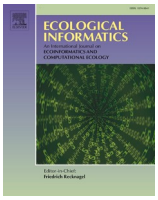
Ecological Informatics 72 (2022) 101894

Contents lists available at ScienceDirect

Ecological Informatics

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



Comparing distance-decay parameters: A novel test under pairwise dependence

Ramiro Martín-Devasa [a,*], Sara Martínez-Santalla [a], Carola Gómez-Rodríguez [b], Rosa M. Crujeiras [c], Andrés Baselga [a]

[a] CRETUS, Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain
[b] CRETUS, Department of Functional Biology (Area of Ecology), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
[c] CITMAga, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

ARTICLE INFO

Keywords:
β-Diversity
Distance-decay
Iberian vertebrates
Pairwise dependence
Parameter comparison
Site-block resampling

ABSTRACT

Distance-decay models fit parametric functions to assess the relationship between similarity and spatial or environmental distance. Despite the widespread use of distance-decay models in ecology and biogeography, no method has been previously developed and validated to assess the significance of differences between the parameters (i.e. intercept and slope) of two distance-decay models. The pairwise autocorrelation of similarity and spatial distance affects the variance of parameter estimates, precluding the use of ordinary *t*-tests. Here, we provide a test statistic (z_{dep}) for the equality of parameters between two distance-decay models. The test can be applied, independently, to any of the model parameters (i.e. intercept and/or slopes). The z_{dep} statistic accounts for pairwise dependence, thus avoiding biases associated to the inflation of degrees of freedom, and it is based on the estimation of parameters' variance using site-block resampling. To validate the z_{dep} statistic, type I and type II errors were empirically evaluated through a simulation study. We simulated six scenarios (three under the null and three under the alternative hypothesis) of distance-decay relationships using different functions: negative exponential, power-law or Gompertz function. We applied the z_{dep} statistic and computed the proportion of rejections of the null hypothesis for $\alpha = 0.01$, 0.05 and 0.1 in each scenario. As a case-study, we also compared distance-decay parameters across several groups of Iberian vertebrates (cyprinids, frogs, lizards and snakes, bats, rodents and carnivores). In the simulation study, the z_{dep} statistic showed a good approximation of the nominal level (α, type I error) and a good statistical power (1 − type II error), the later increasing with sample size, as expected. In Iberian vertebrates, we found significant differences between ectotherms and endotherms, but not within these groups except between cyprinids and other ectotherms. The good performance of the z_{dep} statistic makes it the best option to test for differences in parameters obtained from models fitted from data with pairwise dependence, as distance-decay models. It can also be used beyond distance-decay approaches to compare parameters of any other regression models of pairwise dependent data (as genetic distances, for example).

1. Introduction

The distance-decay of similarity (i.e. the decrease of community similarity with spatial/environmental distance) is a ubiquitous macroecological pattern (Graco-Roza et al., 2022; Nekola and White, 1999; Soininen et al., 2007) that results from the interplay between environmental attributes of the study area and the organisms' dispersal ability and ecological requirements (Morlon et al., 2008; Steinbauer et al., 2012). Distance-decay models are a common approach to analyse the spatial variation of community composition (β-diversity), by assessing the relationship between two pairwise matrices: community similarity and spatial distance. In fact, the shape and slope of such distance-decay relationship can be used to infer the strength of the processes driving the spatial structure of biodiversity patterns, such as environmental filtering and dispersal limitation (Baselga and Gómez-Rodríguez, 2021; Gómez-Rodríguez and Baselga, 2018; Soininen et al., 2007). Moreover, a comparison of distance-decay model parameters (i.e. intercept and slope) across biological groups has proven effective to evidence differences in the role of dispersal limitation within the same geographical context (e.g., Gómez-Rodríguez and Baselga, 2018). However, such comparison of

* Corresponding author.
E-mail address: ramiromaria.martin@usc.es (R. Martín-Devasa).

https://doi.org/10.1016/j.ecoinf.2022.101894
Received 21 April 2022; Received in revised form 29 October 2022; Accepted 30 October 2022
Available online 11 November 2022
1574-9541/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

distance-decay parameters is not straightforward because the variance of parameter estimates is affected by the fact that both variables, i.e. community similarity and spatial distance, are pairwise autocorrelated (i.e. similarity and distance values are the result of the comparison of two biological communities), and hence their values cannot be considered independent (Smouse et al., 1986). While the bias introduced by pairwise autocorrelation has been accounted for in some significance tests (e.g., Mantel test), a statistical test for parameter comparison, equivalent to the *t*-test, has not been developed in a multivariate context.

Various methods have been developed to study the relationship between community similarity and spatial or environmental distances, including semi-variograms (Cressie, 1993), linear regressions (Nekola and White, 1999), or generalised dissimilarity models based on spline regressions (Ferrier et al., 2002, 2007). However, the most widely used methods are (i) the Mantel and partial Mantel tests (Mantel, 1967; Mantel and Valand, 1970; Smouse et al., 1986), which compute the correlation between two distance (or dissimilarity) matrices, and (ii) non-linear regressions, which fit parametric functions, such as the power-law or the negative exponential, to distance-decay data (Nekola and McGill, 2014). Mantel tests have been specifically developed to assess the correlation between distance and/or dissimilarity matrices, and hence are not biased by the structure of distance data. However, its lack of parameters precludes the ecological interpretation of the relationship. On the contrary, in regression-based distance-decay models, the intercept can be interpreted as the expected community similarity at short distances while the slope is the rate at which communities change with distance (Soininen et al., 2007). Therefore, to compare the parameters of two distance-decay models, we need a statistical test that accounts for the pairwise autocorrelation of distance-decay data.

Pairwise autocorrelation arises because both community similarity and spatial or environmental distances are computed by comparing the data observed in two different sites. Thus, each value in a similarity/distance matrix involves two observations (i.e. sites) and each of these observations participates in the computation of more than one of the similarity/distance values (Diserud and Ødegaard, 2007; Smouse et al., 1986). In other words, from N sites we compute (N*N-1)/2 pairwise similarities, so the number of similarity values is larger than the number of independent community observations. This causes the inflation of degrees of freedom (Koenig, 1999) because one value in the similarity matrix does not represent a complete degree of freedom (Dale and Fortin, 2002; Legendre, 1993) (Fig. 1). The inflation of degrees of freedom biases ordinary significance tests, as the F-test, but this bias has been accounted for in tests specifically developed for similarity/distance data. For example, the Mantel test uses permutations to compute the null distribution of the correlation between similarity and distance, thus taking into account pairwise dependence (Koenig, 1999). Beyond correlation, a significance test based on the deviance of non-linear regression models has been recently developed, which uses permutations or site-block resampling to deal with pairwise autocorrelation (Martínez-Santalla et al., 2022). Moreover, and in order to compare the parameters of two distance-decay models, different alternatives have been also used based on permutations or bootstrap (Gómez-Rodríguez and Baselga, 2018; Nekola and White, 1999). However, previous solutions were suboptimal because the resampling procedures did not account for the pairwise dependence of the data. From a practical standpoint, the inflation of degrees of freedom biases the comparison of regression parameters by increasing the rate of false positives (e.g., reporting significant differences in distance-decay parameters when they are not different, inflating Type I error) (Legendre, 1993; Legendre and Fortin, 1989). To overcome this problem, we have designed a specific test to compare the parameters of distance-decay models while controlling for pairwise autocorrelation.

