



# Variation partitioning in double-constrained multivariate analyses: linking communities, environment, space, functional traits, and ecological niches

Ioan Sîrbu<sup>1</sup> · Ana Maria Benedek<sup>1</sup> · Monica Sîrbu<sup>2</sup>

Received: 11 September 2019 / Accepted: 21 July 2021 / Published online: 11 August 2021  
© The Author(s) 2021

## Abstract

Constrained multivariate analysis is a common tool for linking ecological communities to environment. The follow-up is the development of the double-constrained correspondence analysis (dc-CA), integrating traits as species-related predictors. Further, methods have been proposed to integrate information on phylogenetic relationships and space variability. We expand this framework, proposing a dc-CA-based algorithm for decomposing variation in community structure and testing the simple and conditional effects of four sets of predictors: environment characteristics and space configuration as predictors related to sites, while traits and niche (dis)similarities as species-related predictors. In our approach, ecological niches differ from traits in that the latter are distinguished by and characterize the individual level, while niches are measured on the species level, and when compared, they are characteristics of communities and should be used as separate predictors. The novelties of this approach are the introduction of new niche parameters, niche dissimilarities, synthetic niche-based diversity which we related to environmental features, the development of an algorithm for the full variation decomposition and testing of the community–environment–niche–traits–space (CENTS) space by dc-CAs with and without covariates, and new types of diagrams for the results. Applying these methods to a dataset on freshwater mollusks, we learned that niche predictors may be as important as traits in explaining community structure and are not redundant, overweighting the environmental and spatial predictors. Our algorithm opens new pathways for developing integrative methods linking life, environment, and other predictors, both in theoretical and practical applications, including assessment of human impact on habitats and ecological systems.

**Keywords** dc-CA · Trait-based ecology · Ecological niche · Statistical graphics · Mollusk communities

## Introduction

Describing, explaining, and modeling relationships between communities and their environment is the core of many modern ecological studies, aiming at answering a wide range of questions. From a methodological perspective, the number of datasets or matrices used in such studies has steadily increased (Legendre and Legendre 2012; ter Braak et al. 2018). The community matrix (C), usually a

site-by-species table, and the unconstrained ordination analysis was expanded by adding a second table with environmental features (E) (Fig. 1a), used as predictors in constrained ordination (direct gradient) analysis. The addition of a third matrix, with functional traits (T) (Fig. 1b), aiming at testing hypotheses on how biological characteristics of species determine their responses to habitat features, was proposed, and methods of analysis were developed by several authors (Dolédec et al. 1996; Legendre et al. 1997; Dray and Legendre 2008; Lepš 2013; ter Braak et al. 2018; Peng et al. 2021; Pinho et al. 2021). This key approach evaluates processes, functions, and services of ecosystems (Céréghino et al. 2018; Sterk et al. 2013).

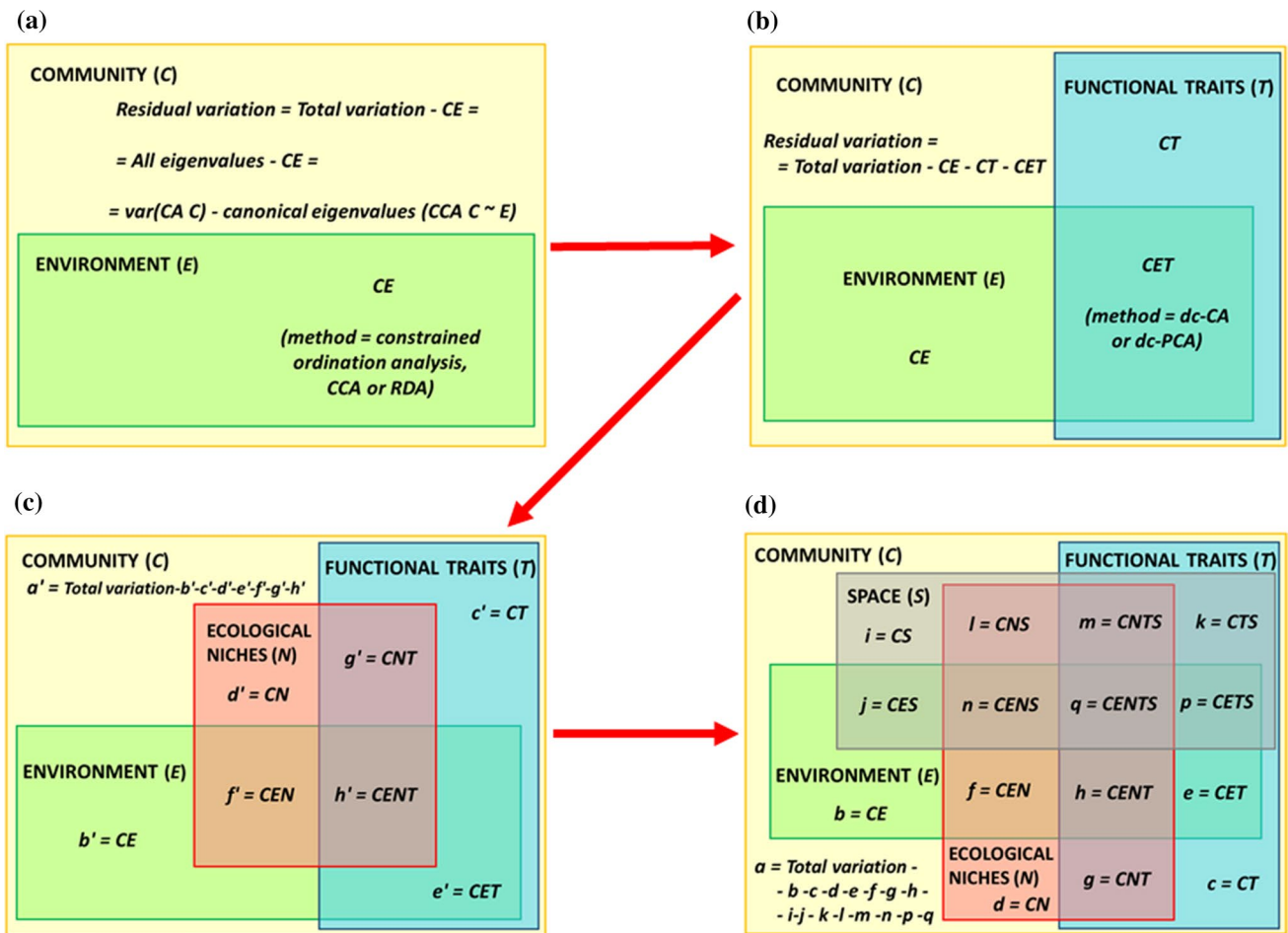
Until recently, correspondence analysis with linear external constraints on both rows and columns (Legendre and Legendre 2012) lacked full mathematical treatment and readily available algorithms and software. ter Braak et al. (2018) provided an explicitly defined mathematical

Communicated by Bryan Brown.

✉ Ioan Sîrbu  
meosirbu@yahoo.com; ioan.n.sirbu@ulbsibiu.ro

<sup>1</sup> Faculty of Sciences, Lucian Blaga University of Sibiu, 5-7 Dr. I. Ratiu St., 550012 Sibiu, Romania

<sup>2</sup> Andrei Şaguna National Pedagogical College, 2 Aleea Turnu Roşu St., 550361 Sibiu, Romania



**Fig. 1** Diagram illustrating methodological evolution (from **a** to **d**) of integrating datasets to analyze relationships between communities and their predictors; (c) and (d) synthesize the gap that our article

aims to fill: defining and analyzing the CENTS (community–environment–niche–traits–space) space by variation decomposition in double-constrained multivariate analyses (VADOC diagrams)

algorithm for double-constrained correspondence analysis (dc-CA) and its linear counterpart, the double-constrained principal components analysis (dc-PCA), which are embedded in the software Canoco since version 5.10 (ter Braak and Šmilauer 2018). An R function for dc-CA was presented by ter Braak et al. (2018), with examples in Peng et al. (2021) and Pinho et al. (2021).

Another component in explaining the variability in community structure is space. One method to extract spatial patterns, which can be used as space predictors (S), from available spatial coordinates is the distance-based Moran eigenvector maps (db-MEM) (Borcard and Legendre 2002), formerly known as principal coordinates of neighbor matrices (PCNM), improved by Dray et al. (2006). Further, incorporating distances from phylogenetic trees became a standard routine to account for the so-called patristic relationships when investigating responses of species and their traits to environment gradients (Desdevises et al. 2003; de Bello et al. 2005, 2017; Pavoine et al. 2011).

Few methods have been developed so far that have taken an integrative approach to the analysis of the four datasets related to predictors of community assembly, and most of them include traits, phylogeny, environment, and space, relating them in various ways. Some of them aim to find spatial patterns in the components of trait diversity attributable to phylogenetic effects and environmental effects in the geographical space (Diniz-Filho et al. 2007). Others remove the phylogenetic signal in traits and the spatial information in environmental variables to associate ‘phylogenetic-free’ traits with ‘space-free’ environment, looking at the unique effects of environment and traits on species composition (Kühn et al. 2009). One of the oldest multivariate methods for trait–environment analysis, the RLQ, is also a three table-related problem (R stands for environment, L for community, and Q for traits matrices) aiming at estimating the parameters, by maximizing the covariance, in a fourth-corner matrix that crosses the habitat characteristics with the traits of species (Dray and Legendre 2008; Legendre and

Legendre 2012). The RLQ is based on co-inertia analysis (Dray et al. 2003) and uses no regression (ter Braak et al. 2018). Pavoine et al. (2011) developed a statistical approach based on RLQ, which analyzes environmental filtering in an explicit geographic and phylogenetic context. However, although the abstract and some figures may suggest that variation partitioning was performed, RLQ is not suited for that goal because it is covariance based and does not use the same scale to measure variance in the different analyses (Peng et al. 2021). Also, RLQ requires a scaling method to make the predictor matrices better comparable and neither tests nor explicitly selects traits or environment predictors. In contrast, with dc-CA one can test and select both traits and environmental variables. In addition, dc-CA is scale invariant, so that no scaling method needs to be selected. It also allows variation partitioning, because it is regression based and its explained variance is in terms of inertia of the abundance table. Peng et al. (2021) suggested the possibility of variation decomposition in dc-CA, and Pinho et al. (2021) partitioned the trait-structured variation (in our approach, this being the blue part in Fig. 1) in neotropical forest composition using dc-CA, showing the unique and shared effects of climate and geographic position. This approach can be used analogously to decompose the environmentally structured variation (the green part in Fig. 1) based on two sets of predictors related to species by transposing the community matrix. Here, we propose an algorithm for decomposing the total variation in community composition (and not only the trait- or environmentally structured variation) into the parts attributable to four sets of predictors, i.e., two related to sites (environment and space) and two related to species (functional traits and ecological niche).

