



Preliminary Survey of the Peat-Bog Hummell Knowe Moss Using Various Numerical Methods

Author(s): R. S. Clymo

Reviewed work(s):

Source: *Vegetatio*, Vol. 42, No. 1/3 (Oct. 31, 1980), pp. 129-148

Published by: [Springer](#)

Stable URL: <http://www.jstor.org/stable/20145799>

Accessed: 23/07/2012 04:48

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Springer is collaborating with JSTOR to digitize, preserve and extend access to *Vegetatio*.

<http://www.jstor.org>

PRELIMINARY SURVEY OF THE PEAT-BOG HUMMELL KNOWE MOSS USING VARIOUS NUMERICAL METHODS*

R.S. CLYMO**

Botany Department, Westfield College, London NW3 7ST, England

Keywords:

Mire, Multidimensional scaling, Ordination, Peat-bog, Principal component analysis, Principal coordinate analysis, Reciprocal averaging, Transformation of data

Introduction

The peat-bog Humell Knowe in northern England is unusual in having a nearly complete *Sphagnum* carpet with abundant *Eriophorum vaginatum*, and *E. angustifolium*, but with little differentiation of hummock or hollow topography. This paper reports a primary survey of the vegetation and stratigraphy, made with strictly limited resources, and intended to direct the planning of further work. A variety of numerical aids were used in the vegetation analysis. The usefulness of these aids was judged by the extent to which they both confirmed what was really obvious to the ecologist-on-the spot, and the extent to which they indicated what further work should be done: a method which does not reveal the really obvious must be suspect, but one which does nothing but reinforce prejudice is equally useless.

* Nomenclature follows Clapham, Tutin & Warburg (1962) – vascular plants; Smith (1978) – *Sphagnum*; Watson (1969) – other bryophytes; and Duncan with James (1970) – lichens.

** The following other members of the Mires Research Group contributed to this survey: D. Bale, R. Bhadresa, D.J. Boatman, O. Clymo, M. Crawley, R.E. Daniels, C. Dickinson, W. Fojt, P. Hayward, P.D. Hulme, E.A. Livett, R. Meade, the late M.C. Pearson, P.J. Shaw, T. Stevenson, J.H. Tallis, P. Wiltshire. I thank them for allowing me to use their results. The analyses and views expressed in this paper are my own. I also thank Mrs. P. Ratnesar for preparing the results for computing and for drawing the diagrams, T. Clymo for typing the text, and Dr J.H. Tallis and Dr K.E. Clymo for commenting on parts of a draft.

The study area

Hummell Knowe Moss (National Grid reference NY 705715; alt. 230 m; long. 2° 28' W, lat. 55° 02' N) is a large blanket mire in the centre of northern England, halfway between Newcastle and Carlisle, and 8 km north of Haltwhistle (Fig. 1). It is one of a group of fine mires known collectively as the Border Mires, which lie in a 30 km band around the central section of the Roman Wall. These include *Butterburn Flow, Broomlee Lough, *Coom Rigg, (Chapman, 1964a, b; 1965), Fozy Moss, Greenlee Lough, *Haining Head Moss, Muckle Moss (Pearson 1960), Robin Rock Flothers and The Wou. (The starred mires

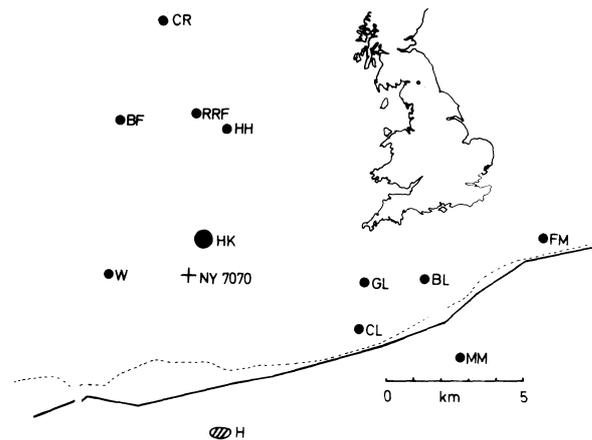


Fig. 1. Location of the Border Mires mentioned in the text. Continuous line, B 6318 road; dashed line, Roman Wall; BF Butterburn Flow; BL Broomlee Lough; CL Crag Lough; CR Coom Rigg; FM Fozy Moss; RRF Robin Rock Flothers; W The Wou. To the south is Haltwhistle (H).

are listed as Grade 1 by Ratcliffe 1977.) Amongst them one may find open water, fen, carr, and bog (much of it apparently undamaged). *Hummell Knowe Moss has a central, slightly domed, bog plain about 0.5 km across (Figs 2, 9) with a nearly complete cover of *Sphagnum magellanicum*, and, in order of abundance, almost constant occurrence, in 25 × 25 cm quadrats, of *Eriophorum vaginatum*, *E. angustifolium*, *Erica tetralix*, *Narthecium ossifragum*, *Andromeda polifolia*, *Vaccinium oxycoccos*, *Drosera rotundifolia* and *Sphagnum tenellum*, with several other species nearly as common. This central area is wet, but with little open water, and is surprisingly free of the hummock and hollow topography, on a scale of about 5 m, which is usually characteristic of such areas. (There are a few hummocks on the moss, not associated with pools, and these may turn out to be of unusual interest.)

Surrounding the central plain (Fig. 2) is vegetation which grades into a variety of marginal communities which border four low (ca. 10–20 m tall) hills to the north-east, north-west and south of the moss. Eastwards and westwards the bog plain slopes downhill, and the surface has obvious drainage channels and eroded peat: *Sphagnum* is less common here, and in the east lichens of the

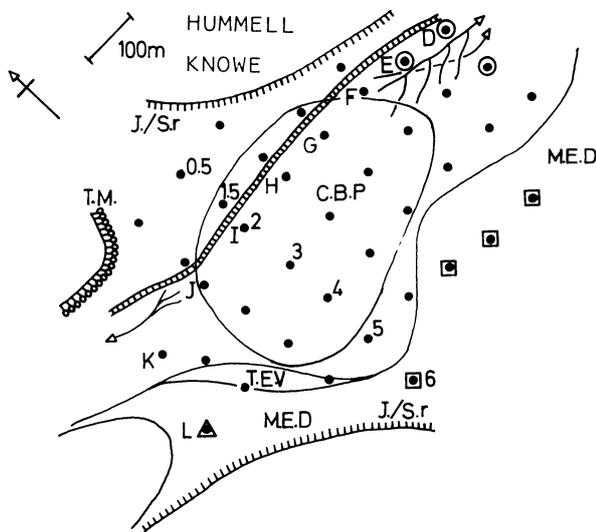


Fig. 2. Isodiametric view of Hummell Knowe bog from the southwest. Filled circles are vegetation sampling sites on a grid (D to L, 0.5, 1.5, 2, 3 . . . 6). A fence runs E-W across the bog. Erosion at east and west ends is associated with drainage channels. Large circle, triangle and square refer to regions of the ordination diagrams. Subjective vegetation types: C.B.P = central bog plain; M.E.D = *Molinia* – *Eriophorum vaginatum* – *Deschampsia flexuosa*; T.E.V = tussocky *E. vaginatum*; T.M = tussocky *Molinia*; J./S.r = *Juncus effusus* – *Sphagnum recurvum*.

genus *Cladonia* are conspicuous. To the north-west and south are communities with little *Sphagnum*, and dominated by *Molinia caerulea*, *Eriophorum vaginatum* and *Deschampsia flexuosa*. To the south-west is a distinctive region which may be flushed by water running off the low hills behind. This region is recognisable by tussocky *Eriophorum vaginatum* on which grow species such as *Potentilla erecta* which are not found elsewhere on the mire centre, though common on the surrounding mineral ground. Between the tussocks is an impoverished bog flora. At the base of the hills to the north and south is a marginal lagg with *Juncus effusus* and *Sphagnum recurvum* common.

The mire continues for more than 500 m to the east (where it is eroded) and to the north, but these areas were not included in the survey. The surrounding area is grazed by sheep, and the hill to the north, Hummell Knowe itself, and the hills to the north-west, have been ploughed and forested. Some of these features are obvious at a glance, others became clear during detailed vegetation recording.

Field methods

The features surveyed and methods used were a compromise between the desirable and the practicable. The survey was made on 20–21 September 1978 by 18 members of the Mires Research Group. Not all were on the site for both days and not all were equally competent taxonomists. The programme was organised to allow rotation of jobs, but also to ensure that the vegetation records were as reliable as possible. The fieldwork occupied, in total, about 200 h.

Because so many people were involved, so little time was available, and so little was known about the study site, a regular sampling pattern was chosen. A reference grid was established. The base-line ran parallel to and 10 m south of the fence which runs east-west across the mire. The grid interval was 100 paces, except that the first line north of the fence was only 50 paces from the base-line. Pacing was used to save time. The accuracy of this grid is about ± 5 paces in the central row, and about ± 10 paces at the northern and southern limits.

The stratigraphy of a complete profile through the peat deposit was examined at one grid intersection (F2) using a Russian-type peat borer (West 1968). Over the whole area the depth of the basal transition from clay to organic remains was established at alternate grid intersections, those in adjacent rows being staggered. Again, the work

was limited by time, and this pattern was chosen to ensure fair cover of most of the site.

The relative height of the mire surface at most grid intersections was measured with a level and staff. A series of closed loops was surveyed. The loop error was usually less than 1 cm.

Vegetation was recorded at most grid intersections in a ladder of 12 contiguous square quadrats, each 25 × 25 cm, on a NE-SW diagonal starting 5 paces SW of the grid point (to avoid trampled areas along grid lines and intersections). A ladder was used for four reasons: so that pattern on a scale up to 3 m would be spanned and might later be revealed by analysis, so that sample site selection was rapid, so that taxonomic advice was readily available, and so that a single recorder could serve as many as four surveyors. The species present in a quadrat were recorded, but no attempt was made to estimate cover, though, as Smart, Meacock & Lambert (1976) very sensibly point out, there is no virtue in recording a feature simply because it is easily recorded. In the present case fourteen people were concerned at various times in vegetation recording, and different observers of the same quadrat often give grossly different (ten-fold) subjective estimates of cover, especially of fine leaved or small but widely dispersed species, which form an important part of the Hummell Knowe mire vegetation. The objective of consistency of identification

was not completely achieved either. It was attempted by having, on the first day, pairs of observers record three adjacent quadrats at each site, with cross reference in any case of uncertain identity. Three of these observers stayed in the vegetation recording party for the whole time and acted as referees. Nevertheless there remain some problems.

1. *Sphagnum magellanicum* could be positively identified when tinged with red, but much of this species, especially where shaded, was green and not consistently distinguishable from *Sphagnum papillosum*. It is likely that *Sphagnum papillosum* is under-recorded and observer biased.

2. The fruticose *Cladonia* species: *C. tenuis*, *C. arbuscula* and *C. impexa* were not consistently distinguishable, so they have been grouped as *C. arbuscula* agg. In general the quadrats contained either no lichens or more than one species, so this should be unimportant.

3. Small leafy liverworts posed two problems. Firstly, they were usually inconspicuous and could be found only by grab sampling within the quadrat. They were therefore under-recorded. Secondly, it is clear from the records that identification was not consistent: at a series of grid points one group consistently recorded *Mylia* whilst another recorded *Mylia anomala*. Such taxa were grouped as 'sp' or 'spp'.

Table 1. Species recorded from sites at Hummell Knowe Moss. Frequency in 0.75 m² at 41 sites.