Here we introduce and evaluate a novel statistic to independently test for the equality of individual parameters of distance-decay models while accounting for the pairwise dependence of the data. This test

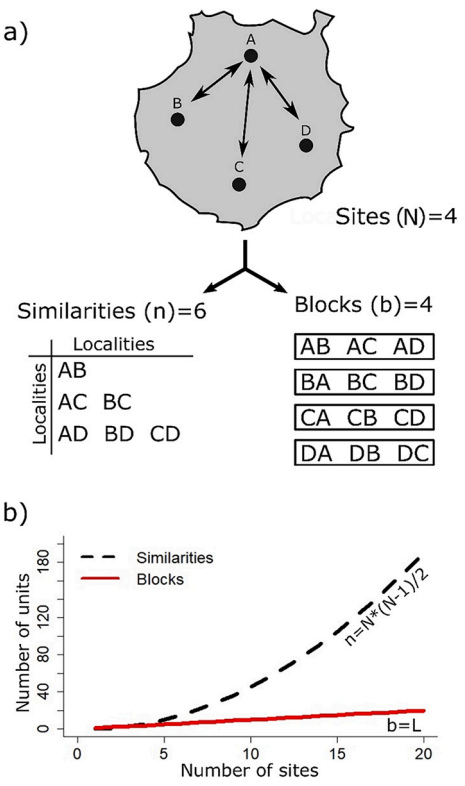


Fig. 1. a) Example showing how pairwise dependence of similarity/distance values results in pseudoreplication and hence the inflation of degrees of freedom. The effective size of the data ($N = 4$) is less than the number of values in the similarity/distance matrix ($n = 6$). Alternatively, in the site-block construction, the effective size of the data (N = 4) and the number or similarity site-blocks (b = 4) are the same, thus avoiding the inflation of degrees of freedom, while keeping the pairwise structure of data. b) Increase with sample size of the disparity between the number of sites or site-blocks and the number of similarities/distance values.

statistic (z_{dep}) is based on the standardised difference of the estimated parameters, analogous to a *t*-test in the sense that the difference between the estimated parameters is divided by the corresponding standard deviation,

$$z_{dep} = \frac{\widehat{\theta_1} - \widehat{\theta_2}}{\sqrt{var(\widehat{\theta}_1) + var(\widehat{\theta}_2) - cov(\widehat{\theta}_1, \widehat{\theta}_2)}} \quad (1)$$

where $\theta_1$and θ_2 are the parameters to compare and $\widehat{\theta_1}$ and $\widehat{\theta_2}$ the corresponding sample estimates. Note that the test is applied to each parameter independently, so a test can be run to assess the difference between intercepts of two distance-decay models (θ_{a1} vs. θ_{a2}), and a different test to assess the difference between their slopes (θ_{b1} vs. θ_{b2}). Also note that the test statistic in (1) is based on the asymptotically normal distribution of each parameter estimate, so z_{dep} is a standardization of the estimates difference. In addition, in contrast to an ordinary t-test, the variance in parameter estimation is computed using block-bootstrap techniques (Kunsch, 1989; Liu et al., 1992) to avoid bias due to the pairwise dependence of similarity indices and distance measures. Block-bootstrap techniques are based on defining the sampling unit as a set of values that present a certain degree of correlation, what is called a block, to mimic the dependence structure of the original data (Lahiri, 1999). In our case of pairwise dependent data, we define a block as the set of values obtained from the comparison of one site with all the others, thus capturing their pairwise dependence structure (Martínez-Santalla et al., 2022). In this case, from N sites N blocks are computed (Fig. 1) avoiding the inflation of degrees of freedom, as the

number of blocks is the same as the number of sites. If we consider the block as the sampling unit for the bootstrap procedure, we ensure that the effective size of the data (N) is equal to the number of observations (N) and, therefore, we avoid inflation-driven biases in the estimation of the variance of parameters.

The site-block resampling procedure we used to estimate the variance of distance-decay model parameters follows the same approach as used by Martínez-Santalla et al. (2022) for a significance test of distance-decay models. In short, resamples of site-blocks (i.e. all similarities or distances involving a given site) are taken with replacement and saved to a similarity block matrix. First, for N sites, N site-blocks are defined by selecting the similarity and distance values involving a given site. These N blocks are saved as columns of two block matrices, one for similarities and another for distances. Second, N block resamples (columns of block matrices) are drawn with replacement, ensuring that the same blocks are sampled for both similarities and distances, and then saved as two resample matrices. Third, resample matrices are initially built with duplicate observations (N * [N−1] values) because each block includes N - 1 values. In this step, the resample matrices are downsized to the original size of the similarity and distance matrices (N * [N - 1] / 2 values), by sampling N * [N - 1] / 2 values from both resample matrices. These values are saved as similarity and distance resample vectors, that will be used to fit the nonlinear distance-decay models and to estimate the model parameters ($\theta_1$and θ_2). Fourth, steps 1 to 3 are repeated B times, and B parameter estimates are saved in two vectors, one for θ_1, and another θ_2. Fifth, the variance of each parameter, and their covariance, are computed from the vectors of B parameter values.

To evaluate the performance of this approach in the estimation of the parameters' variance and hence the adequacy of the z_{dep} statistic, we conducted a calibration study based on simulations to assess whether the z_{dep} statistic is effective in independently detecting differences in individual distance-decay parameters (i.e. either for the intercepts or for the slopes, or for both). This was done by assessing the proportion of false positives (type I error, i.e. significant *p*-value when distance-decay curves were simulated with the same dataset, that is, under the null hypothesis of no differences between distance-decay model parameters) and the proportion of false negatives (type II error; i.e. non-significant p-value when simulated distance-decay curves parameters were different). In order to assess if there is a significant difference between two parameters (for example, between two intercepts or between two slopes), and therefore, to conclude if the null hypothesis (e.g. both distance-decay slopes are identical) should be rejected or not, it is necessary to estimate the distribution of the z_{dep} statistic under the null hypothesis, and to obtain the associated p-value. In our case, as the z_{dep} statistic is the standardised difference of parameters, under the null hypothesis, in which $\theta_1$and θ_2 are assumed to be equal, the statistic z_{dep} follows a N(0,1) distribution. This is because the distribution of the estimators is asymptotically normal, and the linear combination of normal distributions is also normal. In this case, because the z_{dep} statistic is a standardization, it follows a N(0,1).

A non-biased comparison of distance-decay parameters, such as the one provided by the novel z_{dep} test, can be widely applied in biogeographical, ecological and macroecological studies, as evidenced by the variety of studies that have previously sought for analogous analyses. For example, a comparison of distance-decay parameters has been used to assess the differences in distance-decay rates between ecosystems and across scales in fungi (Bahram et al., 2013), between different hosts in parasite communities (Thieltges et al., 2009), between different types of organisms like macro and microinvertebrates (Astorga et al., 2012) or ectotherms and endotherms (Qian and Ricklefs, 2012), between native and non-native species of urban floras (La Sorte et al., 2008), to assess the effects of eutrophication in diatom communities (Goldenberg Vilar et al., 2014) or to compare how grain size and the study extent affects the form of the distance-decay curves (Steinbauer et al., 2012). Moreover, beta diversity and distance-decay studies are informative for biodiversity conservation strategies and environment management (Gossner et al., 2016; Graco-Roza et al., 2022).

In this paper, we introduce the novel z_{dep} statistic to test for the equality of individual parameters between two distance-decay models and exemplify its use with distribution data of vertebrates in the Iberian Peninsula. The main novelty of this approach is the incorporation of an effective method (i.e. site-block resampling) to avoid the inflation of degrees of freedom in a *t*-test like statistic, and the performance of a calibration study. First, we performed a simulation study to analyse the empirical behaviour of the z_{dep} statistic, evaluating its type I and type II errors. For this purpose, we have simulated three scenarios where distance-decay parameters were equal (null hypothesis holds) and three scenarios where the distance-decay parameters were different (alternative hypotheses). This design allows studying the performance of the z_{dep} test with different distance-decay functional forms (i.e. negative exponential, power-law and Gompertz) and in different situations (similar intercept but different slope, different intercept but similar slope, different intercept and slope). Negative exponential and power-law models have been frequently used for distance-decay analyses (Nekola and McGill, 2014), and recently the Gompertz model has been introduced as alternative for sigmoidal distance-decay patterns (Martín-Devasa et al., 2022). Finally, we provide a case-study, applying the z_{dep} statistic to compare distance-decay parameters across several Iberian vertebrate groups (freshwater fishes, frogs, lizards and snakes, bats, rodents and carnivores).

2. Material and methods

2.1. Simulation of virtual biological communities

We performed a simulation study to assess the type I and type II errors of the z_{dep}test. We simulated 300 species' ranges ("virtual regional pool") as circular spatial polygons in a virtual landscape of 3000 × 3000 units. Polygons differed in diameter, as we would expect in species with different dispersal ability and niche breadth, and its centre was allocated randomly into the virtual landscape. As shown by Martín-Devasa et al. (2022), the shape of the distance-decay patterns varies with (i) the size of species ranges, and (ii) the proportion of restricted vs. widespread species. Therefore, to simulate distance-decay curves with different slopes and intercepts, we varied (i) the diameter of polygons representing species ranges and (ii) the proportion of two range size classes (restricted vs. widespread) in several simulation scenarios (see below). Once virtual species ranges were distributed across the landscape, we randomly allocated sampling sites, each representing a "virtual local community". The presence/absence of species in such virtual local communities was computed from the intersection of species ranges (i.e. circular spatial polygons) at each sampling site. For each virtual landscape, we used two sets of random sampling sites, one set with 25 and the other set with 50 sites, to assess how sample size may affect the type I and type II errors of the test. We performed these simulations using the R packages "sp" (Bivand et al., 2013; Pebesma and Bivand, 2005) and "rgeos" (Bivand and Rundel, 2020). Finally, we measured community similarity between sampling sites using Simpson's pairwise similarity index (Baselga, 2010) with the *beta.pair* function of the R "betapart" package (Baselga and Orme, 2012), as well as the Euclidean spatial distance between them. These similarity and spatial distance matrices are the simulated data used in downstream distance-decay analyses.

