



**HAL**  
open science

# **Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of trait patterns in least impacted river reaches**

Bernhard Statzner, Pierre Bady, Sylvain Dolédec, Franz Schöll

## **► To cite this version:**

Bernhard Statzner, Pierre Bady, Sylvain Dolédec, Franz Schöll. Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of trait patterns in least impacted river reaches. *Freshwater Biology*, 2005, 50 (12), pp.2136-2161. <10.1111/j.1365-2427.2005.01447.x>. <halsde-00340340>

**HAL Id: halsde-00340340**

**<https://hal.science/halsde-00340340v1>**

Submitted on 5 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of trait patterns in least impacted river reaches

BERNHARD STATZNER,\* PIERRE BADY,\* SYLVAIN DOLEDEC\* AND FRANZ SCHOLL†

\*CNRS-Ecologie des Hydrosystèmes Fluviaux, Université de Lyon 1, France

†Bundesanstalt für Gewässerkunde, Koblenz, Germany

1. Multiple biological invertebrate traits (e.g. body size, body form, dispersal potential) each described through multiple categories (e.g. small, intermediate or large body size) could serve as indicators of particular types of human impacts on large rivers. The trait composition of natural invertebrate communities is scarcely constrained by taxonomic differences among them, i.e. individual trait categories could be used to discriminate various types of human impacts across large geographic areas, which would require the definition of trait patterns for conditions of relatively low human impact.
2. Using large databases to link 14 biological traits (described through 66 categories) of invertebrate genera to their occurrence in running water reaches with known environmental conditions, we examined the accuracy of various approaches to predict expected trait variation across least impacted river reaches (LIRRs) of Europe in a stepwise analytical procedure. This procedure included Monte Carlo simulations and ultimately the assignment of test-LIRRs (reaches not used in previous analyses) to the previously defined LIRR conditions.
3. Distance from the source was an integrative variable capturing some (but not all) landscape features (e.g. altitude) or habitat variables (e.g. reach shear stress). Correspondingly, the relative abundance of many trait categories changed along 13 European running waters, although particularly the intensity of these changes differed among these 13 running waters.
4. 'Downstream models' (using only distance from the source as predictor) provided the least accurate predictions of expected invertebrate trait patterns when compared with 'landscape models' (using distance from the source in combination with altitude and/or latitude) or 'habitat models' (using reach shear stress, mean annual air temperature and/or pH of the water). Landscape models provided more accurate predictions than habitat models, but the improvement of predictions of expected invertebrate traits patterns obtained using landscape models was negligible in comparison with a simpler 'mean-model' approach, suggesting that defining LIRR conditions through simple descriptions of frequency distributions would be sufficient for the future biomonitoring of large European rivers.
5. To define these LIRR conditions, we used the average of the relative abundance of each trait category from 68 LIRRs ( $\geq 40$  m wide) as expected LIRR values, and computed LIRR frequency distributions that described the deviations of the 68 individual LIRRs from these

expected values. Computing such deviations from the expected LIRR values for 57 test-LIRRs (also  $\geq 40$  m wide), 57 trait categories correctly assigned  $>90\%$  of the test-LIRRs to LIRR conditions if the latter were defined through the entire range of the LIRR frequency distributions. To the  $90\%$ -range enveloped by the LIRR frequency distributions, 42 trait categories correctly assigned  $>80\%$  and 12 categories  $>90\%$  of the test-LIRRs.

6. Using a framework that required no regionalisation of a large geographic area, no modelling of expected values using environmental information and no standardised invertebrate sampling, the performance of our trait approach to assign test-LIRRs to LIRR conditions encourages future assessments of deviations from these defined LIRR conditions in large European river reaches with different types of human impacts.

*Keywords:* benthos, downstream trends, habitat relations, landscape relations, modelling

## Introduction

Despite a long history of developing biomonitoring methods (Statzner *et al.*, 2001a), ecologists are being asked to create new biomonitoring tools in response to recent trends in environmental policies. First, environmental policies are increasingly executed across large geographic areas (Anonymous, 1999a; Ormerod, Pienkowski & Watkinson, 1999). As existing biomonitoring tools are typically developed for smaller geographic areas, this trend requires the redefinition or rescaling of tool(s) before it/they can be applied to other regions (Rosenberg & Resh, 1993; Hill *et al.*, 2000; Wright, Sutcliffe & Furse, 2000). Second, the scientific debate about biodiversity and related ecological functions and management for conservation and restoration of ecosystems (McGrady-Steed, Harris & Morin, 1997; Van der Heijden *et al.*, 1998; Ghilarov, 2000) induced reallocations of considerable financial resources to enhance ecological functions (Anonymous, 1998, 1999a). However, most of the existing biomonitoring tools assess the structural attributes (i.e. taxonomic composition) of communities, whereas tools that assess functional community attributes are rare (Tscharntke, Gathmann & Steffan-Dewenter, 1998). Because relations between structural and functional community attributes are equivocal (e.g. Naeem, 2002), it is important to develop biomonitoring tools that assess functional attributes. In addition, functional community attributes can be compared across large geographic areas even if the structural community attributes differ among them (Statzner *et al.*, 2001a; Bremner, Rogers & Frid, 2003).

For benthic invertebrates of running waters, multi-metric biomonitoring approaches tried to combine functional (i.e. functional feeding groups) and struc-

tural community attributes, but current evidence suggests that the metric 'functional feeding groups' of lotic invertebrates is generally a poor indicator of human disturbances of running waters (e.g. Karr & Chu, 1999). Therefore, Dolédec, Statzner & Bournaud (1999) suggested the use of a number of biological invertebrate traits (e.g. body size, body form, dispersal potential), each described through multiple categories (e.g. small, intermediate or large body size), as each of the individual traits (or each of the more numerous individual trait categories) could serve in a 'multi probe' approach (i.e. having multiple metrics) to discriminate various types of human impact on running waters. Among these biological traits, some describe obvious (e.g. food and feeding habits) or more subtle (e.g. size and the correlated ratios of production/biomass and of production/respiration; see Statzner, 1987; Benke, 1993) ecological functions of lotic invertebrates, whereas others describe functional invertebrate attributes that should be predictably affected by various human impact types (Table 1). The idea to apply this multiple biological trait approach in biomonitoring across large geographical areas was derived from: (i) predictions from habitat templet theory on multiple trait responses of lotic organisms to natural environmental disturbance (Townsend & Hildrew, 1994) and several confirmations of these predictions (e.g. Statzner, Resh & Dolédec, 1994; Townsend, Dolédec & Scarsbrook, 1997) and (ii) demonstrations of the possibility to analyse lotic invertebrate trait patterns in a consistent way across large spatial scales (Statzner *et al.*, 1997, 2001a; Statzner, Dolédec & Hugueny, 2004).

To develop a biomonitoring tool for large European rivers (having a width  $\geq 40$  m) derived from the multiple trait approach, previous studies demonstra-

**Table 1** Examples of predictions on how global human impact or a particular impact type should increase (↑) or decrease (↓) the relative abundance of lotic invertebrate trait categories (see Table 4 in the result section for a detailed description of the traits and categories used in this study)

Impact type	Trait	Category	Rationale
Global	B: Descendants per cycle C: Voltinism	Many ↑ Short ↑	Increase of resilience capacity
Flow increase	A: Body size I: Locomotion K: Body form	Small ↑ Attached ↑ Streamlined ↑	Compensation for action of flow forces
Flow decrease	A: Body size I: Locomotion K: Body form	Large ↑ Swimmer ↑ Spherical ↑	Release from action of flow forces
Siltation	I: Locomotion L: Feeding habits M: Food	Burrower ↑ Deposit-feeder ↑ Fine detritus ↑	Greater penetrability of substrate and greater food availability
Heavy metals	A: Body size N: Respiration	Small ↓ Gills ↓	Greater body surface-volume ratio favours metal uptake per unit volume
Organic pollution	N: Respiration	Aerial ↑	Compensation for O <sub>2</sub> -deficits in the water
Cargo-ship traffic	F: Reproduction G: Parental care	Single individual ↑ Ovoviviparity ↑	Facilitation of foundation of new populations through specimens dispersed by ships

ted that: (i) the biological trait composition of invertebrate communities reliably discriminated reaches with increased overall human impact from less impacted reaches along a large river (French Rhône) flowing through several ecoregions (Dolédéc *et al.*, 1999), (ii) among alternatives to describe the trait composition of invertebrate communities, identifications at the genus level, presence-absence data in the taxa-weighting and the inclusion of alien taxa provided significant discriminations of various levels of human impact across Europe (Gayraud *et al.*, 2003) and (iii) few sample replicates in space or time (across seasons or subsequent years) reliably described the trait composition of invertebrate communities (Bady *et al.*, 2005). Therefore, the next logical step in this development would be the definition of trait patterns at reference conditions of large European rivers, as biomonitoring tools that relate community characteristics to the environmental quality of freshwaters (and other systems) usually use the ‘reference condition approach’ (through comparisons of stressed sites with a group of minimally disturbed sites, Rosenberg & Resh, 1993; Wright *et al.*, 2000). However, reflecting European reality, the databases used in this study included few large river reaches that were almost natural, so our assessment of trait patterns focussed on conditions in best available or least impacted river reaches (LIRRs) of Europe.

Defining trait patterns for LIRRs required the clarification of many questions. For example, a previ-

ous study focussing on smaller European streams demonstrated that only a few trait categories varied significantly with altitude, stream size and latitude (Statzner *et al.*, 2004), suggesting that many of the trait categories in our LIRRs could perhaps be defined using a mean European trait profile (here referred to as the ‘mean-model’ approach) and simple frequency distributions of deviations of category values from the mean profile observed in our LIRRs. In contrast, trait categories that vary significantly with environmental conditions in our LIRRs may require models that predict a substantial part of this variation so that LIRR predictions (i.e. the expected relative abundance of trait categories) could be compared with observations from other least impacted or impacted large river reaches (see Chessman, 1999 or Oberdorff *et al.*, 2002, for a similar approach predicting invertebrate community structure or a fish-based index indicating ‘river health’). Given that the relevance of the environmental description for invertebrates depends on the variables used as well as on the scale at which they are described (see Richards *et al.*, 1997; Hawkins *et al.*, 2000; Johnson, Goedkoop & Sandin, 2004) and corresponding to the data availability, we opted for three environmental regression model types to predict the expected variation of a trait category using LIRR data: (i) a ‘downstream model’ using distance from the source as the independent variable, (ii) in addition to distance from the source, variables such as altitude and latitude were used to integrate landscape features

(in a 'landscape model') that could modify downstream trends in the relative abundance of trait categories and (iii) habitat (i.e. local-scale) variables characterising temperature, near-bed flow and water chemistry of the LIRRs were used in a 'habitat model'.

We had numerous options to test these three alternative approaches, so we organised this paper according to insights obtained from a stepwise analytical procedure. First, we assessed which landscape features and habitat variables were sufficiently described by distance from the source. Second, we searched for common downstream trends in trait category patterns across a large variety of European running waters types. Third, using reaches of intermediate (25 to <40 m wide) to large ( $\geq 40$  m wide) rivers, we compared the predictive power of downstream, landscape and habitat models to that of our mean-model approach. Fourth, we repeated the procedure described in step 3, focussing on the predictive power of the environmental models and using data representing LIRR conditions for large ( $\geq 40$  m wide) European rivers. Fifth, we defined the LIRR trait structure. Sixth, addressing the ultimate task of a biomonitoring tool (see Oberdorff *et al.*, 2002), we validated our model by assessing how many reaches of a different data set were correctly assigned to LIRR conditions using the criteria defined in step 5.

## Methods

Following the suggestions of Gayraud *et al.* (2003) (see also above) we used genus level identifications, presence-absence data in the taxa weighting of the traits (see below) and included alien taxa in the analyses of this study.

### Data availability

For this study, we combined previously described databases on invertebrate abundances (see Gayraud *et al.*, 2003; Statzner *et al.*, 2004) and added information on the occurrence of invertebrate genera for reaches such as the Argens (Giudicelli, Dia & Legier, 1980), Dordogne (Anonymous, 1980), Eau d'Olle (Gay, 1982) and the Upper Elbe [data of the Bundesanstalt für Gewässerkunde (BfG), in charge of the national survey of large German waterways]. As a result, our expanded database on invertebrate occurrences had 2087 cases. Typically, a case represented a reach

sampled more than once per annual cycle, or annual data for a reach sampled more than once in each of several years. The occurrence of macroinvertebrate genera available for these cases excluded Oligochaeta and most Diptera families, because data were not consistently available at the genus level (Gayraud *et al.*, 2003; Statzner *et al.*, 2004). The taxonomic community compositions of the 2087 cases had been assessed using a variety of methods. Statzner *et al.* (2004) showed that the use of different methods resulted in approximately 20% of the variation in the raw abundances and approximately 10% of the variation in ln-transformed abundances of the invertebrates in 384 reaches. In this study, however, this problem should be less important as we used presence-absence data of genera for the taxa-weighting of the traits.

