Bunce, R. G. H.; Barr, C. J.; Whittaker, H. A.. 1983 A stratification system for ecological sampling. In: Fuller, R. M., (ed.) *Ecological mapping from ground, air and space.* Cambridge, Institute of Terrestrial Ecology, 39-46. (ITE Symposium, 10).

A STRATIFICATION SYSTEM FOR ECOLOGICAL SAMPLING

R G H BUNCE, C J BARR AND HEATHER WHITTAKER
Institute of Terrestrial Ecology, Merlewood Research Station, Grange-over-Sands

ABSTRACT

Stratification procedures in ecology have not reached a comparable degree of sophistication to those in other biological disciplines, often relying upon intuition rather than a more systematic approach. Such a reliance upon personal judgement often leads to problems of consistency in interpretation and to the arbitrary selection of sample sites whose relationship to the whole area is insufficiently understood. The methods of stratification used in various branches of ecology are discussed, leading to the description of a system where personal judgement is reduced as far as possible. Strata are produced by objective analysis, from which a small number of randomly selected samples can then be drawn, known to be representative of the whole area. Because of this relationship, detailed studies carried out in these sample areas can be converted to overall estimates which can be subjected to tests of external validity and have statistical errors attached to them. The applications of this system to various types of ecological mapping are discussed.

INTRODUCTION

The advantages of stratified sampling in ensuring economy of expensive field studies have long been established. However, the majority of resource evaluation studies still largely rely upon complete coverage or non stratified random samples. Complete coverage leads to problems of sufficient consistency in recording and identification of precise categories. Stratified systems of sampling are more efficient in providing more accurate estimates with the same number of samples than random procedures and provide the main reason for the development of the system described below.

In contrast, the identification of social strata has proceeded to such an extent that many aspects of the human behaviour of the British people can be relatively accurately predicted from small samples. The measure of the success of such sampling may be seen by the efficiency of the advertising industry in the identification of consumer choices. The analysis of the natural environment is analogous to the sampling of human behaviour in that it involves simplifying multivariate data. In general, stratification is used for sampling in many disciplines but has not been widely applied in ecological science.

It is useful to summarize the approaches adopted by government agencies to various types of ecological surveys. First, the conservation bodies, whether of wildlife or landscape, rely principally upon individual recommendations for identification of key sites - as exemplified by the selective approach used by the Scottish Countryside Commission in setting up a series of scenic heritage areas in Scotland (Countryside Commission for Scotland 1978). Whilst such an approach may work well in finding unique sites, it is difficult to place these in the context of the country as a whole. Much commissioned research concerned with countryside, eg the monitoring of changes in landscape features, requires the identification of areas typical of the country as a whole, but often sites are selected arbitrarily for study, with no assessment of their degree of representation.

The annual reviews by the Ministry of Agriculture, Fisheries and Food (MAFF) and the Department of Agriculture and Fisheries for Scotland (DAFS) give total coverage but rely upon the accuracy of individual farmer's returns. Some stratification is used, however, when studies of different types of farm enterprises are undertaken. The various Forestry Commission censuses use different sampling patterns but do not employ external means of stratification. Similarly, maps of forest potential are based on complete coverage at low levels of detail. The Soil Survey, until recently, was concerned with detailed mapping at a local level, although this has been extended to both a larger scale for the whole country and at 10 km intersects. The Department of the Environment accumulates many types of data associated with ecological matters but invariably relies upon data submitted by component bodies with attendant problems of assessing relative accuracy. For example, as Coppock and Gebbett (1978) stated, there has been no co-ordinated survey of urban land, and the validity of the estimates that have been made is difficult to assess because of the variety of sources from which they were derived. The land utilization surveys of Stamp (1937-47) and Coleman (1961) rely upon complete coverage.

In vegetation science, sampling is usually by random or selective methods. Few attempts at external stratification have been made, with gradient analysis (Whittaker 1973) being comparable to the method described below, although not explicitly using stratification since vegetation and environment are treated as co-ordinates in a two-way diagram. Other ecological studies have used a comparable variety of regimes but examples of independent stratification have not been found. Finally, although technically extremely sophisticated, the analysis of various types of imagery generally attempts complete coverage - for example, there was little discussion of sampling presented in the papers concerned with remote sensing at this symposium.

