Freshwater Biology (1984) **14**, 221–256

A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data

J. F. WRIGHT, D. MOSS,* P. D. ARMITAGE and M. T. FURSE
Freshwater Biological Association, River Laboratory, East Stoke, Wareham, Dorset, and
*Institute of Terrestrial Ecology, Bangor Research Station, Bangor, Gwynedd

SUMMARY. 1. Macro-invertebrate species lists were obtained for 268 sites on forty-one river systems throughout Great Britain by qualitative sampling in spring, summer and autumn. Information on twenty-eight environmental variables was also collated for each site. The sites were ordinated on the basis of their species content using detrended correspondence analysis (DCA) and classified by two-way indicator species analysis (TWINSPAN). Correlation coefficients between ordination scores and single environmental variables indicated that Axis 1 distinguished between types of rivers and Axis 2 reflected variation along the length of rivers. A preliminary classification of sites into sixteen groups has been proposed, together with a key which allows new sites to be classified. Information on the species and environmental features which characterize each group is also presented.

2. Multiple discriminant analysis (MDA) was employed to predict the group membership of the 268 sites using the twenty-eight environmental variables. 76.1% of sites were classified correctly. An independent assessment of predictive ability using forty test sites yielded a 50% success rate. Predictive ability was higher for the classification presented in this paper than in fifteen additional classifications produced using data from single seasons and/or different taxonomic treatments.

3. TWINSPAN and MDA were found to be useful approaches to the classification of running-water sites by their macro-invertebrate fauna and the prediction of community type (as indicated by the occurrence of species in the sites comprising the group) using environmental variables. Extension of the scope of the classification, coupled with the use of additional environmental variables to increase predictive ability, is now in progress.

Introduction

This paper presents the results of the first phase of a project on British rivers which has two major objectives: (i) development of a biological classification of unpolluted running-water sites in Great Britain based on the macro-invertebrate fauna; and (ii) assessment of whether the type of macro-invertebrate community at a site may be predicted using physical and chemical features.

Correspondence: Dr J. F. Wright, Freshwater Biological Association, River Laboratory, East Stoke, Wareham, Dorset BH20 6BB.

Illies & Botosaneanu (1963) and Hawkes (1975) review previous attempts at classifying running-water sites, river zones and entire rivers. Many of the early classifications were developed in Europe and North America for fishery management purposes and divided rivers into zones on the basis of their dominant fish species (Huet, 1949, 1954; Burton & Odum, 1945). However, studies on the invertebrate communities of rivers whose fish fauna was well documented gave conflicting evidence on whether the invertebrate communities were zoned in harmony with the fish (Illies, 1953; Schmitz, 1957), or whether there were gradual transitions in the invertebrate fauna downstream (Maitland, 1966). Illies (1961) suggested a world-wide river-zone classification based on physical criteria and was able to associate the occurrence of particular families of invertebrates with the major zones in his classification. Pennak (1971) also proposed a river-zone classification scheme which could be applied throughout the world, based on thirteen physical, chemical and floristic features, but this proposal has not yet been taken up.

For a classification of British running-water sites based on macro-invertebrates to be of practical value, it must be broad enough to encompass the range of sites which may need to be classified in future, provide an acceptable number of classes and operate at the level of identification of the fauna required by the user. Clearly, different users will have different requirements. The Nature Conservancy Council, with their statutory obligations to conserve good examples of different rivers in Great Britain (Ratcliffe, 1977) and their concern for the species which comprise aquatic communities, may find a classification based on species-level identification of the fauna of greatest value in making decisions on the scheduling of rivers. Likewise, there are occasions when Water Authorities and River Purification Boards may require a species-level classification. However, a classification based on less intensive assessment of the fauna (e.g. presence or absence of particular families of invertebrates) may be of more general application within the water industry where rapid appraisal and classification of sites are required.

Reliable prediction of the type of invertebrate community to be expected at a site, using physical and chemical criteria, would be of considerable value. For example, where rivers are polluted, an ability to predict the type of community if the rivers were unpolluted would be of use as a baseline against which to assess attempts to improve water quality.

The value of both classification and prediction has been apparent for some time (Hawkes, 1975; Peters, 1976; Persoone, 1979). However, projects of this type which require the simultaneous collation of biological, physical and chemical data for a large number of sites, pose major problems with regard to consistency of the data-set. Nevertheless Resh & Unzicker (1975) have argued the case for studies of this type. In view of the limited and better-known invertebrate fauna of British rivers compared with mainland Europe, the availability of physical and chemical data for a wide range of rivers, and advances in multivariate statistical techniques for the ordination and classification of large data-sets, the time appears to be right for a classification exercise in Great Britain.

In this paper, a single classification of running-water sites in Great Britain is presented, based on species lists of macro-invertebrates obtained after sampling in three seasons. We believe that this classification is more reliable than fifteen additional classifications obtained from the same data-base using single-season data and/or different levels of identification. All the classifications are compared by Furse *et al.* (1984) who also examine comparative ability at predicting the fauna from environmental attributes. Armitage *et al.* (1983) examine the behaviour of the Biological Monitoring Working Party (BMWP) score system for water-quality assessment (Chesters, 1980) when applied to the wide range of unpolluted sites available in this project.

Design of sampling programme

Few, if any, major rivers in Great Britain have not been affected to some degree by Man's activities. Foremost amongst changes in the fauna are those due to domestic and industrial effluents and with this in mind the sampling programme was restricted to rivers which, by and large, were free from serious pollution. This ensured that the classification would be based on essentially 'natural' groupings of species.

At the outset, decisions were made on the number of sites which could be examined, how they should be spaced along each river and the frequency with which samples should be taken. Many aquatic invertebrates are seasonal in their occurrence, and it was considered essential to sample in three seasons (spring, summer and autumn) to obtain adequate species information for classification purposes. Verneaux (1976), who examined 240 sites on thirty French river systems, found that the rate of change of invertebrate community composition was greatest nearer the source, and therefore sampling locations were proposed at 5, 10, 20, 30 and 40 km from the source and then at successive 20 km intervals downstream as shown on 1:250,000 O.S. maps. The time available for sample processing indicated that approximately 275 sites could be examined in detail for three seasons.

To overcome the problem of sampling a large number of rivers throughout Great Britain concurrently, the biological sampling programme was planned as a joint effort between the water industry and the Freshwater Biological Association (FBA). With the help of the individual Water Authorities (WA) and River Purification Boards (RPB), a list of approximately 100 rivers was compiled. These were relatively free from pollution, well-documented physically and chemically and were being, or could be, sampled by the relevant WA/RPB. Ultimately, forty-one rivers (+ nine additional tributaries) were chosen for the sampling programme, providing an adequate coverage of most parts of Britain and a wide range of physical and chemical conditions. Choice of specific sites on each river followed the guidelines already stated but, on the advice of the WA/RPB biologists who were familiar with each river, slight adjustments were frequently made to ensure that the sites could be adequately sampled and would encompass the major sections of each river. The 267 locations chosen for sampling on the forty-one river systems are shown in Fig. 1.

Water Authority biologists in England and Wales and Purification Board biologists in Scotland obtained samples of macro-invertebrates at these sites following guidelines provided by the FBA team. All samples were sent to us for sorting and identification of the fauna. The Water Authorities and River Purification Boards also supplied much of the physical data, and all of the chemical data, relating to each biological sampling site. Collation, validation and analysis of the biological, physical and chemical data were carried out by the FBA team in collaboration with the Institute of Terrestrial Ecology, Bangor.

Methods

Data collection

General. Each site comprised a sampling area for macro-invertebrates in the centre of a survey area, whose dimensions were approximately 50 m or 7 channel widths upstream and downstream of the sampling area, whichever distance was the smaller. In each of the three seasons whenever a macro-invertebrate sample was taken in the sampling area, physical data relating to this limited area were also recorded. In one season only (normally summer) physical data were obtained for the survey area as a whole with the intention of recording additional features of the site which would show limited variation with time of year.

Invertebrates—field procedures. The main objective of the invertebrate sampling programme was to obtain the most comprehensive species list for each site, within the limits imposed by the time available. The standard FBA pond-net (900 μm mesh, 230×255 mm frame, 275 mm bag depth) on a 1.5 m handle was recommended for obtaining a qualitative sample of the fauna. Water Authority biologists were requested to use a kick and sweep technique (Furse *et al.*, 1981) over all the major habitats for a total of approximately 3 min within the sampling area which, as far as possible, was typical of the survey area as a whole.

Where the character of the river changed rapidly (e.g. pool and riffle system) sampling was usually confined to the area routinely sampled by the WA/RPB biologists, normally the riffle. However, at three locations which had pool–riffle systems (one on the River Wansbeck and two on the River Forth) biologists took separate pool and riffle samples in each season and these were kept separate throughout the analyses. Hence, 270 sites were sampled, encompassing 267 locations (Fig. 1).

At twelve deep-water sites it was not possible to obtain a representative species list by use of a pond-net and alternative methods were used

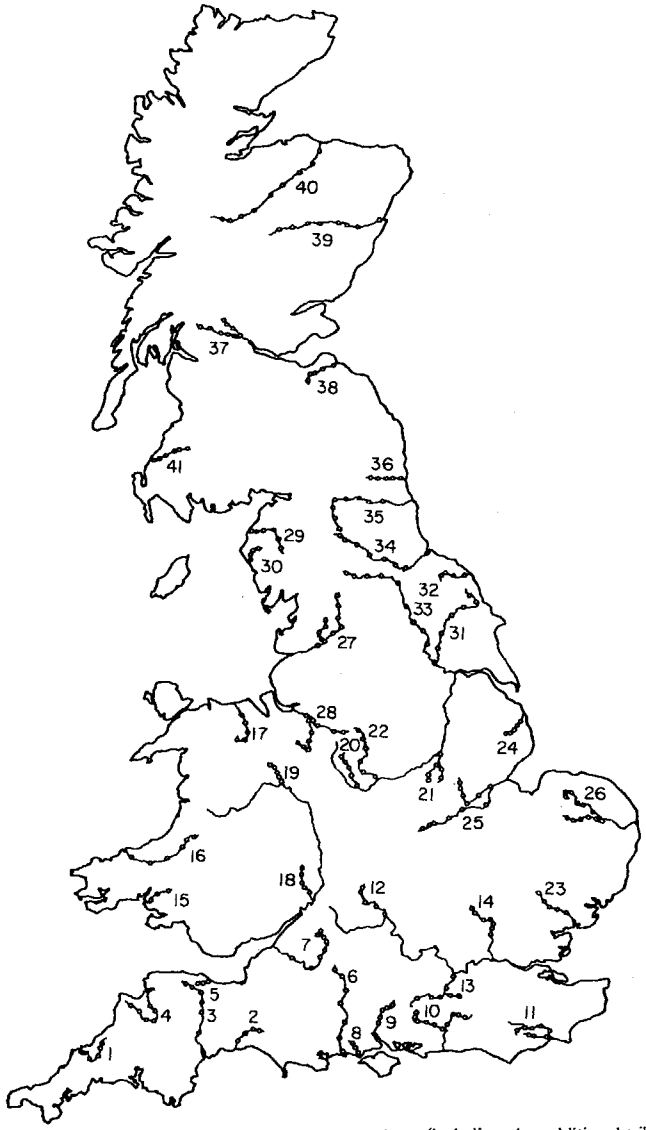

FIG. 1. The 267 locations chosen for sampling on forty-one rivers (including nine additional tributaries) are shown as circles. 1, Camel; 2, Axe; 3, Exe; 4, Torridge; 5, Avill; 6, Avon (Hants); 7, Avon (Bristol); 8, Avonwater; 9, Itchen; 10, Arun and Rother (Sussex); 11, Rother (Kent) and Brede; 12, Evenlode; 13, Wey and Tillingbourne; 14, Lee; 15, Gwendraeth Fach; 16, Teifi; 17, Clwyd; 18, Leadon; 19, Perry; 20, Blithe; 21, Devon and Smite; 22, Dove; 23, Colne; 24, Great Eau; 25, Welland and Glen; 26, Yare and Wensum; 27, Ribble and Hodder; 28, Weaver and Dane; 29, Derwent (NWWA); 30, Ehen; 31, Derwent (YWA); 32, Esk; 33, Ouse; 34, Tees; 35, Tyne (NWA); 36, Wansbeck; 37, Forth and Teith; 38, Tyne (FRPB); 39, Dee; 40, Spey; 41, Stinchar.

instead of, or to supplement, pond-net collections. Air-lift samples were taken at eight sites on the Yorkshire Derwent and Ure, grab samples at two sites on the Forth, and core samples at two sites on the Spey.

Biologists were requested to reduce the volume of the samples by removal of large stones, silt or detritus, if this could be achieved without substantial loss of animals. Samples were preserved in 5% formalin or 70% alcohol before being sent to the FBA. The macro-invertebrate sampling programme took place between spring 1978 and spring 1980.

Invertebrates—laboratory procedures. Each sample was carefully examined for approximately 2 h by placing aliquots in a white flat-bottomed tray and sorting through the material by eye. Since many macro-invertebrates cannot be identified without a binocular or compound microscope, large numbers of specimens were removed to ensure that most, if not all, of the species present in the sample were available for identification. Furse *et al.* (1981) provide additional details of sample processing and of the procedures used to estimate the approximate abundance of families of invertebrates in each sample for use in additional classification exercises.

Identifications were made using the available keys (Armitage, Furse & Wright, 1979). Most macro-invertebrates, except for many dipteran families, could be identified to species. In cases where closely-related species or genera could not be distinguished consistently, they were placed together as a 'group'. A list of the taxa in each 'group' is available from the authors on request.

Identification of the macro-invertebrate fauna was carried out by a team of five scientists at the FBA, Wareham, and this was followed by two substantial checking procedures to ensure both accuracy of identification and consistency in the level of identification. In combining the species lists for three seasons to produce a single listing for the analyses, there was some danger of duplication in the representation of taxa where identification could not be taken to species level in all seasons. Thus, if *Limnodrilus hoffmeisteri* Claparède was identified in one season but only immature and unidentifiable tubificids in another season, the Family Tubificidae was excluded from the full species list since the specimens could have been *L. hoffmeisteri*. The eight-digit code developed by Maitland (1977) for freshwater animals occurring in the British Isles was used as the method for transferring data to the computer prior to the major analyses.

Physical data. Physical data to complement the species lists for each site were obtained from a variety of sources. Table 1 lists all environmental variables used in later analyses together with brief notes on procedures for data collection and a list of abbreviations used in later sections. The distance of each site from the source of the river and the slope of each site were determined from O.S. maps and the Water Data Unit provided further information on the discharge (average daily flow) at each site. The survey area sheet gave information on a number of features, of which altitude, mean channel width, water depth and a measure of the substratum variability were used in later analyses.

