

# Invertebrates in pigeonholes: Evaluating pan-European freshwater typologies in light of their macroinvertebrate communities

Jonathan F. Jupke<sup>1</sup>      Pepe Barquín<sup>2</sup>      Núria Bonada<sup>3</sup>  
Christian K. Feld<sup>4,6</sup>      Peter Haase<sup>5,6</sup>      Kaisa-Leena Huttunen<sup>7</sup>  
Miguel Cañedo-Argüelles<sup>3</sup>      Timo Muotka<sup>7,8</sup>      Riku Paavola<sup>9</sup>  
Edwin T. H. M. Peeters<sup>10</sup>      Leonard Sandin<sup>11</sup>      Denes Schmera<sup>12</sup>  
Philippe Usseglio-Polatera<sup>13</sup>      Ralf B. Schäfer<sup>1</sup>

<sup>1</sup> iES Landau, Institute for Environmental Sciences, University Koblenz-Landau, Fortstrasse 7, 76829 Landau, Germany

<sup>2</sup> Environmental Hydraulics Institute, University of Cantabria, C/ Isabel Torres 15, Parque Científico y Tecnológico de Cantabria, 39011 Santander, Spain.

<sup>3</sup> Grup de Recerca “Freshwater Ecology, Hydrology and Management” (FEHM), Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia, Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona (UB), Diagonal 643, 08028, Barcelona, Catalonia, Spain.

<sup>4</sup> Centre for Water and Environmental Research, University of Duisburg-Essen, 45141 Essen, Germany

<sup>5</sup> Department of River Ecology and Conservation, Senckenberg Research Institute and Natural History Museum Frankfurt, Gelnhausen 65371, Germany

<sup>6</sup> Faculty of Biology, University of Duisburg-Essen, Essen 45141, Germany

<sup>7</sup> Department of Ecology and Genetics, University of Oulu, P.O. Box 3000, FI-90014 Oulu, Finland

<sup>8</sup> Finnish Environment Institute, Freshwater Centre, P.O. Box 413, FI-90014 Oulu, Finland

<sup>9</sup> Oulanka Research Station, University of Oulu Infrastructure Platform, Liikasenvaarantie 134, Kuusamo 93900, Finland

<sup>10</sup> Aquatic Ecology and Water Quality Management Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands

<sup>11</sup> Norwegian Institute for Water Research, Gaustadalléen 21, NO-0349 Oslo, Norway

<sup>12</sup> Centre for Ecological Research, Balaton Limnological Institute, Tihany, Hungary

## Acknowledgements

This study was funded by the GETREAL (Incorporating spatial and seasonal variability in community sensitivity into chemical risk assessment) project, which was funded by the European Chemical Industry Council long-range research initiative (CEFIC-LRi project ECO 50).

We are grateful to the staff that collected and analyzed data for the AQUEM, STAR, WISER and ECOSURV projects and to the various water agencies that participated in the abiotic and biotic data collection.

## 1 Data Availability

Some of the monitoring data sets are freely available. Data from the Ebro Hydrographic Confederation under [www.datossuperficiales.chebro.es:81/WCASF/?rvn=1](http://www.datossuperficiales.chebro.es:81/WCASF/?rvn=1)

Data from Naides under [www.naiades.eaufrance.fr/france-entiere#/](http://www.naiades.eaufrance.fr/france-entiere#/) Data from the RIV-PACS data base under [www.ceh.ac.uk/services/rivpacs-reference-database](http://www.ceh.ac.uk/services/rivpacs-reference-database) Data from the Dutch monitoring agencies under [waterkwaliteitsportaal.nl](http://waterkwaliteitsportaal.nl) Data GLORIC DATA ILLIES DATA EEA

## 2 Abstract

Human actions are altering the composition and functioning of ecological communities globally. This transformation is especially rapid and intensive in freshwater ecosystems. Comprehensive and international regulatory frameworks are necessary to slow down and halt the current loss of biodiversity. These regulatory frameworks, such as the Water Framework Directive (WFD), typically reduce the inherent complexity of their subjects through classification systems that use selected features to classify the subjects into a typology. A recently proposed pan-European typology for freshwater systems has unified the assortment of existing national typologies. However, it lacks biological validation. Using cluster validity analyses, we evaluated whether freshwater macroinvertebrate communities are more similar within than among types. To compare the typology's performance, we conducted the same analyses with three additional freshwater typologies. Beyond the typologies themselves, our analyses enabled us to compare two typologies based on river reaches with those based on ecoregions. Lastly, we provide lists of typical macroinvertebrate taxa for the types of each typologies.

Though performance differed between validity metrics, the newly proposed typology rarely ranked among the best. Overall, regional typologies achieved better scores than those based on river reaches.

However, all typologies failed to achieve high cluster validity scores. SART ZU TYPICAL ASSEMBLAGES Our findings indicate that superior alternatives to the new pan-European rivers typology exist but also that none reliably delineate demarcations between internally homogeneous macroinvertebrate reference communities. Hence, it remains important to further improve freshwater typologies to improve biomonitoring and facilitate conservation efforts.

### 3 Introduction

Freshwater ecosystems provide humanity with clean drinking water and food, they ensure livelihoods for fisheries and farms, and they are places of high recreational and aesthetic value (Youn *et al.* 2014; Béné *et al.* 2016; Börger *et al.* 2021). Combined, these contributions to people, also called ecosystem services, are estimated to have an annual worth of more than four trillion dollars (Costanza *et al.* 2014), which is approximately a twentieth of the global GDP as of 2017 (World Bank 2021). Human activities have reduced biodiversity in freshwater ecosystems (Darwall *et al.* 2018; WWF 2020) and are thereby endangering these systems’ capacities to provide ecosystem services (e.g. Cardinale 2011; Duffy *et al.* 2017; Grizzetti *et al.* 2019).

The Living Planet Index indicates that populations of freshwater vertebrates on average declined by 84% relative to their 1970 levels (but see Leung *et al.* (2020)). Based on the same data, He *et al.* (2019) found a slightly stronger decline (88%) in freshwater megafauna. Declining abundances have also been recorded for invertebrates (e.g. Cumberlidge *et al.* 2009; Baranov *et al.* 2020), however, most studies find stable or ambiguous trends in abundance and richness (Floury *et al.* 2013; Outhwaite *et al.* 2020; van Klink *et al.* 2020) though often associated with strong taxonomic and functional turnover (Jourdan *et al.* 2018; Haubrock *et al.* 2020; Mouton *et al.* 2020). The main stressors to which this situation can be attributed are well known and can be summarized as overexploitation, water pollution, flow modification, habitat degradation and species invasions (Dudgeon *et al.* 2006; Vörösmarty *et al.* 2010). Addressing these problems is complicated by the emergence of new stressors such as microplastics and light pollution (Reid *et al.* 2018) and interactions between co-occurring stressors (Schinegger *et al.* 2012; Birk *et al.* 2020; Comte *et al.* 2021). To prevent unacceptable harm to freshwater ecosystems and secure the supply of ecosystem services, internationally coordinated action and regulatory frameworks are necessary (Darwall *et al.* 2018; Rees *et al.* 2020; Tickner *et al.* 2020).

The Water Framework Directive (WFD) is among the major regulatory frameworks that aim to protect and restore freshwater systems in the European Union. It aims to reach or maintain a ‘good ecological status’ for all water bodies in the European Union (European Commission 2000). Since its implementation in 2000, the WFD has motivated improvements in methods for biomonitoring (Birk *et al.* 2012; Carvalho *et al.* 2019), advanced our knowledge of the ecological status of freshwater systems throughout Europe (EEA 2018) and has stimulated relevant research at the science-policy interface (Reyjol *et al.* 2014). It has not, however, reached its explicit aim, to achieve ‘good ecological quality’ for most surface and groundwater bodies by 2015. As of 2018, only 40% of surface water bodies had reached this status (EEA 2018). Multiple criteria determine the underlying status assessment: biolog-

ical quality elements (benthic invertebrates, phytoplankton, benthic flora and fish fauna), physicochemical variables like nutrient and pollutant concentrations, and hydromorphology.

The member states of the EU use national methods to assess their water bodies. To harmonize the national assessment methods and to ensure a coherent notion of ‘good ecological quality,’ an intercalibration process has been initiated. A recent output of the intercalibration process is a pan-European typology of lentic and lotic freshwater systems (broad river types (BRT), Lyche Solheim *et al.* 2019). While the WFD required its members to create river typologies for the assessment, these typologies differ widely in the number of river types they discern (between 1 and 367) and the variables used to define them (between 2 and 16). The recent effort by Lyche Solheim *et al.* (2019) combined the national typologies and connected the new, broad river types (BRT) with the established intercalibration types (Birk *et al.* 2013; Poikane *et al.* 2014). This pan-European typology has already been used to compare nutrient threshold values across river types (Poikane *et al.* 2019), and in studies examining interaction types of multiple stressors (Birk *et al.* 2020) and the chemical pollution of European surface waters (Posthuma *et al.* 2020).

The typology of Lyche Solheim *et al.* (2019) adopts the type descriptors that are most commonly used in national typologies: catchments size, altitude, geology, region and flow. The first three descriptors are part of the System A approach outlined in the Annex II of the WFD and Lyche Solheim *et al.* (2019) largely follow the thresholds proposed there. While System A includes a ecoregion descriptor that recognizes 25 distinct regions throughout Europe, the region descriptor used by Lyche Solheim *et al.* (2019) only separates between the Mediterranean and the rest of Europe. Flow is also a binary variables (perennial / intermittent) and was only applied to Mediterranean rivers. However, whether this typology based on geographical and environmental attributes matches biological classifications remains unknown. Verdonschot (2006b) suggests that System A-type typologies only partially capture the major drivers or macroinvertebrate distributions in Europe, which he found to be temperature, slope/ flow velocity and river size. The status assessment of the WFD relies on stable type-specific reference conditions. If biological communities vary strongly between reference sites of the same river type erroneous assessments resulting in over- or underprotection are more likely. In extreme cases, where no reliable type-specific reference conditions can be established, e.g. due to high natural variability, a river type might be excluded from the status assessment.

Following the typology of Lyche Solheim *et al.* (2019), we classified observations from a data set of freshwater macroinvertebrates to derive typical taxa for each river type and we evaluated the quality of the classification in comparison to those suggested by other freshwater typologies. Typical taxa are taxa that often occur in a given river type. The

concept differs from indicator taxa, in that the frequency with which they occur in other river types is not considered. These assemblages can be used to define reference conditions or to compare the sensitivity of communities towards stressors.

