

Indicators of ecological change: comparison of the early response of four organism groups to stress gradients

Richard K. Johnson^{1,*}, Daniel Hering², Mike T. Furse³ & Piet F.M. Verdonshot⁴

¹*Department of Environmental Assessment, Swedish University of Agricultural Sciences, P.O. Box 7050 SE-750 07 Uppsala, Sweden*

²*Department of Hydrobiology, University of Duisburg-Essen, D-45117 Essen, Germany*

³*Centre for Ecology and Hydrology, Winfrith Technology Centre, Winfrith Newburgh, DT2 8ZD, Dorchester, Dorset, UK*

⁴*Alterra Green World Research, Freshwater Ecology, Droevendaalsesteeg 3a, NL-36700 AA Wageningen, The Netherlands*

(*Author for correspondence: E-mail: richard.johnson@ma.slu.se)

Received 25 July 2005; accepted 3 October 2005

Key words: early response, streams, bioassessment, monitoring

Abstract

A central goal in monitoring and assessment programs is to detect change early before costly or irreversible damage occurs. To design robust early-warning monitoring programs requires knowledge of indicator response to stress as well as the uncertainty associated with the indicator(s) selected. Using a dataset consisting of four organism groups (fish, macrophytes, benthic diatoms and macroinvertebrates) and catchment, riparian and in-stream physico-chemical variables from 77 mountain and 85 lowland streams we determined the relationships between indicator response and complex environmental gradients. The upper (> 75th percentile) and lower (< 25th percentiles) tails of principal component (PC) gradients were used to study the early response of the four organism groups to stress. An organism/metric was considered as an early warning indicator if the response to the short gradients was more robust (higher R^2 values, steeper slope and lower error) than the null model (organism response to the full PC gradient). For mountain streams, both fish and macrophyte CA scores were shown to exhibit an early warning response to the upper tail of the 1st PC gradient when compared to the null model. Five of the eight metrics showed better response to the upper tail of the 2nd PC gradient compared to the null model, while only one metric (macrophyte CA scores) showed improvement when compared to the lower tail of the 2nd PC gradient. For lowland streams all four organism-groups showed better response (CA scores) to the upper tail of the PC gradient when compared to the null model. Only one metric (fish CA scores) regressed against the lower tail of the 2nd PC gradient was found to be more robust than the PC2 null model. These findings indicate that the nonlinear relationships of organism/metric response to stress can be used to select potentially robust early warning indicators for monitoring and assessment.

Introduction

Humans have altered the landscape of Europe for centuries resulting in a substantial loss of habitats and biodiversity. Aquatic resources in general, and stream habitats in particular, are some of the most threatened on Earth. Recognizing that biodiversity

as well as the functions and services provided by aquatic ecosystems have changed markedly over the years, the European Community recently agreed upon a number of measures to impede degradation and improve quality of inland and coastal waters (European Commission, 2000). One of the innovative aspects of the Water Framework

Directive is the use of multiple indicators (organism groups and metrics) in monitoring and assessment programs. Advocates of the approach argue that the use of multiple indicators increases the probability of detecting change if/when change occurs (a.k.a. multiple lines of evidence). This presumption is based on the premise that the indicators selected are not redundant, but supply complementary information. A second advantage of using multiple indicators to detect ecological impairment is that not only the trajectories, but also the rates of change may differ between indicators selected. Knowledge of how organisms respond to different types of stress can and should be used to design more robust and cost-effective monitoring programs.

Biological response variables are often selected over physical–chemical variables because they represent valued ecosystem attributes such as species richness or ecosystem productivity (e.g., Stevenson et al., 2004). The use of biological variables in European monitoring and assessment programmes has a long history (e.g., Metcalfe, 1989), stemming from the early 1900s when German aquatic ecologists began using the Saprobien Index to assess the effects of organic pollution on streams (Kolkwitz & Marsson, 1902). Although benthic macroinvertebrates are probably the single most common organism group presently used in bioassessment, other groups such as fish and periphyton are being used more frequently. In North America, for example, benthic diatoms, macroinvertebrates and fish are commonly used together to assess the ecological quality of streams (Barbour et al., 1998).

Ideally, the selection of what or which organism group(s) to use in bioassessment should not be arbitrary, but should be based on conceptual models and empirical (e.g., dose–response) relationships that characterize the response of the indicator to the stressor of interest as well as quantify the levels of uncertainty associated with the stressor–response relationship. Organism response to stress varies with a number of abiotic and biotic factors, such as an organism's life history stage. Because responses to environmental stress originate at the biochemical and physiological levels of the organism, changes at the sub-organism-level may provide the earliest warning of possible adverse effects (Johnson et al., 1993).

Unfortunately, knowledge of the normal background variability often limits the use of biochemical and physiological indicators in biomonitoring. The idea of using an indicator that provides an early indication of change has, however, many benefits, not the least being the socio-economic aspects of failing to detect an ecological change early. For example, considerable costs may be incurred if human-induced damage is allowed to proceed undetected.

Organism response to human-induced stress is not always linear, and selection of indicators that respond rapidly at the outset of impairment is one way of determining and quantifying early-on the effects that humans may have on ecosystem integrity. Our working hypothesis is that organisms that show a nonlinear response to the stress will show more rapid response (higher slopes) with moderate than high levels of stress. At high levels of stress, dose–response relationships often show a leveling off (e.g. low variance around the regression line) resulting in typical funnel-shaped response curves. Here we use stress–response relations of four organism groups to determine if the organisms differ in their response to stress. In doing so, we hope to better our understanding of the use of early warning indicators for detecting ecological change if/when it occurs.

Methods

Study sites

Some 162 streams were sampled in 2003 and 2004 as part of the European funded STAR project (Hering & Strackbein, 2002; Furse et al., 2004). Two common stream types (mountain, $n = 77$ and lowland, $n = 85$) are used here to study the response of different organism groups to human-induced stress (Fig. 1). To ensure adequate sampling of stressor gradients, prior to sampling, all sites were pre-classified into five classes of ecological status using physico-chemical and in some instances biological information and/or expert opinion: (i) high (no or only minimal disturbance), (ii) good (slight deviation from high status), (iii) moderate (moderate deviation from high status and significantly more disturbed than good), (iv) poor (major alteration from high status) and