All the remaining fifteen physical variables were obtained from the sample data sheets. Some variables utilized information collected over the three seasons, usually in the form of mean values, but maximum and minimum values from the three seasons were also used, as was date of sampling. The variables obtained for the sampling site at the time of sampling included measures of surface velocity of the water, substratum type, dominant particle size and macrophyte cover, but water depth and width were also recorded for comparison with the complementary measures of depth and width recorded once only on the survey area sheets. This provided an opportunity to assess whether the sampling area was typical of the survey area and also if there was information gain in obtaining records on three separate occasions or whether a single record relevant to the entire survey area was of greater value.

Inevitably, in attempting to acquire physical data on twenty-two variables for 270 sites, some values were missing; these were estimated by reference to information available for other seasons, nearby sites or after discussion with WA/RPB biologists. Comprehensive temperature records could not be obtained for all sites and hence water temperature is not included in the analyses.

Chemical data. Of the 270 biological sampling sites only 137 had coincident chemical sampling programs operated by the WA/RPB chemists. A further thirty-one sites were within 2 km of a chemical sampling site, fifty-one more were less

TABLE 1. The twenty-eight environmental variables used in the analyses together with their abbreviations and brief notes on data collection procedures. Variables whose abbreviations are prefixed L were $\log_{10}$ transformed.

Environmental variable	Abbreviation	Units of measurement or no. categories	Footnote
Information from maps			
Distance of site from source	LKM	km	1
Slope of site	LSLOPE	m km^{-1}	1
Altitude of site	LALT	m	2
Discharge category for site	DISCH	9 categories	3
Information from survey area sheet (completed once)			
Mean channel width of survey area	LMEANW	m	4
Depth category of survey area	DEPTHC	5 categories	5
Substratum heterogeneity in survey area	SUBSTNO	1–7	6
Information from sample data sheet (completed in three seasons)			
Date of sampling – spring	DAY 1	1–365	
Date of sampling – summer	DAY 2	1–365	
Date of sampling – autumn	DAY 3	1–365	
Mean width of water in sample area	LMWIDTH	m	7
Mean depth of water in sample area	LMDEPTH	cm	8
Maximum surface velocity of water in sample area	MAXVEL	5 categories	9
Minimum surface velocity of water in sample area	MINVEL	5 categories	10
Modal/median surface velocity of water in sample area	MEDVEL	5 categories	11
Mean substratum in sample area	MSUBST	phi	12
Minimum dominant particle size in sample area	MINDOM	7 categories	13
Maximum dominant particle size in sample area	MAXDOM	7 categories	14
Modal/median dominant particle size in sample area	MEDDOM	7 categories	15
Maximum percentage macrophyte cover in sample area	MAXMAC	%	16
Minimum percentage macrophyte cover in sample area	MINMAC	%	17
Mean percentage macrophyte cover in sample area	MEANMAC	%	18
Chemical data provided by the water industry			
pH	pH		19
Dissolved oxygen	O2B	mg O_2 l^{-1}	19
Total oxidized nitrogen (Nitrate + Nitrite)	LTON	mg N l^{-1}	19
Chloride	LCL	mg Cl l^{-1}	19
Dissolved orthophosphate	LORPH	mg P l^{-1}	19
Total alkalinity	ALK	mg $CaCO_3$ l^{-1}	19

1. Obtained from 1:250,000 O.S. maps.
2. Obtained from maps by WA/RPB biologists.
3. Provided by Water Data Unit as average daily flow. $1 \leq 0.31$ cumecs; $2 \leq 0.62$; $3 \leq 1.25$; $4 \leq 2.5$; $5 \leq 5.0$; $6 \leq 10.0$; $7 \leq 20.0$; $8 \leq 40.0$; $9 \leq 80.0$.
4. Normally assessed using occurrence of bankside vegetation, etc.
5. Depth in over 50% of survey area. 1, <25 cm; 2, <50; 3, <100; 4, <200; 5, ≥200.
6. No. distinct areas dominated by a different substratum type: clay, silt, sand, gravel; >gravel; bedrock; macrophyte.
7. Mean width from three sampling dates.
8. Mean of nine depth readings: ¼, ½ and ¾ width × three sampling dates.
9, 10, 11. Maximum, minimum and modal/median categories respectively from three sampling dates. Categories: $1 \leq 10$ cm s^{-1}; $2 \leq 25$; $3 \leq 50$; $4 \leq 100$; $5 > 100$.
12. Mean phi values from three sampling dates weighted by % composition of the substratum in each season and based on four phi values estimated by eye: boulders/cobbles (phi −7.75); pebbles/gravel (−3.25); sand (2.0); silt/clay (8.0).

than 10 km away but no chemical samples were available for the remaining fifty-one sites.

Chemical data were obtained for all available determinands as annual means, maximum and minimum values, standard deviations and the number of samples. Data on thirty-three determinands, not all of which were available for each site, were transferred to computer. A detailed examination of the available data with the objective of assessing the importance of different determinands led to the selection of just six for use in the major analyses: pH, dissolved oxygen, total oxidized nitrogen (nitrate + nitrite), chloride, dissolved orthophosphate and total alkalinity. Estimates of the six determinands were then obtained for all sites which lacked data, using information from the nearest relevant site on the river or seeking the advice of the WA/RPB chemist.

Data analysis

Ordination and classification procedures. Ordination of the sites was carried out using detrended correspondence analysis (DCA) (Hill & Gauch, 1980) which is an updated version of reciprocal averaging (RA) (Hill, 1973). RA has been successfully used for the ordination of mire communities (Daniels, 1978) and compares favourably with principal components analysis and polar ordination (Gauch, Whittaker & Wentworth, 1977; Culp & Davies, 1980). However, DCA avoids the 'arch' or 'horseshoe' effect and compression of axis ends which occurs in RA and is an improvement on both RA and non-metric multidimensional scaling (Hill & Gauch, 1980; Gauch, Whittaker & Singer, 1981). DCA was implemented as a FORTRAN program called DECORANA (Hill, 1979a) and an option to downweight rare species in proportion to their frequency was used to avoid the tendency of individual samples with rare species to distort the analysis (Hill, 1979a).

Sites were classified using two-way indicator species analysis, TWINSPAN (Hill, 1979b), an improved version of indicator species analysis (Hill, Bunce & Shaw, 1975). TWINSPAN, the FORTRAN program used to carry out two-way indicator species analysis, classifies both samples and species and constructs ordered two-way tables to exhibit the relationship between them as clearly as possible. The program also constructs a key to the sample classification by identifying one or more 'differential' species which are particularly diagnostic of each division in the classification. The key can therefore be used to classify new sites without the need to reclassify all sites. TWINSPAN has recently been compared with a number of alternative hierarchical classification procedures (both agglomerative and divisive) by Gauch & Whittaker (1981). They conclude that TWINSPAN is usually the best general purpose method, especially when the data-set is complex, noisy, large or unfamiliar. They argue strongly in favour of the use of a divisive strategy in which the analysis may be stopped after a limited number of divisions (e.g. 2, 4, 8, 16 or 32 groups generated). This approach allows the division of site groups to be terminated at a level considered appropriate by the investigator and requires less computation than do agglomerative techniques.

Correlation of site groupings with environmental features. Multiple discriminant analysis (MDA) (=canonical variate analysis) was used to relate the site groupings to the environmental data. The SPSS version of MDA was used (Klecka, 1975) and all twenty-eight environmental variables listed in Table 1 were taken together, in either arithmetic or transformed versions, to generate the discriminant functions. This follows the recommendation of Green & Vascotto (1978) who advise against the use of step-wise MDA. The maximum number of discriminant functions which may be derived is either one less than the number of groups generated in the classification or equal to the number of discriminating variables, whichever is the smaller. The number of discriminant functions used in practice was determined by testing the significance of discriminating information

13, 14, 15. Minimum, maximum and modal/median categories respectively from three sampling dates.

7 phi categories:	−9	−6.5	−4.5	−2	2	6.5	9.5
Range of phi values:	−10, −9, −8	−7, −6	−5, −4	−3, −2 −1	0, 1, 2, 3, 4	5, 6, 7, 8	9, 10
Name of particle:	boulders	cobbles	pebbles	gravel	sand	silt	clay

16, 17, 18. Maximum, minimum and mean % cover respectively from three sampling dates.
19. Mean of all determinations available for one year.

not already accounted for by the earlier functions. Testing was achieved by transforming Wilk's lambda into a chi-squared statistic (Klecka, 1975). This technique has been criticized by Harris (1975), who recommended the greatest characteristic root test; however, this only tests the significance of the first function. At present Wilk's lambda, or partitioned U-test, is probably the best guide to the number of functions to derive. Only those functions statistically significant at $P \leq 0.05$ were used. The numbers of sites in each TWINSPAN group were known from the sample classification and these probabilities of group membership were used in allocating sites to TWINSPAN groups using the environmental data. The ordinations, classifications and multiple discriminant analyses were all carried out on the Cambridge University IBM 370/165 computer.

Results

An early TWINSPAN analysis, using spring data only, indicated that one site in the lower reaches of the River Rother in Kent (Site 64 in Table 5) occupied an extreme position in the ordinations and was placed in a group by itself at Level 3 (eight groups). This was because the site had a higher concentration of chloride than any other site owing to tidal influence, and its fauna reflected this fact. The lowest site on the River Weaver (Site 180) was found to have a very limited fauna (Table 5), being downstream of a sewage works and industrial effluent discharges. These sites were excluded from the analyses, as suggested for extreme outliers by Hill & Gauch (1980), because they were also outside the specification of sites for the project, i.e. unpolluted freshwater sites. This left a data-set which included a total of 599 taxa recorded across 268 sites (265 locations). All subsequent analyses reported in this paper were performed on the combined seasons' (spring, summer and autumn) species lists for each site.

Ordination

Fig. 2 presents an Axis 1 by Axis 2 DCA ordination plot of the samples from the 268 sites with the position of each site indicated by the TWINSPAN group in which it occurred. Further variation in the data-set could be displayed using Axes 3 and 4, but much of the major variation is shown on these first two axes and particularly along Axis 1. The eigenvalues can be thought of as an expression of the variance accounted for by each axis and are as follows: Axis 1=0.283; Axis 2=0.078; Axis 3=0.054; Axis 4=0.046.

Correlation coefficients between the ordination scores for Axes 1–4 and the environmental variables listed in Table 1 are given in Table 2. Variables which were not normally distributed were subjected to $\log_{10}$ transformation prior to the calculation of correlation coefficients. On Axis 1 the highest correlations observed were those with mean substratum (MSUBST: $r=0.779$) and alkalinity (ALK: $r=0.739$). In contrast, the highest correlations on Axis 2 were found with variables which relate to distance downstream, i.e. log of distance from source (LKM: $r=-0.781$) and discharge category (DISCH: $r=-0.740$). It would appear that, whereas variation along Axis 1 may display the variation between different types of rivers, Axis 2 relates more to position along the length of the river. Correlations between environmental variables and Axes 3 and 4 were much lower than those observed on Axes 1 and 2, the highest correlations with Axes 3 and 4 being log of total oxidized nitrogen (LTON: $r=-0.431$) and discharge (DISCH: $r=0.302$) respectively.

A summary of the range of Axis 1 scores recorded both within and between the forty-one rivers and their nine additional tributaries can be seen in Table 3. The ordination scores have been divided into eleven equal-sized classes, and the rivers ranked according to the lowest site class present. Most rivers have a limited range of Axis 1 scores, which supports the view that this axis separates types of rivers. Rivers in the top half of the table (i.e. one or more sites having Axis 1 scores of 100 or less) occur in Scotland, northern England (including the Dane and Dove which drain the southern Pennines), Wales and the south-west. Rivers in the bottom half of the table (no sites with Axis 1 scores of 100 or less) occur in the Midlands, East Anglia, the south and the south-east. Only two rivers occupy slightly anomalous positions in relation to this crude division. The Avonwater in Hampshire and the Yorkshire Derwent are both associated with the upper half of the table because of the upland character of their top sites, which flow through the New Forest and off the North York Moors respectively. Although some 'spatey' rivers of the north and west included occasional slow-

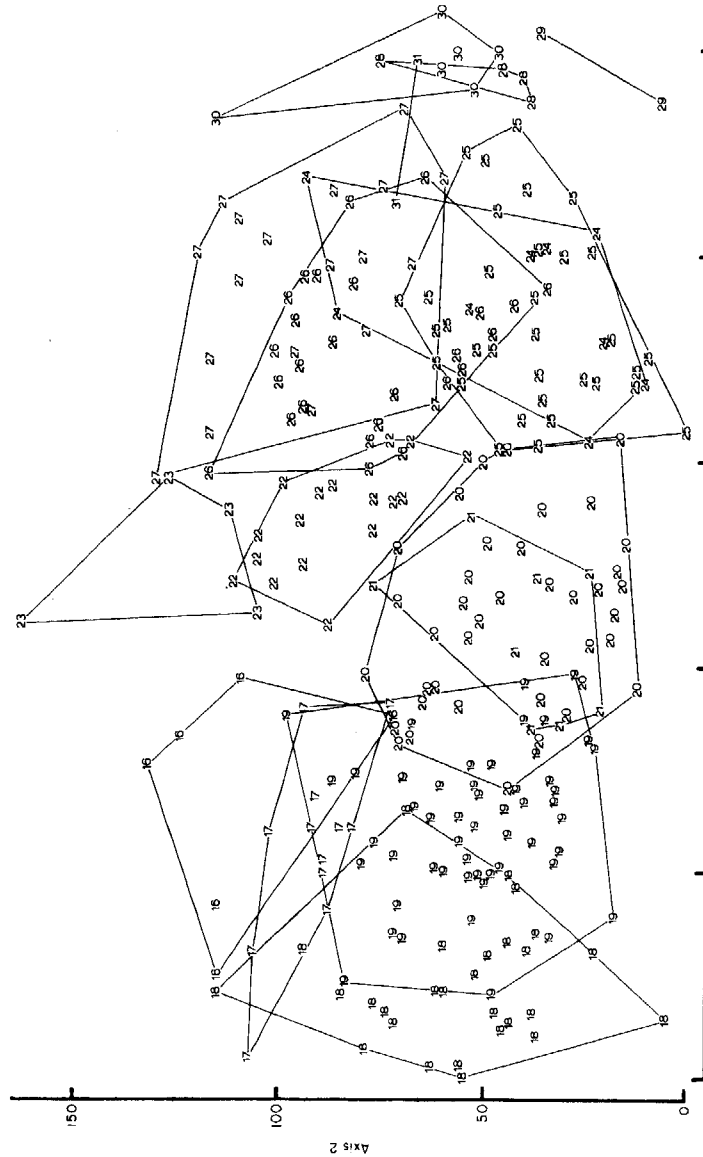

FIG. 2. DCA ordination plot (Axis 1 *v.* Axis 2) of 268 sites. Sites are numbered by the Level 4 TWINSPAN group in which they occurred. Polygons enclose all sites in each TWINSPAN group.

