Project 001

Draft Final Report R&D Project ~~010~~ 001

The development of macroinvertebrate indicator keys using TWINSPAN classification: A manual for NRA biologists

WRc plc

November 1992

R&D 001/10/W

THE DEVELOPMENT OF MACROINVERTEBRATE INDICATOR KEYS USING TWINSPAN CLASSIFICATION: A MANUAL FOR NRA BIOLOGISTS

G P Rutt* and C P Mainstone

* Now with NRA Welsh Region, Llanelli

Research Contractor:
WRc plc
Henley Rd Medmenham
PO Box 16 Marlow
SL7 2HD

National Rivers Authority
Rivers House Waterside Drive
Almondsbury Bristol BS12 4UD

NRA Draft Final Report 001/10/W

National Rivers Authority
Rivers House Waterside Drive
Almondsbury Bristol BS12 4UD

Tel: 0454 624400
Fax: 0454 624409

This report is the result of work jointly funded by the National Rivers Authority and the Scotland and Northern Ireland Forum for Environmental Research.

Dissemination status
Internal: Restricted
External: Restricted

Research contractor
This document was produced under R&D Contract 001 by:

WRc plc
Henley Rd Medmenham
PO Box 16 Marlow
SL7 2HD

Tel: 0491 571531
Fax: 0491 579094

WRc Report Nº NR 3195/4206

NRA Project Leader
The NRA's Project Leader for R&D Contract 001:

F Jones - NRA Welsh Region

Additional copies
Further copies of this document may be obtained from Regional R&D Co-ordinators or the R&D Section of NRA Head Office.

001/10/W

CONTENTS

CONTENTS (Continued) Page

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Reductions in river quality due to organic pollution from livestock farming are a concern in the UK and elsewhere. Effective control is hampered by the sheer number and widespread nature of potential polluting sources. Biological indicator keys offer a simple and rapid means of identifying pollution problems, whereupon remedial action can be instigated.

A rapid appraisal technique has been developed under NRA Project Reference 001, using indicator species within the macroinvertebrate community and the occurrence of sewage fungus. Indicator keys have been developed in West Wales for use throughout most of the year, but since species composition varies with the physico-chemical environment, it will be necessary to develop keys using other indicator species in significantly different habitats and/or geographical areas.

This Manual describes the steps necessary for the development of a macroinvertebrate indicator key, using the TWINSPAN (Two-Way INdicator SPecies ANalysis) computer programme. Further details of this R&D Project, which places the rapid appraisal technique in a proposed overall farm pollution control strategy, are given in the associated R&D Report (NRA Reference 001/11/W) and R&D Project Record (NRA Reference 001/9/W).

KEY WORDS

Biological monitoring, pollution control, farm pollution

1. INTRODUCTION

Macroinvertebrate indicator systems based on TWINSPAN classification (Hill 1979) form the basis of the classification element of RIVPACS (Wright *et al* 1984). They have also shown promise as a means of identifying streams subject to surface water acidification (Wade, Ormerod and Weatherley 1989; Rutt, Weatherley and Ormerod 1990) and in the assessment of farm pollution in West Wales (Reynolds 1989). For certain types of survey such systems may have distinct advantages over widely-used biotic indices such as BMWP score and ASPT, and especially over chemical monitoring. They have the following properties:

- Can be used on site, taking only 15-20 minutes.
- Few taxa to identify.
- Can be based upon abundance making the systems more robust.
- Can be used by non-specialists.
- Yield a simple classification for communication of results to pollution control staff.
- Can cover a large number of sites in a day, enabling comprehensive coverage of catchments and identification of pollution sources.

A recent NRA R&D programme (R&D Project 001) has developed such systems for use in the assessment and control of organic pollution from farms (NRA R&D Report ***; Seager, Jones and Rutt 1992; Rutt, Pickering and Reynolds, In press). They are particularly appropriate because there are a very large number of sources for such pollution in the U.K. and they tend to be spread over large areas with poor accessibility (NCC 1991). It is envisaged that areas for the use of indicator systems might be targeted using pollution risk assessment maps produced during the same project by consideration of factors such as organic waste loading, rainfall, topography and soil type (NRA R&D Report 001/11/W). Indicator systems might be used both to identify polluted watercourses and polluting farms and to evaluate the effectiveness of remedial action resulting from a programme of farm visits. Currently indicator systems are available which are applicable to a large part of Wales and Devon and probably to other upland areas in the Midlands, Northern England and Scotland.

This manual describes how to develop an indicator key for a local area using as an example a key designed for use in the detection of farm pollution in Winter and Spring in West Wales. A schematic overview of the procedure is provided (Figure 1.1). Systems designed for farm pollution are also likely to be sensitive to other organic pollutants such as discharges from sewage treatment works and combined sewer overflows. Keys could be developed for other forms of pollution by following the principles described.

Details concerning the use of rapid appraisal techniques in the field for pollution control are given in Appendices A and B.

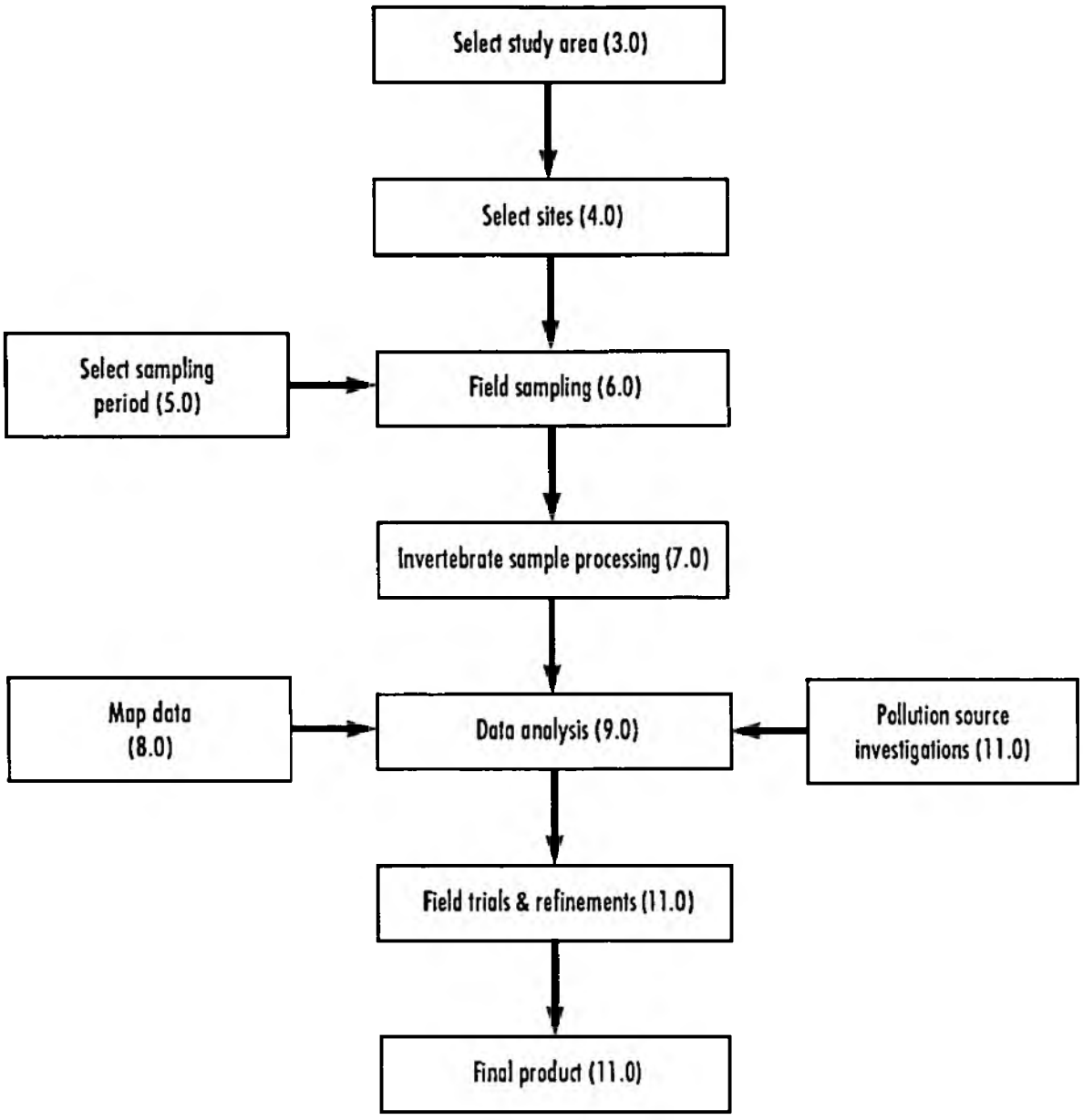


Figure 1.1 Schematic representation of the processes involved in the development of a rapid appraisal system. Numbers in brackets indicate relevant section number in text.

