Check for updates

Received: 20 May 2021 | Revised: 31 January 2022 | Accepted: 12 February 2022

DOI: 10.1111/jbi.14351

Journal of Biogeography WILEY

METHOD

Assessing the nonlinear decay of community similarity: Permutation and site-block resampling significance tests

Sara Martínez-Santalla[1] | Ramiro Martín-Devasa[1] | Carola Gómez-Rodríguez[2] | Rosa M. Crujeiras[3] | Andrés Baselga[1]

[1]CRETUS, Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[2]CRETUS, Department of Functional Biology (Area of Ecology), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[3]CITMAga, Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

Correspondence
Sara Martínez-Santalla, CRETUS, Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain.
Email: saramartinez.santalla@usc.es

Funding information
European Regional Development Fund; Ministerio de Ciencia e Innovación

Handling Editor: Simone Fattorini

Abstract

Aim: Modelling how community similarity decays with spatial distance is a key tool for the study of the processes behind community variation (beta diversity). Distance-decay models are computed from pairwise metrics (i.e. community similarity and spatial distance between localities) and hence suffer from pairwise dependence in the data, precluding the use of standard significance tests. Besides, distance-decay patterns are inherently nonlinear because similarity is bounded between 1 and 0. However, the only standard method to assess model significance under pairwise dependency is the Mantel test, which considers a linear model. To allow the use of nonlinear models in the assessment of distance-decay patterns, we introduce here a nonlinear significance test combining a pseudo-R^2 statistic with either permutations or block-site resampling with replacement.

Location: Global.

Taxon: Applicable to any taxon.

Methods: To assess the performance of the pseudo-R^2 significance test (i.e. type I error and statistical power), we have applied this method to exponential and power-law generalized linear models of simulated distance-decay data. We compared its performance with the one of a linear model (i.e. Mantel test) and illustrated its use with real data.

Results: The pseudo-R^2 significance test has generally good type I error and statistical power, even for highly nonlinear relationships. The Mantel test shows an adequate empirical approximation of type I error, but presents lower statistical power when the distance-decay relationship highly departs from linearity.

Main conclusions: The proposed pseudo-R^2 significance test for nonlinear models is adequate for exponential and power-law models of distance-decay of similarity and should be preferred over the linear Mantel test. Furthermore, this significance test could also be used in other nonlinear distance-based regressions. The novel site-block resampling method could also be applied to other statistical questions related to distance-decay models that cannot be tackled using permutations, such as estimating the variance of model parameters.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2022 The Authors. *Journal of Biogeography* published by John Wiley & Sons Ltd.

KEYWORDS
beta diversity, distance-decay of similarity, Mantel test, significance test, site-block resampling, spatial turnover

1 | INTRODUCTION

The decrease in similarity between biological communities with increasing distance (i.e. distance-decay of similarity) is a macroecological pattern that has been profusely studied in the last two decades. It is used as a fundamental tool for understanding variation in community composition and the processes underlying them (Gómez-Rodríguez & Baselga, 2018; Tornero et al., 2018; Tuomisto et al., 2003; among others). This pattern has also been analysed in applied conservation studies, for example, to identify relevant factors for community turnover in endangered or altered areas (e.g. Draper et al., 2019; Zhang et al., 2019). The most common approach to study distance-decay patterns is to fit a statistical model of the relationship between two pairwise matrices: community similarity and spatial distance. This statistical model allows the estimation of parameters with biological relevance, such as the proportion of variation in community similarity that is explained by distance, or the rate at which community similarity decays with distance.

The search for an adequate model of distance-decay started with the seminal paper by Nekola and White (1999), in which they found that an exponential model (log-transformed similarity linearly regressed on distance) was the best fit to their data. Since then, distance-decay patterns have been mostly assessed with Linear Models (LM) with or without log-transformed similarity and distance (e.g. Nekola & White, 1999) or Generalized Linear Models (GLM) (e.g. Millar et al., 2011; Soininen et al., 2007). More recently, Nekola and McGill (2014) used simulated data to show that the expected functional form of the distance-decay relationship is either a negative exponential function at larger spatial scales or a power-law function at smaller spatial scales. However, even when the most adequate nonlinear functional form is selected, it is not straightforward to correctly assess the significance of the relationship given the inherent pairwise dependence of the data. Significance tests commonly used for LMs and GLMs (i.e. the *F*-test; see Hastie & Pregibon, 1993) should not be applied to distance-decay data because similarity and distance values violate the independence assumption. This results from similarity/distance matrices being constructed by comparing all possible pairs of communities (Legendre et al., 1994; Smouse et al., 1986). Thus, each independent observation (i.e. a biological community) is involved in the calculation of multiple similarity/distance values, which are therefore not independent among them. In the case of linear models, the statistical bias caused by pairwise dependence in distance matrices has been accounted for in the Mantel (Mantel, 1967) test. The Mantel test assesses the independence between two distance variables by comparing a linear statistic (Pearson's correlation) or a nonparametric statistic (Spearman's or Kendall's rank correlation) to the null correlation distribution (test calibration), obtained through sampling-site permutations. When paired with nonparametric correlation statistics, the Mantel test can accommodate the expected nonlinear relationship between community similarity and spatial distance (Dietz, 1983). However, when possible, the use of nonlinear parametric models (e.g. exponential or power-law) should be preferred over plain nonparametric tests. Nonlinear parametric distance-decay models are particularly useful because parameter comparisons between models fitted for different taxa or regions are key to assess the processes behind distance-decay patterns (Gómez-Rodríguez & Baselga, 2018; Peguero et al., 2021; Saito et al., 2015; Soininen et al., 2007; Yang et al., 2021). To our knowledge, there is no explicit significance test available to assess the significance of negative exponential or power-law distance-decay models that considers both the pairwise structure of distance data (in the test calibration strategy) and their functional form (in the test statistic).

