

# get real mzb paper

Jonathan Jupke

März 12, 2021

## Introduction

Freshwater ecosystems contribute to **copious** ecosystem services ranging from provision of clean drinking water and ~~animal proteins~~ to livelihoods, recreation and aesthetics (e.g. **Börger et al. 2021; Béné et al. 2016; Youn et al. 2014**). Combined, these contributions to people are estimated to have an annual worth of more than 4 trillion dollars (Costanza et al. 2014). Despite their crucial role, freshwater ecosystems are currently experiencing a biodiversity crisis (Darwall et al. 2018; WWF 2020).

The Living Planet Index indicates that populations of freshwater vertebrates declined by 84% relative to their 1970 levels on average (**note however that these numbers have been criticized by Leung et al. (2020)**). Based on the same data, He et al. (2019) found **an even** stronger decline (88%) in freshwater megafauna. Negative trends have also been recorded for invertebrates (e.g. **Cumberlidge et al. 2009**), but the emerging patterns ~~seem to be~~ complex (Floury et al. 2013; Baranov et al. 2020; Klink et al. 2020; Outhwaite et al. 2020). Dudgeon et al. (2006) summarized the major threats to freshwater biodiversity as overexploitation, water pollution, flow modification, habitat degradation and species invasions ~~as well as their interactions~~. New stressors such as microplastics and light pollution are emerging (Reid et al. 2018), many with yet to be discerned long-term effects. The situation is further complicated by the fact that most systems are subject to multiple, potentially interacting, stressors (Schinegger et al. 2012; Birk et al. 2020). A decreasing ecological status of freshwater ecosystems can reduce their capacity to provide ecosystem services (Culhane et al. 2019; Grizzetti et al. 2019). Thus, to prevent unacceptable harm to freshwater ecosystems and maintain a **necessary level** of ecosystem services, internationally coordinated actions and regulatory frameworks are necessary (**some ref.**).

**In** the European Union, the Water Framework Directive (WFD) fills this role. Its objective is to protect and improve the quality of freshwater ecosystems and to encourage sustainable use (European Commission 2000). It has led to improvements in **ecological assessments** (**birk 12**) and advanced our knowledge about the ecological status of freshwater systems throughout Europe (Kristensen et al. 2018). **However, it missed its target to achieve good water quality for all surface and groundwater bodies by 2015. As of 2018 only 40% of surface water bodies had a good status** (Kristensen et al. 2018). The status assessment is based on multiple criteria: biological quality elements (benthic invertebrates, phytoplankton, benthic flora and fish fauna), physico-chemical variables like nutrient and pollutant concentrations, and hydromorphology. The water body as a whole is judged by its single worst score. **This one-out-all-out principle hides progress in individual quality elements** (Kristensen et al. 2018) and has been criticized (Carvalho et al. 2019).

While the individual states use national methods to assess the status, the intercalibration process **worked on harmonizing the assessments. A river that would receive good status in one country should receive the same status if it would hypothetically be in another country.**

One recent output of the intercalibration process is a Pan-European typology of lentic and lotic freshwater systems (Lyche Solheim et al. 2019). While the WFD required the participating states to create stream typologies for the assessment **these typologies differ widely in the number and kind of streams they recognize. This made it harder to compare streams in different countries,** and intercalibration types had been established earlier (Birk et al. 2013; Poikane et al. 2014). Many national stream types (42%) were not associated with

any of the intercalibration types and some were linked to several (European Commission 2019). The recent effort by Lyche Solheim et al. (2019), combined the various national classifications and connected the new, broad river types with the already established intercalibration types. This new pan-European typology has already been used to compare nutrient threshold values across river types (Poikane et al. 2019), investigate interaction types of multiple stressors (Birk et al. 2020) and to show that chemical pollution currently limits the ecological status of European surface waters (Posthuma et al. 2020).

Until now it has not been evaluated, whether biological assemblages at least impacted **site** differ between the proposed **river** types **or** whether they are homogeneous within them. The latter part is especially important, because the WFD relies on the reference condition approach, which consists of evaluating a sites ecological status by comparing it with that of least impacted reference sites within the same **stream** type (Wright, Sutcliffe, and Furse (2000)). The spatial stability of reference states is crucial to the validity of **this** approach (Statzner et al. 2001). The typology of Lyche Solheim et al. (2019) largely follows the System A approach of the WFD in which ecoregion, altitude, geology and catchment size are used to classify river segments. Earlier studies (e.g. Verdonschot 2006b) **suggest** that other variables might be more strongly related to the distribution of different taxa and Borgwardt et al. (2019) showed that one of the proposed **categories** (RT1-Very Large rivers) could be further divided into seven different types with distinct assemblages of benthic invertebrates.

In this paper, we evaluate whether assemblages of freshwater macroinvertebrates at least impacted sites follow the classifications proposed by Lyche Solheim et al. (2019). **To this end we evaluate it's merits as a classification of sampling sites in comparison with other classification schemes based on several criteria.** Additionally we provide lists of very common (henceforth typical) taxa and indicator genera for the individual river types. The difference between the **latter two categories** is that good indicators are often rare species (Caceres and Legendre (2009)) **which are not well suited as expected taxa for reference states.** Thus while the latter might be indicative of a certain river type the former will be closer to what one might reasonably expect to find, when sampling ~~within a certain river type.~~

## Material and Methods

### Preparation of macroinvertebrate data

We **collected** a **large** data set of macroinvertebrate ~~samples~~ throughout rivers in Europe. The data included xxx samples (Figure xxx) from yyy years. We used different criteria to identify reference sites. The criterion depended on the data set, as some data sets provided detailed information on environmental conditions while other solely included biological information. If no environmental information where included we identified reference sites using the catchment landcover. Sampling sites from catchments in which the combined area of urban and agricultural land was below 10% of the catchment area where considered as reference sites. We used the Catchment Characterisation and Modelling database for geospatial data on rivers and catchments (Vogt et al. 2007) and the Corine Landcover data ~~from 2018~~ for landcover information (CLC 2018). We provide detailed information on variables and thresholds used to determine ~~the~~ reference sites in table S1. We removed non-reference sites because differences between stream types are more pronounced between reference sites (Verdonschot 2006a) and **because we wanted** to compared the typologies capability to create spatially stable reference assemblages. Sites were assigned the stream type of the nearest river segment or removed **of** this segment was further away than 500 m. This step further reduced the number of sites to xxx.

The data originated from different sources and required adjustments to ensure taxonomic consistency. First we used the gbif data **base** (**ref**) to check wether any of our species were considered synonyms and **changed** their names to accepted name suggeted in GBIF. We used different approaches for the different aims of this study. To compare the different typologies with one another and to derive indicator taxa we only used genus level information. Species level observations were transformed to their genera and observations from **lower taxonomic levels** were removed from the data set. To derive the typical assemblages consistency between river types was not as important since mostly information from the respective river type would be used. Hence we decided to use varying taxonomic levels instead of a fixed one. This way some taxa might be

represented at family level and others at species level, depending on the resolution of the data. *Oligochaetes*, for example, are usually only determined to subclass level. In a setting with one common taxonomic level (e.g. genus) they would have to be omitted if this level would be higher than subclass. By using taxon- and river type specific levels that take low-resolved taxa into account, we can thus use more of the data. However this poses the challenge of finding an optimal level for each taxon given a dataset and river type.

