Stressor prioritisation in riverine ecosystems: Which environmental factors shape benthic invertebrate assemblage metrics?

Andrea Sundermann, Miriam Gerhardt, Heike Kappes, Peter Haase

Ecological Indicators 27 (2013) 83–96

Contents lists available at SciVerse ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Stressor prioritisation in riverine ecosystems: Which environmental factors shape benthic invertebrate assemblage metrics?

Andrea Sundermann[a,b,*], Miriam Gerhardt[a,b], Heike Kappes[c], Peter Haase[a,b]

[a] Senckenberg Research Institute and Natural History Museum Frankfurt, Department of River Ecology and Conservation, Clamecystrasse 12, 63571 Gelnhausen, Germany
[b] Biodiversity and Climate Research Centre, Senckenberganlage 25, 60325 Frankfurt am Main, Germany
[c] Naturalis Center for Biodiversity, Postbus 9517, NL-2300 RA Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 18 July 2012
Received in revised form
30 November 2012
Accepted 2 December 2012

Keywords:
Biological indicators
Land use practices
Local habitat structure
Restoration
Water quality

ABSTRACT

Aquatic ecosystems are amongst the most heavily altered ecosystems and exhibit a disproportional loss of biodiversity. Numerous stressors, such as nutrient enrichment, contaminant pollution, sedimentation and alterations in stream hydrology and habitat structure, account for these losses. Understanding these forces is of utmost importance to prevent riverine ecosystems from further deterioration and to provide helpful insights for restoration practices. In the present study, we analyse the response of biological indicators to a large number of environmental factors. For this, benthic invertebrate assemblages from 83 sites in Germany were described based on 25 metrics from four different metric types. The condition of the sites was described using 27 environmental factors: 13 for water quality, 4 for land use in the catchment and 10 for local scale habitat structure. The relative importance of single environmental predictors or predictor combinations for benthic invertebrate assemblages was analysed with single and multiple linear regression models. The results for the latter models were statistically supported via a bootstrap approach. The models revealed the importance of water quality and catchment-scale land use in explaining benthic invertebrate assemblages; in particular, chloride, oxygen, total organic carbon and the amounts of artificial surfaces and arable land were the most important predictors. Models including solely structural variables such as plan form, bank structures and substrate diversity had lower goodness of fit values than those for other variables. Regarding the four different assemblage metric types, functional metrics had on average lower goodness of fit values than composition/abundance, richness/diversity and sensitivity/tolerance metrics. Among the richness/diversity metrics, however, the model results for the Shannon–Wiener and Simpson diversity indices and evenness were poor. Our results show that catchment-related factors and water quality were of overriding importance in shaping biodiversity patterns and causing species loss. In contrast, structural degradation at a local scale was not the most significant stressor. This finding might explain why structural restoration at a reach scale often yields a low benefit–cost ratio and may be considered to represent inappropriate investment prioritisation.

1. Introduction

Conserving nature and preserving biodiversity represent major challenges facing today's ecologists. As freshwater ecosystems are amongst the most heavily altered ecosystems and display a disproportional loss of biodiversity (Geist, 2011), it is of utmost importance to understand the relative importance of the driving forces of species loss. Freshwater organisms face numerous stressors, such as nutrient enrichment, contaminant pollution, sedimentation and alterations in stream hydrology and habitat structure (Allan, 2004). Among these stressors, the recently most significant and widespread in Europe are expected to be diffuse water pollution from intensive land use and physical degradation of water ecosystems (EU Commission, 2007). Because diffuse water pollution is more difficult to pinpoint and often less discernible, restoration efforts commonly focus on the most obvious constraint, which often is physical degradation. Consequently, restoration projects commonly aim at reconstructing a channel form that is similar to a historic form or that of a least-disturbed reference site (Palmer, 2009). For example, rivers have been re-meandered or re-braided (Jähnig and Lorenz, 2008; Lorenz et al., 2009); dead wood has been donated to structure river beds (Sundermann et al., 2011a); or boulders that once were removed to ease timber floating have been replaced to enhance habitat diversity (Nilsson et al., 2005). Thus, restoration efforts are often carried out at the reach scale, extending for only several hundred metres up to a few kilometres in length. Also, restoration efforts are often focused on

* Corresponding author at: Clamecystrasse 12, 63571 Gelnhausen, Germany. Tel.: +49 6051 61954 3124; fax: +49 6051 61954 3118.
E-mail addresses: Andrea.Sundermann@senckenberg.de (A. Sundermann), Miriam.Gerhardt@senckenberg.de (M. Gerhardt), heike.kappes@naturalis.nl (H. Kappes), Peter.Haase@senckenberg.de (P. Haase).

1470-160X/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ecolind.2012.12.003

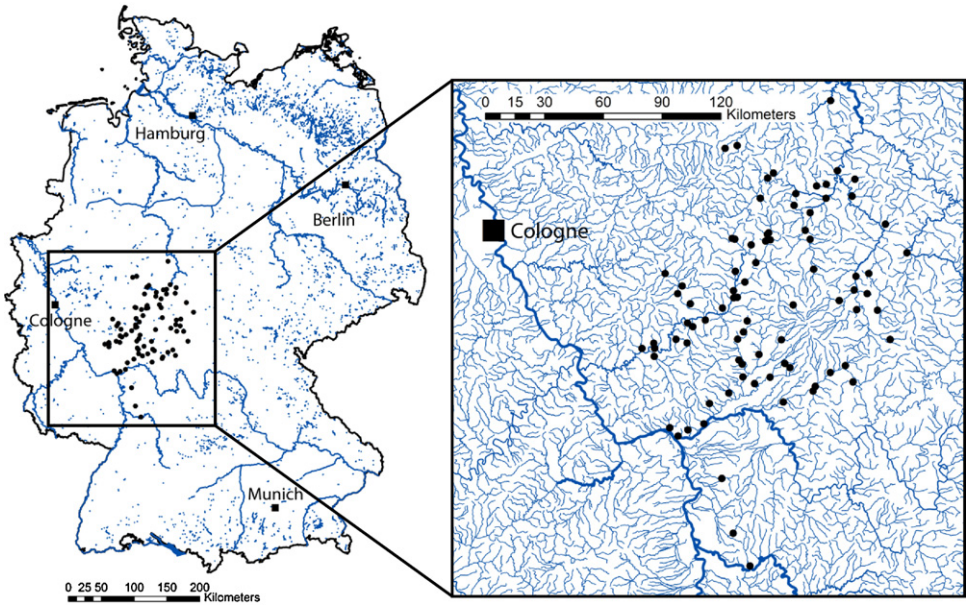

Fig. 1. Location of the 83 investigated streams in Germany.

reach scales, simply because of costs and policy difficulty. Yet, it has been speculated that water quality or other large-scale factors are of overriding importance (Palmer et al., 2010; Sundermann et al., 2011a,b; Haase et al., 2013). If this is the case, structural restoration at a reach scale is anticipated to have a low benefit–cost-ratio and may be considered to represent an inappropriate mode of investment prioritisation. Indeed, only a minority of reach-scale restoration projects in Europe and in the US significantly have enhanced biodiversity, and the underlying mechanism is subject to controversial discussions (Blakely and Harding, 2005; Kowalik and Ormerod, 2006; Larson et al., 2001; Ledger and Hildrew, 2005; Palmer et al., 2010; Sundermann et al., 2011a; Suren and McMurtrie, 2005).

Consequently, the gap in knowledge related to the relative importance of anthropogenic stressors in shaping freshwater communities needs to be bridged. The purpose of this study is to investigate which environmental predictors predict species composition and abundance structure of benthic invertebrate assemblages and whether water quality or land use in the catchment, are of overriding importance compared to local scale habitat structure. In this context, we will analyse 25 biological indicators from four metric types and test how closely they are related to environmental predictors. In addition, we will analyse whether single environmental predictors or a combination of various predictors explain benthic invertebrate assemblages best.

To this end, we analyse the species composition and abundance structure of benthic invertebrate assemblages at 83 sites in highland rivers in Germany. The investigated dataset encompasses a large stressor gradient, as it contains reference sites as well as severely modified sites. Determining the relative importance of environmental predictors in riverine ecosystems will help us to prevent riverine ecosystems from further deterioration, identify flaws in river restoration concepts and will provide helpful insights for optimising restoration projects in the future.

2. Materials and methods

In the present study, 83 fine substrate dominated siliceous highland rivers in Germany were investigated (Fig. 1). The amount of precipitation in the study area ranges from 650 and 950 mm per year with maxima during summer months. The climate is moderate with generally warm summers and mild winters. All sites are in second to fourth-order streams (Strahler system) at elevations between 119 and 304 m above sea level with small catchment areas between 10 and 189 km^2. The benthic invertebrate assemblages were recorded at each site. In addition, water quality was characterised at each site using 13 physicochemical variables; 10 variables were compiled describing the degree of structural degradation at a site; and 4 variables were employed to characterise the land use in the catchment.

2.1. Benthic invertebrate assemblages

The samples originated from routine surface water surveys according to the protocol for collecting samples in river monitoring programs to assess the ecological status of rivers in Germany (Haase et al., 2004). The samples were collected from March to July in 2004 to 2008. The sampling method is based on sampling microhabitats according to their coverage at the sampling site (multi-habitat sampling). All microhabitats in a 100-m-long stream section are recorded in 5% coverage intervals, and each "sampling unit" (25 cm × 25 cm) is sampled using a handnet (mesh size: 0.5 mm) via the kick sampling method. A complete sample is comprised of 20 sampling units which are pooled for further analysis (total sampling area of 1.25 m^2). The organisms are sorted from the sediments in the laboratory and identified according to the "Operational Taxalist for Running Water in Germany" (Haase et al., 2006, http://www.fliessgewaesserbewertung.de/en/download/bestimmung/). The mean number of taxa and individuals collected and identified from each sample was 30.3 ± 11.4 and 1749 ± 1239. To describe the benthic invertebrate assemblages in the samples, 25 metrics were calculated for each site. These metrics can be divided into four metric types: composition/abundance, richness/diversity, sensitivity/tolerance and function (Hering et al., 2004) (Table 1). As the three sensitivity/tolerance metrics are not self-explanatory, they will be explained here. The river-type-specific multi-metric index (MMI) is a German national metric that describes the general degradation of a site. For each river type, the MMI is composed of three to five metrics scaled to values between zero and one, with class boundaries at scoring intervals of 0.2 (Böhmer et al., 2004). The Biological Monitoring Working Party (BMWP) System was set up in Great Britain to recommend a biological classification system

Table 1
Analysed metrics.

