Deliverable D6.6

Version:	1.0
Date:	08/03/13 (amended 03/04/13)
Author:	UDE
Dissemination status:	PU
Document reference:	Deliverable_D6.6

Manuscript on the comparison of biodiversity response to environmental stressors
(M39)

STATUS: DRAFT MANUSCRIPT VERSION FOR SUBMISSION TO A PEER-REVIEWED JOURNAL

Project acronym:	BIOFRESH
Project name:	Biodiversity of Freshwater Ecosystems: Status, Trends, Pressures, and Conservation Priorities
Call and Contract:	FP7-ENV-2008-1
Grant agreement no.:	226874
Project Duration:	01/11/2009 – 30.04.2014 (54 months)
Co-ordinator:	Leibniz-Institute of Freshwater Ecology and Inland Fisheries at Forschungsverbund Berlin e.V., Germany
Partners:	RBINS, Royal Belgian Institute of Natural Sciences, Belgium BOKU, Universität für Bodenkultur Wien, Austria ICLARM, International Center for Living Aquatic Resources Management, Malaysia IRD, Institut de Recherche pour le Développement, France UDE, Universität Duisburg-Essen, Germany IUCN, International Union for Conservation of Nature, Switzerland UOXF.AC, Oxford University, UK UB, Universitat de Barcelona, Spain UFZ, Helmholtz Zentrum für Umweltforschung, Germany UCL, University College of London, UK EAWAG, Eidgenössische Anstalt für Wasserversorgung, Abwasserreinigung und Gewässerschutz, Switzerland UCBL, Université Claude Bernard - Lyon 1, France UPS, Université Paul Sabatier- Toulouse 3, France ECOLOGIC, Ecologic GmbH Institut für Internationale und Europäische Umweltpolitik, Germany EC-ERC, Commission of the European Communities - Directorate General Joint Research Centre, Italy UD, University of Debrecin, Hungary NRM, Naturhistoriska riksmuseet, Sweden CFKK, Center za kartografijo favne in flore, Slovenia

This project is supported by funding from the specific programme 'Cooperation', theme 'Environment (including Climate Change)' under the 7th Research Framework Programme of the European Union.

BIOFRESH

Biodiversity of Freshwater Ecosystems: Status, Trends, Pressures, and Conservation Priorities

Project no. 226874

Large scale collaborative project

D6.6: Manuscript on the comparison of biodiversity response to environmental stressors

Deliverable number	D6.6
Deliverable name	Manuscript on the comparison of biodiversity response to land use
WP no.	WP6
Lead Beneficiary (full name and Acronym)	University of Duisburg-Essen (UDE)
Nature	Manuscript (written report)
Delivery date from Annex I (proj. month)	**M39**
Delivered	Yes
Actual forecast delivery date	2013-03-08
Comments	**Draft manuscript version** This report deviates from the descriptions made in the DoW (p. 37: “Key component: Comparison of response signatures”) insofar that: • a focus had to be put on European data; the data used for this report were mainly derived from previous European research projects • a focus had to be put on land use/cover as the main stressor, as this is the only stressor variable available for all ecosystems considered at the European scale; other stressor variables will be analysed in context of D6.7 (due in Month 51)

Project funded by the European Commission within the Seventh Framework Programme Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Name of the Authors	Name of the Partner	Logo of the Partner
Christian K. Feld (CKF) Sonja Stendera (SS) Sebastian Birk (SB) Daniel Hering (DH)	UDE (lead)	UNIVERSITÄT DUISBURG ESSEN
Martin Kernan (MK) Jorge Salgado Bonnet (JSB)	UCL	UCL
Michael Gerisch (MG)	UFZ	HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH – UFZ
Florian Malard (FM) David Eme (DE)	UCBL	UCB Lyon 1
Florian Pletterbauer (FP)	BOKU	BOKU IHG

Summary

Worldwide, freshwater ecosystems are increasingly threatened by human-induced environmental impacts due to multiple land and water uses, with often severe adverse effects on freshwater ecosystem's ecological status and biodiversity. Thereby, land use alone is known to impose a multifaceted cocktail of stressors on the freshwater environment, for example, eutrophication, toxication, sediment pollution, physical habitat degradation.

The aim of this study is to compare the ecological implications of land use among a series of different—lotic as well as lentic—freshwater ecosystems, with a focus on biodiversity. Therefore, we generated a comprehensive database of freshwater organisms of five different ecosystem types: rivers (fish, macroinvertebrates, macrophytes), lakes (fish, macroinvertebrates, zooplankton), floodplains (molluscs, carabid beetles, floodplain vegetation), ponds (amphibians, macrozoobenthos) and groundwaters (crustaceans). Up to six biodiversity metrics (richness, Simpson and Shannon index, Pielou's evenness, taxonomic distinctness and endemicity/rareness) were generated using either the data on taxa occurrence, taxon presence/absence or abundance, if available, and were related to land use data derived from global and continental databases. In addition, variable sets of geographical (latitude, longitude, altitude) and climatic variables (actual evapotranspiration, mean annual air temperature and precipitation) were used for ecosystem-specific analysis in order to account for natural (geo-climatic) patterns in the spatial distribution of biodiversity.

The quantification of the proportions of variance in the different biodiversity metrics explained by land use and geo-climatic descriptors was deployed by generalised linear and additive regression modelling (GLM, GAM) and boosted regression trees (BRT).

Our results indicate that (natural) geo-climatic descriptors explained the majority of variance in all biodiversity metrics tested and in all ecosystem types, thus revealing a strong natural gradient in the data. This natural gradient, to some extent, interacts with land use and renders pure land use effects, i.e. the effects corrected for the natural gradient, on freshwater biodiversity comparatively weak (mostly <10% explained deviance). Despite the weak role of land use, clear and consistent thresholds of artificial (urban and industrial) land cover were found in river and groundwater systems, where biodiversity decreased at coverages >10%.

Among the metrics tested, species richness revealed the strongest link to both geo-climatic and land use in many cases. In addition, species endemicity/rareness (i.e. the relative spatial distribution of species among all sites/spatial entities expressed as the cumulated relative frequency of species within a community) performed comparatively well in river and groundwater systems.

Our results confirm strong effects of natural geo-climatic variables in the variance of freshwater biodiversity at pan-European and regional scales. These are likely to mask or interfere with other environmental effects, for example, with human-induced land use impact as addressed here.

Introduction

Worldwide, freshwater biodiversity continues is being threatened by multiple human impacts at the local as well as the continental and global scale (MEA, 2005; Vörösmarty et al., 2010). For example, among the multiple sources of environmental impact, almost all large rivers worldwide, with only a few exceptions, are fragmented by dams, which alter natural flows (Poff et al., 2007) and reduce the connectivity between upstream and downstream parts of entire catchments (see also Tockner et al., 2009 for a European perspective). Further, agricultural and other land uses can enhance nutrients and thus are often found to adversely affect the water quality of freshwater ecosystems (e.g. Allan, 2004).

In particular, land use is assumed to impose a serious driver of ecosystem change (e.g. Allan 2004; MEA, 2005; Feld, 2012), since land use is often linked to a multi-faceted complex of environmental impacts (stressors). Besides the eutrophication effects already mentioned before, intensive agricultural land use is often reported to promote erosion (Allan, 2004) and thus often results in fine sediment pollution of rivers, lakes and coastal waters situated within or downstream of agricultural landscapes. Further, the coverage of urbanised and industrialised area within catchments is linked to high amounts of impervious area (Paul and Meyer, 2001) and can severely modify discharge regimes in river or renewal of groundwater aquifers. Large cities are often located in the vicinity of rivers and lakes and can have severe impacts on these freshwater systems, for example, through treated and untreated wastewater pollution, intensive recreational activities and navigation. All these impacts can result from multiple land uses and are likely to change not only the environmental conditions of nearby freshwater ecosystems, but primarily also its ecological status and biodiversity (MEA, 2005; Feld, 2012; Feld, in revision).

While ecological and biodiversity effects of certain land uses are comparatively well known from studies of terrestrial ecosystems, much less empirical knowledge is available from freshwater ecosystems. Recent freshwater studies have mainly addressed lotic systems (Marchant, 2007; Feld, 2012) and floodplains (Gallardo et al., 2011; Gerisch et al., 2012), but have left lentic systems (lakes, ponds) and sub-terrestrial groundwater systems largely unconsidered.

For this study, we thus aimed to address the response of biodiversity to land use in several freshwater ecosystems. For the first time, to our knowledge, a comparison was done between response patterns of diversity in lentic and lotic ecosystems, including groundwater systems. Further, our comparison addressed several organism groups (fauna and flora) in part of the ecosystems considered, which enabled us to compare organism group-specific response patterns within several ecosystems.

The overall objective of this study was twofold: first, to quantify the relationship between land use and selected biodiversity metrics (expressed as variance in biodiversity explained by land use) and second, to compare the strength of the relationships found among organism groups (within ecosystems) and between organism groups (across ecosystems).

3 Generation of stressor data and biodiversity metrics

3.1 Land use

CORINE-derived proportional land use data (http://www. corine.dfd.dlr.de/papers_de.html) were used to quantify the relationship between human land use pressures due to, for example intensified agriculture or urbanisation, and freshwater biodiversity. CORINE data are based on satellite imagery (Landsat 7) and are available as vector or raster data for Europe at comparable resolutions. While the raster data are generalised and available at a minimum resolution of 100 x 100 m, vector maps provide a finer resolution of 25 x 25 m or better and, thus, were used in this study (see Tormos et al., 2011 for a critical discussion of spatial resolution in this context).

GlobCover land use (http://due.esrin.esa.int/globcover/) was used for the analysis of groundwater crustacean diversity instead, since CORINE does not sufficiently cover Eastern Europe. GlobCover Land Cover v2 is a global land cover map at 10 arc second (300 meter) resolution. For this study, we mainly focussed on CORINE level 1, i.e. the broadest classification level (e.g. arable, artificial areas) as this level allows for comparisons with GlobCover land use categories.

Ecosystem type-specific spatial scales were defined (Table 1) and delineated in a GIS system. The resulting areas (circles, polygons) were then projected on a layer containing the land use/cover information. Hence for each area considered, the proportion of CORINE/GlobCover land use types was available.

3.2 Biodiversity metrics

The compositional and structural diversity of biological assemblages was addressed by a total of six biodiversity metrics: taxon richness, taxon endemicity/rareness, taxonomic distinctness, Simpson’s and Shannon’s index and Pielou’s evenness. The former three are calculable based on binary taxon lists (presence/absence data) and the latter three address taxon density and thus require abundance values in addition. While taxon richness is a frequently-used indicator of ecosystem degradation in many freshwater studies (e.g. Birk et al. 2012), taxonomic distinctness has attracted increasing attention in recent years (e.g. Heino et al., 2007; Marchant et al., 2007; Gallardo et al., 2011; Simaika and Samways, 2011), since it adds a unique aspect to the taxonomic aspect of biodiversity not covered by taxon richness (e.g. Gallardo et al., 2011). Taxonomic distinctness refers to the mean taxonomic dissimilarity of any pair of taxa within a community along a Linnean phylogenetic tree (species, genus, family, order, class, phylum) (Warwick and Clarke, 1998). For instance, three species of the same genus are taxonomically less distinct as three species belonging to three different orders. An index of taxonomic endemicity has been defined by Crisp et al. (2001) and Linder (2001) and describes the summed 'rareness' of all species within a community (see Tisseuil et al., 2012 for a recent example of its application). The index is calculated as the sum of species present in a drainage basin weighted by the inverse of the number of drainage basins where the species occurs divided by the total number of species in the drainage basin. The index thus measures the proportion of 'endemics' in a drainage basin and is corrected for the richness effect. Here, the index was also applied to site-related (river) data, which is why it was named an index of rareness rather than of diversity in that case.

The three abundance-based diversity metrics (Shannon and Simpson index and evenness) are widely used to address the dominance structure and equality of community members.

3.3 Data analysis

In order to quantify the relationship between land use impact and biodiversity, we applied two different methods of analysis: generalised regression modelling (GRM) and boosted regression tree modelling (BRT). GRM encompass Generalised Linear Modelling (GLM) and Generalised Additive Modelling (GAM), both of which are applicable to non-parametric data, i.e. non-normal distributions either of input variables or of error distributions (residuals). While GLM is linked to predefined distributions of the response variable (linear, Poisson, Gaussian, logistic) and thus requires *a priori* knowledge about the type of relationship between predictor and response variables, GAM is a distribution-free regression technique. In GAM, the response distribution is actually identified using smoother techniques, for example smoothing by running mean values, neighbourhood smoothing or Kernel smoothing (see Franklin 2009 for a summary of smoothing techniques). Yet a fundamental prerequisite of all GRM methods is full data coverage, i.e. missing values are not allowed – neither among the predictor nor among the response variables.

This is different in BRTs, which allows of using incomplete predictor variables with many missing values. Furthermore, this technique allows of using categorical (nominal, ordinal) and continuous predictor variables in parallel, without any prerequisites of their distributions (Elith et al. 2008). Hence, BRT constitutes a relatively robust and flexible modelling technique.

BRT modelling is a machine learning method and applies multiple regression trees to the data, with the objective to maximise the deviation (variance) explained by the resulting model (Elith et al. 2008). Further, this method allows of identifying response patterns along environmental gradients (Cutler et al., 2007) and thus is capable of detecting potential thresholds at which the response variable either dramatically increases of declines (e.g. Clapcott et al., 2012; Feld 2012).

4 Ecosystem type-specific biodiversity response to land use

4.1 Rivers

4.1.1 Study sites and data

We used part of the WISER river database that contains biological and abiotic monitoring data from eight European ecoregions (Moe et al., 2012). Our subset comprised 1221 river sites distributed among two major ecoregion types in France, Germany and Austria: mountains (i.e. Western and Central Mountains acc. to Illies, 1978) and lowlands (i.e. Western and Central Plains). For each site, biological, geo-climatic and land use data was acquired.

Biological data originate from monitoring surveys targeting three organism groups: fishes, benthic invertebrates and aquatic macrophytes. The field methodology and field protocols followed the national monitoring standards (see Dahm et al., 2012 and Feld, 2012 for a summary). The number of sites differs between organism groups and ecoregion type (Table 2), since the monitoring surveys did not cover all organism groups at all sites in both ecoregion types. The resulting taxa lists per country were combined and manually adjusted for researcher-dependent bias (e.g. due to differences in the determination level achieved in two countries). Species-level was achieved for fishes and macrophytes, while genus level was used for benthic invertebrates; this level is the target level in France. The biodiversity metrics were then calculated based on the final adjusted taxa lists.

Selected geographic and climate variables (hereafter named geo-climatic variables) were used in the models to account for natural patterns in the distribution of biodiversity values (Table 2).