Andromeda polifolia	37	Erica tetralix	37	Pohlia nutans	3
Aulacomnium palustre	8	Eriophorum angustifolium	38	Polytrichum alpestre	7
Barbilophozia sp.	1	E. vaginatum	41	P. commune	8
Calluna vulgaris	30	Festuca ovina	1	Potentilla erecta	2
Calypogeia sp(p?)	25	Galium saxatile	2	Ptilidium ciliare	1
Campylopus atrovirens	2	Gymnocolea inflata	6	Rhynchospora alba	1
Carex sp.	2	Hypnum cupressiforme	8	Sphagnum sp.	1
C. panicea	3	Lepidozia setacea	20	* S. capillifolium	21
Cephalozia bicuspidata	23	Leucobryum glaucum	4	S. cuspidatum	27
Cladonia arbuscula agg.	2	Lophozia sp.	5	S. magellanicum	35
				S. palustre	8
C. furcata	4	Lophocolea sp.	2	S. papillosum	8
C. impexa	5	Mnium hornum	2	S. recurvum	8
C. pyxidata	1	Molinia caerulea	7	* S. rubellum	4
C. tenuis	7	Mylia sp(p?)	32	S. squarrosum	7
C. uncialis	10	Nardia scalaris	1	S. tenellum	34
Deschampsia flexuosa	5	Narthecium ossifragum	36	S. subnitens	2
Dicranella cerviculata	1	Odontoschisma sphagni	34	Splachnum ampullaceum	1
Dicranum sp.	9	Parmelia physodes	2	Trichophorum cespitosum	34
Drosera rotundifolia	32	Plagiothecium undulatum	6	Vaccinium myrtillus	1
Empetrum nigrum	2	Pleurozium scheberi	7	V. oxycoccum	36

* *S. rubellum*, a small form of *S. capillifolium*, is probably not a distinct species (M.O. Hill in Smith 1978) but it was recorded separately.

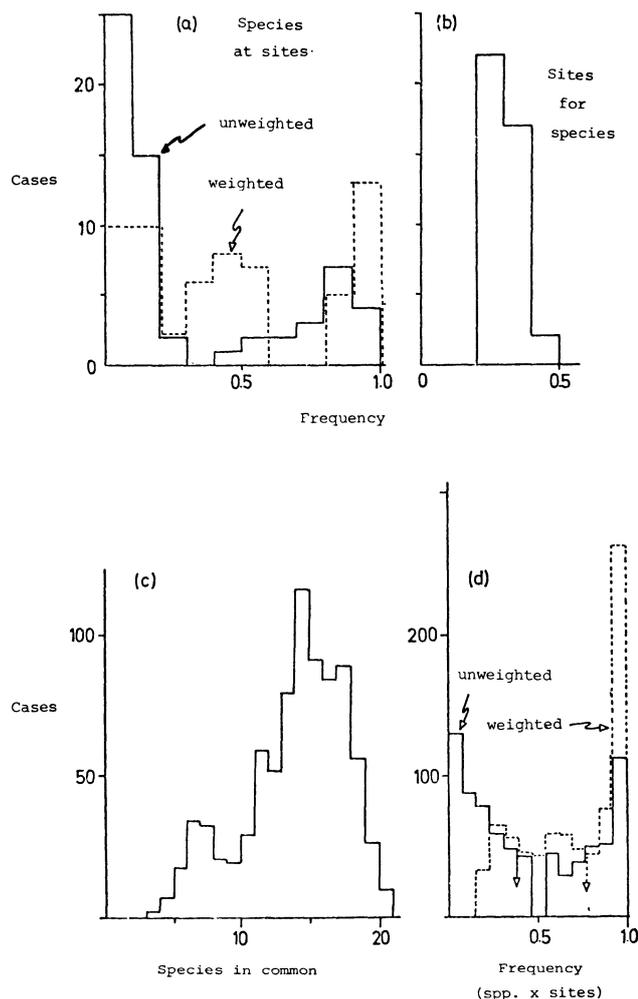


Fig. 3. Distribution of: (a) frequency (at 41 sites) of species; (b) frequency (out of 61 species) at sites; (c) species in common for the (820) comparisons of each site with all others; (d) frequency of all species (out of 12 quadrats at each site, scaled to 0 to 1) at all sites. Only the 777 non-zero values (out of 2501 total) are shown. Dashed lines show values after hyperbolic weighting, with $b = 0.3$ (see text). Vertical arrows show median.

A total of 61 species was recorded at the *41 sites (492 quadrats, total area 30.7 m²). Of these, 10 species occurred at one site only (out of 41) and 28 at fewer than five sites (Table 1, Fig. 3). The number of species at a site ranged from 14 to 29, and overall 31 % of the possible species-at-site occurrences were observed, indicating an unusually

* Only 40 separate sites, but D2 was (accidentally) repeated. The second survey was made by a different group of observers, working in poor light and in a hurry, at the end of the day. The two sets of records have been kept and are both shown in the Figs, to illustrate what is probably the maximum 'recording error'.

uniform vegetation; an impression confirmed by the distribution of the number of species common to pairs of sites (Fig. 3).

After the first day of vegetation surveying it was possible to recognise, tentatively and subjectively, the major vegetation types. The approximate position of the boundary of these types was recorded on the second day, and is shown in Fig. 2.

Finally a simple survey of water chemistry was made. Water samples were squeezed from *Sphagnum*, or taken from runnels, and pH and electrical conductivity measured on the spot.

Numerical methods

We needed to summarise the results, but had no clear idea of the number of dimensions which would be needed, nor of the most useful proximity measure. The number of sites (41) imposed no serious restrictions, but we did want to experiment with means for giving greater or less weight to local site-site relationships. For these reasons we chose ordination rather than classification.

Transformation

The analyses may be considered in stages (Fig. 4). Transformation is used here to describe any manipulation which changes the field records before the analysis proper. A transformation which can be applied to a datum in isolation is here called a weighting (e.g. log, square-root). One which is applied in the context of the values in a row or column is called a standardisation (e.g. to rank, or to unit range or variance, or to zero mean). Transformations needing the complete matrix of values are here called 'adjustments'. The only one used here is Swan's (1970) method for assessing 'degree of absence'.

The importance of weighting is well known: practical ecologists have usually attached more importance to species presence than to amount, as demonstrated by the traditional cover scales, and there is some evidence (e.g. Smartt, Meacock & Lambert 1976) to support this view. For computing purposes it is convenient to have simple continuous functions with one or two parameters which can be varied to adjust the boost given to rarer species (cf. van der Maarel 1979). Some which have been used, for example log functions (Jensén 1978), are awkward insofar as they are undefined if a datum is zero, and would cause unwanted distortion should it be less than 1 %. But

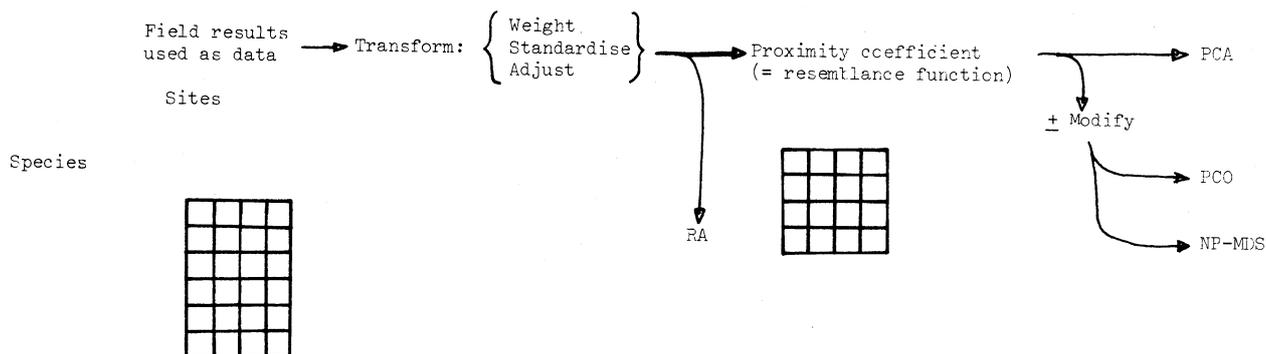


Fig. 4. Stages in an ordination. RA = reciprocal averaging; PCA = principal components analysis; PCO = principal coordinate analysis; NP-MDS = non-parametric multidimensional scaling.

designing such functions is simple. Suppose the original datum, x , can range from 0.0 to 1.0, and the weighted value, $w = f(x)$, is to span the same range. Choose a function, $y = f(x)$, which goes through the point (0,0) and which rises steeply at first but then more slowly. Let $y(1)$ be the value at $x = 1$. Then $w = y/y(1)$ will go through both (0,0) and (1,1). Some functions, such as $w = x^{1/a}$ do this

anyway. Table 2 and Fig. 5 show three such single-parameter functions. Given the subjectivity of choice this is probably two more than are necessary, and it is doubtful if there is any good reason for using a two-parameter function. Of the functions in Fig. 5 the root function gives the greatest boost to rare species, the hyperbolic next most, and the exponential least.

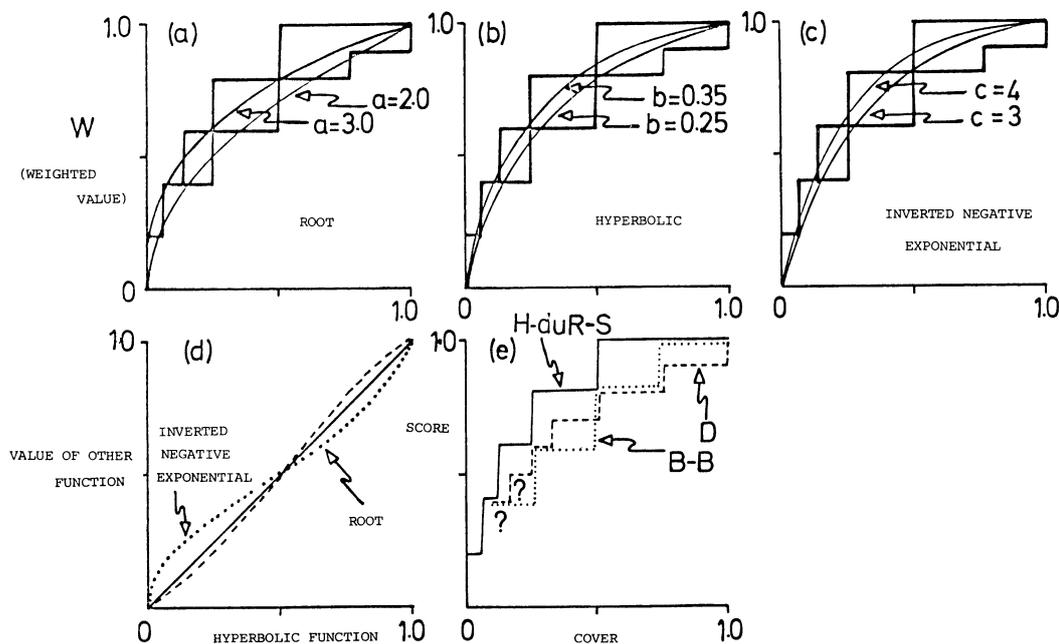


Fig. 5. Three simple weighting functions for two parameter values: (a) $w = x^{1/a}$; (b) $w = x(b+1)/(b+x)$; (c) $w = (1 - e^{-cx})/(1 - e^{-c})$; (d) comparison of these functions with a, b, c such that $x = 0.188, w = 0.5$; (e) traditional cover scales (Hult-Du Rietz-Sernander, Domin, Braun-Blanquet). The lower ends of the D and B-B scales are not unambiguous. The envelope formed by these scales is reproduced in (a), (b), (c).

Table 2. Weighting functions, $w = f(x)$; The range of w and x is 0.0 to 1.0.