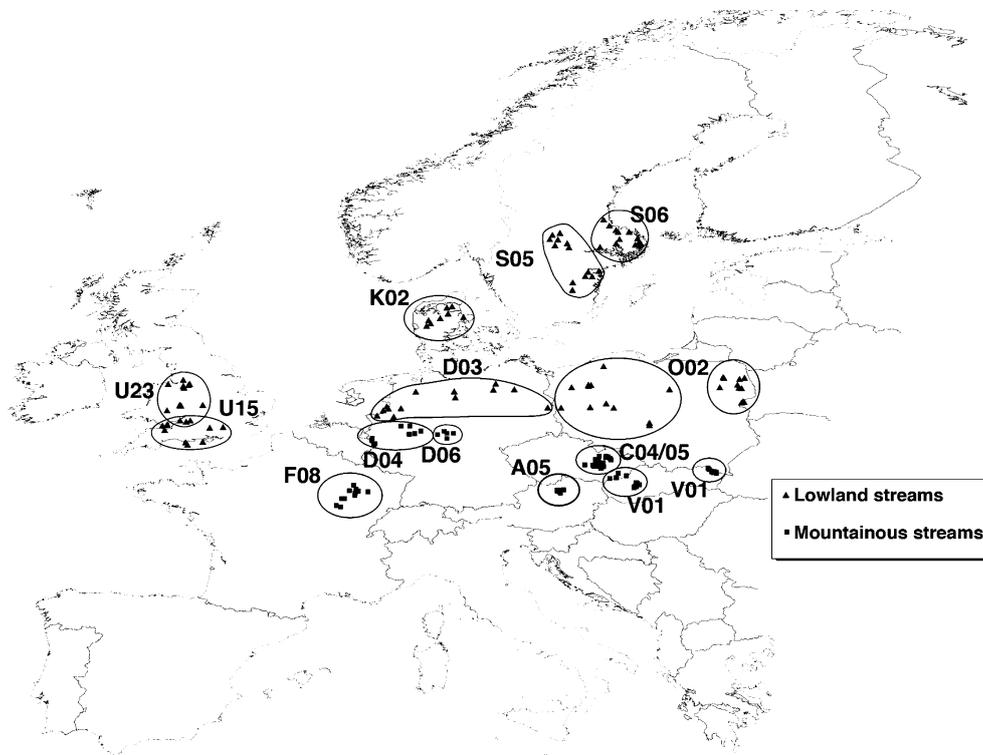


Figure 1. Location of streams and types sampled as part of the EU-STAR project.

(v) bad (severe alteration from high status). For each stream type, 10 and usually 12 sites were sampled of which, in each set, at least three were pre-classified as of high status and the rest spread more or less evenly over the other four quality status classes. In addition, a large number of physical, chemical and geographical variables were sampled or obtained (e.g., using GIS) for each of the streams sampled (Furse et al., 2004). The substratum of each of the sampling sites was classified (in percent) according to seven inorganic size classes (i.e., from silt/clay $< 6 \mu\text{m}$ to large cobbles, boulders and blocks $> 40 \text{ cm}$) and 10 organic classes/fractions such as the amount of algae (macro & micro), vegetation (aquatic submerged and emergent and living parts of terrestrial plants) and detritus (e.g., woody debris). A number of physical-chemical metrics representative of nutrient status (nitrogen and phosphorus fractions) and acidity (pH) status, as well as oxygen conditions (BOD₅), were measured for each site. Catchment and riparian land use/type was classified according to 16 classes (e.g., forest type,

cropland, pasture, urbanization). A number of measures of stream hydrology and morphology were recorded such as mean annual discharge valley and channel form (seven classes) and stream width and depth using the RHS survey technique (Raven et al., 1998). Also included in hydromorphological classification were measures of bank and bed fixation and the number of debris dams and woody debris in and along the stream channel.

Biological samples

Four organism groups were sampled at each stream site; namely, fish, macroinvertebrates, macrophytes and periphytic diatoms. A brief description of the sampling method used is given here, for more detailed information refer to the STAR website (www.eu-star.at).

Fish were normally sampled by electric fishing in accordance with the procedures set out in CEN prestandard PrEN 14011. Fishing was undertaken along two runs of a stop-netted area on a single occasion in late summer or early autumn. The

recommended sampling length was 10× the stream width, with a minimum of a 100 m stream length sampled. The fish variables recorded were number of species, life history stage (young of the year per species), density (number of fish per m²) and assessment of the degree of infestation of external parasites or other diseases.

Benthic macroinvertebrates were sampled in spring and either summer or autumn using a Surber sampler or by standardized kick-sampling with a handnet (area 625 cm², mesh size 500 μm). Generally the sampling section consisted of 20–50 m in length in small (1–100 km²) and 50–100 m in length in medium (100–1000 km²) sized streams. Each sampling site encompassed the whole width of the stream and was deemed to be representative of a minimum area surveyed (i.e., 500 m of length or 100× average width). Before sampling, the sampling site was first classified according to the coverage of all microhabitats with at least 5% cover. A multi-habitat sampling strategy was then adopted that reflected the proportion of different habitat types present at each stream site. Each complete sample consisted of 20 sample units of dimensions 25 cm × 25 cm. These sampling units were proportionally situated in all microhabitats with >5% coverage. The 20 sample units resulted in ca. 1.25 m² of stream bottom being sampled. Each composite sample was preserved with formalin (4% final concentration) or 95% ethanol to a final concentration of ca. 70%. Macroinvertebrates were sorted (subsampling with the target of 700 individuals) and identified (usually to genus/species).

Macrophytes were sampled using a single survey in late summer or early autumn. Macrophytes included higher aquatic plants, vascular cryptogams, bryophytes as well as groups of algae. A 100 m stream length was surveyed in each stream by wading, walking along the bank or by boat according to the MTR method described by Holmes et al. (1999). All macrophyte species were recorded as well as the percent cover of the overall macrophyte growth. Submerged vegetation was observed using a glass-bottom bucket. If identification was uncertain, representative samples were collected and identified later.

Periphytic diatoms were sampled from hard (usually cobbles or pebbles) or soft (sand/silt)

substratum or macrophytes. Wherever possible periphyton samples were collected within the same sampling area as benthic macroinvertebrates. In brief, a minimum of five cobbles were arbitrarily selected at each site (the combined exposed surface area comprised ca. 100 cm²). The stones were individually placed in a plastic tray and 100–200 ml of distilled or filtered water added to the tray. The upper part of the stone substratum was washed using a toothbrush, and the dislodged material was decanted into a sample bottle and a composite sample was preserved (using formalin or Lugol's iodine solution) if the sample could not be processed within 24 h. Submerged macrophytes and parts of emergent ones were collected, placed in a wide-mouth 1-l container, ca. 100–200 ml of distilled/filtered water added and the container was shaken vigorously for about 60 s. A 250 ml aliquot of the sample was decanted to a sample bottle and preserved as above if not analyzed within 14 h. Mineral sediments were sampled using a glass tube submerged in the sediment and extracting sediment and interstitial water. Replicate samples were collected until volume of ca. 200 ml was obtained. Light microscopy was used to identify the living and dead diatom cells. The diatom species were counted (a minimum of 300 diatom valves) and identified to species at 400× and 1000× magnification.