2. RESOURCE IMPLICATIONS

2.1 Materials

2.1.1 Fieldwork and sample processing

The following equipment is required for fieldwork and sample processing, all of which should be readily available to NRA biologists.

Waders
Kicksampling net (1 mm mesh)
Metre rule
Large sample pots for macroinvertebrates
Small sample pots for 'sewage fungus' samples
Fixative/preservative such as Formalin or IMS
Sorting trays
Forceps
Low-power binocular microscope
High-power microscope (for confirmation of sewage fungus)
Identification keys

2.1.2 Data storage and analysis

The following items of software are required for data storage and handling:

TWINSPAN, which can be run on a microcomputer or Mainframe.
A simple statistical package e.g. MINITAB

2.2 Staff resources

If the recommended fifty sites are sampled, it is estimated that the following amounts of staff time would be required to develop an indicator key:

Site selection and route planning	3 Man-days
Fieldwork	8 Man-days
Sample processing	30 Man-days
Data Input	3 Man-days
Data Analysis	2 Man-days
Field trials and adjustments	2 Man-days
Total	48 Man-days

Note

1. If electrofishing data is required to provide validation for the classification, fieldwork time must be increased by about 40 man-days assuming a team of three covering three sites per day.

2. If spot samples are taken to provide chemical information on the sites, allowance must be made for the costs of laboratory analysis.

3. SELECTION OF SAMPLING AREA

3.1 Method

The study area chosen should as far as possible correspond to the area in which the indicator key is intended to be used. Such an area should not contain river systems which support fundamentally different faunas due to large natural variation in factors such as hardness or gradient.

3.2 Example

The study area selected in West Wales lay principally within the county of Dyfed and extended to the Western extremity of Wales, North as far as Aberystwyth and East as far as Llandovery (Figure 3.1). The majority of the land surface is gently undulating with few areas rising above 400 m. It is underlain principally by Palaeozoic sediments mainly of Ordovician and Silurian age with some Cambrian, Devonian and Carboniferous rocks in the extreme south and west. The principal land use is dairy and beef farming with some sheep rearing on the higher, less fertile soils.

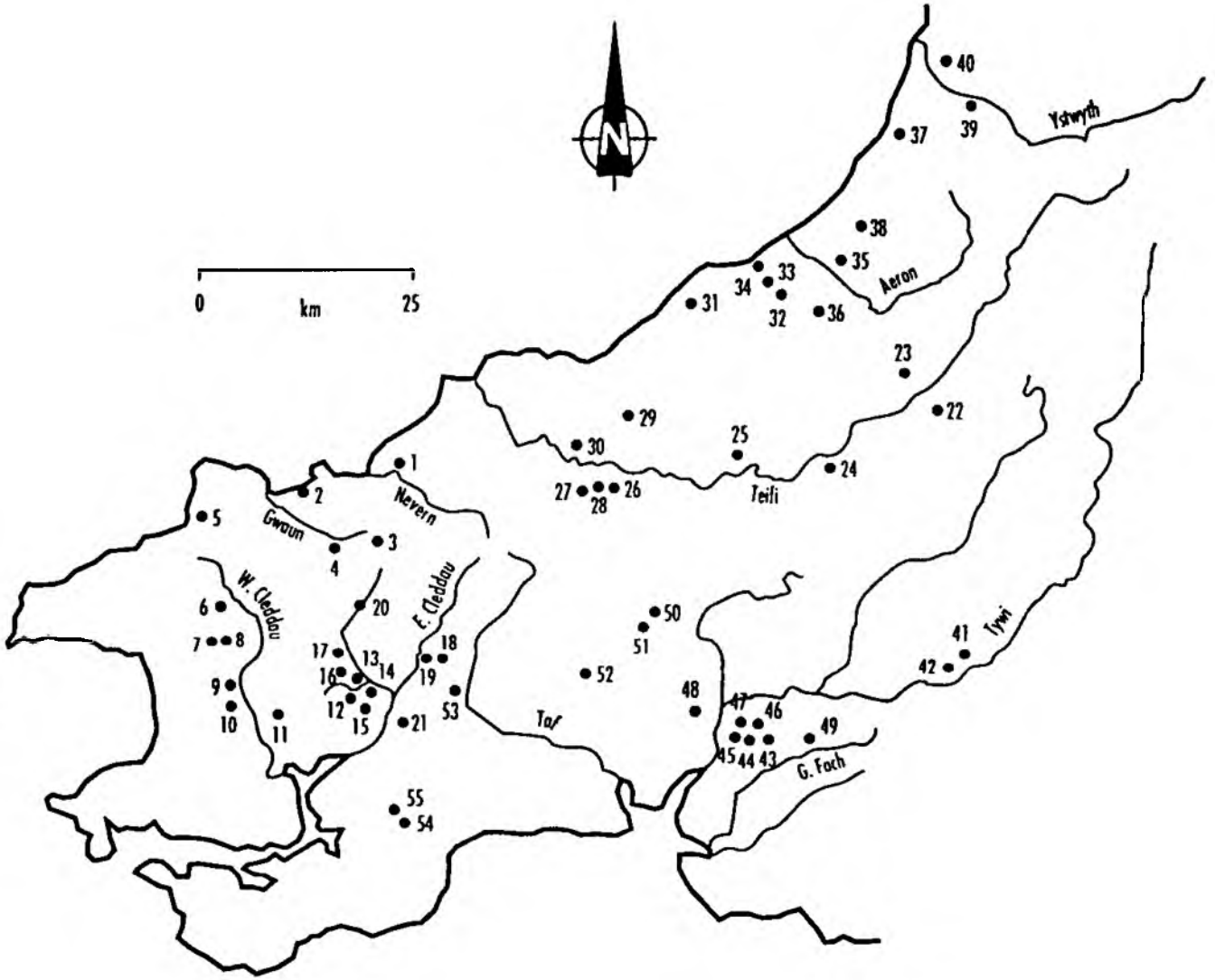


Figure 3.1 Map of West Wales showing the stream sites sampled for the development of the indicator key and their relationship to major river systems

4. SITE SELECTION

4.1 Method

It is recommended that around fifty sites are sampled. Careful liaison with local pollution control staff will be necessary and sites should be chosen so as to ensure:

- A range of pollution impact (from clean sites to grossly polluted).
- An even distribution within the chosen area of study so as to include minor local variations in fauna.
- Restriction to sites not suffering from other forms of pollution.
- Restriction to sites which do not suffer from periodic drought or similar influences.

4.2 Example

After discussion with NRA pollution control staff, 55 sites on streams in West Wales were selected for survey (Figure 3.1). Sites were chosen to include a wide range of pollution impact, pollutant type (e.g. parlour washings, lagoon overflow, silage effluent) and geographical locations. Sites were restricted to those likely to support populations of salmonids and unlikely to be affected by other sources of pollution found in rural catchments e.g. sewage treatment works, sheep dippers. Streams were mostly of first, second or third order (Strahler 1957). Stream widths were in the range 1.0-6.0 m, most sites being between 2.0 and 4.5 m.

5. CHOICE OF SAMPLING PERIOD

5.1 Method

Due to seasonal variation in macroinvertebrate fauna, sampling should be carried out so as to develop a key for a particular period of the year. Work to date indicates that keys could be developed for winter/spring (December-May) and for the summer (July-September), periods when faunal composition is relatively stable.

5.2 Example

Fieldwork in West Wales was carried out between 27 February and 6 April 1992. Following field-testing, it appears that the key is applicable to the period December to May inclusive.

6. FIELD SAMPLING

6.1 Basic procedure at each site

The following procedure should be followed at each sampling site:

1. Select a suitable sampling habitat which is consistent for all sites sampled (in fast-flowing streams this will be a riffle).