Name	$y =$	$w =$	$w = 0.5$	for (0.1875, 0.5)
1. Root	$x^{1/a}$	$x^{1/a}$	$a = -\ln(x)/0.693$	$a \approx 2.416$
2. Hyperbolic	$\frac{1}{1+b/x}$	$\frac{x(b+1)}{x+b}$	$b = \frac{x}{1-2x}$	$b = 0.30$
3. Inverted negative exponential	$1 - e^{-cx}$	$\frac{1 - e^{-cx}}{1 - e^{-c}}$	No simple solution	$c \approx 3.545$

The Hummell Knowe vegetation records contained an unusually large number of species with high frequency (Fig. 3), and many of low frequency. This pattern, if unchanged, would produce analyses dominated by the common species, so the hyperbolic weighting was used, with $b = 0.3$, giving species with frequency 0.19 a weighted value of 0.5. The effect of this on the distribution of weighted frequency is shown in Fig. 3.

The value of standardisation is less clear. Some proximity coefficients (for example the correlation coefficient) include standardising to mean = 0 and variance = 1 and have large effects in some types of analysis – PCA for example. It seems clear however that there is no single criterion which can be applied (Noy-Meir 1971, 1973, Noy Meir, Walker & Williams 1975) to determine what standardisation to use, and that in many cases uncentred data are preferable to centred. Some standardisations increase the importance of rare species, which seems desirable (as already argued), but transformations do this more flexibly. Our interest was chiefly in site-site relations, which might argue for a standardisation by site, but the data as a whole showed unusually little difference in species richness between sites (Fig. 3), so in the end it seemed best to use no overt standardisation.

Swan (1970) devised an adjustment for assessing ‘degree of absence’ of a species. This is certainly a problem with any metric proximity measure. Swan’s adjustment seems to have been applied to field data only rarely, so we tested it on the Hummell Knowe results.

Ordination

Four types of ordination were used (Fig. 4): principal components analysis (PCA), reciprocal averaging (RA), non-parametric multidimensional scaling (NP-MDS) and principal coordinates analysis (PCO).

Reciprocal averaging (Hill 1973) is likely to bend artificial data sets based on vegetation responding to a linear gradient, but it usually keeps the ends apart to give an arch in two dimensions, rather than the convoluted ‘horse-shoe’ of PCA. Comparison of these two methods is often made unfairly, because all axes of an RA plot are usually scaled to the same range, whereas those of a PCA plot are

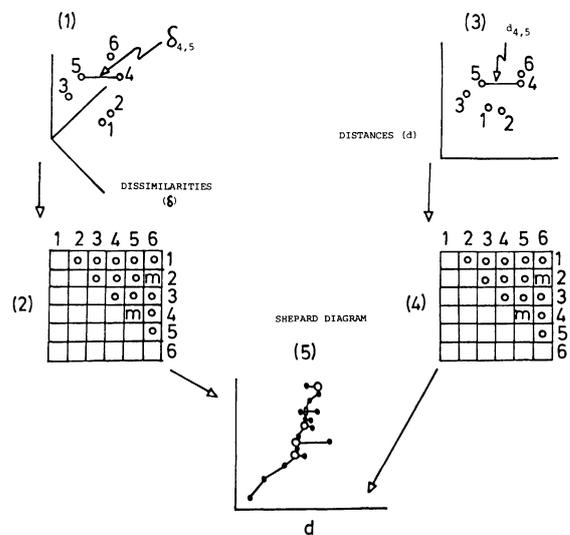


Fig. 6. Non-parametric multidimensional scaling. (1) dissimilarities, δ , in original space, between six points, example $\delta_{4,5}$; (2) matrix of δ 's. At least some values may be missing, m, or may be ignored; (3) 'new space' with smaller (or equal) number of dimensions to original space. There are the same number of distances, d , in new space as dissimilarities in original space; (4) matrix of d 's, missing values as in dissimilarity matrix; (5) Shepard diagram with δ plotted against d . The aim is to adjust the points in new space so that the Shepard diagram is monotonic, or, if that is impossible, to minimise the sum of squares of deviations (on the d axis) from a monotonic line. This means that the numerical values of δ are not needed, only their rank order.

commonly scaled in proportion to the square root of the eigenvector, so that curvature is less pronounced. An estimate of the size of the eigenvector in RA is easily got (Hill 1973) from the direct algorithm. The effect of using this scaling is shown in Fig. 12. Reciprocal averaging is relatively cheap to compute, needs little computer store, and may be used on very large sets of data.

Non-metric multidimensional scaling is not yet so well known, but has several advantages, both in theory and in practical tests (Anderson 1971, Dale 1975, Austin 1976, Prentice 1977 and Fasham 1977). It was devised (for use by psychologists) by Kruskal (1964) following earlier work by Shepard and others. Kruskal's account is a model of clarity, and is the basis of Fig. 6.

For simplicity, suppose that six points (sites, individuals, cases etc.) are characterised by three attributes (characters, species, variates etc.). At the top left of Fig. 6 these points are shown (1) in their three dimensional character space. A matrix, usually triangular, of proximity coefficients may be calculated (2). These are commonly dissimilarities, δ . Now suppose there is a 'new space' (3) but with a smaller number of dimensions. The same number of points – six – may be placed at random in this space, as shown at the top right of Fig. 6. Inter-point distances, d , may be calculated and arranged in a distance matrix (4). The rule for calculating d may be the same as that for δ , or it may be different. Now plot a graph (5) of δ against d . This graph is called the Shepard diagram. If a line, linking the points, can be drawn that runs always up and to the right (or always down and to the right) then the arrangement of points in the new space is, in this special sense, perfect. The line is monotonic: the rank order of d is identical to that of δ . If the rank order is not identical (as it is not in Fig. 6) then measure the imperfection by calculating a 'badness of fit':

$$\sum_{i < j} (d_{ij} - \hat{d}_{ij})^2 \quad (1)$$

where d_{ij} is the distance in new space between points i and j , and \hat{d}_{ij} is the point with the same δ on the current monotonic line. (The location of \hat{d} is defined later.) The badness of fit may be standardised, so that it is dimensionless and may be compared with the value in other sets of data, by dividing by $\sum (d_{ij}^2)$. This standardised 'badness of fit' measure is a squared quantity (c.f. variance). Taking the square root gives a standard-deviation-like measure of badness of fit called the stress, S :

$$S = \left[\frac{\sum_{i < j} (d_{ij} - \hat{d}_{ij})^2}{\sum_{i < j} (d_{ij})^2} \right]^{\frac{1}{2}} \quad (2)$$

The positions of d are defined as being those which, for a given position of points in new space, are monotonic and minimise S .

One or more points may now be moved in the new space, and a new S calculated. The process is repeated, and the positions of points in new space which best represent the points in the original higher dimensional space are those with the smallest stress S .

The 'best' arrangement must be found by an iterative method, though the work may be reduced by choosing starting positions in new space which are close to those for minimum stress. A more serious difficulty is that there is no guarantee that the solution is the absolute best: it may be a local rather than the global minimum. The results shown here are based on cases where at least three attempts from different starting positions have located very similar final positions in new space.

Three technical choices of important details were made:

(1) Primary treatment of ties was used: if $\delta_1 = \delta_2$ then the rank order of d_1 and d_2 does not matter.

(2) The standardisation used was that shown above, rather than

$$\sum_{i < j} (d_{ij} - \bar{d})^2,$$

where \bar{d} is the mean of all the d 's.

(3) Local rather than global scaling, in the sense used by Sibson (1972a), was preferred. This throws the emphasis on individual site relations.

(4) The distances in the new space were assumed to be Euclidean.

An essential feature of this method is that the values of δ_{ij} need never be known, only their rank order. In psychology, the results of experiments are often of this form: people record that they prefer one thing to another, so the rank order of results in the matrix of δ 's is known, but that is all. Psychologists call this matrix 'data', whereas ecologists usually calculate it from 'field data'. It is the discovery of metric coordinates in new space from non-metric (rank order) information in old space which has caused the method to become known as 'non-metric' MDS. As Orlóci (1978) points out however this may mislead ecologists, who first calculate what may be a metric proximity coefficient, whose metric properties they then ignore. To avoid this confusion it seems preferable to call the method 'non-parametric' or NP-MDS. Because only the rank order of δ 's is used, any transformation of δ should be without effect on the positions in new space. Coefficients of similarity may be used as easily as dissimilarities.

Apart from flexibility in choice of proximity coefficient

(which may be a mixed blessing) NP-MDS has three important and useful features:

(1) Some, perhaps many, of the δ 's may be 'missing'. This simply reduces the complexity of the Shepard diagram and reduces the amount of computing. Of course, if for any one site there are fewer δ 's than dimensions in the new space then the point will float in new space. Ecological data matrices do not usually have missing δ 's, but it may be useful to force some δ 's to vanish if only to reduce the amount of computing. This makes it practicable to use NP-MDS on larger data sets: present programs are limited to about 2000 δ 's, which would restrict one to about 45 sites for a full matrix or about 65 for a triangular matrix. If instead one requires only that there be, say, 15 δ 's for each site, then it becomes possible to deal with about 130 sites. An obvious way of removing δ 's is to choose first those which have no species in common, because it is just these for which the proximity measure will cause most distortion. Thinning can be extended to those sites with only 1, 2 . . . n species in common. If further thinning is necessary then three obvious ways of achieving it are to remove the smallest, largest or random δ 's, always preserving a minimum number of values for each site. Only the first stage of thinning of the Hummell Knowe data is illustrated. This again emphasises local relationships.

(2) There is no fixed functional relation between δ and d . As Kruskal (1964) shows NP-MDS may be used to recover the function even though the δ 's are nothing more than rank ordered. To the ecologist this lack of a fixed relation is potentially valuable because it makes the choice of proximity coefficient less important, or, more usefully, means that the coefficient may be chosen for reasons other than those dictated by the ordination method. In practice the scatter of points on the Shepard diagram (e.g. Fig. 15) is often close to a simple curve or even nearly a straight line (Fig. 11). In such cases polynomial MDS (PO-MDS) gives nearly the same result as NP-MDS, but with less computation. In PO-MDS however the *values* of the δ 's are used.

(3) The axes in new space are not ordered, so that the temptation to interpret one as 'the most important' and to try to equate the *axis* with variation in one particular environmental variable or complex can be more easily resisted. The arrangement of points in new space *can* be rotated to principal components, and this has been done in this work, but only to make comparison of diagrams simpler.

There remains the problem of how to choose the number

of dimensions for the 'new space'. Suppose separate analyses are made with 5, 4, . . . 1 dimensions. It may be that as the number of dimensions is decreased there comes a stage where the stress increases sharply. The Shepard diagram may at that point show a clearly branched pattern too. One may then take the number of dimensions needed to keep the stress low as being in some sense the natural dimensionality of the data. The whole matter is complex however (Kruskal & Wish 1977).

The last of the ordination methods used is principal coordinates analysis (Gower 1966, Krzanowski 1969). This will accept most proximity measures, and has sometimes been described as generalised PCA. It is more useful perhaps to consider PCA as a special case of PCO, using Pythagorean distance as the proximity measure. If the Shepard diagram of NP-MDS approximates a straight line then one might expect PCO to give much the same result as NP-MDS, but with much less computing; but whether or not the Shepard diagram *is* nearly straight will not be known until it is computed, though the occurrence in PCO of many or large negative latent roots may warn of non-linearity.

Proximity coefficients

For RA there is no choice. For PCA the choice between using the dispersion matrix or correlation matrix might be considered a choice of coefficient. Most authors who have tested PCA advocate use of the dispersion matrix so we used that.

For PCO the choice is wider, but as Williamson (1978) points out, incidence (presence-absence) data require the use of the Manhattan (or city-block) metric, m , which may be imagined as the distance along the edge of a hypercube:

$$m_{jk} = \sum_{i=1}^n |x_{ij} - x_{ik}| \quad (4)$$

where m_{jk} is the dissimilarity between individuals j and k , and x_{ij} is the value (0 or 1) of the i -th character (of n) of individual j . The site data at Hummell Knowe, after weighting, approached the character of incidence data, so this coefficient was used. It was tried on unweighted data too.