The lack of independence in distance and similarity matrices due to its pairwise structure constitutes the first challenge for designing a significance test of nonlinear distance-decay models, which compares the nonlinear pattern found by the model to what would be expected if similarity and distance were independent from each other (null hypothesis). Most significance tests (e.g. the *F*-test) have a fundamental assumption of units of observation being independent that is violated by the pairwise structure of distance matrices, causing an artificial inflation of degrees of freedom, known as pseudoreplication, that increases with sample size (Hurlbert, 1984) (Figure 1). Pseudoreplication can increase the type I error of significance tests, which can lead to falsely considering the distance-decay model as significant. In the context of distance-decay models, pseudoreplication arises because from N sites, we can compute $N * (N - 1)/2$ similarities. To avoid the inflation of type I error, the Mantel test estimates the null distribution of correlation between similarity and distance through permutations of the sampling sites in the similarity matrix. It then contrasts the observed correlation against its null distribution and calculates the p-value as the proportion of cases in which the null correlation is higher than the observed correlation. The Mantel test constitutes a robust significance test for linear models, although it may be biased when the spatial structure of data is not accounted for (Guillot & Rousset, 2013; Legendre et al., 2015; Legendre & Fortin, 2010). Moreover, a departure from linearity in the relationship may lower its ability to reject the null hypothesis when it is false (low statistical power) (Diniz-Filho et al., 2013; Zeller et al., 2016). This limitation is particularly important in distance-decay models, as we know beforehand that nonlinear models are the expected functional forms (Nekola & McGill, 2014). More fundamentally, the Mantel test is not equivalent to testing the significance of a nonlinear distance-decay model since it does not compare the fitted model with its null hypothesis, but instead uses a linear or nonparametric correlation.

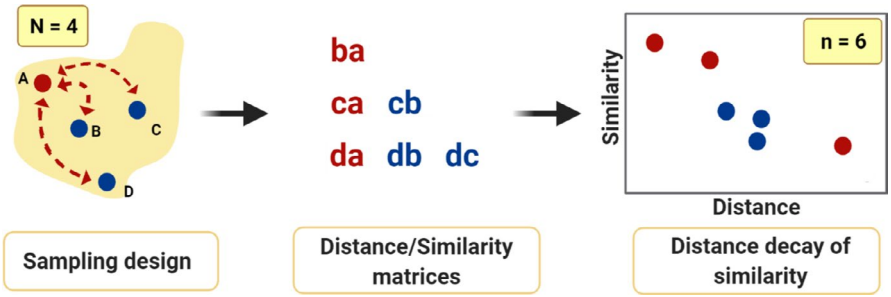

FIGURE 1 Pseudoreplication in distance and similarity data. Sample size artificially inflates after computing similarity and distance matrices, as four sites (i.e. communities) result in six similarity values (i.e. for N sites we have $N * (N - 1)/2$ similarities or distances). Another evidence of the existence of pseudoreplication and pairwise dependence in distance data is given by the fact that site A is involved in the computation of all the similarity and distance values marked in red (i.e. $N - 1$ values). Created with BioRender.com

Effectively detecting nonlinear patterns (i.e. having high statistical power or, in other words, low type II error) is the second challenge for designing a significance test of exponential and power-law models of distance-decay. A possible approach is to use a GLM's pseudo-R^2 statistic, which is suitable for nonlinear models, and estimates how much of the null variation in the data is explained by a fitted, validated model (McCullagh & Nelder, 1989). Thus, in a GLM context, the pseudo-R^2 compares the fit of the model of interest with a null model in which similarity and distance are independent and hence there is no distance-decay relationship. If we follow the rationale of the Mantel test and combine the pseudo-R^2 statistic with permutations of sampling sites in the response matrix, it could be possible to estimate its null distribution under conditions of pairwise dependence. Another option would be to use a resampling method that accounts for the pairwise structure of the data to generate independent resamples (with replacement) of similarity and distance data. In this case, the goal is to preserve the pairwise structure of the data, and therefore resamples should be taken as blocks of similarities involving any given site (site-block resampling). Although for a significance test, both methods would likely yield identical results, a site-block resampling adapted to distance-decay data could be extended to other inference questions affected by pairwise dependence, such as variance estimation of model parameters or testing for differences between fitted models. The pseudo-R^2 value has been previously used in similar contexts, for instance combined with randomizations that do not preserve the pairwise structure of the data (Gómez-Rodríguez & Baselga, 2018), in nonlinear regressions of distance matrices (Lichstein, 2007), or used both as a significance test and a tool for model selection (Generalized Dissimilarity Modelling; Ferrier et al., 2007) of nonlinear distance-decay models. However, to our knowledge, no study has formally assessed the performance of a permutation or block-site resampling significance test for nonlinear distance-decay models.

Distance-decay models have been extensively used in community ecology and biogeography (see, e.g., Nekola & White, 1999, Soininen et al., 2007 and the hundreds of citing papers). From first principles, we know that the relationship between community similarity and spatial (or environmental) distance must be nonlinear (Nekola & McGill, 2014). Therefore, the lack of a standard procedure to assess the significance of nonlinear distance-decay models is a major gap in the biogeographers' toolkit. In this paper, we assess the performance of a pseudo-R^2 statistic in combination with permutations or site-block resampling as a significance test for exponential and power-law models of distance-decay. We propose two strategies to estimate the null distribution of the pseudo-R^2 statistic while considering the pairwise structure of distance-decay data: (i) permutations of sampling sites in the similarity matrix (equivalent to the permutations applied in a Mantel test, but associated here with the pseudo-R^2 statistic of a nonlinear model) or (ii) a site-block resampling method based in selecting as resampling blocks all similarities or distances that involve a given site. We then assess the ability of the pseudo-R^2 significance test to detect the absence of distance-decay (type I error) and its presence (statistical power) using simulations and, finally, we apply the test to three real data examples. This new methodology constitutes an appropriate alternative for nonlinear parametric models, and should be preferred over the Mantel test, as the latter assesses linear or nonparametric relationships.