2. Record the approximate stream width and average depth.

3. Record the substratum composition in the habitat sampled under the following categories:-Bedrock, Boulders/cobbles (> 6 cm), Pebbles/gravel (2 mm-6 cm), Sand Silt/clay.

4. Select five large stones in the riffle and determine the approximate percentage of the surface (above and below) which is covered with growth of 'sewage fungus' (Curtis 1969). Record a mean cover value for above and below stones. If no fungus is discovered check five more large stones and instream vegetation for presence/absence of growth. Retain small quantities of the growth for later confirmation under a microscope.

5. Collect a sample of the benthic fauna by kicksampling for one minute in the riffle (Furse *et al* 1981). Samples should be preserved at the laboratory the same day using IMS (Industrial Methylated Spirits) or 4% Formalin (if storage is likely to be prolonged).

6.2 Additional procedures

6.2.1 Chemical samples

Method

Water quality data can be valuable in validating an indicator key. However, such data are relatively costly to collect and an ideal regime of continuous monitoring at all sites or several spot samples in the weeks prior to the invertebrate sampling would be prohibitively expensive. A single spot sample taken on the day of invertebrate sampling will still yield valuable information especially for more 'stable' determinands such as hardness. It should be taken immediately the sampling habitat has been selected and prior to kicksampling.

Example

In the West Wales survey a single spot water sample was taken at each site. It was later analysed for a range of standard sanitary determinands such as dissolved oxygen, Biochemical Oyygen Demand (BOD), inorganic nutrients and suspended solids.

6.2.2 Electrofishing

Method

Electrofishing data may be a further source of information to be used to validate a key. Electrofishing should take place after all other data has been collected from a site, although if nets are employed they should be installed first of all.

Example

In the West Wales study, lengths of stream in the range 30-60 m were electrofished semi-quantitatively to assess the populations of trout (*Salmo trutta*) and salmon (*Salmo salar*). One run was employed and fishing was carried out between riffles without the use of stop nets. Fish were identified and measured before being returned to the stream. All age classes were pooled to give a minimum total trout density, calculated per 100 m^2. Minimum population estimates derived in this way have been found to correlate well with the results of quantitative sampling based on catch-depletion (Strange, Aprahamian and Winstone 1989).

7. INVERTEBRATE SAMPLE PROCESSING

7.1 Method

7.1.1 Sorting

Samples should be processed at the laboratory. Once the formalin or IMS has been rinsed away, samples should be sorted in white trays. For rare taxa with abundance less than ten, all individuals should be removed from the sample. More abundant taxa can be sub-sampled by removing individuals from randomly selected sections of marked sorting trays, and estimating abundance by multiplying according to the proportion of sample sorted.

7.1.2 Identification

Identification should be to species or genus level except for time-consuming taxa, i.e. chironomids, oligochaetes and simuliids. Samples identified only to family level will provide workable keys but it has been found that the extra information obtained by greater taxonomic penetration yields a more discriminating system. The level of identification must also be geared to whatever biotic indices are to be used for validating the key. BMWP score/ASPT (Armitage *et al* 1983) require a lower degree of resolution than Trent Biotic Index or the Chandler Index.

7.2 Example

In the West Wales study, samples were generally identified to species or genus level. BMWP Score and ASPT were calculated for each site.

8. DATA DERIVED FROM MAPS

For each site, altitude, catchment area, distance from source and stream gradient should be estimated from Ordnance Survey 1:50 000 scale maps.

9. DATA HANDLING

9.1 TWINSPAN classification

9.1.1 Data Format

Method

TWINSPAN (Hill 1979) is a FORTRAN program which requires a specific data format, an example of which is given below.

```
(I2, 2(I2,I3))
1 1 20 3  5
1 4 56
2 1  8 2 21
0

SPEC 1  SPEC 2  SPEC 3  SPEC 4
SAMP 1  SAMP 4
```

This signifies that sample 1 contains species 1 with abundance 20, species 3 with abundance 5, and species 4 with abundance 56. Sample 2 contains species 1 with abundance 8 and species 2 with abundance 21. Note the 0 which terminates the data list. Data may be typed into a suitable format in a data editor or output into a suitable TWINSPAN format from a data base.

Example

For the study in West Wales, data was stored on an ORACLE data base known as BAETIS which has been developed by NRA Welsh Region (NRA 1992). Data is output from this system in the following format:

Welsh spring key

```
01(I3,I4,I7)
1  5   109
1  7     1
1  9    14
etc.
```

Percentage sewage fungus cover above stones was included as a species in the data set.

9.1.2 Running TWINSPAN

Method

TWINSPAN is a multivariate classification technique which has been widely used in freshwater ecology (e.g. Wright *et al* 1984; Ormerod and Edwards 1987). Both mainframe and P.C. versions of the FORTRAN program are available. TWINSPAN operates by splitting an original data set into 2, 4, 8 etc. groups at successive levels of division, based on the similarity of the invertebrate community at each site. For each division, the output lists indicator taxa that distinguish between site groupings. For example, part of an output might take the form:

```
DIVISION  1 (N= 55)       I.E. GROUP *
EIGENVALUE 0.210 AT ITERATION  2
INDICATORS, TOGETHER WITH THEIR SIGN
RHIT SEMI2(-)  OLIGOCHAETA3(+)
MAXIMUM INDICATOR SCORE FOR NEGATIVE GROUP  -1
MINIMUM INDICATOR SCORE FOR POSITIVE GROUP  0
ITEMS IN NEGATIVE GROUP  2  (N=  22) .....
ITEMS IN POSITIVE GROUP  3  (N=  33) .....
```

This indicates that for Division 1, the original data set is split into two groups of 22 and 33 sites. The indicators are >9 *Rhithrogena semicolorata* and > 99 *Oligochaetes*, scoring -1 and +1 respectively. If the score for a site is -1 it will be classified into the negative group, if 0 or 1 it will go into the positive group. An indicator key can be readily constructed using this information (Figure 9.1). A variety of TWINSPAN options may be invoked either by running the program interactively or via a separate 'options' file (Hill 1979). Analysis should proceed as follows:

1. When running TWINSPAN on a data set for the first time, the following options should be set:

   - Cut-off levels for 'pseudo species' (i.e. species that are counted as different indicators at different levels of abundance) should be set to the logarithmic abundance categories: 1, 10, 100, 1000, 10 000.

   - Maximum level of divisions should be set to 6 to reduce the volume of output.

   - Indicator potential of 'cut-off' levels should be set to 0, 1, 1, 1, 1, to prevent 'pseudospecies' category 1 (1-9 individuals) appearing as indicators. This prevents species that 'drift' readily from becoming important indicators in the key.

   - All other options should be set to the default values.