The concept of ecological niche is considered of paramount importance in understanding life–environment relationships (Krebs 2014) while being also subject to a wide array of concepts and measurements, but sometimes affected by ambiguity or equivocal use (Peterson and Soberón 2012; Soberón et al. 2017; Melo-Merino et al. 2020). In multivariate analyses, niches were mostly defined by functional traits (Violle and Jiang 2009; Kearney et al. 2010). Thus, traits and niches have been used interchangeably, or traits have been considered a proxy for assessing niches and included as such in constrained analyses. In our approach, we consider that niches, especially their overlap (or similarity), should be used separately from traits because they relate to different aspects concerning the species and communities. The difference in the species niches characterizes the outcome of their phylogeny, adaptations, and relationships with both biotic and abiotic factors. This synthetic information is contained in the niche dissimilarity matrix and cannot be sufficiently represented by individual (and often related) traits. We consider that another difference is that traits are

distinguished and measured on the individual level, but niches can be characterized only at the species level and can be understood when comparing populations (species), revealing their meaning at the community level.

This paper defines and methodologically highlights the CENTS space, the acronym coming from Community—Environment—Niche—(functional) Traits—Space (Fig. 1). The first objective was to define, measure, and partition the CENTS space (Fig. 1d). Because spatial variability may not always be of interest, we also considered the CENT space (Fig. 1c) and, without a second set of species predictors, the CEST space. We propose an algorithm that combines three existing and widely used methods, the canonical correspondence analysis (CCA) or its linear counterpart, the redundancy analysis (RDA), dc-CA or its linear counterpart, dc-PCA, and the variation partitioning (VP) procedure (Borcard et al. 1992; ter Braak et al. 2018; ter Braak and Šmilauer 2018). The algorithm aims to disentangle and quantify the overlapping effects of E–S and T–N variable groups on C. Analyzing the CENTS space will considerably strengthen and link the fields of functional and structural ecology, thereby generating new insights and leading to a better understanding and quantification of the mechanisms underlying species–trait–environment patterns. In addition, the algorithm we developed may be used for other predictor data tables, such as a table with ecological indicator values or with phylogenetic relationships, and it also may be extended to include more than two data tables for sites or species.

Our second objective was to summarize how species relate to resources and their availability in the environment, synthesize this information in a standardized way, and use these novel measures to apply the algorithm mentioned above, including an N data table, measuring the ecological niche features of the species. For this goal, we proposed a new standardized metric of niche complementarity (dissimilarity) for both categorical and continuous resources, which also account for the availability of resources in the environment. We used this metric to define and measure the species' uniqueness and one more aspect of the community diversity, the niche-based diversity (ND). We explored relationships between diversity measures and environment predictors, highlighting the use of ND in human impact assessment.

## Materials and methods

### Algorithm for variation decomposition and testing of CENT and CENTS space

To explain and decompose the variation in species composition, the algorithm uses the following tables:

- i.  $C$ , the sites-by-species (occurrence, abundance or dominance) table, representing community structure. To study whether traits (or niche or both) can explain variation in species composition, one also needs the transpose of  $C$ , denoted by  $C'$ .  
All the other data tables considered in our algorithm include predictors (explanatory variables):
- ii.  $E$ , the sites-by-environmental variables data table with values of environment descriptors of interest for each site in  $C$ .
- iii.  $T$ , the species-by-traits data table with values of functional traits of each species in table  $C$ . Categorical traits can be expanded to sets of dummy (1/0) variables if needed. If the traits are a mix of distinct types (e.g., numeric, nominal, multi-choice, circular), the matrix  $T$  can be replaced by a species-by-species matrix of distances, for instance, using the Gower distance.
- iv.  $N$ , a species-by-niche metrics table or a species-by-species table that describes the ecological niches of species. In other studies,  $N$  may contain tolerances to environment factors, bioindicator responses, or other features (except for traits). The niche overlap metrics may be based on any categorical or continuous resources (details in the section on niche measures).
- v.  $S$ , a table describing the spatial configuration of the sites. It is based on coordinates or other variables indicating the position of the sites (e.g., latitude, longitude, altitude, depth), which are subject to a db-MEM analysis with forward selection (FWS) of predictors. In other studies, if  $S$  is a rectangular matrix and certain spatial descriptors (e.g., altitude) are of interest, it might be used as such.

Using these five data tables, we propose an algorithm for linking them and exploring their relationships.  $C$  plays the central role (of response variables), and there are two groups of predictors: group 1, comprising  $E + S$ , which apply to sites, being related indirectly to species through their distribution in sites, and group 2, comprising  $T + N$ , directly related to species and indirectly to sites through the structure of the communities they shelter.

In our algorithm, we used dc-CA because  $C$  is a compositional table and the length of the gradient was greater than 3, but when the axes are shorter, or the data are not compositional, the linear form (dc-PCA) might be used instead.

To apply dc-CA, the data sources that are not yet rectangular matrices (units by variables), such as (dis)similarity matrices, need first to be preprocessed to become rectangular. For this, the principal coordinates analysis (PCO) is used. Also, as dc-CA is a regression method, the number of row and column predictor variables should be kept small compared to the number of sites and species, respectively. To limit the number of predictors, we propose to use FWS of variables. In our algorithm, all variables within each group are denoted by the subscript ( $E_{all}$ ,  $S_{all}$ ,  $T_{all}$ , and  $N_{all}$ ), while the selected variables by FWS, are referred to without subscript.

Table 1 shows the preliminary steps needed when space predictors are included (CENTS space), the niche data is a (dis)similarity matrix, the number of environmental variables is large compared to the number of sites, and the number of traits is large compared to the number of species. When this is not the case, these preliminary steps (or some of them) are skipped.

**Table 1** Preliminary steps for selecting variables and preparing the data for the algorithm of double-constrained analysis with variation partitioning

Step	Method	Response variables	Predictors group 1 (E, S)	Predictors group 2 (T, N)	Output and meaning
1	PCO	DHS			$N_{all}$ , i.e., PCO scores for the niche
2	VP dbMEM FWS	$C$	$E_{all}$ and Coord		$S_{all}$ , i.e., spatial db-MEM eigenvectors (PCO scores), selection of $S$ and $E$ , and testing their simple effects
3	dc-CA FWS	$C$	$E_{all}$	$T_{all}$	Selection of $T$ and verification of selection (or reaching consensus) for $E$
4	dc-CA FWS	$C$	$E_{all}$	$N_{all}$	Selection of $N$ and verification of selection (or reaching consensus) for $E$

$E$  environmental variables,  $S$  space variables,  $T$  traits,  $N$  niche variables,  $C$  community composition matrix,  $PCO$  principal coordinates analysis,  $dbMEM$  distance-based Moran Eigenvector Maps,  $DHS$  matrix of the standardized measure of niche dissimilarity between species,  $Coord$  geographical coordinates,  $VP$  variation partitioning procedure,  $FWS$  forward selection procedure,  $dc-CA$  double-constrained correspondence analysis. Subscript *all* indicates that all predictors of a group are included in the analysis. Terms with no subscript refer to the variables selected during FWS

## Estimating and testing the variation parts

Traditional direct ordination method (canonical ordination analysis) with one set of predictors (e.g., E) decomposed the green part in Fig. 1a into two parts: variation in species composition (C) explained by the environmental predictors (CE) and residual variation (variation of C not explained by E). The last is calculated by the difference between the total variation or all eigenvalues in C, assessed by a correspondence analysis of C (CA C), from which the explained variance (the canonical eigenvalues) obtained in the CCA of C constrained by E is subtracted. If two groups of predictors are used (tables with explanatory variables), either for species, for sites, or both (Fig. 1b shows the dc-CA parts of explained variance in C, using T and E as predictors), with the residual variation this gives four parts. Besides the explained variance assessment, the effects of the predictors—simple, conditional (unique), and shared—are also tested for their significance, against the null hypothesis that their effect is random. We have used the Monte Carlo permutation test, with 999 unrestricted permutations. Analogously, there are eight parts in CENT space analysis, which we coded with letters from *a'* to *h'* (Fig. 1c), but only some parts (or sum of parts) could be tested. The full decomposition of CENTS space gives 16 variation parts coded with letters from *a* to *q* (Fig. 1d).

In the following text, we refer to double-constrained analyses by their name (e.g., dc-CA C~), where ~ is followed by the tables with the selected predictors. The symbol  $\times$  is used to separate two groups of predictors, the first related to sites (E or S or both) and the second to species (T or N or both), while + indicates the use of variables belonging to both data tables in the group. Predictors from a table used as covariates are denoted by placing table abbreviation after the | (i.e., vertical line) symbol. For instance, the expression dc-CA C~(E+S) $\times$ (T|N) means that a double-constrained analysis is performed on the species-by-sites data table (C), with sites constrained by variables from combined E and S, and species constrained by T, while variables in N table act as covariates for the effects of T.

The decomposition of the community space using dc-CA is based on the idea that the canonical (double-constrained) eigenvalues resulting from this analysis represent the intersection between the effects of the environment (i.e., E+S) and traits (i.e., T+N) on species composition. In the VP, the explained variation is decomposed in the two conditional effects of the two predictors (or, more often, groups of predictors) and their overlap. Simple effects can be measured and tested using dc-CAs without covariates, while the conditional effects can be measured and tested using the analyses with covariates. For instance, when considering only the CENT space, analyses dc-CA C~(E) $\times$ (T), dc-CA C~(E) $\times$ (N), dc-CA C~(E) $\times$ (T+N) are used to separate

the variation explained by T and N, when C is also constrained by E, and to test the simple effects. The explained variation in dc-CA C~(E) $\times$ (T+N) is the total explained variation (given by the sum of the conditional effect of T, N, and the shared variation), i.e.,  $e' + f' + h'$  (Fig. 1c). In dc-CA C~(E) $\times$ (T), the explained variation is the simple effect of T (conditional effect of T and the overlap, i.e.,  $e' + h'$ ), and similarly, in dc-CA C~(E) $\times$ (N), it is the simple effect of N (conditional effect of N and the overlap, i.e.,  $f' + h'$ ). The significance for each of the simple effects is given by the double-constrained test on all ordination axes performed during the analyses without covariates. When only one dominant gradient is expected in the data, the significance of the test on the first axis will be considered. To test the conditional effects, we use dc-CA C~(E) $\times$ (T|N) (giving the conditional effect of T and its significance, i.e.,  $e'$ ) and dc-CA C~(E) $\times$ (N|T) (giving the conditional effect of N and its significance, i.e.,  $f'$ ). The overlap ( $h'$ ) cannot be tested (Table 2).

To measure all the elements of the community space and to test all the parts that may be tested, we not only need the synthetic results of the dc-CAs, but also the results of the two CCA (or RDA in dc-PCA) testing the constraints either on sites (with E or S or both as predictors, i.e., the canonical eigenvalues for site-related predictors only) or on species (with T or N or both as predictors, i.e., the canonical eigenvalues for species-related predictors only). These also allow the additional testing of some CENT (and CENTS) element sums. The variation parts of the community space can be calculated from various combinations of dc-CAs with covariates and without covariates, used in turn. The unadjusted percentages of explained variation ( $R^2$ ) will then be calculated by dividing the various variation parts by the total variation in the community composition, extracted from the species-by-sites table in the unconstrained CA (or PCA). The adjusted  $R^2$  can also be calculated (Peres-Neto et al. 2006; for details concerning dc-CA see ter Braak and Šmilauer 2018 and the online support site for Canoco 5.1).

