Ecological Indicators 187 (2026) 114960

Contents lists available at ScienceDirect

Ecological Indicators

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

Combining machine learning with expert knowledge to classify anthropogenic pressures on Alpine rivers using benthic invertebrates

Francesca Vallefuoco [a,*], Stefano Larsen [b], Pietro Franceschi [b], Valentina Dallafior [c], Walter Bertoldi [d], Guido Zolezzi [d], Maria Cristina Bruno [b]

[a] Institute for Alpine Environment, Eurac Research, Viale Druso 1, 391000 Bolzano, Italy
[b] Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38098 San Michele all'Adige, Italy
[c] Environmental Protection Agency of the Autonomous Province of Trento, Water Quality Unit, Trento, Italy
[d] Department of Civil, Environmental and Mechanical Engineering, University of Trento, I-38123 Trento, Italy

ARTICLE INFO

Keywords:
Macroinvertebrates
Bioindicators
Random forest
Expert knowledge
WFD monitoring
Anthropogenic stressors

ABSTRACT

Streams and rivers are among the most threatened ecosystems globally, as they integrate processes and stressors across local and regional scales. Although research on multiple-stressor effects is growing rapidly, field assessments remain challenging due to complex interactions between natural and anthropogenic drivers. Combining expert knowledge with machine learning can support decision-making for classifying dynamic ecosystems, overcoming data-limitations. We aimed to assess how expert-based classifications of anthropogenic pressures match data-driven classification based on benthic macroinvertebrates taxonomic and functional composition by applying a Random Forest (RF) approach. We used data typically available at Environment Agencies in charge of implementing WFD requirements. Based on stressor-specific indices and local practitioners' expertise, 160 stream sites in Trentino (NE Italy) were classified according to hydrological, morphological, and chemical alterations, including pristine conditions. Most sites were a-priori classified as impacted by one or more stressors and 16% as unimpacted. RF classification matched the expert-based classification only partially, confirming macroinvertebrate sensitivity to hydro-morphological alterations. Functional traits were less informative than taxonomic features in discriminating between water quality alteration and pristine conditions, and in detecting pollution when additional hydrological stressors (e.g., reduced flow or hydropeaking) were present. However, both functional and taxonomic features reliably detected water pollution even where other hydro-morphological alterations occurred. These findings suggest that macroinvertebrate-based indicators commonly used to assess the ecological status of mountain streams capture overall waterbody stress but have limited ability to discriminate among different stressor types. The proposed approach can be replicated and upscaled to broader regions, offering valuable insights for the management of Alpine running waters.

1. Introduction

As freshwater habitats and biodiversity continue to display the most rapid rates of degradation (Tickner et al., 2020), the need for cost-effective and sensitive indicators of environmental stress is more urgent than ever. This task is particularly challenging in running waters where organisms are affected by a range of natural and anthropogenic disturbances that interact over a range of spatial and temporal scales. Although our understanding of the effects of multiple stressors on natural systems has grown (Orr et al., 2024; Birk et al., 2020; Ormerod et al., 2010), most assessments relied on manipulative approaches in controlled environments, which may limit ecological realism (Spears et al., 2021). In addition, most of the currently adopted biological indicators used to classify the ecological integrity of natural systems lack stressors-specific sensitivity (e.g., Friberg, 2014), limiting our ability to identify and disentangle the relative influence of different stressors.

The use of biological indicators has a long tradition in freshwater science and management, where fish and benthic macroinvertebrates have been widely used to define the ecological integrity of water bodies. In Europe, macroinvertebrates represent a key biological quality element used to guide the classification of running waters in line with the Water Framework Directive (WFD 2000/60/EC; European Parliament, 2000). However, benthic organisms are predominantly influenced by water physical-chemical quality, and most derived metrics thus

* Corresponding author.

https://doi.org/10.1016/j.ecolind.2026.114960
Received 13 January 2026; Received in revised form 30 April 2026; Accepted 11 May 2026

reflect their specific sensitivity to organic pollution (Friberg, 2014). This issue is particularly relevant for Alpine streams and rivers, where anthropogenic modification often results in drastic changes in the hydrology and channel morphology, without necessarily affecting water quality (Comiti, 2012; Bruno et al., 2010). The uncritical use of biological indicators to guide ecological classifications and identify threats can lead to ill-advised management actions, as emphasised in studies conducted in the Alpine environment (Larsen et al., 2019a; Laini et al., 2018; Quadroni et al., 2017), despite improvements in sampling methods have been recently proposed to assess the effect of stressor removal through river restoration (Devreux et al., 2025).

More effort is needed to examine the combined and individual effects of hydrological and morphological alterations on Alpine streams, and the extent to which water quality influences the identification of potential bioindicators to complement currently adopted metrics. However, quantitative information about the degree to which individual river reaches are affected by specific stressors is often scarce, limiting the ability to model biotic responses. On this theme, the importance of local expert knowledge has gained increasing recognition in applied ecology as a valuable complement to data-based approaches and a resource in decision-making processes (Bélisle et al., 2018; Martin et al., 2012; Perera et al., 2012), which can also overcome typical resources or data limitations (Hemming et al., 2018). In evaluating ecological conditions of habitats, local environmental practitioners can often complement quantitative data with qualitative estimates of the types and levels of stressor affecting ecosystems (Drescher et al., 2013). In fact, despite an increasing number of studies addressing multiple stressor effects in aquatic systems (Lemm et al., 2021; Birk et al., 2020; Ferreira et al., 2019), assessments of stressor-specific effects under complex, multi-pressure conditions at the reach scale remain limited also due to data availability constraints (He et al., 2023; Spears et al., 2021). To address this gap, in the specific case of stream ecosystems, we argue that combining expert-based classification of ecological status with a data-driven approach can help identify aspects of community structure or function that are associated with specific stressors and multiple disturbances that can support ecological assessment. In this context, machine learning methods and classification trees can provide an effective means to model complex ecological interactions, enhancing the detection of stressor-specific responses and identifying the main features involved (Cutler et al., 2007). In fact, when it comes to data-driven classification approaches, machine learning methods are particularly useful in ecological research, being able to model non-linear relationships in high-dimensional data with often correlated variables. These methods are increasingly adopted in applied ecology and in river research (Ho and Goethals, 2022), for example to model habitat suitability (e.g., Theodoropoulos et al., 2018; Vezza et al., 2014; Poff and Zimmerman, 2010; Özesmi et al., 2006;), water quality (e.g., Politikos et al., 2024; Zanoni et al., 2022; Singh et al., 2009), and patterning abiotic-biotic relationships (e.g., Yao et al., 2026; Krtolica et al., 2025; Gabriels et al., 2007; Chon et al., 2001; Lek et al., 1996).

The main goal of this study is to evaluate whether a-priori expert-based stressors classification of Alpine river reaches matched a-posteriori classification based on benthic communities using a random forest (RF) approach. Specifically, we aimed to assess the capacity of macroinvertebrate taxonomic and functional composition to discriminate between different types and combinations of anthropogenic pressures by testing several binary classifications in order to: i) determine which type of stressor (or combination thereof), including water quality, hydrological and morphological alterations, can be reliably detected based on macroinvertebrate assemblages; ii) identify which taxonomic and functional features were more informative for the classification process.

To address those objectives, six pairwise comparisons were designed to test whether macroinvertebrate composition could be used to detect: 1) any type of alteration, independently of their nature; 2) hydromorphological alteration; 3) water quality alteration (i.e. pollution); 4) the effects of water pollution in reaches that are also affected by both hydrologic and morphologic alteration; 5) the effects of water pollution on reaches that are affected only by hydrologic alteration; 6) the effects of water pollution on reaches affected by morphologic alterations. These specific comparisons were examined as they represent typical combinations of stressors in Alpine streams (Ferreira et al., 2019; Chiogna et al., 2016; Schinegger et al., 2012). Our study takes an original approach whereby we use a large suite of taxonomic (abundances of taxa) and functional (ecological traits) features to predict a-priori classified river reaches based on local expert knowledge and qualitative physico-chemical and hydro-morphological indicators.

2. Methods

2.1. Study area

The study area covers the main river catchments of Trentino Province (Fig. 1), North-Eastern Italy, characterised by Alpine and mountain territory with glacial U-shaped valleys and a mix of sedimentary carbonate (mainly dolomite) and metamorphic and igneous rocks (gneiss). Only 24% of the total 6200 km^2 surface of the region lays below 1000 m a.s.l., with altitudes ranging from 65 m at Garda Lake to 3767 m at Mount Cevedale. Due to this elevation gradient, rivers display torrential regime, strongly oxygenated waters and generally low temperatures (<10 °C). The main water use is for hydroelectric production (91.2%), where the withdrawn water is entirely returned after use, without consumption, followed by aquaculture (3.4%), agricultural use (3.1%) and civil use (1.6%) (RSA, 2020). As a consequence of the strong hydroelectric water use and the presence of a large number of reservoirs (i.e., 96 dams with reservoir volumes greater than 5000 cubic metres and 23 large hydropower plants, not including the small, run-of-the-river hydropower plants across the territory), the hydrological regime of most rivers is altered (Larsen et al., 2019a; Zolezzi et al., 2009). Contrary to the widespread hydro-morphological alterations, most rivers display good water quality, with only 4 out of 412 monitored rivers not reaching good chemical status according to the WFD (RSA, 2020).