Metric type	Name of metric	Short name	Min	Max	Mean	References
ca	Ephemeroptera, Plecoptera and Trichoptera percentage of abundance	EPT%	3.7	71.2	33.9	
rd	Number of taxa	#Taxa	11	62	30	
rd	Number of genera	#Genera	11	52	26	
rd	Number of families	#Families	9	34	20	
rd	Number of Ephemeroptera, Plecoptera and Trichoptera	#EPT	1	36	12	
rd	Number of Ephemeroptera, Plecoptera, Trichoptera, Coleoptera, Bivalvia and Odonata	#EPTCBO	2	44	16	
rd	Simpson diversity index	Simpson	0.44	0.96	0.84	Simpson (1949)
rd	Shannon-Wiener diversity index	Shannon	0.94	3.48	2.44	Shannon and Weaver (1949)
rd	Evenness	Evenness	0.35	0.97	0.73	Magurran (1983)
st	Multi metric index	MMI	0.00	0.88	0.3	Böhmer et al. (2004)
st	Biological Monitoring Working Party	BMWP	25	218	97	Hawkes (1998)
st	Average Score Per Taxon	ASPT	3.13	7.52	5.4	Armitage et al. (1983)
f	Percentage of active filter feeders	ActFilFeeder	0.0	31.8	4.2	Moog (1995), Schmedtje & Colling (1996), Hering et al. (2004)
f	Percentage of gatherers and collectors	GathColl	8.2	63.7	25.6	
f	Percentage of grazers and scrapers	GrazScra	2.2	50.6	25.6	
f	Percentage of passive filter feeders	PasFilFeeder	0.0	60.0	8.2	
f	Percentage of predators	Predator	1.0	27.3	11.3	
f	Percentage of shredders	Shredder	0.4	49.4	13.3	
f	Percentage of xylophagous taxa	Xylophagous	0.0	4.3	0.3	
f	Rao diversity: Reproduction	Reproduction	1.3	13.7	7.7	Tachet et al. (2000)
f	Rao diversity: Dispersal	Dispersal	1.3	6.3	3.6	Botta-Dukát (2005)
f	Rao diversity: Resistance	Resistance	0.1	4.9	2.2	
f	Rao diversity: Locomotion	Locomotion	0.9	9.0	5.2	
f	Rao diversity: Feeding Type	FeedingType	2.3	13.7	5.6	
f	Rao diversity: Substrate Preference	SubstratePref	2.7	11.4	7.8	

Metric types: ca, composition/abundance metrics; rd, richness/diversity metrics; st, sensitivity tolerance metrics and f, functional metrics. Based on all 83 sites, the minimum (min), maximum (max) and mean values are given for each metric.

for use in national river pollution surveys (Hawkes, 1998). The BMWP score equals the sum of the tolerance scores of all benthic invertebrate families in a sample. When the BMWP score is high, the water quality is good. The Average Score Per Taxon (ASPT) is the BMWP score of the sample divided by the number of scoring families that contributed to the BMWP score (Armitage et al., 1983). These metrics were calculated with ASTERICS, Version 3.01 (http://www.fliessgewaesserbewertung.de/download/berechnung).

We use the R 2.15.0 statistical software package (R Dvelopment Core Team, 2012) to calculate Rao's quadratic entropy (Botta-Dukát, 2005) as a measure of diversity for six functional traits. Each of the six traits were described by several categories (listed in brackets): reproduction (ovoviviparity, isolated eggs, clutches, asexual reproduction), dispersal (aquatic passive, aquatic active, aerial passive, aerial active), resistance (eggs/statoblasts, cocoons, diapause/dormancy, none), locomotion (flier, surface swimmer, full water swimmer, crawler, burrower, interstitial, temporarily attached, permanently attached), substrate preference (flags/boulders/cobbles/pebbles, gravel, sand, silt, macrophytes, microphytes, twigs/roots, organic detritus/litter, mud), and feeding type (absorber, deposit feeder, shredder, scraper, filter-feeder, piercer, predator, parasite) (Tachet et al., 2000).

Reference sites as well as heavily degraded sites were sampled, as documented by the large gradient in metric values obtained (Table 1). Prior to analysis, the metric variables were transformed (counts: $\sqrt{x}$, percentages: $\sqrt{(x/100)}$) to make their sampling variance more equitable.

2.2. Water quality (physicochemical variables)

Data on the following physicochemical variables were available for all investigated sites: ammonium, chloride, total nitrogen, nitrate, orthophosphate, total phosphorus and total organic carbon (TOC), calcium, magnesium, total hardness, electrical conductivity, water temperature and oxygen. The variables were usually measured monthly between the years 2002 and 2007. An average of 32.2 ± 16.7 recordings was performed for each variable at each site. The lowest number of recordings per variable and site was 10. For most variables, the negative impact on an organism increases with increasing concentrations. For all of these variables, we calculated the 90th percentile for analysis. This value was used instead of the maximum to give less weight to outliers or single spikes. However, the opposite pattern was observed for oxygen: the lower the concentration, the higher the stress on the majority of benthic invertebrates. Thus, we used the 10th percentile for the analysis of the oxygen data. All physicochemical variables were transformed ($\log(x+1)$) as necessary prior to analyses to approach normality.

Threshold values expected to be relevant for benthic invertebrates were predetermined for the following physicochemical variables by the German Working Group of Water Issues of the Federal States and the Federal Government (LAWA, 1998): ammonium ($\leq 0.38\,mg\,l^{-1}$), chloride ($\leq 100\,mg\,l^{-1}$), nitrate ($\leq 11\,mg\,l^{-1}$), orthophosphate ($\leq 0.3\,mg\,l^{-1}$), oxygen ($\geq 6\,mg\,l^{-1}$) and TOC ($\leq 5\,mg\,l^{-1}$) (direction of the threshold indicates optimal conditions).

2.3. Local scale habitat structure

The river habitat survey method of LAWA (Kamp et al., 2007) was used to assess the small-scale habitat structure of a site. This method considers a 100-m-long stream section for which a total of 26 variables are investigated, such as erosion, flow diversity, bank stabilisation, constructions, substratum type, cross-section form, vegetation and land use in the floodplain (Kamp et al., 2007). Each of the 26 variables is assigned to one of the following six categories: plan form, longitudinal profile, bed structures, cross-section, bank structures and floodplain corridor (for details, see Kamp et al., 2007). Each variable in the six categories can take integer values between 1 (undisturbed) and 7 (completely disturbed).

Moreover, information regarding the variation in river width and river depth, riparian land use and substrate diversity was compiled for each site. Variations in river width and river depth were estimated based on a 5-step scale, where 5 refers to high variability, similar to what occurs under reference conditions, and 1 refers to very low variability. This estimation was performed by the same person who sampled the benthic invertebrates at the corresponding site. Additionally, riparian land use was recorded at each site by noting which land use categories were present within a 100-m stream section. The land use categories were divided into three groups accounting for land use intensity: (a) native forest and fallow; (b) grassland, timber and parks; and (c) urban areas and agricultural crop land. The three groups were assigned corresponding scores of (a) +2, (b) +1 and (c) 0. If the land use at a site belonged to only one of the three groups, the corresponding score was assigned to this site. When the land use at a site belonged to two or all three groups, the mean values were calculated.

To describe the substrate diversity at a site, Shannon–Wiener diversity (Shannon and Weaver, 1949) was calculated based on estimation of the coverage of microhabitats (see Section 2.1).

2.4. Catchment-scale land use

We calculated the proportion of Corine Land Cover (CLC2000, www.eea.europa.eu/) classes in the catchment as follows: (1) artificial surfaces (CLC class 1), (2) arable land and permanent crops (CLC classes 2.1 and 2.2), (3) pastures and heterogeneous agricultural areas (CLC classes 2.3 and 2.4) and (4) forest and other "natural" cover (CLC classes 3–5). Prior to the analysis, the metric variables were transformed ($\sqrt{(x/100)}$) to make their sampling variance more equitable.

2.5. Analysis

Environmental predictors (water quality, land use and habitat structure) were tested for collinearity (Spearman Rank Correlation Test). Moreover, a principal component analysis (PCA) was calculated in order to test whether the number of environmental predictors can be reduced by summing them up in PCA axis 1 and possibly also PCA axis 2. To describe the range in the conditions of the analysed sites, the values for all 83 sites were plotted separately for the selected environmental predictors (Fig. 2). Regression models were calculated using each biocoenotic metric as the dependent variable and the environmental predictors (water quality, land use and habitat structure) as independent variables. To determine the regression model that fits most of the combinations best, we calculated a set of 27 most common regression models for each combination, e.g. exponential, log probit, logistic and reciprocal models. This resulted in a total of 14,175 models being examined (25 metrics $\times$ 21 environmental predictors $\times$ 27 regression models). A list of the calculated regression models is available as supporting information in the online version of the article. The relationship between environmental predictors and single metrics was analysed by calculating simple linear regression models. To predict metric results from a combination of environmental predictors, we used multiple linear regression (MLR) models. In a first step, a "full model" was calculated. By default, this model includes (a) all variables for water quality, (b) all variables for land use or (c) all variables for habitat structure. In a second step, we aimed at developing a simpler, yet adequate model with a reduced number of predictors ("reduced model"). For this end, we used an automated selection procedure to select only the most significant variables for the reduced model; specifically we conducted backwards elimination and considered the significance of the Akaike Information Criterion (AIC). From hereafter the resulting model is called "best" model. When using AIC as model selection method, it is important to show how well the "best" model compares to alternative or competing models. It is often the case that the best model is hardly distinguishable from competing models. In order to overcome these flaws, the AICc (the small sample unbiased AIC) was recorded in the backwards elimination process. Competing models were identified by comparing the AICc values with that of the best model. Almost as good models were defined as those with $\Delta\text{AICc} \leq 2$ as recommended by Burnham and Anderson (2001). All variables which remained in the competing models were also recorded. Often a slight change in the data might lead to a different set of variables which remain in the model. To account for this, a bootstrap of the entire stepwise regression procedure was accomplished. For each model $b = 1000$ bootstrap resamples of the 83 sites were calculated to assess the frequency and occurrence of each of the environmental predictors in the fit of the reduced models.