This selection is considered to sufficiently account for potential geographical (and country-specific) aspects as well as for longitudinal patterns in biodiversity along the river continuum (catchment size) and eventually for specific climatic differences within the Western and Central ecoregions (temperature, precipitation).

Human-induced land use impact was derived from CORINE land cover data following the basic methodology as described in the previous chapter. Therefore, polygons were automatically delineated in ArcGIS X (ESRI Inc., Redlands, California), encompassing the entire river network upstream of a given site and using digital elevation models to identify the margins of the river network. The catchment delineations were spot-checked for accuracy and finally combined with CORINE land cover data to derive the proportional land use data for the respective catchment areas.

Here, we focussed on those land use categories that either reflect considerable human impact in the ecoregion types considered (intensive agriculture [% arable land] and urban and industrial areas [% artificial area]) or represent rather extensive forms of land use (% pasture and forest) and natural land cover (% water bodies and wetlands) (see Supplementary Material Table S1).

4.1.2 Data analysis

We applied Boosted Regression Tree modelling to identify the relationship between natural (geo-climatic) and land use descriptor variables and the six biodiversity metrics. BRT proved to be advantageous over Generalised Regression Modelling (GRM) techniques in particular for two reasons: i) because missing values occurred in our data, which can be easily handled by BRT, but not by regression modelling and ii) because both natural and land use variables were correlated (multi-collinear).

We first ran BRTs using geo-climatic variables only in order to identify the natural pattern in the biodiversity data. The residuals of these analyses (i.e. the variance in the biodiversity metrics not explained by the geo-climatic variables) were then used in a second BRT together with land use variables to analyse the relationship between land use and biodiversity. This approach allowed of distinguishing two important results. First, we were able to partition the (non-targeted) variance in the biodiversity metrics attributable to geo-climatic variables from that (targeted) variance attributable to land use. And second, we were able to compare the relationship of different biodiversity metrics to both variable groups.

Yet, in contrast to classical variance partitioning, where unique, common and unexplained variances are distinguished (e.g. Feld and Hering, 2007), we neglected the potential effects of other sources of environmental deterioration in our study, for example of physico-chemical pollution and physical habitat modification. Both may interact with land use and hence in part may determine the patterns attributed to land use in this study. However, as the major objective in this entire study was to compare land use effects on aquatic biodiversity across multiple ecosystems, we considered this a minor flaw.

All data preparation and analysis was run in R 2.15.2 (R Development Core Team, 2012), using the libraries ‘gbm’ (v.0.7-2, Ridgeway, 2010) and ‘dismo’ (v.0.7-2, Hijmans et al., 2011) to run the BRTs.

4.1.3 Results

Relationship between the metrics of biodiversity

The cross-comparison of biodiversity metrics revealed similar patterns for all organism groups and is exemplarily shown for river macroinvertebrates in Figure 1. Accordingly, four different aspects of biodiversity are represented by the six metrics tested in this study: richness, dominance structure, rareness and distinctness. Richness was only moderately correlated with Shannon and Simpson diversity and rareness, and weakly correlated with the remaining metrics. Yet, slightly higher correlations were found between richness and the three measures of equity (Simpson and Shannon diversity and Pielou’s evenness) for fish and macrophytes. Contrastingly, Shannon and Simpson diversity and evenness were highly correlated with each other, showing that the aspect of equity (i.e. the even distribution of taxa abundances) is pretty much reflected by all three measures. Rareness only moderately correlated with richness, which suggests that the aspect of rare taxa within communities adds a separate aspect of diversity to the community’s overall richness. Finally, taxonomic distinctness was weakly correlated with all other metrics, thus revealing that the pairwise similarity of taxa along a Linnean phylogenetic tree also adds a unique aspect of diversity not covered by any other of the tested metrics.

General relationship of geo-climatic and land use variables and biodiversity

For all organism groups and in both ecoregion types, the geo-climatic variables alone explained a considerable amount of variance in most biodiversity metrics, and significantly more variance than land use variables (*U*-Test: $p < 0.01$ for all organism groups within both ecoregion types). The variance explained by the natural descriptors ranged 33.3–83.7% (mean ± SD: 57 ± 16.5%) in mountainous and 2–82% (mean ± SD: 43.1 ± 21.2%) in lowland rivers, contrasted to ranges of 0–33.3% (mean ± SD: 7.6 ± 9.4%) and 0–12.1% (mean ± SD: 2.4 ±

4.3%), respectively, detected for land use variables in both ecoregion types. With fishes, geo-climatic variables accounted for significantly more variance at mountain sites (mean ± SD: 62.1 ± 15%) compared to lowland sites (mean ± SD: 30.4 ± 17.7%) (*U*-Test: $p < 0.01$). These differences were less pronounced and insignificant for the other organism groups.

Comparison of biodiversity metrics in their response to geo-climatic and land use variables

Averaged over all organism groups and ecoregion types, the strongest relationship of both geo-climatic and land use variables was found for taxon richness (Figure 2). Nearly 70% of the variance in this metrics was explained by the geographical and climatic descriptors, contrasted to 8.3% of the remaining variance explained by land use. Also Shannon's diversity and rareness were closely related to geo-climatic variables, which explained more than 50% of the variance in these metrics. With all biodiversity metrics, however, less than 10% of the remaining variance (not explained by the geo-climatic descriptors) was explained by land use. Thus, in summary we found a considerably weak relationship of land use to the biodiversity metrics tested.

If separated by organism group and ecosystem type, however, the overall view was slightly different. The variance in the diversity measures explained by land use particularly increased with fishes and macroinvertebrates at mountain sites (Figure 3). Here, land use explained a third of the variance in macroinvertebrate rareness and almost a sixth of the variance in macroinvertebrate genus richness and distinctness. Also for fish richness and Shannon and Simpson diversity, between 15 and 20% of the variance were explained by land use, as opposed to average values between 3 and 8% across ecosystem types and organism groups. The variances explained by land use in the diversity of lowland assemblages, however, was negligible (Figure 3).

Thresholds for land use impact

As part of the outcome of the BRT analysis, Partial Dependence Plots (PDPs) display the individual response of each dependent variable (here: biodiversity metrics) along the environmental descriptors (here: land use variables). Although only little variance in the six biodiversity metrics was explained by land use in general, the PDPs detected largely consistent thresholds for some land use variables at mountain sites (plots not shown here). For instance, at as little as 10% artificial areas in the catchment (urban, industrial), fish and benthic invertebrate diversity often abruptly decreased, except for macroinvertebrate taxonomic distinctness, which increased with increasing artificial land cover. Less consistent, however, were the results for forest: while many macroinvertebrate diversity metrics increased between 10 and 40% catchment land cover as forest, fish diversity started to decrease, at values around 30%.

At lowland sites, the results were largely indistinct and thus are not further presented here.

4.1.4 Discussion

Relationship between the metrics of biodiversity

Among the six biodiversity metrics tested in this study, we found four different aspects of diversity addressed: taxonomic richness, equity (Shannon and Simpson diversity and Pielou's evenness), rareness and distinctness. These aspects were identified to be only weakly or moderately correlated with each other and thus address unique facets of taxonomic diversity. However, the three measures of taxonomic equity were highly correlated and thus in part redundant, which is in line with the findings of other studies (e.g. Gallardo et al., 2011). Our results are also supported by previous analyses using a different lowland dataset (CKF; unpublished data), which identified the same four aspects to be largely uncorrelated in macroinvertebrate data.

While the non-relation of taxonomic richness and distinctness has already been stated by Warwick and Clarke (1998) in their original paper on their index of taxonomic distinctness, the weak relation of rareness (or endemism according to Tisseuil et al., 2012) to the other diversity metrics has not yet been addressed in previous studies. Our results suggest that, although moderately correlated with richness (Figure 1), the index of rareness addresses another aspect of diversity, i.e. the relationship of rare and common taxa within a community. For instance, two communities may have the same number of species, but one may be characterised by many (rare) habitat specialists, while the other may contain more common (cosmopolitan) species. Hence, the separation of both aspects provides more insight into a community's diversity.

General relationship of geo-climatic and land use variables and biodiversity

Freshwater biodiversity is considered to be severely threatened by multiple (human-induced) pressures on freshwater ecosystems (MEA, 2005). Among them, intensive forms of land use (agriculture, urbanisation) impose strong impacts on rivers worldwide (MEA, 2005), because land use is often linked to river pollution (eutrophication by N and P fertiliser, pesticides, untreated domestic and industrial waste water), river fragmentation (water table regulation by dams, hydropower generation, flood protection upstream of urban areas), siltation (due to surface erosion on adjacent floodplains) and water withdrawal (for irrigation or drinking water supply). Although not all of these pressures can be assumed to have affected the sites investigated for this study, it is remarkable that less than ten percent of the variance in the biodiversity metrics was attributable to land use after the exclusion of geo-climatic effects.

Apparently, the latitudinal, longitudinal and altitudinal pattern in our data was partly concordant with a geographical land use gradient. This is supported by BRTs that were run with land use variables and original diversity metric values (i.e. the values uncorrected for geo-climatic effects; results not shown here). A preliminary data mining, however, revealed only slight correlations of percent land use as arable, artificial, forest and pasture with latitude, longitude and altitude (Spearman's correlation coefficient $r < 0.3$ for all combinations). Further, leaving such geo-climatic effects unconsidered would have notably flawed our results, as then we would have wrongly interpreted the natural patterns as being human-induced land use impacts.

A potential solution to this problem might have been to split the data into geo-climatically more homogeneous subsets, for instance by ecoregion (instead of ecoregion type) or country. Such a regional approach may probably reduce the variance explained by geo-climatic variables, as this gradient would then be shorter. On the other hand, however, this would inevitably impose an analytical drawback, since the individual subsets would contain fewer sites. Furthermore, many individual (regional) results may have rendered the large-scale comparison across several countries impossible, which is why we decided to follow the analytical approach as presented here.

Irrespective of the consistent and strong role of geo-climatic descriptors in the BRT models, we think that the weak role of land use was primarily attributable to the generally weak performance of biodiversity metrics with the detection of land use impacts on freshwater ecosystems. In a recent study on the relationships between land use gradients and structural and functional indicators of stream ecological integrity, Clapcott, et al. (2012) found macroinvertebrate and fish richness to be among the weakest indicators of land use effects. Death and Collier (2010) found taxonomic distinctness to be negatively related to natural (forest) land cover, because forested stream sections were dominated by less distinct taxa (mainly Ephemeroptera, Plecoptera and Trichoptera), while open streams were colonised by more distinct taxa (Oligochaeta, Mollusca, Crustacea, Arthrooda). Feld et al. (in revision) were able to show that a gradual species turnover along an environmental impact gradient renders many biodiversity metrics (richness, Shannon diversity, evenness, functional diversity, taxonomic distinctness) weak indicators of this impact. The authors found many degradation-tolerant taxa (e.g. Mollusca, Crustacea, Diptera) to be more frequent and abundant under increasing impact levels, thus replacing more sensitive taxa (e.g. Ephemeroptera, Plecoptera, Trichoptera, some Coleoptera). As a consequence, taxon richness, equity and distinctness remained largely constant along the impact gradient in their study.

Comparison of biodiversity metrics in their response to geo-climatic and land use variables

Averaged over three organism groups and two ecoregion types, taxon richness showed the strongest relation to both geo-climatic descriptors and land use (Figure 2), while land use in particular was found to explain the variance in the richness of mountain fishes and macroinvertebrates best (Figure 3). These findings suggest that taxon richness is still a useful indicator of the impact of environmental change on biodiversity, irrespective of the shortcomings of this metric reported by previous studies (e.g. Dornelas, 2010; Clapcott, et al., 2012; Feld et al., in revision). Yet, the high descriptive power of the seven geo-climatic variables used in our study also suggest that richness patterns require cautious interpretation if compared over regional and larger spatial scales.

Interestingly, also the index of macroinvertebrate rareness performed comparatively well at mountain sites, where land use explained 33% of the variance in this metric. We believe that this is primarily due to the high overall richness of macroinvertebrates in this ecoregion type, which is partly also connected to a high richness of relatively sensitive and often rare taxa (e.g. Ephemeroptera, Plecoptera, Trichoptera). This may also explain, why rareness poorly performed with fishes and macrophytes, both of which were much less rich and taxonomically diverse in our dataset (Supplementary Figure S1).

Although the results of our study revealed a generally weak explanatory role of land use variables in the biodiversity metrics tested, as part of the BRT standard output, the Partial Dependence Plots (PDPs) detected thresholds for some land use variables at which the diversity either abruptly increased or decreased. Of course, due to the weak descriptive performance of land use, the interpretation of the thresholds has to be cautious. But notably, the thresholds that were detected, for example, for % artificial were consistently around 10% and suggest that above this threshold, the diversity of fish and macroinvertebrates decreased. Thereby, our threshold is similar to the values reported by others (Feld, 2012; see also Allan, 2004 for a review). We therefore may infer that the relationships found here between land use and riverine biodiversity are interpretable to some degree, even if the overall explanatory power of land use was found to be extraordinarily low.

Two fauna biodiversity metrics showed consistently weak relationships with geo-climatic descriptors in both ecoregion types: Simpson diversity and evenness. This finding suggests that among the four aspects of diversity covered by our metric selection, the equity of the distribution of taxa was least related to geographic and climatic gradients in our data. The high correlation of Simpson diversity with evenness (Figure 1) confirms that Simpson diversity accounts for evenness much stronger than the other metrics.

With macrophyte metrics, the least relation to geo-climatic descriptors was found for taxonomic rareness and distinctness in both ecoregion types. In part, this may be owed to the comparatively low overall richness of macrophytes in both ecoregion types (Supplementary Figure S1), which do not allow for long gradients in these two metrics. On the other hand, however, the proportions of variance in both macrophyte metrics explained by geo-climatic descriptors are not much lower than those detected for the fauna. Rather, the remaining macrophyte metrics revealed a much stronger relationship to geo-climatic descriptors, compared to the fauna. Hence, the six biodiversity metrics did not reveal consistent patterns of response to geo-climatic changes, but revealed considerable differences between the fauna and the flora.

4.4 Wetlands and floodplains (MG)

4.4.1 Study sites and data

We conducted an extensive literature review in the Web of Science to obtain data on floodplain biodiversity and floodplain stressors. The following keywords were used, querying references not older than 1990 and only covering Europe: "floodplain*" "flood plain" OR "river*" OR "wetland*" AND "Carabid*" OR "Ground beetles" OR "Plant*" OR "Vegetation" OR "Mollusc*". Some of these articles provided information also for other taxonomical groups, which we included as well in our database.

To be selected, papers must include both the taxonomic and environmental information for the respective sites. We collected information on various types of human and environmental stressor variables, mainly hydrological, hydromorphological, land-use, and climatic variables. In cases where species lists where given for each site, but environmental information was provided only for site clusters, we pooled the species data of each environmental cluster to be able to link both data types. From a total of 102 papers, only 78 fulfilled our requirements and were used to compile the information in a database.