This study included 160 sampling stations distributed across 90 streams monitored by the Environmental Protection Agency of the Autonomous Province of Trento (APPA), covering an elevational range from 65 to 1720 m a.s.l. (Fig. 1). Most of the sampling monitoring sites (i.e., 54 monitoring stations) refer to the Trentino section of the Adige River, the second-longest Italian river (catchment in Trentino: 948 km^2), and its main tributaries: the Noce River (1367 km^2, 19 stations) on the right, the Avisio River (940 km^2, 19 stations) and the Fersina River (170 km^2, 11 stations) on the left; other rivers are the Sarca River (1268 km^2 whole catchment area included in the Trentino Province, with 31 stations) and the Trentino section of the Brenta River (618 km^2, with 15 stations). A few stations were located in the small Trentino section of the Chiese River (409 km^2, 6 stations), Vanoi River (237 km^2, 3 stations), Cismon River (209 km^2, 1 station) and Astico River (80 km^2, 1 station). Each sampling site was characterised according to stream typology, distance from the source, Strahler number, and hydroecoregion type (Table 1).

2.2. Macroinvertebrate sampling and processing

Macroinvertebrates were collected by the APPA as part of their institutional monitoring programs at fixed monitoring stations. For this study, we selected only macroinvertebrate samples collected for at least three seasons, to account for seasonal variability. The resulting dataset included 160 stations (Fig. 1), each surveyed one to three times across different seasons over a 10-year period (i.e., from 2009 to 2019), generally at intervals of 3–6 years, resulting in sparse temporal replication. Overall, 1049 macroinvertebrate samples were included. Macroinvertebrate sampling followed the multi-habitat proportional approach defined in the AQEM protocol (Hering et al., 2004) and in the Italian ISPRA Manual n.111/2014 (ISPRA, 2014): a total area of 1 m^2

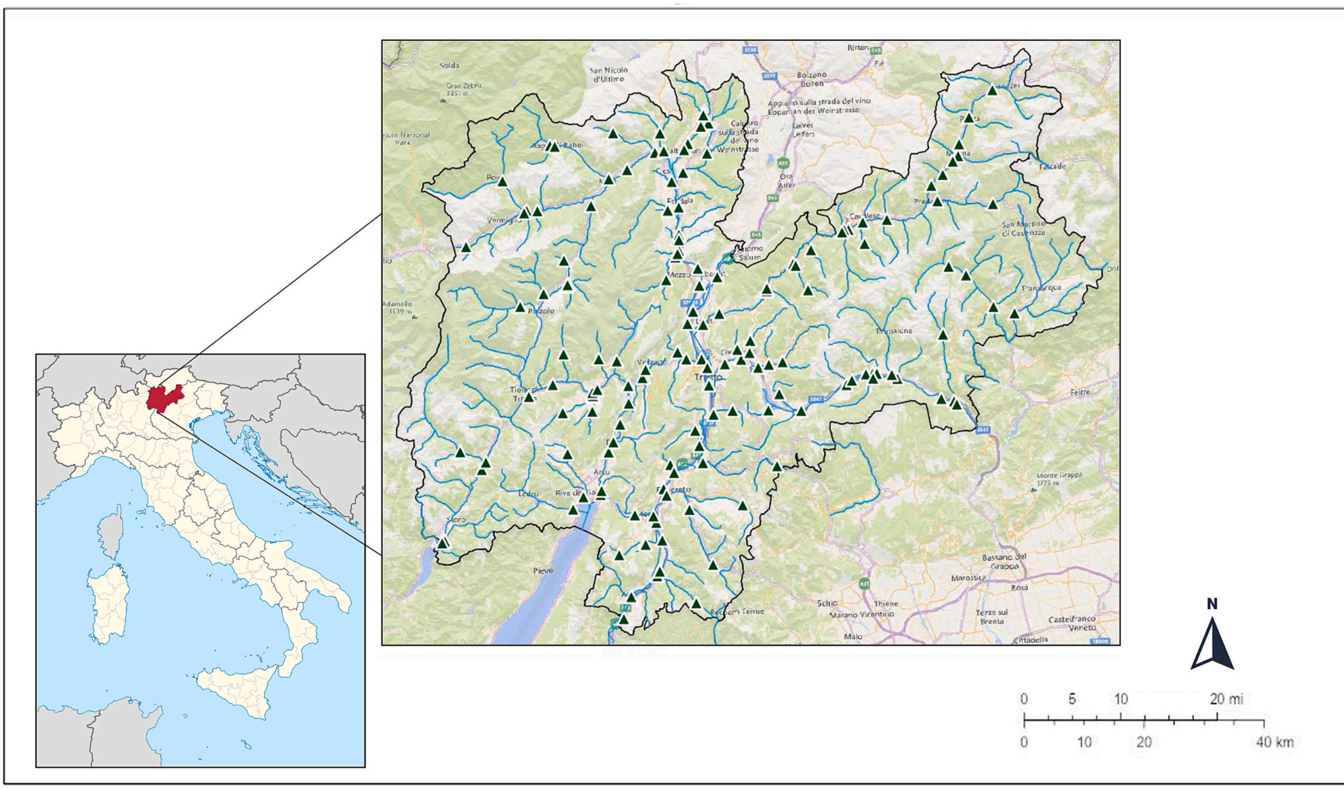

Fig. 1. Map of the main river networks of Trentino Province (Italy) and the 160 sampling sites monitored by APPA included in the analysis.

was sampled by collecting 20 Surber (23 × 22 cm, mesh size 500 µm) replicates within a 20–50 m reach, distributed across the different riverbed microhabitats in proportion to their relative occurrence and substrate diversity. Ephemeroptera and Plecoptera were identified at the genus level, while the other taxa were identified at the family level. Only taxa observed in more than 10 samples were included in the analyses (see Table S1). The final database included 80 taxa for a total of 1,997,621 specimens.

2.3. Functional traits selection and processing

Trait information was retrieved from www.freshwaterecology.info (Schmidt-Kloiber and Hering, 2015) and refers to regional preference (i.e., zonal and altitude preference), habitat preference (e.g., flow, temperature and microhabitat/substrate preference) and other life-history characteristics (e.g., feeding type, maximal size, locomotion type, longevity, life cycle, respiration, reproductive strategy), totalling 13 traits and 67 modalities (Table S2) available for 65 out of the 80 sampled taxa. The association scores of each trait modalities were standardised between 0 and 1 so that the sum for each trait was equal to 1 (Darr et al., 2014). To derive community-level trait proportions, affinities were weighted by species abundance using the functcomp function in FD package (Laliberté et al., 2014).

2.4. Anthropogenic stressors classification

To classify each sampling station according to the presence of specific anthropogenic alterations, we combined qualitative indicators (where available) and the expert opinion of personnels of the Water Quality Unit of APPA in charge of assessing the Ecological Status in the context of the WFD 2000/60/EC. We focused on alteration of natural river hydrology (hydropeaking and/or minimum vital flow conditions), morphology (presence of artificial embankments, weirs, dam or channelization) and water quality (physico-chemical parameters), and their interaction. The LIMECO index (*Livello di Inquinamento dai Macrodescrittori per lo stato ecologico;* according to the Italian law D.M. 260/2010*)*, a multi-metric indicator of water pollution based on threshold levels for the concentration of oxygen, ammonia, nitrate and total phosphorus (Azzellino et al., 2015), was used to identify water quality alterations (hereafter referred as pollution). The MQI (Morphological Quality Index, Rinaldi et al., 2013), which estimates stream morphological alteration with respect to reference state, was used to identify morphological alterations. Both indices range over five quality classes (high, good, moderate, poor and bad) and reaches with values equal to or lower than moderate (i.e., moderate, poor and bad) were identified as polluted and/or morphologically altered. Finally, the IARI index (*Indice di Alterazione del Regime Idrologico*, Bussettini et al., 2011), largely inspired to the well-known *Indicators of Hydrologic Alteration* (Richter et al., 1996) estimates three levels of flow regime conditions (i.e., high, good and bad) based on their alteration relative to reference conditions, using streamflow series on a daily or monthly scale. Reaches with ‘bad’ flow regimes were classified as hydrologically altered.

These quality indicators were available for a subset of the streams and refer to spatial river units that typically coincide with morphologically and hydrologically homogeneous reaches, where the macroinvertebrate sampling stations were located. These reaches often coincided with those portions of rivers formally classified as a “water body” within the application of the WFD requirements at a national level. To complement and further validate the classification of these stream reaches based on the upstream or local presence of the key stressors, we involved local environmental practitioners from the APPA. Anthropogenic impacts at reaches lacking complete qualitative indicators were evaluated using practitioners' knowledge of local and upstream conditions, considering pressures and impacts at the water body scale according to WFD. Each stream reach was classified as pristine or affected by one or more stressors as described in Table 2. For example, the presence of sewage treatment plants, high levels of *E. coli* bacteria, fish aquaculture, intensive agriculture or farms in the

Table 1
Environmental variables recorded at each sampling site and included as predictors (i.e., features) in the RF classifier models.