For NP-MDS the choice of coefficient is very wide. Prentice (1977) argues for the use of one particular coefficient (the sum of minima) because it matches the geometry assumed for distances in new space. But NP-MDS requires only the *rank order* of dissimilarities; and the results often, in practice, depend rather little on the choice of

coefficient. To test this point three coefficients were compared:

- (1) the Manhattan metric
- (2) another Minkowski metric, the 'Euclidean distance'

$$d_{jk} = \left[\sum_{i=1}^c (x_{ij} - x_{ik})^2 \right]^{\frac{1}{2}} \quad (5)$$

- (3) the non-metric Mountford coefficient, θ , defined by

$$e^{\theta a} + e^{\theta b} = e^{\theta(a+b-c)} + 1$$

where a is the number of species at site i , b is the number at site j , and c is the number of species common to both sites. This coefficient (Mountford 1962) is based on an explicit model of species abundance – the log series (Williams 1964) – and for that model the coefficient is independent of

sample size. This property is not particularly useful in the present case, but in other cases where there is need to compare records made from samples of different or unknown area it may be very useful (Holloway & Jardine 1968, Clymo and Régnell unpublished). It is important therefore to know which ordination methods may usefully be used with it. In the Hummel Knowe work θ was converted to a dissimilarity, ϕ , using:

$$\phi = \ln 2 - \theta$$

Modification of the proximity coefficients

The proximity matrix may be modified in various ways. If one starts from the assumption that the proximity measure

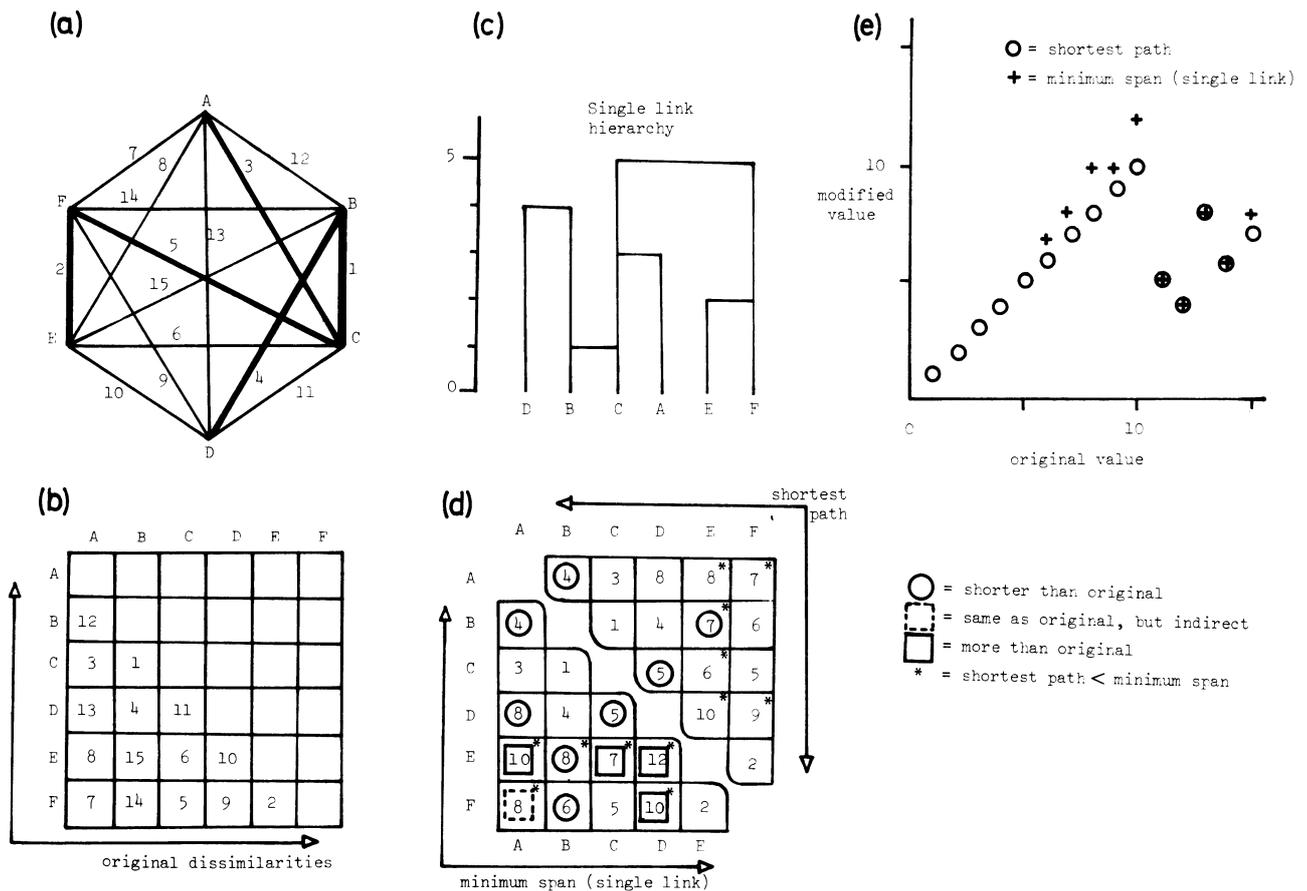


Fig. 7. Modifications to proximity coefficient matrix. (a) graph of network of proximity coefficients linking points A to F; (b) matrix representation of the network; (c) single-link hierarchy, found by using first the shortest link, then the next shortest, and so on; the corresponding structure – the minimum spanning tree – which is the shortest total length of links needed to interconnect all points, is shown by thick lines in (a); (d) single link (minimum spanning tree distance) in lower triangle; shortest path by any route, not just those of the minimum spanning tree, in upper triangle; (e) original and modified values by minimum spanning tree (single link) and shortest path.

is best known for closely similar sites, it becomes logical to try to replace the larger or dubious or unknown dissimilarities by measures compounded from smaller dissimilarities. One approach is to use the minimum (shortest) spanning tree, which is closely connected to the single-link (nearest neighbour) hierarchy. This is attractive because, there is a highly efficient algorithm, with cost a small multiple of n^2 , (Sibson 1972b) for computing single-link structure. Some of the modified elements are smaller than the original, some are unchanged, but some may be larger than the original: Fig. 7 shows an artificially extreme case of this. Some of the paths may have rather a large number of segments – DBCFE in the example. These extremes may be avoided using another approach: the shortest path (Fig. 7). The modified elements are always equal to or smaller than the original ones, and elements which are missing or which have been removed because they have few or no species in common are calculated directly from those which are left – provided there is no disjunction. (Dissimilarities such as the Manhattan and Euclidean distance are already the shortest paths, so this technique is only of use in such cases if some values are to be ignored.) This is the approach used by Williamson (1978). There is a very simple algorithm, shown in Fig. 8, but the computing time is related to n^3 . An intermediate approach not used here, is to replace by shortest paths, only those elements which are missing.

Essentially, as Fig. 7 shows, these approaches suppose that larger dissimilarities are over-estimated, or at least that a large dissimilarity is evidence of unreliability. Shortest path compensates to some extent, the minimum spanning tree to a greater extent.

Computation

Three programs were used for the bulk of the computing – DSD (manipulates and summarises data matrices, calculates RA and proximity coefficient matrices and modifies these); GENSTAT (a general statistical program, written at Rothamsted Experimental Station, and used here for PCA and PCO); and KYST2 (for NP-MDS and PO-MDS, from Bell Telephone Laboratories).

Results

Structure and stratigraphy

The general shape of the bog and the depth of the peat may

- Matrix **C**, elements C_{ij} for $i, j = 1, n$.
1. $C_{ii} = 0; i = 1, n$.
 2. $C_{jj} = \infty$ for missing or invalid elements
 3. $k = 0$
 4. $k = k + 1$
for all $i \neq k$ such that $C_{ik} \neq \infty$
for all $j \neq k$ such that $C_{kj} \neq \infty$
do $C_{ij} = \min [C_{ij}, (C_{ik} + C_{kj})]$
 5. Stop if
any $C_{ii} < 0$: no solution possible
 $k = n$: solved, C_{ij} are the shortest paths
else go to 4

Fig. 8. Shortest path algorithm (from Christofides 1975, after Floyd 1962).

be seen in Fig. 9. The surface is a roughly elliptical eccentric dome about 300 m by 600 m, the centre of which is ca 2 m higher than the edges. Around the perimeter the surface slopes more steeply into valleys to east, west and north. The highest measured point was at F4. The bog has formed in a basin, which may run out in a valley to the north. The southern edge seems to be particularly steep sided, with peat depth increasing from ca 3 m, to ca 10 m over not more than 100 m. The rocks of this region are part of the Carboniferous Limestone, and contain interbedded

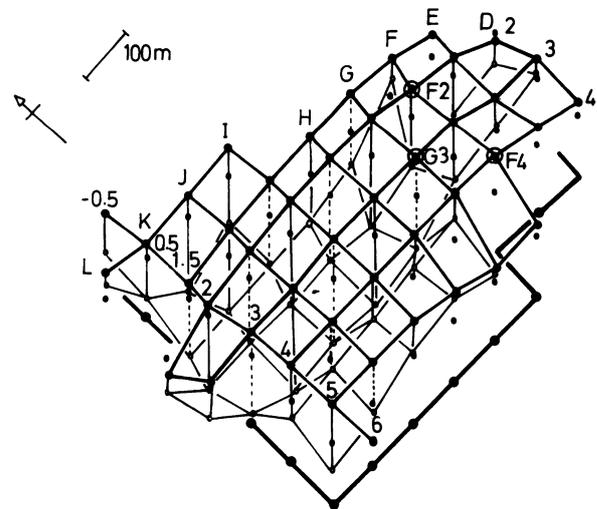


Fig. 9. Surface height and peat depth of Hummell Knowe Moss. The grid interval is 100 paces (except 50 paces between line 1.5 and 2). Surface is shown by large filled circles. Below (except at D2, L2) the small filled circles define a horizontal plane. Peat depth is shown by small unfilled circles – dashed lines indicate interpolated depths. The lower thick line defines the main survey area. Stratigraphy at F2; highest point at F4; greatest peat depth at G3.

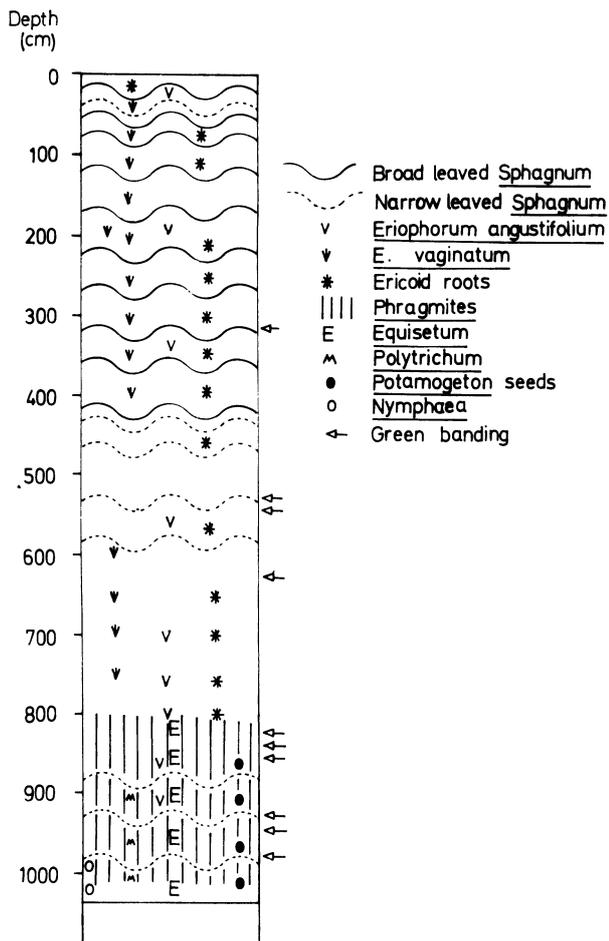


Fig. 10. Stratigraphy at site F2 (Fig. 9).

sandstones, the whole dipping gently to the south. Differential erosion has caused a north-south saw tooth profile, and it is presumably in one of these depressions that the peat has formed.