The conclusion is that stratification is virtually unused. Moreover, the majority of national statistics do not have estimates of the errors attached to them and they have no external tests of validity. Advances in technology in the evaluation of resources have not therefore proceeded in parallel with advances in statistical procedures or data handling through computers.

THE LAND CLASSIFICATION SYSTEM

Many types of resource mapping attempt to derive homogeneous units, but in an ecological context these rarely exist. Heterogeneity is an integral feature of the natural world and any classification system should be designed to deal with it. Analogies may be drawn with taxonomy where many units can be readily recognized as distinct species but certain groups are difficult to differentiate. In the ecological field, difficulties are encountered for the same reason - continuous variation. Whilst the environment of Britain generally varies imperceptibly from one area to another, the objective of a land classification system is to produce classes that match the patterns that are present by helping to define boundaries. The more appropriate the analysis and data used, the better the classes will fit the natural patterns and the more efficiently they will be used as strata for sampling ecological parameters. The subjective observation of intercorrelations between ecological factors is the basis for the interpretation of ecological patterns in the field, and the land classification formalizes such intuitive ideas. The data necessarily consist of a large number of factors to adequately express the environment - single factors, such as altitude, are not useful on a national basis as they hold insufficient correlations - for example, an altitude of 300 m in Devon has different environmental implications from the same height in the north of Scotland.

The main objective of the system is to provide strata for ecological sampling in Britain. It is designed to reduce personal judgement to a minimum, although appropriate numerical methods need to be selected. The other 2 main objectives are to obtain a representative sample of the land use, vegetation and soils of Britain at a given time and to provide a basis for monitoring ecological change. Earlier studies are described by Bunce *et al.* (1975) and Bunce and Smith (1978).

The land classification has 3 phases:

1. Land classification, based on environmental parameters.

2. Ecological characterization.

3. Prediction.

The first phase sets up classes corresponding to the 'x' axis of a regression, ie they act as independent variables. The classes are then used to make ecological observations, equivalent to the 'y' axes value of regression, ie the dependent variables. Intercorrelations between the 2 sets of factors can then be used to make predictions, the validity of which can be assessed according to independent, external criteria.

Land classification

The classes were determined by analysis of 282 attributes recorded from 1228 1 km squares, in a grid at 15 x 15 km intersections. Indicator Species Analysis (ISA) (Hill *et al.* 1975), a divisive, polythetic method of analysis, was used. This method divides the data hierarchically into classes that are relatively homogeneous, identifying indicator attributes at each stage. The data relate to climate, topography, geology and human artefacts recorded from maps. The analysis was used to define 32 classes. The indicator attributes alone can be used to assign any square in Britain to its appropriate class. In this way, the original 282 attributes were reduced to 76. A further 4800 squares have been assigned in this way, by using the key provided by ISA, in order to provide an estimate of the overall breakdown of each 1 km square in Britain to a land class. Increasing the sample number has been shown to modify only slightly the proportion of squares in each land class within a region, but does progressively increase the geographical resolution.

Ecological characterization

Eight squares were drawn at random from each of the 32 land classes. Vegetation data were recorded from 5 random quadrats in each square, with soil data being recorded from pits dug in the centre of each quadrat. Linear quadrats were also placed along hedgerows, running water and roadsides. The patterns of land use throughout the square were mapped, as well as other ecological information such as hedgerow length and woodland composition. The presence of other features, eg the breeds of livestock, was also recorded. The field procedure has subsequently been modified for differing objectives, eg for planning purposes, by the Highland Regional Council.

Prediction

The values from the squares are used to calculate the mean figures per land class for factors such as area of barley or hedgerow length. As the numbers of squares in Britain belonging to each land class are known, estimates can be obtained for any factor by summing the land classes for Britain as a whole. Descriptions of 2 land classes are given in Figures 1 and 2 to show the range of information that can be incorporated. Further estimates of land use potential or additional criteria can be assessed from the information contained in the series of sample squares. Predictive maps can also be made to indicate high concentrations of particular factors. Further information on the methodology is provided by Bunce *et al.* (1981a & b), and Benefield and Bunce (1982).

