near infrared) from a winter scene and two bands (red and middle infrared) from a summer scene, with no need for alternatives. The Sobell edge detector ensures that the appropriate seedpoints are selected away from parcel-edges. Segmentation still aims to ensure a field-by-field segmentation, with a minimum of within-field division, except where this represents real heterogeneity at the subclass level. However, finer thresholds have sometimes been set to separate different zones within urban and semi-natural zones. Post-segmentation generalisation and boundary pixel rejection follows the rules outlined by Fuller *et al.* (1999c).
Once acceptable segmentations have been achieved, vector versions are created in a GIS database, and land parcel attributes, required in later analyses, are attached.

4.2 OSNI Vector Data

LCM2000 production in Northern Ireland aimed, if appropriate, to use Ordnance Survey of Northern Ireland (OSNI) vector data. Data assessments have shown that OSNI parcel sizes are often too small for Landsat spatial resolutions. Various options were considered: i. using image sharpening procedures (as in Jersey (Smith & Fuller 1998 and 'in press')) based on high spatial resolution panchromatic data; ii. using the OSNI polygons solely to create an edge image prior to image segmentation; or iii. using the same methodology as for the rest of the UK and attaching resulting land parcel labels back to the OSNI polygons, once the classification is complete.

Image sharpening procedures, using high spatial resolution panchromatic data, would have allowed mapping of parcels <0.1 ha. IRS panchromatic data, with 5.6 m pixels in scenes 140 km square, would cover Northern Ireland in 5-6 sections, and would offer a relatively cost-effective way of producing and analysing high resolution data of Northern Ireland. SPOT panchromatic images with 10 m pixels only give 60 km square coverage, requiring maybe 16 scenes. Extra scenes, in either case, would have meant extra costs of data and processing. Use of IRS might have doubled the production time; SPOT might have increased it tenfold. However, cloud-free IRS panchromatic data only cover about 10% of Northern Ireland, preventing further consideration. Use of SPOT data was excluded on the grounds both of high cost as well as poor image availability.

Further analysis of land parcel data (provided by the Department of Agriculture for Northern Ireland) showed in fact that 6-8% of the area of Northern Ireland farmland is in parcels below minimum mappable size. Conversely of course this means that over 90% of fields are mappable as individual entities. The small size of some fields does not mean they would be mapped into the wrong cover class: they would simply be aggregated with neighbouring fields, of the same spectral class, during segmentation. In fact, much bigger fields may also be aggregated at segmentation, if they are spectrally similar to a neighbour and with no defining boundary feature.

The conclusion regarding classification of Northern Ireland was that segmentation would be operated as elsewhere in the UK. The image-based segments, once classified, could readily be called upon to label the OSNI polygons, with high levels of accuracy. That could be done afterwards, by any one of the agencies involved.

5. TRAINING THE CLASSIFIER

Training for per-parcel analyses has been objectively based on land parcel cover data recorded in the field and subsequently related to image segments (Fuller *et al.* 1999c). Classification results from one image have 'rolled over' onto overlapping unclassified scenes. Thus, areas of overlap have been used to locate near-identical polygons (>80% overlap) on the neighbouring scene and to pick up *a priori* class labels for use in classification. Probability rules have rejected the use of labels which were attached with low probability ($P < 85\%$). Supplementary training data have also been added by image interpretation, checked against external map data (e.g. in definition of urban areas, woodlands, water bodies and saltmarshes). Training has attempted to identify at least 5 areas per subclass (usually many more), with all examples broadly similar in size, such that none dominates the training set; if possible, there should be at least 30 pixels per training area.

Classification has used the training polygons to derive statistical measures of reflectances, in each chosen band and for each spectral subclass. CLEVER-Mapping shrinks the polygon when extracting raster data, to avoid edge pixels and to ensure the use of 'pure' core pixels of a single cover type. The shrinkage is a dynamic process whereby the required amount of shrinkage is reduced in small parcels with an inadequate number of core pixels. The number of pixels extracted and classified and the shrinkage achieved are reported and stored on the land parcel for future reference. The review of the spectral characteristics of the training areas using 'image chips' (Fuller *et al.* 1999c) to represent the remotely sensed data has proved particularly powerful in helping to refine the training sets. It allows separation of spectral subclasses and comparisons between Broad Habitats to ensure that the subclass selection achieves maximum spectral separability and that the statistical algorithm, which requires normally distributed datasets, is operated within the prescribed rules (Kershaw & Fuller 1992). The procedure is contributing substantially to the intended quality improvements.

6. CLASSIFICATION

6.1 Maximum likelihood classification

The classification procedure uses the same maximum likelihood algorithm as does a normal per-pixel classification. However, it applies the procedure to the polygon, using mean statistics for the raster data within the polygon to select the most likely class in statistical terms. CLEVER-Mapping in IGIS records the probabilities for the top five subclass options, usually covering >90% of the probability distribution. This information has proved particularly powerful in later knowledge-based corrections. Per-pixel classifications are being made as a 'standalone' product and also to record the natural heterogeneity associated with land parcels.

6.2 Knowledge-based correction

Knowledge-based correction (KBC) procedures (Groom & Fuller 1996, Fuller *et al.* 1999c) are being tailored to particular circumstances. They are being used to identify and re-label land parcels which are classified with low confidence and / or with classes out of their natural context. Correction on a per-parcel basis allows more subtle and complex rules to be applied, providing a powerful tool for improving classification accuracy. The following KBC rules have been applied:

1. Where a subclass is allocated with a probability of <50%, the other subclass probabilities have been summed to see if any other Broad Habitat is more appropriate (i.e. its subclasses cumulatively take a greater percentage of the overall probability).
2. Coastal masking has been used, just as in 1990, to preclude the recording of 'inland beaches' and 'urban areas' on the shoreline (Groom & Fuller 1996).
3. The DTM has been used with a threshold to identify erroneous areas of 'Fen, marsh and swamp' off the floodplain.
4. Bare ground in the context of coniferous plantations, evidently felled, is recorded as such.
5. Arable land surrounded by urban cover is corrected to the suburban class.

A number of procedures will be used universally rather than being applied per-scene. Peatland drift will refine the classifications of bogs and heaths. Soils or geological maps will be needed to distinguish acid, calcareous and neutral grasslands; their availability needs investigation after it has become apparent that drift data do not generally distinguish soil acidity. CORINE Land Cover GB will be used to double-check the orchard class, often confused with grass, scrub or woodland. All contextual corrections have been recorded in the GIS database, to include the input and output classes of any changed polygons, and the rule(s) used in their alteration.

7. MAP OUTPUTS

Samples of output appear in the last Report. Further examples are available on the CS2000 website. This report gives samples of per-pixel and per-parcel classifications which have not appeared in earlier reports (Figure 5). The example illustrates well the much cleaner appearance of the per-parcel classification; and it is clear that these will provide valuable information on landscape structure of areal features and their boundaries. The Report also includes an example of segmentation in upland (Figure 6). There had previously been concerns that segmentation would fail in unenclosed land though, in practice, it has operated successfully in thousands of square kilometres of unenclosed semi-natural lowland (e.g. coasts, wetlands, heaths and forests). Figure 6 shows that this success applies in upland heaths and grass moors, with every prospect of accurate per-segment classification.

