Science of the Total Environment 931 (2024) 172850

Contents lists available at ScienceDirect

Science of the Total Environment

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

Improving the performance of macroinvertebrate based multi-metric indices by incorporating functional traits and an index performance-driven approach

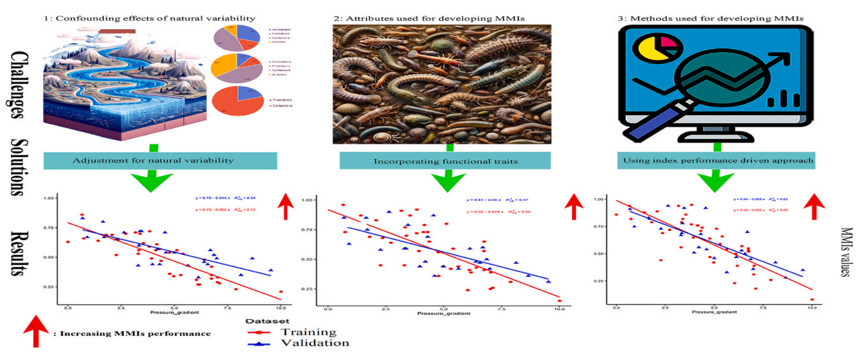
Ali Reza Esmaeili Ofogh [a,d], Eisa Ebrahimi Dorche [a], Sebastian Birk [b,c], Pejman Fathi [a,d], Mojgan Zare Shahraki [a,d], Andreas Bruder [d,*]

[a] *Department of Natural Resources, Isfahan University of Technology, Isfahan 84156-83111, Iran*
[b] *University of Duisburg-Essen, Faculty of Biology, Aquatic Ecology, Universitätsstraße 5, 45141 Essen, Germany*
[c] *Centre for Water and Environmental Research, University of Duisburg-Essen, Universitätsstraße 5, 45141 Essen, Germany*
[d] *Institute of Microbiology, University of Applied Sciences and Arts of Southern Switzerland, 6850 Mendrisio, Switzerland*

HIGHLIGHTS

- Multi-metric indices (MMIs) track anthropogenic effects on freshwater ecosystems.
- Adjustment for natural environmental gradients improves MMIs performance.
- Incorporating functional traits improves MMIs performance and interpretation.
- Index performance-driven MMIs outperform metric performance-driven MMIs.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Functional diversity
Biomonitoring
Spatial variability
Karun River
Middle East
Anthropogenic pressures

ABSTRACT

Human-driven multiple pressures impact freshwater ecosystems worldwide, reducing biodiversity, and impacting ecosystem functioning and services provided to human societies. Multi-metric indices (MMIs) are suitable tools for tracking the effects of anthropogenic pressures on freshwater ecosystems because they incorporate various biological metrics responding to multiple pressures at different levels of biological organization. However, the performance and applicability of MMIs depend on their metrics' selection and their calibration against natural environmental gradients. In this study, we aimed to unravel *i*) how incorporating functional trait-based metrics affects the performance of MMIs, *ii*) how disentangling the natural environmental gradients from anthropogenic pressures effects affects the performance of MMIs, and *iii*) how the performance of MMIs developed using a metric performance-driven approach compares with MMIs developed using an index performance-driven approach. We carried out a field survey measuring abiotic and biotic variables at 53 sites in the Karun River basin (Iran) in 2018. For functional trait-based metrics, we used 15 macroinvertebrate traits and calculated community-weighted mean trait values and functional diversity indices. We used random forest modeling to

* Corresponding author.
E-mail address: andreas.bruder@supsi.ch (A. Bruder).

https://doi.org/10.1016/j.scitotenv.2024.172850
Received 15 September 2023; Received in revised form 20 April 2024; Accepted 26 April 2024
Available online 28 April 2024

account for the effect of natural environmental gradients on each metric. Based on our results, incorporating functional traits increased the MMI performance significantly and facilitated ecological interpretation of MMIs. Both taxonomic and functional components of macroinvertebrate assemblages co-varied strongly with natural environmental gradients, and accounting for these covariations improved the performance of MMIs. Finally, we found that index performance-driven MMIs performed better in terms of precision, bias, sensitivity, and responsiveness than metric performance-driven MMIs.

1. Introduction

Freshwater ecosystems are rich in biodiversity: they cover only 1 % of the earth's surface, but support ~10 % of species diversity (Collen et al., 2014). However, anthropogenic pressures can drastically affect biodiversity and ecosystem functions and services provided by freshwaters (Cardinale et al., 2012; Darwall et al., 2018; Reid et al., 2019; Smeti et al., 2019). To conserve the least impacted and restore the impacted freshwater ecosystems, managers and stakeholders need tools to efficiently quantify the biodiversity status of freshwaters and to assess the magnitude of threats. However, most monitoring programs at national and continental levels have been developed in and for Europe and the United States of America, leaving freshwater ecosystems in other regions, often with high levels of endemism, exposed to human pressures with unknown consequences for their biodiversity (Eriksen et al., 2021; Poikane et al., 2020).

Monitoring of freshwater ecosystem condition was originally based on measuring the physical and chemical variables of water and habitat (Horton, 1965). More recent assemblage-based assessments developed based on different organism groups (i.e., macroinvertebrates, fishes, diatoms, macrophytes) provide complementary tools for assessing river and stream ecosystem conditions (Buss et al., 2015; Masese et al., 2013). Multi-metric indices (MMIs) combine different attributes of the biological assemblages, are (usually) sensitive to anthropogenic pressures, and consequently provide a comprehensive assessment of river ecosystem condition (Schoolmaster et al., 2012). The methods used to develop MMIs have improved over the last decades; however, some challenges remain particularly in integrating ecological principles to develop robust, reliable and informative MMIs. These challenges include *i*) identifying the most efficient attributes of biological assemblages to track the effects of anthropogenic pressures on freshwater ecosystem conditions (Ruaro et al., 2020), *ii*) considering confounding effects between natural environmental gradients and anthropogenic pressure gradients on attributes of the biological assemblages (Cao et al., 2007), *iii*) improving the applicability and the efficiency of methods used to select a suite of ecologically non-redundant metrics that best predict the responses of biological assemblages to anthropogenic pressure gradients (Stoddard et al., 2008), and *iv*) developing approaches to quantify pressure impacts on ecosystem functioning (Berger et al., 2018; Mondy and Schuwirth, 2017).

MMIs were mostly developed based on taxonomic descriptions of aquatic communities (Vadas et al., 2022). However, recent studies showed that functional traits can have high potential for informing on the effects of anthropogenic pressures on biodiversity and ecosystem functioning (Verberk et al., 2013). Environmental conditions filter species locally based on their functional traits (Bêche and Statzner, 2009), i.e., the phenotypic attributes defining the performance and fitness of organisms in relation to environmental conditions (Violle et al., 2007). Functional trait-based metrics therefore have the potential to provide mechanistic understanding of responses of biological assemblages to anthropogenic pressures. For example, the presence of certain functional traits in macroinvertebrate communities, such as the ability to attach to a certain substrate or to break down organic matter, can be used to develop metrics that reflect the ability of ecosystems to provide crucial functions and services (Juvigny-Khenafou et al., 2021). Such metrics can then be combined into multi-metric indices that provide a more comprehensive assessment of stream ecosystem condition than any one metric could provide alone (Elosegi et al., 2017). Despite this great potential of functional traits (Hering et al., 2006; Stoddard et al., 2008), they are yet rarely used in MMIs (but see Bolding et al., 2020; Chen et al., 2019). As a result, we do not know if and by how much the inclusion of functional traits can effectively improve MMIs and, in turn, our knowledge of the effects of anthropogenic pressures on the ecological integrity of freshwater ecosystems (Larson et al., 2021).

A significant challenge facing MMIs utilized in ecosystem management is the variability in assemblage structure across natural environmental gradients, which can overlap with gradients of anthropogenic pressures (Álvarez-Cabria et al., 2010; Ding et al., 2017). Co-variation of natural environmental variables and anthropogenic pressures can reduce our ability to detect the effects of the latter on biological assemblage structure (Mackey and Currie, 2001). Taking account of natural environmental gradients through regional classifications such as ecoregion, basin, or physiographic regions has the potential to enhance the performance of MMIs (Chen et al., 2014; Hawkins et al., 2000). However, some studies have shown limited benefits of regional classifications in mitigating natural gradients in assemblages (Denison et al., 2021). Therefore, integrating regional classifications with local environmental conditions could assist in distinguishing between the impacts of natural environmental gradients and anthropogenic pressures, leading to improved MMI performance (Cao et al., 2007; Mazor et al., 2016).

Another way to enhance the effectiveness of MMIs in detecting ecological impairments involves the selection and combination of individual metrics. Traditional MMIs are based on selecting the most responsive non-collinear metrics from different trophic, taxonomic, and habitat groups, or tolerance classes, and combining them (referred to as the metric performance-driven approach; Hering et al., 2006; Stoddard et al., 2008). However, van Sickle (2010) demonstrated that selecting metrics solely based on their individual performance does not always result in the best overall MMI performance. He suggested that even metrics with weak associations to anthropogenic pressures could contribute to the overall predictive power of the MMI if they explain variance in pressures not captured by other metrics. van Sickle (2010) proposed a method of developing potential MMIs by randomly selecting metrics and then choosing the best MMIs based on their overall performance (referred to as the index performance-driven approach; Bolding et al., 2020).

In this research, we created and compared the performance of MMIs using various combinations of metrics that describe different aspects of benthic macroinvertebrate communities in streams and rivers of the Karun River basin in western Iran. Freshwater ecosystem resources in Iran, as in other regions, are crucial due to water scarcity and the socio-economic importance of these ecosystems. The rapid degradation of freshwater ecosystems underscores the immediate need for appropriate tools for assessing and managing freshwater ecosystem condition. Based on the state of research in this field described above, we anticipated that the performance of MMIs would improve by: *i*) incorporating functional trait-based metrics, *ii*) distinguishing natural environmental gradients from anthropogenic pressure effects, and *iii*) adopting an index performance-driven approach.

2. Material and methods

2.1. Study area

The present study was conducted in wadable sections of streams and rivers of the Karun River basin (described in detail in Esmaeili Ofogh et al., 2023). The results of Esmaeili Ofogh et al. (2023) indicated that autumn was the most suitable season for the development of multimetric indices in the Karun. Thus, in this study we used the water quality and physical habitat quality data collated in a companion study (Fathi et al., 2022) and new biotic data describing benthic macroinvertebrate assemblages (the same samples as in Esmaeili Ofogh et al. (2023) but identified to higher taxonomic resolution) from 53 sites sampled in autumn 2018 (Fig. 1).