Predictor variables	Code	Definition	% Sites within each predictor
TYPES OF ALPINE STREAMS	SS	rhithral (headwaters-fed mainly by rain and snowmelt runoff)	91
	GH	kryal (streams-fed by glacial meltwater)	3
	SR	krenal (groundwater-fed streams, mountain areas or high elevation)	2
	AS	spring (groundwater-fed streams, low elevations, piedmont or valleys)	1
	IN	Intermittent river	3
DISTANCE FROM THE SOURCE	1	<5 km	16
	2	5–25 km	58
	5	25–75 km	17
	4	75–150 km	3
	5	>150 km	3
	6	Intermittent meandering rivers (sinuous or confined)	3
SEASON	Winter		25
	Spring		31
	Summer		27
	Autumn		17
ELEVATION RANGE	<400 m		44.5
	400-800 m		31
	>800 m		24.5
STRAHLER NUMBER	1		44.5
	2		33
	3		15
	4		2.5
	5		5
HYDROECOREGION TYPE	02	Southern Pre-Alps and Dolomites; calcareous mountains with massive and carbonated rocks	64
	03	Inner Alps; alpine high mountains with crystalline mountains	36

Table 2
List of pressure type classification and the corresponding number of APPA macroinvertebrate monitoring stations and records associated with each pressure type.

PRESSURE TYPE (Code)	Definition of pressure types	N stations	N records
N	Pristine; non altered	27	181
H	Unpolluted but hydrologically altered	15	90
M	Unpolluted but morphologically altered	36	235
HM	Unpolluted but hydro-morphologically altered	28	183
P	Polluted but not hydro-morphologically altered	13	67
PH	Polluted AND hydrologically altered	6	48
PM	Polluted AND morphological altered	18	119
PHM	Polluted AND hydro-morphologically altered	19	126
	Total number	162	1049

surrounding area of the sampling site were examined to assess water quality and, if deemed appropriate, were considered as polluted sites. Similarly, for hydrological or morphological stressors, we recorded the upstream presence of hydroelectric plants, embankments or artificial barriers as reported by the APPA experts.

2.5. Random forest (RF) analysis

To discriminate among stream sites affected by specific or combined stressor scenarios based on the benthos community structure, independent RF models were developed using macroinvertebrate taxa abundances (measured as number of individuals per m^2) and functional composition (as community weighted means, i.e., the proportion of each trait category weighted by the relative abundance of taxa) as predictor variables. In keeping in line with the RF literature (e.g., Sahli, 2020; Buskirk et al., 2018), we refer to the predictor variables as “features”. In both cases, a set of key environmental variables were also included as predictor features (Table 1). These variables describe characteristics related to elevation, stream gradient, stream order and season, and were included not only because they directly and indirectly influence benthic organisms, but also because they often covary with the key stressors investigated. For example, lower distance from the source and steep gradient typical of headwater reaches, may be associated with good water quality but also presence of dams (Buchanan et al., 2022).

The six pairwise stressor comparisons tested are listed in Table 3; these comparisons were designed to assess the models' ability to discriminate between specific stressors (or lack thereof) and identify the predictor features contributing the most. The first binary classification compared altered vs non-altered sites (ALL-N), where the first group included all altered stream reaches affected by one or a combination of stressors and the second group included only pristine sites. This provided an overall test for identifying the presence of any type of stressors, regardless of type or combination. The second (HM-N) and third (P-N) comparison tested the ability to respectively identify the presence of hydro-morphological and water quality alterations (pollution), as compared to non-impacted reaches. The last classifications aimed at assessing the ability of the model to identify water quality alterations (pollution) in reaches also affected by other stressors, including both hydro- and morphological (PHM-HM), only hydrological (PH-H) and only morphological (PM-M) alterations.

For each binary classification, the Random Forest workflow included 100 repeated random training-test splits to build the RF classifier, accuracy evaluation against null models, optimisation of model parameters, assessment of classification significance, implementation of balanced subsets (model-b), and extraction of variable importance. The median prediction accuracy, along with the 1st and 99th percentiles, was computed to describe the variability across the 100 models. Prediction accuracy was calculated also for a set of ‘null’ models based on a randomised labelling approach, in which group identities were shuffled across sites before each training-test split, while preserving the original number of observations per category (i.e. the unbalanced design),

Table 3
List of pairwise comparisons tested (pressure type code defined in Table 2). Alteration (ALL) = all combinations of stressors (i.e., P, PH, PM, HM, PHM, H, M). Model-a includes all records, while model-b only watershed-balanced data.

Binary classification		code	N cases Model-a	N cases Model-b
1	Alteration vs No alteration (ALL-N)	ALL	868	584
		N	181	158
2	HyMo vs No alteration (HM-N)	HM	183	183
		N	181	170
3	Polluted vs No alteration (P-N)	P	67	67
		N	181	144
4	Polluted & HyMo vs HyMo (PHM-HM)	PHM	126	126
		HM	183	144
5	Polluted & Hydro vs Hydro (PH-H)	PH	48	37
		H	90	48
6	Polluted & Morpho vs Morpho (PM-M)	PM	119	101
		M	235	114

resulting in 100 independent null models for each pairwise comparison. Thus, 'null' RF classifications were used to estimate the distribution of predictive accuracy values expected by chance, accounting for differences in sample size across categories and ensuring comparability with the RF classification models trained on the real labels (i.e., the actual stressor classification, considered as 'real' models). RF analysis was performed using the tuneRF function from the randomForest package (Liaw and Wiener, 2002) in R statistical software (R Development Core Team, version 4.3.0). In the classification algorithms, we included the optimisation of the mtry parameter (i.e., the number of variables randomly sampled as candidates at each split), selecting for each iteration the value with the minimum stabilised estimate of error rate, while keeping the number of trees fixed at 500 (Oshiro et al., 2012). For each classification, model performance was considered significantly accurate if the prediction accuracy intervals from the real models did not overlap with the confidence interval of the null models' accuracy distribution.

The use of 'null' model approach, whereby we shuffled the observed sample labels, also controlled for any inherent bias to data unbalance (i.e. the null models were equally unbalanced). However, to further minimise the influence of the unbalanced number of observations between groups in some pairwise binary classifications (e.g., many observations from watersheds with altered reaches and few in pristine conditions), all RF models were repeated with a specific subset of observations. This more balanced RF is referred to as model-b and contrasts with model-a, which retains the original number of observations derived from the combination of presence and absence of the stressors of interest (Table 3). Specifically, in model-b we only included samples from watersheds where both categories of the binary classification were present and represented more than 10% of all records in the given watershed. In other words, to limit bias associated with watershed identity, we excluded cases where a given watershed only contributed to one of the two categories included in the RF classifier. In addition, during the implementation of model-b tests, we measured the importance assigned to each variable by the RF through the Mean Decrease Gini (MDG) index, which describes the quality of the split of a node on a variable, averaged over all trees (Han et al., 2016). To identify the most relevant environmental and taxonomic/functional features driving RF classification, the first 15 most important features were extrapolated, reporting their ranking over all 100 models and their association with specific pressures based on their mean relative values (i.e., abundance for taxa and traits, and actual values for environmental features), calculated as the relative difference between group means and normalised by the maximum value.

Finally, to allow comparison of the predictive ability of different models, and to estimate which stressors were better identified, we derived a non-parametric effect size. Specifically, we did not focus on the absolute model accuracy, but rather on comparing the observed model predictions relative to its 'null' control (randomised) model. We used an adjusted effect size (d) that accounts for differences in the number of observations in each model; the non-parametric effect size was calculated as the median difference between the real and null models, standardised by the pooled Median Absolute Deviation (MAD) as follow:

$$PMADxy = \sqrt{\frac{(n_x - 1)MAD_y{}^2 + (n_y - 1)MAD_y{}^2}{n_x + n_y - 2}} \quad (1)$$

$$d = \frac{My - Mx}{PMADxy} \quad (2)$$

where n is the number of observations and x and y refers to the real and null labelling models (M) respectively. This non-parametric effect size specifically quantifies the model's performance relative to their null predictions, thus accounting for differences in the number of observations and unbalancing in the models.

3. Results

3.1. Model performance

The results of all Random Forest predictions, including the median and the 1st and 99th percentiles of model accuracies, are shown in Fig. 2. For each pairwise comparison, we assumed that the classifier model had a good performance when the 1st–99th percentile interval of the real (observed) model accuracy did not overlap with that derived from the null model. In a perfectly balanced comparison, prediction accuracy from the null, random labelling models should centre around 50%.

Results from each binary classification are presented in detail below:

1) ALL-N. The RF models significantly discriminated pristine sites from those affected by any stressors (regardless of type or combination). This test included a notable imbalance in group representation (i.e., altered reaches were more frequent than pristine ones), which was evident in the performance of the null models (i.e., the median prediction accuracy under random labelling was approximately 80.1%, with values ranging from 72.0% to 86.3%, 1st to 99th percentile, respectively). Predictions were accurate in both model-a and the more balanced model-b, although this was true only with models including taxonomic features. Predictions from models with functional features were accurate only in the case of the more balanced model-b.
2) HM-N. Reaches specifically affected by hydro-morphological alterations were consistently well discriminated from pristine sites, in both balanced and unbalanced models, and including both taxonomic and functional features. In this comparison, model-b was well balanced with null accuracy values around 50%.
3) P-N. The discrimination between pristine, non-altered reaches and those exclusively affected by water pollution was satisfactory only for the balanced model-b that included macroinvertebrate taxonomic features. Accuracy of RF models with functional features were not significantly better than null models.
4) PHM-HM. This binary model tested the ability to discriminate reaches affected by a combination of hydro-morphological and water quality alterations from those 'only' affected by hydro-morphological alterations. In other words, it examined if water quality alterations could be identified on top of other stressors. The accuracy of RF models was consistently better than null predictions, for both taxonomic and functional features.
5) PH-H. Among reaches with hydrologic alterations, the RF model was able to significantly identify those also affected by water quality alteration but only considering taxonomic features.
6) PM-M. Among reaches with morphological alterations, the RF model successfully identified those also affected by water quality alteration. Predictions were significantly accurate only considering the balanced model-b, with both taxonomic and functional features.