2.2. Simulation scenarios for hypothesis testing

2.2.1. Null hypothesis

We defined our null hypothesis as the equality of a given parameter (intercept or slope) between two distance-decay models. In this case, we simulated the two parameters being equal between both models. To do so, we simulated one virtual regional pool and a single set of virtual local communities, from which we computed one similarity matrix and one spatial distance matrix. To build two distance-decay curves with equal

parameters, it is necessary to obtain two bootstrap samples from the same similarity/distance matrices to ensure that both curves followed the same model. Besides, to preserve the pairwise structure of the distance-decay data in the two samples, we performed a site-block resampling of similarity/distance values, with site-blocks defined as the set of values derived from the same observation (Martínez-Santalla et al., 2022). In sum, the two distance-decay models were created by site-block resampling the initial simulated similarities and distances.

Three scenarios were designed to study the performance of the z_{dep} statistic under the null hypothesis:

- Scenario #1: It consisted of 20 widespread species that occupied all the territory, with polygon diameters sampled from a uniform distribution U [4000, 6000], and 280 spatially-restricted species with polygon diameters sampled from a U [300, 900]. This scenario is designed to simulate a distance-decay curve with rapid initial decay and incomplete turnover (Fig. 2).
- Scenario #2: It consisted of 20 widespread species that did not occupy all the territory, with polygon diameters sampled from a U [3000, 4000], and 280 spatially-restricted species with polygon diameters sampled from a U [600, 1200]. This scenario is designed to simulate a distance-decay curve with a decrease of similarity values from the shortest distances, resulting in complete turnover at large distances and therefore values of zero similarity between very distant sites (Fig. 2).
- Scenario #3: It consisted of 50 widespread species that did not occupy all the territory, with polygon diameters sampled from a U [3000, 4000], and 250 spatially-restricted species with polygon diameters sampled from a U [1500, 2000]. This scenario is designed to simulate a distance-decay curve with high similarity at short spatial distances and a marked decay from medium to large distances (Fig. 2).

2.2.2. *Alternative hypothesis*

The alternative hypothesis states that the parameters of two distance-decay models (i.e. intercept or slope) are different. In each scenario, we simulated two different virtual communities and, using the same set of sampling sites, we obtained two different similarity matrices and a spatial distance matrix. Thus, we obtained two datasets by combining each similarity matrix with the spatial distance matrix. Then, we computed the distance-decay model for each dataset. Simulating two different datasets, named DS_1 and DS_2, ensures that their distance-decay curves are also different.

Three scenarios were designed to study the performance of the z_{dep} statistic under the alternative hypothesis:

- Scenario #4: DS_1- It consisted of 300 species with polygon diameters sampled from a U [3000, 4000]. DS_2- It consisted of 200 widespread species with polygon diameters sampled from a U [4000, 6000] and 100 spatially-restricted species with polygon diameters sampled from a U [3000, 4000]. This scenario is designed to simulate distance-decay curves with different decay rates and similar, but not identical, initial similarity (Fig. 2).
- Scenario #5: DS_1- It consisted of 20 widespread species with polygon diameters sampled from a U [4000, 6000] and 280 spatially-restricted species with polygon diameters sampled from a U [1500, 2000]. DS_2- It consisted of 20 widespread species with polygon diameters sampled from a U [3000, 4000] and 280 spatially-restricted

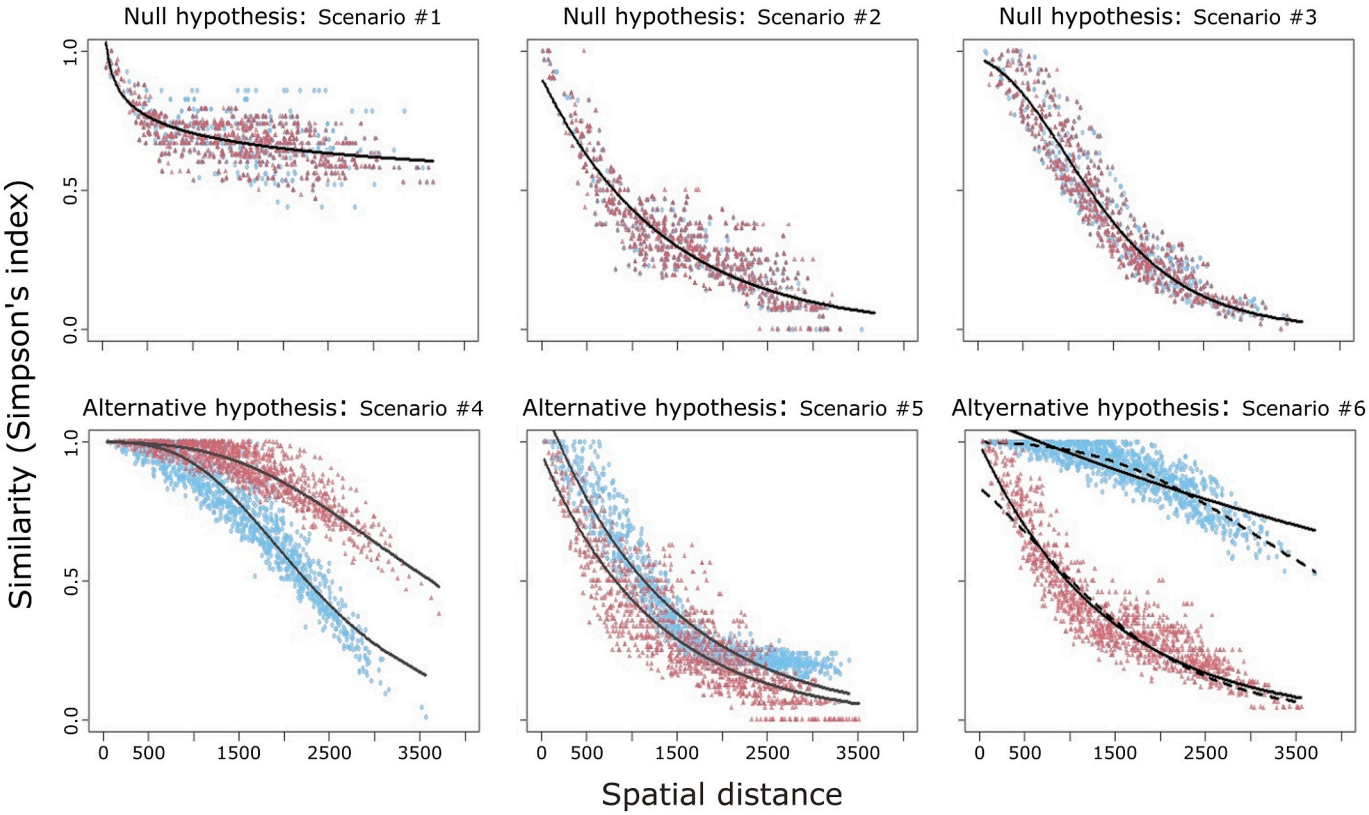


Fig. 2. Example of distance-decay simulations for the calibration of the z_{dep} statistic under the null (top row) and alternative (bottom row) hypotheses with 50 sampling points. Colours (red and blue) represent the two datasets to compare. For the null hypotheses, each dataset was obtained using site-block resampling from a single similarity and distance matrix. Curves represent the adjustment of the model that best fit the data, as detailed in Table 1. In scenario #6, where different models were selected for each dataset, continuous lines represent an exponential model and dashed lines a Gompertz model. Each curve is described by two parameters (θ_a and θ_b, e.g., the intercept and slope of a negative exponential model, respectively), so the z_{dep} test can be used to compare any of them between two datasets (i.e. θ_{a1} vs. θ_{a2}, and/or θ_{b1} vs. θ_{b2}). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

species with polygon diameters sampled from a U [200,1200]. This scenario is designed to simulate distance-decay curves with similar, but not identical, decay rates and different initial similarity (Fig. 2).

