Ecological Indicators 151 (2023) 110335

Contents lists available at ScienceDirect

Ecological Indicators

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

Original Articles

Selection of an adequate functional diversity index for stream assessment based on biological traits of macroinvertebrates

Gonzalo Sotomayor [a,b,*], Henrietta Hampel [b], Raúl F. Vázquez [b,c], Marie Anne Eurie Forio [a], Peter L.M. Goethals [a]

[a] Ghent University, Faculty of Bioscience Engineering, Department of Animal Sciences and Aquatic Ecology, Coupure Links 653, 9000 Ghent, Belgium
[b] Universidad de Cuenca, Facultad de Ciencias Químicas, Laboratorio de Ecología Acuática (LEA), Av. 12 de abril S/N, Cuenca, Ecuador
[c] Universidad de Cuenca, Facultad de Ingeniería, Departamento de Ingeniería Civil, Av. 12 de abril S/N, Cuenca, Azuay, Ecuador

ARTICLE INFO

Keywords:
GAMM
Random forest
Functional diversity
Streams
Macroinvertebrates
Water quality

ABSTRACT

Functional diversity (FD) is useful for the evaluation of freshwater ecosystems. The FD of macroinvertebrate families for river water quality (WQ) assessment in the Paute River Basin (PRB), Ecuador, was investigated. Macroinvertebrate samples and data about 26 physical, chemical, microbiological and hydro-geomorphological variables were available. Literature-based biological traits were allocated as scores to the macroinvertebrates data through fuzzy coding. The Generalised Additive Mixed Model (GAMM) was used to assess the performance of six FD indices using the referred 26 WQ descriptive variables. The best performing GAMM led to selecting the index based on functional dendrograms including the species community pool (wFDc) as the most suitable to characterise FD in the PRB. The sub-basins of the PRB were grouped in three classes applying Average Linkage Clustering (ALC) and using wFDc. The Random Forest (RF) algorithm was used with a global efficiency of 89% to assess the ALC clusters consistency and pre-identify the significant WQ descriptive variables, explaining most of the FD variability. The Kruskal-Wallis test was then applied to refine the outcomes of the previous analysis. Twelve WQ descriptive variables were finally identified as the best discriminant predictors for FD, including the riparian vegetation, electric conductivity, dissolved oxygen, total hardness, faecal coliforms and pH. It is believed that the implemented approach successfully assessed the stream WQ status of the PRB upon selecting a suitable macroinvertebrate FD index; as such, it could be applied to other tropical basins for WQ assessment.

1. Introduction

Good water quality (WQ) is a key component of sustainable socio-economic development (Bartram & Ballance, 1996). However, rivers and lakes are increasingly being contaminated worldwide by anthropic activities, and water use has increased drastically with economic development and population growth. Both facts have resulted in a severe decline in surface WQ and, as such, many countries, both developed and developing ones, currently face threats to their water security (Liu and Zhang, 2019). In this context, an important aspect, especially for developing countries, is the establishment of adequate monitoring programmes to build management plans and policies for successfully reducing the negative effects of human impacts (Azab, 2012).

The use of variables measuring different physical, chemical and biological features of freshwater ecosystems is a useful way to assess their health. From the ecohydrological point of view, it is necessary to understand the links between freshwater biota, e.g., benthic macroinvertebrates, and physical–chemical factors, to support policy for the sustainable use of water resources (Loinaz, 2012). Benthic macroinvertebrates have key functional importance for water bodies. Among other functions, they act in nutrient cycling and energy flow within food webs and may serve as a food source for other invertebrates, fish, and birds (Wallace & Webster, 1996; Buss et al., 2015). In the past two decades, interest in the functional diversity (FD) of benthic macroinvertebrates, has increased because FD helps to understand the relationships between community structure and the functioning of aquatic ecosystems facing various types of disturbance (Bruno et al., 2016). FD is indicated by the diversity of taxa's traits in ecosystems

* Corresponding author at: Ghent University, Faculty of Bioscience Engineering, Department of Animal Sciences and Aquatic Ecology, Coupure Links 653, 9000 Ghent, Belgium.

E-mail addresses: gonzalo.sotomayor@ugent.be (G. Sotomayor), hennihampel@gmail.com (H. Hampel), raulfvazquezz@yahoo.co.uk (R.F. Vázquez), marie.forio@ugent.be (M.A.E. Forio), peter.goethals@ugent.be (P.L.M. Goethals).

https://doi.org/10.1016/j.ecolind.2023.110335
Received 9 December 2022; Received in revised form 4 May 2023; Accepted 5 May 2023
Available online 10 May 2023

(Schleuter et al., 2010). Thus, macroinvertebrate traits such as maximum body size, respiration forms, morphological specialisations for attachment to substrate, osmoregulatory organ types, and feeding habits should provide information about the local environmental conditions (Reynaga et al., 2020). Hence, relevant studies have been carried out on benthic macroinvertebrates and their FD through biological traits (Brown et al., 2018; Forio et al., 2018; Arenas-Sánchez et al., 2021). In general, a range of FD indices is available to assess the ecological integrity of aquatic ecosystems (Mouchet et al., 2010; Pla et al., 2012). These indices allow the comparison among aquatic communities in different types of environments (Koperski, 2019). However, for making the appropriate selection of FD indices for biomonitoring, the key step is to determine which FD index performs best, for which, the explanatory power and their statistical validity must be well assessed (Petchey & Gaston, 2006; Mouchet et al., 2010; Mason et al., 2013).

The Paute River Basin (PRB), one of the most important hydrological systems of Ecuador owing to its significant hydroelectric potential (Salazar & Rudnick, 2008; Castillo and Álvarez, 2014), has been relatively well studied in terms of the relationships between environmental features and macroinvertebrate taxonomical aspects (Turcotte & Harper, 1982; Holguin-Gonzalez et al., 2013; Sotomayor, 2016; Vimos-Lojano et al., 2016; Herrera & Burneo, 2017; Jerves-Cobo et al., 2018; Jerves-Cobo et al., 2020; Sotomayor et al., 2020; Vázquez et al., 2020). However, for this basin, FD research has been conducted, focusing on (i) specific sub-basins and not on the whole basin; and (ii) only the analysis of functional feeding groups (FFG). Hereafter, Iñiguez-Armijos et al. (2018) investigated the influence of replacing native forests with pastures on the FFG in an upper sub-basin; Vimos-Lojano et al. (2020) assessed macroinvertebrates and their FFG in two pristine micro-catchments; Jiménez et al. (2021) carried out a FFG study in a cattle-raising impacted sub-basin; and Sotomayor et al. (2021) analysed the use of the family taxonomic resolution of macroinvertebrates to detect changes of benthic community assemblages in the PRB. Therefore, further evaluation is needed considering the whole study basin, using complementary FD approaches to the ones applied in prior research, aiming at assessing whether environmental changes in the freshwater ecosystems of the study basin are reflected. For instance, the assessment of the robustness of different FD indices applied to the data of the PRB is a very important aspect that should be implemented.

Henceforth, the main objective of the current study was to assess the stream WQ status of the PRB upon the selection of a suitable macroinvertebrate FD index. The specific objectives of this research were: (i) identifying and selecting a benthic macroinvertebrate FD index that could reveal changes in community assemblages along abiotic stress gradients in the study site; (ii) grouping the PRB sub-basins as a function of their FD for management purposes; and (iii) identifying the most significant WQ descriptive variables that explain most of the variability of benthic macroinvertebrates and their respective FD.

2. Materials and methods

2.1. Study area

The study basin is situated in the south of Ecuador (Fig. 1) and has an area of 6442 km^2. The basin elevation ranges between 410 and 4687 m above sea level (a.s.l.). The slope varies from 25% to 50%. The lower temperatures correspond to the western Andes range with a mean daily value of about 6 °C (at about 3500 m a.s.l.), while the warmest areas are situated in the Amazonian-influenced valleys and subtropical zones, with a mean daily value of 24 °C; nevertheless, a remarkable diurnal amplitude is observed (Morris, 1985; Celleri et al., 2007). Due to the notable altitudinal gradient, mean annual rainfall oscillates in intensity and duration, with the lowest value of 660 mm at the centre of the basin and the highest observed value exceeding 3400 mm near the basin outlet. On the other hand, stations located at higher elevations (above 3000 m a.s.l.) receive between 1000 and 1400 mm (Celleri et al., 2007).

The PRB is rich in biodiversity and includes important extents of protected areas (Fig. 1). The most renowned are the Cajas National Park, located in the western, upper, extreme of the basin, which is a Ramsar-Convention wetland site, and the Sangay National Park, located at the north-eastern extreme; both recognised by the United Nations Educational, Scientific and Cultural Organization as World Heritage Sites.

Two major cities, namely Cuenca and Azogues are in the basin with approximately 500,000 and 33,850 inhabitants, respectively (according to the 2010 national census). In general, pollution in the study basin includes domestic wastewater, agricultural runoff, animal husbandry and industrial effluents (Da Ros, 1995; Sotomayor et al., 2018).

2.2. Sampling WQ descriptive variables

The information employed in the current research was gathered by the former Ecuadorian National Secretary of Water (SENAGUA) - Santiago River Hydrographic Demarcation (DHS) at 64 sampling sites within the study basin (Fig. 1) throughout a 5-year monitoring period

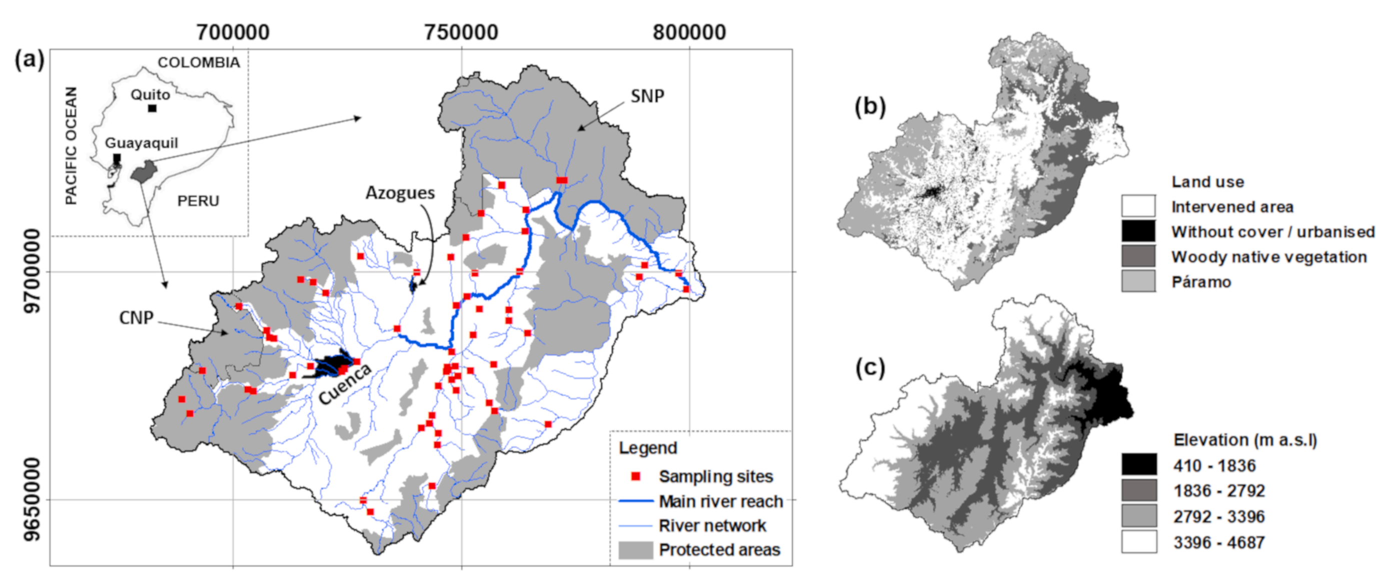

Fig. 1. (a) The Paute River basin in continental Ecuador, its two largest cities (Cuenca and Azogues) and the location of the 64 sampling sites; (b) land use cover; and (c) the digital elevation model of the basin (CNP = Cajas National Park; SNP = Sangay National Park).

(2008, 2010–2013). The provided database includes WQ information about physical, chemical, hydrological, geomorphological, and microbiological descriptive variables and benthic macroinvertebrates families. The descriptive variables were: aluminium (Al^{3+}), ammonium-nitrogen (NH_4^+-N), cadmium (Cd^{3+}), copper (Cu^{2+}), chloride (Cl^-), fluoride (Fl^-), iron (Fe^{2+} & Fe^{3+}), nitrate-nitrogen (NO_3^--N), lead carbonate ($PbCO_3$), pH, potassium (K^+), sodium (Na^+), total alkalinity (TALK), total hardness (TH), total phosphorus (P-tot), dissolved oxygen (DO), 5-day biochemical oxygen demand (BOD_5), faecal coliforms (FC), river slope (Slp), Shreve river order (Shreve), elevation (Elev.), electric conductivity (EC), total solids (TS), turbidity (TU), water temperature (WT) and the fluvial habitat index (FHI) of the Environmental Protection Agency, which is useful for assessing streambed and riparian habitats (Barbour et al; 1999). A multiparametric probe measured DO, EC, pH and WT in situ (model U52G10M, Horiba Ltd. Kyoto, Japan). For assessing the remaining physical, chemical and microbiological parameters, water samples were collected with sterile containers and later transported in iced coolers for laboratory analysis following standard methods (American Public Health Association, 1998).

