# FAME WP6/7 UK/ECOREGION 18 SPATIAL APPROACH

# RIVER AND FISH COMMUNITY TYPOLOGY

# METRIC SELECTION AND ANALYSIS

**November 2003**

**R.A.A. Noble[1], I. G. Cowx[1] & A. Starkie[2]**

[1] Hull International Fisheries Institute

[2]Environment Agency (National Fisheries Technical Team, WFD Technical Specialist)

CONTENTS

## LIST OF FIGURES

# LIST OF TABLES

# 1. UK FIDES CONTENTS

## 1.1. Numbers and geographic distribution of sites

The fish community data used to populate UK-FIDES were sourced from the Environment Agency's national Fisheries Monitoring Programme. The majority of the sampling occasions were taken from the 2002 and 2001 sampling programme. However, the dataset from the South Anglian area (Essex) comprises many sampling occasions for each site, dating back to 1985. UK FIDES contains data for 1070 sites, comprising a total of 2141 fishing occasions. Unfortunately no data were available for South Wales or the southwest of England (Devon & Cornwall). Due to the Environment Agency only being responsible for England and Wales no data were available for Scotland or Northern Ireland (Ecoregion 17).

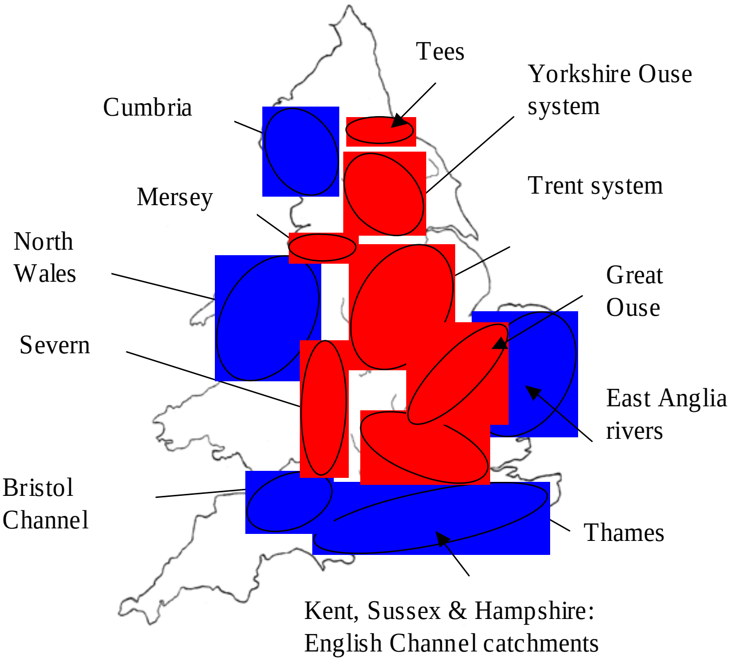

**Figure 1.1** Major river basins and regions represented in UK FIDES.

## 1.2. Environmental impacts

Of the 1070 sites in UK FIDES 1017 had complete sets of impact scores for the major anthropogenic impact criteria (Figure 1.2 and Table 1.1).

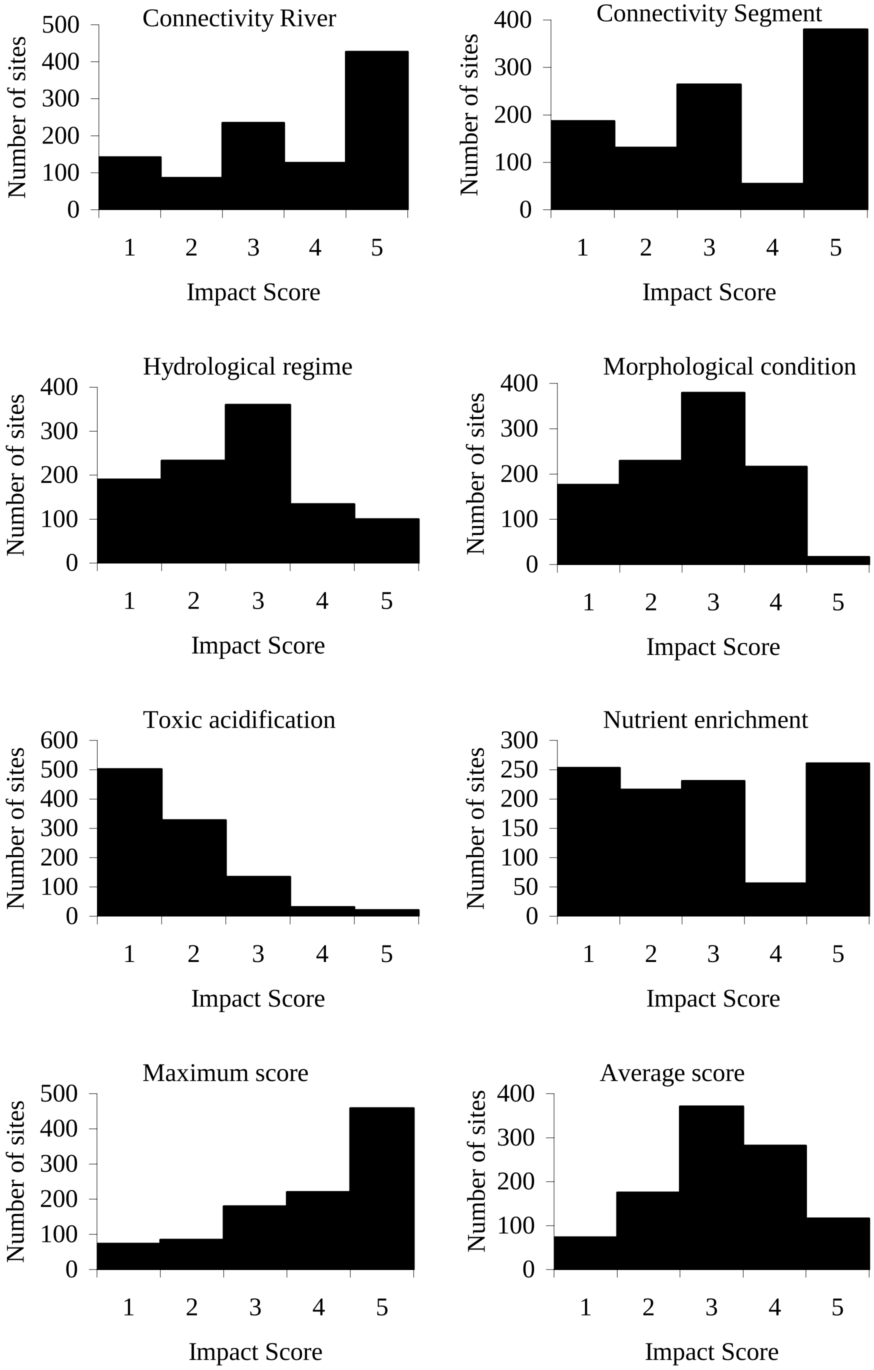

**Figure 1.2** Frequency of impact scores for the six key anthropogenic impact criteria of sites in UK FIDES.

**Table 1.1** Number of sites versus maximum impact score for the major anthropogenic impacts, both including and excluding connectivity_river.

| Max Impact Score | Inc. Connectivity River | Ex. Connectivity River |
|---|---|---|
| 1 | 73 | 79 |
| 2 | 85 | 123 |
| 3 | 180 | 222 |
| 4 | 220 | 176 |
| 5 | 459 | 417 |

Analysis of environmental impacts identified 201 fishing occasions that passed the original impairment criteria for determining the potential calibration dataset (Section 1.4).

## 1.3. Catch characteristics

Catch data was characterised generally having low species diversity with only between 4 and 9 species recorded in the majority of samples. Very few samples recorded more than 12 species (Figure 1.3). Catches generally amounted to less than 150 individual in the 1st run, with about 50% of samples comprising less than 100 fish and 30% of samples comprising <50 fish (Figure 1.4).

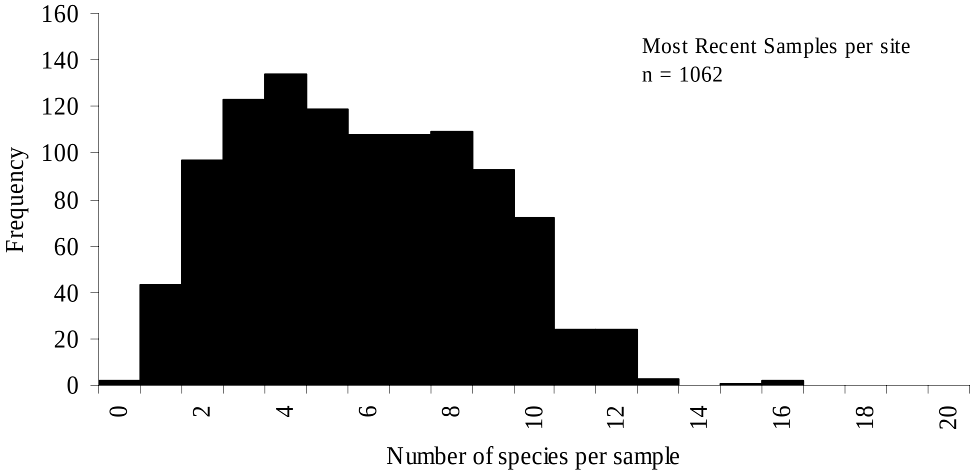

**Figure 1.3** Frequency distribution of the total number of species caught per fishing occasion.

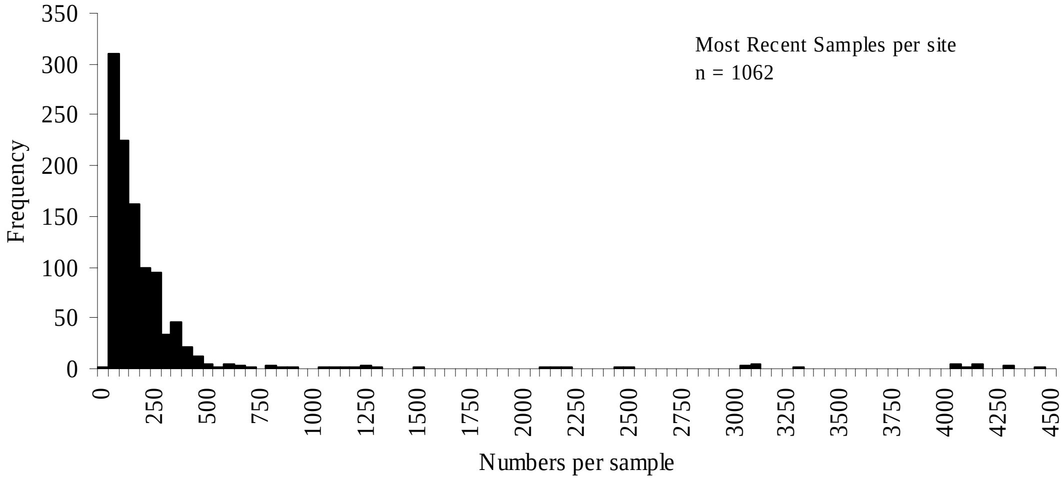

**Figure 1.4** Frequency distribution of the total number of fish caught per fishing occasion.

*Anguilla anguilla* and *Rutilus rutilus* were the most common species in the most recent samples per site in the UK dataset, each occurring in 52% of fishing occasions. The frequent occurrence of roach, however, reflects the large number of lowland sites reported by the Southeast Anglia area of the Environment Agency. Where present, *Barbatula barbatula, Cottus gobio, Phoxinus phoxinus, Salmo salar* and *Gasterosteus aculeatus* exhibited the highest mean abundances of around 1000 to 2000 individuals per hectare. Maximum abundances per site of over 10000 per hectare were found for eight species including *A, anguilla, R. Rutilus, S. salar* and *Salmo trutta fario*. Extremely high abundances (>35000 n.ha$^{-1}$) of *B. barbatula, C. gobio, P. phoxinus,* and *G. aculeatus* were recorded on some fishing occasions (Table 1.2). Only 24 taxa (22 native plus two alien) plus two varieties of hybrid were observed to occur in more than 1% of sites, whilst 13 taxa (including four alien and three "transitional zone" species) were only recorded in <1% of sites.

## 1.4. Calibration dataset

Initial analysis of the UK dataset identified that even when the exclusion criteria for calibration data were set at maximum impact scores of 1 or 2 (excluding Connectivity) the calibration dataset failed to represent the range of fish communities which occur in England and Wales. In particular, few sites representative of "coarse" fish communities were retained in the dataset. To overcome this an initial two-level selection process was established:

1) All sites with maximum impact scores of 1 or 2 (excluding Connectivity);
2) Additional sites with maximum impact scores of 3 – excluding sites that were deemed to be impacted "salmonid" sites (exclusion based on species composition and gradient).

A total of 417 sites were retained in the initial calibration dataset following this selection process. The initial calibration dataset was used in the initial stages of the typology analysis, however, after an initial cluster analysis (Section 2.2.1) the calibration dataset was restricted further by removing the most impacted sites from each of the major clusters in the dendrogram. This was done to minimise the chance of impacted sites affecting the calibration of each river type, especially for types where sufficient low impairment data were available. This restricted the calibration dataset to 312 samples.

*Salmo trutta fario* was the most common species in the trimmed calibration dataset, reflecting the generally higher impact status of many of the lowland sites in the UK FIDES dataset. *Salmo trutta fario, C. gobio, S. salar, B. barbatula* and *P. phoxinus* had the highest mean CPUE values of between 1000 and 2500 individuals per hectare. Where present the mean abundance of *Platichthys flesus* was also high (around 5000 per ha). *Leuciscus leuciscus, R. rutilus, Leuciscus cephalus* and *Gobio gobio* were the most common lowland species present in the trimmed calibration dataset (Table 1.3). Only three alien species, *Cyprinus carpio, Oncorhynchus mykiss* and *Lepomis gibbosus,* occurred in calibration samples. Two notable native species were missing from the calibration dataset, *Cobitis taenia* and *Carassius carassius,* but this is probably due both to their restricted distribution in the UK and in the case of *C. taenia* sampling difficulties.

**Table 1.2** Summary of catch statistics for the most recent samples from each site in the UK dataset (n = 1062). Mean abundance data are calculated as the average abundance where the species is present.

| | % Occurrence | Average | Max |
|---|---|---|---|
| *Anguilla anguilla* | 52.8 | 303 | 20409 |
| *Rutilus rutilus* | 52.5 | 383 | 10000 |
| *Salmo trutta fario* | 49.3 | 905 | 18628 |
| *Leuciscus cephalus* | 47.1 | 171 | 5393 |
| *Leuciscus leuciscus* | 44.4 | 163 | 2584 |
| *Gobio gobio* | 41.6 | 216 | 9000 |
| *Cottus gobio* | 41.1 | 1577 | 37454 |
| *Esox lucius* | 41.1 | 35 | 373 |
| *Barbatula barbatula* | 39.5 | 1207 | 74075 |
| *Perca fluviatilis* | 39.4 | 75 | 1850 |
| *Phoxinus phoxinus* | 34.7 | 1487 | 37454 |
| *Salmo salar* | 18.4 | 2166 | 16214 |
| *Gasterosteus aculeatus* | 15.1 | 1357 | 74075 |
| *Abramis brama* | 10.7 | 31 | 362 |
| *Thymallus thymallus* | 7.3 | 130 | 867 |
| *Scardinius erythrophthalmus* | 6.4 | 29 | 140 |
| *Tinca tinca* | 6.3 | 14 | 61 |
| *Barbus barbus* | 6.1 | 35 | 449 |
| *Alburnus alburnus* | 6.0 | 151 | 864 |
| *Cyprinus carpio* | 5.4 | 29 | 215 |
| *Lamprey all* | 5.2 | 414 | 4167 |
| *Gymnocephalus cernuus* | 3.3 | 24 | 175 |
| *Platichthys flesus* | 2.2 | 531 | 5840 |
| *Rutilus rutilus x Abramis brama* | 2.1 | 12 | 54 |
| *Oncorhynchus mykiss* | 1.7 | 67 | 364 |
| *Rutilus rutilus x Scardinius erythrophthalmus* | 1.2 | 14 | 69 |
| *Carassius auratus* | 0.6 | 9 | 21 |
| *Cobitis taenia* | 0.6 | 255 | 925 |
| *Carassius carassius* | 0.5 | 11 | 32 |
| *Chelon labrosus* | 0.5 | 23 | 35 |
| *Blicca bjoerkna* | 0.4 | 11 | 30 |
| *Gobiidae* | 0.4 | 100 | 240 |
| *leuciscus idus* | 0.2 | 16 | 21 |
| *Pungitius pungitius* | 0.2 | 10 | 16 |
| *Dicentrarchus labrax* | 0.1 | 8 | 8 |
| *Lepomis gibbosus* | 0.1 | 16 | 16 |
| *Pleuronectes platessa* | 0.1 | 34 | 34 |
| *Salvelinus fontinalis* | 0.1 | 20 | 20 |
| *Sander lucioperca* | 0.1 | 18 | 18 |

**Table 1.3** Summary of catch statistics in the trimmed calibration dataset. Mean abundance data are calculated as the average abundance where the species is present.

| | % Occurrence | Average | Max |
|---|---|---|---|
| *Salmo trutta fario* | 70.5 | 1287 | 18628 |
| *Anguilla anguilla* | 56.1 | 550 | 20409 |
| *Cottus gobio* | 51.6 | 1688 | 20409 |
| *Barbatula barbatula* | 43.3 | 1335 | 20409 |
| *Salmo salar* | 39.7 | 2446 | 16214 |
| *Leuciscus leuciscus* | 36.9 | 142 | 1700 |
| *Phoxinus phoxinus* | 35.6 | 1859 | 19231 |
| *Rutilus rutilus* | 35.6 | 374 | 4518 |
| *Leuciscus cephalus* | 34.0 | 161 | 1036 |
| *Gobio gobio* | 31.1 | 187 | 1450 |
| *Esox lucius* | 27.6 | 35 | 373 |
| *Perca fluviatilis* | 23.1 | 101 | 1850 |
| *Thymallus thymallus* | 13.5 | 142 | 728 |
| *Gasterosteus aculeatus* | 12.5 | 620 | 6098 |
| *Barbus barbus* | 6.1 | 30 | 92 |
| *Lamprey all* | 5.1 | 393 | 2611 |
| *Abramis brama* | 4.8 | 38 | 200 |
| *Scardinius erythrophthalmus* | 3.8 | 30 | 96 |
| *Alburnus alburnus* | 3.2 | 161 | 864 |
| *Gymnocephalus cernuus* | 2.6 | 60 | 175 |
| *Cyprinus carpio* | 1.9 | 45 | 154 |
| *Oncorhynchus mykiss* | 1.6 | 45 | 79 |
| *Cobitis taenia* | 1.3 | 375 | 925 |
| *Gobiidae* | 1.0 | 126 | 240 |
| *Platichthys flesus* | 0.6 | 5054 | 5840 |
| *Tinca tinca* | 0.6 | 6 | 7 |
| *Blicca bjoerkna* | 0.3 | 30 | 30 |
| *Lepomis gibbosus* | 0.3 | 16 | 16 |
| *Pleuronectes platessa* | 0.3 | 34 | 34 |

The calibration dataset excluded proportionately more lowland sites and was reflected in the shift in distribution of the numbers of species per sample (Figure 1.5). The majority of samples in the calibration dataset comprised 2 to 5 species and few samples contained more than 10 species.

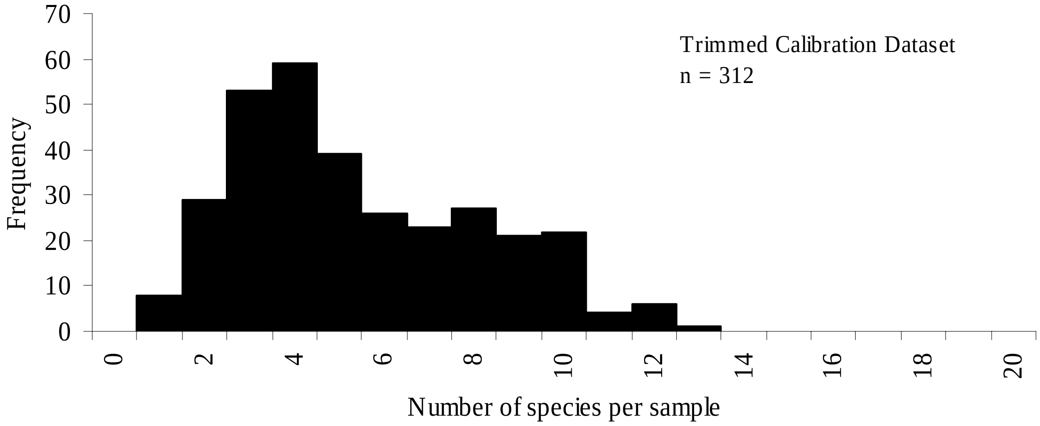

**Figure 1.5** Frequency distribution of the total number of species caught per fishing occasion in the calibration dataset.