2 | MATERIALS AND METHODS

2.1 | Simulated and real data

To assess the performance of a significance test based on a pseudo-R^2 statistic for exponential and power-law models of distance-decay of community similarity, our aim was to approximate (1) its type I error, that is, the probability of rejecting the null hypothesis when it is true and (2) its statistical power, that is, the probability of rejecting the null hypothesis when it is false. To quantify these, we simulated data of biological communities across space under two hypotheses (Figure 2):

1. Null (H_0): Absence of relationship between community similarity and spatial distance between communities (i.e. no distance-decay pattern). All species have the same probability of being present at any site, and the species composition of a community does not depend on its spatial location. Thus, spatial distance and

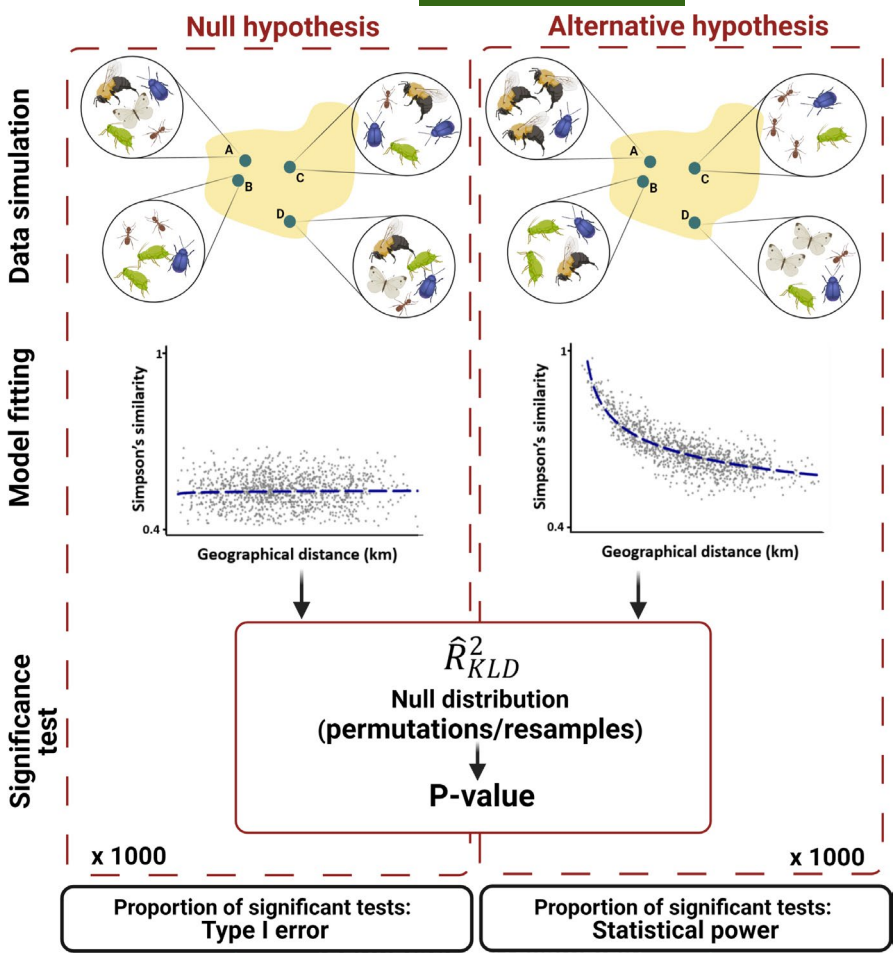

FIGURE 2 Pipeline of the simulations used in this study for assessing the performance of the pseudo-R^2 test (R^2_{KLD}) as a significance test of exponential and power-law models of distance-decay of similarity. Created with BioRender.com

biological similarity are independent, so the null model and the distance-decay model have the same fit to the data.

2. Alternative (H_1): Nonlinear distance-decay of similarity between communities. Species are spatially aggregated and the species composition of a community is correlated with the composition of other communities close to it. Thus, spatial distance is a predictor of community similarity, so the distance-decay model fits the data better than the null model.

To control for the effect of sample size in the error of the pseudo-R^2 test, we simulated communities in $N = \{10, 50, 100\}$ sites with a total of $S = 150$ species across them.

For the simulation of communities under the null hypothesis, N random spatial coordinates were generated from a uniform distribution with a range between 0 and 9000. Then, independently from these coordinates, we simulated species presence–absence matrices of dimension $N \times S$ (N sites and S species) using a binomial random variable with a different probability of success for each species. This probability was taken from the probabilities of a simulated lognormal distribution, to ensure that there is a high proportion of species with narrow distributions and only a few widely distributed species, thus simulating a common pattern in natural communities (Gaston, 1996).

For the simulation of communities under the alternative hypothesis, we generated Gaussian Random Fields (GRF) using grf from the 'geoR' package v.1.7-5.2.2 (Ribeiro Jr et al., 2020) in R with an exponential covariance function. This function allowed us to simulate S species as spatially correlated Gaussian variables with values in N random spatial points. The exponential covariance of the GRF forces the resemblance between Gaussian variables (species) observed in two sampling points to decrease with the distance separating them, such a covariance function is given by:

$$C(h) = \sigma^2 e^{\left(-\frac{h}{\varphi}\right)} \quad (1)$$

where φ is the range parameter that controls the strength of spatial autocorrelation, σ^2 is the partial sill, h is the distance between two points and $C(h)$ is the covariance between variables in those two points. For higher values of φ, the similarity between community composition will decrease more slowly with increasing distance and vice versa. To obtain $N \times S$ presence–absence matrices, as required by many similarity indexes, we transformed the Gaussian variables to binomial data using a different probability of success for each species. This probability was taken from the probabilities of a simulated lognormal distribution, to ensure that there is a high proportion of species with narrow distributions and only a few widely distributed species, as in the case of the null model. We simulated three scenarios under the alternative hypothesis with varying degrees of spatial autocorrelation in the species composition of sampled sites ($\varphi = \{300, 5000, 40{,}000\}$), and $\sigma^2 = 1$.

Finally, we selected three sets of presence/absence matrices with different decay of similarity rate to illustrate the application of the test to real data: (1) Barro Colorado Island's trees (BCI), with 50 one-hectare plots and 225 species (Condit et al., 2002; through 'vegan' package v.2.5-4 in R, Oksanen et al., 2019); (2) *Trechus* beetles in Southern Europe, with 16 countries and 253 species; and (3) *Amara* beetles in Northern Europe, with 18 countries and 135 species. *Trechus* and *Amara* datasets were taken from Löbl and Smetana (2003) and countries were separated into Northern and Southern subsets following Gómez-Rodríguez and Baselga (2018), as these two European regions present markedly different distance-decay patterns.