TABLE 2. Correlation coefficients between ordination scores for Axes 1–4 and environmental variables. Table 1 provides further information on each variable.

Variable	Axis 1	Axis 2	Axis 3	Axis 4
LKM	0.051	−0.781	0.095	0.242
LSLOPE	−0.611	0.503	0.049	−0.135
LALT	−0.571	0.341	0.052	−0.089
DISCH	−0.281	−0.740	0.259	0.302
LMEANW	−0.246	−0.695	0.267	0.202
DEPTHC	0.312	−0.427	0.273	0.166
SUBSTNO	0.248	0.031	−0.067	−0.024
DAY1	−0.133	−0.026	−0.086	0.083
DAY2	−0.009	−0.013	−0.176	0.001
DAY3	−0.019	0.036	−0.193	−0.059
LMWIDTH	−0.230	−0.676	0.269	0.268
LMDEPTH	0.298	−0.475	0.357	0.144
MAXVEL	−0.407	−0.059	−0.254	0.131
MINVEL	−0.507	−0.173	−0.195	−0.091
MEDVEL	−0.462	−0.177	−0.250	−0.116
MSUBST	0.779	0.070	0.137	−0.078
MINDOM	0.695	0.057	0.153	−0.057
MAXDOM	0.725	0.060	0.125	−0.074
MEDDOM	0.719	0.076	0.108	−0.089
MAXMAC	0.531	0.016	−0.354	−0.175
MINMAC	0.295	−0.069	−0.226	−0.127
MEANMAC	0.499	−0.033	−0.328	−0.178
pH	0.430	0.037	−0.372	−0.160
O2B	−0.364	−0.260	0.072	−0.067
LTON	0.718	0.204	−0.431	−0.148
LCL	0.688	0.169	−0.409	0.135
LORPH	0.623	0.086	−0.416	0.003
ALK	0.739	0.156	−0.356	−0.252

flowing sites in their middle and lower reaches which were more typical of the south and east, sites of upland character did not occur along the course of lowland rivers. Examples of the former include Site 259 (Table 5) on the River Spey (Letter *a* of Table 3) at the outflow of the Insh Marshes and Site 88 on the River Teifi at Tregaron Bog (b), together with the three samples taken from pools in the River Wansbeck (c) and River Forth (d). The influence of brackish water (e), through saline intrusion, as seen in the Dane, or through occasional tidal influence as on the Rother and Brede, also draws sites to the right-hand side of the ordination. However, the chloride concentrations and Axis 1 scores for these three sites were not so extreme as to warrant their exclusion from the classification. Of the fifty rivers listed, the Swale/Ure/Ouse system shows the greatest range of Axis 1 scores. Although the river retains its upland character at the six sites on the Swale, its confluence with the Ure and later with the Nidd results in the lower three sites having a substantially different invertebrate community, which is reflected in the ordination score on Axis 1.

Classification

Since two-way indicator species analysis is a divisive technique, the division of the samples into progressively smaller groupings may be continued for as long as seems profitable. Fig. 3 presents a dendrogram of the classification produced by TWINSPAN to Level 4, when sixteen groups of sites had been generated. There was considerable variation in the number of samples contributing to each group and from Level 2 onwards, Group 7 and the groups generated from it were all small.

Fig. 2 shows the limits of the sixteen TWINSPAN groups on the Axis 1 by Axis 2 ordination plot. In practice, TWINSPAN makes use of more than the two planes presented here since, after each division, the samples in each of the two groups are ordinated again before the

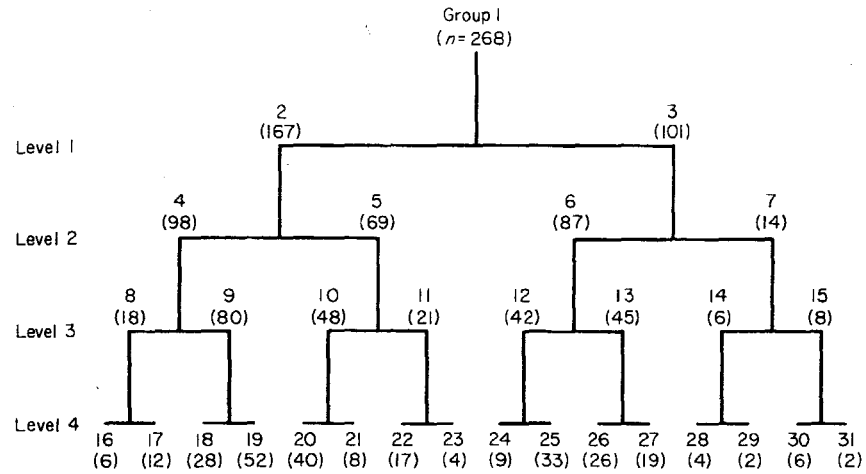

FIG. 3. Classification of the sites by TWINSPAN to Level 4. The number of sites in each group is indicated, in parentheses, below each group number.

next division takes place. Nevertheless, some of the separation between the sixteen groups can be demonstrated. In very general terms, the change from upland to lowland sites occurs from left to right and the upstream to downstream sequence from top to bottom of the figure.

In Table 4 we present an abbreviated version of a species × TWINSPAN group matrix which lists only those taxa which occurred with a frequency of 80% or more in at least one group. Of the 599 taxa used to produce the classification, this table presents just 130, and an indication of their frequency of occurrence in each group is shown using four frequency categories. The percentage frequency of occurrence in the entire data-set (268 samples) is also given so that the widespread species may be distinguished from those which occur commonly in only one or two groups. The sequence for the species list was obtained directly from the TWINSPAN ordered two-way table. Most taxa are identified to species but some, particularly Diptera, are identified to genus or species 'groups' and in four cases family identifications only are available (Hydridae, Lumbricidae, Enchytraeidae and Ceratopogonidae).

Some of the groups, genera and families (e.g. *Thienemannimyia* gp, *Polypedilum* sp., *Micropsectra/Tanytarsus*, *Orthocladius/Cricotopus*, Enchytraeidae and Ceratopogonidae) have a wide occurrence across the TWINSPAN groups at Category 4 (frequency of occurrence >75–100%) simply because they include a number of different species which, individually, are likely to have a more restricted occurrence. Plecoptera, and many Ephemeroptera, Trichoptera and Diptera are characteristic of the 'upland' groups (left-hand side of Table 4), whereas Oligochaeta, Gastropoda and additional species/genera of Diptera dominate the lowland side, along with many other major groups, including Tricladida, Hirudinea, Crustacea and Hemiptera. Of the sixteen TWINSPAN groups, most have twenty or more taxa listed in Table 4 which occur in over 75% of the samples (i.e. Category 4). However, Groups 27 and 24, with just seventeen and eighteen taxa respectively in Category 4 are lower, and Group 27 in particular includes a number of species which are common in organically enriched sites. From a total of 268 sites used in the classification, only seventeen had Biological Monitoring Working Party (BMWP) scores (Chesters, 1980) in the range 50–100, as calculated on the combined season family list. All other sites had a score exceeding 100. Seven of the scores in the range 50–100

TABLE 3. Range of Axis 1 ordination scores recorded at 268 sites on the forty-one main rivers plus nine additional tributaries included in the study. See text for explanation of letters a–e.

River system	Name	Axis 1 ordination score										
		0–25	26–50	51–75	76–100	101–125	126–150	151–175	176–200	201–225	226–250	251–275
39	Dee	5	3									
1	Camel	1	1	2								
27B	Hodder	1		3								
5	Avill	1		2								
35	Tyne (Northumbrian WA)	4		3	1							
41	Stinchar	1	3	1	1							
3	Exe	1	2	2	1							
34	Tees	1	2	2	1	1						
29	Derwent (North West WA)	1		3	1	1						
27A	Ribble	1		2	2	2						
40	Spey	3	4	1					1[a]			
33	Ouse	2	2		1	1			1	2		
30	Ehen		2	2								
17	Clwyd		2		3							
38	Tyne (Forth RPB)		1		3	1						
37A	Forth		1		1	3	1	1[d]	1[d]			
28B	Dane		1			2		1[e]				
37B	Teith			3								
32	Esk			3	1							
4	Torridge			1	5							
22	Dove			4	1	1						
16	Teifi			1	3	1	1[b]					
36	Wansbeck			1	4			1[c]				
15	Gwendraeth Fach				2	1						
8	Avonwater				1		2					
31	Derwent (Yorkshire WA)				1			2	4			
2	Axe					2	2					
13B	Tillingbourne					1	1					
10B	Rother (Sussex)					4	1	1				
20	Blithe					2	2	1				
6	Avon (Hants)					1		6				
24	Great Eau					1	1	1		1	1	
11A	Rother (Kent)					1	1	1			1[e]	
12	Evenlode						2	2	1			
9	Itchen						1	3	1			
19	Perry						1	1	2	1		
26B	Wensum						1	1	1	3		
11B	Brede						2			1[e]		
7	Avon (Bristol)						3	1	2	1	1	
21A	Devon						1		1		1	
25B	Glen						1	1		1	2	1
13A	Wey							4	1			
10A	Arun							1	1	2		
21B	Smite							1	1	1		
18	Leadon							3		2		
23	Colne							1	1	1	2	
28A	Weaver							2	3		1	
26A	Yare							1		4		1
14	Lee								5		2	
25A	Welland								4		2	

TABLE 4. The 130 'species' which occur with a frequency of 80% or more in at least one TWINSPAN group. Their frequency of occurrence in every TWINSPAN group is categorized as follows: 1 = >0–25%; 2 = >25–50%; 3 = >50–75%; 4 = >75–100%. The percentage frequency of each species in the entire data-set (268 samples) is also given.

'Species'	Percentage frequency in entire data-set	TWINSPAN groups (and number of sites per group)															
		16 (6)	17 (12)	18 (28)	19 (52)	20 (40)	21 (8)	22 (17)	23 (4)	24 (9)	25 (33)	26 (26)	27 (19)	28 (4)	29 (2)	30 (6)	31 (2)
Leuctra inermis Kempny	20.5	3	3	4	2	1											
Leuctra hippopus (Kempny)	13.1	4	1	2	1	1	1										
Protonemura meyeri (Pictet)	23.9	4	3	3	2	1	2	1									
Amphinemura sulcicollis (Stephens)	40.7	4	4	4	4	2	2	1	1				1				
Chloroperla torrentium (Pictet)	28.4	4	3	4	2	1	3	1	1		1						
Glossosoma sp.	36.6	1	3	3	4	2	3										
Plectrocnemia conspersa (Curtis)	4.5	4		1	1			1	2								
Hydropsyche instabilis (Curtis)	8.2	2	4	1	1	1						1					
Simulium monticola gp	29.5	3	4			2	2										
Baetis muticus (L.)	40.3	2	4	3	3	3	3	2				1					
Rhithrogena semicolorata gp	50.0	2	4	4	4	3	4	2				1	1				
Ecdyonurus sp.	52.6	3	4	4	4	4	4	2			1						
Leuctra fusca (L.)	57.5	4	4	3	4	4	4	4	3	1	1	1	1				
Isoperla grammatica (Poda)	51.1	3	4	4	4	3	4	2	3	1	1	1	1				
Hydraena gracilis Germar	32.1	4	4	2	3		4	1									
Esolus parallelepipedus (Müller)	48.5	3	4	4	4	3	4	1		2	1	1					
Sericostoma personatum (Spence)	40.3	4	4	2	3	2	4	3	2	1	1		1				
Atherix ibis (Fabricius)	35.8		2		3	3	4	1			1						
Baetis niger (L.)	10.1	2		1	1	1	4	2			1						
Taeniopteryx nebulosa (L.)	23.1		2	1	1	3	4	1	2	1	1	1					
Nais alpina Sperber	44.8	3	2	4	4	2	4	2		1	1	2	1			2	
Sperchon hibernicus gp	19.4		1	1	2	2		1	4		1	1	1				
Caenis rivulorum Eaton	50.7	3	4	3	4	4	3	3		2	1	1					
Rhyacophila dorsalis (Curtis)	63.1	4	4	4	4	4	3	4	4	1	2	2					
Agapetus sp.	44.0	2	4	2	3	3	2	3	1		1	1	1				
Hydropsyche siltalai Döhler	69.0	3	4	4	4	4	3	4	1	1	2	3	1				
Eloeophila sp.	18.3	4		1	1	1	2	2	2	1	1	1	1				
Wiedemannia gp	42.2	3	4	3	3	3	2	2	2		1	2	1				
Heptagenia sulphurea (Müller)	27.2		1	2	2	3	4	1		1	1	1					
Limnius volckmari (Panzer)	76.1	4	4	4	4	4	4	4	3	3	4	3	2		2		
Potamophylax gp	34.3	4	3	1	2	2	2	4	2	1	2	1	2				
Ancylus fluviatilis Müller	72.8	2	4	3	4	4	4	4	3	1		4	2	1		1	
Enchytraeidae	70.1	4	3	4	4	4	4	3	1	3	3	2	2			2	4
Lumbricidae	67.2	4	4	4	3	3	3	3	3	3	2	3	3	3	2	2	2
Baetis rhodani (Pictet)	87.3	4	4	4	4	4	4	4	4	1	4	4	3	2		2	
Ephemerella ignita (Poda)	79.5	4	4	3	4	4	4	4	4	2	4	4	1	1			
Orectochilus villosus (Müller)	44.4	3	1	1	3	3	4	3		2	3	1	1				2
Polycentropus flavomaculatus (Pictet)	63.4	2	3	4	4	4	4	3		3	4	1	1		4		
Hydropsyche pellucidula (Curtis)	70.9	4	4		4	4	4	4		1	4	3	1				
Tipula montium gp	48.1	2		2	3	3	4	2	4	1	2	2	2			2	
Dicranota sp.	67.2	4	4	4	3	3	4	4	4	1	3	3	3				
Brillia modesta (Meigen)	51.1	4	4	1	3	3	3	4	4	2	1	3	3	1			
Polypedilum sp.	80.6	4	4		4	4	3	4	2	4	3	2	3		2	4	2
Rhyacodrilus coccineus (Vejdovsky)	60.4	3	3	1	3	4	3	4	3	2	3	4	2	2	2	1	2
Stylodrilus heringianus Claparède	74.6	3	4	3	4	4	4	4	4	3	4	3	2	1			
Lebertia porosa Thor	51.9	2	1	2	3	4	4	2	3	2	4	2	1	1	2		
Baetis scambus gp	72.0	3	4	3	4	4	4	4			4	3	2	1	2		
Eukiefferiella sp.	89.9	4	4	4	4	4	4	4	4	2	4	4	3		2		2
Halesus gp	32.1	3	3	1	1	2	4	4		1	2	2	1				
Nais elinguis Müller	45.1	1	2	1	3	3	3	2	3		1	3	3	1		3	4
Lumbriculus variegatus (Müller)	58.2	4	3	1	3	3	4	3	3	4	3	2	2	3	4	3	4
Elmis aenea (Müller)	90.3	4	4	4	4	4	4	4	4	3	4	4	4			2	2
Oulimnius tuberculatus (Müller)	78.7	4		3	4	4	4	3	3	4	4	4	3	2	4	3	4
Thienemannimyia gp	89.6	4	4	3	4	4	4	4	4	4	4	4	4	3		4	
Potthastia longimana gp	61.2	2	4	2	3	3	3	2	3	3	3	3	2	1		4	2
Rheocricotopus sp.	54.5	3	3		2	3	4	3	4	3	3	2	4	2			4
Orthocladius/Cricotopus	96.6	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Microtendipes sp.	58.2	2	1	2	3	3	4	2		4	4	2	2	3		3	2
Micropsectra/Tanytarsus	93.7	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Simulium ornatum gp	69.4	4	4	1		4	3	4	4	1		4	3			1	
Chelifera gp	34.7	4	3	1		2	1	2	3	1	2	3	2				
Nais communis gp	26.5	1	1	1	2	1	3	2	1	1	1	2	2	2	2	4	2
Hygrobates longiporus Thor	9.0			1		1	4			2	1	1		1	2		
Lymnaea peregra (Müller)	63.8	2	2	2	3	3	3	3	2	1	3	4	4	4	4	4	4
Gammarus pulex (L.)	80.6	4	4	1	4	4	4	4	4	4	4	4	4	1		3	
Hydroptila sp.	64.2	2	1	2	2	4	4	4	3	3	4	4	2	2		2	2
Ceratopogonidae	81.7	3	3	1	4	4	4	4	4	4	4	4	3	4	2	4	4
Heterotrissocladius sp.	7.8	4	1	1	1		1	1		3	1	1					
Rheotanytarsus/Paratanytarsus	74.3	2	3	2	4	3	4	3	1	3	4	4	4	3	4	4	2
Potamopyrgus jenkinsi (Smith)	69.4	2	3	1	3	4	3	4	3	3	4	4	3	2		2	4
Hygrobates fluviatilis (Ström)	55.2	2	2	1	2	3	3	4	4	1	4	3	3	2	2	1	
Agraylea multipunctata Curtis	12.3		1		1	2	1	1		1	1	1			4	2	
Macropelopia sp.	35.8	3	2	1	1	1	1	4	4	3	2	2	3	1			2

