CENTRE FOR ECOLOGY AND HYDROLOGY

(NATURAL ENVIRONMENT RESEARCH COUNCIL)

Report to Directorate General XVI (European Commission)

DGXVI Contract

GENERALISING THE LAND COVER MAP OF GREAT BRITAIN TO CORINE LAND COVER BY SEMI-AUTOMATED MEANS

Final Report

N. J. Brown, F. F. Gerard, R. M. Fuller and D. S. Edmundson

Centre for Ecology & Hydrology

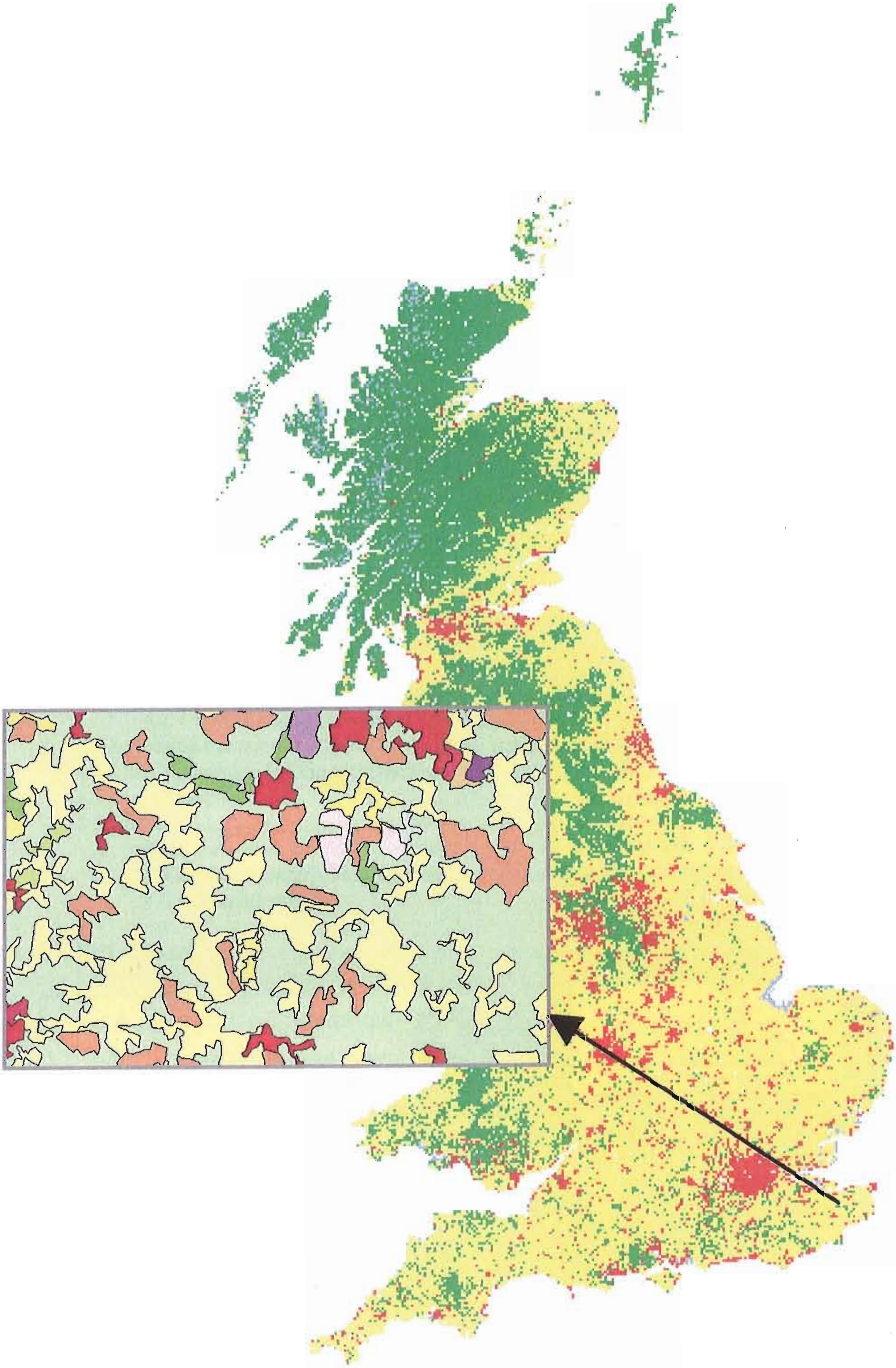
European Commission Directorate General XVI

DETR Department of the Environment, Transport and the Regions

European Environment Agency | Putting Information To Work

Centre for Ecology and Hydrology
Monks Wood
Abbots Ripton
Huntingdon
Cambridgeshire PE17 2LS

December 1999



CONTENTS

List of Figures **Page**

List of tables

1. EXECUTIVE SUMMARY

The European Commission (EC) programme 'Co-ordination of Information on the Environment' (CORINE) includes a major project to map the land cover and land use of member states.

The CORINE Land Cover Map, generally produced in the late 1980s - early 1990s by visual interpretation and manual digitising, shows 44 cover types, in vector format (i.e. as digital map outlines) at 1:100 000 scale, with a minimum mappable unit of 25 ha.

The work was jointly funded by the Directorate-General XVI for regional policy and cohesion (DGXVI) of the European Commission (EC) and the Department of Environment, Transport and the Regions UK (DETR). The work was carried out by the Institute of Terrestrial Ecology, UK (ITE).

The generalisation procedure and the final CORINE outputs were evaluated by the Technical Advisory Unit of the European Topic Centre/Land Cover.

The Land Cover Map of Great Britain (LCMGB) of 1988-90 gives a raster (i.e. grid-based) map which records cover on 25 m cell size, identifying 25 cover-types, with a minimum mappable unit of 0.125 ha, showing landscape patterns at field-by field scale.

The CORINE Land Cover Map of Great Britain has been created by semi-automated generalisation methods, developed in 1994 by ITE, and adapted to run in the ARC/View GIS. It has successfully used as input the LCMGB, together with exogenous data from many other sources. The procedure has retained a level of manual interpretation for land uses and input from conventional cartographers. The CORINE map and data are produced to 3 levels of detail.

Conversion has been completed for all of Great Britain. The data have also been imported into the ITE Countryside Information System for analysis in combination with other environmental parameters, at the 1km level.

The procedure has been validated against standard CORINE manual methods with considerable success. Outputs for Arable, Pastural, Upland and Marginal landscapes have been evaluated. An overall 86% correspondence for GB is in accord with the intended 85% overall 'accuracy' required of CORINE Land Cover mapping.

In all map generalisation processes, some elements of the source information are changed, or perhaps lost altogether. This particularly affects small map objects. Some rarer classes have suffered losses during the generalisation. Major classes have remained largely unchanged.

The very similar appearance of the manual & automated outputs, and the high level of agreement in cover statistics for test sites, demonstrate that the procedure has achieved the desired output, and that the CORINE Land Cover Map of Great Britain conforms with the CORINE requirements.

2. INTRODUCTION

The European Environment Agency (EEA) was launched by the European Union (EU) in 1993 with a mandate to co-ordinate and put to strategic use, information of relevance to the protection and improvement of Europe's environment. The Agency carries out its tasks in co-operation with a European Information and Observation Network (EIONET). EIONET consists of national networks, organised by the Agency to help it retrieve information, and produce efficient and timely information on Europe's environment. To execute particular tasks, institutions or organisations have been contracted as European Topic Centres (ETC). There are today ETCs for Air Emissions, Air Quality, Catalogue of Data Sources, Inland Waters, Land Cover, Marine & Coastal Environment, Nature Conservation and Soil. The ETC on Land Cover (ETC/LC) was established to provide accurate data on land cover in Europe, corresponding to needs across a wide range of applications.

A key activity of ETC/LC has been the completion of a European-wide inventory of land cover in 44 classes. This takes the form of a digital cartographic product, at a scale of 1:100 000. The inventory has been compiled, mostly in the late 1980s - early 1990s, using methods developed within an experimental programme called CORINE (Co-ordinating Information on the European Environment). The development was undertaken by the Environment Directorate (DGXI) of the European Commission between 1980 and 1985. A major task of ETC/LC has been to develop and complete the Land Cover database begun within the CORINE programme. Today the CORINE land cover data base is operationally available for the greater part of the 3.5 million km² covered by the European Union and progress is being made, through the PHARE programme, in the production of maps to CORINE standards for the former Soviet Union states.

The Land Cover Map of Great Britain (LCMGB) 1990, was made by the Institute of Terrestrial Ecology (ITE) using remotely sensed data. The British land cover map differs from CORINE in several respects, including its spatial resolution, the land cover classes mapped and the method of production:

- The CORINE Land Cover Map has generally been produced by visual interpretation of hard copy satellite images followed by manual digitising to give computer maps which show 44 cover types, as digital map outlines, in vector format at 1:100 000 scale, with minimum mappable units of 25 ha.

- The LCMGB is a raster or grid-based product which records 25 cover-types, on 25 m grid, with minimum mappable units of 0.125 ha, showing landscape patterns at the field-by-field scale.

A pilot study (Fuller & Brown, 1996) successfully demonstrated semi-automated procedures, which can be used to convert the LCMGB to CORINE specifications. These procedures involved generalisation from the 25m resolution, reassignment of LCMGB classes to the CORINE categories, generation of CORINE mosaic classes from heterogeneous regions, and the use of knowledge-based operations to add relevant land use information. More details of the original pilot study are described in section 3 of this report.

Conversion ensured that CORINE land cover data for Britain are calibrated against the existing national map and against the proven ground reference data available from Countryside Survey 1990 (Barr *et al.* 1993). The approach also ensures that information about CORINE land cover in Britain is entirely consistent with the national dataset and with the large number of uses to which these data have already been put.

The work was jointly funded by the Directorate-General XVI for regional policy and cohesion (DGXVI) of the European Commission (EC) and the Department of Environment, Transport and the Regions UK (DETR). The work was carried out by the Institute of Terrestrial Ecology, UK (ITE).

The generalisation procedure and the final CORINE outputs were evaluated by the Technical Advisory Unit of the European Topic Centre/Land Cover, represented by Cristina Seabra of CNIG, Portugal.

The project had an Advisory Group that has been a two way process; to assure deliverables were on time and of the quality required, and to advise or seek advice if and when problems arose. The project's Advisory Group members were:

Dr Andrew Stott	Department of the Environment, Transport and Regions
Dr Michael Albas	Commission of the European Communities, DGXVI
Dr Chris Steenmans	European Environment Agency
Dr Vanda Perdigao	Commission of the European Communities, Joint Research Centre
Dr Stuart Gardiner	Scottish Office
Cristina Seabra	Centro Nacional Informacao Geografica (CNIG), Portugal.

3. METHODS

The generalisation methodology used to produce CORINE Land Cover (CLC) was initially developed as part of a feasibility study in 1994. It is a semi-automated procedure and comprised the following main steps:

- Removal of very small parcels <2ha;
- Extraction of 25 ha parcels with direct CORINE equivalence;
- Clustering of smaller parcels;
- Analysis and classification of mosaic parcels;
- Assignment of remaining small parcels to the most appropriate neighbouring class;
- Overlay onto the satellite images to check outputs;
- Use of 'exogenous' data and expert interpretation to identify CORINE land uses;
- Smoothing of polygon boundaries.

The operation of these methods was documented in the ITE report to the Department of the Environment (Fuller & Brown, 1994) and further described in published papers by Fuller & Brown (1996) and Brown *et al.* (1996) with methodological refinements outlined by Gerard *et al.* (1996).

The procedure, initially designed to use ARC/Info and ARC/Grid functions, was built into ARC/View processes as a sequence of scripts which produce consistent outputs, with the minimum of interactions. However, results required inspection and fine-tuning of methods, making some stages iterative. Other stages, especially the interpretation of CORINE's land use classes (e.g. airports, recreation areas) involved some interactive editing. A significant change to the original procedure was to do the interpretation of the land use classes at an earlier stage so that they could be preserved during the remainder of the processing.

The above stages were applied to subsections based on 100 km square tiles of the LCMGB (with overlaps between subsections to ensure satisfactory subsequent edge matching). The following stages apply to output tiles:

- Aggregation: 100 km tiles of the CORINE Land Cover of Britain were assembled into continuous, edge-matched vector data for all of Britain

- Validation: samples of output maps were validated by comparison with equivalent samples mapped by conventional CORINE methods (see Results section), and by comparisons with ground reference data including those from Countryside Survey 1990.

- Measurements: the differences between input and output maps, i.e. the effects of generalisation, were measured; also summary statistics were provided, giving tabulations of CORINE statistics for Britain, England, Scotland and Wales (see Results section).

- Output of 1 km data: Map data were summarised as cover per class per 1 km square and built into a PC-based Countryside Information System (CIS).

- Copies of final digital data, in both vector and raster form, were provided to DETR, EC and members of the CLCGB Advisory Group.

- Hard copy, such as maps of the whole of GB and posters describing the product, are available from ITE and have been provided to a number of organisations and individuals such as DETR and the land cover Topic Centre.

- All data have been archived at ITE, both within computer systems and on CD-ROMs within fire safes.

- Web pages have been constructed, first for use during the generalisation process as a way of showing progress; and secondly, for future information and dissemination of the product. http://mwnta.nmw.ac.uk/ite/Monk/CORINE/homecorine.html

3.1 Methodological improvements and additions

The following paragraphs describe how the original generalisation procedures, which were designed for use in ARC/Info, have been adapted to run in ARC/View. This adaptation to ARC/View incorporated improvements in methodology as well as efficiency. A detailed listing of the ARC/View functions used for each generalisation step is included in Annex 7.7. A customisation of the standard ARC/View menu interface was created (see Figure 3 Annex 7.6), and this was used throughout the generalisation procedures. This interface included new buttons which activated 'avenue' scripts written specifically for the project.

ARC/View handles Grid files in the same general manner as ARC/Grid. The original process involved generalising a given LCMGB tile using a combination of ARC/Info and ARC/Grid procedures on a Unix workstation. All processes were command-driven by typing in set functions at each stage of the conversion. The generalisation also involved five intermediate conversions from raster-to-vector and *vice versa* which were tedious, time consuming and had a weakness in the limited number of vectors that could be handled at any one time by ARC/Info. As the test area was relatively small in the first instance, it was not limited by the number of polygons, so the generalisation process ran satisfactorily. However, an increase in the test area to 50 km square of land, containing a greater variety of LCMGB land classes, resulted in a total number of polygons exceeding the then 100 000 polygon limit of ARC/Info. This in turn led to a need for segmenting the tile into a number of smaller sections, which introduced subsequent problems of edge matching.

ARC/View has much the same functionality as ARC/Grid but its object-oriented macro programming language 'Avenue' allows the user more control over functions (such as grid zonal functions), and the linking together of various functions into sub-routines. The sub-routines can also be accessed via customised menu systems. Examples of these are shown in the Annex.

The initial changes made to the generalisation system were aimed at keeping the data in Grid format in order to avoid the 100 000 polygon limit. This speeded up the generalisation process by reducing the need for conversion between raster and vector formats and between systems. It did not compromise the quality of outputs, or their correspondence with CLC; indeed, the refinements helped to improve these.

The 25 LCMGB classes were simplified to use only the broad types needed to define CLC classes and mosaics.

A document further defining mapping criteria for generalisation to CORINE (Steenmans, 1997) has been produced by the EEA during the GB generalisation process. The generalisation procedures have been further adapted to ensure the output results comply with these updated criteria.

3.2 Use of 'exogenous' data and expert interpretation to identify CORINE land use classes

Certain CORINE classes could not be defined directly by computer algorithms. Features such as Port Areas, Airports, Sport and Leisure facilities, Dump Sites and Mineral Extraction Sites required extra information. Several sources of information have been used in producing an appropriate representation of these classes.

One source was a Bartholomew's point data set. This contained locational information about a range of features required for the CORINE map, including golf courses and airports. This dataset was designed to be used at 1:250 000, and when overlayed on to the LCMGB, the point features occasionally showed positional inaccuracies or displacements of up to 200m. However, on the majority of occasions, the feature in question could be easily defined in the satellite imagery, based on its typical shape or pattern.

Other information sources included locations of mines and mineral workings, from the British Geological Survey (BGS) and land cover information for Scotland derived from aerial survey. Throughout this process continual reference was made to the Ordnance Survey 1:50 000 maps of the late 1980s and early 1990s.

The GIS procedure to deal with these features was, essentially, an interactive one using all the data and documentation available. The preliminary generalised data were displayed on the screen, in combination with the satellite imagery together with any combination of the appropriate digital and paper information described above. The interpreter digitised the outline of the land use feature, such as an airport boundary, using the CORINE mapping criteria as a guide. Generally this boundary interpretation took place at a visual scale of about 1:20 000 to ensure optimum accuracy.

This step was the first of two separate manual checking stages within the generalisation process. It was first carried out at at this early stage, so that the land use parcels could be protected from shape erosion erosion, and also contribute to the later growing and mosaicing procedures.

The generalisation process then continued with the overlay of these manually interpreted vectors, prior to removing polygons less than 2 hectares. These digitised polygons were then given priority throughout the generalisation process. An example of each of the CORINE land use classes treated in this way are shown in Figure 2 below. This process took an average of about 2.5 days per 100km tile and produced over 5000 land use parcels. The CORINE classes that were constructed in this way were less than 2% of the map area. Figure 1 shows their distribution in a relatively complex square in south east England.