3.2. Variable predictor importance

Variable importance for the classification process was reported only for model-b (Fig. 3), as this was the model with the best performance in most comparisons. Overall, our findings indicate that environmental variables such as Strahler order, elevation range and distance from the source had a strong influence in discriminating between types of anthropogenic alterations. Among macroinvertebrate taxa, the most frequently important predictors included the Ephemeroptera genus *Baetis*, the Plecoptera genera *Nemoura* and *Protonemura*, and the Diptera families Chironomidae and Athericidae. Additionally, habitat parameters such as zonation, substrate type, and temperature preference, as well as biological traits related to locomotion and feeding habits, were identified as relevant in the classification process. The top predictor variables identified by the RF algorithm for discriminating between each pairwise classification are discussed below:

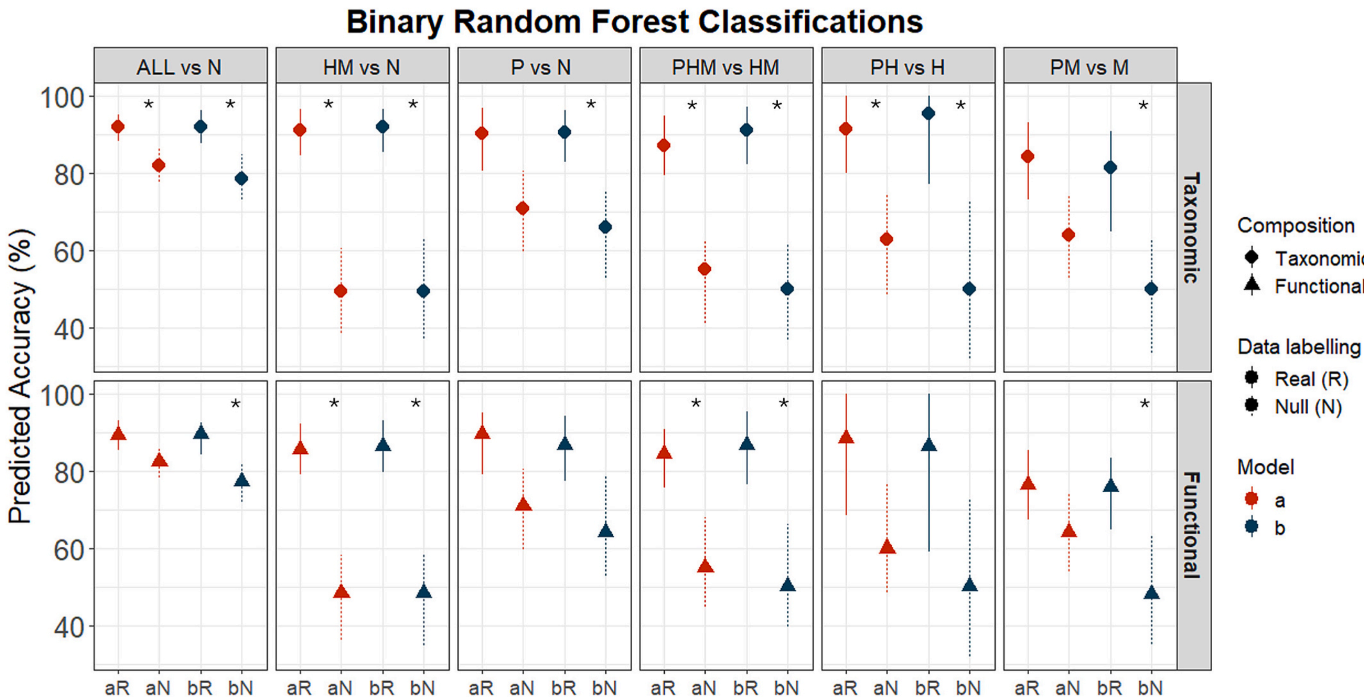

Fig. 2. Median and interval accuracies ranging from the 1st to the 99th percentile, based on predictions from 100 RF models built for each binary classification (acronyms defined in Table 3); * indicates intervals that do not overlap.

1) ALL-N: the Coleoptera Hydraenidae and the Plecoptera *Protonemura* sp. and *Isoperla* sp. were the main taxa associated with pristine sites, whereas *Baetis* sp. and Chironomidae were generally more abundant in the altered sites, typically located farther from the source compared to the pristine sites. In functional composition-based models, sites were more likely to be classified as altered when located at lower elevations in the epipotamal barbel region, characterised by high discharges, wide channels, and low slopes. We observed that gatherers/collectors, crawling or flying taxa (i.e., locomotion modality = other), preferring algal substrates and warm temperature range (>18 °C) were key predictors of altered sites. Conversely, pristine sites were predominantly associated with the uppermost stream sections, typified by fast-flowing, well-oxygenated waters and coarse substrates typical of upper-trout/epirhithral reaches.
2) HM-N: Hydraenidae, the Diptera Athericidae and *Protonemura* sp. were top indicators of pristine sites (where they were more abundant) relative to hydro-morphologically altered ones. Pristine reaches also hosted predators, cold stenothermic taxa, preferring a small cold temperature range (below 10 °C), and taxa living in the upper-trout region with moderate to high current. Low elevation, wide rivers and greater distance from the source were consistently identified as key environmental features in both taxonomic and function-based classifications for detecting hydro-morphological alterations. The mayfly genera *Baetis* and *Serratella*, and the families Ancylidae (Gastropoda) and Erpobdellidae (Hirudinea), as well as gatherers and eurytherm taxa were more abundant in the HM sites.
3) P-N: In addition to the taxa associated with non-altered sites identified in the previous pairwise classifications, the Ephemeroptera *Rhithrogena* sp. was also among the most important taxa for discriminating water-polluted from pristine sites. Conversely, polluted sites were characterised by the Diptera Simuliidae and Chironomidae, *Baetis* sp., the Trichoptera Hydropsychidae, the Oligochaeta Naididae and Lumbriculidae. Stream Strahler order was also ranked as a key predictor. In the functional-based classification, although not statistically significant, polluted sites tended to support warm stenothermic, sessile, and bivoltine taxa adapted to lowland waters.
4) PHM-HM: Pollution in hydro-morphologically altered sites was detected using stream order, elevation gradient, and the presence of Crustacea Asellidae, Gammaridae, Naididae, and *Baetis* sp. The functional-based classification indicated that taxa with a preference for fine to medium gravel substrates, cold temperature ranges, moving as sprawlers, and feeding as grazers or scrapers were mostly associated with HM sites. PHM-altered sites were dominated by taxa adapted to lower elevations, warmer longitudinal zonation, and swimming locomotion.
5) PH-H: pollution in hydrologically altered sites was detected reliably only in taxonomic-based models, which identified the Ephemeroptera *Habroleptoides*, *Baetis* sp., Simuliidae, and Coleoptera Elmidae as important predictors. Simuliidae and Lumbricidae (Oligochaeta) families were more abundant in PH-altered sites, typically associated with lower elevations and greater distance from the source. Regarding stream typology within this subset of data (i.e., rhithral (SS) and kryal (GH) waters) glacier-fed streams (GH) were not associated with the PH category, as reflected both in the original dataset and in the model predictions.
6) PM-M: environmental features such as distance from the source, Strahler and stream type were the most relevant in discriminating water pollution among morphologically altered sites, using both taxonomic and functional data. In particular, the sampling sites included in this comparison were only glacier-fed (GH) and intermittent (IN) streams, which were found to be exclusively associated with PM and M respectively. Gammaridae, *Baetis* sp. and Elmidae, and eucrenal (spring region), metapotamal (bream region) and shredder organisms were identified as the most relevant features for taxonomic and functional models. *Baetis* genera, Hydropsychidae, Simulidae and Chironomidare, taxa preferring potamal regions and bivoltine taxa were more frequent in PM-altered sites.