## 4 Methods

### 4.1 Preparation of macroinvertebrate data

We compiled a data set of macroinvertebrate samples from rivers throughout Europe (Table 1). The data included 152572 samples taken between 1968 and 2020. We assigned the BRT-type of the closeted river segment to each sample or removed the sample if this segment was further away than 500 m. Removing samples with a distance larger than 500 m to the next segment reduced the number of samples to 76621 (50% of all samples). Next, we removed all samples from sites that were considerably impaired. The exclusion criterion depended on the data set, as some included information on environmental conditions while others solely included biological information. When environmental information was lacking, we identified impaired sites using the catchment land cover. Sampling sites from catchments where the combined area of urban and agricultural land exceeded 20% of the catchment area were considered impaired (Malaj *et al.* 2014; Schreiner *et al.* 2016; Szöcs *et al.* 2017; Wolfram *et al.* 2021). We removed these potentially impaired sites because differences between river types are more pronounced between least-impaired sites (Verdonschot 2006a) and because we wanted to compare the typologies' capability to delineate reference assemblages with low spatial variability. Removing impaired sites further decreased the number of samples to 23284 (15% of all samples). Lastly, we only considered the 21169 samples that were taken in or after the year 2000 (14% of all samples). The spatial distribution of the final sites selection is shown in Figure 1. We provide maps of the spatial distribution of removed sites in the Supplementary Materials.

An overview of our analysis is provided in Figure ??.

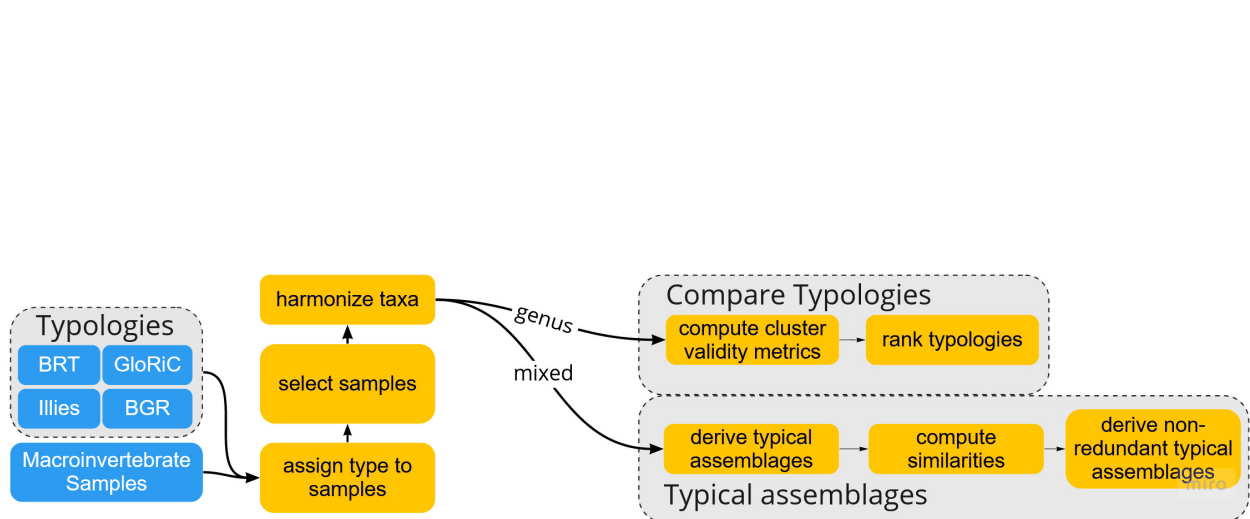


Table 1: Data sets used in the analysis. The exclusion criterion land use indicates

Data set	Area	Time span	Number of Samples	Exclusion criterion	Reference
Monitoring data from German federal agencies	Germany	1968 - 2013	65211	Land use	Bhowmik and Schäfer (2015), Berger et al. (2018), Le et al. (2021)
Ebro Hydrographic Confederation	Spain	2004 - 2019	3668	IASPT* < 4.5 and Land use	Escribano et al (2018)
Naiades	France	2002 - 2020	27052	BDI <sup>†</sup> < 14.5 and Land use	
Cantabria	Spain	2005	55	Land use	Alvarez et al. (2010), Alvarez et al. (2011)
Picos de Europa	Spain	2015 - 2017	24	Land use	
Ecosurv	Hungary	2005	491	Land use	Schmera and Baur (2011), Schmera et al. (2012)
RIVPACS database	UK	1978 - 2004	2504	original site selection	Clark et al. (2003), Turley et al. (2014)
STAR	Europe	2002 - 2003	91	classification in data set and Land use	Verdonschot (2006a), Johnson et al. (2007)
WISER	Europe	2000 - 2008	2565	Land use	Lyche-Solheim et al. (2013)
AQUEM Sweden	Sweden	2000	150	Land use	Verdonschot (2006a)
Koutajoki drainage basin	Finnland	2000 - 2013	322	original site selection	Huttunen et al. (2017)
Monitoring data from the RCS national network managed by the French water agencies	France	2007 - 2013	2694	qualitative ratings of multiple anthropogenic stressors	Mondy and Usseglio-Polattera (2013), Alric et al. (2021)
Monitoring data from Dutch regional water authorities	Netherlands	1978 - 2017	44702	Land use	Peters et al. (2013)
Monitoring data from the German Federal state of Hesse	Germany	2005 - 2018	3044	Land use	Haubrock et al. (2020)

\* Iberian Average Score per Taxon; see Leunda et al. (2009) or Munné and Prat (2009)

<sup>†</sup> Biological Diatom Index; see Lenoir and Coste (1996)



The resulting data set had sufficient data to represent 13 of 20 broad river types. We considered our samples representative for their river type, if we had more than 20 sites, which were distributed evenly across the instances of the river type. Maps of the sampling locations for the individual river types are provided in the Supplementary Materials. We had insufficient samples of the river types: BRT6, 7, 12, 13, 17, 19 and 20, consequently, these were omitted from the further analysis.

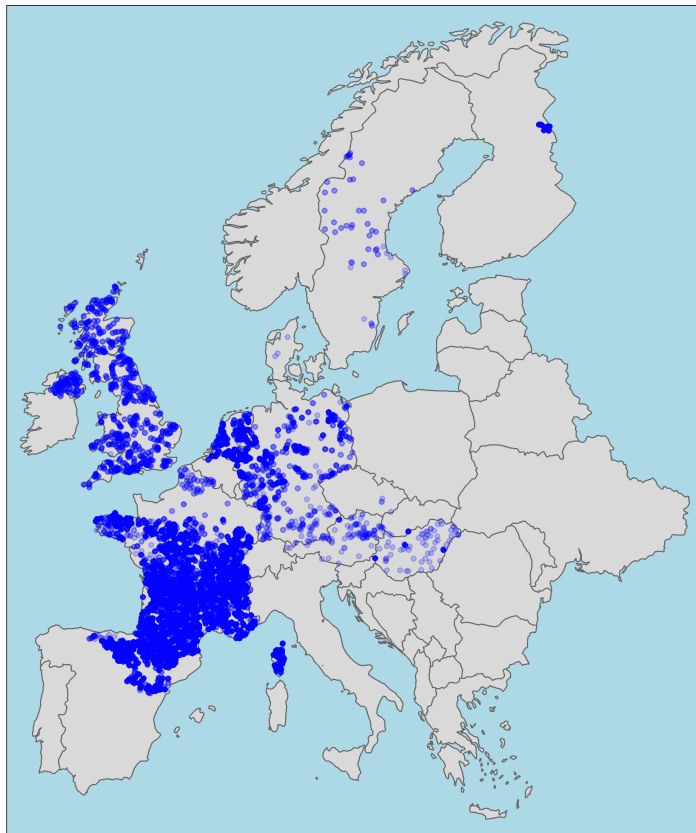


Figure 1: Map of final site selection. Each dot represents one sampling event.

The data originated from different sources and required adjustments to ensure taxonomic consistency. We used the Global Biodiversity Information Facility (GBIF) to replace synonyms with accepted names and to assign a taxonomic level to each observation. To integrate the different data types, we transformed all data to presence-absence. Next, we restricted observations to the classes Clitellata, Insecta, Malacostraca, Bivalvia and Gastropoda, as they occurred in all data sets. When harmonizing taxonomies across data sets, there is a trade-off between comparability among data sets and taxonomic resolution. If we employ a single taxonomic level (e. g. genus) for all taxa across river types, we have to agglomerate observations with finer resolutions and omit those with coarser resolutions. The resulting data set would be appropriate for comparisons between river types because taxa have consistent

representations across river types. If we vary the taxonomic resolution between river types and taxa, we artificially increase the dissimilarity between river types but attain the best available representation for the individual river types. To compare freshwater typologies (see Comparison of typologies), we used the former approach and harmonized all observations at the genus level, omitting all observations with coarser taxonomic levels. This resulted in a data set with 786 genera. To established typical assemblages (see Typical assemblages), we used the latter approach. Typical assemblages were derive independently for each type, which allowed us to set an appropriate taxonomic level per river type and taxon. For each taxon, we calculated the percentage of observations represented at each finer taxonomic resolution. If this value was below 50 %, we harmonized all observations to this level. If it was above 50%, we repeated the procedure with the next finer level. We also included all observations with coarser than optimal resolutions. The resulting data set contained 370 taxa represented at the species level, 558 taxa represented at genus level and 186 taxa represented at family or coarser taxonomic levels.

## 4.2 Comparison of typologies

### 4.2.1 General approach

Lyche Solheim *et al.* (2019) proposed two versions of the BRT typology. BRT20 differentiates between 20 different river types, some of which are very rare. To reduce the number of broad river types, Lyche Solheim *et al.* (2019) merged rare types with common types that matched the rare type in at least two type descriptors. The resulting typology (BRT12) has twelve types. As previous publications have used both (e.g. Posthuma *et al.* (2020) used BRT12 and Poikane *et al.* (2019) used BRT20) we consider both.

To evaluate the two versions of BRT, we compared them with the three other freshwater typologies: i) the k-means typology of the global river classification framework (GloRiC, Ouellet Dallaire *et al.* 2019); ii) the freshwater ecoregions proposed by Illies (1978); and iii) the biogeographical regions proposed by the European Environmental Agency (BGR, EEA 2016). These typologies can be categorized into: reach-based typologies (BRT and GloRiC) and regional typologies (Illies and BGR). Reach-based typologies classify single river reaches. The instances of individual types are not spatially contiguous and can be far apart. They are commonly very close to instances of different types. By contrast, regional typologies classify large contiguous areas and there is only one instance per type. Different types are only adjacent at the boundaries between two types. As with BRT before, we retained those types for which we had more than 20 samples. See table 2 for the remaining types in each typology.

For each typology, we computed four cluster validity metrics, each emphasizing a different aspect of cluster quality: the average silhouette width, the Calinski-Harabasz index, an indicator value score and the classification strength. We used the Jaccard distance between sites as a distance metric.

As a reference frame for these for metrics, we created two additional typologies. To have an upper bound of what to expect, we clustered the sampling sites with an agglomerative clustering method. This typology is constrained neither by environment nor by space but only by the biological data. Thus, we expect it to delineate more clearly between different biological assemblages than the other typologies and to achieve better cluster validity scores. We will refer to this typology as the biological typology. All agglomerative clustering methods begin with as many clusters as observations and then they iteratively combine the most similar observations into clusters until only one cluster remains. The approach differ in how this similarity is computed. Lance & Williams (1967) clarified that many approaches can be expressed as different parametrizations of one method: flexible beta clustering. We compared different computed biological typologies for different number of types with a  $\beta$  parameters of 0.25. Based on two cluster validity metrics, we decided to use nine types. We refer the interested reader to the supplementary materials for more detailed information.