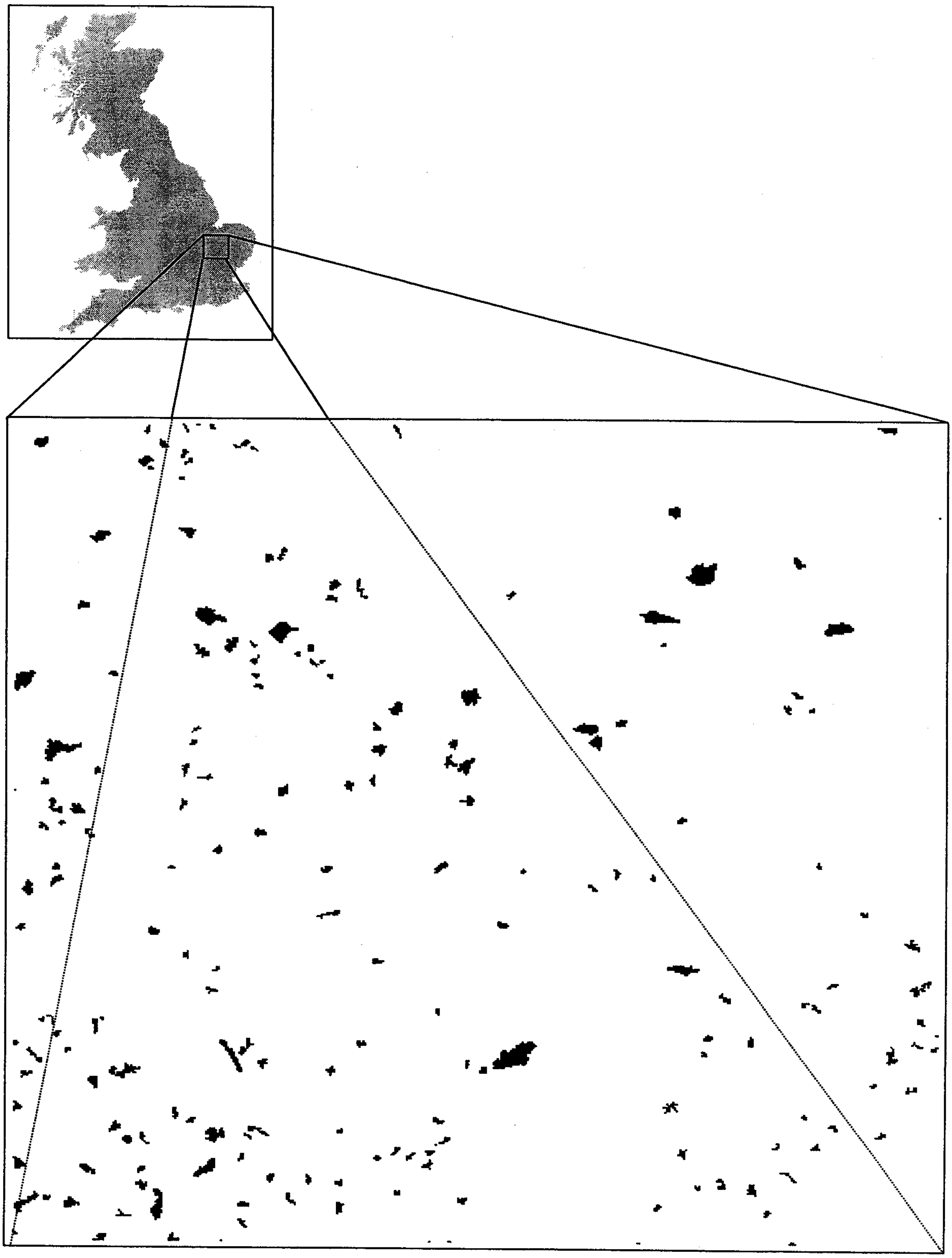


Figure 1. Distribution of 'land use' parcels within a single100km square.

CORINE class 1.2.1
Industrial - East London

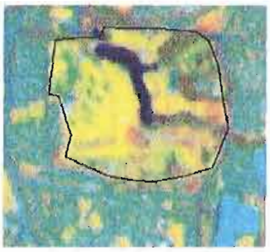
CORINE class 1.2.2
Road/Rail - M25/M2 junction

CORINE class 1.2.3
Port area - Tilbury docks

CORINE class 1.3.1
Mineral excavation - quarry in South Wales

CORINE class 1.3.3
Construction - residential area in Southwater (East Sussex)

CORINE class 1.4.1
Green urban - central Watford

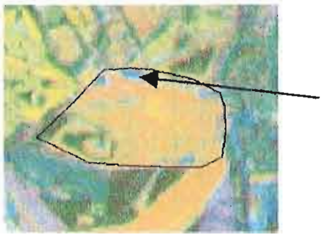


CORINE class 1.4.2
Leisure - historic house and estate

CORINE class 1.4.2
Leisure - Golf course, North London

CORINE class 2.2.2
Fruit growing - hop fields in Kent

CORINE class 1.2.4
Airport - Gatwick



Note: the blue (buildings) are included here because they are part of the infrastructure of the race course i.e. grandstands etc.

CORINE class 1.4.2
Leisure - Epsom race course

Note: different colour schemes have been used in the background images shown.

Figure 2. Examples of each of the CORINE land use classes

3.3 Removal of small parcels and extraction of 25 ha parcels with CORINE equivalence

Instead of using ARC/Info 'eliminate' or 'dissolve' to remove polygons < 2 ha, the 'region group' and 'nibble' commands were used in ARC/View to return a similar result in Grid format (see Figure 3).

Note:

The function **'REGION GROUP'** groups all adjacent orthogonal pixels of the same class into regions (GRID B) (see Figure 3). Conditional rules are then applied on the regions based on their size (i.e. number of pixels). If regions are less than a given size (i.e. in this case 2 hectares) all pixels in the region are given a null value (corresponding to 0 in Figure 3). If regions are greater than a given area they retain the class of the original grid (i.e.GRID A). The result is a mask grid (GRID C) which can be used in the nibble process.

The function 'NIBBLE' replaces the pixels of a Grid (GRID A) that were given null values in the Mask Grid (GRID C) with the values of their nearest neighbours (see Figure 3).

These procedures (see Annex) are similar to the ones described in detail in Brown *et al.* (1996) for ARC/Grid.

Figure 3. The regiongroup and nibble process showing classes in (A) formed into regions (B)

GRID A

8	8	8	5	5	5	5
8	8	8	5	5	5	5
8	8	6	6	6	5	5
1	8	8	6	6	5	5
1	1	1	6	1	5	5
1	1	1	1	1	5	5

GRID B

1	1	1	3	3	3	3
1	1	1	3	3	3	3
1	1	4	4	4	3	3
2	1	1	4	4	3	3
2	2	2	4	2	3	3
2	2	2	2	2	3	3

GRID C

8	8	8	5	5	5	5
8	8	8	5	5	5	5
8	8	0	0	0	5	5
1	8	8	0	0	5	5
1	1	1	0	1	5	5
1	1	1	1	1	5	5

GRID D

8	8	8	5	5	5	5
8	8	8	5	5	5	5
8	8	8	5	5	5	5
1	8	8	1	5	5	5
1	1	1	1	1	5	5
1	1	1	1	1	5	5

3.4 Clustering of smaller parcels

Growing polygons is a GIS procedure, which was originally developed, in the pilot study for the generalisation of 'outliers' (Brown & Fuller 1996). By growing and shrinking regions, small isolations in grid zone boundaries can be 'filled in' (see Figure 5), and small clusters of regions can be merged (see Figure 4). This process, involved 'cost distance mapping' and 'neighbourhood sum' functions. It has remained essentially unaltered except that the process was fully automated into the new ARC/View procedures. One click on the generalisation menu button executed the growing/shrinking procedure on all classes and created a series of grid layers, each of which contained the results of a particular class. Running this process in ARC/View reduced the time required to grow a single class in a 100 km square tile from well over 2 hours to around 10 minutes. This was significant, considering that all 25 classes in the grid needed to be processed.

Parcels were then allowed to grow by a greater increment than initially planned. Thus, more of the small parcels belonging to typically fragmented cover classes were allowed to join, to form parcels of greater than 25 ha (see Figure 4). The rules defining the sequence of priority in which classes were merged after growing and shrinking were reviewed and small changes implemented. The rules now used class statistics on parcel sizes and fragmentation.

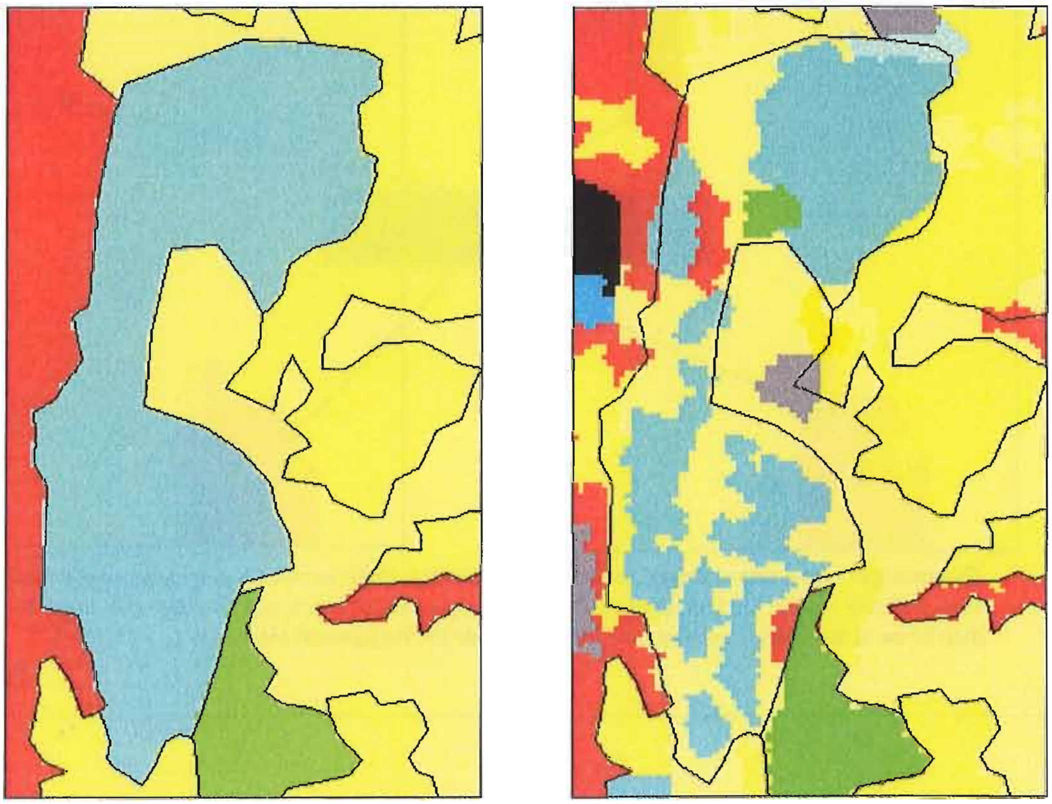
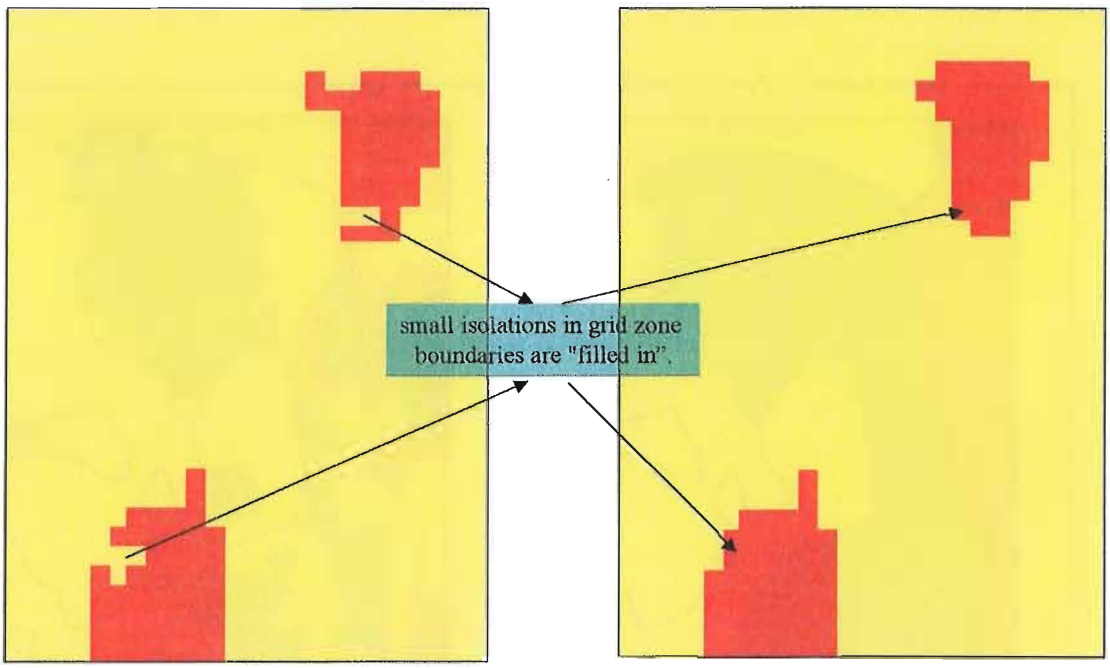


Figure 4. Aggregation of small patches

The process of dealing with unwanted boundary isolations and indentations was also considered during the final stage of boundary smoothing. In tests carried out during the generalisation work it was found that the standard 'smoothing of boundaries', such as the Douglas & Peuker (1973) methods employed by ARC/Info generalisation functions, did not produce a satisfactory result. Too many corners were unnecessarily smoothed, and there was a general lack of control of what happened during the smoothing process. Two requirements needed to be addressed: first the removal of the stepped effect of pixel edges, when producing the final vector output, and secondly, the retaining of significant angular boundary patterns, where their removal adversely affected the final generalised map product. These global smoothing routines proved not to be able to cope with both these issues successfully, possibly because of the large degree of generalisation that was necessarily imposed on our data. In fact, the stepped effect is largely invisible when the data are viewed at their target display scale of 1:100 000. At this scale the 25m edge of a pixel is 0.25mm on the screen. However, as soon as a user begins to zoom in, which users inevitably do, the stepped edge become apparent, hence the requirement for smoothing. Figure 5 below shows some initial removal of indentations but not the smoothing of the boundary. Although a number of smoothing options were tested, it was found that the smoothing carried out on these data, when using the ARC/View 'convert to shape file' function gave the best results. This was the method finally used for producing output vector data. This had the added advantage of being a standard function allowing for easy conversion by future users to and from vector and raster format within the ESRI software. Examples of smoothed output can be seen in Figures 4 and 7 and Annex Figure 1.



This example is at a very enlarged scale (approximately 1:1000) in order to make the pixel boundary shapes clearly visible. In fact, in this example, the upper red polygon is an isolated one, much less than 25 ha in size, and is subsequently dissolved into the background class.

Figure 5

3.5 Analysis and classification of mosaic parcels

The majority of changes made to the generalisation process were in the final stage of the analysis and classification of 'mosaic' or composite polygons. In the original process, this was partly an interactive stage, involving the user visually determining the composition of mosaics – a laborious method suffering from the problems of human subjectivity and error.

Using zonal geometry and zonal statistics functions in ARC/View it is possible to calculate the exact composition of mosaics in terms of % class composition. The results are stored in a table and queried, using the CORINE rules, to determine the correct land class code, and thus the CORINE code. The whole process of mosaics was automated. A schematic overview of the procedure is shown in Figure 6.

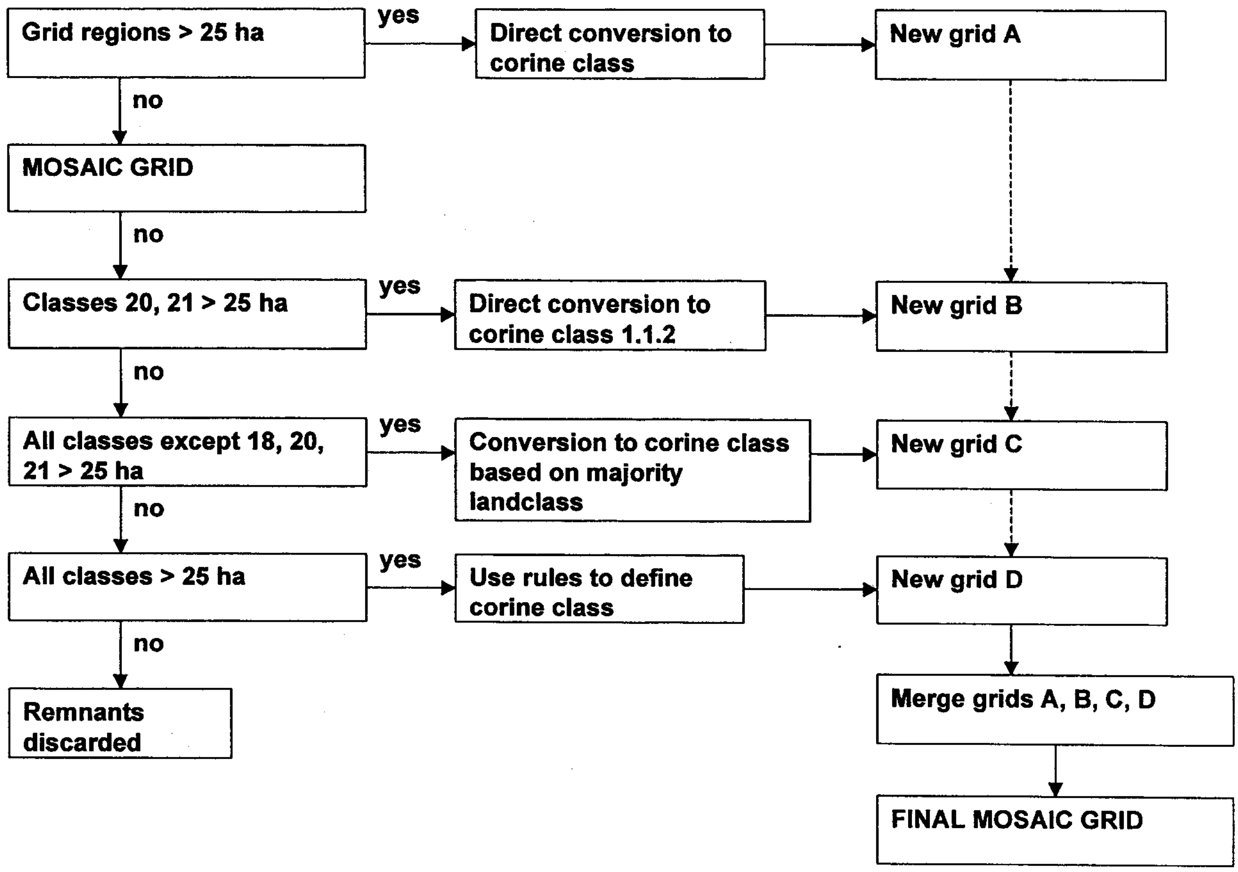


Figure 6. Schematic overview of the mosaic process.