On average, the monitoring sites were visited five times per year. Some were sampled more often, because they were located either at highly polluted sites or, on the contrary, at unaltered environmental (i.e., reference) locations. As a result, a WQ database was developed for the $n_{sp} = 64$ monitoring sites with $n_{rep} = 348$ sampling replicates of $n_v = 26$ WQ descriptive variables, resulting in a total of $n_{obs} = n_{rep} \times n_v = 9048$ observations, represented as $x_{i,j}$, where $i = 1, 2, \ldots, n_v$ and $j = 1, 2, \ldots, n_{rep}$ (Fig. 2). Table 1 list the main statistical properties of each one of the studied WQ descriptive variables.

2.3. Sampling and analysis of benthic macroinvertebrates

Benthic macroinvertebrates samples were collected at each of the 64 monitoring sites (Fig. 1) following the strategy proposed by Von Ellenrieder (2007). At each sampling site, a 20 m long reach was selected. Three evenly-spaced transects were delineated throughout this river reach. Macroinvertebrate samples were collected by kick-sampling, using a standard D-frame net (25 cm aperture; 0.5 mm mesh) (Jacobsen et al., 1997). Sampling was carried out for three minutes, encompassing all existing microhabitats characterised by different depths, substrates and water velocities. The three samples from each transect were pooled together, and sampling continued by visually inspecting (for about 20 min) the substrate and aquatic vegetation to collect the tightly clinging taxa that may have not been dislodged by kick-sampling (Roldán, 2003). Macroinvertebrate samples were preserved in 70% ethanol and sorted with the use of a stereomicroscope. $n_{fam} = 59$ families were identified and grouped into $n_{ord} = 18$ superior taxonomic groups (in its great majority orders).

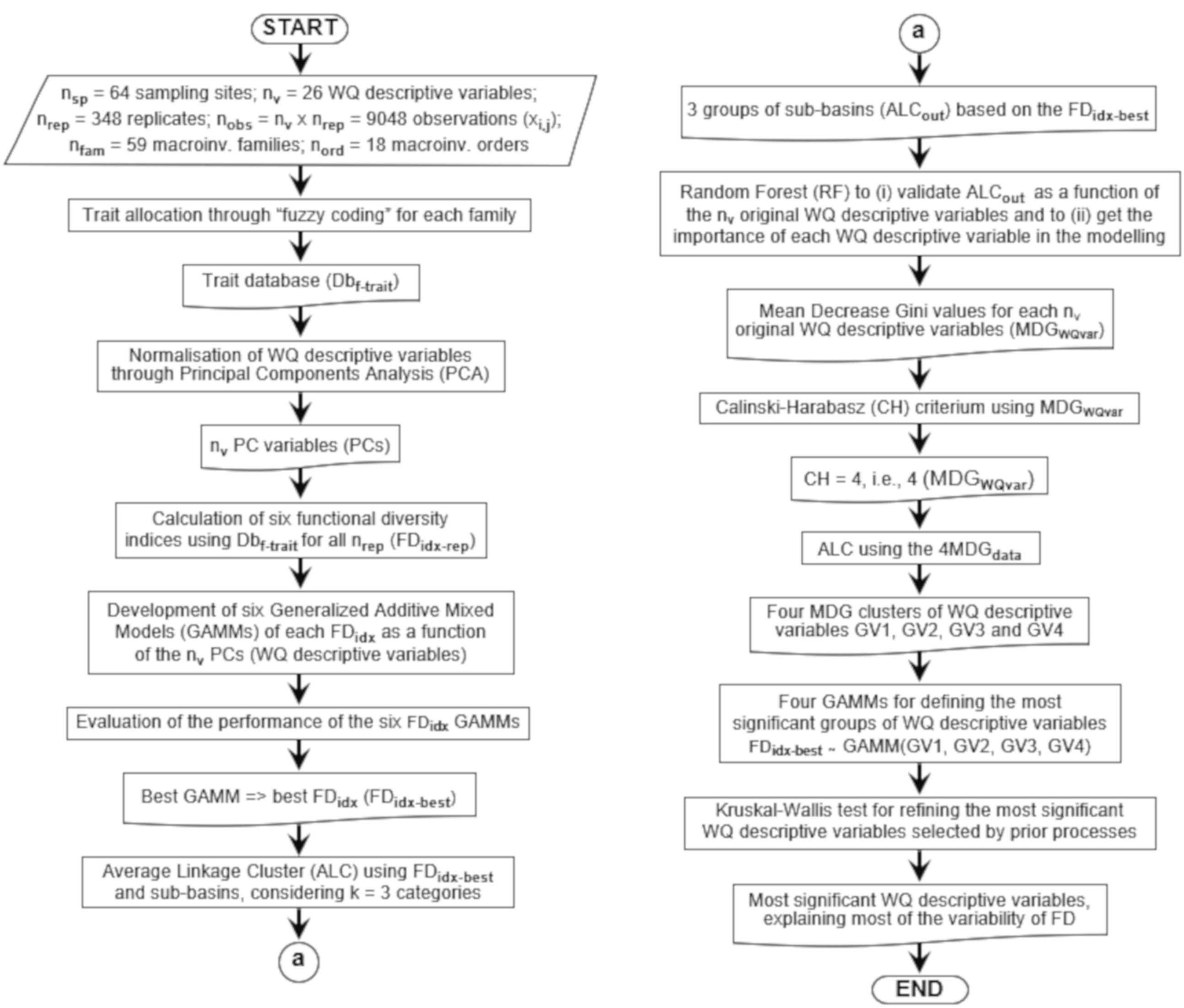

Fig. 2. Flowchart of the methodological steps implemented in the current study.

Table 1
Main statistical properties of the water quality (WQ) descriptive variables that were monitored throughout years 2008 and 2010–2013 in the Paute River basin (PRB), Ecuador (Fig. 1). Legend: Al^{3+} = aluminium; BOD_5 = 5-day biochemical oxygen demand; Cd^{3+} = cadmium; Cl^- = chlorides; Cu^{2+} = cooper; DO = dissolved oxygen; EC = electric conductivity; FC = faecal coliforms; Fe^{2+} & F^{3+} = iron; FHI = fluvial habitat index; Fl^- = fluorides; K^+ = potassium; NH_4^+-N = ammonium-nitrogen; Na^+ = sodium; NO_3^--N = nitrate-nitrogen; $PbCO_3$ = lead; P-tot = total phosphorus; Shreve = river order calculated with the Shreve method; Slp = slope; TALK = total alkalinity; TH = total hardness; TS = total solids; TU = turbidity; WT = water temperature; STD = standard deviation; and m a.s.l. = meters above sea level.

WQ Parameters		Mean	Median	STD	Range
Category	Parameter				
Habitat quality	FHI	130.10	129.00	28.50	71.0–184.0
Hydrogeomorphological	Elev (m a.s.l.)	2334.60	2325.40	773.50	433.7–3766.6
	Shreve	363.90	63.00	1062.30	1–126642.0
	Slp (%)	7.30	7.30	5.60	0.1–29.5
Microbiological	FC (bacteria 100^{-1} ml^{-1})	4962.00	1600.00	6338.10	0–16000.0
Physicochemical	Al^{3+} (mg L^{-1})	0.10	0.00	0.20	0–1.6
	BOD_5 (mg L^{-1})	7.70	3.80	9.00	0–55.8
	Cd^{3+} (mg L^{-1})	0.00	0.00	0.00	0–0.6
	Cl^- (mg L^{-1})	5.20	0.70	26.00	0–363.9
	Cu^{2+} (mg L^{-1})	0.00	0.00	0.10	0–1.3
	DO (mg L^{-1})	6.80	6.90	0.70	4.1–9.8
	EC (μS cm^{-1})	128.60	76.80	193.70	3.0–1810.0
	Fe^{2+} Fe^{3+} (mg L^{-1})	0.20	0.00	0.50	0–3.7
	Fl^- (mg L^{-1})	1.40	0.40	6.40	0–67.9
	K^+ (mg L^{-1})	3.00	0.40	14.60	0–227.1
	Na^+ (mg L^{-1})	6.00	3.50	11.90	0–112.9
	NH_4^+-N (mg L-1)	0.80	0.00	1.50	0–15.0
	NO_3^--N (mg L-1)	0.60	0.10	2.10	0–20.8
	$PbCO_3$ (mg L^{-1})	0.00	0.00	0.10	0–0.9
	pH	7.60	7.60	0.70	5.3–9.4
	P-tot (mg L^{-1})	0.60	0.20	1.00	0–4.8
	TALK (mg L^{-1})	14.40	0.10	23.00	0–77.8
	TH (mg L^{-1})	37.30	28.10	51.50	0–657.0
	TS (mg L^{-1})	2.00	0.00	10.20	0–116.0
	TU (NTU)	18.50	1.00	82.90	0–1136.8
	WT (°C)	14.70	14.10	3.30	8.7–23.5

2.4. Allocation of biological traits

Seven biological trait categories were selected (Table 2). They are well-known as good indicators for catching compositional variations due to ecohydrological stress (Tomanova & Usseglio-Polatera, 2007; Ding et al., 2017; Reynaga et al., 2020). A biological trait database at the family level ($Db_{f\text{-}trait}$) was constructed (Fig. 2) based on the affinity of each family for trait categories using affinity scores within a fuzzy coding procedure (Chevenet et al., 1994) upon the information gathered from Sotomayor et al. (2021) for the PRB. The affinity scores, provided by Sotomayor et al. (2021), vary in the [0, 3] interval, where a score value of 0 indicates that there is no affinity of the taxon for the trait category and a score of 3 indicates total affinity.

2.5. Assessing functional diversity indices

Six functional diversity indices (FD_{idx}) were calculated using $Db_{f\text{-}trait}$ for each one of the replicates ($FD_{idx\text{-}rep}$), i.e., (1) Functional diversity based on dendrograms (FDc) (Petchey & Gaston, 2006); (2) Functional diversity based on dendrograms including the taxa community pool (wFDc) (Pla et al., 2008); (3) Functional dispersion (FDis) (Laliberte & Legendre, 2010); (4) Functional richness (FRic) (Villéger et al., 2008); (5) Rao's quadratic diversity index (Rao) (Rao, 1982); and (6) Rao's quadratic diversity index relative to maximum (rRao) (Pavoine et al., 2005). For FDc and wFDc, both dendrograms were computed using the Ward clustering algorithm and Euclidean distance (Roa-Fuentes et al., 2022). All these indices are dimensionless. Casanoves et al., (2010) provide a complete description of these FD indices. In this regard, some authors claim that species-resolution identification is required when using diversity metrics such as FD indices (Dalu et al., 2017). However, in line with other similar studies (Ferraro & Cole, 1995; Corbi & Trivinho-Strixino, 2006; Mueller et al., 2013; Bo et al., 2020; Eriksen et al., 2021), Sotomayor et al. (2021) found that benthic macroinvertebrates families are suitable for streams bioassessment in the PRB using FD indices.

2.6. Statistical analysis

To identify the adequate FD_{idx} a comparative analysis was performed between the six FD indices. In line with prior research (Gallardo et al., 2011; Koperski, 2019), the multiple regression method Generalised Additive Mixed Models (GAMMs; Lin & Zhang, 1999) was applied to relate each FD_{idx} with the 26 WQ descriptive variables as independent predictors. The GAMM with the best fit was considered the best FD_{idx} ($FD_{idx\text{-}best}$) for assessing the health of the streams within the PRB. A normalisation process through a Principal Component Analysis (PCA) of WQ descriptive variables was carried out to perform the comparative analysis between the six GAMMs.

2.6.1. Modelling FD indices with Generalised Additive Mixed model (GAMM)

PCA transformed the original independent variables into new synthetic variables (PCs) that are mutually orthogonal and provide maximum information without being redundant. The number of PCs equalled the number of independent original variables ($n_v = 26$) (Einax et al., 1997). The PCs were used as independent variables in GAMMs, a process known as Principal Component Regression (PCR) (Varmuza & Filzmoser, 2010) that has been widely used (Çamdevýren et al., 2005; Ul-Saufie et al., 2011; Haque et al., 2013; Shuman et al., 2020) to create a mathematical framework for comparing the six GAMMs.

GAMMs are flexible regression tools that are increasingly used in water resources studies to assess spatial and temporal trends in complex monitoring data sets (Gardner, 2007; Mellor & Cey, 2015; Wood et al., 2017; Iddrisu et al., 2017). A GAMM provides a method for fitting non-linear covariate effects via the inclusion of smooth functions to represent a more complex relationship between predictors (i.e., PCs) and the response variable ($FD_{idx\text{-}rep}$). The repeated measurements on sampling sites (i.e., n_{rep}) induce a structure in the data that violates the

Table 2
Traits and categories used in the study.