Analyses

Environmental gradients and organism response

Principal components analysis (PCA) was used to construct complex stress gradients by reducing the dimensionality of the physical–chemical, hydro-morphological and land use/type characteristics for each of the sites. Most environmental variables were log₁₀ or arcsine square-root transformed before the analyses to approximate normal distributions. To test the early response of the different organism groups to stress, the upper and lower tails of the PC gradients were used as ‘short’ environmental gradients. The short environmental gradients were constructed by using the two tails of the first two PC axes for the mountain and lowland streams (Fig. 2). The 75th-percentile was arbitrarily selected as the cutoff for ‘best available’ sites and the 25th-percentile as the cutoff for ‘perturbed’ sites, resulting in four environmental

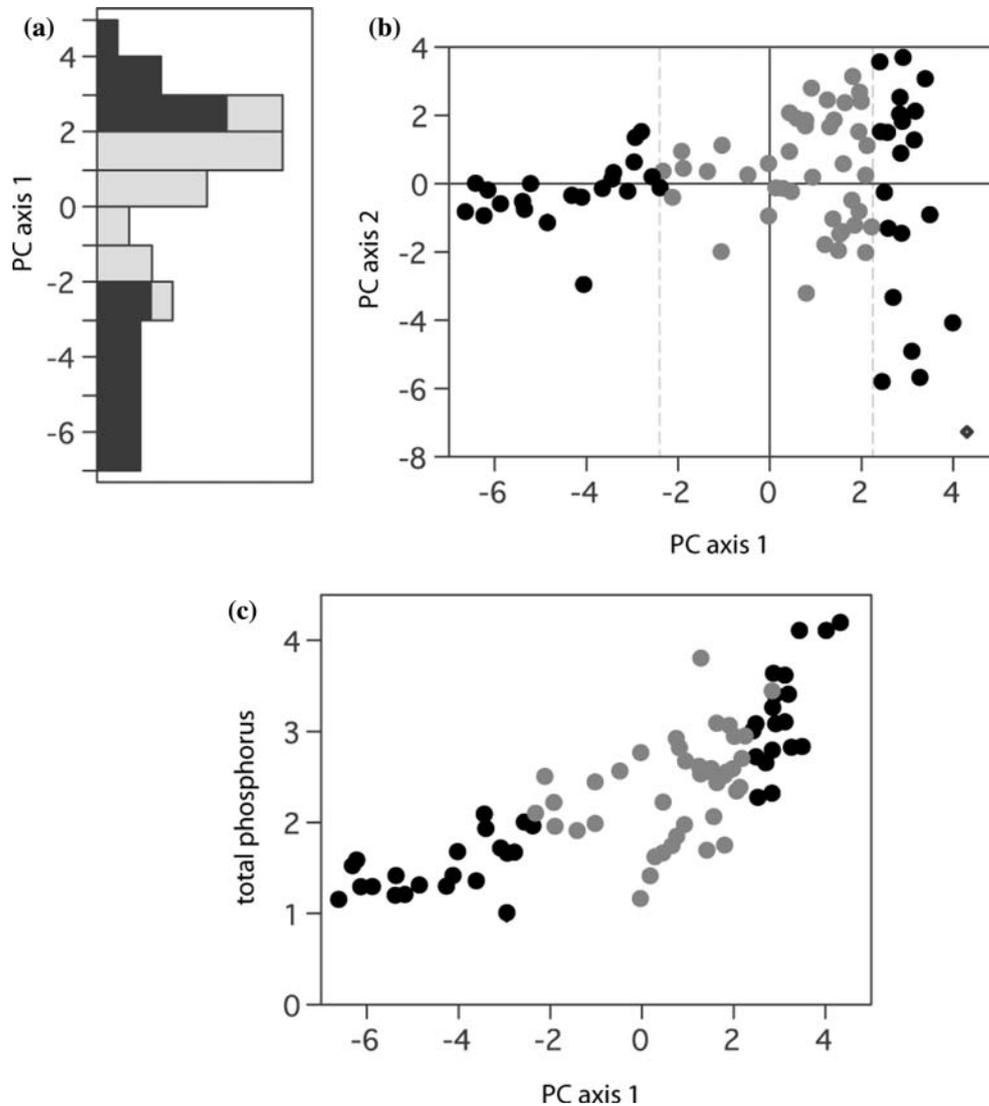


Figure 2. Example of the selection of upper and lower tail sites of the 1st and 2nd principal component gradients (PC axis 1 and PC axis 2, respectively). (a) Distribution of PC axis 1 scores, (b) PC axis 1 plotted against PC axis 2 showing the two tails of the distribution that were used in short-gradient regression analyses, (c) relationship between stream log total phosphate concentration and PC axis 1.

datasets or gradients. Sites with PC-axes scores between the 25th- and 75th-percentiles were omitted from the analyses.

Two biological metrics (correspondence scores and Hill's N2-diversity) were used to compare the response of the different organism groups to stress. To obtain correspondence scores, fish, macroinvertebrate, macrophyte and diatom abundances were ordinated separately for the two stream types using correspondence analysis (ter Braak, 1988, 1990). Correspondence analysis was run on

square-root transformed species abundance, with the downweighting of rare taxa option invoked. The ordination scores on the first CA axis (CA1) and Hill's N2-diversity (Hill, 1973) were used as dependent variables and related to environmental stress gradients.

Linear regression was used to determine the relationship between the two metrics for the four organism groups and their response to the four short gradients (Fig. 3). For the null model we used the PC1 and PC2 gradients for all mountain

($n = 77$) and lowland ($n = 85$) streams. Regression results of the response of the four organism groups to the two tails of the PC gradients were then compared to the null model. Coefficients of variation (adjusted R^2), slope, error (RMSEP) and

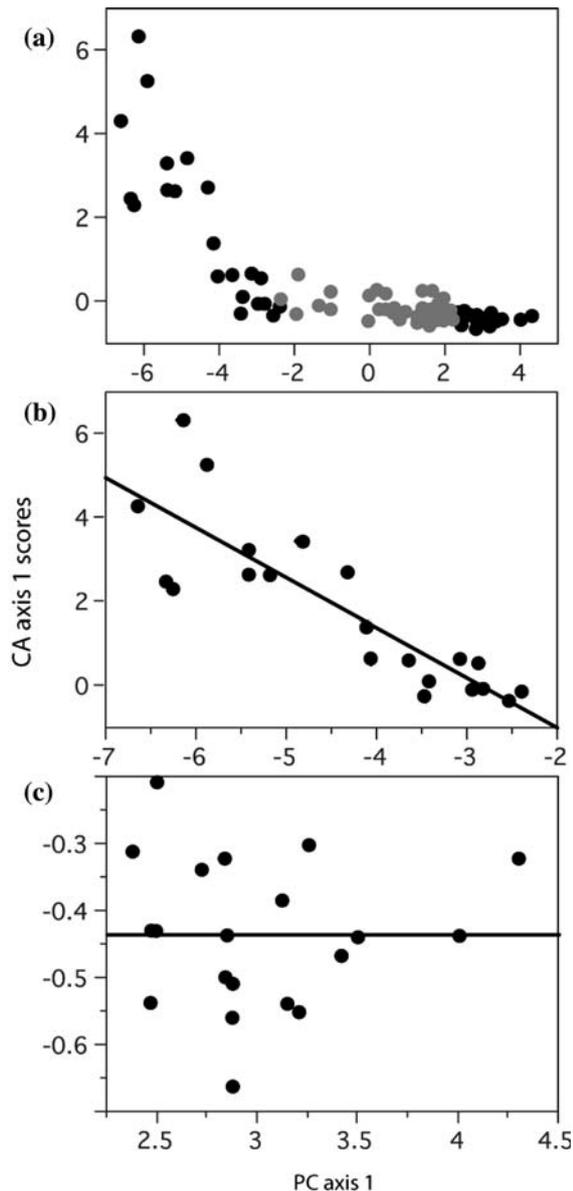


Figure 3. Example of benthic diatom response (CA axis 1 scores) to the 1st principal component gradient (PC axis 1, representing nutrient enrichment) for lowland streams. (a) Response using all stream sites, (b) response using upper tail (> 75th percentile of PC axis 1 gradient) sites, (c) response using lower tail (< 25th percentile of PC axis 1 gradient) sites.