2.2 | Data processing and distance-decay models

Once we had presence/absence matrices and coordinates for every dataset, both simulated and real, we computed spatial distance and community similarity triangular matrices (semimatrices). Then we fitted distance-decay models to the semimatrices to test the significance of the models with the pseudo-R^2 test.

For distance semimatrices (D_x), we calculated Euclidean distances between coordinates of the simulated sampling sites. In the BCI dataset, distance was computed between the UTM coordinates from BCI.env in the R 'vegan' package. In *Amara* and *Trechus* datasets, distance was computed between the centroids of territories. Lastly, we obtained similarity semimatrices (Sim_y) by applying the Simpson's index using the function beta. pair from R package 'betapart' v.1.5.4 (Baselga et al., 2021; Baselga & Orme, 2012).

We chose GLMs over LMs as for fitting exponential and power-law models to our distance-decay data because their regression parameters are estimated considering the untransformed similarities. Thus, the residuals used to compute the pseudo-R^2 statistic in a GLM are untransformed, avoiding the possible bias caused by directly log-transforming similarity and modifying or removing 0-similarity values (Millar et al., 2011; O'Hara & Kotze, 2010). We fitted exponential GLMs by choosing the logarithm as link function and Gaussian errors. For the power-law GLMs, we did the same but transformed the geographical distance using the logarithm. We constructed null GLMs by regressing similarity on its own mean. All models were computed using the glm2 function from the 'glm2' package in R, v.1.2.1 (Marschner, 2011).

2.3 | The pseudo-R^2 test: R^2_{KLD}

To test the significance of exponential and power-law models of distance-decay, we use the pseudo-R^2 measure based on the Kullback–Leibler divergence described by Cameron and Windmeijer (1997). Our choice of a pseudo-R^2 statistic for this significance test is based on the facts that (1) its null hypothesis is the independence between similarity and distance, (2) it can be calculated for exponential and power-law models and (3) it estimates the proportion of deviance in the null model that is explained by the fitted model, thus serving as a measure of explained variation. Its formulation is:

$$R^2_{KLD} = \frac{D\left(y\hat{\mu}_0\right) - D\left(y\hat{\mu}_1\right)}{D\left(y\hat{\mu}_0\right)} \quad (2)$$

in which $D\left(y\hat{\mu}_0\right)$ is the deviance of the null model (intercept-only model) and $D\left(y\hat{\mu}_1\right)$ is the deviance of the distance-decay model. In a GLM context, the deviance of a model is a measure of how well the model fits the data, with a lower deviance value implying a better fit (McCullagh & Nelder, 1989). Therefore, a larger value of r^2_{KLD} implies a better fit to the data by the distance-decay model. To estimate the statistic's value in our data, $\hat{R}^2_{KLD}$, we obtained the deviance values of the fitted models using the deviance function from the 'stats' package v.3.5.0 (R Core Team, 2018).

For data with no pairwise dependence, this test is equivalent to an *F*-test (Hastie & Pregibon, 1993). However, for similarity and distance data, the distribution of the statistic under the null hypothesis is not known because cases are not independent (Legendre et al., 1994). Therefore, to obtain the p-value of the $\hat{R}^2_{KLD}$ test, we have to estimate its expected distribution under the null hypothesis via resampling methods. The key feature to account for the dependence between pairwise similarities is to consider all the similarities or distance values related to a single site as the basic unit for resampling. These sets of data (i.e. all similarity/distance values in which a particular site/community is involved) are hereinafter referred as site-blocks. Thus, we estimated the null distribution of R^2_{KLD} (Figure 2) using (i) site-block permutations, which permute together all the similarity or distance values related to a single site (site-blocks) and (ii) site-block resampling, which samples with replacement the site-blocks in the similarity and distance matrices and then randomizes within-block values in the similarity matrix (Figure 3).

2.3.1 | Site-block permutations

We used the following procedure to obtain permutations of sampling sites in the similarity matrix (Sim_y), making similarity and distance (D_x) independent (Figure 3a):

1. We reordered Sim_y by row and column (Sim_y*) without modifying D_x.
2. We fitted a GLM to Sim_y* and D_x, and a null GLM to Sim_y*; then computed both deviance values and calculated R^2_{KLD}*.
3. We repeated 1–2 for B permutations.

2.3.2 | Site-block resampling

Here, in contrast to the permutation procedure, resamples of site-blocks (i.e. all similarities or distances involving a given site) are taken