2.2. Sample collection and analysis

We measured dissolved oxygen, electrical conductivity, and pH in-situ. We took water samples in triplicate at a depth of 10–15 cm from a riffle at each site to measure physical and chemical variables in the lab (Table S.1) using standard methods (APHA, 2017). Physical habitat quality assessment was conducted based on the framework established by Barbour et al. (1999). The methodology used for this assessment is described in detail in Fathi et al. (2022). Ten physical habitat variables were scored by a panel of six experts. To enhance the repeatability of our method, we implemented the signal-to-noise ratio (S/N) test, as proposed by Kaufmann et al. (1999). This test enabled us to evaluate the reproducibility of physical habitat features, ensuring that any differences observed between sites were attributed to variations in stream conditions rather than discrepancies in expert judgment.

We collected ten macroinvertebrate samples per site with a kick net (30 × 30 cm D-frame, 250 μm mesh, covering 0.09 m^2 of streambed each) along 200 m of river length, with habitats (riffles, runs, and pools) sampled in proportion to their occurrence (Barbour et al., 1999). All samples collected from one site were combined into a single composite sample. Samples were then transferred into sealed plastic containers and preserved in 10 % buffered formalin. After separating the macroinvertebrates from other material, all individuals per sample were counted and identified to the lowest possible level (mostly genus) using standard taxonomic keys (Bouchard and William, 2004; Kriska, 2013; Peckarsky et al., 1990; Tachet et al., 2010).

2.3. Defining the anthropogenic pressure gradient

We used sites in least disturbed conditions to define reference conditions. We used a Principal Component Analysis (PCA) to reduce

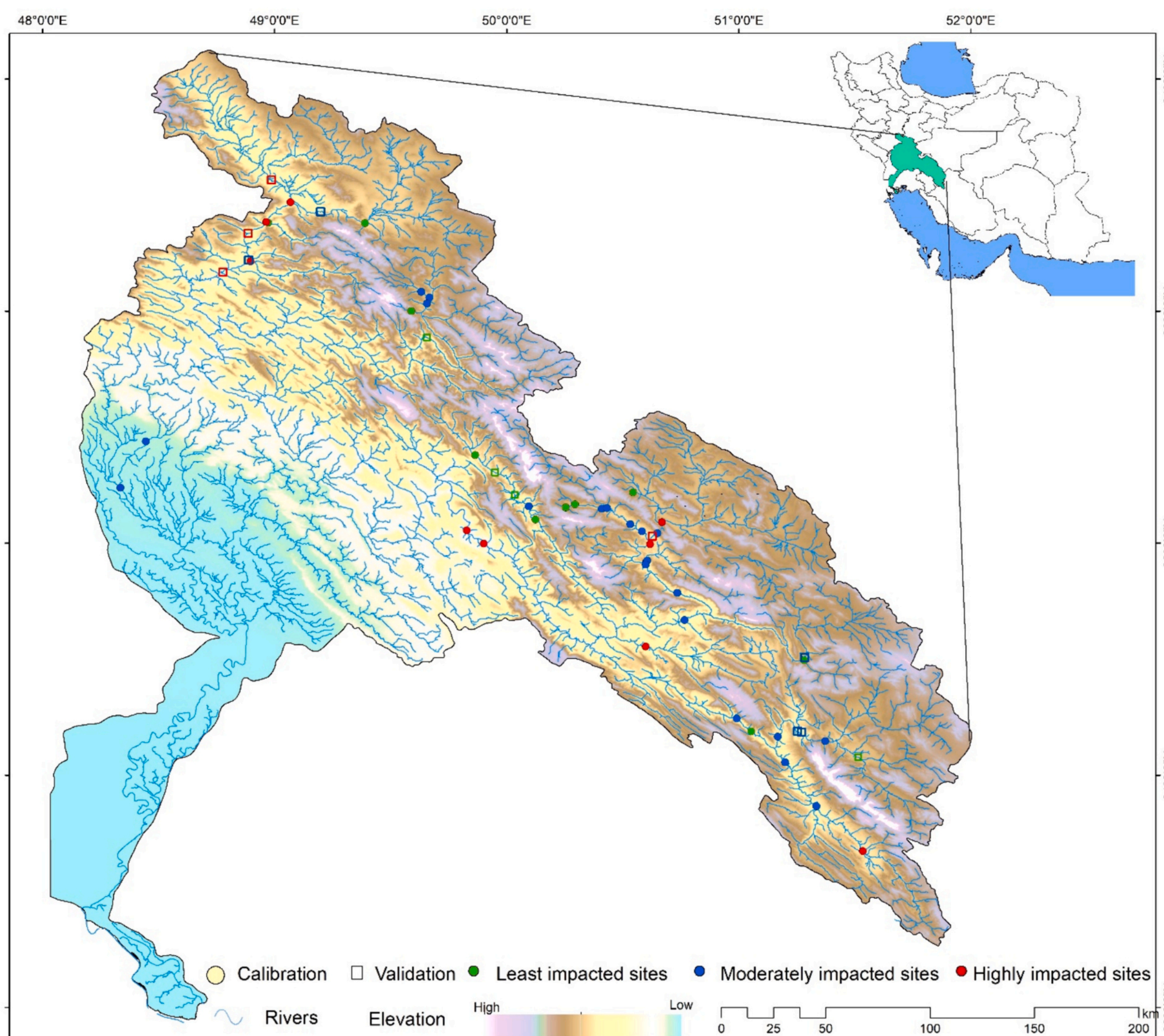


Fig. 1. Location of the sampling sites in the Karun River basin, Iran and their classification based on abiotic parameters.

dimensionality and increase the interpretability of patterns of physical and chemical variables and physical habitat characteristics (Zare Shahraki et al., 2021). PCA axis 1 (PC1) was selected as the anthropogenic pressure gradient to characterize least impacted, moderately impacted, and highly impacted sites (for details see Esmaeili Ofogh et al., 2023).

2.4. Candidate taxonomic and functional metrics

In this study, we used *a priori* metrics selection, based on the expected response of metrics to anthropogenic pressures. Specifically, we calculated 259 metrics using taxonomic information, which were grouped into four categories: tolerance/intolerance, taxonomic diversity indices, taxa richness, and community composition (Table S.2). These categories are intended to capture distinct attributes of macroinvertebrate assemblages and give insights into their biotic integrity (Barbour et al., 1999; Stoddard et al., 2008). The tolerance/intolerance category describes the sensitivity of assemblages to anthropogenic pressures and is divided into very sensitive, sensitive, tolerant, and very tolerant. The community composition category describes the relative abundance of different taxonomic groups (e.g., family) in the assemblage. The taxa richness category describes the number of different taxa, while the taxonomic diversity indices category takes into account the distribution of abundance among taxa and taxonomic relatedness between different taxa in a community.

For calculating functional metrics, we selected 15 functional traits based on their response mechanism to pressure gradients and separated them into categories of life history, ecological preferences, morphological, and dispersal traits (Larson et al., 2021). Each trait was described by two to nine modalities, for a total of 79 modalities (Table S.3). We derived information on traits from different sources including peer-reviewed literature (Tachet et al., 2010; Tomanova and Usseglio-Polatera, 2007; Usseglio-Polatera et al., 2000) and online databases (Schmidt-Kloiber and Hering, 2022). We accounted for intraspecific variability in morphological, habitat and behavioral attributes of macroinvertebrate communities using a fuzzy coding procedure to assign a score to each modality of the trait (Chevenet et al., 1994). The species traits were scored on a 0-to-x scale with "0" indicating "no affinity" to "x" indicating "high affinity" of the taxon for the modality. Subsequently, for each 'taxon $\times$ trait', the preference scores were normalized into a distribution of relative usage frequencies. This was achieved by dividing the preference scores of the species across the categories of a fuzzily-coded attribute by their total sum.

To quantify functional diversity, we followed four steps. First, we described the functional strategies of species using different sets of functional traits. Second, we calculated the functional dissimilarity between species. We used Gower's distance as our dissimilarity metric due to its ability to accommodate different types of traits (numerical and categorical traits) and its tolerance towards missing values (Maire et al., 2015). Third, we built a functional space based on functional dissimilarity (Maire et al., 2015; Villéger et al., 2008). To identify the best functional space, we computed several functional spaces using different clustering methods such as UPGMA, WPGMA, WPGMC, or Ward's method (Petchey and Gaston, 2002), as well as multidimensional functional spaces using PCA (Villéger et al., 2008). Fourth, we selected the best functional space based on how well it represents the initial functional dissimilarity between species (Maire et al., 2015).

We used the *functcomp* function in the FD package in R (Laliberté et al., 2014) to calculate the community-weighted mean (CWM) of each trait modality in the assemblage as a measure of functional composition (99 functional composition metrics; Table S.4). Moreover, we used the *dbFD* function to calculate functional diversity indices (i.e., functional richness [FRic], functional evenness [Feve], functional redundancy [FR], functional dispersion [Fdis], and Rao's entropy [RaoQ]; Laliberté et al., 2014) separately for different functional metric categories and for all traits (25 functional diversity indices, Table S.4).

2.5. Environmental predictor variables

To test the effects of natural environmental gradients on the structure of macroinvertebrate assemblages and consequently on taxonomic and functional metrics, we selected and measured physiographic and bio-climatic variables as predictor variables in each site (Table S.5). The physiographic variables of altitude, river channel slope, Strahler stream order and wetted river width were measured in the field or from maps. Bio-climatic variables, including annual averages (e.g., mean annual temperature, annual precipitation), seasonality (e.g., annual range in temperature and precipitation), and extreme or limiting environmental factors (e.g., temperature of the coldest and warmest month, and precipitation of the wet and dry quarters) were obtained from the CHELSA (Climatologies at High resolution for the Earth's Land Surface Areas) database with a spatial resolution of 30 arc sec (Karger et al., 2017).

2.6. MMI development

We partitioned the data into a training and a validation dataset, using the former to develop MMIs and the latter as an independent dataset to test the MMIs (Bolding et al., 2020). To maintain a gradient of pressures (least, moderate, and high) in both datasets, we randomly selected four sites from reference sites (7 %) and nine sites from impacted sites (16 %) for the validation dataset. The initial steps for metric screening (including range test, sensitivity test and adjustment for natural environmental gradients) were the same between metric performance-driven and index performance-driven approaches (Fig. 2; Bolding et al., 2020). Metrics with low variability between sites are not able to detect natural variability or different levels of anthropogenic pressure (Hering et al., 2006). Thus, we first excluded metrics with very small ranges and/or with similar values at most sites (i.e., if the value of the first quartile equaled that of the third quartile). Second, we retained metrics that correlated with Spearman $R > 0.4$ ($p < 0.05$) to the pressure gradient (van Sickle, 2010). Third, we used predictor variables that are unaltered by direct human activities (the physiographic and bio-climatic variables) to determine expected metric values for each site (Chen et al., 2019). We used random forest (RF) modelling to calculate metric responses to natural environmental gradients. We selected RF models because of their ability to handle non-linear relationships, to model interactions between predictor variables, to conduct bootstrap subsamples, and because their are insensitive to correlated predictor variables and overfitting (Cutler et al., 2007; Chen et al., 2014). We used residuals (observed minus predicted metric values) for further calculations when RF models explained >10 % of the variation in metric values in reference sites; otherwise, we retained the raw metric values (Hawkins et al., 2010).