8. VALIDATION

8.1 Basic checks

Initial checks have used ground reference datasets to check provisional classifications during the iterative process of classification, review and re-classification. The annotated images from

the field reconnaissance have formed the first choice for checking. OS maps have also highlighted discrepancies in those classes which they depict (eg urban, water and woodland). Various external datasets have also been used to check final results. SNH have passed on various maps of peatland. CCW have provided grid-referenced examples of Phase I cover types. DETR have provided digitised urban boundaries for knowledge-based corrections and validation. An example for Cambridge overlay is in Figure 7. The main differences are that the DETR boundaries include airfields and similar, green, open areas of 'developed land'; also the DETR boundaries omit small settlements and exclude recent development. Most of the differences seen in Figure 7 are attributable to these circumstances, rather than LCM2000 errors.

8.2 Preliminary validation

Segments visited in the field reconnaissance (which may include some used as training areas) have been checked on Map outputs to provide an early assessment of quality. Results have been weighted according to class coverage per scene. Overall correspondences range between 87-93% at Broad Habitat level, prior to the final stages of knowledge-based correction.

The measure admittedly includes some circularity in derivation, in that a proportion of validation parcels were also training parcels. These were used because rarer and/or dissected cover types, with all polygons used for training, would otherwise be omitted from checking. It was recognised that this would bias the result. Thus the accuracy measure can only be taken as an indication. Nonetheless, it is a valuable measure, prior to full and independent validation through CS2000 field data; it is also easily applied as a check during production.

8.3 Objective validation / calibration

Objective validation will use the independent sample of land cover data provided by the CS2000 field survey of 1 km squares (as in Fuller *et al.* 1998). The first 12 squares have just been delivered and methodological developments are being made (Figure 8). Initial tests will involve the following:

1. Vector intersection of original field survey vector data with the same dataset resampled via a 25 m raster: the vectors will be converted to a 25 m raster, then transformed back to vector format but retaining the serrated outlines of the 25 m grid. This will measure what proportion of non-correspondence in the next test can be explained solely by virtue of the comparing a continuously variable vector format with vectors based on 25 m raster pixels.
2. Vector intersection of field survey vectors with the LCM2000 vector segments.
3. Raster intersection of field and LCM2000 data, both recorded on a 25 m grid format.
4. Comparison of summary % cover data from field and LCM2000.

The first three outputs will generate correspondence matrices, one matrix per square, which could in operation be combined in various combinations, for example, by:

1. ITE Land Class,
2. Environmental zone,
3. Image section,

4. Summer-winter, summer-only and winter-only outputs,
5. Country and region.

Once these results have been generated for the 10 sample squares, a Technical Advisory group meeting will be convened to discuss the longer term objectives, the proposed methods and the likely consequences for accuracy statements and for integrated uses of field and satellite data.

9. THE PRODUCTION SCHEDULE

Production has involved processing scenes in parallel from initial atmospheric corrections through geo-registration, illumination correction, co-registration of summer-winter data and segmentation stages. The map (Figure 9) shows the state of processing. The revised GANNT records progress in scene-pair equivalents (Figure 10). The number of scene pairs has been extended on the GANNT to reflect the 61 image sections scheduled for processing.

Equipment, class finalisation, methods refinements, field reconnaissance and the **image search** were reported in earlier reports and there is no need to update the details.

Image purchases have brought summer-winter coverage to 90%. **Image analysis** has lagged behind the planned schedule with the equivalent of 13 scene-pairs processed, compared with an expected 18 pairs. This has been due to the extra number of sections made from different combinations of summer / winter data based on primary and substitute images. As outlined earlier, this increased number of sections will continue to apply, such that at least 61 sections will result. The GANNT has been modified to reflect this number, the likely processing rate and the resultant timetable for completion.

Validation methods development / testing has continued with the recent delivery of sample field data, and plans are advancing for the analyses; **validation / calibration** is clearly dependent upon the results of trial comparisons and discussions in the Technical Advisory Group.

Change detection trials will be delayed while efforts concentrate on production and validation. The creation of **summary 1 km data** and **export to CIS** will be tackled once mapping is complete - the procedures are straightforward and have been used previously for LCMGB and CORINE GB. The **assessment of widespread Broad Habitats** envisages two types of assessment: the first, assessment of the effectiveness of mapping the selected Broad Habitats is complete in subjective terms and demonstrably effective, but objective assessment will require the results of validation. The second phase assessment will be an evaluation of the distribution and pattern throughout the UK, including regional and national statistics.

The production of **Quarterly statements** and **Reports** has been as planned. **Technical meetings** have taken place to schedule, though the November meeting was cancelled as no pressing technical issues needed discussion. A meeting on validation is currently pending, awaiting initial results from the 1 km comparisons. **Six-monthly Consortium meetings** have been rescheduled, with Consortium approval, to take account of repetition of activities through Advisory Group meetings and Technical Meetings. A Consortium meeting due in March will be considered at the Advisory Group meeting.

Publication has started with acceptance for a paper on CLEVER-Mapping to the *International Journal of Remote Sensing* (Smith & Fuller, in press). **Preparation of general display and launch materials** was envisaged for later in the proceedings; currently material is displayed on the CS2000 web-site, this is being regularly updated with sample outputs, and various display outputs are available in hard and soft copy formats; presentations on LCM2000 are becoming regular events.

10. CONCLUSIONS

- Summer-winter TM images cover 90% of England and Wales. It is hoped that winter 1999-2000 will complete summer-winter coverage for full UK mapping.
- Software refinements and other procedural developments have been formalised and are now successfully built into routine operations.
- The pre-processing and segmentation developments of earlier in the project are now demonstrably operational and their refinement is producing time savings at post-classification stages.
- Sample field reconnaissance data have been collected for *c.* 85% of Britain and the remainder will be covered in early summer.
- The field data are being used to train the classifier, with the field-mapping of land parcels quickly and objectively transferred onto the image segments.
- Training data are being reviewed using display image 'chips', a novel procedure which has allowed the operators to refine high quality training sets very quickly.
- In IGIS CLEVER-Mapping, the maximum likelihood classification, uniquely, is recording the probabilities for the top five subclass options.
- A combination of internal context, external spatial data and class probabilities is being used in knowledge-based correction.
- Validation has been designed as a two-stage process: first, using training-polygons to check the results per-polygon; second, using field survey data to fully and objectively calibrate results, offering a 'translation' between detailed field classes and the generalised target classes of LCM2000. The latter process is just starting.
- Overall production has involved processing scene-pairs in parallel, from initial atmospheric corrections through geo-registration, illumination correction, co-registration of summer-winter data and segmentation stages. Final classifications are available for 13 scenes with the 14th nearing the KBC stage; preliminary classifications are available for a further scene. Of the remainder, four scene pairs are in the pre-processing stream, progressing through the various stages of pre-processing, such that at any one time there is a segmented image ready for classification as the previous KBC is concluded. All other images which have been purchased have been entered on the system and checked for coverage and quality.
- Initial classifications of widespread Broad Habitats have shown that, in principle, the Broad Habitat classification can be achieved very successfully, with further sub-division of Broad Habitat classes to meet wider user objectives. Checks show an accuracy of 90 % assessed relative to training polygons, prior to 'universal' KBC procedures.
- Overall success will only become clearly apparent once validation, based on field survey squares, starts to provide an objective and independent measure of success, starting in spring.

- The map in Figures 3c, the graph of potential coverage (Figure 4) and the GANNT (Figure 10) show that 80% of Britain will be mapped by the launch, and that 99% will be completed by the end of March 2001 completion date; the remaining infill will be concluded by the end of May 2001.
- Every effort is being made to keep to or improve this timetable. However, LCM2000 is a GIS which will see routine use for probably a decade and more. It is considered that a delay of a few months in completion, while unfortunate for the purposes of the launch, will have relatively little impact on operational uses.
- There will be ample demonstrable output for the launch including the 'national' picture, set on the LCMGB backdrop, plus local and regional detail from 1998-99, both as maps and data.
- The timescale was discussed with the Advisory group on 2 March 2000. A proposed revision to the milestones (Annex C to the contract) is given in Annex I to this report.