- Scenario #6: DS_1- It consisted of 200 widespread species with polygon diameters sampled from a U [4000, 6000] and 100 spatially-restricted species with polygon diameters sampled from a U [3000, 4000]. DS_2- It consisted of 20 widespread species with polygon diameters sampled from a U [3000, 4000] and 280 spatially-restricted species with polygon diameters sampled from a U [600,1600]. This scenario is designed to simulate distance-decay curves with different initial similarity and decay rates (Fig. 2).

2.3. Distance-decay models

We performed non-linear regressions of pairwise community similarities over spatial distances to compute distance-decay models. We fit a negative exponential, a power-law and a Gompertz function using the *nlsLM* function of the "minpack.lm" package (Elzhov et al., 2016). We selected the model with the best fit based on their AIC. In the scenarios under the alternative hypothesis, when a different model was selected for each dataset (DS1 and DS2), we compared the parameters in both types of models. Model parameters were the intercept and slope in the case of negative exponential and power-law functions, and the position parameter and slope in the case of the Gompertz function (Martín-Devasa et al., 2022).

2.4. z_{dep}: Calibration procedure

Type I error (rejecting the null hypothesis when it is true) was evaluated by applying the z_{dep} statistic in 1000 simulations of the null hypothesis scenario, thus obtaining 1000 *p*-values, and computing the proportion of them being less than the nominal values of α (either $\alpha = 0.1$, $\alpha = 0.05$ and $\alpha = 0.01$), i.e. proportion of rejections under the null hypothesis. These α values were chosen because they are the most commonly used significance levels.

The type II error (not rejecting the null hypothesis when it is false) was evaluated with a similar procedure but, in this case, applying the z_{dep} statistic in 1000 simulations of the alternative hypothesis scenario. This provides the empirical proportion of rejections under the alternative hypotheses, i.e. the statistical power of the test (1-type II error). Under the null hypothesis, the proportion of rejections should be very similar to the significance levels (α), indicating that the type I error is the one assumed for each selected significance level. Under the alternative hypothesis, the test should have a proportion of rejections close to 1 (maximum power), which would indicate that the null hypothesis has been rejected in all cases. The procedure followed for this calibration can be found in Appendix A.

2.5. Application of the z_{dep} to real data

We applied the z_{dep} statistic to assess independently the difference in initial similarity (intercept) and in the rate of species turnover (slope) among distance-decay models of several vertebrate orders in the Iberian Peninsula. We obtained the distribution data of freshwater fishes, amphibians, reptiles, and terrestrial mammals from the IUCN red list (https://www.iucnredlist.org), accessions 2020 (amphibians, reptiles and mammals) and 2021 (freshwater fishes). Among freshwater fishes, we selected the order Cypriniformes, among amphibians the order Anura (frogs), among reptiles the order Squamata (snakes and lizards) and among mammals the orders Chiroptera (bats), Rodentia (rodents) and Carnivora (carnivores). For each order, we built a presence/absence matrix for the Iberian Peninsula in a 1°x1° cell grid. The grid was constructed using the "sp" package, with functions *spsample* to place the cell centroids and *SpatialPixelsDataFrame* to obtain the final grid. The community similarity among cells was computed with Simpson's pairwise similarity index, and the spatial distance among them as the geodesic distance between centroids using the *geodist* function of the "geodist" package (Padgham and Sumner, 2020). As in the simulation study, we fitted the negative exponential, power-law and Gompertz functions with the *nlsLM* function and selected the best fitting model based on the AIC statistic.

3. Results

To evaluate the type I and type II errors of the z_{dep} test, we performed a simulation study, assessing its performance in three null hypothesis and three alternative hypothesis scenarios. The best fitting models for each distance-decay simulation are shown in Table 1. The best supported model only differed between datasets (DS1 and DS2) in scenario #6. Thus, we subdivided this scenario into scenario 6.1, in which we fitted a Gompertz function to both DS1 and DS2 distance-decay curves and, scenario 6.2, in which we fitted an exponential function to both DS1 and DS2 distance-decay curves.

3.1. Calibration of the z_{dep} test

The z_{dep} test showed good approximation to the nominal levels of $\alpha = 0.1$, $\alpha = 0.05$ and $\alpha = 0.01$ independently of model function (negative exponential, power-law and Gompertz) and sample size ($n = 25$ vs. $n = 50$) (Fig. 3, Table S1). This indicates that, under the null hypothesis (equality of parameters), the probability of type I error is the one assumed by the value of α. Under the alternative hypothesis, the test showed low type II error (i.e. the proportion of significant *p*-values was close to 1). As expected, type II error was slightly larger when the estimated values of the parameters were similar, but not identical (i.e. intercepts in scenario 4 and slopes in scenario 5). In other words, in the scenarios designed to have one parameter very similar and the other quite different (scenarios 4 and 5), the observed type II errors for each parameter reflected this contrast (Fig. 3, Table S2). The type II error also decreased with sample size, being lower with $n = 50$ than with $n = 25$.

3.2. Real data

We compared the distance-decay patterns of Cypriniformes, Anura (frogs), Squamata (lizards and snakes), Chiroptera (bats), Rodentia (rodents) and Carnivora (carnivores) in the Iberian Peninsula by testing the equality of parameters of their distance-decay curves. In all cases, except in the cyprinids, the exponential function was the best fitting model and was thus used for the comparison of distance-decay parameters across taxa. The cyprinids had significantly different intercept and slope than any other group, showing the steepest decay in community similarity (Fig. 4). No difference in distance-decay parameters was found between ectotherm orders (frogs and squamates). Similarly, mammal orders (endotherms) showed all similar intercepts and slopes, with only the carnivores having a marginally significant ($p < 0.1$) flatter slope than bats and rodents (Table 2). On the contrary, marked differences were observed between ectotherm and endotherm groups. The frogs and squamates' slopes were significantly different from the ones of bats, rodents and carnivores, being steeper in the ectotherm's groups. However, only bats and frogs showed significant differences in the

Table 1
Models best supported by the data (i.e. smallest AIC value) for each of the simulated distance-decay curves, based on the AIC.

Null hypotheses		Alternative hypotheses		
Scenario	Model	Scenario	Model DS1	Model DS2
1	Power-law	4	Gompertz	Gompertz
2	Negative exponential	5	Negative exponential	Negative exponential
3	Gompertz	6	Negative exponential	Gompertz