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolind.2012.12.003.

Besides calculating models for either (a) all variables for water quality, (b) all variables for land use or (c) all variables for habitat structure, we aimed to calculate MLR models including (d) environmental predictors from all three groups (water quality, land use and habitat structure). However, to avoid problems due to overparameterisation (Freedman and Pee, 1989; Pietrobon et al., 2004), we reduced the number of environmental predictors to only those taken into account in at least five of the reduced models calculated under (a), (b) or (c).

3. Results

3.1. Collinearity of predictors

Some of the environmental predictors showed high collinearity, e.g., combinations of calcium, chloride, electrical conductivity, total hardness and magnesium (Spearman's rho ≥ 0.85, Table 2). To reduce the number of variables, calcium, electrical conductivity, total hardness and magnesium were excluded from the analyses. As a representative of this group of variables, chloride remained. Total phosphorus and total nitrogen were also excluded from further analyses, as these predictors are highly correlated with orthophosphate and nitrate (Spearman's rho > 0.89). The decision to retain chloride, orthophosphate and nitrate in the analysis was based on the assumption that these variables have a more direct effect on benthic invertebrates than the other variables. A PCA was calculated, including the remaining variables. PCA axis 1 (28.5%) and PCA axis 2 (13.4%) accounted for only a fairly small amount of the total variance. Correlation of variables with PCA axis 1 was highest for plan form (-0.30). Thus, the low percentage of variance explained and the low correlation of variables with PCA axis 1 did not suggest a further reduction in the number of variables nor did the results suggest for using PCA axis 1 as a surrogate variable in further analyses.

Table 2
Collinearity between predictors: correlations (Spearman's rho) are shown below the diagonal, and *p*-values are shown above the diagonal. The level of significance has been adjusted according to the high number of pairwise comparisons. ***$p < 0.0001$, **$p < 0.001$, *$p < 0.01$. Bold: Predictors that showed high collinearity (Spearman's rho > 0.85) and were consequently excluded from further analyses.

	Water quality							Physicochemical variables					
	Ammonium	**Calcium**	Chloride	**Electrical conductivity**	**Total hardness**	**Total phosphorus**	**Magnesium**	**Total nitrogen**	Nitrate	Orthophosphate	Oxygen	Total organic carbon	Water temperature
Ammonium			***	*		**		*		***	*	***	
Calcium	0.20		***	***	***		***	***	***		***	**	*
Chloride	0.43	0.67		***	***	***	***	***	***	***	***	***	***
Electrical conductivity	0.30	**0.92**	**0.86**		***	*	***	***	***	*	***	***	***
Total hardness	0.18	**0.98**	0.65	**0.92**			***	***	***		***	**	**
Total phosphorus	0.36	0.24	0.42	0.34	0.27			*	*	***	*	***	***
Magnesium	0.05	0.82	0.55	0.79	**0.91**	0.24		***	***		***	**	**
Total nitrogen	0.32	0.53	0.45	0.53	0.53	0.33	0.42		***	*	***	***	
Nitrate	0.23	0.56	0.43	0.54	0.54	0.31	0.42	**0.99**		*	**	***	
Orthophosphate	0.41	0.25	0.45	0.35	0.26	**0.89**	0.21	0.34	0.32		**	***	**
Oxygen	−0.34	−0.45	−0.47	−0.54	−0.47	−0.35	−0.47	−0.42	−0.40	−0.36		***	***
Total organic carbon	0.47	0.36	0.55	0.49	0.39	0.66	0.38	0.47	0.43	0.63	−0.60		***
Water temperature	0.26	0.35	0.64	0.52	0.37	0.42	0.38	0.14	0.11	0.36	−0.54	0.56	
Artificial surfaces	−0.01	0.42	0.59	0.55	0.43	0.36	0.39	0.01	−0.51	0.22	−0.20	0.01	0.55
Arable land	0.31	0.69	0.42	0.66	0.71	0.46	0.68	0.63	0.67	0.40	−0.52	0.48	0.23
Pastures	−0.18	−0.52	−0.44	−0.55	−0.50	−0.26	−0.42	−0.31	−0.33	−0.29	0.41	−0.20	−0.24
Forest	−0.12	−0.57	−0.33	−0.54	−0.59	−0.46	−0.58	−0.41	−0.44	−0.29	0.40	−0.39	−0.25
Plan form	−0.03	0.42	0.43	0.48	0.44	0.32	0.45	0.16	0.18	0.25	−0.09	0.24	0.30
Longitudinal profile	−0.02	0.31	0.24	0.31	0.36	0.37	0.41	0.01	0.03	0.26	−0.07	0.21	0.30
Cross-section	0.02	0.47	0.30	0.45	0.47	0.19	0.43	0.23	0.25	0.07	−0.21	0.12	0.27
Bed structures	0.06	0.39	0.30	0.39	0.41	0.37	0.40	0.15	0.15	0.33	−0.14	0.30	0.28
Bank structures	−0.11	0.41	0.19	0.38	0.41	0.16	0.39	0.19	0.23	0.06	−0.07	0.16	0.11
Floodplain corridor	0.07	0.21	0.23	0.20	0.16	0.19	0.08	0.20	0.22	0.09	−0.01	0.09	0.05
Variance river width	−0.08	−0.20	−0.22	−0.21	−0.20	−0.07	−0.24	0.00	0.03	−0.04	0.17	−0.17	−0.43
Variance river depth	−0.09	−0.28	−0.30	−0.31	−0.30	−0.17	−0.31	−0.17	−0.15	−0.12	0.20	−0.32	−0.39
Riparian land use	−0.09	−0.08	−0.23	−0.16	−0.07	−0.17	−0.02	−0.11	−0.08	−0.05	0.13	−0.15	−0.16
Substrate diversity	0.01	−0.01	0.02	0.05	0.02	−0.01	0.07	0.05	0.06	0.02	−0.17	0.07	0.00

	Land use			Habitat structure Structural variables								
Artificial surfaces	Arable land	Pastures	Forest	Plan form	Longitudinal profile	Cross-section	Bed structures	Bank structures	Floodplain corridor	Variance river width	Variance river depth	Riparian land use	Substrate diversity
	*												
**	**	**	**	***	*	***	**	**					
**	**	**	*	***		*	*				*		
**	**	**	**	***	*	***	**	**			*		
**	**	**	**	***	**	***	**	**			*		
*	**		**	*	**		**						
**	**	**	**	***	***	***	**	**			*		
	**	*	**										
**	**	*	**										
	**	*	*				*						
	**	**	**										
	**		**				*				*		
**			*	*	*		*			***	**		
		*		***	***	*	*						
0.14		**	**	*		**	*	*					
−0.31	−0.60						*						
−0.34	−0.70	0.13				**							
0.49	0.28	−0.26	−0.27		***	***	***	***	**		**		
0.43	0.21	−0.13	−0.27	0.70		***	***	***	*	*	**		
0.29	0.39	−0.27	−0.39	0.62	0.42		***	***	***		*		
0.29	0.34	−0.33	−0.22	0.71	0.70	0.43		***	*	**	*		
0.21	0.31	−0.28	−0.24	0.59	0.45	0.64	0.56		***		*		
0.21	0.21	−0.23	−0.10	0.41	0.31	0.45	0.32	0.79			*	***	
−0.23	−0.08	0.20	0.04	−0.26	−0.31	−0.21	−0.38	−0.26	−0.21		***		
−0.26	−0.15	0.19	0.06	−0.39	−0.40	−0.29	−0.29	−0.34	−0.34	0.66			
−0.28	−0.08	−0.02	0.16	−0.24	−0.05	−0.21	−0.12	−0.15	−0.47	0.13	0.24		
0.02	0.09	0.00	0.01	−0.09	−0.07	−0.05	−0.09	−0.07	−0.13	0.09	−0.03	−0.07	

Table 3
Differences in the goodness of fit (given in as a percentage) between the linear model and any of the other 27 models. For description of metric types, see Table 1.