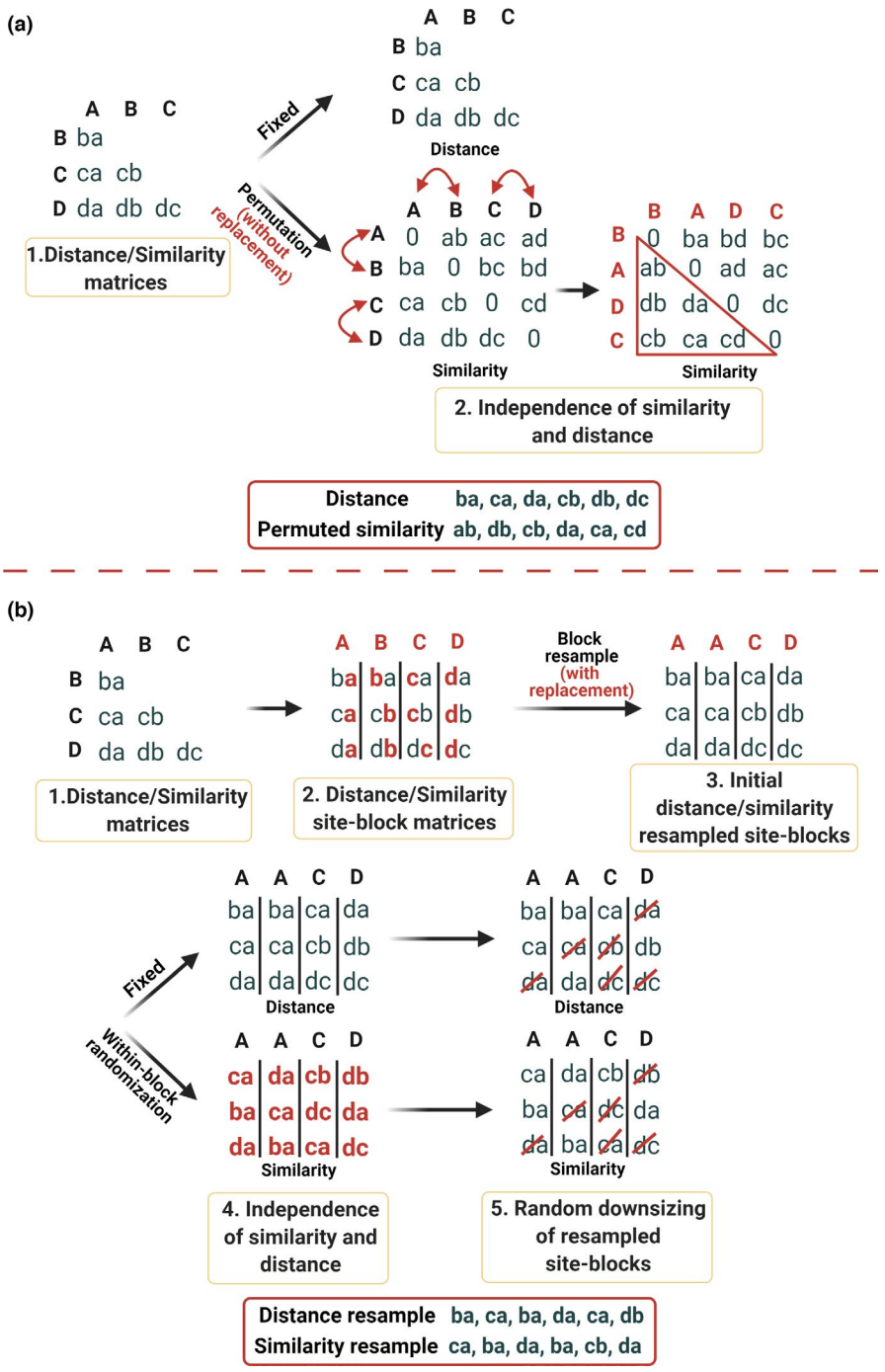

FIGURE 3 Schematic representation of the permutation (a) and site-block resampling (b) procedures to estimate the null distribution of the pseudo-R^2 statistic (R^2_{KLD}) while considering the pairwise structure of distance-decay data. Created with BioRender.com

with replacement and saved to a similarity block matrix. To obtain resamples under the null hypothesis, within-block values need to be randomized in the similarity block matrix. The procedure was (Figure 3b) as follows:

1. For N sites, site-blocks were defined by selecting the values of D_x and Sim_y involving a given site. We saved those N blocks as columns of two block matrices, one for each variable (D_x and Sim_y).
2. We drew N block resamples (columns of block matrices) with replacement, ensuring that the same blocks were sampled for both D_x and Sim_y, and saved them as two resample matrices.
3. To force Sim_y to be independent from D_x (null hypothesis of the deviance test), without changing the block structure, we randomized the Sim_y resample matrix by column (i.e. values within each block resample were randomized in Sim_y).
4. Initially, resample matrices are built with duplicate observations ($N * [N - 1]$ values) because each block includes $N - 1$ values. In this step, we downsized the resample matrices 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 were saved as D_x^* and Sim_y^* resample vectors.
5. We fitted a GLM to D_x^* and Sim_y^* and a null GLM to Sim_y^*, computed both deviance values and calculated $R^2_{KLD}{}^*$.

6. We repeated steps 1–5 B times.

Once we had the null distribution estimated by any of these two methods, we obtained the *p*-value of the test as the proportion of times when $R^{2}_{KLD}{}^{*}$ was larger than $\widehat{R}^{2}_{KLD}$.

2.4 | Assessing the performance of the significance test

To assess the performance of R^{2}_{KLD} (Equation 2) in combination with site-block permutations or site-block resampling procedures for significance testing of distance-decay patterns, we used simulated data to estimate their respective probabilities of (i) rejecting the null hypothesis when it is true (type I error) and (ii) rejecting the null hypothesis when it is false (statistical power). For all the simulation scenarios, the significance level was 0.05. For comparison, we also estimated Mantel test's type I error using the mantel() function from 'vegan' package as well as the type I error of the *F*-test (type I ANOVA as implemented by the anova function from 'stats' package). This allowed us to assess how the inflation of type I error in the *F*-test was corrected by our pseudo-R^2 test and the Mantel test.

To estimate type I error and statistical power, we simulated data under the null and the alternative hypothesis, respectively. We then applied R^{2}_{KLD} with 2500 site-block permutations or site-block resamples, the Mantel test with 2500 permutations, and the *F*-test, and repeated the simulation procedure 1000 times (Figure 2). The expected type I error rate was the same as the significance level, 0.05, which would indicate that the statistic is correctly calibrated. As for statistical power, we expected that the adequacy of the chosen model to the simulated distance-decay relationship would condition the performance of R^{2}_{KLD}. In $\varphi = 5000$ and $\varphi = 40{,}000$ scenarios, the decay of similarity is extended across all the simulated distance range and should be easily detected by all the studied models (exponential, power-law and the linear model of the Mantel test). In contrast, in the scenario with a weaker spatial correlation, $\varphi = 300$, there is a rapid initial decay of similarity from over 0.9 to around 0.4, but for most of the distance range (approximately 95% of the maximum distance), the pattern is close to the null hypothesis. As such, it should be more difficult to detect with the pseudo-R^2 test than the other two simulated scenarios of distance-decay and hence allows to check its performance with weaker patterns (Figure S1).

Finally, we illustrated the application of the pseudo-R^2 test and the Mantel test in the real datasets. In these three scenarios (BCI, *Amara* and *Trechus*), we also compared the fit of linear, exponential and power-law distance-decay models using the estimated R^{2}_{KLD} values to showcase the differences between them.

We conducted all the simulations and statistical work using R v.3.5.0 (R Core Team, 2018). We ran the simulations in the FINISTERRAE supercomputer, a service facilitated by the Centro de Supercomputación de Galicia (CESGA).