We established the optimal level with a hierarchical approach. First, we removed all observations from Phyla and Classes that were not present in all datasets. ~~We assumed that these represented differences in sampling rather than in communities.~~ The classes Clitellata (Annelida), Insecta, Malacostraca (Arthropoda), Bivalvia and Gastropoda (Mollusca) remained. In the following, a higher taxonomic level refers to levels with higher resolution, i.e. species is the highest taxonomic level and kingdom the lowest. For each taxon, we calculated the percentage of observations represented at each higher level. For example, 4.12% of observations from the order *Lepidoptera* are at the species level, 74.77% at the genus level, 7.75% at the family level, and 13.35% at the order level. Now given a threshold X, we hold a taxon to be optimally represented at a certain taxonomic level if less than X% are represented by higher levels. For example, *Lepidoptera* would be represented on order level if  $X > 4.12\% + 74.77\% + 7.75\% = 86.64\%$ . We used 50% as a threshold.

After removing data that did not meet our criteria, we did not have sufficient data to adequately represent all river types. We considered for a river type insufficient, if it contained to few sites ( $< 10$ ) or was spatially clustered. Maps of the sampling locations for the individual river types are provided in the Supplementary materials. The river types we were able to represent are: RT1, RT2, RT3, RT4, RT8, RT9, RT10, RT11, RT14, RT15, RT16 and RT18. Samples from the remaining river types were dropped.

## Comparison of typologies

To evaluate the two typologies proposed by Lyche Solheim et al. (2019) we compared it with the three other typologies: i) the k-means typology of the global river classification framework (GloRiC, Ouellet Dallaire et al. 2019); ii) the ecoregions proposed by Illies (1978); and iii) the biogeographical regions proposed by the European Environmental Agency (European Environmental Agency (EEA) 2016). As mentioned before only genus-level data were used for this analysis. We also ensured that every type in every category was represented by at least ten sites. Observations that failed to meet this criterion were removed until only adhered classes remained. See table Sx for the remaining types per typology and the respective number of samples.

In order to be able to judge the results we created two additional types. As an upper bound we created a classification of biological data using flexible beta clustering (ref) with the  $\beta$  parameter set to 0.625 and 13 groups. As a lower bound we created 100 random partitions of the data. For each partition we first drew the number of classes as a random variable from the interval between xxx and yyy. Then we assigned each observation randomly to one of the groups.

We calculated x cluster quality metrics for the typologies: the average silhouette width, the Calinski-Harabasz index, an indicator value score, the partition analysis ratio, Indicator Species Analysis Minimizing Intermediate Constancies and the within cluster sum-of-squares.

The average silhouette width [ASW; Kaufmann and Rousseeuw (1990)] is computed as  $ASW = \frac{1}{n} \sum_{i=1}^n \frac{a_i - b_i}{\max(a_i, b_i)}$ , where  $a_i$  is the average dissimilarity of point  $i$  to its own cluster and  $b_i$  is the average dissimilarity of point  $i$  to the closest other cluster.

Positive values indicate, that on average points are more similar to observations from their own cluster than to those of the most similar one.

The Calinski-Harabasz Index (CH, Caliński and Harabasz 1974) is computed as  $\frac{BGSS}{WGSS} \times \frac{n-k}{k-1}$  where BGSS is the between group sum-of-squares, WGSS is the within group sum of squares,  $k$  is the number of clusters and  $n$  the number of observations. High values indicate, that variation within clusters is smaller than between clusters. As the second term controls for the degrees of freedom, it can be understood as an analog to the F-Statistic. The algorithm assumes Euclidean data, but good performance with a similar metrics was shown for binary data in the context of fMRI-scans (Dimitriadou et al. 2004).

The indicator value score (IVS) is based on the indicator value (IndVal) proposed by Dufrêne and Legendre (1997). The IndVal itself will be explained in detail below, here we only note that we used 15000 permutations to compute  $p$ -values and controlled the family-wise error rate using Holm's step down procedure (Westfall and Young 1993) using a significance level of 0.05. We controlled the error rate within typologies and not between them, so the number of tests is equal to the number of taxa not the product of taxa and typologies as this would have increased the necessary number of permutations beyond feasibility. IVS is the fraction of taxa that are statistically significant indicators for a class. Higher scores indicate a better classification.

The Indicator Species Analysis Minimizing Intermediate Constancies (ISAMIC, Roberts 2020a) measures the constancy of a taxon within clusters. A taxon has a high constancy if it occurs in all or no samples from a cluster. It is minimized if it occurs in half of the samples. The ISAMIC is computed as:  $\frac{\sum_{i=1}^M 2 \times (\sum_{k=1}^K |c_{ik} - 0.5| / K)}{M}$ , where  $M$  is the number of taxa and  $c_{ik}$  the constancy of taxon  $i$  in cluster  $k$ . The statistics has been used to evaluate clusters before (e.g. Aho, Roberts, and Weaver 2008; Lötter, Mucina, and Witkowski 2013; Roberts 2015).

For ISAMIC and Indval, we assume especially high values for the biological classification, as the same data that are tested, were also used to derive the classification (Caceres and Legendre 2009)

The Partition analysis (PARTANA, Roberts 2019a) computes the ratio of within cluster similarity to between cluster similarity. Here similarities were derived from a Jaccard's distance matrix. The algorithm has shown good performance in Aho, Roberts, and Weaver (2008) and is conceptually similar to the classification strength which is the difference between mean within cluster similarities and mean between cluster similarities. The latter approach has been extensively used to evaluate stream classifications (e.g. Van Sickle 1997; Van Sickle and Hughes 2000; Gerritsen, Barbour, and King 2000). The within cluster sum-of-squares (wcsc)

is computed as  $wcsc = \frac{\sum_{k=1}^K \frac{\sum_{i=1}^{n_k} \sum_{j=1}^{n_k} d_{ij}^2}{n_k}}{2}$ , where  $d_{ij}$  is the dissimilarity of observations  $i$  and  $j$  and  $n_k$  is the number of observation in cluster  $k$ . If  $d$  is a Euclidean distance this is equal to the objective function of k-means clustering.

Based on these six cluster criteria, each typology was assigned a score. We will use these scores to evaluate the overall performances of typologies. The typology that performed best in some metric (i.e. minimizing within cluster sum-of-squares or maximizing all the other criteria) received 5 points, the second 4, the third 3, the fourth 2 and the fifth 1 point. Differences smaller than 5% of the range between biological and random partitions were regarded as ties. If two classes were tied, they both received the point for the position reduced by 0.5. For example, if two typologies are tied for the first place, both receive a score of 4.5. A three way tie, was settled by assigning all three classes the middle score. So if three classes are tied and lie at positions 2,3 and 4, each is assigned 3 points.

## Indicator and typical assemblages

We used the indicator value approach of Dufrêne and Legendre (1997) to identify indicator genera. For presence absences data such as we used, it can be understood as the product of two quantities A and B. For our purposes, A is the relative number of observations of taxon  $i$  that are within river type  $j$ . It was originally described as specificity (Dufrêne and Legendre 1997) but is better understood as concentration (Podani and Csányi 2010) because it is independent of the total number of clusters. B is the relative frequency with which species  $i$  occurs in stream type  $j$ . The maximum score is assigned to a species which only occurs in one stream type ( $A=1$ ) and occurs in all samples of that stream type ( $B=1$ ). Here, we used the group-equalized version of the IndVal that accounts for the different number of observations in each river type. The statistical significance of the IndVal Statistic is assessed with a permutation test which shuffles the river type that sites belong to. We used xxx permutations. As we conduct one test per genus ( $n =$  ) we needed to account for multiple testing and did so using Holm's step-down procedure (Westfall and Young 1993). This procedure ranks tests by their  $p$ -values in ascending order. The first  $p$ -value is divided by the number of tests (here the number of taxa,  $M$ ), the second by  $M-1$ , the third by  $M-2$  and so on until a  $p$ -value exceeds the significance level after the division. We used 0.05 as significance level.