11. REFERENCES

LCM2000 Reports

Fuller R.M., Gerard, F.F., Hill, R.A., Smith, G.M., Thomson, A.G., 1998a. *Countryside Survey 2000 - Part III. Module 7. Land Cover Map 2000. First Interim Report incorporating the Second Quarterly Progress Report. CSLCM/Interim1.* Unpublished ITE report to the LCM2000 Consortium.

Fuller R.M., Gerard, F.F., Hill, R.A., Smith, G.M., Thomson, A.G., 1998b. *Countryside Survey 2000 - Part III. Module 7. Land Cover Map 2000. First Progress Report.* Unpublished ITE report to the LCM2000 Consortium.

Fuller R.M., Smith, G.M. & Hill, R.A. 1998c. Land Cover Map 2000. In: *Countryside Survey 2000 Report CSJMT 7/4. Part III. Module 7.* Unpublished report to the Joint Management team of Countryside Survey 2000. 1p.

Fuller R.M., Smith, G.M. Hill, R.A. & Sanderson, J.M. 1999a. Land Cover Map 2000. In: Hornung, M. (ed.) *Countryside Survey 2000: Second Integrated Progress Report.* Unpublished ITE report to the DETR.

Fuller R.M., Smith, G.M., Sanderson, J.M., Gerard, F.F., Hill, R.A., Thomson, A.G., 1999b. *Countryside Survey 2000 - Part III. Module 7. Land Cover Map 2000. Second Interim Report incorporating the Fourth Quarterly Progress Report. CSLCM/Interim2.* Unpublished ITE report to the LCM2000 Consortium.

Fuller R.M., Smith, G.M., Sanderson, J.M., Hill, R.A., Thomson, A.G. & Clarke, R.T. 1999c. *Countryside Survey 2000 - Part III. Module 7. Land Cover Map 2000. Third Interim Report incorporating the Sixth Quarterly Progress Report, CSLCM/Int3/Prog6.* Unpublished ITE report to the LCM2000 Consortium.

Fuller R.M., Smith, G.M., Sanderson, J.M., Hill, R.A., Thomson, A.G. & Hall, M.W. 1999d. *Countryside Survey 2000. Module 7. Land Cover Map 2000. Seventh Quarterly Progress Report, CSLCM/Prog7.*

Fuller, R.M., 1999e. Module 7. Land Cover Map 2000. 3pp. In: Hornung, M. (ed.) *Countryside Survey 2000 Third Integrated Progress Report.* Unpublished ITE report to the DETR.

Fuller, R.M., Smith, G.M. & Hill, R.A., 1998d. *Countryside Survey 2000 Report CSJMT 8/6. Land Cover Map 2000.* Unpublished report to the Joint Management team of Countryside Survey 2000. 2p.

Fuller, R.M., Smith, G.M. & Hill, R.A., 1999f. *Countryside Survey 2000. Land Cover Map 2000. The Land Cover Map Of Northern Ireland. An Extension To Land Cover Map 2000. Preliminary Report.* Unpublished ITE report to the Northern Ireland Members of LCM2000 Consortium (and provided to the full LCM2000 UK Consortium).

Fuller, R.M., Smith, G.M., Hill, R.A., Thomson, A.G. & Gerard F.F. 1998e. Module 7. Land Cover Map 2000. In: Hornung, M. (ed.) *Countryside Survey 2000 First Integrated Progress Report.* Unpublished ITE report to the DETR.

Smith, G.M., Fuller R.M., Sanderson, J.M., Hill, R.A. & Thomson, A.G., 1999g. Land Cover Map 2000: Fifth Quarterly Progress Report. CSLCM/Prog5.

Other references

Fuller, R.M., Groom, G.B. & Jones, A.R. 1994. The Land Cover Map of Great Britain: an automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering & Remote Sensing.* **60**, 553-562.

Fuller, R.M., Wyatt, B.K. & Barr, C.J. 1998. Countryside Survey from ground and space: different perspectives, complementary results. *Journal of Environmental Management,* **54**, 101-126.

Groom, GB. & Fuller, R.M. 1996. Contextual correction: techniques for improving land cover mapping from remotely sensed images. *International Journal of Remote Sensing*. **17**, 69-89.

JNCC. 1999. *Definitions of Broad Habitats (Draft)*. Unpublished document dated 06/10/99.

Kershaw, C.D. & Fuller, R.M. 1992. Statistical problems in the discrimination of land cover from satellite images: a case study in lowland Britain. *International Journal of Remote Sensing*, 13, 3085-3104.

Smith, G.M., & Fuller, R.M., 1998, *CLEVER-Mapping of land cover in Jersey. Final Report.* Institute of Terrestrial Ecology Report to the States of Jersey.

Smith, G.M., & Fuller, R.M., In press (accepted). Multi-sensor, high resolution, knowledge-based per-parcel classification of land cover: an example in the Island of Jersey. Accepted by the *International Journal of Remote Sensing.*

Table 1. Widespread Broad Habitats listed against LCM2000 target classes, subclasses and variants

widespread Broad Habitats	LCM Target	Subclass /	Variant	Code
25. Oceanic seas	Sea / Estuary			We Ws
13. Standing water/canals *14. Rivers and streams*	Water (inland)			W
18. Littoral rock 19. Littoral sediment	Littoral rock and sediment	rock mud sand saltmarsh	 ungrazed grazed saltmarsh	Lr Lm Ls Lsm Lsg
16. Supra-littoral rock 17. Supra-littoral sediment	Supra-littoral rock and sediment	rock/ shingle Dune with shrub	 	Sr Sh Sd Sds
12. Bog	Bog	Bog (shrub) Bog (grass/herb)	 grass (*Molinia*) herbaceous bog	Bo Bm Bh
10. Dwarf shrub heath	Dwarf shrub heath	Closed Open	'wet' heath 'dry' heath gorse 	Hw H Hd Hg Hb Hbg Hga
15. Montane habitats	Montane habitats (bare/heath)			Zh Z
1. Broad-leaved woodland	Broad-leaf wood (deciduous/	Scrub/shrub Trees 	Deciduous Evergreen	Dn Ds D Dm De
2. Coniferous woodland	Coniferous woodland	Standing Felled	 	C Cl Cn Cf
4. Arable & horticulture	Arable and horticulture	Crops* Grass leys Bare setaside/fallow Orchard/permanent crops	 	A* Agl Ase Ado
5. Improved grassland	Improved grassland	Agricultural/ managed) grass Grass, semi- improved/reverting	Mown / Grazed Hay/silage Grazing marsh Setaside (grass/weeds)	Gi Gih Gsm Gss
6. Neutral 7. Calcareous 8. Acid 9. Bracken	Natural & semi-natural grasslands & bracken	Neutral Calcareous Acid Bracken	Unimproved neutral Rough acid *Nardus* Grass with *Juncus* 	Gu Gn Gr Gc Ga Gm Gj Gbr
11. Fen, marsh and swamp	Fen, marsh, swamp	Swamp Fen/marsh	 	Fw Fm
27. Built up areas, gardens	Built up areas, gardens	Suburban/rural developed Continuous Urban	 residential/commercial industrial	Us U Ui
26. Inland rock	Inland Bare Ground		despoiled natural	Id Ib
	Subdivisions of A where possible	spring* winter*	e.g. Aw, Ab, Ar Ao, Am, Ap, Aq, As, Ac, Ah, Aba	