VALIDATION

Comparison with independent estimates

The figures published by MAFF and DAFS provide a convenient test of the sampling system and the estimates agree reasonably with them, particularly for crops with large areas. For example, the DAFS/MAFF figure for wheat (for 1977/78) was 1.16×10^6 ha, whereas the land classification figure was 1.11×10^6 ha, and for barley were 2.31×10^6 ha and 2.19×10^6 ha respectively. A complete table is given by Bunce *et al.* (1981a). Other published data show similar patterns of comparability with estimates from the land classification. The area of urban land estimated by Best (1976) was 20 517 km^2, whereas a preliminary estimate from the land classification was 22 242 km^2. Urban growth was estimated by Best (1976) at 148 km^2/year, whereas a trial study using the classification suggested that it continued at an average rate of 149 km^2/year between 1971 and 1981. A third example results from the MAFF study of improvable land carried out in 1980 in the north of England. Measurements were made of the area of Agricultural Land Classification (ALC) grades 4 and 5 for the entire region during the study and were subsequently repeated using a sample number of squares drawn from the land classification procedure. From the ALC map the proportions were 31.0% and 15.6% respectively, and from the land classification 26.1% and 15.7%.

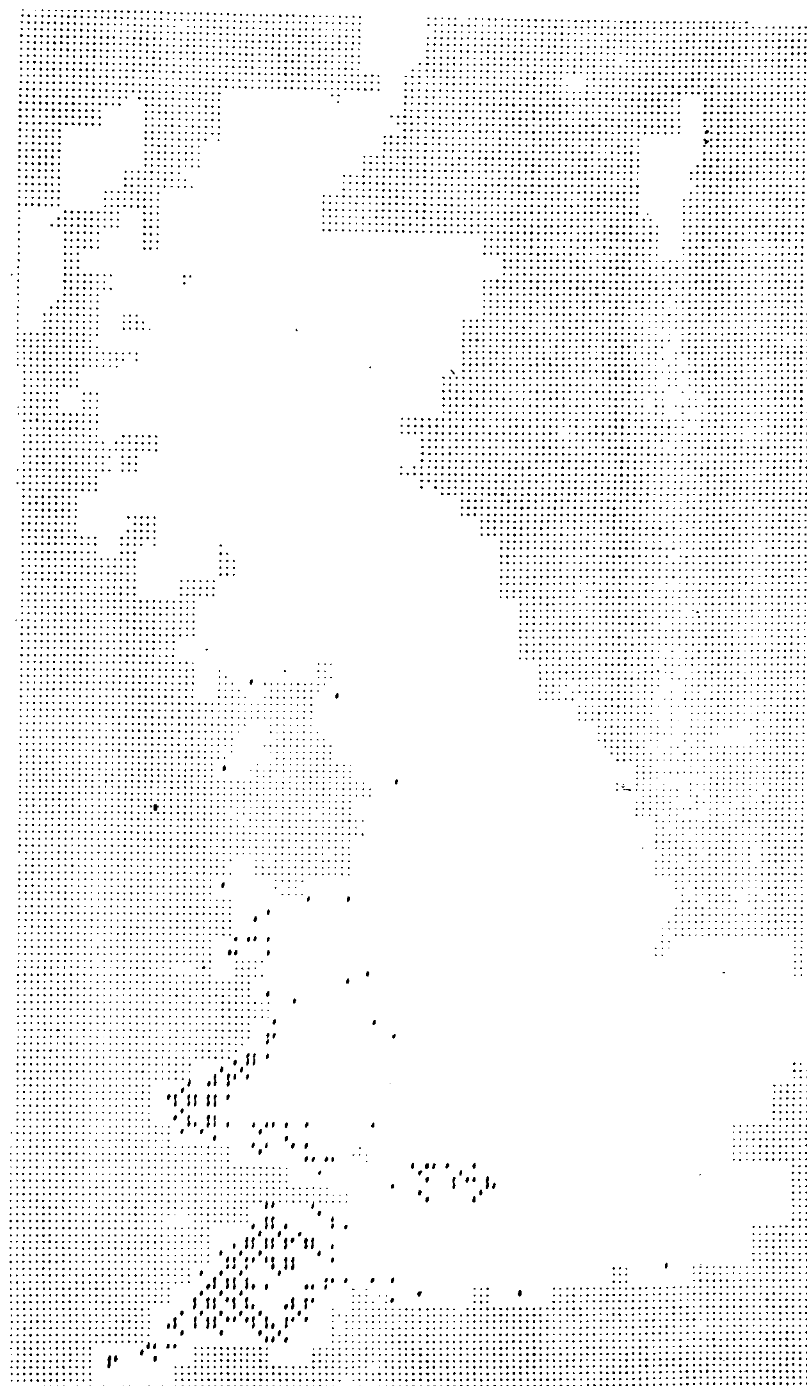
Standard errors

Studies have commenced using the sample squares to calculate first the variances for the individual land classes and secondly those combined for Great Britain as a whole. Preliminary estimates show standard errors of the mean values to vary between 14.3% and 48.0%. Further work is needed on the distribution patterns within the data and on the efficiency of stratified as opposed to random samples. The effects of increasing the sample size as well as the theoretical background to the estimation of errors also need investigation.

Correlation

Various types of correlation analyses can be carried out to test the overall links between the environmental classification and the various sets of survey data. For example, the mean values of the reciprocal averaging ordination axes of the initial environmental data used to produce the land classification were calculated for each class. Component analysis was used to extract component values from the matrix provided by the mean values of the land use