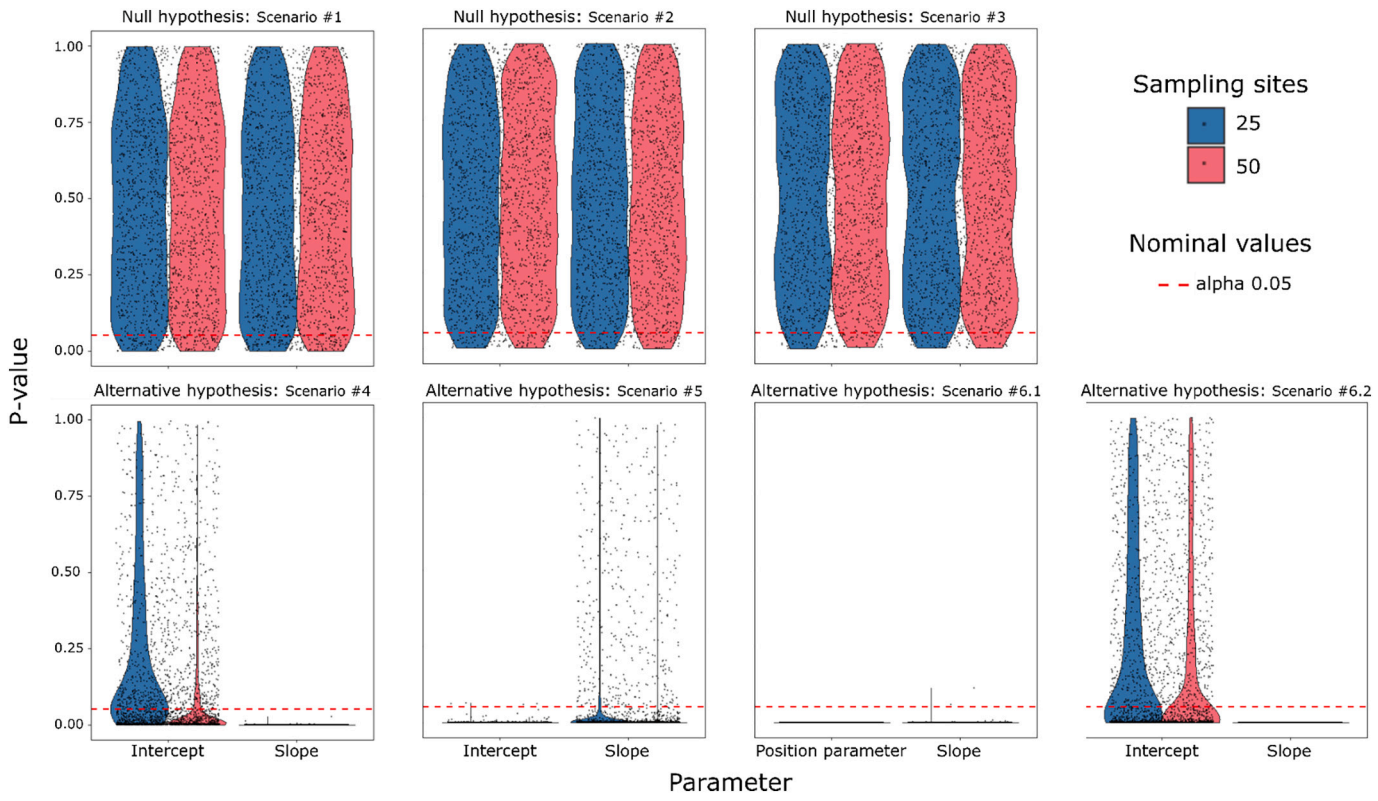


Fig. 3. Distribution of p-values across 1000 replicates in all simulated scenarios, comparing the results for the two parameters (intercept/position and slope) and with the two sample sizes (n = 25 and n = 50). Note that under the null hypothesis, the p-values follow a uniform distribution, so the proportion of rejections equals the nominal α level, for any selected value of α. This proves that the test avoids the inflation of degrees of freedom caused by pairwise dependence. Red dotted lines represent α = 0.05. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

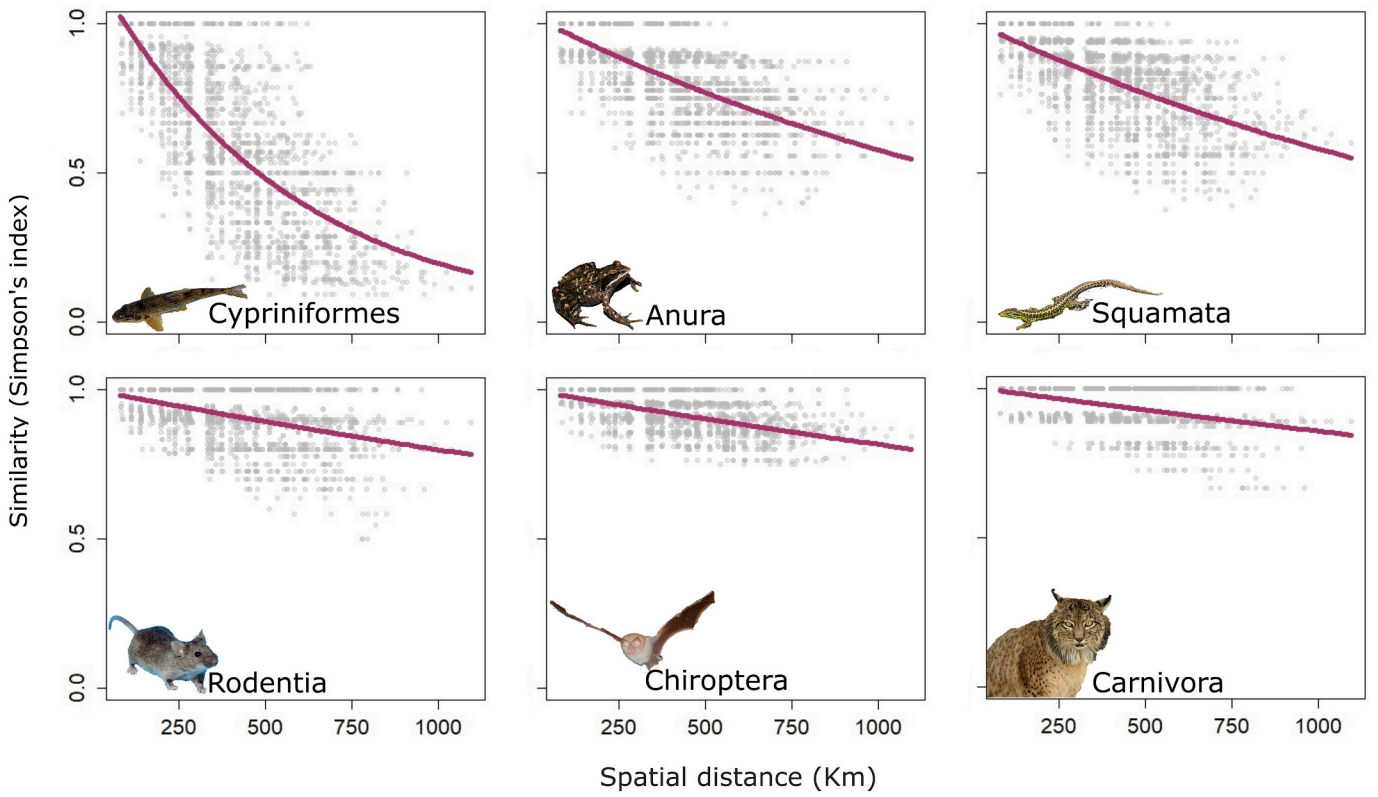


Fig. 4. Distance-decay models (negative exponential) of Iberian vertebrates' orders (obtained with 63 sampling points). Vertebrate pictures were downloaded from Wikimedia Commons (under Creative Commons licence, authors David Perez, Benny Trapp and Carlosblh).

Table 2
P-values of the z_{dep} statistic in the comparison of the intercept and slope between distance-decay models of different groups of Iberian vertebrates (Anu = Anura, Squ = Squamata, Chi = Chiroptera, Rod = Rodentia, Car = Carnivora). Significant values are shown in bold.

Intercept						Slope					
	Cyp	Anu	Squ	Chi	Rod		Cyp	Anu	Squ	Chi	Rod
Anu	**<0.01**					Anu	**<0.01**				
Squ	**<0.01**	0.34				Squ	**<0.01**	0.72			
Chi	**<0.01**	**0.02**	0.45			Chi	**<0.01**	**<0.01**	**<0.01**		
Rod	**<0.01**	0.09	0.57	0.88		Rod	**<0.01**	**<0.01**	**<0.01**	0.53	
Car	**<0.01**	0.12	0.84	0.40	0.61	Car	**<0.01**	**<0.01**	**<0.01**	0.07	0.08

intercept (Table 2). Parameter values can be found in the Supplementary Material Table S3.

4. Discussion

Our results show the good performance of the z_{dep} statistic for the comparison of parameters between distance-decay models, a widely-used statistical approach in the assessment of biodiversity patterns (Nekola and White, 1999; Morlon et al., 2008; Soininen et al., 2007). To overcome the inflation of degrees of freedom inherent to this type of regression models, the z_{dep} statistic implements a site-block resampling approach to estimate the variance of model parameters. The site-block resampling ensures that pairwise dependence is preserved in the data while the number of samples used to compute parameter variance (i.e. blocks) is equal to the effective sample size of the original data (i.e. number of sites), avoiding the inflation of degrees of freedom. Site-block resampling has also proven effective in significance tests for the non-linear relationships between community similarity and spatial distance (Martínez-Santalla et al., 2022). Here, we use site-block resampling for a different problem, i.e. a test for the equality of the parameters between two different distance-decay models. The test can be applied, independently, to any of the model parameters (i.e. intercept and/or slopes). Our simulation study under the null hypothesis shows that the type I error of z_{dep} was consistently similar to the nominal α values. Simulations under the alternative hypothesis evidenced that the z_{dep} test also has good statistical power (low type II error). The low Type I error does not seem to be affected by the mathematical function used to fit the curve, so the z_{dep} test can be used to compare parameters of the most frequently used distance-decay models (power-law, negative exponential and Gompertz function). This is an important advantage as the functional form of the distance-decay curve can change depending on the study scale (Nekola and McGill, 2014) and the species range sizes (Martín-Devasa et al., 2022). As expected, statistical power increases with sample size and decreases when parameter estimates are close, a common behaviour in most tests (Steidl et al., 1997). In the case study with real data, the z_{dep} statistic has also proven useful to reveal differences in the patterns of species turnover with spatial distance among different taxa of Iberian vertebrates.