Where heterogeneous parcels have been classified into CORINE mosaic classes, additional criteria have been incorporated, which include information on parcel sizes and fragmentation, using the rules defined by the EEA manuscript (Steenmans 1997).

3.6 Assignment of remaining small parcels to the most appropriate neighbouring class

In the feasibility study, the small parcels that remained after the classification of mosaic parcels were dissolved into the adjacent parcels using a combination of 'nearest-neighbour' and 'cost-distance' rules. The 'nearest-neighbour' rule tended to split the parcel into sections, which are dissolved into the adjacent parcels, creating new boundaries. In cases where a small parcel had more thematic affinity with one of its neighbouring parcels, this approach was not appropriate. The new procedure included a form of thematic generalisation. The removal of small parcels was first based on their thematic content. A small parcel was joined with the neighbouring parcel that had the highest thematic similarity or affinity. For instance a small parcel of mixed woodland would be joined to a neighbouring area of deciduous woodland rather than a neighbouring area of pasture. If the small parcel had no thematic affinity with any of its neighbours, or it was an island surrounded by a single cover type, then 'nibble' was applied.

3.7 Aggregation

Aggregation or edge-matching of the generalised UK sub-sections has been achieved by making the adjacent sub-sections overlap before they were aggregated. This permitted automatic checking of the overlap area as well as interactive analysis for features straddling the boundary between two sub-sections.

3.8 Final manual check

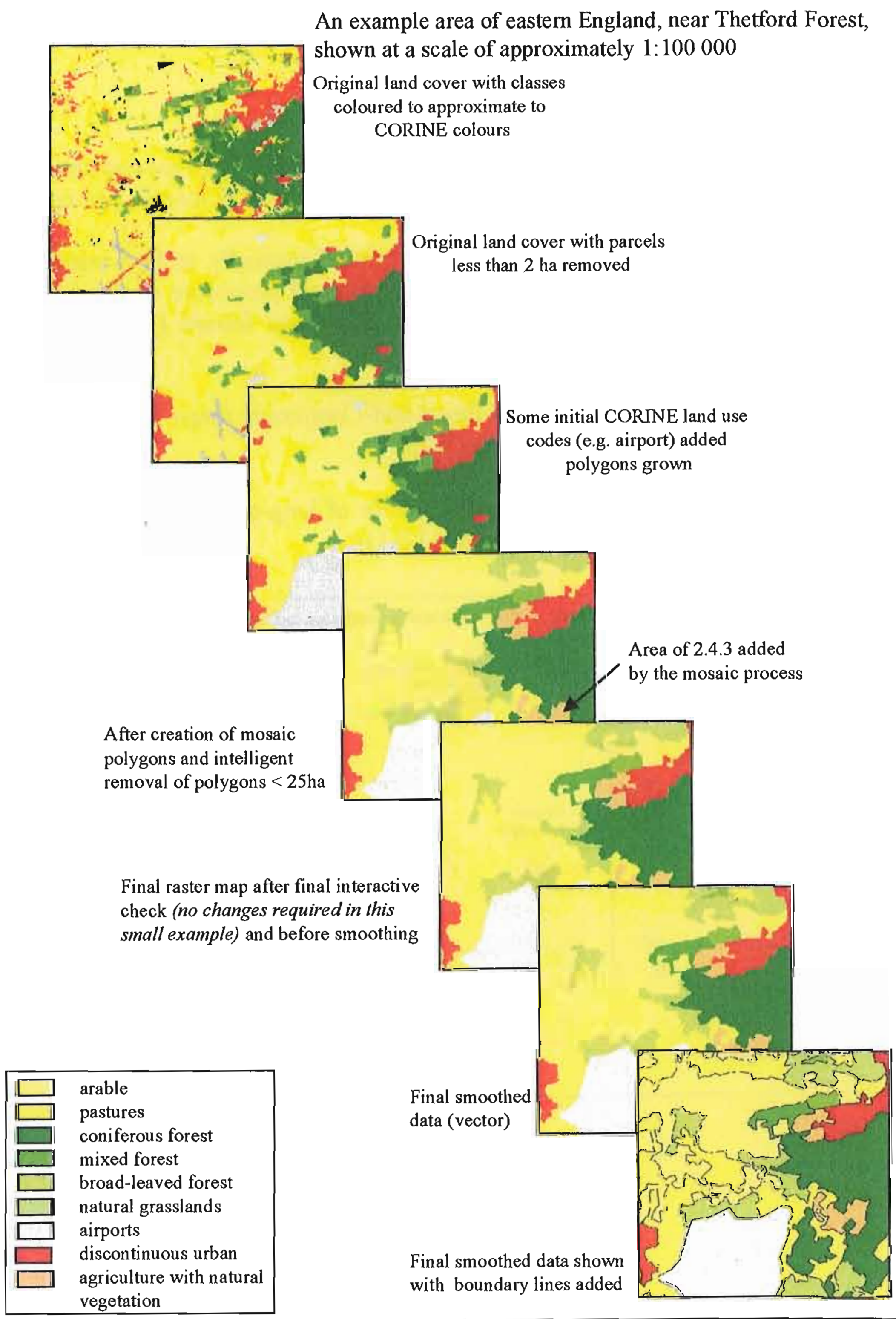
A manual check was now carried out on each of the 100km tiles. This was carried out by a different member of the team to the person who had worked on a tile at its earlier interactive stage. The main checks at this point were to identify the occasional cases of mis-classification, created during the semi-automated procedures. This may have been an over-enlargement of an urban area, due to the 'growing shrinking' process, or perhaps, the creation of a narrow linear feature, that did not meet the 100m wide rule for CORINE linear features. Sometimes the mosaic procedure had mis-classified an arable area as discontinuous urban or vice versa. This was also an opportunity to take 'a cartographers view' of the output, especially in relation to display at the 1:100 000 scale. This checking stage varied in length according to the complexity and type of landscape the tile was situated in, but took an average of about 1.5 days per 100km tile.

The raster data are then available for conversion to vector using the Arc/View 'convert to shape file' function. This provided the best vector 'smoothed' result when viewed at the 1:100 000 scale.

3.9 Changes in class cover caused by separate stages in the generalisation process

The semi-automated generalisation process imposes changes on the data at each stage. Table 2 in the Annex gives an indication of these effects within a single 100km square (TL in SE England), and Figure 7 illustrates the effect through a with example area from this square.

Figure 7. Changes through selected stages of the generalisation process



4. RESULTS

Conversion has been completed for all of GB. The very similar appearance of the manual & automated outputs (see examples in this section), and the high level of agreement in cover statistics for test sites (see Annex Tables 3 to 7), demonstrate that the procedure has achieved the desired output, and that the CORINE Land Cover Map of Great Britain conforms with the CORINE requirements.

Many types of statistical and visual comparison have been carried out during the generalisation process. Statistical analyses are contained in the Annex to this report. Some results have been reproduced in this section, such as in Table 2. Many visual examples are displayed. Figure 1 in the Annex shows a variety of examples from around Great Britain. These emphasise the differing types of landscape across the country and how they are represented by the CORINE Land Cover Map at various scales.

4.1 Evaluation and calibration

The standard procedure developed to validate the CORINE Land Cover map is described in CORINE technical guide. The method proposes use of extensive field surveying and aerial photography to check the accuracy of the CORINE Land Cover map product. The Land Cover Map of Great Britain is based on satellite imagery dated between 1989 and 1990. There is no substantial set of aerial photography available for that period and the data for 508 Countryside Survey 1990 (1 km^2) field samples, are inadequate for a validation exercise as described in the technical guide.

To assess the differences between automated and manual outputs, five test sites were selected and were plotted at 1:1 000 000 scale onto A3 size paper giving study areas of 43 km x 29 km. The sites were chosen to cover the four main landscapes in GB, defined for reporting Countryside Survey 1990, (i.e. Arable, Pastural, Upland and Marginal landscapes - see Barr *et al.*, 1993). The location of these test sites is shown in Figure 2 in Annex Section 7.5. Whereas any one site may comprise predominantly one landscape, there is some intermixing within the rectangular study areas. Test site results were disaggregated in order to analyse the results by landscape type and then to provide appropriate weighting according to the national extents of the landscapes. Table 1 lists the percentage of land covered by each of the four landscape types in GB. Together, the validation areas contain a variety of CORINE land-use and land-cover classes.

Table 1. Percentage of land covered by each of the four landscape types in GB

Landscape type	% coverage
Arable landscape	34
Pastural landscape	29
Upland landscape	21
Marginal landscape	16

4.2 Assessing the differences between automated and manual output for each test site

The CORINE manual interpretations were compared with the CORINE outputs from the semi-automated procedure. Comparisons were carried out, both visually (i.e. qualitatively) and through correspondence tables (i.e. quantitatively). Here we have included the more significant results from the assessment and the visual output. More details are available in the Annex or in the separate report to DGXVI (Gerard *et al* 1999).

Interpretation of CORINE classes was based on the same Landsat TM images which were used in the original per-pixel classifications. The interpreter was a member of the original LCMGB production team. This meant that he had good familiarity with the images, including the summer-winter composites and familiarity with the LCMGB classes and their visual interpretation (through training the image classifier). However, the interpreter had not practised long term operational mapping of CORINE cover; he was not fully familiar with all the test sites; he had difficulties in the interpretation of borderline classes, in the exact delineation of polygons, in the operation of the 25 ha minimum mappable unit rule and in the visual interpretation of mosaic classes.

The manual output is not a 'ground truth' dataset, and it incorporates a 10-15% error that would be normal for any CORINE interpretation. Its production was intended to identify substantial differences in the initial automated procedure, rather than to measure the 'absolute accuracy' of its output product.

Where differences have been identified, action has been taken to rectify them during the generalisation process. An example of this was the identification of a relatively high loss of small 'isolated' parcels during the 'growing shrinking stage'. This lead to a revision of these functions to allow more isolations to join together, forming a patch of over 25 ha that could be retained in its own right (see methods section for details).

The overall correspondence achieved between manual and semi-automated procedure was 83 %. This is believed to represent the correspondence of two maps (the CORINE land cover of GB and the manually interpreted map) each with *c.* 90% accuracy. All the correspondence tables produced for this validation exercise are reproduced in the Annex to this report.

Once complete, including manual corrections to automated outputs, the CORINE map of GB was further evaluated by comparing the class statistics of the LCMGB with those of the CORINE Land Cover map; and also evaluating the changes in class statistics caused from one generalisation step to the other. Results from these two exercises are shown in Annex Table 2 and Figure 7 (Methods section).

The final interactive checking took place after this ground reference comparison had occurred. It covered the whole of GB and was designed to deal with key discrepancies evident from the earlier comparisons. Table 2 indicates the changes in class coverage within one of the test areas (England arable landscape). Figure 8 shows a visual comparison of a selected section of this test area.

4.2.1 Arable landscape (Cambridgeshire)

The Cambridgeshire test site consists almost entirely of arable landscape with small pockets of pastural landscape. The visual quality of the automated output is excellent, closely matching the manual version (Figure 8). It maintains the more complex outlines of the original input, but this detail is not excessive and does not merit further generalisation. Indeed, it could be argued that the manual output oversimplifies outlines.

The correspondence matrix appears in Table 3 (Annex). It shows that overall the correspondence for this test site was 875 per thousand pixels, a good result. In recording this correspondence, it should be recognised that some deficiencies were noted in both the manual and automated product, which were addressed for all of GB, during subsequent checking stages. The comparison was repeated for this test site after the final interactive check, and gave a correspondence of 881 per thousand pixels.

The indicated level of accuracy exceeds normal expectations and gives considerable faith in the automated procedures. A closer look at class-level correspondence highlights the specific differences. However, it must be recognised that rarer classes (at least those which are rare in this study area), cannot be adequately assessed. Minor boundary differences for few smaller parcels may cause major quantitative differences; also the chance inclusion/exclusion of just a single polygon can cause a large percentage difference.

Some significant comparisons were:

- The manual method had been more generous in its inclusion of pasture and amenity grassland around villages as part of the *Discontinuous urban* (1.1.2). The final manual edit rectified this.

- The land use classes (1.2.* and 1.4.*) were captured manually in both methods. The two methods gave very similar results, except the fully manual procedure overlooked the need to record 1.4.1, and it scored a lower land use for *Sport and leisure* (1.4.2). Again this was increased during the final manual edit.

- The dissected patterns of *Broad-leaved forest* and *Coniferous forest* have led to mismatches caused by minor geometric shifts between the maps.

- The manual interpretation missed a significant area of long-term setaside, causing confusion between *Arable* (2.1.1) and *Natural grasslands* (3.2.1).

Figure 8. Comparison of outputs for a sample section of the arable landscape test site (Cambridgeshire), and the effect of final interactive editing.

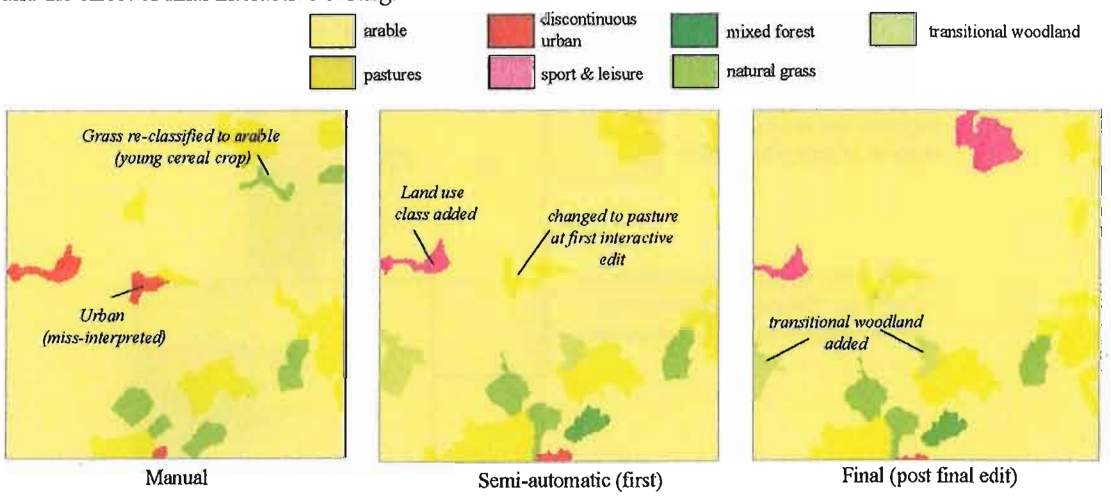


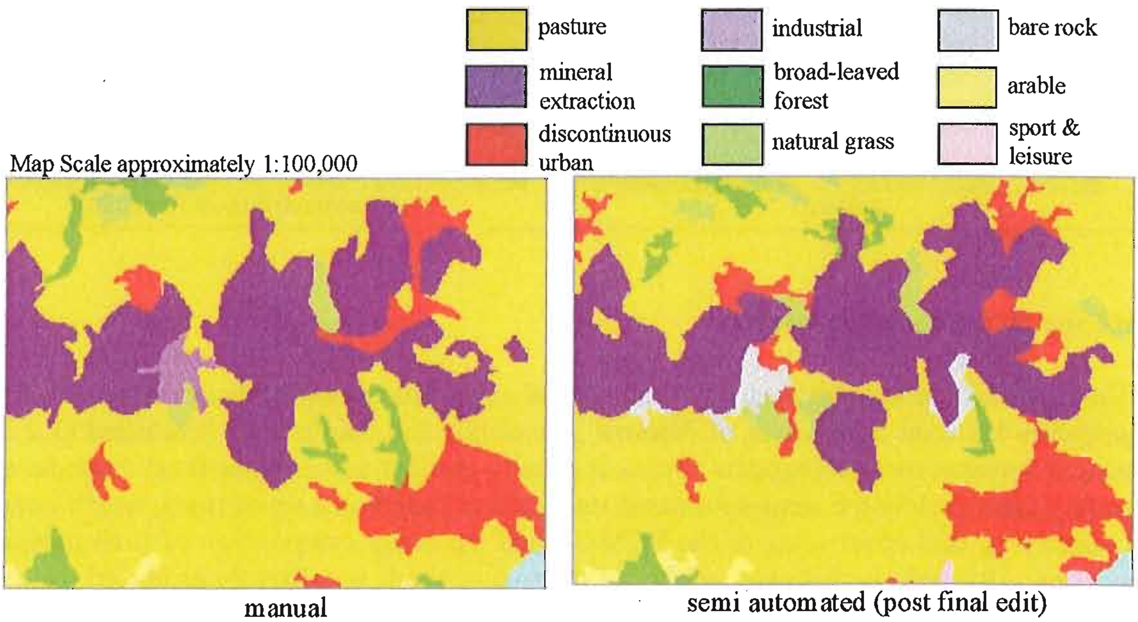
Table 2. Comparison for the test site showing changes in the total area per class within the arable landscape resulting from the final interactive edit stage.