for each of the 32 land classes. Similarly, component values were extracted from the mean proportions of soil groups in the 32 classes. The correlation coefficients from the mean axis scores from the land classification were (both 30 df) 0.823 with the first component of the land use data and 0.923 with the first component of the soils data ($P \leqslant 0.001$ in both cases). The procedure followed is comparable with that described by Fourt *et al.* (1971) and demonstrates the very high correlations present between the analysis of map data and subsequent field survey.

The conclusion to be drawn from the various validation exercises is that the 256 squares represent an adequate sample for many purposes, but that sample numbers from some classes need to be increased when examining particular parameters. The necessary sample number for a required accuracy can be calculated from the preliminary survey.



LAND CLASS SIX

TOPOGRAPHY

Mean max. altitude (m)	....	147
Mean min. altitude (m)	....	75
Altitude class (mean percentage area)	0- 76m.........	21
	77- 198m.........	68
	199- 488m.........	9
	489-1189m.........	-
Slope (°)	....	4

CLIMATE

Mean min. temp. January	(°)..	2.8
Mean max. temp. July	(°)..	12.3
Mean soil deficit	(mm)..	6.0
Mean annual rainfall	(mm)..	11.1
Mean snowfall	(days)..	12.9
Duration bright sunshine	(hrs)..	6.2

SOILS

Mean pH	5.0
Mean loss on ignition (%)	15.3

Percentage of total area

Brown earths......................	45.0
Rendzinas........................	5.0
Gleys............................	22.5
Gleyed brown earths..............	10.0
Brown podsolic soils.............	-
Rankers..........................	2.5
Calcareous brown earths...........	-
Peaty podsols....................	-
Podsols..........................	12.5
Peaty gleys......................	2.5
Peats............................	-

LAND USE

Percentage of total area

Wheat............................	-
Barley...........................	13.7
Other Crops......................	2.9
Horticulture.....................	0.2
Leys.............................	27.1
Permanent grass..................	30.0
Rough pasture....................	-
Bracken..........................	0.3
Rushes...........................	0.9
Moorland.........................	-
Peatland.........................	-
Mountain grass...................	1.3
Woodland.........................	10.8
Cliffs/sand/mud..................	-
Built-up.........................	11.6