When available, the environmental characteristics of the cases (required for the downstream, landscape or habitat models, or for the assessment of human impacts) were described using the data sources of the invertebrate occurrences. We expanded this information using other sources providing data on geographical co-ordinates and simple physical reach characteristics such as altitude, mean reach width and mean reach slope [ARCVIEW, <http://www.esrifrance.fr>; map data, <http://www.ign.fr>; map material (1/50 000 or more detailed scales) of the BfG]. According to Leopold, Wolman & Miller (1964), reach depth ( $d$ , in metres) is typically related to reach width ( $w$ , in metres) by  $d = 0.041 w^{0.84}$ . We used this relation and available data on reach width to calculate the hydraulic radius ( $R$ : cross-sectional area/wetted perimeter, in metres) of the reaches, and combined with reach slope ( $S$ ), gravitational acceleration ( $g$ ,  $\text{m s}^{-2}$ ), and water density ( $\rho$ ,  $\text{kg m}^{-3}$ ) calculated the mean shear stress ( $\tau_0$ ,  $\text{N m}^{-2}$ ) per reach (see Leopold *et al.*, 1964) as  $\tau_0 = \rho g R S$ . We used average air temperature in January and July to characterise reach temperature (<http://www.sage.wisc.edu/atlas>, see also New, Hulme & Jones, 1999) and included basic water quality parameters such as pH, alkalinity, nutrients, indicators of organic pollution and heavy metals to describe reach chemistry (e.g. <http://www.eaufrance.com>; Anonymous, 1999b; Löffler, Wolff & Bergemann, 1999; Braun, 2001). Finally, we expanded descriptions of reach modifications of discharge and channel morphology (e.g. dams, canalisation; using a database of the BfG and <http://www.ign.fr>) and of

cargo-ship traffic (indicated by the transported freight in tons year<sup>-1</sup>) through the reaches (<http://www.elwis.de>; <http://www.binnenschiff.de>; Anonymous, 2002; and assuming ‘no traffic’ for reaches that were obviously too shallow for cargo-ships to pass).

Because of inconsistent data availability, the lack of environmental information was a major constraint for the analyses of this study. For example, of the 2087 cases in our invertebrate occurrence database, 95.7% had information on distance from the source, 72.0% on distance from the source, altitude and latitude, and 59.7% on shear stress, air temperature and pH. In addition, data were also lacking for variables indicative of the degree of human impact. Therefore, we described organic pollution in classes using class limits defined by the French water authorities (see Table 2 for ‘very good’ and ‘good’ quality classes) using the five-day biochemical oxygen demand (BOD<sub>5</sub>). If BOD<sub>5</sub> data were not available, we used NH<sub>4</sub><sup>+</sup> concentrations to describe organic pollution, and if information on BOD<sub>5</sub> and NH<sub>4</sub><sup>+</sup> were not available, we used NO<sub>3</sub><sup>-</sup> concentrations (Table 2). Similarly, we described heavy metal pollution using classes (Table 2) and used the highest class observed to describe the overall heavy metal pollution of our cases. If data describing organic or heavy metal pollution were not available for a case, but the author of the invertebrate data rated the studied conditions as unpolluted, we assigned the case to class 1 for both organic and heavy metal pollution.

The database on the biological invertebrate traits, described in detail by Gayraud *et al.* (2003), contains affinity scores (ranging from ‘zero’ for no affinity of a

species to a given trait category to ‘three’ for a high affinity to that trait category, see Chevenet, Dolédec & Chessel, 1994) for 66 categories of 14 traits (see Table 4 in our result section) of invertebrate species occurring in large European rivers. These affinity scores of species were used to derive similar scores for genera (the taxonomic level considered here) by averaging the affinity scores of species belonging to a given genus. The genus scores were then rescaled so that their sum for a given trait equalled ‘one’. To obtain the proportion of trait categories in a community requires that the traits are linked to the taxonomic structure (e.g. weighting the traits using taxa abundance or biomass, see Bremner *et al.*, 2003; Johnson *et al.*, 2004). Hence, we weighted the affinity scores of the genera for the 66 trait categories by multiplying the affinity scores of each genus with its presence–absence, and added these values by trait category. The latter trait category values were rescaled to sum to ‘one’ for each trait and each case included in our invertebrate occurrence database.

Because the biological trait profiles of our invertebrate genera were aggregates of trait profiles of species occurring in large rivers, it was uncertain if these genera profiles could be used to describe downstream trends in trait categories if smaller headwater streams were included, as other species than those occurring in large rivers could represent a given genus in the headwaters. Using 18 categories of five traits (maximal size, voltinism, locomotion and attachment to substrate, feeding habits, respiration technique) and 248 genera of our large river database that were replicated in a trait database considering biological information from all types of European freshwaters (see details in Usseglio-Polatera *et al.*, 2000), co-inertia analysis (see Dolédec & Chessel, 1994) showed that the 248 genera were not identically but similarly described in both databases (*RV*-coefficient: 0.39; simulated significance of co-structure:  $P < 0.001$ ). This finding suggested that the use of trait profiles of large river taxa in smaller headwater streams should not constrain assessments of downstream trends in trait categories.

#### Data selection and analyses

The description of data availability in the previous section indicated that careful data selection was a prerequisite for our analyses. As a starting point, we

**Table 2** Definition of ‘very good’ (class 1) and ‘good’ (class 2) water quality used by the French water authorities for organic and heavy metal pollution (Anonymous, 2000) and used by us to select least impacted river reaches (LIRRs)

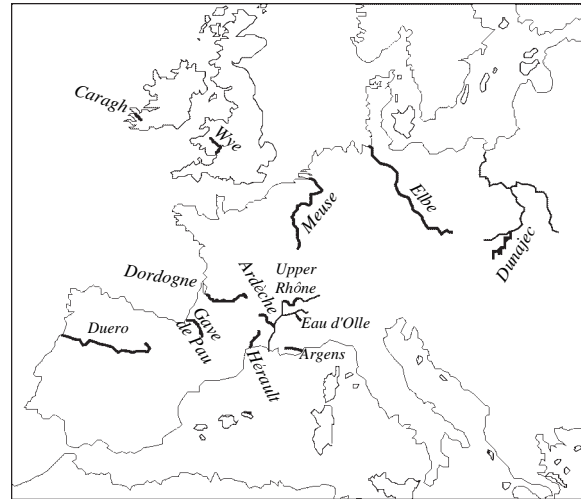
Variable	Class 1*	Class 2*
BOD <sub>5</sub> (mg L <sup>-1</sup> )	≤3	>3–6
NH <sub>4</sub> <sup>+</sup> (mg L <sup>-1</sup> )	≤0.1	>0.1–0.5
NO <sub>3</sub> <sup>-</sup> (mg L <sup>-1</sup> )	≤2	>2–10
Cd (µg L <sup>-1</sup> )	≤0.05	>0.05–0.44
Cu (µg L <sup>-1</sup> )	≤1.29	>1.29–12.90
Hg (µg L <sup>-1</sup> )	≤0.07	>0.07–0.70
Ni (µg L <sup>-1</sup> )	≤6.90	>6.90–20.00
Pb (µg L <sup>-1</sup> )	≤5.77	>5.77–14.33
Zn (µg L <sup>-1</sup> )	≤6.86	>6.86–68.67

\*Class limits defined according to European and French directives, literature analyses values and expert opinion.

decided that in each of the six steps of our analytical procedure described in Table 3, only one of replicated annual averages for a given case was used, as the environmental conditions as well as the trait composition of the invertebrate communities should vary little among years (see Bady *et al.*, 2005).

*Step 1: relating landscape features and habitat variables to distance from the source.* Our database contained 287 reaches for which all of the information given in Table 3 was available. We analysed these data using normalised Principal Component Analysis (PCA) to assess correlations among the environmental variables.

*Step 2: searching for downstream trends in trait categories.* We selected 13 running waters that were spread across Europe (Fig. 1), differed in altitude, slope and



**Fig. 1** The 13 running waters selected to search for downstream trends in 66 invertebrate trait categories. Note that we over-emphasised the extension of shorter running waters on this map.

**Table 3** Selection and characteristics of the data used in our stepwise analyses. Note that selection criteria were often dictated by data availability (see text)

Step of analyses	Data used
Step 1: relating landscape features and habitat variables to distance from the source	287 reaches; co-ordinates (decimal degrees): 10.0W*–20.9E, 41.0–53.4N; distance from the source (km): 0.05–1041; altitude (m a.s.l.): 1–1300; reach width (m): 0.3–800; reach slope (–): 0.00001–0.140; shear stress (N m <sup>-2</sup> ): 0.3–143.7; January air temperature (°C): –6–7; July air temperature (°C): 13–23; mean annual air temperature [0.5 × (January + July); °C]: 3.5–14.5; air temperature range (July – January; °C): 9–21; pH: 6.0–9.5
Step 2: searching for downstream trends in trait categories	13 stream and river systems, see Fig. 1 and text for further details
Step 3: comparing the predictive power of downstream, landscape and habitat models to that of the mean-model approach	80 reaches; co-ordinates: 10.0W–20.9E, 42.4–52.3N; distance from the source: 24–777; altitude: 1–790; reach width: 25–500; shear stress: 0.7–45.2; mean annual air temperature: 3.5–14.5; pH: 6.4–9.0; cargo-ship traffic: 0; organic pollution class: 1–2; heavy metal pollution class: 1–2
Step 4: comparing the predictive power of landscape models to that of the mean-model approach using LIRRs	68 reaches (see Fig. 1 for those without the indication of a country) of the Ain (France; 1 reach), Ardèche (12), Caraghs (1), Dordogne (1), Drome (France; 1), Dunajec (11), Esla (Spain; 2), Gave de Pau (9), Hérault (3), Lahn (Germany; 8), Loire (France; 3), Meuse (9), Rhône (4), and Wye (3); co-ordinates: 10.0W–20.9E, 42.4–52.3N; distance from the source: 24–777; altitude: 10–790; reach width: 40–430; cargo-ship traffic: 0; organic pollution class: 1–2 (51 reaches in class 1); heavy metal pollution class: 1–2 (23 reaches in class 1); genus richness: >10; proportion of genera with alien species: <0.1
Step 5: defining trait patterns for LIRR conditions	Data used in step 4
Step 6: assessing test-LIRRs using the criteria defined in step 5	57 reaches of the Ardèche (4), Gard (France; 1), Gave de Pau (1), Hérault (3), Lahn (27), Loire (8), Meuse (9), and Rhône (4); co-ordinates: 0.8W–8.1E, 43.4–50.4N; distance from the source: 60–777; altitude: 10–230; reach width: 40–500; cargo-ship traffic: 0; organic pollution class: 1–2 (13 reaches in class 1); heavy metal pollution class: 1–2 (8 reaches in class 1)

\*Western longitude transformed into negative values in the analyses.

length and represented a large range of running water types (e.g. small versus large, alpine versus coastal, temperate versus Mediterranean). The smaller streams [e.g. Argens (using only unpolluted reaches), Caragh, Eau d'Olle] represented almost natural conditions, whereas intermediate running waters (e.g. Ardèche, Dordogne, Dunajec, Wye) had some human impact (particularly discharge modifications) but were considered as natural and regional representatives in our database. The larger rivers were even more impacted by humans, but we addressed this problem by selecting data from a period when this impact was deemed minimal. For example, for the Upper Rhône, we selected data from the 1970s, when most of the river was not yet regulated (Perrin, 1978). Likewise, for the Elbe, we selected data from the late 1990s, when restoration measures planned after the German re-unification were implemented. Despite the resulting environmental improvements (Anonymous, 2001), the middle and particularly the lower section of the Elbe remained an important waterway for cargo-ships, and the lower Meuse also had cargo-ship traffic.

To describe the downstream trends of trait categories, we regressed the proportion of a given trait category in a linear relation to distance from the source (*km*), assessed the effects of individual running waters through the categorical variable *river* and included the interaction between them (i.e.  $km \times river$ ). We repeated the analyses adding the square of the distance from the source ( $km^2$ ) to the independent variables, to check if such polynomials would increase the variance explained by the models. To illustrate such relations for some selected trait categories, we modelled the trait-distance relation individually for each running water (if  $km \times river - P < 0.05$ ) or by analysis of covariance (if  $km \times river - P \geq 0.05$ ).

*Step 3: comparing the predictive power of downstream, landscape and habitat models to that of the mean-model approach.* Steps 1 and 2 of our analyses suggested to model trait responses in regressions using distance from the source (downstream models); distance from the source, altitude and latitude (landscape models); and shear stress, mean annual air temperature and pH (habitat models). Thus, for step 3, we needed a set of reaches having all this environmental information available and, at the same time, were least impacted by human activities. Including reaches of intermediate

size (with a reach width between 25–40 m), 80 reaches satisfied these criteria (Table 3).

We tested the raw data and various transformations of these environmental variables to identify, for each variable, the data format that provided the least skewed and flattest (using kurtosis) distribution of the variable, which was then consistently used in simulations of their predictive power in regression models. In addition to appropriate data transformations, we assessed alternative modelling strategies to assure that the best predicting alternative was used in comparisons with the mean-model approach.

The downstream regression models were *a priori* defined (through the relation between the trait categories and distance from the source), whereas the general form of landscape and habitat regression models had to be defined for each of the trait categories. With three potential independent variables and three potential first-order interactions among them in each of the landscape and habitat models, our goal was to build multiple regression models using the fewest independent variables to predict as much as possible of the variability of the trait responses (a strategy recommended by Graham, 2003, as the uncontrolled inclusion of many independent variables into multiple regression models leads to poor predictions, e.g. Gozlan *et al.*, 1999). Therefore, for each trait category, we used all 80 reaches in an interactive stepwise procedure to select the fewest and most significant (using partial *P*-values) terms to define the general form for simpler (ignoring interaction terms) and more complex regressions (including interaction terms if these were significant, i.e. their partial *P*-value was  $< 0.05$ ). If none of the three independent variables was significant ( $P \geq 0.05$ ), we defined the general form of the regression using the variable that tended to explain most of the variability of a given trait category. Correspondingly, we also examined downstream regression models that tended to explain trait responses.

We used Monte Carlo simulations (Manly, 1991) to test the predictive power of the various model types by iterating: (i) random assignments of half of the 80 reaches each to a predictor and a test set, (ii) estimations of the parameters for each previously defined general regression form as well as calculations of the mean of each trait category proportion (corresponding to our mean-model approach) using the 40 predictor reaches and (iii) predictions for the

40 test reaches (using the regression parameters estimated from the predictor set and inserting the corresponding independent variables from the test set, or using the mean of the trait category proportion in the predictor reaches). Thus, these iterations provided predictions for the test reaches that could be compared with the observed trait data. Preliminary iterations using simpler or more complex regressions and the mean-model approach indicated that the mean absolute error (average of  $|\text{observed} - \text{predicted}|$ ) and the variance of this absolute error (i.e. the two criteria used to assess predictive power) stabilised after 10–20 iterations. Therefore, we used 100 iterations in the simulations, assigning predictor and test reaches for all 66 trait categories through one random draw per iteration.