Value	Land Cover Class	Manual	Interim	Final
111	urban	19.1	22.2	19.1
112	discontinuous urban	260.8	200.7	221.2
121	industrial	7.2	7.3	13.5
122	road/rail network			1.2
124	airports	49.4	51.4	55.3
131	mineral excavation	25.6	25.4	25.8
133	construction sites	7.5	7.6	5.4
141	green urban		9.3	4.0
142	sport & leisure	21.0	27.0	43.7
211	arable	3823.2	3673.5	3680.1
231	pastures	561.0	709.0	673.3
242	complex cultivation patterns		51.6	39.2
311	broad-leaved forest	52.3	33.8	33.8
312	coniferous forest	4.7	6.1	6.1
313	mixed forest		3.0	3.0
321	natural grasslands	9.6	20.0	19.9
324	transitional woodland		2.4	2.4
332	bare rock		3.9	1.8
512	water bodies	54.5	43.5	51.5

4.2.2 The pastural landscape (Devon)

The Devon test site is mainly made up of pastural landscape. Visual comparison shows a good match between the two products with a similar overall distribution pattern of the cover types present. The overall correspondence achieved between the two methods for the pastural landscape site was 863 pixels per thousand (Annex Table 4).

Figure 9. Comparison of outputs for a sample section of the pastural landscape test site (Devon)
This is at the location of a large mineral excavation, with its accompanying built infrastructure, within a pastural landscape.



This test site is dominated by *Pastures* (2.3.1), many small areas of *Broad-leaved forest* (3.1.1), and very large *Mineral extraction* sites (1.3.1). There are a few fields of arable crops, and these are generally found as single fields surrounded by pasture.

Some significant comparisons were:

- Annex Table 4 shows that the main discrepancies between the manual and semi-automated methods were caused by mis-matches between *Pasture* (2.3.1) and *Arable* (2.1.1), *Broad-leaved forest* (3.1.1), *Natural grassland* (3.2.1) or *Complex cultivation patterns* (2.4.2).

- In this landscape, the size of the arable areas was often near the 25 ha limit. As a result, the manual interpretation sometimes failed to identify arable areas that were marginally larger than the 25 ha limit, whilst in other cases, arable areas which were marginally smaller than the 25 ha, were included.

- The confusion between pasture and natural grassland simply reflected the difficulty in defining absolute boundaries, based on spectral data, in a continuum of grassland management regimes.

4.2.3 Marginal landscape (West Scotland and Wales)

Two sites were used for the comparison within marginal landscape. The test sites actually contain areas of marginal, pastural and upland landscape. The Welsh area (Figure 10a) has steep and narrow built-up valleys with coal mines (*Mineral extraction sites*: 1.3.1) and slag heaps (*Dump sites*: 1.3.2). Both of these required manual interpretation whereas the Scottish site (Figure 10b) contained few CORINE 1.*.* classes that required manual digitising. When assessed visually, the CORINE maps from semi-automated and manual product show a good match across the area. The percentage correspondence achieved for these test areas was 836 per thousand for Scotland and, initially, 650 per thousand for Wales (Annex Table 6). This was the lowest correspondence level achieved. The less reliable interpretations were addressed during subsequent checking stages throughout GB, and the comparison re-run for the Wales test site, achieving a new correspondence of 780 per thousand, shown in red in Annex Table 6 (see resultant changes in Figure 10b).

Figure 10a. Comparison of outputs for a sample section of the marginal landscape test site (Wales)

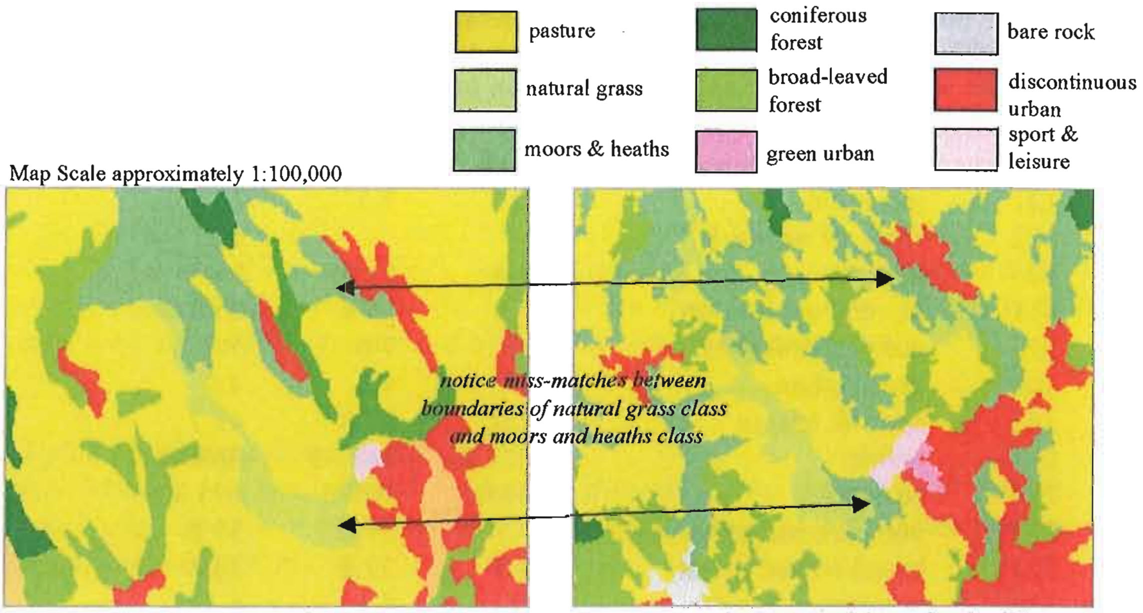
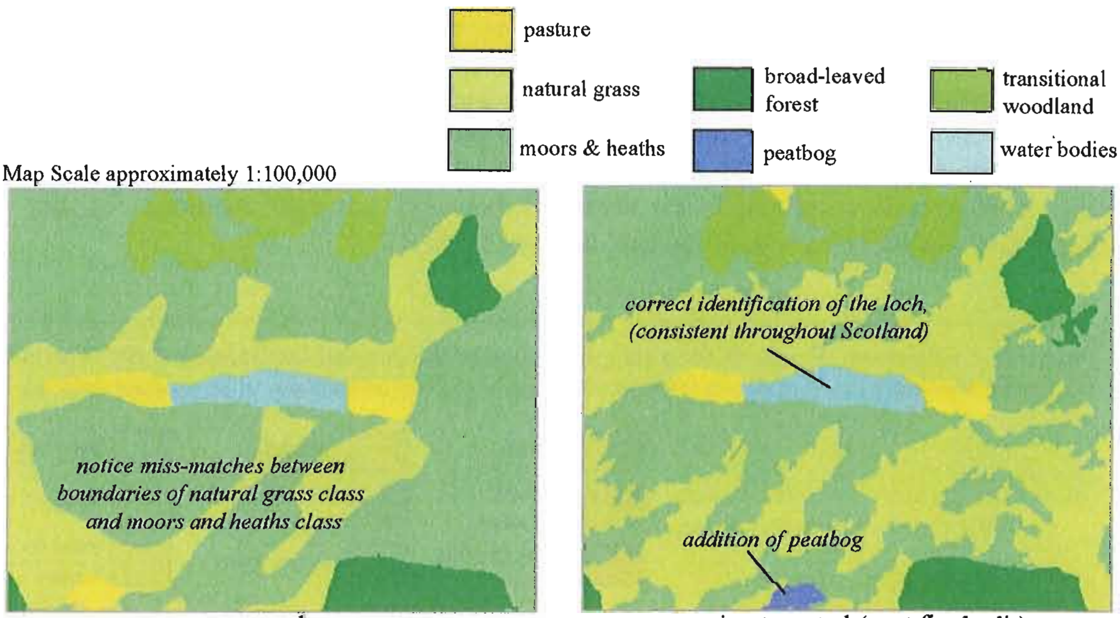


Figure 10b. Comparison of outputs for a sample section of the marginal landscape test site (Scotland)



Some significant comparisons were:

- The main differences between the manual and semi-automated methods were caused by the mismatches between boundaries of *Natural grasslands* and *Moors and heathland* (3.2.2). The transition between natural vegetation types is usually gradual, resulting in fuzzy boundaries. The boundary identified by the semi-automated method was based on a set of rules, which interrogate the underlying land cover class of the LCMGB and the cover composition of each polygon less than 25 ha. Although, in this case, neither of the two methods consistently produced the 'correct' answer, the semi-automated approach had the advantage of being objective in the way it identified boundaries between natural vegetation types.

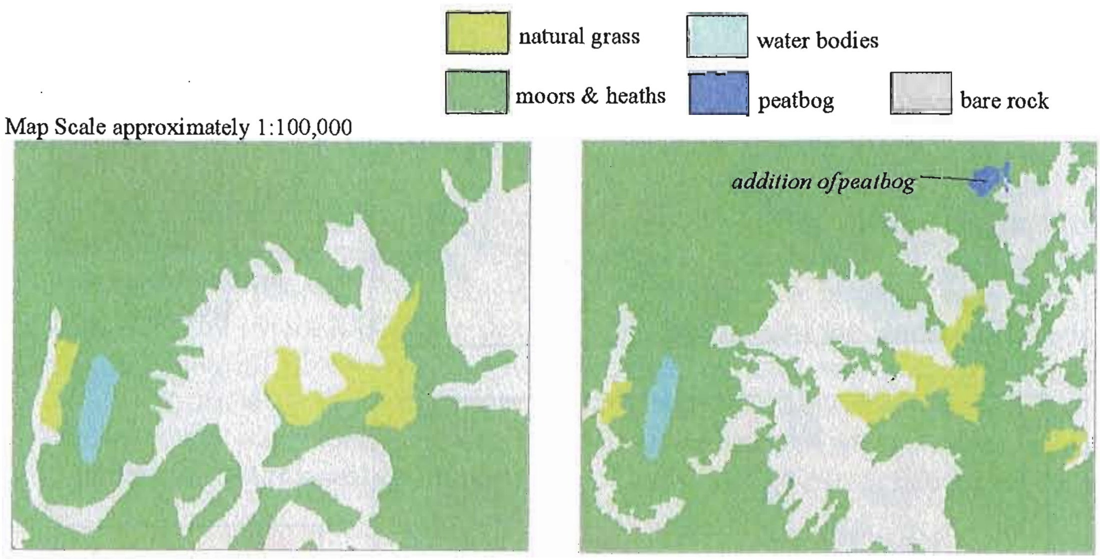
- In the Welsh site the complicated nature of the landscape with the narrow valleys, and the intricate mixture of natural vegetation types produced many polygons as thin slivers. The hilltops and slopes show an intricate pattern of heather moor, semi-natural grasslands and improved pastures with small pockets of woodland and smaller copses. A minor spatial shift between the two maps (manual and semi-automated) created some lower correspondence for these polygons, although the spatial pattern within the maps were similar.

- In the Scottish site, for the correct identification of the CORINE classes 3.1.1 (*Coniferous forest*) and 3.1.2 (*Broad-leaved forest*), manual digitising was necessary to ensure the inclusion of 'recently logged areas', which on the LCMGB are identified as bare. The many lochs in the landscape were successfully generalised by the semi-automated procedure.

4.2.4 Upland landscape (East Scotland)

The east Scotland test site is mainly upland landscape. Unlike the marginal test site of west Scotland, it contained fewer cover classes and the spatial distribution of classes was less intricate and complex (see Figure 11).

Figure 11. Comparison of outputs for a sample section of the upland landscape test site



manual

semi automated (post final

Some significant comparisons were:

- The manual and semi-automated output compare well visually.

- The overall correspondence for the test site was 911 pixels per thousand.

- The main difference between the two results was caused by mismatches between *Natural grasslands* and *Moors and heath* (Annex Table 7).

4.3 Assessing the overall differences between automated and manual output

The selected test areas, especially the ones representing the marginal landscape, included a proportion of landscape types other than the target one. This needs to be taken into account when calculating the overall correspondence results per landscape, and when assessing the correspondence for all of Britain, since the percentage of land covered by each of the four landscape types in GB is different for each landscape type (Table 1). Tables 3 and 4 below shows these calculations prior to the final GB editing stage.

Table 3. The correspondence results: total number of polygons and pixels for the five test sites.

Test site	Corresp. (pixels/1000)	Total n° polygons	Total n° pixels ($x\ 10^3$)
Arable, Cambridgeshire	875	322	1224000
Pastural, Devon	863	236	1201635
Upland, east Scotland	911	77	1200884
Marginal, west Scotland	836	278	1200884
Marginal, Wales	650	421	1229936

Table 4. Results found after assigning the test site pixels to the four landscape types.

Landscape type	Corresp. (pixels/1000)	Total n° pixels ($x\ 10^3$)
Arable	863	1490375
Pastural	807	1712610
Upland	890	1635504
Marginal	714	1192451

Note. Further details of these results were described in the report on Evaluation of semi-automated procedure (Gerard *et al.* 1999).

The results for the landscape types show a small reduction in correspondence ranging from 10 to 60 pixels/1000. This was due mainly to the contribution of data from the Welsh test site, where results were generally poorer. The reduction in correspondence is particularly high for the pastural landscape (60 pixels/1000). The pastural landscape is also the landscape type to which the highest proportion of pixels from the two marginal test sites was re-assigned (*c.* 5×10^8 pixels). Moreover, most of these re-assigned pixels originate from the marginal test site of Wales, which contains many long thin polygons. The lower correspondence values, achieved for these thin polygons, has resulted in the lowering of the overall correspondence of the pastural landscape.

An overall correspondence matrix was calculated from the correspondence matrices for the four landscape types by weighting their contribution according to national coverage as follows:

$$C_{GB} = P_a C_a + P_p C_p + P_u C_u + P_m C_m$$

where C_{GB} is the overall correspondence, P_* is the proportion of land covered by landscape * and C_* is the correspondence within the landscape *.

The calculated value of C_{GB} was 83%. The LCMGB is estimated to be 80%-85% accurate. Supposing (for the moment) that CORINE from manual interpretation were to have had an accuracy of say 90%-100%, and that the semi-automated generalisation procedure perfectly mimics the manual generalisation process, we would expect an overall percentage correspondence to be 72%-85%. The overall correspondence achieved after the final interactive stage was 83%, nearly at the maximum likely value. It is evident that generalisation is removing 'noise' in the data, producing an output which, in CORINE land cover terms, is more accurate than the input. An 83% correspondence might be expected when comparing automated and manual versions, each with >90% accuracy.

Given that the manual mapping, used to check the correspondence, is probably no better than 90% correct (a 'best guess' as we have no 'ground truth' data to prove or disprove that assessment), the value suggests that the automated product may have achieved a similar level of 'accuracy'.

An additional statistical check of the correspondence data was carried out. This was the calculation of the Kappa Coefficient. This is used as a measure of association to test the degree of agreement (reliability or precision) in classification. The result is a value between 0 and 1, with 1 representing a complete agreement. Landis and Koch (1977) state that a Kappa value 0.395 could be termed 'marginally good', 0.6 is 'good' result for this test. The values obtained for the 5 test sites are shown in Table 4 below.

Table 4. Kappa coefficient values for each of the five landscape sites.

Landscape site	Kappa Coefficient
Arable land	0.72
Pastural land	0.75
Upland	0.72
Marginal land (Wales)	0.65
Marginal land (Scotland)	0.79

The achievement of 81% correspondence for marginal land shows the difficulty of mapping in this complex landscape, with its small-scale patterns of highly variable land use. It is clear that pasture cover types, especially those which are also dissected into many small units at, near, or below the 25 ha minimum mappable unit, are highly variable in their interpretation between the manual and semi-automated methods. The 81% probably represents the overlap between two products both less accurate than the national average, each perhaps 85-90% correct. Thankfully, the marginal landscape is the least extensive of the four types, so its contribution of error to the national total does not substantially impact upon the overall result.

The 81% result for pastural landscape reflects the complexity of this landscape, though such complexity is less than that of the marginal type. The result could still imply the overlap between two products each with 90% accuracy.

The correspondences for arable landscape (86%), and upland landscape (89%) was better than might have been expected. The results suggests that, in these simpler landscapes, the generalisation to CORINE format is removing erroneous 'noise' in the per-pixel classification with the greatest effect.

4.4 CORINE Simplification and The Countryside Information System (CIS)

The CORINE map and data are produced to 3 levels of detail. The examples described in this report have, so far, concentrated at the level 3 information, that is, the fullest detail in resolution and class terms (44 classes in Europe). Figure 12 below illustrates the variation in detail at the three different levels.

The data have also been imported into the ITE Countryside Information System (CIS) for analysis in combination with other environmental parameters, at the 1km level. The CIS is a Microsoft Windows-based program developed to give policy advisers, planners and researchers easy access to spatial information about the British countryside. The CIS uses the Windows operating system and follows normal Windows conventions. The CIS allows users to:

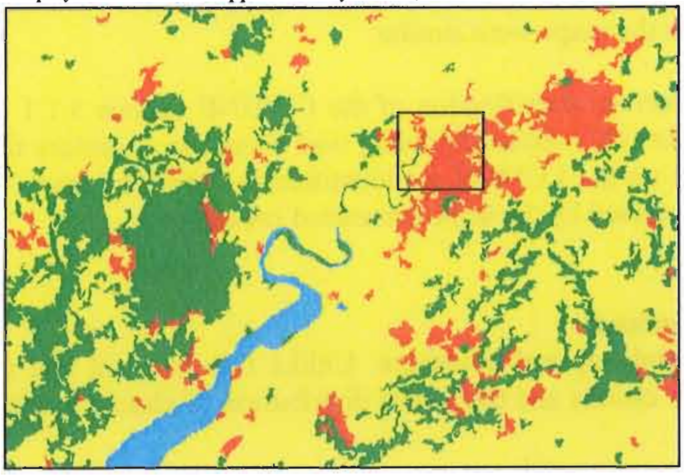
- Combine, analyse and present a comprehensive range of environmental data
- Characterise geographical regions in terms of a range of environmental parameters
- Define and select geographical areas with different environmental features
- Produce maps, tables and charts
- Export and import data to and from file

CIS uses the kilometre squares of the National Grid as a framework. Sample data provided with the CIS are derived from the Countryside Survey 1990 (CS1990) and two previous Institute of Terrestrial Ecology surveys of 1978 and 1984. The data include information on land cover, land use, landscape linear features such as hedges and ditches, plus information of plant species distributions. The sample surveys were conducted using the ITE Land Classification which uses combinations of mapped environmental data (such as geology, climate and topography) to allocate land to one of 32 different classes in GB. Sample-based observations can then extrapolated using the Land Classification.