We present the application of our algorithm using Canoco 5.12 software (ter Braak and Šmilauer 2018). At present, this resource has certain advantages over other softwares. It performs dc-CA with and without covariates, with optional forward selection of predictors, testing unique and shared effects, performs dimensionality tests, and illustrates results by double-constrained ordination diagrams. In addition, for a double-constrained analysis, Canoco also presents the intermediate results, i.e., of the two CCAs (or PCAs), testing the constraints on sites and separately on species. We also propose an alternative, Canoco independent algorithm, without covariates and CCAs included as separate steps, given in the Online Resource. This algorithm was inspired by the R-code for dc-CA variation partitioning in Pinho et al. (2021).

**Table 2** Algorithm for variation partitioning of CENT space relating communities (C) to the environment (E), traits (T), and ecological niches (N)

Step	Analysis	Estimates	Significance tests ( $p$ )
0	CA C	Total variation = All_eg	
1.1'	dc-CA C~(E)×(T N)	$e' = \text{ceg}[(E) \times (T N)]$ $c' = \text{ceg}[T N] - e'$	$\{b' + e' + f' + h'\}$ $\{c' + e'\}$ $\{e'\}^{\text{dc}}$
1.2'	dc-CA C~(E)×(N T)	$f' = \text{ceg}[(E) \times (N T)]$ $d' = \text{ceg}[N T] - f'$	$\{d' + f'\}$ $\{f'\}^{\text{dc}}$
1.3'	dc-CA C~(E)×(T+N)	$h' = \text{ceg}[(E) \times (T+N)] - e' - f'$ $b' = \text{ceg}[E] - \text{ceg}[(E) \times (T+N)]$ $g' = \text{ceg}[T+N] - \text{ceg}[(E) \times (T+N)] - c' - d'$ $a' = \text{All\_eg} - b' - c' - d' - e' - f' - g' - h'$	$\{c' + d' + e' + f' + g' + h'\}$ $\{e' + f' + h'\}^{\text{dc}}$
1.4'	dc-CA C~(E)×(T)		$\{c' + e' + g' + h'\}$ $\{e' + h'\}^{\text{dc}}$
1.5'	dc-CA C~(E)×(N)		$\{d' + f' + g' + h'\}$ $\{f' + h'\}^{\text{dc}}$

The tilde~stands for canonical ordination analysis, followed by the predictors placed in round brackets, separated by the×symbol, for discriminating between predictors related to sites (E) and those linked to species (T or N or both). The total variation (all eigenvalues, coded as All-eg) is given by the correspondence analysis of C (CA C). Canonical eigenvalues are coded as ceg, and the predictors used are enclosed in brackets []. Vertical bar | separates the covariates (to the right of it), the lowercase letters correspond to the variation parts given in Fig. 1c. Significance tests are related to the variation parts given in braces {} for the CCA and  $\{\}^{\text{dc}}$  (with superscript) for the dc-CA, also using the formerly mentioned letters

The algorithm for the decomposition of CENT space and testing its variation parts are explicitly given in Table 2, and the algorithm for CENTS is shown in Table 3. In some studies, ecological niche data are not available or of no research interest; therefore in the Online Resource, we provide the algorithm for CEST (community–environment–space–traits), with and without covariates, and illustrates the results using a novel graphical display, because the classical Venn diagrams proved to be inefficient. We call the visualization of these variation partitioning results VADOC diagrams (a term derived from ‘variation partitioning in double-constrained ordination analyses with multiple predictor tables diagrams’). In every algorithm, step 0 represents the unconstrained correspondence analysis of C (CA C), which returns the total variation in the response variables table, meaning the total eigenvalues.

The algorithm for decomposition of the CENTS space (Table 3) comprises five steps (1–5) in addition to step 0. In the first two steps, the variation in the community constrained by the unique effects of E and S (with S covariate in step 1 and E covariate in step 2) is decomposed. Step 3 decomposes the variation in the community constrained both by E and S, while in the last steps, the community is constrained either by E (step 4) or S (step 5). Each step has five substeps (#.1–#.5), homologous with the five steps (any or each denoted by #). Thus, the first two substeps are represented by dc-CAs in which the unique effects of T and N (T in substep #.1 and N in substep #.2) are included. Substep #.3 are dc-CAs with T+N as species-related predictors, while in

the last two substeps T (substep #.4) and N (substep #.5) are included separately. For the calculation of CENTS elements, one needs only 9 of the total 25 substeps (in Table 3 one of the possible combinations of subsets is given), but the rest are needed to test various variation parts. For CENT space, one step with five substeps is needed (1.1'–1.5'), with only three substeps used to calculate variation parts (Table 2).

The formulas for calculating the variation parts involve the sum of canonical eigenvalues (coded as ceg) of the constrained analysis of predictor tables enclosed in brackets []. When only one group of predictors is used, the analysis is a CCA, and if both groups are present (separated by×), the eigenvalues refer to the dc-CA. For instance,  $b' = \text{ceg}[E] - \text{ceg}[(E) \times (T+N)]$  (step 1.3' in Table 2) means that the variation fraction termed  $b'$  in the CENT (Fig. 1c) is given by the subtraction of the sum of canonical eigenvalues of the double-constrained analysis involving E and (T+N) as predictors (dc-CA C~E×(T+N)), from the sum of canonical eigenvalues of the community constrained by the environmental variables (CCA C~E). Although the variation parts may be calculated and tested using several different combinations of analyses, only one is given in the proposed algorithm to avoid redundancy.

## Niche metrics

To explain the structure and functions of communities, independently of their species' traits, to link the five matrices described in Fig. 1d and describe the niches' properties in

**Table 3** Algorithm for variation partitioning of CENTS space relating communities (C) to the environment (E), space (S), traits (T), and ecological niches (N)

Step	Analysis	Estimates	Significance tests ( <i>p</i> )
0	CA C	Total variation = All_eg	
1.1	dc-CA C~(E S)×(T N)	$e = \text{ceg}[(E S) \times (T N)]$	$\{b + e + f + h\}$ $\{c + e + k + p\}$ $\{e\}^{\text{dc}}$
1.2	dc-CA C~(E S)×(N T)	$f = \text{ceg}[(E S) \times (N T)]$	$\{d + f + l + n\}$ $\{f\}^{\text{dc}}$
1.3	dc-CA C~(E S)×(T + N)	$b = \text{ceg}[E S] - \text{ceg}[(E S) \times (T + N)]$ $h = \text{ceg}[E S] - b - e - f$	$\{c + d + e + f + g + h + k + l + m + n + p + q\}$ $\{e + f + h\}^{\text{dc}}$
1.4	dc-CA C~(E S)×(T)		$\{c + e + g + h + k + m + p + q\}$ $\{e + h\}^{\text{dc}}$
1.5	dc-CA C~(E S)×(N)		$\{d + f + g + h + l + m + n + q\}$ $\{f + h\}^{\text{dc}}$
2.1	dc-CA C~(S E)×(T N)	$k = \text{ceg}[(S E) \times (T N)]$	$\{i + k + l + m\}$ $\{k\}^{\text{dc}}$
2.2	dc-CA C~(S E)×(N T)	$l = \text{ceg}[(S E) \times (N T)]$	$\{l\}^{\text{dc}}$
2.3	dc-CA C~(S E)×(T + N)	$i = \text{ceg}[S E] - \text{ceg}[(S E) \times (T + N)]$ $m = \text{ceg}[S E] - i - k - l$	$\{k + l + m\}^{\text{dc}}$
2.4	dc-CA C~(S E)×(T)		$\{k + m\}^{\text{dc}}$
2.5	dc-CA C~(S E)×(N)		$\{l + m\}^{\text{dc}}$
3.1	dc-CA C~(E + S)×(T N)	$c = \text{ceg}[T N] - \text{ceg}[(E + S) \times (T N)]$ $p = \text{ceg}[T N] - c - e - k$	$\{b + e + f + h + i + j + k + l + m + n + p + q\}$ $\{e + k + p\}^{\text{dc}}$
3.2	dc-CA C~(E + S)×(N T)	$d = \text{ceg}[N T] - \text{ceg}[(E + S) \times (N T)]$ $n = \text{ceg}[N T] - d - f - l$	$\{f + l + n\}^{\text{dc}}$
3.3	dc-CA C~(E + S)×(T + N)	$j = \text{ceg}[E + S - \text{ceg}[(E + S) \times (T + N)] - b - i$ $g = \text{ceg}[T + N] - \text{ceg}[(E + S) \times (T + N)] - c - d$ $q = \text{ceg}[(E + S) \times (T + N)] - e - f - h - k - l - m - n - p - q$ $a = \text{All\_eig} - b - c - d - e - f - g - h - i - j - k - l - m - n - p - q$	$\{e + f + h + k + l + m + n + p + q\}^{\text{dc}}$
3.4	dc-CA C~(E + S)×(T)		$\{e + h + k + m + p + q\}^{\text{dc}}$
3.5	dc-CA C~(E + S)×(N)		$\{f + h + l + m + n + q\}^{\text{dc}}$
4.1	dc-CA C~(E)×(T N)		$\{b + e + f + h + j + n + p + q\}$ $\{e + p\}^{\text{dc}}$
4.2	dc-CA C~(E)×(N T)		$\{f + n\}^{\text{dc}}$
4.3	dc-CA C~(E)×(T + N)		$\{e + f + h + n + p + q\}^{\text{dc}}$
4.4	dc-CA C~(E)×(T)		$\{e + h + p + q\}^{\text{dc}}$
4.5	dc-CA C~(E)×(N)		$\{f + h + n + q\}^{\text{dc}}$
5.1	dc-CA C~(S)×(T N)		$\{i + j + k + l + m + n + p + q\}$ $\{k + p\}^{\text{dc}}$
5.2	dc-CA C~(S)×(N T)		$\{l + n\}^{\text{dc}}$
5.3	dc-CA C~(S)×(T + N)		$\{k + l + m + n + p + q\}^{\text{dc}}$
5.4	dc-CA C~(S)×(T)		$\{k + m + p + q\}^{\text{dc}}$
5.5	dc-CA C~(S)×(N)		$\{l + m + n + q\}^{\text{dc}}$

The tilde~stands for canonical ordination analysis, followed by the predictors placed in round brackets, separated by the × symbol, for discriminating between predictors related to sites (E or S or both) and those linked to species (T or N or both). The total variation (all eigenvalues, coded as All-eg) is given by the correspondence analysis of C (CA C). Canonical eigenvalues are coded as ceg, and the predictors used are enclosed in brackets []. Vertical bar | separates the covariates (to the right of it), the lowercase letters correspond to the variation parts given in Fig. 1d. Significance tests are related to the variation parts given in braces {} for the CCA and {}<sup>dc</sup> (with superscript) for the dc-CA, also using the formerly mentioned letters

terms of dissimilarities, differences, and diversity, we also developed some new niche metrics. These are based on a niche overlap index we developed earlier (Sirbu 2006),

which includes both the resource use by species and their availability in the environment and varies between 0 (completely distinct niches) and 1 (identical niches). Since, for

our purposes, we found no suitable measure in the literature for fulfilling these requirements, we defined the  $HS_{ij}$ , as the standardized niche overlap or similarity index ( $H$  is in honor of Hurlbert, while  $S$  stands for Smith) between the species  $i$  and  $j$ :

$$HS_{i,j} = \frac{\sum_k a_k \sqrt{p_{i,k} p_{j,k}}}{\sqrt{(\sum_k a_k p_{i,k})(\sum_k a_k p_{j,k})}} \quad (1)$$

where  $i$  and  $j$  are indices of summation from 1 to  $s$  (number of species in the community),  $k$  is the index of summation from 1 to  $n$  (number of resources exploited),  $p_{ik}$  is the proportion of the species  $i$  using resource  $k$ , and  $a_k$  is the proportion of resources in the environment. In Sîrbu (2006),  $a_k$  represents the fraction of habitat category  $k$  within the total number of surveyed habitats.