We regressed mean absolute error and error variance of simpler versus more complex landscape and habitat regression models (using all trait categories having this information available) to compare the predictive power of simpler and more complex regressions, and focussed on the type having the greatest predictive power in subsequent analyses. In these, we used all trait categories to regress mean absolute error and error variance obtained by downstream, landscape and habitat models on mean absolute error and error variance obtained by the mean-model approach (using linear or piecewise linear regression, see Wilkinson, Blank & Gruber, 1996). The regression slopes provided an overall (for all trait categories together) assessment of the predictive power of downstream, landscape and habitat models in comparison to the mean-model approach. In addition, outliers in these regressions indicated trait categories that were better predicted by a given approach.

*Step 4: comparing the predictive power of landscape models to that of the mean-model approach using LIRRs.* Applying selection criteria on human impact as in step 3, excluding cases with a reach width <40 m and leaving enough reaches in this selection for the test-LIRRs required for step 6, we selected reaches as LIRRs that were geographically spread across Europe and tried to balance the weight (i.e. number of reaches) of individual rivers as much as possible. After these selections, 82 reaches remained as potential candidates for our LIRRs. For these, we visually inspected the frequency distribution of both genus richness and proportion of genera

with alien species (see Table 2 in Gayraud *et al.*, 2003, for a list of alien species). The lower tail of the distribution on genus richness (at  $\leq 10$  genera) and the upper tail of the distribution on the proportion (at  $\geq 0.1$ ) of aliens were removed as we assumed that the data in these tails were the least representative for LIRR conditions. Consequently, our LIRRs included 68 reaches (see Table 3, step 4). Among these, only 16 had no cargo-ship traffic and organic/heavy metal pollution in class 1, thus representing the most natural reaches. Among these 16 reaches, however, there were still reaches with discharge modifications. Using these 68 LIRRs, we repeated the analytical procedures described in step 3 to compare the predictive power of landscape models to that of the mean-model approach.

*Step 5: defining trait patterns for LIRR conditions.* Step 4 of our analyses suggested to use the mean-model approach for the definition of trait patterns at LIRR conditions. For this purpose, we used the same data as in step 4 (Table 3). For individual trait categories, we defined the LIRR conditions as frequency distributions of the difference between the observed values in the 68 reaches and an expected value (i.e. the average of all 68 reaches; thus, the distribution of observed – expected). Half of the sum of the absolute difference between observed ( $p_j$ ) and expected ( $p_{exp}$ ) values for all categories ( $i$ ) of a trait provided the dissimilarity per trait for the  $j$  reaches (i.e.  $0.5 \sum |p_{ij} - p_{iexp}|$ ), and averaged over all traits, for the overall trait profile (see Spellerberg, 1991, p. 132). These dissimilarities were also described as frequency distributions. With 66 categories, 14 traits and the mean of the 14 traits, the 81 LIRR frequency distributions as expected did not follow a common (e.g. normal) law. Therefore, we defined these distributions using both their entire range or, after removal of extreme values (see step 5 in the result section), the range enveloping 90% of the reaches.

*Step 6: assessing test-LIRRs using the criteria defined in step 5.* Assessing whether test-LIRRs were correctly assigned to LIRR conditions using the criteria defined in step 5, we used all reaches not used as LIRRs in steps 4 and 5. In comparison with the LIRRs, these 57 test-LIRRs were not as widely dispersed across Europe (Table 3, step 6) and included reaches having  $\leq 10$  genera or a proportion of genera with alien species  $\geq 0.1$ . Constraining the test-LIRRs to  $> 10$  genera and  $< 0.1$  genera with alien species would

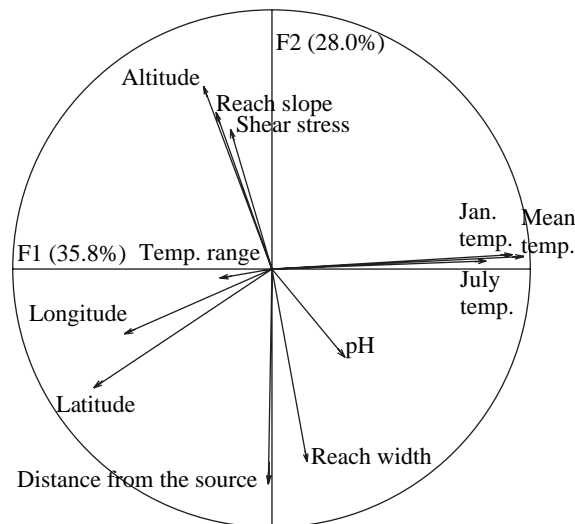
have resulted in only 40 test-LIRRs available. For the 57 test-LIRRs, we compared the observed trait patterns with those expected from the LIRRs to obtain the test-LIRR frequency distributions for individual trait categories, traits and all traits together as described in step 5. Comparing the frequency distributions of the test-LIRRs with those of the LIRRs provided a simple measure for the correct assignment of the test-LIRRs.

*Software use.* We used ADE-4 (Thioulouse *et al.*, 1997 and ADE-4 release 2001) for the normalised PCA and R (the free version of S-Plus) for the Monte Carlo simulations. All other analyses were carried out with SYSTAT 10.

## Results

### *Step 1: relating landscape features and habitat variables to distance from the source*

The first two axes of the normalised PCA on the environmental characteristics of 287 smaller, intermediate and larger running water reaches explained 63.8% of the data variability. Reach width correlated positively with distance from the source (Fig. 2). Altitude, reach slope and shear stress correlated negatively with distance from the source (Fig. 2). Therefore, distance from the source promised to be an



**Fig. 2** Correlations (illustrated by normalised Principal Component Analysis) among environmental characteristics of 287 European running water reaches. F1 and F2 show the variance explained by the first and second axis, respectively.

integrative variable capturing landscape features such as altitude or habitat variables such as shear stress. In contrast, variables such as latitude and air temperature (mean, January, July) would not be sufficiently described by distance from the source (Fig. 2). Therefore, in the next step, we assessed if distance from the source was able to explain downstream trends in trait categories along running waters spread across Europe.

### *Step 2: searching for downstream trends in trait categories*

The downstream trends in the 66 trait categories along the 13 running waters that differed considerably in geographic location (see Fig. 1), altitude, slope, length and other characteristics (see above) are given in Table 4.

Comparing the coefficient of determination between linear ( $km$ ) and polynomial ( $km, km^2$ ) relations of the trait categories with distance from the source suggested that polynomial relations explained a little more of the variation of trait responses than linear relations for few categories (B4: >3000 descendants per cycle; G5: cemented aquatic eggs; H5: dispersal potential >10 000 m; I4: temporarily attached to the substrate; J2: low body flexibility; L2: shredder; L6: passive filter-feeder; M6: feeding on plants >10 mm; see Table 4). Therefore, we focussed on linear trait relations here.

Although many of the trait categories tended to change linearly with  $km$ , only 10 changed significantly with  $km$  (G4: cemented terrestrial eggs; H2 and H3: dispersal potential >10–100 or >100–1000 m; I2: crawler; K1: streamlined body form; L3: scarper; M2 and M3: feeding on detritus >1–10 or >10 mm; M4: feeding on plants  $\leq 1$  mm; N3: plastron respiration; see Table 4). However, for nine of these 10 categories running water effects (*river*) were clearly more significant than  $km$  (all but K1: streamlined body form). In addition, the interaction  $km \times river$  was more significant than  $km$  in half of the 10 categories (H2: dispersal potential >10–100 m; I2: crawler; M2 and M3: feeding on detritus >1–10 or >10 mm; N3: plastron respiration; see Table 4). Within most of the traits, the significance levels for  $km$ , *river* and/or  $km \times river$  differed considerably among the trait categories (Table 4). Thus, although the categories of a given trait were not independent of each other (they added

**Table 4** Common slope (positive or negative) of the relative abundance of trait categories in relation to distance from the source (*km*), partial *P*-values for distance from the source (*km*), running water effects (assessed through the categorical variable *river*) and their interaction (*km* × *river*) and coefficient of determination ( $R_1^2$ ) all for linear distance regressions (*km* untransformed). The coefficient of determination ( $R_2^2$ ) for regressions including a squared distance term ( $km^2$ ) in addition to the previous independent variables is also given. Number of reaches: 165 in 13 running waters (see Fig. 1).

Biological trait/category	Common slope	<i>km</i> ( <i>P</i> -value)	<i>river</i> ( <i>P</i> -value)	<i>km</i> × <i>river</i> ( <i>P</i> -value)	$R_1^2$	$R_2^2$
<b>A: Maximal size (mm)</b>						
A1: ≤5	–	0.19	<10 <sup>-7</sup>	0.02	0.63	0.63
A2: >5–10	–	0.66	<10 <sup>-7</sup>	0.11	0.67	0.68
A3: >10–20	+	0.19	0.02	0.02	0.63	0.63
A4: >20–40	+	0.90	<10 <sup>-4</sup>	0.01	0.70	0.71
A5: >40	+	0.68	0.01	0.12	0.52	0.52
<b>B: No. of descendants per reproductive cycle</b>						
B1: ≤100	+	0.92	<10 <sup>-5</sup>	<10 <sup>-3</sup>	0.75	0.75
B2: >100–1000	+	0.43	<10 <sup>-4</sup>	<10 <sup>-3</sup>	0.69	0.69
B3: >1000–3000	–	0.69	<10 <sup>-3</sup>	0.02	0.64	0.64
B4: >3000	–	0.19	0.16	0.12	0.47	0.53
<b>C: Voltinism</b>						
C1: ≤Bivoltine	–	0.20	<10 <sup>-4</sup>	<10 <sup>-3</sup>	0.51	0.52
C2: Univoltine	–	0.34	<10 <sup>-4</sup>	<10 <sup>-2</sup>	0.74	0.74
C3: ≥Semivoltine	+	0.13	<10 <sup>-3</sup>	<10 <sup>-2</sup>	0.78	0.78
<b>D: No. of reproductive cycles per individual</b>						
D1: 1	–	0.80	<10 <sup>-6</sup>	<10 <sup>-3</sup>	0.79	0.79
D2: 2	–	0.54	<10 <sup>-5</sup>	<10 <sup>-2</sup>	0.63	0.64
D3: >2	+	0.38	<10 <sup>-3</sup>	<10 <sup>-2</sup>	0.80	0.81
<b>E: Life duration of adults (days)</b>						
E1: ≤1	+	0.26	0.96	0.01	0.68	0.69
E2: >1–10	–	0.34	0.16	0.02	0.64	0.65
E3: >10–30	–	0.56	<10 <sup>-3</sup>	<10 <sup>-2</sup>	0.64	0.64
E4: >30–365	–	0.98	0.19	0.25	0.49	0.50
E5: >365	+	0.38	<10 <sup>-5</sup>	<10 <sup>-3</sup>	0.72	0.72
<b>F: Reproductive method</b>						
F1: Single individual	–	0.35	0.29	0.20	0.57	0.57
F2: Hermaphroditism	+	0.12	<10 <sup>-3</sup>	<10 <sup>-3</sup>	0.73	0.74
F3: Male and female	–	0.19	<10 <sup>-3</sup>	<10 <sup>-3</sup>	0.73	0.74
<b>G: Parental care</b>						
G1: Bud production	–	0.45	0.04	0.15	0.65	0.66
G2: Isolated eggs	–	0.37	<10 <sup>-5</sup>	0.01	0.71	0.72
G3: Eggs in vegetation	–	0.20	0.11	0.81	0.35	0.36
G4: Cemented terrestrial eggs	–	<10 <sup>-3</sup>	<10 <sup>-7</sup>	<10 <sup>-3</sup>	0.51	0.51
G5: Cemented aquatic eggs	+	0.45	<10 <sup>-3</sup>	<10 <sup>-3</sup>	0.50	0.58
G6: Ovoviviparity	+	0.16	0.23	<10 <sup>-2</sup>	0.81	0.83
<b>H: Dispersal potential in the water (m)</b>						
H1: ≤10	+	0.63	0.06	0.08	0.61	0.63
H2: >10–100	+	0.02	<10 <sup>-6</sup>	<10 <sup>-3</sup>	0.69	0.69
H3: >100–1000	–	0.02	<10 <sup>-7</sup>	0.06	0.78	0.78
H4: >1000–10000	–	0.69	<10 <sup>-4</sup>	0.03	0.65	0.66
H5: >10000 (also by ships)	+	0.85	<10 <sup>-3</sup>	<10 <sup>-7</sup>	0.79	0.89
<b>I: Locomotion and attachment to substrate</b>						
I1: Swimmer	+	0.68	<10 <sup>-2</sup>	<10 <sup>-2</sup>	0.45	0.48
I2: Crawler	–	0.04	<10 <sup>-5</sup>	<10 <sup>-3</sup>	0.78	0.78
I3: Burrower	+	0.32	<10 <sup>-2</sup>	<10 <sup>-7</sup>	0.62	0.63
I4: Temporarily attached	+	0.06	<10 <sup>-5</sup>	<10 <sup>-2</sup>	0.55	0.67
I5: Permanently attached	+	0.76	0.02	<10 <sup>-5</sup>	0.72	0.78
<b>J: Body flexibility (°)</b>						
J1: None (≤10)	+	0.78	0.02	0.11	0.54	0.55
J2: Low (>10–45)	–	0.38	<10 <sup>-8</sup>	<10 <sup>-8</sup>	0.67	0.75