p -values were used to compare the response of the four organism groups to the PC gradients. Coefficients of variation were used as a measure of the precision, slope provided an estimate of the magnitude of change and the root mean square error of the prediction was an estimate of the standard deviation of the random error associated with the response model.

All tests were performed using the statistical program JMP (version 3.1) (SAS, 1994).

Results

The two stream types studied here, small, shallow mountain streams and medium-sized, lowland streams, differed regarding a number of physico-chemical variables. Mountain streams were generally situated at higher altitude (mean 337 m a.s.l.) and had smaller catchments (mean = 57 km²) compared to lowland streams (mean 57 m a.s.l. and 199 km²) (Table 1). Moreover, mountain streams were often situated in forested catchments (e.g., mean = 58% forest), whilst lowland streams had more of their catchments classified as cropland (mean = 30%) or pasture (mean = 15%). The two stream types also differed regarding the predominant substratum type; cobbles and coarse gravel were most common types of substratum (38 and 23%, respectively) in mountain streams, whereas lowland streams had a high frequency of soft-bottom substratum (sand = 36%). Clear differences were also noted regarding nutrient concentrations. For example, total phosphorus concentrations were on average > 5× higher in lowland (mean TP = 1091 mg/l) compared to mountain (mean TP = 193 mg/l) streams.

Environmental gradients

The first two axes of principal component analysis explained ca. 30% of the variation in catchment land use/cover and physico-chemical variables in mountain and lowland streams (Table 2). The primary environmental gradient for mountain (explained 17.3% of the variance) and lowland (20.4%) streams was related to catchment land use and nutrient concentration. For example, total phosphate was positively correlated (eigenvector

Table 1. Selected physico-chemical and catchment characteristics of mountain and lowland streams

	Mountain (n = 77)	Lowland (n = 85)
Altitude (m a.s.l.)	337 ± 84	57 ± 60
Catchment area (km ²)	57 ± 77	199 ± 199
<i>Catchment classification (%)</i>		
Urban (sum)	0.14 ± 0.24	5.9 ± 13
Forest (sum)	58 ± 23	36 ± 28
Native deciduous	32 ± 33	3.13 ± 5.82
Native coniferous	11 ± 17	3.9 ± 10
Cropland	24 ± 26	30 ± 28
Pasture	8.4 ± 16	15 ± 22
<i>Substratum (%)</i>		
Large cobbles, boulders (> 40 cm)	6.9 ± 18	5.8 ± 13
Coarse blocks, cobbles (> 20–40 cm)	12.7 ± 16	7.9 ± 13
Cobbles (> 6–20 cm)	38 ± 23	12 ± 20
Coarse gravel (> 2–6 cm)	23 ± 18	13 ± 21
Fine gravel (> 0.2–2 cm)	9.3 ± 12	13 ± 23
Sand (> 6 µm–2 mm)	7.5 ± 15	36 ± 37
Silt (< 6 m)	1.6 ± 5	9 ± 25
<i>Physico-chemistry</i>		
pH	7.9 ± 0.57	7.55 ± 0.40
Conductivity (µS/cm)	315 ± 261	390 ± 236
BOD5 (mg O ₂ /l)	2.25 ± 1.58	2.58 ± 1.50
Ammonium (mg NH ₄ /l)	0.166 ± 0.360	321 ± 1918
Nitrate (mg NO ₃ /l)	9.45 ± 9.77	13 ± 14
Total phosphate (µg TP/l)	193 ± 270	1091 ± 2747

Mean ± 1 standard deviation.

loadings = 0.24 and 0.26 for mountain and lowland streams, respectively) and % total forest in the catchment (−0.20 and −0.24, respectively) was negatively correlated with the 1st PC axis. The 2nd PC axis explained another 11.8% (mountain) or 10.3% (lowland) of the variance, and was seemingly related to habitat quality (e.g., number of debris dams) and/or hydromorphological alteration.

Dividing the PC gradients into shorter environmental gradients resulted in clear differences in mean values and ranges of a number of environmental variables (Table 3). For the 1st PC gradient, the most marked among-group differences were related to differences in land use/cover and nutrients. Mountain streams in the best available PC1 group (upper tail) had on averaged 14% of their catchments classified as pasture (range =

0–80%) compared to 2.5% (range = 0–20%) for streams in the perturbed PC1 group (lower tail). Nutrient concentrations also varied between the two groups. For mountain streams total phosphate averaged 65 µg/l (range = 50–182 µg/l) for sites in the PC1 best available group compared to 431 µg/l (range = 30–1270 µg/l) for sites in the PC1 perturbed group. In contrast to mountain streams, lowland streams exhibited stronger gradients in percent catchment land use classified as cropland. Streams in the best available group had on average 7.6% (range 0–40%) of their catchments classified as cropland compared to 39% (range = 0–80%) for streams in the perturbed group. Total phosphate averaged 41 µg/l (range = 8.6–127 µg/l) for streams in the best available group compared to 3315 µg/l (range = 186–15430 µg/l) for streams in the perturbed group.

The 2nd PC gradient was interpreted as being either related to habitat quality or alterations in hydromorphology or both (concomitant changes in habitat/hydromorphology). For both mountain and lowland streams the percentage of substratum classified as coarse blocks and cobbles or coarse gravel changed markedly between the upper (best available) and lower (perturbed) tails of the PC gradient. For mountain streams, coarse blocks and cobbles substratum averaged 14% cover (range = 0–60%) for streams in the best available group compared to 2.4% (range = 0–10%) for sites in the perturbed group. For lowland streams, coarse gravel substratum in the best available group averaged 26% cover (range = 0–80%) compared to 5% (range = 0–50%) for sites in the perturbed group.