2.6.1. Development of metric performance-driven MMIs

Using the metric performance-driven approach, we developed four MMIs: natural-gradients-unadjusted taxonomic MMI (MMI_unadj_tax), natural-gradients-adjusted taxonomic MMI (MMI_adj_tax), natural-gradients-unadjusted combined (taxonomic + trait-based metrics) MMI (MMI_unadj_comb), and natural-gradients-adjusted combined MMI (MMI_adj_comb) (Fig. S.1). For each MMI, we followed a stepwise procedure for core metric selection including responsiveness test, redundancy analysis, final metrics scoring, and MMIs development and performance analysis (Fig. S.1; for detailed methods see Esmaeili Ofogh et al., 2023).

2.6.2. Development of index performance-driven MMIs

Using the index performance-driven approach, we developed four MMIs to: *i*) compare the performance of natural environmental gradients adjusted MMIs with unadjusted MMIs, and *ii*) compare the performance of MMIs developed using taxonomic metrics with those developed using both taxonomic and functional metrics. We utilized the pool of taxonomic and functional metrics unadjusted to natural environmental

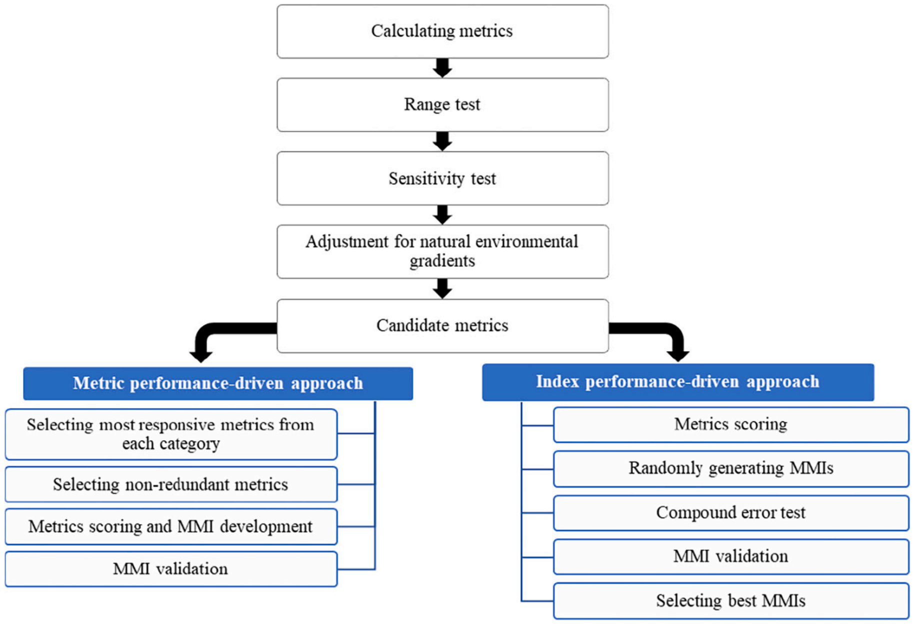


Fig. 2. Flowchart of the steps for the development of an MMI following the metric performance-driven approach and the index performance-driven approach, respectively (modified from Bolding et al., 2020).

gradients to develop the unadjusted taxonomic MMI (MMI_unadj_tax_4) and the taxonomic-functional MMI (MMI_unadj_comb_5). Similarly, we utilized the pool of taxonomic and functional metrics adjusted for natural environmental gradients to develop the adjusted taxonomic MMI (MMI_adj_tax_4) and the taxonomic-functional MMI (MMI_unadj_comb_5). To compare the performance of MMIs developed using a different number of metrics, we randomly selected two metrics from each aforementioned pool and developed four additional MMIs (MMI_unadj_tax_8, MMI_adj_tax_8, MMI_unadj_comb_10, and MMI_adj_comb_10; for more detail see Fig. S.2).

For the development of each index performance-driven MMI, we developed 55,000 MMIs by randomly selecting metrics from the pool of scored metrics that passed the range and sensitivity test (Fig. S.2). If any of the included metrics exhibited a pair-wise Spearman correlation coefficient with a value of $R > 0.85$, or if the average pair-wise Spearman correlation coefficient among the metrics' residuals from the regressions on pressure gradients exceeded 0.60, then the MMI was considered unsuitable as there was a risk of amplifying noise (Bolding et al., 2020). Finally, the three best randomly developed MMIs were selected based on the t-score (highest t-scores for least impacted sites *vs.* highly impacted sites). The values of these three MMIs were regressed against the anthropogenic pressure gradient using the training dataset, and the MMI with the highest R^2 was selected as the final MMI (Fig. S.2).

2.6.3. MMI testing and validation

The performance of the MMIs was tested in four steps. First, we compared the coefficient of variation (CV) of MMI values observed at reference sites to estimate MMI precision. The lower the CV, the more precise the MMI (Chen et al., 2014). Second, to evaluate MMI bias, we used multiple linear regression to assess whether the variation in reference site MMI values was systematically associated with natural variation of environmental variables. Third, we used box-plots of reference and impacted site MMI values to visually assess responsiveness. In addition, we ran independent Student's *t*-tests to test the ability of each candidate metric to differentiate between the least, moderately and highly impacted sites. Fourth, we measured MMI sensitivity as the proportion of impacted sites that would be inferred as non-reference with 10th percentiles of values at reference sites as threshold values (Cao et al., 2007; Paulsen et al., 2008; Stoddard et al., 2008).

For the validation of developed MMIs, we linearly regressed MMI values created from the validation dataset against the pressure gradient. If the *p*-value from a regression was <0.05, the MMI was considered validated. If more than one MMI had $p < 0.05$, we used the Akaike Information Criterion (AIC) to identify the best supported MMI of this candidate set. Moreover, we used non-nested likelihood ratio tests to determine whether adding more metrics to the MMIs (selecting two metrics from each category of metrics instead of one) would make the MMIs significantly more accurate in detecting the effects of anthropogenic pressure (Lewis et al., 2011). All statistical analyses were conducted in R (R Development Core Team, 2022).

3. Results

3.1. Macroinvertebrate assemblages' structure

In autumn 2018, 156,882 individuals from 110 aquatic macroinvertebrate taxa (mostly at genus level) were collected from 53 sites and included in the analysis. The average number of collected individuals per m^2 was 2467. The most abundant taxa were *Hydropsyche* sp., *Labiobaetis* sp., Orthocladiinae, Gammaridae and Chironominae with relative abundances of 10.9, 10.9, 10.8, 8.7, and 7.5 %, respectively, across all samples.

3.2. Defining the pressure gradient

PC1 explained 23.5 % of environmental parameters' variation (Figs. S.3 and S.4). Based on the derived condition class thresholds for the reference sites, 13 sites (25 %) were in the least-impacted conditions. Variable loadings showed that morphological score, electrical conductivity, total nitrogen, and biological oxygen demand were most strongly associated with PC1 (Fig. S.5).

3.3. Adjustment for natural environmental covariation

Natural environmental conditions are highly heterogeneous in the Karun river catchment. The altitude of the sampling sites changed from 67 to 2087 m a.s.l. River channel slope ranged from <0.001 to 8.5 %, annual mean temperature varied from 11.0 to 25.3 °C (Table S.5), and annual mean precipitation ranged from 316 to 886 mm/year (Table S.5). Natural environmental gradients explained >10 % of variation of 84 out of 259 taxonomic-based metrics (32 %) and of 24 out of 104 trait-based metrics, (23 %), respectively. For example, natural environmental gradients explained 62.2 % of the variation of Plecoptera taxa richness and 30.9 % of the variation of Insecta relative percentage of taxa in the reference sites based on RF modelling results.

3.4. Metric performance-driven MMIs

The composition of final metrics selected for MMI development was different between MMIs. For MMI_adj_tax the first two final metrics (relative percentage (%) of tolerant taxa based on BMWP tolerance values, and relative abundance (%) of tubificids) were similar to MMI_unadj_tax. However, the adjusted geographical rarity (Gaston and He, 2010), and taxa richness of Ephemeroptera+Plecoptera outperformed in comparison to unadjusted geographical rarity and taxa richness of Ephemeroptera+Plecoptera metrics and were selected as final metrics (Table S.6). The two final metrics selected from functional composition and functional diversity categories were similar for MMI_unadj_comb and MMI_adj_comb including CWM of species preferably occurring in standing waters but regularly occurring in slow flowing streams and functional richness (FRic) calculated using life history traits (Table S.6).

All MMIs developed using the metric performance-driven approach effectively distinguished reference sites from moderately and highly impacted sites (Fig. 3). Among them, MMI_adj_comb demonstrated superior precision, responsiveness, and sensitivity, while exhibiting the lowest bias and AIC values compared to all other MMIs (Tables 1 and S.7). Incorporating trait-based metrics in the metric performance-driven approach notably enhanced MMI precision (by 4.5 %), responsiveness (as indicated by t-scores for least *vs.* moderately impacted sites = 0.52 and least *vs.* highly impacted sites = 1.24), and sensitivity (by 6.4 %) (Table 1). Moreover, adjusting for natural environmental gradients further augmented precision (by 6.0 %), sensitivity (by 7.7 %), and responsiveness (with t-scores for least *vs.* moderately impacted sites = 0.32 and least *vs.* highly impacted sites = 1.00), while reducing MMI bias (by 4.9 %) when combining taxonomic and functional metrics. Similarly, adjustment for natural environmental gradients increased the precision (by 3.0 %), sensitivity (by 5.1 %), and responsiveness (with t-scores for least *vs.* moderately impacted sites = 0.72 and least *vs.* highly impacted sites = 1.48) of MMI_adj_tax. However, MMI_adj_tax exhibited higher bias compared to MMI_unadj_tax (Table 1).

All MMIs developed based on metric performance-driven approach had strong regressions between MMIs scores and pressure gradients based on the training dataset (Adjusted R^2 ranging from 0.55 to 0.63). Incorporating functional diversity indices in MMIs and adjustment for natural environmental gradients did not affect the MMIs' relationship with the pressure gradients based on the validation dataset (Fig. 4). This was confirmed by the results of the likelihood ratio test that showed that incorporating trait-based metrics and adjustment for natural environmental gradients did not affect MMIs performance in the metric performance-driven approach (Table 2).

3.5. Index performance-driven MMIs

From the pool of 259 unadjusted taxonomic metrics, 194 passed the range test. Of these, 68 metrics passed the sensitivity test. Out of all trait-based unadjusted metrics, 22 metrics passed the range and sensitivity tests. By combining these 22 trait-based metrics with 68 taxonomic metrics, we produced a pool of 100 taxonomic and trait-based unadjusted metrics. From the pool of 259 adjusted taxonomic metrics, 192 passed the range test and 88 of these passed the sensitivity test. 20 trait-based adjusted metrics passed the range and sensitivity tests. By combining these metrics with 88 taxonomic metrics, we produced the pool of 108 taxonomic and trait-based adjusted metrics.