The deepest peat found was 10.7 m at G3. The stratigraphy was examined in the field at F2, where the peat is 10.4 m deep, and is shown in Fig. 10. Immediately above the basal clay is ca 2.5 m of *Phragmites* peat, with seeds of open water species such as *Nymphaea* and *Potamogeton*, and some remains of *Equisetum* and of a small leaved *Sphagnum* and of *Polytrichum*. Taken together these indicate a shallowing lake, probably tending to be acid. Above, there is a fairly sharp transition to amorphous peat with remains of *Eriophorum* and ericaceous roots. Nearer the surface the structure of *Sphagnum* is well preserved. The profile is remarkably similar to that at Coom Rigg, though no brushwood peat was found overlying the *Phragmites*. Clearly one needs more evidence before

concluding anything from the *absence* of a feature: it may have been overlooked. It is clear however that at least part of Hummell Knowe Moss was at one time a lake, and the Moss as a whole is probably therefore partly of a raised bog nature.

Surface-water chemistry

The chemistry of surface waters in September proved rather uniform. In the central area pH ranged from 3.9 to 4.4, and conductivity, after allowing for the contribution from H^+ , from 57 to 134 $\mu S\ cm^{-1}$. At the northern sites pH was higher – 5.5 to 6.0 – but conductivity was no greater, with the single exception of a sample from the lagg at the foot of Hummell Knowe itself, where conductivity was 730 $\mu S\ cm^{-1}$.

Vegetation

In general, the species and their abundance on the central bog plain are remarkably similar to those at Coom Rigg (Chapman 1964a). It is convenient to consider one analysis in some detail first, then to examine the effects, in turn, of transformations; of different methods of ordination; of modifications of the proximity coefficient matrix; and of different proximity coefficients. It becomes clear that there is a 'central' group of 32 sites, which are then re-analysed separately. Finally the species composition of sites in various regions of the ordinations is considered.

Before examining the results it is essential to set standards against which the usefulness of the methods may be assessed. This was done *explicitly*, in writing, before making the analyses. The following groups (Fig. 2) were expected:

- (1) D2, D3, E2 and perhaps E3, associated with erosion at the eastern side.
- (2) E5, F5, G5, I6, associated with the southern margin.
- (3) L4, similar to group (2) but with more species characteristic of the central bog plain.
- (4) Possibly K4, having some similarity to group (3), but not at all floristically distinct from the central bog plain. The real distinction was in the size of tussocks of *Eriophorum vaginatum*. Any useful method must reveal these patterns *and* it should suggest others.

First analysis

All 41 sites were used with hyperbolic weighting and Manhattan dissimilarity in an NP-MDS, shown in Fig. 11(a).

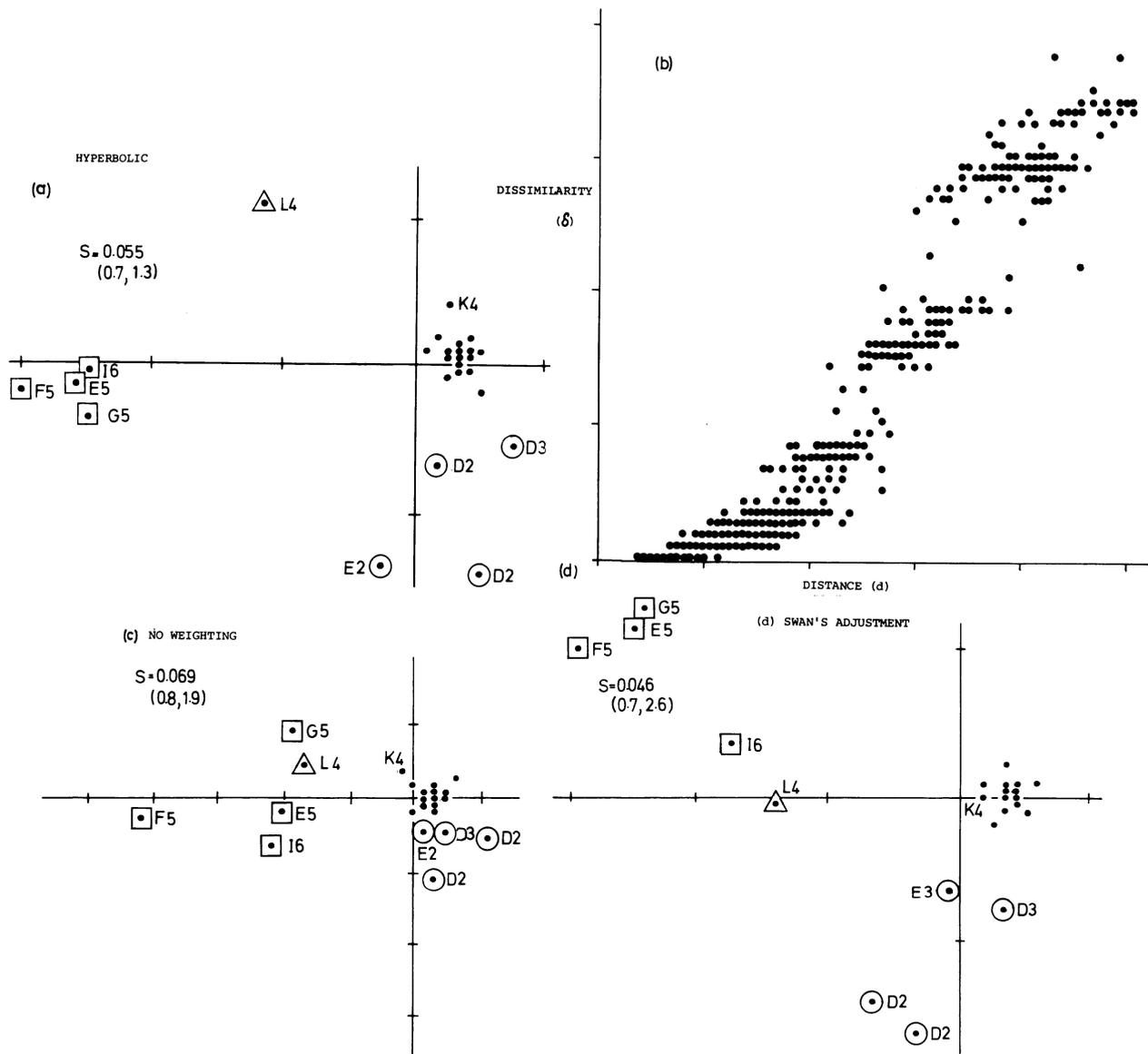


Fig. 11. Ordinations and Shepard diagram of all 41 sites at Hummell Knowe Moss, showing the effect of transformation. All use Manhattan dissimilarity. (a) hyperbolic weighting, NP-MDS; (b) Shepard diagram for (a); (c) no weighting; (d) Swan's (1970) adjustment for degree of absence. S = stress: 0.05 is low. Figures in parentheses are size of stress (relative to that in two dimensions) for three dimensions and for one dimension. In other Figs, with ordinations by RA, PCA and PCO, the values in parentheses are the proportion of variance accounted for by axes 1 to 3. Squares, triangles and circles identify groups of sites referred to in the text and mapped in Figs 2, 17.

The expected groups have appeared with embarrassing clarity: the southern margin group are the most distant from the origin; L4 is clearly separate but between the margin and centre; K4 is rather outside the central group, though were it not expected to be so, the significance of this would be overlooked; and the erosion group is also clearly distinct, though E3 is not separated. The 'replicates'

of D2 are as different as any pair in this group, but they do appear in the same general region.

The Shepard diagram, Fig. 11(b), is not far from linear, and the stress in two dimensions at 6% is low. Adding a third dimension (not illustrated) allows the erosion and marginal groups to differentiate, except for the replicates of D2, but the reduction in stress to 0.7 of that in two

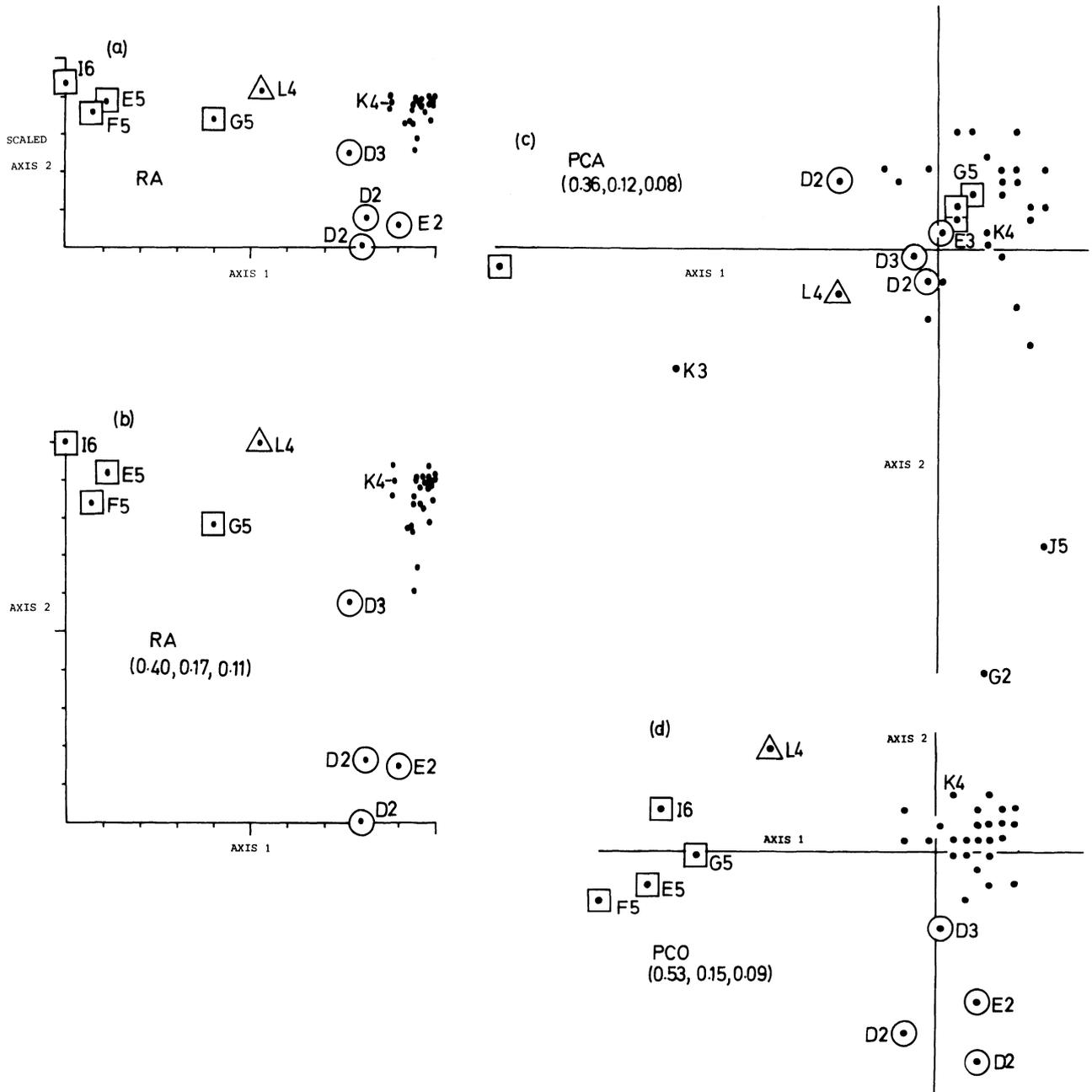


Fig. 12. Ordinations of all 41 sites at Hummell Knowe Moss, showing the effect of type of ordination. All use hyperbolic weighting. (a) RA, second axis scaled in proportion to square root of eigen vector; (b) RA, same results as (a), but second axis unscaled; (c) PCA, dispersion matrix; (d) PCO, Manhattan dissimilarity. Compare with Fig. 11 (a). Other conventions as in Fig. 11.

dimensions is not marked. Nor indeed is the increase in stress in the one dimension analysis.

