Received: 15 November 2022 | Accepted: 6 April 2023

DOI: 10.1111/1365-2435.14344

Functional Ecology | BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

N-dimensional hypervolumes in trait-based ecology: Does occupancy rate matter?

Alex Laini[1] | Thibault Datry[2] | Benjamin Wong Blonder[3]

[1]Dipartimento di Scienze della Vita e Biologia dei Sistemi, Università di Torino, Torino, Italy

[2]INRAE, UR RiverLy, Centre de Lyon-Villeurbanne, Villeurbanne Cedex, France

[3]Department of Environmental Science, Policy, and Management, University of California Berkeley, Berkeley, California, USA

Correspondence
Alex Laini
Email: alex.laini@unito.it

Funding information
ANBI Emilia-Romagna; Autorità di Bacino Distrettuale del Fiume Po; Regione Emilia-Romagna

Handling Editor: Joseph A. Tobias

Abstract

1. Many methods for estimating the functional diversity of biological communities rely on measuring geometrical properties of n-dimensional hypervolumes in a trait space. To date, these properties are calculated from individual hypervolumes or their pairwise combinations. Our capacity to detect functional diversity patterns due to the overlap of multiple hypervolumes is, thus, limited.
2. Here, we propose a new approach for estimating functional diversity from a set of hypervolumes. We rely on the concept of occupancy rate, defined as the mean or absolute number of hypervolumes enclosing a given point in the trait space. Furthermore, we describe a permutation test to identify regions of the trait space in which the occupancy rate of two sets of hypervolumes differs.
3. We illustrate the utility of our approach over existing methods with two examples on aquatic macroinvertebrates. The first example shows how occupancy rate relates to the stability of trait space utilisation due to increased flow intermittency and allows the identification of taxa in regions of the trait space with low occupancy rates. The second example shows how the permutation test based on occupancy rates can detect differences in trait space utilisation due to river morphology variation even with a high degree of overlap among input hypervolumes.
4. Our newly developed approach is particularly suitable in functional diversity analysis when investigating patterns of overlap among multiple hypervolumes. We emphasise the need to consider analyses based on occupancy rate into functional diversity estimation.

KEYWORDS
functional diversity, functional redundancy, functional stability, hypervolume, macroinvertebrates

1 | INTRODUCTION

Functional diversity represents a key component of biodiversity that influences the functioning of ecosystems (Mouchet et al., 2010; Tilman et al., 2014). Community ecologists rely on functional diversity metrics derived from traits as a proxy of the functions provided by biological assemblages to the ecosystems because of the difficulties in measuring functions directly (Carmona et al., 2017; Hatfield et al., 2018). These metrics generally perform better than species richness or abundance in predicting ecosystem functioning

due to their intrinsic link to the mechanism-generating functions (Gagic et al., 2015; Lucas et al., 2021).

Several methods for estimating functional diversity exist, many of them relying on the n-dimensional hypervolume concept (Hutchinson, 1957). The first attempt to define and quantify n-dimensional hypervolumes derived from the convex-hull concept, the smallest convex set of multidimensional space enclosing all points (Cornwell et al., 2006; Villéger et al., 2008). Convex-hulls are sensitive to outliers and do not consider disjunctions or holes in the hypervolume, thus impairing the computation of their volume and their overlap with other hypervolumes (Blonder et al., 2014; Mammola et al., 2021). To address these problems, Blonder et al. (2014, 2018) proposed a method that describes n-dimensional hypervolumes as a set of uniformly distributed random points each associated with a probability density estimate for the underlying dataset, generated via kernel density estimation or other modelling approaches (e.g. support vector machines). Irrespective of the method used, geometrical properties (e.g. volume) of n-dimensional hypervolumes have become established functional diversity metrics (Mammola & Cardoso, 2020; Villéger et al., 2008).

Current metrics based on n-dimensional hypervolumes are limited because they can only handle information about one (alpha diversity measures) or two hypervolumes (beta diversity measures) simultaneously. Size-based metrics, that quantify the amount of trait space (e.g. area or volume) enclosed by a set of species, are calculated individually for each community (Mammola & Cardoso, 2020; Mason et al., 2005; Villéger et al., 2008). On the other hand, position-based metrics, that measure the shift in the multidimensional space (e.g. centroids shift, overlap), are limited to pairwise combinations of hypervolumes (Brown et al., 2020; Guillerme et al., 2020; Loranger et al., 2016). All these metrics do not consider patterns emerging from the overlap of multiple hypervolumes (e.g. tens, hundreds), with some regions potentially showing more overlap than others.

Similar problems arise when analysing functional diversity metrics because common approaches do not quantify the amount of shared trait space by multiple hypervolumes effectively. Most approaches, including regression-based techniques (Maas et al., 2021) and null modelling approaches (Loranger et al., 2016), target the mean response of functional diversity metrics calculated on each hypervolume (e.g. significance of the difference in the mean volume among groups). Moreover, although position-based metrics are well suited to measure the amount of shared trait space, they are limited to pairwise comparisons and become infeasible to calculate (nearly 5×10^5 comparisons for 1000 hypervolumes) and difficult to analyse (pseudo-replication issues) with a high number of hypervolumes. The use of sequential pairwise comparisons could be useful to investigate the overlap of multiple hypervolumes, but it is problematic because sequential operations may introduce noise (e.g. increased bias or variance) into estimates. Lastly, when communities can be grouped according to a grouping factor (e.g. invaded vs. control), some authors calculate functional diversity metrics by pooling data to obtain a single community for each group (e.g. Takács et al., 2021), thus losing information about heterogeneity in trait space utilisation.

To cover this gap, we introduce the concept of occupancy rate as the mean or absolute number of hypervolumes including a point in the trait space. The occupancy rate is intended to reflect the heterogeneity of trait space overlap among multiple hypervolumes, which could be in turn associated with relevant ecological processes. For example, when studying biological communities, regions of the trait space with low overlap could be associated with random dispersal events while regions with high overlap could be related to a core set of functions shared among multiple communities. Furthermore, species in areas of the trait space with low occupancy rates could show low redundancy within a specific group of communities. This is even more important when comparing groups of hypervolumes, for which size- and position-based metrics can return similar values even if the overlap patterns differ among groups (Figure 1).

Here we propose an approach to quantify the occupancy rate through probabilistic n-dimensional hypervolumes (Blonder et al., 2014) and develop a permutation test to evaluate differences in occupancy rates among groups of hypervolumes. We test the performance of the new approach with simulations, and we demonstrate its advantages over existing methods with two examples grounded in the ecology of aquatic macroinvertebrates. These organisms are among the most important indicators of river quality and are strongly dependent on temporal and spatial flow patterns (Bo et al., 2017; Sarremejane et al., 2020). The first example assesses the effect of flow intermittence (period of time during the year with no or lack of flow) accounting for the body size of aquatic macroinvertebrates and shows how occupancy rates relate to environmental gradients and functional redundancy. The second example assesses variations in trait space overlap among macroinvertebrate communities due to river morphology and shows how occupancy rates can highlight among-group differences even with highly overlapped hypervolumes. We compare the results of our approach with those obtained using current functional diversity metrics and discuss their link with community assembly processes, such as dispersal and environmental filtering.

2 | MATERIALS AND METHODS

2.1 | Probabilistic n-dimensional hypervolumes, a brief overview

The occupancy framework builds on the approach proposed by Blonder et al. (2014) and Blonder et al. (2018) to estimate probabilistic n-dimensional hypervolumes. Briefly, these authors provide different methods (hyperbox kernel, gaussian kernel, support vector machine, estimation using arbitrary functions) for building hypervolumes from a set of points in an n-dimensional space. Here we focus on the Gaussian kernel density estimate (KDE) to describe the principles underlying hypervolume constructions, but similar principles apply to the other methods. In a Gaussian KDE, all points contribute to the overall probability density. At first, uniform random points are drawn from hyperellipses surrounding each point and then these points are resampled