CIS also includes a large range of census-based information at national scales, with the 1km x 1km cells of the National Grid used as the basic spatial unit. The core census data are the ITE Land Cover Map of Great Britain. Figure 13 shows an example of CORINE output from this system. The addition of the CORINE Land Cover dataset allows users to compare and contrast the detailed sample-data, the raster-based LCMGB and the CLC product. This effectively integrates the Countryside Survey products into the European framework and, similarly, for GB, provides a detailed perspective to the European CLC product. The existence of CLC in CIS will help its integration into the GB policy, planning and research processes.

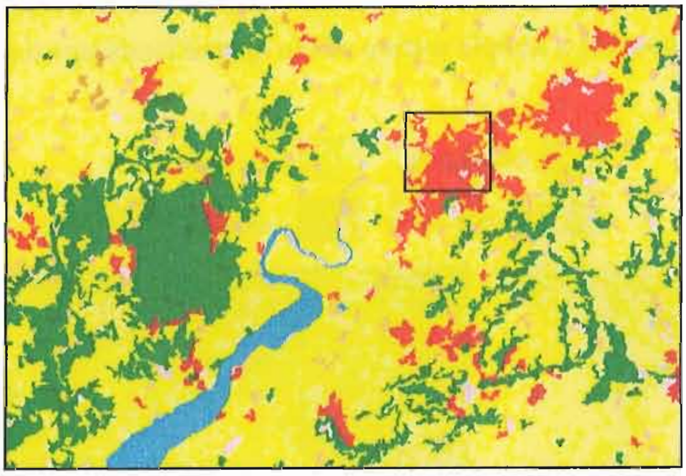
Figure 12. A visual comparison of the ITE CORINE land cover map at CORINE levels 1, 2 and 3. An example from the estuary of the River Severn

displayed at a scale of approximately 1: 250,000

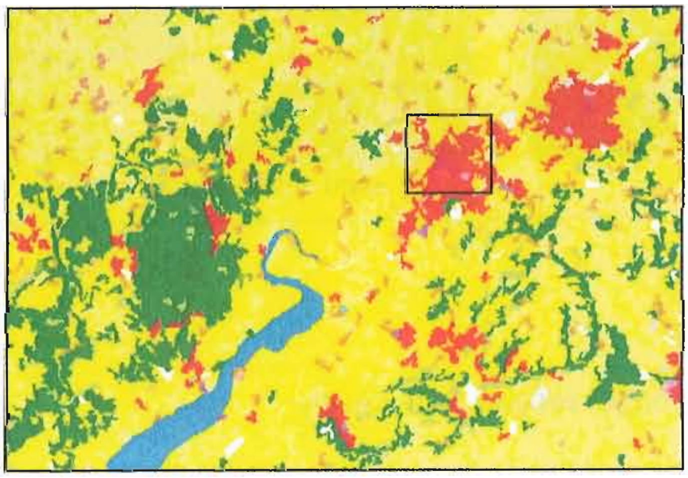


displayed at a scale of approximately 1: 50,000

3 classes present in this area at level one

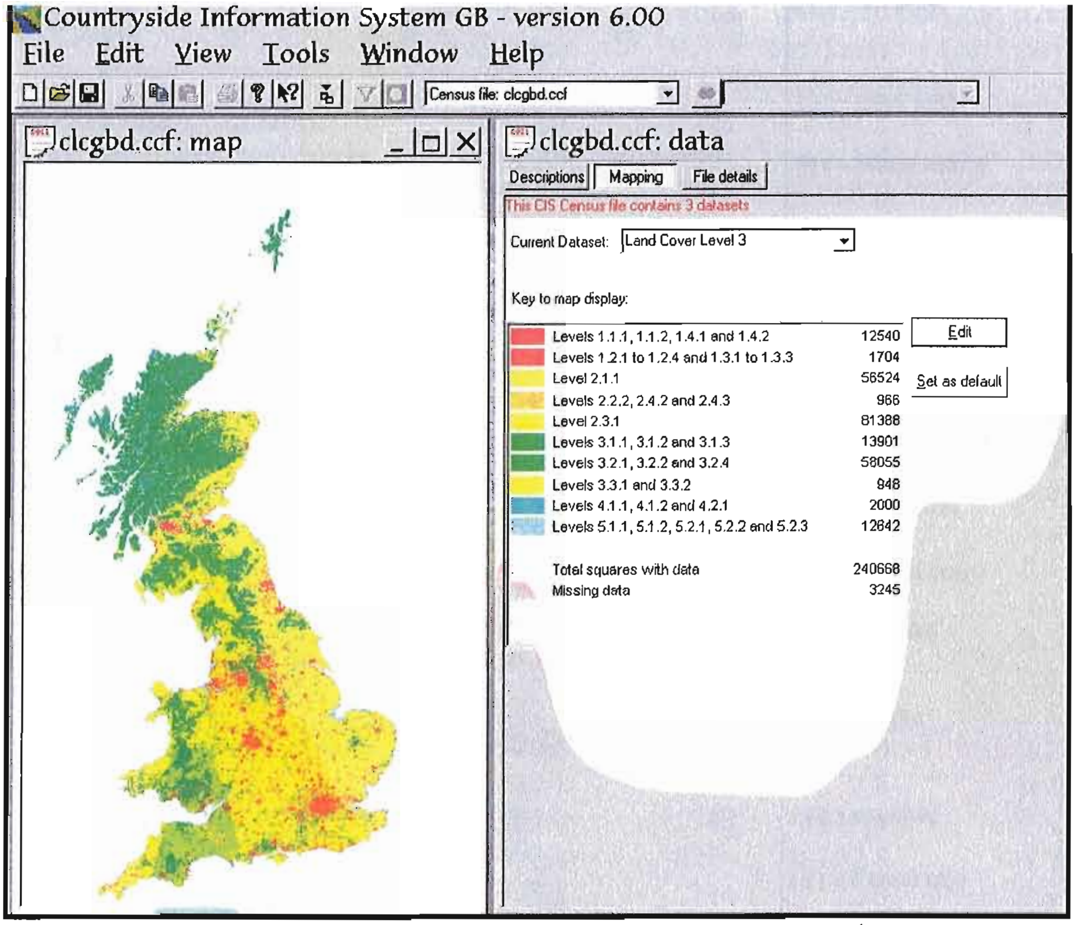


Increased detail at level two



More sub classes in both the urban and arable areas

Figure 13. CORINE land cover shown as the dominant class per 1km square within the Countryside Information System.



4.5 Output map statistics

Table 5 below shows some basic statistics from the CORINE Land Cover map of GB. This table is given in its fuller form in the Annex (Table 1)

Table 5. National statistics from the final CORINE data

Land Cover Class	Total area within GB (sq. kms)	Number of polygons	Largest polygon (sq. kms)	Mean polygon size (sq. kms)
Continuous urban fabric	1113.7	725	2032.2	1.5
Discontinuous urban fabric	10427.4	4801	6723.6	2.2
Industrial or commercial units	1016.9	1118	210.6	0.9
Road and rail networks	41.3	90	15.2	0.5
Port areas	104.3	70	90.1	1.5
Airports	387.1	215	8.8	1.8
Mineral extraction sites	417.8	548	17.5	0.8
Dump sites	60.6	103	2.3	0.6
Construction sites	22.6	34	2.5	0.7
Green urban areas	467.0	735	10.9	0.6
Sport & leisure facilities	1526.5	2422	7.0	0.6
Non-irrigated arable	52798.6	10020	30127.9	5.3
Fruit trees & berry plantations	150.0	238	2.8	0.6
Pastures	75799.4	16959	11893.5	4.5
Complex cultivation patterns	2906.0	6425	17.0	0.5
Agriculture with nat' veg'	710.5	1870	1.1	0.4
Broad-leaved forest	7142.6	9216	60.8	0.8
Coniferous forest	9780.8	4422	269.0	2.2
Mixed forest	389.1	879	7.8	0.4
Natural grasslands	22767.6	11743	1937.4	1.9
Moors & heaths	35697.4	6256	18227.6	5.7
Transitional woodland	479.1	314	22.0	1.5
Beaches, dunes, sands	302.0	214	25.5	1.4
Bare rock	1032.9	1039	31.6	1.0
Inland marshes	39.1	50	3.7	0.8
Peatbogs	2371.8	2333	51.5	1.0
Salt marshes	368.8	281	12.3	1.3
Intertidal flats	1372.6	699	153.1	2.0
Water courses	42.5	20	8.6	2.1
Water bodies	1632.3	1316	75.2	1.2
Coastal lagoons	2.1	4	0.8	0.5
Estuaries	2293.6	90	688.8	25.5

Some points of interest:

1. Major classes have remained largely unchanged in extent as a result of the generalisation.
2. Some rarer classes have suffered significant loss during the generalisation process (this was explained in section 3.2). For instance, the total 369 km^2 for salt marsh in the CORINE map compares to a national figure of 402 km^2 from the source LCMGB. This was a rare class with direct conversion between the two sets of data. The loss was mainly due to small salt marsh polygons being generalised out of the map.

3. Table 1 in the Annex shows comparison of the 'linearity' of classes (shape index); for instance, road and rail networks have a high value of .969 which shows linear, sometimes more complex features. Airports have a low value, indicating their broader, less complex shape.
4. Only 34 construction sites (larger than 25ha) have been identified. Relatively few construction sites are larger than this at a specific point in time. For instance, a new urban development often takes place on a gradual basis. The final 'new area' will be greater than 25ha but it was not all under construction at one time. However, major construction projects are identified, for instance the development of the Channel Tunnel complex, in Folkestone, was under construction at this time.
5. Mean arable patch size is NOT the mean field size. For instance a large area of Fenland England consists of many adjoining arable fields. These are represented as much larger single patches of arable land.
6. Few fruit growing areas (orchards etc) are represented individually, as most fall below the 25 ha limit and will have been subsumed into surrounding arable classes etc.
7. Over 1000 industrial areas are identified, the largest continuous area is over 200 km^2 and is found around the estuary of the River Tees in NE England.
8. 215 airports have been included in the map.

4.6 Effort

Four ITE staff were primarily involved with the creation of the CORINE Land Cover Map of GB Some effort was devoted to managing the project, and for archiving and managing the data input and output.

The map generalisation process itself can be divided into two major components;

1. Computer processing of data through the various stages described in this report.
2. Manual interpretation, checking and editing of the data.

4.6.1 Computer processing

As described earlier, much of the computer processing was carried out in 'batch mode' that is, a suite of processes were run automatically, on varying number of datasets. In the development and testing stages, this may have been on a single set of data, later processing involved larger selections of data, and ultimately some processes were run on the whole GB dataset.

The time taken for the process of 'growing classes' (described in section 3.4) can be divided into two elements. About two days (16 hours) were used to develop the ARC/View script in the Avenue programming language (see Annex 7.8), and test it. Individually, the run time was about 10 minutes for each cover class for each full 100km land tile (part tiles took proportionally less time). This was then followed by a second run using all classes combined, to produce a merged grid. This procedure took about 14 days to complete, mainly running overnight. This was the most computer intensive of the processes done.

Table 6 gives time estimates for the whole generalisation process. However, some processes were run overnight (taking between 1 and 10 hours), or whilst the operator was working on other tasks, such as manual checking of other tiles. Some tasks were run concurrently by a single operator. Similarly, there were significant improvements to working methods during the life of the project, as well as some time spent on developing the GIS functions. This averages about 2.8 days computer time per 100km tile.

Table 6. Estimates of time taken for computer generalisation processes.

Conversion process	Approximate time taken (days)
Import LCMGB data	2
Import exogeneous data	9
Removing parcels < 2ha	6
Growing & merging grids	14
Mosaic, stages 1,2 & 3	9, 4, 5
Code conversion	4
Nibble & regiongroup	11
Removing isolations < 25 ha	4
Convert to vector (shape files)	4
Data management and archive	17
Function development	28
Digitising test sites	4
Correlations for test sites	3
Total	**132 days approx'**

* *Note, some processes were run concurrently, and/or overnight etc, so the estimates do not necessarily represent the amount of manual time that was involved.*

4.6.2 Manual interpretation, checking and editing

Details of the two main manual stages are shown in sections 3.2 and 3.8. In total they took about 4.5 days per 100km tile; about 198 days in total. This also includes some input by the GIS technical advisor.

4.6.3 Overall

The grand total for staff time (not including project management etc) is about 330 days with conversion completed within 1 year. The cost of the conversion was E200k or c. E0.8/km^2. This compares favourably with a figure of E5/km2 generally quoted for CORINE production. Taking account ot the LCMGB production costs, overall costs are <E1000k, somewhat less than CLC production by conventional means which would have been E1200k. The resulting products offered the finer resolution, in terms of minimum mappable unit, required for GB use, plus full compatibility with CLC at the European level, with integration of both products.

5. CONCLUSIONS

- The CORINE Land Cover Map of Great Britain has been created by semi-automated generalisation methods. It has successfully used input from Land Cover Map of Great Britain (LCMGB) 1990, together with exogenous data from many other sources.

- The generalising process has retained a significant level of manual and cartographic interpretation. However, the result has been mostly achieved through automated GIS functions.

- It is believed that the very similar appearance of both manual and automated outputs, the high level of agreement in cover statistics for test sites, and the overall levels of correspondence, demonstrate that the procedure has achieved the desired output and the CORINE Land Cover Map of Great Britain conforms with requirements and can be integrated into the European dataset.

- The procedure has been validated against standard CORINE manual methods with considerable success, and an indication that the target accuracy is exceeded substantially.

- In all map generalisation processes, some elements of the source information are changed, or perhaps lost altogether. This particularly affects small map objects. This effect has been found in the semi-automated generalisation process. Both the statistical and visual results have been described in this final report.

  Major classes have remained largely unchanged.

  Some rarer classes have suffered inevitable loss during the generalisation, with polygons being generalised out of the map.

- Map data have been summarised as cover per class per 1 km square and built into a PC-based information system (CIS).

- Copies of final digital data, in both vector and raster form, have been provided to DETR, DGXVI and members of the CLCGB Advisory Group

- Hard copy, such as maps of the whole of GB and posters describing the product, are being made available from ITE, and these will be provided to a number of organisations and individuals such as DETR and the European Topic Centre on Land Cover.

- The procedure has proved highly cost-effective with conversion costing 17% of the price of new production and the combined cost of producing the automated LCMGB plus the semi-automated conversion to CLCGB being 80-85% of manual CORINE production costs.

6. REFERENCES

Gerard, F.F., Fuller, R.M., Edmundson, D. & Brown, N.J. (1999). Generalising the Land Cover Map of Great Britain to CORINE Land Cover by Semi-automated Means. Evaluation of semi-automated procedure. Report to Directorate general XVI (European Commission).

Barr, C.J., Bunce, R.G.H., Clarke, R.T., Fuller, R.M., Furze, M.T., Gillespie, M.K., Groom, G.B., Hallam, C.J., Hornung, M., Howard, D.C. & Ness, M.J. 1993. *Countryside Survey 1990: main report*. Countryside 1990 Series: Volume 2. Department of the Environment, London.

Bunce, R.G.H., Barr, C.J., Gillespie, M.K. & Howard, D.C. (1996). The ITE Land Classification: Providing an environmental stratification of Great Britain. *Environmental Monitoring and Assessment*, **39**, 39-46.

Fuller R M & Brown. N J. 1994. *A CORINE map of Great Britain by automated means: a feasibility study.* Unpublished Institute of Terrestrial Ecology report to the Department of the Environment.

Fuller, R.M., Groom, G.B. & Jones, A.R. 1994a. 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., Groom, G.B. & Wallis, S.M. 1994b. The availability of Landsat TM images for Great Britain. *International Journal of Remote Sensing.* **15**, 1357-1362.

Fuller, R. M. & Brown, N. 1996. A CORINE map of Great Britain by automated means. Techniques for automatic generalization of the Land cover map of Great Britain. *International Journal of Geographical Information Systems,* **8**, 937-953.

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.

Douglas, D. & Peuker, T. 1973. An algorithm for the reduction of the number of points required to represent a digitised line and its caricature. *Canadian Cartographer*, **10**, 112-122.

Landis, J.R., & Koch, G.C. (1977). The measurement of observer agreement for categorical data, *Biometrics,* **33**,159-174.

Steenmans, C. 1997. Mapping Criteria for CLC Generalisation. Internal report to CLC Topic Centre.

McVay, K.R. 1998. Dissolve: an Avenue script for merging polygons. University of Houston, Clear Lake.