**Table 4** (Continued)

Biological trait/category	Common slope	<i>km</i> ( <i>P</i> -value)	<i>river</i> ( <i>P</i> -value)	<i>km</i> × <i>river</i> ( <i>P</i> -value)	$R_1^2$	$R_2^2$
J3: High (>45)	+	0.58	<10 <sup>-7</sup>	<10 <sup>-3</sup>	0.67	0.70
K: Body form						
K1: Streamlined	-	<10 <sup>-2</sup>	<10 <sup>-2</sup>	0.02	0.70	0.72
K2: Flattened	-	0.18	0.02	<10 <sup>-3</sup>	0.70	0.70
K3: Cylindrical	+	0.71	0.05	<10 <sup>-2</sup>	0.71	0.71
K4: Spherical	+	0.08	<10 <sup>-2</sup>	<10 <sup>-5</sup>	0.64	0.66
L: Feeding habits						
L1: Engulfer	+	0.82	<10 <sup>-2</sup>	0.01	0.59	0.61
L2: Shredder	-	0.83	<10 <sup>-3</sup>	<10 <sup>-2</sup>	0.49	0.58
L3: Scraper	-	<10 <sup>-2</sup>	<10 <sup>-7</sup>	0.05	0.76	0.76
L4: Deposit-feeder	+	0.27	0.06	0.64	0.51	0.51
L5: Filter-feeder, active	+	0.31	0.15	0.03	0.67	0.68
L6: Filter-feeder, passive	+	0.05	<10 <sup>-2</sup>	0.15	0.63	0.67
L7: Piercer	+	0.70	<10 <sup>-3</sup>	0.06	0.52	0.55
M: Food (type and size in mm)						
M1: Detritus ≤1	+	0.31	<10 <sup>-2</sup>	0.57	0.60	0.61
M2: Detritus >1-10	+	0.01	<10 <sup>-6</sup>	<10 <sup>-5</sup>	0.50	0.51
M3: Detritus >10	+	0.04	<10 <sup>-8</sup>	<10 <sup>-5</sup>	0.65	0.67
M4: Plants ≤1	-	<10 <sup>-2</sup>	<10 <sup>-10</sup>	<10 <sup>-2</sup>	0.73	0.74
M5: Plants >1-10	+	0.46	0.01	0.03	0.31	0.31
M6: Plants >10	-	0.70	<10 <sup>-2</sup>	0.02	0.46	0.73
M7: Animals ≤1	+	0.84	0.02	0.15	0.59	0.59
M8: Animals >1-10	-	0.49	0.01	0.71	0.45	0.47
M9: Animals > 10	+	0.29	<10 <sup>-6</sup>	0.25	0.59	0.59
N: Respiration technique of aquatic stages						
N1: Tegument	+	0.13	0.01	0.34	0.38	0.38
N2: Gill	-	0.65	0.17	0.61	0.41	0.41
N3: Plastron	-	0.03	<10 <sup>-10</sup>	<10 <sup>-2</sup>	0.78	0.78
N4: Aerial	-	0.53	<10 <sup>-2</sup>	<10 <sup>-2</sup>	0.32	0.33

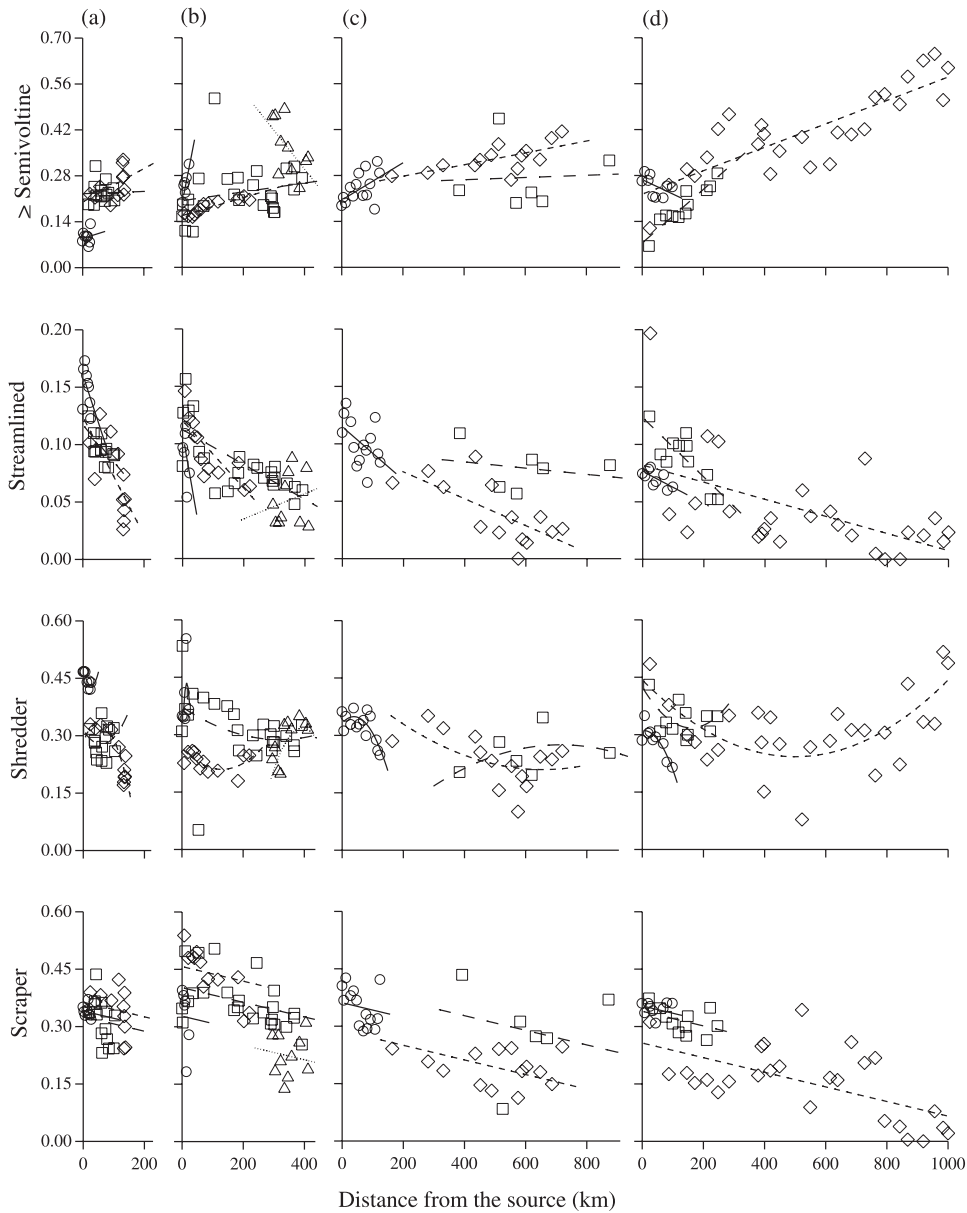
to ‘one’ in each reach), they responded almost independently along and across the rivers included here, suggesting the use of individual models for each of the categories of a given trait to predict expected trait patterns for the definition of LIRR conditions.

For many of the trait categories that were not significantly related to *km*, the partial *P*-values for *river* and *km* × *river* were highly significant (Table 4). Thus, many of the trait categories responded significantly to *km*, but the slope of these responses varied among the running waters.

Four trait categories for which the statistics in Table 4 indicated different responses were selected to illustrate the downstream trends along the 13 running waters (Fig. 3). A voltinism of ≥semivoltine (C3) tended to increase with *km* and was significantly related to *river* and *km* × *river*. Correspondingly, the increase in the relative abundance of ≥semivoltine invertebrates with *km* had a range of slopes in 11 of the running waters, whereas ≥semivoltine forms even

decreased along two (Upper Rhône and Argens; Fig. 3). A streamlined body form (K1) decreased significantly with *km*, *river* was significant, but *km* × *river* was less significant than the former. Correspondingly, the decrease of invertebrates with a streamlined body form with *km* had slopes that were more similar than in the previous example in all but one (Upper Rhône) running water (Fig. 3). The decrease with *km* of shredders (L2) was not significant and a polynomial relation (*km*, *km*<sup>2</sup>) that varied among the running waters appeared to be the most appropriate way to model this category. The resulting downstream patterns varied considerably among the running waters (Fig. 3), and the overall trend in the data was dominated by one river (Elbe). Finally, scrapers (L3) decreased significantly with *km*, *river* was significant but *km* × *river* was marginally significant (*P* = 0.05). Correspondingly, we used analysis of covariance to indicate one common slope for all 13 running waters (Fig. 3). The distribution of data, however, suggested

○ Eau d'Olle (a), Caragh (b), Hérault (c), Argens (d)    ◇ Gave de Pau (a), Wye (b), Meuse (c), Elbe (d)  
 □ Ardèche (a), Dordogne (b), Duero (c), Dunajec (d)    △ Upper Rhône (b)



**Fig. 3** Downstream trends in the relative abundance (proportion) of four selected trait categories in 165 reaches of 13 European running waters (see Fig. 1). These four trait categories represented various types of trait responses indicated by the statistics in Table 4. Regression lines were drawn beyond the range covered by the data to visualise downstream trends along shorter running waters.

that the common slope provided a poor description of the downstream trends in several of the running waters (e.g. Ardèche, Hérault and Wye; Fig. 3).

Ignoring the too impacted Lower Elbe and Lower Meuse and reaches too close to the sources (<100 km), Fig. 3 shows downstream trends interfering in the definition of trait patterns under LIRR

conditions for large European rivers. For this subset of data, the overall downstream trends were less pronounced than in smaller upstream reaches, i.e. distance from the source alone could be not sufficient to explain or predict trait patterns along large rivers spread across Europe. Therefore, in the next step of our analyses, we assessed the predictive power of

downstream, landscape, habitat and mean-model approaches.

*Step 3: comparing the predictive power of downstream, landscape and habitat models to that of the mean-model approach*

Defining the general form of regression models on trait responses using landscape features and habitat variables (see Table 3), alternative choices between simpler (no interaction among independent variables) and more complex (including significant first-order interaction terms) concerned 46 cases (trait categories in either landscape or habitat models). Using the results of the Monte Carlo simulations for all 46 cases, and regressing mean absolute error and the variance of the absolute error of predictions from simpler models ( $y$ ) on predictions from more complex models ( $x$ ) resulted in

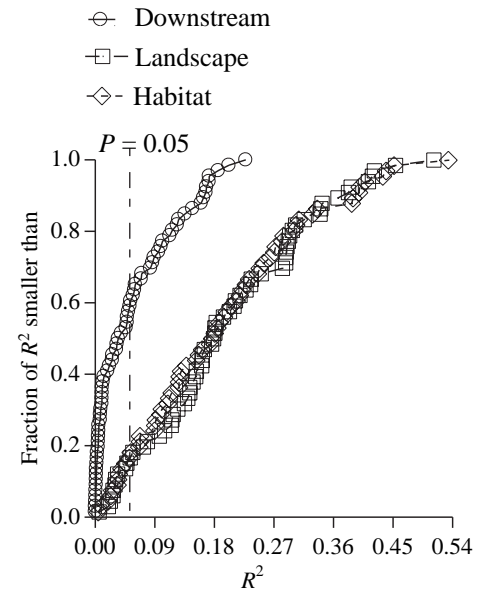
$$y = 1.06(\pm 1SE: 0.01)x \quad (R^2 = 0.996)$$

for the mean absolute error, and

$$y = 0.99(\pm 0.01)x \quad (R^2 = 0.992)$$

for the variance of the absolute error. Thus, simpler models predicted a mean absolute error that was greater than the error predicted by the more complex models, whereas both model types produced a similar variance of the error. Consequently, we ignored simpler models if more complex ones were available in subsequent parts of this section.

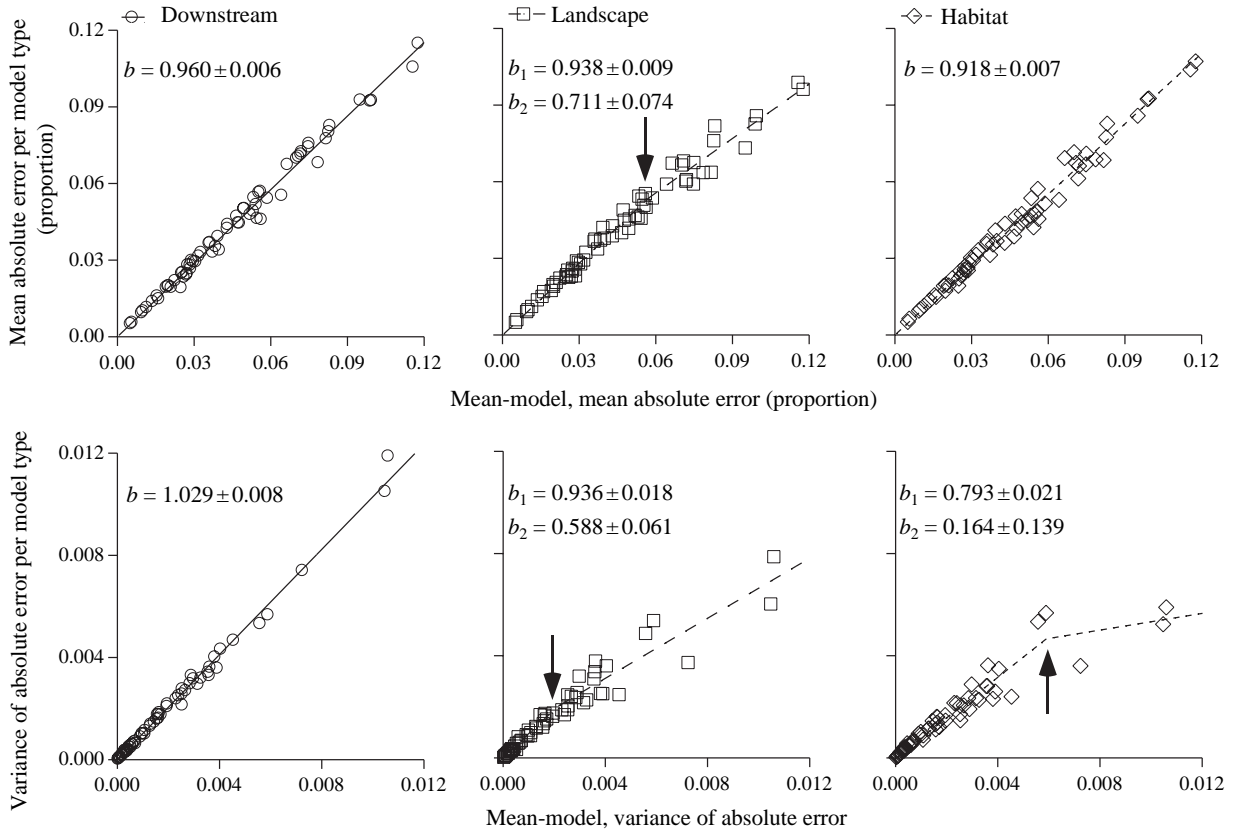
Using all 80 reaches (see Table 3), distance from the source alone significantly ( $P < 0.05$ ) explained variability (at the most approximately 25%) in the relative abundance of approximately 40% of the 66 trait categories (Fig. 4). In comparison, landscape and habitat models were significant for approximately 85% of the trait categories and explained up to approximately 50% of the variability in the relative abundance of some trait categories (Fig. 4). In the significant landscape models, 42 trait categories were significantly related to altitude, 35 to distance from the source and 18 to latitude. Thus, in the landscape models, particularly the combination of altitude with distance from the source increased the explained variability in trait responses in comparison to the downstream models. In the significant habitat models, 40 trait categories were significantly related to shear stress, 36 to mean annual air temperature and 20 to