In its original definition,  $HS$  was applied with categorical resources (e.g., habitat types, food items) or for gradients expressed as discrete series (e.g., time, season, counts of resource). For resources varying continuously along gradients (e.g., weight of food, caloric equivalent, tolerances to physical or chemical parameters of soil or water), the challenge of assessing the availability (the offer of the environment) and their use (requirements of the species) can be addressed with cubic spline interpolation functions and integral calculus (see the Online Resource). This is a novel methodological contribution proposed by us for ecological niches measured and compared along continuous resources.

$HS$  serves as the basis for other measures and is applied in different methods.

We define the *standardized measure of niche dissimilarity* between species  $i$  and  $j$  as:

$$DHS_{i,j} = 1 - HS_{i,j} \quad (2)$$

where  $DHS_{i,i} = 0$ .  $DHS$  gives a symmetrical matrix of niche dissimilarities, having 0 on the main diagonal. We used this metric to further relate the sites-by-species values ( $C$ ) to niche dissimilarities. For this purpose, we made a PCO analysis on the  $DHS$  matrix, saved the species scores on the coordinates axes (defining the  $N_{all}$  data table), and used an FWS procedure within a dc-CA for selecting those axes' scores that best predicted the species composition while accounting also for their response to environmental variables. The result was the  $N$  matrix that entered the algorithm along with  $E$ ,  $S$ , and  $T$ , for defining, measuring, and testing the CENT and CENTS spaces.

Another measure used as a synthetic parameter for community analysis was the mean amount of total difference between the niche of each species and all the others. We define  $MSDHS_i$  as the *mean niche difference or dissimilarity between species  $i$  and the other species within the*

*community*, this being also a measure of the *species uniqueness* (sensu Ricotta et al. 2016):

$$MSDHS_i = \frac{\sum_j DHS_{i,j}}{s} \quad (3)$$

For illustrative purposes, we used  $MSDHS$  as a trait, in a dc-CA, for linking communities to the environment.

We define the *ecological niche-based diversity measure*  $FDen$  as:

$$FDen_h = (1/2) \sum_i \sum_j DHS_{i,j} q_{h,i} q_{h,j} \quad (4)$$

where  $q_{h,i}$  is the relative abundance of species  $i$  in the sample or site  $h$ . This is a modified variant of the Rao quadratic entropy measure (Rao 1982), which considers  $DHS$ . Because the product results in small values, and to standardize the range, we introduced a new variant ( $FDsa$ ) and further used in our applications the *standardized ecological niche-based diversity measure*  $FDsb$  that ranges between 0 and 1:

$$FDsa_h = (1/2) \sum_i \sum_j \sqrt[3]{DHS_{i,j} q_{h,i} q_{h,j}} \quad (5)$$

$$FDsb_h = \frac{\left[ (1/2) \sum_i \sum_j \sqrt[3]{DHS_{i,j} q_{h,i} q_{h,j}} \right] - \min(FDsa_h)}{\max(FDsa_h) - \min(FDsa_h)} \quad (6)$$

We compared the proposed ND indices with several measures of taxonomic diversity (TD): the number of species ( $s$ ), Hill's (N2) diversity measure (Hill 1973 ap. ter Braak and Šmilauer 2018), Shannon (H) entropy measure, and Pielou's evenness index (J), which relates H to the logarithm of species number (Krebs 2014; ter Braak and Šmilauer 2018). To assess FD we used the standardized version (Šmilauer and Lepš 2014; ter Braak and Šmilauer 2018) of Rao functional diversity, but computed it only on those traits selected by forward selection (FWS) in the dc-CA. We performed a PCO on traits using Gower distance and then calculating Rao diversity (FD(Rao)) on those scores (PCOs) and the sites-by-species ( $C$ ) table (Laliberté and Legendre 2010). We suggested and applied several methods to compare these measures and test their possible redundancy. We used our dataset for comparative purposes to show different ways to display these facets of diversity. We performed an RDA with all the described diversity metrics as response variables and  $E$  as predictors to plot and compare TD, FD, and ND. For this analysis, we constructed  $t$  value biplots (ter Braak and Šmilauer 2018) to evaluate the significance of the response of these diversity measures to each of the selected  $E$ . We performed a dc-CA with all the selected space and species predictors and plotted FD(Rao),  $FDen$ , and  $FDsb$  on the double-constrained space separately. We constructed



a dendrogram based on the Pearson correlation coefficients using an average agglomerative method to compare all the diversity measures used in the analyses.

### Case study: freshwater mollusk communities from the middle Olt River

To illustrate the application of the measures and methods described here, we used our data on the habitats' dimension of ecological niches (Sîrbu 2006) and the data from a study on freshwater mollusk community conducted in the middle Olt River basin, Romania. These data concern 23 species of freshwater mollusks from 20 sites. We conducted sampling and data analysis between March and July 2020. The area of reference is a river sector 83 km long, situated in the middle Olt River, which comprises a series of six reservoirs built in the second half of the twentieth century. The area extends from the town of Făgăraș (45.8512° N, 24.9733° E) to the Carpathian gorges (45.5317° N, 24.2721° E). The study aimed to explore responses of native communities to human pressure (hydro-technical works and others), and it will be published in extenso elsewhere.

For the dc-CA and other unimodal analyses, response data (weights in g) were log-transformed (by the relation  $y' = \log(1 + y)$ ) and standardized by sample total; thus, the final response variables were the relative dominance of species in terms of wet weight.

The selected E were flow (denoted Flow), an ordinal variable ranging from 1 (stagnant water, usually near the dams of the reservoirs) to 6 (rapid, turbulent flow in the gorges), an estimation of human impact (Impact), ranging from 1 (close to the natural condition of the river) to 6 (stagnant waters near the dams in the reservoir, with artificial banks and large parts of the substrate covered in concrete or other artificial materials), and the variable measuring the distance from the sampling station to the nearest dam downstream (Dis\_dam, in m).

T were also selected by FWS in the dc-CA analysis. We selected maximal shell size of mollusks (SizeM), an ordinal variable ranging from (1) < 2.5 to (6) > 100 mm, and feeding type (FeedT), a factor with the levels: scraper (SCR), scraper and sediment feeder (SS), sediment feeder (SED), scraper and filter feeder (SF), filter feeder (F), sediment and filter, or suspension and deposit feeders (SEDF). The values for the traits were based on the works of Falkner et al. (2001), Glöer (2019), Killeen et al. (2004), Piechocki and Wawrzyniak-Wydrowska (2016), and Sîrbu and Benedek (2018).

To calculate the niche metrics, we used a dataset of species-by-habitat categories obtained during an earlier survey conducted between 1996 and 2003 (Sîrbu 2006). We used 28 types of habitats as resource descriptors. We considered the ratios of species occurrence in each of these categories as measures of resource use and the proportions of habitat

categories within all the 405 surveyed habitats for assessing the resource availability in the research area.

We used PTC Mathcad 14 and PTC Mathcad Prime 5.0.0.0 (Parametric Technology Corporation) to calculate niche metrics and Canoco 5.12 (ter Braak and Šmilauer 2018) for the multivariate analyses and data visualization. We computed correlations between diversity indices using Hmisc package in R (Harrell 2020) and built the dendrogram in R version 3.6.1 (R Core Team 2019).

## Results

### CENT(S) variation partitioning for the mollusk dataset

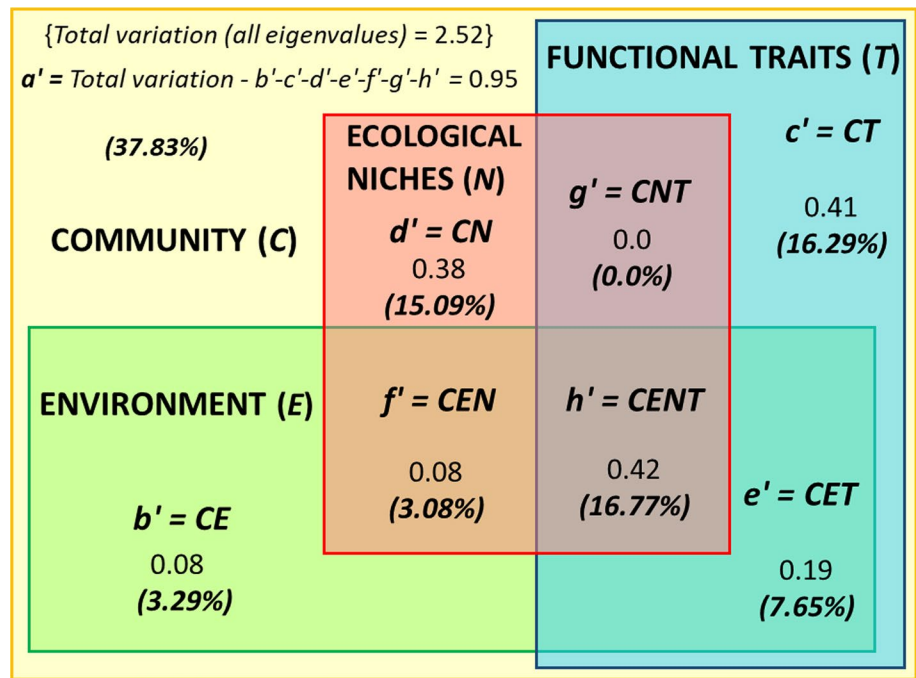
In our proposed algorithm for decomposition of the explained variation in C, the selection of dc-CA steps (Table 2 for CENT and Table 3 for CENTS analyses) depends on the questions asked in the particular study. In our case study, for illustration and discussion, we performed all the analyses and calculated all the variation parts (Fig. 2 for the CENT analysis, Fig. 3 for the CENTS analysis) and tested all the possible parts (or sums of parts) (Table 4). We present the VADOC diagrams in the two variants, the second (Fig. 3) being adapted based on a suggestion made by ter Braak (pers. comm.).