TABLE 4 (*continued*)

'Species'	Percentage frequency in entire data-set	TWINSPAN groups (and number of sites per group) 16 (6)	17 (12)	18 (28)	19 (52)	20 (40)	21 (8)	22 (17)	23 (4)	24 (9)	25 (33)	26 (26)	27 (19)	28 (4)	29 (2)	30 (6)	31 (2)
Anabolia nervosa Curtis	19.4	1	1	1	1	1	4	2		2	2	1	1	1	2		
Synorthocladius semivirens (Kieffer)	22.0	1	2	1		1	2	1	1	2	2	1		1	4	2	
Nais simplex gp	20.1		1	1	1	2	1	1		1	1	2	2	1	2	2	4
Limnodrilus hoffmeisteri Claparède	61.9	2	2			3	4	4	3	4	4	4	4	4	4	4	4
Aulodrilus pluriseta Piguet	40.7	2	1	1	2	2	3	4		4	3	2	2	2	4	3	4
Glossiphonia complanata (L.)	65.7	1	3	1	2	4	4	4	4	4	4	4	4	3	4	2	
Erpobdella octoculata (L.)	65.7	1	1	1	3	3	4	3	2	3	4	4	4	4	4	4	
Asellus meridianus Racovitza	25.4	1	1		1	2	3	2	2	1	2	1	2		2	2	4
Prodiamesa olivacea (Meigen)	44.0	4	2	1	1	2	2	4	4	3	3	3	4	4			
Hygrobates nigromaculatus Lebert	38.8	1	2	1	1	2	2	4	4	3	3	4	2	1		2	
Polycelis nigra gp	34.0		1		1	2	4	1	1	1	3	3	2	2	4	4	
Pisidium subtruncatum Malm	50.7		2		1	2	4	4	3	3	4	4	4	2	4	3	
Pisidium nitidum Jenyns	46.3			1		3	4	3	2	1	4	4	3	1	4	3	
Stylaria lacustris (L.)	41.0			1	1	3	4	3		3	3	3	2	3	4	4	4
Psammoryctides barbatus (Grube)	48.5	1	2		1	2	1	3	2	2	4	4	4	2	4	4	
Helobdella stagnalis (L.)	50.0	2	1	1	2	3	2	2	1	4	4	3	4	2	4	4	2
Crangonyx pseudogracilis (Bousfield)	18.7				1	2		1		2		1	2	4	2	3	4
Baetis vernus Curtis	45.9	2	2	1	1	2		4	1	2	4	4	3	1		2	
Caenis moesta gp	45.9			1	1	3	4	2		3	4	3	2	3	2	3	
Deronectes depressus gp	32.8				1	2	4	2		2	3	2	3	1	2	4	2
Ophidonais serpentina (Müller)	13.4				1	1	1	2		1	1	1	2	1		3	4
Tubifex tubifex (Müller)	23.1		2		1	1	1	2	2	3	1	1	4	1		4	2
Limnodrilus udekemianus Claparède	15.7					1	1			1	1	2	1	2	2	4	
Potamothrix hammoniensis (Michaelsen)	26.1				1	1		2	1	2	3	2	2	4	4	4	4
Chironomus sp.	14.6	2	1	1	1	1		1	1	3	1	1	2	4	2	1	2
Parachironomus sp.	7.5					1	1	1		1	1	1	1	4			2
Paratendipes sp.	19.0	1	1		1	1	1	1	1	2	2	2	3	2		4	
Hydridae	16.8		1	1	1	1	1	1	1	1	2	2	1		4		
Valvata piscinalis (Müller)	22.8				1	1		1		2	3	2	1	1	4	3	2
Physa fontinalis (L.)	24.3		1		1	1	1	1		1	3	2	2	3	4	4	2
Planorbis albus Müller	29.1				1	2	1	1		1	3	3	2	2	4	4	2
Piscicola geometra (L.)	23.1		1		1	1		3		1	3	2	1	4	4	4	2
Asellus aquaticus (L.)	48.9	1	1		1	2	1	3		4	4	4	4	4	4	4	2
Centroptilum luteolum (Müller)	29.9	3	1	1	1	1	2	2		3	3	2	2	3	4	2	2
Ablabesmyia sp.	17.2	2		1	1	1	1	1		4	2		1	2	2	2	
Sphaerium corneum (L.)	44.8		1	1	1	3	1	2		4	4	3	3	4	4	2	
Athripsodes cinereus (Curtis)	27.6				1	2	4	1		4	4	1	2				
Glyptotendipes sp.	8.6		1			1					1	1	1	4			2
Dugesia tigrina (Girard)	4.1									1				2			4
Dugesia polychroa gp	12.7									1	2	1	1	1	4	3	
Valvata cristata Müller	6.7					1				1	1	1			4	2	2
Bithynia tentaculata (L.)	22.4				1	1		1		3	4	1	2	4	2	4	2
Bithynia leachi (Sheppard)	5.2										1		1	3	4	1	
Limnesia maculata (Müller)	4.5									1	1	1		1	4	2	4
Sigara falleni (Fieber)	10.1				1	1	1			1	1		1	2	2	4	4
Procladius gp	27.6	1	1			1	1	1		4	3	2	3	4	2	4	4
Cryptochironomus sp.	21.3				1	1				3	3	1	2	2	4	3	2
Dicrotendipes sp.	12.3					1				3	1	1	3	4	2	4	2
Planorbis vortex (L.)	14.6					1		1			2	2	2		2	4	
Dero digitata (Müller)	2.2														2	2	4
Cloeon dipterum (L.)	6.3										1	1	1	3	2	4	2
Ischnura elegans (Linden)	3.4										1		1	4		2	2
Sigara dorsalis (Leach)	17.9					1	1	1		1	2	1	2	4	4	4	4
Haliplus lineatocollis (Marsham)	13.4				1	1	1	1			1	2	2	2		2	4
Haliplus fluviatilis Aubé	13.4					1	1			1		2	2	2	2	3	4
Acricotopus lucens (Zetterstedt)	2.2							1					1			2	4
Athripsodes aterrimus (Stephens)	9.0										1	2	1		4	1	
No. of 'species' listed in Category 4.		32	35	20	30	29	51	35	22	18	34	26	17	20	35	31	27

occurred at sites placed in Group 27 and a further two were in Group 24. In addition, the Category 4 species in Group 27 are typically of widespread occurrence throughout the data-set and the preponderance of Oligochaeta, Hirudinea, Crustacea, Mollusca and Diptera is compatible with mild enrichment of some of the sites.

Table 5 lists the sites, together with their National Grid references, number of taxa and TWINSPAN group. Also indicated are the seventeen sites used in the classification with BMWP scores in the range 50–100. Some of these sites may have fairly low scores simply because the environment is harsh and incapable of supporting a diverse fauna. This is probably the case for some sites near the headwaters of rivers in the north and west (Sites 159, 199 and 235) but at most other sites in lowland England mild organic enrichment or the influence of brackish water is probably responsible. The overriding impression which emerges on examination of Table 5 alongside Fig. 2 is that, with relatively few exceptions, the sites form a logical sequence down each river system. The high correlation between Axis 2 ordination scores and both distance from source (LKM) and discharge (DISCH) leads us to expect that many of the sites high up in a catchment will have high Axis 2 scores, whereas successive sites further downstream will have low Axis 2 scores. In many rivers, there is also a tendency for a sequence of sites to occupy several TWINSPAN groups whose locations are progressively further to the right of the ordination plot. This is most easily seen in the Swale/Ure/Ouse system which has its source in the Pennines and, after being a 'spatey' upland river for some distance (Groups 18 and 19), assumes a progressively more lowland character (Groups 20 and 24). In contrast, a number of rivers in the west and north retain their upland character for most of their length (Groups 17–19) including the Camel, Avill, Ehen, Esk, Tyne (Northumbrian WA), Dee and Stinchar.

Although the majority of sites conform to a pattern in which successive sites downstream either stay in the same TWINSPAN group, occupy a group with predominantly lower Axis 2 scores, or occupy a group having a more lowland character, approximately 10% of sites break this pattern. Of the rivers regarded as being upland in character for at least part of their course (Table 3) there are relatively few which show unusual sequences and for most of these the reasons are clear. Some result from the occurrence of a site with a relatively slow current speed in relation to sites both upstream and downstream, as in the case of Site 259 on the River Spey and the pool samples on the River Wansbeck and River Forth. Site 183 on the River Derwent (North West WA) breaks a sequence, probably through being immediately downstream of Bassenthwaite. The lowest two sites on the River Torridge (Sites 19 and 20) appear to have a more upland character than sites nearer the source. This may be related to the fact that at these sites the gradient of the river increases once again instead of decreasing downstream.

In examining the lowland rivers (Table 3), there appears to be a slightly greater tendency for sites to fail to conform to the pattern outlined in the previous paragraph, although the sequence of expected groups still holds in a high proportion of cases. There are several possible reasons for this, including Man's influence in altering both the physical and chemical characteristics of lowland rivers. Of all the lowland river systems investigated, the Welland and Glen showed the least pattern in the sequence of TWINSPAN groups occupied by successive sites downstream.

Appendix 1 gives a key to the sixteen groups derived by two-way indicator species analysis. Provided that any new site is sampled with an equivalent expenditure of effort to that used in this project and that the site falls within the range of sites used in the preparation of this classification, it may be keyed through to a group using the differential species presented in the Appendix.

In order to indicate some of the environmental conditions which typify each of the sixteen groups of sites, the mean values of the twenty-eight variables are listed for each TWINSPAN group in Table 6. The more notable features of each group are also briefly reviewed in Table 7. The relationship of the groups listed in Table 7 may be seen by reference to their positions on the TWINSPAN plot (Fig. 2).

Correlation of site groupings with environmental variables. After the classification of sites into a series of groups using two-way indicator species analysis on the species × sites matrix, we assessed whether sites could be predicted to the same

TABLE 5. The 270 sites at which samples were taken together with their grid references, the number of taxa and TWINSPAN groups at level of division 4 (i.e. sixteen groups). *Additional pool samples. †Sites with BMWP scores (combined season family data) in the range 50–100. All other sites used in the classification have scores in excess of 100.