**Fig. 4** Cumulative frequency distribution of the coefficient of determination ( $R^2$ ) for downstream, landscape and habitat models on the relative abundance of 66 trait categories using all 80 reaches (see Table 3) and the more complex general model form (including first-order interaction terms among independent variables if these were significant at  $P < 0.05$ ) of landscape and habitat models.

pH. Thus, in the habitat models, variability of trait responses was frequently explained through the combination of shear stress and temperature.

Comparison of the other three model types with the predictions from the mean-model approach indicated overall (i.e. for all 66 trait categories) that predictions from downstream models scarcely decreased the mean absolute error but scarcely increased the error variance (see slopes indicated in Fig. 5). In contrast, landscape and habitat models improved the predictions on the relative abundance of trait categories in comparison with the mean-model approach (all slopes in Fig. 5 significantly  $< 1$ ). For trait categories with a mean absolute error approximately  $< 0.06$  in predictions from the mean-model approach, predictions from landscape and habitat models similarly reduced the error (slopes: 0.94 versus 0.92). For errors approximately  $> 0.06$  in predictions from the mean-model approach, however, landscape models provided a greater error reduction than habitat models (0.71 versus 0.92). Habitat models (0.79) reduced the error variance more than landscape models (0.94) for categories with  $x$ -values approximately  $< 0.002$  in Fig. 5. Between approximately 0.002 and 0.006 along



**Fig. 5** Comparison of the predictions on the relative abundance of 66 trait categories (mean absolute error of the predictions and the variance of this error obtained by Monte Carlo simulations using models and the 80 reaches considered in Fig. 4) from downstream, landscape and habitat models with those from the mean-model approach. In these comparisons, we used simple or, if significant ( $P < 0.05$ ), piecewise linear regression and indicated regression slopes  $\pm 1$  SE and the discontinuity in piecewise regressions by arrows (there, the slope changed from  $b_1$  to  $b_2$ ).

the  $x$ -axes in Fig. 5, landscape models were better (0.59) than habitat models (0.79). For  $x$ -values approximately  $>0.006$ , predictions from habitat models produced a gentler slope (0.16) than those from landscape models (0.59), but the data scatter (and thus the SE of the slope) was greater (Fig. 5). Thus, the overall reduction of the error variance was similar for both landscape and habitat models, whereas previous considerations of the absolute error suggested that the predictive power of the landscape models was greater than that of the habitat models. Accordingly, we assessed the performance of landscape models in the next step of our analyses.

*Step 4: comparing the predictive power of landscape models to that of the mean-model approach using LIRRs*

Defining the general form of models on trait responses using landscape features and LIRRs, alternative

choices between simpler (no interaction among independent variables) and more complex (including significant first-order interaction terms) concerned 14 trait categories. Using the results of the Monte Carlo simulations for all 14 categories and regressing mean absolute error and the variance of the absolute error of predictions from simpler models ( $y$ ) on predictions from more complex models ( $x$ ) resulted in

$$y = 1.01(\pm 1\text{SE}: 0.01)x \quad (R^2 = 0.998)$$

for the mean absolute error, and

$$y = 0.91(\pm 0.03)x \quad (R^2 = 0.983)$$

for the variance of the absolute error. Thus, predictions from both model types had a similar mean absolute error, whereas predictions from simpler models had a variance of the absolute error that was smaller than that obtained by predictions from more complex models. Correspondingly, we ignored more

complex models if these were available and focussed on simpler ones in subsequent parts of this section.

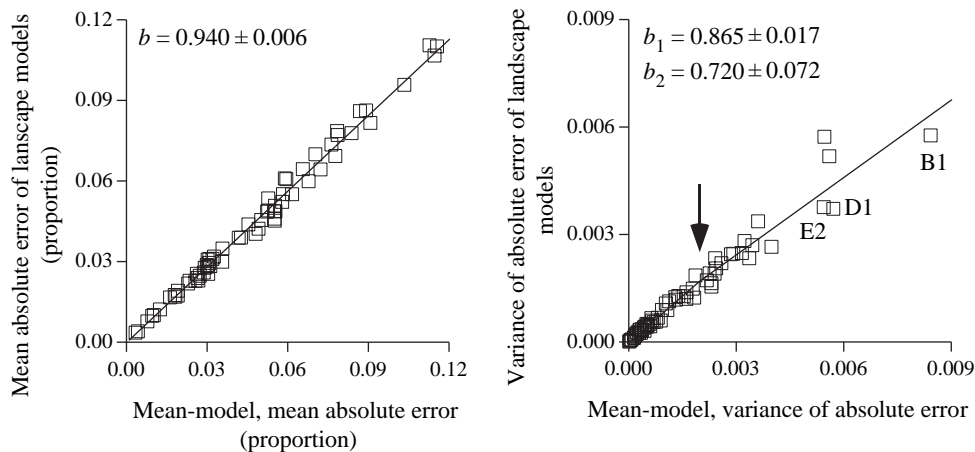
Using all 68 LIRRs, the simpler landscape models were significant for approximately 85% of the 66 trait categories, but explained less variability (all  $R^2 < 0.38$ , approximately 85% of the categories with  $R^2 < 0.27$ ) than the landscape models in the previous step (see Fig. 4). This reduction in the descriptive power of the landscape models in comparison with the previous step had two causes. First, the simpler landscape models (no interaction terms included) used in this step would *a priori* explain less variability in the trait responses than the more complex models (including significant interaction terms) used in the previous step. Second, in contrast to the previous step, we eliminated reaches with a width  $< 40$  m from the analyses in this step, i.e. the more distinct trait responses occurring closer to the source (see step 2) were excluded here. In the significant ( $P < 0.05$ ) simpler landscape models, 30 trait categories were significantly related to altitude, 37 to distance from the source and 22 to latitude. Thus, in contrast to the previous step, distance from the source was more frequently significant than altitude, despite the fact that the longitudinal trait responses occurring closer to the source were excluded in this step.

Overall (i.e. for all 66 trait categories), simpler landscape models using LIRRs improved the predictions on the relative abundance of trait categories in

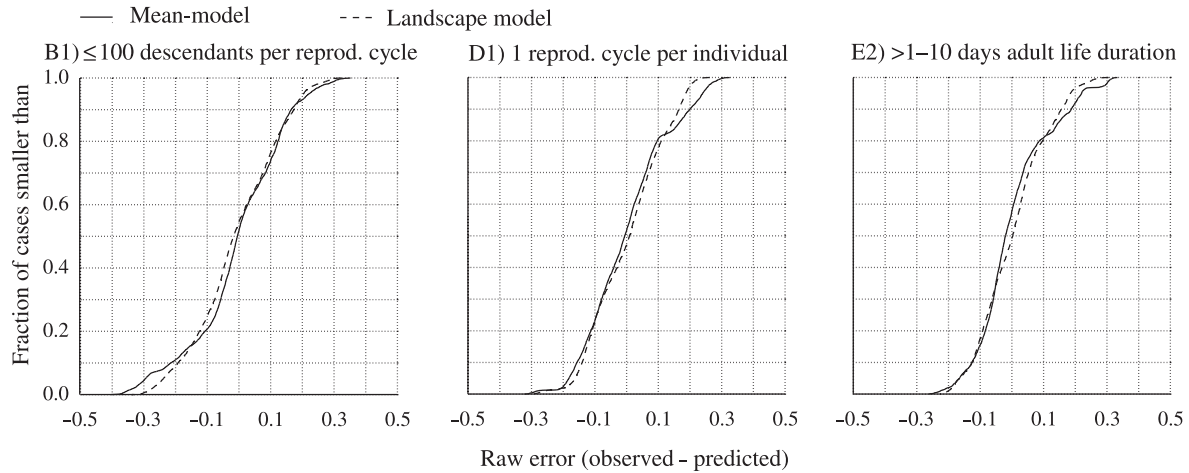
comparison with the mean-model approach (all slopes in Fig. 6 significantly  $< 1$ ). This improvement concerned primarily the error variance, which was particularly reduced for three trait categories (B1:  $\leq 100$  descendants per reproductive cycle, D1: one reproductive cycle per individual and E2: life duration of adults  $> 1-10$  days; see Fig. 6). We checked the raw error of the predictions on these three trait categories obtained by the Monte Carlo simulations by comparing the landscape models and the mean-model approach. A simple illustration of these errors indicated that the mean-model approach produced little more error than the landscape models (see the tails of the distributions in Fig. 7). With such marginal differences in the three trait categories having the greatest reduction in error variance if predicted by landscape models, it would be an inappropriate expense to use such models to predict expected LIRR trait patterns in a biomonitoring routine. Therefore, we defined the expected LIRR conditions using simpler descriptions of frequency distributions of traits and trait categories in the next step.

#### Step 5: defining trait patterns for LIRR conditions

Using the average of the relative abundance of the trait categories from the 68 LIRRs as expected values (see Appendix), frequency distributions of the deviations from these expected values of the individual



**Fig. 6** Comparison of the predictions on the relative abundance of 66 trait categories (mean absolute error of the predictions and the variance of this error obtained by Monte Carlo simulations using 68 LIRRs described in Table 3) from simpler landscape models (ignoring first-order interaction terms among independent variables if these were significant at  $P < 0.05$ ) with those from the mean-model approach. The arrow indicates the discontinuity in the piecewise regression (see Fig. 5 for further details). The landscape models provided the most important reduction in the error variance for three trait categories further illustrated in Fig. 7 (B1:  $\leq 100$  descendants per reproductive cycle; D1: one reproductive cycle per individual; and E2: life duration of adults  $> 1-10$  days).



**Fig. 7** Cumulative frequency distribution of the raw error of predictions from landscape models and the mean-model approach on the three trait categories for which landscape models provided the most important reduction in the error variance (see Fig. 6).

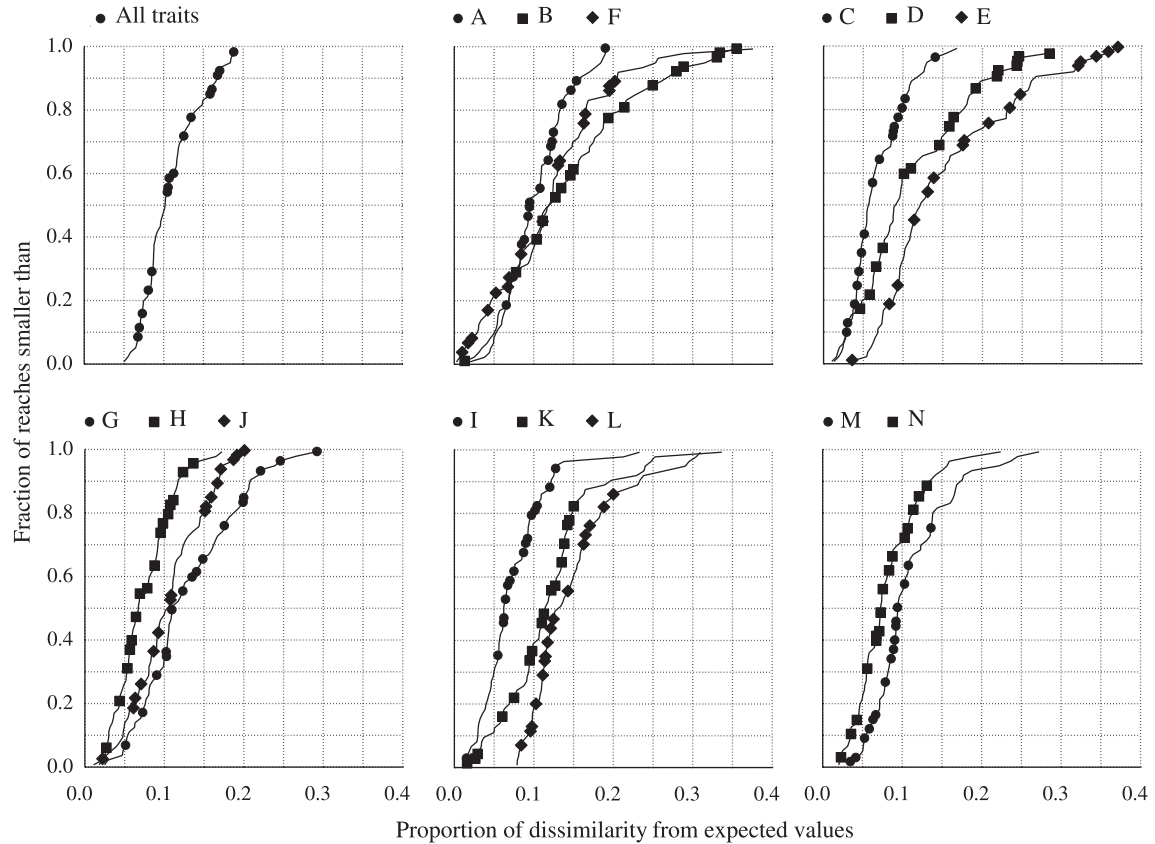
LIRRs provided the simplest definition of our LIRR conditions. Thus, environmental health of other large European river reaches could be assessed by simple comparisons of observations from these other reaches with the distributions obtained for our LIRRs.