Trait	Category
Feeding habits	Collector-Filterer (C-Ft)
	Collector-Gatherer (CG)
	Piercers (Pc)
	Predators (Pr)
	Scrapers (Sc)
	Shredders (Sh)
	Parasite (PA)
Respiration	Tegument (Teg)
	Gill
	Plastron (Pla)
	Spiracle (Spi)
Body form	Streamlined (Str)
	Flattened (Flat)
	Cylindrical (Cy)
	Spherical (Sph)
Maximum body size (mm)	<2.5
	2.5–5
	5–10
	10–20
	20–40
	40–80
	>80
Body flexibility	None (<10°)
	Low (>10° – 45°)
	High (>45°)
Mobility and attachment to the substratum (locomotion)	Flier (Fli)
	Surface swimmer (SS)
	Full water swimmer (FWS)
	Crawler (Cra)
	Burrower (Bur)
	Temporarily attached (TA)
Reproduction	Asexual (As)
	Clutches & cemented (CC)
	Clutches & free (CF)
	Clutches in vegetation (CV)
	Clutches & Terrestrial (CT)
	Isolated eggs & clutches (IEC)
	Isolated eggs & free (IEF)
	Ovoviviparity (Ovi)

assumption of independence between samples. To deal with this, the GAMM method was chosen because it has the advantage of relaxed independence assumptions (Wang & Goonewardene, 2004; Polansky & Robbins, 2013; Ingersoll et al., 2016) by considering in every GAMM an additional random load that in this study was provided using as predictor a sampling site identification label beside the PCs data. To estimate the (population) parameters of each GAMM, the Restricted Maximum Likelihood (REML) estimation process (Gardner, 2007) was used. For further details about GAMMs see Chen (2000), Zahro & Caraka (2017) and Aljoumani et al., (2019).

2.6.2. Choosing the most adequate FD index by selecting the best GAMM-based model

The choice of the most adequate GAMM was based on the use of the square of the Person correlation coefficient (R^2) that is usually used for measuring the congruency between observed and fitted FD values (Iddrisu et al., 2017; Wood et al., 2017) and the examination of histograms of best-fit model residuals (Gardner, 2007; Montgomery et al., 2012; Mellor & Cey, 2015; Aljoumani et al., 2019), besides the implementation of the Shapiro-Wilk test (S-W) (Shapiro & Wilk, 1965) for normality check of model residuals (Hastie & Tibshirani, 1990; Shadish et al, 2014; Mellor & Cey, 2015; Laanaya et al., 2017) considering a 95% confidence level. The GAMM with the highest R^2 and symmetry of residuals was chosen as the most adequate and, as such, the most suitable FD_{idx} for the PRB ($FD_{idx\text{-}best}$). The resulting GAMMs processes were cross-validated using the Venetian blinds method, which is based on the use of segments (splits or groups) of observed data (Ballabio & Consonni, 2013; Mevik & Wehrens, 2015; Rácz et al., 2018). Thus, the number of replicates ($n_{rep} = 348$) was split into 5 groups for cross-validation (i.e., 4 groups of 70 observations for model training and the fifth group of 68 observations for model validation), implying that each GAMM used about 80% of the data for training and 20% for validation. This low number of splits was chosen to aim at avoiding the potential over-estimation of the predictive capability of the model (i.e., "overfitting") (Rani et al., 2019). The same set of replicates of size n_{rep} was used for each GAMM run for the training and validation steps.

2.6.3. Clustering the sub-basins as a function of the selected FD index

Once ($FD_{idx\text{-}best}$) was selected, a hierarchical agglomerative cluster analysis (Sokal & Michener, 1958) through Average Linkage Clustering (ALC) was carried out. The ALC was performed employing the $FD_{idx\text{-}best}$ values through the square Euclidean distance measure (Carter et al., 1989). The ALC used the average of the distance values between pairs of clusters (Frank & Todeschini, 1994) to group the sub-basins as a function of the $FD_{idx\text{-}best}$. The ALC aimed to find clusters of sub-basins in the FD-dimensional space, based on the similarity criterion among them and their sampling sites. Sotomayor et al. (2021) found that for the PRB, a maximum of three WQ classes should be employed when the FD is calculated using family taxonomical identifications. Hence, prior to ALC, the number of clusters was defined as equal to three.

2.6.4. Assessing the ALC outputs using the random forest classification algorithm

As a result of the ALC, three groups of sampling sites were identified, in correspondence with three groups of sub-basins of the PRB. The class label (i.e., 1, 2 and 3, representing high, medium and low FD levels, henceforth termed, C1, C2 and C3, respectively), corresponding to these three sampling sites clusters was used as dependent-categorical variable in the Random Forest (RF) algorithm (Breiman, 2001), as a function of the WQ descriptive variables. RF builds many single non-correlated decision trees (Breiman et al., 1984), i.e., a set of hierarchically organised restrictions or conditions which are successively applied from a root (parent) node to a terminal (or child) node or leaf of a tree to make repeated predictions of the phenomenon. When building each tree, RF uses a training subset that is randomly chosen and then replaced for a number of times equal to the number of trees in the ensemble. It means that each decision tree uses the Bootstrap Aggregation method, i.e., approximately two-thirds of the training samples are employed for prediction while the remaining roughly one-third of the training samples are employed for validation of prediction accuracy. Meanwhile, for each node/split in a decision tree, the RF algorithm selects a random subset of predictors as candidates for splitting. Predictions are built by averaging over the predictions made by each tree in the forest (Louppe et al., 2013).

RF can be used for both regression and classification problems; herein, the focus was on classification tasks. The parameters that were tuned for estimating the RF model were the number of trees (n_{tree}), and the number of variables randomly selected at each node, i.e., mtry (Fox et al., 2017). Herein, for parameterising the RF model, the strategy suggested by Strobl et al. (2007), Strobl and Robillard (2008) was implemented, which basically is a grid search, in which all possible combinations of given discrete parameter regions are evaluated. The RF models/runs were insensitive to the selection of mtry; however, values of n_{tree} greater than the default (i.e., 500) produced more consistent results. Thus, for this study it was adopted the default mtry value (i.e., $\sqrt{n_v}$) and $n_{tree} = 3000$. The RF was implemented through a cross-validation process using 80% of the data for training and 20% for testing. Further, since the generated response variable was imbalanced, equal importance was pre-defined for each class (Khalilia et al., 2011; Brown & Mues, 2012; Larras et al., 2017) to avoid the biggest class winning the most votes (Boulesteix et al., 2012; Rebala et al., 2019). For evaluating whether the RF classification algorithm correctly allocated sampling sites to the three FD classes, the non-error rate (NER) classification measure was used. NER is the average of the sensitivity values of all classes (Ballabio & Consonni, 2013), with the sensitivity of the g-th class (Sn_g) being the model's ability to correctly recognise sampling sites belonging to the g-th class. It is defined as the ratio between n_{gg} and the total number of sampling sites belonging to the g-th class (n_g). The NER provides an overall evaluation of the classification quality. Thus, to evaluate the classifier performance assigning every specific class, the Youden Index (Ψ; Youden, 1950) was calculated (Ballabio et al., 2018). $\Psi = Sn_g - (1 - Sp_g)$, with Sp_g being the specificity. The Sp_g represents the capability of a given classifier to reject samples of the other classes. It is calculated as the ratio of samples, which were not classified in the g-th class over the total number of samples not belonging to the g-th class (Sokolova et al., 2006). The variation range of NER and Ψ is between 0 and 1, with 1 being their optimal value.

2.6.5. Assessing the significant WQ descriptive variables using the RF algorithm

To evaluate the importance of each WQ descriptive variable in predicting the FD classes, i.e., C1, C2 and C3, the Mean Decrease Gini (MDG) was used (Han et al., 2016), which measures the classification impact of variables by totalling the amount of decrease in impurity as the classification is performed (Breiman, 2001, 2002). Once the MDG was calculated for each WQ descriptive variable, the output was a dataset ranked from highest to lowest MDG values but without statistical criteria (e.g., a cut-off threshold) for distinguishing the variables that are significant for explaining the variability of the FD classes. Hence, the (Caliński & Harabasz, 1974) index (CH) was computed to estimate groups of WQ descriptive variables as a function of their MDG. This index is the ratio between the between-cluster variance and the within-cluster variance. The larger the CH value, that is, the larger the between-cluster variance and the smaller the within-cluster variance, the better the data (i.e., WQ descriptive variables) grouping. The highest CH value corresponds to the most compacted group of WQ descriptive variables. Then, an ACL analysis was performed using the MDG data with the predefined number of clusters (NC = 4) based on the highest CH value (herein the CH = 4). Further, four testing/validation processes using GAMMs were employed, corresponding to the four MDG clusters produced by the ALC (henceforth termed groups of variables GV1, GV2, GV3 and GV4). In the GAMMs, the dependent variable was always the $FD_{idx\text{-}best}$ and the independent variables were either: (i) all the groups of WQ descriptive variables, i.e., GV1, GV2, GV3 and GV4 (GAMM-1); (ii) WQ descriptive variables belonging to GV1 (GAMM-2); (iii) WQ descriptive variables belonging to GV1 and GV2 (GAMM-3); and (iv) WQ descriptive variables belonging to GV1, GV2 and GV3 (GAMM-4). For GAMM-1, GAMM-2, GAMM-3 and GAMM-4 no cross-validation was performed, and their random effects were based on Shreve order, elevation and slope.

The analysis for pre-selecting the most significant WQ descriptive variables was based on the calculation of the R^2 goodness of fit statistic upon the comparison of the predictions of either GAMM model with the respective $FD_{idx\text{-}best}$ value. Hence, the GVs (and all their WQ descriptive variables) explaining most of the variability of $FD_{idx\text{-}best}$ were identified upon the consideration of two criteria. The first criterion was based on the comparison of the R^2 values of GAMM-2, GAMM-3 and GAMM-4 regarding the R^2 value achieved by GAMM-1 (i.e., when using all the WQ descriptive variables); that is, by computing the R^2 gain when incorporating a given GV into the $FD_{idx\text{-}best}$ modelling, with respect to the R^2 value associated with GAMM-1. The higher R^2 gains enabled pre-selecting the most significant groups of WQ descriptive variables (in explaining most of the variability of the $FD_{idx\text{-}best}$). Further, the second criterion assumed that every WQ descriptive variable, belonging to a given GV, contributes evenly to the GV explanation of the variability of $FD_{idx\text{-}best}$; the number of variables belonging to every GV was, then, the second criterion considered in this study, complementary to the R^2 criterion, for selecting the GVs explaining most of the variability of $FD_{idx\text{-}best}$.

Further, the non-parametric Kruskal-Wallis test (K-W) (Kruskal & Wallis, 1952), was carried out for enhancing the selection of the most significant WQ descriptive variables by defining whether the pre-selected WQ descriptive variables significantly differ in their three FD classes C1, C2 and C3. Herein, for each WQ descriptive variable, the set of sampling sites and their replicates was divided into the three FD classes, after which the K-W test was carried out with the null hypothesis, H_0, being that the medians within each of the three FD classes are the same. Thus, if the p-value of the K-W test was greater or equal than 0.05, then it was concluded that there was no significant difference amongst the three FD classes at the 95% confidence level. On the contrary, if $p < 0.05$, it was concluded that there was a significant difference. WQ descriptive variables that showed no significant difference among their three FD classes were finally considered of little importance to explain the variability of the FD in the PRB (Sotomayor et al., 2020), even though, they might have been part of a pre-selected group of WQ descriptive variables. Finally, Fisher's least significant difference (LSD) test was used to calculate/visualise intervals around the means of the most significant WQ descriptive variables, as a function of the FD classes for visualisation of the populations whose means are statistically different from each other (Williams & Abdi, 2010).

The FD metrics and the ALC were calculated/implemented with the FDiversity software (Casanoves et al., 2011). The PCA was implemented with MATLAB® using the PCA toolbox version 1.3 (Ballabio, 2015). The GAMMs were calculated with the R® package "gamm4" (Wood & Scheipl, 2020). The RF was executed using the R® package "randomForest" (Liaw & Wiener, 2015) and the CH, K-W, S-W and LSD tests were implemented with MATLAB® using subroutines particularly developed for this study.

3. Results

3.1. Functional structure of the benthic macroinvertebrate communities

The dominant category for each biological trait of the benthic macroinvertebrate communities of the Paute river basin were (Fig. 3): (a) for the feeding habits: the Collector-Gatherers (CG); (b) for respiration: the Gills; (c) for the body form: Cylindrical (Cy); (d) for the maximum body size: range from 5 to 10 mm; (e) for the body flexibility: greater than 45°; (f) for the mobility and attachment to the substratum (locomotion): the Crawler (Cra); (g) for reproduction: the Clutches & Cemented (CC).

3.2. Choosing the most adequate FD index by selecting the best GAMMs

The R^2 index values that were obtained for the GAMMs of FD_{idx}, as a function of the WQ descriptive variables, were: $R^2_{FDc} = 0.29$, $R^2_{wFDc} = 0.27$, $R^2_{Rao} = 0.19$, $R^2_{rRao} = 0.11$, $R^2_{FDis} = 0.18$ and $R^2_{FRic} = 0.21$. Fig. 4 shows the results of the assessment of the distributions of residuals corresponding to every one of the six GAMMs, concretely, the histograms of residuals, plotted together with the respective fitted normal distribution. Even though the residual trends of wFDc, FDc, Rao, rRao and FDis follow a normal distribution, a decreasing trend of p-values obtained in the scope of the Shapiro-Wilk (S-W) test of normality (i.e., a detriment of residual symmetry) was observed. For FRic the residual asymmetry was confirmed ($p < 0.05$), implying that the normal distribution (ND) assumption is false. It should be observed that in the case of the FDis index, the fitted normal distribution of residuals is in the limit of statistical significance ($p = 0.05$). Indeed, visual inspection suggests that the residual normality significance is very weak.