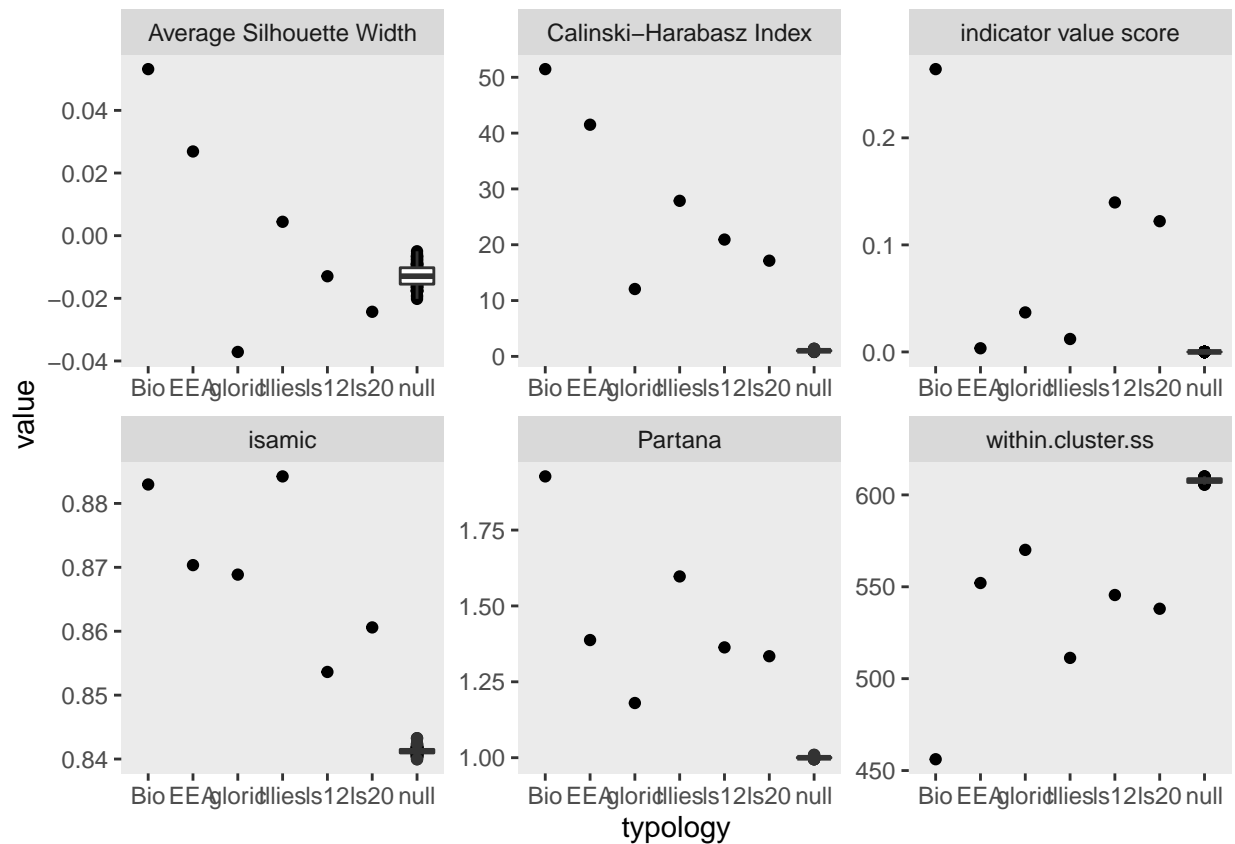
These indicator species provide valuable insight into the communities but miss the ubiquitous generalist species that occur in many different river types (tramp species sensu McGeoch, Van Rensburg, and Botes (2002)). Even if these taxa are common within the river type (high B) they will typically have low concentrations in most river types (low A) and hence low and statistically non-significant indicator values. Hence the indicator assemblages do not represent a typical sample, in these sense that these taxa can reasonably expected to occur in samples of the river type. We derived such typical assemblages by setting explicit thresholds for A and B. As mentioned earlier, the data used to derive typical assemblages have different taxonomic levels and we set different thresholds for different taxonomic levels. All species that occurred in more than 25% of samples of a river type (i.e.  $B > 0.25$ ) were considered typical. The value increased to 50% for genera and 75% for families or lower taxonomical levels. We also included species with very high concentrations ( $A = 0.9$ ) but only if they also occurred in more than 5% of samples, to prevent singeltons and rare species from exerting undue influence. Others have used alterations of the indicator value before, even if this was implicit, Wagner and Edwards (2001) for example considered only the A component. McGeoch, Van Rensburg, and Botes (2002) discusses the advantage of different combinations of A and B.

We used a harmonized trait data-base of European macroinvertebrates to assign traits values to the indicator and typical assemblages. The traits are fuzzy-coded (Chevene, Dolédec, and Chessel 1994), which allow single species to express tendencies towards different trait modalities. This can be appropriate for metaphoetic taxa or if traits are assigned to a genus to represent variation among species. We considered seven traits: feeding mode, respiration strategy, voltinism, locomotion, oviposition, body size and body form. If the traits for a taxon were not available in the data-base we used that of a lower taxonomic level, if available. Taxa that could not be assigned traits were removed from the analysis. We computed mean trait values for each trait and assemblage.

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 and Srinivasan 2021), tidyverse packages (Wickham et al. 2019) and taxize 0.9.98 (Scott Chamberlain and Eduard Szocs 2013; Chamberlain et al. 2020). Geospatial analyses were conducted using sf (Pebesma 2018). Cluster were created and evaluated with fpc (Hennig 2020), indicpecies (Caceres and Legendre 2009), labdsv (Roberts 2019b), optpart (Roberts 2020b).

viusalization are created with ggplot2 (Wickham 2016) and tmap (Tennekes 2018).

## Results



## More ideas

-> Dispersal ability? Sarremejane 2020

## Comparison of classifications

The ASW was highest in the biological classification (0.05) followed by EEA (0.03) and Illies (0). All stream-based typologies had negative ASWs (LS12: -0.01, LS20: -0.02; Gloric: -0.04), that were smaller than or equal to those of the random partitions ( $-0.01 \pm 0.003$ ). The CH was highest in the biological clustering (51.46). Both ecoregion-based approaches had higher CH (EEA: 41.5; Illies: 27.87) than the stream-based typologies (ls12: 20.93, ls20: 17.14, gloric: 12.07). All typologies had higher scores than the random partitions ( $1.01 \pm 0.09$ ).