We considered three organism groups in this study: ground beetles (Carabidae), floodplain vegetation, and snails and bivalves (Mollusca). These groups reflect different ecological and life history adaptations to floodplain dynamism and this are assumed to respond different to the stressor gradients investigated. For each of the group we calculated species richness, Simpson and Shannon diversity, and Pielou's evenness. If abundance data were missing, only species richness was calculable. Our dataset on floodplain biodiversity and stressor variables consisted of 611 different sampling sites in 21 countries, distributed among 51 rivers and 28 catchments. In total, our database comprises 1,268 plant species, 357 ground beetle species, and 208 mollusc species (Table 3).

As the majority of the taxa reported in the literature were species, we decided to use species-only data for our analysis. The sampling protocols for the individual studies were largely comparable, although, for instance, he sampling days of pitfall trapping of ground beetles ranged considerably between the studies. We did not correct for these differences, however, because i) the information provided was insufficient to apply a consistent procedure of standardisation (e.g. indication of broad time-frames and distinct sampling days, pooled taxa lists *vs.* trap-specific taxa lists) and ii) we were not interested in predictive modelling but rather addressed the comparability of patterns in the biodiversity response of different organism groups to land use.

The richness and diversity of species are not only a result of anthropogenic land use and other impacts, but is also strongly linked to natural environmental descriptors, such as climate and altitude. To account for such natural gradients in our data, we included geographic and climate data for each site (Table 4).

CORINE land cover (level 1 classification) was derived within a radius of 5,000 m around each site and calculated as the relative amount of the area occupied by the different land use categories. We used 5,000 m as a compromise threshold for the buffer-size because land use variability in an area of this size should match the dispersal behaviour of both mobile and less mobile organism groups (e.g., carabid ground beetles *vs.* Mollusca). We used level 1 data, as this coarsest level integrates over similar categories in a hierarchical structure (Table 5). Preliminary analyses showed that CORINE level 2 or even fines classification tends to increase the variability in the data, while the number of suitable data points for subsequent analyses decreases.

Among all sites, we removed those land use categories, which sum up to less than 5%. To estimate land use on a local scale, we used the habitat type, which in almost all cases was provided directly in the articles. Prior to the analyses, we also checked the land use categories for collinearity and excluded those with a correlation > 0.6.

4.4.2 Data analysis

We applied partial linear regression to partial out the undesired natural (geo-climatic) effects from the targeted land use effects. This technique was applied to both the diversity metrics and the land use variables, the latter of which also revealed geographical patterns. Land use, to some degree, is a function of altitude, precipitation, and temperature.

A stepwise approach was then applied to each organism group. First, we selected subsets for each organism groups covering only those sites where the targeted taxa were recorded. A Principal Components Analysis (PCA) was applied to the geo-climatic descriptors of each subset and the PCA scores (only axis 1 and 2) were derived for each site. The score represented the main gradients of natural variability.

Linear regression modelling (Simpson and Shannon diversity, evenness) and Generalised Linear Modelling (GLM with Poisson error distribution: only species richness) were then used to analyses the variance in the biodiversity metrics that was explained by the major geo-climatic gradients. Similarly, we also regressed each land use category against both major geo-climatic PCA axes in order to analyse the variance in the land use data that was explained by the natural descriptors.

Eventually, for each organism group, we regressed the residuals of the biodiversity model (i. e. the biodiversity metrics corrected for geo-climatic effects) against those of the corrected land use models. All data met the requirements of simple linear regression, with normally-distributed residuals. The final model selection was done by step-wise removal of non-significant variables until all remaining descriptors were significant. The final models were compared using the AIC and the log-likelihood.

All data preparation, manipulation, analyses, and graphical representations were run in R 2.15.2 (R Development Core Team, 2012).

4.4.3 Results

Overall, the models for molluscs performed best with regard to the explained deviance (mean value over the four diversity metrics: 30%). This was comparably high for Pielou’s evenness and Simpson's diversity. Floodplain vegetation performed less well; on average land use explained only 8% of the residual deviance in vegetation diversity. The models for ground beetles, on average, explained about 18% of the diversity, which places them between molluscs and floodplain vegetation (explained deviation: molluscs > ground beetles > vegetation).

The response of the three organism groups to land use (corrected for geo-climatic effects) was inconsistent and differed notably (Table 6). For example, ground beetle and mollusc diversity decreased with increasing urbanisation, while the diversity of floodplain vegetation partly increased with an increasing amount of artificial surfaces. Vegetation diversity (Simpson's and Shannon's index) and ground beetle richness were found to increase stronger in areas with high amounts of wetland habitats. But mollusc and ground beetle diversity considerably decreased in such habitats. Ground beetle communities were also more diverse in areas with high agricultural activities (usually meadows), whereas the other groups did not respond to this land use category.

The results also show that diversity metrics within taxonomic groups respond differentially to similar land use impacts. For example, vegetation richness was positively related to the amount of semi-natural areas and forests, while Pielou's evenness and Simpson's diversity decreased. Molluscs showed a different response to the amount of wetlands and agricultural areas: species richness increased with both land use categories, but the two diversity metrics showed a negative relationship.

But there are also similarities, as for example species richness and various diversity metrics of ground beetles decrease with increasing amount of artificial surfaces or wetlands.

4.4.4 Discussion

To our knowledge, this is the first study that evaluates the response of multiple taxonomical groups to land use gradients across a variety of different habitats in European floodplain ecosystems. Our main finding is that land use affects biodiversity in an irregular, strongly taxa- and habitat-dependent matter and that the diversity of ecological communities is strongly linked to the intensity of anthropogenic stressors, especially hydrological alteration and urbanisation. Across the taxonomic groups studied, we prove that ecological communities in floodplains respond sensitively to ecological degradation and assume the different life strategies among the taxonomical groups as being responsible for this. We also discuss that relevant diversity signals might be overlooked due to limitations of standard biodiversity measures.

In this study, the degree of urbanisation was found to have the strongest impact on floodplain biodiversity. Together with agricultural land use, (urban) artificial surfaces explained up to 50% of the variance in species diversity. Thereby, intensive land uses often come along with river regulation, which has led to marked habitat losses in and fragmentation of floodplains worldwide, with mainly negative impacts especially on plant diversity (Uowolo et al., 2005; Houlahan et al., 2006). Our results, however, did not support this pronounced impact on the flora, as we found only plant species richness responding significantly to increasing amount of artificial surfaces. Other plant diversity measures were not significantly related to urbanisation or agricultural land use. Also Pretto et al. (2012) found plant species richness to increase along urbanisation gradients and explained that urbanisation causes also the establishment of a large proportion of non-native plant species. They assumed concurrent factors that come along with urbanisation, thus outweighing the natural environmental effects (e.g. climate and lithology) and changing the composition of established communities. Lososova and Lanikova (2010) restricted such findings not only to non-native species, but concluded that although urbanisation generally results in a loss of natural habitats, there are new man-made habitats potentially suitable also for native species. Here we argue similarly, stating that floodplain alterations homogenises habitat conditions and therefore enables several other, often ubiquitous species to establish in these cultivated or even artificial systems.

Contrary to increasing species richness, plant diversity was found to decrease strongly with the amount of forests and semi-natural areas. One should be aware, however, that forests by nature exhibit a completely different species pool than, e.g. grasslands or other semi-natural areas. Therefore, lower plant diversity of sites being more forest rich is not per se a negative signal. It also indicates the intrinsic diversity levels of the ecosystems.

Contrary to plants, urbanisation decreased most of the diversity measures of ground beetles and molluscs, which clearly supports previous studies postulating adverse effects of urbanisation on invertebrate communities. Jones and Leather (2012) reviewed the effects of urbanisation on invertebrates and found an impressive number of studies reporting that different taxonomic groups decreased in species richness and diversity due to fragmentation, isolation, habitat loss, high disturbance levels, surrounding land use, traffic, or pollution. This suggests that direct and indirect effects of urbanisation have larger impacts on the diversity of taxonomic groups occupying higher trophic positions. Johnsen et al. (2013) argued similarly, assuming that actively dispersing predators associated more negatively with urban systems relative to herbivores with passive dispersal. But there are also exceptions to this general adverse effect of urbanisation on invertebrate diversity. A considerably high invertebrate diversity, for example, has been found in habitat corridors or at habitat edges in highly urbanised areas (e.g. Magura et al., 2004). Alike, we found ground beetle evenness in this study to increase with urbanisation, thus suggesting that sufficient habitats and resources were available for ground beetles even under urban impacts.

Although the amount of deviance in biodiversity measures explained by land use was comparably low, this contribution was significant. This was especially pronounced in plant communities, where only approx. 10% of the diversity could be explained by land use. In contrast, regression models of mollusc evenness explained almost 50%, which was exceptionally high. These findings support current knowledge claiming that divergent response patterns of different organism groups to urbanisation depend on major life-strategies of the organisms, their trophic positions and their ability to establish populations in secondary habitats (Vallet et al., 2010; Iikin et al., 2012). However, a close look into the species lists of our dataset also revealed that species composition and ecological properties strongly change along anthropogenic gradients, which cannot be unravelled by standard diversity measures like species richness or Shannon diversity. Hence, our results support other studies claiming that the indicative potential of standard diversity measures is fairly limited (e.g. Feld et al. 2009; Fiedler and Truxa, 2012). The main constraint of using mere species' numbers or abundance-related indices is that they quantify community characteristics, but do not qualify community changes. Consequently, a considerable species turnover along an environmental gradient inevitably remains undetected, if neither species richness nor the species' abundances change considerably.

Moreover, biodiversity standard measures are strongly dependent on sampling effort and thus often biased, if based on different sampling methodology. Our dataset too, comprised studies with varying sampling efforts so that direct comparisons of diversity values between studies should be done with caution and may require additional treatment, such as rarefaction, before they can be used for regression modelling. In addition, geographical gradients may add bias to the data. We found plant species richness to reveal a latitudinal gradient that can be easily explained with altitudinal and climatic changes from northern mountain ranges (low diversity) to southern Mediterranean floodplains (high diversity). Hence, geo-climatic variables can superimpose land use effects, which may require separate analyses of smaller, geo-climatically more homogeneous subsets. Therefore, we aimed to minimise this (untargeted) bias by partialling out such effects from the (targeted) land use effects in our study.

4.5 Groundwater (FM/DE)

4.5.1 Study sites and data

A total of 21,700 occurrences of 1,571 obligate groundwater crustacean species were assembled from the literature, pre-existing databases, and personal collections. This data set was specifically assembled within the framework of BioFresh (Malard et al., 2012). Crustacean groups included in the data set were: Ostracoda, Cladocera, Copepoda, Bathynellacea, Amphipoda, Isopoda, Decapoda and Thermosbeanacea. Species occurrences were projected on a European map and species richness and endemicity were calculated for a total of 701 cells of 100 × 100 km (Figure 4). Endemicity was calculated as the sum of species present in a cell weighted by the inverse of the number of cells in which the species occurs divided by the total number of species in the cell (Tisseuil et al., 2012).

Actual evapotranspiration (AET), mean annual air temperature (T) and habitat heterogeneity (HH) were considered as additional variables to take into consideration the effect of energy, climate and habitat heterogeneity on species diversity (Figure 4). Average values were calculated for each cell. AET and T were obtained from the CGIAR Global High-Resolution Soil-Water Balance (http://www.cgiar-csi.org/data/global-high-resolution-soil-water-balance) and WorldClim Global Climate Data (http://www.worldclim.org), respectively. Habitat heterogeneity was calculated for each cell using the Shannon index ($H = - \Sigma\, pi \times \ln pi$), where pi represented the areal proportion of 13 groundwater habitat categories. Habitat categories were derived from the groundwater habitat map of Europe specifically elaborated within the framework of BioFresh (Cornu et al., 2013).

Land use data were obtained at the scale of Europe from the GlobCover land cover map (http://due.esrin.esa.int/globcover/). The areal proportion in 100 × 100-km cells of the following classes was calculated: urbanized areas, arable land, extensive agriculture, forests, shrubs, open areas (grassland and bare areas), and wetlands. A principal component analysis (PCA) was performed using the proportion of these classes in 526 cells covering the whole Europe except Scandinavia. The first axis of PCA was strongly correlated to the proportion of arable land and extensive agriculture, whereas the proportion of urbanized areas was correlated to axis 4 (Figure 5). Therefore, the areal proportion of agriculture (arable land + extensive agriculture) and the proportion of urbanized areas were retained as human stressors in the analysis of biodiversity patterns at the scale of Europe (Figure 4).

4.5.2 Data analysis

Boosted regression trees (BRT), generalized linear models (GLM) and generalized additive models were used to test for the effect of land uses (agriculture and urbanization) and natural variables (AET, T and HH) on species richness and endemicity. Cells with a land area < 20% were excluded from the analysis as well as cells corresponding to Scandinavia (see Figure 2). Thus, models were performed using a total of 526 cells. Endemicity was arc sin-transformed prior to analysis. GLM and GAM were performed using a negative binomial law for species richness (i.e. count) and a Gaussian law for the arc sin-transformed values of endemicity. Quadratic terms were considered in GLM and interactions among variables were assessed in all models. Cubic regression splines were used as smoothing function in GAM and interactions were tested by means of tensor product (te).

All analyses were performed with R 2.14.1 (R Development Core Team, 2011), using the libraries ‘ade4’ (v.1.5-1, Dray et al., 2007) for multivariate analyses, ‘mgcv’ (v.1.7-22, Wood, 2011) for GAM, MASS (v. 7.3-23, Venables and Ripley, 2002) to fit negative binomial distributions with GLM, and ‘gbm’ (v.0.7-2, Ridgeway, 2010) and ‘dismo’ (v.0.7-2, Hijmans et al., 2011) to perform BRT.

4.5.3 Results

The amount of deviance explained by the models varied from 59.6 to 24.4 %, with the highest deviance being explained by BRT (Table 8). All models explained a higher deviance for species richness than for endemicity.

The proportions of agricultural and urbanized lands were retained as significant variables in all models (Table 9 and 10; see also Supplementary Material, Table S2 for model details in GLM and GAM). Species richness was maximal for low values of urbanized areas (<10%) and decreased with increasing proportion of agriculture (Figure 6). Endemicity decreased with increasing proportions of agriculture and urbanized areas (Figure 7). For species richness, agriculture was found to interact with temperature in GLM and GAM, but this interaction was not apparent in BRT.

The amount of deviance explained by land use variables was typically smaller than that explained by natural variables (Tables 9 and 10). Species richness was primarily explained by AET and habitat heterogeneity (Table 9, Figure 6), with a significant interaction between these two variables (Figure 8, right panel). Regional species richness increased with increasing AET, this increase being typically more pronounced when habitat heterogeneity was high. Similarly, temperature was found to interact with AET: richness peaked when AET was high and temperature was neither cold nor high. (Figure 8, left panel).