3.3. Non-parametric effect size

To quantitatively compare the accuracy of the different RF models,

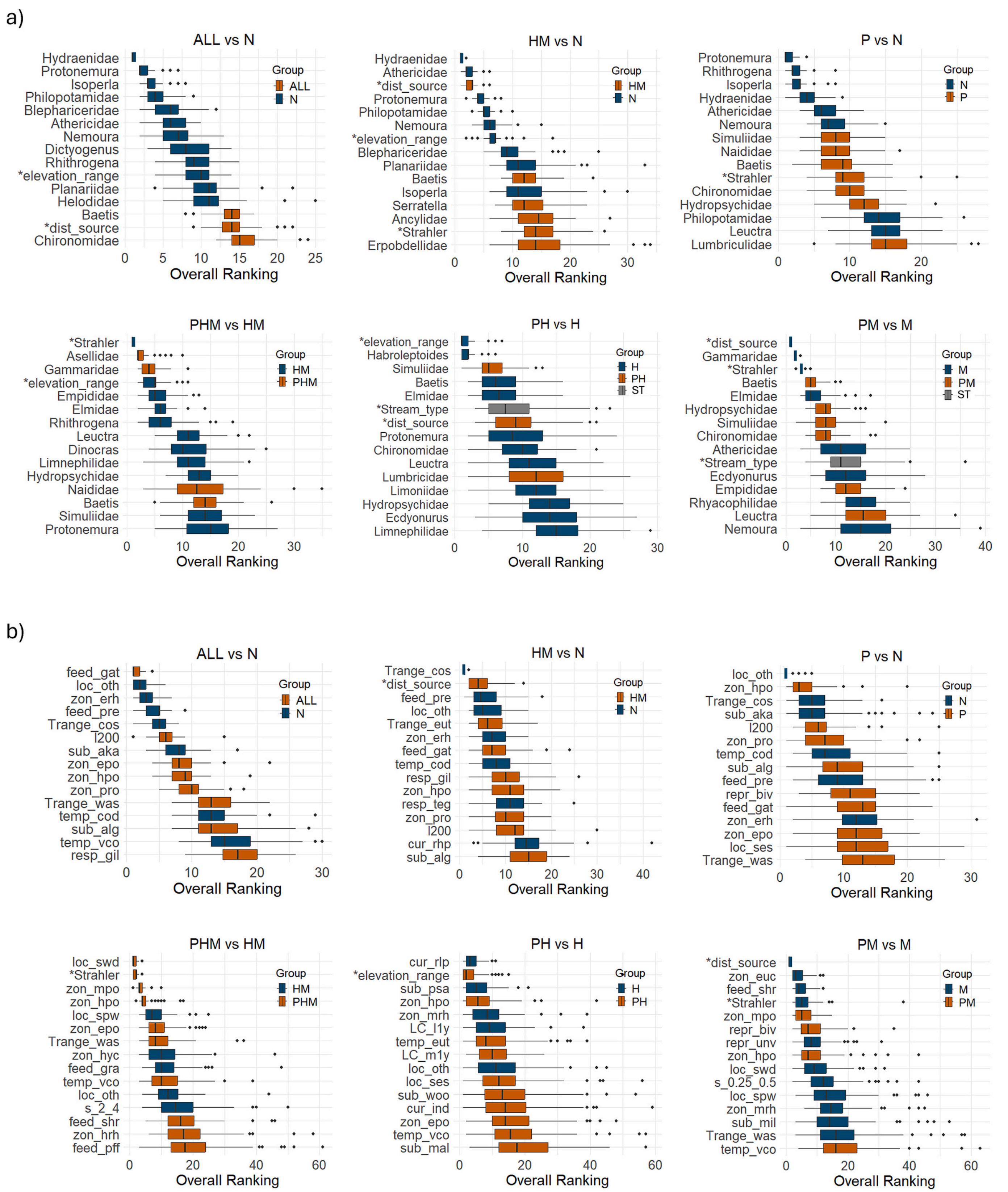

Fig. 3. Top 15 feature importance scores from the RF classifier trained on a more balanced watershed dataset (model-b) for each variable across the six binary classifications (acronyms defined in Table 3): a) Taxonomic composition; b) Functional composition. ST = stream type; stream-type associations for each pressure group are described in Section 3.2.

we used the non-parametric effect size (Fig. 4). This approach was also necessary because some binary tests were unbalanced with null predictions reaching ~80% accuracy. In general, more balanced models tended to perform better than the unbalanced counterpart. This was the case for all models based on macroinvertebrate functional traits composition, and four out of six models based on taxonomic features. Overall, models based on taxonomic composition displayed larger effect sizes than those based on functional traits.

RF models performed particularly well (largest effect size) in detecting hydro-morphological alterations (compared to pristine conditions; HM-N), especially when modelling taxa relative abundances. Large effect sizes were also associated with detection of water pollution among reaches already impacted by hydro-morphological alterations (PHM-HM).

4. Discussion

We sought to understand the degree to which macroinvertebrate composition data (taxonomic abundance and trait proportions) could be used within a RF approach to predict the presence and combination of multiple anthropogenic alterations in a large Alpine catchment. Recent studies have demonstrated the value of RF and other machine-learning classifiers for river ecological assessment, either by improving classification accuracy and identifying key environmental drivers (e.g., Martyszunis et al., 2024; Arrighi and Castelli, 2023), or by focusing on stressor-specific macroinvertebrate responses (e.g., Yao et al., 2026; Krtolica et al., 2025). Building on these approaches, our study reveals that RF models can effectively discriminate stressor types at the reach scale by integrating a priori expert-based assignment of single and combined anthropogenic pressures with environmental indicators of hydro-morphological integrity, physico-chemical water quality, and macroinvertebrate community composition. This approach not only allowed overcoming data limitation on the ecological status of many river reaches (Hemming et al., 2018; Drescher et al., 2013; Martin et al., 2012) but also enabled a detailed investigation of the different sensitivity of macroinvertebrates towards a range of anthropogenic alterations.

Overall, our results provided a rather complex picture whereby the ability of macroinvertebrates to indicate ecological conditions appears contingent on stressor type and their cooccurrence. As the spatial distribution of anthropogenic stressors in river networks inevitably covary with elevation and stream order, these natural watershed-level features were often included as key predictors of river conditions. Elevation, stream order, stream types and distance from source emerged as valuable predictors of the ecological status of Alpine rivers, when measured through its hydro-morphological and water quality components. This reflects the inherent spatial structure of our Alpine river network, where non-impacted sites are more frequent in headwaters. This longitudinal gradient corresponds to increasing anthropogenic pressures such as urbanization, intensive agriculture, river regulation and habitat alterations (Pollice et al., 2020; Larsen et al., 2019b). In our modelling framework, elevation-related variables act as proxies for river zonation, while discrimination among pressure types is supported by consistent taxonomic and functional assemblages, reflecting not only longitudinal patterns but also specific environmental stress responses. Associations observed in macroinvertebrate RF trait-based models were consistent with this altitudinal pattern: taxa adapted to cold, fast-flowing headwaters and traits linked to rheophilic or cold-stenothermic preferences were predominantly associated with pristine conditions, whereas downstream assemblages reflected traits related to warmer, slower-flowing environments and higher anthropogenic pressures. These findings align with the broader understanding of macroinvertebrate responses to multiple stressors. Macroinvertebrate-based metrics are generally considered sensitive to water quality and better indicators of organic pollution than to other stressors such as streamflow or morphological alterations (Friberg, 2014). In fact, several EPT taxa (Ephemeroptera, Plecoptera, and Trichoptera), widely recognised as sensitive indicators of water quality, emerged as key features in the RF classification process, particularly in identifying less impacted conditions, except for *Baetis* species, *Serratella* sp. and Hydropsychidae, which appeared rather tolerant to anthropogenic pressures.

Unexpectedly, we found that neither taxonomic nor functional features were particularly efficient in discriminating between pristine sites and those only affected by water pollution (P-N comparison). Conversely, better classification performance was associated with the ability to detect pollution in sites already altered by both hydro and morphological stressors (PHM-HM comparison), with Asellidae and Naididae particularly abundant at sites with reduced water quality, as also observed in other studies (e.g., Horváthová et al., 2024; O'Callaghan et al., 2019; Bodon et al., 2017). In other words, RF models appeared to better identify water quality alterations among already hydro-morphologically altered reaches as opposed to reaches in pristine conditions. However, the performance of models designed to detect alterations to water quality in stream reaches also affected by hydrological (PH-H) or morphological stressors (PM-M) was not always optimal. Effect sizes suggest that the ability to detect water pollution was slightly higher among reaches affected by hydrological alterations compared to reaches impacted by morphological stressors. Nevertheless, the first binary comparison (PH-H) showed good accuracy only in models based on taxonomic features. The second comparison (PM-M) revealed that the median accuracy of the models was lower than in other pairwise comparisons in both taxonomic and functional composition-based models.

Although Simuliidae are typically regarded as indicators of good chemical water quality due to their preference for well-oxygenated and low-nutrient streams (Tachet et al., 2010; Bonada et al., 2006), in our findings Simuliidae and Lumbricidae emerged as indicator taxa for detecting poor water quality in stream reaches affected by hydrological alterations (PH sites). Under these conditions, the community was dominated by taxa preferring the lower course of rivers with life cycle durations longer than one year and tolerance to a wide temperature range. In contrast, *Baetis* sp., Hydropsychidae, Simuliidae, and Chironomidae appeared more abundant in polluted waters also affected by morphological alterations (PM); the tolerance of these taxa to water quality and habitat stressors has been previously recorded (e.g., Kail et al., 2012; Štangler et al., 2013; Rossaro et al., 2022), although *Baetis* and Hydropsychidae are usually considered sensitive to water quality degradation (e.g. Kaelin and Altermatt, 2016; Vilenica et al., 2022; Tubić et al., 2024). In these sites (PM), the assemblages were characterised by taxa associated with middle and lower lowland river zone, showing bivoltine life cycles and preference for very cold temperatures.

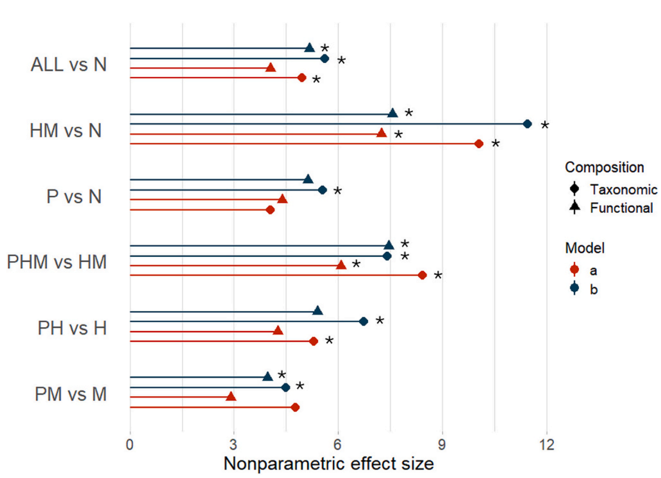

Fig. 4. Nonparametric effect size (d) for the performed pairwise comparisons (acronyms defined in Table 3); *indicates comparisons that are different from the null model, as shown in Fig. 2.