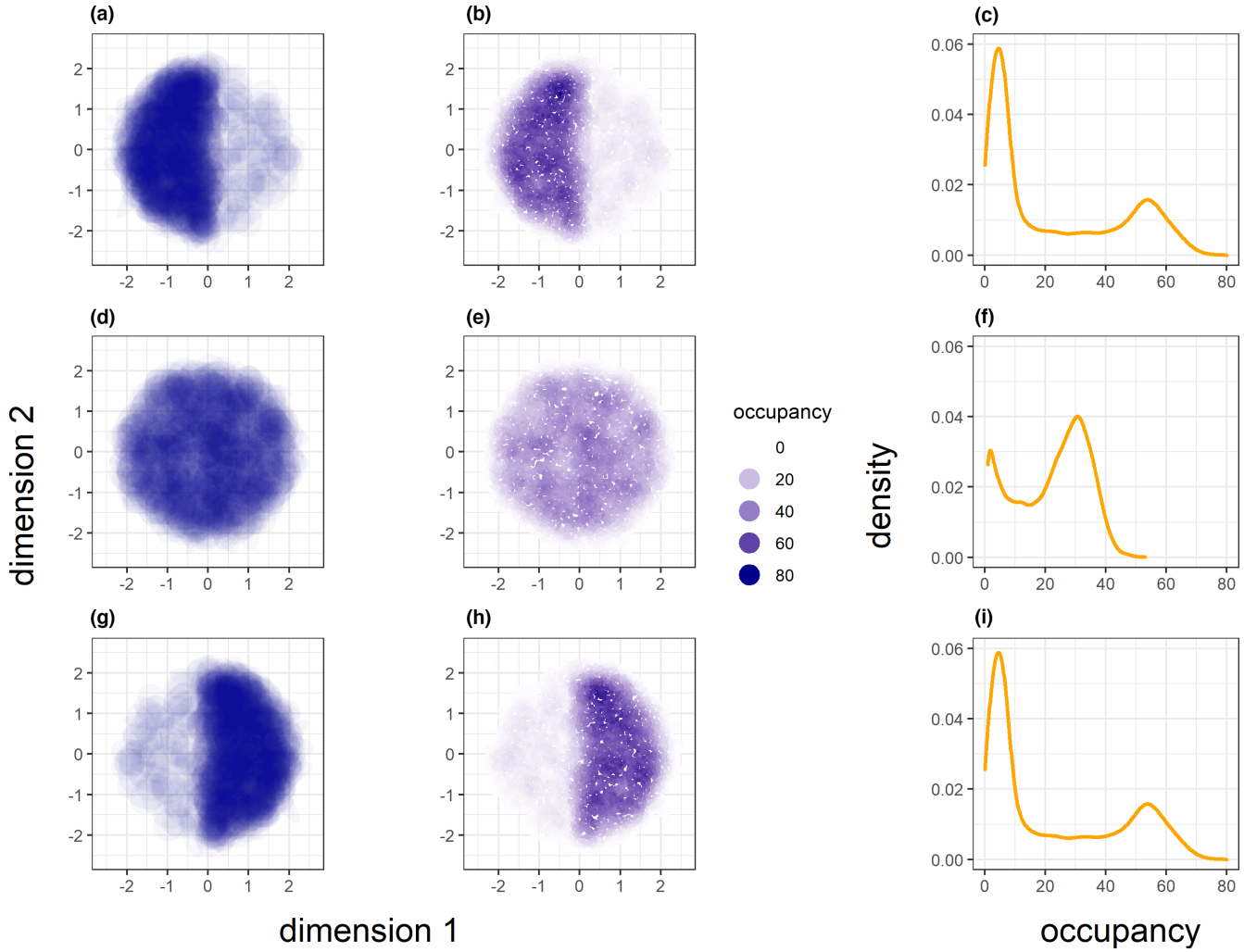


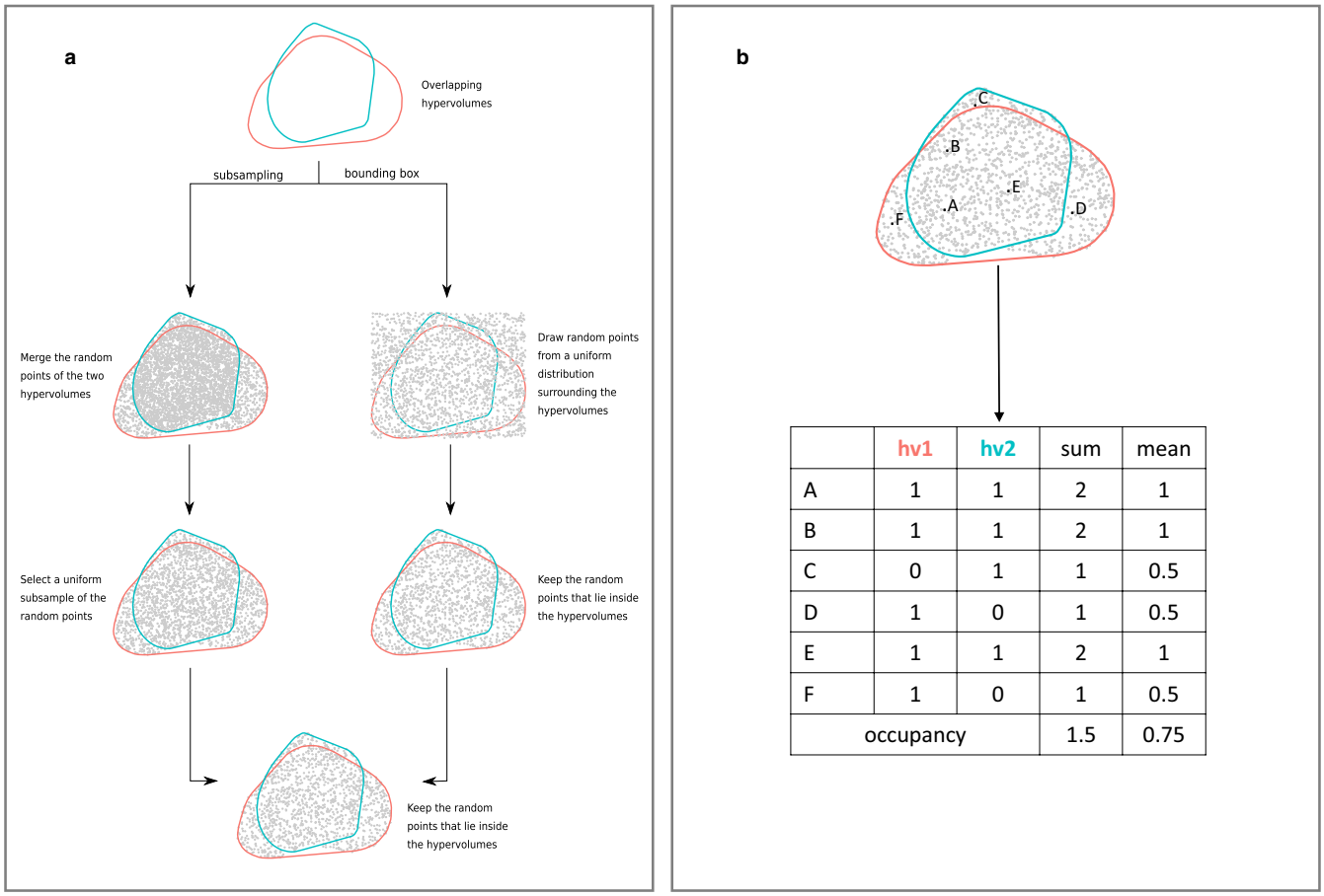
FIGURE 1 Limitations of current approaches to estimate functional diversity. Panels (a, d, g) include 1000 randomly generated ellipses. Mean volume and volume of the union of hypervolumes are equal among panels. Panels (b, e, h) represent occupancy rate estimates based on the probabilistic n-dimensional representation of hypervolumes of panels (a, d, g). Here, the occupancy rate is computed as the number of hypervolumes enclosing a given random point. Panels (c, f, i) represent the smoothed density estimate of occupancy rate values across all random points. Panels (a, d) differ in their overall occupancy rate patterns. Panels (a, g) have similar occupancy rate patterns, but a different trait space utilisation.

to uniform density. The kernel density is estimated for each random point and only the points above a threshold enclosing a certain quantile of the probability or volume are retained to describe the hypervolume (see box 1 in Blonder et al., 2018). Therefore, a probabilistic n-dimensional hypervolume is described by a set of random points from which several properties can be inferred (e.g. volume, position).

2.2 | Methods for calculating occupancy rates

The first step for calculating occupancy rates is to select a set of random points describing the union of $q > 1$ hypervolumes. Two methods are used to accomplish this task. The *subsampling* method joins the random points of the q hypervolumes and then selects a uniformly distributed subset of them. The method *box* creates a bounding box around the union of the q hypervolumes that is then filled with random points drawn from a uniform distribution at a specified density.