Previous contributions have evidenced the relevance of comparing the parameters of distance-decay models across biological groups or biogeographic regions in order to better understand the causal processes behind community turnover. Some of these studies have provided different partial solutions to the problem we address in this paper. For example, Goldenberg Vilar et al. (2014) used a two-sample *t*-test. This approach could be, in principle, suitable to compare regression parameters, but its performance is biased with pairwise dependent data (Rasch et al., 2011). A step forward was introduced by Nekola and White (1999) and Gómez-Rodríguez and Baselga (2018), who used resampling procedures on similarity values to estimate the empirical distribution of either the difference between parameters or the parameters themselves. However, these approaches do not completely solve the problem of Type I error inflation because they do not account for the pairwise dependence of the similarity data during the resampling procedure. In turn, La Sorte and McKinney (2006) developed a resampling method directly on the presence/absence table rather than on the similarity matrix. This method truly accounts for the pairwise dependence of similarity data, but has some important differences with our z_{dep} statistic. First, the observed parameter difference may not be within the bounds of the estimated null distribution, a situation that precludes estimating the probability that the observed parameter difference may be caused by chance (La Sorte and McKinney, 2006). When the value to compare is out of the bounds of the estimated null distribution, this type of permutation tests can result in inaccurate *p*-value estimations (Onghena and May, 1995; Phipson and Smyth, 2010). This is not a concern for the z_{dep} statistic as its distribution under the null hypothesis is known to follow a N(0,1) distribution. The second main difference is the calibration test. In this paper we have evaluated the performance of the z_{dep} statistic in terms of type I and type II errors with a simulation study in the specific context of pairwise dependence, obtaining good results. Although permutation tests also tend to have good performances, to our knowledge there is no type I and II error evaluation for La Sorte and McKinney (2006) permutation test, so its approximation to the nominal level (α) and type II error remains unknown. In general, we think that the specific design of the z_{dep} statistic to account for pairwise dependence, its known distribution under the null hypothesis, and the good performance evidenced by our simulation study make it the best option to perform comparisons of parameters between distance-decay models.

The usefulness of the z_{dep} statistic has been exemplified in the study of distance-decay patterns of Iberian vertebrates. We have compared distance-decay parameters among taxonomic groups given that these parameters can be used to understand how the organisms' dispersal and ecological characteristics influence the spatial turnover of biological communities (Gómez-Rodríguez and Baselga, 2018; Morlon et al., 2008; Soininen et al., 2007). Therefore, the steeper distance-decay slope of the cyprinids is probably reflecting the poor dispersal ability of freshwater fishes in comparison with terrestrial vertebrates (Leroy et al., 2019). This is an expected pattern in freshwater fish because drylands and saltwater are barriers that constrain their distribution to rivers and lakes, which can be considered as islands (Dias et al., 2014; Tedesco et al., 2012). As a consequence, dispersal limitation and therefore spatial distance between sites are major factors driving the composition of cyprinids communities (Baselga and Leprieur, 2015; Drakou et al., 2009). Our results also show that, in the Iberian Peninsula, ectotherms have higher turnover rate than endotherms, a pattern that seems to be the general trend across regions (Buckley and Jetz, 2008; Qian and Ricklefs, 2012). This can be explained by the poor dispersal ability of amphibians and reptiles (Crnobrnja-Isailovic, 2007), and/or by their smaller tolerance to environmental changes compared to endotherms (Buckley et al., 2012). In our particular case, the Iberian Peninsula is a biogeographical heterogeneous territory, with regions at both sides of the Eurosiberian / Mediterranean boundary, and multiple mountain ranges creating steep environmental gradients at relatively small spatial scales. As a result, species turnover in lizards and frogs tends to be more marked than in mammals because their community composition is more tightly related to the environmental differences between regions (Sillero et al., 2009). The lack of difference in distance-decay patterns between frogs and lizards, or among mammal orders is probably related to the spatial extent of the study, which may not be large enough to evidence

differences when dispersal abilities are largely similar between taxonomic orders.

In general terms, besides the spatial extent of the analysis, the statistical power of the z_{dep} statistic could also be affected by any factor influencing the variance of parameter estimates, as in any other statistical test. For example, smaller grain sizes (e.g., 0.5° cells instead of 1° cells), uneven sampling schemes or low sampling effort might increase the variance of parameters, thus reducing the statistical power of z_{dep}. Regarding the calibration of the Type I error rate, it is important to stress that data simulation under the null hypothesis is far from trivial. Initially, our intuition was that we could simulate species distributions in a single virtual landscape and use two independent sets of sampling points to compute two distance-decay models (data not shown). However, such procedure yielded a probability of rejection larger than the nominal α value because in fact this approach does not ensure that the distance-decay curves followed the exact same distribution. In other words, the simulations were deviating from the null hypothesis because, even if the species distributions are the same, the patterns of dissimilarity can be different in different regions of the landscape, so using two independent sets of sampling points could result in two different distance-decay curves. To solve this problem, we based our approach to simulate distance-decay data under the null hypothesis on the idea that we needed to generate two models with the same theoretical parameters but different error. By making a resampling of the same model we can ensure that this requirement is fulfilled. With our method we do not know the real parameters, but we are sure that we are working with the null hypothesis because both models are samples of the same one, ensuring that $f1(x) = \beta + \beta x + \varepsilon 1$, $f2(x) = \beta + \beta x + \varepsilon 2$. Therefore, we can effectively calibrate Type I error.

Here we have illustrated the use of the z_{dep} statistic to compare distance-decay patterns across taxonomic groups. However, the z_{dep} test can be extended to the study of any pattern resulting from the application of pairwise metrics, for example, the relationship between genetic and spatial distances (Wright, 1943). The functional form of the relationship between genetic and spatial distance can also be modelled with parametric functions, and the slope of this relation is used to infer the organisms' dispersal ability (Chust et al., 2016; Gómez-Rodríguez et al., 2020; Kinlan and Gaines, 2003; Lester et al., 2007). Similarly, the z_{dep} can be also applied to data lacking spatial structure, such as temporal turnover or compositional variation with abiotic gradients. Therefore, the z_{dep} statistic can be extended to the study of numerous diversity patterns at various hierarchical levels of biological organization (e.g. Baselga et al., 2015, 2022).

5. Conclusions

The good performance of the z_{dep} statistic in the calibration study and its applicability to real biodiversity data illustrated here with the Iberian vertebrates, make it the best option to assess differences in parameters of regression models of pairwise data, such as models of distance-decay of community similarity. The z_{dep} statistic provides accurate results as it accounts for pairwise dependence, which avoids the inflation of degrees of freedom, and thus provides a good approximation to the nominal α level (i.e. good Type I error). Importantly too, the z_{dep} statistic can be applied to any functional form of distance-decay, including non-linear relationships as the negative exponential, power-law, or sigmoidal (i.e. Gompertz). Because the model parameters of these functions have a biological interpretation (Qian, 2009; Saito et al., 2015; Soininen et al., 2007), assessing their differences across biological groups or biogeographic regions is key for our inferences of the causal processes behind community turnover (Gómez-Rodríguez and Baselga, 2018). In other words, this new method allows assessing whether differences in distance-decay parameters are significant between biological groups or regions, and this opens the opportunity to study whether the biological attributes of biological groups (e.g., dispersal limitation or ecological requirements) or the characteristics of biogeographic regions drive the differences in distance-decay patterns. This type of inferences, based on comparisons between biological groups or biogeographic regions with contrasting attributes, are crucial for understanding the causal processes behind the spatial turnover of biological communities and, therefore, for our ability to predict how these biodiversity patterns would be affected by future changes in the environmental conditions.

Data accessibility statement

The code used in the simulation study is available as Supplementary Material. The z_{dep}statistic is to be implemented in the R package *betapart* (Baselga and Orme, 2012).

Declaration of Competing Interest

None.