# 2. FISH-BASED TYPOLOGY

## 2.1. Freshwater fish distribution in the UK

The formation of freshwater fish communities in the UK and the distributions of native species dates back to the end of the last Ice Age. It is generally accepted that, except for euryhaline species (*Salmo salar, Salmo trutta, Anguilla anguilla* etc.) that entered the freshwaters formed by the retreat of the ice cap by way of the sea, all the primary freshwater fish (stenohaline species) are indigenous only to the river catchments bordering the eastern English Channel and southern North Sea (Wheeler 1977). The stenohaline fish species are believed to have entered England through connections between the major eastern English rivers and the Northern European rivers when they are thought to have converged on a land bridge between UK continental Europe (Regan 1911). It is believed that the majority of stenohaline species were originally restricted to the eastern flowing rivers between the Yorkshire Ouse basin and the Thames basin. However, the native fish fauna of the Thames is thought to be distinct to that of the other catchments in this area (notably the probable original absence of *Cobitis taenia, Blicca bjoerkna* and *Lota lota* from the Thames). Regan (1911) suggested that the Thames may be more similar to the rivers of southern England, which he suggested may have had connection with the rivers of northern France.

Wheeler (1977) suggested that to a certain extent the current distribution of stenohaline species still reflects their origin since the fish fauna of eastern England is much richer than that of the western or Scottish rivers. Further, at the time of Regan's analysis, *Lota lota, Gymnocephalus cernua, Cobitis taenia, Blicca bjoerkna and Barbus barbus* were still restricted to the eastern rivers, whilst *Scardinius erythrophthalmus* and *Carassius carassius* exhibited a similar distribution. Wheeler (1977) suggested that the wider distribution of many species and the patchy distribution of others beyond the eastern rivers (between the Humber and the Thames) was probably due redistribution by humans. Wheeler & Jordan (1990) showed that the distribution of the Barbel (*Barbus barbus*) beyond it's native rivers (Yorkshire Ouse to Thames) was both entirely a result of anthropogenic activity and very recent (after 1890). Barbel was only introduced into the River Severn in 1956 (Hickley & North 1981).

### 2.1.1. *SERCON River regions*

SERCON is a tool developed by Boon *et al.* (1994, 1996, 1997, 2002) to evaluate the conservation importance and status of habitats. The tool comprises some elements that assess and score the "naturalness" of fish communities. Although the most up to date revisions (Maitland 2003) have developed expected species occurrence and native/alien status species lists for the 97 main hydrometric areas in Great Britain (e.g. Severn Basin, Thames Basin etc.), the original version identified 5 regions of the UK (Figure 2.1) in which the fish community could be predicted (Table 2.1).

**Table 2.1** Diagnostic native species for the five geographical regions in the original version of SERCON (Boon *et al.* 1996) (Reproduced from Maitland 2003).

| Zone | Source waters | Head water streams | Middle (lotic) reaches | Middle (lentic) reaches |
|---|---|---|---|---|
| A | No fish | *Salmo trutta fario* | *Salmo salar*<br>*Salmo trutta fario*<br>*Gasterosteus aculeatus* | *Salmo salar*<br>*Salmo trutta fario*<br>*Gasterosteus aculeatus*<br>*Platichthys flesus* |
| B | No fish | *Salmo trutta fario* | *Salmo salar*<br>*Phoxinus phoxinus*<br>*Barbatula barbatula* | *Gasterosteus aculeatus*<br>*Lampetra fluviatilis*<br>*Esox lucius*<br>*Petromyzon marinus* |
| C | No fish | *Salmo trutta fario*<br>*Cottus gobio* | *Salmo salar*<br>*Phoxinus phoxinus*<br>*Barbatula barbatula*<br>*Thymallus thymallus* | *Gasterosteus aculeatus*<br>Lamprey Spp.<br>*Esox lucius*<br>*Rutilus rutilus*<br>*Perca fluviatilis* |
| D | No fish | *Salmo trutta fario*<br>*Cottus gobio* | *Gobio gobio*<br>*Alburnus alburnus*<br>*Barbus barbus*<br>*Leuciscus leuciscus*<br>*Thymallus thymallus* | *Abramis brama*<br>*Leuciscus cephalus*<br>*Rutilus rutilus*<br>*Scardinius erythrophthalmus*<br>*Tinca tinca*<br>*Gymnocephalus cernuus*<br>*Perca fluviatilis* |
| E | No fish | *Salmo trutta fario*<br>*Cottus gobio* | *Gobio gobio*<br>*Alburnus alburnus*<br>*Barbus barbus*<br>*Leuciscus leuciscus*<br>*Thymallus thymallus* | *Blicca bjoerkna*<br>*Abramis brama*<br>*Leuciscus cephalus*<br>*Rutilus rutilus*<br>*Scardinius erythrophthalmus*<br>*Tinca tinca*<br>*Gymnocephalus cernuus*<br>*Perca fluviatilis* |

Note: This scheme indicates for each of the main freshwater habitats in river systems in each geographic area of the UK, the species that characterise that habitat. It considers only those species that are native to each geographic area concerned and most typical for the reaches indicated. Thus, *S. trutta* may occur in many waters in the south of the UK but, unlike northern Britain, are not the most diagnostic species there and thus not listed. Ubiquitous species such as *Anguilla anguilla* are not included in the scheme (Maitland 2003).

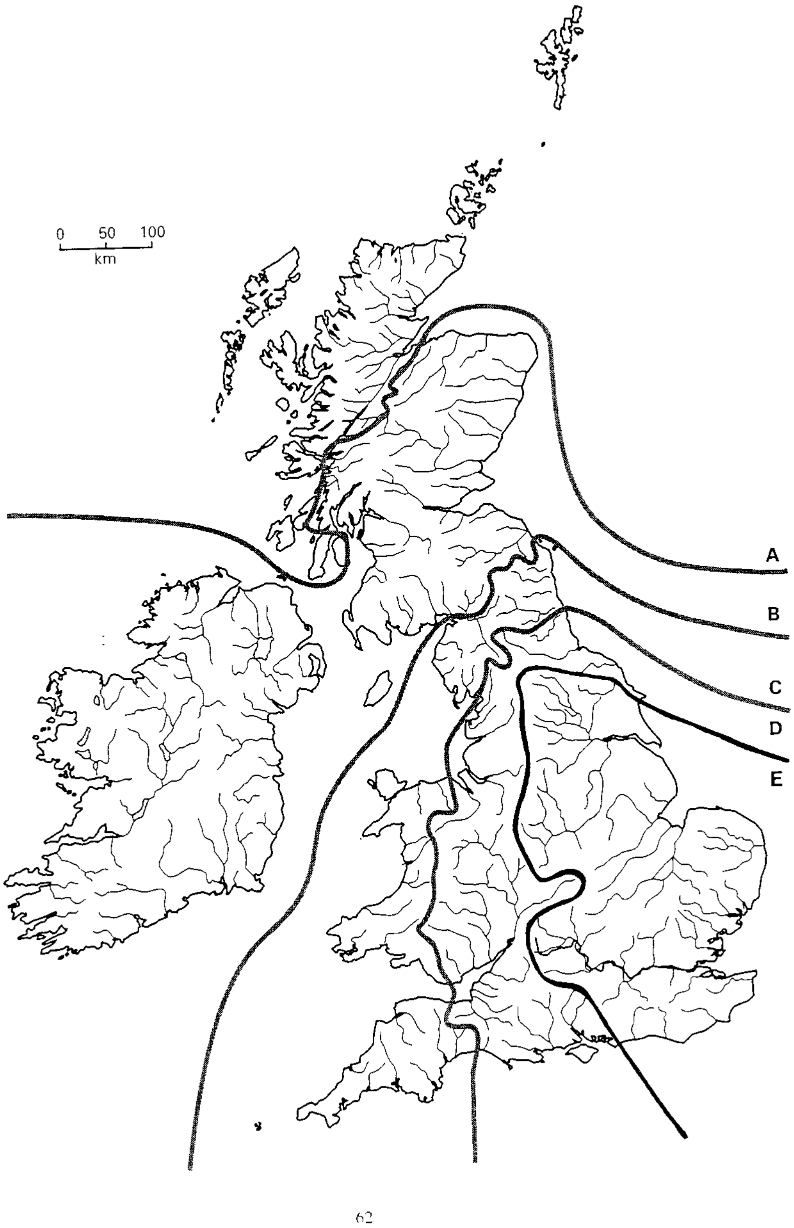

**Figure 2.1** The five fish-regions of the UK as described in the original version of SERCON.

UK FIDES only contains data for 3 of the regions.

- Region E - rivers of the southeast and east coast up to the Humber rivers.
- Region D – relatively diverse group of rivers/regions including south coast rivers (e.g. Hampshire rivers), River Severn, Mersey catchment and the River Tees (Northeast coast river).
- Region C – rivers of extreme southwest or England, Wales, Cumbria and northeast England.

This basic classification system was identified as requiring improvement to allow better determination of conservation status of freshwaters. This was particularly important for many species, which although present in the region were not present in all water bodies. Particularly important in regions C, D and E are the distribution of *B. barbus* (particularly in Region D where it is alien in some rivers) and *T. thymallus*. Additionally the distributions of *C. gobio* and *B. barbatula* are not uniform in all rivers in region C. To address this lists of species (of both native and alien origins) have been described per main hydrometric area (Maitland 2003).

## 2.2. Community-type analysis

### 2.2.1. *Cluster analysis*

Fish abundance data from calibration fishing occasions were entered into a cluster analysis (Ward's method, phi-squared measure) to determine various fish community types represented in UK FIDES (Figure 2.2). Cluster analyses were run using species presence/absence, species abundance (n ha$^{-1}$ 1st run CPUE, both actual and Log10 transformed), percentage abundance and standardised abundance (range of 0 to 1 for each species) to determine which gave the most ecologically appropriate groupings. Although each approach created a major division equivalent to a salmonid / "coarse" fish split, differences were observed between the community measures used in the analysis. Presence absence data and percent abundance data were deemed inappropriate as each of the major clusters formed using showed significant subgroups when actual abundance values were assessed. When untransformed abundance data were used the clustering was driven primarily by species such as *Cottus gobio*, *Barbatula barbatula*, *Phoxinus phoxinus* and *Anguilla anguilla*, which were both relatively common but also had a greater range of abundance than many of the key species in both lowland and upland rivers. Analyses indicated that standardised abundance data were more appropriate for the cluster analysis giving a suitable balance between the presence/absence of a species and their abundance.

Nine major groups were identified within the resulting dendrogram (Figure 2.2) and were explored and statistically tested using ANOVA (Table 2.4). The nine groups were assigned into eight fish community types (Figure 2.2). Three adjustments were made to the dendrogram based on ANOVA analysis of both biotic and abiotic variables for the nine major clusters, together with expert judgement (three adjustments correspond to clusters marked by A and B (comprising two changes) in Figure 2.2).

A. This cluster was determined by very high abundances of minnows, bullhead, stone loach and eels. However, analysis showed that both salmonid and cyprinid fish types occurred within this cluster. Therefore, the type was further divided by an additional cluster analysis on this group of sites to extract the predominantly salmonid types and the types characterised by coarse fish. These site were then grouped with the most similar major type in the dendrogram.

B. Analysis of this cluster indicated that it comprised both a lowland coarse fish and also sites that represented short coastal streams characterised by trout, eels, some coarse fish and coastal species such as flounder. Therefore, this cluster was divided by a further cluster analysis into two community types. The lowland fish cluster derived from this group could not be easily distinguished from one of the other major coarse fish cluster so they were merged together into a single fish community type (Figure 2.2)

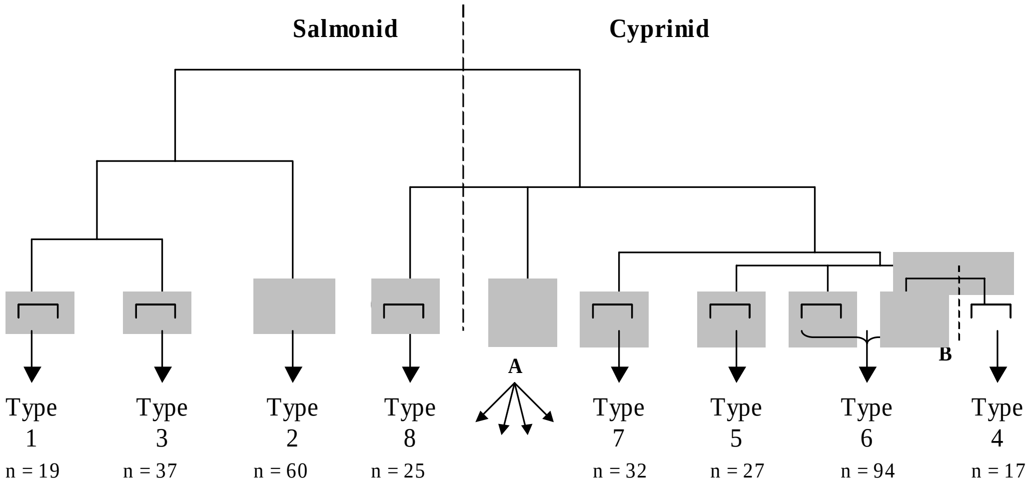

**Figure 2.2** Cluster dendrogram (Ward's method, Phi-squared measure) for fish species abundance (Standardised 1st run CPUE) in UK rivers. Nine major types are identified.

Eight fish types were described on the basis of the cluster analysis each of which were characterised by both species composition and abundance (Table 2.2 and 2.3). Four salmonid types (including one regional type), three predominantly cyprinid types and one semi-transitional type were identified (sites coded by longitudinal zonation i.e. Type 1 highest gradient):

Type 1) *S. trutta fario* zone, *S. salar* may be present but zone may be above natural salmon migration limits.

Type 2) Very high abundance of *S. salar*, high abundance of *S. trutta fario*. Lamprey, *A. anguilla*, *C. gobio, B. barbatula* and *P. phoxinus* may also be present.

Type 3) Very high abundance of *S. trutta fario, S. salar* generally present but not in high abundance, *C. gobio* very abundant, *B. barbatula* also present.

Type 4) Mixed species community type characterised by coastal taxa *A. anguilla, P. flesus, G. aculeatus* and Gobidae. *S. trutta fario* also common.

Type 5) High abundance of *L. cephalus, B. barbus* present and also relatively high in abundance. *E. lucis* and *T. thymallus* also present. Represents upper barbel zone of very large river systems, in particular sampling sites on the main river stem (N.B. Barbel are not present in every river system in the UK).

Type 6) Lowland coarse fish type characterised by *L. cephalus, L. leuciscus, G. gobio, R. rutilus, A. brama* and *E. lucius*.

Type 7) Generally low overall abundance, *R. Rutilus* dominate, *E. lucius* and *A. alburnus,* are relatively abundant. Representative of large lowland rivers, in particular the main river stem, characterised by water depth (generally boat based survey data).

Type 8) Characterised by *S. salar*, *E. lucius* and *T. thymallus* and specifically representative of the chalk rivers of the South Coast of England, in particular the rivers of Hampshire.

**Table 2.2** Mean 1st run CPUE (n ha$^{-1}$) per native fish species in each of the eight major types described by cluster analysis. Distinctive species / species abundance per group are highlighted.

| | Fish type code by longitudinal zonation | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| *A. brama* | | | | 1.7 | 0.6 | 5.0 | 0.1 | 2.1 |
| *A. alburnus* | | 3.1 | | 50.8 | 7.6 | | 10.8 | |
| *A. anguilla* | 82.1 | 116.2 | 1096.1 | 1107.7 | 595.4 | 94.5 | 85.3 | 19.8 |
| *B. barbatula* | 51.6 | 277.9 | 1446.1 | 238.2 | 718.2 | 862.2 | 127.0 | 8.2 |
| *B. barbus* | | | | | 19.4 | 0.4 | | 0.4 |
| *B. bjoerkna* | | | | | | | 0.9 | |
| *C. taenia* | | | | | | 15.9 | | |
| *C. gobio* | 106.4 | 845.0 | 2922.8 | 619.9 | 615.3 | 827.1 | 78.4 | 105.9 |
| *C. carpio* | | | | | | 2.8 | | |
| *E. lucius* | | 0.6 | | | 6.9 | 8.0 | 56.5 | 9.2 |
| *G. aculeatus* | | 1.8 | 2.3 | 655.9 | 10.9 | 107.5 | 43.8 | 41.9 |
| GOBIIDAE | | | | 22.0 | | 0.1 | | |
| *G. gobio* | | 0.6 | | 1.4 | 27.3 | 165.5 | 51.8 | 6.7 |
| *G. cernuus* | | | | | | 4.9 | 0.5 | |
| Lamprey all | | 7.3 | 24.4 | 254.9 | | 6.5 | | |
| *L. gibbosus* | | | | | | | 0.5 | |
| *L. cephalus* | | 1.0 | 2.7 | 2.5 | 72.9 | 135.9 | 49.9 | 20.6 |
| *L. leuciscus* | | 0.9 | | 13.6 | 37.8 | 136.6 | 51.3 | 23.6 |
| *O. mykiss* | 2.4 | | | 3.8 | | 0.1 | 2.5 | 0.9 |
| *P. fluviatilis* | | 1.4 | 0.9 | 1.2 | 6.3 | 63.6 | 21.8 | 10.6 |
| *P. phoxinus* | | 93.9 | 588.4 | 366.1 | 850.7 | 1282.4 | 829.5 | 101.3 |
| *P. flesus* | | | | 594.5 | | | | |
| *P. platessa* | | | | 2.0 | | | | |
| *R. rutilus* | | 4.0 | 6.2 | 35.4 | 50.7 | 309.9 | 266.6 | 55.2 |
| *S. salar* | 455.9 | 4262.5 | 341.3 | 439.8 | 15.9 | 31.9 | | 443.9 |
| *S. t. fario* | 4890.7 | 1296.3 | 2138.8 | 1096.4 | 33.7 | 82.1 | 13.9 | 173.2 |
| *S. erythrophthalmus* | | 0.1 | | 2.6 | | 3.3 | 0.2 | |
| *T. thymallus* | | | 5.7 | | 9.5 | 3.1 | | 207.6 |
| *T. tinca* | | | | | | 0.1 | | |

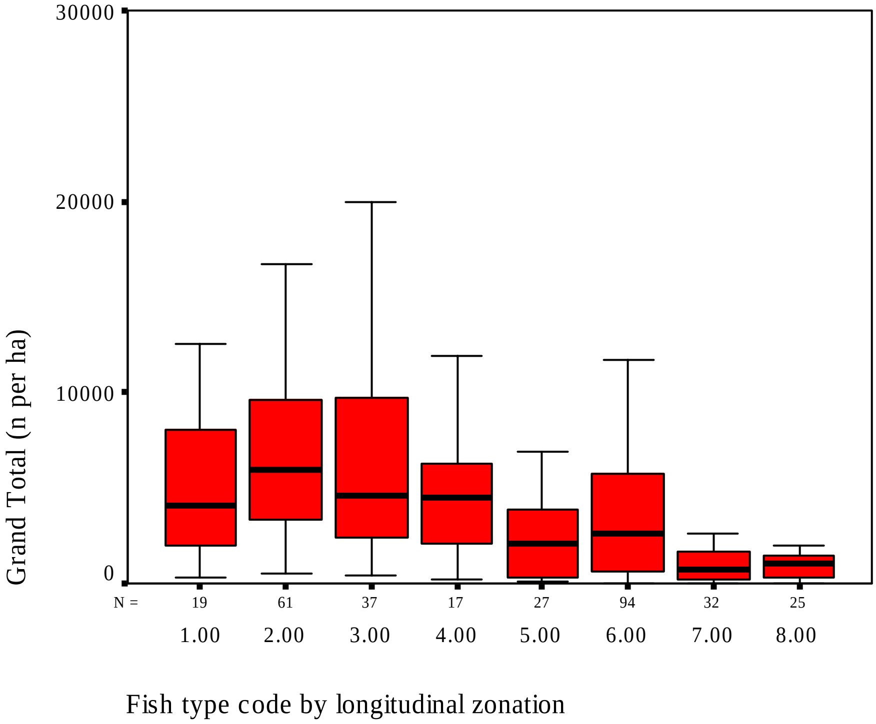

**Figure 2.3** Box plot of total 1st run CPUE for all fish per fish type.

**Table 2.3** Kruskall-Wallis tests of 1st run CPUE of native species per each of the eight groups described by cluster analysis.

| Species/Taxa | Chi-Square | df | Significance |
|---|---|---|---|
| *Abramis brama* | 16.03 | 7 | 0.025 |
| *Alburnus alburnus* | 32.60 | 7 | 0.000 |
| *Anguilla anguilla* | 34.79 | 7 | 0.000 |
| *Barbatula barbatula* | 52.40 | 7 | 0.000 |
| *Barbus barbus* | 147.53 | 7 | 0.000 |
| *Blicca bjoerkna* | 8.75 | 7 | 0.271 |
| *Cottus gobio* | 64.50 | 7 | 0.000 |
| *Cyprinus carpio* | 14.14 | 7 | 0.049 |
| *Esox lucius* | 151.93 | 7 | 0.000 |
| *Gasterosteus aculeatus* | 32.05 | 7 | 0.000 |
| GOBIIDAE | 22.87 | 7 | 0.002 |
| *Gobio gobio* | 104.76 | 7 | 0.000 |
| *Gymnocephalus cernuus* | 13.91 | 7 | 0.053 |
| Lamprey all | 39.96 | 7 | 0.000 |
| *Leuciscus cephalus* | 118.40 | 7 | 0.000 |
| *Leuciscus leuciscus* | 135.01 | 7 | 0.000 |
| *Perca fluviatilis* | 66.57 | 7 | 0.000 |
| *Phoxinus phoxinus* | 81.07 | 7 | 0.000 |
| *Platichthys flesus* | 34.82 | 7 | 0.000 |
| *Rutilus rutilus* | 112.64 | 7 | 0.000 |
| *Salmo salar* | 194.25 | 7 | 0.000 |
| *Salmo trutta fario* | 194.45 | 7 | 0.000 |
| *Scardinius erythrophthalmus* | 13.86 | 7 | 0.054 |
| *Thymallus thymallus* | 216.92 | 7 | 0.000 |
| *Tinca tinca* | 4.65 | 7 | 0.702 |

### 2.2.2. *Abiotic characteristics of fish types*

Types 1, 2 and 3, the three main salmonid types, were characterised by the highest altitudes and gradients (Figures 2.7 and 2.8), whilst groups, 5, 6 and 7 the main cyprinid groups were characterised by low altitudes and gradients. Types 6 and 7, the main cyprinid types, differed both in altitude and gradient with type 7 having predominantly lower altitudes and gradients that Type 6. Additionally, despite the scarcity of depth data, Type 7 was generally characterised by a greater average depth to the river channel (Figure 2.11), indicating that this type reflects the lowland main river stem of large rivers. Type 5 was characterised by large widths (Figure 2.10) and large distances from source (Figure 2.9) and yet were characterised by gradients and altitude of the upper lowland types, indicating that this is the upper lowland type of large rivers, in particular the main river stem of large rivers.