The proportion of dissimilarity of the overall trait profiles (across all 66 categories) of the 68 LIRRs from the expected trait profile ranged between approximately 0.05 and 0.2 (Fig. 8), i.e. across all categories, individual LIRRs differed at the most by approximately 20% from the mean trait profile of all 68 LIRRs. The upper end of the distribution (dissimilarity near 0.2) had only a short tail, and the 16 most natural reaches were spread across the distribution. These patterns suggested that the upper limit of the distribution could serve to discriminate LIRR conditions from human-impacted conditions for the overall trait profile (i.e. impact would be indicated through one side of the distribution). For several individual traits, however, the upper end of the distribution had a relatively long tail and included none of the most natural reaches (e.g. F: reproductive method, I: locomotion and attachment to substrate, K: body form, L: feeding habits, M: food and N: respiration technique of aquatic stages; see Fig. 8). For all these traits, cutting off the upper 10% of the distribution would eliminate a considerable part of these longer tails (see Fig. 8). For consistency, we cut off the upper 10% of all trait distributions, which provided a more conservative alternative to the use of the upper limit of the entire distributions in the assessment of test-LIRR data in step 6.

Although a given type of human impact should shift individual trait categories in a predictable direction (see Table 1), multiple types of human impact could shift individual trait categories in either direction. Therefore, we summarised the frequency distributions of raw observed – expected values for the 68 LIRRs and each trait category by indicating the lower and upper limit of each distribution (Appendix). To obtain a more conservative description of these distributions, we also indicated the lower and upper limit that enveloped 90% of the LIRRs in these distributions (see Appendix for cut off rules). Through comparison with these LIRR frequency distributions, the last step of our analyses had to assess how test-LIRRs would be assigned to LIRR conditions.

#### *Step 6: assessing test-LIRRs using the criteria defined in step 5*

Using the overall trait profile (i.e. all traits together), 53 of the 57 test-LIRRs were correctly assigned to LIRR conditions if the latter were defined through the entire range of the LIRR frequency distribution obtained in step 5, and 49 of the test-LIRRs fell within the 90%-range of the LIRR distribution. Among the eight incorrectly assigned test-LIRRs, six had a genus richness  $\leq 10$  and/or a proportion of genera with alien species  $\geq 0.1$ , i.e. most of the incorrectly assigned reaches did not match the more conservative criteria used to select the LIRR data (see Table 3). Thus, applying these more conservative criteria to a



**Fig. 8** Cumulative frequency distribution of the dissimilarity of all 68 LIRRs in comparison with expected values of the trait categories (see Appendix) for the overall trait profile (all traits) and each individual trait (see Table 4 for trait labels). The positions of the 16 most natural LIRRs (no cargo-ship traffic; organic and heavy metal pollution in class 1: very good quality) are shown with solid symbols.

test-LIRR in subsequent, independent validations of impact assessment rules should assign a greater proportion of test-LIRRs to LIRR conditions.

Using each trait individually, a proportion of 0.87–1 of the 57 test-LIRRs were correctly assigned to LIRR conditions if the latter were defined through the entire range of the LIRR frequency distributions obtained in step 5 (Fig. 9). Not surprisingly, the best assignment was achieved for traits having a relatively long upper tail in the LIRR distribution (compare Figs 8 & 9), such as body form (K), reproductive method (F), respiration technique of aquatic stages (N), locomotion and substrate attachment (I), life duration of adults (E) and number of descendants per reproductive cycle (B). These traits, however, did not systematically change their ranks when we used the 90%-range of the values observed for the LIRRs. For example, most of the test-LIRRs were consistently assigned correctly with trait B (number of descend-

ants per reproductive cycle), whereas the proportion of correctly assigned test-LIRRs with trait N (respiration technique of aquatic stages) dropped to 0.7. These patterns suggested that the upper tail in the distribution of the test-LIRRs had either less (B) or more (N) reaches than that of the LIRRs. For the 90%-range enveloped by the LIRR distributions, eight traits correctly assigned a proportion of >0.8, and one trait of >0.9 of the test-LIRRs (Fig. 9).

Using all trait categories individually, a proportion >0.9 of the test-LIRRs was correctly assigned by 57 of the 66 categories to the entire range of the LIRR frequency distributions (Fig. 9). For the 90%-range enveloped by the LIRR distributions, 42 trait categories correctly assigned a proportion of >0.8, and 12 categories of >0.9 of the test-LIRRs (Fig. 9). In this context, trait categories that failed to assign a substantial proportion of the test-LIRRs to the LIRR conditions presumably will be less useful in future

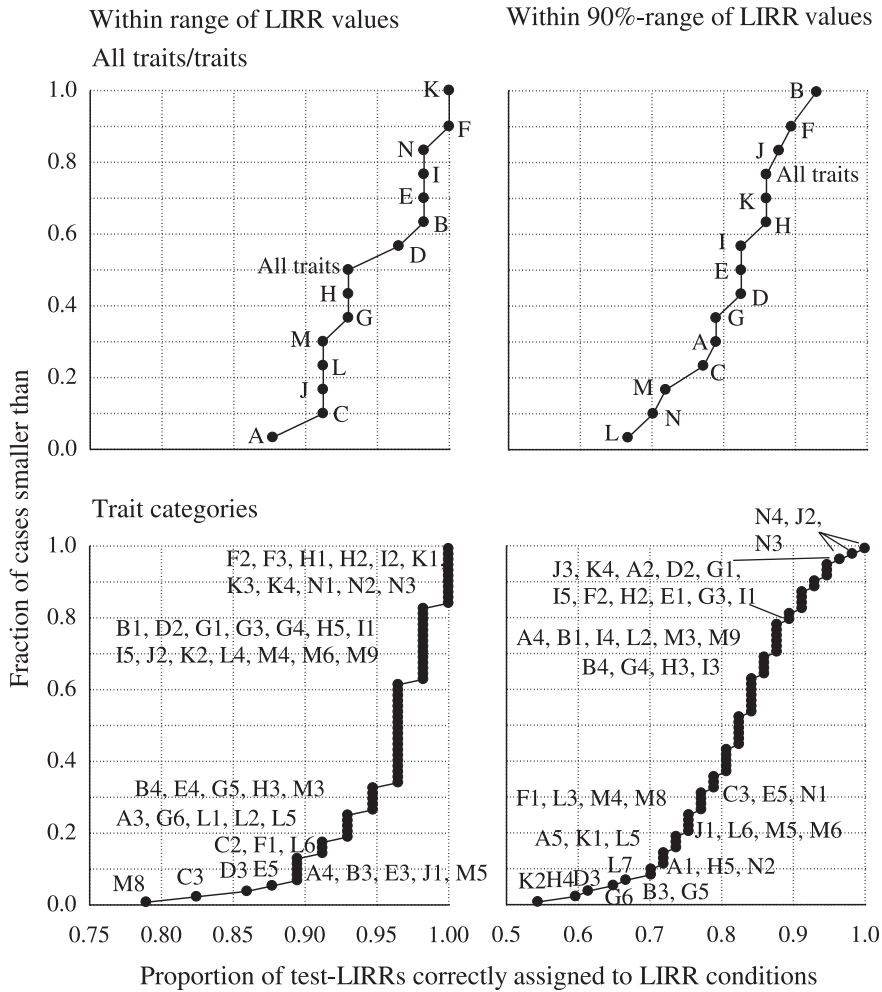


Fig. 9 Cumulative frequency distribution of the proportion of correctly assigned test-LIRRs ( $n = 57$ ) using the two alternative criteria defining the LIRR conditions in step 5 for the overall trait profile (all traits), each individual trait and each individual trait category (see Table 4 for trait and trait category labels).

biomonitoring than categories that provided a substantial proportion of correct assignments (identified by the labels added to Fig. 9). Corresponding to the patterns observed in step 2, individual categories within a given trait provided almost independent results in this step, as some categories of a given trait provided fewer and others more correct assignments of test-LIRRs.

## Discussion

### Downstream patterns

Assessing downstream trends in environmental variables, reach width was positively related to distance

from the source, whereas altitude, reach slope and shear stress were negatively related to distance from the source (Fig. 2). These correlations reflected obvious downstream trends occurring in running waters, with the exception of the shear stress correlation. Typically, the flow forces acting on the bed of running waters increase below the spring source in headwater reaches before they start to drop along downstream reaches (Statzner, Gore & Resh, 1988; Brummer & Montgomery, 2003). With the dominance of larger river reaches in our data, however, these physical changes in headwaters were underrepresented.

Similarly to these environmental variables, many of the trait categories responded significantly to distance from the source ( $km$ ), but the slope of these responses

varied among the running waters. In contrast, Lamouroux, Dolédec & Gayraud (2004) reported often similar slopes for two French river basins (Loire and Rhône) in relations between the proportion of biological invertebrate trait categories and physical reach characteristics (particularly the reach Froude number). These latter results were obtained studying smaller headwater reaches (distance from the source ranging between 4 and 30 km). In comparison, our data covered a larger range of downstream gradients, geographic extension and running water types, which explains why we found more interactions between a variable (*km*) describing physical reach conditions (e.g. shear stress, see Fig. 2) and river systems (the categorical variable *river*) than Lamouroux *et al.* (2004).

Despite the differences among individual running waters illustrated by Fig. 3 (e.g. between Upper Rhône and Meuse), the interaction  $km \times river$  was often significant in our data. Thus, the common slope along all running waters (although often not significant itself) indicated dominant overall downstream trends for many of the trait categories (see Table 4). Among these downstream trends in the trait categories, several could be reasonably associated with longitudinally decreasing shear stress (see Fig. 2), and thus decreasing coarseness of the bottom substrate (or increasing deposition of finer material). Corresponding to our predictions on trait category responses to physical habitat changes (flow in- or decrease, siltation, see Table 1), streamlined body forms and small-sized invertebrates increased with decreasing *km* (i.e. increasing shear stress) (Table 4). These traits constitute alternatives to reduce drag at elevated near-bed flow forces (Statzner, 1988; Vogel, 1994). Spherical, large-sized, swimming forms, burrowers in fine sediments and perhaps ( $P > 0.05$  for both, *km* and  $km \times river$ ) deposit-feeders of fine detritus increased with *km*, also correspondingly to our predictions. In contrast to our predictions (Table 1), however, invertebrates that were temporarily or permanently attached to the bottom substrate increased with *km*.

Discharge variability in streams is also related to *km*. Rapidity and height of flood responses to rainfall in running water reaches decrease with increasing catchment size and decreasing reach slope (Gordon, McMahon & Finlayson, 1992). Thus, with decreasing *km*, the harshness of flood disturbances increases.

Corresponding to predictions by Townsend & Hildrew (1994) on traits conferring resilience or resistance to organisms along an increasing disturbance gradient, lower *km*-values were associated with smaller maximal size, more descendants per reproductive cycle, shorter life cycles, fewer reproductive cycles per individual, less parental care (such as ovoviviparity) and more streamlined or flattened body forms (Table 4). In contrast to predictions of Townsend & Hildrew (1994), however, higher mobility or body flexibility were not associated with lower *km*-values (Table 4).

Finally, trends in our data corresponded to predictions on downstream trends by Vannote *et al.* (1980), e.g. that shredders decreased (although, in our case, feeders of coarse detritus increased), and forms feeding on finer detritus increased (Table 4). In contrast to predictions of Vannote *et al.* (1980), however, scrapers of small plants did not change in a hump-shaped pattern, but monotonously decreased with *km* (note the similarity of  $R^2$ -values of the linear and polynomial relation for L3 and M4 in Table 4).

Such equivocal support of predictions on trait responses relates to trade-offs that have been repeatedly acknowledged by ecologists since the 1920s: similar habitats may have taxa with different combinations of biological traits, as there are different solutions to living under given environmental constraints (Statzner, Hildrew & Resh, 2001b). As a consequence, various studies examining the patterns of multiple traits of lotic invertebrates in the context of environmental constraints reported similar responses in some and different responses in other traits (e.g. Usseglio-Polatera, 1994; Richards *et al.*, 1997; Townsend *et al.*, 1997; Snook & Milner, 2002; Lamouroux *et al.*, 2004; Mérigoux & Dolédec, 2004; Statzner *et al.*, 2004). This variation would interfere in the definition of trait patterns at reference conditions beyond that for European LIRRs.

Ignoring the patterns observed in Lower Elbe and Lower Meuse, which were more impacted by human activities than the other reaches included here, Fig. 3 provided the first illustration of relatively natural downstream trends in trait categories in very different European running water types. Our findings suggested that the relative abundance of trait categories changes gradually along running waters, in contrast to the distinct changes in the taxonomic composition of lotic invertebrate communities that are associated

with sudden downstream changes in physical conditions (e.g. from a steeply sloped mountain channel to a gently sloped alluvial channel, see Statzner & Higler, 1986; Grubaugh, Wallace & Houston, 1996).