As a lower bound of what to expect, we created 100 random typologies. Differences between types are purely random. There should be no systematic differences between the types. We hence expect these random typologies to receive lower cluster validity scores than all the other tested typologies. For each random typology, we first drew the number of types as a random variable from the interval between the lowest number of types in any of the typologies (6 in BGR) and the highest number (14 in BRT20 and GloRiC). Then we assigned each sample randomly to one of the groups.

#### **4.2.2 Cluster validity metrics**

Here, we provide a short primer on each cluster validity metric we used. More extended information is provided in the Supplementary Materials. The average silhouette width (ASW, Kaufmann & Rousseeuw 1990) is the average of all silhouette widths. A silhouette width is the difference between the average distance of sample  $i$  to all samples from the most similar type and the average distance of sample  $i$  to all other samples from the same type. The silhouette width is scaled so it lies between -1 and 1. Values close to 1 indicate that most samples are more similar to samples from their own type than to those of other types. Values around zero suggest that the sample lies between two types and values close to -1 signal that the sample is more similar to samples from another type than to those from its own. In addition to the commonly applied ASW, we also computed two modified versions, which

summarize the distance in  $a_i$  and  $b_i$  with the generalized mean instead of the arithmetic mean (Lengyel & Botta-Dukát 2019). The generalized mean has one modifiable parameter,  $p$ , and for  $p = 1$  the modifiable mean is equal to the arithmetic mean. As we increase  $p$  the weight of outlying samples increases. We used  $p = -\infty$  which corresponds to the minimum distance to compute  $ASW_{min}$  and  $p = \infty$  which returns the maximum distance to compute  $ASW_{max}$ .

The Calinski-Harabasz Index (CH, Caliński & Harabasz 1974) is the scaled ratio of the distances between types and the distances within types. Higher values indicate that distances between types are larger than those within types and thus that samples are more similar to samples from their own type than to samples from other types.

The indicator value score (IVS) is based on the indicator value (IndVal) proposed by Dufrêne & Legendre (1997). IVS is the fraction of taxa that are statistically significant indicators (at a significance level of 0.01) for some type of a typology. High scores show that the commonness and specificity of taxa differ between types. Therefore higher scores indicate better typologies.

Lastly, we computed the classification strength (CS, Van Sickle 1997). Classification strength is the difference between mean within cluster similarity ( $\overline{W}$ ) and mean between cluster ( $\overline{B}$ ) similarity. We calculated similarities as  $1 - \text{Jaccard distance}$ . The CS ranges between 0 ( $\overline{W} = \overline{B}$ ) and 1 ( $\overline{B} = 0$ ), where higher values indicate a stronger classification. CS has been used in similar analyses (e.g. Gerritsen *et al.* 2000; Van Sickle & Hughes 2000; Vasconcelos *et al.* 2013), which enables us to directly compare our results with those of previous studies.

We ranked the typologies based on each of these four cluster validity metrics, from 1 to 5 relating to the worst and best performances respectively. Differences smaller than 5% of the range between the biological and the mean of the random typologies were regarded as ties. When typologies were tied, they all received the mean of their untied scores. For example, if three typologies are tied for the first place, instead of 5, 4 and 3 points, each receives a score of 4.

### 4.3 Typical assemblages

We established typical assemblages for all five freshwater typologies using the data set with mixed taxonomic levels. We define a typical taxon as one that frequently occurs in a type and typical assemblages as the lists of these taxa. Typical taxa differ from indicator taxa which usually need to be common and type-specific. As a measure of commonness, we used the percentage of samples from a type that included a taxon. We considered species as typical if they occurred in more than 20% of samples. Coarser taxonomic levels needed to be

more common to be typical. Genera needed to occur in at least 33% of samples and families or coarser taxonomic levels in at least 66%. Having derived typical assemblages that way, we assessed how similar they were within typologies. Low similarities show that the common taxa differ between types, whereas high similarities imply that the same taxa frequently occur irrespective of type. As similarity metric, we used the Jaccard similarity. When the similarity between type A and type B is 0.5, half of the taxa in type A’s typical assemblage also occur in type B. Note that these similarities are asymmetric since they relate to the number of taxa in a typical assemblage. If type B’s typical assemblage has more taxa, the similarity between B and A will be lower than that between A and B.

Lastly, we derived non-redundant typical assemblages. Two typical assemblages were considered redundant when their similarities exceeded 0.7 both ways. We iteratively combined the most similar river types until no redundant types remained. For example, the BRT20-types BRT8, 9 (mid-altitude siliceous rivers) were redundant and combined into BRT8\_9. Subsequently, BRT8\_9 was found to be redundant with BRT16 (glacial rivers) so these types were combined to BRT8\_9\_16. This way, we can evaluate whether the combinations of types from BRT20 to BRT12 are justified by biological homogeneity. In the Supplementary Materials, we provide lists of all complete and non-redundant typical assemblages.

## 4.4 Data and Software

As geospatial representation of the BRT typology, we used the data provided by Globevnik (2019) (Accessed 11.12.2019). The version of the GloRiC typology we used is available under <https://www.hydrosheds.org/page/gloric> (Accessed 30.07.2019). Illies freshwater ecoregions were queried from <https://www.eea.europa.eu/data-and-maps/data/ecoregions-for-rivers-and-lakes> and the EEA biogeographical regions from <https://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3> (both Accessed 08.01.2020). We used the Catchment Characterization and Modeling database (Vogt *et al.* 2007) and the Corine Landcover data from 2018 (CLC 2018) to derive the relative area of land use types in catchments.

All computations were conducted in the R Statistical Environment v. 4.0.3 (R Core Team 2020). Data were prepared using data.table 1.14.0 (Dowle & Srinivasan 2021), tidyverse packages (Wickham *et al.* 2019) and taxize 0.9.98 (Scott Chamberlain & Eduard Szocs 2013; Chamberlain *et al.* 2020). Geospatial analyses were conducted using sf (Pebesma 2018). Clusters were created and evaluated with fpc (Hennig 2020), indicpecies (Caceres & Legendre 2009), labdsv (Roberts 2019), optpart (Roberts 2020). Generalized silhouette widths were computed with the R functions provided in the supplementary materials of

Lengyel & Botta-Dukát (2019). Figures and maps were created with ggplot2 (Wickham 2016) and tmap (Tennekes 2018).

## 5 Results

### 5.1 Comparison of classifications

Overall Illies received the highest score, with 18 of 20 possible points, followed by BGR (14.5), BRT12 (12.5), GloRiC (11) and BRT20 (7.5).

The CH was higher in regional typologies than in reach-based typologies and higher in BGR than in Illies' ecoregions. BRT12 had a higher score than BRT20 and GloRiC had the lowest CH.

Most typologies had similar CS, approximately 0.04. Only Illies' ecoregions exceeded 0.05 with a score of 0.07.

The Illies' ecoregions had the largest IVS closely followed by GloRiC. These two were very close to the biological classification. The three other typologies had similar IVS around 0.09.

Increasing the emphasis on compactness (i.e.  $p$ ) decreased the silhouette widths (Figure 2). All typologies had positive  $ASW_{min}$ , BGR and Illies' ecoregions were close to the biological typology. Reach-based typologies had lower  $ASW_{min}$  than regional ones and GloRiCs was slightly higher than the similar scores of BRT12 and BRT20. With  $p = 1$ , only the ASW of BGR exceeded zero and the random typology. In the remaining typologies, the difference between region- and reach-based typologies shrunk. Illies' ecoregions had a slightly higher ASW than BRT12, which had a higher ASW than BRT20. GloRiC had the lowest ASW. The  $ASW_{max}$  is negative for all typologies, including the biological. The absolute value only declined markedly for BGR, Illies ecoregions and GloRiC, so that BRT12 had the highest  $ASW_{max}$ . It was followed by BGR, BRT12, Illies' ecoregions and GloRiC.

### 5.2 Typical assemblages

The mean similarity in BRT-type typical assemblages was higher than in the other three typologies (Figure 3).

For both BRT-type typical assemblages the mean similarity was above 0.5, approaching 0.75 in BRT12. The mean similarity for the other three typical assemblages all lay around 0.2.

While found high similarities ( $>0.75$ ) in the typical assemblages of all five typologies, only for the BRT-type typologies did we find complete overlap (a similarity of 1).

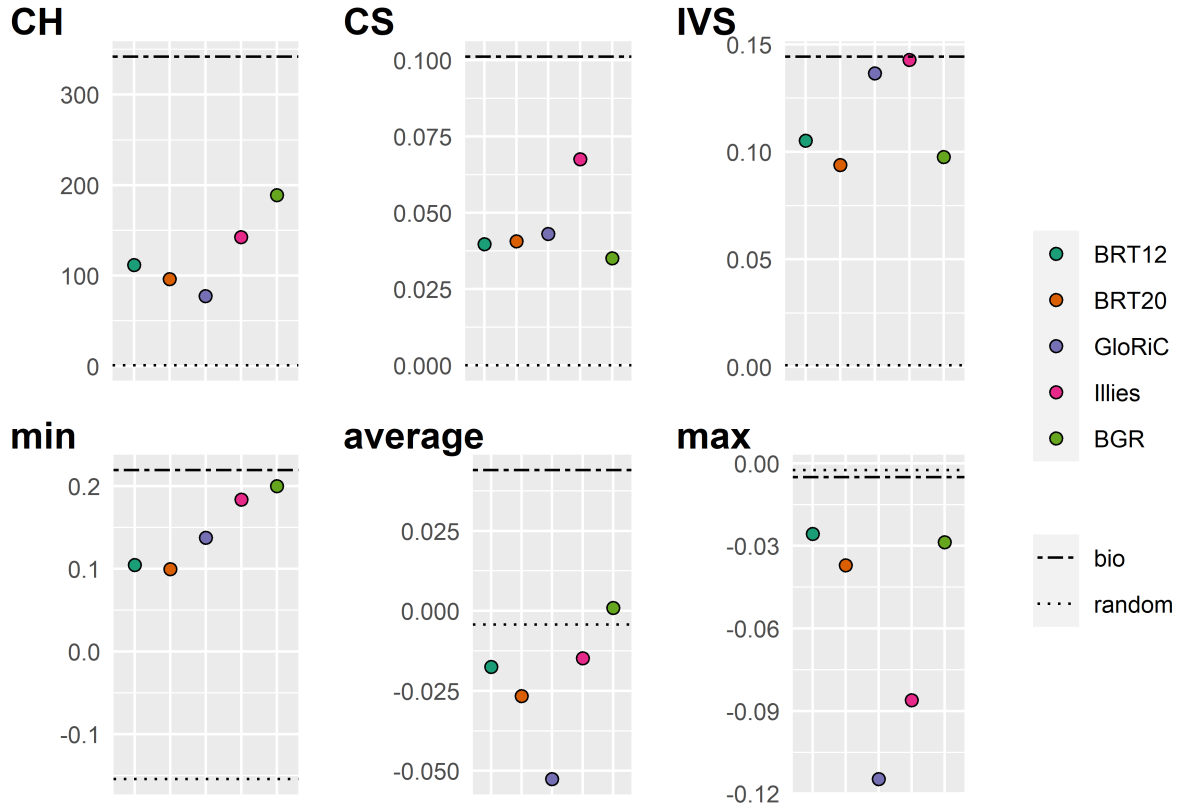


Figure 2: Cluster validity criteria for freshwater typologies. (a) Calinski Harabasz, Classification Strength and Indicator Value score of all eight freshwater typologies. (b) Generalized Silhouette width over the degrees  $p$  of the mean. The dashed horizontal line indicates a score of zero.