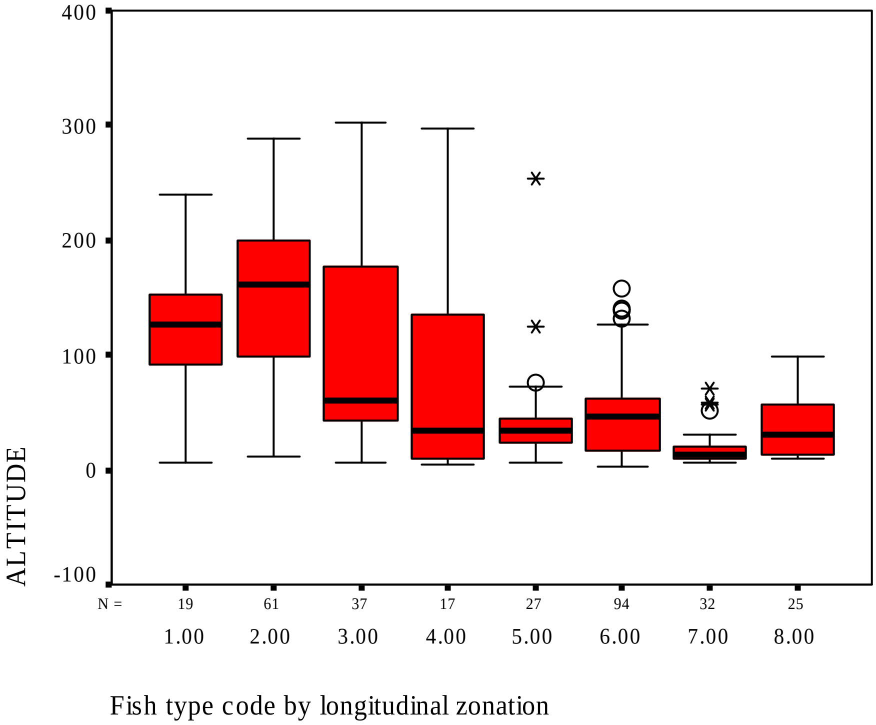

**Figure 2.4** Box plot of site altitude (m) per fish type.

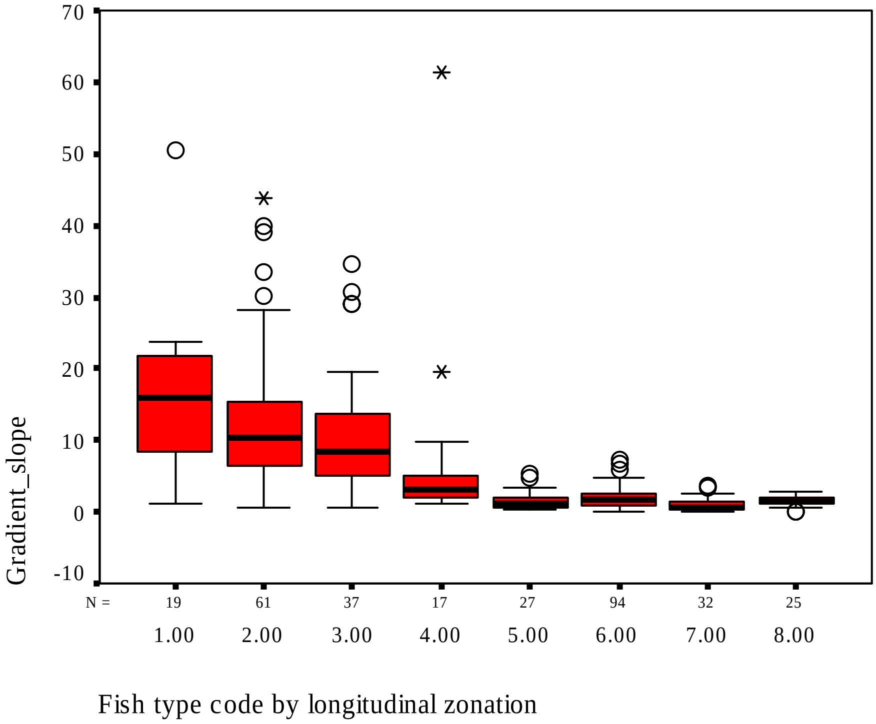

**Figure 2.5** Box plot of site gradient slope per fish type.

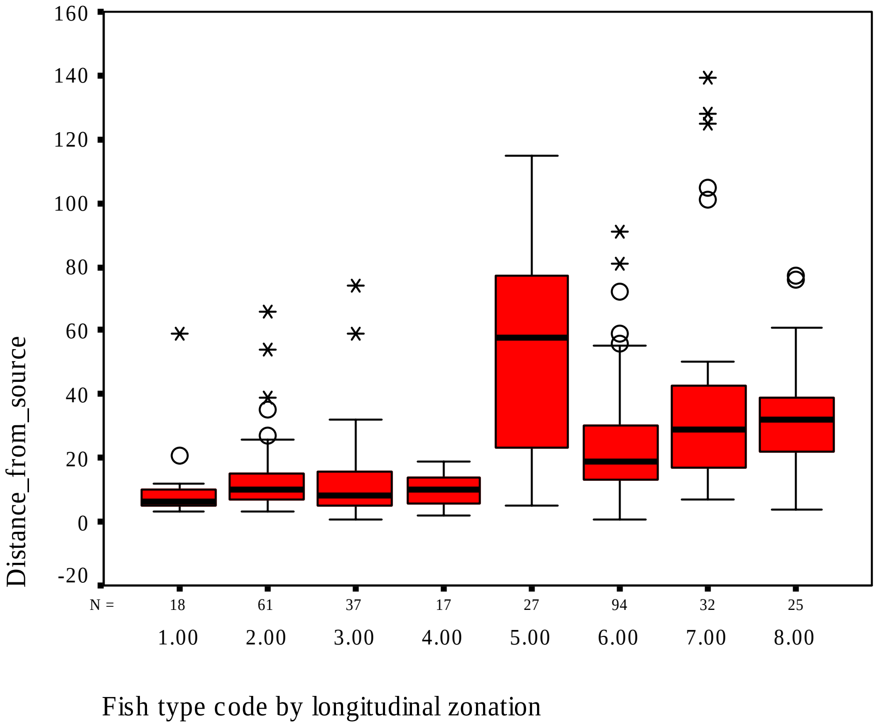

**Figure 2.6** Box plot of site distance from source per fish type.

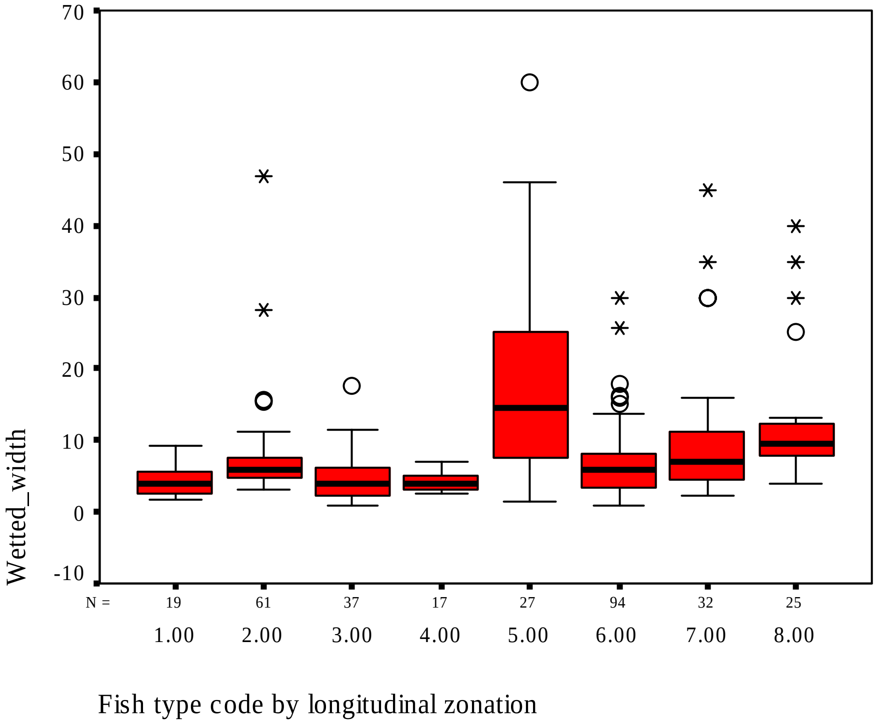

**Figure 2.7** Box plot of site wetted width (m) per fish type.

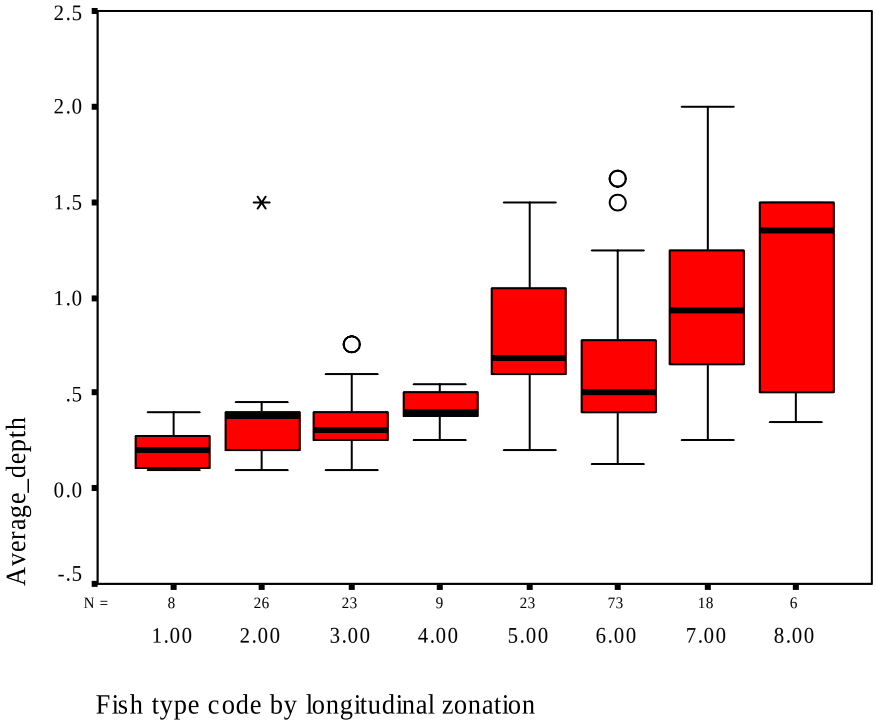

**Figure 2.8** Box plot of site average depth (m) per fish type.

Analysis of the distribution of sites between the main river and geographic regions of the UK indicated that some of the fish types present are important regional types. Type 8, the grayling/salmon/pike type, was predominantly found in the calcareous spring fed rivers of Hampshire on the south coast of England.

## 2.3. River basin characterisation – Abiotic typologies

### 2.3.1. *UK National WFD Typology*

The national typology for rivers to be adopted by the UK is still in preparation by the UK Rivers Technical Advisory Group (UK RIVERS TAG). However, it is almost certain that the national typology will follow the WFD Typology scheme A, and act as a reporting typology. This typology will be tested against each of the biotic communities representing Ecological Quality Elements to determine the typology's ability to predict community types and to detect any significant regional deviations. It is probable that the UK FAME partners will be asked to test this typology for fish communities using the FAME spatial approach. However, at this stage data for Scotland, which are suitable for analysis with the FAME approach, are yet to be sourced.

### 2.3.2. *Integrating abiotic characteristics to the fish community typology*

The distribution of each of the main abiotic site characteristics (example box plots Figures 2.4 – 2.8) and geographical distribution of sites between river regions in each of the sub-clusters identified in the species abundance dendrogram was examined. Many of the abiotic variables used to describe the rivers were significantly correlated (Table 2.4), especially the correlations between the site characteristics (e.g. gradient, altitude, catchment size etc.) and the geographic location of the site within the UK (e.g. coastal location of estuary). Strong correlations between these variable reflect the differences between the east and west coast of Britain, with the majority of large lowland rivers (e.g. Thames, Trent, Yorkshire Ouse, Great Ouse) flowing to the east coast and into the North Sea whilst west coast rivers, with the exception of the Rivers Severn and Mersey, are generally characterised as short with higher gradients and altitudes. Discriminant analysis was undertaken to determine the abiotic characteristics which best discriminated between the eight fish types identified by cluster analysis.

The abiotic variables, with the exception of the categorical variables, were Log10 transformed as they exhibited non-normal distributions and failed homogeneity of variance tests between the eight fish types. The abiotic characteristics of each fish types exhibited significant differences (Table 2.5)

**Table 2.4** Spearman rank correlations between abiotic variables. (Correlations significant at p <0.01 are flagged in bold type, grey shading indicates correlation > 0.5)

| | *Easting* | *Northing* | *West Coast rivers* | *East Coast rivers* | *English Channel rivers* | *Size of catchment* | *Altitude* | *Gradient_slope* | *Distance from source* | *Distance_to_mouth* | *Mean annual °C* | *Mean January °C* | *Mean July °C* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Northing* | **-0.46** | | | | | | | | | | | | |
| *West Coast rivers* | **-0.71** | **0.33** | | | | | | | | | | | |
| *East Coast rivers* | **0.69** | -0.02 | **-0.67** | | | | | | | | | | |
| *English Channel rivers* | 0.07 | **-0.57** | **-0.24** | **-0.40** | | | | | | | | | |
| *Size of catchment* | **0.35** | -0.08 | **-0.38** | **0.46** | **-0.11** | | | | | | | | |
| *Altitude* | **-0.60** | **0.45** | **0.40** | **-0.28** | **-0.24** | **-0.38** | | | | | | | |
| *Gradient_slope* | **-0.56** | **0.29** | **0.46** | **-0.50** | 0.01 | **-0.72** | **0.60** | | | | | | |
| *Distance from source* | **0.29** | 0.01 | **-0.32** | **0.44** | **-0.18** | **0.94** | **-0.38** | **-0.67** | | | | | |
| *Distance_to_mouth* | **-0.31** | **0.41** | **0.15** | 0.04 | **-0.28** | 0.03 | **0.70** | **0.13** | 0.00 | | | | |
| *Mean annual °C* | **0.54** | **-0.83** | **-0.34** | **0.17** | **0.40** | **0.22** | **-0.47** | **-0.40** | **0.11** | **-0.32** | | | |
| *Mean January °C* | **0.34** | **-0.74** | **-0.22** | -0.02 | **0.48** | **0.13** | **-0.36** | **-0.32** | 0.03 | **-0.15** | **0.87** | | |
| *Mean July °C* | **0.76** | **-0.68** | **-0.43** | **0.50** | 0.03 | **0.29** | **-0.49** | **-0.48** | **0.23** | **-0.35** | **0.81** | **0.51** | |
| *Wetted_width* | **0.23** | **0.09** | **-0.30** | **0.41** | **-0.19** | **0.74** | **-0.24** | **-0.49** | **0.72** | 0.02 | 0.02 | **-0.10** | **0.16** |

**Table 2.5** Tests of equality of group means for Log10 transformed abiotic variables between each fish type.

| | Wilks' Lambda | F | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Log Distance from source | 0.644 | 23.841 | 7 | 302 | 0.000 |
| Log Distance to mouth | 0.914 | 4.075 | 7 | 302 | 0.000 |
| Wading sites | 0.719 | 16.850 | 7 | 302 | 0.000 |
| Calcareous geology | 0.892 | 5.219 | 7 | 302 | 0.000 |
| Log Catchment size | 0.575 | 31.841 | 7 | 302 | 0.000 |
| Log Altitude | 0.633 | 25.033 | 7 | 302 | 0.000 |
| Log Gradient +1 | 0.374 | 72.287 | 7 | 302 | 0.000 |
| Log width | 0.738 | 15.355 | 7 | 302 | 0.000 |
| Hampshire Rivers | 0.692 | 19.169 | 7 | 302 | 0.000 |

A stepwise discriminant analysis was undertaken using critical values of Wilk’s Lambda to determine exclusion or inclusion of an abiotic variable at each step of the analysis. Critical values of F = 3.84 and F = 2.71 were used to determine variable entry or removal respectively. A six-step analysis was undertaken and included gradient, Hampshire rivers, distance to mouth, altitude and catchment size and significant discriminant variables (Table 2.7 and 2.8). Given the scarcity of depth data, the variable “wading sites” (i.e. using the variable wading or boat) was used to retain as many sites in the analysis as possible. This was possible, as wading sites must have an average depth of less than 0.7 to 1m for health and safety reasons and therefore gave some reflection of site depth.

**Table 2.6** Abiotic variables included in a stepwise discriminant analysis for the nine fish types.

| Step | | Tolerance | F to Remove | Wilks' Lambda |
|---|---|---|---|---|
| 1 | Log Gradient +1 | 1.000 | 72.287 | |
| 2 | Log Gradient +1 | 0.988 | 66.032 | 0.692 |
| | Hampshire Rivers | 0.988 | 15.858 | 0.374 |
| 3 | Log Gradient +1 | 0.931 | 58.464 | 0.512 |
| | Hampshire Rivers | 0.983 | 16.086 | 0.298 |
| | Log width | 0.940 | 11.154 | 0.273 |
| 4 | Log Gradient +1 | 0.910 | 47.032 | 0.371 |
| | Hampshire Rivers | 0.983 | 15.584 | 0.241 |
| | Log width | 0.918 | 12.079 | 0.226 |
| | Wading sites | 0.942 | 9.733 | 0.217 |
| 5 | Log Gradient +1 | 0.883 | 49.566 | 0.335 |
| | Hampshire Rivers | 0.738 | 16.845 | 0.216 |
| | Log width | 0.914 | 11.165 | 0.195 |
| | Wading sites | 0.942 | 9.699 | 0.190 |
| | Log Distance to mouth | 0.715 | 5.970 | 0.176 |
| 6 | Log Gradient +1 | 0.658 | 13.599 | 0.179 |
| | Hampshire Rivers | 0.727 | 16.619 | 0.189 |
| | Log width | 0.913 | 11.019 | 0.171 |
| | Wading sites | 0.920 | 7.634 | 0.160 |
| | Log Distance to mouth | 0.334 | 8.921 | 0.164 |
| | Log Altitude | 0.314 | 5.963 | 0.155 |

The first abiotic discriminant function was strongly correlated with the gradient of the site (canonical correlation = 0.872) and (Table 2.8) explained 68.9% of the variance in the discriminant model (Table 2.7). The variables included in the next steps further discriminate individual types. Function 2, which is positively correlated with the variable “Hampshire Rivers” distinguishes Type 8, a type characterised by *S. salar* and *T. thymallus*, which is mainly driven by the South coast Chalk streams of Hampshire. Function 3, which is strongly positively correlated with distance from river source and river width distinguishes the lowland river types, especially Type 5 the upper barbel zone in large rivers. Function 4 distinguished between the two major lowland types, correlating with “wading sites” it distinguished Type 7 which was characterised by deeper average depth that the other lowland type, Type 6 (Table 2.8).

**Table 2.7** Eigenvalues and canonical correlations for the five discriminant functions.

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 2.181 | 68.9 | 68.9 | 0.828 |
| 2 | 0.554 | 17.5 | 86.3 | 0.597 |
| 3 | 0.240 | 7.6 | 93.9 | 0.440 |
| 4 | 0.123 | 3.9 | 97.8 | 0.331 |
| 5 | 0.045 | 1.4 | 99.2 | 0.207 |
| 6 | 0.025 | 0.8 | 100.0 | 0.157 |

a First 6 canonical discriminant functions were used in the analysis.