All MMIs developed based on the index performance-driven approach successfully discriminated reference sites from moderately and highly impacted sites (Fig. 5). Based on the bias test, the amount of variation of MMIs scores explained by natural environmental variables was near zero for all index-based MMIs (Table 2). Adjustment for the natural environmental gradient increased the precision (except for MMI_adj_tax_4), sensitivity and responsiveness of all MMIs (Table 2). Index performance-driven MMIs developed by combining taxonomic

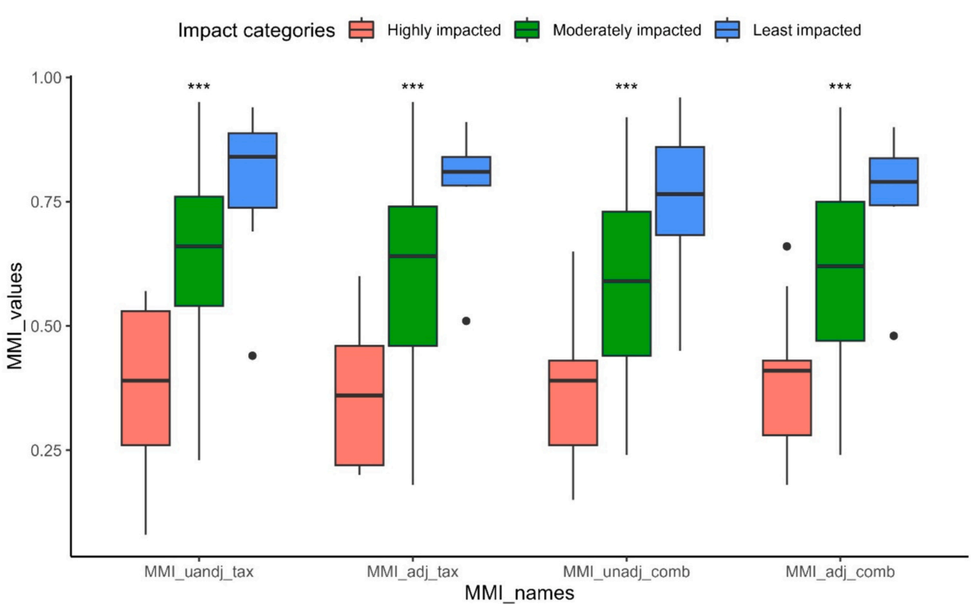


Fig. 3. Ranges of the four MMIs developed by metric performance-driven approach at least, moderately and highly impacted sites (MMI_unadj_tax: MMI developed using unadjusted taxonomic metrics; MMI_adj_tax: MMI developed using adjusted taxonomic metrics; MMI_unadj_comb: MMI developed using unadjusted taxonomic and functional metrics; MMI_adj_comb: MMI developed using adjusted taxonomic and functional metrics). Boxes show the interquartile (25th–75th percentiles), horizontal lines the median, points the outliers, and range bars the maximum and minimum values excluding outliers (Kruskal-Wallis test result, ***: *p*-value < 0.001).

Table 1

Comparison of MMIs performance in terms of precision (coefficient of variation [CV]), bias (adjusted R^2: % of variation among reference site MMI values explained by multiple regression model), sensitivity (% of priory classified impacted sites considered impaired), and responsiveness (Student's t-value between least, moderate and highly impacted sites). MMI_unadj_tax: MMI developed using unadjusted taxonomic metrics; MMI_adj_tax: MMI developed using adjusted taxonomic metrics; MMI_unadj_comb: MMI developed using unadjusted taxonomic and functional metrics; MMI_adj_comb: MMI developed adjusted taxonomic and functional metrics; MMI_unadj_tax_4: MMI developed using 4 unadjusted taxonomic metrics; MMI_unadj_tax_8: MMI developed using 8 unadjusted taxonomic metrics; MMI_adj_tax_4: MMI developed using 4 adjusted taxonomic metrics; MMI_adj_tax_8: MMI developed using 8 adjusted taxonomic metrics; MMI_unadj_comb_5: MMI developed using 5 unadjusted taxonomic and functional metrics; MMI_unadj_comb_10: MMI developed using 10 unadjusted taxonomic and functional metrics; MMI_comb_fun_5: MMI developed using 5 adjusted taxonomic and functional metrics; MMI_comb_10: MMI developed using 10 adjusted taxonomic and functional metrics.

MMIs type	MMIs name	Precision	Bias	Sensitivity	Responsiveness		
					Least *vs.* moderate	Moderate *vs.* high	Least *vs.* high
Metric performance-driven MMIs	MMI_unadj_tax	39	7.45^{ns}	66.7	3.52***	3.96***	6.77***
	MMI_adj_tax	36	10.82*	71.8	4.24***	3.67***	8.25***
	MMI_unadj_comb	27	11.42*	79.49	3.54***	3.65***	6.58***
	MMI_adj_comb	21	6.52^{ns}	87.18	3.86***	3.72***	7.58***
Index performance-driven MMIs	MMI_unadj_tax_4	13.51	0.00^{ns}	93.33	5.60***	2.36*	5.60***
	MMI_unadj_tax_8	13.18	0.42^{ns}	90	5.07***	2.72*	5.07***
	MMI_adj_tax_4	20.14	0.00^{ns}	93.33	4.26***	4.00***	7.73***
	MMI_adj_tax_8	10.60	0.00^{ns}	93.33	5.88***	2.34*	7.84***
	MMI_unadj_comb_5	21.26	0.00^{ns}	73.33	4.62***	3.84***	7.60***
	MMI_unadj_comb_10	17.75	0.27^{ns}	90	5.33***	3.46**	8.42***
	MMI_adj_comb_5	14.05	0.00^{ns}	90	6.31***	3.47**	9.14***
	MMI_adj_comb_10	12.56	0.04^{ns}	93.33	5.91***	3.48**	9.40***

^{ns} p-Value > 0.05.
* p-Value < 0.05.
** p-Value < 0. 01.
*** p-Value < 0.001.

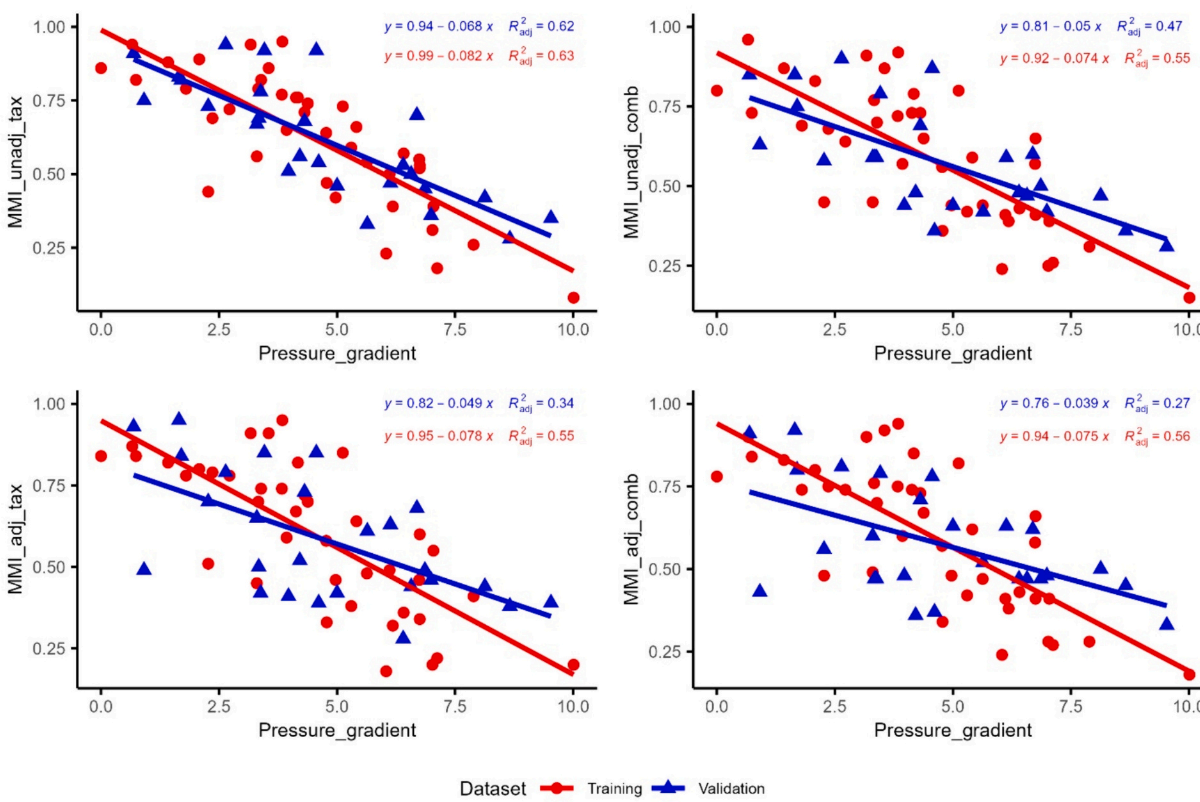


Fig. 4. Linear regression of metric performance-driven MMIs scores versus pressure gradient, for training datasets (solid red circles, solid red lines) and validation datasets (solid blue triangles, solid blue lines). (MMI_unadj_tax: MMI developed using unadjusted taxonomic metrics; MMI_adj_tax: MMI developed using adjusted taxonomic metrics; MMI_unadj_comb: MMI developed using unadjusted taxonomic and functional metrics; MMI_adj_comb: MMI developed using adjusted taxonomic and functional metrics).

and trait-based metrics had higher precision (lower CV), sensitivity and responsiveness than MMIs based on taxonomic metrics (Table 2).

Similar to metric performance-driven MMIs, all MMIs developed based on the index performance-driven approach had very strong relationships between MMIs scores and the pressure gradient based on the training dataset (Fig. 6). However, the strength of the relationship between MMIs scores and pressure gradient was substantially higher in MMIs developed using taxonomic and functional diversity metrics (R^2

Table 2
Results of likelihood-ratio tests comparing the MMIs, describing the MMIs' ability to detect the change in pressure gradient.

	LRT[a]	*P*-value	Df[b]
Metric performance-driven MMIs			
MMI_unadj_tax *vs.* MMI_adj_tax	1.95	0.74	6
MMI_unadj_comb *vs.* MMI_adj_comb	4.73	0.78	8
MMI_unadj_tax *vs.* MMI_unadj_comb	8.32	0.21	6
MMI_adj_tax *vs.* MMI_adj_comb	11.10	0.08	6
Index performance-driven MMIs			
MMI_unad_tax_4 *vs.* MMI_unadj_tax_8	7.64	0.11	6
MMI_adj_tax_4 *vs.* MMI_adj_tax_8	16.61	0.002	6
MMI_unadj_comb_5 *vs.* MMI_unadj_comb_10	2.58	0.77	7
MMI_adj_comb_5 *vs.* MMI_adj_comb_10	4.54	0.47	7

[a] LRT: likelihood of ratio test.
[b] Df: degree of freedom.

ranging from 0.44 to 0.63) than those developed based only on taxonomic metrics (R^2 ranging from 0.35 to 0.46) based on the validation dataset (Fig. 5). Adjustment for natural environmental variables decreased the strength of the relationship between MMIs sores and pressure gradients (Fig. 5). The likelihood ratio tests indicated that adding more metrics to MMIs just increased the performance of MMI_adj_tax_8 in comparison to MMI_adj_tax_4 (Table 2). Based on the AIC, the MMI_adj_comb_5 provided the best fit among the index performance-driven MMIs (Table S.7).