The mean size of typical assemblages was similar in all typologies. The mean size over all typologies was 28.4 taxa. BGR had the smallest typical assemblages on average with 25.3 taxa and Illies the largest with 30.8 taxa. The differences are more pronounced in the sizes of individual typical assemblages. The two smallest typical assemblage were from the high-altitude calcareous/ mixed rivers (RT15) of BRT20 with then taxa and the Mediterranean Region of BGR with eleven. The two largest typical assemblages were both from types of Illies freshwater ecoregions: Ireland and Northern Ireland as well as England with 66 and 62 taxa respectively. Note that the two types were also similar to one another and combined for the non-redundant typical assemblages.

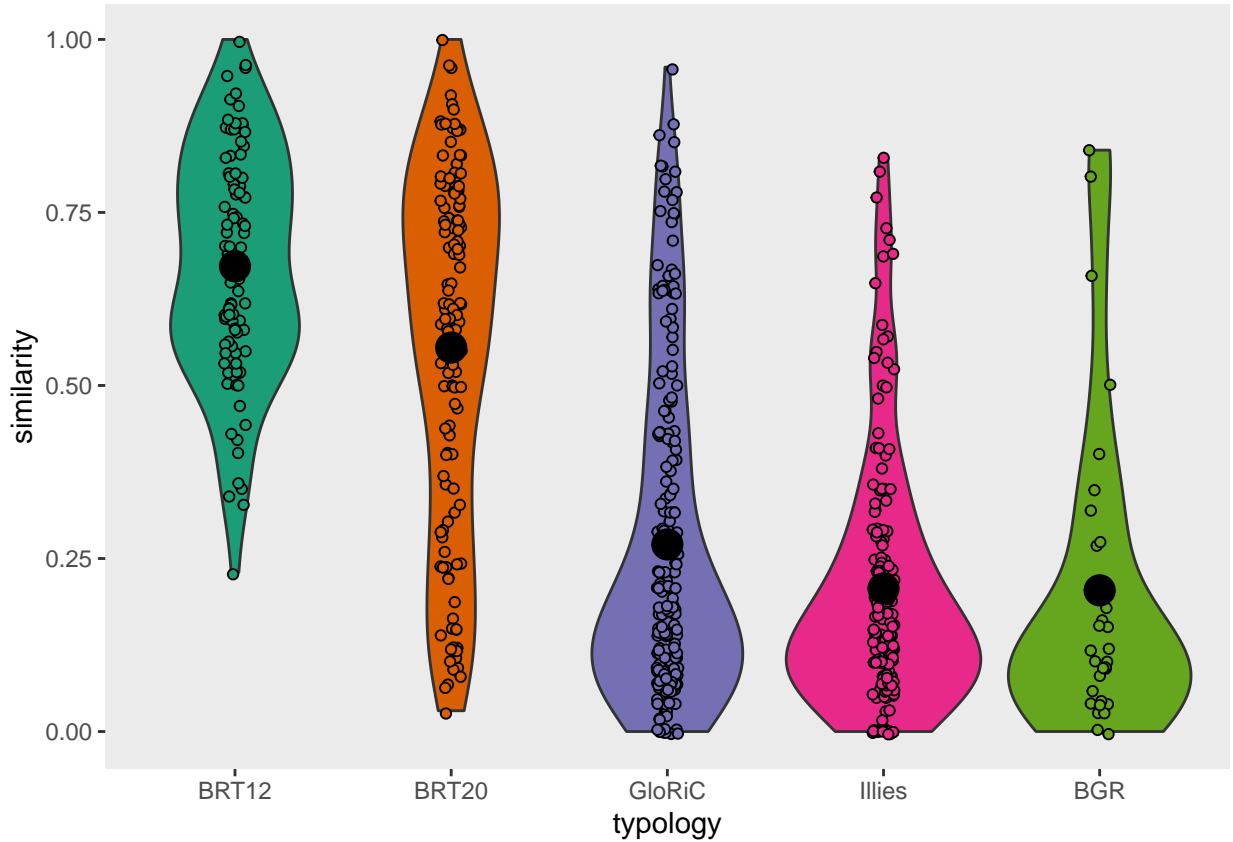


Figure 3: Similarity between typical assemblages for five freshwater typologies. The colored dots are similarities between two types and the violins represent their frequency distribution. The bigger black dots are the arithmetic means.

Based on these typical assemblages, we found redundant types in all typologies except for BGR.

In BRT12, we combined the river types 2, 3, 4, 5, 6 and 7 which include very small to large lowland rivers as well as calcareous mid-altitude rivers. We also combined the two river types 8 and 9 - siliceous mid-altitude rivers. In BRT20, we found the river types 8, 9 and



16 (mid-altitude siliceous and glacial rivers) to be redundant. In addition, we combined the river types 2, 3, 4, 5, 10 and 11. They represent very small to large lowland rivers and mid-altitude calcareous and mixed rivers. In GloRic we combined two river types, 4 and 6. Unfortunately, Ouellet Dallaire *et al.* (2019) do not provide an interpretation for their types beyond the numbers. For Illies, we merged the two types England and Ireland and Northern Ireland as well as Western plains and Western highlands.

The number of taxa in all typical assemblages of a typology varied considerably between typologies. The typical assemblages included 5.7% of taxa for BRT12, 7% for BRT20, 12.5% for GloRic 11.1% for Illies and 8.3% for BGR.

On average, taxa occurred in 4.4 typical assemblages in BRT12, 4.1 in BRT20, 2.6 in GloRic, 2.2 in Illies and 1.5 in BGR.

The two taxon that occurred in the most typical assemblages were both families: Chironomidae (49 of 56 typical assemblages) and Simuliidae (40). The most common genera were the mayfly *Baetis* (38), the caddisfly *Hydropsyche* (37) and the gammarid *Gammarus* (35). The most common species were two mayflies *Serratella ignita* (34) and *Baetis rhodani* (14).

## 6 Discussion

We used four cluster validity metrics to compare the partitioning of macroinvertebrate samples proposed by six different freshwater typologies. Three typologies were reach-based (BRT12, BRT20 and GloRic) and two were regional (Illies and BGR). As a reference frame for the cluster validity scores, we also computed a biological typologies of the samples and 100 random typologies. As we expected, the cluster validity scores of the five freshwater typologies lay in the interval between biological and random typologies in most cases. The only exceptions were silhouette widths (Figure 2).

Overall, the freshwater ecoregions proposed by Illies (1978) performed best in our assessment. The BGR, BRT12, GloRic and BRT20 followed in that order. Illies' freshwater ecoregions had the highest IVS and classification strength. While GloRic is a close second in the IVS, it has the lowest scores in CH and ASW. Between the two BRT typologies, BRT12 always received better or equal scores and had the highest  $ASW_{max}$  of all typologies.

Average silhouette widths were remarkably low for all typologies. Only BGR had a positive ASW and a higher score than the random typologies. The negative ASWs indicate that most communities are on average more similar to communities from other types than from their own.  $ASW_{min}$  gives the largest weight to the separation of clusters without considering compactness. It clearly shows a stronger separation between the types of regional typologies

than between those of reach-based ones. It is also at this level that all typologies have positive ASWs and that the difference to random partitions is largest. The differences between regions and reaches and between typologies and null-partitions decrease with increasing  $p$ . The negative  $ASW_{max}$  indicate that for each of the typologies, there is some community within their type that is more dissimilar to it, than the most dissimilar community from the next closest type. Overall, we see that the typologies discriminate between different different macroinvertebrate faunas. They all have CH, CS and IVS larger than purely random typologies. However, most types seem to have very similar, neighboring types. The assignment of individual communities to a specific type is therefore far from clear, as communities vary as strongly within types as among neighboring ones. We had insufficient data to evaluate most of the combinations made to arrive at BRT12 from BRT20. Many of the types that were merged are very rare (e.g. BRT7, BRT20) However, our analysis lends support to the combination of BRT14, 15 and 16 to a single high-altitude river type. We found the typical assemblages of BRT15 and 16 to be redundant. BRT14 possibly only differed from those two types due to the coarse taxonomic resolution of our data for this type. As BRT12 generally performed better than BRT20 in this study, the combinations made by Lyche Solheim *et al.* (2019) seem to be justified. We suggest that future studies should rather employ the aggregated than the full typology. However, given the low performance of reach-based typologies in general, we advise caution when assuming spatially homogeneous reference communities within them.

Our analyses further suggest that large redundancies remain between the typical assemblages of the broad river types. We found four different mid-altitude river types and two low altitude types to have very similar typical taxa. These redundant river types extended across gradients in altitude, geology and river size, though they omit very large or high altitude rivers. This might indicate several things. More finely resolved taxonomic data could be necessary to show differences between these types. Potentially, the BRT typology with twenty types is too coarse. Our analysis identified mostly those taxa as common that are sufficiently euryoecious that their distribution is largely independent from the features that delineate the different river types. Within the individual river types only few type-specific taxa were common. This could indicate that even within a single river type environmental conditions vary sufficiently to limit the distribution of more stenooecious species. It is also possible that common anthropogenic stressors homogenized the assemblages (Mondy & Usseglio-Polatera 2014). We removed impaired samples from impaired sites to prevent this but the definition of reference conditions and their identification after the fact and at large spatial scales, are difficult (Herlihy *et al.* 2008; Pardo *et al.* 2012; McNellie *et al.* 2020).

The typical assemblage of very large river (BRT1) harbors many genera that include invasive species like *Dreissena polymorpha*, *Dikerogammarus villosus*, *Corbicula fluminea* and *Potam-*

*ogyrrus antipodarum*. The presence of these taxa within the typical assemblage thus highlights the importance of neozoa in very large rivers (Leitner *et al.* 2020). The typical assemblage for the large combined river type BRT2\_3\_8\_9\_10\_11 mainly includes Trichoptera (6), Ephemeroptera (6) and Elmidae (4) genera. These genera include several species that are typical for the rhithral, such as *Esolus parallelepipedus*, *Elmis aenea* and *Hydropsyche pellucidula*. In addition, several taxa that are more common in mid-altitude rivers are included like Rhithrogena or Ecdyonurus. The typical assemblage of siliceous and organic high altitude rivers (BRT14) only contains families or lower taxonomic levels. This is most likely what leads to the differentiation from the combined high altitude type BRT15\_16. The combined high altitude river type (BRT15\_16) has a higher taxonomic resolution and contains the largest number of Plecoptera genera of any typical assemblage (4), among them Isoperla, which is common in high-altitude rivers. Lastly, the typical assemblage of Mediterranean mid-altitude rivers also exclusively consists of families or lower taxonomic levels. Unique to this typical assemblage is only the family Gerridae.