Site code	Name of river	National Grid ref.	No. of taxa	TWINSPAN group
1	R. Camel	SX 104 827	80	17
2	R. Camel	SX 088 778	63	19
3	R. Camel	SX 065 715	89	19
4	R. Camel	SX 015 685	81	19
5	R. Axe	ST 457 053	94	22
6	R. Axe	ST 402 060	78	20
7	R. Axe	ST 326 025	109	20
8	R. Axe	SY 262 953	91	20
9	R. Exe	SS 791 407	65	16
10	R. Exe	SS 853 383	90	19
11	R. Exe	SS 912 342	84	19
12	R. Exe	SS 930 245	96	19
13	R. Exe	SS 948 153	80	19
14	R. Exe	SX 929 984	77	20
15	R. Torridge	SS 324 178	112	20
16	R. Torridge	SS 399 126	98	20
17	R. Torridge	SS 470 061	115	20
18	R. Torridge	SS 542 064	85	20
19	R. Torridge	SS 543 143	80	19
20	R. Torridge	SS 499 185	81	19
21	R. Avill	SS 925 395	66	17
22	R. Avill	SS 960 430	96	17
23	R. Avill	SS 984 432	85	17
24	Western Avon	SU 071 585	74	26
25	Western Avon	SU 132 558	86	25
26	R. Avon (Hants)	SU 163 437	82	25
27	R. Avon (Hants)	SU 129 330	86	20
28	R. Avon (Hants)	SU 163 174	92	25
29	R. Avon (Hants)	SU 149 035	91	25
30	R. Avon (Hants)	SZ 158 933	89	25
31	Sherston Avon	ST 880 873	82	22
32	Tetbury Avon	ST 915 893	73	22
33	Bristol Avon	ST 943 862	92	25
34	Bristol Avon	ST 965 831	66	20
35	Bristol Avon	ST 947 758	92	25
36	Bristol Avon	ST 922 681	68	25
37	Bristol Avon	ST 856 609	82	27
38	Bristol Avon	ST 689 677	77	28
39	Avonwater	SZ 250 996	80	16
40	Avonwater	SZ 292 961	114	22
41	Avonwater	SZ 307 941	93	22
42	Candover Brook	SU 565 345	86	26
43	R. Itchen	SU 523 325	89	22
44	R. Itchen	SU 481 282	112	22
45	R. Itchen	SU 470 233	96	22
46	R. Itchen	SU 466 175	138	25
47	R. Rother (Sussex)	SU 773 273	85	22
48	R. Rother (Sussex)	SU 769 260	94	22
49	R. Rother (Sussex)	SU 783 234	100	22
50	R. Rother (Sussex)	SU 863 226	87	20
51	R. Rother (Sussex)	SU 935 213	90	20
52	R. Rother (Sussex)	TQ 034 178	106	20

TABLE 5 (*continued*)

Site code	Name of river	National Grid ref.	No. of taxa	TWINSPAN group
53	R. Arun	TQ 187 292	102	27
54	R. Arun	†TQ 139 309	56	27
55	R. Arun	TQ 072 307	89	27
56	R. Arun	TQ 046 229	134	25
57	R. Dudwell	TQ 655 224	98	22
58	R. Brede	†TQ 765 163	48	27
59	R. Brede	TQ 783 177	75	22
60	R. Brede	†TQ 903 179	45	31
61	R. Rother (Kent)	TQ 720 262	106	20
62	R. Rother (Kent)	TQ 771 243	113	25
63	R. Rother (Kent)	TQ 850 270	111	31
64	R. Rother (Kent)	TQ 936 237	59	–
65	R. Evenlode	SP 202 312	78	26
66	R. Evenlode	SP 220 281	102	26
67	R. Evenlode	SP 274 197	74	26
68	R. Evenlode	SP 366 173	64	20
69	R. Evenlode	SP 448 102	88	25
70	R. Tillingbourne	TQ 130 470	79	22
71	R. Tillingbourne	TQ 053 479	64	22
72	R. Wey	SU 756 417	95	25
73	R. Wey	SU 873 437	82	26
74	R. Wey	SU 947 438	90	20
75	R. Wey	TQ 005 532	80	25
76	R. Wey	TQ 070 613	67	25
77	R. Mimram	TL 193 207	81	26
78	R. Mimram	TL 208 180	91	26
79	R. Mimram	TL 282 134	64	26
80	R. Lee	TL 365 143	86	26
81	R. Lee	TL 384 076	91	28
82	R. Lee	TL 374 044	83	25
83	R. Lee	TQ 374 983	116	28
84	Gwendraeth Fach	SN 543 163	133	16
85	Gwendraeth Fach	SN 460 139	148	21
86	Gwendraeth Fach	SN 419 077	126	20
87	R. Teifi (Tyfi)	SN 749 659	74	19
88	R. Teifi (Tyfi)	SN 684 628	125	21
89	R. Teifi (Tyfi)	SN 642 547	114	21
90	R. Teifi (Tyfi)	SN 523 454	141	21
91	R. Teifi (Tyfi)	SN 373 403	107	21
92	R. Teifi (Tyfi)	SN 217 437	145	21
93	R. Clwyd	SJ 040 488	96	16
94	R. Clwyd	SJ 109 519	85	17
95	R. Clwyd	SJ 124 571	71	17
96	R. Clwyd	SJ 091 658	100	17
97	R. Clwyd	SJ 060 719	93	20
98	R. Leadon	SO 688 481	84	27
99	R. Leadon	SO 697 404	66	26
100	R. Leadon	†SO 701 332	53	26
101	R. Leadon	SO 730 307	66	26
102	R. Leadon	SO 770 270	81	26
103	R. Perry	SJ 347 302	85	26
104	R. Perry	SJ 374 294	84	26
105	R. Perry	SJ 396 245	76	25
106	R. Perry	SJ 422 210	84	26

TABLE 5 (*continued*)

Site code	Name of river	National Grid ref.	No. of taxa	TWINSPAN group
107	R. Perry	SJ 439 171	79	20
108	R. Blithe	SJ 942 435	55	23
109	R. Blithe	SJ 975 393	68	20
110	R. Blithe	SK 024 334	101	25
111	R. Blithe	SK 048 259	94	20
112	R. Blithe	SK 109 190	95	20
113	R. Smite	†SK 690 262	48	27
114	R. Smite	SK 697 333	98	27
115	R. Smite	SK 773 427	65	27
116	R. Devon	SK 822 315	73	22
117	R. Devon	SK 812 390	82	25
118	R. Devon	SK 785 511	64	25
119	R. Dove	SK 084 665	93	17
120	R. Dove	SK 121 598	80	17
121	R. Dove	SK 146 504	100	17
122	R. Dove	SK 115 392	91	20
123	R. Dove	SK 163 312	90	20
124	R. Dove	SK 268 270	69	20
125	Stambourne Brook	TL 759 384	95	27
126	R. Colne	TL 798 323	87	27
127	R. Colne	TL 867 289	86	25
128	R. Colne	TL 921 272	81	25
129	R. Colne	TL 997 256	71	29
130	Great Eau	TF 332 779	82	23
131	Great Eau	TF 370 768	89	23
132	Great Eau	TF 403 777	92	26
133	Great Eau	TF 425 826	73	26
134	Great Eau	TF 452 867	102	30
135	R. Glen	SK 985 286	81	27
136	R. Glen	†SK 990 248	47	30
137	R. Glen	TF 019 177	62	26
138	R. Glen	TF 068 112	65	26
139	R. Glen	TF 156 190	100	30
140	R. Glen	TF 235 260	80	30
141	R. Welland	SP 697 864	75	26
142	R. Welland	†SP 750 883	60	27
143	R. Welland	†SP 779 923	56	30
144	R. Welland	SP 914 976	59	26
145	R. Welland	TF 007 063	77	26
146	R. Welland	TF 228 106	94	30
147	R. Wensum	TF 885 240	83	23
148	R. Wensum	TF 881 282	74	25
149	R. Wensum	TF 964 273	65	25
150	R. Wensum	TG 005 202	86	25
151	R. Wensum	TG 052 178	88	25
152	R. Wensum	TG 161 137	84	25
153	R. Blackwater	TF 952 068	71	27
154	R. Blackwater	TF 987 039	76	26
155	R. Yare	†TG 045 058	61	24
156	R. Yare	TG 108 084	87	25
157	R. Yare	TG 190 082	106	25
158	R. Yare	TG 236 060	76	29
159	R. Hodder	†SD 702 590	44	18
160	R. Hodder	SD 715 524	85	19

TABLE 5 (*continued*)

Site code	Name of river	National Grid ref.	No. of taxa	TWINSPAN group
161	R. Hodder	SD 658 479	76	19
162	R. Hodder	SD 697 411	75	19
163	Gayle Beck	SD 785 803	58	18
164	R. Ribble	SD 806 726	65	19
165	R. Ribble	SD 806 614	87	19
166	R. Ribble	SD 851 551	98	20
167	R. Ribble	SD 775 466	93	20
168	R. Ribble	SD 715 387	60	19
169	R. Ribble	SD 662 356	79	20
170	R. Dane	SJ 930 636	86	17
171	R. Dane	SJ 849 637	70	20
172	R. Dane	SJ 715 674	63	20
173	R. Dane	†SJ 660 734	40	27
174	R. Weaver	†SJ 561 499	57	27
175	R. Weaver	SJ 620 470	83	27
176	R. Weaver	SJ 659 451	65	26
177	R. Weaver	SJ 651 536	74	27
178	R. Weaver	†SJ 665 610	46	27
179	R. Weaver	†SJ 657 729	50	28
180	R. Weaver	SJ 628 750	9	–
181	R. Derwent (NWWA)	NY 255 176	59	18
182	R. Derwent (NWWA)	NY 243 260	85	19
183	R. Derwent (NWWA)	NY 200 321	70	20
184	R. Derwent (NWWA)	NY 116 307	77	19
185	R. Derwent (NWWA)	NY 046 304	84	19
186	R. Derwent (NWWA)	NY 009 293	76	19
187	R. Ehen	NY 068 159	60	19
188	R. Ehen	NY 014 130	51	19
189	R. Ehen	NY 012 125	62	19
190	R. Ehen	NY 007 061	56	19
191	R. Derwent (YWA)	SE 942 910	90	16
192	R. Derwent (YWA)	SE 988 848	87	25
193	R. Derwent (YWA)	SE 892 795	95	25
194	R. Derwent (YWA)	SE 790 715	118	25
195	R. Derwent (YWA)	SE 710 555	122	25
196	R. Derwent (YWA)	SE 697 424	109	24
197	R. Derwent (YWA)	†SE 683 287	66	24
198	R. Esk	NZ 663 062	66	16
199	R. Esk	†NZ 685 085	60	19
200	R. Esk	NZ 762 076	89	19
201	R. Esk	NZ 869 082	80	19
202	R. Swale	NY 885 015	54	18
203	R. Swale	SD 933 978	68	18
204	R. Swale	SE 046 985	65	18
205	R. Swale	NZ 146 007	62	18
206	R. Swale	SE 319 918	83	19
207	R. Swale	SE 398 759	109	20
208	R. Ure	SE 467 621	91	24
209	R. Ouse	SE 556 552	95	24
210	R. Ouse	SE 591 455	95	24
211	R. Tees	NY 762 338	53	18
212	R. Tees	NY 814 288	43	18
213	R. Tees	NY 931 259	46	18
214	R. Tees	NZ 042 172	61	19
215	R. Tees	NZ 178 163	69	19

TABLE 5 (*continued*)

Site code	Name of river	National Grid ref.	No. of taxa	TWINSPAN group
216	R. Tees	NZ 288 101	65	19
217	R. Tees	NZ 346 114	73	20
218	South Tyne	NY 758 372	49	18
219	South Tyne	NY 717 459	50	18
220	South Tyne	NY 683 554	49	18
221	South Tyne	NY 674 617	48	18
222	South Tyne	NY 781 643	59	19
223	South Tyne	NY 910 659	70	19
224	R. Tyne (NWA)	NY 990 641	78	19
225	R. Tyne (NWA)	NZ 111 643	72	19
226	R. Wansbeck	NY 996 844	87	19
227	R. Wansbeck	NZ 053 842	91	19
228	R. Wansbeck	NZ 119 850	82	19
229	R. Wansbeck	NZ 174 858	98	20
230	R. Wansbeck	NZ 236 862	74	20
231	R. Wansbeck	*NZ 237 862	79	22
232	R. Teith	NN 628 078	97	19
233	R. Teith	NN 668 045	67	19
234	R. Teith	NN 723 013	69	19
235	Water of Chon	†NN 435 035	40	19
236	R. Forth	NN 507 014	76	19
237	R. Forth	NS 599 974	111	21
238	R. Forth	NS 669 960	89	21
239	R. Forth	*NS 669 960	64	24
240	R. Forth	NS 710 956	81	20
241	R. Forth	*NS 710 956	50	24
242	R. Forth	NS 770 955	107	20
243	R. Tyne (FRPB)	NT 378 618	70	17
244	R. Tyne (FRPB)	NT 413 689	75	19
245	R. Tyne (FRPB)	NT 459 690	86	20
246	R. Tyne (FRPB)	NT 513 733	78	20
247	R. Tyne (FRPB)	NT 593 772	75	20
248	R. Dee	NO 061 896	36	18
249	R. Dee	NO 143 915	50	18
250	R. Dee	NO 271 944	57	18
251	R. Dee	NO 385 965	37	18
252	R. Dee	NO 557 980	94	19
253	R. Dee	NO 608 973	57	18
254	R. Dee	NO 719 964	70	18
255	R. Dee	NJ 904 023	54	19
256	R. Spey	NN 522 947	48	18
257	R. Spey	NN 614 943	65	18
258	R. Spey	NN 708 980	61	18
259	R. Spey	NH 838 062	80	24
260	R. Spey	NH 946 188	90	18
261	R. Spey	NJ 038 264	60	18
262	R. Spey	NJ 183 388	68	19
263	R. Spey	NJ 283 452	46	18
264	R. Spey	NJ 343 610	65	18
265	R. Stinchar	NX 395 956	55	18
266	R. Stinchar	NX 321 957	69	19
267	R. Stinchar	NX 272 937	44	19
268	R. Stinchar	NX 204 899	48	19
269	R. Stinchar	NX 140 858	60	19
270	R. Stinchar	NX 089 825	65	19

TABLE 6. Mean values of twenty-eight environmental variables for each of the sixteen TWINSPAN groups. Variables whose abbreviations are prefixed L have been back transformed. Further details are given in Table 1.