Metric	Metric type	Water quality Physicochemical variables							Land use			
		Ammonium	Chloride	Nitrate	Orthophosphate	Oxygen	Total organic carbon	Water temperature	Artificial surfaces	Arable land	Pastures	Forest
EPT%	ca	13.1	9.7	1.8	4.1	0.7	7.0	0.6	0.2	1.9	7.2	3.8
ActFilFeeder	f	5.3	2.2	0.6	3.6	0.7	6.2	2.1	1.8	1.0	1.6	2.9
GathColl	f	3.8	0.4	0.4	2.0	0.2	1.1	0.8	0.6	1.9	2.4	1.0
GrazScra	f	16.8	5.3	0.7	4.3	1.9	1.7	1.5	0.3	1.8	6.7	3.6
PasFilFeeder	f	0.5	1.5	0.7	0.9	2.7	1.8	3.1	0.2	5.8	5.1	3.3
Predators	f	0.3	1.5	2.4	0.4	0.0	1.0	0.4	0.2	0.9	0.1	0.5
Shredders	f	0.7	0.4	2.2	0.7	0.2	0.4	0.5	0.9	0.4	1.0	0.5
Xylophagous	f	5.8	1.8	0.5	0.2	0.3	0.1	1.0	0.8	0.3	1.6	0.5
Reproduction	f	3.1	0.8	0.1	2.6	0.4	0.1	1.0	0.7	1.8	3.0	0.7
Dispersal	f	8.1	2.7	2.8	0.3	3.0	0.8	6.1	6.2	1.6	2.1	2.8
Resistance	f	1.6	11.0	2.7	8.0	8.1	5.8	7.2	5.9	7.6	3.9	7.6
Locomotion	f	1.3	7.0	3.3	0.6	4.6	2.3	3.3	3.6	4.4	2.8	5.8
FeedingType	f	9.2	3.0	1.7	5.3	0.0	5.1	3.8	3.7	0.8	0.4	1.3
SubstratePref	f	1.2	0.3	2.8	4.5	0.3	0.5	0.5	0.5	0.1	2.9	1.0
#EPT	rd	17.0	7.7	1.1	2.5	5.6	1.7	4.3	2.9	2.4	4.0	7.5
#EPTCBO	rd	16.6	7.3	1.2	2.0	4.7	0.7	4.3	4.8	2.1	8.0	4.2
#Families	rd	10.5	6.4	1.2	1.6	1.2	0.8	3.1	4.6	1.0	3.2	0.0
#Genera	rd	11.0	5.6	1.0	0.9	1.3	0.1	4.0	7.4	2.0	4.1	0.0
#Taxa	rd	12.6	5.9	0.9	1.5	1.4	0.3	3.8	8.4	0.8	3.2	0.1
Shannon	rd	8.3	7.0	2.3	5.0	0.2	0.9	0.1	0.7	2.4	2.5	1.0
Evenness	rd	1.3	1.8	1.5	3.1	0.3	0.8	0.1	0.0	2.1	0.8	1.2
Simpson	rd	3.9	3.6	4.7	3.4	0.1	1.3	1.3	0.9	2.6	1.5	2.6
ASPT	st	10.1	5.5	1.2	0.7	1.6	1.2	1.4	2.6	0.8	3.3	0.4
BMWP	st	15.5	6.9	1.5	3.3	3.1	1.1	4.5	5.4	1.0	4.9	0.8
MMI	st	12.9	11.1	0.3	6.1	2.1	4.6	0.1	0.0	0.0	4.0	0.1

Habitat structure Structural variables									
Plan form	Longitudinal profile	Cross-section	Bed structures	Bank structures	Floodplain corridor	Variance river width	Variance river depth	Riparian land use	Substrate diversity
0.4	1.1	0.1	0.1	0.3	0.1	3.5	2.8	0.3	0.3
0.8	2.8	0.6	0.1	0.3	0.6	5.0	2.5	0.4	2.3
0.8	0.4	0.7	0.4	1.6	0.1	3.5	2.7	0.0	1.7
0.6	0.1	0.7	1.7	0.7	0.2	0.6	1.9	0.2	0.5
2.5	2.7	2.0	3.4	1.7	1.5	0.8	1.1	0.0	2.9
0.3	0.6	2.8	0.9	0.0	2.4	0.7	0.3	1.5	1.7
2.0	0.3	1.2	0.4	0.2	0.4	0.5	0.6	0.0	0.7
0.5	1.1	1.6	0.2	0.2	1.6	0.3	0.0	1.6	0.5
1.0	2.1	0.7	0.4	0.6	4.9	3.6	2.3	1.5	0.8
4.1	0.1	0.7	3.5	1.8	2.0	0.8	1.9	1.3	6.7
7.5	0.0	3.2	6.1	1.6	1.3	3.8	0.8	3.1	1.1
2.1	0.4	1.3	2.6	0.9	0.5	0.1	0.3	0.0	1.6
1.8	1.5	0.8	0.7	0.8	2.3	3.1	4.6	1.5	1.5
0.4	1.7	0.1	1.1	0.8	0.5	1.2	2.6	0.8	0.5
0.4	0.1	1.1	0.3	0.6	0.2	3.0	2.1	0.1	1.5
0.3	0.2	1.2	0.8	0.8	0.4	2.2	2.2	0.1	1.7
1.9	1.4	1.5	3.4	2.0	0.6	1.1	0.1	0.3	1.9
2.1	1.2	1.3	3.2	1.8	0.6	1.0	1.1	0.4	2.2
3.0	1.6	1.3	2.7	1.7	0.6	0.1	1.2	0.2	2.3
0.3	1.3	0.7	2.4	0.7	0.2	0.2	1.8	0.1	2.4
0.8	0.4	0.4	1.0	0.4	0.1	0.4	0.5	0.1	1.2
0.6	0.8	0.5	1.4	0.3	0.0	0.3	1.6	0.0	2.3
0.5	0.1	0.1	0.6	0.2	0.2	1.7	0.3	0.0	1.3
1.6	0.3	1.2	3.3	1.6	0.5	1.7	1.1	0.1	2.1
0.6	0.0	0.0	0.4	0.3	0.8	3.0	0.5	0.0	0.8

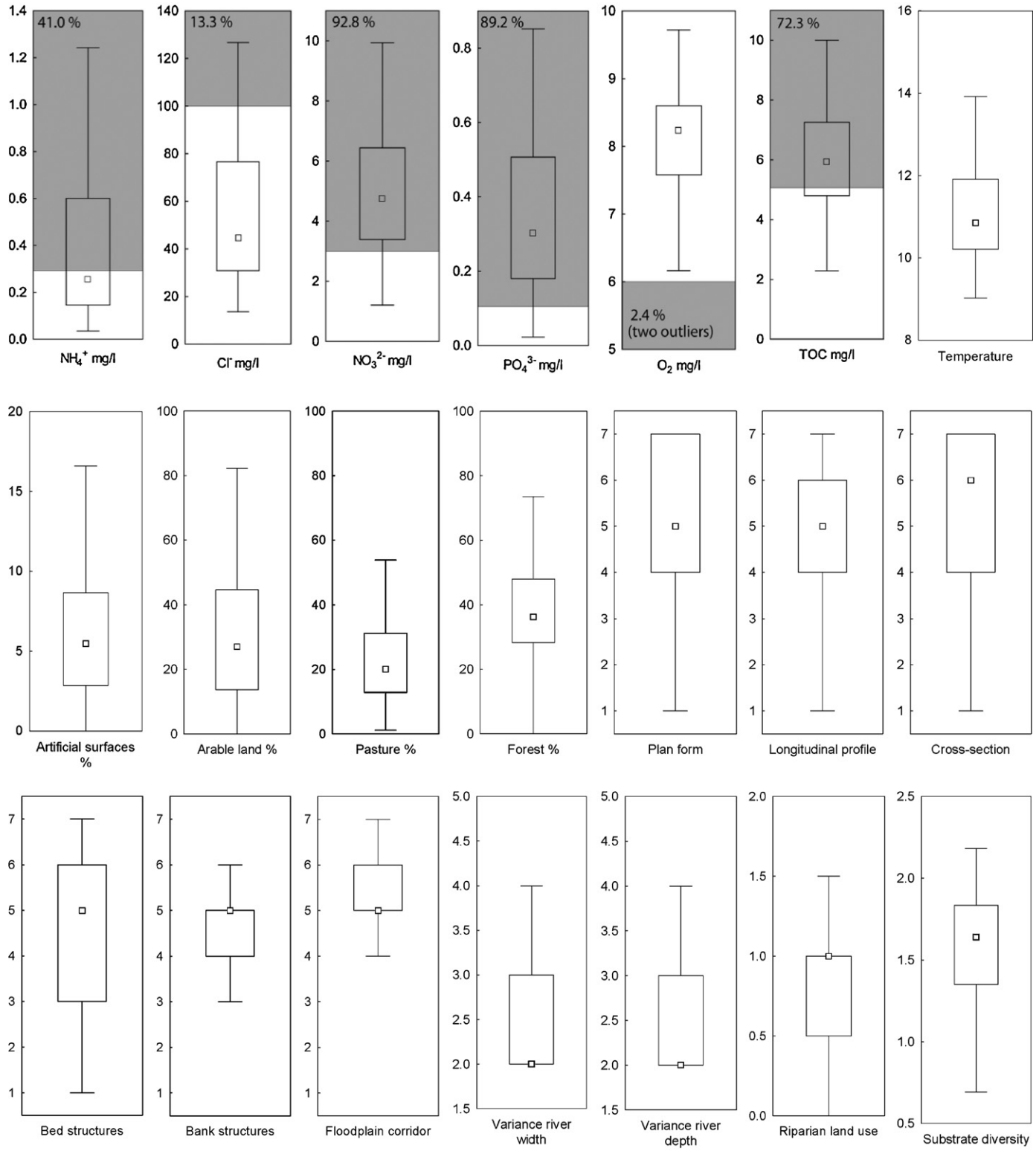

Fig. 2. Box plots show the range in the conditions of the analysed sites ($N = 83$) for environmental predictors that will be considered for further analysis. Values that exceed predetermined threshold values are shaded in grey. The percentages indicate the amount of sites exceeding predetermined limits. The box-plots do not consider outliers and extreme values.

3.2. Conditions of the analysed sites

The box plots shown in Fig. 2 demonstrate the large quality gradient covered by the analysed sites: regarding all variables for which threshold values have been established, at least some of the sites meet reference condition requirements (except for bank structures and floodplain corridors), meaning that the impact is negligibly low, whereas other sites are considerably impacted and fail to meet predetermined limit values. Due to this wide spectrum, we assumed that the gradient of each analysed environmental predictor was sufficiently long to cause differences in species composition and abundance structure of benthic invertebrate

Table 4
Goodness of fit (R^2 values) for simple linear regression (SLR) models. The levels of significance (p) for the regression models are represented by different shades of grey; ranging from dark to light grey: $p < 0.001$, $p < 0.01$, $p < 0.05$. White: Results were not significant ($p \geq 0.05$). Bold values indicate a negative relationship between a metric value and an environmental variable. For description of metric types, see Table 1.