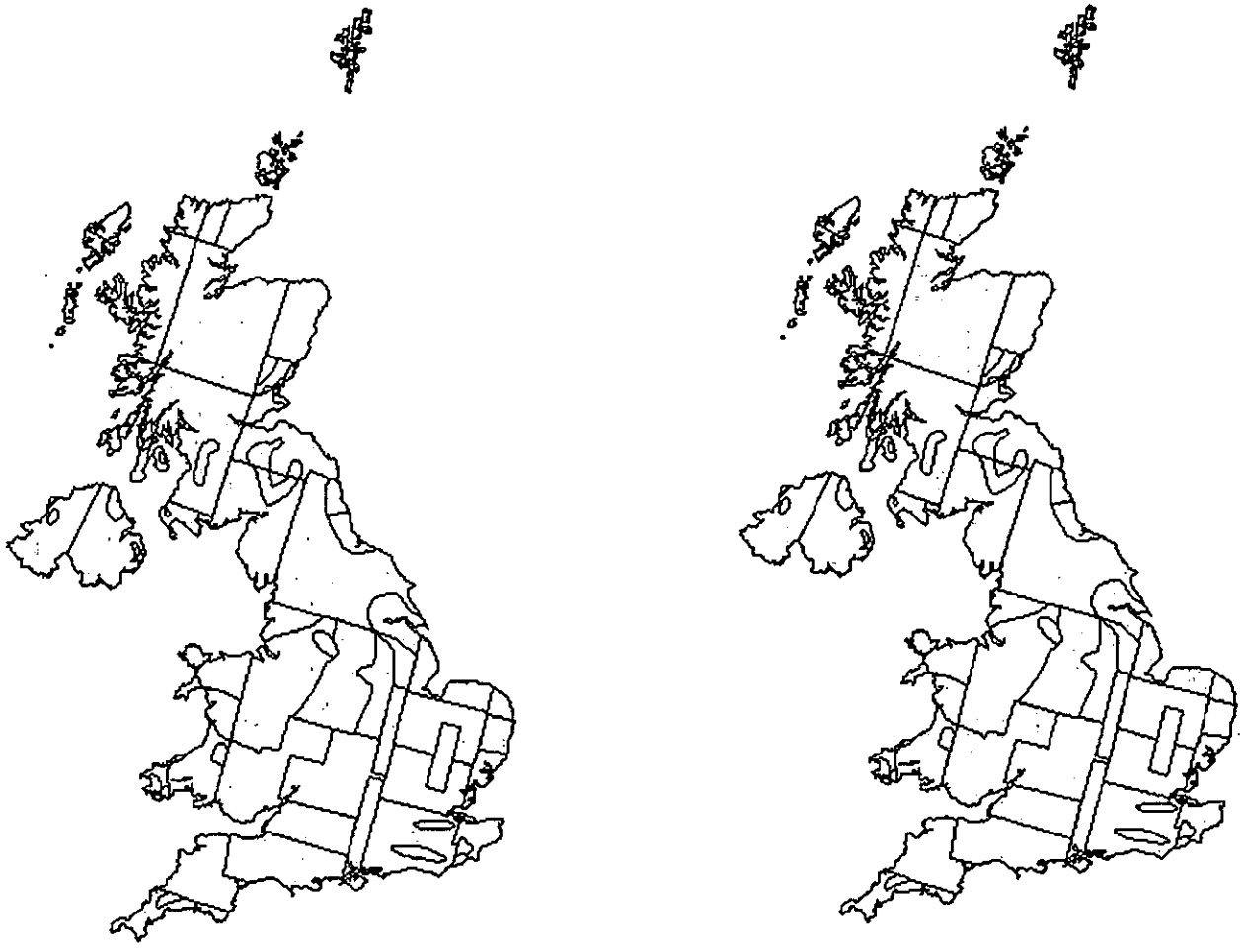


Figure 1: a. areas of the UK covered by both summer and winter images; b. areas covered by summer or winter images

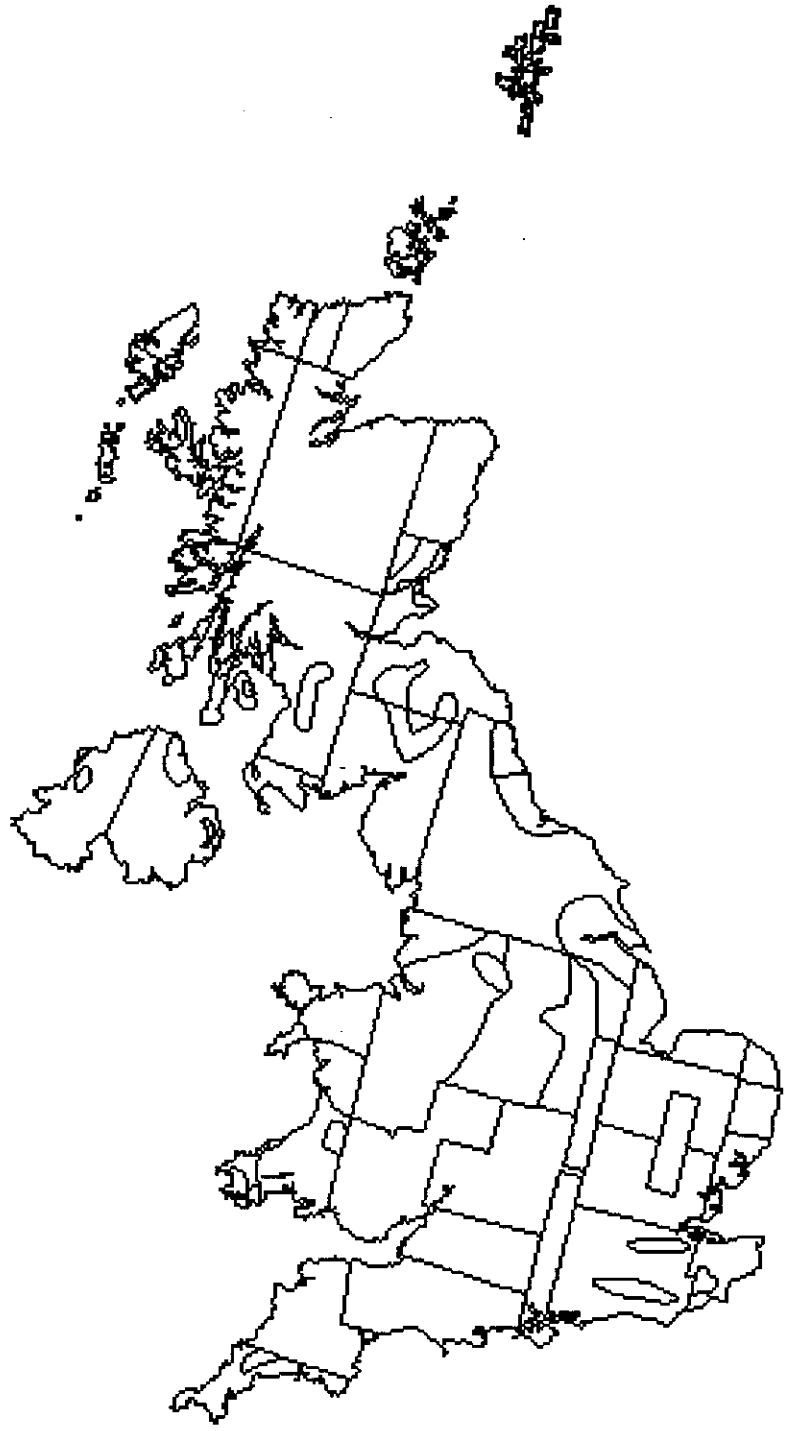


Figure 2. The likely mosaic of image-sections which will make up Land Cover Map 2000.