TWINSPAN group	LKM	LSLOPE	LALT	DISCH	LMEANW	DEPTHC	SUBSTNO	DAY 1	DAY 2	DAY 3	LM-WIDTH	LM-DEPTH	MAX-VEL	MIN-VEL
16	5.90	7.37	113.4	1.33	4.66	1.00	1.83	117	207	277	4.82	13.33	3.50	2.33
17	8.97	6.37	104.5	2.83	5.93	1.25	2.17	128	216	294	5.82	16.15	4.33	3.50
18	20.57	5.62	194.2	5.86	19.15	2.07	1.54	130	213	290	19.32	32.25	4.00	3.43
19	27.50	3.06	54.7	5.81	18.18	1.87	1.65	127	220	296	15.15	28.50	4.02	3.42
20	30.73	1.44	34.6	5.10	14.59	1.95	2.05	131	213	299	13.35	24.11	4.32	3.52
21	29.41	1.01	36.0	6.25	18.49	2.75	2.50	109	193	270	19.41	36.32	4.25	3.50
22	9.62	2.76	36.7	2.41	6.19	1.47	2.88	135	227	308	5.46	20.45	3.65	2.76
23	3.94	5.48	63.7	1.00	2.26	1.00	2.50	116	201	297	2.02	13.51	4.00	3.00
24	64.73	0.40	13.7	7.11	28.27	4.44	2.44	113	197	271	27.80	169.82	2.22	1.44
25	29.15	0.88	23.0	4.42	10.24	2.52	2.52	117	210	288	9.18	44.12	3.76	2.91
26	12.44	1.60	46.4	2.31	5.62	1.81	2.00	132	226	304	5.37	25.43	3.77	2.92
27	10.45	1.64	39.0	2.74	3.55	1.63	2.58	116	209	287	3.70	17.42	3.68	2.47
28	49.89	0.67	11.7	5.75	22.41	4.25	1.75	120	206	301	20.42	135.55	1.50	1.25
29	37.47	1.02	4.5	3.50	9.95	2.50	3.00	109	237	284	8.73	55.74	3.00	1.50
30	30.10	0.80	7.6	2.83	10.43	3.33	2.17	132	228	293	10.22	76.70	1.83	1.17
31	24.68	0.55	3.5	4.00	13.42	4.00	1.50	150	240	309	13.41	145.28	2.00	1.00

TWINSPAN group	MED-VEL	M-SUBST	MIN-DOM	MAX-DOM	MED-DOM	MAX-MAC	MIN-MAC	MEAN-MAC	pH	O2B	LTON	LCL	LORPH	ALK
16	2.67	−5.10	−6.79	−5.67	−6.42	5.0	2.7	3.7	7.34	10.9	1.79	21.35	0.122	82.5
17	3.58	−5.03	−6.73	−5.31	−6.06	17.0	7.7	11.8	7.63	10.2	2.78	20.54	0.165	117.4
18	3.75	−6.36	−7.29	−6.50	−7.08	4.7	0.6	2.5	7.32	11.4	0.34	9.39	0.022	39.5
19	3.79	−5.79	−6.65	−6.12	−6.49	12.7	4.0	7.7	7.34	11.4	0.87	15.79	0.038	50.9
20	3.97	−4.04	−5.59	−3.56	−4.88	40.8	13.1	25.8	7.70	11.0	3.00	25.94	0.193	123.4
21	3.87	−3.58	−5.03	−3.31	−4.22	41.4	23.0	32.1	6.96	11.0	1.15	14.16	0.048	47.1
22	3.29	−1.46	−3.55	−0.31	−2.26	48.1	18.1	33.2	7.64	10.6	4.26	27.34	0.266	144.5
23	3.25	−2.60	−3.25	−3.25	−3.25	21.2	5.7	13.1	7.81	10.5	7.12	29.79	0.123	190.7
24	2.00	4.32	2.14	6.42	4.00	4.4	0.8	2.4	7.37	10.2	1.74	20.30	0.075	110.3
25	3.39	−0.11	−1.45	1.45	−5.60	51.7	19.2	37.1	7.85	10.4	5.40	35.08	0.407	200.6
26	3.42	−0.01	−1.92	1.86	0.37	68.1	22.3	42.1	7.87	9.9	6.70	33.43	0.222	230.9
27	2.95	−0.95	−2.17	0.36	−1.25	40.8	8.2	25.7	7.56	9.1	6.70	57.58	0.473	189.3
28	1.25	1.85	2.00	2.00	2.00	5.7	3.0	4.1	7.95	10.2	7.77	75.35	1.064	236.8
29	2.00	3.30	1.62	3.25	2.25	75.0	10.5	42.7	7.93	10.6	7.52	61.80	0.571	234.0
30	1.33	5.24	5.25	6.83	5.83	80.8	22.0	52.9	8.09	10.6	6.09	51.65	0.172	211.5
31	1.00	0.91	1.50	1.50	1.50	85.0	67.5	75.0	7.63	11.4	2.64	370.51	0.183	95.3

TABLE 7. The location and some important features of the sixteen groups of sites generated by two-way indicator species analysis (Fig. 3). Definition of 'upland' and 'lowland' is as used for the interpretation of Table 3.

Group (no. sites)	Mean no. taxa (range)	Location(s)	
16 (n = 6)	88.3 (65–133)	Upland England, Wales	Top sites on upland rivers, typically small, steep and at high altitude with coarse substratum. Little macrophyte and low alkalinity. (Includes top site on Avonwater.)
17 (n = 12)	84.3 (66–100)	Upland England, Wales and Scotland	Top and also upstream sites on upland rivers. Similar in character to Group 16.
18 (n = 28)	55.2 (36–90)	Northern England, Scotland	Top and upstream sites on upland rivers but extending further downstream in the Dee and Spey which retain their upland character. Typically high altitude sites with higher discharge and lower alkalinity than Groups 16 and 17. Restricted fauna.
19 (n = 52)	72.7 (40–97)	Upland England, Wales and Scotland	Mainly sites in the middle and lower reaches of upland rivers whose upstream sites occur in Groups 16–18.
20 (n = 40)	88.0 (63–126)	Upland and lowland England, Wales and Scotland	Middle and lower reach sites on rivers whose upper sites occur in Groups 16–19 and also 21–23. Compared to Group 19, these sites have a lower slope, less coarse substratum, higher macrophyte cover and a higher alkalinity.
21 (n = 8)	122.5 (89–148)	Wales and Scotland	A small group distinguished from Group 20 by lower alkalinity but a richer invertebrate community.
22 (n = 17)	88.2 (64–114)	Lowland England	Upper and middle reach sites with a higher alkalinity than Groups 16–21. (Includes pool site on Wansbeck in upland England.)
23 (n = 4)	77.2 (55–89)	Lowland England	A small group dominated by top sites with a high alkalinity.
24 (n = 9)	79.0 (50–109)	Upland and lowland England and Scotland	Lower reaches or 'pool' sites typically deep with a small mean substratum particle size and little macrophyte present.
25 (n = 33)	90.7 (64–138)	Lowland England	Sites in the middle and lower reaches of lowland rivers (but including middle reach sites on Yorkshire Derwent) having a high alkalinity.
26 (n = 26)	76.6 (53–102)	Lowland England	Sites in the upper and more particularly the middle reaches of lowland rivers with a high alkalinity and high macrophyte cover.
27 (n = 19)	71.9 (40–102)	Lowland England	Typically sites in the upper and middle reaches of lowland rivers with a high alkalinity.
28 (n = 4)	83.5 (50–116)	Lowland England	Four sites in the lower reaches of lowland rivers, typically deep as in Group 24, but with a higher alkalinity.
29 (n = 2)	73.5 (71–76)	Eastern England	Two sites in the lower reaches of rivers with a high alkalinity and high macrophyte cover.
30 (n = 6)	79.8 (47–102)	Eastern England	Six further sites in the lower reaches of rivers with a high alkalinity and macrophyte cover also, but deeper and with a smaller mean substratum particle size than Group 29.
31 (n = 2)	78.0 (45–111)	Lowland England	Two deep water sites in the S.E. with a high chloride level (brackish) and high macrophyte cover.

groups using environmental data only. Multiple discriminant analysis (MDA) was employed for this purpose using the twenty-eight environmental variables listed in Table 1, and two separate analyses were carried out. In the first, an attempt was made to predict the TWINSPAN group of each of the 268 sites in the entire data-set using the environmental data. Although this provided a measure of predictive ability within the present data-set, it could not be regarded as an independent assessment of performance. This could only be achieved by the use of a set of unclassified sites. Hence, in the second analysis forty sites were removed at random, the remaining 228 sites reclassified using TWINSPAN, and MDA repeated on the 228 samples to obtain predictive equations which could be used to classify the forty test sites. The predictions were then compared with the results obtained by keying out each site using the macro-invertebrate community.

Prediction of TWINSPAN groups for the 268 sites. Table 8 presents the outcome of using MDA on the environmental data to predict TWINSPAN group membership. Although in previous sections emphasis has been placed on the classification at Level 4 (i.e. sixteen groups), the results for Levels 1–5 are presented here for comparative purposes. The number of significant discriminant functions ($P \leq 0.05$) are also given for each level of the analysis. The procedure for determining the group into which a site is placed using environmental features depends upon determining the position of the site in relation to all the groups in discriminant space and assessing the highest probability of group membership for the site. Where the highest probability is for the same group as determined on biological criteria, we regarded the prediction as correct. Table 8 also includes the percentage of sites for which the prediction is incorrect but where group membership based on biological criteria corresponds to the second most probable group determined by MDA. At Level 1, when all sites are allocated to one or other of two groups, the environmental data places 91.4% of sites in the group in which they are placed by TWINSPAN. By the time there are sixteen different groups into which each site may be placed, MDA on the environmental data still classifies 76.1% of sites correctly. In addition, in a high proportion of the cases where the prediction is incorrect, the correct group is listed as the second most probable on environmental grounds.

Examination of the standardized canonical discriminant function coefficients for the first four discriminant functions at Level 4 (Table 9) enables an assessment to be made of the relative contribution made by each of the environmental variables used in the analysis. In Function 1, alkalinity, ALK, has the highest coefficient (0.553), followed by the log of slope, LSLOPE (−0.331). Correlation of environmental variables with ordination scores also demonstrated the importance of alkalinity ($r=0.739$) along the first axis of variation (Table 2). In Function 2 the highest coefficients recorded are mean % macrophyte cover, MEANMAC, (0.872), the log of chloride, LCL (−0.545), minimum % macrophyte cover, MINMAC (−0.528) and the log of mean water depth in sample area, LMDEPTH (−0.445). To understand why these variables are important at Level 4, it is helpful to examine a plot of the group centroids on the first two discriminant functions (Fig. 4). Each group centroid is obtained by averaging the MDA scores for the sites within a particular TWINSPAN group and represents the most typical location of a site from that group in the discriminant function space. Comparison of the group means on the first two functions indicates how far apart the groups are along the most important of the eight axes used in this analysis. Group means for alkalinity (Table 6) are presented alongside each group centroid to demonstrate the strong relation with Function 1. The most notable feature on Function 2 is the extent to which Group 31 is an outlier, and this is relevant to an understanding of the environmental variables which show high coefficients on Function 2. Reference to Table 6 indicates that Group 31 has a much higher percentage macrophyte cover and chloride concentration than any other TWINSPAN group. These variables have therefore been of paramount importance in discriminating between Group 31 and all other groups, although they appear to be of less value in discriminating between the remaining groups. In contrast, water depth in sample area, which is used in its logarithmic form in MDA, and is ranked as the fourth most important variable on Function 2, shows a clear pattern of change through the group centroids along Axis 2 (Fig. 4).

Although alkalinity and mean depth have been used as a demonstration of some of the differences between TWINSPAN group centroids, Table 9 indicates that most of the environmental variables used in the analysis are of some value for discrimination between groups. Of the twenty-eight variables under consideration, all but five were ranked as being one of the ten most important variables in at least one of the first four discriminant functions. The five variables which were found to be of little value were SUBSTNO, DAY 1, DAY 2, DAY 3 and MAXVEL. Date of sampling in

TABLE 8. The percentage of the 268 sites predicted to the correct TWINSPAN group using multiple discriminant analysis on twenty-eight environmental variables. Results are presented for each of five levels of division in the classification together with the number of significant discriminant functions used. In addition, the percentage of sites in which the correct group is given as the second most probable on environmental grounds, is also presented.

	Level of classification				
	1	2	3	4	5
No. of TWINSPAN groups	2	4	8	16	32
No. of significant discriminant functions	1	3	5	8	10
Percentage of correct predictions	91.4	84.3	79.1	76.1	73.1
Percentage of sites in which correct group is the second most probable	8.6	14.2	15.3	15.3	17.6

TABLE 9. Standardized canonical discriminant function coefficients of environmental variables for Level 4 of TWINSPAN. Numbers in parentheses indicate rankings of the variables in each discriminant function.

Variable	Function 1	Function 2	Function 3	Function 4
LKM	−0.128 (16)	0.037 (27)	0.170 (13)	0.303 (9)
LSLOPE	−0.331 (2)	0.197 (15)	−0.172 (12)	−0.110 (20)
LALT	−0.250 (7)	0.025 (28)	0.299 (6)	−0.280 (10)
DISCH	−0.286 (3)	−0.065 (25)	0.117 (19)	−0.514 (5)
LMEANW	−0.280 (4)	0.142 (20)	−0.009 (28)	0.721 (1)
DEPTHC	0.180 (11)	−0.324 (6)	0.240 (9)	−0.103 (21)
SUBSTNO	0.133 (15)	0.109 (23)	0.142 (16)	0.033 (27)
DAY 1	0.072 (24)	−0.192 (16)	−0.108 (20)	0.178 (16)
DAY 2	−0.088 (22)	−0.150 (19)	−0.176 (11)	−0.211 (13)
DAY 3	0.116 (17)	0.131 (21)	−0.097 (22)	−0.043 (25)
LMWIDTH	0.184 (10)	0.044 (26)	0.117 (18)	−0.328 (8)
LMDEPTH	−0.019 (28)	−0.445 (4)	−0.034 (26)	0.039 (26)
MAXVEL	0.086 (23)	0.230 (12)	0.094 (23)	0.011 (28)
MINVEL	−0.033 (27)	0.285 (8)	−0.157 (15)	−0.193 (14)
MEDVEL	−0.049 (26)	0.179 (17)	0.263 (7)	0.564 (3)
MSUBST	0.252 (6)	−0.110 (22)	0.658 (2)	0.173 (17)
MINDOM	0.094 (21)	0.274 (9)	−0.546 (3)	−0.127 (19)
MAXDOM	0.196 (9)	0.265 (10)	−0.074 (24)	0.162 (18)
MEDDOM	−0.058 (25)	−0.222 (13)	0.304 (5)	−0.069 (22)
MAXMAC	0.104 (20)	−0.399 (5)	−0.164 (14)	0.536 (4)
MINMAC	0.141 (14)	−0.528 (3)	−0.028 (27)	0.044 (24)
MEANMAC	−0.114 (18)	0.872 (1)	−0.233 (10)	−0.059 (23)
pH	−0.218 (8)	−0.075 (24)	−0.048 (25)	0.711 (2)
O2B	−0.108 (19)	−0.160 (18)	−0.251 (8)	0.367 (6)
LTON	0.166 (12)	0.314 (7)	0.340 (4)	−0.257 (11)
LCL	0.166 (13)	−0.545 (2)	−0.750 (1)	−0.216 (12)
LORPH	0.258 (5)	0.231 (11)	0.125 (17)	0.179 (15)
ALK	0.533 (1)	0.216 (14)	0.101 (21)	−0.354 (7)
% variance explained	55.15	13.27	6.85	5.63

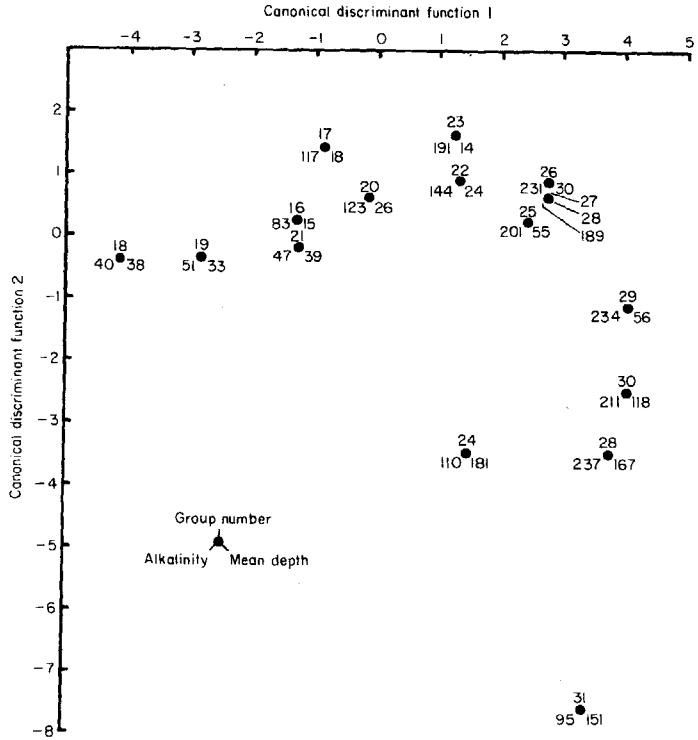

FIG. 4. The position of the centroids of the sixteen TWINSPAN groups in discriminant space obtained by MDA on twenty-eight environmental variables. Group means for alkalinity and mean depth are also given.

each of the three seasons was expected to be of limited value in this analysis, which combines the results of three seasons, but was retained for uniformity with additional analyses performed on single season data-sets (Furse *et al.*, 1984).