An inclusion test is used to determine if each random point lies within one or more of the q hypervolumes. The inclusion test is based on a recursive partitioning tree data structure to efficiently search for inclusion within a hyperellipse, implemented in the unexported function evalfspherical() of the package hypervolume (Blonder et al., 2014) for the R statistical software (R Core Team, 2021). Results of the inclusion test are converted into a binary matrix (1,0) with the number of columns and rows equal to the number of hypervolumes being considered and the number of random points, respectively. Within the binary matrix, 1 means that a hypervolume includes a given random point and 0 that it does not. For each random point, the occupancy rate is calculated by applying a function to each row (e.g. mean) of the binary matrix. The entire process for calculating occupancy rates is described in Figure 2. Besides the



	hv1	hv2	sum	mean
A	1	1	2	1
B	1	1	2	1
C	0	1	1	0.5
D	1	0	1	0.5
E	1	1	2	1
F	1	0	1	0.5
occupancy			1.5	0.75

FIGURE 2 Rationale of the occupancy rate framework. (a) Two methods are implemented to find a set of uniformly-distributed random points to describe the union of two or more hypervolumes. Once the set of random points has been obtained, the occupancy rate is calculated as a function (e.g mean) of the number of hypervolumes enclosing each random point. (b) Calculation example for occupancy rates based on sum and mean as summary statistic.

calculation of occupancy rates, the volume of the union of the input hypervolumes is calculated as the weighted sum of the volumes of input hypervolumes, where weights are proportional to the degree of overlap of the target hypervolume to the other hypervolumes. Both the *subsampling* and *box* methods introduce biases that cause a slight change in the volume of the input hypervolumes. To check whether the re-computed volumes are consistent with the original ones two measures of goodness of fit are calculated, the mean absolute error and the root mean square error (RMSE). Occupancy rates can be calculated for groups of observations such as those emerging from repeated measures over space, time or treatments.

2.3 | Between-group differences in occupancy rates

We developed a permutation test to evaluate differences in the occupancy rate of each random point between two groups of hypervolumes, as described further in Chen et al. (2023). For each pairwise group comparison, original hypervolumes are randomly assigned to one of the two groups under comparison. The test itself is performed by counting the number of times the observed differences are smaller or greater than those expected by chance, or by combining them for obtaining a two-tailed test.

2.4 | Simulations

2.4.1 | Comparison of the performances of the subsample and box methods

We compared the speed and accuracy of the *subsample* and *box* methods for different volumes of input hypervolumes. We randomly generated 5 hyperspheres from 100 randomly generated points at different radii (1, 5, 15) and for dimensions ranging from 1 to 5 using the package uniformly (Laurent, 2022). The performances of the two methods were evaluated by (i) recording the time needed to complete the analysis and (ii) calculating the normalised root mean square error (NMRSE) between input and re-constructed hypervolumes. NMRSE allows the comparison of the performances at different number of dimensions and was calculated as the RMSE divided by its standard deviation. The *box*

method was tested at two box densities (500 and 5000, the latter being the default value in the hypervolume package). We used the mean as the summary statistics for this and the other simulations because the sum provided analogous results with the simulation settings used in this work.

2.4.2 | Behaviour of the metric to increasing volume variability of input hypervolumes

We simulated the effect of an increasing variability of hypervolume volumes on occupancy rate estimation. To do this, we randomly generated numbers following a truncated normal distribution with a mean of five and an increasing standard deviation (from 0.1 to 4.1 by a step of 0.5 for a total of nine levels of standard deviation). For each level of standard deviation, we simulated 10 values representing the volume of the hypersphere. To build hypervolumes, we first derived the radius r of the hypersphere from the simulated volumes and then we randomly generated 1000 points from a uniform distribution within a hypersphere of radius r and origin at 0 with the package uniformly. These points were used as the input to build hypervolumes with a Gaussian KDE and default settings. We then calculated occupancy rates for each simulated level of standard deviation and for five dimensions (1–5) both with the *subsample* and the *box* methods and compared the results with the expected results. Seven box densities (500, 1000, 2500, 5000, 7500, 1000) were tested for the *box* method. The comparison with the expected results was performed using the slope and the R^2 of a linear regression model. The slope was used to evaluate the deviation of the observed from the expected result (with a perfect match the slope is equal to 1), while the R^2 to evaluate the accuracy of the method used (*subsample* or *box*).

2.4.3 | Behaviour of the permutation test

We tested the performance of the permutation test on two groups of simulated n-dimensional spheres. The first group consisted of 10 hypervolumes simulated by drawing 1000 random points from a hypersphere of radius 1 centred at −1 with an increasing number of dimensions (e.g. two dimensions: $x = -1$, $y = -1$; three dimensions: $x = -1$, $y = -1$, $z = -1$) using the package gena (Potanin, 2022). The second group was simulated in the same way as the first except that hypervolumes were centred at 1. By construction, all the hypervolumes should have the same volume (e.g. for two-dimensional hypervolumes a volume of 3.14), although deviations are expected due to the random nature of probabilistic n-dimensional hypervolume estimation. The deviation from the expected volume is expected to increase with an increasing number of dimensions because the same number of random points (1000) was used for hypervolume construction. Our newly developed approach was applied with and without specifying groups using both the *subsample* and *box* methods. Moreover, we tested the effect of an increasing box density on the results of the *box* methods.

2.5 | Case studies

2.5.1 | Example 1: Does flow intermittence favour small-bodied species of aquatic macroinvertebrates?

We evaluated the effect of flow intermittence on the body size of aquatic macroinvertebrates along a flow intermittence gradient. Evidence exists that small body size promotes survival at intermittent sites (Arias-Real et al., 2022; Piano et al., 2020) because reduced dimensions allow fast development and population growth (Bonada et al., 2007) and resistance in the hyporheic zone (Bêche et al., 2006). At flow resumption, species with small body size can quickly recolonize intermittent sites. A gradual decrease in the volume of hypervolumes built on body size is, thus, expected with increasing distance from upstream, colonist-source sites. After some months from flow resumption, a complete recolonisation even from larger-bodied species can occur and, thus, the trait space volume is expected to be comparable between perennial and intermittent sites. Therefore, we hypothesise a decrease in the mean occupancy rate with increasing distance from the upstream perennial source because of reduced volume at most intermittent sites at flow resumption. Moreover, we expect a decrease of the occupancy rate in the regions of the trait space occupied by moderate/large-bodied taxa with increasing flow intermittence.

To test our hypothesis, we used a subset of the dataset described in Datry (2012). Briefly, macroinvertebrates were collected from 10 sites distributed longitudinally across 20 km in the Albarine River (Eastern France). Across 2009 and 2010, samples were collected in autumn (October 2009 and November 2010; <3 weeks after flow resumption) and spring (April 2009 and 2010; 3–5 months after flow resumption) using standard approaches. The dataset includes presence–absence data of one perennial site and nine intermittent sites for a total of 40 observations (10 sites × 4 sampling dates). Annual flow intermittence, defined as the percent of the year that a site was non-flowing, spans from 0% (perennial site) to 36.1% (most intermittent site; Table 1). Flow intermittence increases with distance downstream along the intermittent section due to transmission losses into the underneath groundwater (Datry, 2012).

The trait space was built from the maximum body size of aquatic macroinvertebrates included in the DISPERSE database (Sarremejane et al., 2020). Maximum body size in the DISPERSE database is represented as fuzzy coded values divided into 7 categories (<0.25 cm, ≥0.25–0.5 cm, ≥0.5–1 cm, ≥1–2 cm, ≥2–4 cm, ≥4–8 cm, ≥8 cm). Fuzzy coding describes the affinity of a taxon for different categories of a given trait and each taxon can have affinities for more than one category (Chevene et al., 1994). For each category, affinity scores span from 0 (no affinity) to 3 (maximum affinity). For each taxon, affinity scores were divided by the row sum prior to the analysis. Affinity scores were then averaged when

TABLE 1 Performances of the two methods (subsample and box) for calculating occupancy rates for varying number of dimensions (2–5) and volumes (1, 5 and 15). The box method was tested at two box densities, 500 and 5000. Both the running time (seconds) and the normalised root mean square error (NMRSE) are reported.