#### *Downstream, landscape, and habitat models versus the mean-model approach*

Downstream, landscape features and finally habitat attributes should have increasing influence on lotic invertebrates (see Richards *et al.*, 1997; Hawkins *et al.*, 2000; Johnson *et al.*, 2004). Thus, we expected that the accuracy of these three model types to predict variation in the trait categories would increase in the same order (i.e. downstream, landscape and habitat models). Alternatively, we expected that in comparison with the relatively low variation of trait patterns observed in smaller European streams (Statzner *et al.*, 2001a, 2004), the relative abundance of trait categories would vary even less among the LIRRs of large European rivers. This latter expectation addressed the fact that changes in taxonomic composition are more pronounced in headwaters than in downstream sections (Statzner & Higler, 1986), and that the mechanisms controlling these patterns differ more among headwaters than larger rivers (Gomi, Sidle & Richardson, 2002). Correspondingly, the data plotted in Fig. 3 suggested that changes in the relative abundance of trait categories could be greater along headwater streams than along large rivers. Thus, focussing on large rivers would result in less variation in trait patterns, so that simple descriptions of the deviations of traits from a mean trait profile (i.e. the mean-model approach) would be sufficient to define LIRR conditions for large European rivers.

Although distance from the source was an integrative variable capturing some (but not all) landscape features and habitat variables (Fig. 2), and although the relative abundance of many trait categories changed along running waters (Table 4), particularly the intensity (and sometimes the direction of the slope) of these changes differed among the latter. Therefore, corresponding to our expectation, downstream models provided the worst predictions of expected invertebrate trait patterns when compared with landscape or habitat models. In contrast to our expectation, however, landscape models provided better predictions of expected invertebrate traits patterns than habitat models. This latter result was

perhaps caused by differences in the data quality, as information on the variables used in the landscape models was easier to obtain and more reliable than information on the variables used in the habitat models. Typically, invertebrate responses are determined predominantly by conditions prevailing at the local scale in natural running waters (e.g. Johnson *et al.*, 2004).

In a purely statistical context, landscape models significantly reduced the mean error and the variance of this error in predictions of expected trait patterns compared with the alternative mean-model approach. However, the amount of the reduction of the mean error and the error variance achieved by these landscape models was negligible. One could speculate that inaccurate descriptions of the landscape variables could have reduced the predictive power of the landscape models. However, reliable information on these variables (distance from the source, altitude, latitude) was easily extracted from our data sources. Thus, the trait patterns varied so little across our LIRRs that we could use simple descriptions of frequency distributions to define LIRR conditions. This relatively high trait stability observed across large European LIRRs confirmed results from smaller natural or almost natural European streams (Statzner *et al.*, 2001a, 2004). Currently, there are no convincing explanations for this trait stability across Europe; perhaps, the latter reflects the uniform action of stream-system specific, local physical factors that affect many traits in a similar way across large geographic areas (Statzner *et al.*, 2004). Whatever the cause of this trait stability, it will obviously facilitate the definition of reference conditions for European running waters.

#### *Assignment of test-LIRRs to LIRR conditions*

The correct assignment of cases to predefined conditions (e.g. reference versus impacted) is the ultimate criterion to assess the performance of a biomonitoring approach. Focussing our analyses on least impacted conditions, we evaluated this performance by assigning test-LIRRs to independently defined LIRR conditions. Depending on the definition of the latter, many traits and their categories correctly assigned >90% of the test-LIRRs to the entire range of the LIRR frequency distributions, or >80% of them to the 90% range enveloped by the LIRR frequency distributions.

For running waters, reference conditions, and benthic macroinvertebrates, such performances in the correct assignment of cases to predefined conditions have been assessed for a variety of biomonitoring approaches. Multiple studies classified reference reaches with similar taxonomic characteristics of the invertebrate communities into groups and used environmental predictors (local, landscape and/or regional) to develop predictive models for such reference groupings. The correctly predicted assignments of reference reaches to such groupings (often based on internal validation tests) depended on factors such as the number of reference reaches used to build the predictive model, the number of seasons considered or the environmental predictors included in the model (Reynoldson & Wright, 2000; Reece *et al.*, 2001). Although these factors caused some variation, and natural variation within such groupings interferes with such assignments, predictions from such models often achieved approximately 70% of correct reach assignments to reference groupings in Europe (Moss *et al.*, 1999; Alba-Tercedor & Pujante, 2000; Wright, 2000; Heino *et al.*, 2002), North America (Reynoldson, Rosenberg & Resh, 2001), and Australia/New Zealand (Davies, Norris & Thoms, 2000; Simpson & Norris, 2000; Joy & Death, 2003). Using selected biological indices developed to indicate acidification or eutrophication in Swedish streams, Sandin & Johnson (2000) were able to assign 75–95% of the reference reaches correctly to a reference (defined by water chemistry).

In comparison with the performance of these biomonitoring approaches, our trait approach assigned thus as many or even more of our test-LIRRs to LIRR conditions, and did so within a framework that required no regionalisation of a large geographic area, no modelling of expected values using environmental information and no standardised invertebrate sampling. Two of the possible explanations for this performance have been already mentioned before: first, the environmental conditions varied perhaps too little among our LIRRs to have major effects on trait patterns, and second, uniform action of stream-system specific, local physical factors perhaps affect many traits in a similar way across large geographic areas. In addition, reducing the taxonomic diversity of invertebrates (in our LIRRs: approximately 250 genera) to 66 trait categories, each of these typically shared by more than one genus, reduced the overall

variance in our large-scale data set (Charvet *et al.*, 2000), which should have facilitated the assignment of cases to predefined conditions in comparison with taxonomy-based biomonitoring approaches.

### Outlook

For large European rivers, it has been shown that invertebrate traits: (i) discriminated overall human impact on river reaches (Dolédéc *et al.*, 1999), (ii) significantly discriminated various levels of human impacts using relatively simple taxonomic information (presence–absence of genera) (Gayraud *et al.*, 2003), (iii) can be accurately described by few sample replicates in space or time (Bady *et al.*, 2005) and (iv) assigned a substantial proportion of test-LIRRs to LIRR conditions (this study). The ultimate question, however, is how much the stable trait patterns observed under reference conditions will change under human impact. Consequently, the next step in the development of a biomonitoring tool for large European rivers using invertebrate traits has to assess deviations of large river reaches with different types of human impacts from the here defined LIRR conditions. If such deviations exist and adequately discriminate a given type of human impact, it would require a collective effort of European ecologists working on large rivers to develop this invertebrate trait approach towards an operational biomonitoring tool. From experience we gained assembling information in our databases, this task could be achieved by collectively sharing existing European data on the: (i) biological traits of invertebrates, (ii) taxonomic structure of invertebrate communities and (iii) environmental conditions of the large river reaches inhabited by these communities.

### Acknowledgments

We thank all researchers and editors who published results in a format that we could use them for our databases, the many colleagues that provided personal information to fill data gaps (all mentioned in one of our previous papers on the subject), and Nicolas Lamouroux, two anonymous referees and Richard K. Johnson for comments that improved the manuscript. This work was supported by the Bundesministerium für Bildung und Forschung (Germany) (project FKZ0330029 'Biologische Merkmale

von Flußwirbellosen als Basis einer überregionalen Bewertung ökologischer Funktionsfähigkeit’).

## References

- Alba-Tercedor J. & Pujante A.M. (2000) Running-water biomonitoring in Spain: opportunities for a predictive approach. In: *Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques* (Eds J.F. Wright, D.W. Sutcliffe & M.T. Furse), pp. 207–216. Freshwater Biological Association, Ambleside.
- Anonymous (1980) *Etude Hydrobiologique de la Dordogne*. Cemagref Institute, Lyon.
- Anonymous (1998) Special report No 3/98 concerning the implementation by the commission of EU policy and action as regards water pollution accompanied by the replies of the commission. *Official Journal of the European Communities*, **C191**, 2–44.
- Anonymous (1999a) Council decision of 25 January 1999 adopting a specific programme for research, technological development and demonstration on energy, environment and sustainable development (1998–2002). *Official Journal of the European Communities*, **L064**, 58–77.
- Anonymous (1999b) *Gewässergüteatlas der Bundesrepublik Deutschland*. Länderarbeitsgemeinschaft Wasser, Berlin.
- Anonymous (2000) Système d’évaluation de la qualité de l’eau des cours d’eau. *Les Etudes des Agences de l’Eau*, **64**, 1–21.
- Anonymous (2001) *Wasserwirtschaft in Deutschland*. Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, Berlin.
- Anonymous (2002) *Statistique Annuelle de la Navigation Intérieure 2001*. Ministère de l’équipement, des transports et du tourisme, Paris.
- Bady P., Dolédec S., Fesl C., Gayraud S., Bacchi M. & Schöll F. (2005) Use of invertebrate traits for the biomonitoring of European large rivers: the effects of sampling effort on genus richness and functional diversity. *Freshwater Biology*, **50**, 159–173.
- Benke A.C. (1993) Concepts and patterns of invertebrate production in running waters. *Verhandlungen der Internationalen Vereinigung für Limnologie*, **25**, 15–38.
- Braun M. (2001) *Zahlentafeln der Physikalisch-chemischen Untersuchungen des Rheinwassers und des Schwebstoffs 1998*. Internationale Kommission zum Schutz des Rheins, Koblenz.
- Bremner J., Rogers S.I. & Frid C.L.J. (2003) Assessing functional diversity in marine benthic ecosystems: a comparison of approaches. *Marine Ecology Progress Series*, **254**, 11–25.
- Brummer C.J. & Montgomery D.R. (2003) Downstream coarsening in headwater channels. *Water Resources Research*, **39**, 1294, doi:10.1029/2003WR001981.
- Charvet S., Statzner B., Usseglio-Polatera P. & Dumont B. (2000) Traits of benthic macroinvertebrates in semi-natural French streams: an initial application to biomonitoring in Europe. *Freshwater Biology*, **43**, 277–296.
- Chessman B.C. (1999) Predicting the macroinvertebrate faunas of rivers by multiple regression of biological and environmental differences. *Freshwater Biology*, **41**, 747–757.
- Chevenet F., Dolédec S. & Chessel D. (1994) A fuzzy coding approach for the analysis of long-term ecological data. *Freshwater Biology*, **31**, 295–309.
- Davies N.M., Norris R.H. & Thoms M.C. (2000) Prediction and assessment of local stream habitat features using large-scale catchment characteristics. *Freshwater Biology*, **45**, 343–369.
- Dolédec S. & Chessel D. (1994) Co-inertia analysis: an alternative method for studying species-environment relationships. *Freshwater Biology*, **31**, 277–294.
- Dolédec S., Statzner B. & Bournaud M. (1999) Species traits for future biomonitoring across ecoregions: patterns along a human-impacted river. *Freshwater Biology*, **42**, 737–758.
- Gay C. (1982) Les communautés benthiques d’un torrent des Alpes française: l’Eau d’Olle (Isère). *Travaux du Laboratoire d’Hydrobiologie Grenoble*, **71/72/73**, 7–31.
- Gayraud S., Statzner B., Bady P., Haybach A., Schöll F., Usseglio-Polatera P. & Bacchi M. (2003) Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of alternative metrics. *Freshwater Biology*, **48**, 2045–2064.
- Ghilarov A.M. (2000) Ecosystem functioning and intrinsic value of biodiversity. *Oikos*, **90**, 408–412.
- Giudicelli J., Dia A. & Legier P. (1980) Etude hydrobiologique d’une rivière de région méditerranéenne, l’Argens (Var, France). *Bijdragen tot de Dierkunde*, **50**, 303–341.
- Gomi T., Sidle R.C. & Richardson J.S. (2002) Understanding processes and downstream linkages of headwater systems. *BioScience*, **52**, 905–916.
- Gordon N.D., McMahon T.A. & Finlayson B.L. (1992) *Stream Hydrology*. Wiley & Sons, Chichester.
- Gozlan R.E., Mastrorillo S., Copp G.H. & Lek S. (1999) Predicting the structure and diversity of young-of-the-year fish assemblages in large rivers. *Freshwater Biology*, **41**, 809–820.
- Graham M.H. (2003) Confronting multicollinearity in ecological multiple regression. *Ecology*, **84**, 2809–2815.
- Grubaugh J.W., Wallace J.B. & Houston E.S. (1996) Longitudinal changes of macroinvertebrate commu-