Overall, regional approaches that assign large contiguous areas to a single type were a better summary of our extensive data set of invertebrate occurrences than reach-scale typologies. The constraints of our study may have played in favor of regional approaches. The taxonomic resolution of our data only allowed inferences at the genus level. Potentially, species-level data would have highlighted bigger differences at smaller scales and hence been better summarized by typologies with smaller mapping units. Along the same lines, Moog *et al.* (2004) showed that higher taxonomic resolution was necessary to delineate smaller ecoregions and the same logic could expand to reach-scale typologies. Similarly, Verdonshot (2006a) showed that differences in classification strength were noticeable between ‘best-available’ and family-level data on a pan-European scale. However, the difference was small and their analysis should not be extrapolated to infer if this difference also exists between species- and genus-level data. Hawkins *et al.* (2000) note that, across taxa groups and typologies, higher taxonomic resolution does not always lead to greater classification strength.

Additionally, the likelihood of assigning an observation to the wrong type due to inaccuracies in the spatial data are lower for regional typologies, where such mistakes can only happen toward the margins between types. We reduced this error by using a more stringent distance-to-river criterion for the selection of sites than similar studies. For example, Irving *et al.* (2018) used a cutoff of 3 km to assign streamflow gauging stations to rivers. Very small rivers are largely missing from the river network representations used by GloRiC and the BRT as provided in Globovnik (2019). Hence, samples from these rivers are either omitted from the analysis if they are too far removed from the next river or assigned to the next river segment. In most cases, this is likely but not certainly a reach of a similar type. New and more highly resolved river networks (e.g. Lin *et al.* 2021) might provide more accurate

mapping of river types and ameliorate these problems.

The patterns we observed can also be interpreted in the light of metacommunity theory. Our analysis of generalized silhouette widths showed that the separation between communities of different BRT types decreases when we give more weight to outliers. The separation declines up to a point where it is lower than in the random typologies. This pattern could be explained by strong dispersal limitation leading to divergent communities in comparable environments. The spatial distance between the instances of different types in reach-scale typologies is often small. Typically, landscapes and catchments are made up of mosaics of different river types. In positing that assemblages vary between instances of different reach-scale types, we implicitly assume that dispersal is strong enough for species to track variation in local conditions (Leibold *et al.* 2004) but not so strong that mass effect can overrule environmentally induced patterns (Mouquet & Loreau 2003). If taxa are strongly limited by dispersal e.g. through historical and anthropogenic dispersal barriers (Leibold *et al.* 2010; Belletti *et al.* 2020) or through the sheer distance between sites, they are unable to reach potentially favorable sites. In our case, stronger dispersal limitation would lead to larger differences between sites of the same type. Dispersal limitation leads to a pattern known as distance decay, which is often found at large spatial scales (Nekola & White 1999; Morlon *et al.* 2008) but varies in its magnitude between taxa, realms and degrees of latitude (Graco-Roza *et al.* 2021). Studies have found both, invertebrate communities that are predominantly structured by environment (e.g. Heino *et al.* 2012; Landeiro *et al.* 2012) and those that are structured by space (Mykrä *et al.* 2007; Astorga *et al.* 2012). However, even for microbes, which under the Bass-Becking hypothesis (“everything is everywhere but the environment selects”) were long considered to have cosmopolitan distributions, the role of dispersal limitation, at least at large spatial scales, is becoming evident (Telford *et al.* 2006; Lindström & Langenheder 2012; Soininen 2012).

The better performance of regional typologies might also be taken to support the notion, that the composition of aquatic macroinvertebrate communities are adapted to changes in large-scale rather than small-scale variables. Many studies have investigated this question before. Among them are studies that have studied the explanatory potential of ecoregions, which might be considered as categorical combinations of large-scale environmental conditions like climate, geology and altitude. Verdonschot & Nijboer (2004) used the data compiled during the AQEM project to investigate potential drivers of macroinvertebrate community composition. They found that the distribution of macroinvertebrates follows Illies’ freshwater ecoregions and that large-scale variables like geology explain most of the variation. However, using an extended data set, that included data from both AQUEM and STAR, Verdonschot (2006b) did find ecoregions to be a rather weak predictor of invertebrate community composition. Similarly, ecoregions explained a negligible amount of

variance in community composition in Swedish boreal rivers (Johnson *et al.* 2004). The review of Hawkins *et al.* (2000) finds that ecoregion-based approaches usually outperform catchment-scale classifications. However, the classification strength was generally low and lay between 0.07 and 0.16, which slightly exceeds the ones we found (between 0.04 to 0.07) but is in agreement with several other studies (e.g. Snelder *et al.* 2004; Heino & Mykrä 2006; Mykrä *et al.* 2009; Vasconcelos *et al.* 2013). The interplay between local and regional control cannot be considered detached from the previous discussion of dispersal limitation since the degree to which assemblages are regulated regionally or locally likely depends on the magnitude of dispersal (Ryabov & Blasius 2011). Synthesis of these findings is impeded by the fact that many of the studies relied on combinations of variance partitioning and eigenfunction analyses that have been shown to be flawed (Gilbert & Bennett 2010; Smith & Lundholm 2010; Tuomisto *et al.* 2012). An alternative method to evaluate the importance of dispersal processes for community assembly was recently proposed (Vilmi *et al.* 2020) but has yet to be extensively tested and applied.

Typologies based on environmental variables cannot account for the effects of biotic interactions on species distributions. While these received much attention in the early days of macroecology (Diamond 1975; Connor & Simberloff 1983) the focus shifted towards environmental filters (e.g. Whittaker & others 1970; Pearson & Dawson 2003) assuming that biotic interactions are only important at local scales. Their role for macroecology has been revisited (e.g. Wisz *et al.* 2013) partly due to the realization that many studies can not differentiate between environmental filtering *sensu stricto* and biotic interactions (Kraft *et al.* 2015). Propelled by a series of new methods (e.g. Popovic *et al.* 2019; Ovaskainen & Abrego 2020) several studies have recently investigated the effect of biotic interactions on larger biogeographical patterns. While some indicate, that the role of environmental conditions is more important than that of biotic interactions (Elo *et al.* 2021), others suggest that biotic interactions modulate the response to the abiotic environment (Abrego *et al.* 2021).

<– Ways forward –> While there were clear and systematic differences between typologies, all cluster validity metrics had comparatively low scores. It is beyond the scope of the current study to assess whether this is a problem of the implementation or a general issue, i.e. if it is theoretically possible to create environmental typologies that capture large amounts of variation between communities or if niche processes or environmental filters *sensu stricto* (Kraft *et al.* 2015) generally do not explain enough variance. The latter would constitute a serious problem for the reference state approach as this would entail that spatial stability of reference communities (Statzner *et al.* 2001) should generally not be assumed. However, since several studies have found typologies to explain a significant amount of variation (e.g. Lorenz *et al.* 2004; Johnson *et al.* 2007) we deem issues with the implementation more likely.

The river typology of the WFD must be optimized not only for macroinvertebrates but for all four biological quality elements (macroinvertebrates, diatoms, fish and macrophytes) (Dodkins *et al.* 2005). Future studies should assess whether the conclusions drawn from this study are supported by results for the other taxa groups. River classifications must be understood as models Loveland & Merchant (2004). Alongside our growing knowledge about freshwater systems, we should strive to further improve them, as to increase their value for biomonitoring and to render them more effective tools for conservation. Our study has shown that the typologies proposed by Lyche Solheim *et al.* (2019) should not be the end of this effort.

## 7 Supplementary Materials

### 7.1 Distribution of removed sampling sites

In table 1, *Land use* refers to the 20% cutoff for relative area of agriculture or urban areas in a catchment. This criterion was applied wherever no other information was available. IASPT is short for Iberian Average score per taxon which refers to the average Iberian biological monitoring working party score (IBMWP, between 1 and 10) of a sample. The IASPT has been shown to be a reliable indicator of water quality (e.g. Leunda *et al.* 2009; Munné & Prat 2009). The Biological Diatom Index (BDI) (Lenoir & Coste 1996) is a standardized way to evaluate the quality of freshwater systems in France using diatoms.

In two of the data sets (RIVPACS and Koutajoki) the sampling sites were selected to represent reference conditions so we used all of the sites. For the data from the RCS national network several qualitative metrics were available. The measured stressors were: organic matter, nitrogen compounds (except nitrates), nitrates, phosphorous compounds, suspended matter, organic micropollutants (other), mineral micropollutants (metal), pesticides and polycyclic aromatic hydrocarbons (PAH). Each was rated on a scale of high, good, intermediate, poor and bad. All sites for which two or more categories were rated as intermediate or one more were rated as either poor or bad was removed from the analysis.

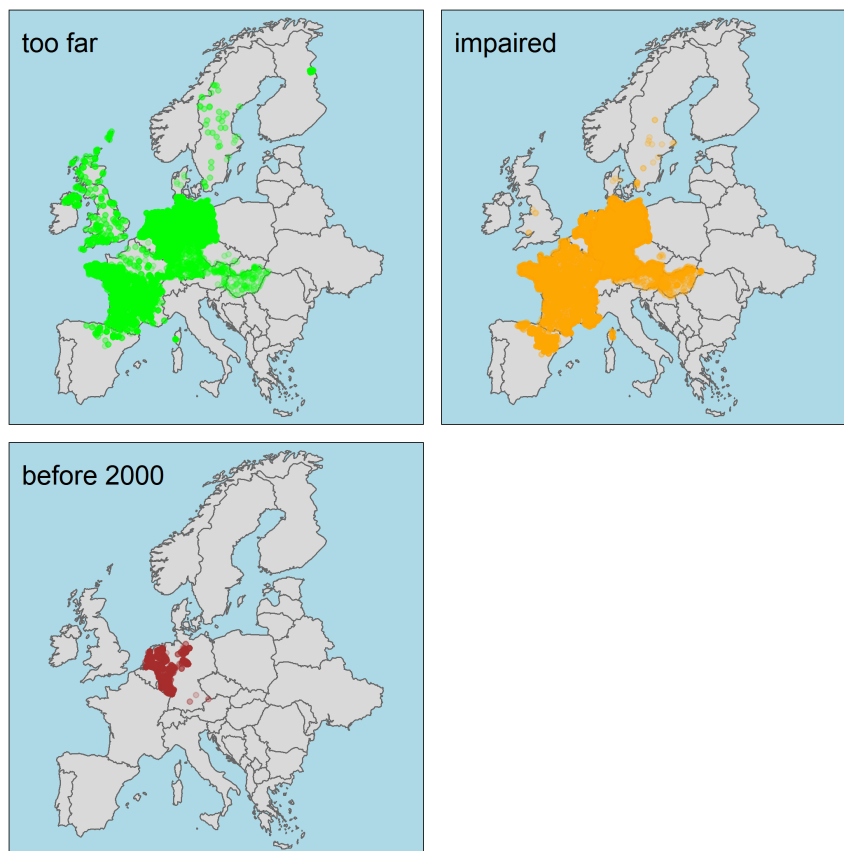


Figure 4: Maps of removed sampling sites. The first map (too far) shows sites that were more than 500 meters removed from the closeded BRT river segment. The second map (impaired) shows sites that were categorized as impaired. The third map (before 2000) shows sites that were sampled before 2000.