Based on the precision, bias, responsiveness, and sensitivity tests and AIC, the index performance-driven MMIs had a significantly better fit than the metric performance-driven MMIs (Tables 1 and S.7). The index performance-driven MMIs also explained more of the variance of the pressure gradient in the training and validation datasets (Figs. 4 and 6).

4. Discussion

Anthropogenic activities significantly affect the abiotic conditions of stream ecosystems, leading to changes in the taxonomic and functional diversity and composition of their benthic macroinvertebrate assemblages (Bêche and Statzner, 2009). MMIs provide a useful framework to combine functional (e.g. functional feeding groups) and structural assemblage attributes to assess anthropogenic pressure effects on the biological integrity of communities. The purpose of this study was to assess the performance of MMIs to detect ecological impairment of freshwater ecosystems by incorporating functional traits, adjusting for natural environmental gradients, and evaluating different combinations of metrics. Recent studies highlighted the potential of functional traits in quantifying and predicting the effects of anthropogenic pressures on the condition of freshwater ecosystems and their biota, and to inform on the mechanisms of impacts (Culp et al., 2011; Rodil et al., 2013; van den Brink et al., 2011). However, the application of functional traits in routine biomonitoring is scarce (Chen et al., 2019; Ding et al., 2017). Moreover, confounding effects of natural spatial and temporal variability on macroinvertebrate assemblage structure have the potential to greatly decrease the ability of MMIs to detect ecological impairment (Chen et al., 2014; Denison et al., 2021). We compared the performance of two common approaches for developing MMIs, i.e., the more traditional metric performance-driven approach *vs.* the index performance-driven approach, utilizing taxonomic and functional metrics of macroinvertebrate assemblages sampled from streams of the Karun River basin, Iran.

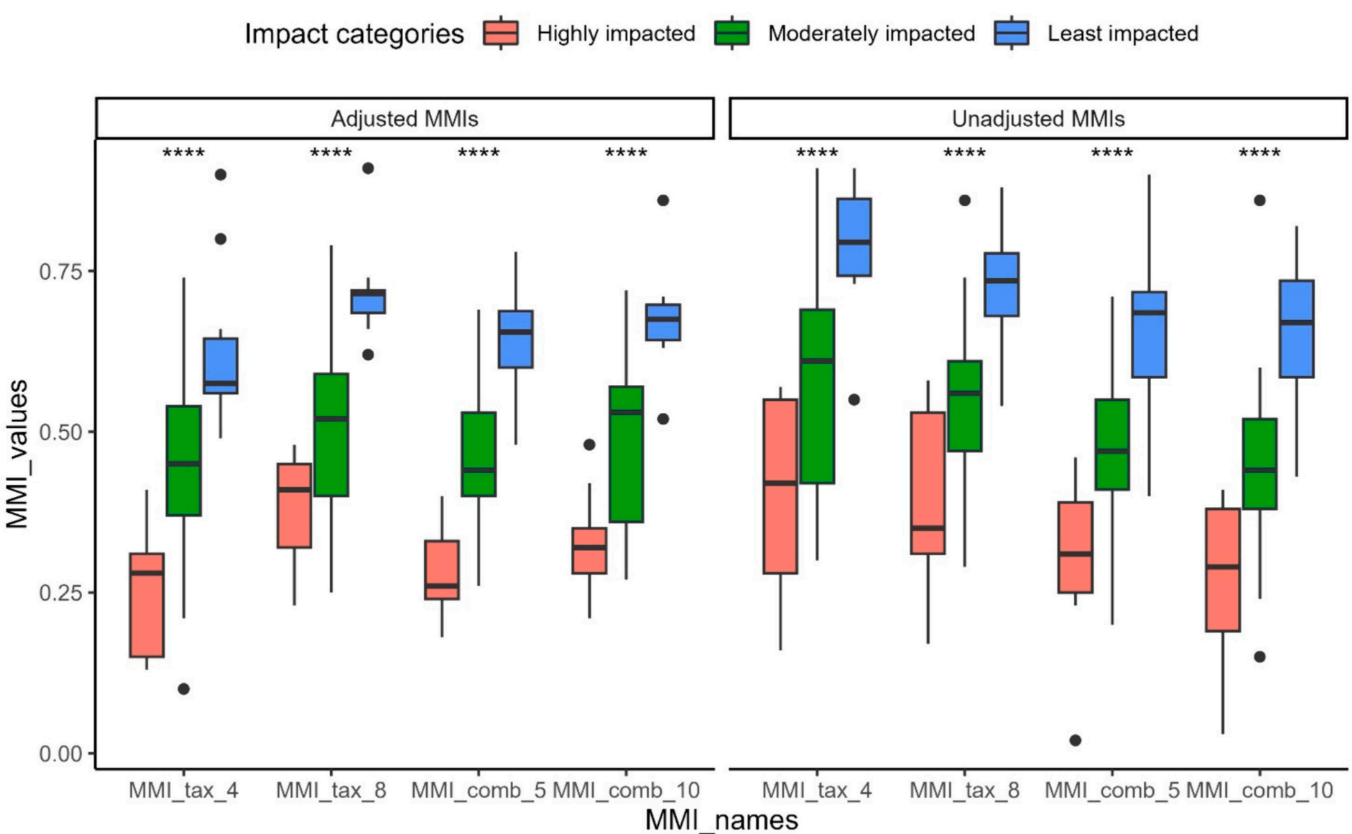


Fig. 5. Ranges of the index performance-driven MMIs at least, moderately and highly impacted sites. Boxes show the interquartile (25th–75th percentiles), horizontal lines the median, circles the outliers, and range bars the maximum and minimum values excluding outliers (Kruskal-Wallis test result, ****: p-value < 0.001). (MMI_ tax_4: MMI developed using 4 taxonomic metrics; MMI_ tax_8: MMI developed using 8 taxonomic metrics; MMI _comb_5: MMI developed using 5 taxonomic and functional metrics; MMI _comb_10: MMI developed using 10 taxonomic and functional metrics).

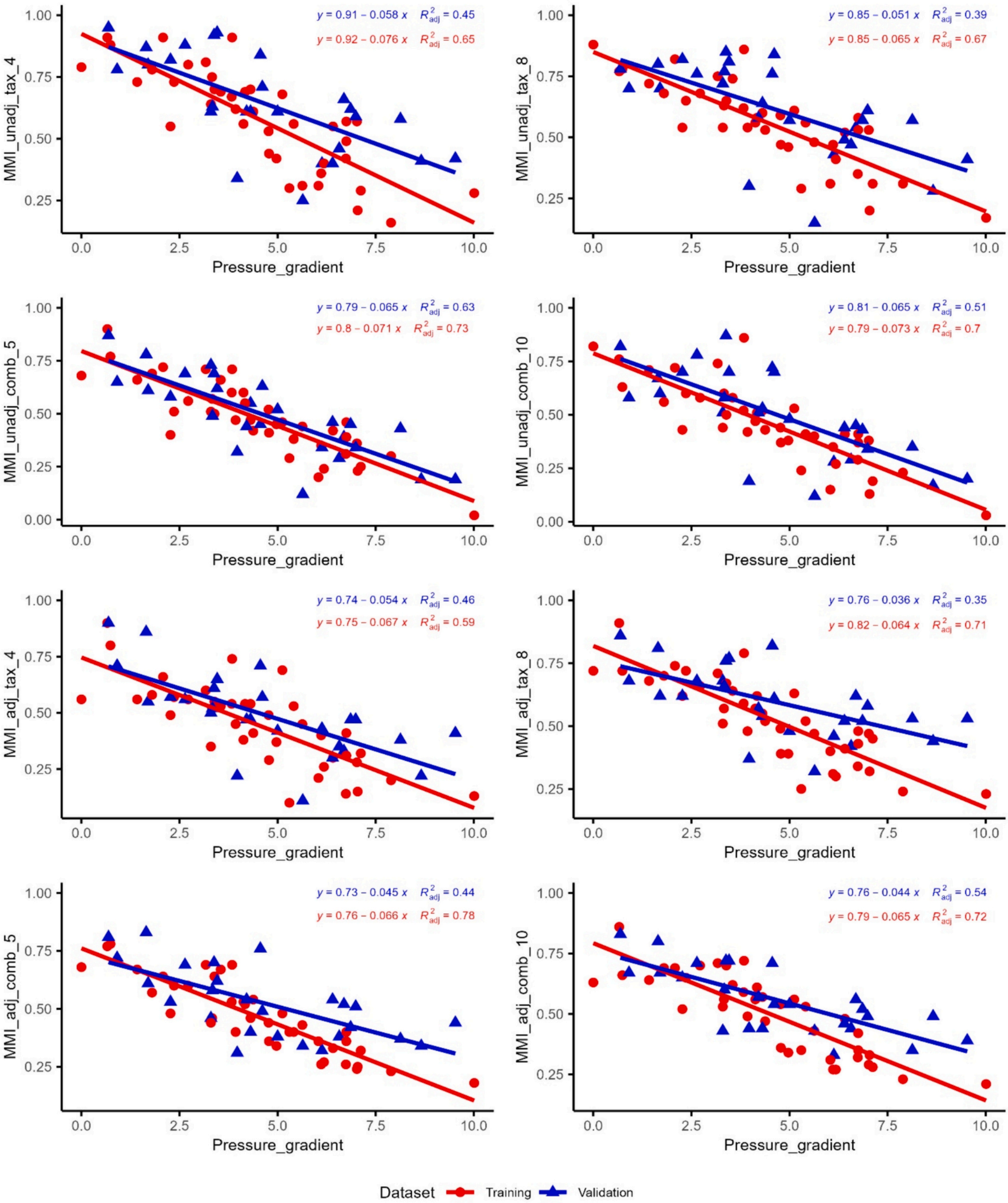


Fig. 6. Linear regression of index performance-driven MMIs scores versus the pressure gradient, for training datasets (solid red circles, solid red lines) and validation datasets (solid blue triangles, solid blue lines). (MMI_unadj_tax_4: MMI developed using 4 unadjusted taxonomic metrics; MMI_unadj_tax_8: MMI developed using 8 unadjusted taxonomic metrics; MMI_adj_tax_4: MMI developed using 4 adjusted taxonomic metrics; MMI_adj_tax_8: MMI developed using 8 adjusted taxonomic metrics; MMI_unadj_comb_5: MMI developed using 5 unadjusted taxonomic and functional metrics; MMI_unadj_comb_10: MMI developed using 10 unadjusted taxonomic and functional metrics; MMI_adj_comb_5: MMI developed using 5 adjusted taxonomic and functional metrics; MMI_adj_comb_10: MMI developed using 10 adjusted taxonomic and functional metrics).