The performance of the GAMM-based model of FDc was regarded as the best one, although the associated R^2 value ($R^2_{FDc} = 0.29$) is only marginally better than the respective performance index for wFDc (reflected by $R^2_{wFDc} = 0.27$). Further, the range of variation of the residuals associated with the GAMM-based model of wFDc (Fig. 3a) is narrower than the respective range of variation of residuals associated with the GAMM-based model of FDc (Fig. 4c). Finally, the normality test suggested the best fit to a normal distribution (p-value = 0.89) of the

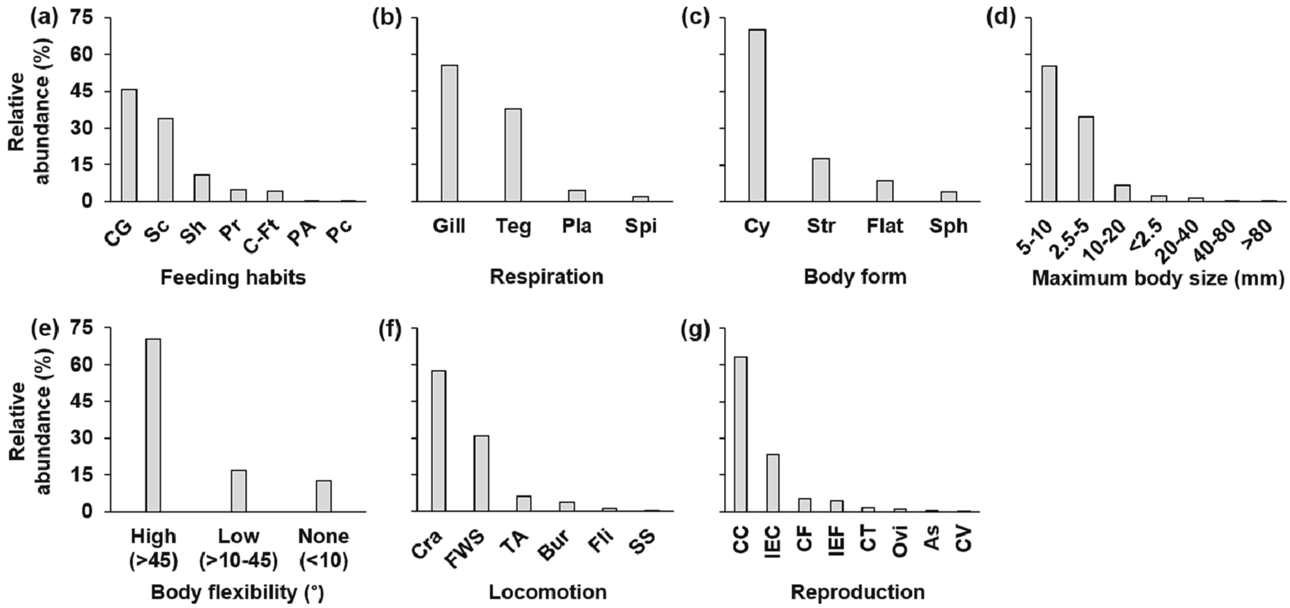

Fig. 3. Relative abundance as a function of the following traits: (a) feeding habits; (b) respiration; (c) body form; (d) maximum body size; (e) body flexibility; (f) mobility and attachment to the substratum (locomotion); and (g) reproduction. C-Ft = Collector-Filterer; CG = Collector-Gatherer; Pc = Piercers; Pr = Predators; Sc = Scrapers; Sh = Shredders; PA = Parasite; Teg = Tegument; Pla = Plastron; Spi = Spiracle; Str = Streamlined; Flat = Flattened; Cy = Cylindrical; Sph = Spherical; Fli = Flier; SS = Surface swimmer; FWS = Full water swimmer; Cra = Crawler; Bur = Burrower; TA = Temporarily attached; As = Asexual; CC = Clutches & cemented; CF = Clutches & free; CV = Clutches in vegetation; CT = Clutches & Terrestrial; IEC = Isolated eggs & clutches; IEF = Isolated eggs & free and Ovi = Ovoviviparity.

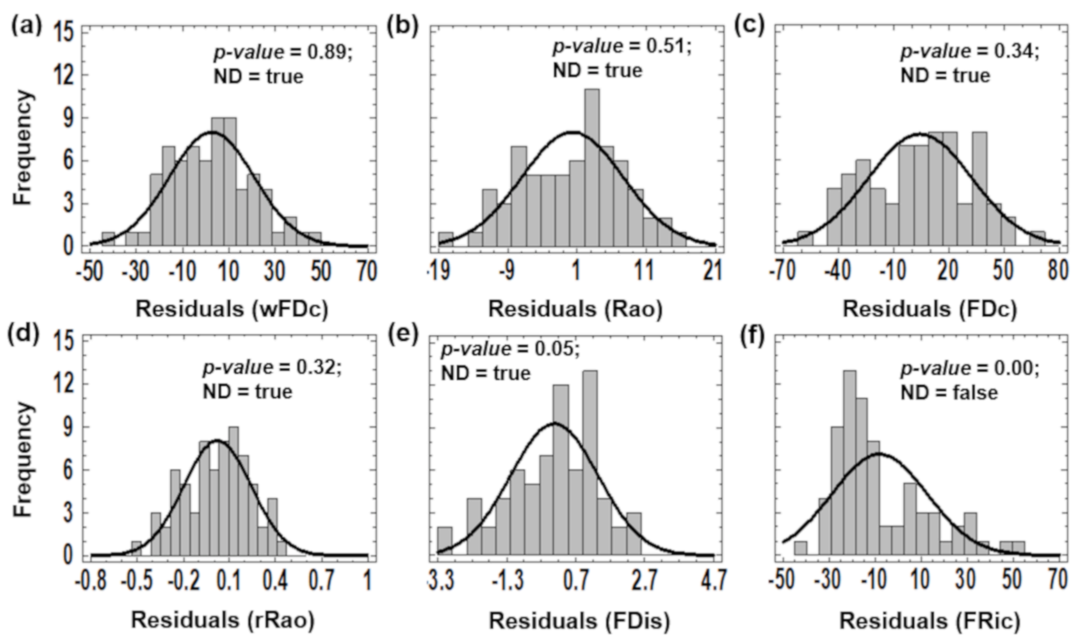

Fig. 4. Histograms of residuals of the GAMMs of every one of the six functional diversity indexes: (a) functional diversity based on dendrograms including the species community pool; (b) Rao's quadratic diversity index; (c) functional diversity based on dendrograms; (d) Rao's quadratic diversity index relative to maximum; (e) functional dispersion; and (f) functional richness. The fitted normal distribution was plotted on top of the histogram of the residuals. The p-value of the Shapiro-Wilk test for normality is given for each set of residuals (with a significance level, 1 – α = 0.95). ND = normal distribution.

residuals associated to the wFDc model (Fig. 4a). Considering these, wFDc was chosen as the most adequate FD index for bioassessment in the PRB. The wFDc index varied between 0.0 and 112.4 with an average value of 38.6 and a standard deviation (STD) of 21.5.

3.3. Clustering the sub-basins as a function of the selected FD index

Fig. 5a shows the dendrogram produced by the ALC test with the three well-differentiated sub-basin groups C1, C2 and C3 (i.e., high, medium, and low FD levels, respectively) as a function of the selected index, wFDc. The dendrogram includes the explicit identification of the subbasins of the study basin by means of a code that is indicated in the dendrogram between parentheses. Hereafter, Fig. 5b illustrates the spatial distribution of the respective sub-basins including the sub-basin identification codes. Coloured sub-basins included in Fig. 5b indicate no-data availability; as such, they do not have an identification code. Fig. 5c

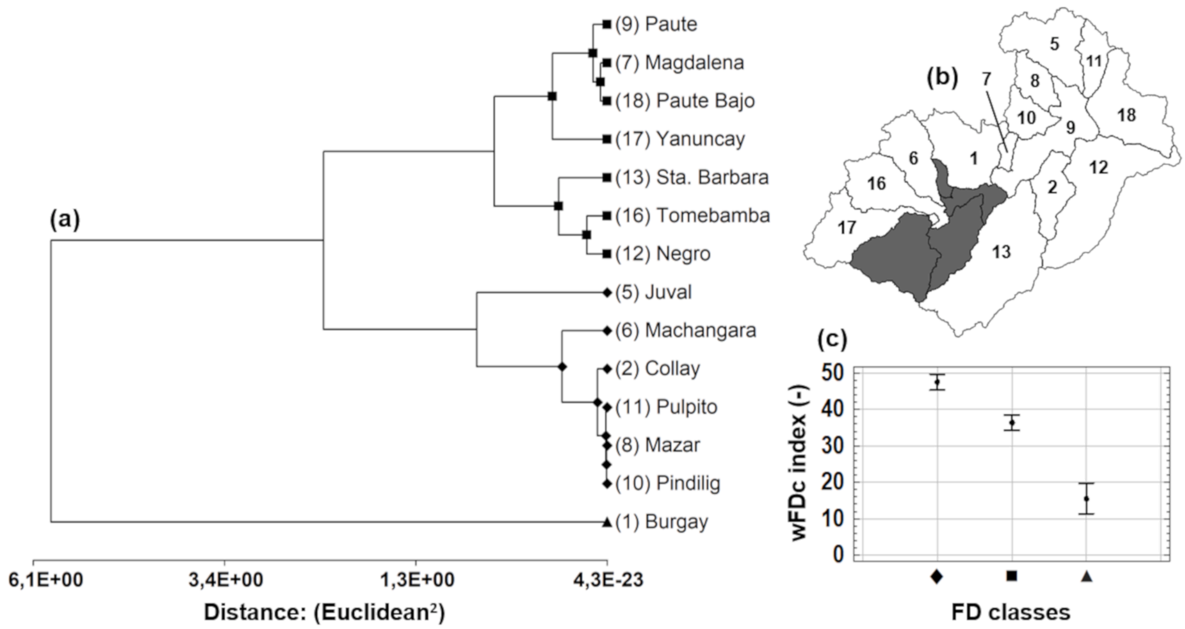

Fig. 5. (a) Dendrogram of cluster analysis showing three main groups of sub-basins as a function of the functional diversity (FD) index based on dendrograms including the species community pool (wFDc) and a sub-basin identification code (depicted between parentheses); (b) map of sub-basins identification codes used in the dendrogram (sub-basins in colour denote no-data availability); and (c) mean wFDc index values and respective Fisher's-based intervals for each FD group of sub-basins. The symbology used in (c) matches the respective symbology used in (a), i.e., ◆ = C1; ■ = C2 and ▲ = C3.

shows the outcome of Fisher's least significant difference test for the wFDc dataset as a function of the three (FD) sub-basins groups. The symbology used for the (FD) sub-basins classes C1, C2 and C3, is matching the respective symbology used in the dendrogram, which enables linking the results of the ALC analysis (Fig. 5a) with the respective results of Fisher's test (Fig. 5c). The average values of wFDc and the respective Fisheŕs interval limits are 47.4 ± 2.2 (C1); 36.3 ± 2.1 (C2); and 15.5 ± 4.2 (C3). The respective wFDc standard deviation values are 21.4 (C1); 18.7 (C2); and 9.3 (C3).

Regarding the performance of the ALC outputs using the RF algorithm, the overall classification achieved a very acceptable NER value of 0.89, whilst the assessment of the performance of the RF method classifying every specific group of sub-basins revealed relatively low differences between the groups, with C3 having the lowest Ψ value, i.e., $\Psi_{C3} = 0.77$ followed by $\Psi_{C1} = 0.85$ and $\Psi_{C2} = 0.95$, resulting from values of sensitivity (Sn_g) and specificity (Sp_g) greater than 0.7 which suggest the acceptability of the RF algorithm (i.e., $Sn_{gC1} = 0.91$, $Sp_{gC1} = 0.94$; $Sn_{gC2} = 0.97$, $Sp_{gC2} = 0.97$; $Sn_{gC3} = 0.80$, $Sp_{gC3} = 0.97$).

3.4. Assessing the most significant WQ descriptive variables

The CH values corresponding to the variation of the number of clusters between 2 and 4 were respectively 88.4 (NC = 2), 105.2 (NC = 3) and 174.4 (NC = 4). Thus, the optimal NC value was defined as 4. Fig. 6 shows the outcome of the ALC analysis, namely, the MDG distribution for the WQ descriptive variables included in this study, as well as

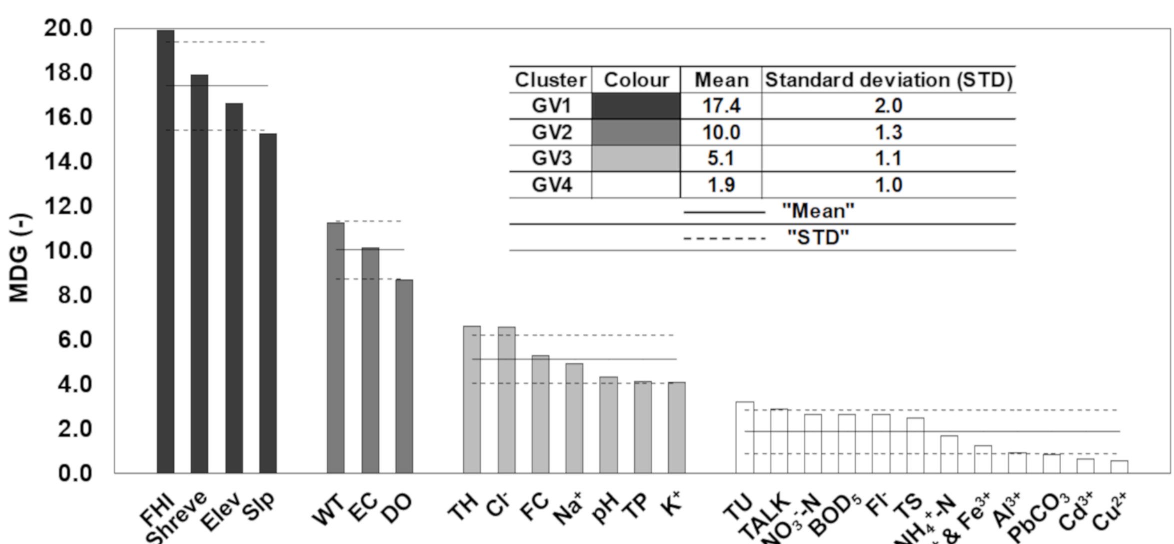

Fig. 6. Distribution of the Mean Decrease Gini (MDG) index for the study water quality descriptive variables and the respective four clusters (GV1, GV2, GV3 and GV4) identified as a function of MDG, together with the MDG mean and standard deviation (STD) values.

the respective four groups of WQ descriptive variables (GV1, GV2, GV3 and GV4). In the figure, every one of these groups is characterised by a MDG mean and a standard deviation. The MDG mean decreases from GV1 to GV4. Every group contains a different number of WQ descriptive variables. The first two groups contain a low number of variables and the last two groups a higher number of variables. When considering only the MDG values, the GV4 variables are the less informative in terms of explaining the variability of FD in the study basin.