**Table 2.8** Canonical correlation structure matrix between individual abiotic variables and discriminant functions.

| Variable | Function | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Log Gradient +1 | 0.872 | 0.013 | -0.006 | 0.311 | 0.266 | 0.268 |
| Log Catchment size[a] | -0.493 | 0.157 | 0.440 | 0.127 | -0.217 | -0.157 |
| Hampshire Rivers | -0.211 | 0.657 | -0.635 | 0.277 | -0.182 | 0.100 |
| Calcareous geology[a] | 0.137 | 0.264 | 0.068 | 0.005 | 0.004 | -0.123 |
| Log width | -0.236 | 0.380 | 0.758 | 0.290 | -0.366 | -0.085 |
| Log Distance from source[a] | -0.483 | 0.078 | 0.512 | 0.059 | -0.243 | -0.056 |
| Wading sites | 0.313 | 0.432 | 0.040 | -0.744 | 0.176 | 0.359 |
| Log Distance to mouth | -0.021 | -0.093 | 0.477 | 0.066 | 0.862 | -0.126 |
| Log Altitude | 0.490 | 0.162 | 0.257 | 0.022 | 0.677 | -0.457 |

Grey shading indicates largest absolute correlation between each variable and any discriminant function.

[a] This variable not used in the analysis.

**Table 2.9** Standardised canonical discriminant function coefficients for each variable and discriminant function

| | Function | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Log Distance to mouth | -0.778 | -0.016 | 0.254 | 0.406 | 1.109 | 0.966 |
| Wading sites | 0.101 | 0.480 | 0.177 | -0.769 | 0.083 | 0.464 |
| Log Altitude | 0.610 | 0.419 | -0.124 | -0.293 | -0.080 | -1.590 |
| Log Gradient +1 | 0.666 | -0.074 | 0.101 | 0.638 | 0.021 | 0.808 |
| Log width | -0.002 | 0.516 | 0.775 | 0.278 | -0.385 | 0.059 |
| Hampshire Rivers | -0.339 | 0.804 | -0.499 | 0.463 | 0.345 | 0.175 |

Comparison of discriminant function plots for the trimmed calibration dataset grouped by both the types predicted in the cluster analysis with that by the groups predicted in the discriminant analysis indicate that there is considerable overlap and noise in the abiotic characteristics of the identified fish types. This can be particularly seen by the comparison between Figure 2.9 and Figure 2.10.

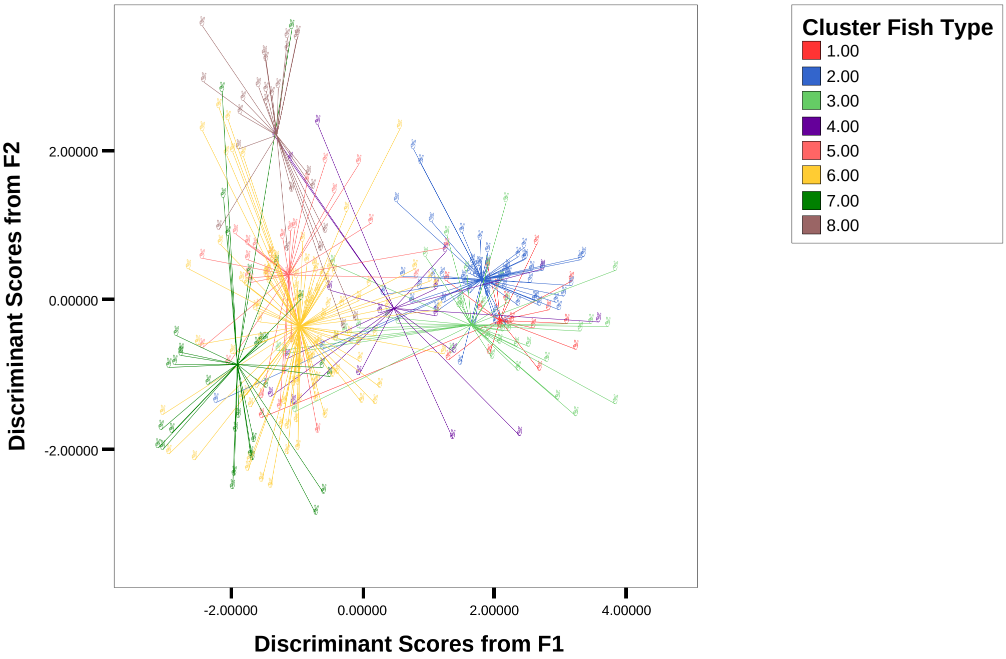

**Figure 2.9** Plot of discriminant scores for the first and second discriminant function for calibration sites. Types are grouped based on the type predicted by the cluster analysis

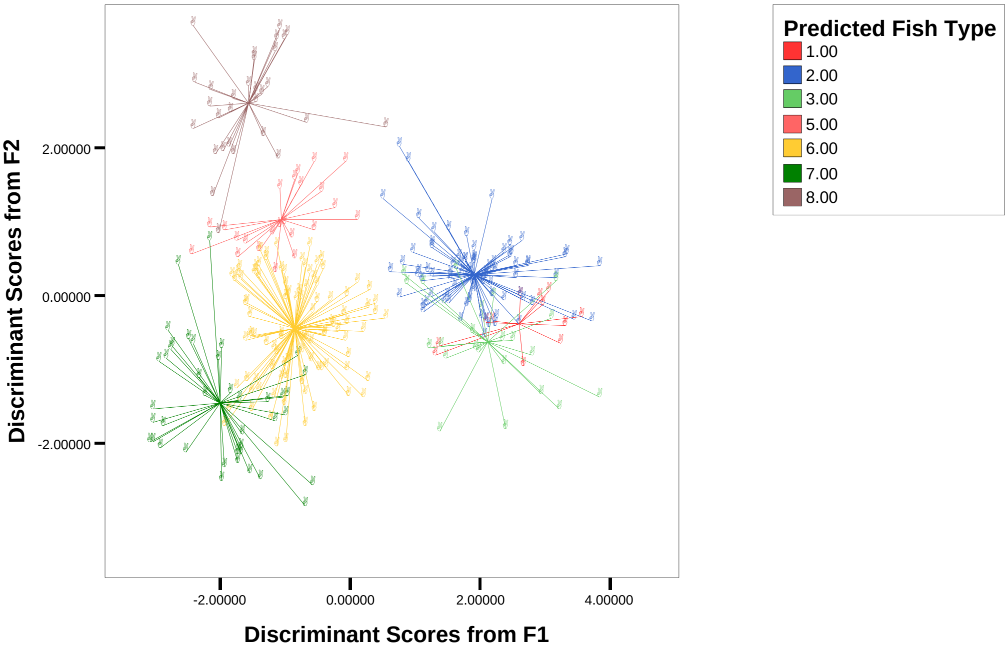

**Figure 2.10** Plot of discriminant scores for the first and second discriminant function for calibration sites. Types are grouped based on the type predicted by the discriminant analysis.

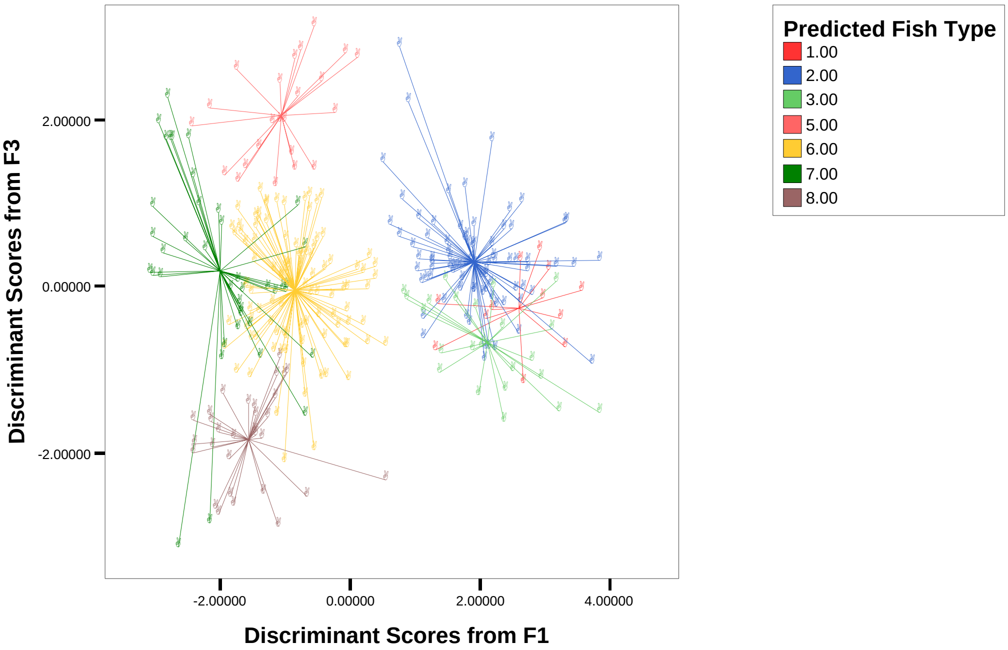

**Figure 2.11** Plot of discriminant scores for the second and third discriminant function for calibration sites. Types are grouped based on the type predicted by the discriminant analysis.

Due to the unevenness of cluster size, predictions in discriminant analysis were weighted using the original group size in the calibration dataset. The type prediction by discriminant analysis indicated that more sites were placed into the dominant types 2 and 6 (Table 2.10). In particular in the trimmed calibration dataset the numbers of sites in Type 2 increased from 61 to 83. In contrast the numbers of sites predicted to be members of types 1, 3, 4 and 5 were less than in the actual cluster grouping. In particular the discriminant analysis did not predict any of the calibration dataset to belong to Type 4.

**Table 2.10** Number of sites per fish type as predicted by cluster analysis and the subsequent discriminant analysis.

| Fish Types | Cluster (n) | Discriminant (n) |
|---:|---:|---:|
| 1 | 19 | 12 |
| 2 | 61 | 83 |
| 3 | 37 | 23 |
| 4 | 17 | 0 |
| 5 | 27 | 23 |
| 6 | 94 | 100 |
| 7 | 32 | 40 |
| 8 | 25 | 30 |

Analysis of the distribution of actual fish types between the types predicted by the discriminant analysis indicates that many of the actual community types are spread between different predicted groups (Figure 2.12). The majority of type 2 sites are retained in type 2 by the discriminant model although some are predicted to be type 1 or 3. The prominent feature of the spread of fish types is the prediction that a large proportion of type 3 (brown trout and bullhead type) sites are predicted in the analysis to be type 2 (upland salmon type). Additionally, far fewer sites are placed into type 3 by the discriminant analysis than are placed in the cluster grouping. Type 4 is not predicted at all by the discriminant analysis, however this is only a minor type with few representative samples.

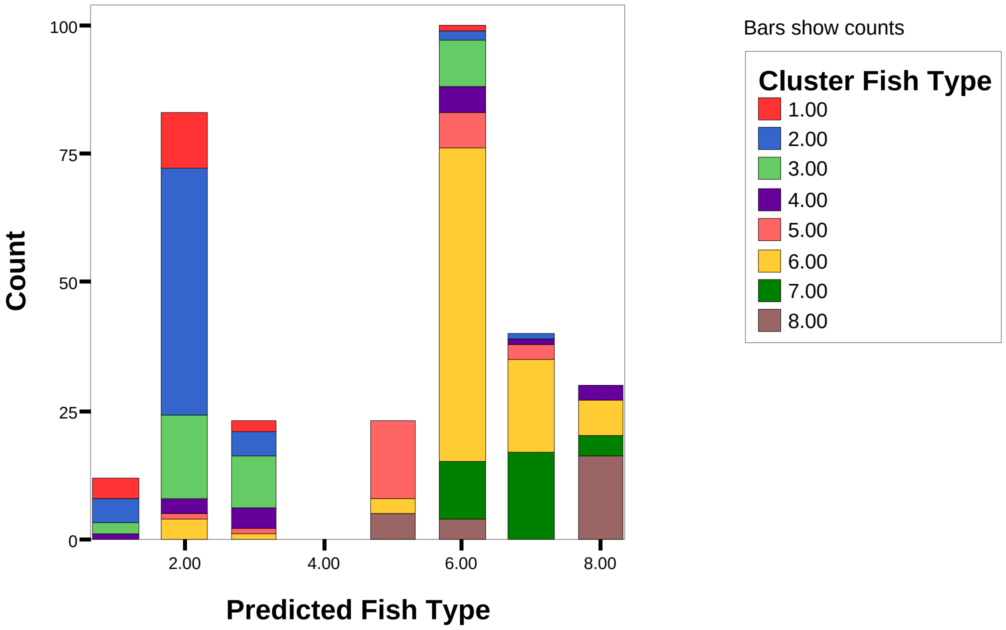

**Figure 2.12** Distribution of fish types from the cluster analysis amongst the predicted fish types from the discriminant analysis.

# 3. ASSESSING REFERENCE AND DEGRADED CONDITIONS

## 3.1. Overview of approach

### 3.1.1. *Designation of degraded sites by discriminant analysis*

The distribution of impacted sites between the eight fish types by the abiotic discriminant model reflected the distribution of sites in the calibration dataset. Types 2, 6 and 7 were the dominant types with few sites belonging to types 4 and 1 (Table 3.1). The numbers of sites predicted for Type 1 was too low for further analysis.

**Table 3.1** Numbers of impacted sites per fish type as designated by discriminant analysis compared with numbers of calibration sites per fish type.

| | Fish Type | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Calibration | 19 | 61 | 37 | 17 | 27 | 94 | 32 | 25 |
| Impacted | 8 | 80 | 27 | 5 | 20 | 266 | 298 | 35 |

Analysis of the impacted sites predicted to belong to each type indicated that the typology and discriminant analysis struggled to define the sites that fell into Type 4. This type was excluded from further analyses.

Within each fish type the impact status of sites were unevenly distributed. The low number of sites in impact categories 1 (presumed high status) and 5 (presumed bad status) didn't allow the potential for description of high quality reference condition for most types. Additionally the low number of sites of presumed bad status did not allow the use of the suggested modified EQR (Spatial Approach guidelines v2.0). Therefore, any EQR could only be standardised using the calibration point alone.

### 3.1.2. *Overview of metric selection procedure*

Only data from the most recent fishing occasion at each site were used to screen, select, score and test the metrics/potential index. As an initial step all most recent samples were used to screen the potential metrics. All potential metrics (metrics all, native and sentinel) were screened using five approaches:

- Spearman rank correlation between metric values and impacts. Individual impact criteria and average impact scores were correlated with the metrics. Averages of impacts were calculated for both all main impacts (i.e. the hydrological, morphological and water quality criteria and multi-scale connectivity) and for the physicochemical site impacts only (i.e. connectivity excluded).
- Box-plots were produced for each metric per five impact classes based on the sums of impact scores. The five classes were equivalent to the five classes for each individual impairment criteria i.e. Class 1 equals sum of impacts = 5 (based on 5 impact criteria), Class 2 equalled sum of impacts = 6 –10 etc.
- Scatter plots of metrics versus total impact scores.
- Statistical tests (comparisons of means t-tests, not assuming equal variance) for each metric between calibration and impacted sites per fish type.
- Logistic regression to predict calibration sites per type.

From this the best metrics per guild type (Overall community characteristics, Tolerance, Habitat, Reproduction, Longevity, Feeding, Migration) were selected for further analysis and scoring. For the next stage in the analysis the number of fishing occasions per river type was assessed to identify restrictions in data availability for metric and impact analysis. Where possible the available dataset was split into two parts, one to develop the index and one to test the index.

### 3.1.3. *Standard approach to IBI model development*

Analysis of the spread of metrics within and between impact classes with each type showed that there was considerable noise within metrics, even for those which showed some trend with degradation. The biological noise precluded the use of traditional metric scoring procedures (percentile boundaries, individual metric scoring and summation of score). Therefore, discriminant analysis was used to assess differences between impact classes and to create a predictive Biological Integrity Class model. Three discriminant approaches were assessed for their predictive ability:

1) All candidate metrics entered into a stepwise procedure.
2) Only metrics pre-selected by previous methods entered into a stepwise procedure.
3) Pre-selected metrics entered into a one-step discriminant analysis.

For most fish types the use of a one step discriminant model with pre-selected metrics exhibited the best prediction rate of impact class.

## 3.2. Type 1 Upper Trout Zone

Insufficient sites were available to screen metrics or develop an index for this type.

## 3.3. Type 2 Salmon & Trout Zone

Screening of the metrics for Type 2 identified few metrics with strong correlations with impact criteria. The scarcity of biomass records within Type 2 precluded the use of biomass metrics in the analysis. Generally all guild-based community metrics exhibited correlations of <0.5 with overall impact scores. Only the density of long migratory species was strongly correlated to impact. However, this guild only represents *Salmo salar* and Lamprey species (note sea trout rarely recorded in surveys). This is reflected by the strong correlations for the metrics for Atlantic salmon as a sentinel species.

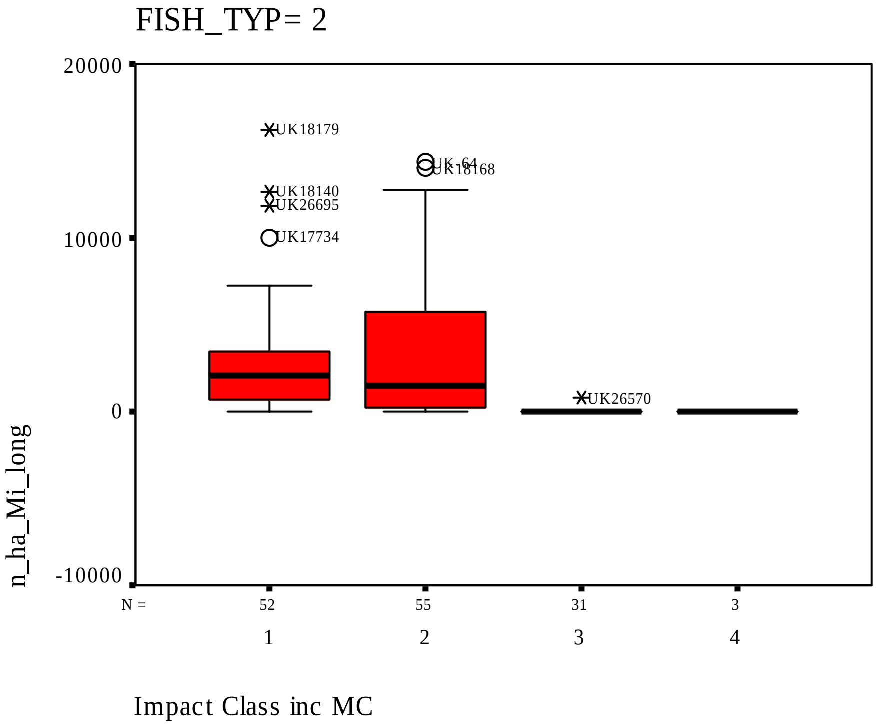

**Figure 3.1** Box plot of metric values for n ha$^{-1}$ of *Salmo salar* per total impact category in river type 2.

For Type 2 stronger correlation was identified between the metrics and total impact scores when multiscale connectivity was included. This reflects the strong correlation between the metrics (particularly Atlantic salmon) and the individual connectivity criteria. It is apparent that in this river type connectivity impacts are the most influential on overall river quality. This distribution of impact classes for each of the individual impact criteria also reflects that in most cases these upland streams have low impact ratings for hydrological, morphological and water quality conditions.

*Salmo salar* and *Salmo trutta fario* are the key species in community-type 2 with the density of both exhibiting a negative correlation with impact. The density of both these species had the major influence over the most responsive guild-based metric. The Density of *S. salar* exhibited the highest correlation with impact, especially connectivity impacts. Salmon were generally absent from sites of moderate or worse quality. However, there was considerable overlap in the spread of density values recorded at sites of presumed high or good status. The densities of salmon recorded in sites of presumed good status ranged from almost zero to around 15000 per hectare, which completely covered the range of values recorded at sites of presumed high status.

The density of trout also exhibited considerable overlap between the impact classes from presumed high to poor status (Figure 3.3).