The total variation in the community composition was 2.52 (sum of all eigenvalues), all predictors explaining 68%. The remaining 32% was the residual variation ( $a$ ), which cannot be explained by any of the selected predictors. Simple effects (variance in community composition explained by groups of predictors) were all significant ( $p$  on all axes < 0.05), similar and higher in T (explained variation is 40.5%) and N (34.8%), followed by E (30.8%) and S (11.5%). The strongest conditional (unique) effects were those of T and N (14.7% for T and 13.9% for N), compared to those of E (4.9%) and S (3.9%). Also, the lack of overlap between N and T in the absence of environmental and spatial predictors (their shared effect being null) clearly indicates that niche and traits, at least in our study, are anything but redundant, and adding niche data is beneficial for the model and for understanding the underlying mechanisms of communities structural and functional responses. In our study, if the niche data are removed, the unexplained variation in C rises from about one-third (32%) to about half (45.9%). This variation is also an argument in favor of using independent data on resource use of species in community ecology.

### Niche-based diversity indices

MSDHS, used as a functional trait in a dc-CA with E, explained 7.25% of the variation (6.58% adjusted) in the

**Fig. 2** VADOC diagram showing decomposition of the variation in community composition explained by environment, traits, and niches. Letters *b'* to *h'* represent the parts of variation explained by different combinations of predictor groups, *h'* is the CENT (acronym from community–environment–niche–traits) overlap space, and *a'* represents the residual variation that the selected predictors cannot explain. The figures represent the components of variation in the mollusk communities in our case study, given both as the absolute variation explained and as percents of the total variation in community composition. 20 communities and 3 samples from each site (60 samples in total) were analyzed



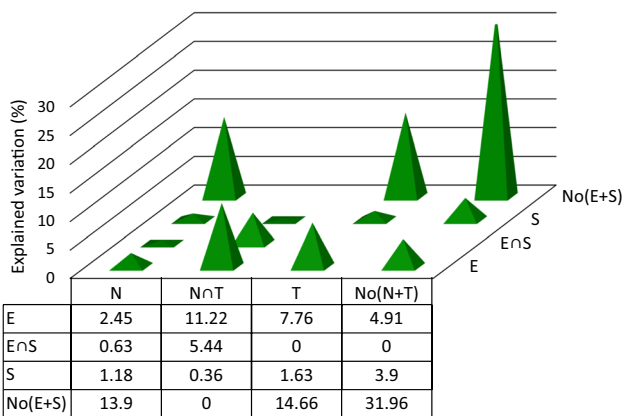
species–environment relationship, its effect being significant ( $p = 0.006$  on all axes). This result shows that, if needed, some niche metrics can be used as a proxy for traits, in this case showing how distinctness of niches relates to environment predictors.

In the RDA with TD, FD, and ND as response variables and E as predictors, the adjusted explained variation in the diversity measures was 34.4%, and the axes

were significant (pseudo-F on all axes = 4.3,  $p = 0.013$ ). The simple effect was significant for all three environmental variables, and it was strongest for Impact (43.3%,  $p = 0.006$ ). However, conditional effects of Flow and the distance to the nearest downstream placed dam were not significant, showing their redundancy with the human pressure in explaining the mollusk diversity in our study area. FDSb and, marginally, *s* had a significant response (positive in this case) to Impact (Fig. 4).

In the dendrogram of diversity measures (Fig. 5), the two ND were not grouped, in contrast to the TD of heterogeneity (N2, H, and J). FDen joined at a greater distance the TD group, while FDSb was closely linked to *s*, and these two were further linked to FD(Rao). This way of illustrating relationships might help distinguish the meaning of the three different types of diversities (taxonomic, functional, and niche based) and also argues against their redundancy.

Another way to visualize the relationships of ND with the other variables is to build a dc-CA diagram with E, T, N, and S, and in that space to project one of these measures by either a loess function or by a GLM or GAM (the last two methods also allowing for testing the model significance), resulting in a contour plot with the isolines in the double-constrained space. We illustrated the variation of FDSb using a loess function (Fig. 6), the coefficient of determination being  $R^2 = 51.8\%$ . FDSb increased toward the average values of the environmental predictors (the intersection of axes) and decreased with Impact and Flow. By contrast, the FDen increased along the first double-constrained axis, showing a negative relationship with Impact, and the classical FD(Rao)



**Fig. 3** Another type of VADOC diagram, showing decomposition of the variation in community composition (C) explained by environment (E), space (S), traits (T), and niches (N). The figures represent the components of variation in the mollusk communities in our case study, given in percents of the total variation in community composition (all eigenvalues = total variation in  $C = 2.52$ ). Predictors are identified by letters in caps,  $\cap$  means the shared effects of two predictors, while No stands for none of the predictors included in parentheses. 20 communities and 3 samples from each site (60 samples in total) were analyzed

**Table 4** Significance test results for the variation partitioning analyses

Step	dc-CA analysis	Significance tests ( <i>p</i> all axes) from CCA	Significance test ( <i>p</i> all) from dc-CA
<b>I CENT SPACE</b>			
1.1'	C~(E)×(TIN)	{ <i>b'</i> + <i>e'</i> + <i>f'</i> + <i>h'</i> } = 0.001 { <i>c'</i> + <i>e'</i> } = 0.058	{ <i>e'</i> } = 0.004
1.2'	C~(E)×(NIT)	{ <i>d'</i> + <i>f'</i> } = 0.061	{ <i>f'</i> } = 0.382
1.3'	C~(E)×(T+N)	{ <i>c'</i> + <i>d'</i> + <i>e'</i> + <i>f'</i> + <i>g'</i> + <i>h'</i> } = 0.005	{ <i>e'</i> + <i>f'</i> + <i>h'</i> } = 0.001
1.4'	C~(E)×(T)	{ <i>c'</i> + <i>e'</i> + <i>g'</i> + <i>h'</i> } = 0.024	{ <i>e'</i> + <i>h'</i> } = 0.01
1.5'	C~(E)×(N)	{ <i>d'</i> + <i>f'</i> + <i>g'</i> + <i>h'</i> } = 0.001	{ <i>f'</i> + <i>h'</i> } = 0.001
<b>II CENTS SPACE</b>			
1.1	C~(EIS)×(TIN)	{ <i>b</i> + <i>e</i> + <i>h</i> + <i>f</i> } = 0.001 { <i>c</i> + <i>e</i> + <i>k</i> + <i>p</i> } = 0.058	{ <i>e</i> } = 0.004
1.2	C~(EIS)×(NIT)	{ <i>d</i> + <i>f</i> + <i>l</i> + <i>n</i> } = 0.061	{ <i>f</i> } = 0.54
1.3	C~(EIS)×(T+N)	{ <i>c</i> + <i>d</i> + <i>e</i> + <i>f</i> + <i>g</i> + <i>h</i> + <i>k</i> + <i>l</i> + <i>m</i> + <i>n</i> + <i>p</i> + <i>q</i> } = 0.005	{ <i>e</i> + <i>f</i> + <i>h</i> } = 0.001
1.4	C~(EIS)×(T)	{ <i>c</i> + <i>e</i> + <i>g</i> + <i>h</i> + <i>k</i> + <i>m</i> + <i>p</i> + <i>q</i> } = 0.024	{ <i>e</i> + <i>h</i> } = 0.001
1.5	C~(EIS)×(N)	{ <i>d</i> + <i>f</i> + <i>g</i> + <i>h</i> + <i>l</i> + <i>m</i> + <i>n</i> + <i>q</i> } = 0.001	{ <i>f</i> + <i>h</i> } = 0.001
2.1	C~(SIE)×(TIN)	{ <i>i</i> + <i>k</i> + <i>l</i> + <i>m</i> } = 0.088	{ <i>k</i> } = 0.343
2.2	C~(SIE)×(NIT)		{ <i>l</i> } = 0.271
2.3	C~(SIE)×(T+N)		{ <i>k</i> + <i>l</i> + <i>m</i> } = 0.192
2.4	C~(SIE)×(T)		{ <i>k</i> + <i>m</i> } = 0.231
2.5	C~(SIE)×(N)		{ <i>l</i> + <i>m</i> } = 0.158
3.1	C~(E+S)×(TIN)		{ <i>e</i> + <i>k</i> + <i>p</i> } = 0.002
3.2	C~(E+S)×(NIT)		{ <i>f</i> + <i>l</i> + <i>n</i> } = 0.326
3.3	C~(E+S)×(T+N)	{ <i>b</i> + <i>e</i> + <i>f</i> + <i>h</i> + <i>i</i> + <i>j</i> + <i>k</i> + <i>l</i> + <i>m</i> + <i>n</i> + <i>p</i> + <i>q</i> } = 0.001	{ <i>e</i> + <i>f</i> + <i>h</i> + <i>k</i> + <i>l</i> + <i>m</i> + <i>n</i> + <i>p</i> + <i>q</i> } = 0.001
3.4	C~(E+S)×(T)		{ <i>e</i> + <i>h</i> + <i>k</i> + <i>m</i> + <i>p</i> + <i>q</i> } = 0.001
3.5	C~(E+S)×(N)		{ <i>f</i> + <i>h</i> + <i>l</i> + <i>m</i> + <i>n</i> + <i>q</i> } = 0.001
4.1	C~(E)×(TIN)	{ <i>b</i> + <i>e</i> + <i>f</i> + <i>h</i> + <i>j</i> + <i>n</i> + <i>p</i> + <i>q</i> } = 0.001	{ <i>e</i> + <i>p</i> } = 0.004
4.2	C~(E)×(NIT)		{ <i>f</i> + <i>n</i> } = 0.385
4.3	C~(E)×(T+N)		{ <i>e</i> + <i>f</i> + <i>h</i> + <i>n</i> + <i>p</i> + <i>q</i> } = 0.001
4.4	C~(E)×(T)		{ <i>e</i> + <i>h</i> + <i>p</i> + <i>q</i> } = 0.001
4.5	C~(E)×(N)		{ <i>f</i> + <i>h</i> + <i>n</i> + <i>q</i> } = 0.001
5.1	C~(S)×(TIN)	{ <i>i</i> + <i>j</i> + <i>k</i> + <i>l</i> + <i>m</i> + <i>n</i> + <i>p</i> + <i>q</i> } = 0.036	{ <i>k</i> + <i>p</i> } = 0.3
5.2	C~(S)×(NIT)		{ <i>l</i> + <i>n</i> } = 0.141
5.3	C~(S)×(T+N)		{ <i>k</i> + <i>l</i> + <i>m</i> + <i>n</i> + <i>p</i> + <i>q</i> } = 0.009
5.4	C~(S)×(T)		{ <i>k</i> + <i>m</i> + <i>p</i> + <i>q</i> } = 0.014
5.5	C~(S)×(N)		{ <i>l</i> + <i>m</i> + <i>n</i> + <i>q</i> } = 0.006

Codes of the explained variation parts are given in Fig. 2 for CENT and Fig. 3 for CENTS analyses

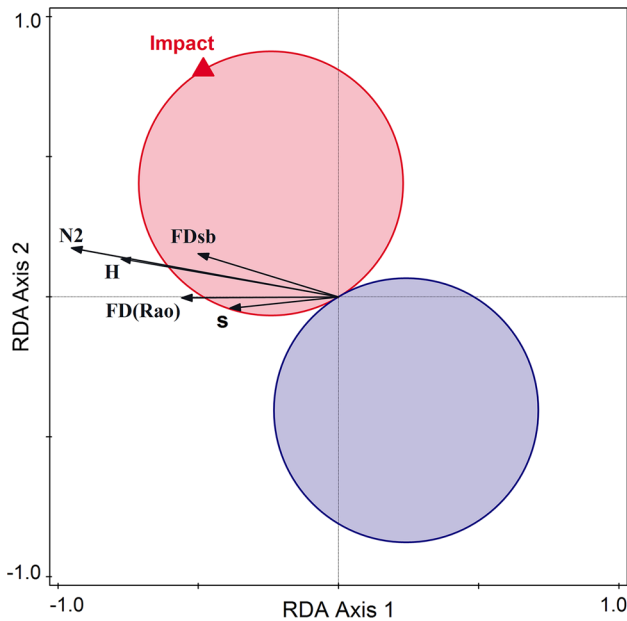
showed a linear increase with the reduction of Flow and the increase of Impact (figures not shown here).