7. ANNEX

7.1 Contents

Figure 1. selections from GB using the CORINE land cover map

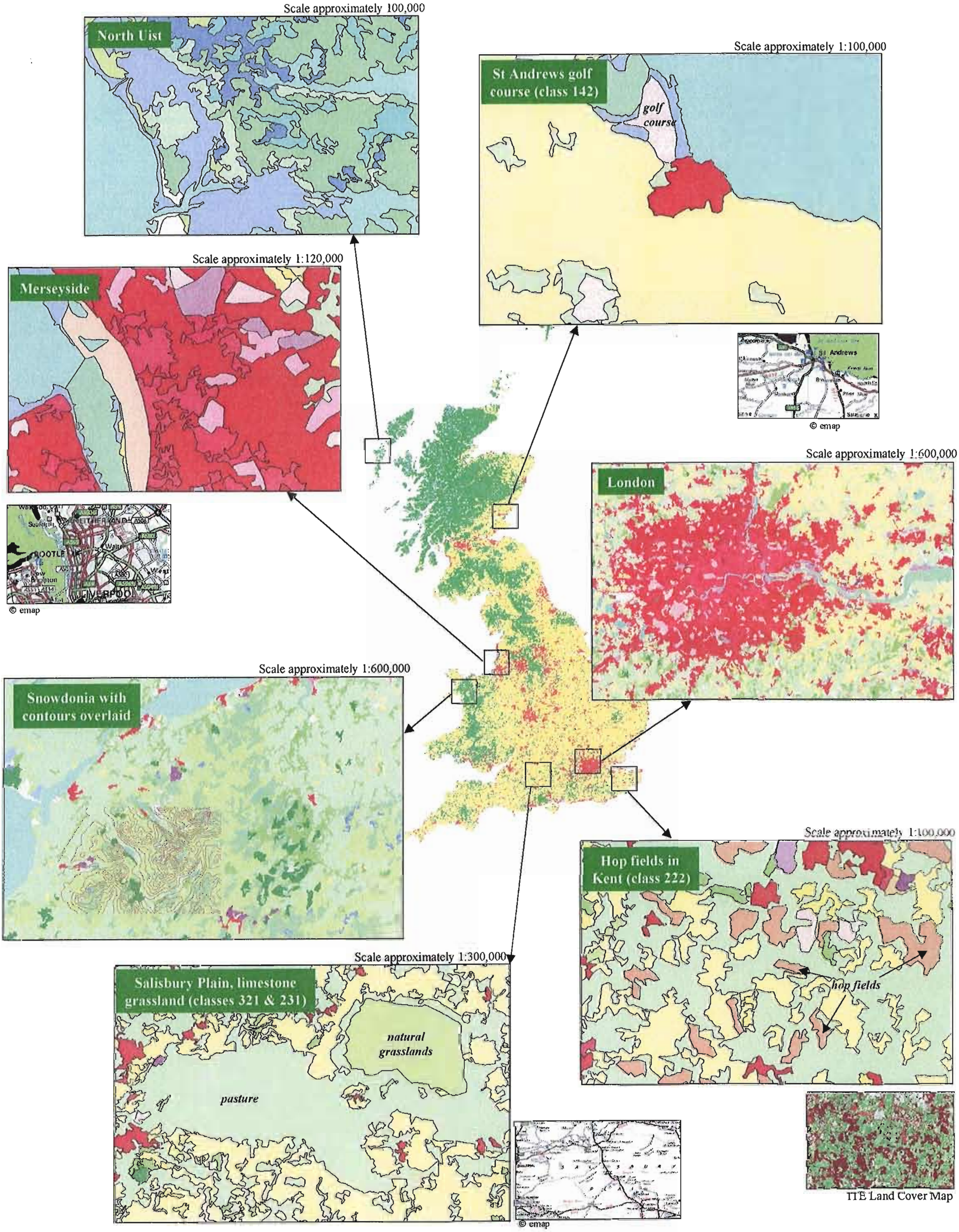


Table 1. National statistics from the CORINE land cover map of GB

corine	class description	Area sq kms	no. polygons	largest patch	mean patch size	mean boundary	shape index
111	urban	1113.7	725	2032.2	1.5	11.8	0.61
112	discontinuous urban	10427.4	4801	6723.6	2.2	13.4	0.49
121	industrial	1016.9	1118	210.6	0.9	5.4	0.47
122	road/rail	41.3	90	15.2	0.5	5.6	0.97
123	port areas	104.3	70	90.1	1.5	11.3	0.60
124	airports	387.1	215	8.8	1.8	8.0	0.35
131	mineral	417.8	548	17.5	0.8	5.0	0.52
132	dump sites	60.6	103	2.3	0.6	4.3	0.58
133	construction	22.6	34	2.5	0.7	5.2	0.62
141	green urban	467.0	735	10.9	0.6	5.9	0.73
142	sport & leisure	1526.5	2422	7.0	0.6	4.8	0.61
211	arable	52798.6	10020	30127.9	5.3	22.3	0.34
222	fruit & berry	150.0	238	2.8	0.6	5.1	0.64
231	pastures	75799.4	16959	11893.5	4.5	22.6	0.40
242	complex cultivation	2906.0	6425	17.0	0.5	6.0	1.06
243	agriculture with nat veg	710.5	1870	1.1	0.4	5.3	1.11
311	broad leaved forest	7142.6	9216	60.8	0.8	8.5	0.87
312	coniferous forest	9780.8	4422	269.0	2.2	10.8	0.39
313	mixed forest	389.1	879	7.8	0.4	5.4	0.98
321	natural grasslands	22767.6	11743	1937.4	1.9	13.7	0.56
322	moors & heaths	35697.4	6256	18227.6	5.7	23.4	0.33
324	transitional woodland	479.1	314	22.0	1.5	7.7	0.40
331	beaches	302.0	214	25.5	1.4	12.6	0.71
332	bare rock	1032.9	1039	31.6	1.0	8.7	0.70
411	inland marshes	39.1	50	3.7	0.8	6.8	0.69
412	peatbogs	2371.8	2333	51.5	1.0	10.0	0.78
421	salt marshes	368.8	281	12.3	1.3	11.0	0.67
423	intertidal flats	1372.6	699	153.1	2.0	14.8	0.60
511	water courses	42.5	20	8.6	2.1	23.8	0.89
512	water bodies	1632.3	1316	75.2	1.2	7.7	0.49
521	coastal lagoons	2.1	4	0.8	0.5	6.5	0.96
522	estuaries	2293.6	90	688.8	25.5	48.6	0.15

Table 2. Changes in class extent (hectares) through individual stages of the generalisation process.
Data for one 100km square (TL SE England)

Class CORINE (*LCMGB*)	**LCMGB original**	**After removal of polygons of less than 2 hectares**	**After growing polygons together**	**After creation of mosaic polygons and removal of polygons of less than 25ha**	**After first Interactive creation & editing of land use polygons**	**Final map - after final interactive check and edit**	**%-change from original to final**	**Change in the % of square covered by this class**
Arable (2.1.1)	23868	25551	25216	27044 *gained by absorption of many of the <25ha polygons from other classes*	27291 *main gains here were from incorrectly classified disc urban class*	27290	+14.3	+3.2
Broad-leaved forest (deciduous) (3.1.1)	1285	1144 *many small patches lost in this class in this 100km square*	1148	739	733	733	-42.9	-5.52
-gone to mixed(3.13)				total 38	36	36		
-gone to green urban (1.4.1)	This class includes some tree cover estimated from typical examples at about 20% of whole polygon for this square			total 27(5)	60(12)	60(12)		
-gone to sport (1.4.2)	This class includes some tree cover estimated from typical examples at about 10% of whole polygon for this square			total 271(27)	353(35)	360(36) *identification of missing sports venue's*		
-gone to transitional woodland (32.4)				total 16	16	16		
-gone to fruit/berry (2.2.2)	This class includes some tree cover estimated from typical examples at about 60% of whole polygon for this square			total 25(15)	42(24)	42(24)		
Coniferous forest (3.1.2)	585	634 *gained from internal small patches of other classes*	644	640	640	640	+9.4	+0.55
All trees? In squares with larger forest areas, these classes gain rather than lose	1870	1778	1792	1480	1496	1497	-19.9	-3.37
Salt marshes (4.2.1)	51	52	53	47	47	47	-3.9	-.04

7.5 Results from validation exercise

The earlier report to DGXVI (Gerard *et al*, 1999) gives a comprehensive analysis of the statistical comparisons carried out within the validation exercise. These are summarised in the results section of this report, but they have been modified to take account of the improved results obtained after the final editing process was done. Tables 3 to 7 below show the detailed results. Figure 2 below shows the location of the five test sites.

Figure 2. Location of the fives test sites used for the evaluation.

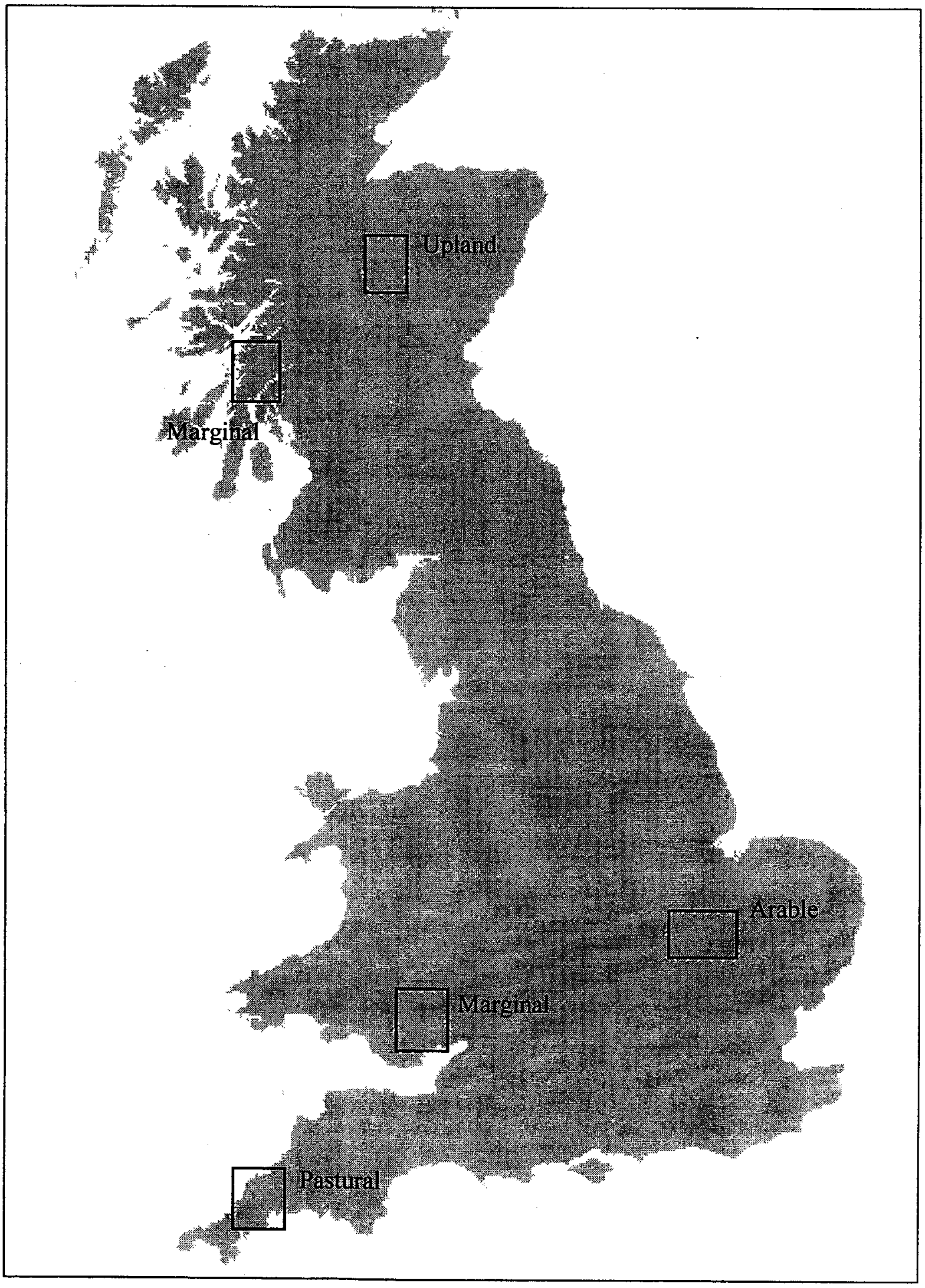


Table 3. Correspondence per 1000 pixels, between manual and automated CORINE interpretations (validation for Arable) note improved correspondence in post edit figures (see in red)

MANUAL (columns) / AUTOMATED (rows)

CORINE class	111 - cont urb	112 - discont	121 - indust	124 - airport	131 - mineral	133 - constr	141 - green urb	142 - recreat	211 - arable	231 - pasture	242 - mixed cult	243 - agric/nat	311 - broadlf	312 - conifer	313 - mixed woodl	321 - nat grass	324 - trans woodl	332 - bare rock	512 - wtrbody	cover / 1000	% auto corr.
111 - cont urb	**2.8**	1.4							0.0	0.3										4.5	**61**
112 - discont	0.9	**28.9**		0.2	0.0	0.0		0.0	9.2	1.6			0.2						0.0	41.0	**71**
121 - indust		0.0	**1.5**						0.0	0.0										1.5	
124 - airport		0.1		**9.0**					1.2	0.2										10.5	**86**
131 - mineral	0.0	0.0			**4.9**				0.2	0.1									0.1	5.2	
133 - constr		0.0				**1.3**			0.2	0.0										1.5	
141 - green urb		0.0							0.3	1.5										1.9	
142 - recreat		0.7						**3.2**	1.0	0.7										5.5	**58**
211 - arable		9.2		0.3	0.0	0.0		0.3	**717.0**	17.3			3.3	0.2		0.3			1.6	749.5	**96**
231 - pasture	0.2	11.5		0.0	0.2	0.2		0.8	39.2	**91.1**			0.6			0.0			1.1	144.8	**63**
242 - mixed cult	0.1	1.0		0.2	0.1				7.2	1.2			0.1	0.0		0.2			0.4	10.5	
243 - agric/nat									0.3	0.1			0.0						0.2	0.5	
311 - broadlf		0.1							1.1	0.0			**5.5**	0.2						6.9	**80**
312 - conifer									0.3	0.1			0.4	**0.4**						1.2	**32**
313 - mixed woodl									0.1				0.3	0.2						0.6	**32**
321 - nat grass		0.2							2.4	0.1						**1.4**			0.0	4.1	**34**
324 - trans woodl									0.1	0.0			0.4							0.5	
332 - bare rock				0.3					0.5											0.8	
512 - wtrbody		0.1			0.0				0.6	0.4			0.0						**7.8**	8.9	**88**
cover / 1000	3.9	53.3	1.5	10.1	5.2	1.5		4.3	780.9	114.6			10.7	1.0		2.0			11.1	1000.0	
% manual corr.	**71**	**54**	**1**	**1**	**1**	**87**		**74**	**92**	**79**			**51**	**41**		**71**			**70**		
After interactive edit	61	64	1	92	92	68		75	92	76			52	41		71			81		

Total pixels interpreted= 1224000000

Total matching = 874.7 pixels / 1000

Total matching = 881.0 pixels / 1000

Table 4.Correspondence per 1000 pixels, between manual and automated CORINE interpretations(validation for pastural in Cornwall)

AUTOMATED	MANUAL																			
Corine Value	112 - discont	121 - indust.	123 - port areas	124 - airports	131 - mineral	142 - recreat.	211 - arable	231 - pasture	242 - cultivate	243 - agric/nat	311 - broadlf	312 - conifer	313 - mixed forest	321 - nat grass	322 - moors/heath	332 - bare rock	423 - intert. flats	522 - estuaries	523 - sea/ocean	cover / 1000	% auto corr.
112 - discont	26.5	0.3	0.0	0.3	1.1	0.2	0.3	8.9		0.4	0.1			0.2			0.2		0.1	38.6	69
121 - indust	0.2	1.2						0.0									0.0			1.4	83
123 - port areas	0.0		0.4			0.0											0.0			0.4	88
124 - airports	0.2			2.4				0.4			0.0									2.9	80
131 - mineral	0.8	0.1			22.3			1.7		0.2	0.0			0.1						25.2	89
142 - recreat	0.4		0.0			1.5	0.0	3.1		0.6	0.0						0.1		0.0	5.8	26
211 - arable	0.0			0.0		0.4	11.0	12.6		0.3	0.4			0.1			0.1		0.0	25.0	44
231 - pasture	4.9	0.1	0.1	0.5	1.0	0.6	8.1	582.0		1.4	14.2	0.1	0.3	1.7			1.7	0.0	0.3	617.0	94
242 - cultivate	0.4			0.5	0.0	0.0	2.0	16.0	[illegible]	0.4	0.3						0.3		0.1	20.0	
243 - agric/nat	0.0							2.7		[illegible]	0.4			0.0			0.1	0.0	0.0	3.2	
311 - broadlf	0.2			0.0	0.1		0.3	20.8		0.0	24.2	0.2	1.0	0.6			0.4	0.0		47.8	51
312 - conifer								0.2			0.1	[illegible]	0.3							1.4	49
313 - mixed forest													[illegible]								
321 - nat grass	0.1	0.1			0.3		0.2	6.5			0.1			5.7						12.9	44
322 - moors/heath	0.0							0.1						0.2	[illegible]					0.4	
332 - bare rock		0.6			0.7		0.4	0.8			0.1			0.1		[illegible]				2.7	
423 - intert. flats	0.7		0.0			0.5	0.0	3.6			0.0						9.8	0.1	0.0	14.8	66
522 - estuaries								0.0									0.4	0.6		1.1	58
523 - sea/ocean	0.1		0.0			0.1	0.3	1.6		0.0							2.3	0.0	174.9	179.3	98
cover / 1000	34.4	2.3	0.5	3.6	25.5	3.3	22.6	660.9		3.4	40.2	1.0	1.7	8.6			15.7	0.9	175.4	1000.0	
% manual corr.	77	52	69	65	87	45	49	88			60	71		66			63	71	100		

Total pixels interpreted = 1201635000 Total matching = 863 pixels / 1000

Table 5.Correspondence per 1000 pixels, between manual and automated CORINE interpretations (validation for Scotand marginal)