4.1. Performance of taxonomic and combined MMIs

Incorporating functional trait-based metrics increased the performance of both metric performance-driven and index performance-driven MMIs. Probably, the functional metrics included in MMIs explained an important part of macroinvertebrate assemblage variability in response to pressure gradient that is not reflected in taxonomic metrics. Assemblage variability detected by functional metrics may include impacts on food acquisition and locomotion (Mouchet et al., 2010), size structure change (Dulvy et al., 2004), etc. that are only indirectly reflected in taxonomic metrics.

The use of functional traits in biomonitoring approaches can also provide insight into the mechanisms of pressure impacts on stream ecosystems and thereby provide a powerful tool to improve biomonitoring and management lessons derived from it (Culp et al., 2011; Menezes et al., 2010). When exposed to pressures, shifts in the distribution of functional traits can indicate the mechanisms by which the pressures affect communities (e.g., alteration in microhabitat quality), but also which traits respond to the pressures and which infer resistance. For instance, in our study, three modalities of flow current preferences were selected as final metrics in MMIs. Increasing anthropogenic pressure was associated with slow flow velocity, which increased the relative abundance of species that prefer these conditions (i.e., limnophile and rheo- to limnophile species). Inversely, species bound to habitats with high currents (rheobionts) had higher abundance in the least disturbed sites where these conditions were more prevalent. Anthropogenic activities, such as global warming, water withdrawal, land cover alteration, agricultural tile drainage, channel alteration, and impoundments have altered the natural flow regime in the Karun River (United Nations, 2013). The selection of current preference modalities as final metrics may be related to the direct (i.e., exposure to flow) effects of flow velocity reduction in impacted sites or indirect ones (e.g., changes in habitat structure, resource availability, and water temperature regimes; Bruder et al., 2017), jointly favoring species that usually have high relative abundance in low-velocity habitats (Lytle and Poff, 2004).

Another appealing feature of the trait-based metrics compared to taxonomic metrics is the potential of the former to generalize across biogeographic regions because traits are independent of biogeography, whereas taxonomy is not (Brown et al., 2018; Statzner and Bêche, 2010). Furthermore, information on trait composition (but not taxonomy) can be linked to ecosystem functions (e.g., ecosystem processes, food-web dynamics; Gessner and Chauvet, 2002), which are often at the heart of management and conservation efforts (Menezes et al., 2010). However, the functional composition and diversity metrics selected for MMIs in our study differed from those used in other studies (Chen et al., 2019; Pascal et al., 2012). This difference might be due to a combination of *i*) variations in the prevalent types and intensity of anthropogenic pressures present in different ecosystems (Vörösmarty et al., 2010), *ii*) the composition of MMIs, as the inclusion of one metric in MMIs could influence the selection of other metrics (Schoolmaster et al., 2012), and *iii*) inconsistency in single trait-pressure gradient relationships due to correlations among traits, correlations among environmental variables, strength of biotic interactions, and methods used for trait characterization (Hamilton et al., 2020). Differences in MMI composition among studies require an objective metric selection procedure as applied in this study.

Several factors define how well functional diversity indices respond to anthropogenic pressures, including the types of functional traits used to calculate functional diversity (Larson et al., 2021). We calculated five functional diversity indices (functional richness, functional evenness, functional dispersal, functional redundancy, and Rao's entropy) based on life history, ecological, morphological and dispersal activity trait sets. Among the calculated functional diversity indices, functional richness calculated using life history traits decreased along the pressure gradient and was selected for MMIs. The strong relationship of functional richness with the pressure gradient in the Karun might be related to the filtering effects of anthropogenic pressures in moderately and highly impacted sites. Truchy et al. (2022) also found that the presence of multiple anthropogenic pressures restricted the colonization of specific life forms (e.g., taxa that do not have resistance form, and taxa with aquatic larval stages) in impaired habitats, consequently reducing functional richness.

4.2. Adjustment for natural environmental gradients

The natural environment serves as a selective force, filtering species based on their adaptive traits to specific habitat conditions (Southwood, 1977; Townsend et al., 1997). This natural process leads to variations in the structure of biotic communities along the river continuum, from headwaters to estuaries, driven by pronounced gradients in habitat conditions (Vannote et al., 1980). Understanding these ecological dynamics is crucial for accurately assessing anthropogenic impacts on aquatic ecosystems. Our study revealed that the heterogeneity of sites, particularly in climatic and physiographic variables related to altitude, significantly influenced the taxonomic and functional structure of macroinvertebrate assemblages. MMIs adjusted for natural environmental gradients outperformed unadjusted ones, further reinforcing the importance of considering environmental variability in ecological assessments (Chen et al., 2014; Moya et al., 2011; Vander Laan and Hawkins, 2014).

While ecoregions may have advantages as classification units in bioassessment programs and can help mitigate the impact of natural environmental gradients on MMI performance, they may be deemed too general for specific biomonitoring purposes (Hawkins et al., 2000). Ecoregions typically cover large geographical areas with varied environmental conditions and ecological communities, potentially concealing subtle variations in ecological processes and species distributions crucial for certain biomonitoring objectives. Consequently, directly incorporating natural environmental gradients into models to predict local assemblages is considered more effective than relying on landscape classifications (Hawkins et al., 2010), as our comparison has also demonstrated.

4.3. Index performance-driven approach vs. metric performance-driven approach

Subjectivity in assigning metrics to different categories, in the number of metrics used for MMIs, and the choice of methods for considering redundancy between final metrics reduced the consistency of metric performance-driven MMIs (Hering et al., 2006; Stoddard et al., 2008). Moreover, combining the most responsive metrics does not necessarily produce the most responsive MMIs (van Sickle, 2010). We compared the performance of metric performance-driven and index performance-driven MMIs. Taxonomic and combined MMIs developed based on both approaches successfully discriminated least impacted sites from moderately and highly impacted sites (Figs. 4 and 6). However, index performance-driven MMIs indeed outperformed metric performance-driven MMIs in terms of precision, bias, responsiveness and sensitivity (Table 1).

The higher performance of index performance-driven MMIs might be due to the incorporation of metrics that are not individually strongly related to pressure gradients but contribute useful information on macroinvertebrate assemblage responses to the anthropogenic pressure gradients (Bolding et al., 2020). Metrics such as the relative abundance of Pulmonata, relative abundance of tolerant individuals using BMWP TVs, tolerant taxa richness using BMWP TVs, McIntosh index (heterogeneity of a sample in geometric terms; McIntosh, 1967), and relative abundance of Ephemerellidae showed weaker relationships with pressure gradients compared to other metrics selected based on metric performance-driven MMIs' approach. Despite this, they were included as final metrics in the index performance-driven MMIs because their inclusion led to improved overall MMI performance. Ultimately, what managers are most concerned about is the sensitivity of an index, which is a combination of responsiveness and precision. Our findings indicate that in terms of distinguishing between reference sites and impacted sites, index performance-driven MMIs were more effective than metric performance-driven MMIs. Therefore, in certain cases, the index performance-driven approach can offer a more advanced biomonitoring tool compared to traditional metric performance-driven approaches (Bolding et al., 2020).

We also examined the effectiveness of MMIs using a combination of four versus eight metrics. By conducting compound error tests to control for correlations between the final metrics, we anticipated that having more metrics would enhance the sensitivity of MMIs to pressure gradients by amplifying signals from correlated metrics related to pressure gradients and minimizing errors from uncorrelated metrics. However, this was not consistently observed except in the case of MMI_unadj_comb_10 (combining ten metrics), which showed superior sensitivity and responsiveness compared to MMI_unadj_comb_5 (combining five metrics). In other MMIs, the responses of final metrics may reflect a mix of natural gradients and pressure gradients, resulting in intricate responses that either magnify or diminish ecological change signals depending on the specific conditions of application. Therefore, increasing the number of final metrics could potentially heighten the correlated errors within MMIs by emphasizing the influence of natural gradients, consequently leading to a decrease in MMI performance.

All the MMIs developed in this study had a significant relationship with anthropogenic pressure gradients based on the validation dataset. The high consistency in the developed MMIs' response to the pressure gradient based on the validation dataset indicates the high potential of implementing the developed MMIs for tracking impairment in the Karun and similar systems. The relationship between MMIs' scores and the pressure gradient based on the validation dataset was higher for index performance-driven MMIs (R^2 in the range of 0.35 to 0.63) than metric performance-driven MMIs (R^2 in the range of 0.27 to 0.62). This finding was consistent with the MMIs' performance discussed above.

5. Conclusions

In conclusion, our study highlights the value of robust biomonitoring tools to assess and address the multitude of anthropogenic pressures impacting freshwater ecosystems and their communities. We recommend initiating biomonitoring programs using the adjusted taxonomic-functional MMIs (MMI_adj_comb_5) in the Karun basin and elsewhere and suggest further validation and refinement through experiences and data gained in long-term monitoring campaigns. Embracing the index performance-driven approach holds potential for advancing and improving biomonitoring practices, ultimately supporting informed decision-making and sustainable management of freshwater ecosystems. By employing evidence-based approaches and collaborative efforts, we can strive towards healthier and more resilient freshwater environments for current and future generations.

Funding

The research was supported by the Isfahan University of Technology (IUT) and the Swiss Leading House for South Asia and Iran (ZHAW). In addition, the Iranian Ministry of Energy, Iran Water and Power Resources Development Company, and the Department of Environment of Iran provided in-kind contributions during the fieldwork. A.R. Esmaeili Ofogh acknowledges the support received from a Swiss Government Excellence Scholarship during his PhD-exchange, which facilitated data analysis and the writing of this manuscript at the University of Applied Sciences and Arts of Southern Switzerland (SUPSI).

CRediT authorship contribution statement

Ali Reza Esmaeili Ofogh: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Eisa Ebrahimi Dorche:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Sebastian Birk:** Writing – review & editing, Supervision. **Pejman Fathi:** Writing – review & editing, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mojgan Zare Shahraki:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Andreas Bruder:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andreas Bruder reports financial support was provided by Swiss Leading House for South Asia and Iran. Ali Reza Esmaeili Ofogh reports financial support was provided by Swiss Government Excellence Scholarship.

Data availability

Data will be made available on request.

Acknowledgments

We extend our gratitude to S. Asadollah, E. Motaghi, H. Ebrahimi, N. Kheradmand, and H. Fattehi for their invaluable assistance during field and lab work. We also acknowledge the contribution of A. Rezvani for producing the maps used in this research. Their support and dedication were essential to the success of this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.172850.

References

Álvarez-Cabria, M., Barquín, J., Antonio Juanes, J., 2010. Spatial and seasonal variability of macroinvertebrate metrics: do macroinvertebrate communities track river health? Ecol. Indic. 10, 370–379. https://doi.org/10.1016/j.ecolind.2009.06.018.