Transformations

Using no transformation, Fig. 11(c), produces no marked

change from Fig. 11(a) in the sense that one can see the resemblance very clearly, but the *differences* are important in that if one had no prior expectation then interpretation would be much more awkward. Probably a single linear gradient would be suggested with F5 at one end and D2 at

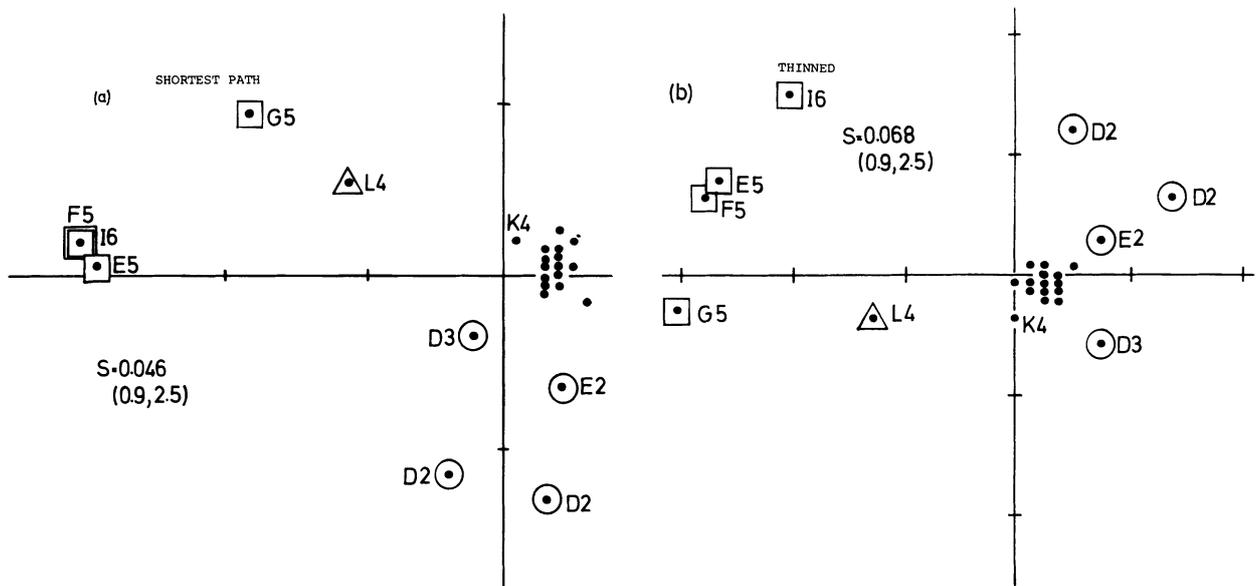


Fig. 13. NP-MDS ordinations of all 41 sites at Hummell Knowe Moss, showing the effect of modifying the dissimilarities. Both ordinations use hyperbolic weighting and Manhattan dissimilarity. (a) dissimilarities between sites with fewer than 8 species in common replaced by shortest paths; (b) data thinned by ignoring dissimilarities between sites with fewer than 8 species in common (this reduces the number of dissimilarities from 820 to 709). Compare with Fig. 11(a). Other conventions as in Fig. 11.

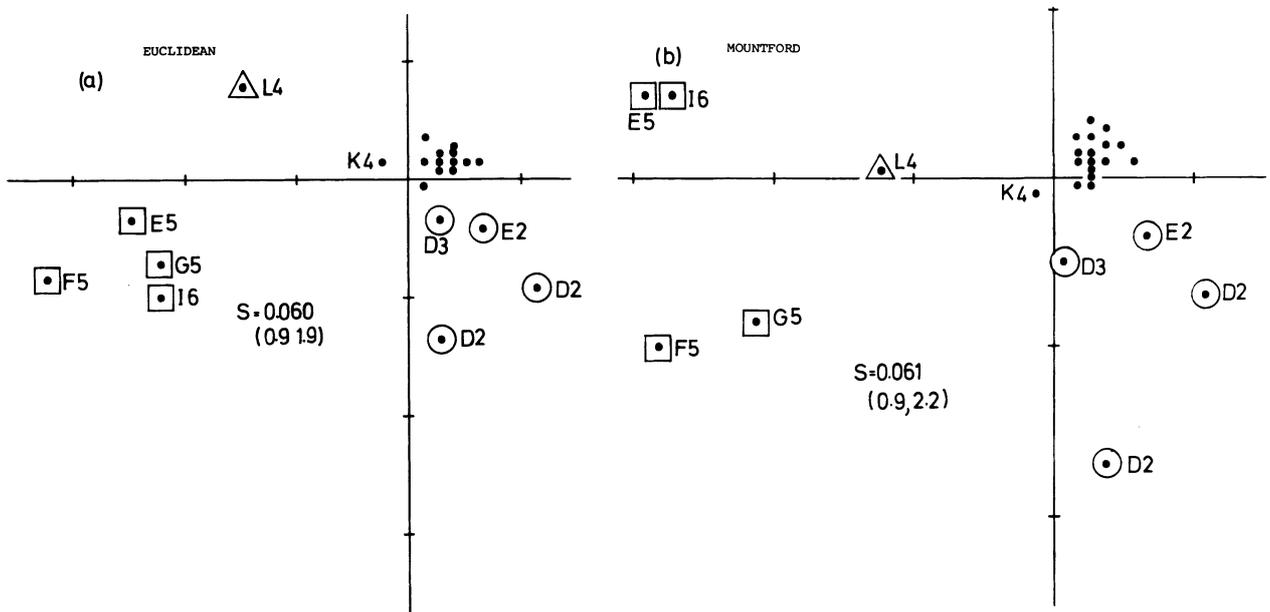


Fig. 14. NP-MDS ordinations of all 41 sites at Hummell Knowe Moss, showing the effect of different dissimilarity coefficients. (a) hyperbolic weighting, Euclidean distance; (b) modified Mountford's coefficient, ϕ (see text). Compare with Fig. 11(a). Other conventions as in Fig. 11.

the other. This is not ecological sense, and results from the small effect of the rarer species. The stress is about 7%, and it is clear that low stress *per se* is no guide to ecological usefulness (although high stress should always leave one dubious about an interpretation). Swan's adjustment produces a pattern, Fig. 11(d), which is arguably superior to Fig. 11(a) in that the grouping of the D2 'replicates' is closer, and that I6, which is physically between the marginal sites E5, F5, and G5 but closer to the south western lagg, is placed on the ordination diagram close to the intermediate L4 site. (But see the discussion of Fig. 14.) The stress is lower than in Fig. 11(a), and two dimensions give a considerable improvement over one, though three do not produce much further improvement.

Ordination methods

The results of analyses on hyperbolically weighted data are shown in Fig. 11(a) and Fig. 12. Reciprocal averaging produces a pattern similar to NP-MDS, but makes I6 an extreme and places G5 close to L4. The adjacent K4 is not an outlier of the central group. As with Fig. 11(c), the similarity to Fig. 11(a) is obvious, but interpretation of Fig. 12(a), (b) – the scaled and unscaled plots – would be more awkward than that of Fig. 11(a) or 11(d). The PCA, Fig. 12(c), accounts for 48% of the variance on the first two axes, but is ecologically quite misleading. The PCO Fig. 12(d), was made on the Manhattan dissimilarity, and produces results similar to that of NP-MDS, though K4 is less distinct. The first two axes account for 68% of the variance, and the analysis, given the proximity coefficient matrix as data, costs about 15% of the three NP-MDS ones.

Modifications of the proximity matrix

Again, hyperbolically weighted Manhattan dissimilarity was used, with NP-MDS, so the results in Fig. 13 may be compared with those in Fig. 11(a). No pair of sites had fewer than two species in common, so it is unlikely that shortest path calculations will have much effect. For illustration, it was assumed that the dissimilarity for site pairs with fewer than eight species in common was unknown, and this was calculated from the shortest path, Fig. 13(a). This may be compared with the result, Fig. 13(b), got when thinning the data by simply ignoring such cases (about 10% of the total). The shortest path result is very similar to Fig. 11(a). The thinned result has spread the erosion and marginal groups out, as might be expected,

because comparisons are now more dependent on relations between very similar sites.

Proximity coefficient

Again, NP-MDS of Manhattan dissimilarity, Fig. 11(a), may be compared with Euclidean distance, Fig. 14(a), and the modified Mountford coefficient, Fig. 14(b). Euclidean dissimilarity gives a very similar result to the Manhattan dissimilarity, as might be expected. The Mountford dissimilarity gives the now familiar pattern but it separates E5 and I6 from F5 and G5 which lie between them. This unexpected result proves to be caused by the occurrence of a few species such as *Potentilla erecta* and *Galium saxatile* in E5 and I6 and a few (different) species such as *Carex panicea* in the other pair. More interestingly, it draws attention to the greater abundance of *Polytrichum commune* in E5 and I6, and of *Sphagnum recurvum* in F5 and G5, a feature previously overlooked. These differences have obvious ecological interpretations.

The use of NP-MDS and PCO with the non-metric Mountford's dissimilarity are explored further in Fig. 15. The shortest path (for all data, whether or not they have species in common) is shown in Fig. 15(a). The separation of E5 and I6 from F5 and G5 is not obvious, though the order is still there. The site D3 is now less conspicuously separated: it is on the edge of the eroded area and still contains most of the central bog species, though they are infrequent. The Shepard diagram is markedly curved, and the stress is only 3%, with little change in three dimensions, but a big increase in one dimension. The effect of thinning, by ignoring all pairs for which the species list of one is identical to or a subset of the other is shown in Fig. 15(c). The 60% remaining pairs, now emphasising *differences* between sites, still give the familiar pattern, though as with the thinning in Fig. 13(b), the erosion sites are now less distinctly separated as a group.

The result with PCO, containing as it does two large negative latent roots, is, as expected, ecologically misleading. All the extreme sites, of all kinds are grouped together, but apparently equally extreme are F1.5, F2, G1.5, H1.5 at one side and H5, K2 at the other. That these groups are floristically similar is reasonable – the first group are adjacent, and the second group are in the same subjective vegetation zone (Fig. 2). The relation between these groups and the 'extreme' erosion and marginal groups is not simple however. The whole diagram seems, by comparison with all the others (except the unhelpful PCA) to be warped and differentially stretched.