## 7.2 Flexible beta clustering

Flexible beta clustering has four parameters but it is common to only specify one,  $\beta$ , while the other three parameters are derived from it.

This section describes in greater detail how we computed the flexible beta clustering of the sampling sites. We used flexible beta clustering as it has been shown to be robust to various potential problems like outliers or noisy data (Milligan 1989). We employed the common parameterization of  $(\alpha_1 = \alpha_2 = \alpha, \beta = (1 - \alpha)/2, \gamma = 0)$  and varied beta in ten steps from -1 to 1 (-1, -0.77, -0.55, -0.33, -0.11, 0.11, 0.33, 0.55, 0.77, 1). We cut the resulting dendograms at each split between 2 and 30 final nodes. For each of the resulting 290 clusterings we calculated the average silhouette, width, cophenetic correlation coefficient and the cluster membership entropy.



## 7.3 Cluster validity metrics

### 7.3.1 Average silhouette width

The silhouette width  $s(i)$  is defined as follows:

1. Compute  $a_i$ , the average dissimilarity between sample  $i$  and all other samples from the same type
2. Among all other clusters  $C$ , find the cluster than minimizes  $d(i, C)$  the average distance of  $i$  to all observations of  $C$ . This cluster is  $b$ ;  $b_i = \min_C(d(i, C))$ .
3.  $s(i) = \frac{b_i - a_i}{\max(a_i, b_i)}$
4. The average silhouette width is the arithmetic mean of all silhouette widths.  $ASW = \frac{1}{n} \sum_{i=1}^n s(i)$

Lengyel & Botta-Dukát (2019) recently proposed a generalized version of the ASW. The classic ASW employs the arithmetic average to compute  $a_i$  and  $b_i$ , which assumes that spherical clusters are optimal. Using a generalized mean instead, we can flexibly adjust our validity metric to put a stronger emphasis on compactness ( $a_i$ ) or separation ( $b_i$ ). The generalized mean of degree  $p$  ( $M^p$ ) is computed as:

$$M^p(\mathbf{x}) = \left( \frac{1}{n} \sum_{i=1}^n x_i^p \right)^{1/p}$$

The generalized mean can take the value of common summary statistics such as the minimum ( $p = -\infty$ ), the maximum ( $p = \infty$ ) or the harmonic mean ( $p = -1$ ). For example, for  $p = -\infty$  the silhouette width is the difference between the minimum distance of observation  $i$  to any other observation from the same type and the minimum distance from that observation to any observation from the next closest type. This perspective excludes outliers and values separation over compactness. The weighting shifts towards compactness as we increase  $p$ . We evaluated the silhouette width for  $p \in \{-\infty, -2, -1, 1, 2, \infty\}$ . If not further specified, ASW refers to the common average silhouette width (i.e.  $p = 1$ ) in the remainder of the text.

### 7.3.2 Calinski Harabasz Index

The Calinski-Harabasz Index (CH, Caliński & Harabasz 1974) is computed as

$$CH = \frac{BGSS}{WGSS} \times \frac{n - k}{k - 1}$$

where  $BGSS$  is the squared sum of distances between group centroids and the overall centroid (between group sum-of-squares),  $WGSS$  is sum of squares of distances between observations of one group (within group sum-of-squares),  $k$  is the number of clusters. High values indicate that variation within types is smaller than between types. As the second term controls for the degrees of freedom, here determined by the number of clusters, it can be understood as an analog to the F-Statistic. The algorithm assumes Euclidean data but good performance with a similar metric was shown for binary data in the context of fMRI-scans (Dimitriadou *et al.* 2004).

### 7.3.3 Indicator Value Statistic

The IndVal itself will be explained in detail below. Here, we only note that we used 999 permutations to compute  $p$ -values and ignored the family-wise error rate.

Both indicator and typical assemblages were derived for BRT20. We used the IndVal approach of Dufrêne & Legendre (1997) to identify indicator taxa. For this analysis, we used *data genus* which consists of genus-level presence-absence data. The IndVal can be understood as the product of the two quantities  $A$  and  $B$ . For our purposes,  $A$  is the relative number of observations of taxon  $i$  that are within type  $j$ . It was initially described as specificity (Dufrêne & Legendre 1997) but is better understood as concentration (Podani & Csányi 2010) because it is independent of the total number of types.  $B$  is the relative frequency with which species  $i$  occurs in type  $j$ . The maximum score is assigned to a species that only occurs in one type ( $A = 1$ ) and occurs in all samples of that type ( $B = 1$ ). Here, we used the group-equalized version of the IndVal which accounts for the fact that the number of samples differs between types. The statistical significance of the IndVal statistic was assessed with a permutation test that computes IndVal values for random permutations of sites and types and compares the observed IndVal against this empirical distribution. We used  $2 * 10^5$  permutations. We controlled the family-wise error rate with Bonferroni's correction (Bonferroni 1935). This correction is conservative and has low statistical power (García 2004; Nakagawa 2004). Hence, we regard the hypothesis, that a genus is an indicator for a broad river type, strongly supported by our data if we find the indication to be statistically significant. We used 0.05 as the base significance level.

Table 2: Types of different typologies that were used in the analyses.

Typology	Types
BRT12	RT1, RT2, RT3, RT4, RT5, RT6, RT7, RT8, RT9, RT10, RT11
BRT20	RT1, RT2, RT3, RT4, RT5, RT8, RT10, RT11, RT14, RT15, RT16, RT17, RT18
GloRiC	3, 4, 6, 7, 8, 14, 15, 17, 18, 22, 24, 26, 29, 30
Illies	Alps, Central highlands, Central plains, England, Fenno-scandian shield, Hungarian lowlands, Ibero-Macaronesian region, Ireland and Northern Ireland, Italy and Corsica, Pyrenees, Western highlands, Western plains
BGR	Alpine Bio-geographical Region, Atlantic Bio-geographical Region, Boreal Bio-geographical Region, Continental Bio-geographical Region, Mediterranean Bio-geographical Region, Pannonian Bio-geographical Region

## References

- Abrego, N., Roslin, T., Huotari, T., Ji, Y., Schmidt, N.M., Wang, J., *et al.* (2021). Accounting for species interactions is necessary for predicting how arctic arthropod communities respond to climate change. *Ecography*, 1–12.
- Alric, B., Dézerald, O., Meyer, A., Billoir, E., Coulaud, R., Larras, F., *et al.* (2021). How diatom-, invertebrate- and fish-based diagnostic tools can support the ecological assessment of rivers in a multi-pressure context: Temporal trends over the past two decades in France. *Science of the Total Environment*, 762, 143915.
- Alvarez-Cabria, M., Barquin, J. & Juanes, J.A. (2010). Spatial and seasonal variability of macroinvertebrate metrics: Do macroinvertebrate communities track river health? *Ecological Indicators*, 10, 370–379.
- Alvarez-Cabria, M., Barquin, J. & Juanes, J.A. (2011). Macroinvertebrate community dynamics in a temperate european atlantic river. Do they conform to general ecological theory? *Hydrobiologia*, 658, 277–291.
- Astorga, A., Oksanen, J., Luoto, M., Soininen, J., Virtanen, R. & Muotka, T. (2012). Distance decay of similarity in freshwater communities: Do macro- and microorganisms follow the same rules? *Global Ecology and Biogeography*, 21, 365–375.
- Baranov, V., Jourdan, J., Pilotto, F., Wagner, R. & Haase, P. (2020). Complex and nonlinear climate-driven changes in freshwater insect communities over 42 years. *Conservation Biology*, cobi.13477.
- Belletti, B., Garcia de Leaniz, C., Jones, J., Bizzi, S., Börger, L., Segura, G., *et al.* (2020). More than one million barriers fragment Europe’s rivers. *Nature*, 588, 436–441.
- Berger, E., Haase, P., Schäfer, R.B. & Sundermann, A. (2018). Towards stressor-specific macroinvertebrate indices: Which traits and taxonomic groups are associated with vulnerable and tolerant taxa? *Science of the Total Environment*, 619, 144–154.
- Béné, C., Arthur, R., Norbury, H., Allison, E.H., Beveridge, M., Bush, S., *et al.* (2016). Contribution of fisheries and aquaculture to food security and poverty reduction: Assessing the current evidence. *World Development*, 79, 177–196.
- Bhowmik, A.K. & Schaefer, R.B. (2015). Large scale relationship between aquatic insect traits and climate. *PLoS one*, 10, e0130025.
- Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., *et al.* (2012). Three hundred ways to assess Europe’s surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. *Ecological Indicators*, 18,

31–41.

- Birk, S., Chapman, D., Carvalho, L., Spears, B.M., Andersen, H.E., Argillier, C., *et al.* (2020). Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. *Nature Ecology & Evolution*.
- Birk, S., Willby, N.J., Kelly, M.G., Bonne, W., Borja, A., Poikane, S., *et al.* (2013). Inter-calibrating classifications of ecological status: Europe’s quest for common management objectives for aquatic ecosystems. *Science of the Total Environment*, 454–455, 490–499.
- Bonferroni, C.E. (1935). Il calcolo delle assicurazioni su gruppi di teste. *Studi in onore del professore salvatore ortu carboni*, 13–60.
- Börger, T., Campbell, D., White, M.P., Elliott, L.R., Fleming, L.E., Garrett, J.K., *et al.* (2021). The value of blue-space recreation and perceived water quality across europe: A contingent behaviour study. *Science of The Total Environment*, 145597.
- Caceres, M.D. & Legendre, P. (2009). Associations between species and groups of sites: inindices and statistical inference. *Ecology*, 90, 3566–3574.
- Caliński, T. & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3, 1–27.
- Cardinale, B.J. (2011). Biodiversity improves water quality through niche partitioning. *Nature*, 472, 86–89.
- Carvalho, L., Mackay, E.B., Cardoso, A.C., Baattrup-Pedersen, A., Birk, S., Blackstock, K.L., *et al.* (2019). Protecting and restoring Europe’s waters: An analysis of the future development needs of the Water Framework Directive. *Science of the Total Environment*, 658, 1228–1238.
- Chamberlain, S., Szoecs, E., Foster, Z., Arendsee, Z., Boettiger, C., Ram, K., *et al.* (2020). *Taxize: Taxonomic information from around the web*.
- Clarke, R.T., Wright, J.F. & Furse, M.T. (2003). RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. *Ecological modelling*, 160, 219–233.
- CLC. (2018). CORINE land cover. <https://land.copernicus.eu/pan-european/corine-land-cover>.
- Comte, L., Grantham, T. & Ruhi, A. (2021). Human stabilization of river flows is linked with fish invasions across the USA. *Global Ecology and Biogeography*, 725–737.
- Connor, E.F. & Simberloff, D. (1983). Interspecific competition and species co-occurrence patterns on islands: Null models and the evaluation of evidence. *Oikos*, 455–465.