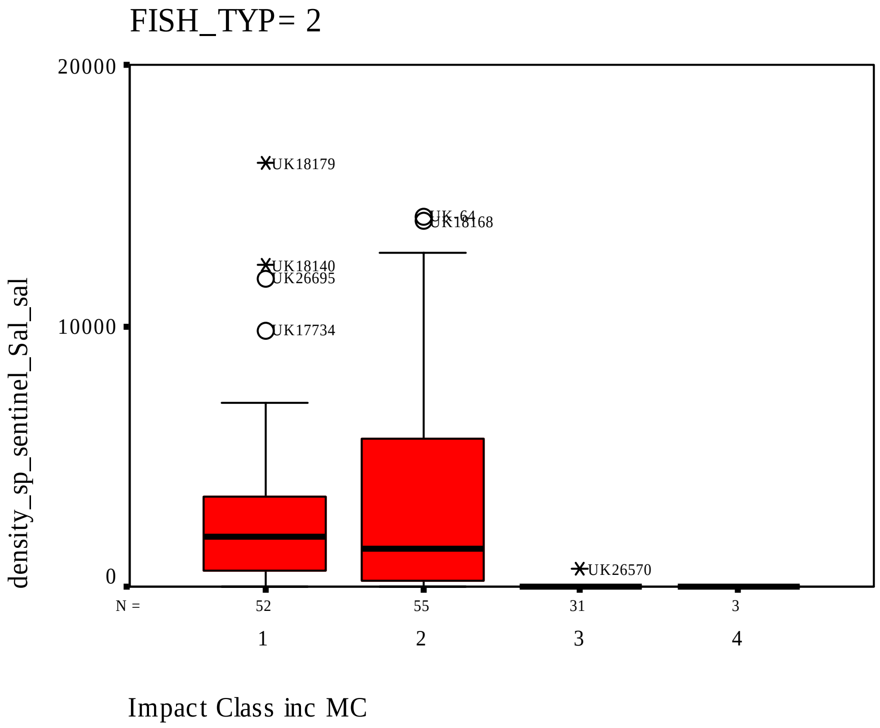

**Figure 3.2** Box plot of metric values for n ha$^{-1}$ of *Salmo salar* per total impact category in river type 2.

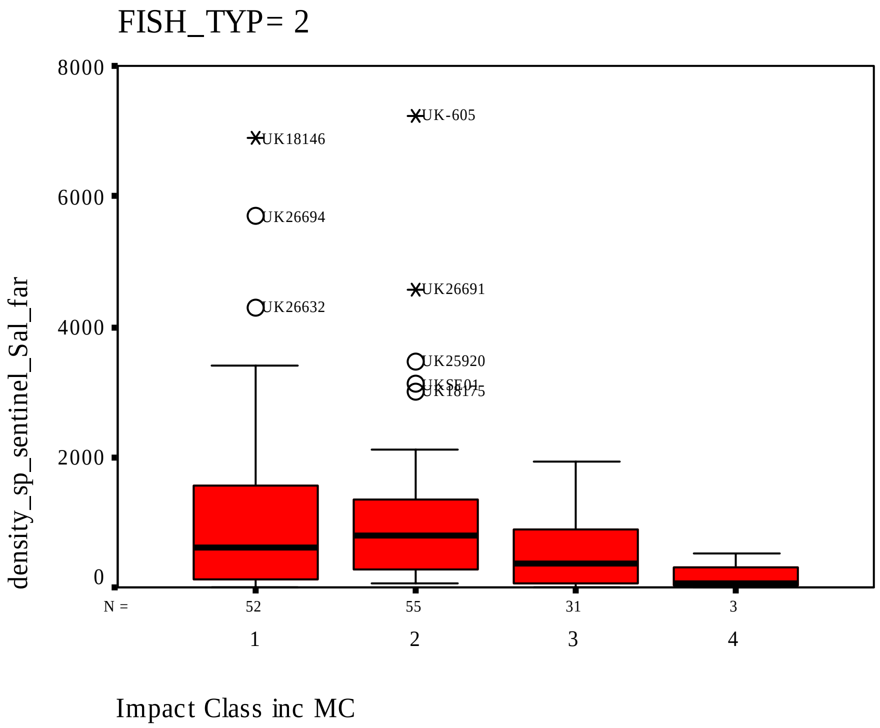

**Figure 3.3** Box plot of metric values for n ha$^{-1}$ of *Salmo trutta fario* per total impact category in river type 2.

The generally low correlation coefficients of the guild-based metrics, together with their correlation with the abundance of salmon and trout, restricted their use as suitable metrics in a multi-metric index. Despite the strong correlation coefficients between the abundance of trout and salmon and the impact variables the considerable range of values recorded and the overlaps in distribution between the impact classes poses considerable problems for setting reference and class boundary values using the approach defined in the guidelines for the spatially-based approach. Analysis of the distribution of the metrics values around their median values indicated that the use of the median as a reference value of a standardised EQR together with the 10th percentile as a class boundary for each impact class did not have sufficient resolution for a five-tier assessment scheme.

Analyses within Type 2 were restricted to three impact classes; high, good and moderate or worse. Fifty sites were randomly selected for evaluation purposes and were excluded from model development. A suite of pre-selected metrics were entered into a one step discriminant analysis (Tables 3.2 to 3.5). Two discriminant functions were derived reflecting the abundance of salmonids and long migratory species. The second function was correlated to the % contribution of 0+ trout to the community, effectively discriminating between high and good status (Figure 3.6).

**Table 3.2** Average group statistics for the pre-selected metrics entered into discriminant analysis for Type 2.

| Metric type | Metric | Impact class for discriminatory analysis | | | |
|---|---|---|---|---|---|
| | | High | Good | Moderate or worse | Total |
| Community | Density_sp_all | 5677.79 | 9235.52 | 1631.82 | 5869.37 |
| Tolerance | n_sp_Intol | 2.21 | 2.38 | 1.42 | 2.04 |
| | n_ha_Intol | 5009.18 | 6238.94 | 1340.66 | 4390.80 |
| Habitat | n_ha_Hab_wc | 4294.58 | 5750.74 | 735.55 | 3799.06 |
| | n_ha_Hab_rh | 5571.13 | 9134.00 | 1524.89 | 5764.67 |
| Reproduction | perc_sp_Re_lith | 81.34 | 87.60 | 76.39 | 82.33 |
| | n_ha_Re_lith | 5571.13 | 9134.00 | 1518.84 | 5762.85 |
| Longevity | perc_sp_Lon_sl | 15.87 | 22.16 | 35.16 | 24.18 |
| | n_ha_Lon_ll | 10.24 | 0.00 | 19.16 | 8.87 |
| | perc_nha_Lon_ll | 2.63 | 0.00 | 2.87 | 1.66 |
| Feeding | n_sp_Fe_insev | 2.18 | 2.36 | 1.42 | 2.02 |
| | n_ha_Fe_insev | 4964.97 | 6238.38 | 1340.66 | 4377.25 |
| Migration | perc_sp_Mi_long | 46.21 | 33.86 | 1.32 | 27.77 |
| | perc_nha_Mi_long | 50.03 | 44.24 | 1.68 | 33.15 |
| Sentinel | density_sp_sentinel_Sal_sal | 2958.37 | 3150.24 | 17.18 | 2147.48 |
| | presence_0plus_Sal_sal | 0.68 | 0.62 | 0.03 | 0.46 |
| | density_0plus_Sal_sal | 1379.08 | 1647.42 | 4.13 | 1070.90 |
| | perc_0plus_Sal_sal | 44.95 | 44.34 | 0.63 | 31.34 |
| | density_sp_sentinel_Sal_far | 1162.76 | 1176.78 | 542.84 | 981.37 |
| | presence_0plus_Sal_far | 0.55 | 0.70 | 0.42 | 0.57 |
| | density_0plus_Sal_far | 615.76 | 527.76 | 96.16 | 424.13 |
| | perc_0plus_Sal_far | 35.03 | 47.16 | 11.16 | 32.64 |

Grey shaded metrics failed the initial tolerance test and were not included in the discriminant analysis.

**Table 3.3** Eigenvalue, variance explanation and canonical correlation of the discriminant functions for Type 2 model.

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 2.405 | 89.0 | 89.0 | 0.840 |
| 2 | 0.296 | 11.0 | 100.0 | 0.478 |

**Table 3.4** Wilks' Lambda and significance of discriminant functions for Type 2 model.

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 through 2 | 0.227 | 169.224 | 38 | 0.000 |
| 2 | 0.772 | 29.547 | 18 | 0.042 |

**Table 3.5** Standardized canonical discriminant function coefficients for Type 2 model.

| Metric | Function | |
|---|---:|---:|
| | 1 | 2 |
| Density_sp_all | 3.576 | 2.659 |
| n_sp_Intol | 1.711 | 0.267 |
| n_ha_Intol | -1.964 | -1.055 |
| n_ha_Hab_wc | -2.987 | -3.309 |
| n_ha_Lon_ll | -0.127 | -0.494 |
| perc_sp_Lon_sl | 0.148 | 0.031 |
| n_sp_Fe_insev | -1.362 | -0.372 |
| perc_nha_Mi_long | -0.174 | -0.156 |
| density_sp_sentinel_Sal_sal | 2.297 | 2.246 |
| presence_0plus_Sal_sal | 0.073 | 0.491 |
| density_0plus_Sal_sal | 0.033 | -0.104 |
| perc_0plus_Sal_sal | 0.293 | -0.351 |
| density_sp_sentinel_Sal_far | 1.627 | 0.048 |
| presence_0plus_Sal_far | -0.381 | 0.352 |
| density_0plus_Sal_far | -0.639 | 1.020 |
| perc_0plus_Sal_far | 0.126 | -1.349 |
| perc_nha_Lon_ll | 0.524 | 0.663 |
| perc_sp_Mi_long | 1.111 | 0.468 |
| perc_sp_Re_lith | 0.334 | -0.285 |

**Table 3.6** Discriminant function model, structure matrix. Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions for Type 2.

| | Function | |
|---|---:|---:|
| | 1 | 2 |
| perc_sp_Mi_long | 0.679 | 0.229 |
| perc_nha_Mi_long | 0.487 | -0.073 |
| presence_0plus_Sal_sal | 0.452 | -0.097 |
| perc_0plus_Sal_sal | 0.403 | -0.180 |
| density_sp_sentinel_Sal_sal | 0.284 | -0.185 |
| perc_sp_Lon_sl | -0.248 | -0.128 |
| density_0plus_Sal_sal | 0.219 | -0.211 |
| density_0plus_Sal_far | 0.175 | 0.000 |
| density_sp_sentinel_Sal_far | 0.159 | -0.086 |
| perc_0plus_Sal_far | 0.256 | -0.408 |
| n_ha_Hab_wc | 0.287 | -0.404 |
| n_ha_Re_lith[a] | 0.181 | -0.400 |
| n_ha_Hab_rh[a] | 0.181 | -0.400 |
| Density_sp_all | 0.182 | -0.397 |
| n_sp_Fe_insev | 0.378 | -0.393 |
| n_sp_Intol | 0.378 | -0.380 |
| n_ha_Fe_insev[a] | 0.276 | -0.348 |
| n_ha_Intol | 0.272 | -0.344 |
| perc_sp_Re_lith | 0.122 | -0.327 |
| presence_0plus_Sal_far | 0.117 | -0.293 |
| perc_nha_Lon_ll | -0.038 | 0.217 |
| n_ha_Lon_ll | -0.083 | 0.211 |

a This variable not used in the analysis.

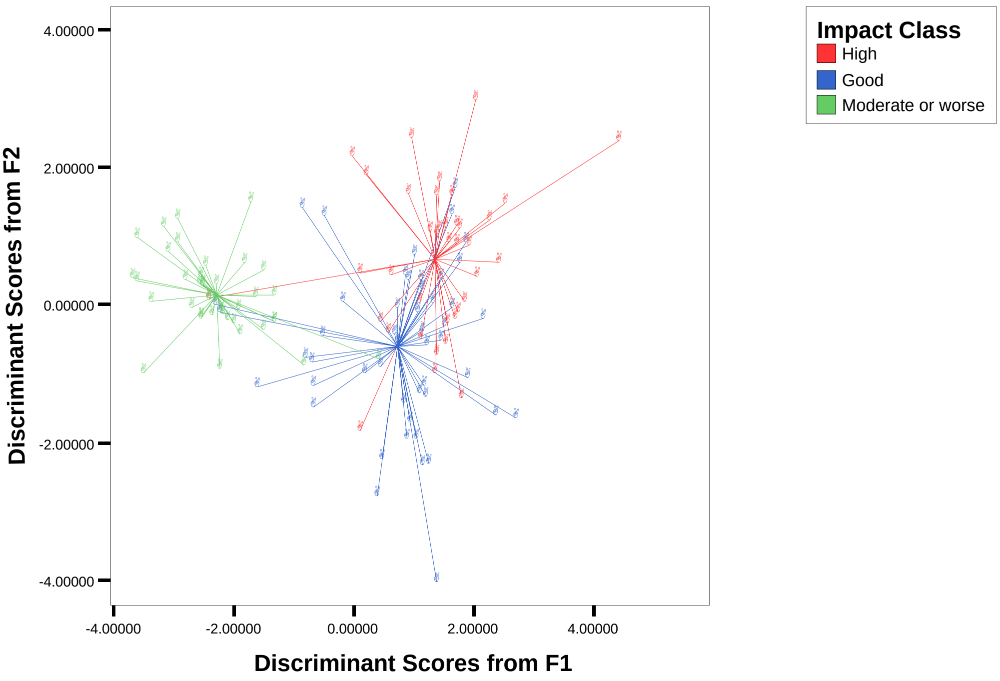

**Figure 3.4** Discriminant plot for sites used in the discriminant model to determine impact status for Type 2.

The discriminant model had good predictive power to distinguish the three impact classes. Within the model development dataset 95% of presumed moderate sites were predicted into the same class, and 71% of high quality sites were predicted to be high quality (Figure 3.5). Few sites of presumed impact status higher than moderate were predicted to be in moderate condition.

Within the evaluation dataset prediction rates and matches were similar to the model development dataset. Only 13% of presumed good quality sites were predicted to be moderate in quality and no moderate sites were predicted to be better than they were presumed to be (Figure 3.6).

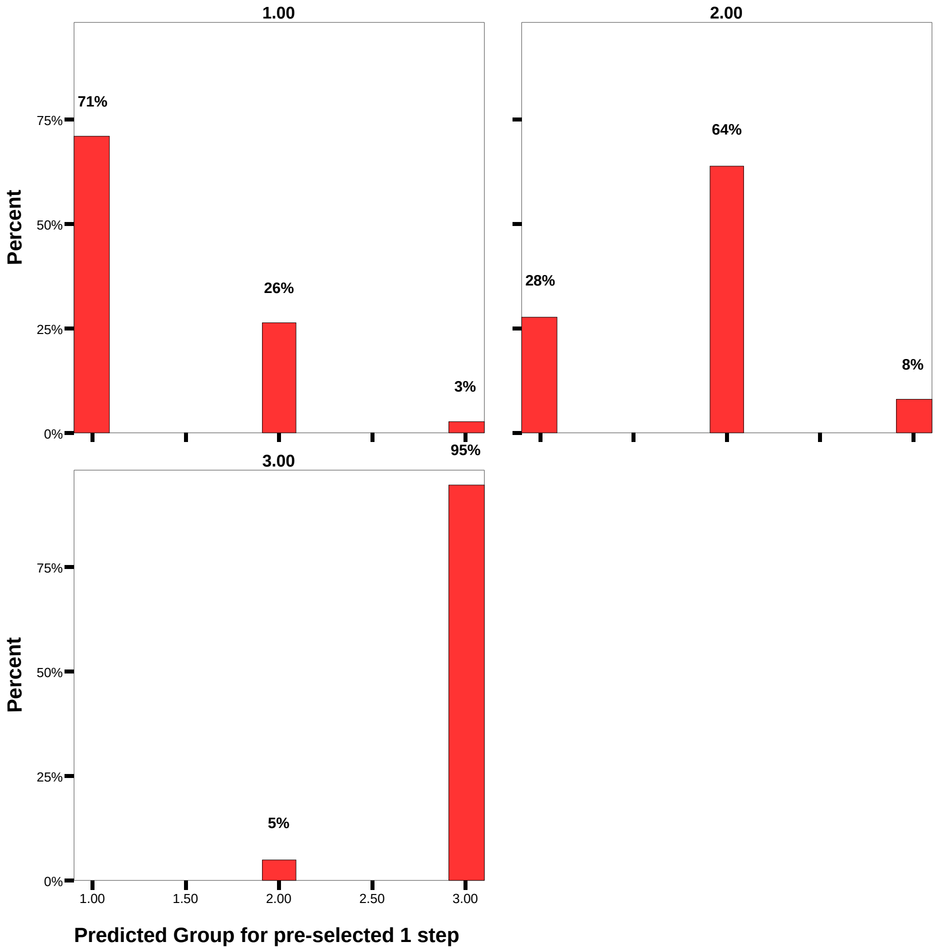

**Figure 3.5** Matching of predicted status to presumed ecological status for sites used in the discriminant analysis for Type 2. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

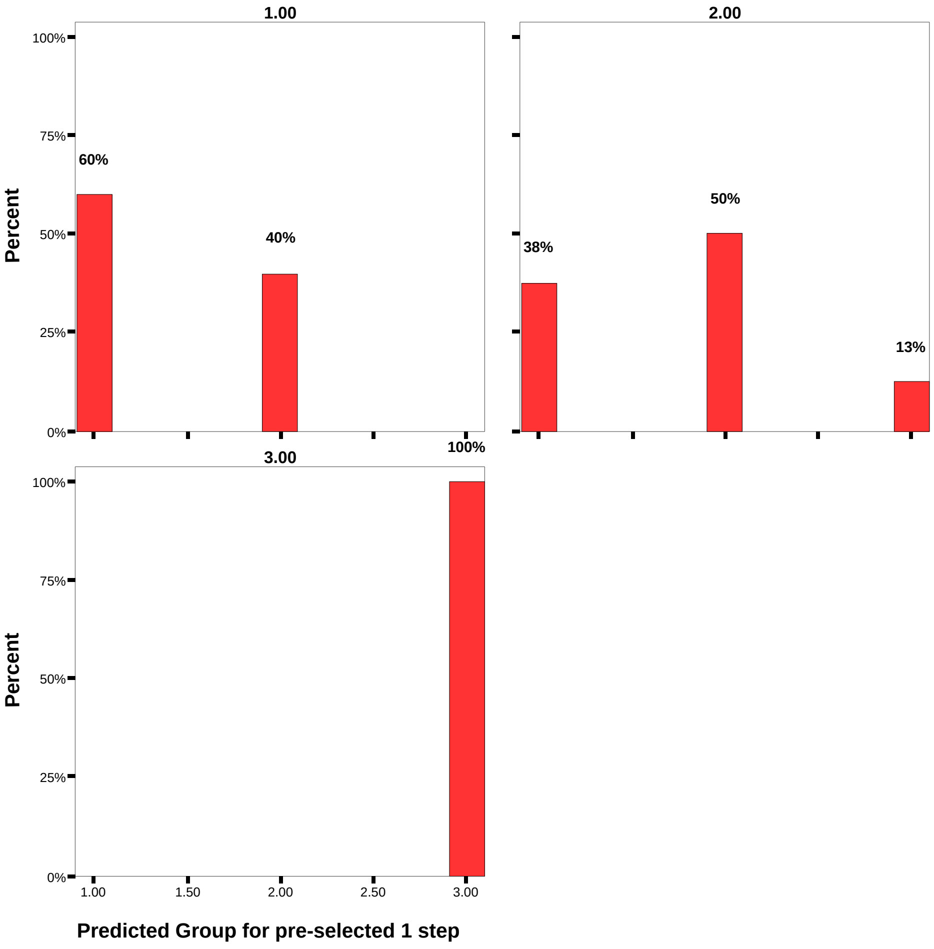

**Figure 3.6** Matching of predicted status to presumed ecological status for 30 randomly selected sites used to evaluate the discriminant analysis for Type 2. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

## 3.4. Type 3 Lower Trout Zone

The scarcity of biomass records within this type precluded the assessment of biomass metrics. Analyses were limited to three presumed status classes; high, good and moderate. The sample size was too small to allow a separate evaluation dataset.

**Table 3.7** Average group statistics for the pre-selected metrics entered into discriminant analysis for Type 3.

| Metric type | Metric | Impact class for discriminant analysis | | | |
|---|---|---|---|---|---|
| | | High | Good | Moderate | Total |
| Community | N_sp_all | 3.31 | 3.47 | 3.88 | 3.48 |
| | Density_sp_all | 5185.13 | 7527.44 | 3792.63 | 6404.85 |
| Tolerance | perc_sp_Intol | 68.88 | 65.31 | 42.88 | 63.27 |
| Habitat | perc_sp_Hab_rh | 76.38 | 74.81 | 58.75 | 73.08 |
| | n_sp_Hab_eury | 0.63 | 0.92 | 1.63 | 0.93 |
| Reproduction | perc_sp_Re_lith | 76.38 | 74.47 | 54.50 | 72.32 |
| | perc_sp_Re_phyt | 0.00 | 0.61 | 4.25 | 0.93 |
| | n_ha_Re_phyt | 0.00 | 2.78 | 12.63 | 3.35 |
| Feeding | perc_nha_Fe_omni | 0.00 | 1.31 | 21.75 | 3.68 |
| Sentinel | density_sp_sentinel_Sal_sal | 413.19 | 192.44 | 129.25 | 242.88 |
| | presence_0plus_Sal_sal | 0.13 | 0.22 | 0.13 | 0.18 |
| | density_0plus_Sal_sal | 88.31 | 79.92 | 37.75 | 76.53 |
| | perc_0plus_Sal_sal | 11.44 | 14.33 | 3.63 | 12.13 |
| | density_sp_sentinel_Sal_far | 1792.25 | 1962.89 | 1166.88 | 1811.25 |
| | presence_0plus_Sal_far | 0.44 | 0.67 | 0.50 | 0.58 |
| | density_0plus_Sal_far | 307.75 | 948.28 | 125.50 | 667.77 |
| | perc_0plus_Sal_far | 21.50 | 26.89 | 16.25 | 24.03 |
| | density_sp_sentinel_Thy_thy | 0.00 | 5.86 | 0.00 | 3.52 |
| | presence_0plus_Thy_thy | 0.00 | 0.03 | 0.00 | 0.02 |
| | density_0plus_Thy_thy | 0.00 | 3.44 | 0.00 | 2.07 |
| | perc_0plus_Thy_thy | 0.00 | 1.64 | 0.00 | 0.98 |

Grey shaded metrics failed the initial tolerance test and were not included in the discriminant analysis.