### Discussion

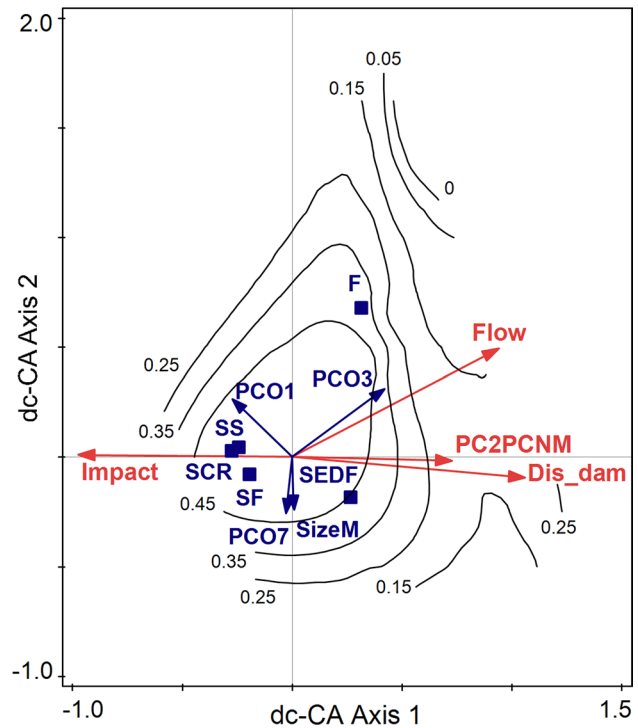
To expand the methodological framework of life–environment analyses, we modified the original three-table dc-CA method to introduce a five-table ordination algorithm (Fig. 1). Our approach combines, for the first time, the variability associated with the environment and locations where a species occurs and the variability associated with the traits and ecological niche of species within sites, incorporating

niche predictors, separately and independently from traits. It also allows the decomposition of the variation in community composition, something that has been also done recently for the CEST space (Pinho et al. 2021). We show additional procedures for variation decomposition and testing in a five-matrix system, namely the CENTS space. In addition, various parts of the variation were tested and illustrated by a novel visualizing technique, the VADOC diagrams, which we present in two variants.

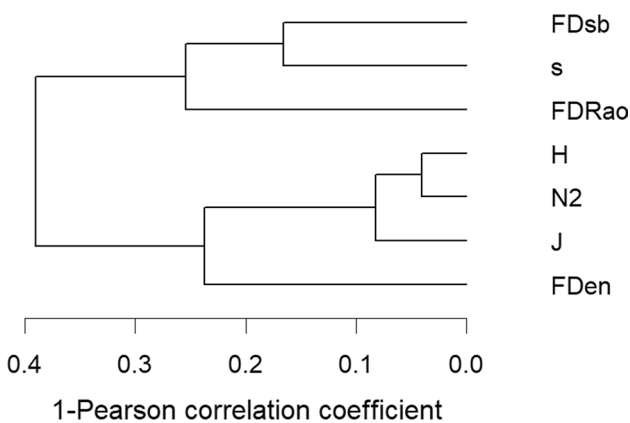
Our study reveals that evolutionary and adaptive signals may be of prevailing importance, compared to environmental and geographical predictors (the last categories being used more frequently), in explaining the quantitative structure of



**Fig. 4** t-Value biplot (Van Dobben circles) showing the relationships between human pressures (Impact, as predictor) and the diversity measures (as response variables): number of species (s), N2—Hill’s diversity measure, H—Shannon entropy measure, FD(Rao)—Rao’s measure of functional diversity, FDsb—standardized niche-based diversity measure. The pink circle indicates the area of significant positive response of diversity measures to impact, and the blue circle the area of significant negative response. This is an attribute plot of the RDA between the diversity measures and the impact, distance to the dam and the flow as predictors; their simple effects are all significant, while the conditional effect is significant only for the Impact ( $R^2=43.31\%$ ,  $P\text{-adj}=0.006$ ). 20 communities and 3 samples from each site (60 samples in total) were analyzed



**Fig. 6** Double-constrained correspondence analysis (dc-CA) with all selected predictors and traits, with FDsb niche-based standardized diversity plotted by loess functions. Impact intensity of human impact, Flow flow, Dis\_dam distance to the nearest dam downstream, SizeM maximal shell size of mollusks, SCR scraper, SEDF sediment and filter feeder, SF scraper and filter feeder, SS scraper and sediment feeder, PC2PCNM the second spatial eigenvector, selected from db-MEM, PCO1, PCO3, and PCO7 first, third, and seventh PCOs’ axes scores selected by forward selection from the axes resulting in principal coordinates analysis (PCO) on niche dissimilarities (the original DHS matrix). 20 communities and 3 samples from each site (60 samples in total) were analyzed



**Fig. 5** Dendrogram on the complement of Pearson correlation coefficients between diversity measures, built by the average agglomerative linkage method (diversity measures: FDsb standardized niche-based diversity, s number of species, FDRao Rao functional diversity, H Shannon entropy, N2 Hill’s diversity, J Pielou’s evenness index, FDen ecological niche-based diversity measure). 20 communities and 3 samples from each site (60 samples in total) were analyzed

ecological communities. Using niche measures separately from traits might bring new insights into community assembly and reveal the adaptive features acquired by species during their evolution. Although usually used interchangeably, we argue that niche and traits descriptors can and should be treated as independent sources of information, and on a practical level, as different matrices in double-constrained analyses. Traits are characteristics identified mostly on the individual level, while niche metrics are features revealed on the species level, adding meaning and expanding the use of integrative methodology applied in ecology.

Peng et al. (2021) found dc-CA useful in assessing links between multiple stressors and ecosystem health, something we also emphasize by expanding its use and addressing the possibility to assess pressures, for instance, by measuring the significance and intensity of the relationship between the niche-based diversity and human impact.

Looking at spatial patterns may be important because the spatial distribution of species is directly influenced by

environmental variables, while functional traits are expected to be indirectly related to space through their relationship with the environment (Pavoine et al. 2011). Therefore, if spatial signals are found in the traits but these are not correlated with the environment, some key environmental processes (present or past) might have been overlooked.

In their paper on trait–environment relationships of benthic macroinvertebrate communities sampled from the Danube River, Peng et al. (2021) chose not to separate the spatial component from the environmental one because, in their river dataset, most variables were highly correlated with the distance from the source, with the feeble possibility to separate their effects. In our study case, although the dataset also comes from sampling sites along a river, we chose to include the spatial predictors because the impact of anthropic activities, mainly dam building, alters the natural succession of environmental conditions along the river.

The environmental filtering hypothesis predicts that within a community, the environment selects species with similar traits (Grime 2006). Environmental filtering is thought to be a major mechanism structuring communities. However, our ability to accurately infer filtering based on community composition data was recently questioned by Tucker and Cadotte (2013), as competition can give rise to the same patterns as those caused by environmental filtering. Although distinguishing and testing the causes behind community assembly need manipulative studies of communities, including in the analyses of community structure, the ecological niche alongside species traits in correlation with the environment (and space) could help us understand the ecological and evolutionary processes that lead to certain species composition. In our approach to community assembly, we consider, and our study shows, that reducing ecological niches to an attribute characterized or derived from functional traits (Violle and Jiang 2009; Kearney et al. 2010) depletes the study of relationships between communities and their environment of an essential multivariate space of explanatory factors. Comparing the species through their niches may add new insight into evolutionary and adaptive constraints that link the structure and functions of ecological systems. Many studies have revealed that constrained analyses that relate groups of interacting populations to external factors (ter Braak and Šmilauer 2018; Šmilauer and Lepš 2014) need to be completed by adding information on space variability (Dray et al. 2006, 2012; Peres-Neto and Legendre 2010; Sharma et al. 2011), functional traits (Dolédec et al. 1996; Dray and Legendre 2008; Céréghino et al. 2018; ter Braak et al. 2018; Xu et al. 2019), and phylogenetic relationships (Desvignes et al. 2003; de Bello et al. 2005), to build and test a framework for understanding the underlying mechanisms of the responses of ecological systems to environmental changes, the laws that govern processes and offered services, and to predict their trajectories. We do not

provide new methodology, but use existing knowledge and methods to step forward. Considering that traits and niche metrics characterize different ecological levels, introducing the niche in multivariate analyses aiming at linking communities to their environment means also testing for (dis) similarities between species concerning the use and partitioning of resources, adaptations, and interactions, i.e., for the intersection between their evolution and ecology.

If niche parameters are not available or desired,  $N$  can be a table of ecological indicator values for species, such as Ellenberg values in vascular plants, Grime's CRS plant strategies, or saprobity values in diatoms (e.g., ter Braak 1987; ter Braak and van Dam 1989). This possibility extends the potential ecological area of application of the paper (ter Braak, pers. comm.).

The MSDHS characterizes the mean difference between the niche of a species and all the others. This niche-based index represents a novel measure of the species-facet level of redundancy (Ricotta et al. 2016) and can be used to evaluate derived measures, such as the *mean functional dissimilarity or vulnerability*. Further, MSDHS should be used for characterizing the community level of niche-based redundancy and the uniqueness of the functions associated with each species (*functional uniqueness*) by adapting the measures described by Ricotta et al. (2016). The derived measures should replace the *functional dissimilarities* with the *niche dissimilarities*, as described in Eqs. 4–6. MSDHS may also enter new analyses (of multiple regression, for instance) as a synthetic response variable or, as shown before, to be considered as a trait within a dc-CA. De Cáceres et al. (2011) provide a framework for calculating resource niche metrics (space resource) using the distance-based PCO approach. Instead, we have chosen a direct ordination method, but all the metrics introduced by De Cáceres et al. (2011) can be easily calculated on the indices and used in the methods we have presented here.