Endemicity was primarily related to and increased with temperature (Figure 7). Endemicity also increased with AET for AET values > 600 mm.

4.5.4 Discussion

The present study provides the first comprehensive analysis of diversity patterns in the European groundwater fauna at a spatial resolution finer than biogeographical regions (Hof et al., 2008) or countries (Gibert and Culver, 2005; Stoch and Galassi, 2010). The most striking features of groundwater crustacean diversity patterns emerging from this study can be summarized as follows. First, species richness shows a hump-shaped latitudinal pattern with a peak at latitudes of ca. 42–46°, whereas endemicity markedly decreased north of the 46th latitude. Second, a significant proportion of spatial variance in species richness (59.6%) and endemicity (38.8%) could be explained using a relatively small number of explanatory variables. Third, patterns of species richness and endemicity are predominantly shaped by natural (geo-climatic) variables although the effect of human stressors expressed in this study as the proportion of urbanized areas and agriculture was statistically significant.

Species richness primarily responded to variation in energy (e.g. AET) and habitat heterogeneity. This finding is consistent with the widespread idea that the number of species in groundwater should primarily be restricted by the amount of energy and number of available habitats (Gibert et al., 1994; Malard et al., 2009). However, at this stage of our analysis, which was essentially correlative, we can hardly determine the precise way by which energy and habitat heterogeneity influence the processes that ultimately control the number of species in a place: speciation, extinction and dispersal (Currie et al., 2004; Wiens, 2011). Culver et al. (2006) hypothesized that the ridge of high biodiversity in the subterranean, terrestrial fauna might correspond to regions of low extinction rate, where productivity remained high and did not experience major decreases in recent geological times. Alternatively, according to the biotic interaction hypothesis (Schemske, 2002; Currie et al., 2004), the rate of speciation may also be higher in resource-rich and heterogeneous areas, in which the portion of variation in the taxa fitness due to abiotic factors decreases relative to biotic interactions.

There was a latitudinal threshold at ca 46° N above which endemicity markedly decreased with increasing latitude. This latitudinal pattern was best explained by variation in actual mean annual air temperature, which is linearly correlated to mean annual groundwater temperature in the heterothermic zone (Mermillod et al., 2013). This suggests that the proportion of rare species in a community is determined by physiological constraints imposed by present-day climatic conditions. Yet rather we suggest that the observed latitudinal pattern reflects the selective effect of long-term climatic oscillations for generalism and vagility (Dynesius & Jansson, 2000). If area, through population size, largely determines the extinction rate (Hugueny et al., 2011), endemic species must have endured severe loss in northern regions of cold Pleistocene climates. This historical effect can persist for longer among groundwater species showing low dispersal ability, because it is less likely to be overwritten by subsequent dispersal phases (Foulquier et al., 2008; Hof et al., 2008). This implies that our analysis of patterns of endemicity would likely explain a much greater proportion of deviance if it incorporates variables accounting for historical variability of climatic conditions (Leprieur et al., 2011; Baselga et al., 2012).

Despite the low amount of deviance explained by land use variables, we found consistent response curves of richness and endemicity to human stressors. Both species richness and endemicity in groundwater decreased with higher proportion of urbanized areas and agriculture, a finding that was entirely consistent with that reported for rivers (see section 4.1). Species richness exhibited a threshold effect whereas the response of endemicity was linear. A number of site-specific studies suggested that human disturbance resulting from urbanization and/or agriculture may decrease the number of obligate groundwater taxa via groundwater warming (Foulquier et al., 2011), anoxification (Malard et al., 1996), or interference competition with epigean taxa, whose penetration at depth into the surface is promoted by increased organic matter flux (Datry et al., 2005). Our results indicated that endemicity rather than species richness might be a better indicator of human disturbance, perhaps because narrowly-distributed species have narrower niches than widely-distributed species. A next important step for gaining better insight into the relationships between biodiversity and human disturbances would be to incorporate metrics that retain the identity of species (e.g. beta diversity) and/or their functional traits.

5 Comparison of biodiversity response patterns across ecosystems (meta analysis of response strengths and thresholds)

For this draft manuscript, three different freshwater ecosystems have been considered, with different organism groups in focus, so that the comparison of results requires cautious interpretation. Not only were the spatial scales different between groundwaters (100 x 100 km grid), rivers (catchments, mainly 10–500 km^2) and wetlands (circles of 5 km radius around each site), but also the biological data. While mainly species occurrence data (but not proven absences) were available for groundwaters, we hold abundance data for all organism groups considered in the remaining ecosystem types.

Despite these differences, however, several findings deserve consideration, as they revealed some level of concordance among ecosystem-specific results. First of all, a considerably high proportion of the variance in freshwater diversity was explained by (natural) geo-climatic variables. We infer that, in the realm of broad-scale studies (like ours), natural descriptors tend to explain most of the variance in diversity and potentially may interfere with (or mask) the effects of other environmental descriptors, such as land use in our study. The partitioning of natural effects from land use effects thus, was of paramount importance in our study.

Second, although the average variance in species richness and endemicity/rareness was low (< 10%; compare Table 9 and 10, Figure 2), both metrics showed a notable relation to land use, in particular with benthic invertebrates in mountain rivers (Figure 3).

Third, species richness was the biodiversity metric best related to land use in river and groundwater systems. This is notable, because mere richness is often considered a weak indicator of diversity-environment relationships (Feld et al., 2011; Pereira et al., 2013), since this metric does not account, for example, for species equity (or dominance structure) and its functional or phylogenetic diversity (Feld et al., in revision). This finding may be a result of the spatial scale of our study, which was rather broad compared to that of Feld et al. (in revision). We assume that measures of equity (such as Shannon diversity or Pielou’s evenness) are less relevant at continental or sub-continental levels, because the underlying paradigm of evenness (i.e. evenly distribution of members within a community) does not apply to the broad scale. That is, similar to alpha diversity, evenness refers to the level of a single community, while the concept (or paradigm) underlying the measure of evenness may not be transferable to broader spatial scales. Rather, adaptations might be necessary also for evenness, such as the consideration of beta or gamma diversity, when it comes to the comparison of diversity indices across broader scales.

Fourth, the thresholds found for artificial (urban) land use impact on macroinvertebrate diversity in rivers and crustacean diversity in groundwaters were similar. In both ecosystem types, the metrics decreased (mostly abruptly) at levels as low as 10% artificial area (less pronounced for groundwaters). A general trend of diversity decrease was also detectable for % as arable in both ecosystem types, but no distinct threshold was detectable based on our data. The results support previous findings (see e.g. Paul and Meyer, 2001 for a review) and confirm that in particular the proportion of artificial (i.e. urbanised and industrialised) area in the catchment of freshwater ecosystems is a good proxi of the ecological implications linked to urbanisation, for instance, due to the modification of discharge regimes in largely impervious landscapes.

And fifth, macroinvertebrate and mollusc diversity were stronger related to land use in rivers and floodplains, respectively, compared to fish and carabid beetles in both ecosystem types. Hence, among the fauna groups, the immobile taxa were stronger related to land use than the mobile ones. This suggests that rather immobile taxa are better indicators of land use, probably because their ability to escape adverse conditions due to this stressor is limited. Yet, further investigation is necessary to test this hypothesis.

6 Acknowledgements

This work was supported by the European Commission through the BioFresh project (“Biodiversity of Freshwater Ecosystems: Status, Trends, Pressures, and Conservation Priorities”), FP7-ENV-2008, Contract no. 226874. The European groundwater crustacean dataset was assembled as part of a collaborative project involving the Universities of Ljubljana, Slovenia (M. Zagmajster & C. Fiser), L'Aquila, Italy (D. Galassis & F. Stoch) and Lyon, France (F.M., D.E., and P. Marmonnier). The natural environmental variables used for the analysis of groundwater biodiversity patterns were provided by J.-F. Cornu at Muséum National D'Histoire Naturelle, UMR BOREA IRD-CNRS, Paris, France.

7 References

Allan J.D. (2004) Landscapes and riverscapes: the Influence of land use on stream ecosystems. Annual Review of Ecology, Evolution and Systematics, 35, 257–284.

Baselga A., Gomez-Rodrıguez C., Lobo J.M. (2012) Historical legacies in world amphibian diversity revealed by the turnover and nestedness components of beta diversity. PLoS ONE 7, 323–341. doi:10.1371/journal.pone.0032341.

Birk S., Bonne W., Borja A., Brucet S., Courrat A., Poikane S., Solimini A., van de Bund W., Zampoukas N., Hering D. (2012) Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecol Ind 18: 31–41. doi:10.1016/j.ecolind.2011.10.009.

Clapcott J.E., Collier K.J., Death R.G., Goodwin E.O., Harding J.S., Kelly D., Leathwick J.R., Young R.G. (2012) Quantifying relationships between land-use gradients and structural and functional indicators of stream ecological integrity. Freshwater Biology, 57, 74–90.

Cornu J.-F., Eme D., Malard F. (2013) The distribution of groundwater habitats in Europe. Hydrogeology Journal (in press).

Crisp, D.T., Laffan, S., Linder, H.P. & Monro, A. (2001) Endemism in the Australian flora. Journal of Biogeography, 28, 183–198.

Culver D.C., Deharveng L., Bedos A., Lewis J.L., Madden M., Reddell J.R., Sket B., Trontelj P., White, D. (2006) The mid-latitude biodiversity ridge in terrestrial cave fauna. Ecography, 29, 120–128.

Currie D.J., Mittelbach G.G., Cornell H.V., Field R., Guégan J.-F., Hawkins B.A., Kaufman D.M., Kerr J.T., Oberdorff T., O'Brien E., Turner J.R.G. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecology Letters, 7, 1121–1134.

Cutler D.R., Edwards T.C. Jr, Beard K.H., Cutler A., Hess K.T., Gibson J. et al. (2007) Random forests for classification in ecology. Ecology, 88, 2783–2792.

Dahm, V., Hering, D., Nemitz, D., Graf, W., Schmidt-Kloiber, A., Leitner, P., Melcher, A., Feld, C.K. (2013) Effects of physico-chemistry, land use and hydromorphology on three riverine organism groups: a comparative analysis with monitoring data from Germany and Austria. Hydrobiologia, 704, 389–415.

Datry T., Malard F., Gibert, J. (2005) Response of invertebrate assemblages to increased groundwater recharge rates in a phreatic aquifer. Journal of the North American Benthological Society, 24, 447–460.

Death R.G. & Collier K.J. (2010) Measuring stream macroinvertebrate responses to gradients of vegetation cover: when is enough? Freshwater Biology, 55, 1447–1464.

Dornelas, M. (2010) Disturbance and change in biodiversity. Phil. Trans. R. Soc. B, 365, 3719–3727. doi: 10.1098/rstb.2010.0295.

Dray S., Dufour A.B., Chessel D. (2007): The ade4 package-II: Two-table and K-table methods. R News. 7(2): 47–52.

Dynesius M., Jansson R. (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proceedings of the National Academy of Sciences of the United States of America, 97, 9115–9120.

Elith J., Leathwick J.R. & Hastie T. (2008) A working guide to boosted regression trees. Journal of Animal Ecology, 77, 802–813.

Feld C.K. (2012) Response of three lotic assemblages to riparian and catchment-scale land use: implications for designing catchment monitoring programmes. Freshwater Biology. DOI: 10.1111/fwb.12077.

Feld C.K., Hering D. (2007) Community structure or function: effects of environmental stress on benthic macroinvertebrates at different spatial scales. Freshwater Biology, 52, 1380–1399.

Feld C.K., de Bello F., Dolédec S. (in revision): Species and trait replacements imply weak response of biodiversity to hydromorphological alteration: a case study using lowland river invertebrate communities.

Feld C.K., Martins da Silva P., Sousa J.P., deBello F., Bugter R., Grandin U., Hering D., Lavorel S., Mountford O., Pardo I., Pärtel M., Römbke J., Sandin L., Jones K.B., Harrison P.A. (2009) Indicators of biodiversity and ecosystem services: a synthesis across ecosystems and spatial scales, Oikos, 118, 1862–1871. doi: 10.1111/j.1600-0706.2009.17860.x.

Feld, C.K., Birk, S., Bradley, D.C., Hering, D., Kail, J., Marzin, A., Melcher, A., Nemitz, D., Pedersen, M.L., Pletterbauer, F., Pont, D., Verdonschot, P.F.M. & Friberg, N. (2011) Chapter 3 - From Natural to Degraded Rivers and Back Again: a Test of Restoration Ecology Theory and Practice, Advances in Ecological Research, 44, 119–209.

Fiedler K., Truxa C. (2012) Species richness measures fail in resolving diversity patterns of speciose forest moth assemblages. Biodiversity and Conservation, 21, 2499–2508. doi: 10.1007/s10531-012-0311-5.

Foulquier A., Malard F., Lefebure T., Gibert J., Douady C.J. (2008) The imprint of Quaternary glaciers on the present-day distribution of the obligate groundwater amphipod Niphargus virei (Niphargidae). Journal of Biogeography, 35, 552–564.

Foulquier A., Malard F., Mermillod−Blondin F., Montuelle B., Dolédec S., Volat B., Gibert J. (2011). Surface water linkages regulate trophic interactions in a groundwater food web. Ecosystems, 14, 1339–1353.

Franklin, J. (2009) Mapping Species Distributions: Spatial Inference and Prediction. Cambridge: Cambridge University Press, 320 pp.

Gallardo B., Gascón S., Quintana X., Comín F.A. (2011) How to choose a biodiversity indicator - Redundancy and complementarity of biodiversity metrics in a freshwater ecosystem. Ecological Indicators, 11, 1177–1184. doi:10.1016/j.ecolind.2010.12.019.

Gerisch M., Agostinelli V., Henle K., Dziock F. (2011) More species, but all do the same: contrasting effects of flood disturbance on ground beetle functional and species diversity. Oikos 121: 508–515. doi:10.1111/j.1600-0706.2011.19749.x.

Gibert J., Culver D.C. (2005) Diversity Patterns in Europe. In: Encyclopedia of Caves (Eds D.C. Culver & W.B. White). Elsevier Academic Press, New York : 196–201.

Gibert J., Stanford J.A., Dole-Olivier M.J., Ward, J.V. (1994) Basic attributes of groundwater ecosystems and prospects for research. In J. Gibert, D. L. Danielopol and J. A. Stanford (editors). Groundwater Ecology. Academic Press, San Diego, California: 7–40.

Heino J., Mykrä H., Hämäläinen H., Aroviita J., Muotka T. (2007) Responses of taxonomic distinctness and species diversity indices to anthropogenic impacts and natural environmental gradients in stream macroinvertebrates. Freshwater Biology 52, 1846–1861. doi:10.1111/j.1365-2427.2007.01801.x.