Data availability

We only used public data and the code of our simulation procedure has been shared in the Attach File step

Acknowledgments

We are grateful to an anonymous reviewer and the associate editor for valuable comments and suggestions that greatly improved this paper. This work was supported by the Spanish Ministry of Science and Innovation and the European Regional Development Fund (ERDF) through grants PID2020-112935GB-I00, PID2020-116587GB-I00, and a FPU scholarship (Ref.: FPU17/03016; Ministry of Education) to R.M.D.

Appendix A. Test calibration

The performance of the z_{dep} statistic was evaluated using simulated distance-decay data to build null and alternative hypothesis scenarios in which we assessed which model function (negative exponential, power-law or Gompertz) was the best-supported to describe the relationship between community similarity and spatial distance. Once the best-supported model was fitted to the data, we used the z_{dep} statistic to perform parameters comparison. The complete procedure is as follows:

1. Simulate two datasets (S1, D1) and (S2, D2), where S and D are the community similarity and spatial distance matrices, respectively, in a null or alternative hypothesis scenario.
2. Fit the best-supported model (Table 1) to estimate the distance-decay parameters of each curve: $\widehat{\theta_{a1}}$, $\widehat{\theta_{b1}}$ and $\widehat{\theta_{a2}}$, $\widehat{\theta_{b2}}$, respectively. The first subindex stands for the parameter (a = intercept or position parameter, and b = slope), while the second subindex stands for the two models that are being compared (model 1 or model 2).
3. Obtain the value of the z_{dep} statistic (Eq. 1) for the comparisons θ_{a1} vs. θ_{a2}, and θ_{b1} vs. θ_{b2}. To compute the variance in the parameters' estimations, required for obtaining the denominator of the test statistic, we used 300 resamples of a site-block resampling.
4. Repeat steps 1–31,000 times, obtaining 1000 values of the test statistic, namely $z_{dep.\ k}$, with $k = 1, \ldots, 1000$.

5. Compute the corresponding 1000 *p*-values as 2*P(|z| > $|z_{dep.\ k}|$), being $z_{dep.\ k}$ the observed value of the statistic in each of the *k* samples computed in step 4, and z a N (0,1) distribution.
6. Calculate the proportion of null hypotheses rejections in each of the *k* samples with a nominal level of α = 0.1, α = 0.05 and α = 0.01.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2022.101894.

References

Astorga, A., Oksanen, J., Luoto, M., Soininen, J., Virtanen, R., Muotka, T., 2012. Distance decay of similarity in freshwater communities: do macro- and microorganisms follow the same rules? Glob. Ecol. Biogeogr. 21 (3), 365–375. https://doi.org/10.1111/j.1466-8238.2011.00681.x.

Bahram, M., Kõljalg, U., Courty, P.-E., Diédhiou, A.G., Kjøller, R., Põlme, S., Ryberg, M., Veldre, V., Tedersoo, L., 2013. The distance decay of similarity in communities of ectomycorrhizal fungi in different ecosystems and scales. J. Ecol. 101 (5), 1335–1344. https://doi.org/10.1111/1365-2745.12120.

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. Glob. Ecol. Biogeogr. 19 (1), 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x.

Baselga, A., Gómez-Rodríguez, C., 2021. Assessing the equilibrium between assemblage composition and climate: a directional distance-decay approach. J. Anim. Ecol. 90 (8), 1906–1918. https://doi.org/10.1111/1365-2656.13509.

Baselga, A., Leprieur, F., 2015. Comparing methods to separate components of beta diversity. Methods Ecol. Evol. 6 (9), 1069–1079. https://doi.org/10.1111/2041-210X.12388.

Baselga, A., Orme, C.D.L., 2012. Betapart: an R package for the study of beta diversity. Methods Ecol. Evol. 3 (5), 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x.

Baselga, A., Gómez-Rodríguez, C., Vogler, A.P., 2015. Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes: multi-hierarchical macroecology. Glob. Ecol. Biogeogr. 24 (8), 873–882. https://doi.org/10.1111/geb.12322.

Baselga, A., Gómez-Rodríguez, C., Araújo, M.B., Castro-Insua, A., Arenas, M., Posada, D., Vogler, A.P., 2022. Joint analysis of species and genetic variation to quantify the role of dispersal and environmental constraints in community turnover. Ecography 2022 (5). https://doi.org/10.1111/ecog.05808.

Bivand, R., Rundel, C., 2020. Rgeos: Interface to geometry engine—open source ('GEOS'). https://CRAN.R-project.org/package=rgeos.

Bivand, R.S., Pebesma, E., Gomez-Rubio, V., 2013. Applied Spatial Data Analysis with R, Second edition. Springer, NY. https://asdar-book.org/.

Buckley, L.B., Jetz, W., 2008. Linking global turnover of species and environments. Proc. Natl. Acad. Sci. 105 (46), 17836–17841. https://doi.org/10.1073/pnas.0803524105.

Buckley, L.B., Hurlbert, A.H., Jetz, W., 2012. Broad-scale ecological implications of ectothermy and endothermy in changing environments: Ectothermy and endothermy. Glob. Ecol. Biogeogr. 21 (9), 873–885. https://doi.org/10.1111/j.1466-8238.2011.00737.x.

Chust, G., Villarino, E., Chenuil, A., Irigoien, X., Bizsel, N., Bode, A., Broms, C., Claus, S., Fernández de Puelles, M.L., Fonda-Umani, S., Hoarau, G., Mazzocchi, M.G., Mozetič, P., Vandepitte, L., Veríssimo, H., Zervoudaki, S., Borja, A., 2016. Dispersal similarly shapes both population genetics and community patterns in the marine realm. Sci. Rep. 6 (1), 28730. https://doi.org/10.1038/srep28730.

Crnobrnja-Isailovic, J., 2007. Cross-section of a refugium: Genetic diversity of amphibian and reptile populations in the Balkans. In: Phylogeography of Southern European Refugia. Springer, pp. 327–337.

Dale, Mark R.T., Fortin, Marie-Josée, 2002. Spatial autocorrelation and statiatical test in ecology. Écoscience 9 (2), 162–167. https://doi.org/10.1080/11956860.2002.11682702.

Dias, M.S., Oberdorff, T., Hugueny, B., Leprieur, F., Jézéquel, C., Cornu, J.-F., Brosse, S., Grenouillet, G., Tedesco, P.A., 2014. Global imprint of historical connectivity on freshwater fish biodiversity. Ecol. Lett. 17 (9), 1130–1140. https://doi.org/10.1111/ele.12319.

Diserud, O.H., Ødegaard, F., 2007. A multiple-site similarity measure. Biol. Lett. 3 (1), 20–22. https://doi.org/10.1098/rsbl.2006.0553.

Drakou, E.G., Bobori, D.C., Kallimanis, A.S., Mazaris, A.D., Sgardelis, S.P., Pantis, J.D., 2009. Freshwater fish community structured more by dispersal limitation than by environmental heterogeneity. Ecol. Freshw. Fish 18 (3), 369–379. https://doi.org/10.1111/j.1600-0633.2009.00354.x.

Elzhov, T.V., Mullen, K.M., Spiess, A.-N., Bolker, B., 2016. Minpack.Lm: R Interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. https://CRAN.R-project.org/package=minpack.lm.

Ferrier, S., Drielsma, M., Manion, G., Watson, G., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in Northeast New South Wales. II. Community-level modelling. Biodiversity and Conservation 11, 2309–2338.

Ferrier, S., Manion, G., Elith, J., Richardson, K., 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Divers. Distrib. 13 (3), 252–264. https://doi.org/10.1111/j.1472-4642.2007.00341.x.

Goldenberg Vilar, A., van Dam, H., van Loon, E.E., Vonk, J.A., van Der Geest, H.G., Admiraal, W., 2014. Eutrophication decreases distance decay of similarity in diatom communities. Freshw. Biol. 59 (7), 1522–1531. https://doi.org/10.1111/fwb.12363.

Gómez-Rodríguez, C., Baselga, A., 2018. Variation among European beetle taxa in patterns of distance decay of similarity suggests a major role of dispersal processes. Ecography 41 (11), 1825–1834. https://doi.org/10.1111/ecog.03693.

Gómez-Rodríguez, C., Miller, K.E., Castillejo, J., Iglesias-Piñeiro, J., Baselga, A., 2020. Disparate dispersal limitation in *Geomalacus* slugs unveiled by the shape and slope of the genetic–spatial distance relationship. Ecography 43 (8), 1229–1240. https://doi.org/10.1111/ecog.05142.