Organism/metric response to stress

Three of the four organism groups showed a significant response to the 1st PC (null model) gradient for mountain streams; the exception being diatoms, which did not show a significant response to this stressor gradient (Table 4). Coefficients of variation for the various null model regressions varied markedly among the organism groups. Macroinvertebrates showed the strongest response, with R^2 values for CA scores and N2-diversity of 0.422 and 0.259, respectively, followed by macrophytes (0.306 and 0.247) and fish (0.118 and 0.142). Comparison of organism-group response along the short gradients to the null model showed only two

Table 2. Eigenvectors (loadings) of physico-chemical, substratum and catchment land use/cover

	Mountain streams		Lowland streams	
	PC1	PC2	PC1	PC2
Eigenvalue	6.9	4.7	8.8	4.4
Percent	17.3	11.8	20.4	10.3
Cum percent	17.3	29.1	20.4	30.7
<i>Eigenvectors</i>				
Total forest	-0.20	0.02	-0.24	-0.03
Total urban	-0.10	0.00	0.15	-0.13
Wetland (mire)			-0.19	-0.02
Open grass/bushland	-0.11	0.06	0.04	0.24
Standing water			-0.20	-0.07
Cropland	0.32	0.01	0.19	-0.13
Pasture	-0.11	-0.06	0.17	0.23
Clear-cutting			-0.26	-0.01
Shading at zenith (foliage cover)	-0.19	0.20	0.04	0.18
Average width of woody riparian vegetation (m)	-0.21	0.17	-0.05	0.06
no. of debris dams	0.02	0.29	0.11	0.22
no. of logs	-0.08	0.24	0.13	0.16
Shoreline covered with woody riparian vegetation left	-0.22	0.27	0.07	0.15
No bank fixation	-0.16	0.03	-0.02	0.26
No bed fixation	-0.08	0.08	-0.05	0.30
Stagnation	0.11	0.20	0.09	-0.13
Straightening	0.20	0.10	0.12	-0.25
Hygropetric	0.08	-0.06	-0.16	-0.02
Large cobble	0.01	-0.08	-0.26	-0.03
Coarse blocks	-0.17	-0.15	-0.26	-0.06
Cobbles	-0.25	0.07	-0.21	0.03
Coarse gravel	0.11	0.05	-0.03	0.13
Fine gravel	0.18	0.16	0.09	-0.16
Sand	0.15	0.03	0.20	0.00
Silt	0.19	0.18	0.01	0.09
Submerged macrophytes	0.09	-0.27	-0.05	-0.16
Emergent macrophytes	0.10	-0.08	0.07	-0.19
Xylal	0.02	0.24	0.12	0.22
CPOM	-0.03	0.13	0.13	-0.29
FPOM	0.13	0.13	0.14	-0.27
pH	0.17	0.20	0.19	0.07
Conductivity	0.28	-0.02	0.29	0.02
BOD5	0.12	0.23	-0.03	-0.05
Ammonium	-0.10	0.31	0.12	-0.31
Nitrite	-0.17	0.31		
Nitrate	0.24	0.10	0.18	0.22
<i>ortho</i> -Phosphate	0.25	0.11	0.28	0.01
Total phosphate	0.24	0.21	0.26	-0.02

Only variables with loadings >0.15 on either PC1 or PC2 are shown, and loadings >0.15 are shown in bold.

Table 3. Selected physico-chemical variables and catchment characteristics of the PC gradient-ends for mountain and lowland streams

	Upper tail PC1	Lower tail PC1	Upper tail PC2	Lower tail PC2
<i>Mountain streams</i>				
<i>n</i>	19	20	19	19
Altitude (m a.s.l.)	388 (250–534)	309 (174–485)	295 (160–430)	346 (220–485)
Catchment area (km ²)	25 (10–95)	117 (16–450)	138 (23–450)	30 (9.3–63)
Native deciduous forest (%)	46 (0–100)	16 (0–50)	34 (0–80)	18 (0–80)
Native coniferous forest (%)	7.4 (0–70)	12 (0–40)	3.2 (0–40)	22 (0–60)
Cropland (%)	1.6 (0–10)	56 (10–90)	39 (0–90)	36 (0–80)
Pasture (%)	14 (0–80)	2.5 (0–20)	7.9 (0–50)	2.1 (0–20)
Coarse blocks (%)	28 (0–55)	7.5 (0–60)	14 (0–60)	2.4 (0–10)
Cobbles (%)	52 (25–95)	15 (0–45)	30 (0–95)	36 (5–60)
Coarse gravel (%)	10 (0–20)	29 (0–80)	25 (0–80)	34 (5–60)
Conductivity (μS/cm)	156 (69–272)	592 (118–1662)	504 (90–1662)	367 (134–710)
Nitrate (mg/l)	2.9 (0.63–11)	18 (4.4–45)	9.7 (0.74–23)	15 (1.5–38)
Total phosphate (μg/l)	65 (50–182)	431 (30–1270)	175 (20–910)	347 (91–1270)
<i>Lowland streams</i>				
<i>n</i>	21	20	21	21
Altitude (m a.s.l.)	88 (2–261)	50 (7.5–180)	47 (0–120)	77 (4–239)
Catchment area (km ²)	236 (45–1139)	147 (8.8–459)	103 (8.8–413)	301 (57–883)
Native deciduous forest (%)	0 (0–0)	6 (0–20)	9.5 (0–30)	1.0 (0–10)
Native coniferous forest (%)	0 (0–0)	0.5 (0–10)	4.8 (0–60)	6.7 (0–30)
Cropland (%)	7.6 (0–40)	39 (0–80)	22 (0–70)	41 (0–80)
Pasture (%)	0 (0–0)	32 (0–80)	38 (0–70)	5.2 (0–40)
Coarse blocks (%)	21 (0–40)	0.3 (0–5)	0 (0–0)	5 (0–50)
Cobbles (%)	23 (5–65)	0 (0–0)	8.1 (0–50)	6.2 (0–50)
Coarse gravel (%)	11 (0–30)	8.3 (0–80)	26 (0–80)	5 (0–50)
Conductivity (μS/cm)	143 (24–375)	652 (122–1022)	484 (205–879)	455 (26–1022)
Nitrate (mg/l)	2.1 (0.04–23)	19.6 (0.2–45)	22 (6.6–45)	6.6 (0.04–41)
Total phosphate (μg/l)	41 (8.6–127)	3315 (186–15430)	1784 (45–13201)	1841 (19–15430)

Upper tail (best available) = sites above the 75th-percentile; lower tail (perturbed) = sites below the 25th-percentile. Mean values and in parenthesis min and max values.

groups that responded significantly to the upper tail of the PC1 gradient (best available of the PC gradient), and none of the groups showed a significant response to the lower (perturbed) tail of the PC1 gradient. Both fish and macrophyte CA scores indicated an early response. Coefficients of variation for fish increased from 0.118 to 0.306, and the slope changed from -0.1376 to -1.188 when CA scores were regressed against the upper tail of the PC1 gradient. For macrophyte CA scores, the R^2 increased only marginally (from 0.306 to 0.359), but the slope changed from -0.3904 to -4.583 .