- Costanza, R., De Groot, R., Sutton, P., Van der Ploeg, S., Anderson, S.J., Kubiszewski, I., *et al.* (2014). Changes in the global value of ecosystem services. *Global environmental change*, 26, 152–158.
- Cumberlidge, N., Ng, P.K., Yeo, D.C., Magalhães, C., Campos, M.R., Alvarez, F., *et al.* (2009). Freshwater crabs and the biodiversity crisis: Importance, threats, status, and conservation challenges. *Biological Conservation*, 142, 1665–1673.
- Darwall, W., Bremerich, V., De Wever, A., Dell, A.I., Freyhof, J., Gessner, M.O., *et al.* (2018). The Alliance for Freshwater Life: A global call to unite efforts for freshwater biodiversity science and conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 28, 1015–1022.
- Diamond, J.M. (1975). Ecology and evolution of communities. In: *Ecology and evolution of communities* (ed. Cody, J.M., M. L. & Diamond). Harvard University Press, pp. 342–444.
- Dimitriadou, E., Barth, M., Windischberger, C., Hornik, K. & Moser, E. (2004). A quantitative comparison of functional MRI cluster analysis. *Artificial Intelligence in Medicine*, 31, 57–71.
- Dodkins, I., Rippey, B., Harrington, T.J., Bradley, C., Chathain, B.N., Kelly-Quinn, M., *et al.* (2005). Developing an optimal river typology for biological elements within the Water Framework Directive. *Water Research*, 39, 3479–3486.
- Dowle, M. & Srinivasan, A. (2021). *Data.table: Extension of ‘data.frame’*.
- Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., *et al.* (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews*, 81, 163–182.
- Duffy, J.E., Godwin, C.M. & Cardinale, B.J. (2017). Biodiversity effects in the wild are common and as strong as key drivers of productivity. *Nature*, 549, 261–264.
- Dufrêne, M. & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological monographs*, 67, 345–366.
- EEA. (2016). Biogeographical Regions.
- EEA. (2018). European waters assessment of status and pressures 2018. *EEA Report*.
- Elo, M., Jyrkänkallio-Mikkola, J., Ovaskainen, O., Soininen, J., Tolonen, K.T. & Heino, J. (2021). Does trait-based joint species distribution modelling reveal the signature of competition in stream macroinvertebrate communities? *Journal of Animal Ecology*.

- Escribano, N., Oscoz, J., Galicia, D., Cancellario, T., Durán, C., Navarro, P., *et al.* (2018). Freshwater macroinvertebrate samples from a water quality monitoring network in the iberian peninsula. *Scientific data*, 5, 1–6.
- European Commission. (2000). DIRECTIVE 2000/60/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Union*, L327.
- Floury, M., Usseglio-Polatera, P., Ferreol, M., Delattre, C. & Souchon, Y. (2013). Global climate change in large European rivers: Long-term effects on macroinvertebrate communities and potential local confounding factors. *Global Change Biology*, 19, 1085–1099.
- García, L.V. (2004). Escaping the Bonferroni iron claw in ecological studies. *Oikos*, 105, 657–663.
- Gerritsen, J., Barbour, M.T. & King, K. (2000). Apples, oranges, and ecoregions: On determining pattern in aquatic assemblages. *Journal of the North American Benthological Society*, 19, 487–496.
- Gilbert, B. & Bennett, J.R. (2010). Partitioning variation in ecological communities: Do the numbers add up? *Journal of Applied Ecology*, 47, 1071–1082.
- Globevnik, L. (2019). Broad typology for rivers and lakes in Europe for large scale analysis.
- Goodwin, C.N. (1999). Fluvial classification: Neanderthal necessity or needless normalcy. *Wildland Hydrology*, 5230, 229–236.
- Graco-Roza, C., Aarnio, S., Abrego, N., Acosta, A.T., Alahuhta, J., Altman, J., *et al.* (2021). Distance decay 2.0—a global synthesis of taxonomic and functional turnover in ecological communities. *bioRxiv*.
- Grizzetti, B., Liqueste, C., Pistocchi, A., Vigiak, O., Zulian, G., Bouraoui, F., *et al.* (2019). Relationship between ecological condition and ecosystem services in European rivers , lakes and coastal waters. *Science of the Total Environment*, 671, 452–465.
- Haubrock, P.J., Pilotto, F. & Haase, P. (2020). Do changes in temperature affect EU Water Framework Directive compliant assessment results of central European streams? *Environmental Sciences Europe*, 32.
- Hawkins, C.P., Norris, R.H., Gerritsen, J., Hughes, R.M., Jackson, S.K., Johnson, R.K., *et al.* (2000). Evaluation of the Use of Landscape Classifications for the Prediction of Freshwater Biota : Synthesis and Recommendations. *Journal of the North American Benthological Society*, 19, 541–556.

- He, F., Zarfl, C., Bremerich, V., David, J.N.W., Hogan, Z., Kalinkat, G., *et al.* (2019). The global decline of freshwater megafauna. *Global Change Biology*, 25, 3883–3892.
- Heino, J., Grönroos, M., Soininen, J., Virtanen, R. & Muotka, T. (2012). Context dependency and metacommunity structuring in boreal headwater streams. *Oikos*, 121, 537–544.
- Heino, J. & Mykrä, H. (2006). Assessing physical surrogates for biodiversity: Do tributary and stream type classifications reflect macroinvertebrate assemblage diversity in running waters? *Biological Conservation*, 129, 418–426.
- Hennig, C. (2020). *Fpc: Flexible procedures for clustering*.
- Herlihy, A.T., Paulsen, S.G., Van Sickle, J., Stoddard, J.L., Hawkins, C.P. & Yuan, L.L. (2008). Striving for consistency in a national assessment: The challenges of applying a reference-condition approach at a continental scale. *Journal of the North American Benthological Society*, 27, 860–877.
- Huttunen, K.-L., Mykrä, H., Oksanen, J., Astorga, A., Paavola, R. & Muotka, T. (2017). Habitat connectivity and in-stream vegetation control temporal variability of benthic invertebrate communities. *Scientific Reports*, 7, 1–9.
- Illies, J. (1978). *Limnofauna europaea*. Fischer Stuttgart.
- Irving, K., Kuemmerlen, M., Kiesel, J., Kakouei, K., Domisch, S. & Jähnig, S.C. (2018). A high-resolution streamflow and hydrological metrics dataset for ecological modeling using a regression model. *Scientific data*, 5, 1–14.
- Johnson, R.K., Furse, M.T., Hering, D. & Sandin, L. (2007). Ecological relationships between stream communities and spatial scale: Implications for designing catchment-level monitoring programmes. *Freshwater Biology*, 52, 939–958.
- Johnson, R.K., Goedkoop, W. & Sandin, L. (2004). Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes. *Freshwater Biology*, 49, 1179–1194.
- Jourdan, J., O'Hara, R.B., Bottarin, R., Huttunen, K.L., Kuemmerlen, M., Monteith, D., *et al.* (2018). Effects of changing climate on European stream invertebrate communities: A long-term data analysis. *Science of the Total Environment*, 621, 588–599.
- Kaufmann, L. & Rousseeuw, P. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley&Sons.
- Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S. & Levine, J.M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, 29, 592–599.



- Lance, G.N. & Williams, W.T. (1967). A general theory of classificatory sorting strategies: 1. Hierarchical systems. *The computer journal*, 9, 373–380.
- Landeiro, V.L., Bini, L.M., Melo, A.S., Pes, A.M.O. & Magnusson, W.E. (2012). The roles of dispersal limitation and environmental conditions in controlling caddisfly (trichoptera) assemblages. *Freshwater Biology*, 57, 1554–1564.
- Le, T.D.H., Schreiner, V.C., Kattwinkel, M. & Schäfer, R.B. (2021). Invertebrate turnover along gradients of anthropogenic salinisation in rivers of two German regions. *Science of the Total Environment*, 753, 141986.
- Leibold, M.A., Economo, E.P. & Peres-Neto, P. (2010). Metacommunity phylogenetics: Separating the roles of environmental filters and historical biogeography. *Ecology letters*, 13, 1290–1299.
- Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., *et al.* (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology letters*, 7, 601–613.
- Leitner, P., Borgwardt, F., Birk, S. & Graf, W. (2020). Multiple stressor effects on benthic macroinvertebrates in very large European rivers – A typology-based evaluation of faunal responses as a basis for future bioassessment. *Science of The Total Environment*, 143472.
- Lengyel, A. & Botta-Dukát, Z. (2019). Silhouette width using generalized mean—A flexible method for assessing clustering efficiency. *Ecology and Evolution*, 9, 13231–13243.
- Lenoir, A. & Coste, M. (1996). Development of a practical diatom index of overall water quality applicable to the french national water board network. In: *International symposium, volksbildungsheim grilhof vill, AUT, 17-19 september 1995*. Universität Innsbruck, pp. 29–43.
- Leunda, P.M., Oscoz, J., Miranda, R. & Ariño, A.H. (2009). Longitudinal and seasonal variation of the benthic macroinvertebrate community and biotic indices in an undisturbed Pyrenean river. *Ecological Indicators*, 9, 52–63.
- Leung, B., Hargreaves, A.L., Greenberg, D.A., McGill, B., Dornelas, M. & Freeman, R. (2020). Clustered versus catastrophic global vertebrate declines. *Nature*.
- Lin, P., Pan, M., Wood, E.F., Yamazaki, D. & Allen, G.H. (2021). A new vector-based global river network dataset accounting for variable drainage density. *Scientific Data*, 1–9.
- Lindström, E.S. & Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. *Environmental Microbiology Reports*, 4, 1–9.

- Lorenz, A., Feld, C.K. & Hering, D. (2004). Typology of streams in Germany based on benthic invertebrates: Ecoregions, zonation, geology and substrate. *Limnologica*, 34, 379–389.
- Lorenz, A., Feld, C.K. & Hering, D. (2004). Typology of streams in Germany based on benthic invertebrates: Ecoregions, zonation, geology and substrate. *Limnologica*, 34, 379–389.
- Loveland, T.R. & Merchant, J.M. (2004). Ecoregions and ecoregionalization: geographical and ecological perspectives. *Environmental management*, 34 Suppl 1, 1–13.
- Lyche Solheim, A., Austnes, K., Globevnik, L., Kristensen, P., Moe, J., Persson, J., *et al.* (2019). A new broad typology for rivers and lakes in Europe: Development and application for large-scale environmental assessments. *Science of the Total Environment*, 697, 134043.
- Lyche-Solheim, A., Feld, C.K., Birk, S., Phillips, G., Carvalho, L., Morabito, G., *et al.* (2013). Ecological status assessment of european lakes: A comparison of metrics for phytoplankton, macrophytes, benthic invertebrates and fish. *Hydrobiologia*, 704, 57–74.
- Malaj, E., Ohe, P.C. von der, Grote, M., Kühne, R., Mondy, C.P., Usseglio-Polatera, P., *et al.* (2014). Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale. *Proceedings of the National Academy of Sciences*, 111, 9549–9554.
- McNellie, M.J., Oliver, I., Dorrough, J., Ferrier, S., Newell, G. & Gibbons, P. (2020). Reference state and benchmark concepts for better biodiversity conservation in contemporary ecosystems. *Global Change Biology*, 26, 6702–6714.
- Milligan, G.W. (1989). A Study of the Beta-Flexible Clustering Method. *Multivariate Behavioral Research*, 24, 163–176.
- Mondy, C.P. & Usseglio-Polatera, P. (2013). Using conditional tree forests and life history traits to assess specific risks of stream degradation under multiple pressure scenario. *Science of the Total Environment*, 461-462, 750–760.
- Mondy, C.P. & Usseglio-Polatera, P. (2014). Using fuzzy-coded traits to elucidate the non-random role of anthropogenic stress in the functional homogenisation of invertebrate assemblages. *Freshwater Biology*, 59, 584–600.
- Moog, O., Schwalb, A.N., Ofenböck, T. & Gerritsen, J. (2004). Does the ecoregion approach support the typological demands of the EU ‘Water Framework Directive’? *Hydrobiologia*, 516, 21–33.
- Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., *et al.* (2008). A general framework for the distance–decay of similarity in ecological communities. *Ecology*