Gossner, M.M., Lewinsohn, T.M., Kahl, T., Grassein, F., Boch, S., Prati, D., Birkhofer, K., Renner, S.C., Sikorski, J., Wubet, T., Arndt, H., Baumgartner, V., Blaser, S., Blüthgen, N., Börschig, C., Buscot, F., Diekötter, T., Jorge, L.R., Jung, K., Allan, E., 2016. Land-use intensification causes multitrophic homogenization of grassland communities. Nature 540 (7632), 266–269. https://doi.org/10.1038/nature20575.

Graco-Roza, C., Aarnio, S., Abrego, N., Acosta, A.T.R., Alahuhta, J., Altman, J., Angiolini, C., Aroviita, J., Attorre, F., Baastrup-Spohr, L., Barrera-Alba, J.J., Belmaker, J., Biurrun, I., Bonari, G., Bruelheide, H., Burrascano, S., Carboni, M., Cardoso, P., Carvalho, J.C., Soininen, J., 2022. Distance decay 2.0 – a global synthesis of taxonomic and functional turnover in ecological communities. Glob. Ecol. Biogeogr. 31 (7), 1399–1421. https://doi.org/10.1111/geb.13513.

Kinlan, B.P., Gaines, S.D., 2003. Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology 84 (8), 2007–2020.

Koenig, W., 1999. Spatial autocorrelation of ecological phennomena. Tree 14 (1), 22–26.

Kunsch, H.R., 1989. The jackknife and the bootstrap for general stationary observations. Ann. Stat. 1217–1241.

La Sorte, F.A., McKinney, M.L., 2006. Compositional similarity and the distribution of geographical range size for assemblages of native and non-native species in urban floras. Divers. Distrib. 12, 679–686. https://doi.org/10.1111/j.1366-9516.2006.00276.x.

La Sorte, F.A., McKinney, M.L., Pyšek, P., Klotz, S., Rapson, G.L., Celesti-Grapow, L., Thompson, K., 2008. Distance decay of similarity among European urban floras: the impact of anthropogenic activities on β diversity. Glob. Ecol. Biogeogr. 17 (3), 363–371. https://doi.org/10.1111/j.1466-8238.2007.00369.x.

Lahiri, S.N., 1999. Theorical comparisons of block bootstrap methods. Ann. Stat. 27 (1), 386–404.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74 (6), 1659–1673.

Legendre, P., Fortin, M.-J., 1989. Spatial pattern an ecological analisis. Vegetatio 80, 107–138.

Leroy, B., Dias, M.S., Giraud, E., Hugueny, B., Jézéquel, C., Leprieur, F., Oberdorff, T., Tedesco, P.A., 2019. Global biogeographical regions of freshwater fish species. J. Biogeogr. 46 (11), 2407–2419. https://doi.org/10.1111/jbi.13674.

Lester, S.E., Ruttenberg, B.I., Gaines, S.D., Kinlan, B.P., 2007. The relationship between dispersal ability and geographic range size. Ecol. Lett. 10 (8), 745–758. https://doi.org/10.1111/j.1461-0248.2007.01070.x.

Liu, R.Y., Singh, K., et al., 1992. Moving blocks jackknife and bootstrap capture weak dependence. Exploring the Limits of Bootstrap 225, 248.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27 (2), 209–220.

Mantel, N., Valand, R., 1970. A tehnique of nonparametric multivariate analysis. Biometrics 26 (3), 547–558.

Martín-Devasa, R., Martínez-Santalla, S., Gómez-Rodríguez, C., Crujeiras, R.M., Baselga, A., 2022. Species range size shapes distance-decay in community similarity. Divers. Distrib. 28 (7), 1348–1357. https://doi.org/10.1111/ddi.13550.

Martínez-Santalla, S., Martín-Devasa, R., Gómez-Rodríguez, C., Crujeiras, R.M., Baselga, A., 2022. Assessing the nonlinear decay of community similarity: permutation and site-block resampling significance tests. J. Biogeogr. 49 (5), 968–978. https://doi.org/10.1111/jbi.14351.

Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., Valencia, R., Green, J.L., 2008. A general framework for the distance-decay of similarity in ecological communities. Ecol. Lett. 11 (9), 904–917. https://doi.org/10.1111/j.1461-0248.2008.01202.x.

Nekola, J.C., McGill, B.J., 2014. Scale dependency in the functional form of the distance decay relationship. Ecography 37 (4), 309–320. https://doi.org/10.1111/j.1600-0587.2013.00407.x.

Nekola, J.C., White, P.S., 1999. The distance decay of similarity in biogeography and ecology. J. Biogeogr. 26 (4), 867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x.

Onghena, P., May, R.B., 1995. Pitfalls in computing and interpreting randomization test p values: a commentary on Chen and Dunlap. Behavior Research Methods, Instruments & Computers 27 (3), 408–411.

Padgham, M., Sumner, M.D., 2020. geodist: Fast Dependency-Free Geodesic Distance Calculations. Version 0.0.7. https://CRAN.R-project.org/package=geodist.

Pebesma, E.J., Bivand, R.S., 2005. Classes and methods for spatial data in R. R News 5 (2), 9–13. https://CRAN.R-project.org/doc/Rnews/.
Phipson, B., Smyth, G.K., 2010. Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. Stat. Appl. Genet. Mol. Biol. 9 (1) https://doi.org/10.2202/1544-6115.1585.
Qian, H., 2009. Beta diversity in relation to dispersal ability for vascular plants in North America. Glob. Ecol. Biogeogr. 18 (3), 327–332. https://doi.org/10.1111/j.1466-8238.2009.00450.x.
Qian, H., Ricklefs, R.E., 2012. Disentangling the effects of geographic distance and environmental dissimilarity on global patterns of species turnover. Glob. Ecol. Biogeogr. 21 (3), 341–351. https://doi.org/10.1111/j.1466-8238.2011.00672.x.
Rasch, D., Kubinger, K.D., Moder, K., 2011. The two-sample t test: pre-testing its assumptions does not pay off. Stat. Pap. 52 (1), 219–231. https://doi.org/10.1007/s00362-009-0224-x.
Saito, V.S., Soininen, J., Fonseca-Gessner, A.A., Siqueira, T., 2015. Dispersal traits drive the phylogenetic distance decay of similarity in Neotropical stream metacommunities. J. Biogeogr. 42 (11), 2101–2111. https://doi.org/10.1111/jbi.12577.
Sillero, Brito, Skidmore, Toxopeus, 2009. Biogeographical patterns derived from remote sensing variables: the amphibians and reptiles of the Iberian Peninsula. Amphibia-Reptilia 30 (2), 185–206. https://doi.org/10.1163/156853809788201207.
Smouse, P.E., Long, J.C., Sokal, R.R., 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst. Zool. 35, 6.
Soininen, J., McDonald, R., Hillebrand, H., 2007. The distance decay of similarity in ecological communities. Ecography 30 (1), 3–12. https://doi.org/10.1111/j.0906-7590.2007.04817.x.
Steidl, R.J., Hayes, J.P., Schauber, E., 1997. Statistical power analysis in wildlife research. The Journal of Wildlife Managment 61 (2), 270–279.
Steinbauer, M.J., Dolos, K., Reineking, B., Beierkuhnlein, C., 2012. Current measures for distance decay in similarity of species composition are influenced by study extent and grain size. Glob. Ecol. Biogeogr. 21 (12), 1203–1212. https://doi.org/10.1111/j.1466-8238.2012.00772.x.
Tedesco, P.A., Leprieur, F., Hugueny, B., Brosse, S., Dürr, H.H., Beauchard, O., Busson, F., Oberdorff, T., 2012. Patterns and processes of global riverine fish endemism: global riverine fish endemism patterns. Glob. Ecol. Biogeogr. 21 (10), 977–987. https://doi.org/10.1111/j.1466-8238.2011.00749.x.
Thieltges, D.W., Ferguson, M.A.D., Jones, C.S., Krakau, M., de Montaudouin, X., Noble, L. R., Reise, K., Poulin, R., 2009. Distance decay of similarity among parasite communities of three marine invertebrate hosts. Oecologia 160 (1), 163–173. https://doi.org/10.1007/s00442-009-1276-2.
Wright, S., 1943. Isolation by distance. Genetics 28 (2), 114–138.