Metric		Water quality Physicochemical variables							Land use				Habitat structure Structural variables									
	Metric type	Ammonium	Chloride	Nitrate	Orthophosphate	Oxygen	Total organic carbon	Water temperature	Artifical	Arable land	Pasture	Forest	Plan form	Longitudinal profile	Cross-section	Bed structures	Bank structures	Floodplain corridor	Variance river width	Variance river depth	Riparian land use	Substrate diversity
EPT%	ca	0.0	0.2	.	**0.2**	0.2	**0.2**	0.2	0.0	**0.3**	**0.2**	**0.2**	0.0	0.0	0.0	0.1	.	.	0.0	0.0	.	.
ActFilFeeder	f	.	0.0	.	.	0.0	0.0	.	.	.	.	.	0.0	0.0	.	.	.	.	0.0	.	.	.
GathColl	f	0.0	.	.	0.0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GrazScra	f	0.0	0.0	.	0.0	0.2	0.1	0.1	.	**0.2**	**0.1**	0.0	.	.	.	0.0	.	.	0.0	.	.	.
PasFilFeeder	f	.	0.0	.	.	.	.	0.0	0.0	.	.	.	0.0	.	.	.	.	.	.	.	.	.
Predators	f	.	.	.	.	.	.	.	.	.	.	0.0	.	.	.	.	.	.	.	.	.	.
Shredders	f	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Xylophagous	f	.	0.0	.	.	0.0	0.0	.	.	.	.	.	0.0	.	.	.	.	.	.	.	.	.
Reproduction	f	0.0	.	.	0.0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Dispersal	f	0.0	.	.	.	.	.	.	**0.0**	.	.	.	.	.	.	.	.	.	.	.	.	.
Resistance	f	.	**0.1**	.	.	.	.	.	**0.1**	**0.0**	.	0.0	**0.1**	.	.	**0.0**	.	.	.	.	.	.
Locomotion	f	.	**0.0**	.	.	0.0	.	.	.	.	.	0.0	.	.	.	.	.	.	.	.	.	.
FeedingType	f	0.0	.	.	0.1	.	.	0.1	0.0	.	.	.	.	.	.	.	.	.	.	.	.	.
SubstratePreff	f	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
#EPT	rd	0.0	**0.3**	0.0	**0.2**	0.2	**0.3**	**0.2**	**0.1**	**0.3**	**0.2**	**0.1**	0.0	0.0	0.0	0.1	.	.	0.0	0.1	.	.
#EPTCBO	rd	0.1	**0.3**	0.0	0.2	0.2	**0.3**	**0.2**	**0.1**	**0.2**	**0.2**	**0.1**	0.0	.	0.0	0.0	.	.	0.0	0.1	.	.
#Families	rd	.	**0.2**	.	0.1	0.1	0.1	0.2	.	0.0	.	0.0	0.0	0.0	.	0.0	.	.	0.0	0.1	.	.
#Genera	rd	.	**0.2**	.	0.1	0.1	0.2	0.2	**0.2**	**0.1**	**0.1**	0.1	0.0	0.0	.	0.0	.	.	0.0	0.1	.	.
#Taxa	rd	.	**0.2**	.	0.1	0.1	0.2	0.2	**0.2**	**0.1**	**0.1**	0.1	0.0	.	.	0.0	.	.	0.0	0.1	.	.
Shannon	rd	.	0.1	0.0	0.0	0.1	0.2	0.2	**0.1**	**0.1**	**0.1**	0.1	0.0	.	.	0.0	.	.	.	.	.	.
Evenness	rd	.	.	.	.	.	0.0	0.0	0.0	**0.1**	0.0	**0.1**	.	.	.	.	.	.	.	.	.	.
Simpson	rd	.	0.0	.	.	0.1	0.1	0.1	.	0.0	.	0.1	.	.	.	.	.	.	.	.	.	.
ASPT	st	0.0	**0.3**	0.1	0.1	0.3	**0.3**	**0.3**	**0.2**	**0.2**	**0.2**	**0.2**	0.0	0.0	0.0	0.1	.	.	0.0	0.1	.	.
BMWP	st	0.0	**0.3**	0.0	0.1	0.2	**0.3**	**0.2**	**0.2**	**0.2**	**0.2**	**0.1**	0.0	0.0	0.0	0.0	.	.	0.0	0.1	.	.
MMI	st	.	**0.2**	0.0	0.2	0.3	**0.3**	**0.3**	**0.1**	**0.2**	**0.1**	**0.2**	0.1	0.1	0.0	0.1	.	.	0.0	0.1	.	.
Mean R^2 per metric type	f	0.0	0.0	0.0	0.0	0.2	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	rd	0.0	0.2	0.0	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	st	0.0	0.3	0.0	0.1	0.2	0.3	0.3	0.1	0.2	0.2	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0
Mean R^2 for all variables	ca	0.20							0.24				0.05									
	f	0.03							0.03				0.02									

assemblages. Thus, the next step was to find an appropriate model to describe the relationship between the environmental predictors and benthic invertebrate assemblages, as described by metric values.

3.3. Selecting the most appropriate regression model

The simplest type of regression model is the linear model. However, in a statistical sense, it is not always the best model. Therefore, the degree to which any of the other 27 models was able to better fit the benthic invertebrate assemblage-based metric values was calculated (Table 3). The comparison showed that taking all environmental predictors into account, the fit of any of the other 27 regression models was on average only 2.1 ± 2.59% better than that of the linear model. Thus, the linear regression model was considered to be appropriate to fit all environmental predictors.

3.4. Fit of simple linear regression (SLR) models

Regarding single environmental predictors, the goodness of fit (R^2 values) was highest for chloride, oxygen, total organic carbon, water temperature and arable land, especially for composition/abundance, richness/diversity and sensitivity/tolerance metric types. However, the best models only rarely achieved R^2 values higher than 0.30. Thus, only some of the SLR models were able to explain more than 30% of the variability in the dataset, and the predictability of the metric results based on single environmental variables was limited. On average, the goodness of fit for the SLR models was better for physicochemical variables and land use than for structural variables, and this difference was significant for the composition/abundance, richness/diversity and sensitivity/tolerance metric types (Mann Whitney U-Test, all $p < 0.01$) (compare "Mean goodness of fit (R^2) for all variables and metric types combined" in Table 3, lower block). The functional metrics

showed the lowest R^2 values on average for both physicochemical and structural variables (Table 4).

3.5. Fit of multiple linear regression (MLR) models

Independent of whether the models took into account (a) all variables for water quality, (b) all variables for land use, (c) all variables for habitat structure or (d) the combination of the most relevant environmental predictors, the goodness of fit (R^2 values) for the MLR models was higher on average than for the SLR models (Table 5). The models presented higher R^2 values when they included physicochemical variables or catchment-scale land use (results under (a), (b) and (d)) than when they included reach-scale structural variables alone. The three different approaches to select the most important variables, lead to a slightly different selection of variables. For example the results for the metric EPT% under (a) revealed that the variables chloride, oxygen and TOC were selected for the best model. In addition orthophosphate and temperature were chosen in competing models, and the bootstrap revealed that all five variables were of relevance (Table 5). For many metrics, three physicochemical variables remained in the reduced models. These variables were chloride, oxygen and total organic carbon (TOC). Regarding land use (results under b), the best model revealed that the amounts of artificial surfaces and arable land were of importance for many metrics. However, the results for the competing models and the bootstrap approach documented that the amounts of pastures and forest were also of importance. Among the structural variables, three variables were of relevance: bed structures, variance in river depth and bank structures. However, when including the combination of all environmental predictors by default (d), almost none of the structural variables remained in the best models. The bootstrap approach documented the relevance of variance in river depth for the richness/diversity and sensitivity/tolerance metrics. Regarding the four different metric types, functional metrics performed less well on average than composition/abundance, richness/diversity and sensitivity/tolerance metrics. Among the richness/diversity metrics, however, the model results for the Shannon–Wiener and Simpson diversity indices as well as evenness were poor or even non-significant.

4. Discussion

The aim of the present study was to investigate whether benthic invertebrate assemblages in streams are determined by local habitat structure or by water quality and other large-scale factors, such as land use in the catchment. For this end, we chose a correlative approach, which is often applied (Campos-Aranda, 2011; Wagenhoff et al., 2011; Seeboonruang, 2012) and which is less

Table 5
(a–d) Goodness of fit (R^2) and level of significance (p) for the multiple linear regression models. By default, the full model includes (a) all variables for water quality, (b) all variables for land use, (c) all variables for habitat structure or (d) a subset of environmental predictors (variables included in at least two of the reduced models calculated under (a), (b) or (c)). Variables that are not included by default are shaded in grey. For a description of metric types, see Table 1. n.s. indicates that the results of calculated models were not significant. Variables remaining in the reduced models (red model) after backwards elimination are indicated by the corresponding level of significance: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Variables which remained in almost as good model as the reduced model (applying the AICc approach) are indicated by "0". Variables which were chosen after the bootstrap approach are indicated by the box.

a)