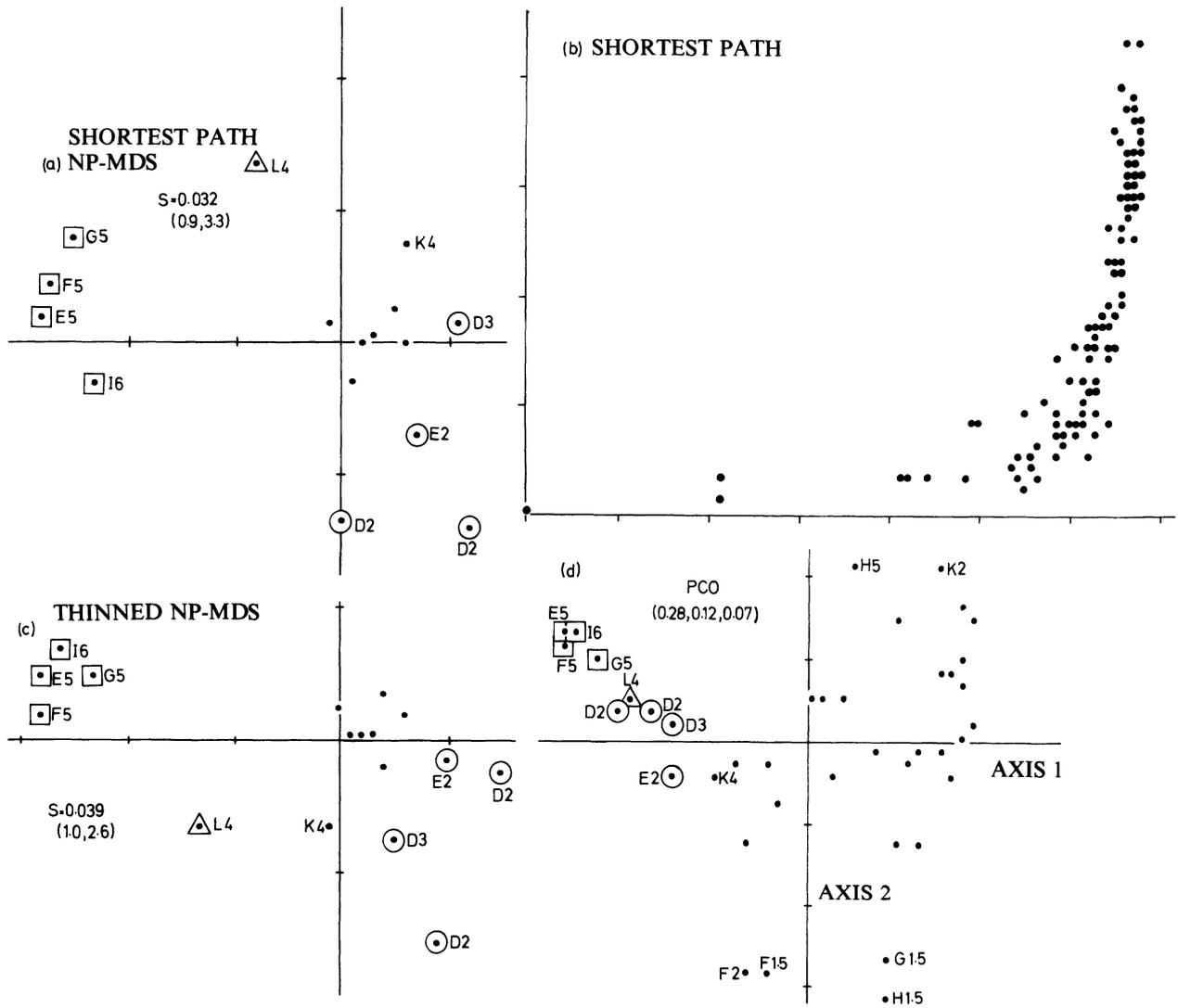


Fig. 15. Ordinations and Shepard diagram of all 41 sites at Hummell Knowe Moss, using the modified Mountford coefficient, ϕ (see text). (a) NP-MDS dissimilarities replaced by shortest paths, whether with species in common or not; (b) Shepard diagram for (a); (c) NP-MDS, ignoring dissimilarities between sites which have identical species lists, or between sites one of which is a subset of the other. (This reduces the number of dissimilarities from 820 to 495); (d) PCO. Compare with Fig. 11(a). Other conventions as in Fig. 11.

Analyses of the 32 'central' sites

So far the methods have in the main been tested for their ability to reveal the ecologically obvious. The next stage was to look for further unobvious patterns. To do this, the outlier erosion and marginal sites were removed (but leaving K4 which should therefore appear as an outlier in the new analyses) and the remaining 32 sites were re-analysed. The results of analyses of hyperbolically weighted data are shown in Fig. 16. One must expect random 'noise' to be more prominent once outliers have been

removed, and the stress of 17% for the NP-MDS (Manhattan dissimilarity) reflects this. Nevertheless the NP-MDS and PCO (same dissimilarity) gave a generally similar pattern, with the exception perhaps of J0.5, which had a relatively impoverished flora, though neither its stress (NP-MDS) nor residual (PCO) were notably large. The RA is considerably different, though the same outliers appear, and when plotted on a map of the bog surface, c.f. Fig. 17, produced no obvious patterns. Finally PCO on the shortest path replacements for pairs of sites with fewer than 12 species in common (emphasising comparisons of

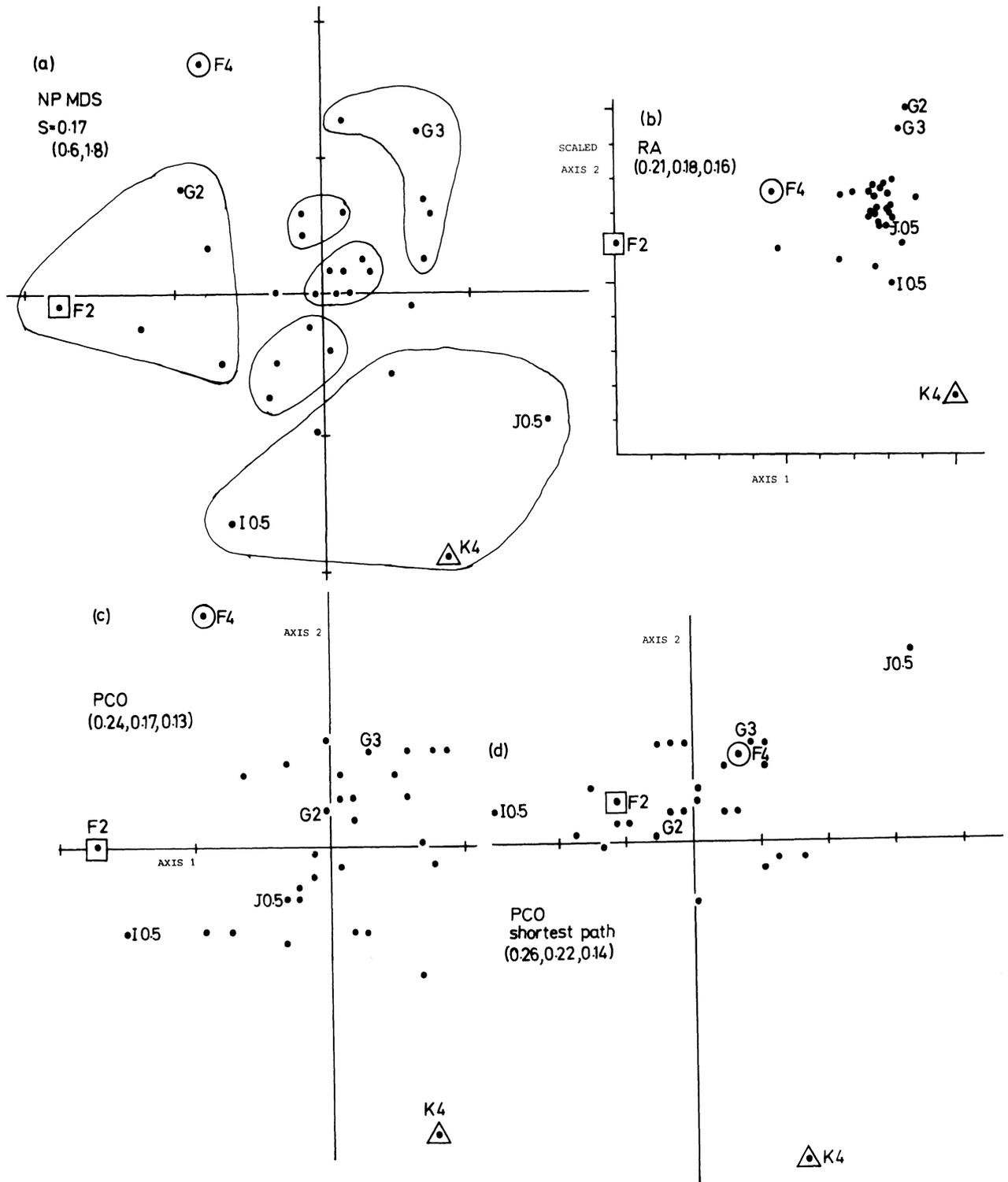


Fig. 16. Ordination of the 32 'central' sites at Hummell Knowe Moss. All use hyperbolic weighting, and all but (b) use Manhattan dissimilarity. (a) NP-MDS, subjective groups outlined for species composition in Fig. 18; (b) RA; (c) PCO; (d) PCO, dissimilarities between sites with fewer than 12 species in common replaced by shortest paths. Large symbols identify sites, or groups of sites, referred to in the text.

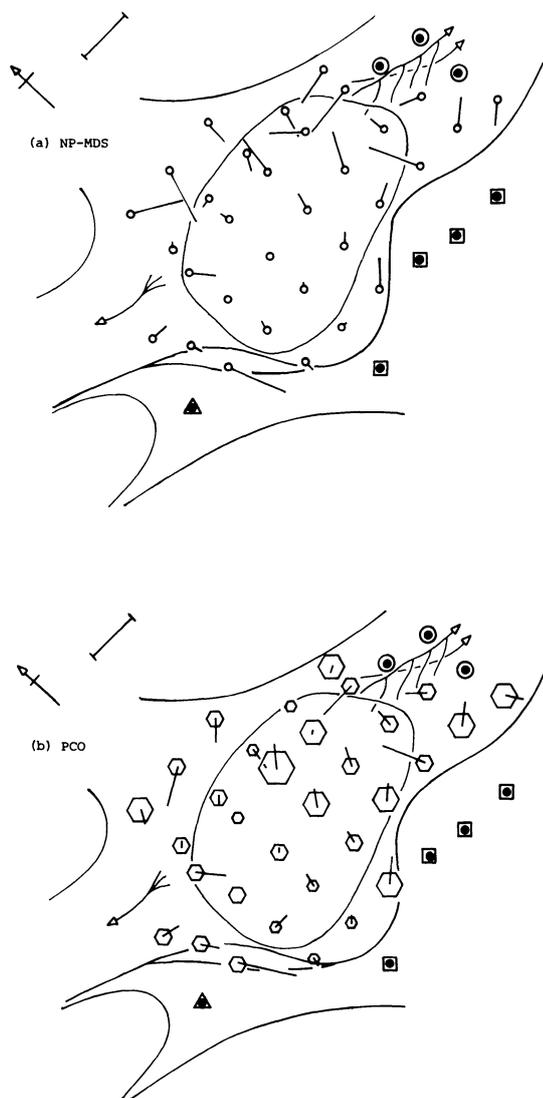


Fig. 17. Ordination position plotted on isodiometric map of surface of Hummell Knowe Moss. (a) NP-MDS of Fig. 16(a); (b) PCO of Fig. 16(c). The length and direction of line indicates the distance and direction from the origin of the ordination. Circles, squares and triangles as in Fig. 2. Hexagons in (b) show residual at that point: larger hexagon = larger residual.

closely similar sites) confirms the uncertainty about site J0.5, and pushes G2 and F2 toward the centre, indicating that their peripheral position in other analyses is owed to superficial resemblance to the peripheral sites.

The positions of the sites in the NP-MDS and PCO ordinations are plotted in Fig. 17 on a map of the bog surface. Tentatively one may identify for further investigation:

(1) An 'erosion' component pointing 'westerly' on the ordinations, and including D4, E3, F1.5, F2 and, surprisingly, G2. It is characterised by lower frequency, or loss, of species compared to the 'central' area, and is carried further in the 'erosion' group D2, D3, E2 already removed from the analysis. The species composition is shown in Fig. 18. Particularly noticeable are the reduction and eventual loss of *Sphagnum magellanicum*, *Drosera rotundifolia*, *Eriophorum angustifolium* and *Lepidozia setacea* (but not *Mylia* or *Odontoschisma sphagni*), and the gain or increase in *Sphagnum capillifolium* and lichens.