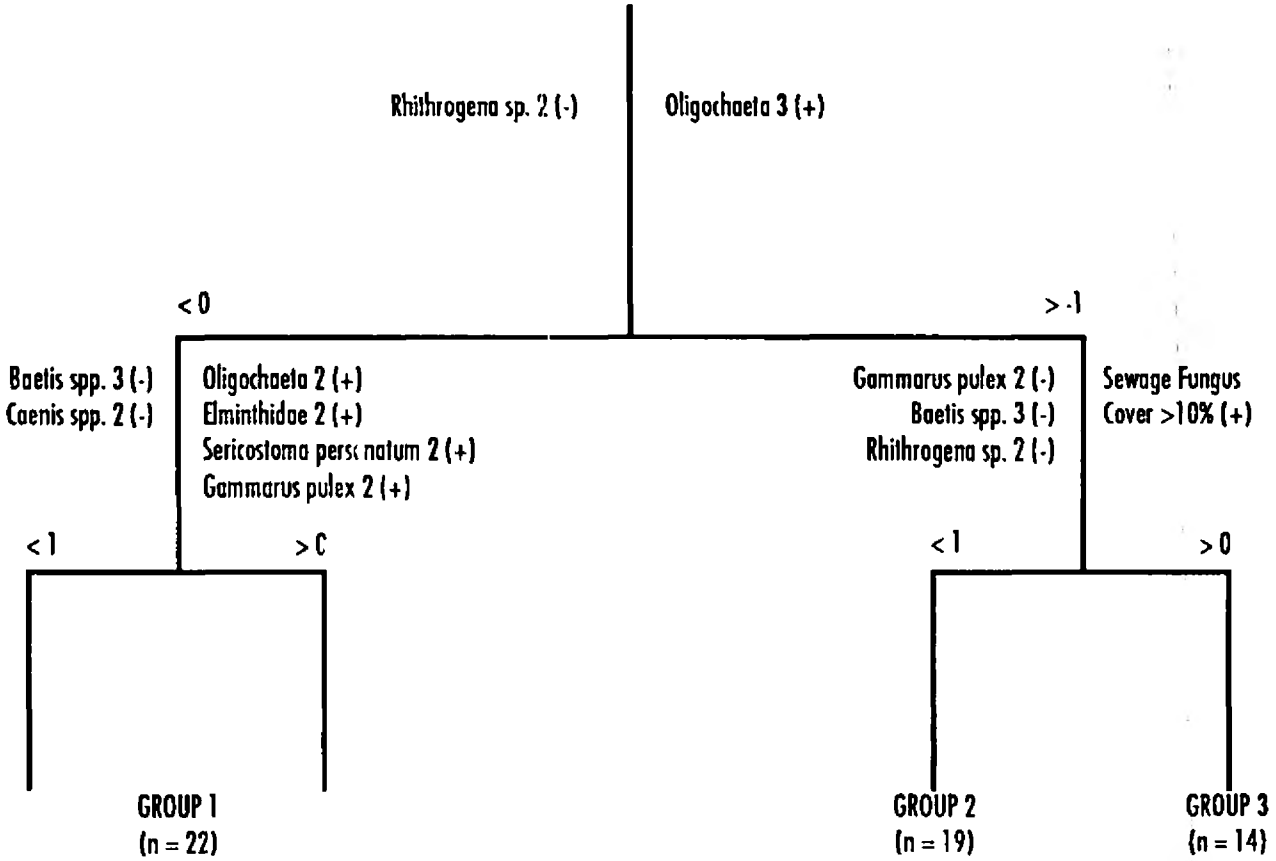


Figure 9.1 TWINSPAN indicator key derived from data from West Wales. The numerals after the taxon names indicate log abundance category (1 = 1-9, 2 = 10-99, 3 = 100-999)

2. The output should be examined to see whether there is an even pattern of division. If groups containing only one or two sites appear at division level One or Two, these outliers should be excluded from further analysis using the 'omission of samples' option.

3. When an even pattern of division has been obtained at the first two levels of division, the resultant classification will remain relatively stable despite subsequent exclusion of potential indicators.

4. It is now necessary to determine whether any taxa which are difficult to identify need to be masked out as potential indicators. At this point, the groupings obtained should be tested for differences in the degree of pollution, as indicated by biotic indices (e.g. BMWP Score and ASPT) and chemical data if available (see Section 6.2). Usually major differences will only be apparent down to the second level of division. Any groups which appear similar in degree of pollution should be considered together and their indicators ignored.

5. If taxa which are not readily enumerated in the field appear as indicators of a meaningful division they should be masked out by invoking the option to omit species as potential indicators.

6. TWINSPAN is re-run, eliminating impractical indicators until a practical indicator system is produced.

Example

Initial TWINSPAN classification of the data from West Wales produced an even pattern of division so that there was no need to exclude any of the 55 sites. Comparison of BMWP score and ASPT between groups at the second level of division (four groups) indicated that two of the groups could be combined, as TWINSPAN had forced a division when there was little ecological difference to justify one. Several inconvenient indicators such as *Leuctra*, *Chironomidae*, *Potamopyrgus jenkinsi*, *Amphinemeura* and *Chloroperla* were masked out as potential indicators to yield a simple key (Figure 9.1). Taxa to the left of each division score -1, taxa to the right +1 ; the net total for a site indicates whether it will be placed to the left or right of each division.

9.1.3 Properties of the TWINSPAN groups

Method

In order to determine whether the derived TWINSPAN groups reflect different degrees of organic pollution, the following analyses may be carried out:

1. Between-group differences in a range of biotic and abiotic variables may be examined by analysis of variance, or by non-parametric tests such as Kruskal-Wallis if preferred (Sokal and Rohlf 1981).

2. The distribution of different invertebrate taxa between the stream groups may be investigated using X^2 tests (Sokal and Rohlf 1981). Analyses may be carried out at different levels of abundance e.g. *Baetis* spp. (1) (one or more individuals in a one-minute kick sample), *Baetis* spp. (2) (10 or more individuals) and *Baetis* spp. (3) (100 or more individuals).

Example

For the key developed for West Wales, all analyses were carried out using the MINITAB statistical package. There were highly significant differences ($p < 0.001$) between the three groups in several pollution-dependent variables such as BMWP score, ASPT, ammoniacal nitrogen and minimum total trout density (Table 9.1). There were also highly significant differences ($p < 0.001$) in other variables not directly related to farm pollution, such as pH, conductivity and hardness. These differences are likely to reflect the fact that the cleaner streams tended to drain catchments with poorer soils and softer waters where dairy and beef farming are likely to be less prevalent. Less significant differences ($p < 0.05$) in distance from source and catchment area may reflect the fact that farm pollution tends to have the greatest impact on small streams (Howells and Merriman 1986).

The results of the X^2 tests (Table 9.2) supported the hypothesis of a relationship between stream group and degree of pollution. A number of pollution-sensitive taxa (including several stoneflies, mayflies and caddis) showed preference for Groups 1 or Groups 1 and 2, whilst a few tolerant taxa such as the leech *Helobdella stagnalis*, oligochaetes and chironomids showed a preference for Groups 2 and 3.

Table 9.1 Relationships between TWINSPAN groups from data set SPP1 and biotic and environmental variables

Variable (units)	Group 1 (n=22)	Group 2 (n=19)	Group 3 (n=14)	F	p
Pollution dependent variables					
BMWP	127 (101-154)	102 (75-130)	64 (38-91)	23.95	<0.001***
ASPT	6.3 (6.1-6.6)	5.6 (5.1-6.0)	5.4 (4.7-6.2)	17.64	<0.001***
Oxygen ($mg\ l^{-1}$)	11.0 (9.9-12.2)	10.7 (10.1-11.3)	10.0 (9.2-10.8)	5.68	0.006**
BOD ($mg\ l^{-1}$)	0.9 (0.2-1.7)	1.5 (0.8-2.1)	1.4 (1.1-1.8)	4.42	0.017*
Ammoniacal N + ($mg\ l^{-1}$)	0.02 (0.01-0.06)	0.12 (0.03-0.47)	0.11 (0.02-0.54)	12.04	<0.001***
Solids + ($mg\ l^{-1}$)	6.1 (3.7-10.1)	7.4 (4.7-11.5)	7.9 (4.8-13.2)	1.52	0.229
Min Trout + (per 100 m^2)	14 (8-25)	6 (2-18)	2 (1-5)	20.08	<0.001***
Pollution independent variables					
Altitude (m)	85 (28-141)	50 (19-81)	82 (22-142)	2.87	0.065
Slope (%)	3.3 (0-7.4)	1.5 (0.5-2.5)	2.7 (1.2-4.3)	2.37	0.103
Width (m)	3.2 (2.0-4.4)	2.8 (1.7-3.9)	2.4 (1.5-3.2)	2.89	0.065
Dist. from source (Km)	5.0 (2.5-7.6)	4.0 (1.5-6.5)	2.6 (1.2-3.9)	4.91	0.011*
Catchment Area (Km^2)+	6.7 (3.3-13.4)	4.6 (2.0-10.6)	3.1 (1.4-6.8)	4.30	0.019*

Table 9.1 continued

Variable (units)	Group 1 (n=22)	Group 2 (n=19)	Group 3 (n=14)	F	p
pH	7.1 (6.6-7.5)	7.4 (7.2-7.7)	7.4 (7.3-7.5)	7.93	0.001***
Conductivity ($uS\ cm^{-1}$)	153 (115-192)	234 (147-321)	175 (133-216)	9.55	< 0.001***
Hardness ($mg\ l^{-1}$)	48 (32-64)	77 (49-105)	57 (40-75)	10.05	< 0.001***

Notes: + Denotes variable was log-transformed prior to analysis
Group values are means with standard deviation ranges (+ & - 1SD) in brackets
F-statistic (F) and probability values (p) are from analysis of variance
NB Chemical data are based on the results of a single spot sample