MANUAL

AUTOMATED

CORINE class	112 - cont urb	142 - recreat	231 - pasture	311 - broadlf	312 - conifer	321 - nat grass	322 - moors/heath	324 - trans wood	331 - beaches	332 - bare rock	412 - peat bogs	512 - wtrbody	522 - estuaries	523 - sea/ocean	cover / 1000	% auto corr.
112 - cont urb	0.9		0.3			0.4			0.1						1.8	52
142 - recreat		[illegible]														
231 - pasture	0.1	0.2	33.6			7.6	3.7				0.1	0.1		0.6	46.0	73
311 - broadlf			0.1	[illegible]		0.1	0.4				0.3				0.9	
312 - conifer	0.0	0.1	1.7		223.7	5.2	4.4		0.0			0.5		0.3	236.0	95
321 - nat grass	0.1	0.0	6.0			168.9	44.5		0.1		0.0	0.3		0.7	220.7	77
322 - moors/heath	0.1		9.3			57.2	248.7		1.2		1.5	0.8		0.9	319.8	78
324 - trans wood			0.1			0.2	2.5	29.1						0.0	31.9	91
331 - beaches	0.5		0.0			0.1			1.3					0.3	2.2	58
332 - bare rock			0.0			0.2				[illegible]					0.3	
412 - peat bogs						1.2	1.3				0.0	0.0			2.5	1
512 - wtrbody			0.8			1.3	1.0					29.3			32.4	91
522 - estuaries													[illegible]			
523 - sea/ocean	0.1		0.4			3.0	1.1		0.5			0.5		99.9	105.5	95
cover / 1000	1.9	0.2	52.3		223.7	245.6	307.6	29.1	3.3		2.0	31.6		102.7	1000.0	
% manual corr.	**50**		**64**		**100**	**69**	**81**	**100**	**39**		**1**			**97**		

Total pixels interpreted = 1200884375 **Total matching = 836 pixels / 1000**

Table 6.Correspondence per 1000 pixels, between manual and automated CORINE interpretations (validation for Wales marginal) note improved correspondence in post edit figures (see in red)

MANUAL

AUTOMATED

Corine Value	111 - cont urb	112 - discont	121 - indust.	122 - road and rail	123 - port areas	131 - mineral	132 - dump sites	142 - recreat.	211 - arable	231 - pasture	242 - cultivate	243 - agric/nat	311 - broadlf	312 - conifer	313 - mixed forest	321 - nat grass	322 - moors/heath	332 - bare rock	412 - peat bogs	421 - salt marshes	512 - water bodies	522 - estuaries	cover / 1000	% auto corr.
111 - cont urb	**3.7**	2.1	0.7							0.2			0.0			0.1						0.2	6.9	**53**
112 - discont	1.8	**78.0**	3.7	0.3	0.0	1.1	0.0	0.2	0.4	12.1	0.0	0.7	1.4	0.2	0.2	2.1	1.2	0.0			0.0	0.0	103.3	**76**
121 - indust		2.1	**4.4**						0.2	0.5		0.1	0.0			0.1	0.2					0.1	7.7	**57**
122 - road/rail			0.1	**0.0**						0.1													0.2	**4**
123 - port areas	0.0	0.0			**1.7**					0.0												0.1	1.9	**93**
131 - mineral		0.2	0.0			**2.3**				0.5						0.1	0.1						3.0	**74**
132 - dump sites							**0.3**			0.0													0.3	**98**
142 - recreat		0.1	0.0					**3.1**	0.1	4.1			0.1	0.0		0.4	0.3	0.1					8.3	**38**
211 - arable		1.0	0.0			0.4			**8.7**	5.5	0.3	0.1	0.1	0.0	0.0	0.5						0.0	16.8	**52**
231 - pasture	0.2	12.4	1.0	0.0	0.0	0.9	0.0	0.7	4.0	**347.2**	0.3	3.2	8.7	2.1	2.5	14.7	10.5	0.2			0.1	0.3	409.0	**85**
242 - cultivate	0.2	5.5	0.6		0.4	0.2		0.0	2.0	7.0		0.8	0.4	0.4	0.2	0.8	0.0	0.3				0.2	18.9	**0**
243 - agric/nat		0.3					0.0			0.9			0.3	0.0		0.1							1.6	
311 - broadlf	0.0	3.9	0.1	0.3		0.3	0.3	0.1	0.9	40.9	0.1	1.0	**33.2**	14.5	11.4	5.2	1.4				0.1		113.8	**29**
312 - conifer		0.1				0.0			0.2	2.5		0.0	1.8	**11.4**	2.1	0.8	0.4				0.2		19.4	**59**
313 - mixed forest		0.1						0.0	0.1	1.9			1.4	1.2	**1.9**	0.4	0.1						7.1	**26**
321 - nat grass	0.1	2.3	0.3			1.4		0.7	0.1	49.3	0.2	0.3	6.8	3.2	2.9	**76.5**	9.6	0.6			0.2		154.3	**50**
322 - moors/heath	0.1	2.6	0.5	0.1		1.7		0.2	0.1	17.7		0.7	5.5	4.6	1.1	10.6	**70.0**				0.1		115.4	**61**
332 - bare rock		0.8	0.0			1.3				0.4			0.0	0.1		0.1	0.2						3.0	7
412 - peat bogs										0.0						0.2							0.2	
421 - salt marshes			0.0						0:0	0.0												0.2	0.3	
512 - water bodies	0.0	0.1	0.0		0.4					0.1			0.0	0.2			0.0				**3.1**	0.3	4.2	**1**
522 - estuaries	0.0		0.1		0.1				0.0	0.0												**4.2**	4.4	**95**
cover / 1000	6.0	111.6	11.7	0.6	2.6	9.5	0.7	4.9	16.7	490.9	0.9	6.9	59.8	37.9	22.2	112.5	93.9	1.3			3.7	5.6	1000.0	
% manual corr.	**61**	**70**	**37**	**2**	**67**	**24**	**41**	**63**	**52**	**71**		**15**	**56**	**30**	**8**	**68**	**74**				**4**	**76**		
After final interactive edit	53	75	55	61	93	74	98	38	59	85			29	59	26	50	61				79	83		

Total pixels interpreted = 1229935625

Total matching = 650 pixels / 1000

Total matching = 780 pixels / 1000

Table 7.Correspondence per 1000 pixels, between manual and automated CORINE interpretations (validation for Uplands)

MANUAL (columns) / **AUTOMATED** (rows)

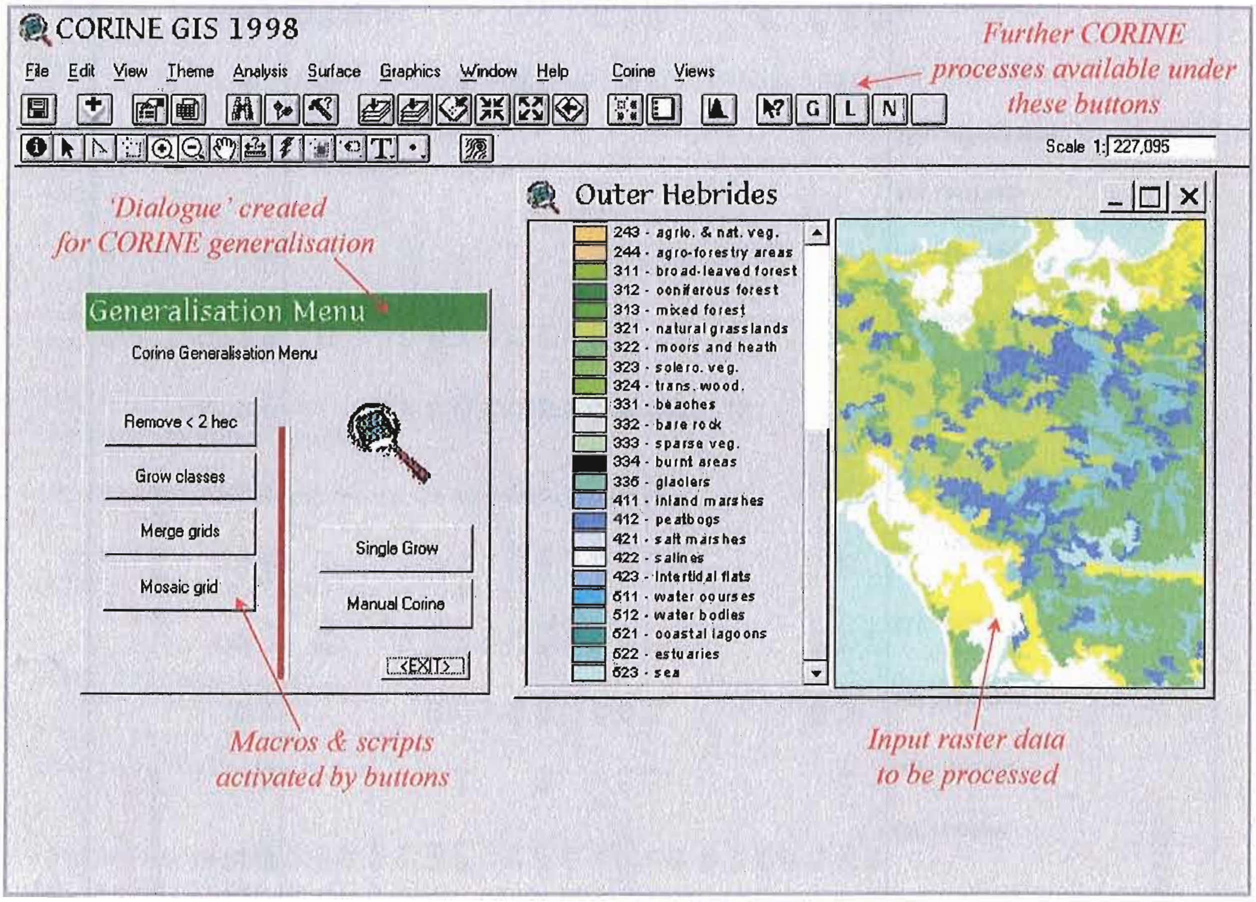
CORINE class	142 - recreat.	231 - pasture	242 - complex Cult	243 - agri/nat	311 - broadlf	312 - conifer	321 - nat grass	322 - moors/heath	324 - trans wood	332 - bare rock	412 - peat bogs	512 - wtrbody	cover / 1000	% auto corr.
142 - recreat.		0.0					0.0	0.9		0.0			0.9	
231 - pasture		6.8					5.5	1.0				0.0	13.3	
242 - complex Cult							0.0						0.0	
243 - agri/nat							0.0						0.0	
311 - broadlf							0.3	0.0					0.3	
312 - conifer		0.2				69.1	1.1	3.1				0.2	73.6	94
321 - nat grass		0.3					27.6	14.0		0.8		0.0	42.7	65
322 - moors/heath		1.2					32.5	783.5		12.3		1.1	830.4	94
324 - trans wood		0.1					0.7	0.2					1.0	
332 - bare rock							0.5	9.8		20.5		0.2	30.9	
412 - peat bogs								2.5		0.0			2.5	
512 - wtrbody		0.0					0.1	0.6		0.0		3.6	4.4	82
cover / 1000		8.6				69.1	68.4	815.3		33.6		5.1	1000.0	
% manual corr.		80				100	40	96		61		71		

Total pixels interpreted = 1200884376

Total matching = 911 pixels / 1000

7.6 ARC/View customized interface

Figure 3. An example screen shot from an ARC/View generalisation process



7.7 Avenue scripts

This is a selection of the major Avenue scripts created for the CORINE generalisation process. Some are completely new scripts offering unique functionality. Some scripts created linked together processes already available in the ARC/View 'Spatial Analyst Extension'. Our scripts were embedded into the customised interface shown in Annex Figure 3. We have reproduced here the most significant new scripts and identified what standard 'Spatial Analyst' functionality has been used. A number of minor scripts were created which were used alongside these at various points in the processes, but these have not been included in this annex.

The code for these scripts has not been edited for inclusion in this Annex. These are straight copies from the avenue script.

7.7.1 Removing parcels < 2ha

Script name - 25hec~z0.ave

```
av.getProject.SetWorkDir("/users/seol/dse/datadirs/datavla/tempdir".AsFilenam
e)
'Title:     Get View and Active Grid Theme
```

```
theView = av.GetActiveDoc
'theDisplay = theView.GetDisplay
theGridTheme = theView.GetActiveThemes.Get(0)
theOriginal = theGridTheme.GetGrid

'-----------------------------------------------------------------------

'Title:     RegionGroup

theRGroup = theoriginal.RegionGroup(TRUE, FALSE, Nil)

if (theRGroup.HasError) then
return NIL
end

'make the grid

name = "Regiongrid"
pre = "regiong"

aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
theRgroup.Rename(aFN)

if (theRgroup.HasError) then return NIL end

' create a theme
gthm = GTheme.Make(theRgroup)

' set name for theme
gthm.SetName(name)

' add theme to the View
theView.AddTheme(gthm)


regionTheme=gthm
if (regionTheme = NIL) then exit end
if (regiontheme.Is(GTHEME)) then
    regiontheme = regiontheme.GetGrid
else
exit
end


thergrid = regiontheme


'-----------------------------------------------------------------------



Tg1 = thergrid.test("([count] < 400)")

if (tg1.HasError) then
return NIL
end

'make the grid

name = "testgrid1"
pre = "tg1"

aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
tg1.Rename(aFN)
```

```
if (tgl.HasError) then return NIL end

' create a theme
gthm = GTheme.Make(tgl)

' set name for theme
gthm.SetName(name)

' add theme to the View
theView.AddTheme(gthm)


testTheme=gthm
if (testTheme = NIL) then exit end
if (testtheme.Is(GTHEME)) then
    testtheme = testtheme.GetGrid
else
exit
end


thetestgrd = testtheme


'-------------------------------------------------------------------

'thetestGrd2 = (thetestgrd = 1.AsGrid).SetNull(theoriginal)
newGrid = (theTestgrd = 1).Con(999.AsGrid,theOriginal)

if (newgrid.HasError) then
return NIL
end

'make the grid

name = "newgrd2"
pre = "newg2"

aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
newgrid.Rename(aFN)

if (newgrid.HasError) then return NIL end

' create a theme
gthm = GTheme.Make(newgrid)

' set name for theme
gthm.SetName(name)

' add theme to the View
theView.AddTheme(gthm)


newTheme=gthm
if (newTheme = NIL) then exit end
if (newtheme.Is(GTHEME)) then
    newgtheme = newtheme.GetGrid
else
exit
end


newg2 = newgtheme
```

```
'-------------------------------------------------------------------

thetestGrd2 =  (newg2 = 999.AsGrid).SetNull(newg2)

if (thetestgrd2.HasError) then
return NIL
end

'make the grid

name = "thetestgrd2"
pre = "thetestg2"

aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
thetestgrd2.Rename(aFN)

if (thetestgrd2.HasError) then return NIL end

' create a theme
gthm = GTheme.Make(thetestgrd2)

' set name for theme
gthm.SetName(name)

' add theme to the View
theView.AddTheme(gthm)


testgrd2Theme=gthm
if (testgrd2Theme = NIL) then exit end
if (testgrd2theme.Is(GTHEME)) then
    test2theme = testgrd2theme.GetGrid
else
exit
end


thetgrid2 = test2theme


'-------------------------------------------------------------------


nibblegrid = theoriginal.nibble(thetgrid2, TRUE)

if (nibblegrid.HasError) then
return NIL
end

'make the grid

name = "nibblegrid"
pre = "nibbleg"

aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
nibblegrid.Rename(aFN)

if (nibblegrid.HasError) then return NIL end

' create a theme
gthm = GTheme.Make(nibblegrid)

' set name for theme
```

```
gthm.SetName(name)

' add theme to the View
theView.AddTheme(gthm)
```

7.7.2 Growing & merging grids (3 scripts)

<u>Script name</u> - andrew.ave

```
'START OF GROWING PROCESS


'-----------------------
'source to cost
'-----------------------
'av.getProject.SetWorkDir("/users/seol/dse/datadirs/datav1a/tempdir".AsFilena
me)

theView = av.GetActiveDoc
thePrj = theView.GetProjection
if (thePrj.IsNull) then
  hasPrj = false
else
  hasPrj = true
  thePrj = theView.GetProjection
end

theThemes = theView.GetactiveThemes

if ((thethemes.count < 1) or (thethemes.count > 1))  then
  MsgBox.Error("Please make only one grid theme active","*** Error ***")
  return nil
end

' display the source name of the grid for each grid theme
for each t in theView.GetActiveThemes

' obtain cellsize
if (t.Is(GTheme)) then
  cellSize = t.GetGrid.GetCellSize
else

  ' obtain extent and cell size if not set
  ae = theView.GetExtension(AnalysisEnvironment)
  box = Rect.Make(0@0,1@1)
  cellSize = 1
  if ((ae.GetExtent(box) <> #ANALYSISENV_VALUE) or (ae.GetCellSize(cellSize)
<> #ANALYSISENV_VALUE)) then
    ce = AnalysisPropertiesDialog.Show(theView, TRUE, "Output Grid
Specification")
    if (ce = NIL) then return NIL end
    ce.GetCellSize(cellSize)
    ce.GetExtent(box)
  end
end

' obtain type of statistic, neighborhood,

staType =  #GRID_STATYPE_MAX
theNbr = NbrHood.Make

' run function
if (t.Is(GTHEME)) then
  g = t.GetGrid
```

```
  r = g.FocalStats(staType, theNbr, FALSE)
  r = r.int
else
exit
end

'make the grid

name = "NbrSum"
pre = "nbhdsum"

aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
r.Rename(aFN)

if (r.HasError) then return NIL end

' create a theme
gthm = GTheme.Make(r)

' set name for theme
gthm.SetName(name ++ "of class" ++ g.GetName)

' add theme to the View
theView.AddTheme(gthm)

end
```

Script name - mergegrids.ave

```
av.getProject.SetWorkDir("/users/seol/dse/datadirs/datavla/tempdir".AsFilenam
e)

theView = av.GetActiveDoc
thePrj = theView.GetProjection
if (thePrj.IsNull) then
  hasPrj = false
else
  hasPrj = true
  thePrj = theView.GetProjection
end

theThemes = theView.GetThemes

if (thethemes.count < 1)  then
  MsgBox.Error("Please make than one grid theme active","*** Error ***")
  exit
end

thegridlist = list.make
count = 0
' display the source name of the grid for each grid theme
for each t in theView.GetThemes

  if (t.Is(GTHEME)) then
    ' get grid and display source name
    count = count + 1
    if(count = 1) then
    thefirstgrid = t.getgrid
    thefirstgrid = thefirstgrid.int
    else
    theGrid = t.GetGrid
    thegrid = thegrid.int
    thegridlist.add(thegrid)
    end
  else
    return nil
```

```
  end
end

mergeFN = av.getProject.GetWorkDir.MakeTmp("merge","")
'allocationFN = av.getProject.GetWorkDir.MakeTmp("allocation","")

mergeFN = thefirstgrid.Merge(theGridList)

'mergeFN = mergeFN.int

if (mergeFN.HasError) then
return NIL
end

theGtheme = Gtheme.Make(mergeFN)
theGtheme.SetName("merge")
theview.AddTheme(theGtheme)
theGtheme.SetActive(true)
theview.getwin.activate
av.GetProject.SetModified(True)
'---------------------------
'msgbox.info(""+count.AsString+"","")
```

<u>Script name - Dissolve</u>

Description: Merges adjacent polygons which have the same value for a specified field in the attributes table. This is required where merged grids have created two polygons adjacent to each other with the same CORINE class.