**Table 3.8** Eigenvalue, variance explanation and canonical correlation of the discriminant functions for Type 3 model.

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 1.036 | 68.9 | 68.9 | 0.713 |
| 2 | 0.467 | 31.1 | 100.0 | 0.564 |

**Table 3.9** Wilks' Lambda and significance of discriminant functions for Type 3 model.

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 through 2 | 0.335 | 54.164 | 32 | 0.009 |
| 2 | 0.682 | 18.973 | 15 | 0.215 |

**Table 3.10** Standardized canonical discriminant function coefficients for Type 3 model.

| Metrics | Function 1 | 2 |
|---|---:|---:|
| N_sp_all | -0.669 | -2.296 |
| Density_sp_all | -0.500 | 0.444 |
| perc_sp_Intol | -1.163 | -0.385 |
| perc_sp_Hab_rh | -1.636 | -10.488 |
| n_sp_Hab_eury | 1.045 | 2.821 |
| perc_sp_Re_lith | 3.269 | 12.555 |
| perc_sp_Re_phyt | -0.750 | 2.434 |
| perc_nha_Fe_omni | 1.644 | -0.860 |
| density_sp_sentinel_Sal_sal | -0.469 | -0.317 |
| presence_0plus_Sal_sal | 1.753 | 0.797 |
| density_0plus_Sal_sal | 0.184 | 0.394 |
| perc_0plus_Sal_sal | -1.538 | -0.413 |
| density_sp_sentinel_Sal_far | 0.815 | -0.097 |
| presence_0plus_Sal_far | 0.536 | 0.303 |
| density_0plus_Sal_far | -0.632 | 0.769 |
| perc_0plus_Sal_far | -0.248 | -0.419 |

**Table 3.11** Discriminant function model, structure matrix. Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions for Type 3.

| | Function 1 | 2 |
|---|---:|---:|
| perc_nha_Fe_omni | .590 | -.201 |
| n_sp_Hab_eury | .339 | .065 |
| perc_sp_Intol | -.316 | .054 |
| perc_sp_Re_phyt | .313 | -.049 |
| perc_sp_Re_lith | -.292 | .082 |
| perc_sp_Hab_rh | -.247 | .069 |
| N_sp_all | .101 | .018 |
| density_sp_sentinel_Sal_far | -.099 | .098 |
| density_0plus_Sal_sal | -.066 | .007 |
| density_0plus_Sal_far | -.078 | .316 |
| presence_0plus_Sal_far | -.004 | .315 |
| Density_sp_all | -.070 | .194 |
| presence_0plus_Sal_sal | -.024 | .178 |
| density_sp_sentinel_Sal_sal | -.107 | -.174 |
| n_ha_Re_phyt[a] | -.013 | .159 |
| perc_0plus_Sal_far | -.070 | .147 |
| perc_0plus_Sal_sal | -.094 | .109 |

a This variable not used in the analysis.

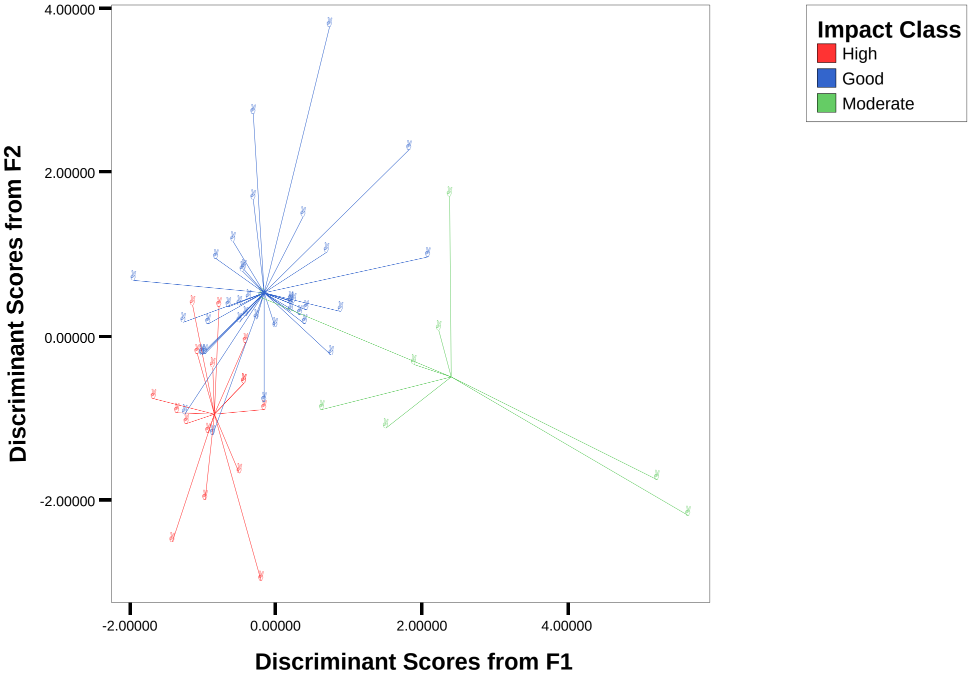

**Figure 3.7** Discriminant plot for sites used in the discriminant model to determine impact status for Type 3.

The discriminant model allowed reasonable distinction between the presumed impact classes. In the case of high or good status sites few were predicted to be of lower than the presumed status and no sites of high status were predicted as moderate status (Figure 3.8).

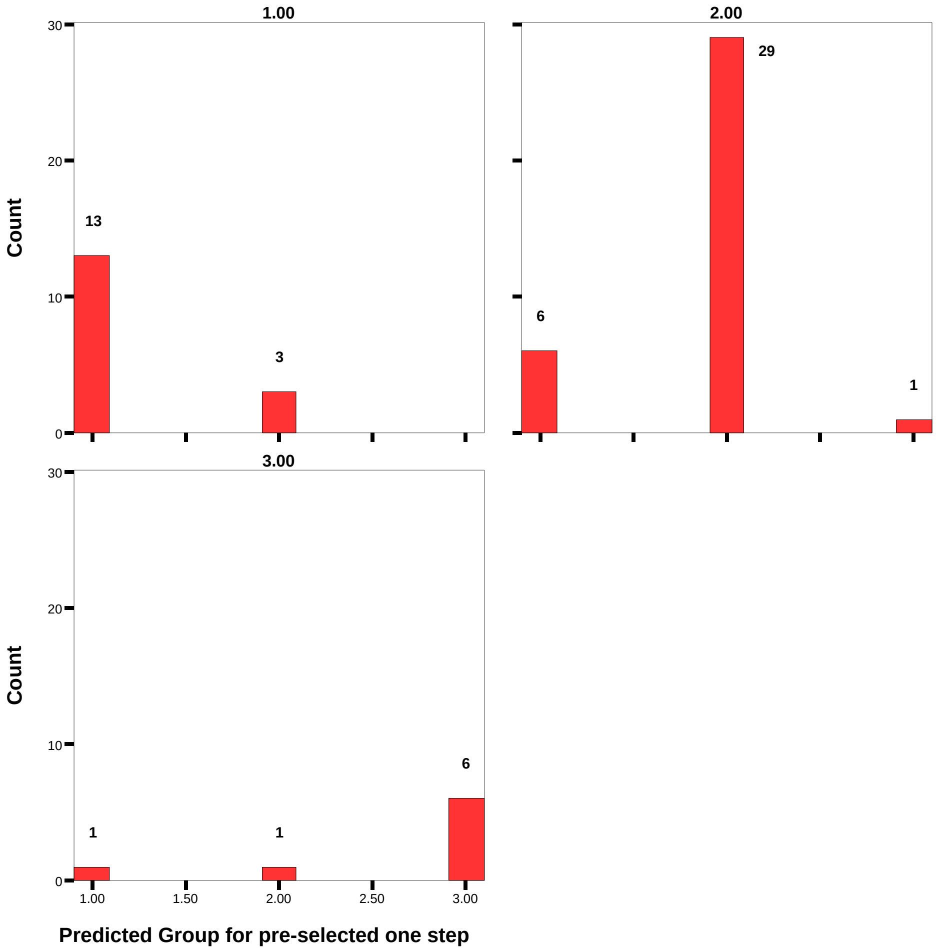

**Figure 3.8** Matching of predicted status to presumed ecological status for sites used in the discriminant analysis for Type 3. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

## 3.5. Type 4 Small Coastal Streams

There were insufficient sites within this type for further analysis.

## 3.6. Type 5 Upper Barbel Zone of Large Rivers

The scarcity of biomass records within this type precluded the assessment of biomass metrics. Analyses were limited to two presumed status classes; good or better and moderate or worse. The sample size was too small to allow a separate evaluation dataset.

**Table 3.12** Average group statistics for the pre-selected metrics entered into discriminant analysis for Type 5.

| Metric | Impact Class for analysis | | |
|---|---|---|---|
| | Good or better | Moderate or worse | Total |
| N_sp_all | 8.32 | 8.00 | 8.16 |
| Density_sp_all | 3720.73 | 1926.43 | 2844.44 |
| perc_sp_Intol | 23.91 | 19.76 | 21.88 |
| perc_nha_Intol | 17.00 | 18.29 | 17.63 |
| perc_nha_Hab_wc | 47.77 | 70.48 | 58.86 |
| perc_nha_Hab_b | 52.23 | 28.38 | 40.58 |
| perc_nha_Hab_rh | 78.77 | 66.19 | 72.63 |
| n_ha_Re_lith | 2870.27 | 1195.10 | 2052.16 |
| n_sp_Re_phyt | 0.45 | 0.90 | 0.67 |
| n_ha_Re_phyt | 7.86 | 11.48 | 9.63 |
| perc_nha_Lon_ll | 13.23 | 29.24 | 21.05 |
| n_ha_Fe_pisc | 7.86 | 10.57 | 9.19 |
| perc_nha_Fe_omni | 12.32 | 44.76 | 28.16 |
| perc_sp_Mi_long | 17.18 | 4.81 | 11.14 |
| perc_nha_Mi_long | 15.00 | 4.14 | 9.70 |
| n_ha_Mi_potad | 95.77 | 132.29 | 113.60 |
| density_sp_sentinel_Bar_bab | 19.45 | 7.10 | 13.42 |
| density_sp_sentinel_Leu_cep | 69.27 | 77.19 | 73.14 |
| density_sp_sentinel_Leu_leu | 45.45 | 100.00 | 72.09 |
| density_sp_sentinel_Rut_rut | 60.82 | 101.33 | 80.60 |
| density_sp_sentinel_Sal_sal | 19.18 | 0.10 | 9.86 |
| density_sp_sentinel_Sal_far | 36.41 | 14.29 | 25.60 |
| presence_0plus_Sal_far | 0.00 | 0.05 | 0.02 |
| density_0plus_Sal_far | 0.00 | 0.95 | 0.47 |
| perc_0plus_Sal_far | 0.00 | 0.62 | 0.30 |
| density_sp_sentinel_Thy_thy | 6.36 | 43.14 | 24.33 |
| presence_0plus_Thy_thy | 0.14 | 0.00 | 0.07 |
| density_0plus_Thy_thy | 2.50 | 0.00 | 1.28 |
| perc_0plus_Thy_thy | 9.73 | 0.00 | 4.98 |

Grey shaded metrics failed the initial tolerance test and were not included in the discriminant analysis.

**Table 3.13** Eigenvalue, variance explanation and canonical correlation of the discriminant functions for Type 5 model.

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 9.616 | 100.0 | 100.0 | 0.952 |

**Table 3.14** Wilks' Lambda and significance of discriminant functions for Type 5 model.

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 | 0.094 | 68.508 | 24 | 0.000 |

**Table 3.15** Standardized canonical discriminant function coefficients for Type 5 model.

| Metric | Function 1 |
|---|---:|
| N_sp_all | 0.601 |
| Density_sp_all | 1.083 |
| perc_sp_Intol | -1.214 |
| perc_nha_Intol | 0.627 |
| perc_nha_Hab_wc | 1.861 |
| perc_nha_Hab_b | 2.724 |
| perc_nha_Hab_rh | 2.020 |
| n_ha_Re_lith | -0.877 |
| n_sp_Re_phyt | 0.705 |
| n_ha_Re_phyt | -5.712 |
| perc_nha_Lon_ll | -0.514 |
| n_ha_ Fe_pisc | 5.311 |
| perc_nha_Fe_omni | 3.284 |
| perc_sp_Mi_long | -0.189 |
| perc_nha_Mi_long | -1.257 |
| n_ha_Mi_potad | 5.795 |
| density_sp_sentinel_Bar_bab | -2.127 |
| density_sp_sentinel_Leu_cep | -6.975 |
| density_sp_sentinel_Leu_leu | -0.399 |
| density_sp_sentinel_Rut_rut | 1.077 |
| density_sp_sentinel_Sal_sal | 0.691 |
| presence_0plus_Thy_thy | -0.132 |
| density_0plus_Thy_thy | -0.080 |
| perc_0plus_Thy_thy | 0.255 |

The discriminant function was linked to both the presence of long migratory species and the contribution of omnivores to the community (Table 3.16).

**Table 3.16** Discriminant function model, structure matrix. Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions for Type 5.

| Metric | Function<br>1 |
|---|---:|
| perc_sp_Mi_long | -.251 |
| perc_nha_Fe_omni | .236 |
| perc_nha_Mi_long | -.176 |
| perc_nha_Hab_b | -.171 |
| density_sp_sentinel_Sal_far[a] | -.168 |
| perc_nha_Hab_wc | .158 |
| perc_nha_Lon_ll | .121 |
| n_sp_Re_phyt | .108 |
| n_ha_Re_lith | -.106 |
| density_sp_sentinel_Bar_bab | -.096 |
| density_sp_sentinel_Sal_sal | -.094 |
| density_sp_sentinel_Thy_thy[a] | .090 |
| presence_0plus_Thy_thy | -.090 |
| perc_0plus_Thy_thy | -.087 |
| perc_nha_Hab_rh | -.085 |
| density_0plus_Thy_thy | -.079 |
| Density_sp_all | -.078 |
| perc_0plus_Sal_far[a] | -.058 |
| density_0plus_Sal_far[a] | -.058 |
| presence_0plus_Sal_far[a] | -.058 |
| density_sp_sentinel_Leu_leu | .051 |
| n_ha_Mi_potad | .039 |
| perc_sp_Intol | -.036 |
| n_ha_Re_phyt | .032 |
| density_sp_sentinel_Rut_rut | .024 |
| n_ha_Fe_pisc | .024 |
| N_sp_all | -.018 |
| perc_nha_Intol | .012 |
| density_sp_sentinel_Leu_cep | .010 |

a This variable not used in the analysis.

The discriminant function enabled a 100% accurate prediction rate within the dataset available (Figure 3.9).

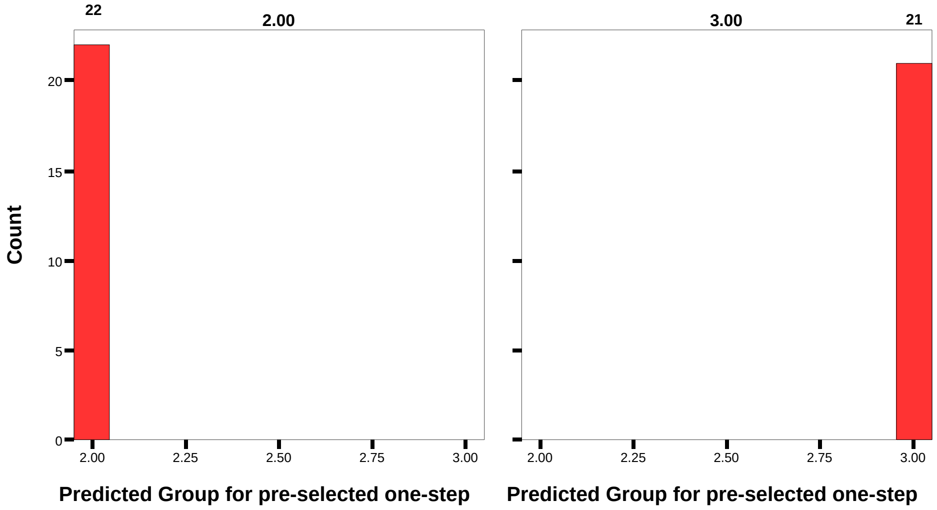

**Figure 3.9** Matching of predicted status to presumed ecological status for sites used in the discriminant analysis for Type 5. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

## 3.7. Type 6 Upper Lowland Coarse Fish Zone

Sufficient data were available within this type to assess both abundance and biomass based metrics. Four ecological status classes were available; good, moderate, poor and bad, although the majority of sites were from the good, moderate and poor status classes. Sufficient sample were available to allow 100 sites to be randomly selected as an evaluation dataset.

**Table 3.17** Average group statistics for the pre-selected metrics entered into discriminant analysis for Type 6.

| | | Impact Class inc MC | | | | |
|---|---|---|---|---|---|---|
| | | Good | Moderate | Poor | Bad | Total |
| Community | Density_sp_all | 6839.84 | 3786.42 | 1847.57 | 649.00 | 3690.49 |
| Tolerance | n_ha_Intol | 1795.77 | 1258.44 | 614.30 | 2.64 | 1094.09 |
| | perc_nha_Tol | 16.79 | 17.28 | 35.65 | 52.09 | 25.71 |
| Habitat | n_ha_Hab_b | 3550.98 | 1966.04 | 914.22 | 193.64 | 1892.25 |
| | n_ha_Hab_rh | 6228.47 | 3418.94 | 1365.53 | 292.55 | 3218.58 |
| Reproduction | n_ha_Re_lith | 6054.11 | 3262.14 | 1192.27 | 116.91 | 3050.23 |
| | perc_nha_Re_phyt | 1.00 | 2.22 | 5.57 | 7.36 | 3.41 |
| Longevity | perc_nha_Lon_ll | 9.40 | 15.06 | 32.52 | 28.45 | 20.77 |
| | n_ha_Lon_sl | 4305.11 | 2386.11 | 846.99 | 188.09 | 2195.34 |
| Feeding | n_ha_Fe_insev | 1795.56 | 1254.50 | 613.47 | 2.64 | 1092.46 |
| | perc_nha_Fe_pisc | 0.32 | 1.96 | 4.64 | 5.00 | 2.69 |
| | perc_nha_Fe_omni | 17.61 | 21.56 | 54.77 | 66.73 | 35.14 |
| Migration | perc_nha_Mi_potad | 4.68 | 8.11 | 14.48 | 10.36 | 9.72 |
| Sentinel | density_sp_sentinel_Bar_bab | 0.28 | 6.90 | 5.46 | 0.00 | 4.33 |
| | biom_sp_sentinel_Bar_bab | 0.01 | 4.85 | 0.86 | 0.00 | 1.89 |
| | density_sp_sentinel_Leu_cep | 149.86 | 127.88 | 168.36 | 37.36 | 144.10 |
| | biom_sp_sentinel_Leu_cep | 61.76 | 79.85 | 47.38 | 13.30 | 59.86 |
| | density_sp_sentinel_Leu_leu | 117.47 | 128.24 | 201.99 | 67.55 | 149.94 |
| | biom_sp_sentinel_Leu_leu | 9.58 | 28.09 | 12.20 | 4.13 | 16.26 |
| | density_sp_sentinel_Rut_rut | 318.47 | 141.69 | 327.57 | 278.36 | 262.80 |
| | biom_sp_sentinel_Rut_rut | 16.71 | 17.55 | 24.03 | 21.31 | 19.93 |
| | density_sp_sentinel_Sal_far | 125.67 | 186.67 | 50.81 | 2.64 | 111.43 |

Grey shaded metrics failed the initial tolerance test and were not included in the discriminant analysis.

**Table 3.18** Eigenvalue, variance explanation and canonical correlation of the discriminant functions for Type 6 model.

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 0.803 | 79.4 | 79.4 | 0.667 |
| 2 | 0.144 | 14.3 | 93.7 | 0.355 |
| 3 | 0.063 | 6.3 | 100.0 | 0.244 |

**Table 3.19** Wilks' Lambda and significance of discriminant functions for Type 6 model.

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 through 3 | 0.456 | 165.434 | 57 | 0.000 |
| 2 through 3 | 0.822 | 41.331 | 36 | 0.249 |
| 3 | 0.940 | 12.955 | 17 | 0.739 |

**Table 3.20** Standardized canonical discriminant function coefficients for Type 6 model.

| Metric | Function<br>1 | <br>2 | <br>3 |
|---|---|---|---|
| Density_sp_all | -7.847 | -8.737 | 3.841 |
| n_ha_Intol | -0.881 | -1.534 | 3.951 |
| perc_nha_Tol | 0.203 | 0.181 | 0.155 |
| n_ha_Hab_b | 2.176 | 3.076 | -5.450 |
| n_ha_Hab_rh | 4.261 | 3.493 | 4.047 |
| perc_nha_Re_phyt | 0.201 | -0.732 | -0.393 |
| perc_nha_Lon_ll | -0.176 | -0.609 | 0.150 |
| n_ha_Lon_sl | 2.184 | 3.258 | -5.944 |
| perc_nha_Fe_omni | 0.973 | -0.167 | -0.787 |
| perc_nha_Mi_potad | 0.018 | 0.884 | 0.551 |
| density_sp_sentinel_Bar_bab | 0.092 | 0.066 | 0.177 |
| biom_sp_sentinel_Bar_bab | -0.030 | 0.395 | -0.210 |
| density_sp_sentinel_Leu_cep | 0.060 | -0.051 | 0.178 |
| biom_sp_sentinel_Leu_cep | -0.258 | -0.132 | -0.280 |
| density_sp_sentinel_Leu_leu | .331 | 0.412 | 0.655 |
| biom_sp_sentinel_Leu_leu | -0.141 | 0.547 | -0.351 |
| biom_sp_sentinel_Rut_rut | 0.180 | 0.346 | 0.301 |
| density_sp_sentinel_Sal_far | 0.207 | 0.631 | -0.615 |
| perc_nha_Fe_pisc | 0.119 | 1.106 | 0.538 |

Three discriminant functions were derived. The first function was positively correlated with the contribution of omnivores and tolerant species to the overall community abundance. The function was also positively correlated to contribution of phytophils and piscivores to the overall community abundance (Table 3.21).