All four organism-groups showed a significant response to the 2nd PC (null model) gradient. The

strongest relationship was found for macroinvertebrate CA scores ($R^2 = 0.475$, $p < 0.0001$) and macrophyte ($R^2 = 0.435$, $p < 0.0001$) and fish ($R^2 = 0.311$, $p < 0.001$) diversity. Fish CA scores and macroinvertebrate and diatom diversity were also significantly related to the 2nd PC gradient, albeit weakly (R^2 value < 0.16). Comparison of organism-group response using the upper tail of the PC2 gradient with the null model showed that R^2 values and/or regression slopes of all organism groups (and four of the six regressions) increased, indicating a significant early warning response. Neither macrophyte nor diatom CA scores were significantly related to the 2nd PC (null model)

Table 4. Summary statistics for regression of organism group CA scores and N2-diversity and PC gradients for mountain streams

	Fish		Macrophytes		Macroinvertebrates		Diatoms	
	CA axis 1 scores	Diversity	CA axis 1 scores	Diversity	CA axis 1 scores	Diversity	CA axis 1 scores	Diversity
<i>PC1 gradient (null model)</i>								
<i>n</i>	72	72	58	58	76	76	76	76
<i>R</i> ²	0.118	0.143	0.306	0.247	0.259	0.422	-0.013	-0.012
RMSEP	0.937	1.236	1.577	7.674	0.921	8.550	0.931	5.326
Slope	-0.138	0.201	-0.390	1.650	0.211	-2.802	-0.009	0.077
<i>p</i> value	0.0018	0.0006	0.0001	0.0001	0.0001	0.0001	0.8333	0.7416
<i>PC1 upper tail (values > 75th percentile)</i>								
<i>n</i>	19	19	12	12	19	19	19	19
<i>R</i> ²	0.306	-0.059	0.359	0.031	0.107	-0.057	-0.056	-0.004
RMSEP	0.730	0.718	2.425	2.799	0.190	8.530	1.163	3.385
Slope	-1.188	-0.012	-4.583	2.495	0.184	-0.887	-0.145	-1.793
<i>p</i> value	0.0082	0.9529	0.0234	0.2358	0.0933	0.8507	0.8211	0.3439
<i>PC1 lower tail (values < 25th percentile)</i>								
<i>n</i>	19	19	19	19	20	20	20	20
<i>R</i> ²	0.099	0.124	0.146	-0.052	-0.024	0.048	-0.039	-0.049
RMSEP	0.952	1.887	0.237	10.482	1.530	9.433	0.565	6.881
Slope	-0.288	0.623	-0.079	0.562	0.188	-2.166	-0.049	0.380
<i>p</i> value	0.102	0.0766	0.0595	0.7497	0.4628	0.1781	0.6019	0.7399
<i>PC2 gradient (null model)</i>								
<i>n</i>	72	72	58	58	76	76	76	76
<i>R</i> ²	0.142	0.311	-0.015	0.435	0.475	0.053	-0.012	0.163
RMSEP	0.923	1.108	1.907	6.644	0.775	10.942	0.930	4.845
Slope	0.186	-0.358	-0.038	-2.449	-0.341	1.324	-0.015	-1.015
<i>p</i> value	0.0006	0.0001	0.7161	0.0001	0.0001	0.0254	0.7559	0.0002
<i>PC2 upper tail (values > 75th percentile)</i>								
<i>n</i>	19	19	18	18	19	19	19	19
<i>R</i> ²	0.206	0.057	0.248	0.500	0.352	0.146	0.220	0.029
RMSEP	0.705	1.587	0.192	7.159	1.021	10.211	1.005	6.931
Slope	0.370	-0.504	0.112	-6.898	-0.738	4.535	0.545	1.888
<i>p</i> value	0.0291	0.1671	0.0205	0.0006	0.0044	0.0596	0.0247	0.2326
<i>PC2 lower tail (values < 25th percentile)</i>								
<i>n</i>	16	16	17	17	19	19	19	19
<i>R</i> ²	-0.068	-0.032	0.485	-0.065	0.050	0.111	-0.040	-0.048
RMSEP	0.963	1.001	0.259	1.773	0.113	10.022	0.466	3.222
Slope	0.053	-0.199	-0.224	0.052	-0.033	-3.816	0.055	-0.285
<i>p</i> value	0.842	0.4765	0.0011	0.8935	0.1802	0.0898	0.5858	0.6818

Values shown in bold text are significant ($p < 0.05$).

gradient, whereas relatively strong relationships (R^2 values of 0.248 and 0.220, respectively) were noted when these metrics were regressed using the upper tail of the PC2 gradient (slopes increased from -0.0383 to 0.1119 for macrophyte CA scores and from -0.0154 to 0.5447 for diatom CA

scores). Only one metric, macrophyte CA scores, showed a significant relationship using the lower tail of the PC2 gradient ($R^2 = 0.485$, $p = 0.0011$).

Five of the eight metrics showed a significant response to the 1st PC (null model) gradient for lowland streams (Table 5). The strongest

relationship was found between diatom CA scores and the stress gradient ($R^2 = 0.606$, $p < 0.0001$), followed by macrophyte CA scores ($R^2 = 0.366$, $p < 0.0001$). Although the slopes of the other three regressions were significant, the relations were relatively weak (R^2 values < 0.181). Comparison of organism groups/metric response of the upper tail of the PC1 gradient (best available sites) and the null model showed that four of the eight relationships were significant. The response of two organism groups, in particular, improved suggesting that these organism groups/metrics might be considered as early warning indicators of stress: R^2 values for fish and macroinvertebrate CA scores increased from 0.049 to 0.327 and from 0.148 to 0.565 and the slopes changed from -0.089 to 0.522 and from 0.116 to 0.170, respectively. Diatom CA scores showed only a modest increased response (R^2 value increased from 0.606 to 0.724), and the R^2 value for macrophyte CA scores was actually lower (0.366–0.290) compared to the null model. However, the slopes of both relationships increased markedly from -0.355 to -1.188 for diatoms and from 0.1401 to -1.115 for macrophytes. None of the metrics showed significant relationships using the perturbed sites.

Three of the four organism groups showed a significant response to the 2nd PC (null model) gradient, however R^2 values were low (< 0.177). Neither of the two diatom metrics showed a significant response to this stressor gradient. Comparison of organism/metric response using the short gradients with the null model revealed no significant relationships using the best available sites. The relationship between fish CA scores using the lower tail of the PC2 gradient (perturbed sites) was slightly better than the null model; R^2 values increased from 0.177 to 0.273 and slopes changed from -0.2272 to -0.3041 .