Assessing FD in ecological communities is very promising for studying the response of diversity to environmental gradients and the effects of diversity on ecosystem functioning (Lepš et al. 2006) as well as on ecosystem functions and services (de Bello et al. 2010, 2016; Cadotte et al. 2011; Boersma et al. 2016; Carmona et al. 2016). Defining and measuring relationships between the different types of diversity is a complex task (Mouchet et al. 2010; de Bello et al. 2010, 2016; Pavoine and Bonsall 2011; Lepš 2013; Carmona et al. 2016). We have shown that both FDen and FDsb (as measures of ND) are not redundant with other diversity measures, including FD(Rao), but explain in a complementary manner the responses of communities to environmental gradients. Thus, they should not be used as a proxy for each other, at least not a priori. In our case study, FDsb proved to be more sensitive to human impact than other diversity measures, suggesting that ND could be a valuable tool in

environmental monitoring and ecological management and as an indicator when evaluating success in conservation projects.

There is a growing interest in using FD for developing new methods for environmental monitoring and human impact evaluation and mitigation (Cianciaruso et al. 2009; Laliberté and Legendre 2010; Cadotte et al. 2011; Boersma et al. 2016) and for life cycle assessment (Ahmed et al. 2018). We agree with all these and add that ND measures should be used in the future, since they show the same possibilities and are not redundant with the classical FD measures. Although relying on (and inspired by) the Rao entropy measure, their levels of reference and meaning are different.

The methods presented here are useful when applied to tables with high variability of data (biological and environmental). For this reason, we selected for illustrative purposes data from a heterogeneous river sector, where the structure of the mollusk communities reflects the diversity of habitats (ranging from flowing to stagnant waters in the reservoirs), environmental conditions, and sources of human impact. The study group is also heterogeneous, including gastropods and bivalves of different origins (natives, native invasive, allochthonous invasive) with various traits.

### Expanding further the framework

As Pavoine et al. (2011) suggested for their method using RLQ, our algorithm can also be extended to address other key issues in ecology, such as the interspecific relationships between plants and herbivores or pollinators. In this case, the traits and niche of herbivores or pollinators will be related to the traits and niche of plants, with the possibility of accounting for the direct effect of environment and space on plants and their indirect effect on herbivores.

Phylogeny is considered another source of variation and explanation in the relationship between species and environment regardless of whether traits are included or not in the analysis (Desdevises et al. 2003; Šmilauerová and Šmilauer 2007; Pavoine and Bonsall 2011; de Bello et al. 2017; Xu et al. 2019). Our integrative approach could be further expanded, including the sixth matrix of patristic distances or genetic distances, depending on the available data and research questions. This expansion could be done by combining our algorithm with that proposed by Desdevises et al. (2003) and using specific tests and related methods (Campbell et al. 2011).

In effect, the same principles and methods of variation partitioning of two predictor tables might be used for three or more tables. Our  $2 \times 2$  variation partitioning algorithm in dc-CA can be extended to a  $3 \times 3$  algorithm by adding, for instance, invasive species as site descriptors and phylogeny as species descriptors. Further, this extension could be generalized into an  $(n \times k)$  variation partitioning, where  $n$  tables

are related to the sites and  $k$  to the species. Besides, our illustration of results may be used directly to interpret both simple and conditional effects of various kinds of predictors on the community structure, relying on the ability to test the fractions of explained variance. Besides, if a non-trivial experimental design is involved, the permutation scheme in dc-CA can be easily adapted to suit and reflect the sampling protocol.

In dc-CA, the number of both environmental variables (space predictors included) and traits (niche predictors included) needs to be significantly smaller than the number of cases and species (ter Braak and Šmilauer 2018). Therefore, our algorithm is not suitable for datasets with many predictors and few cases (sites and species). An alternative may be to replace the dc-CA with a double-asymmetrical co-correspondence analysis (CoCA, which considers unimodal responses to underlying gradients) or its linear counterpart, the co-inertia analysis (CoIA), which can be used with many predictors and few objects (cases or samples). Although non-symmetrical CoCA can be done by existing R packages and functions, the double-asymmetrical CoCA and CoIA still have to be developed.

Another possible expansion is to step outside the two-dimensional table of the double-constrained method and adopt a multi-level approach, such as an  $n$ -dimensional constrained analysis. By this expansion, we propose the theoretical and methodological development toward an ‘omni-spaces explanatory ecology of communities’: by this, meaning the consideration and use of an unlimited number and types of explanatory data tables, constraining either sites or species (rows and columns), in any number of response data sets (multiple interacting or linked communities).

Unlike environmental variables and traits, which have a straightforward meaning, the significance and underlying mechanisms of niche overlap may vary greatly, depending on how the niche is defined and what data are used to evaluate it. Here, we have illustrated a relatively simple case study where we included the use and partitioning of habitat types, related mainly to life–environment interactions. When food resources or interspecific relationships are considered, the focus will be on biotic processes. In complex studies, considering different aspects of the niche, several niche overlap matrices can be combined, calculating the centroids, the mean, or another measure of central tendency of the similarities for each pair of species.

### Conclusions

Including ecological niche metrics, especially niche dissimilarities, in the CENTS model, as a separate data table, adds the functional interaction between the species to the model that accounts for distribution, space variability,

environmental predictors, and biological traits, and sometimes also for phylogenetic relationships. This can lead to new theoretical and applied research objectives. We have shown that one can use an existing three-table ordination method to analyze relationships among five or more datasets. Maybe the future will show that for challenging the complex environmental issues related to overpopulation, resource exploitation, degradation of habitats, invasive species, pollution, and loss of ecological functions, new methods have to be developed for linking more datasets from different informational spaces. The multivariate approach is already mature; new needs arise by linking and integrating multi-spaces, each of them being a system of multivariate datasets linked within and among the spaces. Searching for solutions to such complex issues will no longer result in predicting a structure or a group of functions, but tracing and modeling complex trajectories of ecological systems simultaneously along time, space, and other gradients. This will also mean a complex and multidisciplinary approach and the development of new mathematical tools for assisting the needs of ecological research. In this frame, we completed only a small step but showed the urge to move forward.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00442-021-05006-6>.

**Acknowledgements** This paper was written within the project financed by Lucian Blaga University of Sibiu Research Grants LBUS-IRG-2017-03. We express our gratitude to those who have guided and supported our work: Jan Lepš and Stéphane Dray in the first stages, Robert K. Rose, who commented on the first draft and assisted with English usage, Petr Šmilauer and Cajo ter Braak during the entire development of the paper, the latter also sharing ideas and unpublished materials. Special regards to the handling editor Bryan L. Brown who had the wisdom and patience for supporting our work during the difficult transition from the first version of the manuscript to the present shape while acting also as a professional adviser. We are also thankful to the Editor in Chief Joel Trexler, to the Editorial Assistant Jessica Sanchez and to Rini Sharon Jeyaraj for their support during the review process, and to the anonymous referees who helped improve the paper. Legal permit for field surveys was provided by the Sibiu subsidiary of the National Agency for Protected Natural Areas in Romania, to which we also express sincere regards.

**Author contribution statement** IS and AMB contributed equally to this work; IS elaborated the ideas, concepts, niche measures and statistical graphics, IS and AMB conceived the research, developed the algorithms, analyzed the data, and wrote the revisions, AMB and IS wrote the original draft, IS and MS collected and processed the data from the field. All authors commented on previous versions of the manuscript and read and approved the final manuscript.

**Funding** This study was funded by Lucian Blaga University of Sibiu research grants LBUS-IRG-2017-03.

**Availability of data and material** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request and will be soon available on a digital repository.

**Code availability** The statistical software R is freeware and Canoco is licensed, and its license is the personal property of the corresponding author. Mathcad is licensed, and we used a Mathcad Education – University Edition Subscription license in 2020.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethics approval** All applicable institutional and national guidelines for the care and use of animals were followed, as well as those of the EU Council Directive 86/609/EEC on experimental use of animals. Legal permits for field work in protected areas were provided by the Sibiu subsidiary of the National Agency for Protected Natural Areas in Romania.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmed DA, van Bodegom, Tukker A (2018) Evaluation and selection of functional diversity metrics with recommendations for their use in life cycle assessments. *Int J Life Cycle Assess* 24:485–500. <https://doi.org/10.1007/s11367-018-1470-8>
- Boersma KS, Dee LE, Miller SJ, Bogan MT, Lytle DA, Gitelman AI (2016) Linking multidimensional functional diversity to quantitative methods: a graphical hypothesis-evaluation framework. *Ecology* 97:583–593. <https://doi.org/10.1890/15-0688>
- Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol Model* 153:51–68
- Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73:1045–1055
- Cadotte MW, Carscadden K, Mirotnick N (2011) Beyond species: functional diversity and the maintenance of ecological processes and services. *J Appl Ecol* 48:1079–1087. <https://doi.org/10.1111/j.1365-2664.2011.02048.x>
- Campbell V, Legendre P, Lapointe FJ (2011) The performance of the congruence among distance matrices (CADM) test in phylogenetic analysis. *Evol Biol* 11:64. <https://doi.org/10.1186/1471-2148-11-64>
- Carmona CP, de Bello F, Mason NWH, Lepš J (2016) Traits without borders: integrating functional diversity across scales. *Trends Ecol Evol* 31:382–394. <https://doi.org/10.1016/j.tree.2016.02.003>
- Cérèghino R, Pillar VD, Srivastava DS, de Omena PM, MacDonald AAM, Barberis IM, Corbara B, Guzman LM, Leroy C, Bautista