Hijmans R.J., Phillips S., Leathwick J. & Elith J. (2011) dismo: Species Distribution Modeling. R Package Version 0.7-2. Available at: http://CRAN.R-project.org/package=dismo.

Hof C., Brändle M., Brandl, R. (2008) Latitudinal variation of diversity in European freshwater animals is not concordant across habitat types. Global Ecology and Biogeography, 17, 539–546.

Houlahan J.E., Keddy P.A., Makkay K., Findlay C.S. (2006) The effects of adjacent land use on wetland species richness and community composition. Wetlands, 26, 79–96. doi: 10.1672/0277-5212(2006)26[79:TEOALU]2.0.CO.

Hugueny B., Movellan A., Belliard J. (2011) Habitat fragmentation and extinction rates within freshwater fish communities: a faunal relaxation approach. Global Ecology and Biogeography, 20, 449–463.

Ikin K., Knight E., Lindenmayer D.B., Fischer J., Manning A.D. (2012) Linking bird species traits to vegetation characteristics in a future urban development zone: implications for urban planning. Urban Ecosystems, 15, 961–977. doi: 10.1007/s11252-012-0247-2.

Illies J. (1978) Limnofauna Europaea. Gustav Fischer, Stuttgart.

Johnson P.T.J., Hoverman J.T., McKenzie V.J., Blaustein A.R., Richgels K.L.D. (2013) Urbanization and wetland communities: applying metacommunity theory to understand the local and landscape effects. Journal of Applied Ecology, 50, 34–42. doi: 10.1111/1365-2664.12022.

Jones E.L., Leather S.R. (2012) Invertebrates in urban areas: A review. European Journal of Entomology, 109, 463–478.

Leprieur F., Tedesco A.P., Hugueny B., Beauchard O., Dürr H.H., Brosse S., Oberdorff, T. (2011) Partitioning global patterns of freshwater fish beta diversity reveals contrasting signatures of past climate changes Ecology Letters, 14, 325–334.

Linder H.P. (2001) Plant diversity and endemism in sub-Saharan tropical Africa. Journal of Biogeography, 28, 169–182.

Lososova Z., Lanikova D. (2010) Differences in trait compositions between rocky natural and artificial habitats. Journal of Vegetation Science, 21, 520–530. doi: 10.1111/j.1654-1103.2009.01160.x.

Magura T., Tothmeresz B., Molnar T. (2004) Changes in carabid beetle assemblages along an urbanisation gradient in the city of Debrecen, Hungary. Landscape Ecology, 19, 747–759. doi: 10.1007/s10980-005-1128-4.

Malard F. (2012) Filling the gap: distribution database of European groundwater crustaceans. BioFresh Newsletter 4: 7. http://www.freshwaterbiodiversity.eu/tl_files/downloads/.

Malard F., Boutin C., Camacho A., Ferreira D., Michel G., Sket B., Stoch F. (2009) Diversity patterns of stygobiotic crustaceans across multiple spatial scales in Europe. Freshwater Biology, 54, 4, 756–776.

Malard F., Plenet S., Gibert, J. (1996). The use of invertebrates in ground water monitoring: a rising research field. Ground Water Monitoring and Remediation, 16, 103–113.

Marchant R. (2007) The use of taxonomic distinctness to assess environmental disturbance of insect communities from running water. Freshwater Biology 52, 1634–1645. doi:10.1111/j.1365-2427.2007.01785.x.

MEA (Millennium Ecosystem Assessment) (2005). Ecosystems and Human Well- being: Synthesis. Island Press, Washington, DC.

Mermillod-Blondin F., Lefour C., Lalouette L., Renault D., Malard F., Simon L., Douady C.J. (2013). Thermal tolerance breadths among groundwater crustaceans living in a thermally constant environment. Journal of Experimental Biology, in press.

Moe, S.J., Schmidt-Kloiber, A., Dudley, B.J. & Hering, D. (2012) The WISER way of organising ecological data from European rivers, lakes, transitional and coastal waters. Hydrobiologia, 704, 11–28.

Paul M.J., Meyer J.L. (2001) Streams in the urban landscape. Annual Review of Ecology and Systematics 32, 333–365.

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Hoft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M. & Wegmann, M. (2013) Essential Biodiversity Variables. Science, 339, 277–278.

Poff N.L., Olden J.D., Merritt D., Pepin D. (2007) Homogenization of regional river dynamics by dams and global biodiversity implications. Proceedings of the National Acadamy of Sciences USA 104: 5732–5737.

Pretto F., Celesti-Grapow L., Carli E., Brundu G., Blasi C. (2012) Determinants of non-native plant species richness and composition across small Mediterranean islands. Biological Invasions, 14, 2559–2572. doi: 10.1007/s10530-012-0252-7.

R Development Core Team (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org. (Version: 2.14.1)

R Development Core Team (2012) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org. (Version: 2.15.2)

Ridgeway G. (2010) gbm: Generalized Boosted Regression Models R Package Version 1.6-3.1. Available at: http://CRAN.Rproject. org/package=gbm.

Schemske D. (2002). Tropical diversity: patterns and processes. In: Ecological and Evolutionary Perspectives on the Origins of Tropical Diversity: Key Papers and Commentaries (eds Chazdon, R. & Whitmore, T.). University of Chicago Press, Chicago, IL.

Simaika, J.P. & Samways, M.J. (2011) Comparative assessment of indices of freshwater habitat conditions using different invertebrate taxon sets. Ecological Indicators, 11, 370–378.

Stoch F., Galassi, D.M.P. (2010) Stygobiotic crustacean species richness: a question of numbers, a matter of scale. Hydrobiologia, 653, 217–234.

Tisseuil, C., Cornu, J.-F., Beauchard, O., Brosse, S., Darwall, W., Holland, R., Hugueny, B., Tedesco, P.A. & Oberdorff, T. (2012) Global diversity patterns and cross-taxa convergence in freshwater systems. Journal of Animal Ecology, 82, 365–376.

Tormos T., Kosuth P., Durrieu S., Villeneuve B. & Wasson J.G. (2011) Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. Physics and Chemistry of the Earth, 36, 549–559.

Uowolo A.L., Brinkley D., Adair E.C. (2005) Plant diversity in riparian forests in northwest Colorado: Effects of time and river regulation. Forest Ecology and Management, 218, 107–114. doi: 10.1016/j.foreco.2005.07.003.

Vallet J., Daniel H., Beaujouan V., Roze F., Pavoine S. (2010) Using biological traits to assess how urbanization filters plant species of small woodlands. Applied Vegetation Science, 13, 412–424. doi: 10.1111/j.1654-109X.2010.01087.x.

Venables W. N., Ripley B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Vörösmarty C.J., McIntyre P.B., Gessner M.O., Dudgeon D., Prusevich A., Green P., Glidden S., Bunn S.E., Sullivan C.A., Liermann C.R., Davies P.M. (2010) Global threats to human water security and river biodiversity. Nature, 467, 555–561.

Warwick R.M., Clarke K.R. (1995) New “biodiversity” measures reveal a decrease in taxonomic distinctness with increasing stress. Marine Ecology Progress Series, 129, 301–305.

Wiens J.J. (2011) The causes of species richness patterns across space, time, and clades and the role of ecological limits. The Quaterly Review of Biology, 86, 75–96.

Wood S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73, 3–36.

8 Tables

Table 1: Spatial scale and shape of areas for which land use was derived.

Ecosystem type	Spatial scale	Area shape
River (sites)	Entire catchment of a site	Polygon
Rivers (catchments)	(Sub)catchments (CCM2[1])	Polygon
Lakes	Buffer around each lake (buffer size to be added here in the final version)	Polygon
Wetlands	Buffer zone around each site (5 km radius)	Circle
Groundwaters	Cells of 100 x 100 km^2	Polygon

[1] CCM2 catchments to be explained here in the final draft.

Table 2: Summary of geographic and climatic variables used in the boosted regression models to partial out the natural patterns in the biodiversity metrics. (N = number of sites available for each subset)

Mountains	Fish (N = 273)			Invertebrates (N = 470)			Macrophytes (N = 115)		
Variables	Range	Mean	SD	Range	Mean	SD	Range	Mean	SD
Annual mean temperature (°C)	5.4–13.8	8.1	1.2	5.4–13.8	8.0	1.0	6.3–13.8	8.3	1.1
Annual precipitation (mm)	531–1322	796	162	531–1322	796	161	531–1207	769	156
Catchment area (km^2)	10–7862	406	1019	10–47142	293	2997	10–17731	367	1591
Latitude (dec. degree)	42.81–51.41	48.8	1.60	42.81–51.61	49.53	1.60	42.81–51.41	49.5	1.9
Longitude	0.74–15.81	11.11	4.10	0.74–15.81	10.40	3.46	0.74–15.16	9.10	3.15
Altitude (m)	250–810	395	128	250–810	372	112	250–810	369	105

Lowlands	Fish (N = 317)			Invertebrates (N = 751)			Macrophytes (N = 379)		
Variables	Range	Mean	SD	Range	Mean	SD	Range	Mean	SD
Annual mean temperature (°C)	7.9–14.6	9.3	1.0	7.5–14.6	9.3	0.8	7.9–14.6	9.3	0.9
Annual precipitation (mm)	492–1083	725	117	473–1083	723	120	473–1083	712	124
Catchment area (km^2)	10–3130	225	374	10–46306	678	3260	10–18665	504	1638
Latitude (dec. degree)	43.41–54.81	50.61	2.01	43.41–54.81	51.45	1.61	43.41–54.81	51.3	1.9
Longitude	-3.89–15.76	7.90	4.01	-3.89–15.76	8.17	3.11	-3.89–14.96	8.21	3.50
Altitude (m)	3–249	122	72	1–249	79	74	3–249	91	68

Table 3: Summary statistics of four biodiversity metrics (mean [SD]) and three organism groups considered in the floodplain analyses. NA indicates biodiversity metrics not calculable for the organism groups.

Biodiversity metric	Floodplain vegetation	Carabid ground beetles	Mollusca
No. of taxa	1,268	357	208
Richness	35.1 (42.8)	35.7 (22.7)	14.2 (SD to be added)
Taxonomic distinctness	NA	NA	NA
Endemicity	NA	NA	NA
Shannon	2.51 (1.09)	2.37 (0.79)	2.36 (0.49)
Simpson	0.84 (0.15)	0.81 (0.14)	0.81 (0.14)
Evenness	0.77 (0.21)	0.69 (0.17)	0.69 (0.17)

Table 4: Geo-climatic (natural) variables used for the floodplain analyses to account for geographical patterns in the data.

Variable	Range	Mean	SD
Annual mean temperature (°C/10)	0.2–17.9	10.47	33.87
Maximum temperature of the warmest month	17.2–32.7	24.79	33.87
Minimum temperature of the coldest month	-17.2–7.9	-16.28	45.52
Annual precipitation (mm)	226.3–1336.0	680.89	189.65
Precipitation of wettest quarter	87.9–443.0	232.78	77.71
Precipitation of driest quarter	12.2–241.0	112.89	48.11
Latitude	36.84–65.21	48.48	5.56
Longitude	-8.67–25.84	9.69	8.71
Altitude (m)	-2.00–1093.75	131.74	170.36

Table 5: CORINE land cover information (level 1) used in this study. 'Group' refers to the organism group-specific analyses with consideration of the respective land use category after testing for collinearity (Pl = plants; Ca = carabid ground beetles; Mo = Mollusca).

Land cover category	Group
Artificial surfaces	Pl, Ca, Mo
Agricultural areas	Ca, Mo
Forest and seminatural areas	Pl, Mo
Wetlands	Pl, Ca, Mo
Water bodies	Pl, Ca, Mo

Table 6: Regression model summary of the relationships between land use stressors and floodplain organism diversity. Diversity metrics are coded Mo_ for molluscs, Ca_ for ground beetles and Pl_ for floodplain vegetation. Significance at p <0.05: *, p <0.01: ** and p <0.001: ***; b.s. = borderline significant).

Biodiversity metric (No. of sites)	Land use categories in the final model	Regression formula	adjusted R^2	AIC	Residual deviance (% deviance explained)	Model significance
Mo_Richness (n=132)	AGRICULTURE ** WETLANDS ***	Y=0.14+0.002x+0.12x	0.13	-76.8	1.7 (14.8%)	**
Mo_Shannon (n=53)	AGRICULTURE * WETLANDS **	Y=0.68-0.01x-0.48x	0.15	68.6	9.7 (18.2%)	**
Mo_Simpson (n=53)	FOREST ** WETLANDS ***	Y=-0.07+0.004x-0.25x	0.33	-43.0	1.2 (35.6%)	***
Mo_evenness (n=53)	ARTIFICIAL *** AGRICULTURE ** WETLANDS ***	Y=0.63-0.11x-0.005x-0.41x	0.46	-26.5	1.6 (49.4%)	***
Ca_Richness (n=132)	ARTIFICIAL *** WETLANDS ***	Y=1.79-3.43x-1.03x	0.18	527.7	396.1 (19.0%)	***
Ca_Shannon (n=126)	ARTIFICIAL *** AGRICULTURE * WETLANDS **	Y=0.31-1.09x+0.58x-0.55x	0.18	278.6	62.2 (19.8%)	***
Ca_Simpson (n=126)	ARTIFICIAL * AGRICULTURE ***	Y=-0.11 – 0.16x+0.32x	0.13	-3.4	6.7 (14.8%)	***
Ca_evenness (n=126)	AGRICULTURE ** WATER BODIES *	Y=0.48+0.25x-0.23x	0.18	-30.0	5.5 (19.1%)	***
Pl_Richness (n=306)	ARTIFICIAL *** FOREST *** WETLANDS *	y=-1.31+1.44x +0.28x + 0.13x	0.19	-183.5	9.5 (13.7%)	***
Pl_Shannon (n=204)	FOREST *	Y=0.37-0.68x	0.02	629.1	253.3 (4.9%)	*
Pl_Simpson (n=204)	FOREST *** WETLANDS (b.s.)	Y=0.37-0.19x+0.26x	0.15	-129.3	1.8 (6.4%)	**
Pl_evenness (n=204)	FOREST **	Y=0.75-0.29x	0.06	-144.1	1.5 (7.2%)	**

Table 7: Natural variables used in the analysis of groundwater biodiversity patterns.

Variables	Range	Mean	SD
Actual evapotranspiration (mm)	502.38	501.33	98.97
Mean annual air temperature (°C)	9.89	25.03	5.23
Habitat heterogeneity	0.87	0.49	0.20
Variable n	701	701	701

Table 8: Deviance explained by boosted regression trees (BRT), generalized linear models (GLM) and generalized additive models (GAM).