3 | RESULTS

In simulations under the null hypothesis, all statistics (exponential R^{2}_{KLD}, power-law R^{2}_{KLD} and Mantel's *r*) had distributions with a mean close to 0 (Figure 4), reflecting the equivalence under this scenario between the fitted models and the null model. Both permutation and site-block resampling strategies used to estimate the null distribution of R^{2}_{KLD} avoided the inflation of type I error. Their estimated type I error was close to 0.05 and did not increase with the number of sites in the similarity matrix (Figure 4). The Mantel test showed an equally good performance regarding type I error. On the contrary, the *F*-test failed to return an accurate type I error rate and its inflation increased with sample size (Figure S2).

In simulations under the alternative hypothesis, for the scenarios with high spatial autocorrelation in community composition ($\varphi = 5000$ and $\varphi = 40{,}000$), most R^{2}_{KLD} estimated values were over 0.25, meaning that exponential and power-law models based on spatial distance were able to explain a considerable proportion of variation in community similarity. However, when spatial autocorrelation in community composition was low ($\varphi = 300$), the values of $\widehat{R}^{2}_{KLD}$ were generally under 0.05, which implies that the power-law and exponential models do not explain a higher proportion of variability than the null model (Figure 4; Figures S3 and S4). This had consequences in the statistical power, as this scenario ($\varphi = 300$) is in fact very similar to the null hypothesis. In consequence, in simulations under the alternative hypothesis, the R^{2}_{KLD} significance test, based either on site-block permutations or site-block resampling, was generally able to reject the null hypothesis for all combinations of spatial autocorrelation and sample size (statistical power > 0.9), except for the distance-decay scenario with the weakest spatial autocorrelation ($\varphi = 300$) (Figure 4; Figures S3 and S4). On the contrary, when spatial autocorrelation was weak ($\varphi = 300$), the estimated statistical power increased with sample size, especially for the power-law R^{2}_{KLD}, although all tests performed poorly for extremely low sample sizes ($N = 10$), as expected. The negative exponential R^{2}_{KLD} and the Mantel test, which do not accommodate a rapid initial distance-decay, always presented a lower statistical power than the power-law model (Figure 4; Figure S4).

Finally, R^{2}_{KLD} test was able to detect significant ($\alpha = 0.05$) distance-decay patterns in the three real data cases, independently of whether the model applied was negative exponential or power-law; as the Mantel test did for linear models (Figure 5; Table S1). Overall, we found that negative exponential, power-law and linear models were not equivalent in their fit to real data; especially for the *Trechus* dataset, which had the most pronounced distance-decay. The power-law model fitted best the data (highest $\widehat{R}^{2}_{KLD}$ values) in BCI and *Trechus* datasets, although the amount of explained variation was much higher in the southern European *Trechus* beetles than in the tree communities of Barro Colorado Island. In turn, the negative exponential model (and similarly the linear model) provided a better fit for the northern European *Amara* beetles.

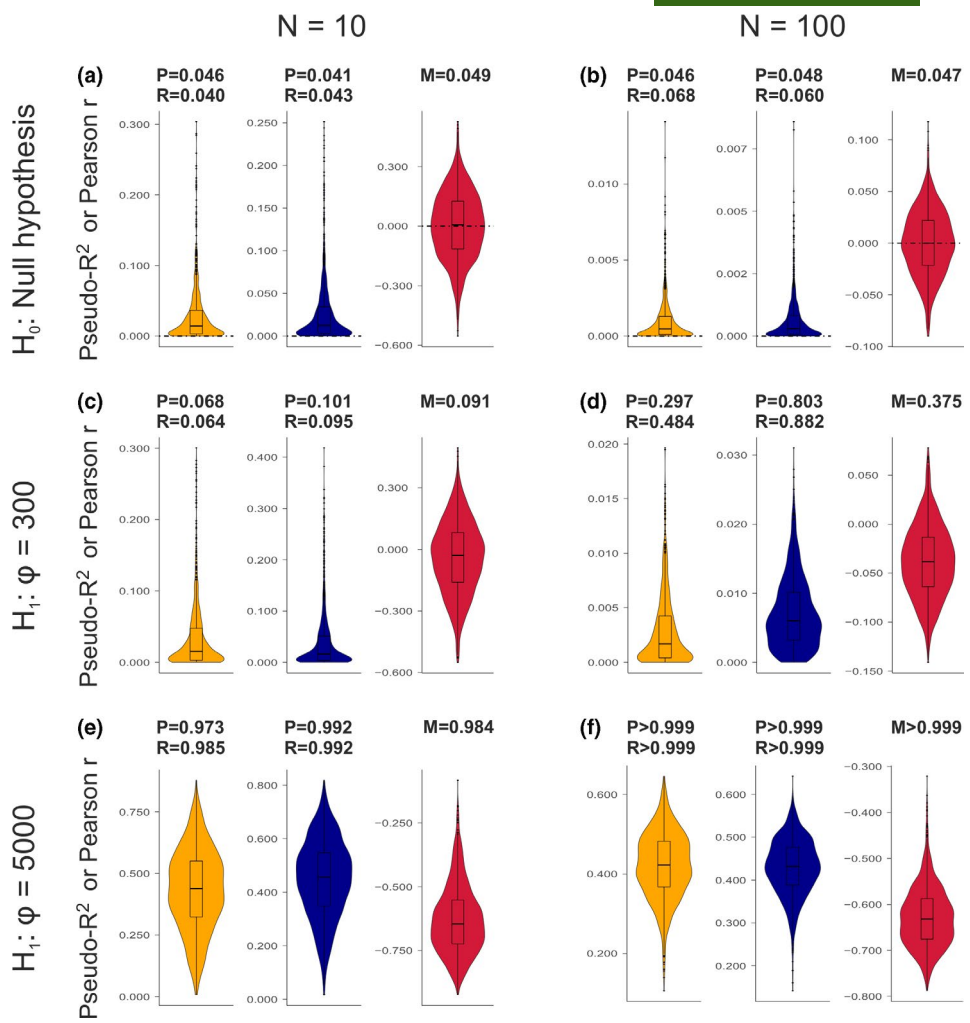

FIGURE 4 Proportion of significant tests across 1000 simulations for the site-block permutations (P), site-block resampling (R) and Mantel test (M). Simulations were run either under the null hypothesis (a, b) or the alternative hypotheses (c, d for $\varphi = 300$; and e,f for $\varphi = 5000$), and for a small ($N = 10$ in a, c, e) and large ($N = 100$ in b, d, f) number of sites. The proportion of significant tests in simulations under the null hypothesis is type I error, while the proportion of significant tests in simulations under the alternative hypothesis is statistical power. Estimated distribution of the statistics (R^2_{KLD} or Pearson r), type I error and statistical power are shown for negative exponential (yellow), power-law (blue) and linear Mantel models (red). Results for the $\varphi = 40{,}000$ scenario, as well as for $N = 50$ under all simulation scenarios are provided in the supplementary material (Figures S3 and S4)