No. of dimensions	Radius	Running time			NRMSE		
		Subsample	Box 500	Box 5000	Subsample	Box 500	Box 5000
2	1	2.81	0.42	0.71	0.31	0.30	0.14
2	5	2.91	1.25	9.83	0.18	0.09	0.08
2	15	2.74	8.53	78.5	0.04	0.08	0.09
3	1	10.75	1.30	2.26	0.36	0.11	0.03
3	5	11.1	3.78	24.1	0.20	0.06	0.07
3	15	10.7	21.2	213	0.18	0.05	0.05
4	1	44.2	4.7	6.7	0.37	0.16	0.05
4	5	48.5	10.6	59.8	0.33	0.05	0.03
4	15	52.1	54.8	495	0.34	0.15	0.14
5	1	166	15.9	21.5	0.45	0.12	0.07
5	5	205	35.3	184	0.57	0.11	0.09
5	15	192	161	1496	0.34	0.07	0.06

multiple traits were available for a given taxon (e.g. when genus-level information was available for traits but family-level information was present in the community composition dataset). The effect of trait averaging on the results was also assessed. We obtained the trait space with a classical multidimensional scaling (MDS) on the mixed-variables coefficient of distance, an algorithm suitable for fuzzy-coded data (Pavoine et al., 2009). For each of the 40 observations, hypervolumes were obtained with a Gaussian kernel density from the first two dimensions of the trait space. A fixed bandwidth calculated using the Silverman estimator on all the MDS coordinates was used to estimate all the hypervolumes. The mean occupancy rate was then calculated for each site using the *box* method at 5000 box density. To test our hypotheses, Pearson correlation between the percentage of intermittence and the mean occupancy rate was calculated. The correlation of the percentage of intermittence with the mean volume of input hypervolumes and the volume of the union of input hypervolumes was also calculated. Moreover, we estimated a linear mixed-model to test the relationship between the volume of individual hypervolumes and flow intermittence using site as the random intercept. The mixed model was estimated with the lmerTest package (Kuznetsova et al., 2017).

We then searched for taxa in regions of the trait space with low occupancy rate. To do this, we calculated the mean occupancy rate in the proximity of the MDS coordinate of each species within a circle with a radius of 0.029.

2.5.2 | Example 2: Does river morphology affect flow velocity preferences of aquatic macroinvertebrates?

The second example explored differences in the occupancy rate among mesohabitats differing in flow velocity using a trait space based on the flow velocity preferences of aquatic macroinvertebrates. Aquatic macroinvertebrates of running waters exhibit a wide range of flow preferences, that is, some taxa prefer habitats with low or null flow velocity while others prefer habitats with high flow velocity. Mesohabitats are geomorphic and hydraulic units, generally at 10^{-1}–10^{-3} m scale, that have distinct combinations of current velocity, depth and substrate (Belletti et al., 2017). Mesohabitats with higher flow velocity (e.g. riffle, glides) generally host different communities compared to those with low or null flow velocities (e.g. pool, backwater, isolated pond; Laini, Burgazzi, et al., 2022). We, thus, expect communities within the same mesohabitat to occupy different areas of the trait space due to river morphology.

We used a dataset of 150 samples collected in three braided rivers of northern Italy. Mesohabitat as well as the abundance of each taxon were recorded for each sample. Taxonomic resolution spanned from genus to family, because of difficulties in identifying larvae to species level. This is a common practice for macroinvertebrates and functional index calculation because most traits are conserved within higher taxonomic levels (Gutiérrez-Cánovas et al., 2015). Details about the methods used to collect the data are described in (Laini, Burgazzi, et al., 2022). The trait space was obtained from macroinvertebrate flow preferences according to Tachet et al. (2010). Flow preferences are coded into four categories (null, slow, moderate and fast) and each taxon can have an affinity for more than one category. Affinity scores, spanning from 0 (no affinity) to 3 (maximum affinity), were divided by the row sum prior to the analysis. Scores were averaged when multiple traits were available for a given taxon (e.g. genus-level information for traits and family-level information for the community composition dataset). The effect of trait averaging was also assessed. We obtained the trait space with an MDS as in the first example. Hypervolumes were obtained for each sample with a Gaussian

kernel density from the first two dimensions of the trait space. A fixed bandwidth calculated using the Silverman estimator on all the MDS coordinates was used to estimate all the hypervolumes and kernel density estimation was weighted by the abundance of each taxon.

We estimated a linear mixed-model to test the difference in the volume of input hypervolumes among mesohabitat using river as the random intercept. Pairwise differences among mesohabitats were then statistically evaluated with a post hoc test performed with the emmeans package (Lenth, 2022). We calculated both the mean volume of input hypervolumes and the volume of the union of input hypervolumes for each group. We calculated occupancy rates with the *box* method and a box density of 5000. To test our hypothesis that communities within the same mesohabitat occupy different regions of the trait space compared to other mesohabitats we used the newly developed permutation test. Taxonomic and trait information was managed with the R package biomonitoR (Laini, Guareschi, et al., 2022). Plots were obtained with the ggplot2 package (Wickham, 2009).

3 | RESULTS

3.1 | Simulation results

The *subsample* and *box* methods differed in their overall performances. The *box* method performed better than the *subsample* method by a magnitude of 1.02–12.3 in reconstructing input hypervolumes for each number of dimensions based on the NRMSE value (Table 1). The accuracy of the *box* method at the two box densities tested was similar, except for the simulation at volume 1. Within simulations with the same number of dimensions, the running time for the *subsample* method was similar irrespective of the volume tested. On the contrary, the *box* method showed an increased running time when increasing the volume of input hypervolumes. The *subsample* method was generally faster than the *box* method for volumes greater than 5, although time depended on the box density (500 faster than 5000) and the number of dimensions.

Overall, both methods showed good performances in detecting patterns when input hypervolumes differed in their volume variability (Table 2). The R^2 of the model fit between expected and observed occupancy rates was close to or greater than 0.9 for all the simulations, except for the *subsample* method in 4 and 5 dimensions. However, both methods estimated lower occupancy rates than expected, as indicated by the slope of the regression. The accuracy of the two algorithms in reconstructing occupancy rates decreased with increasing the number of dimensions. The accuracy was lower for the *subsample* than for the *box* method.

Both *subsample* and *box* methods were successful in detecting between-group differences in occupancy rates. After the permutation test, the observed volume was close to the expected volume (Table 3). However, the simulation exercise showed that the two methods, particularly the *box* method, tend to overestimate the volume of the union of input hypervolumes both when groups are specified or not. This problem increases with increasing the number of dimensions and affects in turn the occupancy rate estimates.

The same results obtained with the mean are expected for the sum because the matrix from which both statistics are calculated is the same. However, caution is needed when using unbalanced datasets where groups differ for the number of observations. In such cases, using sum as summary statistics could return results that are driven by a different sample size rather than by differences in functional diversity.

3.2 | Effects of flow intermittence on the body size of aquatic macroinvertebrates

The first two MDS axes explained 43.0% and 31.0% of the overall variation for a total of 74.0%. The mean volume and standard deviation of the n-dimensional hypervolumes as well as the volume of the union of input hypervolumes are reported in Table 4. Existing metrics were related to intermittency. The relationship between the volume of input hypervolumes and intermittence was significant ($F_{1,38}=10.6$, $p=0.002$). The correlation of intermittence with the mean volume of input hypervolumes and the volume of the union of input hypervolumes were −0.79 ($p=0.007$; Figure 3a) and 0.78 ($p=0.008$, Figure 3b), respectively.

The correlation of intermittence with the mean occupancy rate was −0.83 ($p=0.003$; Figure 3c), indicating that the extent of the

TABLE 2 Accuracy of the approaches for calculating occupancy rates, obtained from simulated data. Slope and R^2 of a linear regression between expected and observed values are reported for a number of dimensions ranging from 2 to 5 and for the subsample and the box methods. For the box method, results obtained at multiple box densities (500, 1000, 2500, 5000, 7500, 10,000) are shown.

No. of dimensions	Metric	Subsample	500	1000	2500	5000	7500	Box 10,000
2	R^2	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	R^2	0.98	0.96	0.97	0.97	0.97	0.97	0.97
4	R^2	0.82	0.94	0.93	0.93	0.94	0.94	0.93
5	R^2	0.48	0.88	0.87	0.87	0.88	0.87	0.87
2	Slope	0.82	0.85	0.86	0.85	0.85	0.85	0.85
3	Slope	0.67	0.73	0.75	0.74	0.74	0.73	0.74
4	Slope	0.43	0.57	0.58	0.57	0.58	0.58	0.57
5	Slope	0.26	0.43	0.42	0.42	0.42	0.42	0.41

TABLE 3 Comparison of the results obtained with the subsample and the box methods. Two disjoint groups of 10 hypervolumes generated from n-dimensional hyperspheres with an increasing number of dimensions (from 2 to 5) were compared (sub = subsample method; box = box method). For each dimension multiple statistics are provided: (1) mean, minimum and maximum volumes of input hypervolumes; (2) the volume of the union of all hypervolumes (all) and the volumes of union of hypervolumes of the two groups (group 1 and 2); (3) occupancy rates resulting from the union of all the hypervolumes and occupancy rates for the two group of hypervolumes and (4) the volume after permuting the labels of hypervolumes (not available without a grouping factor).