NATIVE SPECIES

Percentage cover of major species

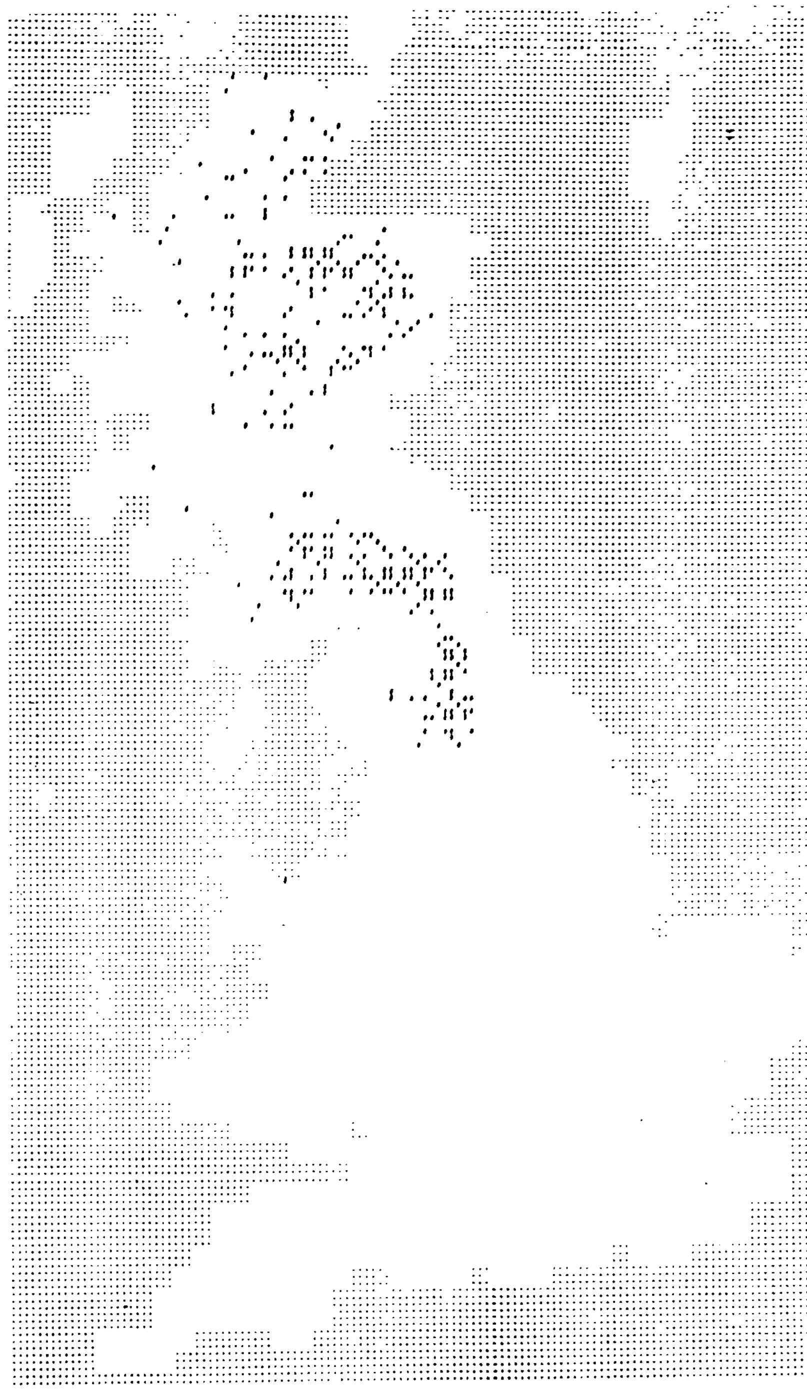
Perennial rye grass.....	20.0
Ling heather............	0.1
Common bent.............	6.5
Purple moor grass.......	-
Yorkshire fog...........	6.4
White clover............	3.0
Cocksfoot...............	4.0
Matgrass................	-
Bracken.................	1.5
Crested dogstail........	1.8
Italian rye grass.......	2.5
Timothy.................	3.1
Deer grass..............	-
Sheeps fescue...........	0.6
Creeping bent...........	3.4