4 | DISCUSSION

In this paper, we assessed the performance of a significance test, based on a deviance-based pseudo-R^2 (R^2_{KLD}), for exponential and power-law models of distance-decay of similarity. Our results show that the R^2_{KLD} significance test performs well for nonlinear distance-decay patterns, using either site-block permutations or site-block resampling to estimate the distribution of R^2_{KLD} under the null hypothesis. In other words, both block-site permutations and block-site resampling accounted for the pairwise dependence of the similarity and distance values, thus overcoming the inflation of the type I error that would arise if standard significance tests were used. We thus provide a method with several major advantages for the study of distance-decay patterns: (i) it provides a measure of goodness-of-fit, the pseudo-R^2 statistic, which is easily interpretable as the amount of variation in community similarity explained by spatial distance (or alternative predictors with pairwise structure, e.g. climatic distance); (ii) it provides a robust estimate of statistical significance, with good type I error rate and statistical power; and (iii) it allows the implementation of Generalized Linear Models to fit the distance-decay relationship, with all the advantages derived from GLMs, such as model-comparison-based AIC or the estimation of model parameters with biological relevance (i.e. the slope of the distance-decay) that can be compared between different regions or taxonomic groups.

We have found that the R^2_{KLD} significance test performed well under absence and presence of distance-decay, independently of the calibration strategy used to estimate its null hypothesis (either permutations or site-block resampling). That is, type I error was low and statistical power was high, either with permutations or site-block resampling strategies. The Mantel test performed well to detect distance-decay patterns when the decay pattern was clearly

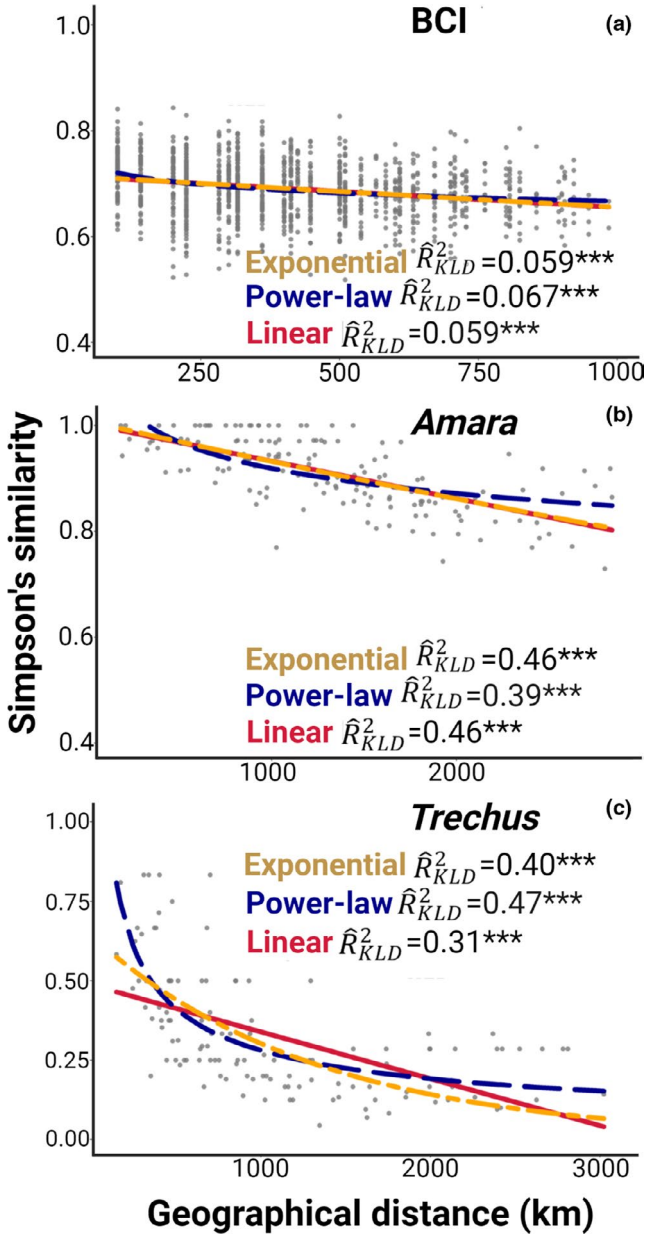

FIGURE 5 Empirical distance-decay models for (a) Barro Colorado Island's trees (BCI), (b) *Trechus* beetles in southern Europe and (c) *Amara* beetles in northern Europe. Pseudo-R^2 values are provided for different models (i.e. negative exponential, power-law and linear) *p*-value. The R^2_{KLD} significance test was applied to exponential and power-law models and the Mantel test to linear models. *** = <0.001

present across the full range of distance and could be approximately accommodated by a linear model. However, the Mantel test suffered from low statistical power when the decay pattern was curved and less marked (data simulated with $\varphi = 300$). It should be stressed that we know a priori that linear models are not appropriate for similarity/dissimilarity indices, as these are bounded between zero and one, leading to nonlinear, asymptotic relationships with spatial or environmental distances. As shown by Diniz-Filho et al. (2013) and Zeller et al. (2016), the Mantel test can fail to detect a distance-decay relationship when this relationship is far from linear. This issue with the Mantel test is exemplified by the distance-decay scenario simulated with the weakest spatial autocorrelation ($\varphi = 300$). In this scenario, similarity decays steadily at very short distances and reaches the lower asymptote fast (i.e. the decay pattern is strongly nonlinear). As expected, the power-law model R^2_{KLD} detects more efficiently the distance-decay pattern (lower type II error) than the exponential R^2_{KLD} or the linear model (Mantel test) since the power-law function can accommodate steep initial distance-decay. The potential misfit of the Mantel test as a significance test when modelling highly nonlinear patterns can also be observed in empirical data. For example, in the most pronounced and less linear of the real distance-decay data examples (*Trechus*), we found differences between the models' fit with the power-law model having the best value of explained deviance. These two results showcase the relevance of introducing in the community ecology/macroecology toolkit a significance test for nonlinear distance-decay models. Indeed, testing the significance of a linear and nonlinear model is not equivalent since they fit the data differently. As a result, the use of a linear significance test for a nonlinear model could lead to not detecting the distance-decay pattern that we are really modelling.