Measure	Group	2_D_sub	2_D_box	3_D_sub	3_D_box	4_D_sub	4_D_box	5_D_sub	5_D_box
Input volume	Mean	3.82		6.34		9.37		12.2	
	Min	3.75		6.11		9.02		11.9	
	Max	3.92		6.49		9.68		12.5	
Volume	All	8.36	8.49	13.6	14.7	20.4	23.9	27.7	35.7
	Group 1	4.12	4.23	6.80	7.34	10.2	12.1	13.9	17.8
	Group 2	4.24	4.27	6.82	7.34	10.2	11.9	13.9	18
Mean occupancy rate	All	0.46	0.45	0.47	0.43	0.46	0.39	0.44	0.34
	Group 1	0.92	0.90	0.93	0.86	0.92	0.78	0.88	0.68
	Group 2	0.91	0.90	0.93	0.87	0.92	0.78	0.88	0.68
Volume after permuting	All	–	–	–	–	–	–	–	–
	Group 1	3.97	4.00	6.58	6.49	9.87	9.77	13.2	12.5
	Group 2	4.08	4.01	6.63	6.52	9.9	9.57	13.2	12.6

TABLE 4 Results of different methods for estimating functional diversity, obtained from a trait space based on the body size of aquatic macroinvertebrates (mean volume = mean volume of the input hypervolumes; union volume = volume of the union of input hypervolumes). Four observations for each site were available ($n = 40$). Flow intermittence (period of time during the year with no or lack of flow) is also reported.

Site	Intermittence	Mean volume	Union volume	Mean occupancy rate
T1	0.00	1.80 ± 0.02	2.04	0.88
T2	0.53	1.77 ± 0.06	2.03	0.87
T3	5.26	1.80 ± 0.01	2.01	0.90
T4	1.78	1.78 ± 0.03	2.00	0.88
T5	16.5	1.76 ± 0.04	2.04	0.86
T6	20.8	1.79 ± 0.04	2.00	0.89
T7	28.8	1.62 ± 0.16	2.11	0.77
T8	32.9	1.68 ± 0.11	2.08	0.81
T9	32.9	1.53 ± 0.35	2.08	0.74
T10	36.1	1.67 ± 0.14	2.12	0.79

period without flow affects the body size of macroinvertebrates. Traits averaging had a negligible effect on the relationship of flow intermittence with the mean occupancy rate and the mean volume of input hypervolumes, while it affected its relationship with the volume of the union of input hypervolumes (Appendix S1).

Areas with low occupancy rates were on the left side of the MDS ordination, corresponding to moderate and large-bodied taxa (Figure 3). Taxa in areas with low occupancy rates were *Gammarus* (sites T8, T9, T10), Limnephilidae (site T8), Erpobdellidae and *Niphargus* (site T7). These taxa have their highest body size value in the ranges ≥1–2 cm (*Gammarus*, Limnephilidae) and ≥2–4 cm (Erpobdellidae, *Niphargus*). Again, trait averaging had a negligible effect on the results (Appendix S1).

3.3 | Effect of river morphology on aquatic macroinvertebrates

The first two MDS axes explained 61.0% and 27.0% of the overall variation for a total of 88.0%. The mean volume and standard deviation of n-dimensional hypervolumes were 0.65 ± 0.12, 0.70 ± 0.18, 0.75 ± 0.12, 0.72 ± 0.12, 1.0.68 ± 0.15 for isolated pool, pool, backwater, glide and riffle. The effect of mesohabitat on the mean volume was significant ($F_{4143.7} = 3.13$, $p = 0.019$). According to the post hoc test backwater had a significantly greater volume than the isolated pool ($p < 0.015$), although this result could be an artefact of traits averaging (Appendix S1). The volume of the union of input hypervolumes for each mesohabitat was similar (isolated pool = 1.37, pool = 1.38, backwater = 1.26, glide = 1.47 and riffle = 1.40). The mean occupancy rate was similar among mesohabitats, with isolated pool, pool, backwater, glide and riffle showing values of 0.47, 0.50, 0.59, 0.49 and 0.49. Patterns obtained by averaging traits were similar to those obtained by resampling traits (Appendix S1).

Although mean occupancy rates did not highlight differences among mesohabitats, patterns of occupancy rates within the trait space differed among mesohabitats (Figure 4). The permutation test captured this pattern and showed that some mesohabitats had significantly higher values than others according to their position in the trait space (Figure 5). The permutation test, thus, detected heterogeneity in trait space utilisation driven by river morphology even with a high degree of overlap in the input hypervolumes. The volume of the significant fraction was greater when comparing mesohabitats with different hydraulic characteristics (Table 5). Traits averaging had a negligible

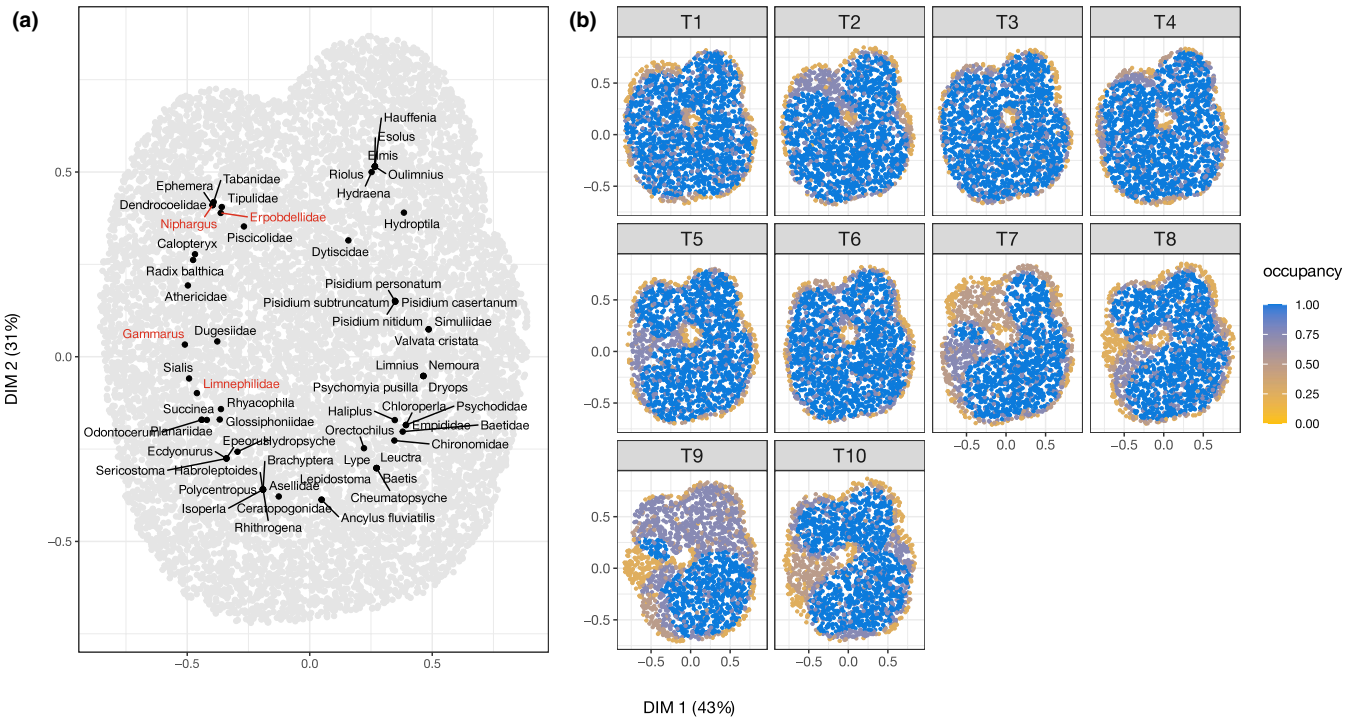


FIGURE 3 Results of the occupancy rate framework applied to 40 hypervolumes to test the effect of flow intermittency on the body size of aquatic macroinvertebrates. Panel (a) shows taxa coordinates obtained with a metric MDS. Taxa in red are those that lies in areas with low occupancy rates of sites T7, T8, T9 and T10. Panel (b) visualises the occupancy rate of each studied site. Occupancy rate can range from 0 to 1 and represent the mean number of hypervolumes enclosing a given random point (0 = no hypervolume within a site includes a random point, 1 = all the hypervolumes within a site include a given random point).