The R^2 model performance index, characterising the predictions of the GAMMs, and used as a first criterion for selecting the GVs that are explaining most of the variability of wFDc, adopted the following values: 0.29 (GAMM-1), 0.22 (GAMM-2), 0.24 (GAMM-3) and 0.26 (GAMM-4). Thus, in proportion to the R^2 value obtained when all the WQ descriptive variables are included in the GAMM (i.e., 0.29), GV1 explained about 75.9% of the total variability of the wFDc, GV2 explained about 6.9% of this variability, GV3 about 6.9% and GV4 about 10.3%. The second criterion for the selection of the WQ descriptive variables that are explaining most of the wFDc variability, namely, the number of variables belonging to every GV, showed that every one of the 4 variables of GV1 explained about 19.0% of the variability of wFDc, each one of the 3 variables of GV2 about 2.3%, each one of the 7 variables of GV3 about 1.0%, and each one of the 12 variables of GV4 about 0.9%. Thus, the simultaneous consideration of both criteria leads to the conclusion that the 12 variables of GV4 are less significant for explaining most of the variability of wFDc. Furthermore, the K-W test indicated that the mean values of Na^+ and K^+ (i.e., members of GV3) for the three FD classes C1, C2 and C3, were not significantly different (p-values > 0.05); thereby, they were considered non-significant in explaining most of the variability of the FD in the study basin. Thus, 12 WQ descriptive variables from GV1, GV2 and GV3 were finally considered relevant in explaining the variability of FD in the PRB.

Hereafter, Fig. 7 shows the results of Fisheŕs test for these 12 most significant WQ descriptive variables as a function of the three FD classes. Every plot of the figure depicts the mean values and respective Fisher's-based intervals of the variables. For some WQ descriptive variables, clear trends are observed. For instance, for FHI ($FHI_{C1} = 142.1$, $FHI_{C2} = 125.7$ and $FHI_{C3} = 103.8$) and slope (Fig. 7a and 7d) values decline from C1 to C3. For dissolved oxygen ($DO_{C1} = 6.87$ mg L^{-1}; $DO_{C2} = 6.99$ mg L^{-1} and $DO_{C3} = 6.11$ mg L^{-1}) the C3 value was lower (Fig. 7g). Contrarily, values for electric conductivity ($EC_{C1} = 76.76$ μs cm^{-1}; $EC_{C2} = 137.52$ μs cm^{-1} and $EC_{C3} = 281.99$ μs cm^{-1}; Fig. 7f), total hardness ($TH_{C1} = 28.9$ mg L^{-1}; $TH_{C2} = 38.9$ mg L^{-1}; $TH_{C3} = 62.2$ mg L^{-1}; Fig. 7h), chlorides ($Cl^-_{C1} = 1.0$ mg L^{-1}; $Cl^-_{C2} = 4.9$ mg L^{-1}; $Cl^-_{C3} = 21.8$ mg L^{-1}; Fig. 7i), faecal coliforms ($FC_{C1} = 4835.7$ bacteria 100^{-1} ml^{-1}; $FC_{C2} = 3669.7$ bacteria 100^{-1} ml^{-1}; $FC_{C3} = 10656.9$ bacteria 100^{-1} ml^{-1}; Fig. 7j) and pH ($pH_{C1} = 7.4$; $pH_{C2} = 7.6$; $pH_{C3} = 7.8$; Fig. 7k), increase from C1 to C3. No specific patterns were observed for the rest of the variables; for instance, TP show similar average values between C1 and C3 ($TP_{C1} = 0.68$ mg L^{-1}; $TP_{C2} = 0.46$ mg L^{-1} and $TP_{C3} = 0.72$ mg L^{-1}; Fig. 7l).

4. Discussion

FDc and wFDc (Pla et al., 2008; Petchey & Gaston, 2006) were the most adequate indices to detect community changes according to the output of the GAMMs. These FD indices are defined as the total branch lengths of the functional dendrogram that can be constructed from the information of taxa functional traits. However, the main difference between FDc and wFDc is that FDc does not account for taxa weight in the community abundance, while wFDc is derived from a weighted dissimilarity matrix before creating the dendrogram (Casanoves et al., 2010; Seiz, 2015). It is likely that because of the inclusion of the relative abundance of benthic macroinvertebrates in the calculation of wFDc, the respective GAMM performed better. Gusmao et al. (2016) also reported that the inclusion of the abundance of macroinvertebrates in FD indices helps to distinguish better between contaminated and non-contaminated sites. On the other hand, FDis, FRic, Rao, and rRao, also incorporate in their calculation a measure of the relative abundance of taxa. However, the outperformance of wFDc, in the scope of the GAMM based modelling, suggests that its dendrogram-based origin is the main aspect for a more reliable FD assessment. Past studies (Brown & Milner, 2012; Martínez et al., 2013) reported more adequate FD assessments when using FDc than when using the Rao index. These findings indicate that in some cases (such as in the current research) the dendrogram-based indices are better for detecting community changes along a stress/pollution gradient. Other studies used the wFDc successfully, e.g., in paleolimnology (Nevalainen et al., 2015; Nevalainen & Luoto, 2017); for monitoring the restoration success in wetlands (Howard, 2019); to assess the FD of riparian forests (Lozanovska et al., 2018); in studies of tropical forest succession (Lohbeck et al., 2012); and for assessing the effect of increasing soil nutrient loads in FD of plants (Helsen et al., 2014). Nevertheless, to quantify/characterise the FD of freshwater macroinvertebrate assemblages, the Rao index is the most frequently used index (Schmera et al., 2017) and, to the best of our knowledge, there are no similar studies using the wFDc index of macroinvertebrates in an ecohydrological framework.

Sotomayor et al., (2021) found that the family taxonomic resolution, used in this study, was sufficient to assess 3 FD (i.e., WQ related) classes in the PRB. Thus, the ALC, using the wFDc index, identified three well-defined sub-basins groups, i.e., C1, C2 and C3 with high, medium, and low wFDc classes, respectively (Fig. 5). These groups are congruent with the findings of previous research, carried out in the same study basin, that reported similar monitoring sites grouping, based on the assessment

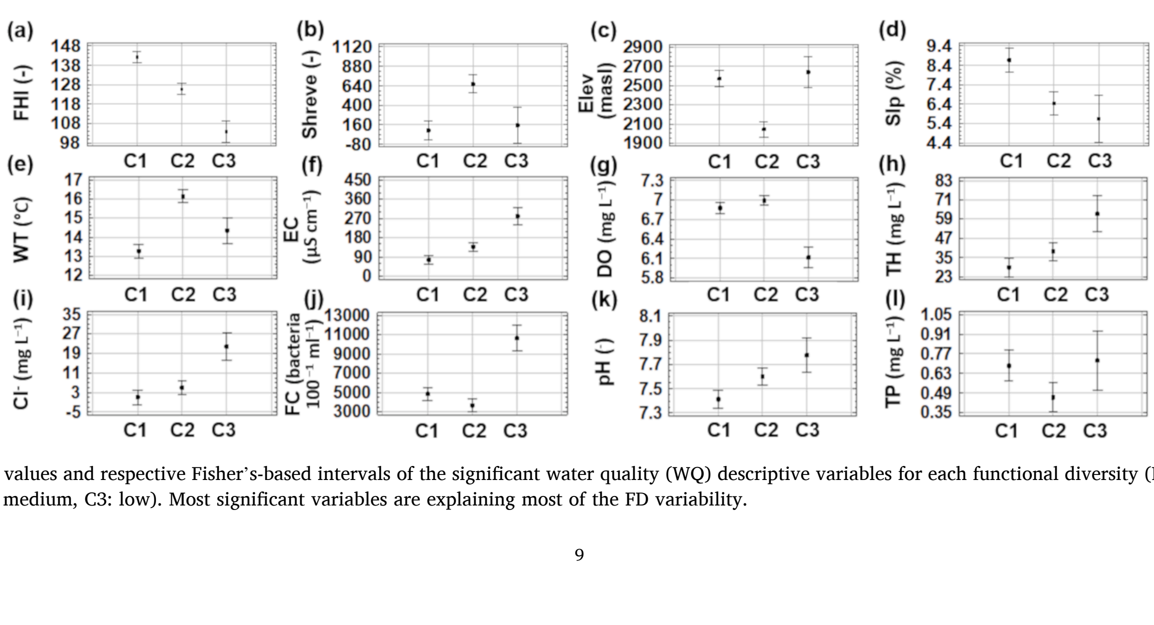

Fig. 7. Mean values and respective Fisher's-based intervals of the significant water quality (WQ) descriptive variables for each functional diversity (FD) class (i.e., C1: high, C2: medium, C3: low). Most significant variables are explaining most of the FD variability.

of ecological degradation and pollution gradient. For C3, i.e., the Burgay sub-basin (Fig. 5), Da Ros (1995), Pauta-Calle and Chang-Gómez (2014), Sotomayor (2016) and Sotomayor et al., (2018), Sotomayor et al. (2020) reported degraded WQ conditions. Domestic and industrial wastewater discharge, extensive agriculture, cattle ranching, and the loss of native vegetation cover are the main anthropogenic threats that cause the surface WQ pollution in this sub-basin and the subsequent loss of benthic macroinvertebrates taxa (Sotomayor et al., 2020). Both, C1 and C2 sub-basins groups (Fig. 5) are consistent with the findings of former studies (i.e., Sotomayor, 2016; Sotomayor et al., 2020). The sub-basins of C1 (Fig. 5) are in less anthropised parts of the study basin; on the contrary, the subbasins of C2, despite having some parts of their extents being protected areas, may be regarded as degraded (Fig. 1 and Fig. 5). The sub-basins of C2 were considered by Sotomayor et al., (2020) as hydrological systems that are on the way to a more serious surface WQ degradation. These findings emphasise the correspondence between taxonomic and functional approaches (Zhang et al., 2019).

The NER index revealed that, overall, 89% of the replicates were correctly classified by the RF method, which is considered a very good model performance (Xu et al., 2014). The performance of the RF method classifying every specific group was $\Psi_{C1} = 0.85$, $\Psi_{C2} = 0.95$ and $\Psi_{C3} = 0.77$. Since these Ψ values are greater than 0.6, the RF test can be considered as acceptable (Chen et al., 2015). Ψ_{C3}, being the lowest value, is likely to be connected to the fact that the data set is imbalanced as a result of having only one sub-basin in C3 (i.e., Burgay), whilst C1 and C2 have six and seven sub-basins, respectively. The latter implies that there is a low number of replicates in class C3 in comparison with the other classes. Although RF is assumed to be robust when dealing with imbalanced data sets in a classification framework (Khalilia et al., 2011; Brown & Mues, 2012; Larras et al., 2017), in this study, the imbalanced data condition seems to have resulted in the lower Ψ_{C3} value. Nevertheless, despite the latter, the current RF performance may be still considered acceptable in terms of the overall and per-class performance indexes, NER and Ψ.

The different tests that were implemented identified twelve significant WQ descriptive variables that were clustered in 3 groups (Fig. 6 and Fig. 7) as a function of their importance to explain the variability of FD of macroinvertebrates (Fig. 5). The most important variable was the fluvial habitat index (FHI) that declined from C1 to C3. This finding is congruent with other studies where the good conditions of riparian ecosystems and the streambed heterogeneity have been identified as key factors to maintain higher FD of benthic macroinvertebrates (de Castro et al., 2017; Riis et al., 2020). In general, the riparian zones ensure important processes such as allochthonous organic matter inputs (Magliozzi et al., 2020), fine sediments retention (Turunen et al., 2017), maintaining biodiversity (Firmiano et al., 2021) and mitigating the effects of anthropogenic pressure (Valera et al., 2019; Forio et al., 2022). Furthermore, the riparian forests create high diversity of benthic habitats in-streambed promoting adequate FD levels (Reid et al., 2010). This finding shows the crucial importance of riparian zones and streambed heterogeneity to maintain the macroinvertebrates FD in streams of the study basin.