Sites only affected by hydro-morphological alterations (HM) were clearly discriminated from pristine sites, showing the highest effect size, with several taxa such as Hydraenidae and *Protonemura* sp. virtually absent from hydro-morphologically altered locations. This suggests that macroinvertebrates could show sensitivity to stressors other than water quality issues, at least in small mountain streams. For example, in addition to EPT taxa, RF models consistently identified several Coleoptera families such as Hydraenidae and Elmidae, and Diptera families including Blephariceridae and Athericidae as important features in pairwise comparisons targeting hydro-morphological pressure detection. These taxa are typically associated with riffles and fast-running waters of mountain streams, inhabit clean, cool, well-oxygenated streams, and they could be a potentially valuable bioindicators due to their sensitivity to both physical and chemical alterations (Compin and Céréghino, 2003). This outcome of our study appears relevant also in the light of the findings of Devreux et al. (2025), who showed that, provided a suitable multi-scale invertebrate sampling method is adopted, macroinvertebrates could show sensitivity to morphological alterations, as it resulted from a comparison between a near-natural and a restored reach in a braided rivers, with higher taxa and trait diversity in the natural reach.

Some important aspects related to differences between taxonomic and functional metrics were uncovered. Trait-based approaches are often considered more efficient than taxonomic metrics as indicators of environmental changes (Menezes et al., 2010; Berger et al., 2018), because anthropogenic stressors select taxa based on adaptive traits, so that community trait profiles can track the specific source of impairment (Dolédec and Statzner, 2008). However, in this study the functional features of macroinvertebrate communities were not discriminating between stressor types better than the taxonomic features. Functional metrics showed some limitations in the ability to distinguish between different stressors categories compared to taxonomic approaches, especially where hydro-morphological alterations were compared to pristine sites (HM-N), and as previously discussed, in the presence of water pollution under hydrological pressure (PH-H). This may be partly due to traits information was not available for every taxa observed in the study, which might have overlooked the presence of sensitive taxa unique to either pristine or altered conditions (Mathers et al., 2024; Schmera et al., 2017; Poff et al., 2006; Dolédec et al., 2006), thus reducing the discriminatory power of the RF classifier based on the functional approach.

Another caveat worth discussing regards the balance of the analytical design that can limit quantitative comparison of predictive accuracy among the different RF classifications (Aria et al., 2021). We observed that data unbalance influenced the discriminatory power of the models, with more balanced models (model-b) performing better than the unbalanced counterparts (model-a). Such bias may have particularly influenced the predictive power of the classifications ALL-N and P-N, which included a small number of sites affected by water pollution. More specifically, our study area included few sites classified as P (water pollution only), relative to those affected by other or a combination of stressors, which is typical for Alpine rivers (Larsen et al., 2019a; Ferreira et al., 2019; Chiogna et al., 2016; Schinegger et al., 2012).

As a final remark, we highlight that the present study has been conducted by analysing datasets that should be available at many regional Environmental Protection Agencies, at least in Europe, as they are part of the institutional, regular monitoring that is required to fulfil WFD-compliant national environmental legislations. The proposed approach that integrates RF analysis and expert knowledge on relevant local and upstream stressors for each macroinvertebrate sampling station could be profitably upscaled within the Alpine region and expanded towards other bio-geographical regions, with high potential to significantly improve our present ability to assess the role of hydro-morphological, pollution stressors, and of restoration and mitigation measures, on river systems.

5. Conclusions

Compared to approaches that assess correlations between individual or multiple pressures and biotic indices (e.g. richness, diversity), our approach enabled the integration of several relevant environmental and biotic features into machine learning algorithms. These included stream characteristics as well as individual taxa abundances and functional traits used as predictive features. The used dataset is representative of what should be available at regional Environment Agencies in EU countries addressing the WFD requirements, which facilitates its replicability. This approach allowed an evaluation of the effectiveness of benthic invertebrates in distinguishing specific stressors, and the identification of a set of taxa and traits potentially useful to develop metrics capable of detecting the combined effects of multiple, co-occurring, environmental pressures. However, data availability and the resolution of trait information remain as possible constraints limiting the effectiveness of functional approaches to disentangle multiple stressor types in freshwater ecosystems. Investigating stressor-specific indicators in multi-stressor contexts is particularly relevant in montane and Alpine ecosystems in light of the ongoing climate and environmental changes (Becquet et al., 2022; Hotaling et al., 2017; Khamis et al., 2014). Our findings highlight the potential of combining machine learning methods with expert knowledge to appraise the degree to which macroinvertebrate assemblages can help discriminate between multiple forms of anthropogenic alteration, especially where rivers are predominantly affected by hydro-morphological regulation rather than chemical pollution (Calabrese et al., 2020; Khamis et al., 2014).

CRediT authorship contribution statement

Francesca Vallefuoco: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Stefano Larsen:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Pietro Franceschi:** Writing – review & editing, Methodology, Conceptualization. **Valentina Dallafior:** Writing – review & editing, Data curation. **Walter Bertoldi:** Writing – review & editing, Supervision. **Guido Zolezzi:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Maria Cristina Bruno:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was partly supported by funding from the Institute for Alpine Environment of Eurac Research and by the Italian Ministry of Education, University and Research (MIUR) through the 'Departments of Excellence' grant 2018–2022 (Law n. 232/2016) awarded to Prof. Guido Zolezzi. The authors gratefully acknowledge the Environmental Protection Agency of the Autonomous Province of Trento for providing data and technical support. The authors thank the Department of Innovation, Research University and Museums of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2026.114960.

Data availability

Data will be made available on request.

References

Aria, M., Cuccurullo, C., Gnasso, A., 2021. A comparison among interpretative proposals for random forests. Mach. Learn. Appl. 6, 100094. https://doi.org/10.1016/j.mlwa.2021.100094.

Arrighi, C., Castelli, F., 2023. Prediction of ecological status of surface water bodies with supervised machine learning classifiers. Sci. Total Environ. 857, 159655. https://doi.org/10.1016/j.scitotenv.2022.159655.

Azzellino, A., Canobbio, S., Çervigen, S., Marchesi, V., Piana, A., 2015. Disentangling the multiple stressors acting on stream ecosystems to support restoration priorities. Water Sci. Technol. 72 (2), 293–302. https://doi.org/10.2166/wst.2015.177.

Becquet, J., Lamouroux, N., Condom, T., Gouttevin, I., Forcellini, M., Launay, B., Rabatel, A., Cauvy-Fraunié, S., 2022. Macroinvertebrate distribution associated with environmental variables in alpine streams. Freshw. Biol. 67 (10), 1815–1831. https://doi.org/10.1111/fwb.13977.

Bélisle, A.C., Asselin, H., LeBlanc, P., Gauthier, S., 2018. Local knowledge in ecological modeling. Ecol. Soc. 23 (2). https://doi.org/10.5751/ES-09949-230214.

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.

Birk, S., Chapman, D., Carvalho, L., Spears, B.M., Andersen, H.E., Argillier, C., Hering, D., 2020. Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. Nat. Ecol. Evol. 4 (8), 1060–1068. https://doi.org/10.1038/s41559-020-1216-4.

Bodon, M., Gaiter, S., Costa, S., 2017. La componente biologica nella valutazione delle acque sorgive: considerazioni sulla validità di questo indicatore dopo 25 anni di sperimentazione in Liguria. Biol. Ambient. 31, 49–56.

Bonada, N., Prat, N., Resh, V.H., Statzner, B., 2006. Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. Annu. Rev. Entomol. 51 (1), 495–523. https://doi.org/10.1146/annurev.ento.51.110104.151124.

Bruno, M.C., Maiolini, B., Carolli, M., Silveri, L., 2010. Short time-scale impacts of hydropeaking on benthic invertebrates in an Alpine stream (Trentino, Italy). Limnologica 40 (4), 281–290. https://doi.org/10.1016/j.limno.2009.11.012.

Buchanan, B.P., Sethi, S.A., Cuppett, S., Lung, M., Jackman, G., Zarri, L., Rahm, B.G., 2022. A machine learning approach to identify barriers in stream networks demonstrates high prevalence of unmapped riverine dams. J. Environ. Manag. 302, 113952. https://doi.org/10.1016/j.jenvman.2021.113952.

Buskirk, T.D., Kirchner, A., Eck, A., Signorino, C.S., 2018. An introduction to machine learning methods for survey researchers. Surv. Pract. 11 (1). https://doi.org/10.29115/SP-2018-0004.

Bussettini, M., Lastoria, B., Braca, G., Mariani, S., Venturelli, S., 2011. Regime idrologico (in Italian). In: Implementazione della Direttiva 2000/60/CE. Analisi e valutazione degli aspetti idromorfologici. Versione 1.1. ISPRA, Manuali e Linee Guida. ISBN: 978-88-448-0439-8, (accessed 09 January 2026). https://www.isprambiente.gov.it/contentfiles/00010100/10147-analisi-e-valutazione-degli-aspetti-idromorfologici-agosto-2011.pdf.

