ECOBASE OF BRITAIN: status report on a digital data base of Britain

David P. Bickmore
Royal College of Art, London, SW7
and
Dr. Graham J. Witty
NERC Experimental Cartography Unit
London, SW7

1. The Experimental Cartography Unit (ECU) is part of the British Natural Environment Research Council - an official body which embraces a number of internationally known institutes concerned with geology, hydrology, oceanography, ecology, etc. This particular project was initiated in 1976 by the Institute of Terrestrial Ecology (ITE) who required ECU to digitise the 1:250,000 Ordnance Survey map of the country. In doing so, they stated:

2. "Basically the requirement is a simple one. We expect increasingly to be concerned in ITE in the characterisation of the ecological resources of Britain for a variety of customers. Much of this characterisation will be on a national or regional scale rather than a purely local scale, and we will need to present the results of the surveys and analyses of existing information in a cartographic format which is both convincing and flexible. Furthermore, it is almost impossible to predict the regions for which characterisation will be required. We believe, therefore, that first priority should be given for a national coverage at 1:250,000 scale with about as much detail as is currently given in OS maps at that scale.

Certainly this complete coverage should be available within two to three years if we are to meet the requirements of our customers."

"We readily concede, of course, that it would be nice to have more detail than can be obtained at the 1:250,000 scale, but we do not want a situation in which the time for completion of the survey disappears over the horizon - excellent though the few maps produced may be. We would prefer a complete coverage with a limited amount of detail and precision."

"Our need for this digitisation is becoming urgent. Not only are we concerned with the ecological survey of Britain, but we have extensive studies of the land use of the uplands, and of coppice woodlands in the south of England. Our pollution studies will require cartographic presentation and we are currently planning a project on energy production of ecological systems."

3. At this same time ECU was already engaged in another large digitising operation for the Institute of Geological Sciences (IGS) in order to produce a series of Geochemical atlases at 1:250,000 scale. This project started in the north of Scotland: it is proceeding southwards at a somewhat slower rate than that envisaged by the ecologists on account of the massive amount of sampling data involved. The geochemical project also required that topographic base map information should be digitised at 1:50,000 scale. In addition, updated geological compilations were made at 1:50,000 scale and these were also digitised.

4. Under these circumstances ECU argued that Ecobase should consist of material derived from the 1:50,000 scale as that came into existence; the rest of the country should be digitised at 1:250,000 scale being replaced, bit by bit, as further 1:50,000 data became available. The present coverage of the two scales is shown (fig. 1) with the larger scale over most of the Highlands and Islands of Scotland - circa 20% of Britain. Thus, digital data covering the whole of Britain has been available since mid-1979.

5. It must be emphasised at once that this digital data does not represent the totality of the published OS maps either at 1:50,000 or 1:250,000 scales. The objectives of Ecobase are manifestly different from those of topographic maps which anyhow exist as printed maps already and obviously do not require reconstituting digitally at substantial extra cost.

The maps elements that have been digitised for Ecobase are:

1. Coastlines (high water mark) including islands.
2. Rivers, canals and lakes.
3. Motorways, A class and B class roads.
4. Railways.
5. Woodlands.
6. Urban areas.

Additionally, ECU is "importing" digital data from other government agencies for:

7. Contours (at 100m interval)
8. Administrative boundaries (to ward/parish level).

Note that items 1-6 are broadly common as between 1:50,000 and 1:250,000, though the larger scale gives a positional accuracy of +/-10m against +/-75m for the smaller scale.

6. The following notes illustrate some of the issues involved in such an exercise.

(a) Coastlines We plan to use the same coastline data throughout relying on software (program REDUCE) to smooth data and get rid of intermediate co-ordinates within each line segment that are visually "unnecessary" for smaller scale plots. However, the smaller islands are still left with some line segments and are liable to show up, if only perhaps as dots, at smaller scales. The only solution here seems to be that of measuring the areas of every island and sieving out the smaller ones as appropriate for the scale - and for the purpose - of the map. This area measurement is a somewhat tedious and lengthy process even when run in batch mode.

(b) River classification Initially all rivers and lakes at both scales were tagged with single feature codes. However, to enable a coherent 1:250,000 water pattern to be produced from a data base which has topographic input at variable scales, a recoding exercise has been necessary. Table 1 below gives details of this.

(c) Roads Each class of road carries a feature code which distinguishes it so that the selection of particular classes for particular scales is straightforward. In a few selected areas we have also digitised minor roads, tracks and "rights of way": these lower categories may have some value in themselves and they have also been used to identify areas of inaccessibility (e.g., 500 metres from any track or road) a factor which may have ecological significance.

(d) Railways These have been treated as a single class of feature

(though "disused railways" are also separately digitised). We are aware that smaller scale maps may make it necessary to add a code to those lines which, on operational grounds; are of greater importance.