**Table 3.21** Discriminant function model, structure matrix. Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions for Type 6.

| | Function | | |
|---|---:|---:|---:|
| | 1 | 2 | 3 |
| perc_nha_Fe_omni | 0.803 | -0.123 | 0.092 |
| perc_nha_Lon_ll | 0.530 | 0.143 | 0.432 |
| perc_nha_Tol | 0.465 | -0.179 | -0.240 |
| perc_nha_Re_phyt | 0.281 | 0.042 | -0.036 |
| perc_nha_ Fe_pisc | 0.261 | 0.134 | 0.067 |
| n_ha_Lon_sl | -0.228 | -0.218 | 0.024 |
| n_ha_Re_lith | -0.227 | -0.226 | 0.035 |
| n_ha_Intol | -0.148 | -0.080 | 0.076 |
| n_ha_Fe_insev | -0.148 | -0.080 | 0.076 |
| biom_sp_sentinel_Leu_leu | -0.088 | 0.469 | -0.034 |
| density_sp_sentinel_Rut_rut | 0.037 | -0.372 | 0.217 |
| biom_sp_sentinel_Bar_bab | -0.051 | 0.337 | -0.084 |
| density_sp_sentinel_Sal_far | -0.204 | 0.281 | -0.062 |
| Density_sp_all | -0.231 | -0.255 | 0.060 |
| n_ha_Hab_rh | -0.229 | -0.230 | 0.042 |
| density_sp_sentinel_Bar_bab | 0.020 | 0.198 | 0.085 |
| n_ha_Hab_b | -0.187 | -0.196 | 0.057 |
| biom_sp_sentinel_Leu_cep | -0.115 | 0.191 | 0.093 |
| density_sp_sentinel_Leu_leu | 0.127 | 0.066 | 0.548 |
| density_sp_sentinel_Leu_cep | 0.010 | -0.043 | 0.534 |
| perc_nha_Mi_potad | 0.295 | 0.191 | 0.368 |
| biom_sp_sentinel_Rut_rut | 0.084 | 0.002 | 0.104 |

a This variable not used in the analysis.

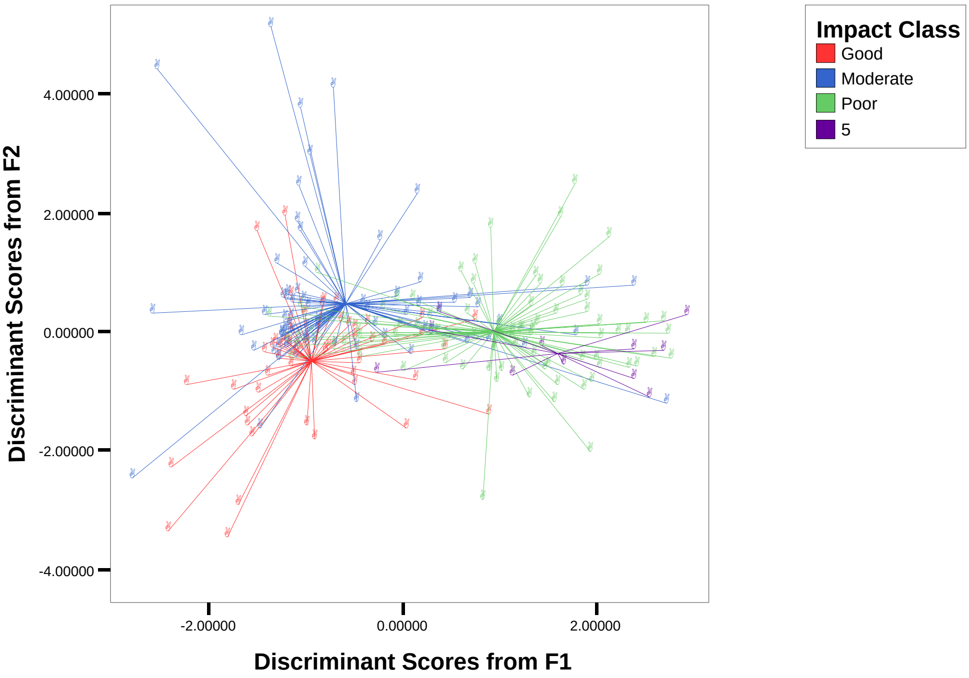

**Figure 3.10** Discriminant plot for sites used in the discriminant model to determine impact status for Type 6.

Within the model development dataset the predictive power of the model was reasonable with correct predictions of good and bad status generally matching the presumed status. However, prediction rates for moderate and poor status were variable with fairly high proportions of cases being miss-classified. Importantly the model had a higher prediction rate for bad status than occurred in the dataset, especially predicting sites of presumed poor status to be bad.

Within the evaluation dataset the prediction rates were variable, with many sites of presumed moderate status being predicted to be of good status. Also, although no sites of presumed bad status were in the evaluation dataset many sites were predicted to be of bad status, including some of presumed good status (Figure 3.11).

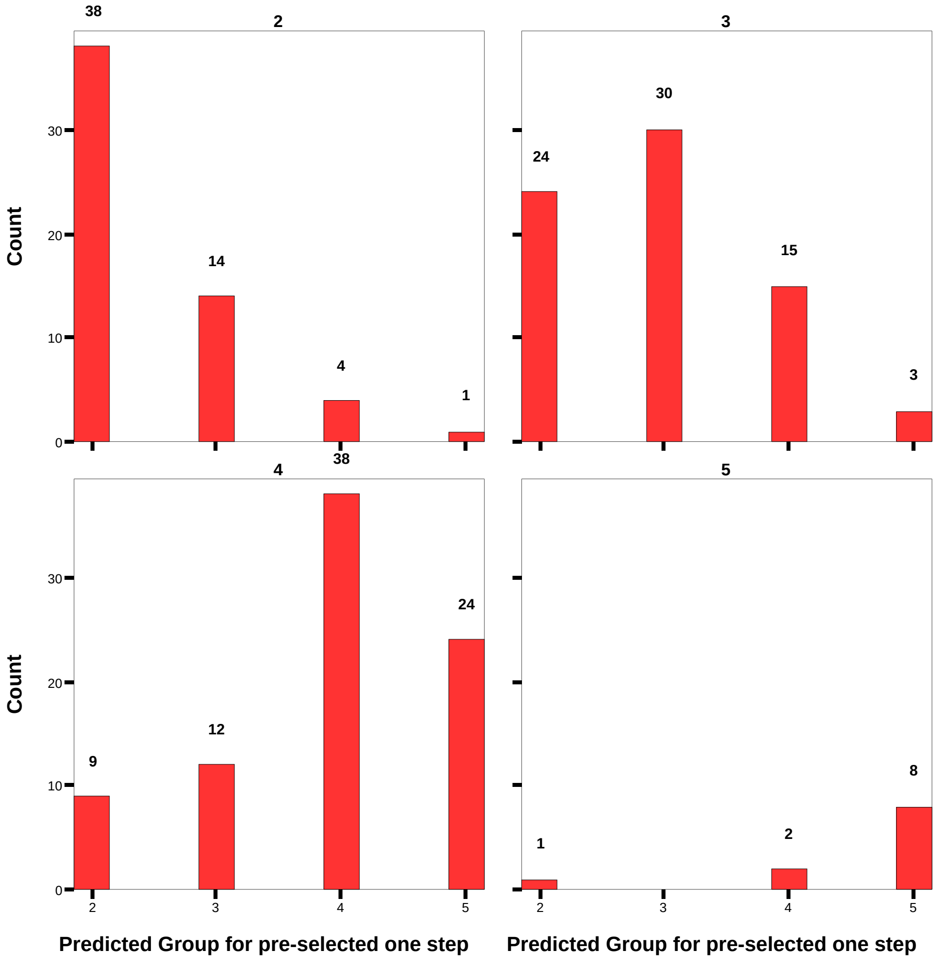

**Figure 3.11** Matching of predicted status to presumed ecological status for sites used in the discriminant analysis for Type 6. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

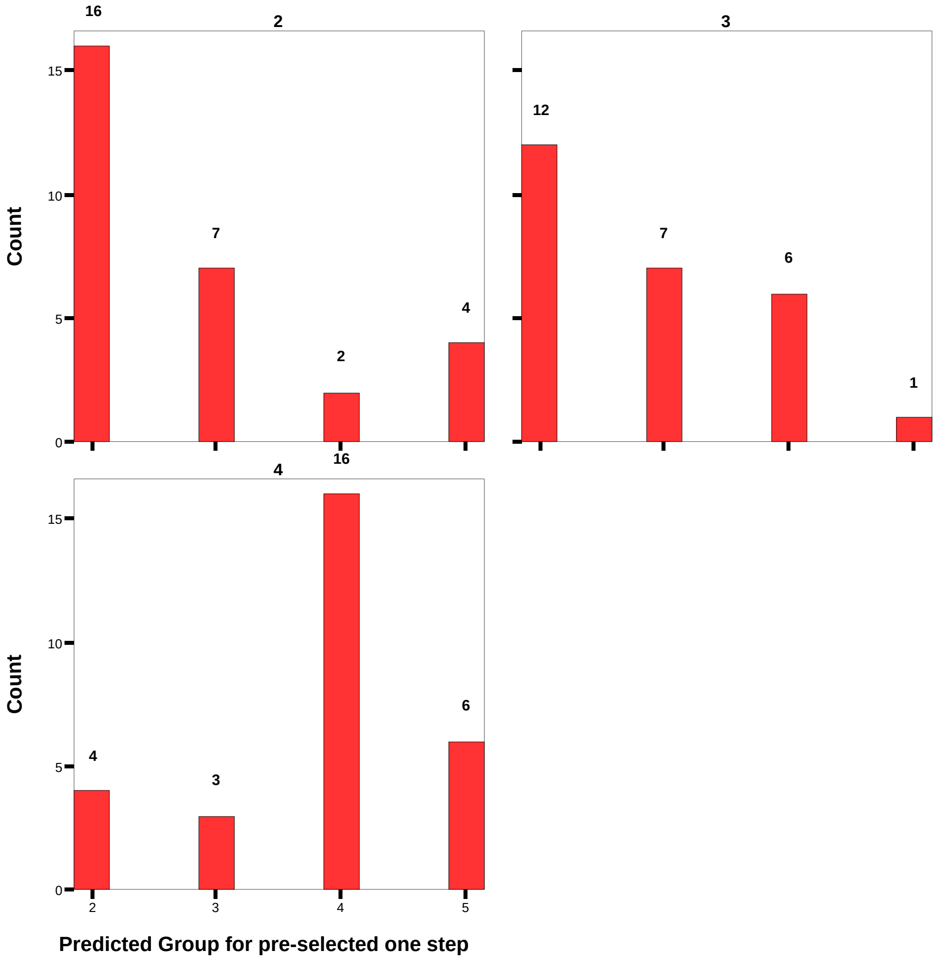

**Figure 3.12** Matching of predicted status to presumed ecological status for 100 randomly selected sites used to evaluate the discriminant analysis for Type 6. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

## 3.8. Type 7 Lower Lowland Coarse Fish Zone (Large Deep Rivers)

Sufficient data were available within this type to assess both abundance and biomass based metrics. Four ecological status classes were available; good, moderate, poor and bad, although the majority of sites were from the moderate and poor status classes. Sufficient sample were available to allow 50 sites to be randomly selected as an evaluation dataset.

**Table 3.22** Average group statistics for the pre-selected metrics entered into discriminant analysis for Type 7.

| | | Impact Class inc MC | | | | |
|---|---|---|---|---|---|---|
| | | Good | Moderate | Poor | Bad | Total |
| Community | N_sp_all | 8.222 | 6.754 | 6.757 | 5.286 | 6.683 |
| | Biom_sp_all | 288.010 | 516.602 | 164.182 | 98.018 | 251.518 |
| Tolerance | perc_nha_Intol | 10.556 | 6.049 | 1.908 | 0.000 | 3.103 |
| | perc_kgha_Tol | 20.222 | 24.492 | 49.796 | 69.286 | 44.033 |
| Habitat | perc_kgha_Hab_wc | 94.889 | 83.672 | 68.724 | 61.905 | 72.856 |
| | perc_sp_Hab_rh | 53.778 | 49.787 | 34.270 | 21.857 | 37.815 |
| | perc_nha_Hab_rh | 65.778 | 44.361 | 27.809 | 16.095 | 32.358 |
| Reproduction | perc_sp_Re_lith | 45.333 | 40.803 | 27.217 | 15.714 | 30.305 |
| | perc_nha_Re_lith | 59.111 | 38.180 | 22.822 | 12.238 | 27.107 |
| | perc_sp_Re_phyt | 14.000 | 14.656 | 18.941 | 17.571 | 17.564 |
| Longevity | perc_nha_Lon_sl | 45.000 | 21.984 | 12.086 | 5.571 | 15.226 |
| Feeding | kg_ha_ Fe_pisc | 62.367 | 262.947 | 33.619 | 23.164 | 91.348 |
| | perc_nha_Fe_insev | 10.556 | 5.721 | 1.743 | 0.000 | 2.918 |
| | perc_nha_Fe_omni | 30.222 | 42.016 | 61.579 | 56.286 | 55.049 |
| Migration | perc_kgha_Mi_long | 0.222 | 7.836 | 15.513 | 27.619 | 14.066 |
| Sentinel | density_sp_sentinel_Bar_bab | 0.000 | 1.459 | 1.987 | 0.000 | 1.609 |
| | biom_sp_sentinel_Bar_bab | 0.000 | 3.299 | 2.096 | 0.000 | 2.139 |
| | density_sp_sentinel_Leu_cep | 127.444 | 31.230 | 67.579 | 15.762 | 56.193 |
| | biom_sp_sentinel_Leu_cep | 101.431 | 57.489 | 25.983 | 8.412 | 35.168 |
| | density_sp_sentinel_Leu_leu | 71.667 | 31.574 | 77.158 | 22.524 | 60.790 |
| | biom_sp_sentinel_Leu_leu | 16.313 | 11.888 | 3.808 | 0.920 | 6.050 |
| | density_sp_sentinel_Rut_rut | 336.556 | 151.049 | 397.046 | 156.667 | 312.280 |
| | biom_sp_sentinel_Rut_rut | 18.948 | 20.950 | 20.846 | 19.534 | 20.689 |
| | density_sp_sentinel_Sal_far | 32.667 | 5.082 | 2.809 | 0.000 | 4.243 |

Grey shaded metrics failed the initial tolerance test and were not included in the discriminant analysis.

**Table 3.23** Eigenvalue, variance explanation and canonical correlation of the discriminant functions for Type 7 model.

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 0.785 | 69.6 | 69.6 | 0.663 |
| 2 | 0.232 | 20.6 | 90.1 | 0.434 |
| 3 | 0.111 | 9.9 | 100.0 | 0.317 |

**Table 3.24** Wilks' Lambda and significance of discriminant functions for Type 7 model.

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 through 3 | 0.409 | 203.834 | 72 | 0.000 |
| 2 through 3 | 0.730 | 71.701 | 46 | 0.009 |
| 3 | 0.900 | 24.091 | 22 | 0.342 |

**Table 3.25** Standardized canonical discriminant function coefficients for Type 7 model.

| | Function | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| N_sp_all | 0.096 | 0.321 | -0.180 |
| Biom_sp_all | -0.135 | -0.545 | -0.392 |
| perc_nha_Intol | 0.803 | -0.166 | -0.536 |
| perc_kgha_Tol | 0.174 | 0.294 | 0.733 |
| perc_kgha_Hab_wc | -0.138 | 0.030 | 0.586 |
| perc_sp_Hab_rh | -0.220 | -0.206 | -0.281 |
| perc_nha_Hab_rh | 0.379 | -0.179 | 0.880 |
| perc_sp_Re_lith | -0.284 | -0.159 | -0.512 |
| perc_nha_Re_lith | -0.263 | 0.327 | -0.009 |
| perc_sp_Re_phyt | 0.057 | 0.139 | -0.304 |
| perc_nha_Lon_sl | -0.258 | 0.188 | -0.082 |
| kg_ha_Fe_pisc | -0.326 | -0.231 | 0.277 |
| perc_nha_Fe_insev | -0.557 | 0.083 | 0.500 |
| perc_nha_Fe_omni | 0.500 | -0.434 | -0.263 |
| perc_kgha_Mi_long | 0.403 | -0.483 | 0.226 |
| density_sp_sentinel_Bar_bab | 0.163 | -0.025 | -0.129 |
| biom_sp_sentinel_Bar_bab | -0.021 | -0.196 | 0.062 |
| density_sp_sentinel_Leu_cep | 0.238 | 0.158 | -0.256 |
| biom_sp_sentinel_Leu_cep | 0.103 | 0.502 | 0.317 |
| density_sp_sentinel_Leu_leu | 0.201 | -0.103 | -0.191 |
| biom_sp_sentinel_Leu_leu | -0.332 | 0.241 | 0.033 |
| density_sp_sentinel_Rut_rut | -0.044 | 0.338 | -0.544 |
| biom_sp_sentinel_Rut_rut | -0.022 | -0.185 | 0.370 |
| density_sp_sentinel_Sal_far | -0.244 | 0.559 | 0.168 |

**Table 3.26** Discriminant function model, structure matrix. Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions for Type 7.

| | Function | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| perc_kgha_Tol | 0.548 | -0.032 | 0.396 |
| kg_ha_ Fe_pisc | -0.525 | -0.476 | 0.014 |
| perc_sp_Re_lith | -0.433 | 0.070 | -0.330 |
| perc_sp_Hab_rh | -0.425 | 0.049 | -0.311 |
| Biom_sp_all | -0.407 | -0.263 | -0.063 |
| perc_nha_Re_lith | -0.392 | 0.257 | -0.127 |
| perc_nha_Hab_rh | -0.379 | 0.237 | -0.136 |
| perc_nha_Fe_omni | 0.369 | -0.066 | -0.309 |
| perc_kgha_Hab_wc | -0.331 | 0.115 | -0.062 |
| perc_nha_Fe_insev | -0.301 | 0.174 | -0.020 |
| perc_nha_Intol | -0.295 | 0.153 | -0.035 |
| biom_sp_sentinel_Leu_leu | -0.255 | 0.057 | -0.026 |
| biom_sp_sentinel_Leu_cep | -0.228 | 0.152 | -0.039 |
| perc_sp_Re_phyt | 0.145 | 0.021 | -0.127 |
| density_sp_sentinel_Sal_far | -0.187 | 0.501 | 0.076 |
| perc_nha_Lon_sl | -0.375 | 0.388 | -0.040 |
| density_sp_sentinel_Leu_cep | 0.046 | 0.360 | -0.296 |
| N_sp_all | -0.107 | 0.262 | -0.394 |
| perc_kgha_Mi_long | 0.276 | -0.148 | 0.358 |
| density_sp_sentinel_Leu_leu | 0.091 | 0.149 | -0.288 |
| density_sp_sentinel_Rut_rut | 0.119 | 0.154 | -0.285 |
| density_sp_sentinel_Bar_bab | 0.023 | -0.036 | -0.227 |
| biom_sp_sentinel_Bar_bab | -0.042 | -0.082 | -0.140 |
| biom_sp_sentinel_Rut_rut | -0.001 | -0.013 | -0.027 |

a This variable not used in the analysis.

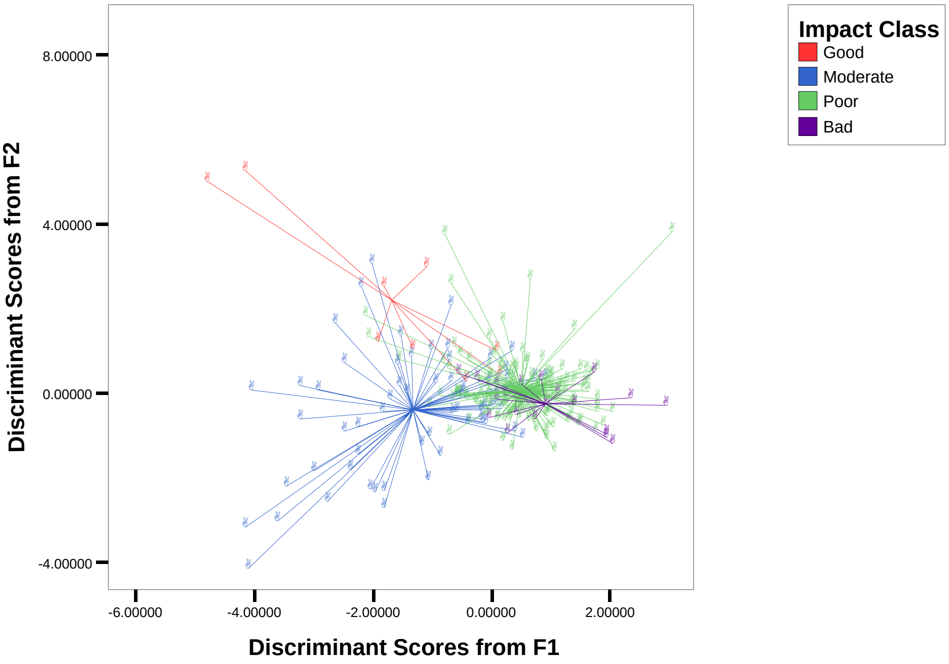

**Figure 3.13** Discriminant plot for functions 1 and 2 for sites used in the discriminant model to determine impact status for Type 7.