Table 9.2 Frequency of occurrence of selected invertebrate taxa in the three TWINSPAN groups generated by analysis of the SPP1 data set

TAXON	Group 1 (n=22)	Group 2 (n=19)	Group 3 (n=14)	X^2	p
Taxa associated with Group 1					
Leuctra spp.	*****	**	***	17.7	<0.001
Protonemoura spp.	***	*	*	10.0	<0.01
Chloroperla spp.	*****	**	***	13.1	<0.01
Hydropsyche spp.	*****	**	**	21.8	<0.001
Sericostoma personatum	****	*	*	16.2	<0.001
Brachyptera risi (2)	***	**	**	6.9	<0.05
Isoperla grammatica (2)	***	*	*	7.9	<0.02
Hydropsyche spp. (2)	***	*		17.2	<0.001
Rhithrogena sp. (3)	**	*		10.9	<0.01
Amphinemeura spp.	****	**	**	8.9	<0.02
Perlodes microcephala	**	*		7.2	<0.05
Leuctra spp. (2)	***	*	*	13.7	<0.01
Chloroperla spp. (2)	***	*		17.3	<0.001
Rhithrogena sp.	*****	****	****	6.6	<0.05
Isoperla grammatica	*****	***	**	8.9	<0.02
Taxa associated with Groups 1 and 2					
Ecdyonurus spp.	****	***	*	15.3	<0.001
Caenis spp.	***	*		7.1	<0.05
Hydraena gracilis	****	***	*	17.2	<0.001
Rhithrogena sp. (2)	*****	***		32.0	<0.001
Elminthidae (2)	***	**		8.9	<0.02
Gammarus pulex	****	*****	***	13.9	<0.01
Elminthidae	*****	****	**	10.8	<0.01
Rhyacophila dorsalis	*****	****	**	15.3	<0.001
Plectrocnemia spp.	***	***		9.3	<0.01
Pisidium spp.	***	***	**	9.4	<0.01
Nemoura spp.	***	****	*	10.6	<0.01
Potamopyrgus sp. (2)	**	***	*	8.8	<0.02
Gammarus pulex (2)	**	****	*	11.3	<0.01
Ancylus fluviatilis	****	***	*	12.0	<0.01
Dicranota spp.	*****	***	***	7.9	<0.02
Baetis spp. (3)	****	***	*	11.6	<0.01

Table 9.2 continued

TAXON	Group 1 (n=22)	Group 2 (n=19)	Group 3 (n=14)	X^2	p
Taxa associated with Group 2					
Asellus spp.	*	**	*	6.8	<0.05
Habrophlebia fusca		***		26.1	<0.001
Potamopyrgus spp.	***	****	***	6.0	<0.05
Taxa associated with Groups 2 and 3					
Helobdella stagnalis	*	***	**	10.4	<0.01
Oligochaeta (3)	*	*****	***	27.1	<0.001
Chironomidae (3)	*	****	**	10.9	<0.01
Oligochaeta (2)	***	*****	*****	14.6	<0.001
Taxa not showing Group association					
Gyrinidae	***	**	*	5.9	N.S.
Glossiphonia complanata	*	**	**	5.5	N.S.
Oligochaeta	*****	*****	*****	1.4	N.S.
Baetis spp.	*****	*****	*****	0.0	N.S.
Chironomidae	*****	*****	*****	0.0	N.S.
Lumbricidae	***	***	***	0.4	N.S.
Brachyptera risi	****	****	****	0.7	N.S.
Limnephilidae	****	***	***	4.5	N.S.
Paraleptophlebia spp.	**	**	*	1.6	N.S.
Erpobdella octoculata	**	***	*	5.0	N.S.
Chironomidae (2)	****	*****	*****	5.5	N.S.
Simuliidae (2)	***	****	***	1.3	N.S.
Simuliidae	****	*****	*****	3.3	N.S.
Ceratopogonidae	**	***	***	1.8	N.S.
Baetis spp. (2)	*****	*****	****	4.0	N.S.

Notes: Asterisks indicate percentage occurrence
(* 1-20% of sites in group; ** 21-40; *** 41-60%; **** 61-80%; ***** 81-100%).
For certain taxa, abundance categories are treated as different taxa i.e., (2) indicates >9 individuals in a 1 minute kick sample, (3) > 99.
X^2 values and associated probabilities are given.

10. INVESTIGATION OF POLLUTION SOURCES

10.1 Method

Investigation of pollution sources is another way of testing the validity of the TWINSPAN groupings. The ideal approach would be for pollution control to conduct a concurrent programme of farm visits to all premises upstream of each site to identify sources of pollution. As this is unlikely to be feasible, pollution sources might be identified by walking the catchment upstream of apparently polluted sites on the day of sampling or as soon as possible after the sampling programme had been completed. Local pollution control staff could also provide anecdotal information.

10.2 Example

Between 24 April and 17 May 1990, most of the streams found to be affected by organic pollution in the survey in West Wales were investigated by walking the catchments upstream of the sampling sites. Further information was obtained from NRA Pollution Control staff (Table 10.1).

Table 10.1 Characteristics of stream sites in West Wales arranged by TWINSPAN group

(The sources of pollution thought to be responsible for the observed biological impacts are given)

Site	Stream	NGR	BMWP	ASPT	Sewage fungus(%)	Min. trout density ($100\ m^{-2}$)	Possible pollution source
Group 1							
3	Afon Cwmau	SN037339	154	6.7	0	17.3	
6	W. Cleddau Trib	SM933276	168	6.5	0	13.2	
7	Nant y Coy Brook	SM921242	100	6.3	0	30.7	
9	Camrose Brook	SM939191	103	6.1	0	16.3	
13	Deepford Brook	SM049200	139	6.0	52	3.9	Fish Kill in April 1988
19	Afon Rhydabil	SN107232	119	6.3	0	16.7	
20	Syrfnwy	SN047269	126	6.3	0	5.1	
22	Nant Eiddig	SN593452	148	6.4	0	10.1	
23	Nant Creuddyn	SN567492	148	6.2	0	16.5	
24	Afon Iar	SN500414	170	6.5	0	19.9	
25	Afon Cerdin	SN421415	78	6.0	0	9.8	
29	Afon Dulais	SN315467	142	6.2	0	15.9	
31	Afon Soden	SN373568	109	6.4	0	5.7	
35	Nant Cilcennin	SN500600	110	6.5	0	12.9	
36	Afon Feinog	SN466565	120	6.3	0	20.5	
37	Afon Carrog	SN562719	98	6.5	0	12.8	
38	Afon Arth	SN541630	157	6.8	0	14.7	
39	Nant Adal	SN624749	136	6.5	0	18.0	
40	Nant Paith	SN604787	101	6.3	0	36.9	

Table 10.1 continued

Site	Stream	NGR	BMWP	ASPT	Sewage fungus(%)	Min. trout density ($100\ m^{-2}$)	Possible pollution source
48	Nant y Ci	SN386187	107	5.6	0	2.8	
51	Pontgarreg Fach	SN316275	108	6.4	152	21.9	Leaking lagoon
55	Cresswell Trib.	SN096077	155	6.5	0	8.4	
Group 2							
5	Aberbach Stream	SM895361	147	5.7	<1	12.5	Yard run off
8	Nant y Coy Brook	SM922242	116	5.8	803	44.8	Chronic silage effluent discharge
10	Knock Brook	SM938191	81	4.8	803	14.9	Whey spread to land
11	Fenton Brook	SM973174	97	5.7	0	0	Fish kill in 1989 and yard runoff
14	Deepford Brook	SN072198	89	4.9	953	5.9	Various - intensive dairying catchment
15	Cotland Mill	SN054193	118	5.9	0	12.9	Possible yard runoff
16	Holmes Stream	SN042208	78	4.9	703	13.4	Inadequate storage of slurry
30	Afon Hirwaun	SN258424	73	5.6	0	4.5	Variety of intermittent discharges
32	Drywi	SN445585	81	5.8	5	3.2	Uncertain - history of inputs
34	"	SN438593	104	5.8	0	1.5	Uncertain - history of inputs