Calabrese, S., Mezzanotte, V., Marazzi, F., Canobbio, S., Fornaroli, R., 2020. The influence of multiple stressors on macroinvertebrate communities and ecosystem attributes in northern Italy pre-Alpine rivers and streams. Ecol. Indic. 115, 106408. https://doi.org/10.1016/j.ecolind.2020.106408.

Chiogna, G., Majone, B., Paoli, K.C., Diamantini, E., Stella, E., Mallucci, S., Bellin, A., 2016. A review of hydrological and chemical stressors in the Adige catchment and its ecological status. Sci. Total Environ. 540, 429–443. https://doi.org/10.1016/j.scitotenv.2015.06.149.

Chon, T.S., Kwak, I.S., Park, Y.S., Kim, T.H., Kim, Y., 2001. Patterning and short-term predictions of benthic macroinvertebrate community dynamics by using a recurrent artificial neural network. Ecol. Model. 146 (1–3), 181–193. https://doi.org/10.1016/S0304-3800(01)00305-2.

Comiti, F., 2012. How natural are Alpine mountain rivers? Evidence from the Italian Alps. Earth Surf. Process. Landf. 37 (7), 693–707. https://doi.org/10.1002/esp.2267.

Compin, A., Céréghino, R., 2003. Sensitivity of aquatic insect species richness to disturbance in the Adour – Garonne stream system (France). Ecol. Indic. 3, 135–142. https://doi.org/10.1016/S1470-160X(03)00016-5.

Cutler, D.R., Edwards Jr., T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88 (11), 2783–2792. https://doi.org/10.1890/07-0539.1.

Darr, A., Gogina, M., Zettler, M.L., 2014. Functional changes in benthic communities along a salinity gradient–a western Baltic case study. J. Sea Res. 85, 315–324. https://doi.org/10.1016/j.seares.2013.06.003.

Devreux, L., Belletti, B., Dolédec, S., Marmonier, P., Chapuis, M., 2025. Morpho-ecological restoration in braided rivers: the need to account for geomorphological heterogeneity. River Res. Appl. 41 (9), 1888–1903. https://doi.org/10.1002/rra.70033.

Dolédec, S., Statzner, B., 2008. Invertebrate traits for the biomonitoring of large European rivers: an assessment of specific types of human impact. Freshw. Biol. 53, 617–634. https://doi.org/10.1111/j.1365-2427.2007.01924.x.

Dolédec, S., Phillips, N., Scarsbrook, M., Riley, R.H., Townsend, C.R., 2006. Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. J. N. Am. Benthol. Soc. 25 (1), 44–60. https://doi.org/10.1899/0887-3593(2006)25[44:COSAFA]2.0.CO;2.

Drescher, M., Perera, A.H., Johnson, C.J., Buse, L.J., Drew, C.A., Burgman, M.A., 2013. Toward rigorous use of expert knowledge in ecological research. Ecosphere 4 (7), 1–26. https://doi.org/10.1890/ES12-00415.1.

European Parliament, 2000. Parliament and Council Directive 2000/60/EC establishing a Framework for Community Action in the Field of Water Policy. Official Journal of the European Union PE-CONS (3639/3631/3600 REV 3631), Brussels, Belgium.

Ferreira, T., Globevnik, L., Schinegger, R., 2019. Water Stressors in Europe: New Threats in the Old World. In: Multiple Stressors in River Ecosystems: Status, Impacts and Prospects for the Future. Elsevier, Amsterdam, The Netherlands, pp. 139–155. https://doi.org/10.1016/B978-0-12-811713-2.00008-X.

Friberg, N., 2014. Impacts and indicators of change in lotic ecosystems. WIREs Water 1, 513–531. https://doi.org/10.1002/wat2.1040.

Gabriels, W., Goethals, P.L., Dedecker, A.P., Lek, S., De Pauw, N., 2007. Analysis of macrobenthic communities in Flanders, Belgium, using a stepwise input variable selection procedure with artificial neural networks. Aquat. Ecol. 41 (3), 427–441. https://doi.org/10.1007/s10452-007-9081-7.

Han, H., Guo, X., Yu, H., 2016. Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In: 2016 7th IEEE International Conference on Software Engineering and Service Science (Icsess). IEEE, pp. 219–224. https://doi.org/10.1109/ICSESS.2016.7883053.

He, F., Arora, R., Mansour, I., 2023. Multispecies assemblages and multiple stressors: synthesizing the state of experimental research in freshwaters. Wiley Interdiscip. Rev.: Water 10 (3), e1641. https://doi.org/10.1002/wat2.1641.

Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F., Wintle, B.C., 2018. A practical guide to structured expert elicitation using the IDEA protocol. Methods Ecol. Evol. 9 (1), 169–180. https://doi.org/10.1111/2041-210X.12857.

Hering, D., Moog, O., Sandin, L., Verdonschot, P.F., 2004. Overview and application of the AQEM assessment system. Hydrobiologia 516, 1–20. https://doi.org/10.1023/B:HYDR.0000025255.70009.a5.

Ho, L., Goethals, P., 2022. Machine learning applications in river research: trends, opportunities and challenges. Methods Ecol. Evol. 13 (11), 2603–2621. https://doi.org/10.1111/2041-210X.13992.

Horváthová, T., Lafuente, E., Bartels, J.-A., Wallisch, J., Vorburger, C., 2024. Tolerance to environmental pollution in the freshwater crustacean Asellus aquaticus: a role for the microbiome. Environ. Microbiol. Rep. 16 (3), e13252. https://doi.org/10.1111/1758-2229.13252.

Hotaling, S., Finn, D.S., Joseph Giersch, J., Weisrock, D.W., Jacobsen, D., 2017. Climate change and alpine stream biology: progress, challenges, and opportunities for the future. Biol. Rev. 92 (4), 2024–2045. https://doi.org/10.1111/brv.12319.

ISPRA, 2014. Metodi biologici per le acque superficiali interne - 2010. Protocollo di campionamento e analisi dei macroinvertebrati bentonici dei corsi d'acqua guadabili. ISPRA, Manuali e Linee Guida 111/2014. ISBN: 978-88-448-0651 MLG_111_2014_Metodi_Biologici_acque.pdf (accessed 09 January 2026).

Kaelin, K., Altermatt, F., 2016. Landscape-level predictions of diversity in river networks reveal opposing patterns for different groups of macroinvertebrates. Aquat. Ecol. 50, 283–295. https://doi.org/10.1007/s10452-016-9576-1.

Kail, J., Arle, J., Jähnig, S.C., 2012. Limiting factors and thresholds for macroinvertebrate assemblages in European rivers: empirical evidence from three datasets on water quality, catchment urbanization, and river restoration. Ecol. Indic. 18, 63–72. https://doi.org/10.1016/j.ecolind.2011.09.038.

Khamis, K., Hannah, D.M., Clarvis, M.H., Brown, L.E., Castella, E., Milner, A.M., 2014. Alpine aquatic ecosystem conservation policy in a changing climate. Environ. Sci. Pol. 43, 39–55. https://doi.org/10.1016/j.envsci.2013.10.004.

Krtolica, I., Kamenko, I., Paunović, M., Raković, M., Atanacković, A., Talanov, M., Popović, N., 2025. Use of artificial neural networks (ANNs) to assess xenobiotics in a river catchment using macroinvertebrates as bioindicators. Aquat. Toxicol. 286, 107466. https://doi.org/10.1016/j.aquatox.2025.107466.

Laini, A., Bolpagni, R., Cancellario, T., Guareschi, S., Racchetti, E., Viaroli, P., 2018. Testing the response of macroinvertebrate communities and biomonitoring indices under multiple stressors in a lowland regulated river. Ecol. Indic. 90, 47–53. https://doi.org/10.1016/j.ecolind.2018.02.051.

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. Retrieved from cran.r-project.org/web/packages/FD.

Larsen, S., Bruno, M.C., Zolezzi, G., 2019a. WFD ecological status indicator shows poor correlation with flow parameters in a large Alpine catchment. Ecol. Indic. 98, 704–711. https://doi.org/10.1016/j.ecolind.2018.11.047.

Larsen, S., Bruno, M.C., Vaughan, I.P., Zolezzi, G., 2019b. Testing the river continuum concept with geostatistical stream-network models. Ecol. Complex. 39, 100773. https://doi.org/10.1016/j.ecocom.2019.100773.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S., 1996. Application of neural networks to modelling nonlinear relationships in ecology. Ecol. Model. 90 (1), 39–52. https://doi.org/10.1016/0304-3800(95)00142-5.

Lemm, J.U., Venohr, M., Globevnik, L., Stefanidis, K., Panagopoulos, Y., van Gils, J., Posthuma, L., Kristensen, P., Feld, C.K., Mahnkopf, J., Hering, D., Birk, S., 2021. Multiple stressors determine river ecological status at the European scale: towards an integrated understanding of river status deterioration. Glob. Chang. Biol. 27, 1962–1975. https://doi.org/10.1111/gcb.15504.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R news 2 (3), 18–22. ISSN 1609–3631.

Martin, T.G., Burgman, M.A., Fidler, F., Kuhnert, P.M., Low-Choy, S., McBride, M., Mengersen, K., 2012. Eliciting expert knowledge in conservation science. Conserv. Biol. 26 (1), 29–38. https://doi.org/10.1111/j.1523-1739.2011.01806.x.

Martyszunis, A., Loga, M., Przeździecki, K., 2024. Using machine learning for the assessment of ecological status of unmonitored waters in Poland. Sci. Rep. 14 (1), 24509. https://doi.org/10.1038/s41598-024-74511-4.