- FO, Romero GQ, Trzcinski MK, Kratina P, Debastiani VJ, Gonçalves AZ, Marino NAC, Farjalla VF, Richardson BA, Richardson MJ, Dézerald O, Gilbert B, Petermann J, Talaga S, Piccoli GCO, Jocqué M, Montero G (2018) Constraints on the functional trait space of aquatic invertebrates in bromeliads. *Funct Ecol* 32:2435–2447. <https://doi.org/10.1111/1365-2435.13141>
- Cianciaruso MV, Batalha MA, Gaston KJ, Petchey OL (2009) Including intraspecific variability in functional diversity. *Ecology* 90:81–89. <https://doi.org/10.1890/07-1864.1>
- de Bello F, Berg MP, Dias ATC, Diniz-Filho JAF, Götzenberger L, Hortal J, Ladle RJ, Lepš J (2005) On the need for phylogenetic ‘corrections’ in functional trait-based approaches. *Folia Geobot* 50:349–357. <https://doi.org/10.1007/s12224-015-9228-6>
- de Bello F, Lavergne S, Meynard CN, Lepš J, Thuiller W (2010) The partitioning of diversity: showing Theseus a way out of the labyrinth. *J Veg Sci* 21:992–1000. <https://doi.org/10.1111/j.1654-1103.2010.01195.x>
- de Bello F, Carmona CP, Lepš J, Szava-Kovats R, Pärtel M (2016) Functional diversity through the mean trait dissimilarity: resolving shortcomings with existing paradigms and algorithms. *Oecologia* 180:933–940. <https://doi.org/10.1007/s00442-016-3546-0>
- de Bello F, Šmilauer P, Diniz-Filho JAF, Carmona CP, Lososová Z, Herben T, Götzenberger L (2017) Decoupling phylogenetic and functional diversity to reveal hidden signals in community assembly. *Methods Ecol Evol* 8:1200–1211. <https://doi.org/10.1111/2041-210X.12735>
- De Cáceres M, Sol D, Lapidra O, Legendre P (2011) A framework for estimating niche metrics using the resemblance between qualitative resources. *Oikos* 120:1341–1350. <https://doi.org/10.1111/j.1600-0706.2011.19679.x>
- Desdés Y, Legendre P, Azouzi L, Morand S (2003) Quantifying phylogenetically structured variation. *Evolution* 57:2647–2652. <https://doi.org/10.1111/j.0014-3820.2003.tb01508.x>
- Diniz-Filho JAF, Bini LM, Rodriguez MA, Rangel TFLVB, Hawkins BA (2007) Seeing the forest for the trees: partitioning ecological and phylogenetic components of Bergmann’s rule in European Carnivora. *Ecography* 30:598–608. <https://doi.org/10.1111/j.2007.0906-7590.04988.x>
- Dolédec S, Chessel D, ter Braak CJF, Champely S (1996) Matching species traits to environmental variables: a new three-table ordination method. *Environ Ecol Stat* 3:143–166
- Dray S, Legendre P (2008) Testing the species traits-environment relationships: the fourth-corner problem revisited. *Ecology* 89:3400–3412. <https://doi.org/10.1890/08-0349.1>
- Dray S, Chessel D, Thioulouse J (2003) Co-inertia analysis and the linking of ecological tables. *Ecology* 84:3078–3089
- Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbor matrices (PCNM). *Ecol Model* 196:483–493. <https://doi.org/10.1016/j.ecolmodel.2006.02.015>
- Dray S, Péllissier R, Couternon P, Fortin MJ, Legendre P, Peres-Neto PR, Bellier E, Bivand R, Blanchet FG, De Cáceres M, Dufour AB, Heegaard E, Jombart T, Munoz F, Oksanen J, Thioulouse J, Wagner HH (2012) Community ecology in the age of multivariate multiscale spatial analysis. *Ecol Monogr* 82(3):257–275. <https://doi.org/10.1890/11-1183.1>
- Falkner G, Obdržálek P, Castella E, Speight MCD (2001) Shelled Gastropoda of Western Europe. Verlag der Friedrich-Held-Gesellschaft, München
- Glöer P (2019) The freshwater Gastropods of the West-Palaearctic. Volume I, Fresh- and brackish waters except spring and subterranean snails. Identification key, Anatomy, Ecology, Distribution. Neustadt/Holstein. [www.muchowdruck.de](http://www.muchowdruck.de)
- Grime JP (2006) Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. *J Veg Sci* 17:255–260. <https://doi.org/10.1111/j.1654-1103.2006.tb02444.x>
- Kearney M, Simpson SJ, Raubenheimer D, Helmuth B (2010) Modelling the ecological niche from functional traits. *Philos Trans R Soc B* 365:3469–3483. <https://doi.org/10.1098/rstb.2010.0034>
- Killeen JJ, Aldridge DC, Oliver PG (2004) Freshwater Bivalves of Britain and Ireland. Occasional Publication 82, Field Studies Council and AIDGAP: Preston Mountford, UK
- Krebs C (2014) Ecological methodology. Niche measures and resource preferences, chap 14, 3rd edn, pp 597–653. [www.zoology.ubc.ca/~krebs/downloads/krebs\\_chapter\\_14\\_2017.pdf](http://www.zoology.ubc.ca/~krebs/downloads/krebs_chapter_14_2017.pdf)
- Kühn I, Nobis MP, Durka W (2009) Combining spatial and phylogenetic eigenvector filtering in trait analysis. *Global Ecol Biogeogr* 481:745–758. <https://doi.org/10.1111/j.1466-8238.2009.00481.x>
- Laliberté E, Legendre P (2010) A distance-based framework for measuring functional diversity from multiple traits. *Ecology* 91:299–305. <https://doi.org/10.1890/08-2244.1>
- Legendre P, Legendre L (2012) Numerical ecology, Third English. Elsevier, Oxford
- Legendre P, Garzin R, Harmelin-Vivien M (1997) Relating behaviour to habitat: solutions to the fourth corner problem. *Ecology* 78:547–562
- Lepš J (2013) Diversity and ecosystem function. In: van der Maarel M, Franklin J (eds) *Vegetation ecology*, Second. Wiley Blackwell, Hoboken, pp 308–346. <https://doi.org/10.1002/9781118452592>
- Lepš J, de Bello F, Lavorel S, Berman S (2006) Quantifying and interpreting functional diversity of natural communities: practical considerations matter. *Preslia* 78:481–501
- Melo-Merino SM, Reyes-Bonilla H, Lira-Noriega A (2020) Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. *Ecol Model*. <https://doi.org/10.1016/j.ecolmodel.2019.108837> (Article 108837)
- Mouchet MA, Villéger S, Mason NWH, Moullot D (2010) Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. *Funct Ecol* 24:867–876. <https://doi.org/10.1111/j.1365-2435.2010.01695.x>
- Pavoine S, Bonsall MB (2011) Measuring biodiversity to explain community assembly: a unified approach. *Biol Rev* 86:792–812. <https://doi.org/10.1111/j.1469-185X.2010.00171.x>
- Pavoine S, Vela E, Gachet S, de Béclair G, Bonsall M (2011) Linking patterns in phylogeny, traits, abiotic variables and space: a novel approach to linking environmental filtering and plant community assembly. *J Ecol* 99:165–175. <https://doi.org/10.1111/j.1365-2745.2010.01743.x>
- Peng FJ, ter Braak CJF, Rico A, den Brink PJV (2021) Double constrained ordination for assessing biological trait responses to multiple stressors: a case study with benthic macroinvertebrate communities. *Sci Total Environ* 754:142171. <https://doi.org/10.1016/j.scitotenv.2020.142171>
- Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. *Glob Ecol Biogeogr* 19:174–184. <https://doi.org/10.1111/j.1466-8238.2009.00506.x>
- Peres-Neto PR, Legendre P, Dray S, Borcard D (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecol* 87:2614–2625
- Peterson AT, Soberón J (2012) Species distribution modeling and ecological niche modeling: getting the concepts right. *Nat Conserv* 10:102–107. <https://doi.org/10.4322/natcon.2012.019>
- Piechocki A, Wawrzyniak-Wydrowska B (2016) Guide to freshwater and Marine Mollusca of Poland. Bogucki Wydawnictwo Naukowe, Poznań
- Pinho et al (2021) Functional biogeography of Neotropical moist forests: trait-climate relationships and assembly patterns of tree communities. *Glob Ecol Biogeogr* 30:1430–1446. <https://doi.org/10.1111/geb.13309>



- Rao CR (1982) Diversity and dissimilarity coefficients: a unified approach. *Theor Popul Biol* 21:24–43. [https://doi.org/10.1016/0040-5809\(82\)90004-1](https://doi.org/10.1016/0040-5809(82)90004-1)
- Ricotta C, de Bello F, Moretti M, Caccianiga M, Cerabolini BEL, Pavoine S (2016) Measuring the functional redundancy of biological communities: a quantitative guide. *Methods Ecol Evol* 7:1386–1395. <https://doi.org/10.1111/2041-210X.12604>
- Sharma S, Legendre P, De Cáceres M, Boisclair D (2011) The role of environmental and spatial processes in structuring native and non-native fish communities across thousands of lakes. *Ecography* 34:762–771. <https://doi.org/10.1111/j.1600-0587.2010.06811.x>
- Șirbu I (2006) Aspects concerning the distribution and ecology of the freshwater molluscs from the Romanian Inner Carpathian Basin. *Heldia* 6:115–134
- Șirbu I, Benedek AM (2018) Trends in Unionidae (Mollusca, Bivalvia) communities in Romania: an analysis of environmental gradients and temporal changes. *Hydrobiologia* 810:295–314. <https://doi.org/10.1007/s10750-017-3173-8>
- Šmilauer P, Lepš J (2014) Multivariate analysis of ecological data using Canoco 5, 2nd edn. Cambridge University Press, Cambridge
- Šmilauerova M, Šmilauer P (2007) What youngsters say about adults: seedling roots reflect clonal traits of adult plants. *J Ecol* 95:406–413
- Soberón J, Osorio-Olvera L, Peterson T (2017) Diferencias conceptuales entre modelacion de nichos y modelacion de areas de distribucion. *Revista Mexicana De Biodiversidad* 88:437–441. <https://doi.org/10.1016/j.rmb.2017.03.011>
- Sterk M, Gort G, Klimkowska A, van Ruijven J, van Teeffelen AJA, Wamelink GWW (2013) Assess ecosystem resilience: Linking response and effect traits to environmental variability. *Ecol Indic* 30:21–27. <https://doi.org/10.1016/j.ecolind.2013.02.001>
- ter Braak CJF (1987) Unimodal models to relate species to environment. Agricultural Mathematics Group, Wageningen. <http://edepot.wur.nl/201452>
- ter Braak CJF, Šmilauer P (2018) Canoco reference manual and user's guide: software for ordination (version 5.10). Microcomputer Power, Ithaca
- ter Braak CJF, van Dam H (1989) Inferring pH from diatoms: a comparison of old and new calibration methods. *Hydrobiologia* 178:209–223. <https://doi.org/10.1007/BF00006028>
- ter Braak CJF, Šmilauer P, Dray S (2018) Algorithms and biplots for double constrained correspondence analysis. *Environ Ecol Stat* 25:171–197. <https://doi.org/10.1007/s10651-017-0395-x>
- Tucker CM, Cadotte MW (2013) Unifying measures of biodiversity: understanding when richness and phylogenetic diversity should be congruent. *Divers Distrib* 19:845–854. <https://doi.org/10.1111/ddi.12087>
- Violle C, Jiang L (2009) Towards a trait-based quantification of species niche. *J Plant Ecol* 2:87–93. <https://doi.org/10.1093/jpe/rtp007>
- Xu J, Dang H, Wang M, Chai Y, Guo Y, Chen Y, Zhang C, Yue M (2019) Is phylogeny more useful than functional traits for assessing diversity patterns under community assembly processes? *For-ests*. <https://doi.org/10.3390/f10121159> (Article 1159)

## Software

- Canoco 5.12 is registered software and can be acquired from Microcomputer Power, 111 Clover Lane Ithaca, NY 14850 USA. <https://www.microcomputerpower.com>
- Harrell Jr FE with contributions from Dupont C. and many others (2019) Hmisc: Harrell Miscellaneous. R package version 4.2–0. <https://CRAN.R-project.org/package=Hmisc>
- PTC Mathcad 14 and Mathcad Prime 5.0 are registered software developed by the Parametric Technology Corporation. <https://www.ptc.com/en/products/mathcad>. <https://www.mathcad.com/>
- R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Support site: <http://www.canoco5.com/>