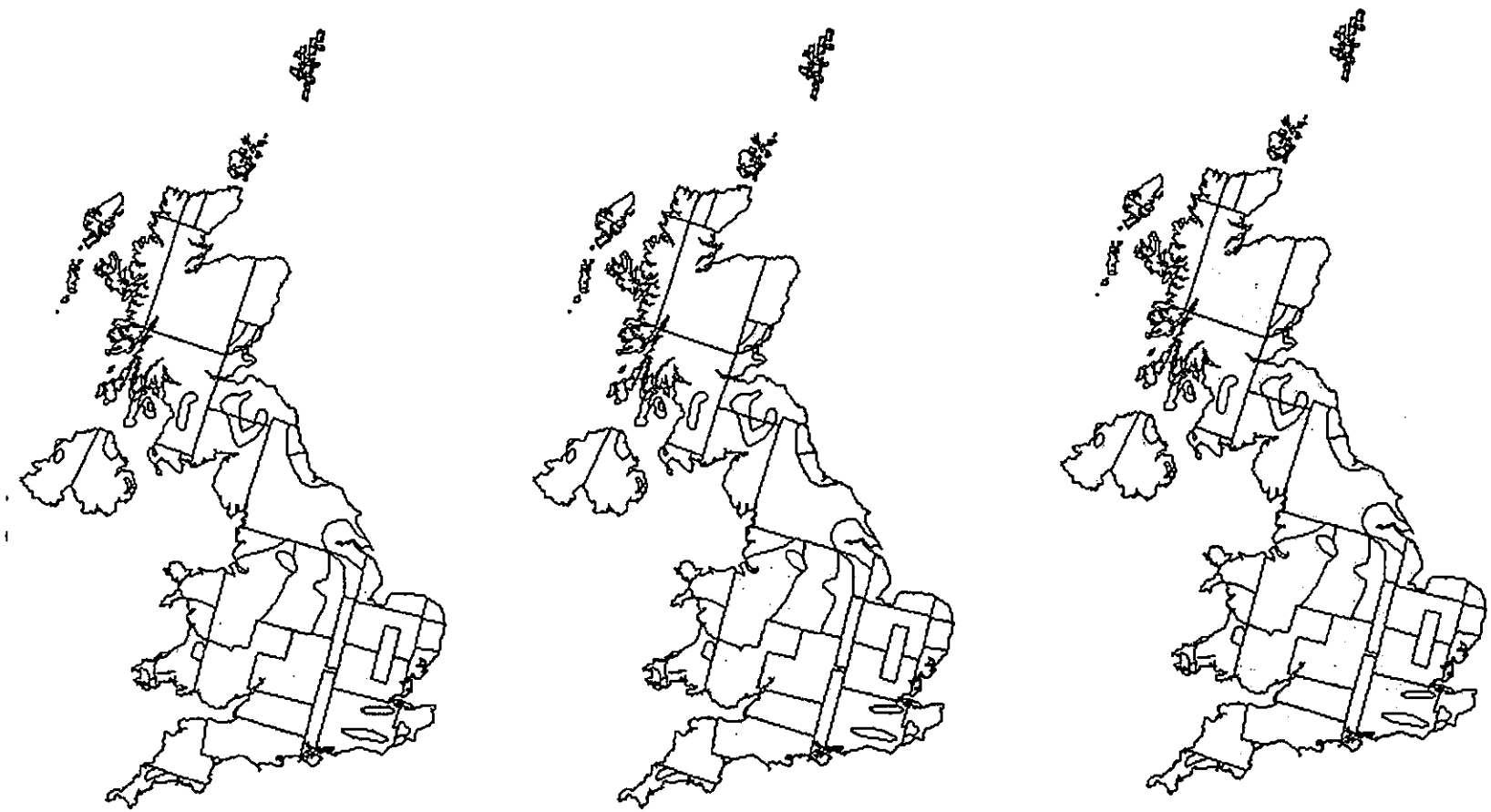


Figure 3: a. expected coverage by the time of the CS2000 launch if processing progresses systematically, infilling cloud holes before moving to new scenes; b. the sections of imagery which would be needed to complete coverage of England and Wales in time for the launch; c. the sections to be processed if the aim is to maximise coverage by the time of the launch.

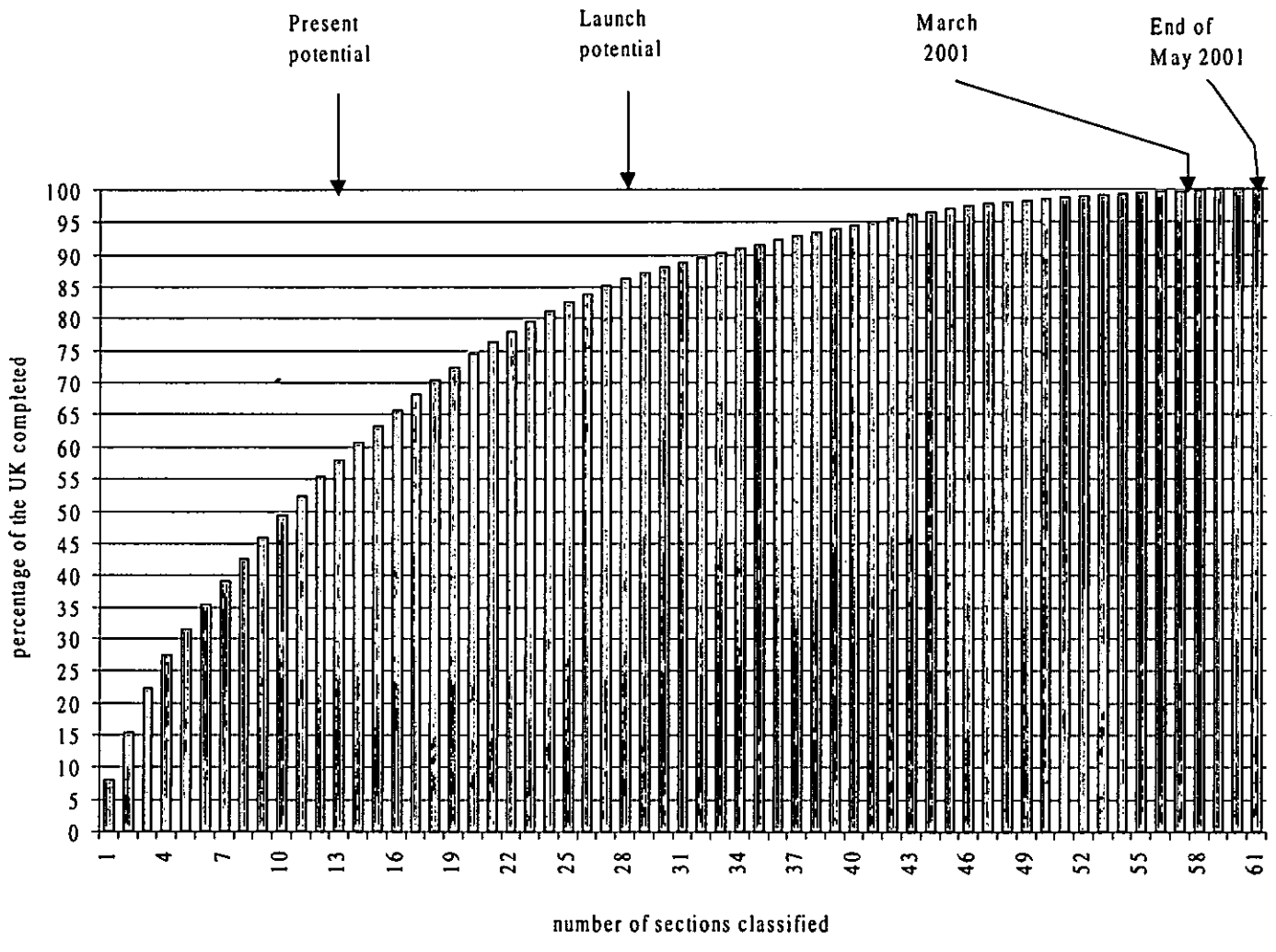


Figure 4. Cumulative mapped area, if image sections were processed in order of size from largest to smallest, from the beginning.