The ecoregion-based approaches had the same IV-score (0.01) which was higher than that of the biological classification (0.26) which respectively was higher than that of the stream-based typologies (LS: 0.12; GloRic: 0.04). The score of LS was similar to that of the biological classifications and higher than that of GloRic. The random partitions all had IV-scores of 0. The ISAMIC was highest in the biological classification (0.883) followed by Illies (0.884) and Gloric (0.869). EEA and LS had the lowest scores (0.87 and 0.861, respectively). Relative to the biological and the random partitioning ( $0.841 \pm 0.001$ ), all four typologies had similar scores. The Partana score was highest for the biological classification (1.93) followed by Illies (1.6). EEA (1.39) and LS (1.33) had very similar scores. Gloric's scores was the lowest (1.18) but higher than that of the random partitions ( $1 \pm 0$ ). The within cluster sum-of-squares was lowest in the biological clustering (456.16) and

highest in the random partitions ( $607.85 \pm 1.16$ ). Of the typologies, Illies (511.3) had the lowest SS followed by LS (538.05), EEA (552.05), and GLORIC (570.09).

Overall Illies received the highest score, with 24 of 30 possible points, followed by EEA (22.5), LS12 (18), LS20 (14.5) and Gloic (11).

-> Maybe I should have a closer look at the results of the single LS clusters

## Indicator genera

Of the 610 total genera, 226 were selected as statistically significant indicators by our analysis. Most were indicative of one or two river types (202, 89%) but three were indicative of more than four types. These taxa were Rhyacophila (indicative of 7), Leuctra (6), and Baetis (5). Baetis was the most common genus occurring in 1230 of 1481 sites followed by Hydropsyche (1020), Gammarus (1016), Rhyacophila (930) and Leuctra (872). The number of indicative genera varied strongly between river types. RT4 had the most indicative genera (71) and RT14 and RT15 had the least with five and six respectively. The mean number of genera was 26 with a standard deviation of 18. Of the 71 genera that were indicative of RT4, 40 were unique to RT4. Dikergammarus in RT1 has the highest overall indicator value with 0.73.

Table xx shows all indicator genera and table yy

## Typical assemblages

Based on the criteria outline in the material and methods we combined the river types RT4 and 5 as well as RT10 and 11. The former type now represents lowland, calcareous streams irrespective of size and the latter represents Mid-altitude calcareous streams irrespective of size.

82 of 1432 taxa were included in the typical assemblages. Most were species (53), followed by genus (22) and family or lower level (7). On average, taxa occurred in 2.4 typical assemblages. Chironomidae was the only taxon that was typical for all river types. The most common genus, Baetis, occurred in eight river types. Three species tied for most common: *Elmis aenea*, *Baetis rhodani* and *Ancylus fluviatilis*. Each occurred in five typical assemblages. The average size of typical assemblages was 18 with a standard deviation of 7.33. The three largest typical assemblages are associated with the lowland river types RT4\_5 (30), RT2 (29) and RT3 (26). RT9 has the smallest typical assemblage with nine taxa. The mean taxonomic resolution was 2.06 with a standard deviation of 0.36. The three lowland river types also had the highest taxonomic resolutions (2.48, 2.47 and 2.40). RT9 had an above average taxonomic resolution (2.28) but the assemblages with the lowest resolution (RT14 and 15, 1.63) also had small typical assemblages (11 and 12). Most taxa in the typical assemblages were common (i.e.  $B > 0.25$ ). Only 14 taxa (7.1%) were added because of their high concentration. All of these 14 were species and their were associated with either RT4\_5 (5) or RT1 (9). The only taxon that would have been included in a typical assemblage due to both criteria (i.e.  $A > 0.9$  and  $B > 0.25$ ) was *Echinogammarus ischnus* which belongs to the typical assemblage of RT1.

Overall, the typical assemblages are smaller than the indicator assemblages. The size of a river types typical assemblage showed a moderate correlation (0.47) to the size of its indicator assemblage.

-> what are the strongest indicators, discuss them for each.

## Traits of Indicator assemblages

The figure X shows the community weighted mean traits of the indicator assemblages. As traits were fuzzy-coded they express the tendency of the hypothetical average indicator species towards expressing a certain trait modality. For improve legibility, we will henceforth refer to the indicator and typical assemblages with the name of their associated river type. Hence, in the following RT01 does not refer to a specific river type but rather the assemblages associated with it.



Shredding was most pronounced in assemblages of high altitude river types (RT14, 15, 16). In these assemblages, filterers and predators were rare or absent. In the lowland assemblages, shredding was more common in smaller rivers. Parasitic feeding strategies were comparatively rare and only occurred in RT1 and 4. In RT1, two genera expressed this modality: *Unio* and *Sisyrh*. Both are only parasitic in their larval stages (QUELLEN). RT4 included taxa that are parasitic in some life stages (*Unio*, *Unionicola* and *Anodonta*) as well as obligate parasites (*Hemiclepsis* and *Theromyzon*). The predation feeding strategy was most pronounced in RT5 (0.49) and high in assemblages of small lowland rivers. In assemblages of larger and high altitude rivers it was less pronounced and absent in RT14. The prevalence of herbivory varied between assemblages with a minimum of 0.12 in RT3 to a maximum of 0.46 in RT11. The modality was pronounced in assemblages of mid- and high-altitude river types ( $> 0.3$ ) with the exception of RT8 and RT9. The prevalence of gathering varied between 0.1 and 0.2 (exceptions were RT5 with 0.05 and RT10 with 0.26) and tended to be lower in assemblages of smaller rivers (cf. 2,3 and 4,5). Filtering was most pronounced in RT1 where it was the dominant feeding strategy. In other assemblages it played a minor role. Most affinities lied below 0.1 and the feeding strategy was absent from RT15 and 16.

The breathing strategy was divided into three categories: gills, spiracles and tegument. In most assemblages (10 out of 13) teguments were the dominant breathing strategy. Respiration through gills was also common and the prevalent strategy in the three remaining assemblages (RT1, 8 and 11). Breathing with spiracles was absent from four assemblages (RT3, 9, 14 and 16) and very rare in two (RT1 and 2). This modality was most common in RT5 where it was more prevalent than breathing with gills. In all assemblages, except RT1, univoltinism was more common than semi- or multivoltinism. In RT1, multivoltinism was most common (0.52). In assemblages of lowland rivers, multivoltinism was more pronounced than semivoltinism, but the latter became more common in assemblages of mid- and high-altitude rivers. Semivoltinism was more common than multivoltinism in RT16 and the latter was absent from RT14. Crawling was the most common type of locomotion in all assemblages. RT1 had the lowest affinity to the modality (0.41). This was not due to very high scores in one other locomotion strategy but rather to high values in all others (approx. 0.2 each). Swimming was pronounced in assemblages of lowland calcareous stream (RT4,5) but noticeably less in assemblages of lowland siliceous streams (RT2,3). It was least pronounced in RT14. Sessility was the least common strategy in most assemblages (10 out of 13) but never absent. The scores tended to be lower in assemblages of mid- and high-altitude streams but were low in RT2 (0.05) and high in RT8 and 9 (0.11; 0.16). Burrowing was more common in assemblages of large streams than in those of small streams with the exception of RT11. In all indicator assemblages except RT1, aquatic oviposition was most pronounced. It was the only modality in RT14 and 15. Ovivivipary was the most common type in RT1 and moderately common (approx 0.2) in assemblages of other lowland streams with the exception of RT5. It was absent in RT 8, 9, 14, 15 and 18. Terrestrial oviposition was uncommon in most assemblages and absent in RT 3, 10, 14, 15 and 16. It was most common in assemblages of mid-altitude siliceous streams (RT8: 0.12, RT9: 0.08). Larger body size was more common in assemblages of lowland rivers than in those of mid- or high-altitude rivers. It was the most common modality in RT1 and 3 but absent from RT14. Medium size was pronounced to varying degrees. It was rare in RT1 (0.08) and the most common modality in RT5, 9 and 16. Small size was common in all assemblages (always  $> 0.3$ ), least pronounced in RT1 (0.33) and most pronounced in RT 14 (0.69). In all assemblages except RT1, the most pronounced modality for body form was cylindrical. In RT1, flattened was more pronounced (0.51) than cylindrical (0.35), while both other modalities (spherical and streamlined) had scores below 0.1. Spherical was the least common modality. Scores above 0.1 occurred only in RT3 (0.17) and RT4 (0.14). The modality was absent from RT9, 14, and 16. Besides RT1, flattened body forms were common in RT16 (0.29) and RT5 (0.2) and rare in RT3, 8 and 15. Streamlined bodies were common in RT5, 14 and rare in RT1, 2, 3, 11 and 18.