Several hydro-geomorphological factors are important to explain the variability of the FD in the study basin. Shreve order, elevation, and slope are parameters that belong to the first Average Linkage Clustering (ALC) group of Mean Decrease Gini (MDG), together with FHI. These results are consistent with the findings of other studies, which have suggested that hydro-geomorphological factors like those mentioned can modulate/select different types of traits (Magliozzi et al., 2019), as such, influencing FD. The trend of Shreve order and elevation is the lack of significant difference between C1 and C3 (Fig. 7b and 7c). Thus, the river order and the altitudinal range are factors likely linked to the biogeographical circumstances/distribution of some traits in the study basin but are not necessarily related to low or high FD.

The second ALC group of WQ descriptive variables is formed by WT, EC and DO (Fig. 6). WT exhibits the lowest average value in C1 (Fig. 7e), which is explained by the reported higher values of FHI. Thus, the good condition of riparian ecosystems is a key WT regulator due to shade incidence on the stream channel (Beschta, 1997; Kalny et al., 2017; Roon et al., 2021). Sites with light availability (open canopy), in C2 and C3 sub-basin groups, had high WT values. The importance of WT to explain trends/selection of biological macroinvertebrates traits and thereby their FD has been already reported (Pallottini et al., 2017; Ding et al., 2017). EC exhibited increasing average values from C1 to C3 (Fig. 7f). DO values are similar in C1 and C2, whilst DO_{C3} is lower (Fig. 7g). EC increased at Burgay sub-basin as this hydro-system is enriched by both organic (i.e., FC; Fig. 7j) and inorganic pollution (i.e., TH, Cl^-; Fig. 7h-i). Organic pollution (i.e., FC because of cattle raising and poor sanitation practices) explains the reduction of the DO in C3 because high abundances of faecal coliforms and their metabolic activity have been reported as an important factor for oxygen consumption (De Troyer et al., 2016). A clear association between low FD values and reduced DO concentration has been reported before (Meng et al., 2021), which is in line with the results of the present study. For example, Paz et al., (2022) found that sites under organic pollution, i.e., reflected by low DO concentrations, mostly exhibited macroinvertebrate traits that potentially enabled tolerant species persistence.

Regarding the third group of WQ descriptive variables, average values of TH, Cl^-, FC and pH follow an increasing trend from C1 to C3 (Fig. 7h-k). Heino (2008) showed a negative relationship between the FD of littoral macroinvertebrate communities in lakes and high TH values. Further, the communities in hard water lakes supported some dominating functional groups that resulted in low FD levels. Similarly, this study found higher TH values in the sampling stations with lower FD. Also, high levels of Cl^- have been described as drivers of low Shannon index values (Moyo & Rapatsa, 2016). Taxonomic diversity, represented by the Shannon index, was found to be significantly correlated with the functional diversity of freshwater macroinvertebrates (Bazzanti et al., 2009; Schmera et al., 2017). Regarding pH, Leiva et al., (2022) reported slight increases associated with a deterioration of the FD of stream macroinvertebrates, similar to what was recorded in this study.

Based on the statement that for specific hydrological systems a crucial step is to determine which benthic macroinvertebrate metric/index performs best for WQ assessments, following a statistically sounded approach, this study developed a methodology for choosing a benthic macroinvertebrate FD metric for WQ evaluation. The FD metric selection for freshwater ecosystems is frequently arbitrary and based only on references from other hydrological systems, sometimes very different to the study basin. This could lead to biased conclusions about the ecological integrity status of rivers. This research recommends establishing an initial monitoring period and then implementing a protocol based on sound statistic methods for selecting an adequate FD metric. The proposed protocol also identifies significant WQ descriptive variables, which has the potential of reducing the number of variables to be monitored and, consequently, the monitoring time and related monetary expenses.

5. Conclusions

Using the benthic macroinvertebrates information, the functional trait approach and six functional diversity (FD) indices were applied as a way of studying the ecohydrological features of the Paute River Basin (PRB), Ecuador. The study proposes a statistical protocol to (i) assess and select an FD index; (ii) perform an adequate spatial clustering on the study site as a function of its FD levels, and (iii) choose the significant water quality (WQ) descriptive variables. The Generalised Additive Mixed Model (GAMM) of every one of the six FD indices was developed using 26 WQ descriptive variables (standardized). The GAMM with the best performance identified the community-based weighted FD index (wFDc) as the optimal FD metric for the study basin. The study results suggested that riparian ecosystems are likely to play a key role to maintain the stream's ecological integrity and their macroinvertebrate FD in the PRB. Likewise, hydro-geomorphological (i.e., slope), microbiological (i.e., FC) and chemical (i.e., EC, DO, TH, pH) aspects are likely to be crucial features to influence the macroinvertebrate FD in the study basin. There is a similar spatial clustering at the sub-basin scale between the current research using FD and past taxonomic studies that have been carried out in the study basin. The application of the proposed protocol on other similar hydrological systems may be relevant for assessing the WQ of rivers and categorising sites in terms of FD.

CRediT authorship contribution statement

Gonzalo Sotomayor: Conceptualization, Methodology, Software, Data curation, Formal analysis, Investigation, Writing – original draft. **Henrietta Hampel:** Methodology, Writing – original draft, Writing – review & editing. **Raúl F. Vázquez:** Methodology, Writing – original draft, Visualization, Writing – review & editing. **Marie Anne Eurie Forio:** Methodology, Writing – review & editing. **Peter L.M. Goethals:** Methodology, Writing – review & editing.

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.

Data availability

The data that has been used is confidential.

Acknowledgements

The authors would like to express their gratitude to the former Ecuadorian National Secretariat of Water (SENAGUA) - DHS for making the raw data accessible to the current study. This research was carried out in the context of (1) the Doctoral research of the first author, which is being performed at the Department of Animal Sciences and Aquatic Ecology, Faculty of Bioscience Engineering of the Ghent University, Belgium, supported by the VLIR Network Ecuador; and (2) the research Project "Análisis de la descomposición de la materia orgánica en ecosistemas acuáticos altoandinos del Austro Ecuatoriano", financed partly by the Vice-Presidency of Research (VIUC) of the Universidad de Cuenca, Ecuador.

References

Aljoumani, B., Kluge, B., Sanchez, J., Wessolek, G., 2019. Evaluating the variation of dissolved metals on a highway roadside using a generalized additive mixed model (GAMM). Water Air Soil Pollut. 230 (4), 1–15. https://doi.org/10.1007/s11270-019-4137-6.

Arenas-Sánchez, A., Dolédec, S., Vighi, M., Rico, A., 2021. Effects of anthropogenic pollution and hydrological variation on macroinvertebrates in Mediterranean rivers: A case-study in the upper Tagus river basin (Spain). Sci. Total Environ. 766, 144044 https://doi.org/10.1016/j.scitotenv.2020.144044.

American Public Health Association, 1998. Standard methods for the examination of water and wastewater, 20th edition. Washington, D.C.

Azab, A. (2012). Integrating GIS, Remote Sensing, and Mathematical Modelling for Surface Water Quality Management in Irrigated Watersheds (Delft University). https://doi.org/10.1201/b11797.

Ballabio, D., 2015. A MATLAB toolbox for Principal Component Analysis and unsupervised exploration of data structure. Chemom. Intel. Lab. Syst. 149, 1–9. https://doi.org/10.1016/j.chemolab.2015.10.003. November 2017.

Ballabio, D., Consonni, V., 2013. Classification tools in chemistry. Part 1: linear models. PLS-DA. Anal. Methods 5 (16), 3790–3798.

Ballabio, D., Grisoni, F., Todeschini, R., 2018. Multivariate comparison of classification performance measures. Chemom. Intel. Lab. Syst. 174, 33–44. https://doi.org/10.1016/j.chemolab.2017.12.004. February.

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 Edi). US Environmental Protection Agency, Office of Water, Washington, D.C.. EPA 841-B-99-002.

Bartram, J., & Ballance, R. (1996). Water Quality Monitoring - A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes (First Edit). https://doi.org/10.1159/000170272.

Bazzanti, M., Bella, V.D., Grezzi, F., 2009. Functional characteristics of macroinvertebrate communities in mediterranean ponds (Central Italy): Influence of water permanence and mesohabitat type. Ann. Limnol. 45 (1), 29–39. https://doi.org/10.1051/limn/09005.

Beschta, R.L., 1997. Riparian shade and stream temperature: Perspective. Rangelands 19 (2), 25–28.

Bo, T., Doretto, A., Levrino, M., Fenoglio, S., 2020. Contribution of beta diversity in shaping stream macroinvertebrate communities among hydro-ecoregions. Aquat. Ecol. 54 (4), 957–971. https://doi.org/10.1007/s10452-020-09786-6.

Boulesteix, A.L., Janitza, S., Kruppa, J., König, I.R., 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdiscipl. Rev.: Data Mining Knowledge Disc. 2 (6), 493–507. https://doi.org/10.1002/widm.1072.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1201/9780429469275-8.

Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*.

Breiman, L. (2002). Manual on setting up, using, and understanding random forests v3. 1.

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 (2), 325–333.

Brown, L.E., Milner, A.M., 2012. Rapid loss of glacial ice reveals stream community assembly processes. Glob. Chang. Biol. 18 (7), 2195–2204. https://doi.org/10.1111/j.1365-2486.2012.02675.x.

Brown, I., Mues, C., 2012. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. Expert Syst. Appl. 39 (3), 3446–3453. https://doi.org/10.1016/j.eswa.2011.09.033.

Bruno, D., Gutiérrez-Cánovas, C., Velasco, J., Sánchez-Fernández, D., 2016. Functional redundancy as a tool for bioassessment: A test using riparian vegetation. Sci. Total Environ. 566–567, 1268–1276. https://doi.org/10.1016/j.scitotenv.2016.05.186.

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 (1) https://doi.org/10.1007/s10661-014-4132-8.

Caliński, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Comm. Statist. Simulation Comput. 3 (1), 1–27. https://doi.org/10.1080/03610927408827101.

Çamdevýren, H., Demýr, N., Kanik, A., Keskýn, S., 2005. Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs. Ecol. Model. 181 (4), 581–589. https://doi.org/10.1016/j.ecolmodel.2004.06.043.

Carter, R.L., Morris, R., Blashfield, R.K., 1989. On the partitioning of squared Euclidean distance and its applications in cluster analysis. Psychometrika 54 (1), 9–23. https://doi.org/10.1007/BF02294446.

Casanoves, F., Di Rienzo, J. A., Pla, L., 2010. *User manual F-Diversity: statistical software for the analysis of functional diversity*.

Casanoves, F., Pla, L., Di Rienzo, J.A., Díaz, S., 2011. FDiversity: A software package for the integrated analysis of functional diversity. Methods Ecol. Evol. 2 (3), 233–237. https://doi.org/10.1111/j.2041-210X.2010.00082.x.

Castillo, L. G., Álvarez, M. A, Carrillo, J.M., 2014. Numerical modeling of sedimentation and flushing at the Paute-Cardenillo Reservoir. *ASCE-EWRI. International Perspective on Water Resources and Environment, January 8-10, 2014*, 2–11.

Celleri, R., Willems, P., Buytaert, W., Feyen, J., 2007. Space–time rainfall variability in the Paute Basin, Ecuadorian Andes. Hydrol. Processes: Int. J. 21 (24), 3316–3327. https://doi.org/10.1002/hyp.

Chen, C., 2000. Generalized additive mixed models. Commun. Stat. Theory Methods 29 (5–6), 1257–1271. https://doi.org/10.1080/03610920008832543.

Chen, F., Xue, Y., Tan, M.T., Chen, P., 2015. Efficient statistical tests to compare Youden index: accounting for contingency correlation. Stat. Med. 34 (9), 1560–1576. https://doi.org/10.1002/sim.6432.

Chevenet, F., Dolédec, S., Chessel, D., 1994. A fuzzy coding approach for the analysis of long-term ecological data. Freshw. Biol. 31 (3), 295–309. https://doi.org/10.1111/j.1365-2427.1994.tb01742.x.

Corbi, J.J., Trivinho-Strixino, S., 2006. Influence of taxonomic resolution of stream macroinvertebrate communities on the evaluation of different land uses. Retrieved from Acta Limnol. Bras. 18 (4), 469–475. http://www.csa.com/partners/viewrecord.php?requester=gs&collection=ENV&recid=7395281.

Da Ros, G., 1995. *La contaminación de aguas en Ecuador: una aproximación económica* (A. Yala, Ed.). Instituto de Investigaciones Económicas, Pontificia Universidad Católica del Ecuador.

Dalu, T., Wasserman, R.J., Tonkin, J.D., Alexander, M.E., Dalu, M.T.B., Motitsoe, S.N., Manungo, K.I., Bepe, O., Dube, T., 2017. Assessing drivers of benthic macroinvertebrate community structure in African highland streams: an exploration using multivariate analysis. Sci. Total Environ. 601-602, 1340–1348.

de Castro, D.M.P., Dolédec, S., Callisto, M., 2017. Landscape variables influence taxonomic and trait composition of insect assemblages in Neotropical savanna streams. Freshw. Biol. 62 (8), 1472–1486. https://doi.org/10.1111/fwb.12961.

De Troyer, N., Mereta, S.T., Goethals, P.L.M., Boets, P., 2016. Water quality assessment of streams and wetlands in a fast growing east African City. Water (Switzerland) 8 (4), 1–21. https://doi.org/10.3390/w8040123.

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.

Einax, J. W., Zwanziger, H. W., & Geiss, S. (1997). Chemometrics in environmental analysis. In *Water Management*. https://doi.org/10.1002/352760216X.

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. Ind. 126 (January) https://doi.org/10.1016/j.ecolind.2021.107609.