Table 10.1 continued

Site	Stream	NGR	BMWP
41	Nant Coch	SN661252	75
42	Gurrey Fach	SN632231	94
43	Nant Cwmffrwd	SN445165	127
44	" "	SN443165	111
45	" "	SN422174	133
47	Nant y Glaston	SN422175	68
49	Gwendraeth Trib	SN491162	97
53	Afon Rhydbennau	SN150196	167
54	River Creswell	SN096076	93
Group 3			
1	River Gammon	SN083400	32
2	Aberbach stream	SM997386	60
4	Pontfaen Brook	SN027329	99

ASPT	Sewage fungus(%)	Min. trout density ($100\ m^{-2}$)	Possible pollution source
5.0	10	1.1	Deliberate discharges from lagoons
5.5	0	4.1	Silage effluent and Yard runoff
6.0	3	0	Leaking slurry stores
5.8	5	0	Lagoon leakage and yard runoff
6.1	1	7.9	Parlour and dairy washings
4.8	853	4.2	Dungstead washings
5.4	803	32.3	Farm identified - source uncertain
6.4	1	10.3	
5.8	0	1.5	Sewage treatment works
5.3	100	2.7	Spreading of slurry, whey etc.
5.5	100	4.0	Yard washings and whey from pig farm
6.6	80	2.5	Parlour washings and silage effluent

Table 10.1 continued

Site	Stream	NGR	BMWP	ASPT	Sewage fungus(%)	Min. trout density ($100\ m^{-2}$)	Possible pollution source
12	Churchill Brook	SN050197	53	4.4	60	11.5	Overflowing slurry lagoon
17	Slade Brook	SN037224	55	5.0	30	6.1	Dairy and yard washings and road runoff
18	Afon Rhydabil	SN116232	86	6.1	0	0	Poorly contained parlour washings
21	E. Cleddau trib	SN119181	55	5.0	20	1.2	Slurry spreading and yard runoff ?
26	Durog	SN292382	106	5.9	0	0	Uncertain - history of pollution
27	Durog	SN277375	108	6.0	80	0	Whey spread to land
28	"	SN283378	59	5.9	100	0	Whey spread to land
33	Drywi	SN440592	55	5.5	0	1.5	Uncertain - history of pollution
46	Nant y Glaston	SN425175	18	3.6	95	0	Dungstead washings
50	Pontgarreg Fach	SN321280	49	5.4	80	0	Farm identified - source uncertain
52	Afon Fenni	SN251221	65	5.9	30	0	Fish Kill in 1989 and lagoon overflow

Notes: Subscripts for Sewage Fungus cover denote changes in group occasioned by modifications to key

11. PRODUCING A PRACTICAL SYSTEM

11.1 Method

It is strongly recommended that a newly developed TWINSPAN key should be tested under field conditions to identify possible improvements to make the system more practical and accurate. Non-specialist staff (Pollution Control Officers or bailiffs) who may use the key should be involved in such trials.

11.2 Example

Following short field trials in January and February 1991, it was felt necessary to incorporate sewage fungus at two extra stages in the system. These modifications eliminated the need for invertebrate assessments in cases of gross pollution and ensured that sites cannot be classified into Group 1 if sewage fungus is found to be present in visible quantities at a site, either above or below stones. Sewage fungus is a definite indicator of organic pollution and its presence even at low abundance should prevent sites being classified in the unpolluted group. The possibility of using the family Heptageniidae as an indicator rather than Rhithrogena was also investigated so as to ease identification in the field. Examination of raw data from the 55 sites showed that this would not lead to any differences in classification, so the simplification was made. The key was then prepared as a flow-chart for more extensive trials (Figure 11.1). When tested at 146 sites in 15 sub-catchments in February/March 1991, 96 sites fell into the intermediate Group 2; this Group was subsequently split so as to provide better discrimination in the mid-range of pollution impact (Figure 11.2). The validity of this final version of the key was confirmed by analysis of variance of biotic indices for these sites (Table 11.1).

Table 11.1 Relationship between pollution groups indicated by revised indicator key for West Wales (five groups) and biological quality

	n	BMWP Score	ASPT
Group 1	28	123 (104-142)	6.4 (5.9-6.9)
Group 2a	30	122 (102-142)	6.3 (5.8-6.7)
Group 2b	40	89 (58-120)	5.5 (4.7-6.3)
Group 2c	23	76 (47-106)	5.4 (4.7-6.1)
Group 3	19	59 (26-92)	4.6 (3.5-5.8)
F		27.30	23.17
p		<0.001	<0.001

Notes: Group values are means with standard deviation ranges in brackets.
The F statistic (F) and probability value (p) are from analysis of variance.

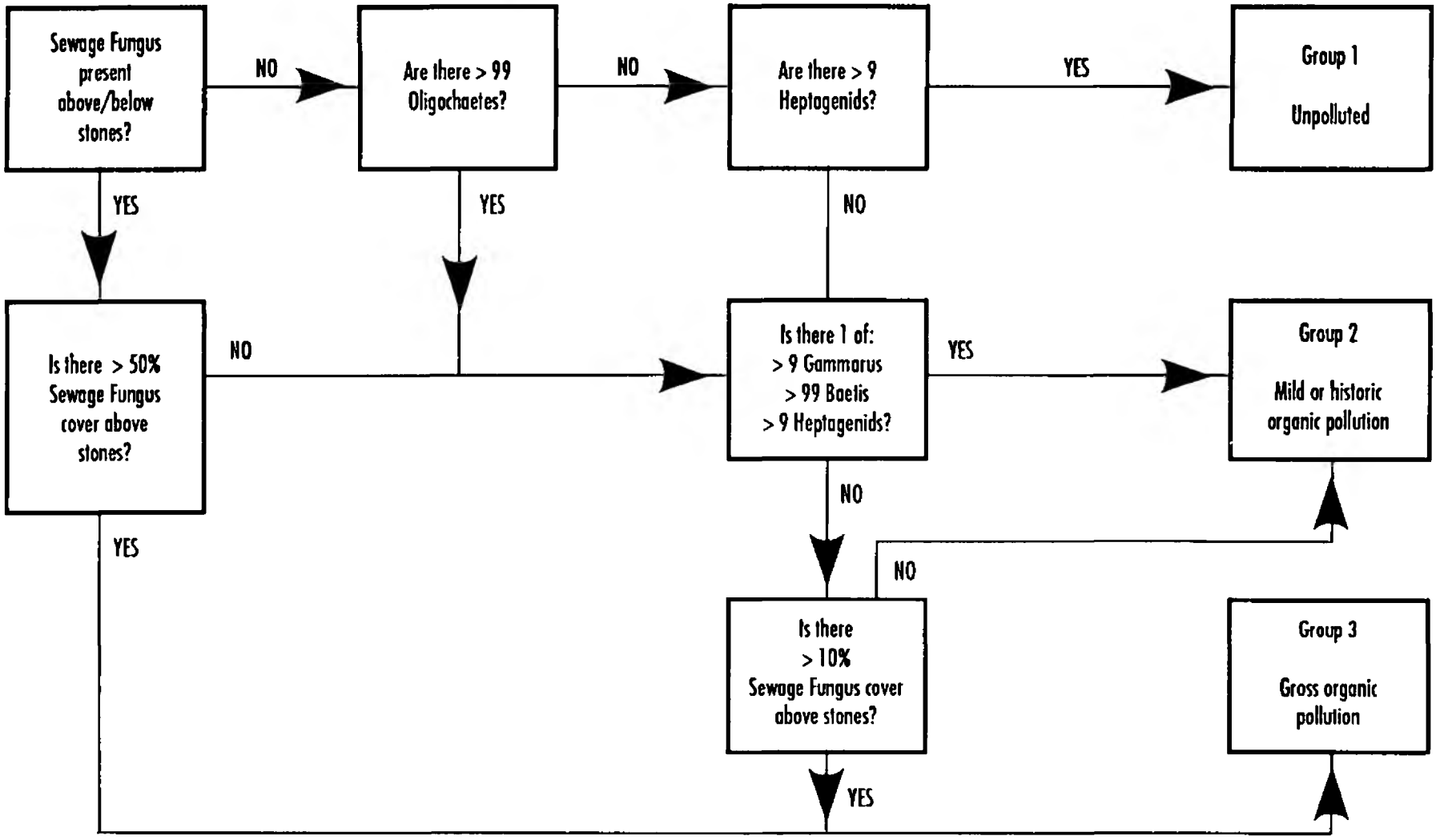


Figure 11.1 Adaption of the TWINSPAN key for West Wales produced as a flow chart. Starts in the top left-hand corner