The community-weighted mean traits of typical assemblages differed noticeably from those of the indicator assemblages (LINK TO FIGURE). As in the indicator assemblages, shredding was more pronounced in assemblages of smaller rivers and higher altitudes. Its highest values were in RT9 and RT15 (0.24 both) but it was not dominant in either. The lowest values were in RT1 (0.13), RT4\_5 (0.11) and RT14 (0.12). The range of affinities towards this modality was small. The same is true of predation. It was highest in RT14 (0.31), where it was the dominant feeding strategy, and lowest in RT1 (0.13), RT8 (0.14) and RT10\_11 (0.14). No general patterns were observed. Parasitism was the rarest feeding strategy. It was absent in seven of



eleven assemblages and very rare in the other four (0.01 in RT8, 9, 14 and 15). It only occurred in assemblages of mid- or high-altitude river types. The dominant feeding strategy in most assemblages was herbivory. It was most pronounced in assemblages of the mid-altitude river types RT8 (0.48), 9 (0.4) and 10\_11 (0.4). It was rarest in the assemblages of the two high-altitude river types RT14 (0.12) and 15 (0.17). Gathering was present in all assemblages. It was least prevalent in the typical assemblages of the two siliceous river types, RT3 (0.12) and RT8 (0.14) and most common RT14 (0.29) and RT18 (0.3). Filtering is still the most pronounced modality in RT1 (0.31) and least common in RT9 (0.01).

In all typical assemblages, breathing with gills or teguments were the common modes of breathing. Community mean weighted traits ranged between 0.3 and 0.6, with one high (0.73 for gills in RT1) and low (0.27 for teguments in RT1) exception. Breathing through spiracles was the least common modality in all typical assemblages and absent in RT1 and RT14. It was most common in RT10\_11 (0.21).

RT1 was the only river type in which univoltinism was not the dominant reproduction strategy. Instead, multivoltinism was most common (0.56). In all assemblages univoltinism ranged between 0.55 (RT14) and 0.73 (16). Multivoltinism was least common in RT16 and relatively common ( $>0.3$ ) in RT4\_5, RT14, RT15 and RT18. Semivoltinism is the least common modality in all typical assemblages. The maximal affinity toward this modality was found in RT16 (0.12). Crawling was the dominant locomotion modality in all typical assemblages. Affinities ranged between 0.4 (RT1) and 0.67 (RT16). Swimming was most pronounced in RT1 (0.27) and ranged between 0.1 and 0.18 in all other typical assemblages. Sessile organisms were least common in RT9 (0.05) and most common in RT14 (0.22). For most assemblages the affinities ranged between 0.1 and 0.2. The affinity towards burrowing was also highest in RT14 (0.23) but lowest in RT15 (0.11). Terrestrial oviposition was missing from the typical assemblages. Aquatic oviposition was the dominant modality in all typical assemblages and ranged between 0.6 (RT1) and 0.98 (RT16). Ovovivipary was most common in RT1 (0.4), RT4\_5 (0.23) and RT15 (0.18). In most assemblages, small or medium-sized organisms made up most of the typical assemblages. Together, these two modalities exceeded 0.8 in all assemblages except RT1. In most, small size was more common than medium size but the opposite was true in RT15 and they were equally common in RT3. In RT1, large organisms dominated the typical assemblage (0.52) and small organisms were rare (0.15).

Most taxa in the typical assemblages had a cylindrical body shape. The affinity towards this modality exceeded 0.5 in all typical assemblages, except RT1 (0.38), RT10\_11 (0.47) and RT 18 (0.46) and was most pronounced in RT14 (0.86). It was the dominant modality in all typical assemblages, except RT1, which had a larger affinity towards flattened body shapes (0.44). Flattened had the second highest affinity in all typical assemblages, except RT14. Streamlined bodies are more common than spherical ones in assemblages. The latter is rare in most types and absent from RT14 and RT15.

## Discussion

We compared the partitioning of aquatic invertebrate samples based four different freshwater typologies. Two were stream-based (xx and yy) and two were ecoregion-based (xx and yy). To better understand the metrics we used for the comparison, we also computed a biological clustering of the samples and 100 random partitions of the data. We expected the former to always outperform and the latter to always be outperformed by the environment-based typologies. These expectations were met, with the exceptions of stream-based typologies performing worse than random in terms of ASW and ecoregion-based approaches better than biological clusters in terms of IV-scores.

Overall Illies typology performed the best in our assessment. It had, or was tied for, the best scores in four out of six criteria. Second best was EEA 3 out of

Discuss TA and Indval species for each river type. Potamon species in RT2 and 4 Ind: Acrolorus, Anisus, Viviparus Potamon species in RT1 Ind : Theodoxus

alpine species : Rhabdiopteryx indicator for RT15 is adjusted to high altitude rivers (according to Lorenz et al 2004 ) also Rhithrogena for RT14/ 15 (through Rhithrogena alpestris), Rhyacophila (through Rhyacophila torrentium and vulgaris) for all mid- and high altitude types except RT14

Which species occurred everywhere? Tramp species. McGeoch.

Summary TA similarity Summary TA Traits

Summary Indval similarity Summary Indval Traits

Other papers that compare or evaluate clusters. Other papers that computed indvals

Flip in RT14 from no predation to dominant feeding strategy.

Relative ranking of typologies differs through the metrics.

Based on the data set we compiled and the stream/ ecoregion types we evaluated, Illies performed best, and LS was the better stream-based typology.

Why are landscape level typologies more effective here? Spatial signal, dispersal limitations, historical distribution. -> Verdonschot and Nijboer (2004) also endorse illies1978limnofauna and the effect of large scale variables in general. However, using a expended data set, that included both AQUEM and START data, Verdonschot (2006b) did not ecoegions to be a strong predictor of invertebrate community composition. Similarly, ecoregions explained a negligible amount of variance in community composition in Swedish boreal streams (Johnson, Goedkoop, and Sandin 2004) Hawkins et al. (2000) also dont like Lorenz, Feld, and Hering (2004) like Johnson et al. (2007) large scale factors most important single but when grouped local more important

More studies on local vs regional: pro regional Urban et al 2006, DD Corkum, 1989; Richards, Johnson & Host, 1996; Roth, Allan & Erickson, 1996; Lammert & Allan, 1999(from Johnson 2007)

pro local Johnson et al 2007

Streams too large? See small waterbodies are important paper by ... and add the new stream system by ...