Metric	Metric type	Full model R^2	Full model p	Red. model R^2	Red. model p	Water quality Physicochemical variables: Ammonium	Chloride	Nitrate	Orthophosphate	Oxygen	TOC	Temperature
EPT%	ca	0.47	***	0.47	***	.	*	.	0	*	*	0
ActFilFeeder	f	0.21	**	0.20	***	0	0	.	.	***	.	0
GathColl	f	0.18	*	0.16	**	*	.	.	*	.	0	.
GrazScra	f	0.24	**	0.23	***	.	.	.	.	*	*	.
PasFilFeeder	f	0.10	n.s.	0.08	n.s.	.	.	.	.	.	.	.
Predators	f	0.04	n.s.	0.02	n.s.	.	.	.	.	.	.	.
Shredders	f	0.07	n.s.	0.03	n.s.	.	.	.	.	.	.	.
Xylophagous	f	0.14	n.s.	0.11	**	.	0	.	.	0	0	.
Reproduction	f	0.21	*	0.18	**	0	.	0	***	.	0	.
Dispersal	f	0.16	n.s.	0.11	**	**	*	.	0	.	.	.
Resistance	f	0.12	n.s.	0.09	*	0	**	.	.	.	0	.
Locomotion	f	0.08	n.s.	0.04	n.s.	.	.	.	.	.	.	.
FeedingType	f	0.20	*	0.20	**	.	.	.	.	.	.	**
SubstratePref	f	0.04	n.s.	-	-	.	.	.	.	.	.	.
#EPT	rd	0.54	***	0.53	***	.	***	.	.	**	**	.
#EPTCBO	rd	0.54	***	0.53	***	.	***	.	0	**	***	.
#Families	rd	0.39	***	0.37	***	0	***	.	.	*	0	.
#Genera	rd	0.42	***	0.40	***	0	***	.	.	*	0	.
#Taxa	rd	0.42	***	0.40	***	0	***	.	.	*	*	.
Shannon	rd	0.25	**	0.21	***	.	.	*	.	0	0	*
Evenness	rd	0.09	n.s.	0.04	n.s.	.	.	.	.	.	.	.
Simpson	rd	0.13	n.s.	0.09	**	.	.	.	.	.	**	0
ASPT	st	0.55	***	0.53	***	.	***	.	.	***	***	0
BMWP	st	0.51	***	0.50	***	.	***	.	.	**	**	0
MMI	st	0.51	***	0.51	***		.	.	*	**	.	*

Table 5 (*continued*)

b)

Metric	Metric type	Full model R^2	Full model p	Red. model R^2	Red. model p	Land use: Artificial surfaces	Land use: Arable land	Land use: Pastures	Land use: Forest
EPT%	ca	0.44	***	0.42	***	**	***	0	0
ActFilFeeder	f	0.07	n.s.	0.04	n.s.	.	.	.	.
GathColl	f	0.06	n.s.	0.04	n.s.	.	.	.	.
GrazScra	f	0.23	***	0.23	***	0	**	0	0
PasFilFeeder	f	0.08	n.s.	0.07	n.s.	.	.	.	.
Predators	f	0.06	n.s.	0.05	*	.	0	.	*
Shredders	f	0.05	n.s.	0.03	n.s.	.	.	.	.
Xylophagous	f	0.04	n.s.	0.03	n.s.	.	.	.	.
Reproduction	f	0.04	n.s.	0.03	n.s.	.	.	.	.
Dispersal	f	0.03	n.s.	-	-	.	.	.	.
Resistance	f	0.07	n.s.	0.06	*	*	0	0	0
Locomotion	f	0.15	*	0.15	*	**	**	*	0
FeedingType	f	0.05	n.s.	0.04	n.s.	.	.	.	.
SubstratePref	f	0.08	n.s.	0.07	n.s.	.	.	.	.
#EPT	rd	0.46	***	0.46	***	**	***	0	0
#EPTCBO	rd	0.40	***	0.40	***	**	**	0	0
#Families	rd	0.33	***	0.32	***	***	***	0	0
#Genera	rd	0.34	***	0.33	***	***	***	0	0
#Taxa	rd	0.33	***	0.32	***	***	***	0	.
Shannon	rd	0.21	***	0.20	***	0	0	0	***
Evenness	rd	0.09	n.s.	0.09	**	0	.	0	**
Simpson	rd	0.13	*	0.05	n.s.	.	.	.	.
ASPT	st	0.45	***	0.44	***	***	***	0	0
BMWP	st	0.41	***	0.40	***	***	***	0	0
MMI	st	0.40	***	0.40	***	***	***	0	0

complex than many other multivariate statistical methods. One of the first steps in analysing the data was to find a mathematical function that most closely fits our data. The results clearly showed that the most simple model – the linear model – performed better than other non-linear models. This justified our approach to assume the dependent variable (in our case the metric) to be a linear function of the independent environmental variable. This might be unexpected as nonlinearity is often associated with ecological data (e.g. Dodds et al., 2010). In our case, however, the transformation of the data was sufficient to provide a linear relationship.

In our study, we did not just aim at finding the one and only model which fits the data best. One should be very cautious about attributing a level of importance to the variables that are included in a single stepwise regression. The reason is that slight changes in the data can by chance lead to a very different set of variables that will predict equally well (Chernick and LaBudde, 2011). We accounted for this by taking the results for the best model after subset regression into account and by also considering the results for competing (almost as good) models as well as the results for the bootstrap approach. This procedure lets us identify the most important environmental variables which explain the species composition and abundance structure of benthic invertebrate assemblages.

The results of this study indicate the relevance of water quality and land use in the catchment (Table 5a, b and d). The variables total organic carbon (TOC), oxygen and chloride as well as the land-use variables remained in many of the reduced models. High TOC usually result from anthropogenic inputs, such as from fertilisers, pesticides, surfactants and solvents, either related to their direct use or from inefficient sewage treatment plants (Visco et al., 2005). Oxygen is physically linked to water temperature but is also correlated with the concentrations of degradable substances, including TOC, and nutrients such as nitrate or phosphate, as eutrophication may result in excessive growth of algae and macrophytes followed by a reduction in oxygen as decay occurs. Chloride may reach surface waters from many sources, including agricultural runoff, wastewater from industries, effluent from wastewater treatment plants and road salting. Sources and correlates of TOC, oxygen and chloride, are related to large-scale rather than to small-scale factors. Therefore, these three variables can be considered a good proxy for the type and intensity of anthropogenic land use in a catchment (see Theodoropoulos and Iliopoulou-Georgudaki, 2010 for comparison). This conclusion was also reflected by our results. When considering land use variables alone (Table 5b), the artificial surface and arable land use variables were included in the majority of the reduced models, with the exception of functional metrics, which were not selected. However, in the most complex MLR models (Table 5d), artificial surfaces and arable land use appear to be of less importance than in the model presented in Table 5b, and thus, information regarding the anthropogenic impact in the catchment appears to be included within physicochemical variables.

The results of our study indicate that water quality and land use in the catchment, explain the species composition and abundance structure, rather than local scale habitat structure. This contradicts the assumption that benthic invertebrate assemblages strongly respond to local habitat structure. For instance, Sponseller et al. (2001) showed that benthic invertebrate indices were most closely related to land-cover patterns evaluated at a 200 m sub-corridor scale, suggesting that small-scale, streamside

Table 5 (*continued*)

c)

Metric	Metric type	Full model R^2	Full model p	Red. model R^2	Red. model p	Habitat structure: Planform	Longitudinal profile	Cross-section	Bed structures	Bank structures	Floodplain corridor	Variance river width	Variance river depth	Riparian land use	Substrate diversity
EPT%	ca	0.20	n.s.	0.14	**	.	.	.	*	.	.	.	0	0	.
ActFilFeeder	f	0.25	*	0.21	**	**	0	0	*	.	0	**	.	.	0
GathColl	f	0.15	n.s.	0.09	*	.	.	.	.	*	.	*	.	.	.
GrazScra	f	0.14	n.s.	0.12	*	.	.	.	*	.	.	.	.	.	.
PasFilFeeder	f	0.11	n.s.	0.08	*	.	.	.	.	.	.	.	.	.	.
Predators	f	0.06	n.s.	0.04	n.s.	.	.	.	.	.	.	.	.	.	.
Shredders	f	0.06	n.s.	-	-	.	.	.	.	.	.	.	.	.	.
Xylophagous	f	0.18	n.s.	0.15	*	.	.	*	0	*	.	.	0	.	0
Reproduction	f	0.10	n.s.	0.05	n.s.	.	.	.	.	.	.	.	.	.	.
Dispersal	f	0.17	n.s.	0.12	*	**	.	.	.	.	.	.	.	.	.
Resistance	f	0.15	n.s.	0.12	**	**	.	.	.	.	.	*	.	.	.
Locomotion	f	0.06	n.s.	0.05	n.s.	.	.	.	.	.	.	.	.	.	.
FeedingType	f	0.11	n.s.	0.04	n.s.	.	.	.	.	.	.	.	.	.	.
SubstratePref	f	0.11	n.s.	0.08	*	.	*	.	*	.	.	.	.	.	.
#EPT	rd	0.25	*	0.22	***	.	.	.	*	.	.	.	**	.	.
#EPTCBO	rd	0.23	*	0.17	***	.	.	0	0	0	.	.	**	.	.
#Families	rd	0.23	*	0.18	***	0	.	.	.	0	.	.	**	.	.
#Genera	rd	0.20	n.s.	0.15	**	.	.	.	0	0	.	.	**	.	.
#Taxa	rd	0.20	n.s.	0.16	**	.	.	.	0	0	.	.	**	.	.
Shannon	rd	0.11	n.s.	0.05	*	.	.	.	*	.	.	.	0	.	.
Evenness	rd	0.04	n.s.	-	-	.	.	.	.	.	.	.	.	.	.
Simpson	rd	0.07	n.s.	0.03	n.s.	.	.	.	.	.	.	.	.	.	.
ASPT	st	0.23	*	0.18	***	.	.	.	*	.	.	.	*	.	.
BMWP	st	0.24	*	0.21	**	.	.	0	0	0	.	.	**	.	0
MMI	st	0.25	*	0.20	***	.	.	.	*	.	.	.	**	.	.

development effectively alters the structure of species assemblages. Similar results were found by Hunsaker and Levine (1995), Sandin (2009) and Gombeer et al. (2011), who described benthic invertebrate assemblages as being especially sensitive to a number of local habitat factors. Also, there are several recent multi-stressor studies (Matthaei et al., 2010; Townsend et al., 2008; Wagenhoff et al., 2011) highlighting that deposited fine sediment is a pervasive stressor in running waters. Sediment load is another local factor altering habitat quality, as a high sediment load might lead to streambed colmation (Brunke, 1999). This would particularly impair burrowing species and may reduce habitat availability, especially if hard substrates, such as gravel, cobbles or boulders, are covered by a thin layer of fine sediments.