(e) Woodland This is an undifferentiated category with no reference to species at this stage. Obviously those areas where the source material was at 1:50,000 scale will, in general, show more small bits of woods: however, woods are relatively infrequent in Highland Scotland and this geographical fact balances the smaller scale source used so far for the rest of Britain. The problem of selecting woods at smaller scales is that of selecting any other "islands" and will probably have to depend on area measurement.

It is worth noting that woods adjoining roads or rivers have been digitised by a single line segment carrying feature codes for both.

(f) Urban areas OS maps show a tint that defines built-up areas and it is the outline of this that has been digitised: this tint obviously does not cover isolated buildings. The problem of selecting only major built-up areas for showing on small scale maps is very similar to those of selecting woods and islands.

(g) Contours The imported data set for contours was originally prepared more as input for a digital terrain model than with the intention of providing, for example, complete sets of closed polygons representing, say, the 100m contour. Thus, the digital contours stop at cliffs and have small gaps and discontinuities which have need of repair before being incorporated in Ecobase. The data base as well as providing a set of contours at 100m interval also includes some form lines at 50m interval where the relative flatness of slope makes this desirable, as well as a liberal scatter of spot heights. To edit this data set - topologically as well as topographically - is a long and expensive task under the ECU system, taking some 30%-50% of the time of redigitising de novo.

It has to be noted that Ecobase contours will, when tidied up, deal only with the land area and have not so far been extended below the zero-contour; nor do they portray the depths of lakes, etc.

(h) Administrative Boundaries This data set is also an imported one (from the Dept. of the Environment): it shows boundaries down to parish/ward level and was originally digitised at one-inch-to-the-mile scale and in relation to boundaries as used in the 1971 population census. Since

much statistical data, e.g., from the population census is published in terms either of these boundaries or of centroids based on them, this boundary data, once incorporated into Ecobase, consequently provides a key for computer mapping of demographic, medical, sociological data, as may be wished.

However, before this data can be properly incorporated into Ecobase two editing functions are necessary on it. The first of these requires the elimination of duplicated boundaries so that the file contains only a single digital record of any line segment but tags common boundaries with the codes of the relevant contiguous areas. It is estimated that this editing process will reduce the size of the data sets by 40%. Secondly, editing is necessary to ensure that where these imported boundaries are coincident with parts of topographical lines, like rivers that are already in the data base, one or other line segment is deleted and the archive appropriately updated so that both/all feature codes are alloted to the particular segment. This is, again, a job for interactive editing, and a long one.

7. While it is not clear whether Ecobase, as it now stands, is a large system or a small one - and perhaps judgements on that will alter over time - it is clear that we have more data than can be accommodated on the 80 megabite disk pack on the ECU's PDP-11/34 computer. This means that fast interactive interrogation of any of the data of all the area is not possible - at least at the resolution of digitising. This is not very surprising but is disappointing. The attractiveness of Ecobase is partly related to the concept of instant retrieval of any area and any combination of features.

8. The 1:250,000 part of the data (1) (water pattern, communications, woodlands and urban areas) occupies 66% of an 80 megabyte disc-pack and comprises 100,870 line segments with a total line length of 554 metres (2). The 1:50,000 part of the data (1), which covers approximately 20% of the land area, occupies a whole 80 megabyte disc and comprises 136,650 line segments with a total line length of 1,012 metres (2). The administrative boundaries and contours to cover the whole area at present occupy a further $1\frac{1}{2}$ disc-packs, although this figure will be reduced when line duplication within the administrative boundary dataset is eliminated.

---

(1) See Figure 1

(2) At 1:250,000 scale

Table 1

Categories of rivers and lakes as shown on Ordnance Survey 1:50,000 maps and as used in Ecobase.

Feature codes 4 or 5 in position 1 denote river or lake respectively.
Feature codes in position 2 give categories of river or lake.

(a) Rivers shown by double lines which are non-parallel.	Both banks plus islands within river are digitised (Code 4,28). Centre lines also entered by eye during interactive editing (Code 4,27).
(b) River shown by double parallel lines.	Centre lines are digitised (Code 4,26).
(c) Thick lines.	Centre lines are digitised (Code 4,29).
(d) Thin lines constituting all other rivers, drains etc.	Lines are digitised (Code 4).

NB By running program LINKER any combination of the above may be selected.

(e) Lakes within a category (a) river system.	Digitised as mapped (Code 5,58).
(f) Ditto for category (b) rivers.	Digitised as mapped (Code 5,56).
(g) Ditto for category (c) rivers.	Digitised as mapped (Code 5,59).
(h) Ditto for category (d) rivers.	Digitised as mapped but excluding those estimated as below 25 square kilometres on ground (ie 10mm^2 at 1:50,000) (Code 5,50). Remainder as mapped (Code 5).