Tendency towards more shredders in high smaller rivers. Matches the predictions of the river continuum concept (Vannote et al. 1980). The latter has been critized for xxx but seems not to be a large problem.

In both kind of assemblages we derived, very large rivers differed most markedly from the other river types. Trends that always hold -> filtering in RT1; voltinism in RT1 ; crawling dominant ; size RT1

Longer generation times were clearly associated with high altitude streams (more semi and less mutlivoltinism) in the indicator assemblages but not in the typical assemblages. This runs counter to results of Finn and Poff (2005) Freshwater Biology

Moog et al 2004: "Finer spatial resolution of the smaller bioregions also required finer taxonomic resolution" Verdonschot (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 not large and their analysis does not allow to infer the loss of using genus instead of species level data. Hawkins et al. (2000) notes that across taxa groups and typologies, higher taxonomic resolution does not always lead to greater classification strength.

While we should expect the the performance of stream typologies to improve with increasing taxonomic resolution, we would not expect this improvement to systematically vary between typologies.

Intraspecific trait variability - (<https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13568>)

The typology of the WFD must be optimized not only for macroinvertebrates but for all four biological elements (macroinvertebrates, diatoms, fish, macrophytes) (Dodkins et al. 2005). Thus these taxa should also be assessed to draw a final conclusion.

A stream typology based on environmental variables cannot account for the effects of biotic interactions. While these received much attention in in the early days of macroecology (ref to Diamond papers) the focus shifted towards environmental filters (see e.g Stendra et al 12 or Johnson et al 2004 for references). Recently however, their role for macroecology as been revisited (see Elo et al 21 and references therein).

Mass effects could dilute effect of stream based typologies - recent studies dont suggest that its important (some paper references in MCE book). Two recently proposed methods test the effects of dispersal limitation (Vilmi et al. 2020) and mass effect (Leboucher et al. 2020) on metacommunities and might be interesting to persue in this context. Both have yet to be comprehensively evaluated.

Dispersal limitation can lead to some species missing from sites that are appropriate for them. As network connectivity strongly impact dispersal in river networks, regional influences are especially important (Tonkin et al 2018 from Cid et al. 2020)

Furthermore it remains unclear whether current ecoregions or stream classes will change uniformly with future climate change or will con- and diverge. Environmental legislative frameworks must acknowledge that climate change will alter their subjects and adapt when necessary (Arneth et al. 2020).

## References

- Aho, Ken, David W. Roberts, and T. Weaver. 2008. "Using geometric and non-geometric internal evaluators to compare eight vegetation classification methods." *Journal of Vegetation Science* 19 (4): 549–62. <https://doi.org/10.3170/2008-8-18406>.
- Arneth, Almut, Yunne-jai Shin, Paul Leadley, Carlo Rondinini, and Elena Bukvareva. 2020. "Post-2020 biodiversity targets need to embrace climate change." *PNAS* 0: 1–10. <https://doi.org/10.1073/pnas.2009584117>.
- Baranov, Viktor, Jonas Jourdan, Francesca Pilotto, Rüdiger Wagner, and Peter Haase. 2020. "Complex and nonlinear climate-driven changes in freshwater insect communities over 42 years." *Conservation Biology*, cobi.13477. <https://doi.org/10.1111/cobi.13477>.
- Béné, Christophe, Robert Arthur, Hannah Norbury, Edward H Allison, Malcolm Beveridge, Simon Bush, Liam Campling, et al. 2016. "Contribution of Fisheries and Aquaculture to Food Security and Poverty Reduction: Assessing the Current Evidence." *World Development* 79: 177–96.
- Birk, Sebastian, Daniel Chapman, Laurence Carvalho, Bryan M. Spears, Hans Estrup Andersen, Christine Argillier, Stefan Auer, et al. 2020. "Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems." *Nature Ecology & Evolution*, no. 1. <https://doi.org/10.1038/s41559-020-1216-4>.
- Birk, Sebastian, N. J. Willby, Martyn G. Kelly, W. Bonne, Angel Borja, Sandra Poikane, and Wouter Van De Bund. 2013. "Intercalibrating classifications of ecological status: Europe's quest for common management objectives for aquatic ecosystems." *Science of the Total Environment* 454-455: 490–99. <https://doi.org/10.1016/j.scitotenv.2013.03.037>.
- Borgwardt, Florian, Patrick Leitner, Wolfram Graf, and Sebastian Birk. 2019. "Ex uno plures – defining different types of very large rivers in Europe to foster solid aquatic bio-assessment." *Ecological Indicators* 107 (August): 105599. <https://doi.org/10.1016/j.ecolind.2019.105599>.
- Börger, Tobias, Danny Campbell, Mathew P White, Lewis R Elliott, Lora E Fleming, Joanne K Garrett, Caroline Hattam, Stephen Hynes, Tuija Lankia, and Tim Taylor. 2021. "The Value of Blue-Space Recreation and Perceived Water Quality Across Europe: A Contingent Behaviour Study." *Science of The Total Environment*, 145597.
- Caceres, Miquel De, and Pierre Legendre. 2009. "Associations between species and groups of sites: inindices and statistical inference." *Ecology* 90 (12): 3566–74. <https://doi.org/10.1890/08-1823.1>.
- Caliński, Tadeusz, and Jerzy Harabasz. 1974. "A Dendrite Method for Cluster Analysis." *Communications in Statistics-Theory and Methods* 3 (1): 1–27.
- Carvalho, Laurence, Eleanor B. Mackay, Ana Cristina Cardoso, Annette Baattrup-Pedersen, Sebastian Birk, Kirsty L. Blackstock, Gábor Borics, 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: 1228–38. <https://doi.org/10.1016/j.scitotenv.2018.12.255>.