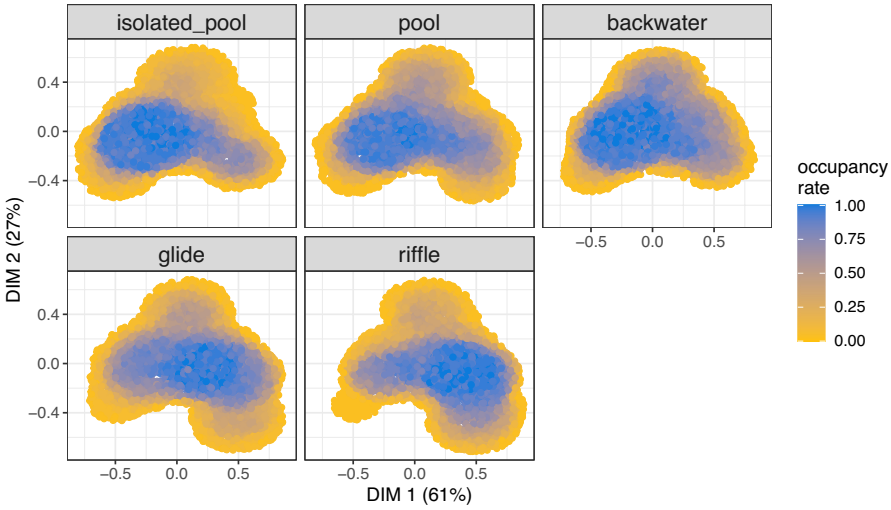


FIGURE 4 Occupancy rates of different mesohabitats calculated on a trait space based on the flow preferences of aquatic macroinvertebrates. Colours are proportional to the number of hypervolumes enclosing a random point. For each mesohabitat, the resulting hypervolume has been calculated from 30 hypervolumes (for a total of 150 hypervolumes). Values range from 0 (no hypervolume enclosing the random point) to 1 (all the hypervolumes of a given mesohabitat encloses the random point).

effect on these results, indicating a strong effect of mesohabitat on flow velocity preferences of macroinvertebrates (Appendix S1). The only exception was for the comparison pool-backwater, which showed a great variability with the resampling approach.

4 | DISCUSSION

N-dimensional hypervolumes represent a flexible and appealing approach for quantifying functional diversity. When analysing multiple overlapping hypervolumes, current methods show limitations because they do not consider heterogeneity in trait space utilisation emerging from the overlap effectively. To overcome these limitations, we developed a new approach based on the occupancy rate concept, the mean or absolute number of hypervolumes that occupy a given region of a multidimensional space. With two examples based on aquatic macroinvertebrates, we showed how occupancy rates relate to environmental gradients and to among-group differences in trait space utilisation that in turn can be associated to relevant ecological processes such as environmental filtering and dispersal.

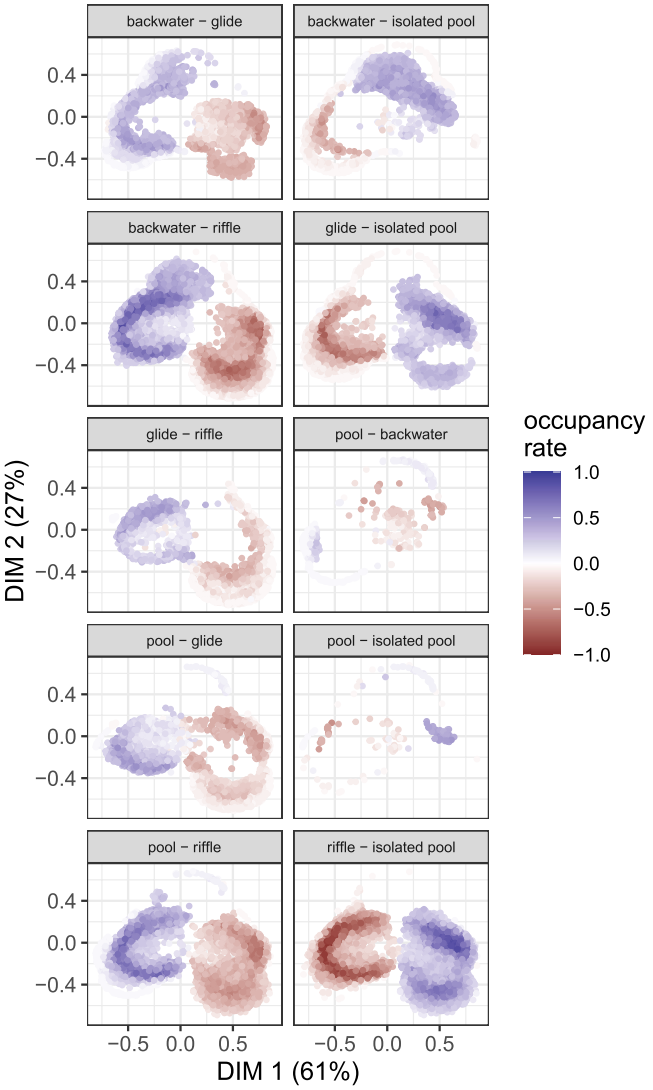


FIGURE 5 Pairwise difference among mesohabitats calculated using the permutation test on occupancy rates performed on a trait space based on the flow preferences of aquatic macroinvertebrates. Differences in occupancy rates were obtained by calculating the difference between the occupancy rates of one hypervolume (e.g. backwater) and those of a second hypervolume (e.g. glide) for each random point. Values can range from 1 to −1, where positive values mean that the occupancy rates of the first hypervolume are higher than those of the second hypervolume while negative values mean the opposite. Colour intensity is proportional to the absolute value of the differences (colours are more intense for 1 and −1). Only significant differences were retained according to the permutation test.

The newly developed approach showed good performance in detecting expected patterns. This is especially true for the *box* method, which showed greater accuracy than the *subsample* method in most simulations. Drawbacks of the *box* method are increased computation times, occupancy rate estimates dependent on the box density and overestimation of volumes in higher dimensions. In our simulations, the dependence of occupancy rate estimates to box density was not an issue, but in real case examples, we suggest testing more box densities to look for the best compromise between running time and accuracy. Overestimation is due to more random points detected as unique to one or the other hypervolumes compared to the *subsample* method probably because points become sparser in the multidimensional space at higher dimensions. The *subsample* method showed poorer performances than the *box* method, likely because it returns a slightly non-uniform subsample of random points when the point density is very heterogeneous in the multidimensional space. However, this method could be a valuable alternative to the *box* method that can be used for exploratory analyses in high dimensions because of the decreased computation time. Some improvements to the basic *subsample* algorithm (not tested in the present work) are proposed in the hypervolume package (e.g. find_optimal_occupancy_thin()). However, these solutions slow down computation times in higher dimensions and may have limited practical application in such cases.

TABLE 5 Volume of the significant fraction of a trait space calculated on the flow preferences of aquatic macroinvertebrates for each pairwise comparison of mesohabitats (i_pool = isolated pools). The significant fraction represents the volume significantly occupied by a first (hv_1) or a second (hv_2) group of hypervolumes and is calculated using a permutation test on occupancy rates. Total volume is calculated as the sum of the results of hv_1 and hv_2.

hv_1	hv_2	hv_1 volume	hv_2 volume	Total volume
Pool	Backwater	0.07	0.04	0.11
Pool	Glide	0.22	0.34	0.56
Pool	Riffle	0.38	0.39	0.77
Pool	i_ pool	0.02	0.01	0.03
Backwater	Glide	0.28	0.22	0.50
Backwater	Riffle	0.49	0.47	0.96
Backwater	i_ pool	0.30	0.14	0.44
Glide	Riffle	0.18	0.29	0.47
Glide	i_pool	0.32	0.31	0.63
Riffle	i_pool	0.45	0.36	0.81

In the first example, size-based metrics and mean occupancy rates calculated on a trait space based on the body size of aquatic macroinvertebrates were related to flow intermittence. These results confirm our hypothesis of a decrease in the stability of trait space utilisation with increasing intermittence values. The lack of surface water acts as an environmental filter that favours species with small body size, fast life cycles and more cycles per year, while filters out long-lived species with moderate to large body size (Arias-Real et al., 2022; Piano et al., 2020). Species with large body size, thus, disappear or become rare at the most intermittent sites, causing the occupancy rate to drop in certain regions of the trait space. The decrease in the occupancy rate results from two distinct processes, the first being the paucity of species with similar body size while the second being the low detection probability of some species as measured from multiple measures at the same site. The paucity of species with similar body size is related to functional redundancy

(species with the same or similar traits) within a single community (Carmona et al., 2016). The detection probability is linked to both the effect of imperfect detection and the effect of environmental variables (MacKenzie et al., 2002). A site with a species with low functional redundancy and low detection probability is, thus, more prone to have low occupancy rates, at least in some regions of the trait space. The added value of the new approach, emerging from this first case study, is as an (i) easy-to-understand metric for evaluating the variation in the overlap of multiple hypervolumes in the trait space and (ii) the identification of species with low functional redundancy and detectability when multiple observations are available.