Ferraro, S.P., Cole, F.A., 1995. Taxonomic level sufficient for assessing pollution impacts on the southern California Bight macrobenthos-revisited. Environ. Toxicol. Chem.: Int. J. 14 (6), 1031–1040.

Firmiano, K.R., Castro, D.M.P., Linares, M.S., Callisto, M., 2021. Functional responses of aquatic invertebrates to anthropogenic stressors in riparian zones of Neotropical savanna streams. Sci. Total Environ. 753, 141865 https://doi.org/10.1016/j.scitotenv.2020.141865.

Forio, M.A.E., Goethals, P.L.M., Lock, K., Asio, V., Bande, M., Thas, O., 2018. Model-based analysis of the relationship between macroinvertebrate traits and environmental river conditions. Environ. Model. Softw. 106, 57–67. https://doi.org/10.1016/j.envsoft.2017.11.025.

Forio, M.A.E., Burdon, F.J., De Troyer, N., Lock, K., Witing, F., Baert, L., De Saeyer, N., Rîșnoveanu, G., Popescu, C., Kupilas, B., Friberg, N., Boets, P., Johnson, R.K., Volk, M., McKie, B.G., Goethals, P.L.M., 2022. A Bayesian Belief Network learning tool integrates multi-scale effects of riparian buffers on stream invertebrates. Sci. Total Environ. 810 https://doi.org/10.1016/j.scitotenv.2021.152146.

Fox, E.W., Hill, R.A., Leibowitz, S.G., Olsen, A.R., Thornbrugh, D.J., Weber, M.H., 2017. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. Environ. Monit. Assess. 189 (7) https://doi.org/10.1007/s10661-017-6025-0.

Frank, I. E., Todeschini, R., 1994. *The Data Analysis Handbook* (Vol. 14). https://doi.org/10.1016/S0922-3487(08)70048-0.

Gallardo, B., Gascón, S., Quintana, X., Comín, F.A., 2011. How to choose a biodiversity indicator - redundancy and complementarity of biodiversity metrics in a freshwater ecosystem. Ecol. Ind. 11 (5), 1177–1184. https://doi.org/10.1016/j.ecolind.2010.12.019.

Gardner, L., 2007. *A Generalized Additive Mixed Effects Modeling (GAMM) Approach to Short-term River Temperature Forecasting for the Fraser River, British Columbia : Model Evaluation and Implications for Salmon Fishery Management*. University of British Columbia.

Gusmao, J.B., Brauko, K.M., Eriksson, B.K., Lana, P.C., 2016. Functional diversity of macrobenthic assemblages decreases in response to sewage discharges. Ecol. Ind. 66, 65–75. https://doi.org/10.1016/j.ecolind.2016.01.003.

Han, H., Guo, X., & Yu, H., 2016. Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, *0*, 219–224. https://doi.org/10.1109/ICSESS.2016.7883053.

Haque, M. M., Rahman, A., Hagare, D., & Kibria, G. (2013). Principal component regression analysis in water demand forecasting: An application to the Blue Mountains, NSW, Australia. 1(1), 49–59.

Hastie, T., & Tibshirani, R. (1990). Generalized additive models. In *Statistical models* (First). https://doi.org/10.1214/ss/1177013604.

Heino, J., 2008. Patterns of functional biodiversity and function-environment relationships in lake littoral macroinvertebrates. Limnol. Oceanogr. 53 (4), 1446–1455. https://doi.org/10.4319/lo.2008.53.4.1446.

Helsen, K., Ceulemans, T., Stevens, C.J., Honnay, O., 2014. Increasing soil nutrient loads of european semi-natural grasslands strongly alter plant functional diversity independently of species loss. Ecosystems 17 (1), 169–181. https://doi.org/10.1007/s10021-013-9714-8.

Herrera, I.A., Burneo, P., 2017. Environmental flow assessment in Andean rivers of Ecuador, case study: Chanlud and El Labrado dams in the Machángara River. Ecohydrol. Hydrobiol. 17 (2), 103–112. https://doi.org/10.1016/j.ecohyd.2017.01.002.

Holguin-Gonzalez, J.E., Boets, P., Alvarado, A., Cisneros, F., Carrasco, M.C., Wyseure, G., Nopens, I., Goethals, P.L.M., 2013. Integrating hydraulic, physicochemical and ecological models to assess the effectiveness of water quality management strategies for the River Cuenca in Ecuador. Ecol. Model. 254, 1–14.

Howard, R., 2019. Evaluating restoration success by tracking the structural and functional recovery of restored, drained, and intact wetlands. *Electronic Thesis and Dissertation Repository*. Retrieved from https://ir.lib.uwo.ca/etd/6122.

Iddrisu, W.A., Nokoe, K.S., Luguterah, A., Antwi, E.O., 2017. Generalized additive mixed modelling of river discharge in the black volta river. Open J. Stat. 07 (04), 621–632. https://doi.org/10.4236/ojs.2017.74043.

Ingersoll, T., Cole, S., Madren-Whalley, J., Booker, L., Dorsey, R., Li, A., Salem, H., 2016. Generalized additive mixed-models for pharmacology using integrated discrete multiple organ co-culture. PLoS One 11 (4), 1–12. https://doi.org/10.1371/journal.pone.0152985.

Iñiguez-Armijos, C., Hampel, H., Breuer, L., 2018. Land-use effects on structural and functional composition of benthic and leaf-associated macroinvertebrates in four Andean streams. Aquat. Ecol. 52 (1), 77–92. https://doi.org/10.1007/s10452-017-9646-z.

Jacobsen, D., Schultz, R., Encalada, A., 1997. Structure and diversity of stream invertebrate assemblages: the influence of temperature with altitude and latitude. Freshw. Biol. 38 (2), 247–261. https://doi.org/10.1046/j.1365-2427.1997.00210.x.

Jerves-Cobo, R., Lock, K., Van Butsel, J., Pauta, G., Cisneros, F., Nopens, I., Goethals, P.L.M., 2018. Biological impact assessment of sewage outfalls in the urbanized area of the Cuenca River basin (Ecuador) in two different seasons. Limnologica 71, 8–28. https://doi.org/10.1016/j.limno.2018.05.003.

Jerves-Cobo, R., Benedetti, L., Amerlinck, Y., Lock, K., De Mulder, C., Van Butsel, J., Cisneros, F., Goethals, P., Nopens, I., 2020. Integrated ecological modelling for evidence-based determination of water management interventions in urbanized river basins: Case study in the Cuenca River basin (Ecuador). Sci. Total Environ. 709 https://doi.org/10.1016/j.scitotenv.2019.136067.

Jiménez, R., Barnuevo, E., Timbe, B., Astudillo, P.X., 2021. El uso de gremios tróficos en macroinvertebrados acuáticos como herramienta de monitoreo en los Altos Andes del Sur de Ecuador. Neotropical Biodiversity 7 (1), 310–317. https://doi.org/10.1080/23766808.2021.1953891.

Kalny, G., Laaha, G., Melcher, A., Trimmel, H., Weihs, P., Rauch, H.P., 2017. The influence of riparian vegetation shading on water temperature during low flow conditions in a medium sized river. Knowl. Manag. Aquat. Ecosyst. (418), 5.

Khalilia, M., Chakraborty, S., Popescu, M., 2011. Predicting disease risks from highly imbalanced data using random forest. BMC Med. Inf. Decis. Making 11 (1). https://doi.org/10.1186/1472-6947-11-51.

Koperski, P., 2019. Phylogenetic diversity of larval Chironomidae (Diptera) in lowland rivers as a potential tool in assessment of environmental quality. Hydrobiologia 836 (1), 83–96. https://doi.org/10.1007/s10750-019-3942-7.

Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47 (260), 583–621.

Laanaya, F., St-Hilaire, A., Gloaguen, E., 2017. Water temperature modelling: comparison between the generalized additive model, logistic, residuals regression and linear regression models. Hydrol. Sci. J. 62 (7), 1078–1093. https://doi.org/10.1080/02626667.2016.1246799.

Laliberte, E., Legendre, P., 2010. A distance-based framework for measuring functional diversity from multiple traits. Ecology 91 (1), 299–305. https://doi.org/10.1890/08-2244.1.

Larras, F., Coulaud, R., Gautreau, E., Billoir, E., Rosebery, J., Usseglio-Polatera, P., 2017. Assessing anthropogenic pressures on streams: A random forest approach based on benthic diatom communities. Sci. Total Environ. 586, 1101–1112. https://doi.org/10.1016/j.scitotenv.2017.02.096.

Leiva, M., Marchese, M., Lorenz, G., Diodato, L., 2022. Functional diversity of benthic macroinvertebrates regarding hydrological and land use disturbances in a heavily impaired lowland river. Limnologica 92 (December), 125940. https://doi.org/10.1016/j.limno.2021.125940.

Lin, X., Zhang, D., 1999. Inference in generalized additive mixed models by using smoothing splines. *J. R. Statist. Soc.: Series B (Statistical Methodology)*, *61*(2), 381–400.

Liu, Y., Zhang, Z. Zhang, F., 2019. *Challenges for Water Security and Sustainable Industrial, Domestic Water Use and Pollution.*.

Liaw, A., Wiener, M., 2015. *randomForest: Breiman and Cutler's Random Forests for Classification and Regression. R package version 4*, p. 14.

Lohbeck, M., Poorter, L., Paz, H., Pla, L., van Breugel, M., Martínez-Ramos, M., Bongers, F., 2012. Functional diversity changes during tropical forest succession. Perspect. Plant Ecol., Evolut. System. 14 (2), 89–96. https://doi.org/10.1016/j.ppees.2011.10.002.

Loinaz, M., 2012. *Integrated ecohydrological modeling at the catchment scale* (Technical University of Denmark). Retrieved from http://orbit.dtu.dk/services/downloadRegister/9891763/Maria_C_Loinazr_PhD_thesis_WWW_Version.pdf.

Louppe, G., Wehenkel, L., Sutera, A., Geurts, P., 2013. Understanding variable importances in forests of randomized trees. Adv. Neural Inf. Proces. Syst. 26, 1–9.

Lozanovska, I., Ferreira, M.T., Aguiar, F.C., 2018. Functional diversity assessment in riparian forests – multiple approaches and trends: A review. Ecol. Ind. 95 (August), 781–793. https://doi.org/10.1016/j.ecolind.2018.08.039.

Magliozzi, C., Usseglio-Polatera, P., Meyer, A., Grabowski, R.C., Alahuhta, J., 2019. Functional traits of hyporheic and benthic invertebrates reveal importance of wood-driven geomorphological processes in rivers. Funct. Ecol. 33 (9), 1758–1770.

Magliozzi, C., Meyer, A., Usseglio-Polatera, P., Robertson, A., Grabowski, R.C., 2020. Investigating invertebrate biodiversity around large wood: Taxonomic vs functional metrics. Aquat. Sci. 82 (4), 1–13. https://doi.org/10.1007/s00027-020-00745-9.

Martínez, A., Larrañaga, A., Basaguren, A., Pérez, J., Mendoza-Lera, C., Pozo, J., 2013. Stream regulation by small dams affects benthic macroinvertebrate communities: From structural changes to functional implications. Hydrobiologia 711 (1), 31–42. https://doi.org/10.1007/s10750-013-1459-z.

Mason, N.W.H., de Bello, F., Mouillot, D., Pavoine, S., Dray, S., Zobel, M., 2013. A guide for using functional diversity indices to reveal changes in assembly processes along ecological gradients. J. Veg. Sci. 24 (5), 794–806.

Mellor, A.F.P., Cey, E.E., 2015. Using generalized additive mixed models to assess spatial, temporal, and hydrologic controls on bacteria and nitrate in a vulnerable agricultural aquifer. J. Contam. Hydrol. 182, 104–116. https://doi.org/10.1016/j.jconhyd.2015.08.010.

Meng, X., Cooper, K.M., Liu, Z., Li, Z., Chen, J., Jiang, X., Ge, Y., Xie, Z., 2021. Integration of α, β and γ components of macroinvertebrate taxonomic and functional diversity to measure of impacts of commercial sand dredging. Environ. Pollut. 269 https://doi.org/10.1016/j.envpol.2020.116059.

Mevik, B., Wehrens, R., 2015. Introduction to the pls Package. *Help Section of the "Pls" Package of RStudio Software*, (Section 7), 1–23.

Montgomery, D., Peck, E.A., Geoffrey, V., 2012. Introduction to Linear Regression Analysis (Fifth). John Wiley & Sons.

Morris, A., 1985. Forestry and land-use conflicts in Cuenca, Ecuador. Mt. Res. Dev. May (1), 183–196.

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 (4), 867–876. https://doi.org/10.1111/j.1365-2435.2010.01695.x.

Moyo, N.A.G., Rapatsa, M.M., 2016. Impact of urbanization on the ecology of Mukuvisi River, Harare, Zimbabwe. Phys. Chem. Earth 92, 14–19. https://doi.org/10.1016/j.pce.2015.09.007.

Mueller, M., Pander, J., Geist, J., 2013. Taxonomic sufficiency in freshwater ecosystems: Effects of taxonomic resolution, functional traits, and data transformation. Freshwater Sci. 32 (3), 762–778. https://doi.org/10.1899/12-212.1.

Nevalainen, L., Luoto, T.P., Manca, M., Weisse, T., 2015. A paleolimnological perspective on aquatic biodiversity in Austrian mountain lakes. Aquat. Sci. 77 (1), 59–69. https://doi.org/10.1007/s00027-014-0363-6.