LANDSCAPE

Barbed wire fences
Hedges managed
Hedges neglected
Hedges on banks
Hedgerow trees
Mature woodland
Gravel streams
Vernacular local
Slate roofs
Corrugated iron on roofs
Tracks
Farmhouse and farms

Figure 1 Description and occurrence of land class 6, gently rolling enclosed country; mainly fertile pastures

LAND CLASS TWENTY-TWO



TOPOGRAPHY

Mean max. altitude (m)	....	430
Mean min. altitude (m)	....	298
Altitude class (mean percentage area)	0- 76m.........	-
	77- 198m.........	2
	199- 488m.........	87
	489-1189m.........	9
Slope	(°)	8

CLIMATE

Mean min. temp. January	(°)..	-0.2
Mean max. temp. July	(°)..	19.0
Mean soil deficit	(mm)..	5.0
Mean annual rainfall	(mm)..	12.8
Mean snowfall	(days)..	51.6
Duration bright sunshine	(hrs)..	4.4

SOILS

Mean pH	4.5
Mean loss on ignition (%)	70.2

Percentage of total area

Brown earths.....................	10.0
Rendzinas........................	-
Gleys............................	12.5
Gleyed brown earths..............	2.5
Brown podzolic soils.............	7.5
Rankers..........................	2.5
Calcareous brown earths..........	-
Peaty podzols....................	15.0
Podzols..........................	-
Peaty gleys......................	32.5
Peats............................	15.0

LAND USE

Percentage of total area

Wheat............................	-
Barley...........................	2.1
Other Crops......................	1.4
Horticulture.....................	-
Leys.............................	0.9
Permanent grass..................	6.5
Rough pasture....................	1.6
Bracken..........................	1.1
Rushes...........................	3.3
Moorland.........................	25.6
Peatland.........................	12.1
Mountain grass...................	-
Woodland.........................	34.9
Cliffs/sand/mud..................	1.1
Built-up.........................	1.6

NATIVE SPECIES

Percentage cover of major species

Perennial rye grass.....	2.8
Ling heather............	28.9
Common bent.............	0.3
Purple moor grass.......	3.9
Yorkshire fog...........	0.8
White clover............	3.3
Cocksfoot...............	0.1
Matgrass................	0.3
Bracken.................	1.8
Crested dogstail........	0.5
Italian rye grass.......	-
Timothy.................	0.3
Deer grass..............	0.3
Sheeps fescue...........	1.1
Creeping bent...........	-

LANDSCAPE

Woodlands over 5 ha.
Stones/rocks
Lines of shrubs

Figure 2 Description and occurrence of land class 22, margins of high mountains, moorlands, often afforested

APPLICATIONS

Many of the potential applications of the system are demonstrated by the project currently being undertaken under a contract with the Department of Energy (Bunce *et al.* 1981c). The main objective of this study is to estimate the area of land in Britain that could be available for wood energy plantations under various constraints, both financial and institutional. A representative series of squares was submitted to other programme participants (cf MacDonald *et al.* 1981) who assessed the net present values of the current agricultural pattern and the potential of conventional forestry, of single stem trees for energy, and of coppice trees for energy. The probability of institutional constraints upon forestry development was defined by examining their occurrence on the base maps. Each square therefore had a series of overlays for which the overlaps were successively calculated to determine which land use or potential land use would have the highest yield in economic terms. The institutional constraints were then superimposed at successive probability levels to examine their effect. Various discount rates and different markets, eg methanol or synthetic natural gas, were then taken in order to investigate the sensitivity of the availability of land to economic factors. Each land class is treated separately, the individual estimates being combined to obtain figures for Britain as a whole. Land classes with high potential can then be defined, regional predictions made and the principal patterns of likely availability described for the whole of Britain.

This project shows how the system can be used to obtain estimates of potential for Britain from a small number of sample squares. The framework of the land classes is being used to assess the sensitivity of the land use patterns in Britain to changes in basic economic assumptions. Pilot studies were used for representative classes to focus the study on areas identified as being of significance so that expensive trials are avoided. It is of particular importance that the estimates are related back to identifiable areas of land rather than synthetic units. These areas can be identified subsequently for experimental field studies.