APHA, 2017. Standard Methods for the Examination of Water and Wastewater, 23rd ed. American Public Health Association. https://doi.org/10.2105/SMWW.2882.219.

Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish, Second. ed. U.S. Environmental Protection Agency; Office of Water, Washington, D.C. EPA 841-B-99-002.

Bêche, L.A., Statzner, B., 2009. Richness gradients of stream invertebrates across the USA: taxonomy- and trait-based approaches. Biodivers. Conserv. 18, 3909–3930. https://doi.org/10.1007/s10531-009-9688-1.

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? Sci. Total Environ. 619–620, 144–154. https://doi.org/10.1016/j.scitotenv.2017.11.022.

Bolding, M.T., Kraft, A.J., Robinson, D.T., Rooney, R.C., 2020. Improvements in multi-metric index development using a whole-index approach. Ecol. Indic. 113 https://doi.org/10.1016/j.ecolind.2020.106191.

Bouchard, J., William, R., 2004. Guide to Aquatic Invertebrates of the Upper Midwest: Identification Manual for Students, Citizen Monitors, and Aquatic Resource Professionals. Water Resources Center, University of Minnesota, St. Paul, Minnesota.

Brown, L.E., Khamis, K., Wilkes, M., Blaen, P., Brittain, J.E., Carrivick, J.L., Fell, S., Friberg, N., Füreder, L., Gislason, G.M., Hainie, S., Hannah, D.M., James, W.H.M., Lencioni, V., Olafsson, J.S., Robinson, C.T., Saltveit, S.J., Thompson, C., Milner, A.M., 2018. Functional diversity and community assembly of river invertebrates show globally consistent responses to decreasing glacier cover. Nat. Ecol. Evol. 2, 325–333. https://doi.org/10.1038/s41559-017-0426-x.

Bruder, A., Salis, R.K., Jones, P.E., Matthaei, C.D., 2017. Biotic interactions modify multiple-stressor effects on juvenile brown trout in an experimental stream food web. Glob. Chang. Biol. 23, 3882–3894. https://doi.org/10.1111/gcb.13696.

Buss, D.F., Carlisle, D.M., Chon, T.S., Culp, J., Harding, J.S., Keizer-Vlek, H.E., Robinson, W.A., Strachan, S., Thirion, C., Hughes, R.M., 2015. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. Environ. Monit. Assess. 187, 4132. https://doi.org/10.1007/s10661-014-4132-8.

Cao, Y., Hawkins, C.P., Olson, J., Kosterman, M.A., 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. J. North Am. Benthol. Soc. 26, 566–585. https://doi.org/10.1899/06-078.1.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., Wardle, D.A., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S., Naeem, S., 2012. Biodiversity loss and its impact on humanity. Nature 486, 59–67. https://doi.org/10.1038/nature11148.

Chen, K., Hughes, R.M., Xu, S., Zhang, J., Cai, D., Wang, B., 2014. Evaluating performance of macroinvertebrate-based adjusted and unadjusted multi-metric indices (MMI) using multi-season and multi-year samples. Ecol. Indic. 36, 142–151. https://doi.org/10.1016/j.ecolind.2013.07.006.

Chen, K., Rajper, A.R., Hughes, R.M., Olson, J.R., Wei, H., Wang, B., 2019. Incorporating functional traits to enhance multimetric index performance and assess land use gradients. Sci. Total Environ. 691, 1005–1015. https://doi.org/10.1016/j.scitotenv.2019.07.047.

Chevenet, F., Dolédec, S., Chessel, D., 1994. A fuzzy coding approach for the analysis of long-term ecological data. Freshw. Biol. 31, 295–309. https://doi.org/10.1111/j.1365-2427.1994.tb01742.x.

Collen, B., Whitton, F., Dyer, E.E., Baillie, J.E.M., Cumberlidge, N., Darwall, W.R.T., Pollock, C., Richman, N.I., Soulsby, A.-M., Böhm, M., 2014. Global patterns of freshwater species diversity, threat and endemism. Glob. Ecol. Biogeogr. 23, 40–51. https://doi.org/10.1111/geb.12096.

Culp, J.M., Armanini, D.G., Dunbar, M.J., Orlofske, J.M., LeRoy Poff, N., Pollard, A.I., Yates, A.G., Hose, G.C., 2011. Incorporating traits in aquatic biomonitoring to enhance causal diagnosis and prediction. Integr. Environ. Assess. Manag. 7, 187–197. https://doi.org/10.1002/ieam.128.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792. https://doi.org/10.1890/07-0539.1.

Darwall, W., Bremerich, V., de Wever, A., Dell, A.I., Freyhof, J., Gessner, M.O., Grossart, H.P., Harrison, I., Irvine, K., Jähnig, S.C., Jeschke, J.M., Lee, J.J., Lu, C., Lewandowska, A.M., Monaghan, M.T., Nejstgaard, J.C., Patricio, H., Schmidt-Kloiber, A., Stuart, S.N., Thieme, M., Tockner, K., Turak, E., Weyl, O., 2018. The Alliance for Freshwater Life: a global call to unite efforts for freshwater biodiversity science and conservation. Aquat. Conserv. 28, 1015–1022. https://doi.org/10.1002/aqc.2958.

Denison, C.D., Scott, M.C., Kubach, K.M., Peoples, B.K., 2021. Integrating regional frameworks and local variability for riverine bioassessment. Environ. Manag. 68, 126–145. https://doi.org/10.1007/s00267-021-01479-6.

Ding, N., Yang, W., Zhou, Y., González-Bergonzoni, I., Zhang, J., Chen, K., Vidal, N., Jeppesen, E., Liu, Z., Wang, B., 2017. Different responses of functional traits and diversity of stream macroinvertebrates to environmental and spatial factors in the Xishuangbanna watershed of the upper Mekong River Basin, China. Sci. Total Environ. 574, 288–299. https://doi.org/10.1016/j.scitotenv.2016.09.053.

Dulvy, N.K., Freckleton, R.P., Polunin, N.V.C., 2004. Coral reef cascades and the indirect effects of predator removal by exploitation. Ecol. Lett. 7, 410–416. https://doi.org/10.1111/j.1461-0248.2004.00593.x.

Elosegi, A., Gessner, M.O., Young, R.G., 2017. River doctors: learning from medicine to improve ecosystem management. Sci. Total Environ. 595, 294–302. https://doi.org/10.1016/j.scitotenv.2017.03.188.

Eriksen, T.E., Brittain, J.E., Søli, G., Jacobsen, D., Goethals, P., Friberg, N., 2021. A global perspective on the application of riverine macroinvertebrates as biological indicators in Africa, South-Central America, Mexico and Southern Asia. Ecol. Indic. 126 https://doi.org/10.1016/j.ecolind.2021.107609.

Esmaeili Ofogh, A.R., Ebrahimi Dorche, E., Birk, S., Bruder, A., 2023. Effect of seasonal variability on the development and application of a novel Multimetric Index based on benthic macroinvertebrate communities – a case study from streams in the Karun river basin (Iran). Ecol. Indic. 146 https://doi.org/10.1016/j.ecolind.2022.109843.

Fathi, P., Ebrahimi Dorche, E., Zare Shahraki, M., Stribling, J., Beyraghdar Kashkooli, O., Esmaeili Ofogh, A., Bruder, A., 2022. Revised Iranian Water Quality Index (RIWQI): a tool for the assessment and management of water quality in Iran. Environ. Monit. Assess. 194, 504. https://doi.org/10.1007/s10661-022-10121-9.

Gaston, K.J., He, F., 2010. Species occurrence and occupancy. In: Magurran, A.E., McGill, B.J. (Eds.), Biological Diversity: Frontiers in Measurement and Assessment. Oxford University Press, Oxford, pp. 141–151.

Gessner, M.O., Chauvet, E., 2002. A case for using litter breakdown to assess functional stream integrity. Ecol. Appl. 12, 498–510. https://doi.org/10.1890/1051-0761(2002)012[0498:ACFULB]2.0.CO;2.

Hamilton, A.T., Schäfer, R.B., Pyne, M.I., Chessman, B., Kakouei, K., Boersma, K.S., Verdonschot, P.F.M., Verdonschot, R.C.M., Mims, M., Khamis, K., Bierwagen, B., Stamp, J., 2020. Limitations of trait-based approaches for stressor assessment: the case of freshwater invertebrates and climate drivers. Glob. Chang. Biol. 26, 364–379. https://doi.org/10.1111/gcb.14846.

Hawkins, C.P., Norris, R.H., Gerritsen, J., Hughes, R.M., Jackson, S.K., Johnson, R.K., Stevenson, R.J., 2000. Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. J. North Am. Benthol. Soc. 19, 541–556.

Hawkins, C.P., Cao, Y., Roper, B., 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshw. Biol. 55, 1066–1085. https://doi.org/10.1111/j.1365-2427.2009.02357.x.

Hering, D., Feld, C.K., Moog, O., Ofenböck, T., 2006. Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. Hydrobiologia 566, 311–324. https://doi.org/10.1007/s10750-006-0087-2.

Horton, R.K., 1965. An index-number system for rating water quality. J. Water Pollut. Control Fed. 37, 300–306.

Juvigny-Khenafou, N.P.D., Piggott, J.J., Atkinson, D., Zhang, Y., Macaulay, S.J., Wu, N., Matthaei, C.D., 2021. Impacts of multiple anthropogenic stressors on stream macroinvertebrate community composition and functional diversity. Ecol. Evol. 11, 133–152. https://doi.org/10.1002/ece3.6979.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P., Kessler, M., 2017. Climatologies at high resolution for the earth's land surface areas. Sci. Data 4. https://doi.org/10.1038/sdata.2017.122.

Kaufmann, P.R., Levine, P., Robison, E.G., Seeliger, C., Peck, D.V., 1999. Quantifying Physical Habitat in Wadeable Streams. United States Environmental Protection Agency (Issue July).

Kriska, G., 2013. Freshwater Invertebrates in Central Europe, a Field Manual, 1st ed. Springer-Verlag Wien. https://doi.org/10.1007/978-3-7091-1547-3.

Laliberté, E., Legendre, P., Shipley, B., 2014. FD: Measuring Functional Diversity From Multiple Traits, and Other Tools for Functional Ecology (R package version 1.0-12).

Larson, E.I., Poff, N.L., Funk, W.C., Harrington, R.A., Kondratieff, B.C., Morton, S.G., Flecker, A.S., 2021. A unifying framework for analyzing temporal changes in functional and taxonomic diversity along disturbance gradients. Ecology 102, e03503. https://doi.org/10.1002/ecy.3503.

Lewis, F., Butler, A., Gilbert, L., 2011. A unified approach to model selection using the likelihood ratio test. Methods Ecol. Evol. 2 (2), 155–162. https://doi.org/10.1111/j.2041-210X.2010.00063.x.