Models	Species richness	Endemicity
BRT	59.6	38.8
GLM	44.4	24.4
GAM	51.9	27.3

Table 9: Relative weight of land use and natural variables in boosted regression trees (BRT) and generalized linear models (GLM) explaining species richness.

Variables	BRT	GLM
AET	59.8	47.3
Habitat	17.6	26.4
Urban	8.5	7.5
Temperature	7.6	6.4
Agriculture	6.5	4.0
AET * Temperature		4.2
Temperature * Agriculture		4.0

Table 10: Relative weight of land use and natural variables in boosted regression trees (BRT) and generalized linear models (GLM) explaining endemicity.

Variables	BRT	GLM
Temperature	52.1	67.0
AET	20.8	18.5
Habitat	12.7	7.4
Agriculture	9.1	2.9
Urban	5.3	4.2

9 Figure captions

Figure 1: Interrelationship of six river macroinvertebrate biodiversity metrics at N = 1,221 river sites in mountain and lowland ecoregions. Red lines indicate lowess smoothers to illustrate the relationship between the metrics. Numbers in the upper right panel show Spearman correlation coefficients.

Figure 2: Mean variance (± SD) in six river biodiversity metrics explained by geo-climatic (grey) and land use variables (black). Mean values were calculated from all organism groups and ecoregion types.

Figure 3: Variance in six biodiversity metrics explained by geo-climatic descriptors and land use in the BRT models. Top row: macroinvertebrates (INV_); middle row: fishes (FI_); bottom row: macrophytes (MP_). Left column: mountain sites; right column: lowland sites.

Figure 4: Maps of Europe showing the dependent variables (species richness and endemicity) and explanatory variables used for the analysis of groundwater biodiversity patterns at the scale of Europe. The hatched area shows cells that were not considered in the analysis.

Figure 5: Correlation circles of principal component analysis performed on the proportion of seven land use classes in 526 100 × 100-km cells covering Europe, except Scandinavia.

Figure 6: Partial Dependence Plots showing the fitted values for species richness along selected environmental gradients.

Figure 7: Partial Dependence Plots showing the fitted values for endemicity along selected environmental gradients.

Figure 8: BRT interaction plot showing the fitted values of species richness and its interactions between temperature and actual evapotranspiration (left panel), and actual evapotranspiration and habitat heterogeneity, respectively (right panel).

Figure S1: Overall richness of the three organism groups at lowland and mountain sites.

10 Figures

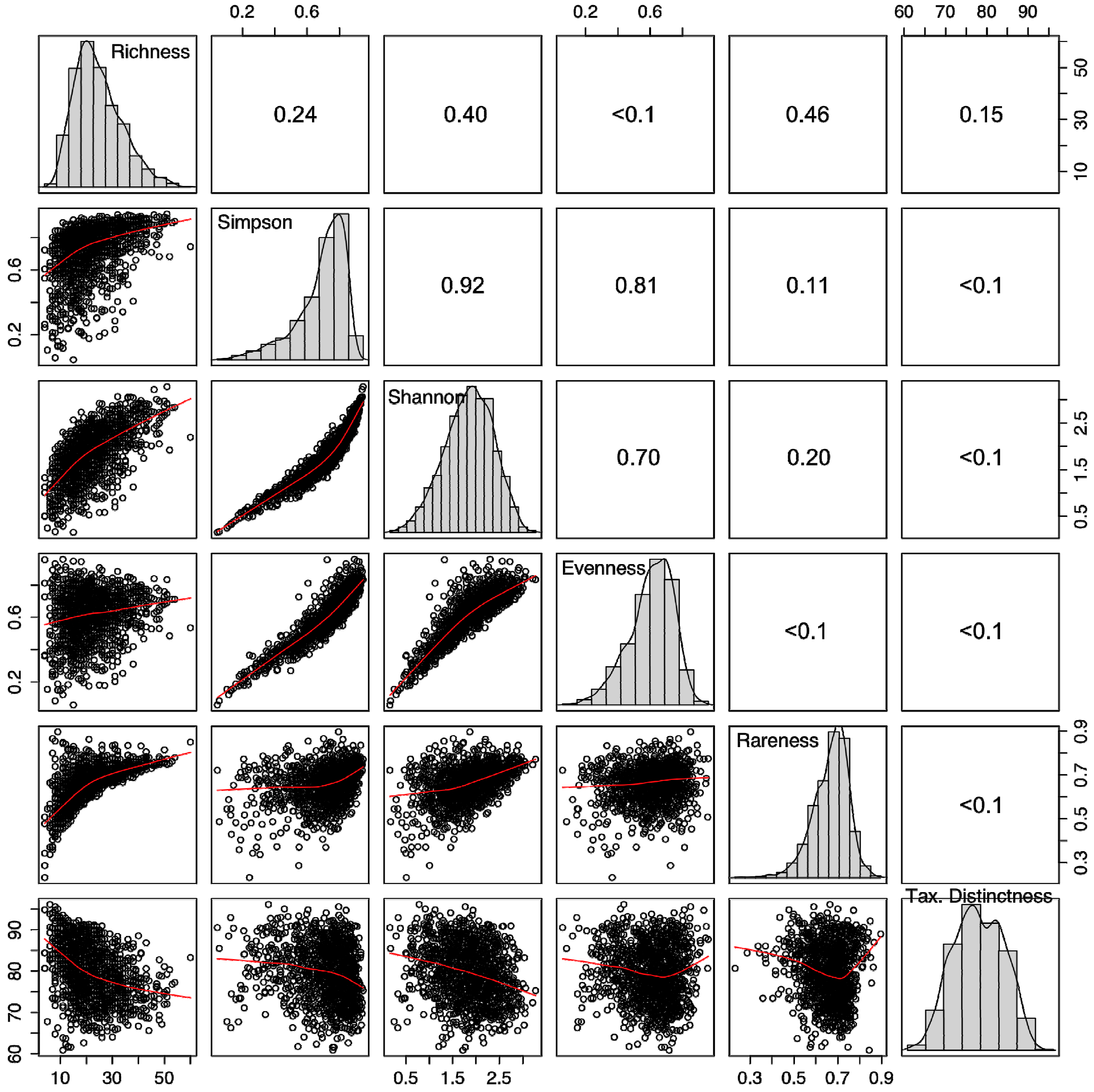


Figure 1

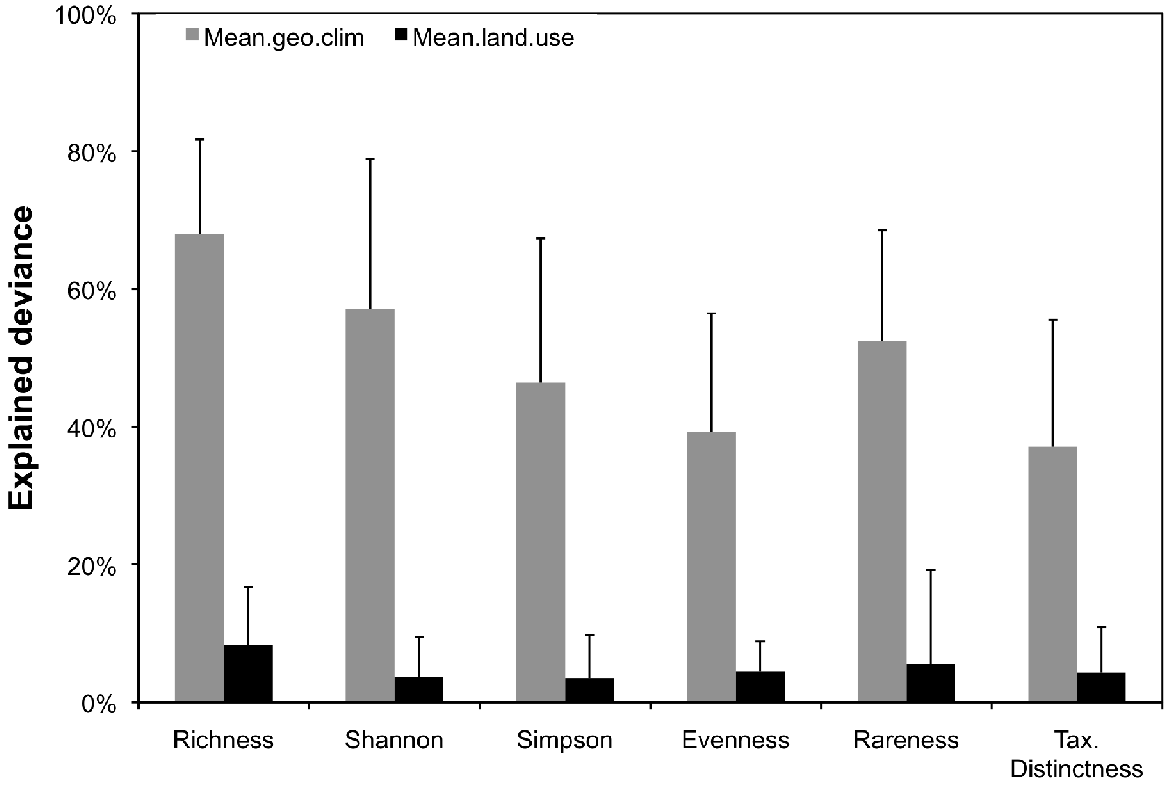

100%
80%
60%
40%
20%
0%
Explained deviance
Mean.geo.clim
Mean.land.use
Richness
Shannon
Simpson
Evenness
Rareness
Tax. Distinctness

Figure 2

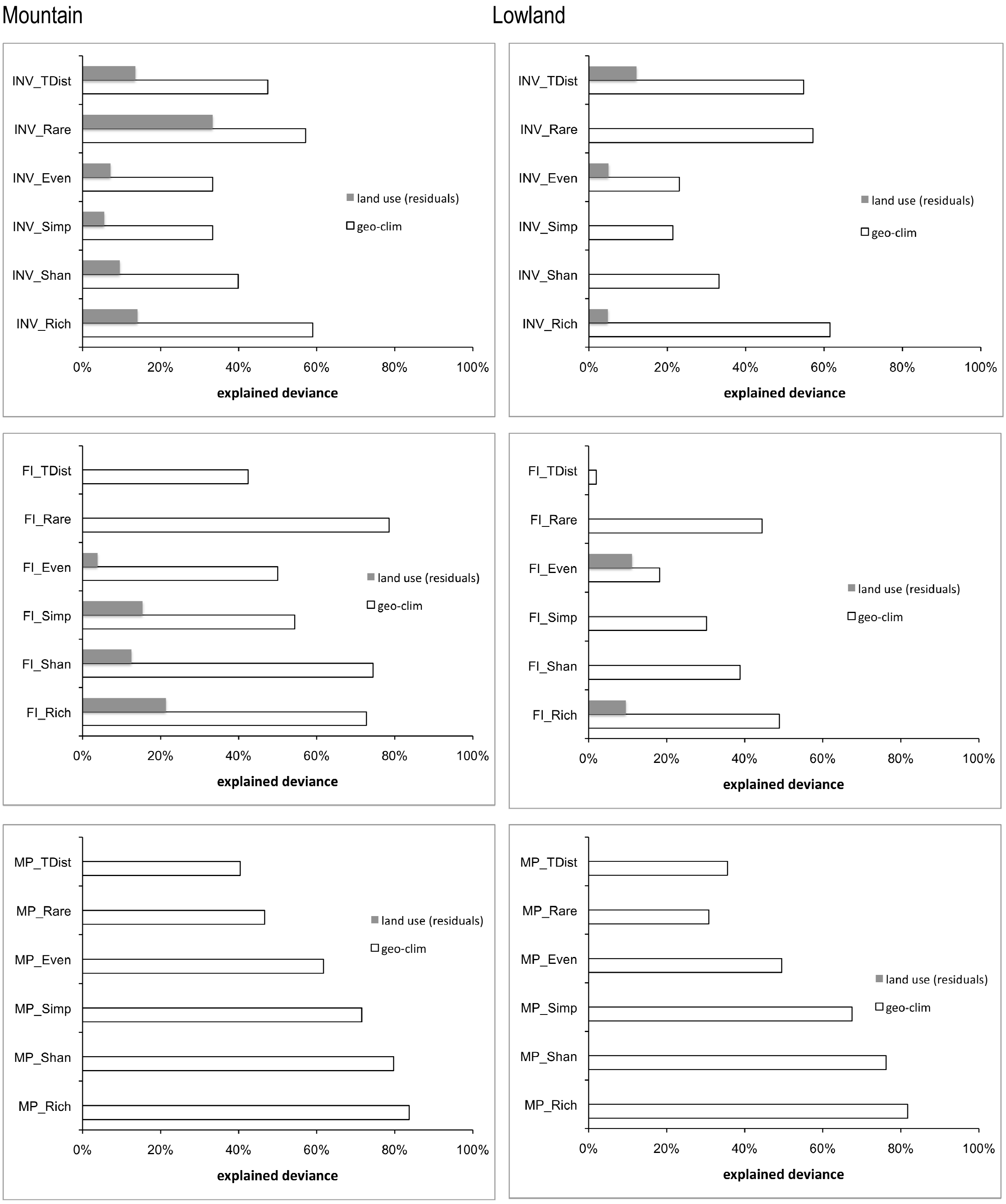

Mountain
Lowland
INV_TDist
INV_Rare
INV_Even
INV_Simp
INV_Shan
INV_Rich
FI_TDist
FI_Rare
FI_Even
FI_Simp
FI_Shan
FI_Rich
MP_TDist
MP_Rare
MP_Even
MP_Simp
MP_Shan
MP_Rich
■ land use (residuals)
☐ geo-clim
0%
20%
40%
60%
80%
100%
explained deviance