- Chamberlain, Scott, Eduard Szoecs, Zachary Foster, Zebulun Arendsee, Carl Boettiger, Karthik Ram, Ignasi Bartomeus, et al. 2020. *Taxize: Taxonomic Information from Around the Web*. <https://github.com/ropensci/taxize>.
- Chevene, François, Sylvain Dolédec, and Daniel Chessel. 1994. “A Fuzzy Coding Approach for the Analysis of Long-Term Ecological Data.” *Freshwater Biology* 31 (3): 295–309.
- Chytry, Milan, Lubomir Tichy, Jason Holt, and Zoltan Botta-Dukat. 2002. “Determination of Diagnostic Species with Statistical Fidelity Measures.” *Journal of Vegetation Science* 13 (1): 79–90.
- CLC. 2018. “CORINE Land Cover.” <https://Land.copernicus.eu/Pan-European/Corine-Land-Cover>.
- Costanza, Robert, Rudolf De Groot, Paul Sutton, Sander Van der Ploeg, Sharolyn J Anderson, Ida Kubiszewski, Stephen Farber, and R Kerry Turner. 2014. “Changes in the Global Value of Ecosystem Services.” *Global Environmental Change* 26: 152–58.
- Culhane, Fiona, Heliana Teixeira, Antonio J. A. Nogueira, Florian Borgwardt, Daniel Trauner, Ana Lillebø, Ger Jan Piet, et al. 2019. “Risk to the supply of ecosystem services across aquatic ecosystems.” *Science of the Total Environment* 660: 611–21. <https://doi.org/10.1016/j.scitotenv.2018.12.346>.
- Cumberlidge, Neil, Peter KL Ng, Darren CJ Yeo, Celio Magalhães, Martha R Campos, Fernando Alvarez, Tohru Naruse, et al. 2009. “Freshwater Crabs and the Biodiversity Crisis: Importance, Threats, Status, and Conservation Challenges.” *Biological Conservation* 142 (8): 1665–73.
- Darwall, William, Vanessa Bremerich, Aaike De Wever, Anthony I. Dell, Jörg Freyhof, Mark O. Gessner, Hans Peter Grossart, 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 (4): 1015–22. <https://doi.org/10.1002/aqc.2958>.
- Dimitriadou, Evgenia, Markus Barth, Christian Windischberger, Kurt Hornik, and Ewald Moser. 2004. “A quantitative comparison of functional MRI cluster analysis.” *Artificial Intelligence in Medicine* 31 (1): 57–71. <https://doi.org/10.1016/j.artmed.2004.01.010>.
- Dodkins, Ian, Brian Rippey, Thomas J. Harrington, Catherine Bradley, Bernadette Ni Chathain, Mary Kelly-Quinn, Martin McGarrigle, Stephanie Hodge, and David Trigg. 2005. “Developing an optimal river typology for biological elements within the Water Framework Directive.” *Water Research* 39 (15): 3479–86. <https://doi.org/10.1016/j.watres.2005.06.008>.
- Dowle, Matt, and Arun Srinivasan. 2021. *Data.table: Extension of ‘Data.frame’*. <https://CRAN.R-project.org/package=data.table>.
- Dudgeon, David, Angela H. Arthington, Mark O. Gessner, Zen Ichiro Kawabata, Duncan J. Knowler, Christian Lévêque, Robert J. Naiman, et al. 2006. “Freshwater biodiversity: Importance, threats, status and conservation challenges.” *Biological Reviews* 81 (2): 163–82. <https://doi.org/10.1017/S1464793105006950>.
- Dufrêne, Marc, and Pierre Legendre. 1997. “Species Assemblages and Indicator Species: The Need for a Flexible Asymmetrical Approach.” *Ecological Monographs* 67 (3): 345–66.
- 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 (1).
- . 2019. “European Overview – River Basin Management Plans. Re- Port from the Commission to the European Parliament and the Council on the Imple- Mentation of the Water Framework Directive (2000/60/EC) and the Floods Directive (2007/60/EC): Second River Basin Management Plans and First Flood Risk Manage- Ment Plans.” *SWD* 296.
- European Environmental Agency (EEA). 2016. “Biogeographical Regions.” <https://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3>.
- Floury, Mathieu, Philippe Usseglio-Polatera, Martial Ferreol, Cecile Delattre, and Yves Souchon. 2013. “Global climate change in large European rivers: Long-term effects on macroinvertebrate communities

- and potential local confounding factors.” *Global Change Biology* 19 (4): 1085–99. <https://doi.org/10.1111/gcb.12124>.
- Gerritsen, Jeroen, Michael T. Barbour, and Kurt King. 2000. “Apples, oranges, and ecoregions: On determining pattern in aquatic assemblages.” *Journal of the North American Benthological Society* 19 (3): 487–96. <https://doi.org/10.2307/1468109>.
- Grizzetti, B., C. Liqueste, A. Pistocchi, O. Vigiak, G. Zulian, F. Bouraoui, A. De Roo, and A. C. Cardoso. 2019. “Relationship between ecological condition and ecosystem services in European rivers , lakes and coastal waters.” *Science of the Total Environment* 671: 452–65. <https://doi.org/10.1016/j.scitotenv.2019.03.155>.
- Hawkins, Charles P., Richard H Norris, Jeroen Gerritsen, Robert M. Hughes, Susan K. Jackson, Richard K. Johnson, and R. Jan Stevenson. 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 (3): 541–56.
- He, Fengzhi, Christiane Zarfl, Vanessa Bremerich, Jonathan N. W. David, Zeb Hogan, Gregor Kalinkat, Klement Tockner, and Sonja C. Jähnig. 2019. “The global decline of freshwater megafauna.” *Global Change Biology* 25 (11): 3883–92. <https://doi.org/10.1111/gcb.14753>.
- Hennig, Christian. 2020. *Fpc: Flexible Procedures for Clustering*. <https://CRAN.R-project.org/package=fpc>.
- Illies, Joachim. 1978. *Limnofauna Europaea*. Fischer Stuttgart.
- Johnson, Richard K., Mike T. Furse, Daniel Hering, and Leonard Sandin. 2007. “Ecological relationships between stream communities and spatial scale: Implications for designing catchment-level monitoring programmes.” *Freshwater Biology* 52 (5): 939–58. <https://doi.org/10.1111/j.1365-2427.2006.01692.x>.
- Johnson, Richard K., Willem Goedkoop, and Leonard Sandin. 2004. “Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes.” *Freshwater Biology* 49 (9): 1179–94. <https://doi.org/10.1111/j.1365-2427.2004.01262.x>.
- Kaufmann, L., and P. Rousseeuw. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley&Sons.
- Klink, Roel van, Diana E. Bowler, K. B. Gongalsky, A. B. Swengel, A. Gentile, and Jonathan M. Chase. 2020. “Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances.” *Science* 420 (April): in press.
- Kristensen, Peter, Caroline Whalley, Fernanda Néry Nihat Zal, T Christiansen, and others. 2018. “European Waters Assessment of Status and Pressures 2018.” *EEA Report*, no. 7/2018.
- Leboucher, Thibault, Juliette Tison-Rosebery, William R. Budnick, Aurélien Jamoneau, Wim Vyverman, Janne Soininen, Sébastien Boutry, and Sophia I. Passy. 2020. “A metacommunity approach for detecting species influenced by mass effect.” *Journal of Applied Ecology* 57: 2031–40. <https://doi.org/10.1111/1365-2664.13701>.
- Leung, Brian, Anna L. Hargreaves, Dan A. Greenberg, Brian McGill, Maria Dornelas, and Robin Freeman. 2020. “Clustered versus catastrophic global vertebrate declines.” *Nature*, no. January. <https://doi.org/10.1038/s41586-020-2920-6>.
- Lorenz, Armin, Christian K. Feld, and Daniel Hering. 2004. “Typology of streams in Germany based on benthic invertebrates: Ecoregions, zonation, geology and substrate.” *Limnologica* 34 (4): 379–89. [https://doi.org/10.1016/S0075-9511\(04\)80007-0](https://doi.org/10.1016/S0075-9511(04)80007-0).
- Lötter, M. C., L. Mucina, and E. T. F. Witkowski. 2013. “The classification conundrum: Species fidelity as leading criterion in search of a rigorous method to classify a complex forest data set.” *Community Ecology* 14 (1): 121–32. <https://doi.org/10.1556/ComEc.14.2013.1.13>.
- Lyche Solheim, A., K. Austnes, Lidija Globevnik, Peter Kristensen, Jannicke Moe, J Persson, Sandra Poikane, Wouter van de Bund, and Sebastian Birk. 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 (September): 134043. <https://doi.org/10.1016/j.scitotenv.2019.134043>.
- McGeoch, Melodie A., Berndt J. Van Rensburg, and Antoinette Botes. 2002. “The verification and application of bioindicators: A case study of dung beetles in a savanna ecosystem.” *Journal of Applied Ecology* 39 (4): 661–72. <https://doi.org/10.1046/j.1365-2664.2002.00743.x>.
- Ouellet Dallaire, Camille, Bernhard Lehner, Roger Sayre, and Michele Thieme. 2019. “A multidisciplinary framework to derive global river reach classifications at high spatial resolution.” *Environmental Research Letters* 14 (2): 024003. <https://doi.org/10.1088/1748-9326/aad8e9>.
- Outhwaite, Charlotte L., Richard D. Gregory, Richard E. Chandler, Ben Collen, and Nick J. B. Isaac. 2020. “Complex long-term biodiversity change among invertebrates, bryophytes and lichens.” *Nature Ecology & Evolution* 4 (3): 384–92. <https://doi.org/10.1038/s41559-020-1111-z>.
- Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.
- Podani, János, and Béla Csányi. 2010. “Detecting indicator species: Some extensions of the IndVal measure.” *Ecological Indicators* 10 (6): 1119–24. <https://doi.org/10.1016/j.ecolind.2010.03.010>.
- Poikane, Sandra, Martyn G. Kelly, Fuensanta Salas Herrero, Jo Anne Pitt, Helen P. Jarvie, Ulrich Claussen, Wera Leujak, Anne Lyche Solheim, Heliana Teixeira, and Geoff Phillips. 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. <https://doi.org/10.1016/j.scitotenv.2019.133888>.
- Poikane, Sandra, Nikolaos Zampoukas, Angel Borja, Susan P Davies, Wouter van de Bund, and Sebastian Birk. 2014. “Intercalibration of Aquatic Ecological Assessment Methods in the European Union: Lessons Learned and Way Forward.” *Environmental Science & Policy* 44: 237–46.
- Posthuma, Leo, Michiel C. Zijp, Dick De Zwart, Dik Van de Meent, Lidija Globevnik, Maja Koprivsek, Andreas Focks, Jos Van Gils, and Sebastian Birk. 2020. “Chemical pollution imposes limitations to the ecological status of European surface waters.” *Scientific Reports*, no. September: 1–20. <https://doi.org/10.1038/s41598-020-71537-2>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reid, Andrea J., Andrew K. Carlson, Irena F. Creed, Erika J. Eliason, Peter A. Gell, P T Johnson, Karen A. Kidd, et al. 2018. “Emerging threats and persistent conservation challenges for freshwater biodiversity.” *Biological Reviews*. <https://doi.org/10.1111/brev.12480>.
- Resh, Vincent H, and David M Rosenberg. 1993. *Freshwater Biomonitoring and Benthic Macroinvertebrates*. 504.4 FRE. Chapman & Hall New York, NY, USA:
- Roberts, David W. 2015. “Vegetation classification by two new iterative reallocation optimization algorithms.” *Plant Ecology* 216: 741–58. <https://doi.org/10.1007/s11258-014-0403-2>.
- . 2019a. *Labdsv: Ordination and Multivariate Analysis for Ecology*. <https://CRAN.R-project.org/package=labdsv>.
- . 2019b. *Labdsv: Ordination and Multivariate Analysis for Ecology*. <https://CRAN.R-project.org/package=labdsv>.
- . 2020a. *Optpart: Optimal Partitioning of Similarity Relations*. <https://CRAN.R-project.org/package=optpart>.
- . 2020b. *Optpart: Optimal Partitioning of Similarity Relations*. <https://CRAN.R-project.org/package=optpart>.
- Schinegger, Rafaela, Clemens Trautwein, Andreas Melcher, and Stefan Schmutz. 2012. “Multiple human pressures and their spatial patterns in European running waters.” *Water and Environment Journal* 26 (2): 261–73. <https://doi.org/10.1111/j.1747-6593.2011.00285.x>.