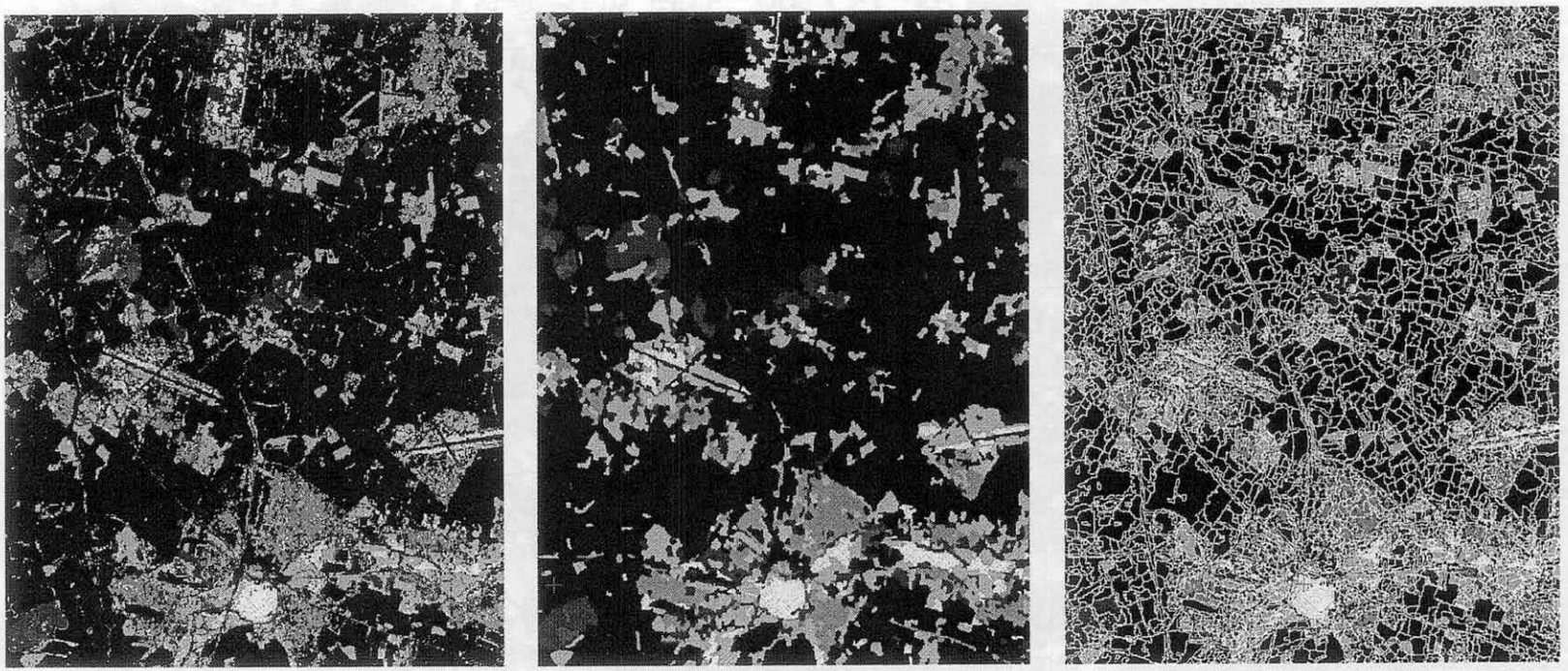


Figure 5. LCM2000 outputs of the Monks Wood area showing Huntingdon and the River Great Ouse valley to the south, with Monks Wood to north-west and Woodwalton Fen north-central. The Figure shows i. a per-pixel classification, ii. a per-segment classification and iii. segments overlaid onto the per-pixel result. This figure is intended to show the complementary nature of results with segments offering greater accuracy, pixels recording within-field heterogeneity and marking the larger linear features, while boundaries serve to provide the spatial-structural relationships which define field sizes and place pixel details in their true context.