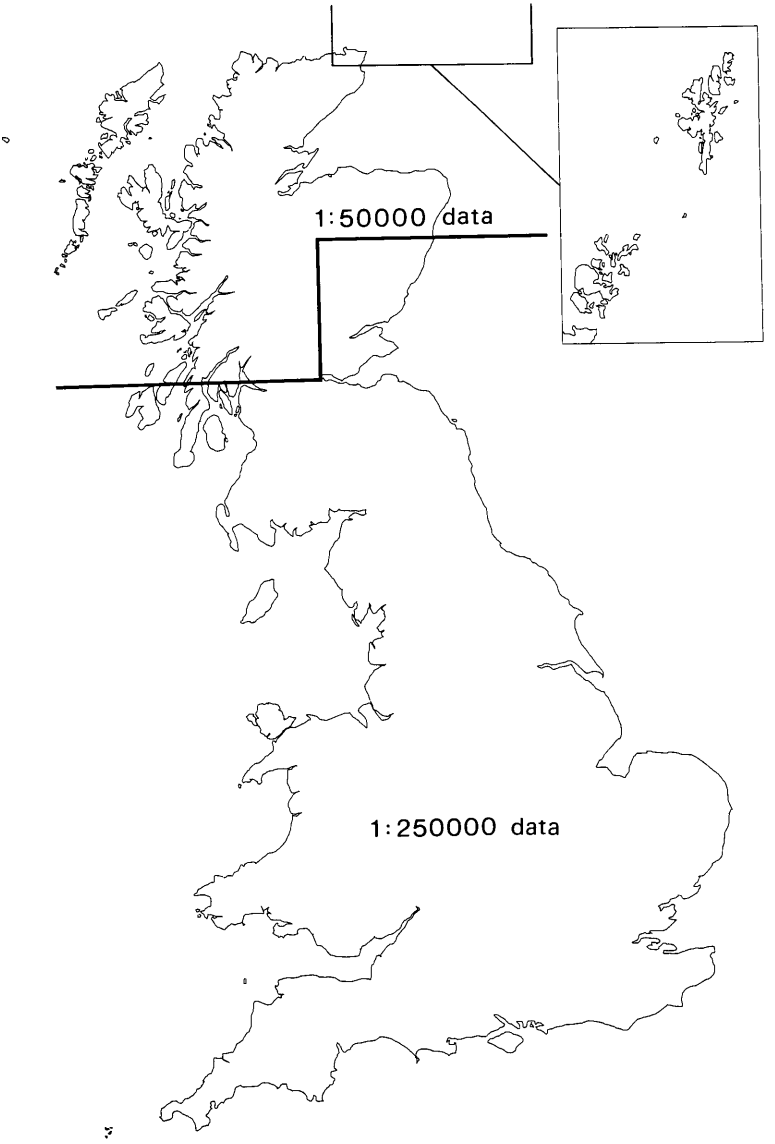

Figure 1
Scale of digitisation of planimetric data

DIRECT PLOTS OF ECOBASE

Figure 2(a)
At 1:2.000.000
Showing coast major rivers and lochs as digitised at 1:50.000

Area of figs 2(b) and 2(c) indicated.

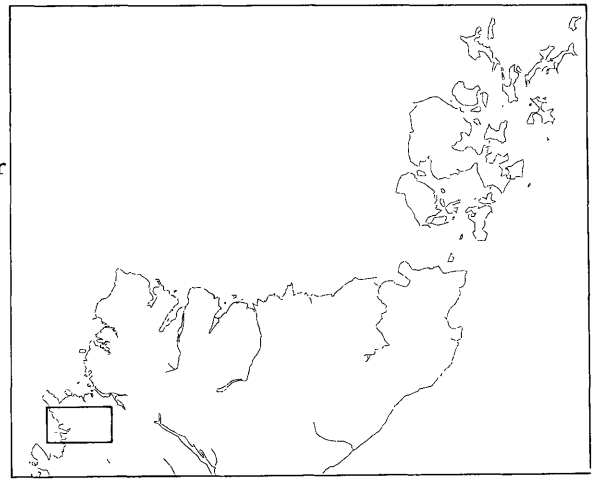

Figure 2(b)
At 1:250.000
Showing coast & three out of four (width) categories of rivers, lochs etc

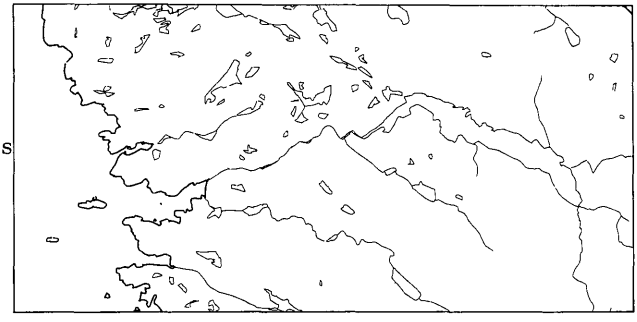

Figure 2(c)
At 1:250 000
Showing coast all rivers and lochs as digitised at 1:50 000