A more detailed assessment of group membership prediction at Level 4 can be seen in Table 10. The percentage of sites correctly predicted is given for each TWINSPAN group (overall success rate 76.1%) and the occurrence of incorrect predictions is also shown. There appears to be no clear relationship between group size and prediction success, or between upland/lowland groups or upstream/downstream groups and prediction success.

Prediction of TWINSPAN groups using forty test sites. This procedure was used as an independent assessment of predictive ability (Pielou, 1977). Although it is more rigorous than the previous test, it does make the assumption that the forty test sites come within the range of the 228 site classification and hence that it is valid to attempt to classify them.

After the removal of forty test sites, chosen at random, the remaining 228 sites were classified using TWINSPAN and the MDA repeated using the environmental data for these sites to obtain predictive equations. The forty test sites were then classified on biological criteria using the TWINSPAN key (Appendix 1) and also on environmental criteria by prediction of TWINSPAN groups using the MDA equations. The results are presented in Table 11. MDA using environmental data for the 228 sites yielded a broadly similar predictive ability to that seen in the main classification exercise (Table 8)

TABLE 10. Prediction of TWINSPAN groups at Level 4 for 268 sites using multiple discriminant analysis

Group	No. of sites	Predicted group membership 16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	% sites correctly predicted
16	6	4			1			1										66.7
17	12	1	8		1	2												66.7
18	28			25	3													89.3
19	52	2		4	44	1	1											84.6
20	40				4	25	2	2			5	1	1					62.5
21	8					1	7											87.5
22	17					1		11			4		1					64.7
23	4		1						2			1						50.0
24	9									8		1						88.9
25	33							3			26	3	1					78.8
26	26					1		1			5	19						73.1
27	19								1		1	1	15	1				78.9
28	4										1			3				75.0
29	2										1				1			50.0
30	6										1	1				4		66.7
31	2																2	100.0
Total no. sites	268																	76.1

TABLE 11. Prediction of TWINSPAN groups for five levels of a 228 site classification using multiple discriminant analysis and ability to predict the correct group for forty test sites. The percentage of sites in which the correct group is the second most probable on environmental criteria is also given. Note that only fifteen and twenty-nine groups were produced at TWINSPAN Levels 4 and 5 because one group at Level 3 contained a single site and hence the group could not be divided further.

	Level of classification 1	2	3	4	5
No. of TWINSPAN groups	2	4	8	15	29
No. of significant discriminant functions	1	3	5	7	9
228 site classification: percentage of correct predictions	93.0	84.2	81.6	67.5	68.4
40 test sites: percentage of correct predictions	95.0	87.5	62.5	50.0	30.0
Percentage of the forty test sites in which correct group is the second most probable	5.0	12.5	27.5	25.0	15.0

although the success rate at Levels 4 and 5 was slightly lower in the 228 site classification.

Prediction of TWINSPAN groups for the forty test sites using environmental data (and assuming that the correct groups were those based on the keying out of the biological samples using the differential species), yielded a slightly higher success rate than the 228 site classification at Levels 1 and 2 but a lower success rate at Levels 3, 4 and 5 (Table 11). Incorrect predictions in which the correct group was the second most probable on environmental criteria were found to account for a high proportion of incorrect predictions as far as Level 4. It seems probable that in attempting to classify new sites on environmental criteria only, having taken measures to ensure that the new sites come within the scope of the 268 site classification, we may expect to have a success rate somewhere between 50% and 76% in prediction at Level 4 (sixteen groups). Furse *et al.* (1984) used additional procedures to demonstrate that even when an incorrect group prediction is made, the similarity between the fauna of the group predicted and the actual fauna at a site is often high enough for the prediction to be of value.

Discussion

This project was conceived as a practical response to the need for a greater understanding of the variation in running-water sites and their macro-invertebrate communities in Great Britain. The difficulties inherent in bringing together existing data-sets collected for differing purposes emphasized the need for a new country-wide data-base of information on the physical, chemical and biological attributes of a wide variety of running-water sites (Peters, 1976).

Hawkes (1975) discussed the value to the water industry of a biological classification coupled to a system of prediction of the type of macro-invertebrate community using physical and chemical features. He emphasized the need for a classification in the field of water quality and river pollution, where existing biological indices of water quality generally disregard natural differences in the biota associated with contrasted sites. In addition, a classification would be of value to the Nature Conservancy Council as a means of checking the adequacy of the range of rivers listed in the Nature Conservation Review (Ratcliffe, 1977). Work on the macrophytes of British rivers is also in progress (Holmes, 1980) and as over 150 of the river lengths sampled by Holmes are coincident with or close to sites sampled in this project we intend to compare and contrast classifications derived independently using invertebrates and macrophytes. Beyond these immediate applications, a comprehensive classification should be helpful in comparing the results of studies on different rivers, thereby leading to a greater understanding of river ecology (Hawkes, 1975; Cushing *et al.*, 1980).

Of the many classifications reviewed by Hawkes (1975), none offer a scheme with widespread application in Great Britain, based on the macro-invertebrate fauna which may be compared with the classification presented in this paper. However, several new studies have now been published. In France, Verneaux (1973, 1976) proposed a scheme consisting of ten typological levels which succeed one another along a theoretical flowing water system from the source to the mouth. By examining 240 sites on thirty rivers as a single system he attempted to integrate information on the invertebrate and fish fauna with physical features, including distance from source, mean width, slope and maximum mean temperature of the hottest month, along a theoretical river system. Each typological level corresponded to a grouping of species (biocoenotype) which did not necessarily belong to the same biocoenosis but which did possess similar ecological characteristics. Methods were then devised for determining the theoretical ecological type of a given station using macro-invertebrates, fish, or physical data. The purpose of this scheme was to ensure that rational water quality objectives and quality criteria were devised for each watercourse based on knowledge of the ecological type.

More recently, Persoone (1979) suggested a method for the biotypological classification of watercourses in the Member States of the European Community which could be implemented by a large group of hydrobiologists with the minimum of equipment or specialist taxonomic training. He proposed the collection of physico-chemical data on just five parameters coupled with information on the macrophytic vegetation and macroscopic fauna over a large number of stream sections. Analyses would then lead to the establishment of a definitive number of habitats and biocoenotypes present in the watercourses of the community. Although the advantages of a simple system applied consistently over a wide area have great appeal, the dangers of having to prejudge the physico-chemical parameters to be recorded should not be underestimated.

In Scotland, Maitland (1979) has undertaken detailed studies on the flora, fauna and environmental features of both standing and running waters in Shetland, the Outer Hebrides and in the Tay catchment. In England, Townsend, Hildrew & Francis (1983) have examined the invertebrate and fish fauna of thirty-four stony riffle stream sites in the Ashdown Forest area of Sussex. They used TWINSPAN to classify their sites and DECORANA and also stepwise multiple regression analysis in an attempt to assess the importance of a variety of physico-chemical factors in determining the distribution of species and the structure of communities. These detailed studies within limited geographical areas contrast with, but also complement our own approach, in which a much wider range of running-water sites are examined in rather less detail in order to describe large-scale pattern.

In North America, Pennak (1979) suggested seven broad lotic categories to be found in the U.S.A. but conceded that there remain many others which are difficult to categorize as they are so variable from place to place. He suggested that the thirteen physical, chemical and biological criteria listed in his earlier paper (Pennak, 1971) be used for characterizing these additional types. Cushing *et al.* (1980) classified thirty-four stream stations (thirty in North America and four in Europe) using fifteen physical and chemical variables but used cluster analysis for the initial grouping of sites followed by multiple discriminant analysis to display the variables which made the greatest contribution to the separation of the groups in discriminant space. Their results demonstrate the value of this approach, but they emphasize that as yet it is not possible to construct a simple and universally useful classification of running waters. We believe that further collaborations of this type involving the use of multivariate techniques to examine the physical and chemical data for a wide variety of sites may be the logical way to integrate the intensive classification exercises now underway in different countries which have very varied macro-invertebrate communities.

Ordination of the species data for the 268 sites in the present study (Fig. 2) failed to reveal clearly separated clusters. This was anticipated because the physical gradients along river systems favour progressively changing assemblages of species from source to mouth (Vannote *et al.*, 1980). Although zones characterized by sharply defined macro-invertebrate communities should not be expected, the entry of a major tributary or a sudden change in the physical attributes or chemical status of the river can lead to changes in the macro-invertebrate community. For example. the substantial physical changes in the characteristics of the River Spey as it passes through the Insh Marshes (Site 259) and the pool samples on the River Wansbeck (231) and River Forth (239 and 241) produced changes in the invertebrate community in a limited reach.

In view of the lack of natural discontinuities between sites, any classification is bound to be somewhat arbitrary. It can, however, still be of value, so long as it is appropriate to the needs of the user. It has already been stressed that the classification is of sites rather than sections of rivers but it is clear from Table 5 that many rivers show progressive changes in TWINSPAN group downstream and that, given sufficient sampling, it would be possible to assign stretches of river to particular TWINSPAN groups. Clearly, this procedure would only be valid in the absence of atypical physical and/or chemical features at unsampled sites which are capable of modifying the invertebrate community to the extent of changing the TWINSPAN group. Where a stretch of river displayed both pools and riffles, it would be necessary to designate the section on riffle samples only. At a later date it may be possible to expand the classification to include a comprehensive series of pool samples.

We are also aware that Man has already had an impact on some of the sites used in the classification. Hawkes (1975) lists pollution, the effects of changes in land use and changes due to impoundments and canalization as some of the most obvious influences on rivers. As far as was practical, polluted sites were excluded from the present classification but in order to be of practical value the scheme cannot be restricted to sites thought to be uninfluenced by Man. This is particularly relevant to lowland Britain.

Although the classification based on combined species data from three seasons is regarded as the most reliable, we have also generated fifteen additional classifications based on different taxonomic treatments and/or sampling regimes (Furse *et al.*, 1984). These alternatives are of considerable importance where limited time or expertise preclude the use of the classification presented in this paper. Furse *et al.* (1984) also discuss the circumstances in which the different classifications are likely to be of greatest value.

Multiple discriminant analysis, already widely used in taxonomy and in some areas of ecology, has rarely been utilized in benthic ecology (King, 1981), although it has been strongly advocated by Green & Vascotto (1978) as a means of analysing the environmental factors controlling patterns of species composition in aquatic communities. Some of the additional uses of MDA are listed by Wiegleb (1981) who, like us, used the technique for predictive purposes. In this study, we achieved 76.1% success in the prediction of group membership in the sixteen group classification using MDA with twenty-eight environmental features as variables. This percentage is higher than the success rate achieved in seven other classifications which

were subjected to the same procedure in Furse *et al.* (1984). A further 15.3% of sites were wrongly predicted but the correct group was listed as the second most probable by MDA. These 'near misses' give substance to the belief that the incorporation of additional environmental variables may improve our current ability to predict macro-invertebrate community type as expressed in the TWINSPAN group.

Many environmental variables are correlated with each other, and there is a danger that importance in influencing macro-invertebrate community type will be falsely ascribed to some factors which are themselves highly correlated with factors of genuine importance. This point is particularly relevant when examining individual correlations between environmental variables and ordination scores. In the present paper, correlation coefficients between environmental variables and ordination scores (Table 2) suggest that substratum type, water chemistry and factors which express position downstream are important, and many of these factors also have high coefficients in the listing of standardized canonical discriminant function coefficients (Table 9). In this study it has not been possible to consider temperature, flow regime or the contribution made by changing food resources available along river systems. However, we hope to incorporate some of these variables in future work. Smith (1981) has put forward procedures for estimating annual and monthly river temperatures using air temperatures only. This may offer a way of obtaining the relevant data. Whereas the number of stations at which river temperatures are monitored regularly in Great Britain only just exceeds seventy, the network of stations monitoring air temperature is very considerable.

Jones & Peters (1977) have also demonstrated that there is a significant relationship between the discharge regime at a site and its invertebrate community. Sites with a very stable discharge regime (e.g. chalk streams) have invertebrate communities which are very different from those characteristic of 'spatey' rivers. The Base Flow Index (BFI) (NERC, 1978), developed by the Institute of Hydrology, is a measure of the proportion of a river's runoff that derives from stored sources, and this index of 'spatiness' will be incorporated in future analyses.

The initial classification is now being enlarged to include more rivers in Scotland, Wales and lowland England and attempts are being made to improve predictive ability by the inclusion of further environmental features. In addition the system must be thoroughly tested before it can be recommended as a practical management tool. However, we believe that the results already available are a convincing demonstration of the value of this approach. Finally, we hope to use the biological and environmental data obtained during this work as a starting point for a more fundamental examination of the factors which determine the macro-invertebrate community at a site.