- nities along an Appalachian stream continuum. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 896–909.
- Hawkins C.P., Norris R.H., Gerritsen J., Hughes R.M., Jackson S.K., Johnson R.K. & Stevenson R.J. (2000) Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. *Journal of the North American Benthological Society*, **19**, 541–556.
- Heino J., Muotka T., Paavola R., Haemaeläinen H. & Koskenniemi E. (2002) Correspondence between regional delineations and spatial patterns in macroinvertebrate assemblages of boreal headwater streams. *Journal of the North American Benthological Society*, **21**, 397–413.
- Hill M.O., Roy D.B., Mountford J.O. & Bunce R.G.H. (2000) Extending Ellenberg's indicator values to a new area: an algorithmic approach. *Journal of Applied Ecology*, **37**, 3–15.
- Johnson R.K., Goedkoop W. & Sandin L. (2004) Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes. *Freshwater Biology*, **49**, 1179–1194.
- Joy M.K. & Death R.G. (2003) Biological assessments of rivers in the Manawatu-Wanganui region of New Zealand using a predictive macroinvertebrate model. *New Zealand Journal of Marine and Freshwater Research*, **37**, 367–379.
- Karr J.R. & Chu E.W. (1999) *Restoring Life in Running Waters*. Island Press, Washington DC.
- Lamouroux N., Dolédec S. & Gayraud S. (2004) Biological traits of stream macroinvertebrate communities: effects of microhabitat, reach and basin filters. *Journal of the North American Benthological Society*, **23**, 449–466.
- Leopold L.B., Wolman M.G. & Miller J.P. (1964) *Fluvial Processes in Geomorphology*. Freeman, San Francisco.
- Löffler J., Wolff S. & Bergemann M. (1999) *Wassergütedaten der Elbe – Zahlentafel 1997*. Arbeitsgemeinschaft für die Reinhaltung der Elbe, Hamburg.
- Manly B.J.F. (1991) *Randomization and Monte Carlo Methods in Biology*. Chapman & Hall, London.
- McGrady-Steed J., Harris P.M. & Morin P.J. (1997) Biodiversity regulates ecosystem predictability. *Nature*, **390**, 162–165.
- Mérigoux S. & Dolédec S. (2004) Hydraulic requirements of stream communities: a case study on invertebrates. *Freshwater Biology*, **49**, 600–631.
- Moss D., Wright J.F., Furse M.T. & Clarke R.T. (1999) A comparison of alternative techniques for prediction of the fauna of running-water sites in Great Britain. *Freshwater Biology*, **41**, 167–181.
- Naeem S. (2002) Ecosystem consequences of biodiversity loss: the evolution of a paradigm. *Ecology*, **83**, 1537–1552.
- New M.G., Hulme M. & Jones P.D. (1999) Representing 20th century space-time climate variability. I. Development of a 1961–1990 mean monthly terrestrial climatology. *Journal of Climate*, **12**, 829–856.
- Oberdorff T., Pont D., Hugueny B. & Porcher J.-P. (2002) Development and validation of a fish-based index for the assessment of “river health” in France. *Freshwater Biology*, **47**, 1720–1734.
- Ormerod S.J., Pienkowski M.W. & Watkinson A.R. (1999) Communicating the value of ecology. *Journal of Applied Ecology*, **36**, 847–855.
- Perrin J.-F. (1978) Signification Ecologique des Peuplements Benthiques du Haute-Rhône Français. PhD Thesis, Université Lyon 1, Villeurbanne.
- Reece P.F., Reynoldson T.B., Richardson J.S. & Rosenberg D.M. (2001) Implications of seasonal variation for biomonitoring with predictive models in the Fraser River catchment, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 1411–1418.
- Reynoldson T.B. & Wright J.F. (2000) The reference condition: problems and solutions. In: *Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques* (Eds J.F. Wright, D.W. Sutcliffe & M.T. Furse), pp. 293–303. Freshwater Biological Association, Ambleside.
- Reynoldson T.B., Rosenberg D.M. & Resh V.H. (2001) Comparison of models predicting invertebrate assemblages for biomonitoring in the Fraser River catchment, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 1395–1410.
- Richards C., Haro R.J., Johnson L.B. & Host G.E. (1997) Catchment and reach-scale properties as indicators of macroinvertebrate species traits. *Freshwater Biology*, **37**, 219–230.
- Rosenberg D.M. & Resh V.H. (Eds) (1993) *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman & Hall, London.
- Sandin L. & Johnson R.K. (2000) The statistical power of selected indicator metrics using macroinvertebrates for assessing acidification and eutrophication of running waters. *Hydrobiologia*, **422/423**, 233–243.
- Simpson J.C. & Norris R.H. (2000) Biological assessment of river quality: development of AUSRIVAS models and outputs. In: *Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques* (Eds J.F. Wright, D.W. Sutcliffe & M.T. Furse), pp. 125–142. Freshwater Biological Association, Ambleside.
- Snook D.L. & Milner A.M. (2002) Biological traits of macroinvertebrates and hydraulic conditions in a glacier-fed catchment (French Pyrénées). *Archiv für Hydrobiologie*, **153**, 245–271.
- Spellerberg I.F. (1991) *Monitoring Ecological Change*. Cambridge University Press, Cambridge, U.K.

- Statzner B. (1987) Characteristics of lotic ecosystems and consequences for future research directions. In: *Potentials and Limitations of Ecosystem Analysis* (Eds E.D. Schulze & H. Zwölfer), pp. 365–390. Springer Verlag, Berlin.
- Statzner B. (1988) Growth and Reynolds number of lotic macroinvertebrates: a problem for adaptation of shape to drag. *Oikos*, **51**, 84–87.
- Statzner B. & Higler B. (1986) Stream hydraulics as a major determinant of benthic invertebrate zonation patterns. *Freshwater Biology*, **16**, 127–139.
- Statzner B., Dolédec S. & Hugueny B. (2004) Biological trait composition of European stream invertebrate communities: assessing the effects of various trait filter types. *Ecography*, **27**, 470–488.
- Statzner B., Gore J.A. & Resh V.H. (1988) Hydraulic stream ecology: observed patterns and potential applications. *Journal of the North American Benthological Society*, **7**, 307–360.
- Statzner B., Hildrew A.G. & Resh V.H. (2001b) Species traits and environmental constraints: entomological research and the history of ecological theory. *Annual Review of Entomology*, **46**, 291–316.
- Statzner B., Resh V.H. & Dolédec S. (Eds) (1994) Ecology of the Upper Rhône River: a test of habitat templet theories. *Freshwater Biology* (Special issue), **31**, 253–554.
- Statzner B., Bis B., Dolédec S. & Usseglio-Polatera P. (2001a) Perspectives for biomonitoring at large spatial scales: a unified measure for the functional composition of invertebrate communities in European running waters. *Basic and Applied Ecology*, **2**, 73–85.
- Statzner B., Hoppenhaus K., Arens M.-F. & Richoux P. (1997) Reproductive traits, habitat use and templet theory: a synthesis of world-wide data on aquatic insects. *Freshwater Biology*, **38**, 109–135.
- Thioulouse J., Chessel D., Dolédec S. & Olivier J.-M. (1997) ADE-4: a multivariate analysis and graphical display software. *Statistics and Computing*, **7**, 75–83.
- Townsend C.R. & Hildrew A.G. (1994) Species traits in relation to a habitat templet for river systems. *Freshwater Biology*, **31**, 265–276.
- Townsend C.R., Dolédec S. & Scarsbrook M.R. (1997) Species traits in relation to temporal and spatial heterogeneity in streams: a test of habitat templet theory. *Freshwater Biology*, **37**, 367–387.
- Tscharntke T., Gathmann A. & Steffan-Dewenter I. (1998) Bioindication using trap-nesting bees and wasps and their natural enemies: community structure and interactions. *Journal of Applied Ecology*, **35**, 708–719.
- Usseglio-Polatera P. (1994) Theoretical habitat templates, species traits, and species richness: aquatic insects in the Upper Rhône River and its floodplain. *Freshwater Biology*, **31**, 417–437.
- Usseglio-Polatera P., Bournaud M., Richoux P. & Tachet H. (2000) Biological and ecological traits of benthic freshwater macroinvertebrates: relationships and definition of groups with similar traits. *Freshwater Biology*, **43**, 175–205.
- Van der Heijden M.G.A., Klironomos J.N., Ursic M., Moutoglou P., Streitwolf-Engel R., Boller T., Wiemken A. & Sanders I.R. (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature*, **396**, 69–72.
- Vannote R.L., Minshall G.W., Cummins K.W., Sedell J.R. & Cushing C.E. (1980) The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences*, **37**, 130–137.
- Vogel S. (1994) *Life in Moving Fluids*, 2nd edn. Princeton University Press, Princeton.
- Wilkinson L., Blank G. & Gruber C. (1996) *Desktop Data Analysis with SYSTAT*. Prentice-Hall, Upper Saddle River, New Jersey.
- Wright J.F. (2000) An introduction to RIVPACS. In: *Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques* (Eds J.F. Wright, D.W. Sutcliffe & M.T. Furse), pp. 1–24. Freshwater Biological Association, Ambleside.
- Wright J.F., Sutcliffe D.W. & Furse M.T. (Eds) (2000) *Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques*. Freshwater Biological Association, Ambleside.

**Appendix** Summary statistics on trait categories defining LIRR conditions based on 68 large European river reaches (see Table 3, step 4 for reach characteristics). For each (1) trait category (see Table 4 for category labels), we indicated (2) the expected category value (average proportion in all 68 LIRRs) and (3) the SE of this mean expected value; the (4) lower and (5) upper limit of the entire frequency distribution of observed minus expected category values for the 68 LIRRs; (6) the number of the 16 most natural LIRRs (no cargo-ship traffic, organic and heavy metal pollution in class 1) in the lower half (<median) of the frequency distribution of observed minus expected category values of the 68 LIRRs; and the (7) lower and (8) upper limit of the frequency distribution of observed minus expected category values that enveloped 90% of the values of the 68 LIRRs. Note that the values enveloping 90% of the reaches were obtained by cutting off the tails of the entire distributions in relation to the distribution of the 16 most natural LIRRs, i.e. for none or one of the most natural LIRRs <median, seven reaches were cut off at the lower tail, for two or three of them <median, six were cut off at the lower and one at the upper tail, and so on; for eight most natural LIRRs <median, we randomly cut off three and four reaches from both tails of the entire distribution.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A1	0.201	0.008	-0.177	0.119	5	-0.087	0.095
A2	0.341	0.008	-0.134	0.190	2	-0.087	0.150
A3	0.331	0.006	-0.100	0.123	14	-0.086	0.078
A4	0.097	0.005	-0.089	0.078	8	-0.059	0.062
A5	0.031	0.004	-0.031	0.141	12	-0.031	0.050
B1	0.401	0.017	-0.353	0.305	11	-0.306	0.170
B2	0.402	0.011	-0.186	0.226	8	-0.113	0.177
B3	0.166	0.012	-0.157	0.237	5	-0.138	0.217
B4	0.031	0.003	-0.031	0.112	8	-0.031	0.050
C1	0.070	0.005	-0.065	0.148	9	-0.052	0.073
C2	0.686	0.008	-0.168	0.140	8	-0.111	0.104
C3	0.245	0.007	-0.113	0.099	10	-0.092	0.084
D1	0.700	0.015	-0.290	0.285	5	-0.165	0.244
D2	0.201	0.010	-0.190	0.204	10	-0.145	0.100
D3	0.099	0.008	-0.099	0.184	12	-0.099	0.092
E1	0.065	0.005	-0.065	0.083	4	-0.058	0.066
E2	0.320	0.014	-0.230	0.298	5	-0.133	0.216
E3	0.136	0.005	-0.073	0.111	3	-0.058	0.111
E4	0.313	0.011	-0.226	0.174	10	-0.182	0.114
E5	0.166	0.011	-0.149	0.202	13	-0.148	0.138
F1	0.044	0.002	-0.031	0.082	7	-0.019	0.029
F2	0.195	0.016	-0.195	0.347	14	-0.195	0.205
F3	0.761	0.016	-0.376	0.200	2	-0.187	0.192
G1	0.013	0.002	-0.013	0.061	5	-0.013	0.022
G2	0.216	0.013	-0.167	0.291	5	-0.129	0.219
G3	0.027	0.003	-0.027	0.075	8	-0.027	0.041
G4	0.020	0.002	-0.020	0.070	8	-0.020	0.026
G5	0.615	0.009	-0.167	0.170	7	-0.097	0.112
G6	0.108	0.010	-0.108	0.230	13	-0.108	0.123
H1	0.207	0.007	-0.112	0.159	6	-0.072	0.090
H2	0.318	0.005	-0.102	0.129	13	-0.075	0.065
H3	0.381	0.007	-0.128	0.144	3	-0.077	0.110
H4	0.090	0.003	-0.048	0.095	11	-0.039	0.033
H5	0.004	0.001	-0.004	0.030	6	-0.004	0.012
I1	0.072	0.004	-0.052	0.095	9	-0.044	0.058
I2	0.693	0.008	-0.196	0.126	7	-0.099	0.101
I3	0.099	0.004	-0.069	0.111	5	-0.051	0.074
I4	0.126	0.005	-0.064	0.124	9	-0.054	0.075
I5	0.009	0.001	-0.009	0.038	6	-0.009	0.025
J1	0.160	0.009	-0.160	0.170	10	-0.124	0.116
J2	0.262	0.011	-0.162	0.200	6	-0.111	0.158
J3	0.578	0.011	-0.191	0.163	5	-0.153	0.133
K1	0.073	0.003	-0.073	0.054	3	-0.046	0.045
K2	0.294	0.009	-0.128	0.252	14	-0.116	0.088
K3	0.490	0.013	-0.290	0.168	2	-0.132	0.146
K4	0.143	0.012	-0.143	0.281	12	-0.138	0.132

Appendix (Continued)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
L1	0.071	0.004	-0.057	0.095	7	-0.050	0.065
L2	0.283	0.008	-0.121	0.126	8	-0.112	0.081
L3	0.294	0.010	-0.252	0.202	8	-0.182	0.111
L4	0.104	0.005	-0.073	0.092	4	-0.052	0.068
L5	0.069	0.006	-0.069	0.126	14	-0.069	0.082
L6	0.069	0.004	-0.057	0.070	4	-0.039	0.045
L7	0.109	0.009	-0.109	0.201	7	-0.089	0.113
M1	0.182	0.004	-0.073	0.066	4	-0.049	0.056
M2	0.094	0.003	-0.045	0.084	9	-0.030	0.036
M3	0.028	0.002	-0.024	0.032	9	-0.018	0.020
M4	0.293	0.008	-0.206	0.137	5	-0.133	0.126
M5	0.091	0.003	-0.047	0.052	8	-0.037	0.036
M6	0.010	0.001	-0.010	0.012	3	-0.008	0.012
M7	0.117	0.005	-0.060	0.107	8	-0.055	0.074
M8	0.131	0.004	-0.061	0.086	11	-0.052	0.067
M9	0.054	0.005	-0.054	0.103	11	-0.054	0.061
N1	0.452	0.009	-0.121	0.224	10	-0.103	0.142
N2	0.461	0.008	-0.157	0.147	8	-0.100	0.089
N3	0.052	0.004	-0.052	0.094	2	-0.052	0.069
N4	0.035	0.004	-0.035	0.082	8	-0.035	0.056