(2) A 'marginal' component pointing 'southerly' or 'south-easterly' on the ordinations, including site K4, and characterised by reduction in frequency of *Narthecium ossifragum*, *Sphagnum tenellum*, *Drosera rotundifolia*, *Trichophorum cespitosum* and *Lepidozia setacea*. The extreme form of this is seen in the marginal sites E5, F5, G5, I6 removed from the analysis, and having none of the first three species, but high frequency (Fig. 18) of *Sphagnum recurvum*, *Polytrichum commune*, *Molinia caerulea* and *Deschampsia flexuosa*.

(3) A 'northerly' or 'north-easterly' component including sites E4, F3, G3, G4, H2, H3, H5 and I2. This is characterised by reduced frequency of *Sphagnum tenellum* ('north-easterly') or *Calluna vulgaris* ('northerly'), small changes in frequency of leafy liverworts, and small increases in *Sphagnum capillifolium* and in lichens. This component may be a slightly drier variant of the 'central' bog type. It may be significant, however, that this component clashes with the 'southerly' one along the line of the fence, so differences in management may be involved.

Species ordinations

These produced two polar groups. One contained the species of the central bog plain, most of high frequency. The other contained species of the margins and erosion, of lower frequency. These ordinations revealed nothing new, so are not shown here.

Discussion

1. NP-MDS, PCO, and to a smaller extent RA, produce patterns which seem ecologically sensible.
2. The PCO and NP-MDS analyses of the 'central' 32 sites have produced interpretable patterns, and not simply random 'noise'.
3. These patterns were not obvious either on the spot, or in the collected results. Two of the three components seem

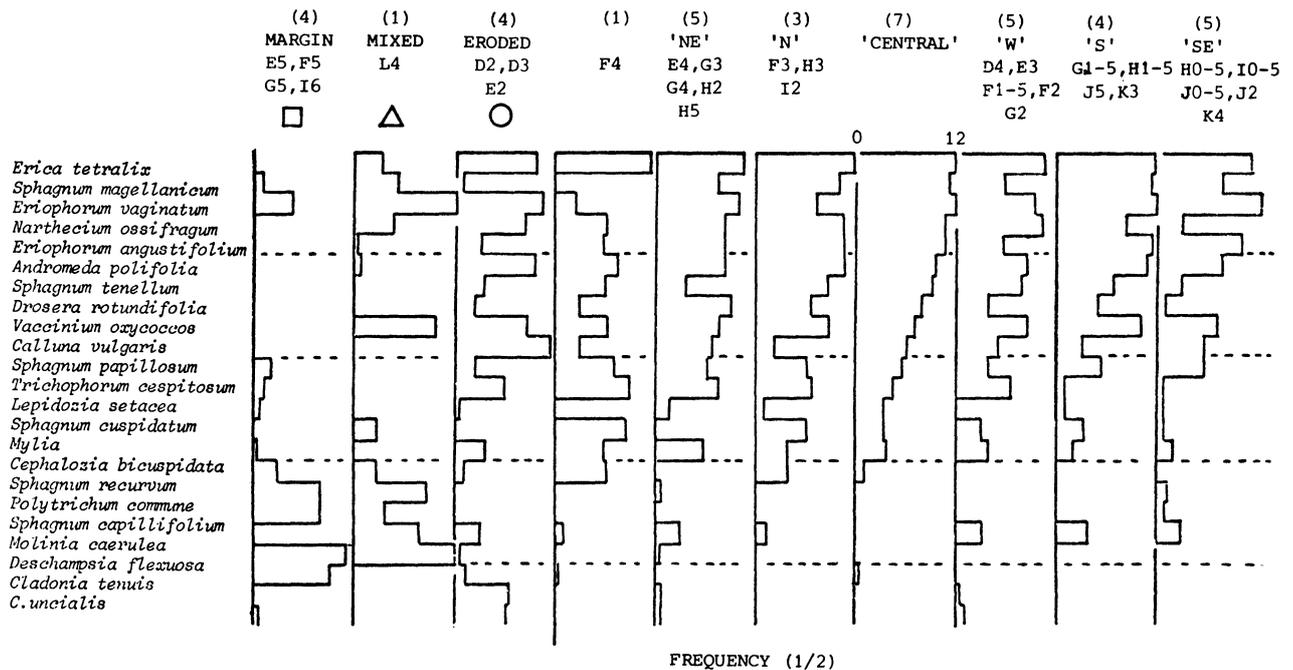


Fig. 18. Frequency of species at subjectively grouped sites. Figures in parentheses are number of sites in group. Leftmost three groups were separated in analysis of all 41 sites (e.g. Fig. 11 (a)). Others are shown in Fig. 16(a). Directions - 'NE', 'N' etc. - refer to position on the ordination diagram, not necessarily on the bog. Species are ordered by frequency in the 'central' group, except the last seven which are arranged to show associations in the first three groups. A complete species list is shown in Table 1.

to be less extreme variants of ones already obvious, but the third is new.

4. The differences between these groups are subtle and in some cases seem to point to particular species in an unexpected way.

5. It is obvious from Fig. 17 that the subjective boundary of the central bog plain is not at all well identified: the whole should be moved to the south-west and reduced at the north and east sides.

There is sufficient here to direct plans, and to that extent the work has succeeded.

Summary

(1) Hummell Knowe Moss is one of several fine and relatively undamaged peat bogs in northern England, close to the border with Scotland.

(2) The central, eccentrically domed, mass of peat is ca 300 x 600 m, and much of this is 7 to 10 m deep.

(3) A single profile has 8 m of bog peat overlying 2.5 m of *Phragmites* peat with seeds of *Potamogeton* and *Nymphaea*. The bog as a whole is probably part 'raised' and part 'blanket', as are others in the area.

(4) Much of the present surface is wet, with *Sphagnum magellanicum*, *Eriophorum* spp., *Andromeda polifolia* etc. There are some eroded areas with less *Sphagnum* and more abundant fruticose lichens (*Cladonia* spp.), and marginal areas with *Molinia caerulea*, *Deschampsia flexuosa*, *Sphagnum recurvum* and *Polytrichum commune*.

(5) Numerical analyses show the importance of data transformation. Of the ordination methods tried, PCA produced unhelpful results, but RA, PCO and NP-MDS were all adequate. PCO can take a variety of dissimilarity measures, but not all produce useful results. NP-MDS is more tolerant, and can also be used to adjust the relative importance of inter-pair dissimilarities in a more flexible way than can PCO.

(6) Further analyses by RA after removal of outliers were not of great use, but those by PCO and NP-MDS revealed clear patterns. Two of the groups of sites were less clearly developed versions of the erosion and marginal vegetation types. The third type may be developed in slightly drier conditions, or may in some places result from differences in management.

(7) The usefulness of the methods in this survey was NP-MDS > PCO > RA > PCA. Computing cost of

NP-MDS, PCO and RA parallels usefulness: more useful, greater cost.

References

- Anderson, A.J.B., 1971. Ordination methods in ecology. *J. Ecol.* 59: 713–726.
- Austin, M.P., 1976. Performance of four ordination techniques assuming three different non-linear species response models. *Vegetatio* 33: 43–49.
- Chapman, S.B., 1964a. The ecology of Coom Rigg Moss, Northumberland. I. Stratigraphy and present vegetation. *J. Ecol.* 52: 299–313.
- Chapman, S.B., 1964b. The ecology of Coom Rigg Moss, Northumberland. II. The chemistry of peat profiles and the development of the bog system. *J. Ecol.* 52: 315–351.
- Chapman, S.B., 1965. The ecology of Coom Rigg Moss, Northumberland. III. Some water relations of the bog system. *J. Ecol.* 53: 371–384.
- Christofides, N., 1975. Graph theory, an algorithmic approach. Academic Press, New York, 400 pp.
- Clapham, A.R., T.G. Tutin & E.F. Warburg, 1962. Flora of the British Isles. 2nd edn. Cambridge University Press. Cambridge, 1591 pp.
- Dale, M.B., 1975. On objectives of methods of ordination. *Vegetatio* 30: 15–32.
- Duncan, U.K. with P.W. James, 1970. Introduction to British Lichens. Buncle, Arbroath, 292 pp.
- Fasham, M.J.R., 1977. A comparison of nonmetric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines and coenoplanes. *Ecology* 58: 551–561.
- Floyd, R.W., 1962. Algorithm 97 – shortest path. *Comm. Assoc. Comput. Mach.* 5: 345.
- Gower, J.C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53: 325–338.
- Hill, M.O., 1973. Reciprocal averaging: an eigenvector method of ordination. *J. Ecol.* 61: 237–249.
- Holloway, J.D. & N. Jardine, 1968. Two approaches to zoogeography: a study based on the distributions of butterflies, birds and bats in the Indo-Australian area. *Proc. Linn. Soc. Lond.*, 179: 153–188.
- Jensén, S. 1978. Influences of transformation of cover values on classification and ordination of lake vegetation. *Vegetatio* 37: 19–31.
- Krzanowski, W.J., 1969. The algebraic basis of classical multivariate methods. *Statistician*. 20 (4): 51–61.
- Kruskal, J.B., 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*. 29: 1–27.
- Kruskal, J.B. & M. Wish, 1977. Multidimensional scaling. Sage Publications. Beverley Hills, California.
- Maarel, E. van der. 1979. Transformation of cover-abundance values in phytosociology and its effects on community similarity, *Vegetatio* 39: 97–114.
- Mountford, M.D., 1962. An index of similarity and its application to classificatory problems. In: P.W. Murphy (ed): *Progress in Soil Zoology*, p. 43–50. Butterworths. London.
- Noy-Meir, I., 1971. Multivariate analysis of the semi-arid vegetation in south eastern Australia. *Proc. Ecol. Soc. Austral.* 6: 159–193.
- Noy-Meir, I., 1973. Data transformations in ecological ordination. I. Some advantages of non-centering. *J. Ecol.* 61: 329–341.
- Noy-Meir, I., D. Walker & W.T. Williams, 1975. Data transformations in ecological ordinations. II. On the meaning of data standardization. *J. Ecol.* 63: 779–800.
- Orlóci, L., 1978. Multivariate analysis in vegetation research. 2nd edn. Junk, The Hague, VIII + 450 pp.
- Pearson, M.C., 1960. Muckle Moss, Northumberland. *J. Ecol.* 48: 647–666.
- Prentice, I.C., 1977. Non-metric ordination methods in ecology. *J. Ecol.* 65: 85–99.
- Ratcliffe, D.A. (ed), 1977. Nature conservation review. Vol. 2. Site Accounts, 320 pp.
- Sibson, R., 1972a. Order invariant methods for data analysis. *J. Roy. Statist. Soc. B.* 34: 311–349.
- Sibson, R., 1972b. SLINK: an optimally efficient algorithm for the single-link cluster method. *Computer J.* 16: 30–34.
- Smartt, P.F.M., S.E. Meacock & J.M. Lambert. 1976. Investigations into the properties of quantitative vegetational data. II. Further data type comparisons. *J. Ecol.* 64: 41–78.
- Smith, A.J.E., 1978. The Moss Flora of Britain and Ireland. Cambridge University Press. Cambridge.
- Swan, J.M.A., 1970. An examination of some ordination problems by use of simulated vegetational data. *Ecol.* 51: 89–102.
- Watson, E.V., 1968. British Mosses and Liverworts. 2nd edn. University Press. Cambridge.
- West, R.G., 1968. Pleistocene geology and biology. Longmans. London.
- Williams, C.B. 1964. Patterns in the balance of nature. Academic Press. London & New York.
- Williamson, M.H., 1978. The ordination of incidence data. *J. Ecol.* 66: 911–920.

Accepted 14 December 1979