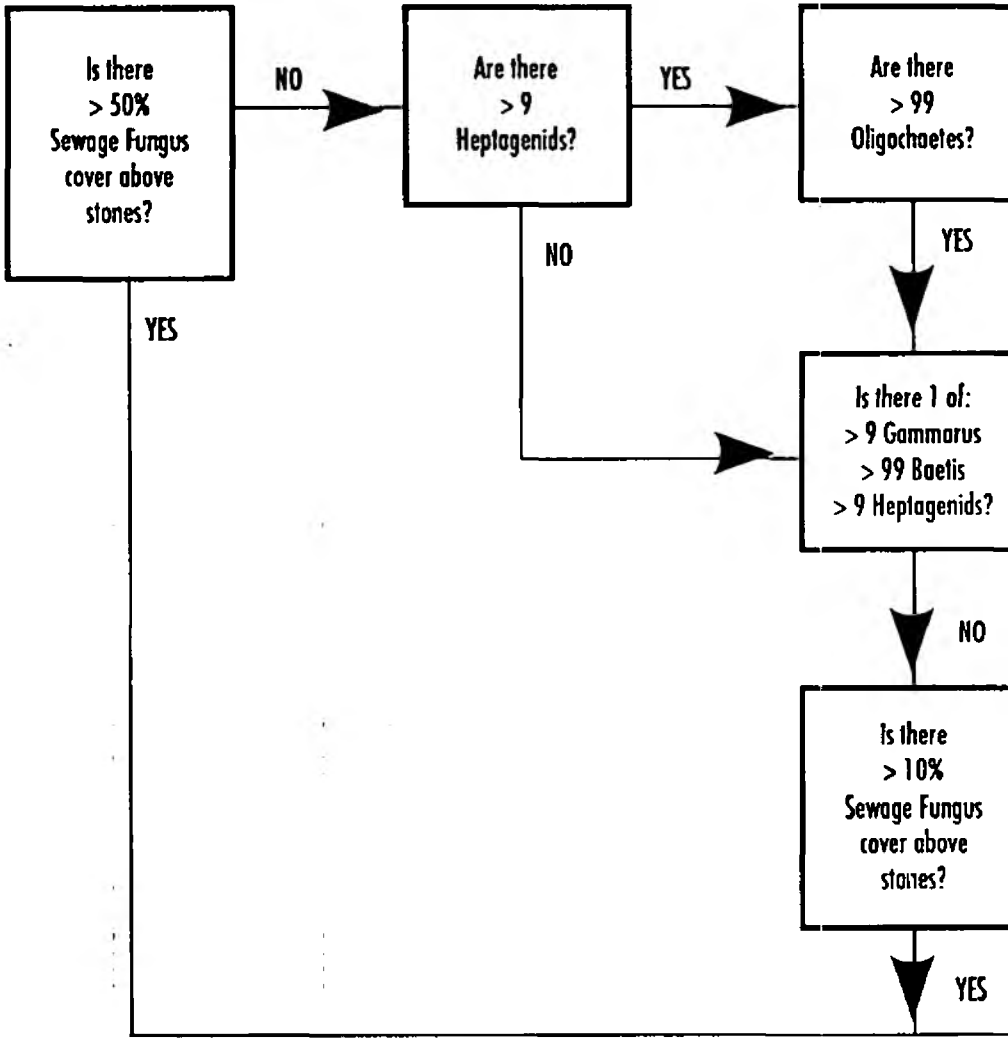
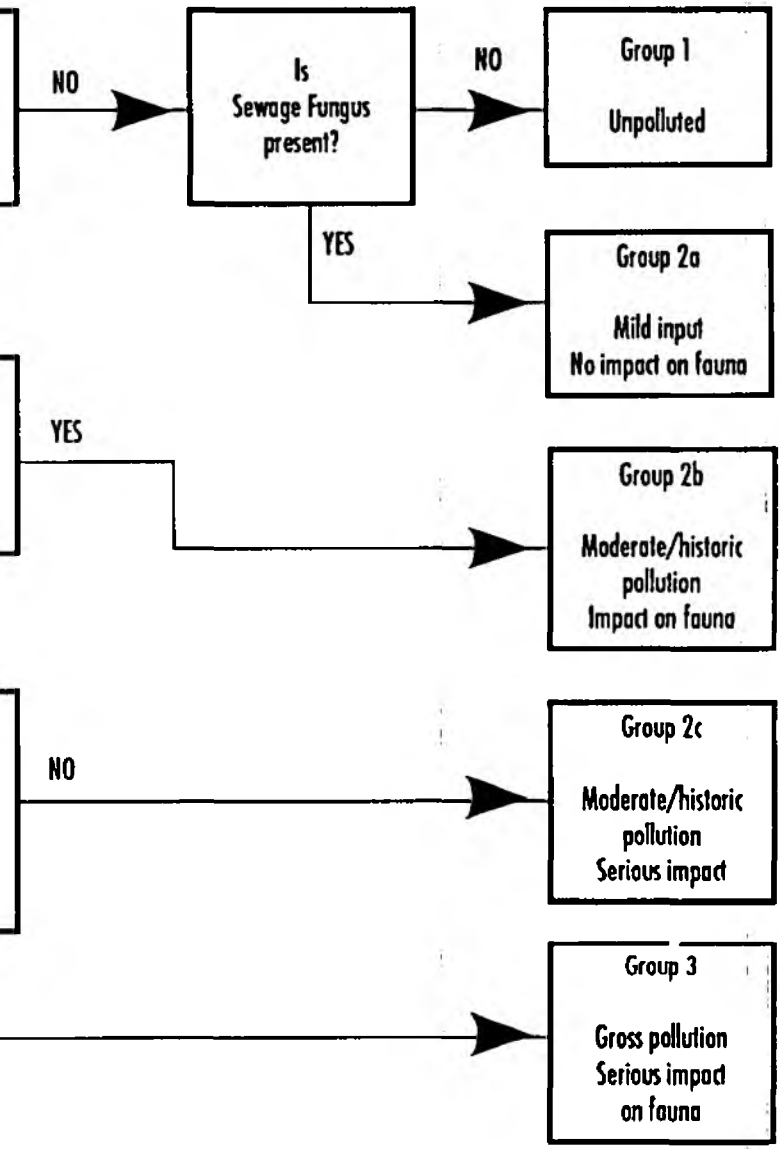


Figure 11.2 Final version of indicator system for winter/spring in West Wales


NO
Is
Sewage Fungus
present?
NO
Group 1
Unpolluted
YES
Group 2a
Mild input
No impact on fauna
YES
Group 2b
Moderate/historic
pollution
Impact on fauna
NO
Group 2c
Moderate/historic
pollution
Serious impact
Group 3
Gross pollution
Serious impact
on fauna

ACKNOWLEDGEMENTS

Staff of the Southwest Area Environmental Appraisal Unit of NRA Welsh Region assisted with the fieldwork carried out in the development of the original indicator system for West Wales. Area Pollution Control staff took part in field trials and made valuable suggestions for improvements.

REFERENCES

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

Curtis, E.J.C. (1969) Sewage Fungus: Its nature and effects, *Water Research*, *3*, 289-311.

Furse, M.T., Wright, J.F., Armitage, P.D. and Moss, D. (1981) An appraisal of pond-net samplers for biological monitoring of lotic macroinvertebrates, *Water Research*, **15**, 679-689.

Hill, M.O. (1979) TWINSPAN - a FORTRAN programme 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.

Howells, W.R. and Merriman, R. (1986) Pollution from agriculture in the area of the Welsh Water Authority. *Effects of land use on freshwaters: agriculture, forestry, mineral exploitation, urbanisation*, edited by J.F. de Le G. Solbe, pp 267-282. Water and Waste Water Technology Series. Ellis Horwood Ltd. Chichester.

NCC (1991) Nature Conservation and pollution from farm wastes. Publicity Services Branch, Nature Conservancy Council, Northminster House, Peterborough.

NRA (1992) NRA BAETIS: Users Guide. Revision 1.01. BAETIS Development Team. NRA Welsh Region, St Mellons, Cardiff.