The second example dealing with flow preferences of macroinvertebrate communities also demonstrated the added value of our approach. Occupancy rates were helpful in showing that the investigated mesohabitats host communities with different flow preferences. Fast-flowing mesohabitats showed significantly higher occupancy rates in the right part of the trait space while the slow-flowing ones preferentially occupied the left side. Similarly to the example on intermittence, flow velocity can act as a filter that affects the structure of macroinvertebrate communities. For example, organisms adapted to high flow velocity are disfavoured in slow-flowing mesohabitats. Although differences in flow preferences between fast and slow-flowing habitats are expected (Extence et al., 1999), we found differences in community preferences even within mesohabitat with similar hydraulic characteristics. This could be due to the drift of organisms from the fast-flowing mesohabitats to the slow-flowing ones when mesohabitats are connected within the main channel (e.g. from riffle to backwater) but not when they are isolated from one another (e.g. isolated pool). Proximity to fast-flowing mesohabitats can induce mass effect (Heino et al., 2017; Shmida & Wilson, 1985), whose detection can be masked when using binary outcomes (a random point is included or not in a given group of hypervolumes) instead of occupancy rates (that take into account the number of hypervolumes that enclose a given random point) on aggregated hypervolumes. Our approach can, thus, detect significant differences in trait space utilisation even with a high degree of overlap of input hypervolumes.

Besides the benefits of evaluating ecological patterns, our approach presents some technical advantages. Although for sake of pedagogy both examples involved hypervolumes based on a two-dimensional trait space, our approach extends to n-dimensional hypervolumes. Therefore, occupancy rates can be calculated on the trait space dimensionality that best suits specific case studies (e.g. after evaluating the functional space quality). Moreover, our approach allows less noisy comparisons of multiple hypervolumes than existing methods, thus avoiding multiplicative errors resulting from applying a pairwise approach multiple times.

5 | CONCLUSIONS

Existing methods for quantifying functional diversity are limited because they do not exploit the information provided by the overlap of multiple hypervolumes effectively. Our framework complements these methods by providing a powerful tool to detect differences in trait space occupancy rate. This is especially useful when a high degree of overlap between groups of hypervolumes can mask underlying ecological processes, for example, when studying the effect of invasive species on native communities or when analysing community assembly processes in different habitats.

Our approach can be technically and conceptually extended further. Although being developed to quantify among group differences in trait space occupancy rates, our approach can be modified to consider continuous variables. Furthermore, mean occupancy rates can be considered as a measure of within-group redundancy (Carmona et al., 2019; de Bello et al., 2007) and can be coupled with null model approaches for uncertainty estimation (Chen et al., 2023; Mammola et al., 2021). We hope that our newly developed framework will contribute to test new hypotheses in functional ecology and to foster new developments in this expanding field of ecology.

AUTHOR CONTRIBUTIONS

Alex Laini conceived the idea, designed the methodology, analysed the data and led the manuscript writing; Thibault Datry contributed to writing and data processing; Benjamin Wong Blonder guided on methodology and contributed to writing.

ACKNOWLEDGEMENTS

Alex Laini was partially supported by the Autorità di Bacino Distrettuale del Fiume Po, within the project 'Habitat modelling at mesoscale for environmental flow evaluation in the Trebbia, Taro and Enza rivers', with funding from Regione Emilia-Romagna and ANBI Emilia-Romagna. We thank Simone Guareschi and Tommaso Cancellario for their insightful comments on an earlier draft of this paper and Gemma Burgazzi for her help in the field.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to report.

DATA AVAILABILITY STATEMENT

Functions for calculating occupancy rates and for performing the permutation test are available in the hypervolume package version 3.1.0 either on CRAN (http://cran.r-project.org/web/packages/hypervolume/index.html) or GitHub (http://github.com/bblonder/hypervolume). Code for replicating all the simulations and examples is archived in the Zenodo Digital Repository https://doi.org/10.5281/zenodo.7823672. Example 1 data can be requested through the Intermittent River Biodiversity Analysis and Synthesis (IRBAS) website (https://irbas.inrae.fr/irbas/irbas). Data for the second example are archived in the Dryad Digital Repository https://doi.org/10.5061/dryad.cjsxksnb9 (Laini et al., 2023).

ORCID

Alex Laini https://orcid.org/0000-0002-3458-7538
Thibault Datry https://orcid.org/0000-0003-1390-6736
Benjamin Wong Blonder https://orcid.org/0000-0002-5061-2385

REFERENCES

Arias-Real, R., Gutiérrez-Cánovas, C., Menéndez, M., & Munoz, I. (2022). Drying niches of aquatic macroinvertebrates identify potential biomonitoring indicators in intermittent and ephemeral streams. *Ecological Indicators*, *142*, 109263.

Bêche, L. A., Mcelravy, E. P., & Resh, V. H. (2006). Long-term seasonal variation in the biological traits of benthic-macroinvertebrates in two Mediterranean-climate streams in California, USA. *Freshwater Biology*, *51*(1), 56–75.

Belletti, B., Rinaldi, M., Bussettini, M., Comiti, F., Gurnell, A. M., Mao, L., Nardi, L., & Vezza, P. (2017). Characterising physical habitats and fluvial hydromorphology: A new system for the survey and classification of river geomorphic units. *Geomorphology*, *283*, 143–157. https://doi.org/10.1016/j.geomorph.2017.01.032

Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, *23*(5), 595–609. https://doi.org/10.1111/geb.12146

Blonder, B., Morrow, C. B., Maitner, B., Harris, D. J., Lamanna, C., Violle, C., Enquist, B. J., & Kerkhoff, A. J. (2018). New approaches for delineating n-dimensional hypervolumes. *Methods in Ecology and Evolution*, *9*(2), 305–319. https://doi.org/10.1111/2041-210X.12865

Bo, T., Doretto, A., Laini, A., Bona, F., & Fenoglio, S. (2017). Biomonitoring with macroinvertebrate communities in Italy: What happened to our past and what is the future? *Journal of Limnology*, *76*, 21–28. https://doi.org/10.4081/jlimnol.2016.1584

Bonada, N., Rieradevall, M., & Prat, N. (2007). Macroinvertebrate community structure and biological traits related to flow permanence in a Mediterranean river network. *Hydrobiologia*, *589*(1), 91–106. https://doi.org/10.1007/s10750-007-0723-5

Brown, M. J. M., Holland, B. R., & Jordan, G. J. (2020). Hyperoverlap: Detecting biological overlap in n-dimensional space. *Methods in Ecology and Evolution*, *11*(4), 513–523. https://doi.org/10.1111/2041-210X.13363

Carmona, C. P., de Bello, F., Mason, N. W. H., & Lepš, J. (2016). Traits without Borders: Integrating functional diversity across scales. *Trends in Ecology & Evolution*, *31*(5), 382–394. https://doi.org/10.1016/j.tree.2016.02.003

Carmona, C. P., de Bello, F., Mason, N. W. H., & Lepš, J. (2019). Trait probability density (TPD): Measuring functional diversity across scales based on TPD with R. *Ecology*, *100*(12), e02876. https://doi.org/10.1002/ecy.2876

Carmona, C. P., Guerrero, I., Morales, M. B., Oñate, J. J., & Peco, B. (2017). Assessing vulnerability of functional diversity to species loss: A case study in Mediterranean agricultural systems. *Functional Ecology*, *31*(2), 427–435. https://doi.org/10.1111/1365-2435.12709

Chen, D., Laini, A., & Blonder, B. (2023). *Statistical inference methods for n-dimensional hypervolumes: Applications to niches and functional diversity*.

Chevene, F., Doleadec, S., & Chessel, D. (1994). A fuzzy coding approach for the analysis of long-term ecological data. *Freshwater Biology*, *31*(3), 295–309.