Mathers, K.L., Robinson, C.T., Hill, M., Kowarik, C., Heino, J., Deacon, C., Weber, C., 2024. How effective are ecological metrics in supporting conservation and management in degraded streams? Biodivers. Conserv. 33 (14), 3981–4002. https://doi.org/10.1007/s10531-024-02933-7.

Menezes, S., Baird, D.J., M. Soares, M.V., 2010. Beyond taxonomy: a review of macroinvertebrate trait-based community descriptors as tools for freshwater biomonitoring. J. Appl. Ecol. 47 (4), 711–719. https://doi.org/10.1111/j.1365-2664.2010.01819.x.

O'Callaghan, I., Harrison, S., Fitzpatrick, D., Sullivan, T., 2019. The freshwater isopod Asellus aquaticus as a model biomonitor of environmental pollution: a review. Chemosphere 235, 498–509. https://doi.org/10.1016/j.chemosphere.2019.06.217.

Ormerod, S.J., Dobson, M., Hildrew, A.G., Townsend, C.R., 2010. Multiple stressors in freshwater ecosystems. Freshw. Biol. 55, 1–4. https://doi.org/10.1111/j.1365-2427.2009.02395.x.

Orr, J.A., Macaulay, S.J., Mordente, A., Burgess, B., Albini, D., Hunn, J.G., Restrepo-Sulez, K., Wilson, R., Schechner, A., Robertson, A.M., Lee, B., 2024. Studying interactions among anthropogenic stressors in freshwater ecosystems: a systematic review of 2396 multiple-stressor experiments. Ecol. Lett. 27 (6), e14463. https://doi.org/10.1111/ele.14463.

Oshiro, T.M., Perez, P.S., Baranauskas, J.A., 2012. How many trees in a random forest?. In: International Workshop on Machine Learning and Data Mining in Pattern Recognition. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 154–168.

Özesmi, S.L., Tan, C.O., Özesmi, U., 2006. Methodological issues in building, training, and testing artificial neural networks in ecological applications. Ecol. Model. 195 (1–2), 83–93. https://doi.org/10.1016/j.ecolmodel.2005.11.012.

Perera, A.H., Drew, C.A., Johnson, C.J., 2012. Experts, expert knowledge, and their roles in landscape ecological applications. In: Expert Knowledge and its Application in Landscape Ecology. Springer, New York, NY, pp. 1–10.

Poff, N.L., Zimmerman, J.K.S., 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. Freshw. Biol. 55, 194–205. https://doi.org/10.1111/j.1365-2427.2009.02272.x.

Poff, N.L., Olden, J.D., Vieira, N.K.M., Finn, D.S., Simmons, M.P., Kondratieff, B.C., 2006. Functional trait niches of north American lotic insects: trait-based ecological applications in light of phylogenetic relationships. J. N. Am. Benthol. Soc. 25 (4), 730–755. https://doi.org/10.1899/0887-3593(2006)025[0730:FTNONA]2.0.CO;2.

Politikos, D., Stefanidis, K., Varlas, G., Papadopoulos, A., Dimitriou, E., 2024. Predicting the ecological quality of rivers: a machine learning approach and a what-if scenarios tool. Environ. Model. Assess. 29 (6), 1059–1077. https://doi.org/10.1007/s10666-024-09980-y.

Pollice, A., Jona-Lasinio, G., Gaglio, M., Blanchet, F.G., Fano, E.A., 2020. Modelling the effect of directional spatial ecological processes for a river network in northern Italy. Ecol. Indic. 112, 106144. https://doi.org/10.1016/j.ecolind.2020.106144.

Quadroni, S., Crosa, G., Gentili, G., Espa, P., 2017. Response of stream benthic macroinvertebrates to current water management in Alpine catchments massively developed for hydropower. Sci. Total Environ. 609, 484–496. https://doi.org/10.1016/j.scitotenv.2017.07.099.

Richter, B.D., Baumgartner, J.V., Powell, J., Braun, D.P., 1996. A method for assessing hydrologic alteration within ecosystems. Conserv. Biol. 10 (4), 1163–1174.

Rinaldi, M., Surian, N., Comiti, F., Bussettini, M., 2013. A method for the assessment and analysis of the hydromorphological condition of Italian streams: the morphological quality index (MQI). Geomorphology 180, 96–108. https://doi.org/10.1016/j.geomorph.2012.09.009.

Rossaro, B., Marziali, L., Boggero, A., 2022. Response of chironomids to key environmental factors: perspective for biomonitoring. Insects 13 (10), 911. https://doi.org/10.3390/insects13100911.

RSA, 2020. Rapporto sullo stato dell'ambiente. Agenzia Provinciale per la protezione dell'ambiente, Autonomous province of Trento. https://www.appa.provincia.tn.it/ocmultibinary/download/48501/543218/96/b122202bf00a760701caa4d8b07d106e.pdf/file/14_RSA+2020_Acqua.pdf (accessed 1 September 2025).

Sahli, H., 2020. An introduction to machine learning. In: TORUS 1–Toward an Open Resource using Services: Cloud Computing for Environmental Data, pp. 61–74. https://doi.org/10.1002/9781119720492.ch7.

Schinegger, R., Trautwein, C., Melcher, A., Schmutz, S., 2012. Multiple human pressures and their spatial patterns in European running waters. Water Environ. J. 26 (2), 261–273. https://doi.org/10.1111/j.1747-6593.2011.00285.x.

Schmera, D., Heino, J., Podani, J., Erős, T., Dolédec, S., 2017. Functional diversity: a review of methodology and current knowledge in freshwater macroinvertebrate research. Hydrobiologia 787, 27–44. https://doi.org/10.1007/s10750-016-2974-5.

Schmidt-Kloiber, A., Hering, D., 2015. An online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences. Ecol. Indic. 53, 271–282. https://doi.org/10.1016/j.ecolind.2015.02.007.

Singh, K.P., Basant, A., Malik, A., Jain, G., 2009. Artificial neural network modeling of the river water quality—a case study. Ecol. Model. 220 (6), 888–895. https://doi.org/10.1016/j.ecolmodel.2009.01.004.

Spears, B.M., Chapman, D.S., Carvalho, L., Feld, C.K., Gessner, M.O., Piggott, J.J., Banin, L.F., Gutiérrez-Cánovas, C., Solheim, A.L., Richardson, J.A., Schinegger, R., 2021. Making waves. Bridging theory and practice towards multiple stressor management in freshwater ecosystems. Water Res. 196, 116981. https://doi.org/10.1016/j.watres.2021.116981.

Štangler, A., Halgoš, J., Beracko, P., 2013. Blackfly (Diptera, Simuliidae) communities and species richness estimation in Carpathian montane streams. Cent. Eur. J. Biol. 8 (7), 681–692. https://doi.org/10.2478/s11535-013-0187-0.

Tachet, H., Richoux, P., Bournaud, M., Usseglio-Polatera, P., 2010. Invertébrés d'eau douce: systématique, biologie, écologie. CNRS éditions, Paris. IBSN: 978–2–271-06945-0.

Theodoropoulos, C., Vourka, A., Skoulikidis, N., Rutschmann, P., Stamou, A., 2018. Evaluating the performance of habitat models for predicting the environmental flow requirements of benthic macroinvertebrates. J. Ecohydraul. 3 (1), 30–44. https://doi.org/10.1080/24705357.2018.1440360.

Tickner, D., Opperman, J.J., Abell, R., Acreman, M., Arthington, A.H., Bunn, S.E., Cooke, S.J., Dalton, J., Darwall, W., Edwards, G., Harrison, I., 2020. Bending the curve of global freshwater biodiversity loss: an emergency recovery plan. BioScience 70 (4), 330–342. https://doi.org/10.1093/biosci/biaa002.

Tubić, B., Andjus, S., Zorić, K., Vasiljević, B., Jovičić, K., Čanak Atlagić, J., Paunović, M., 2024. Aquatic insects (Ephemeroptera, Plecoptera and Trichoptera) metric as an important tool in water quality assessment in hilly and mountain streams. Water 16 (6), 849. https://doi.org/10.3390/w16060849.

Vezza, P., Parasiewicz, P., Spairani, M., Comoglio, C., 2014. Habitat modeling in high-gradient streams: the mesoscale approach and application. Ecol. Appl. 24 (4), 844–861. https://doi.org/10.1890/11-2066.1.

Vilenica, M., Vidaković Maoduš, I., Mihaljević, Z., 2022. The impact of hydromorphological alterations on mayfly assemblages of a mid-sized lowland river in South-Eastern Europe. Insects 13 (5), 436. https://doi.org/10.3390/insects13050436.

Yao, S., Liu, J., Dai, G., Liu, X., Wen, C., Yang, P., Wan, S., Jian, Y., 2026. Decoding the interactive effect of water quality-land use on benthic macroinvertebrate biodiversity in rivers with an interpretable machine learning framework. Glob. Ecol. Conserv., e04161 https://doi.org/10.1016/j.gecco.2026.e04161.

Zanoni, M.G., Majone, B., Bellin, A., 2022. A catchment-scale model of river water quality by machine learning. Sci. Total Environ. 838, 156377. https://doi.org/10.1016/j.scitotenv.2022.156377.

Zolezzi, G., Bellin, A., Bruno, M.C., Maiolini, B., Siviglia, A., 2009. Assessing hydrological alterations at multiple temporal scales: Adige River, Italy. Water Resour. Res. 45. https://doi.org/10.1029/2008WR007266.