In the present study, however, the performance of models including only variables related to local scale habitat structure was poor (Table 5c). This was the case despite the fact that a broad spectrum of variables for habitat structure was included in the analysis, such as variables that considered instream habitat structure (e.g., substrate diversity or variance in river width and depth) as well as bank, riparian or floodplain structure. These results are supported by other studies that have also attributed more influence to catchment land use than to reach-scale habitat structures (Herringshaw et al., 2011; Kappes et al., 2011; Morley and Karr, 2002; Roy et al., 2003; Stephenson and Morin, 2009). The mentioned studies have in common that all study regions have experienced various anthropogenic disturbances, such as agriculture and urbanisation. The latter appears to overwhelm local factors. If this is true, benthic invertebrate assemblages in almost undisturbed catchments should react much more clearly on local factors than it was the case in our study. Therefore, we should not conclude that local factors such as the habitat structure at a study site are not relevant for the species composition and abundance structure. It is rather the land use in the catchment and water quality which seem to be of overriding importance compared to local factors.

Thus, when we attempt to predict and understand the effects of multiple stressors on benthic invertebrate assemblages, which is one of the most important challenges presently facing ecological studies (Tockner et al., 2010), we should concentrate on large-scale factors linked to land use practices in a catchment. These interactions are complex, indicating that we need to understand the functional chains in river ecology, which may allow us to identify the initial stressors from which the negative impact on the benthic invertebrate assemblages emanates. Our results showed that a maximum of 62% of the variance in the analysed metric values was explained by the reduced models (highest value for $R^2 = 0.62$). This finding indicates that in addition to the analysed variables, other (environmental) variables might exist that explain a considerable amount of the remaining variance in the benthic invertebrate assemblages. One aspect might be due to the sampling period of collected samples. The samples were taken in the period of early spring to mid-summer of five years. This sampling regimen might add an extra amount of variance to the data set (between seasons and also between years). Taking the samples within a more finite time period might have enhanced the precision of the predictive models. However, taking many samples in a short time frame is often not easy to realise in water management practice.

Table 5 (*continued*)

d)

Metric	Metric type	Full model R^2	Full model p	Red. model R^2	Red. model p	Water quality Physicochemical variables: Ammonium	Chloride	Orthophosphate	Oxygen	TOC	Temperature	Land use: Artificial surfaces	Arable land	Pastures	Forest	Habitat structure: Bed structures	Bank structures	Variance river depth
EPT%	ca	0.59	***	0.56	***	.	0	0	*	.	**	0	.	***	**	.	.	.
ActFilFeeder	f	0.25	n.s.	0.02	***	.	.	.	***	.	.	.	.	.	.	.	.	.
GathColl	f	0.31	*	0.28	***	*	.	**	.	.	.	.	**	.	*	.	.	.
GrazScra	f	0.39	***	0.37	***	.	.	.	.	.	*	**	.	**	.	.	.	.
PasFilFeeder	f	0.13	n.s.	0.07	n.s.	.	.	.	.	.	.	.	.	.	.	.	.	.
Predators	f	0.13	n.s.	0.08	*	.	.	.	.	.	.	.	.	.	.	.	.	.
Shredders	f	0.12	n.s.	0.07	n.s.	.	.	.	.	.	.	.	.	.	.	.	.	.
Xylophagous	f	0.22	n.s.	0.18	*	.	*	.	*	.	.	.	*	.	.	*	.	.
Reproduction	f	0.16	n.s.	0.14	**	.	.	**	.	.	.	.	.	.	.	.	.	.
Dispersal	f	0.20	n.s.	0.16	*	.	.	.	.	.	.	*	.	.	.	.	.	.
Resistance	f	0.21	n.s.	0.14	**	.	.	.	.	.	.	.	**	.	.	.	.	.
Locomotion	f	0.11	n.s.	0.04	n.s.	.	.	.	.	.	.	.	.	.	.	.	.	.
FeedingType	f	0.28	*	0.25	***	0	.	*	0	.	***	.	.	.	0	0	.	0
SubstratePref	f	0.10	n.s.	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.
#EPT	rd	0.63	***	0.62	***	.	*	0	0	**	.	.	.	**	0	0	0	0
#EPTCBO	rd	0.61	***	0.58	***	.	**	.	**	***	.	.	.	**	.	.	.	.
#Families	rd	0.47	***	0.45	***	.	**	.	.	.	.	.	.	.	.	.	.	**
#Genera	rd	0.48	***	0.44	***	.	**	.	*	.	.	.	.	.	.	.	.	.
#Taxa	rd	0.49	***	0.46	***	.	.	.	.	*	.	.	.	.	.	.	.	.
Shannon	rd	0.31	*	0.25	***	.	.	.	.	*	.	.	.	.	*	.	.	.
Evenness	rd	0.15	n.s.	0.08	**	.	.	.	.	.	.	.	.	.	**	.	.	.
Simpson	rd	0.21	n.s.	0.14	**	.	.	.	.	.	.	.	.	.	*	.	.	.
ASPT	st	0.62	***	0.59	***	.	0	.	**	***	.	**	.	**	0	.	.	.
BMWP	st	0.58	***	0.56	***	.	**	.	0	*	0	.	.	*	0	.	.	*
MMI	st	0.57	***	0.55	***	.	.	*	.	.	*	.	***	.	.	.	.	.

4.1. Performance of different metric types

The assemblage descriptors belonged to four metric types: composition/abundance, richness/diversity, sensitivity/tolerance and function. The results obtained for the functional metrics were especially poor. This was true for the feeding types as well as for functional diversity indices. However, a rising number of studies already showed the high potential of functional metrics to reflect the influence of environmental stressors better than, e.g. composition/abundance metrics (Dolédec et al., 2011; Statzner et al., 2001). Among the richness/diversity metrics, the range of goodness of fit values for the model results was quite large. While models for number of Ephemeroptera, Plecoptera and Trichoptera (#EPT), number of Ephemeroptera, Plecoptera, Trichoptera, Coleoptera, Bivalvia and Odonata (#EPTCBO), and number of taxa, genera and families worked well, the results for Shannon–Wiener diversity, evenness and Simpson diversity were comparatively poor. On the first glance, these results seem to be inconsistent. However, in contrast to the five metrics that only count the number of corresponding taxa, the three last indices also account for the evenness of their abundance. This means that comparatively high values for indices such as Shannon–Wiener diversity can also be achieved under impaired conditions, as long as the remaining taxa exhibit approximately equal abundances. Thus, our results show that unspecific diversity indices tend to be unsuitable for predicting the impact of environmental factors on benthic invertebrate assemblages. These findings are in line with other studies (Patrício et al., 2009; Peet, 1975; Washington, 1984). Therefore, the Shannon–Wiener diversity index is only suitable to a certain extent for detecting changes. We recommend further research to understand the behaviour of ecological indicators in view of their crucial importance for management and protection practices.

4.2. Implications for water management practices

The results of our study highlight the importance of land-use in the catchment and water quality for benthic invertebrate assemblages. Thus, in a multiple stressor environment structural restoration at the reach scale will yield a low benefit–cost ratio and may represent inappropriate investment prioritisation, at least from the perspective of benthic invertebrate assemblages (Jähnig et al., 2011). Of course, this might be different for catchments where structural degradation is the only impact on the organisms. Future restoration projects should therefore not only aim

to improve flow regime and channel form but also require a more holistic approach (Sundermann et al., 2011a,b). Specifically, measures aimed at reducing diffuse pollution, such as via extending riparian corridors over longer river stretches, might have a more positive effect on benthic invertebrate assemblages compared to small-scale habitat improvement. Physical degradation alone may only shape benthic invertebrate communities significantly when superordinate stressors, such as intensive anthropogenic land use practices and related implications, can be neglected due to having a low impact in a system (Kail et al., 2012).

Acknowledgements

This study was financed by the research funding program "LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-oekonomischer Exzellenz" of Hesse's Ministry of Higher Education, Research and the Arts. All data were provided by the Hessisches Landesamt für Umwelt und Geologie, which is greatly acknowledged.