Figure 3

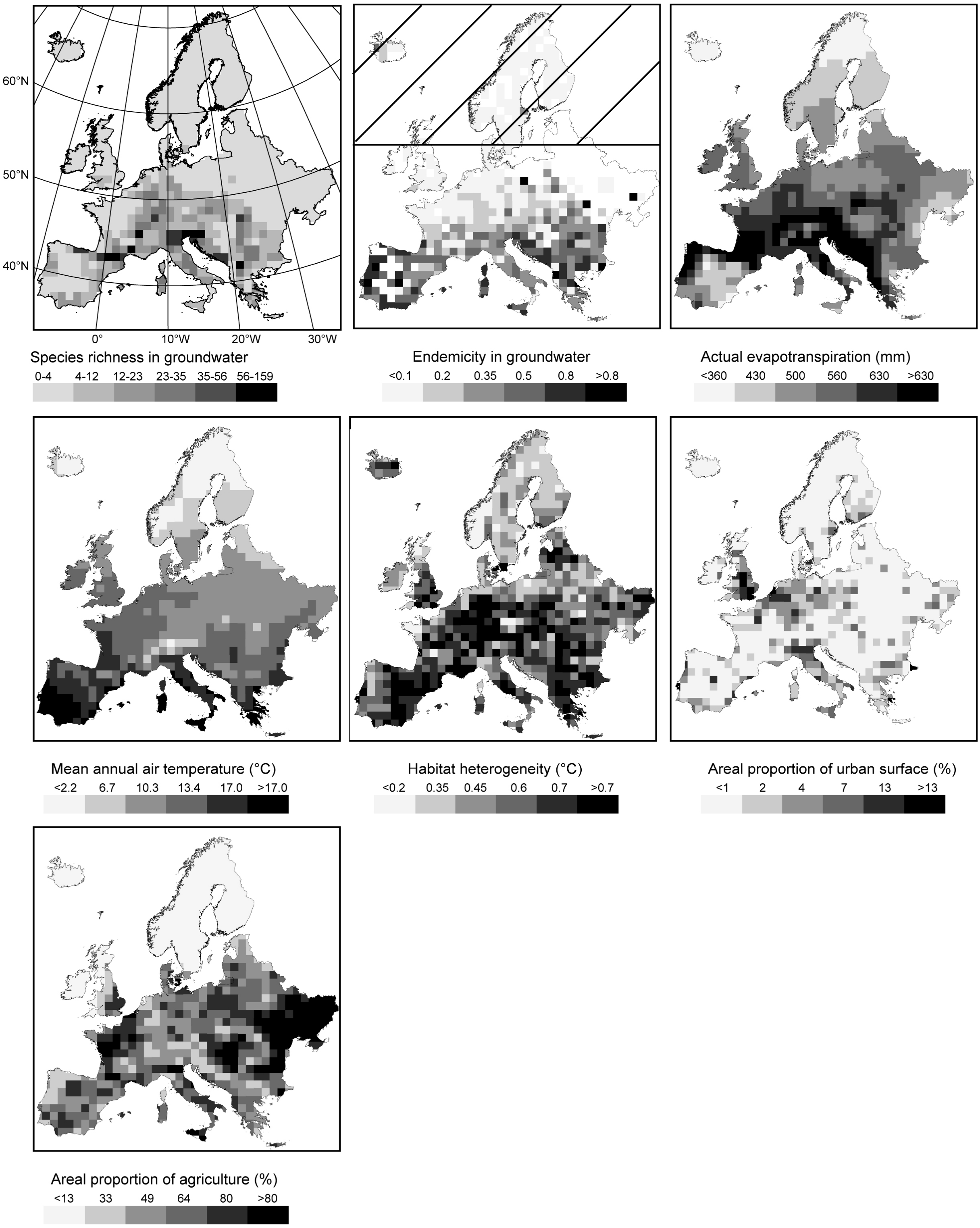

60°N
50°N
40°N
0°
10°W
20°W
30°W
Species richness in groundwater
0-4 4-12 12-23 23-35 35-56 56-159
Endemicity in groundwater
<0.1 0.2 0.35 0.5 0.8 >0.8
Actual evapotranspiration (mm)
<360 430 500 560 630 >630
Mean annual air temperature (°C)
<2.2 6.7 10.3 13.4 17.0 >17.0
Habitat heterogeneity (°C)
<0.2 0.35 0.45 0.6 0.7 >0.7
Areal proportion of urban surface (%)
<1 2 4 7 13 >13
Areal proportion of agriculture (%)
<13 33 49 64 80 >80

Figure 4

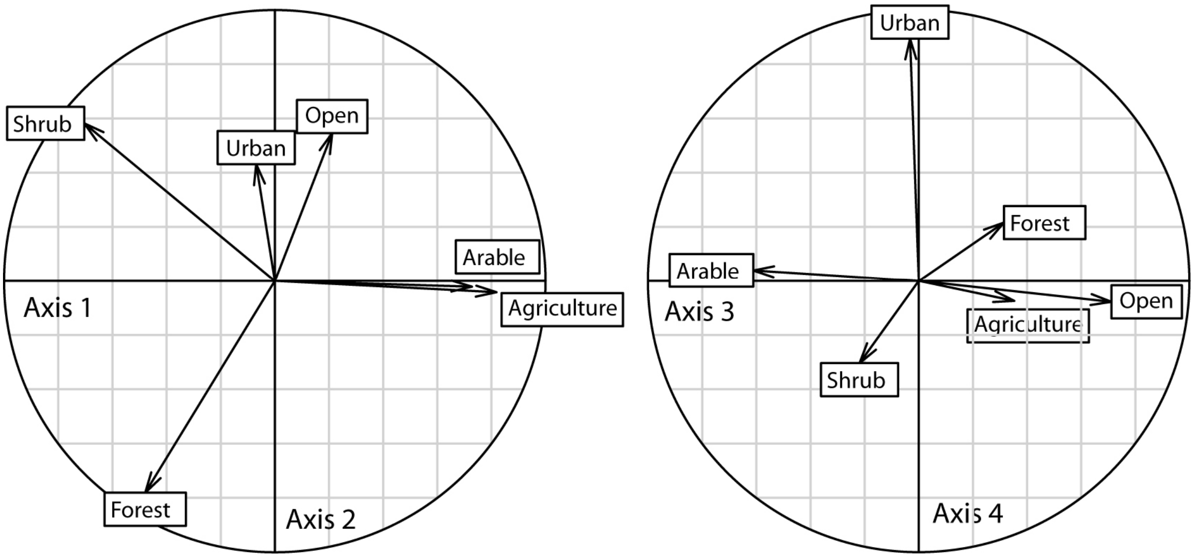

Shrub
Urban
Open
Arable
Agriculture
Axis 1
Forest
Axis 2
Urban
Forest
Arable
Open
Axis 3
Agriculture
Shrub
Axis 4

Figure 5

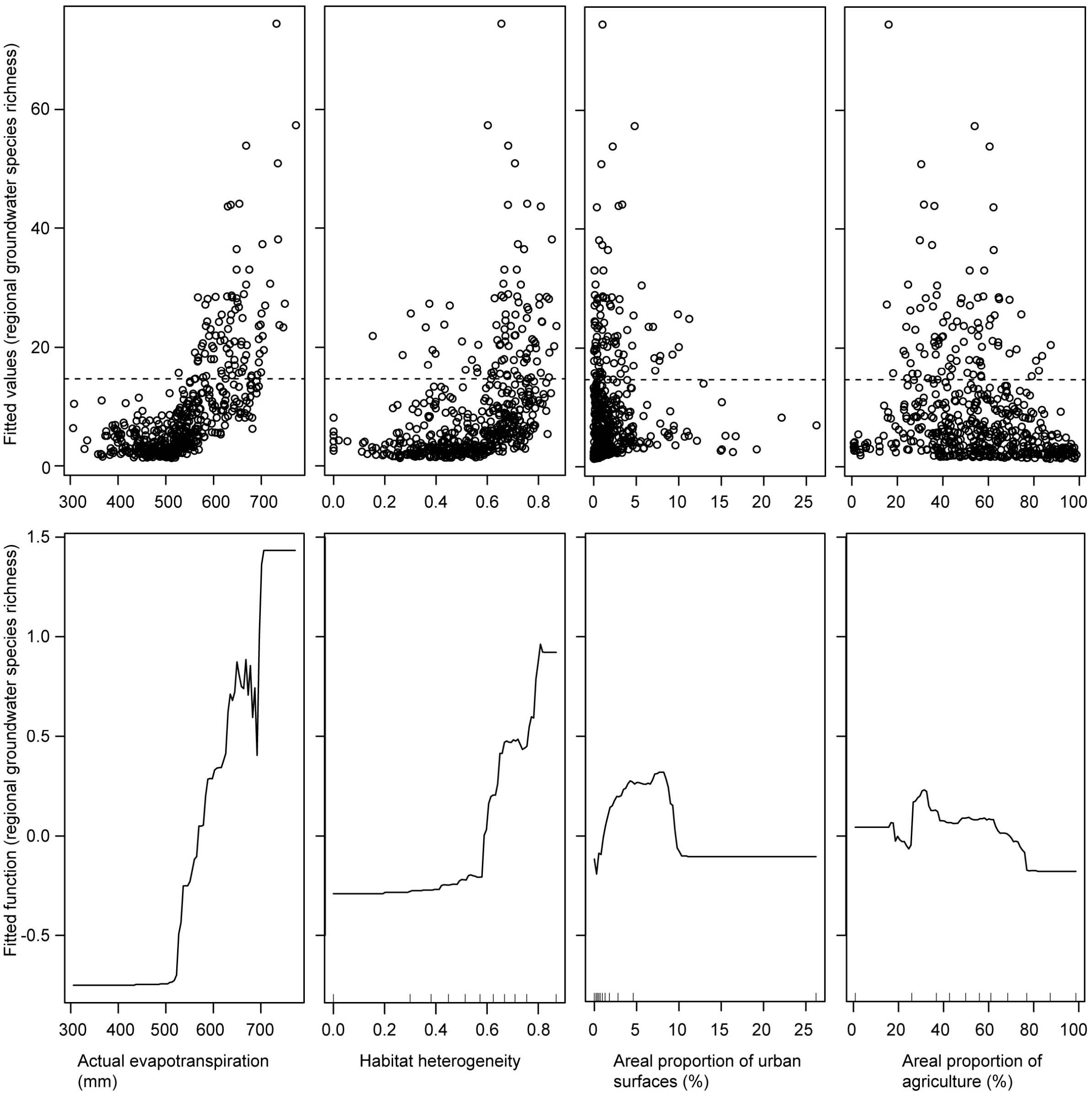

Fitted values (regional groundwater species richness)
0
20
40
60
Fitted function (regional groundwater species richness)
-0.5
0.0
0.5
1.0
1.5
300 400 500 600 700
0.0 0.2 0.4 0.6 0.8
0 5 10 15 20 25
0 20 40 60 80 100
Actual evapotranspiration (mm)
Habitat heterogeneity
Areal proportion of urban surfaces (%)
Areal proportion of agriculture (%)

Figure 6

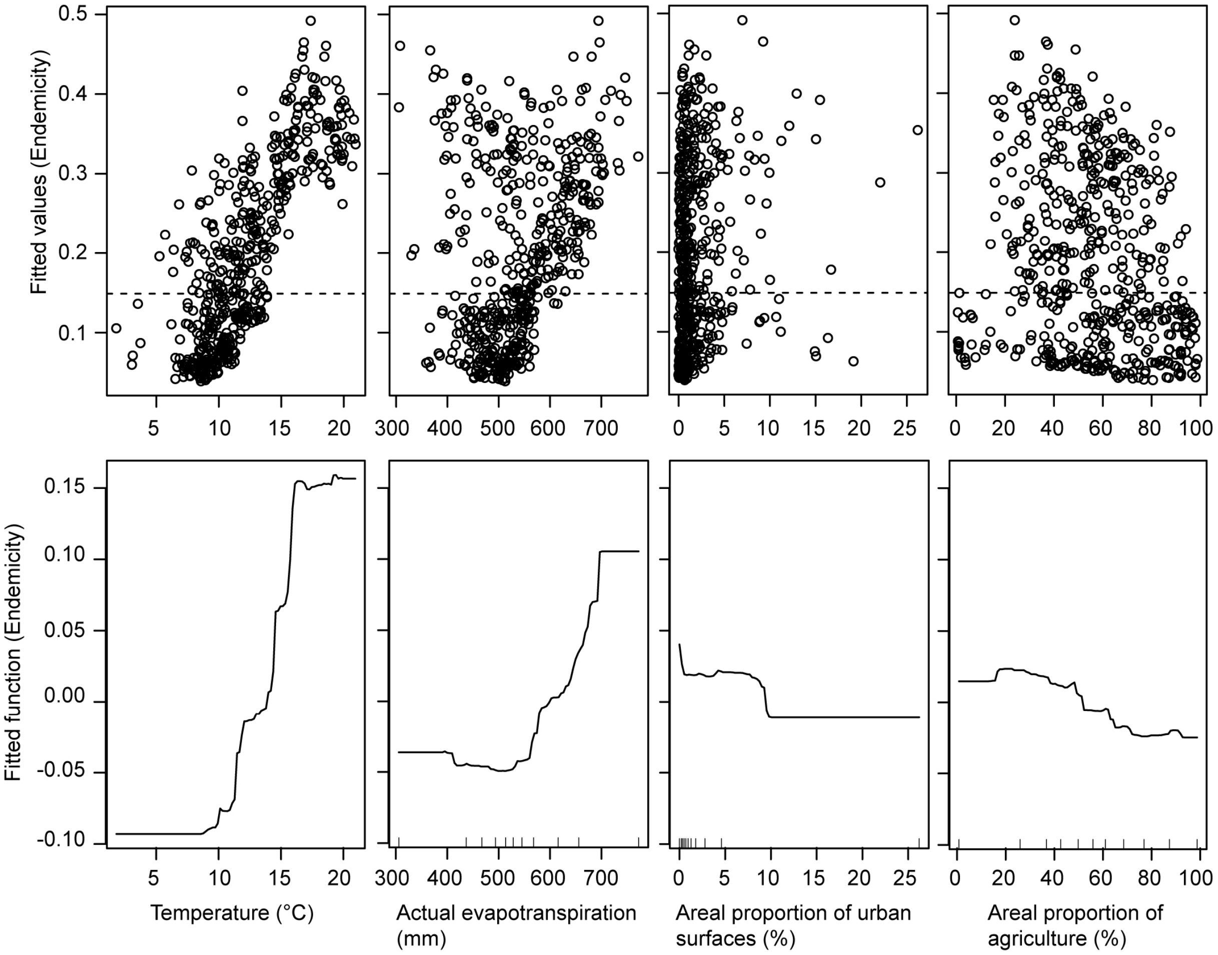

Fitted values (Endemicity)
0.5
0.4
0.3
0.2
0.1
5
10
15
20
300
400
500
600
700
0
5
10
15
20
25
0
20
40
60
80
100
Fitted function (Endemicity)
0.15
0.10
0.05
0.00
-0.05
-0.10
Temperature (°C)
Actual evapotranspiration (mm)
Areal proportion of urban surfaces (%)
Areal proportion of agriculture (%)