This script was available via the ESRI software Web site and was written by:

Kenneth R. McVay
Environmental Geologist
Research Asst. Geology
University of Houston Clear Lake

7.7.3 Mosaic, stages 1, 2 & 3

A selection of standard ESRI ARC/View software were utilised for these processes. The main functions are listed below. These are invoked via the customized CORINE menu buttons within ARC/View (See Annex Figure 1).

```
Name: Spatial Tools Extension.
Name:      Spatial.MajorityFilter
Name: Spatial.ExtractByCount
Name: Spatial.Nibble
Name: RegionGroup
Name: ExtractByCount.
```

7.7.4 Land cover code conversion (2 scripts)

<u>Script name - describe.ave</u>

```
'Title:     Get View and Active Grid Theme
□

□
theView = av.GetActiveDoc
□
```

```
thePrj = theView.GetProjection
□
'theDisplay = theView.GetDisplay
□
theGridTheme = theView.GetActiveThemes.Get(0)
ingrid = theGridTheme.GetGrid
ingrid = ingrid.int

newfinalg = ingrid.isnull.Con(0.asgrid,

(inGrid = 0.AsGrid).Con(0.asgrid,
(inGrid = 1.AsGrid).Con(1.asgrid,
(inGrid = 2.AsGrid).Con(2.asgrid,
(inGrid = 3.AsGrid).Con(3.asgrid,
(inGrid = 4.AsGrid).Con(4.asgrid,
(inGrid = 7.AsGrid).Con(7.asgrid,
(inGrid = 7.AsGrid).Con(7.asgrid,
(inGrid = 7.AsGrid).Con(7.asgrid,
(inGrid = 8.AsGrid).Con(7.asgrid,
(inGrid = 9.AsGrid).Con(7.asgrid,
(inGrid = 10.AsGrid).Con(10.asgrid,
(inGrid = 11.AsGrid).Con(10.asgrid,
(inGrid = 12.AsGrid).Con(7.asgrid,
(inGrid = 13.AsGrid).Con(10.asgrid,
(inGrid = 14.AsGrid).Con(14.asgrid,
(inGrid = 17.AsGrid).Con(17.asgrid,
(inGrid = 17.AsGrid).Con(17.asgrid,
(inGrid = 17.AsGrid).Con(17.asgrid,
(inGrid = 18.AsGrid).Con(18.asgrid,
(inGrid = 19.AsGrid).Con(7.asgrid,
(inGrid = 20.AsGrid).Con(20.asgrid,
(inGrid = 21.AsGrid).Con(21.asgrid,
(inGrid = 22.AsGrid).Con(22.asgrid,
(inGrid = 23.AsGrid).Con(14.asgrid,
(inGrid = 24.AsGrid).Con(17.asgrid,
10.Asgrid)))))))))))))))))))))))))' ie class 25

if (newfinalg.HasError) then

msgbox.info("error","")
return NIL
end
```

Script name - landtonew.ave

```
name = "nwlandcov"
pre = "newland"
aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
newfinalg.Rename(aFN)
if (newfinalg.HasError) then return NIL end

newfina = GTheme.Make(newfinalg)

' set name for theme
newfina.SetName(name)

' add theme to the View
theView.AddTheme(newfina)
landcoverlist = Dictionary.Make( 100 )
□

□
landcoverlist.set (0, "none")
□
```

```
landcoverlist.set (1, "Sea")
□
landcoverlist.set (2 , "Inland Water")
□
landcoverlist.set (3 , "Beach + Coastal Bare")
□
landcoverlist.set (4 , "Saltmarsh")
landcoverlist.set (5 , "Grass Heath, Rough / Marsh Grass,"+nl+"Moorland
Grass, Bracken, Ruderal Weed")
landcoverlist.set (6 , "Mown / grazed turf"+nl+"Meadow / verge / semi-
natural")
landcoverlist.set (10 , "Open Shrub Moor, Dense Shrub Moor"+nl+"Dense Shrub
Heath, Open shrub heath")
landcoverlist.set (14 , "Shrub / Orchard"+nl+"Felled Forest")
landcoverlist.set (15 , "Deciduous Woodland")
landcoverlist.set (16 , "Coniferous Woodland")
landcoverlist.set (17 , "Upland bog"+nl+"Lowland bog")
landcoverlist.set (18 , "Tilled Land")
landcoverlist.set (20 , "Suburban / Rural Development")
landcoverlist.set (21 , "Continuous Urban")
landcoverlist.set (22 , "Inland Bare Ground")
landcoverlist.set (25 , "Sport and Leisure")
landcoverlist.set (26 , "Airport")
landcoverlist.set (27 , "Green urban areas")
landcoverlist.set (28 , "Industrial")
landcoverlist.set (29 , "Mineral")
landcoverlist.set (30 , "Construction")
landcoverlist.set (31 , "Road, rail and associated land")
landcoverlist.set (32 , "Port areas")
landcoverlist.set (33 , "Dump sites")
landcoverlist.set (34 , "Fruit trees and berry plantations")
landcoverlist.set (35 , "Estuary")
landcoverlist.set (36 , "Coastal Lagoon")
landcoverlist.set (37 , "Water courses")
landcoverlist.set (38 , "Inland marshes")
landcoverlist.set (39 , "Intertidal")



corinecoverlist = Dictionary.Make( 100 )

corinecoverlist.set (111 , "Continuous urban fabric")
corinecoverlist.set (112 , "Discontinuous urban fabric")
corinecoverlist.set (121 , "Industrial or commercial units")
corinecoverlist.set (122 , "Road and Rail networks + land")
corinecoverlist.set (123 , "Port areas")
corinecoverlist.set (124 , "Airports")
corinecoverlist.set (131 , "Mineral extraction sites")
corinecoverlist.set (132 , "Dump sites")
corinecoverlist.set (133 , "Construction sites")
corinecoverlist.set (141 , "Green urban areas")
corinecoverlist.set (142 , "Sport and leisure facilities")
corinecoverlist.set (211 , "Non-irrigated arable land")
corinecoverlist.set (212 , "Permanently irrigated land")
corinecoverlist.set (213 , "Rice Fields")
corinecoverlist.set (221 , "Vineyards")
corinecoverlist.set (222 , "Fruit trees and berry plantations")
corinecoverlist.set (223 , "Olive groves")
corinecoverlist.set (231 , "Pastures")
corinecoverlist.set (241 , "Annual crops with permanent crops")
corinecoverlist.set (242 , "Complex cultivation patterns")
corinecoverlist.set (243 , "Agricutural Land with natural vegetation")
corinecoverlist.set (244 , "Agro-forestry areas")
corinecoverlist.set (311 , "Broad-leaved forest")
```

```
corinecoverlist.set (312 , "Coniferous forest")
corinecoverlist.set (313 , "Mixed forest")
corinecoverlist.set (321 , "Natural grasslands")
corinecoverlist.set (322 , "Moors and heathland")
corinecoverlist.set (323 , "Sclerophyllous vegetation")
corinecoverlist.set (324 , "Transitional woodland-scrub")
corinecoverlist.set (331 , "Beaches, dunes, sands")
corinecoverlist.set (332 , "Bare rocks")
corinecoverlist.set (333 , "Sparsely vegetated areas")
corinecoverlist.set (334 , "Burnt areas")
corinecoverlist.set (335 , "Glaciers and perpetual snow")
corinecoverlist.set (411 , "Inland marshes")
corinecoverlist.set (412 , "Peat bogs")
corinecoverlist.set (421 , "Salt marshes")
corinecoverlist.set (422 , "Salines")
corinecoverlist.set (423 , "Intertidal flats")
corinecoverlist.set (511 , "Water courses")
corinecoverlist.set (512 , "Water bodies")
corinecoverlist.set (521 , "Coastal lagoons")
corinecoverlist.set (522 , "Estuaries")
corinecoverlist.set (523 , "Sea and ocean")

'-------------------------------------------------------------------------------
-----------------------

landcover = msgbox.miniyesno ("CORINE value or LCMGB value? - CORINE
Default", true )

if (landcover = true) then

   thecorinevalue = 0
   actualcorine = 0
   while (((thecorinevalue <= 110) or (thecorinevalue >=524)) or
(thecorinevalue.AsString.IsNumber).Not)
   thecorinevalue = MsgBox.Input("CORINE class value","CORINE value
description",theCorinevalue.AsString)

   if (thecorinevalue = NIL) then return nil end

   if (thecorinevalue.IsNumber) then
     thecorinevalue = thecorinevalue.AsNumber
     thecorinelist = list.make

     if ((corinecoverlist.get(thecorinevalue))<>nil) then
     thecorinelist.add(corinecoverlist.get(thecorinevalue))
     'MsgBox.ListAsString( thecorinelist, "Landcover description", "CORINE" )
     msgbox.info("Landcover
Description:"+nl+""+corinecoverlist.get(thecorinevalue)+"","CORINE")
     end

   end
   end

end

if (landcover = false) then

   thecorinevalue = 0
   actualcorine = 0
   while (((thecorinevalue <= 0) or (thecorinevalue >=40)) or
(thecorinevalue.AsString.IsNumber).Not)
   thecorinevalue = MsgBox.Input("Landcover class value","LAndcover value
description",theCorinevalue.AsString)

   if (thecorinevalue = NIL) then return nil end
```

```
    if (thecorinevalue.IsNumber) then
      thecorinevalue = thecorinevalue.AsNumber
      thecorinelist = list.make
      if ((landcoverlist.get(thecorinevalue))<>nil) then
      thecorinelist.add(landcoverlist.get(thecorinevalue))
      'MsgBox.ListAsString( thecorinelist, "Landcover description", "LCMGB" )
      msgbox.info("Landcover
Description:"+nl+""+landcoverlist.get(thecorinevalue)+"","LCMGB")
      end
    end
    end

end
```

7.7.5 Nibble & regiongroup (mainly Spatial Analyst tools)

A standard ESRI ARC/View software were utilised for these processes. The main functionality is listed below. These are invoked via the customized CORINE menu buttons within ARC/View (See Annex Figure 1).

```
Spatial.Thin"
        SourceCode: "'Name: BE.Thin\n'\n'Title: Thin\n'\n'Topic:
Analysis\n'\n'Description: Prompts user for Thin request variables then runs
Thin on\n'the active theme.  The name of the active grid theme and the
variables used are \n'saved in the Comments section of the grid theme created
by running this script.\n'\n'Requires: Spatial Analyst; an active view; an
active grid theme; run from menu\n'or button\n'\n'Created January 16, 1997;
by Bill Eichenlaub\n\ntheView = av.GetActiveDoc\ntheDisplay
av.getProject.SetWorkDir("/users/seol/dse/datadirs/datav1a/tempdir".AsFilenam
e)
'Title:      Get View and Active Grid Theme

theView = av.GetActiveDoc
'theDisplay = theView.GetDisplay
theGridTheme = theView.GetActiveThemes.Get(0)
theOriginal = theGridTheme.GetGrid

'-----------------------------------------------------------------------

'Title:      RegionGroup

theRGroup = theoriginal.int
if (theRGroup.HasError) then
return NIL
end

'make the grid

name = "Reggrid"
pre = "regiong"

aFN = av.GetProject.GetWorkDir.MakeTmp(pre,"")
theRgroup.Rename(aFN)

if (theRgroup.HasError) then return NIL end

' create a theme
gthm = GTheme.Make(theRgroup)

' set name for theme
gthm.SetName(name)
```

```
' add theme to the View
theView.AddTheme(gthm)
```

7.7.6 Removing isolations

<u>Script name - singleton.ave</u>

```
'Title:     Get View and Active Grid Theme
□

□
theView = av.GetActiveDoc
□
theDisplay = theView.GetDisplay
□
theGridTheme = theView.GetActiveThemes.Get(0)
theGrid = theGridTheme.GetGrid
theOriginal = theGridTheme.GetGrid

'-----------------------------------------------------------------------

'Title:     RegionGroup

theRGroup = theGrid.RegionGroup(FALSE, FALSE, Nil)

'-----------------------------------------------------------------------

'Title: Majority filter

theRGrid = theRGroup

theFiltered = theRGrid.MajorityFilter (TRUE, FALSE)

'-----------------------------------------------------------------------

'Title: Test for single pixels

theSingle = theoriginal.MajorityFilter (TRUE, FALSE)

theTestgrd = thergrid.test("([count] = 1)")

newGrid = (theTestgrd = 1).Con(theSingle,theOriginal)

theGTheme = GTheme.Make(newGrid)

if (newgrid.HasError) then
return NIL
end

theView.AddTheme(theGTheme)

theNewT = theView.GetThemes.Get(0)
av.GetProject.SetModified(True)
```

7.7.7 Smoothing

A significant attempt was made to use standard ARC generalisation functionality to give an optimum smoothing of the vectorised raster data for display at the 100 000 scale. This included a test of a range of parameters such as weed tolerances, grain size etc. However the best result was that achieved during the standard ARC/View 'theme - convert to shape file' function. The script below is an example of one that was developed but later abandoned.

<u>Script name - natgrid.ave</u>

```
' Name:  smoothing
'
' Title:  Smoothes selected features
'
' Get the view and its projection

theView = av.GetActiveDoc
thePrj = theView.GetProjection
if (thePrj.IsNull) then
  hasPrj = false
else
  hasPrj = true
  thePrj = theView.GetProjection
end

  theClassName = "polygon"

  ' Get the name for a new theme, from View.Export script

  def = av.GetProject.MakeFileName("theme", "shp")
  newtheme = "natgrid"
  def = FileDialog.Put(def, "*.shp", "New file for simplified "+ newtheme)

  if (def = NIL) then return nil end

  ' Create a new shapefile theme, from View.Export script

  anFTab = ftab.MakeNew (def, polygon)
  anftab.AddFields({Field.Make("id", #FIELD_LONG, 12, 0)})
  anFTab.SetEditable(true)

  tbl = anftab
  shpfld = (tbl.FindField("Shape"))
  if (shpfld.IsVisible.Not) then
    shpfld.SetVisible(shpfld.IsVisible.Not)
    WasNotVisible = TRUE
  else
    WasNotVisible =FALSE
  end

  if (WasNotVisible) then
    shpfld.SetVisible(FALSE)
  end

  ' Find the shape field

  theSField = anFTab.FindField("Shape")

  ' Set up status bar

  av.ShowMsg("Creating new shapes for theme"++ newtheme)

  ' Initialize counters for reporting

 xcount = 0
 count = 0

while (xcount < 7)
 av.SetStatus(100*(xcount/7))
 ycount = 0
 while(ycount < 13)
  r = anftab.addrecord

  ' Process each shape
```

```
    if (theClassName = "Polygon") then
      theOShape = Polygon.MakeNull
    else
      theOShape = Polyline.MakeNull
    end
    anFTab.QueryShape(r,theView.GetProjection,theOShape)

    ' Make a list for collecting the new set of point lists

    theNShape = List.Make

    ' Make a list for collecting the vertices to keep

      theNList = List.Make
      x = xcount * 100000
      y = ycount * 100000
      ULxy = point.make(x,(y + 100000))
      LLxy = point.make(x,y)
      LRxy =point.make((x+100000),y)
      URxy =point.make((x+100000),(y+100000))

        if (hasPrj) then theNList.Add(ULxy).ReturnUnProjected(thePrj) else
theNList.Add(ULxy) end
        if (hasPrj) then theNList.Add(LLxy).ReturnUnProjected(thePrj) else
theNList.Add(LLxy) end
        if (hasPrj) then theNList.Add(LRxy).ReturnUnProjected(thePrj) else
theNList.Add(LRxy) end
        if (hasPrj) then theNList.Add(URxy).ReturnUnProjected(thePrj) else
theNList.Add(URxy) end

      ' Finish up individual lists

      theNShape.Add(theNList)

    ' Finish up the shape

    if (theClassName = "Polygon") then
      theNShape = Polygon.Make(theNShape)
    else
      theNShape = Polyline.Make(theNShape)
    end
    anFTab.SetValue(theSField, r, theNShape)
    idfield = anFTab.FindField("id")
    anftab.SetValue(idfield, r, count)
    count = count + 1

  Ycount = Ycount + 1
 end

 Xcount = Xcount + 1
end

  ' Clear status message

  av.ClearMsg
  av.ClearStatus

  ' Report on counts of old and new theme

  MsgBox.Info("National grid created","grid created!")

  ' Create a theme and add it to the View

  anFTab.SetEditable(false)
```

```
  fthm = FTheme.Make(anFTab)
  theView.AddTheme(fthm)

  fthm.SetActive(true)

' Bring the View to the front

theView.GetWin.Activate
```