References

Allan, J.D., 2004. Landscapes and riverscapes. The influence of land use on stream ecosystems. Annu. Rev. Ecol., Evol. Syst. 35, 257–284.
Armitage, P.D., Moss, D., Wright, J.F., Furse, M.T., 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. Water Res. 17, 333–347.
Blakely, T.J., Harding, J.S., 2005. Longitudinal patterns in benthic communities in an urban stream under restoration. N. Z. J. Mar. Freshw. Res. 39, 17–28.
Böhmer, J., Rawer-Jost, C., Zenker, A., Meier, C., Feld, C., Biss, R., Hering, D., 2004. Assessing streams in Germany with benthic invertebrates: development of a multimetric invertebrate based assessment system for German rivers. Limnologica 34, 416–432.
Botta-Dukát, Z., 2005. Rao's quadratic entropy as a measure of functional diversity based on multiple traits. J. Veg. Sci. 16, 533–540.
Brunke, M., 1999. Colmation and depth filtration within streambeds. Retention of particles in hyporheic interstices. Int. Rev. Hydrobiol. 84, 99–117.
Burnham, K.P., Anderson, D.R., 2001. Kullback–Leibler information as a basis for strong inference in ecological studies. Wildl. Res. 28 (2), 111–119.
Campos-Aranda, D.F., 2011. Transference of hydrologic information through multiple linear regression, with best predictor variables selection. Agrociencia 45, 863–880.
Chernick, M.R., LaBudde, R., 2011. An Introduction to Bootstrap Methods with Applications. John Wiley & Sons, Inc., New Jersey, Hoboken.
Dodds, W.K., Clements, W.H., Gido, K., Hilderbrand, R.H., King, R.S., 2010. Thresholds, breakpoints, and nonlinearity in freshwaters as related to management. J. N. Am. Benthol. Soc. 29, 988–997.
Dolédec, S., Phillips, N., Townsend, C., 2011. Invertebrate community responses to land use at a broad spatial scale: trait and taxonomic measures compared in New Zealand rivers. Freshw. Biol. 56, 1670–1688.
EU Commission, 2007. Towards sustainable water management in the European Union. First stage in the implementation of the Water Framework Directive 2000/60/EC. Commission staff working document. Accompanying document to the communication forum from the commission to the European Parliament and the council, COM (2007).
Freedman, L.S., Pee, D., 1989. Return to note on screening regression equations. Am. Stat. 43, 279–282.
Geist, J., 2011. Integrative freshwater ecology and biodiversity conservation. Ecol. Indicators 11, 1507–1516.
Gombeer, S.C., Knapen, D., Bervoets, L., 2011. The influence of different spatial-scale variables on caddisfly assemblages in Flemish lowland streams. Ecol. Entomol. 36, 355–368.
Haase, P., Lohse, S., Pauls, S., Schindehutte, K., Sundermann, A., Rolauffs, P., Hering, D., 2004. Assessing streams in Germany with benthic invertebrates: development of a practical standardised protocol for macro invertebrate sampling and sorting. Limnologica 34, 349–365.
Haase, P., Schindehütte, K., Sundermann, A., 2006. Operationelle Taxaliste als Mindestanforderung an die Bestimmung von Makrozoobenthosproben aus Fließgewässern zur Umsetzung der EU-Wasserrahmenrichtlinie in Deutschland. http://www.fliessgewaesserbewertung.de/download/bestimmung (accessed 20.12.11).
Haase, P., Hering, D., Jähnig, S.C., Lorenz, A.W., Sundermann, A., 2013. The impact of hydromorphological restoration on river ecological status: a comparison of fish, benthic invertebrates, and macrophytes. Hydrobiologia, http://dx.doi.org/10.1007/s10750-012-1255-1.
Hawkes, H.A., 1998. Origin and development of the Biological Working Party score system. Water Res. 32, 964–968.
Hering, D., Meier, C., Rawer-Jost, C., Feld, C., Biss, R., Zenker, A., Sundermann, A., Lohse, S., Böhmer, J., 2004. Assessing streams in Germany with benthic invertebrates: selection of candidate metrics. Limnologica 34, 398–415.
Herringshaw, C.J., Stewart, T.W., Thompson, J.R., Anderson, P.F., 2011. Land use, stream habitat and benthic invertebrate assemblages in a highly altered Iowa watershed. Am. Midl. Nat. 165, 274–293.
Hunsaker, C.T., Levine, D.A., 1995. Hierarchical approaches to the study of water-quality in rivers. Bioscience 45, 193–203.
Jähnig, S.C., Lorenz, A.W., 2008. Substrate-specific macroinvertebrate diversity patterns following stream restoration. Aquat. Sci. 70, 292–303.
Jähnig, S.C., Lorenz, A.W., Hering, D., Antons, C., Sundermann, A., Jedicke, E., Haase, P., 2011. River restoration success: a question of perception. Ecol. Appl. 21, 2007–2015.
Kail, J., Arle, J., Jähnig, S.C., 2012. Limiting factors and thresholds for macroinvertebrate assemblages in European rivers: empirical evidence from three datasets on water quality, catchment urbanization, and river restoration. Ecol. Indicators 18, 63–72.
Kamp, U., Binder, W., Hölzl, K., 2007. River habitat monitoring and assessment in Germany. Environ. Monit. Assess. 127, 209–226.
Kappes, H., Sundermann, A., Haase, P., 2011. Distant land use affects terrestrial and aquatic habitats of high naturalness. Biodivers. Conserv. 20, 2297–2309.
Kowalik, R.A., Ormerod, S.J., 2006. Intensive sampling and transplantation experiments reveal continued effects of episodic acidification on sensitive stream invertebrates. Freshw. Biol. 51, 180–191.
Larson, M.G., Booth, D.B., Morley, S.A., 2001. Effectiveness of large woody debris in stream rehabilitation projects in urban basins. Ecol. Eng. 18, 211–226.
LAWA (Länderarbeitsgemeinschaft Wasser), 1998. Beurteilung der Wasserbeschaffenheit von Fließgewässern in der Bundesrepublik Deutschland – Chemische Gewässergüteklassifikation. Kulturbuchverlag, Berlin.
Ledger, M.E., Hildrew, A.G., 2005. The ecology of acidification and recovery: changes in herbivore-algal food web linkages across a stream pH gradient. Environ. Pollut. 137, 103–118.
Lorenz, A.W., Jähnig, S.C., Hering, D., 2009. Re-meandering German lowland streams: qualitative and quantitative effects of restoration measures on hydromorphology and macroinvertebrates. Environ. Manage. 44, 745–754.
Magurran, A.E., 1983. Ecological Diversity and its Measurement. Croom Helm, London.
Matthaei, C.D., Piggott, J.J., Townsend, C.R., 2010. Multiple stressors in agricultural streams: interactions among sediment addition, nutrient enrichment and water abstraction. J. Appl. Ecol. 47, 639–649.
Moog, O. (Ed.), 1995. Fauna Aquatica Austriaca – a comprehensive species inventory of Austrian aquatic organisms with ecological data. , 1st ed. Wasserwirtschaftskataster, Bundesministerium für Land- und Forstwirtschaft, Wien.
Morley, S.A., Karr, J.R., 2002. Assessing and restoring the health of urban streams in the Puget Sound basin. Conserv. Biol. 16, 1498–1509.
Nilsson, C., Lepori, F., Malmqvist, B., Hjerdt, N., Helfield, J.M., Palm, D., Oestergren, J., Jansson, R., Brännäs, E., Lundqvist, H., 2005. Forecasting environmental responses to restoration of rivers used as log floatways: an interdisciplinary challenge. Ecosystems 8, 779–800.
Palmer, M.A., 2009. Reforming watershed restoration: science in need of application and applications in need of science. Estuar. Coast. 32, 1–17.
Palmer, M.A., Menninger, H.L., Bernhardt, E., 2010. River restoration, habitat heterogeneity and biodiversity: a failure of theory or practice? Freshw. Biol. 55, 205–222.
Patrício, J., Neto, J.M., Teixeira, H., Salas, F., Marques, J.C., 2009. The robustness of ecological indicators to detect long-term changes in the macrobenthos of estuarine systems. Mar. Environ. Res. 68, 25–36.
Peet, R.K., 1975. Relative diversity indices. Ecology 56, 496–498.
Pietrobon, R., Taylor, M., Guller, U., Higgins, L.D., Jacobs, D.O., Carey, T., 2004. Predicting gender differences as latent variables: summed scores, and individual item responses: a methods case study. Health. Qual. Life Out. 2, 59.
R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria www.R-project.org
Roy, A.H., Rosemond, A.D., Paul, M.J., Leigh, D.S., Wallace, J.B., 2003. Stream macroinvertebrate response to catchment urbanisation (Georgia, USA). Freshw. Biol. 48, 329–346.
Sandin, L., 2009. The effects of catchment land-use near-stream vegetation, and river hydromorphology on benthic macro invertebrate communities in a south-Swedish catchment. Fundam. Appl. Limnol. 174, 75–87.
Schmedtje, U., Colling, M., 1996. Ökologische Typisierung der aquatischen Makrofauna. Informationsberichte des Bayerischen Landesamtes für Wasserwirtschaft 4/96.
Seeboonruang, U., 2012. A statistical assessment of the impact of land uses on surface water quality indexes. J. Environ. Manage. 101, 134–142.
Shannon, C.E., Weaver, W., 1949. The Mathematical Theory of Communication. The University of Illinois Press, Urbana.
Simpson, E.H., 1949. Measurement of diversity. Nature 163, 688.
Sponseller, R.A., Benfield, E.F., Valett, H.M., 2001. Relationships between land use spatial scale and stream macroinvertebrate communities. Freshw. Biol. 46, 1409–1424.
Statzner, B., Bis, B., Doledec, S., Usseglio-Polatera, P., 2001. Perspectives for biomonitoring at large spatial scales: a unified measure for the functional composition of invertebrate communities in European running waters. Basic Appl. Ecol. 2, 73–85.

Stephenson, J.M., Morin, A., 2009. Covariation of stream community structure and biomass of algae invertebrates and fish with forest cover at multiple spatial scales. Freshw. Biol. 54, 2139–2154.
Sundermann, A., Antons, C., Cron, N., Lorenz, A.W., Hering, D., Haase, P., 2011a. Hydromorphological restoration of running waters: effects on benthic invertebrate assemblages. Freshw. Biol. 56, 1689–1702.
Sundermann, A., Stoll, S., Haase, P., 2011b. River restoration success depends on the species pool of the immediate surroundings. Ecol. Appl. 21, 1962–1971.
Suren, A.M., McMurtrie, S., 2005. Assessing the effectiveness of enhancement activities in urban streams. II. Responses of invertebrate communities. River Res. Appl. 21, 439–453.
Tachet, H., Bournaud, M., Richoux, P., Usseglio-Polatera, P., 2000. Invertébrés d'eau douce: systématique, biologie, écologie. CNRS Editions, Paris.
Theodoropoulos, C., Iliopoulou-Georgudaki, J., 2010. Response of biota to land use changes and water quality degradation in two medium-sized river basins in southwestern Greece. Ecol. Indicators 10, 1231–1238.
Tockner, K., Pusch, M., Borchardt, D., Lorang, M.S., 2010. Multiple stressors in coupled river–floodplain ecosystems. Freshw. Biol. 55 (Suppl. 1), 135–151.
Townsend, C.R., Uhlmann, S.S., Matthaei, C.D., 2008. Individual and combined responses of stream ecosystems to multiple stressors. J. Appl. Ecol. 45, 1810–1819.
Visco, G., Campanella, L., Nobili, V., 2005. Organic carbons and TOC in waters: an overview of the international norm for its measurements. Microchem. J. 79, 185–191.
Wagenhoff, A., Townsend, C.R., Phillips, N., Matthaei, C.D., 2011. Subsidy-stress and multiple-stressor effects along gradients of deposited fine sediment and dissolved nutrients in a regional set of streams and rivers. Freshw. Biol. 56, 1916–1936.
Washington, H.G., 1984. Diversity, biotic and similarity indices. Water Res. 18, 653–694.