Figure 7

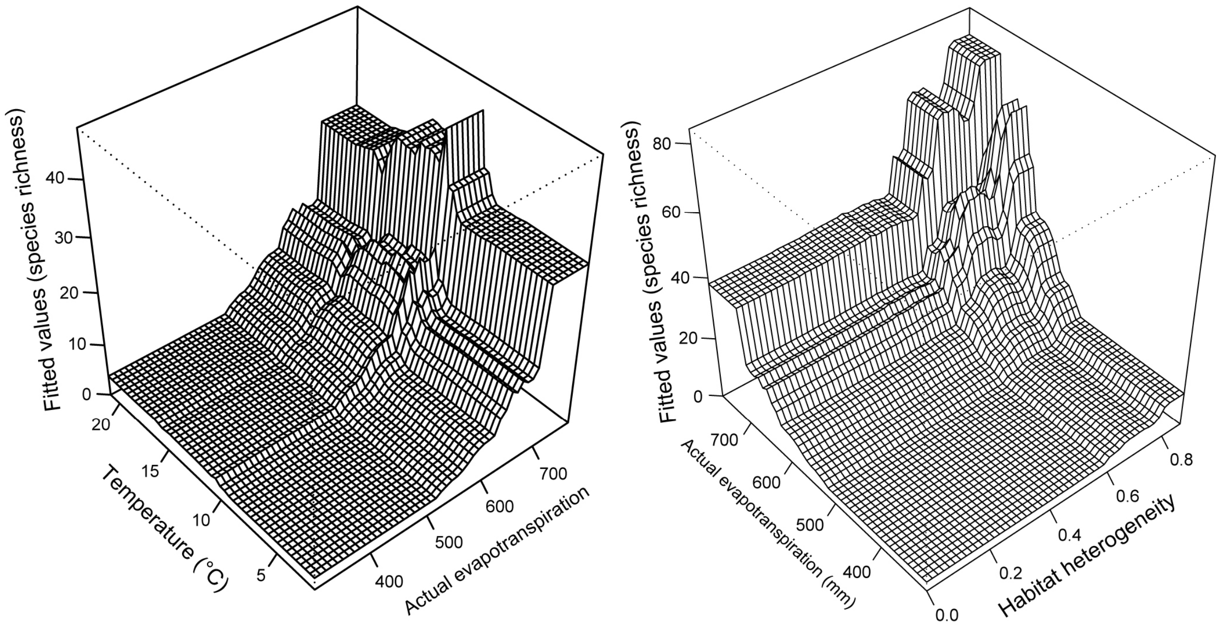

Fitted values (species richness)
0
10
20
30
40
Temperature (°C)
20
15
10
5
Actual evapotranspiration
400
500
600
700
Fitted values (species richness)
0
20
40
60
80
Actual evapotranspiration (mm)
700
600
500
400
Habitat heterogeneity
0.0
0.2
0.4
0.6
0.8

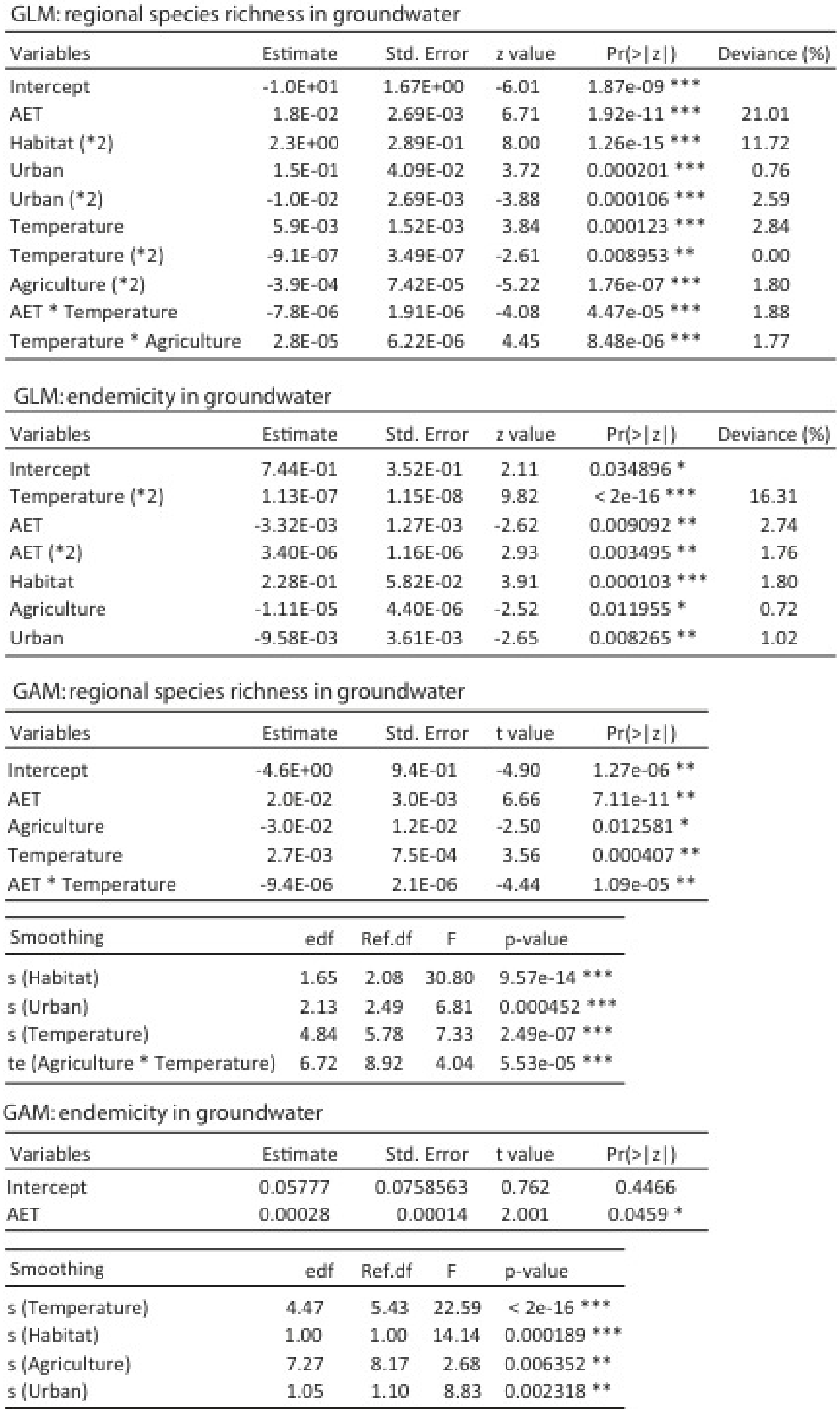
Figure 8

11 Supplementary material

Table S1: Summary of land use variables used to identify the impact of land use on riverine biodiversity.

		Mountain			Lowland		
	Land use	Range	Mean	SD	Range	Mean	SD
INV	Arable	0–100	33	23.9	0–100	60.8	24.2
	Forest	0–100	46.3	22.3	0–100	20.8	18.7
	Artificial	0–86	5.1	6.6	0–100	8.6	10.2
	Pasture	0–80	13.9	14.3	0–100	9.1	12.4
	Water bodies	0–15	0.3	1.1	0–18	0.3	1.5
	Wetlands	0–2	0.04	0.2	0–3	0.05	0.3
FI	Arable	0–93	33.1	23.1	0–96	56.1	25.7
	Forest	0–100	46.3	21.7	0–100	26.2	22.2
	Artificial	0–18	3.6	3.2	0–40	6.4	6.4
	Pasture	0–80	14.6	16.1	0–73	11.1	13.6
	Water bodies	0–7	0.3	1	0–6	0.1	0.5
	Wetlands	0–2	0.06	0.3	0–2	0.03	0.2
MP	Arable	0–97	29.2	24	0–100	58.7	24.4
	Forest	0–100	48.9	24.2	0–100	22.4	19.9
	Artificial	0–25	4.1	4.5	0–80	7.8	9.2
	Pasture	0–71	15.7	16	0–100	10.4	13.4
	Water bodies	0–15	0.2	1.1	0–18	0.4	1.7
	Wetlands	0–2	0.06	0.3	0–3	0.05	0.3

Table S2: Summary statistics of GLM and GAM used to analyse the effects of land use and natural geo-climatic descriptors on the regional species richness and endemicity of European groundwater crustaceans.



GLM: regional species richness in groundwater

Variables	Estimate	Std. Error	z value	Pr(>|z|)	Deviance (%)
Intercept	-1.0E+01	1.67E+00	-6.01	1.87e-09 ***	
AET	1.8E-02	2.69E-03	6.71	1.92e-11 ***	21.01
Habitat (*2)	2.3E+00	2.89E-01	8.00	1.26e-15 ***	11.72
Urban	1.5E-01	4.09E-02	3.72	0.000201 ***	0.76
Urban (*2)	-1.0E-02	2.69E-03	-3.88	0.000106 ***	2.59
Temperature	5.9E-03	1.52E-03	3.84	0.000123 ***	2.84
Temperature (*2)	-9.1E-07	3.49E-07	-2.61	0.008953 **	0.00
Agriculture (*2)	-3.9E-04	7.42E-05	-5.22	1.76e-07 ***	1.80
AET * Temperature	-7.8E-06	1.91E-06	-4.08	4.47e-05 ***	1.88
Temperature * Agriculture	2.8E-05	6.22E-06	4.45	8.48e-06 ***	1.77

GLM: endemicity in groundwater

Variables	Estimate	Std. Error	z value	Pr(>|z|)	Deviance (%)
Intercept	7.44E-01	3.52E-01	2.11	0.034896 *	
Temperature (*2)	1.13E-07	1.15E-08	9.82	< 2e-16 ***	16.31
AET	-3.32E-03	1.27E-03	-2.62	0.009092 **	2.74
AET (*2)	3.40E-06	1.16E-06	2.93	0.003495 **	1.76
Habitat	2.28E-01	5.82E-02	3.91	0.000103 ***	1.80
Agriculture	-1.11E-05	4.40E-06	-2.52	0.011955 *	0.72
Urban	-9.58E-03	3.61E-03	-2.65	0.008265 **	1.02

GAM: regional species richness in groundwater

Variables	Estimate	Std. Error	t value	Pr(>|z|)
Intercept	-4.6E+00	9.4E-01	-4.90	1.27e-06 **
AET	2.0E-02	3.0E-03	6.66	7.11e-11 **
Agriculture	-3.0E-02	1.2E-02	-2.50	0.012581 *
Temperature	2.7E-03	7.5E-04	3.56	0.000407 **
AET * Temperature	-9.4E-06	2.1E-06	-4.44	1.09e-05 **

Smoothing	edf	Ref.df	F	p-value
s (Habitat)	1.65	2.08	30.80	9.57e-14 ***
s (Urban)	2.13	2.49	6.81	0.000452 ***
s (Temperature)	4.84	5.78	7.33	2.49e-07 ***
te (Agriculture * Temperature)	6.72	8.92	4.04	5.53e-05 ***

GAM: endemicity in groundwater

Variables	Estimate	Std. Error	t value	Pr(>|z|)
Intercept	0.05777	0.0758563	0.762	0.4466
AET	0.00028	0.00014	2.001	0.0459 *

Smoothing	edf	Ref.df	F	p-value
s (Temperature)	4.47	5.43	22.59	< 2e-16 ***
s (Habitat)	1.00	1.00	14.14	0.000189 ***
s (Agriculture)	7.27	8.17	2.68	0.006352 **
s (Urban)	1.05	1.10	8.83	0.002318 **

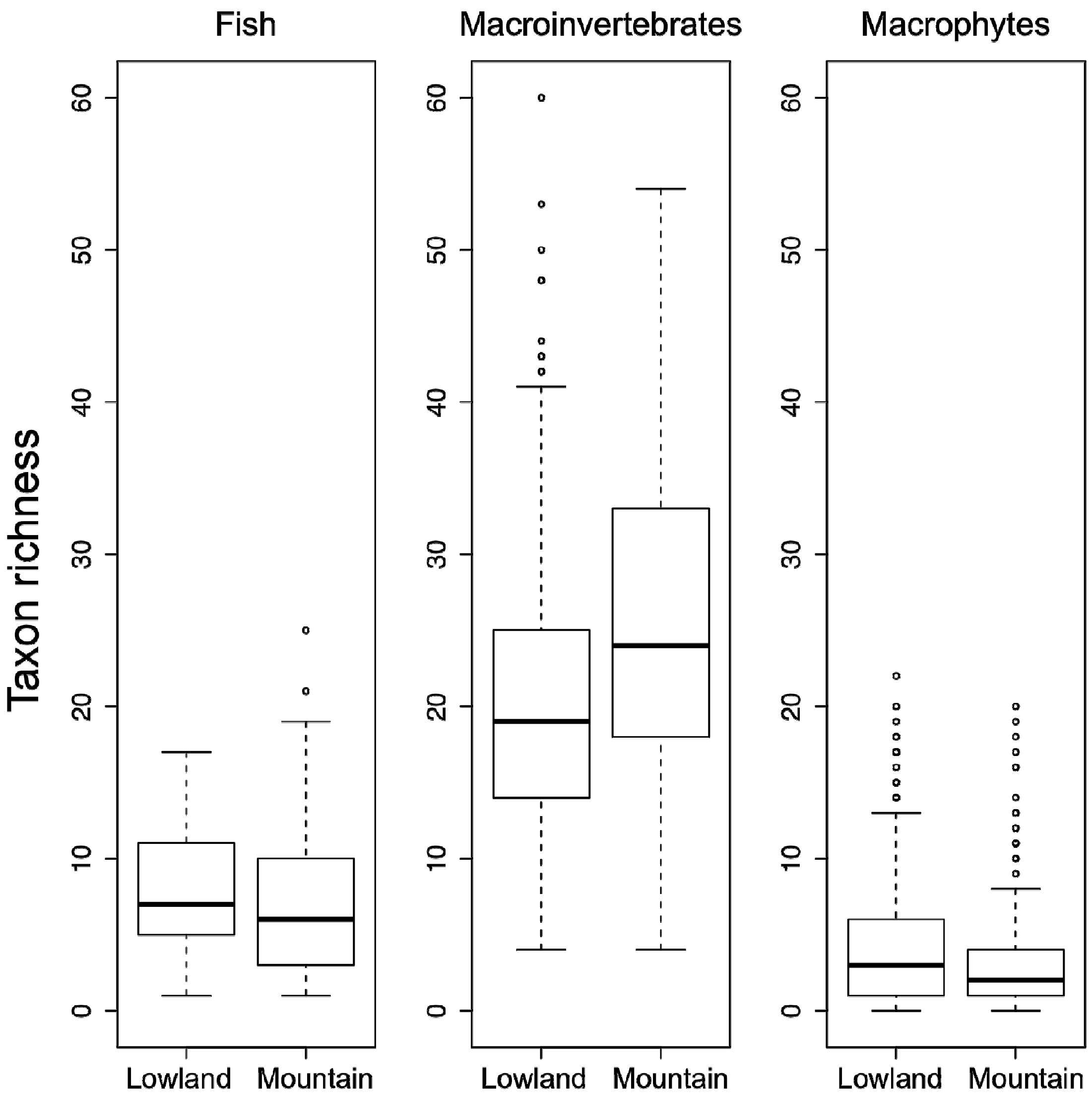

Fish
Macroinvertebrates
Macrophytes
Taxon richness
0
10
20
30
40
50
60
Lowland
Mountain

Figure S1