Nevalainen, L., Luoto, T.P., Pfrender, M., 2017. Relationship between cladoceran (Crustacea) functional diversity and lake trophic gradients. Funct. Ecol. 31 (2), 488–498.

Pallottini, M., Cappelletti, D., Fabrizi, A., Gaino, E., Goretti, E., Selvaggi, R., Céréghino, R., 2017. Macroinvertebrate functional trait responses to chemical pollution in agricultural-industrial landscapes. River Res. Appl. 33 (4), 505–513. https://doi.org/10.1002/rra.3101.

Pauta-Calle, G., & Chang-Gómez, J. (2014). Indices de calidad del agua de fuentes superficiales y aspectos toxicológicos, evaluación del Río Burgay. *MASKANA, I+D+ingeniería*, 165–176.

Pavoine, S., Ollier, S., Pontier, D., 2005. Measuring diversity from dissimilarities with Rao's quadratic entropy: Are any dissimilarities suitable? Theor. Popul Biol. 67 (4), 231–239. https://doi.org/10.1016/j.tpb.2005.01.004.

Paz, L.E., Rodriguez, M., Gullo, B., Rodrigues Capítulo, A., 2022. Impacts of urban and industrial pollution on functional traits of benthic macroinvertebrates: Are some traits advantageous for survival? Sci. Total Environ. 807, 150650 https://doi.org/10.1016/j.scitotenv.2021.150650.

Petchey, O.L., Gaston, K.J., 2006. Functional diversity: Back to basics and looking forward. Ecol. Lett. 9 (6), 741–758. https://doi.org/10.1111/j.1461-0248.2006.00924.x.

Pla, L., Casanoves, F., Di Rienzo, J., 2012. Quantifying Functional Biodiversity. Springer Briefs in Environmental Science.

Pla, L., Casanoves, F., Di Rienzo, J., Fernandez, F., & Finegan, B. (2008). Confidence intervals for functional diversity indices considering species abundance. *In XXIV International Biometric Conference, Dublin*.

Polansky, L., Robbins, M.M., 2013. Generalized additive mixed models for disentangling long-term trends, local anomalies, and seasonality in fruit tree phenology. Ecol. Evol. 3 (9), 3141–3151. https://doi.org/10.1002/ece3.707.

Rácz, A., Bajusz, D., Héberger, K., 2018. Modelling methods and cross-validation variants in QSAR: A multi-level analysis$. SAR QSAR Environ. Res. 29 (9), 661–674. https://doi.org/10.1080/1062936X.2018.1505778.

Rani, M., Marchesi, C., Federici, S., Rovelli, G., Alessandri, I., Vassalini, I., Ducoli, S., Borgese, L., Zacco, A., Bilo, F., Bontempi, E., Depero, L.E., 2019. Miniaturized near-infrared (MicroNIR) spectrometer in plastic waste sorting. Materials 12 (17), 2740.

Rao, C.R., 1982. Diversity and dissimilarity coefficients: A unified approach. Theoritical Popul. Biol. 21 (1), 24–43. https://doi.org/10.13140/RG.2.1.3901.9924.

Rebala, G., Ravi, A., & Churiwala, S. (2019). Random forests. In *An Introduction to Machine Learning* (pp. 77–94). https://doi.org/10.1023/A:1010933404324.

Reid, H.E., Brierley, G.J., Boothroyd, I.K.G., 2010. Influence of bed heterogeneity and habitat type on macroinvertebrate uptake in peri-urban streams. Int. J. Sedim. Res. 25 (3), 203–220. https://doi.org/10.1016/S1001-6279(10)60039-X.

Reynaga, M.C., Nieto, C., Rodríguez, J.S., Izquierdo, A.E., 2020. Biological traits of macroinvertebrates from Puna peatbogs: Patterns along spatial environmental gradients. Freshwater Sci. 39 (1), 137–146.

Riis, T., Kelly-Quinn, M., Aguiar, F.C., Manolaki, P., Bruno, D., Bejarano, M.D., Clerici, N., Fernandes, M.R., Franco, J.C., Pettit, N., Portela, A.P., Tammeorg, O., Tammeorg, P., Rodríguez-González, P.M., Dufour, S., 2020. Global overview of ecosystem services provided by riparian vegetation. Bioscience 70 (6), 501–514.

Roa-Fuentes, L., Villamizar-Peña, L.A., Mantilla-Carreño, J.A., Jaramillo, M.A., 2022. Functional diversity and species diversity in flooded and unflooded tropical forests. Acta Oecol. 114 (May) https://doi.org/10.1016/j.actao.2022.103814.

Roldán, G., 2003. *Bioindicación de la calidad del agua en colombia: Propuesta para el uso del método BMWP Col.* Retrieved from https://books.google.com.co/books?id=ZEjgIKZTF2UC&printsec=frontcover&dq=calidad+del+agua&hl=es-419&sa=X&ved=0ahUKEwi9_4jo8pDNAhWM6CYKHSX4CsAQ6AEIQDAC#v=onepage&q&f=false.

Roon, D.A., Dunham, J.B., Groom, J.D., 2021. Shade, light, and stream temperature responses to riparian thinning in second-growth redwood forests of northern California. PLoS One 16 (2 February), 1–25. https://doi.org/10.1371/journal.pone.0246822.

Salazar, G., Rudnick, H., 2008. Hydro power plants in Ecuador: A technical and economical analysis. *IEEE Power and Energy Society 2008 General Meeting: Conversion and Delivery of Electrical Energy in the 21st Century*, (July), 1–5. https://doi.org/10.1109/PES.2008.4596897.

Schleuter, D., Daufresne, M., Massol, F., Argillier, C., 2010. A user's guide to functional diversity indices. Ecol. Monogr. 80 (3), 469–484. https://doi.org/10.1890/08-2225.1.

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 (1), 27–44. https://doi.org/10.1007/s10750-016-2974-5.

Seiz, A., 2015. *Functional Diversity and Redundancy of Benthic Communities in Mangrove Forests*.

Shadish, W.R., Zuur, A.F., Sullivan, K.J., 2014. Using generalized additive (mixed) models to analyze single case designs. J. Sch. Psychol. 52 (2), 149–178. https://doi.org/10.1016/j.jsp.2013.11.004.

Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika Trust 52 (3, 4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591.

Shuman, T.C., Smiley, P.C., Gillespie, R.B., Gonzalez, J.M., 2020. Influence of physical and chemical characteristics of sediment on macroinvertebrate communities in agricultural headwater streams. Water (Switzerland) 12 (11), 1–17. https://doi.org/10.3390/w12112976.

Sokal, R.R., Michener, C.D., 1958. A Statistical Method for Evaluating Systematic Relationships. University of Kansas, Lawrence.

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In J. G. Carbonell & J. Siekmann (Eds.), *AI 2006: Advances in Artificial Intelligence - 19th Australian Joint Conference onArtificial Intelligence Hobart, Australia, December 4-8, 2006 Proceedings* (pp. 1015–1021). https://doi.org/10.1007/11941439_114.

Sotomayor, G., Hampel, H., Vázquez, R.F., 2018. Water quality assessment with emphasis in parameter optimisation using pattern recognition methods and genetic algorithm. Water Res. 130, 353–362. https://doi.org/10.1016/j.watres.2017.12.010.

Sotomayor, G., Hampel, H., Vázquez, R.F., Forio, M.A.E., Goethals, P.L.M., Helmuth, B., 2021. Implications of macroinvertebrate taxonomic resolution for freshwater assessments using functional traits: The Paute River Basin (Ecuador) case. Divers. Distrib. 28 (8), 1735–1747.

Sotomayor, G., Hampel, H., Vázquez, R.F., Goethals, P.L.M., 2020. Multivariate-statistics based selection of a benthic macroinvertebrate index for assessing water quality in the Paute river basin (Ecuador). Ecol. Ind. 111 (April 2020), 106037 https://doi.org/10.1016/j.ecolind.2019.106037.

Sotomayor, G., 2016. Evaluación de la calidad de las aguas superficiales mediante técnicas de estadística multivariante: Un estudio de caso en la cuenca del Río Paute, al sur de Ecuador (Universidad Nacional de La Plata). Retrieved from http://hdl.handle.net/10915/53959.

Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinf. 8 https://doi.org/10.1186/1471-2105-8-25.

Strobl, R.O., Robillard, P.D., 2008. Network design for water quality monitoring of surface freshwaters: A review. J. Environ. Manage. 87 (4), 639–648. https://doi.org/10.1016/j.jenvman.2007.03.001.

Tomanova, S., Usseglio-Polatera, P., 2007. Patterns of benthic community traits in neotropical streams: Relationship to mesoscale spatial variability. Fundam. Appl. Limnol. 170 (3), 243–255. https://doi.org/10.1127/1863-9135/2007/0170-0243.

Turcotte, P., Harper, P.P., 1982. The macro-invertebrate fauna of a small Andean stream. Freshw. Biol. 12 (5), 411–419. https://doi.org/10.1111/j.1365-2427.1982.tb00635.x.

Turunen, J., Aroviita, J., Marttila, H., Louhi, P., Laamanen, T., Tolkkinen, M., Luhta, P.-L., Kløve, B., Muotka, T., Mateos, D.M., 2017. Differential responses by stream and riparian biodiversity to in-stream restoration of forestry-impacted streams. J. Appl. Ecol. 54 (5), 1505–1514.

Ul-Saufie, A.Z., Yahya, A.S., Ramli, N., 2011. Improving multiple linear regression model using principal component analysis for predicting PM 10 concentration in Seberang Prai, Pulau Pinang. Int. J. Environ. Sci. 2 (2), 403–410.

Valera, C., Pissarra, T., Filho, M., Valle Júnior, R., Oliveira, C., Moura, J., Sanches Fernandes, L., Pacheco, F., 2019. The buffer capacity of riparian vegetation to control water quality in anthropogenic catchments from a legally protected area: A critical view over the Brazilian new forest code. Water (Switzerland) 11 (3), 549.

Varmuza, K., Filzmoser, P., 2010. Introduction to multivariate statistical analysis in chemometrics. In *Applied Spectroscopy* (Vol. 64). https://doi.org/10.1366/000370210791114185.

Vázquez, R.F., Vimos-Lojano, D., Hampel, H., 2020. Habitat suitability curves for freshwater macroinvertebrates of tropical andean rivers. Water (Switzerland) 12 (10), 1–22. https://doi.org/10.3390/w12102703.

Villéger, S., Mason, N.W.H., Mouillot, D., 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology 89 (8), 2290–2301. https://doi.org/10.1890/07-1206.1.

Vimos-Lojano, D. J., Martínez-Capel, F., Hampel, H., Vázquez, R.F., 2016. Hydrological implications at mesohabitat scale on aquatic macroinvertebrates communities of Andean rivers of southern Ecuador: Towards the determination of environmental flow. *21st Century Watershed Technology Conference and Workshop 2016: Improving Quality of Water Resources at Local, Basin and Regional Scales*, *2016-Janua*, 77–80. https://doi.org/10.13031/wtcw.2016015.

Vimos-Lojano, D., Hampel, H., Vázquez, R.F., Martínez-Capel, F., 2020. Community structure and functional feeding groups of macroinvertebrates in pristine Andean streams under different vegetation cover. Ecohydrol. Hydrobiol. 20 (3), 357–368. https://doi.org/10.1016/j.ecohyd.2020.04.004.

Von Ellenrieder, N., 2007. Composition and structure of aquatic insect assemblages of Yungas mountain cloud forest streams in NW Argentina. Rev. Soc. Entomol. Argentina 66, 57–76.

Wallace, J.B., Webster, J.R., 1996. The role of macroinvertebrates in stream ecosystem function. Annu. Rev. Entomol. 41 (1), 115–139. https://doi.org/10.1146/annurev.en.41.010196.000555.

Wang, Z., Goonewardene, L.A., 2004. The use of MIXED models in the analysis of animal experiments with repeated measures data. Can. J. Anim. Sci. 84 (1), 1–11. https://doi.org/10.4141/a03-123.

Williams, L. J., Abdi, H., 2010. Fisher's Least Significant Difference (LSD) Test. In N. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 840–853).

Wood, S., Scheipl, F., 2020. *gamm4: Generalized Additive Mixed Models using "mgcv" and "lme4". R package version 0.2-6*.

Wood, S.A., Atalah, J., Wagenhoff, A., Brown, L., Doehring, K., Young, R.G., Hawes, I., 2017. Effect of river flow, temperature, and water chemistry on proliferations of the benthic anatoxin-producing cyanobacterium Phormidium. Freshwater Science 36 (1), 63–76. https://doi.org/10.1086/690114.

Xu, Z., Huang, G., Weinberger, K. Q., & Zheng, A. X. (2014). Gradient boosted feature selection. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (February 2015), 522–531. https://doi.org/10.1145/2623330.2623635.

Youden, W.J., 1950. Index for rating diagnostic tests. Cancer 3 (1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.

Zahro, J., Caraka, R.E., 2017. Modelling inflation using generalized additive mixed models (GAMM). Int. J. Chem., Mathem. Phys. (IJCMP) 1 (1), 73–139.

Zhang, Y., Cheng, L., Li, K., Zhang, L., Cai, Y., Wang, X., Heino, J., 2019. Nutrient enrichment homogenizes taxonomic and functional diversity of benthic macroinvertebrate assemblages in shallow lakes. Limnol. Oceanogr. 64 (3), 1047–1058. https://doi.org/10.1002/lno.11096.