The principle of applying various constraints to the potential of land in a representative series of samples is widely applicable in environmental impact studies. The effect of various potential changes can be compared overall and those likely to affect the system to a major degree could be identified. In a different type of study, the strata are being used to identify populations for genetic studies. The validity of small numbers of samples allows regular monitoring and the potential of the land classification for following habitat or landscape change is therefore under consideration. It is planned to monitor the loss of agricultural land to urban development in a similar way, with increases in sample number for some classes being incorporated in order to attain the required accuracy. At a regional scale, studies are in progress to set up a local series of representative squares for sampling rural land use, but which at the same time co-ordinate with the national series of land classes. A comparable local study has derived land classes to provide an unbiased stratification for assessing pollution levels and monitoring changes. In all such studies, the potential for change can be examined efficiently by reducing the requirement for field studies and identifying the topics likely to be profitable for future work.

CONCLUSION

Stratification is rarely used in ecological surveys. The system described uses defined environmental strata that have been shown to be relevant to the survey of a range of ecological factors. The data from strata can be subjected to tests of statistical accuracy and external validity. They provide a framework for measuring current and potential land use and enable the sensitivity of the proportions of the land surface of Britain to various changes to be assessed.

REFERENCES

BENEFIELD, C.B. & BUNCE, R.G.H. 1982. *A preliminary visual presentation of land classes in Britain.* (Merlewood research and development paper no. 91). Grange-over-Sands: Institute of Terrestrial Ecology.

BEST, R.H. 1976. The extent and growth of urban land. *The Planner,* 62, 8-11.

BUNCE, R.G.H., MORREL, S.K. & STEL, H.E. 1975. The application of multivariate analysis to regional survey. *J. environ. Manage.,* 3, 151-165.

BUNCE, R.G.H. & SMITH, R.S. 1978. *An ecological survey of Cumbria.* Kendal: Cumbria C.C. and Lake District Special Planning Board.

BUNCE, R.G.H., BARR, C.J. & WHITTAKER, H.A. 1981a. An integrated sysem of land classification. *Annu. Rep. Inst. terr. Ecol. 1980,* 28-33.

BUNCE, R.G.H., BARR, C.J. & WHITTAKER, H.A. 1981b. *Land classes in Great Britain: preliminary descriptions for users of the Merlewood method of land classification.* (Merlewood research and development paper no. 86). Grange-over-Sands: Institute of Terrestrial Ecology.

BUNCE, R.G.H., PEARCE, L.H. & MITCHELL, C.P. 1981c. The allocation of land for energy crops in Britain. In: *Energy from biomass,* edited by W. Palz, P. Chartier and D.O. Hall, 103-109. London: Applied Science.

COLEMAN, A. 1961. The second land use survey: progress and prospect. *Geogrl J.,* 127, 168-186.

COPPOCK, J.T. & GEBBETT, L.F. 1978. *Land use and town and country planning.* (Reviews of U.K. statistical sources 8). London: Pergamon.

COUNTRYSIDE COMMISSION FOR SCOTLAND. 1978. *Scotland's scenic heritage.* Perth: C.C.S.

FOURT, D.F., DONALD, D.G.M., JEFFERS, J.N.R. & BINNS, M.O. 1971. Corsican pine (*Pinus nigra* var *maritima* (Ait.) Melville) in southern Britain: a study of growth and site factors. *Forestry,* 44, 189-207.

HILL, M.O., BUNCE, R.G.H. & SHAW, M.W. 1975. Indicator species analysis, a divisive polythetic method of classification and its application to a survey of native pinewoods in Scotland. *J. Ecol.,* 63, 597-613.

MacDONALD, D.W., BUNCE, R.G.H. & BACON, P.J. 1981. Fox populations, habitat characterization and rabies control. *J. Biogeogr.,* 8, 145-151.

STAMP, L.D. 1937-1947. *The land of Britain: The final report of the Land Utilisation Survey of Britain.* In 92 parts of 9 volumes. London: Geographical Publications.

WHITTAKER, R.H. 1973. *Ordination and classification of communities.* The Hague: Junk.