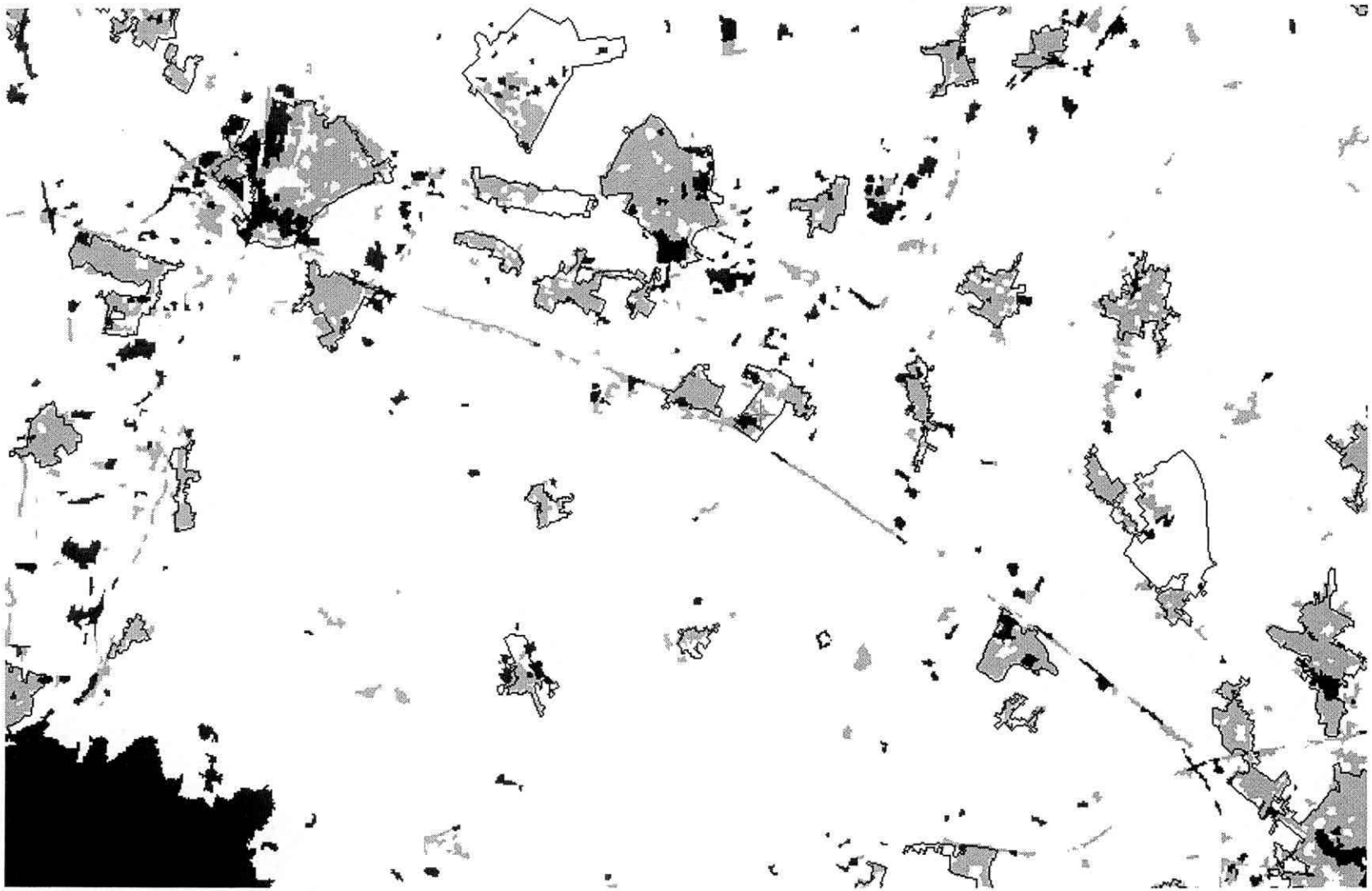


Figure 7. The A14 corridor, Huntingdon to Cambridge from LCM2000, overlaid with DETR urban polygon boundaries; black - urban; dark grey - industrial; light grey - suburban.

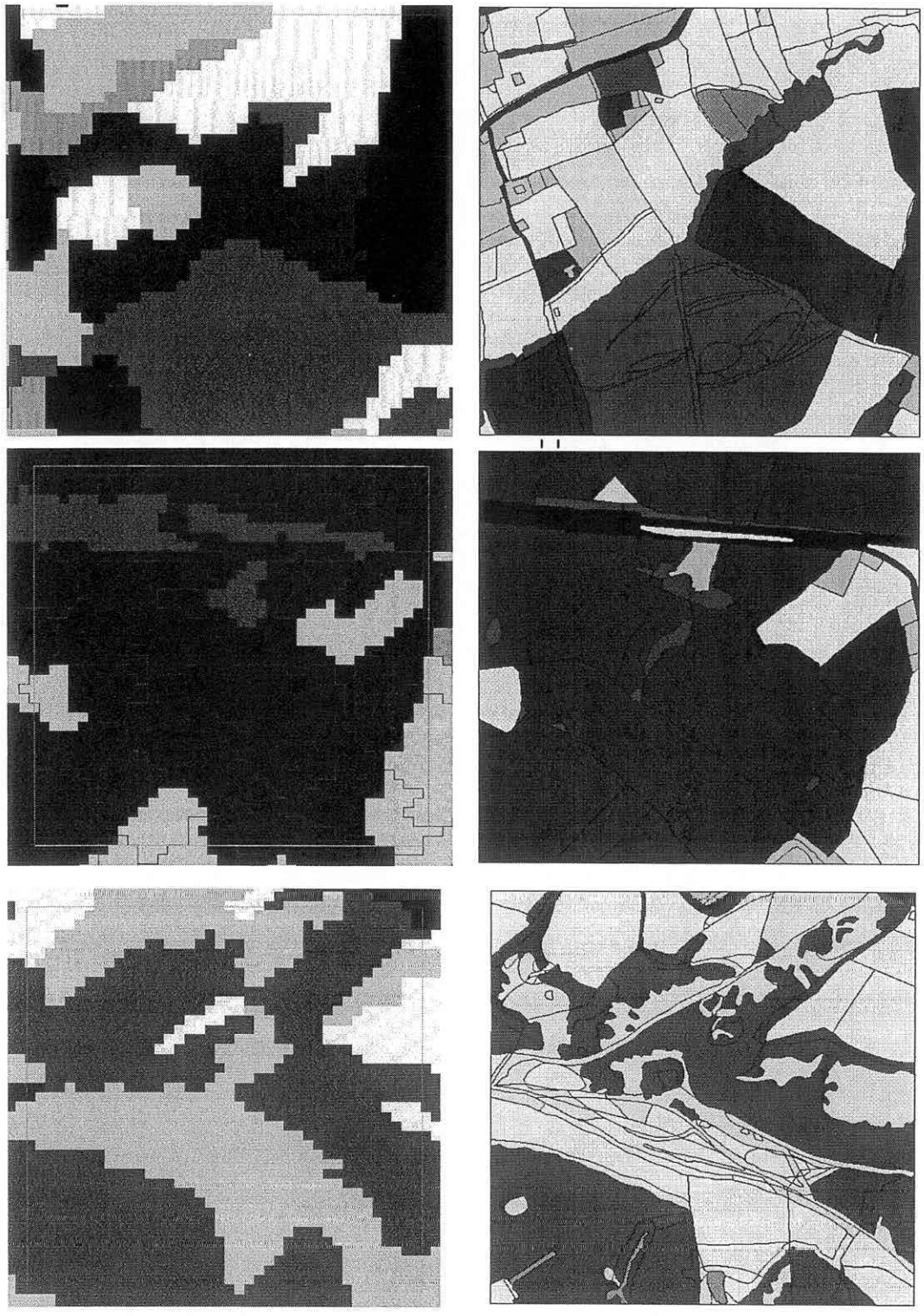


Figure 8. Three examples of CS2000 1 km squares surveyed by CLEVER-Mapping from satellite data (left) and in the field (right). These examples are intended to show some of the spatial-structural issues, rather than specific classes: hence the use of black and white reproduction. The outlines on the satellite maps denote approximate area of 1 km squares, but the maps have not yet been fully co-registered and trimmed to exact spatial extent.