Discussion

Assessing the ecological integrity of running water ecosystems, and being confident that if change occurs it will be detected, is a fundamental objective of most monitoring programmes as well as the underpinning aim of the recently adopted European Water Framework Directive (European Commission, 2000). The major stressors affecting

the integrity of European surface waters are over-exploitation, nutrient enrichment and organic pollution, acidification and alterations of hydrology and morphology (Stanner & Bordeau, 1995). Our results support this view; the two main stress gradients were interpreted as being related to land use and nutrient concentrations (the primary gradient) and alterations in habitat quality and hydromorphology (the secondary gradient). Streams, in particular, are affected (simultaneously) by a multitude of human-generated pressures. For example, agricultural land use can result in several different types of stress, which may singly or in concert affect the structure and function of stream assemblages. For instance, increased runoff due to agricultural activity can result in changes in hydrology, increased siltation and changes in habitat quality/quantity, while inputs of nutrients from agriculture can result in eutrophication effects. The single and combined effects of stress on the organism assemblages inhabiting the ecosystem will vary, depending on the response of the organism to the stress. Here we show that organism response to stress was in some cases asymmetrical (thereby supporting the conjecture that organism-responses are not redundant), and several organism groups/metrics responded differently to the environmental gradients tested here. This information is useful for designing more cost-effective monitoring programs, where the use of early warning indicators can potentially signal change before deterioration is allowed to proceed too far.

Since environmental stress gradients are often correlated (multiple stressors), principal components analysis was used to construct complex stressor gradients. Comparison of the response of the four organism groups to the tails of the PC stressor gradients was used to evaluate organism-specific response to stress. In particular, we were interested in the slope and error of the response within the top end (upper tail) of the environmental gradient, where ecological impairment may be considered as changing from high to lower quality along the gradient. Higher slope and lower error than the null model would imply that the organism can be considered as an early warning indicator for the stressors studied here.

The primary gradient for both mountain and lowland streams was interpreted as representing a

Table 5. Summary statistics for regression of organism group CA scores and N2-diversity and PC gradients for lowland streams

	Fish		Macrophytes		Macroinvertebrates		Diatoms	
	CA axis 1 scores	Diversity	CA axis 1 scores	Diversity	CA axis 1 scores	Diversity	CA axis 1 scores	Diversity
<i>PC1 gradient (null model)</i>								
<i>n</i>	82	82	81	81	71	71	81	81
<i>R</i> ²	0.049	0.030	0.366	-0.011	0.148	0.011	0.606	0.181
RMSEP	1.047	1.638	1.462	5.327	0.828	7.114	0.854	7.849
Slope	-0.089	0.114	0.140	0.083	0.116	-0.380	-0.355	1.269
<i>p</i> value	0.026	0.0662	0.0001	0.6924	0.0006	0.184	0.0001	0.0001
<i>PC1 upper tail (values > 75th percentile)</i>								
<i>n</i>	21	21	19	19	21	21	21	21
<i>R</i> ²	0.327	-0.001	0.290	0.109	0.565	-0.028	0.724	0.054
RMSEP	0.998	1.156	2.281	3.660	0.205	6.927	1.016	4.324
Slope	0.522	-0.184	-1.115	1.109	0.170	0.743	-1.188	1.011
<i>p</i> value	0.004	0.333	0.0102	0.0912	0.0001	0.5102	0.0001	0.1599
<i>PC1 lower tail (values < 25th percentile)</i>								
<i>n</i>	19	19	20	20	20	20	15	15
<i>R</i> ²	0.000	0.003	-0.040	-0.052	-0.056	0.104	-0.073	-0.011
RMSEP	0.921	1.711	1.192	6.325	0.114	7.471	1.332	9.251
Slope	0.414	-0.789	0.275	0.691	0.234	-9.552	0.001	3.685
<i>p</i> value	0.3306	0.3184	0.6131	0.8101	0.9844	0.1291	0.8276	0.3865
<i>PC2 gradient (null model)</i>								
<i>n</i>	82	82	81	81	71	71	81	81
<i>R</i> ²	0.177	-0.013	0.053	0.111	0.089	-0.013	-0.005	0.021
RMSEP	0.974	1.673	1.788	4.995	0.856	7.202	1.362	8.586
Slope	-0.227	0.002	-0.219	-0.871	0.187	0.159	-0.056	-0.737
<i>p</i> value	0.0001	0.985	0.0222	0.0014	0.0067	0.7794	0.4403	0.1059
<i>PC2 upper tail (values > 75th percentile)</i>								
<i>n</i>	21	21	21	21	21	21	21	21
<i>R</i> ²	0.031	0.017	0.112	0.050	0.061	0.008	-0.041	-0.026
RMSEP	0.257	1.428	1.226	3.093	1.053	7.852	0.228	7.495
Slope	-0.115	-0.577	-0.807	-1.544	0.556	2.962	-0.370	-1.829
<i>p</i> value	0.2157	0.2614	0.075	0.1691	0.1468	0.2935	0.6469	0.4931
<i>PC2 lower tail (values < 25th percentile)</i>								
<i>n</i>	20	20	21	21	19	19	21	21
<i>R</i> ²	0.273	-0.053	0.003	-0.053	-0.115	0.051	-0.007	-0.045
RMSEP	0.840	2.071	0.661	7.150	1.263	8.138	0.856	9.490
Slope	-0.304	-0.059	0.083	0.024	0.199	-3.642	0.097	-0.445
<i>p</i> value	0.0106	0.825	0.3148	0.9781	0.6869	0.2714	0.3653	0.7043

Values shown in bold text are significant ($p < 0.05$).

gradient in land use and in-stream nutrient concentrations. Benthic diatoms rely on nutrients (especially P) for growth. Therefore, we expected that diatoms would react strongly to changes in the upper tail of the PC gradient, where nutrients might be limiting (e.g., for lowland streams the

upper tail represented a gradient from 8.6 to 127 $\mu\text{g TP/l}$). Likewise, as many benthic macroinvertebrates (e.g. grazers and scrapers) rely on diatoms for food we might expect a close, albeit weaker, relation between macroinvertebrate community composition and the upper tail of the

PC gradient. Our findings of the response of benthic diatoms and macroinvertebrates to the 1st PC gradient were, however, equivocal. Neither diatom CA scores nor diversity were significantly related to the 1st PC (the null model) gradient for mountain streams and the slopes of the two metrics were not significant when regressed against the short gradients (upper and lower tails) of the 1st PC axis. By contrast, for lowland streams both metrics were significantly related to the 1st PC (null model) gradient. The relation between diatom CA scores and the null model was highly significant (CA scores had an R^2 value of 0.606), and this relation improved when CA score were regressed against the upper tail of the PC gradient ($R^2 = 0.724$). The fit between macroinvertebrate CA scores also improved when regressed against the upper tail of the 1st PC gradient ($R^2 = 0.148$ for the null model compared to 0.565 for the upper tail of the gradient). These findings, in particular the first principle relation between benthic diatom response and the PC (nutrient) gradients, supports the conjecture that benthic diatoms, and to some extent even macroinvertebrates, may be considered as early warning organisms of nutrient enrichment. However, the finding that neither diatoms nor macroinvertebrates showed better improvement when regressed against the upper tail of the 1st PC gradient for mountain streams implies that caution should be exercised when extrapolating these finding to other stream types.