Cornwell, W. K., Schwilk, D. W., & Ackerly, D. D. (2006). A trait-based test for habitat filtering: Convex hull volume. *Ecology*, *87*(6), 1465–1471. https://doi.org/10.1890/0012-9658(2006)87[1465:ATTFHF]2.0.CO;2

Datry, T. (2012). Benthic and hyporheic invertebrate assemblages along a flow intermittence gradient: Effects of duration of dry events. *Freshwater Biology*, *57*(3), 563–574. https://doi.org/10.1111/j.1365-2427.2011.02725.x

de Bello, F., Lepš, J., Lavorel, S., & Moretti, M. (2007). Importance of species abundance for assessment of trait composition: An example based on pollinator communities. *Community Ecology*, *8*(2), 163–170. https://doi.org/10.1556/ComEc.8.2007.2.3

Extence, C. A., Balbi, D. M., & Chadd, R. P. (1999). River flow indexing using British benthic macroinvertebrates: A framework for setting hydroecological objectives. *Regulated Rivers: Research & Management*, *15*(6), 545–574. https://doi.org/10.1002/(SICI)1099-1646(199911/12)15:6<545::AID-RRR561>3.0.CO;2-W

Gagic, V., Bartomeus, I., Jonsson, T., Taylor, A., Winqvist, C., Fischer, C., Slade, E. M., Steffan-Dewenter, I., Emmerson, M., Potts, S. G., Tscharntke, T., Weisser, W., & Bommarco, R. (2015). Functional identity and diversity of animals predict ecosystem functioning better than species-based indices. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801), 20142620. https://doi.org/10.1098/rspb.2014.2620

Guillerme, T., Puttick, M. N., Marcy, A. E., & Weisbecker, V. (2020). Shifting spaces: Which disparity or dissimilarity measurement best summarize occupancy in multidimensional spaces? *Ecology and Evolution*, *10*(14), 7261–7275. https://doi.org/10.1002/ece3.6452

Gutiérrez-Cánovas, C., Sánchez-Fernández, D., Velasco, J., Millán, A., & Bonada, N. (2015). Similarity in the difference: Changes in community functional features along natural and anthropogenic stress gradients. *Ecology*, *96*(9), 2458–2466. https://doi.org/10.1890/14-1447.1

Hatfield, J. H., Harrison, M. L. K., & Banks-Leite, C. (2018). Functional diversity metrics: How they are affected by landscape change and how they represent ecosystem functioning in the tropics. *Current Landscape Ecology Reports*, *3*(2), 35–42. https://doi.org/10.1007/s40823-018-0032-x

Heino, J., Alahuhta, J., Ala-Hulkko, T., Antikainen, H., Bini, L. M., Bonada, N., Datry, T., Erős, T., Hjort, J., Kotavaara, O., Melo, A. S., & Soininen, J. (2017). Integrating dispersal proxies in ecological and environmental research in the freshwater realm. *Environmental Reviews*, *25*(3), 334–349. https://doi.org/10.1139/er-2016-0110

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Laini, A., Burgazzi, G., Chadd, R., England, J., Tziortzis, I., Ventrucci, M., Vezza, P., Wood, P. J., Viaroli, P., & Guareschi, S. (2022). Using invertebrate functional traits to improve flow variability assessment within European rivers. *Science of the Total Environment*, *832*, 155047. https://doi.org/10.1016/j.scitotenv.2022.155047

Laini, A., Datry, T., & Blonder, B. (2023). *Data from: N-dimensional hypervolumes in trait-based ecology: Does occupancy rate matter?* https://doi.org/10.5061/dryad.cjsxksnb9

Laini, A., Guareschi, S., Bolpagni, R., Burgazzi, G., Bruno, D., Gutiérrez-Cánovas, C., Miranda, R., Mondy, C., Várbíró, G., & Cancellario, T. (2022). biomonitoR: An R package for managing ecological data and calculating biomonitoring indices. *PeerJ*, *10*, e14183. https://doi.org/10.7717/peerj.14183

Laurent, S. (2022). *uniformly: Uniform sampling*. R package Version 0.2.0. https://CRAN.R-project.org/package=uniformly

Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means*. R package Version 1.7.2. https://CRAN.R-project.org/package=emmeans

Loranger, J., Blonder, B., Garnier, É., Shipley, B., Vile, D., & Violle, C. (2016). Occupancy and overlap in trait space along a successional gradient in Mediterranean old fields. *American Journal of Botany*, *103*(6), 1050–1060. https://doi.org/10.3732/ajb.1500483

Lucas, J. M., Jonas, J., Laws, A. N., Branson, D. H., Pennings, S. C., Prather, C. M., & Strickland, M. S. (2021). Functional and taxonomic diversity of grasshoppers differentially shape above- and below-ground communities and their function. *Functional Ecology*, *35*(1), 167–180. https://doi.org/10.1111/1365-2435.13682

Maas, B., Brandl, M., Hussain, R. I., Frank, T., Zulka, K. P., Rabl, D., Walcher, R., & Moser, D. (2021). Functional traits driving pollinator and predator responses to newly established grassland strips in agricultural landscapes. *Journal of Applied Ecology*, *58*(8), 1728–1737. https://doi.org/10.1111/1365-2664.13892

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, *83*(8), 2248–2255.

Mammola, S., & Cardoso, P. (2020). Functional diversity metrics using kernel density n-dimensional hypervolumes. *Methods in Ecology and Evolution*, *11*(8), 986–995. https://doi.org/10.1111/2041-210X.13424

Mammola, S., Carmona, C. P., Guillerme, T., & Cardoso, P. (2021). Concepts and applications in functional diversity. *Functional Ecology*, *35*(9), 1869–1885. https://doi.org/10.1111/1365-2435.13882

Mason, N. W. H., Mouillot, D., Lee, W. G., & Wilson, J. B. (2005). Functional richness, functional evenness and functional divergence: The primary components of functional diversity. *Oikos*, *111*(1), 112–118. https://doi.org/10.1111/j.0030-1299.2005.13886.x

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. *Functional Ecology*, *24*(4), 867–876. https://doi.org/10.1111/j.1365-2435.2010.01695.x

Pavoine, S., Vallet, J., Dufour, A.-B., Gachet, S., & Daniel, H. (2009). On the challenge of treating various types of variables: Application for improving the measurement of functional diversity. *Oikos*, *118*(3), 391–402. https://doi.org/10.1111/j.1600-0706.2008.16668.x

Piano, E., Doretto, A., Mammola, S., Falasco, E., Fenoglio, S., & Bona, F. (2020). Taxonomic and functional homogenisation of macroinvertebrate communities in recently intermittent alpine watercourses. *Freshwater Biology*, *65*(12), 2096–2107.

Potanin, B. (2022). *gena: Genetic algorithm and particle swarm optimization*. R package Version 1.0.0. https://CRAN.R-project.org/package=gena

R Core Team. (2021). *A language and environment for statistical computing*. R Foundation for Statistical Computing. http://CRAN.R-project.org/

Sarremejane, R., Cid, N., Stubbington, R., Datry, T., Alp, M., Cañedo-Argüelles, M., Cordero-Rivera, A., Csabai, Z., Gutiérrez-Cánovas, C., Heino, J., Forcellini, M., Millán, A., Paillex, A., Pařil, P., Polášek, M., Tierno de Figueroa, J. M., Usseglio-Polatera, P., Zamora-Muñoz, C., & Bonada, N. (2020). DISPERSE, a trait database to assess the dispersal potential of European aquatic macroinvertebrates. *Scientific Data*, *7*(1), 386. https://doi.org/10.1038/s41597-020-00732-7

Shmida, A., & Wilson, M. V. (1985). Biological determinants of species diversity. *Journal of Biogeography*, *12*(1), 1–20. https://doi.org/10.2307/2845026

Tachet, H., Richoux, P., Bournaud, M., & Usseglio-Polatera, P. (2010). *Invertébrés d'eau douce: Systématique, biologie, écologie (édition revue et augmentée edizione)*. CNRS Editions.

Takács, P., Abonyi, A., Bánó, B., & Erős, T. (2021). Effect of non-native species on taxonomic and functional diversity of fish communities in different river types. *Biodiversity and Conservation*, *30*(8), 2511–2528. https://doi.org/10.1007/s10531-021-02207-6

Tilman, D., Isbell, F., & Cowles, J. M. (2014). Biodiversity and ecosystem functioning. *Annual Review of Ecology, Evolution, and Systematics*, *45*(1), 471–493. https://doi.org/10.1146/annurev-ecolsys-120213-091917

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

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://www.springer.com/us/book/9780387981413

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. Additional figures and tables.

How to cite this article: Laini, A., Datry, T., & Blonder, B. W. (2023). N-dimensional hypervolumes in trait-based ecology: Does occupancy rate matter? *Functional Ecology*, *00*, 1–13. https://doi.org/10.1111/1365-2435.14344