- letters*, 11, 904–917.
- Mouquet, N. & Loreau, M. (2003). Community patterns in source-sink metacommunities. *The american naturalist*, 162, 544–557.
- Mouton, T.L., Tonkin, J.D., Stephenson, F., Verburg, P. & Floury, M. (2020). Increasing climate-driven taxonomic homogenization but functional differentiation among river macroinvertebrate assemblages. *Global Change Biology*, 0–3.
- Munné, A. & Prat, N. (2009). Use of macroinvertebrate-based multimetric indices for water quality evaluation in Spanish Mediterranean rivers: An intercalibration approach with the IBMWP index. *Hydrobiologia*, 628, 203–225.
- Mykrä, H., Aroviita, J., Hämäläinen, H., Karjalainen, S.M., Visuri, M., Riihimäki, J., *et al.* (2009). Utility of a single a priori river typology for reference conditions of boreal macroinvertebrates and diatoms. *Fundamental and Applied Limnology*, 175, 269–280.
- Mykrä, H., Heino, J. & Muotka, T. (2007). Scale-related patterns in the spatial and environmental components of stream macroinvertebrate assemblage variation. *Global Ecology and Biogeography*, 16, 149–159.
- Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology*, 15, 1044–1045.
- Nekola, J.C. & White, P.S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of biogeography*, 26, 867–878.
- Ouellet Dallaire, C., Lehner, B., Sayre, R. & Thieme, M. (2019). A multidisciplinary framework to derive global river reach classifications at high spatial resolution. *Environmental Research Letters*, 14, 024003.
- Outhwaite, C.L., Gregory, R.D., Chandler, R.E., Collen, B. & Isaac, N.J.B. (2020). Complex long-term biodiversity change among invertebrates, bryophytes and lichens. *Nature Ecology & Evolution*, 4, 384–392.
- Ovaskainen, O. & Abrego, N. (2020). *Joint species distribution modelling: With applications in R*. Cambridge University Press.
- Pardo, I., Gómez-Rodríguez, C., Wasson, J.G., Owen, R., Bund, W. van de, Kelly, M.G., *et al.* (2012). The European reference condition concept: A scientific and technical approach to identify minimally-impacted river ecosystems. *Science of the Total Environment*, 420, 33–42.
- Pearson, R.G. & Dawson, T.P. (2003). Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global ecology and biogeography*, 12, 361–371.

- Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10, 439–446.
- Peeters, E.T.M., Zuidam, J.P. van, Zuidam, B.G. van, Van Nes, E.H., Kosten, S., Heuts, P.G., *et al.* (2013). Changing weather conditions and floating plants in temperate drainage ditches. *Journal of Applied Ecology*, 50, 585–593.
- Podani, J. & Csányi, B. (2010). Detecting indicator species: Some extensions of the IndVal measure. *Ecological Indicators*, 10, 1119–1124.
- Poikane, S., Kelly, M.G., Salas Herrero, F., Pitt, J.A., Jarvie, H.P., Claussen, U., *et al.* (2019). Nutrient criteria for surface waters under the European Water Framework Directive: Current state-of-the-art, challenges and future outlook. *Science of the Total Environment*, 695, 133888.
- Poikane, S., Zampoukas, N., Borja, A., Davies, S.P., Bund, W. van de & Birk, S. (2014). Intercalibration of aquatic ecological assessment methods in the european union: Lessons learned and way forward. *Environmental Science & Policy*, 44, 237–246.
- Popovic, G.C., Warton, D.I., Thomson, F.J., Hui, F.K.C. & Moles, A.T. (2019). Untangling direct species associations from indirect mediator species effects with graphical models. *Methods in Ecology and Evolution*, 2019, 1571–1583.
- Posthuma, L., Zipp, M.C., De Zwart, D., Van de Meent, D., Globevnik, L., Koprivsek, M., *et al.* (2020). Chemical pollution imposes limitations to the ecological status of European surface waters. *Scientific Reports*, 1–20.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rees, C.B. van, Waylen, K.A., Schmidt-Kloiber, A., Thackeray, S.J., Kalinkat, G., Martens, K., *et al.* (2020). Safeguarding Freshwater Life Beyond 2020: Recommendations for the New Global Biodiversity Framework from the European Experience. *Conservation Letters*, preprints202001.0212.v1.
- Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T., *et al.* (2018). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*.
- Reyjol, Y., Argillier, C., Bonne, W., Borja, A., Buijse, A.D., Cardoso, A.C., *et al.* (2014). Assessing the ecological status in the context of the european water framework directive: Where do we go now? *Science of the Total Environment*, 497, 332–344.
- Roberts, D.W. (2019). *Labdsv: Ordination and multivariate analysis for ecology*.
- Roberts, D.W. (2020). *Optpart: Optimal partitioning of similarity relations*.

- Ryabov, A.B. & Blasius, B. (2011). A graphical theory of competition on spatial resource gradients. *Ecology Letters*, 14, 220–228.
- Schinegger, R., Trautwein, C., Melcher, A. & Schmutz, S. (2012). Multiple human pressures and their spatial patterns in European running waters. *Water and Environment Journal*, 26, 261–273.
- Schmera, D. & Baur, B. (2011). Testing a typology system of running waters for conservation planning in Hungary. *Hydrobiologia*, 665, 183–194.
- Schmera, D., Baur, B. & Erős, T. (2012). Does functional redundancy of communities provide insurance against human disturbances? An analysis using regional-scale stream invertebrate data. *Hydrobiologia*, 693, 183–194.
- Schreiner, V.C., Szöcs, E., Bhowmik, A.K., Vijver, M.G. & Schäfer, R.B. (2016). Pesticide mixtures in streams of several European countries and the USA. *Science of the Total Environment*, 573, 680–689.
- Scott Chamberlain & Eduard Szocs. (2013). Taxize - taxonomic search and retrieval in R. *F1000Research*.
- Smith, T.W. & Lundholm, J.T. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography*, 33, 648–655.
- Snelder, T.H., Cattaneo, F., Suren, A.M. & Biggs, B.J.F. (2004). Is the River Environment Classification an improved landscape-scale classification of rivers? *Journal of the North American Benthological Society*, 23, 580–598.
- Soininen, J. (2012). Macroecology of unicellular organisms—patterns and processes. *Environmental microbiology reports*, 4, 10–22.
- Statzner, B., Bis, B., Dolédec, S. & Usseglio-Polatera, P. (2001). Perspectives for biomonitoring at large spatial scales: A unified measure for the functional composition of invertebrate communities in European running waters. *Basic and Applied Ecology*, 2, 73–85.
- Szöcs, E., Brinke, M., Karaoglan, B. & Schäfer, R.B. (2017). Large Scale Risks from Agricultural Pesticides in Small Streams. *Environmental Science and Technology*, 51, 7378–7385.
- Telford, R.J., Vandvik, V. & Birks, H.J.B. (2006). Dispersal limitations matter for microbial morphospecies. *Science*, 312, 1015.
- Tennekes, M. (2018). tmap: Thematic maps in R. *Journal of Statistical Software*, 84, 1–39.
- Tickner, D., Opperman, J.J., Abell, R., Acreman, M., Arthington, A.H., Bunn, S.E., *et al.* (2020). Bending the Curve of Global Freshwater Biodiversity Loss: An Emergency Recovery Plan. *BioScience*, XX, 1–13.

- Tuomisto, H., Ruokolainen, L. & Ruokolainen, K. (2012). Modelling niche and neutral dynamics: On the ecological interpretation of variation partitioning results. *Ecography*, 35, 961–971.
- Turley, M.D., Bilotta, G.S., Extence, C.A. & Brazier, R.E. (2014). Evaluation of a fine sediment biomonitoring tool across a wide range of temperate rivers and streams. *Freshwater Biology*, 59, 2268–2277.
- van Klink, R., Bowler, D.E., Gongalsky, K.B., Swengel, A.B., Gentile, A. & Chase, J.M. (2020). Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science*, 368, 417–420.
- Van Sickle, J. (1997). Using Mean Similarity Dendrograms to Evaluate Classifications. *Journal of Agricultural, Biological and Environmental Statistics*, 2, 370–388.
- Van Sickle, J. & Hughes, R.M. (2000). Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. *Journal of the North American Benthological Society*, 19, 370–384.
- Vasconcelos, M.C., Melo, A.S. & Schwarzbald, A. (2013). Comparing the performance of different stream classification systems using aquatic macroinvertebrates. *Acta Limnologica Brasiliensia*, 25, 406–417.
- Verdonschot, P.F.M. (2006a). Data composition and taxonomic resolution in macroinvertebrate stream typology. *Hydrobiologia*, 566, 59–74.
- Verdonschot, P.F.M. (2006b). Evaluation of the use of Water Framework Directive typology descriptors, reference sites and spatial scale in macroinvertebrate stream typology. *Hydrobiologia*, 566, 39–58.
- Verdonschot, P.F.M. & Nijboer, R.C. (2004). Testing the European stream typology of the Water Framework Directive for macroinvertebrates. *Hydrobiologia*, 516, 35–54.
- Vilmi, A., Gibert, C., Escarguel, G., Happonen, K., Heino, J., Jamoneau, A., *et al.* (2020). Dispersal–niche continuum index: a new quantitative metric for assessing the relative importance of dispersal versus niche processes in community assembly. *Ecography*, 1–10.
- Vogt, J., Soille, P., De Jager, A., Rimavičiūtė, E., Mehl, W., Foisneau, S., *et al.* (2007). *A pan-European river and catchment database*.
- Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., *et al.* (2010). Global threats to human water security and river biodiversity. *Nature*, 467, 555.
- Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., *et*

- al.* (2010). Global threats to human water security and river biodiversity. *Nature*, 467, 555.
- Whittaker, R.H. & others. (1970). Communities and ecosystems. *Communities and ecosystems*.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., *et al.* (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4, 1686.
- Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., *et al.* (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88, 15–30.
- Wolfram, J., Stehle, S., Bub, S., Petschick, L.L. & Schulz, R. (2021). Water quality and ecological risks in European surface waters – Monitoring improves while water quality decreases. *Environment International*, 152, 106479.
- World Bank. (2021). <https://data.worldbank.org/indicator/ny.gdp.mktp.cd>.
- WWF. (2020). *Living planet report 2020-bending the curve of biodiversity loss*. World Wildlife Fund.
- Youn, S.-J., Taylor, W.W., Lynch, A.J., Cowx, I.G., Beard Jr, T.D., Bartley, D., *et al.* (2014). Inland capture fishery contributions to global food security and threats to their future. *Global Food Security*, 3, 142–148.