Lytle, D.A., Poff, N.L.R., 2004. Adaptation to natural flow regimes. Trends Ecol. Evol. 19, 94–100. https://doi.org/10.1016/j.tree.2003.10.002.

Mackey, R.L., Currie, D.J., 2001. The diversity-disturbance relationship: is it generally strong and peaked? Ecology 82 (12), 3479. https://doi.org/10.2307/2680166.

Maire, E., Grenouillet, G., Brosse, S., Villéger, S., 2015. How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. Glob. Ecol. Biogeogr. 24, 728–740. https://doi.org/10.1111/geb.12299.

Masese, F.O., Omukoto, J.O., Nyakeya, K., 2013. Biomonitoring as a prerequisite for sustainable water resources: a review of current status, opportunities and challenges to scaling up in East Africa. Ecohydrol. Hydrobiol. 13, 173–191. https://doi.org/10.1016/j.ecohyd.2013.06.004.

Mazor, R.D., Rehn, A.C., Ode, P.R., Engeln, M., Schiff, K.C., Stein, E.D., Gillett, D.J., Herbst, D.B., Hawkins, C.P., 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshw. Sci. 35, 249–271. https://doi.org/10.1086/684130.

McIntosh, R.P., 1967. An index of diversity and the relation of certain concepts to diversity. Ecology 48, 392–404. https://doi.org/10.2307/1932674.

Menezes, S., Baird, D.J., Soares, A.M.V.M., 2010. Beyond taxonomy: a review of macroinvertebrate trait-based community descriptors as tools for freshwater biomonitoring. J. Appl. Ecol. https://doi.org/10.1111/j.1365-2664.2010.01819.x.

Mondy, C.P., Schuwirth, N., 2017. Integrating ecological theories and traits in process-based modeling of macroinvertebrate community dynamics in streams. Ecol. Appl. 27 https://doi.org/10.1002/eap.1530, 1051-0761.

Mouchet, M.A., Villéger, S., Mason, N.W.H., Mouillot, D., 2010. Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. Funct. Ecol. 24, 867–876. https://doi.org/10.1111/j.1365-2435.2010.01695.x.

Moya, N., Hughes, R.M., Domínguez, E., Gibon, F.M., Goitia, E., Oberdorff, T., 2011. Macroinvertebrate-based multimetric predictive models for evaluating the human impact on biotic condition of Bolivian streams. Ecol. Indic. 11, 840–847. https://doi.org/10.1016/j.ecolind.2010.10.012.

Pascal, C., Villeneuve, B., Archaimbault, V., Usseglio-Polatera, P., Mondy, C.P., Villeneuve, B., Archaimbault, V., Usseglio-Polatera, P., 2012. A new macroinvertebrate-based multimetric index (I_2 M_2) to evaluate ecological quality of French wadeable streams fulfilling the WFD demands: a taxonomical and trait approach. Ecol. Indic. 18, 452–467. https://doi.org/10.1016/j.ecolind.2011.12.013.

Paulsen, S.G., Mayio, A., Peck, D.v., Stoddard, J.L., Tarquinio, E., Holdsworth, S.M., van Sickle, J., Yuan, L.L., Hawkins, C.P., Herlihy, A.T., Kaufmann, P.R., Barbour, M.T., Larsen, D.P., Olsen, A.R., 2008. Condition of stream ecosystems in the US: an overview of the first national assessment. J. North Am. Benthol. Soc. 27, 812–821. https://doi.org/10.1899/08-098.1.

Peckarsky, B.L., Fraissinet, R.P., Marjory, P.J.R., Conklin, D.J., 1990. Freshwater Macroinvertebrates of Northeastern North America. Cornell University Press, New York.

Petchey, O.L., Gaston, K.J., 2002. Functional diversity (FD), species richness and community composition. Ecol. Lett. 5, 402–411. https://doi.org/10.1046/j.1461-0248.2002.00339.x.

Poikane, S., Salas Herrero, F., Kelly, M.G., Borja, A., Birk, S., van de Bund, W., 2020. European aquatic ecological assessment methods: a critical review of their sensitivity to key pressures. Sci. Total Environ. 740, 140075 https://doi.org/10.1016/j.scitotenv.2020.140075.

R Development Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T.J., Kidd, K.A., MacCormack, T.J., Olden, J.D., Ormerod, S.J., Smol, J.P., Taylor, W.W., Tockner, K., Vermaire, J.C., Dudgeon, D., Cooke, S.J., 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. Biol. Rev. 94, 849–873. https://doi.org/10.1111/brv.12480.

Rodil, I.F., Lohrer, A.M., Hewitt, J.E., Townsend, M., Thrush, S.F., Carbines, M., 2013. Tracking environmental stress gradients using three biotic integrity indices:

advantages of a locally-developed traits-based approach. Ecol. Indic. 34, 560–570. https://doi.org/10.1016/j.ecolind.2013.06.023.
Ruaro, R., Gubiani, É.A., Hughes, R.M., Mormul, R.P., 2020. Global trends and challenges in multimetric indices of biological condition. Ecol. Indic. 110, 105862 https://doi.org/10.1016/j.ecolind.2019.105862.
Schmidt-Kloiber, A., Hering, D. (Eds.), 2022. www.freshwaterecology.info - The Taxa and Autecology Database for Freshwater Organisms, version 8.0 (accessed on 28.07.2023).
Schoolmaster, D.R., Grace, J.B., Schweiger, E.W., 2012. A general theory of multimetric indices and their properties. Methods Ecol. Evol. 3, 773–781. https://doi.org/10.1111/j.2041-210X.2012.00200.x.
Smeti, E., von Schiller, D., Karaouzas, I., Laschou, S., Vardakas, L., Sabater, S., Tornés, E., Monllor-Alcaraz, L.S., Guillem-Argiles, N., Martinez, E., Barceló, D., López de Alda, M., Kalogianni, E., Elosegi, A., Skoulikidis, N., 2019. Multiple stressor effects on biodiversity and ecosystem functioning in a Mediterranean temporary river. Sci. Total Environ. 647, 1179–1187. https://doi.org/10.1016/j.scitotenv.2018.08.105.
Southwood, T.R.E., 1977. Habitat, the templet for ecological strategies? J. Anim. Ecol. 46, 337–365. https://doi.org/10.2307/3817.
Statzner, B., Bêche, L.A., 2010. Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? Freshw. Biol. 55, 80–119. https://doi.org/10.1111/j.1365-2427.2009.02369.x.
Stoddard, J.L., Herlihy, A.T., Peck, D.V., Hughes, R.M., Whittier, T.R., Tarquinio, E., 2008. A process for creating multimetric indices for large-scale aquatic surveys. J. North Am. Benthol. Soc. 27, 878–891. https://doi.org/10.1899/08-053.1.
Tachet, H., Richoux, P., Bournaud, M., Usseglio-Polatera, P., 2010. Les Invertébrés d'eau douce (NE): Systématique, biologie, écologie. CNRS edition, France.
Tomanova, S., Usseglio-Polatera, P., 2007. Patterns of benthic community traits in neotropical streams: relationship to mesoscale spatial variability. Fundam. Appl. Limnol. 170, 243–255. https://doi.org/10.1127/1863-9135/2007/0170-0243.
Townsend, C.R., Scarsbrook, M.R., Dolédec, S., 1997. Quantifying disturbance in streams: alternative measures of disturbance in relation to macroinvertebrate species traits and species richness. J. North Am. Benthol. Soc. 16, 531–544. https://doi.org/10.2307/1468142.
Truchy, A., Sponseller, R.A., Ecke, F., Angeler, D.G., Kahlert, M., Bundschuh, M., Johnson, R.K., McKie, B.G., 2022. Responses of multiple structural and functional indicators along three contrasting disturbance gradients. Ecol. Indic. 135 https://doi.org/10.1016/j.ecolind.2021.108514.
United Nations, 2013. Inventory of Shared Water Resources in Western Asia. https://doi.org/10.18356/7b8aae96-en.
Usseglio-Polatera, P., Bournaud, M., Richoux, P., Tachet, H., 2000. Biological and ecological traits of benthic freshwater macroinvertebrates: relationships and definition of groups with similar traits. Freshw. Biol. 43, 175–205. https://doi.org/10.1046/j.1365-2427.2000.00535.x.
Vadas, R.L., Hughes, R.M., Bae, Y.J., Baek, M.J., Gonzáles, O.C.B., Callisto, M., Carvalho, D.R. de, Chen, K., Ferreira, M.T., Fierro, P., Harding, J.S., Infante, D.M., Kleynhans, C.J., Macedo, D.R., Martins, I., Silva, N.M., Moya, N., Nichols, S.J., Pompeu, P.S., Ruaro, R., Silva, D.R.O., Stevenson, R.J., Terra, B. de F., Thirion, C., Ticiani, D., Wang, L., Yoder, C.O., 2022. Assemblage-based biomonitoring of freshwater ecosystem health via multimetric indices: a critical review and suggestions for improving their applicability. Water Biol. Secur. 1, 100054 https://doi.org/10.1016/j.watbs.2022.100054.
van den Brink, P.J., Alexander, A.C., Desrosiers, M., Goedkoop, W., Goethals, P.L.M., Liess, M., Dyer, S.D., 2011. Traits-based approaches in bioassessment and ecological risk assessment: strengths, weaknesses, opportunities and threats. Integr. Environ. Assess. Manag. 7, 198–208. https://doi.org/10.1002/ieam.109.
van Sickle, J., 2010. Correlated metrics yield multimetric indices with inferior performance. Trans. Am. Fish. Soc. 139, 1802–1817. https://doi.org/10.1577/t09-204.1.
Vander Laan, J.J., Hawkins, C.P., 2014. Enhancing the performance and interpretation of freshwater biological indices: an application in arid zone streams. Ecol. Indic. 36, 470–482. https://doi.org/10.1016/j.ecolind.2013.09.006.
Vannote, R.L., Minshall, G.W., Cummins, K.W., 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. 37, 130–137. https://doi.org/10.1139/f80-017.
Verberk, W.C.E.P., van Noordwijk, C.G.E., Hildrew, A.G., 2013. Delivering on a promise: integrating species traits to transform descriptive community ecology into a predictive science. Freshw. Sci. 32, 531–547. https://doi.org/10.1899/12-092.1.
Villéger, S., Mason, N.W.H., Mouillot, D., 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology 89, 2290–2301. https://doi.org/10.1890/07-1206.1.
Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E., 2007. Let the concept of trait be functional! Oikos 116, 882–892. https://doi.org/10.1111/j.2007.0030-1299.15559.x.
Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. Nature 467, 555–561. https://doi.org/10.1038/nature09440.
Zare Shahraki, M., Ebrahimi Dorche, E., Fathi, P., Flotemersch, J., Blocksom, K., Stribling, J., Keivany, Y., Beyraghdar Kashkooli, O., Scown, M., Bruder, A., 2021. Defining a disturbance gradient in a middle-eastern river basin. Limnology 91, 125923. https://doi.org/10.1016/j.limno.2021.125923.