Acknowledgments

This study was funded jointly by the Natural Environment Research Council and the Department of the Environment (contract number DGR/480/520). We are particularly grateful for the assistance provided by the Water Authorities (Wales and England) and River Purification Boards (Scotland) without whose cooperation this work would not have been possible. We also wish to thank Linda Bird and John Blackburn for their substantial contribution to the sorting of samples and identification of specimens. Jon Bass and Terry Gledhill identified all specimens of Sphaeriidae and Hydrachnellae respectively. Occasional taxonomic assistance was also provided by R. J. W. Aldridge, G. J. Bird, D. A. Cooling, J. M. Elliott, G. N. Foster, A. G. Irwin, M. P. Kerney, M. Ladle, T. T. Macan, R. W. Martin. L. C. V. Pinder, P. A. Tullett and I. D. Wallace. We thank Ralph Clarke for his contributions to data analysis and Cambridge University for allowing us to use their computing facilities. Richard Mann kindly read and criticized the manuscript. Diana Morton typed the text.

References

Armitage P.D., Furse M.T. & Wright J.F. (1979) A bibliography of works for the identification of freshwater invertebrates in the British Isles. *Occasional Publications of the Freshwater Biological Association*, **5**, 1–18.

Armitage, P.D., Moss D., Wright J.F. & Furse M.T. (1983) The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. *Water Research*, **17**, 333–347.

Burton G.W. & Odum E.P. (1945) The distribution of stream fish in the vicinity of Mountain Lake, Virginia. *Ecology*, **26**, 182–194.

Chesters R.K. (1980) Biological Monitoring Working Party. The 1978 national testing exercise. *Department of the Environment, Water Data Unit Technical Memorandum*, **19**, 1–37.

Culp J.M. & Davies R.W. (1980) Reciprocal averaging and polar ordination as techniques for analyzing macroinvertebrate communities. *Canadian Journal of Fisheries and Aquatic Sciences*, **37**, 1358–1364.

Cushing C.E., McIntire C.D., Sedell J.R., Cummins K.W., Minshall G.W., Petersen R.C. & Vannote R.L. (1980) Comparative study of physical-chemical variables of streams using multivariate analyses. *Archiv für Hydrobiologie*, **89**, 343–352.

Daniels R.E. (1978) Floristic analyses of British mires and mire communities. *Journal of Ecology*, **66**, 773–802.

Furse M.T., Moss D., Wright J.F. & Armitage P.D. (1984) The influence of seasonal and taxonomic factors on the ordination and classification of running-water sites in Great Britain and on the prediction of their macro-invertebrate communities. *Freshwater Biology*, **14**, 257–280.

Furse M.T., Wright J.F., Armitage P.D. & Moss D. (1981) An appraisal of pond-net samples for biological monitoring of lotic macro-invertebrates. *Water Research*, **15**, 679–689.

Gauch H.G. & Whittaker R.H. (1981) Hierarchical classification of community data. *Journal of Ecology*, **69**, 537–557.

Gauch H.G., Whittaker R.H. & Singer S.B. (1981) A comparative study of non-metric ordinations. *Journal of Ecology*, **69**, 135–152.

Gauch H.G., Whittaker R.H. & Wentworth T.R. (1977) A comparative study of reciprocal averaging and other ordination techniques. *Journal of Ecology*, **65**, 157–174.

Green R.H. & Vascotto G.L. (1978) A method for the analysis of environmental factors controlling patterns of species composition in aquatic communities. *Water Research*, **12**, 583–90.

Harris R.J. (1975) *A Primer of Multivariate Statistics*. Academic Press, New York.

Hawkes H.A. (1975) River zonation and classification. *River Ecology* (Ed. B. A. Whitton), pp. 312–374. Blackwell Scientific Publications, Oxford.

Hill M.O. (1973) Reciprocal averaging: an eigenvector method of ordination. *Journal of Ecology*, **61**, 237–249.

Hill M.O. (1979a) *DECORANA—A FORTRAN Program for Detrended Correspondence Analysis and Reciprocal Averaging*. Ecology and Systematics, Cornell University, Ithaca, New York.

Hill M.O. (1979b) *TWINSPAN—A FORTRAN Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of the Individuals and Attributes*. Ecology and Systematics, Cornell University, Ithaca, New York.

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. *Journal of Ecology*, **63**, 597–613.

Hill M.O. & Gauch H.G. (1980) Detrended correspondence analysis: an improved ordination technique. *Vegetatio*, **42**, 47–58.

Holmes N.T.H. (1980) Preliminary results from river macrophyte survey and implications for conservation. *Nature Conservancy Council, Chief Scientist's Team Notes*, **24**, 1–68.

Huet M. (1949) Aperçu des relations entre la pente et les populations des eaux courantes. *Schweizerische Zeitschrift für Hydrologie*, **11**, 333–351.

Huet M. (1954) Biologie, profils en long et en travers des eaux courantes. *Bulletin Français de Pisciculture*, **175**, 41–53.

Illies J. (1953) Die besiedlung der Fulda (insbes. das benthos der salmoniden-region) nach dem jetzigen stand der untersuchungen. *Bericht der Limnologischen Flussstation Freudenthal*, **5**, 1–28.

Illies J. (1961) Versuch einer allgemein biozönotischen Gleiderung der 'Fliessgewasser. *Internationale Revue der Gesamten Hydrobiologie*, **46**, 205–213.

Illies J. & Botosaneanu L. (1963) Problèmes et methodes de la classification et de la zonation écologique des eaux courantes, considerées surtout du point de vue faunistique. *Mitteilungen, Internationale Vereinigung für Theoretische und Angewandte Limnologie*, **12**, 1–57.

Jones H.R. & Peters J.C. (1977) Physical and biological typing of unpolluted rivers. *Water Research Centre, Technical Report*, **41**, 1–48.

King J.M. (1981) The distribution of invertebrate communities in a small South African river. *Hydrobiologia*, **83**, 43–65.

Klecka W.R. (1975) Discriminant analysis. *SPSS. Statistical Package for Social Sciences* (Eds N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner and D. H. Bent), pp. 434–467. McGraw-Hill, New York.

Maitland P.S. (1966) *The Fauna of the River Endrick. Studies on Loch Lomond*, 2. Published for University of Glasgow by Blackie, Glasgow.

Maitland P.S. (1977) *A Coded Checklist of Animals Occurring in Fresh Water in the British Isles*. Natural Environment Research Council.

Maitland P.S. (1979) *Synoptic Limnology. The Analysis of British Freshwater Ecosystems*. Natural Environment Research Council.

NERC (Natural Environment Research Council) (1978) *Low Flow Study Report Number 1*. Institute of Hydrology, Wallingford.

Pennak R.W. (1971) Towards a classification of lotic habitats. *Hydrobiologia*, **38**, 321–334.

Pennak R.W. (1979) The dilemma of stream classification. *Classification, Inventory, and Analysis of Fish and Wildlife Habitat*, pp. 59–66. Biological Services Program, U.S. Fish and Wildlife Service.

Persoone G. (1979) Proposal for a biotypological classification of water-courses in the European Communities. *Biological Indicators of Water Quality* (Eds A. James and L. Evison), pp. 7.1–7.32. Wiley & Sons, New York.

Peters J.C. (1976) Ecological problems in the management of water resources. *Proceedings of the Royal Society of Arts Conference on Recreation and Conservation in Water Areas*, 23 November 1976.
Pielou E.C. (1977) *Mathematical Ecology*. Wiley & Sons, New York.
Ratcliffe D.A. (Ed.) (1977) *A Nature Conservation Review*. Cambridge University Press.
Resh V.H. & Unzicker J.D. (1975) Water quality monitoring and aquatic organisms: the importance of species identification. *Journal of the Water Pollution Control Federation*, **47**, 9–19.
Schmitz W. (1957) Die Bergbach-Zoozönosen und ihre Abgrenzung, dargestellt am Beispiel der oberen Fulda. *Archiv für Hydrobiologie*, **53**, 465–498.
Smith K. (1981) The prediction of river water temperatures. *Hydrological Sciences Bulletin*, **26**, 19–32.
Townsend C.R., Hildrew A.G. & Francis J. (1983) Community structure in some southern English streams: the influence of physico-chemical factors. *Freshwater Biology*, **13**, 521–544.
Vannote R.L., Minshall G.W., Cummins K.W., Sedell J.R. & Cushing C.E. (1980) The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, **37**, 130–137.
Verneaux J. (1973) Cours d'eau de Franche-Comté (massif du Jura). Recherches écologiques sur le réseau hydrographique du Doubs—Essai de biotypologie. *Annales Scientifiques de l'Université de Besançon. 3e Serie, Zoologie*, **9**, 1–260.
Verneaux J. (1976) Fondements biologiques et écologiques de l'étude de la qualité des eaux continentales—Principales méthodes biologiques. *La Pollution des Eaux Continentales. Incidence sur les Biocenoses Aquatiques* (Ed. P. Pesson), pp. 229–285. Gauthier-Villars, Paris.
Wiegleb G. (1981) Application of discriminant analysis on the analysis of the correlation between macrophyte vegetation and water quality in running waters of Central Europe. *Hydrobiologia*, **79**, 91–100.

(Manuscript accepted 1 August 1983)

APPENDIX 1. Key to sixteen groups of running-water sites by means of qualitative two-way indicator species analysis. At each stage in the key add together the scores of all the species which are present at the site being classified. This should be used in conjunction with Fig. 3.

Note: The classification on which this key is based incorporates permanent, unpolluted flowing-water sites in freshwater. Most sites were ⩾5 km from the source and, where pool and riffles alternate, almost all samples were taken from riffles. When attempting to classify new sites ensure that they fall within the scope of the classification and have been subjected to the same level of sampling intensity used for the sites on which the classification is based.

1. *Asellus aquaticus* (L.) (+1) — *Leuctra fusca* (L.) (−1)
   Baetis muticus (L.) (−1) — *Isoperla grammatica* (Poda) (−1)
   Rhithrogena semicolorata gp (−1) — *Esolus parallelepipedus* (Müller) (−1)
   Ecdyonurus sp. (−1) — *Rhyacophila dorsalis* (Curtis) (−1)
   Amphinemura sulcicollis (Stephens) (−1) — *Glossosoma* sp. (−1)

   Score −2 or less, go to 2
   Score −1 or more, go to 3

2. *Pisidium subtruncatum* Malm (+1) — *Amphinemura sulcicollis* (Stephens) (−1)
   Stylaria lacustris (L.) (+1) — *Leuctra inermis* Kempny (−1)
   Glossiphonia complanata (L.) (+1) — *Simulium equinum* (L.) (+1)
   Hygrobates fluviatilis (Ström) (+1) — *Simulium monticola* gp (−1)

   Score 1 or less, go to 4
   Score 2 or more, go to 5

3. *Stylodrilus heringianus* Claparède (−1) — *Elmis aenea* (Müller) (−1)
   Cloeon dipterum (L.) (+1) — *Eukiefferiella* sp. (−1)
   Sigara dorsalis (Leach) (+1) — *Simulium ornatum* gp (−1)

   Score −1 or less, go to 6
   Score 0 or more, go to 7

4. *Hydraena gracilis* Germar (−1) — *Eloeophila* sp. (−1)
   Hydropsyche instabilis (Curtis) (−1) — *Brillia modesta* (Meigen) (−1)
   Drusus annulatus Stephens (−1) — *Prodiamesa olivacea* (Meigen) (−1)
   Halesus gp (−1) — *Simulium ornatum* gp (−1)

   Score −6 or less, go to 8
   Score −5 or more, go to 9

5. *Ecdyonurus* sp. (−1) — *Macropelopia* sp. (+1)
   Caenis moesta gp (−1) — *Prodiamesa olivacea* (Meigen) (+1)
   Esolus parallelepipedus (Müller) (−1) — *Simulium reptans* gp (−1)
   Glossosoma sp. (−1) — *Atherix ibis* (Fabricius) (−1)

   Score −1 or less, go to 10
   Score 0 or more, go to 11

6. *Bithynia tentaculata* (L.) (−1) — *Polycentropus flavomaculatus* (Pictet) (−1)
   Nais elinguis Müller (+1) — *Athripsodes cinereus* (Curtis) (−1)
   Lebertia porosa Thor (−1) — *Synorthocladius semivirens* (Kieffer) (−1)
   Haliplus lineatocollis (Marsham) (+1) — *Microtendipes* sp. (−1)

   Score −3 or less, go to 12
   Score −2 or more, go to 13

7. *Bithynia leachi* (Sheppard) (−1) — *Graptodytes pictus* (Fabricius) (+1)
   Sphaerium corneum (L.) (−1) — *Mystacides azurea* (L.) (−1)
   Pisidium henslowanum (Sheppard) (−1) — *Prodiamesa olivacea* (Meigen) (−1)
   Enchytraeidae (+1)

   Score −2 or less, go to 14
   Score −1 or more, go to 15

8. *Baetis muticus* (L.) (+1) — *Plectrocnemia conspersa* (Curtis) (−1)
   Oulimnius tuberculatus (Müller) (−1) — *Heterotrissocladius* sp. (−1)

   Score −2 or less, ... Group 16
   Score −1 or more, ... Group 17

9. *Potamopyrgus jenkinsi* (Smith) (+1) — *Orectochilus villosus* (Müller) (+1)
   Lumbriculus variegatus (Müller) (+1) — *Ceratopogonidae* (+1)
   Gammarus pulex (L.) (+1) — *Simulium ornatum* gp (+1)

   Score 2 or less, ... Group 18
   Score 3 or more, ... Group 19

10. *Baetis niger* (L.) (+1) — *Anabolia nervosa* Curtis (+1)

    Score 1 or less, ... Group 20
    Score 2 ..... Group 21

11. *Aulodrilus pluriseta* Piguet (−1) — *Hydropsyche pellucidula* (Curtis) (−1)

    Score −1 or less, ... Group 22
    Score 0 ..... Group 23

12. *Pisidium nitidum* Jenyns (+1) — *Cyrnus trimaculatus* (Curtis) (−1)
    Hygrobates fluviatilis (Ström) (+1) — *Hydropsyche pellucidula* (Curtis) (+1)
    Baetis scambus gp (+1) — *Eukiefferiella* sp. (+1)
    Baetis rhodani (Pictet) (+1) — *Simulium ornatum* gp (+1)
    Stictotarsus duodecimpustulatus (Fabricius) (−1)

    Score 2 or less, ... Group 24
    Score 3 or more, ... Group 25

13. *Tubifex tubifex* (Müller) (+1) *Hydroptila* sp. (−1)
Ephemerella ignita (Poda) (−1) *Potthastia longimana* gp (−1)
Hydropsyche pellucidula (Curtis) (−1)

Score −2 or less, ... Group 26
Score −1 or more, ... Group 27

14. Hydridae (+1)

Score 0 ..... Group 28
Score 1 ..... Group 29

15. *Polycelis nigra* gp (−1)

Score −1 ..... Group 30
Score 0 ..... Group 31