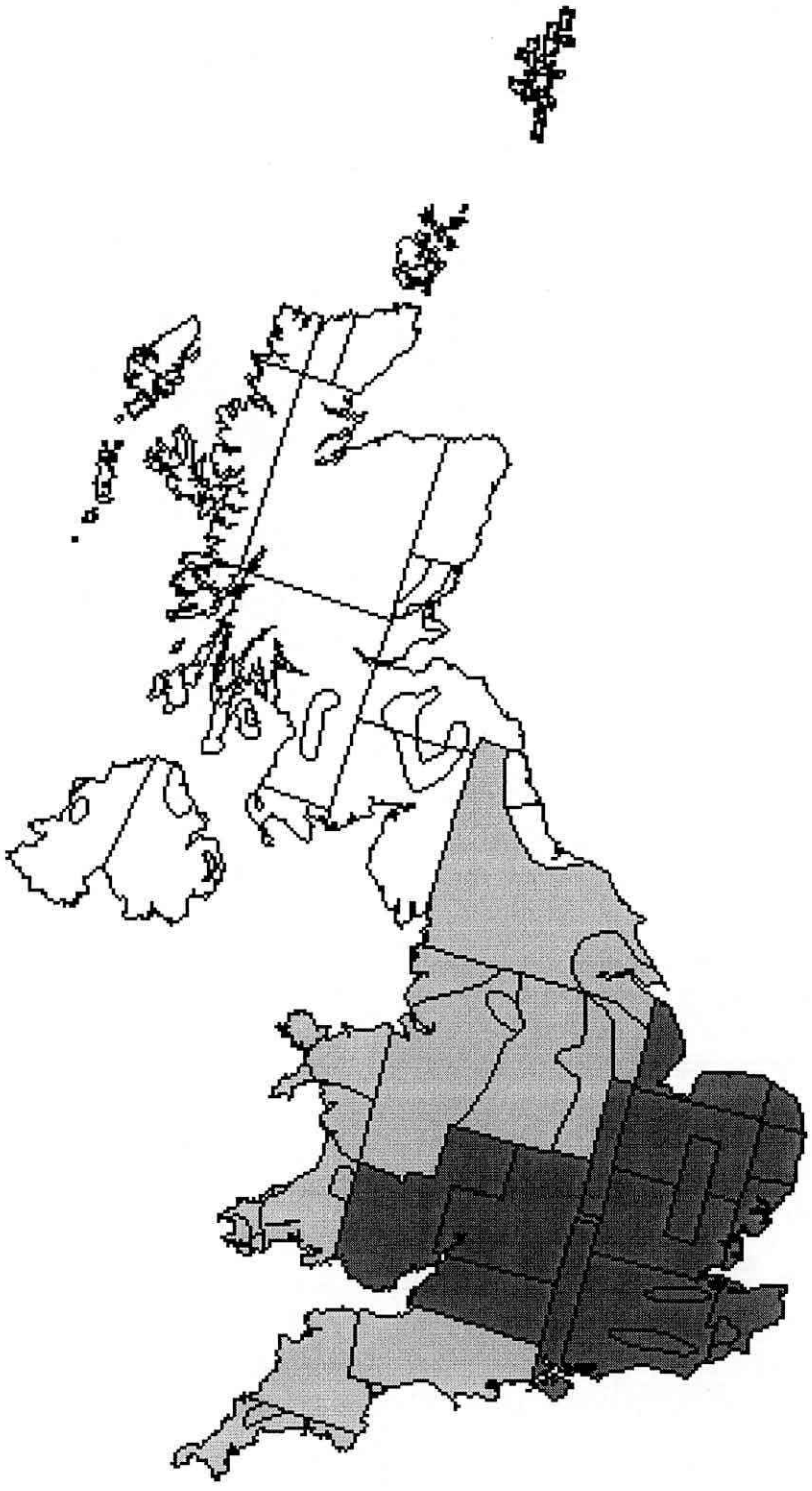
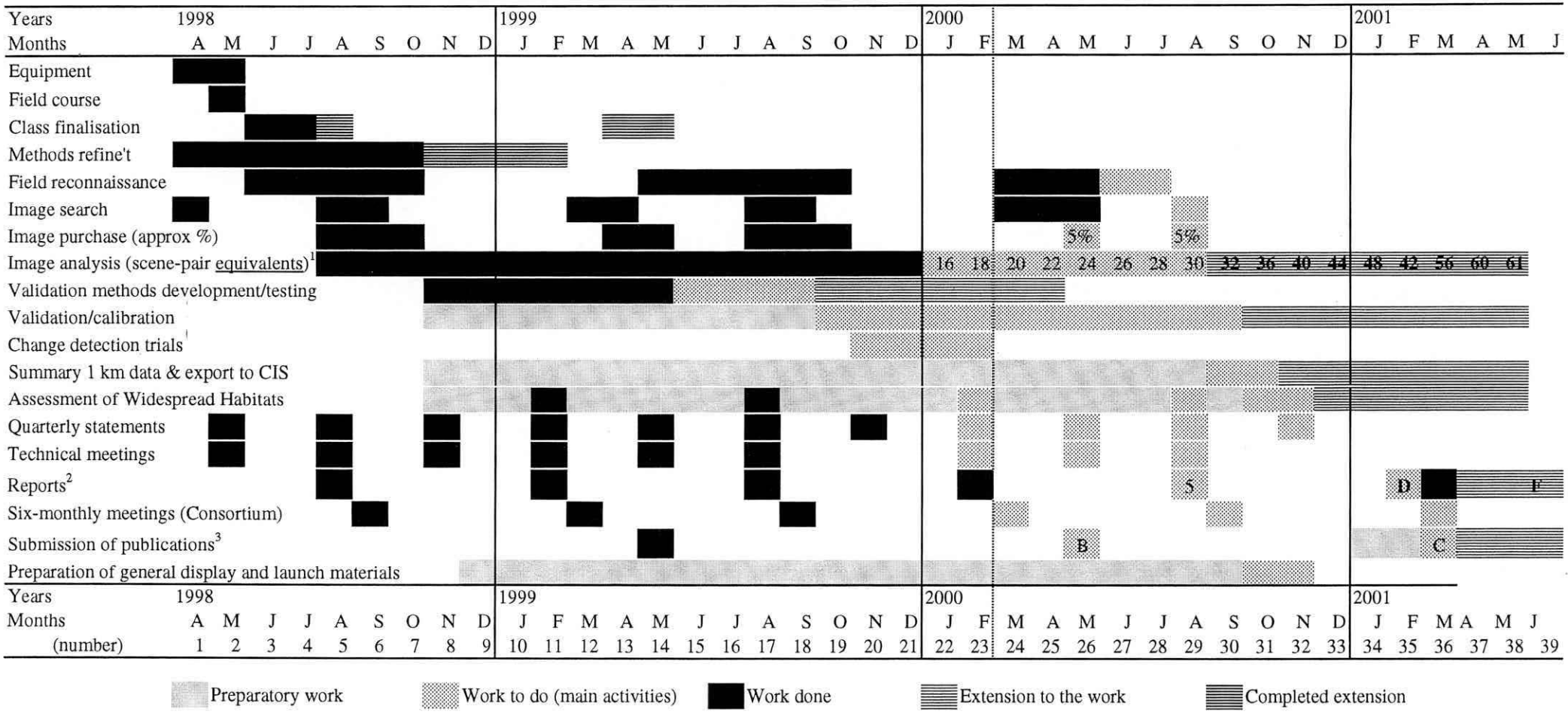


Figure 9. Areas currently classified (dark shading) and in the pre-processing stream (light shading).

Figure 10. GANNT - The revised timetable for Land Cover Map 2000 (UK) and estimated progress in production (to the end of February 2000).



ANNEX I

CS2000 MODULE 7: LAND COVER MAP 2000

MILESTONES AND STAGED PAYMENTS

(REVISED MARCH 1999)

OUTPUT	DESCRIPTION	DATE
First Interim Report	*Completed*	*Sept 1998*
Second Interim Report	*Completed.*	*March 1999*
Third Interim Report	*Routine production completed for 15% of UK, 10% under classification, with a further 25% in pre-processing phases; image acquisitions year 2 assessed; 85% of ground reconnaissance completed, validation tests outlined and reported, reporting categories updated, six-monthly consortium meeting held.*	*Sept 1999*
Fourth Interim Report	*Images purchased and evaluated for 90% of UK; searches instigated for winter 1999-2000 coverage. Routine production completed for 25% of UK; further 35% in processing stream. Field / satellite comparisons started; validation operations identified, tests commenced; technical advisory meeting on validation arranged; test GIS coverages available; test CIS datasets in production.*	*March 2000*
Pre-launch preparations completed and displays/ presentations finalised; Fifth Interim Report	Ground reconnaissance completed. Infill-image acquisitions completed. Production completed for 70% of UK (on target for 80% by the November launch). Validation methods implemented for mapped areas and summary data available for launch. Change detection evaluated with recommendations. Material provided for launch report: illustrative material prepared and presented (illustrative examples to cover all widespread Broad Habitats, 4 UK countries and 6 Environmental Zones). Six-monthly progress meeting held (if deemed appropriate by Consortium). Outlines on above in an additional, fifth, Interim Report.	Sept 2000
Draft Final Report	Launch-objectives completed in previous November. Production completed for 95% of UK. Draft report outlining contents: full details on methods; preliminary results of analyses; outline of illustrative material. Six-monthly progress meeting held (if deemed appropriate by Consortium).	March 2001
Final Report	Final report completed and data delivery completed. Delivery of GIS data, derived products generated (1 km square / CIS summary data). Assessment of Broad Habitats; national, regional and zonal statistics. Report approved for publication. Draft paper submitted.	June 2001

Natural Environment Research Council