Both fish and macrophyte CA scores for mountain streams and fish CA scores for lowland streams showed improved response (higher R^2 values and steeper slopes) compared to the null model. Although macrophyte growth in streams might be expected to be related to increased nutrient concentrations, fish response would not unless there is a bottom-up effect where an increase in diatom biomass results in an increase in macroinvertebrate biomass and subsequently changes in the fish community. For mountain streams we find no support for this conjecture, since neither diatom nor macroinvertebrates were significantly related to the upper tail of the PC gradient. Other factors may, however, be affecting the responses noted here. For example, although we interpreted the primary PC gradients in both mountain and lowland streams to represent nutrient enrichment,

other factors, like characteristics of the riparian foliage (PC1 mountain streams) covary with nutrients along these gradients.

Clear differences were noted not only among the four organism groups studied here, but also between the two metrics used to assess their response to stress. For null model predictions in lowland streams, for example, diversity did not show a significant response for three of the four organism groups (only diatom diversity responses were significant). Conversely, for null model predictions in mountain streams (PC2 gradient) diversity metrics responded more clearly than CA scores (all four for diversity compared to two of four for CA scores). This finding implies that consideration should be given not only to the organism group but also to the metric selected to monitor the effects of the stressor of interest. Recent studies comparing the multiple organism groups and metrics lend support to this finding (e.g., Hering et al., submitted; Johnson et al., 2006).

Evaluating organism–response relations along short environmental gradients revealed interesting findings. We anticipated that benthic diatoms would respond strongly when sites became more impaired (nutrient enriched), and that diatoms would be an appropriate ‘first choice’ indicator for monitoring early changes in nutrient levels. Data from lowland streams strongly support the use of diatoms (and also macroinvertebrates) for monitoring the effects of agricultural land use. However, for mountain streams we found no such support for this relation. Although nutrient concentrations were strongly correlated with the 1st PC gradients for both mountain and lowland streams, other factors may be confounding the nutrient–diatom response signal. Our finding that fish and macrophytes responded to the 1st PC gradient for mountain streams lends support to this conjecture. In summary, our results showed that rates of organism response to the environmental gradients studied here varied among the four groups, implying that certain organisms/metrics can be considered as early warning indicators of ecological change. Selection of organisms that respond more rapidly at the outset of impairment is one way of determining (quantifying) the potential harmful, human-induced effects on ecosystem integrity before degradation is

allowed to proceed to the point where the damage is either too costly or impossible to restore. Another commonly used approach is to 'create' early-warning metrics (or pollution-specific metrics) by weighting taxa according to their tolerance or sensitivity to a known stressor (e.g. the Saprobien index). Clearly, both approaches should be used together in designing robust methods for detecting ecological change.

Acknowledgements

This paper is a result of the EU-funded project STAR (5th Framework Programme; contract number: EVK1-CT-2001-00089). Parts of the data analysis were supported by the EU-funded Integrated Project Euro-limpacs (6th Framework Programme; contract number: GOCE-CT-2003-505540). We are most grateful to all STAR partners having provided data for this analysis.

References

- Barbour, M. T., J. Gerritsen, B. D. Snyder & J. B. Stribling, 1998. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish (2nd edn.). EPA/841/B/98-010 U.S. Environmental Protection Agency Office of Water, Washington, DC.
- European Commission, 2000. Directive 2000/60/EC of the European Parliament and of the Council – Establishing a framework for Community action in the field of water policy, Brussels, Belgium, 23 October 2000.
- Furse, M. T., A. Schmidt-Kloiber, J. Strackbein, J. Davy-Bowker, A. Lorenz, J. van der Molen J. & P. Scarlett, 2004. Results of the sampling programme. A report to the European Commission. Framework V Project STAR (EVK1-CT-2001_00089).
- Hill, M. O., 1973. Reciprocal averaging: an eigenvector method of ordination. *Journal of Ecology* 61: 237–249.
- Hering, D., R. K. Johnson, S. Kramm, S. Schmutz, K. Szoszkiewicz & P. F. M. Verdonschot. Assessment of European rivers with diatoms, macrophytes, invertebrates and fish: a comparative metric-based analysis of organism response to stress, submitted manuscript.
- Hering, D. & J. Strackbein, 2002. STAR stream types and sampling sites. A report to the European Commission. Framework V Project STAR (EVK1-CT-2001_00089).
- Holmes, N. T. H., J. R. Newman, S. Chadd, K. J. Rouen, L. Saint & F. H. Dawson, 1999. Mean Trophic Rank: A users manual. R & D Technical Report No. E 38, Environment Agency, Bristol, UK.
- Johnson, R. K., T. Wiederholm & D. M. Rosenberg, 1993. Freshwater biomonitoring using individual organisms, populations and species assemblages of benthic macroinvertebrates. In Rosenberg, D. M. & V. H. Resh (eds), *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman and Hall, New York, 40–158.
- Johnson, R. K., D. Hering, M. T. Furse & R. T. Clarke, 2006. Detection of ecological change using multiple organism groups: metrics and uncertainty. *Hydrobiologia* 566: 115–137.
- Kolkwitz, R. & M. Marsson, 1902. Grundsätze für die biologische Beurteilung des Wassers nach seiner Flora und Fauna. *Mitteilungen Prüfungsanstalt Wasserversorgung und Abwasserreinigung* 1: 33–72.
- Metcalf, J. L., 1989. Biological water-quality assessment of running waters based on macroinvertebrate communities – history and present status in Europe. *Environmental Pollution* 60: 101–139.
- Raven, P. J., N. T. H. Holmes, F. H. Dawson, P. J. A. Fox, M. Everard, I. R. Fozzard & K. J. Rouen, 1998. River habitat quality – the physical character of rivers and streams in the UK and Isle of Man. *River Habitat Survey Report Number 2*, Environment Agency, Bristol, Scottish Environment Protection Agency, Stirling, Environment and Heritage Service, Belfast, 84 pp.
- SAS., 1994. JMP – Statistics Made Visual, Version 3.1. SAS Institute Inc, Cary, NC, USA.
- Stanner, D. & P. Bordeaux, 1995. Europe's Environment: The Dobbris Assessment. European Environment Agency, Luxembourg 712 pp.
- Stevenson, R. J., R. C. Bailey, M. C. Harrass, C. P. Hawkins, J. Alba-Tercedor, C. Couch, S. Dyer, F. A. Fulk, J. M. Harrington, C. T. Hunsaker & R. K. Johnson, 2004. Designing data collection for ecological assessments. In Barbour, M. -T., S. B. Norton, H. R. Preston & K. W. Thornton (eds), *Ecological Assessment of Aquatic Resources: Linking science to decision making*. SETAC, Pensacola, FL, USA, 55–84.
- ter Braak, C. F. J., 1988. CANOCO – a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal component analysis and redundancy analysis (version 3.15). Agricultural Mathematics Group, Wageningen, The Netherlands.
- ter Braak, C. F. J., 1990. Update Notes: CANOCO Version 3.10. Agricultural Mathematics Group, Wageningen, The Netherlands.