Ormerod, S.J and Edwards, R.W. (1987) The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. *Freshwater Biology*, **17**, 533-546.

Reynolds, N. (1989) The effects of agricultural pollution on the biological quality of headwater streams in West Wales - rapid assessment of agricultural pollution by the use of an indicator key. Unpublished Internal Report Nº PL/EAW/89/3. NRA Welsh Region, St Mellons, Cardiff.

Rutt, G.P., Weatherley, N.S. and Ormerod, S.J. (1990) Relationships between the physicochemistry and macroinvertebrates of British upland streams: the development of modelling and indicator systems for predicting fauna and detecting acidity, *Freshwater Biology*, **24**, 463-480.

Seager, J., Jones, F.H. and Rutt, G.P. (1992) Assessment and Control of farm pollution, *Journal of the Institute of Water and Environmental Management*, **6 (Vol 1)**, 48-54.

Sokal, R.R. and Rohlf, F.J. (1981) Biometry. Freeman, New York.

Strahler, A.N. (1957) Quantitative analysis of water-shed geomorphology, *Transactions of the American Geophysical Society*, **38**, 913-920.

Strange, C.D., Aprahamian, M.W., and Winstone, A.J. (1989) Assessment of a semi-quantitative electric fishing sampling technique for juvenile Atlantic salmon, *Salmo salar* L., and trout *Salmo trutta* L., in small streams *Aquaculture and Fisheries Management*, **20**, 485-492.

Wade K.R., Ormerod S.J. and Weatherley N.S. (1989) Classification and ordination of macroinvertebrate assemblages to predict stream acidity in upland Wales, *Hydrobiologia*, **171**, 59-78.

Wright, J.F., Moss, D., Armitage, P.D. and Furse, M.T. (1984) A preliminary classification of running water sites in Great Britain based on macroinvertebrate species and the prediction of community type using environmental data, *Freshwater Biology*, **14**, 221-256.

APPENDIX A - USING RAPID APPRAISAL IN SURVEYS OF FARM POLLUTION

A1. SITE SELECTION

Sites should be carefully selected within a catchment so as to sample all significant tributaries and pin-point sources of pollution. Start at the bottom of the catchment and work up towards the headwaters. Sample the main stream every 2-3 km and each tributary. If pollution is discovered, continue sampling the affected watercourse, working upstream until the source is located.

A2. PROCEDURE ON SITE

1. On arrival at each site, record site number, date, time, season, method (i.e. one minute kick-sampling), Sampler (initials), watercourse, location and Grid Reference on site record sheet (Appendix B).

2. Select a suitable RIFFLE for sampling.

3. Record approximate width and depth of riffle sampled on the site sheet.

4. Record substratum composition of the riffle on site sheet as:

   - Bedrock %
   - Boulders/cobbles (>6 cm) %
   - Pebbles/gravel (2 mm - 6 cm) %
   - Sand %
   - Silt/cay %

5. Examine five large, submerged stones for sewage fungus growth. Record approximate percentage cover above and below stones on site sheet.

6. If no sewage fungus was evident at step 5, examine the site carefully for presence of sewage fungus which may also grow on vegetation, tree roots etc.

7. Take a one-minute kick sample from the riffle and tip contents into a white sorting tray.

8. Sort through the tray with reference to the indicator flow chart applicable to the area being surveyed (e.g. Figure 11.2) and derive pollution group.

A3. FEEDBACK OF INFORMATION

Pollution control should be informed of the results of each catchment survey by an agreed procedure. Catchment maps (e.g. Figure A3.1), together with lists of polluting farms in order of severity, have proved a suitable format.

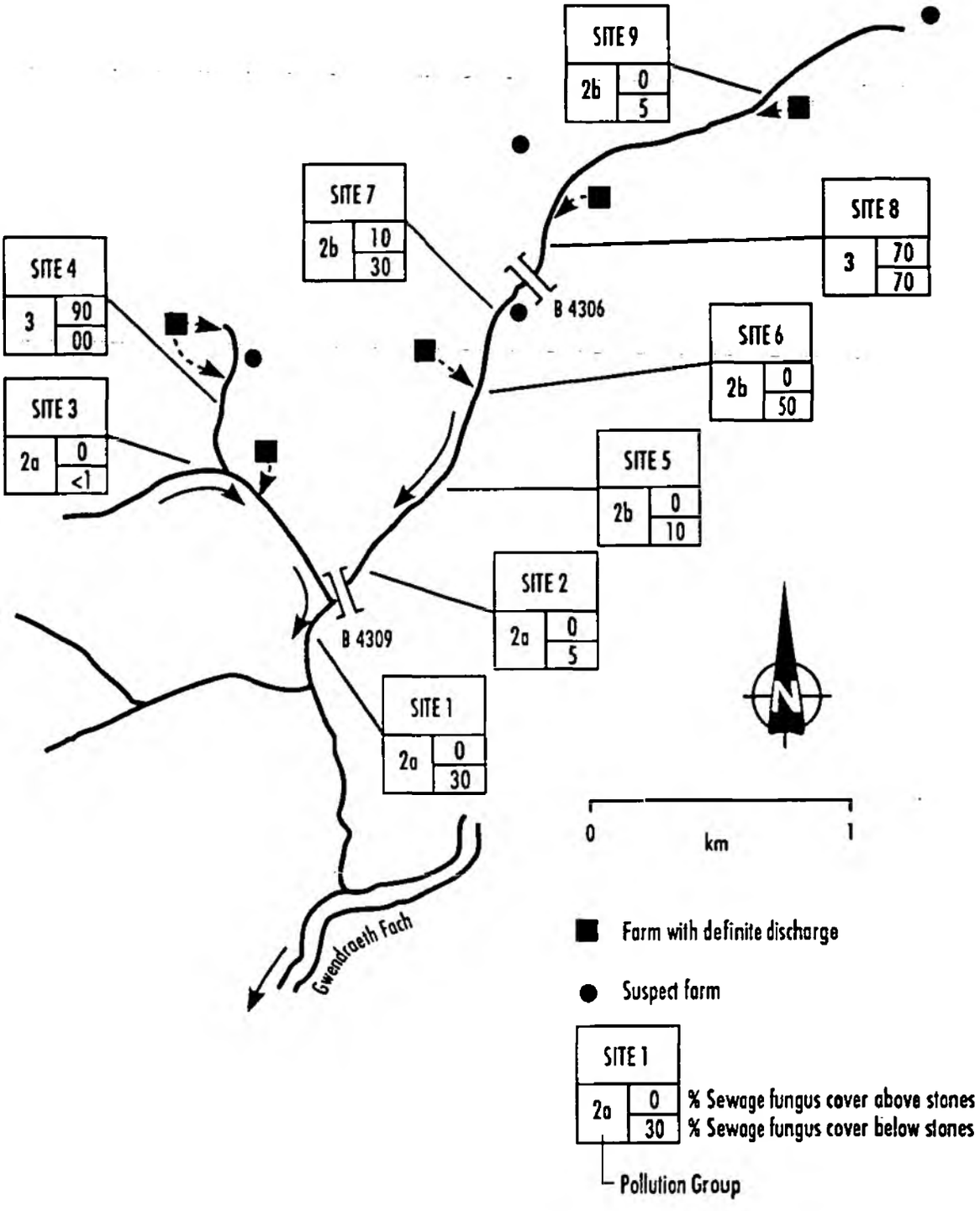


Figure A3.1 Example of a catchment map produced to communicate findings to pollution control staff. Data from a survey of the Nant Rhydw on 14 and 20 February, 1991

APPENDIX B - EXAMPLE OF A FIELD RECORD SHEET

SITE RECORD SHEET FOR RAPID APPRAISAL

B1. SITE DETAILS

Site Nº	Date	Time
Season	Method	Sampler
	Watercourse	Location
	N.G.R.	

B2. DETAILS OF RIFFLE SAMPLED

Width m Depth in sampling area cm

Substratum Composition (in the riffle sampled)

Bedrock	%
Boulders/cobbles (>6 cm)	%
Pebbles/gravel (2 mm-6 cm)	%
Sand	%
Silt/clay	%

B3. SEWAGE FUNGUS COVER

Above stones % Below large stones %

B4. CIRCLE POLLUTION GROUP

1 2a 2b 2c 3

B5. COMMENTS

(e.g. pollution source/foaming etc.)