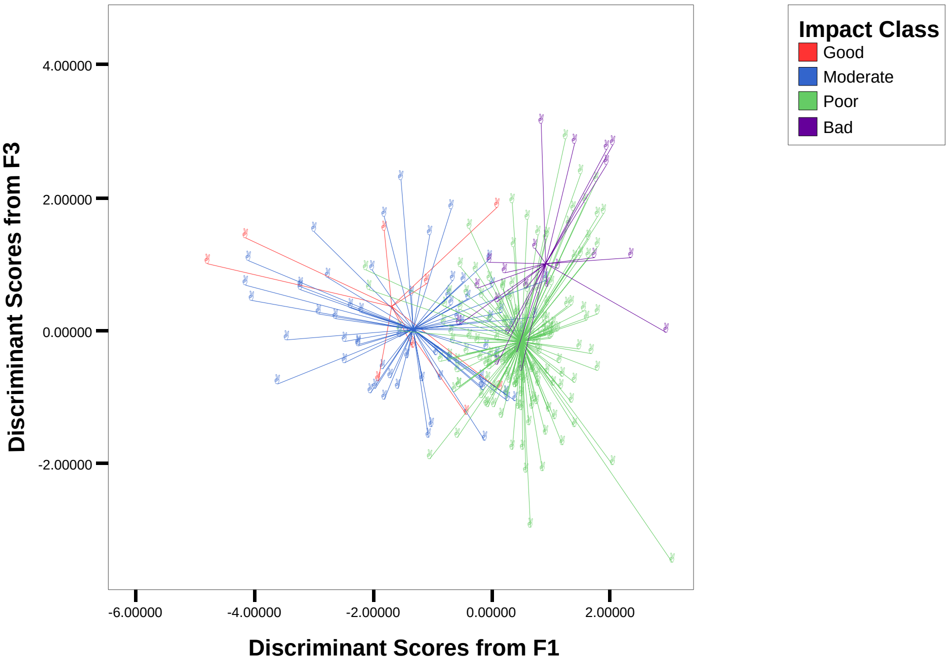

**Figure 3.14** Discriminant plot for functions 1 and 3 for sites used in the discriminant model to determine impact status for Type 7.

The model allowed reasonable prediction of ecological status class with the majority of predictions matching the presumed status (Figure 3.15). Within the model development dataset only few sites were predicted to be in a much poorer state. Within the evaluation dataset prediction rates were similar although the prediction of moderate ecological status was not as powerful (Figure 3.16). This model is probably limited by the relatively low number of sites with presumed good or bad status.

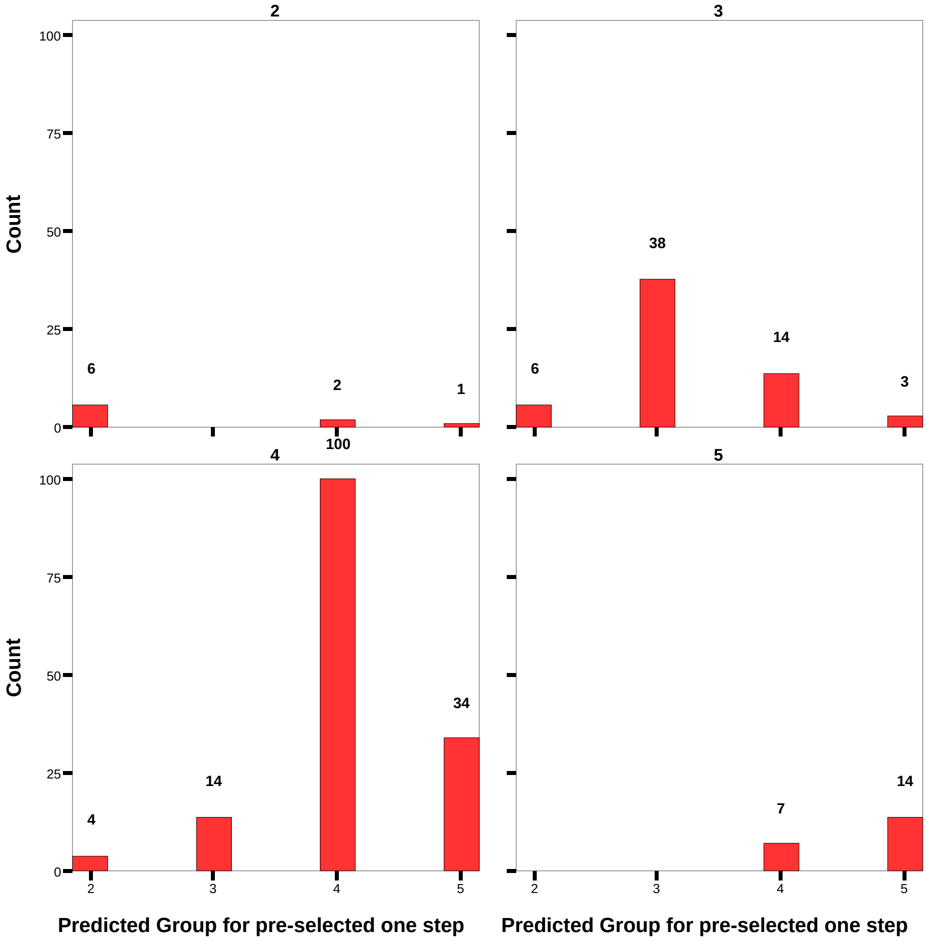

**Figure 3.15** Matching of predicted status to presumed ecological status for sites used in the discriminant analysis for Type 7. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

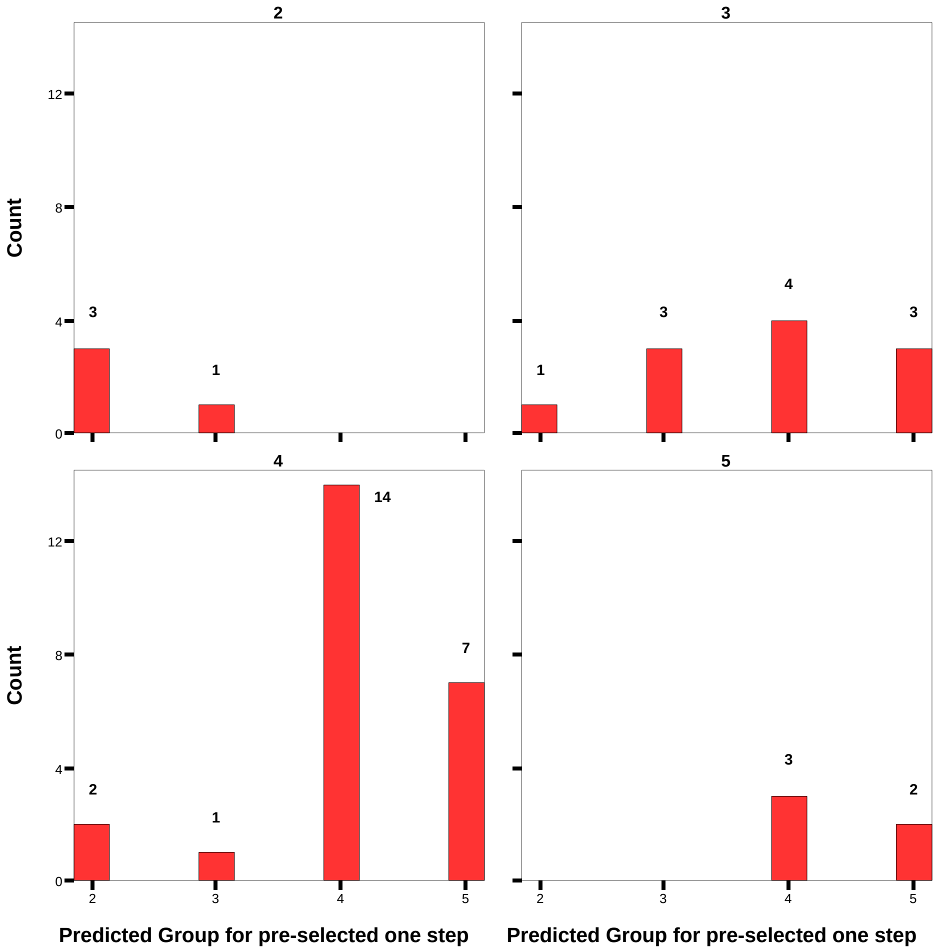

**Figure 3.16** Matching of predicted status to presumed ecological status for 50 randomly selected sites used to evaluate the discriminant analysis for Type 7. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

## 3.9. Type 8 Hampshire Rivers Salmonid Chalk Stream Type

The scarcity of biomass records within this type precluded the assessment of biomass metrics. Additionally the low number of sample precluded the selection of a separate evaluation dataset. Only two presumed class of presumed ecological status could be used; good or better and moderate or worse (Table 3.27).

**Table 3.27** Average group statistics for the pre-selected metrics entered into discriminant analysis for Type 8.

| | | Impact Class inc MC | | |
|---|---|---|---|---|
| | | Good or better | Moderate or worse | Total |
| Community | N_sp_all | 3.44 | 3.77 | 3.63 |
| | Density_sp_all | 1549.04 | 1612.54 | 1586.08 |
| Tolerance | n_ha_Intol | 1296.44 | 754.91 | 980.55 |
| | perc_nha_Tol | 2.96 | 13.29 | 8.98 |
| Habitat | perc_sp_Hab_rh | 85.68 | 76.71 | 80.45 |
| | n_ha_Hab_eury | 86.28 | 457.89 | 303.05 |
| Reproduction | perc_nha_Re_lith | 96.84 | 78.66 | 86.23 |
| | n_ha_Re_phyt | 2.24 | 19.94 | 12.57 |
| Longevity | perc_sp_Lon_ll | 12.80 | 21.14 | 17.67 |
| Feeding | perc_nha_ Fe_pisc | 0.16 | 4.66 | 2.78 |
| | n_ha_Fe_insev | 1296.44 | 754.91 | 980.55 |
| Migration | perc_nha_Mi_potad | 7.64 | 22.80 | 16.48 |
| | perc_sp_Mi_potad | 10.52 | 22.97 | 17.78 |
| Sentinel | density_sp_sentinel_Bar_bab | 0.40 | 0.00 | 0.17 |
| | density_sp_sentinel_Leu_cep | 1.76 | 88.14 | 52.15 |
| | density_sp_sentinel_Leu_leu | 44.24 | 151.43 | 106.77 |
| | density_sp_sentinel_Rut_rut | 4.92 | 425.51 | 250.27 |
| | density_sp_sentinel_Sal_sal | 363.32 | 284.60 | 317.40 |
| | density_sp_sentinel_Sal_far | 755.44 | 332.03 | 508.45 |
| | density_sp_sentinel_Thy_thy | 74.04 | 134.49 | 109.30 |
| | presence_0plus_Thy_thy | 0.00 | 0.03 | 0.02 |
| | density_0plus_Thy_thy | 0.00 | 0.66 | 0.38 |
| | perc_0plus_Thy_thy | 0.00 | 1.69 | 0.98 |

Grey shaded metrics failed the initial tolerance test and were not included in the discriminant analysis.

**Table 3.28** Eigenvalue, variance explanation and canonical correlation of the discriminant functions for Type 8 model.

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 0.723 | 100.0 | 100.0 | 0.648 |

**Table 3.29** Wilks' Lambda and significance of discriminant functions for Type 8 model.

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 | 0.580 | 26.388 | 19 | 0.120 |

**Table 3.30** Standardized canonical discriminant function coefficients for Type 8 model.

|  | Function 1 |
|---|---|
| N_sp_all | 0.327 |
| Density_sp_all | -2.041 |
| n_ha_Intol | -0.625 |
| perc_nha_Tol | -0.246 |
| perc_sp_Hab_rh | -0.249 |
| n_ha_Hab_eury | 6.087 |
| perc_nha_Re_lith | 0.470 |
| n_ha_Re_phyt | -0.557 |
| perc_sp_Lon_ll | 0.268 |
| perc_nha_ Fe_pisc | -0.290 |
| perc_nha_Mi_potad | -0.315 |
| perc_sp_Mi_potad | -0.478 |
| density_sp_sentinel_Bar_bab | 0.345 |
| density_sp_sentinel_Leu_cep | 0.495 |
| density_sp_sentinel_Leu_leu | -0.233 |
| density_sp_sentinel_Rut_rut | -4.649 |
| density_sp_sentinel_Sal_sal | 1.195 |
| density_sp_sentinel_Sal_far | 1.132 |
| presence_0plus_Thy_thy | -0.168 |

The discriminant function developed was related to contribution of trout and other lithophils to the community, together with the contribution of potadromous and tolerant species to the community (Table 3.31).

**Table 3.31** Discriminant function model, structure matrix. Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions for Type 8.

| | Function<br>1 |
|---|---|
| perc_nha_Mi_potad | -0.467 |
| perc_sp_Mi_potad | -0.436 |
| perc_nha_Re_lith | 0.432 |
| density_sp_sentinel_Sal_far | 0.399 |
| perc_nha_Tol | -0.333 |
| n_ha_Re_phyt | -0.329 |
| n_ha_Intol | 0.313 |
| n_ha_Fe_insev[a] | 0.313 |
| perc_nha_Fe_pisc | -0.264 |
| density_sp_sentinel_Leu_cep | -0.246 |
| perc_sp_Hab_rh | 0.237 |
| density_sp_sentinel_Rut_rut | -0.230 |
| perc_sp_Lon_ll | -0.227 |
| density_sp_sentinel_Thy_thy[a] | -0.212 |
| n_ha_Hab_eury | -0.195 |
| density_sp_sentinel_Bar_bab | 0.183 |
| density_sp_sentinel_Leu_leu | -0.159 |
| perc_0plus_Thy_thy[a] | -0.130 |
| presence_0plus_Thy_thy | -0.130 |
| density_0plus_Thy_thy[a] | -0.130 |
| N_sp_all | -0.091 |
| density_sp_sentinel_Sal_sal | 0.061 |
| Density_sp_all | -0.019 |

a This variable not used in the analysis.

The discriminant function enabled reasonable prediction rates within the model development dataset with only a low proportion of sites being miss-classified (Figure 3.17).

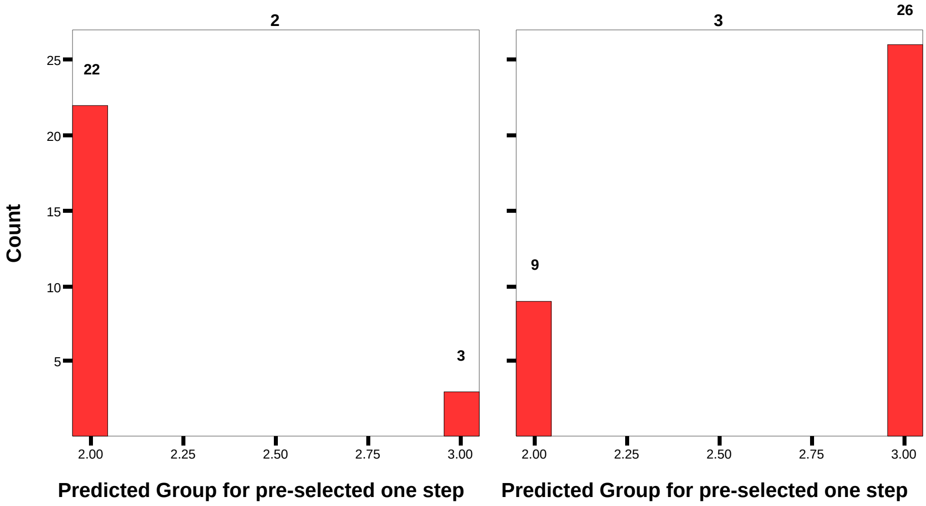

**Figure 3.17** Matching of predicted status to presumed ecological status for sites used in the discriminant analysis for Type 8. The x axis in each plot reflects predicted status (1 = high etc) whilst each plot reflects sites from a particular presumed impact class.

## 4. SYNOPSIS

The analysis of the metrics within the different river types indicated some key considerations for further analysis and the development of a robust IBI for the UK:

- The fish-based and abiotic typology retains some potential errors and difficulty in identifying and assigning sites into distinct typological units. This is probably a function of the diverse nature of rivers in the UK compounded by distribution of the sites collected for FIDES.
- Even within the different river types identified by the fish-based typology considerable noise remains for the metric and species abundance values. The noise is reflected in the moderate correlation coefficients between metrics and impacts and the considerable overlap between metrics values in each of the presumed impact classes.
- Despite the considerable noise within each of the typological units the metrics that, during WP3, were identified to have predictable responses to degradation were generally the ones that exhibited the strongest correlation during the screening process.
- In most types the same group of metrics were identified to exhibit the best correlation with impact.
- Although similar types of metric were found to exhibit the best correlation with impact in different types, the statistic to measure the metric was variable between types. In upland salmonid types, dominated by small sized individuals, the numbers per hectare statistic was generally the best correlated with impact. However, in lowland types (e.g. Types 5 and 6) the percentage statistics of both abundance and biomass were generally the better measures. This is probably due to the difficulties related to sampling large rivers quantitatively.

The considerable noise in the UK dataset causes acute problems for ecological assessment. Few of the metrics within river types exhibited sufficiently strong or distinct trends with impacts severely limiting the number of metrics with low associated errors in impact prediction available for use in an IBI. The low number of metrics with correlation coefficients of >0.4 is severely problematic for traditional IBI development and is compounded by the auto correlation between the best metrics (e.g. intolerant species and insectivorous species are the same suite of species in most river types).

However, the use of a suite of pre-selected metrics in a discriminant analysis did allow the prediction if impact classes within most of the identified fish types. Although the prediction power varied between the fish types the results are promising. In most cases the majority of predictions were valid and mis-classification in the most case were not too severe. The use of discriminant models also allow the use of similar metrics in the same analysis. The use of species, % species, nha$^{-1}$, % nha$^{-1}$ etc. of the same guild metrics enables the analysis to make use of a number of measures which may respond in different ways or over different scales.

It is recognised that there are a number of gaps in the dataset available for analysis in the UK and that the uneven spread of sites between river types restricted the analysis of some types. It is not common for routine monitoring to cover degraded stretches of rivers and hence data for the highly impacted examples are scarce for certain types. Additionally routine monitoring in England and Wales is split into salmonid (Salmon and Trout) and Coarse fish monitoring programmes and is often focussed on spawning and nursery sites to assess recruitment processes. Therefore, it is often uncommon for the main river stem of rivers to be assessed due to sampling difficulties. This results in scarcity of data for certain types e.g. the main river stem barbel zone and the grayling zone. The paucity of data for these types (especially grayling zone types) limited the development of an IBI for these rivers. Further research is need to develop and evaluate properly an IBI approach that meets the requirements of the WFD.

## References

Boon, P.J. (2000) The development of integrated methods for assessing river conservation value. In *Assessing the Integrity of Running Waters*, Jungworth, M., Muhar, S., Schmutz S. (eds). *Hydrobiologia* **422/423**: 413-428.

Boon, P.J. & Howell, D.L. (1997) Defining the quality of fresh waters: theme and variations. In : *Freshwater Quality: Defining the Indefinable?.* Boon, P.J. & Howell, D.L. (eds) The Stationary Office: Edinburgh; 522-533.

Boon, P.J. Holmes, N.T.H., Maitland, P.S., & Rowell, T.A.(1996) *SERCON: System for Evaluating Rivers for Conservation. Version 1 Manual.* Scottish National Heritage Research Survey and Monitoring Report No. 61. Scottish National Heritage: Edinburgh.

Boon, P.J. Holmes, N.T.H., Maitland, P.S., Rowell, T.A.& Davies, J. (1997) A system for evaluating rivers for conservation (‘SERCON’): development, structure and function. In : *Freshwater Quality: Defining the Indefinable?.* Boon, P.J. & Howell, D.L. (eds) The Stationary Office: Edinburgh; 299-326.

Boon, P.J. Holmes, N.T.H., Maitland, P.S., & Fozzard, I.R. (2002) Developing a new version of SERCON (System for Evaluating Rivers for Conservation). *Aquatic Conserv. Mar. Freshw. Ecosyst.* **12**: 439-455.

Hickley, P. & North, E. (1981). An appraisal of anglers' catch composition in the barbel reach of the River Severn. *Proc. 2nd Brit Freshwat. Fish. Conf.,* Univ. Liverpool, pp 94-100.

Maitland, P.S. (2003) *Evaluating the ecological and conservation status of freshwater fish communities in the United Kingdom*. Unpublished Report to Scottish National Heritage.

Regan, C.T. (1974) *The Freshwater Fishes of the British Isles* London: Meuthen. 287pp.

Wheeler, A. (1977) The origin and distribution of the freshwater fishes of the British Isles. *Journal of Biogeography* **4**: 1-24.

Wheeler, A. & Jordan, D.R. (1990). The status of the barbel, *Barbus barbus* (L.) (Teleostei, Cyprinidae), in the United Kingdom. *J. Fish Biol.* **37**, 393-399.

Wheeler, A. & Maitland, P.S. (1973) The scarcer freshwater fishes of the British Isles. I. Introduced species. *Journal of Fish Biology* **5**, 49-68.