- Scott Chamberlain, and Eduard Szocs. 2013. "Taxize - Taxonomic Search and Retrieval in R." *F1000Research*. <https://f1000research.com/articles/2-191/v2>.
- Statzner, Bernhard, Barbara Bis, Sylvain Dolédec, and Philippe Usseglio-Polatera. 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 (1): 73–85. <https://doi.org/10.1078/1439-1791-00039>.
- Tennekes, Martijn. 2018. "tmap: Thematic Maps in R." *Journal of Statistical Software* 84 (6): 1–39. <https://doi.org/10.18637/jss.v084.i06>.
- Tichy, Lubomir, and Milan Chytrý. 2006. "Statistical Determination of Diagnostic Species for Site Groups of Unequal Size." *Journal of Vegetation Science* 17 (6): 809–18.
- Van Sickle, John. 1997. "Using Mean Similarity Dendrograms to Evaluate Classifications." *Journal of Agricultural, Biological and Environmental Statistics* 2 (4): 370–88.
- Van Sickle, John, and R. M. Hughes. 2000. "Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon." *Journal of the North American Benthological Society* 19 (3): 370–84. <https://doi.org/10.2307/1468101>.
- Vannote, Robin L., G. Wayne Minshall, Kenneth W. Cummins, James S. Sedell, and Colbert E. Cushing. 1980. "The River Continuum Concept." *Canadian Journal of Fisheries and Aquatic Sciences* 37 (1): 130–37. <https://doi.org/10.1139/f80-017>.
- Verdonschot, Piet F. M. 2006a. "Data composition and taxonomic resolution in macroinvertebrate stream typology." *Hydrobiologia* 566 (1): 59–74. <https://doi.org/10.1007/s10750-006-0070-y>.
- . 2006b. "Evaluation of the use of Water Framework Directive typology descriptors, reference sites and spatial scale in macroinvertebrate stream typology." *Hydrobiologia* 566 (1): 39–58. <https://doi.org/10.1007/s10750-006-0071-x>.
- Verdonschot, Piet F. M., and Rebi C. Nijboer. 2004. "Testing the European stream typology of the Water Framework Directive for macroinvertebrates." *Hydrobiologia* 516 (1-3): 35–54. <https://doi.org/10.1023/B:HYDR.0000025257.30311.b7>.
- Vilmi, Annika, Corentin Gibert, Gilles Escarguel, Konsta Happonen, Jani Heino, Aurélien Jamoneau, Sophia I. Passy, 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. <https://doi.org/10.1111/ecog.05356>.
- Vogt, Jürgen, P. Soille, A. De Jager, E. Rimavičiūtė, W. Mehl, S. Foisneau, K. Bódis, et al. 2007. *A pan-European river and catchment database*. <https://doi.org/10.2788/35907>.
- Wagner, Helene H, and Peter J Edwards. 2001. "Quantifying Habitat Specificity to Assess the Contribution of a Patch to Species Richness at a Landscape Scale." *Landscape Ecology* 16 (2): 121–31.
- Westfall, Peter H, and S Stanley Young. 1993. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. Vol. 279. John Wiley & Sons.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemond, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wright, John F, David W Sutcliffe, and Mike T Furse. 2000. "Assessing the Biological Quality of Fresh Waters." *RIVPACS and Other Techniques*. Freshwater Biological Association, Ambleside, England.
- WWF. 2020. *Living Planet Report 2020-Bending the Curve of Biodiversity Loss*. World Wildlife Fund.



Youn, So-Jung, William W Taylor, Abigail J Lynch, Ian G Cowx, T Douglas Beard Jr, Devin Bartley, and Felicia Wu. 2014. "Inland Capture Fishery Contributions to Global Food Security and Threats to Their Future." *Global Food Security* 3 (3-4): 142–48.