The methods proposed here are subject of some limitations. First, we would like to explicitly acknowledge that a pseudo-R^2 value does not substitute the necessary diagnostic checking of the model's assumptions (such as homoscedasticity or normally distributed residuals). Its use as a measure of explained variation must be conditioned on these assumptions being statistically validated, or the model being adequate for the data of interest and the research question by other means (e.g. in Nekola & McGill, 2014), as advised in the work of Nakagawa and Schielzeth (2013). Additionally, the pseudo-R^2 statistic might be overly optimistic for small sample sizes as an estimation of explained variation and its value should be corrected accordingly (Heinzl et al., 2005).

As a second limitation, the spatial structure of non-spatial distances (i.e. climatic distances) might pose additional problems. The use of R^2_{KLD} combined with site permutations or site-block resampling as a significance test could be applied to other types of nonlinear distance-decay models, in which the predictors are not spatial distances but distances between climatic variables, time, etc (e.g. Astorga et al., 2012 and Saito et al., 2015, with environmental distances; Matsuoka et al., 2016, with temporal distances). However, the use of Mantel tests with structured distance variables (e.g. a spatially structured climatic variable) has been criticized by various authors because neither the original Mantel nor the partial Mantel test (Smouse et al., 1986) are able to overcome the effect of this structure (Guillot & Rousset, 2013; Legendre et al., 2015; Legendre & Fortin, 2010). When applied to structured predictor variables other than spatial distances, R^2_{KLD} will very likely present a similar bias. A possible solution for spatially structured variables could be a modification of the test following the work of Crabot et al. (2019) with the Mantel test. These authors proposed a method based on spatially

constrained randomizations to correct the inflation of its type I error due to the spatial autocorrelation of the variables. Furthermore, the site-block resampling that we describe in this work could be useful for resolving questions about distance-decay models that may be affected by the paired structure of the data. Some examples of this application would be estimating the variance of regression parameters for model comparison tests or addressing AIC and uncertainty in model selection while considering the real structure of distance-decay data (Fieberg et al., 2020).

In conclusion, our results show the adequacy of using a deviance-based pseudo-R^2, R^2_{KLD}, in combination with site permutations or site-block resampling as a significance test and a measure of explained variation for exponential and power-law distance-decay models. The methodology we propose here constitutes an accurate nonlinear option that should be preferred over the commonly used linear or nonparametric Mantel test. Additionally, both the R^2_{KLD} statistic and the proposed site-block resampling method have the potential to be adapted for different questions and approaches to the study of nonlinear distance-decay patterns, such as assessing the effect of non-spatial factors (e.g. climatic variables) on community similarity or estimating the variance of model parameters (e.g. the rate at which similarity decays with distance).

ACKNOWLEDGEMENTS

We thank Xavier Picó, Christine Meynard, Simone Fattorini and two anonymous reviewers for insightful comments and suggestions to earlier versions of the manuscript. This research was supported by the Spanish Ministry of Science and Innovation and the European Regional Development Fund (ERDF) through grants CGL2016-76637-P and PID2020-112935GB-I00, PID2020-116587GB-I00 and a FPI scholarship to S.M.-S (BES-2017-081643). R.M.D. was supported by the Spanish Ministry of Education through a FPU scholarship (Ref.: FPU17/03016; Ministry of Education). No permits were needed to carry out our research.

CONFLICT OF INTEREST

All authors declare that they have no conflict of interest for this work.

DATA AVAILABILITY STATEMENT

The code for data simulation, functions for significance tests based on permutation and site-block resampling, and code for assessing Type I error and statistical power of the tests is available in File S1 in Supporting Information.

ORCID

Sara Martínez-Santalla https://orcid.org/0000-0002-1190-7332
Ramiro Martín-Devasa https://orcid.org/0000-0002-3890-5003
Carola Gómez-Rodríguez https://orcid.org/0000-0002-2019-7176
Rosa M. Crujeiras https://orcid.org/0000-0002-3907-8951
Andrés Baselga https://orcid.org/0000-0001-7914-7109

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?: Decay of similarity in freshwater communities. *Global Ecology and Biogeography*, *21*(3), 365–375. https://doi.org/10.1111/j.1466-8238.2011.00681.x

Baselga, A., & Orme, C. D. L. (2012). Betapart: An R package for the study of beta diversity: *Betapart package*. *Methods in Ecology and Evolution*, *3*(5), 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Baselga, A., Orme, C. D. L., Villéger, S., De Bortoli, J., Leprieur, F., & Logez, M. (2021). Betapart: Partitioning beta diversity into turnover and nestedness components. R package version 1.5.4. http://CRAN.R-project.org/package=betapart

Cameron, A. C., & Windmeijer, F. A. G. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics*, *77*(2), 329–342. https://doi.org/10.1016/S0304-4076(96)01818-0

Condit, R., Pitman, N., Leigh, E. G., Chave, J., Terborgh, J., Foster, R. B., … Hubbell, S. P. (2002). Beta-diversity in tropical Forest trees. *Science*, *295*(5555), 666–669. https://doi.org/10.1126/science.1066854

Crabot, J., Clappe, S., Dray, S., & Datry, T. (2019). Testing the Mantel statistic with a spatially-constrained permutation procedure. *Methods in Ecology and Evolution*, *10*(4), 532–540. https://doi.org/10.1111/2041-210X.13141

Dietz, E. J. (1983). Permutation tests for association between two distance matrices. *Systematic Zoology*, *32*, 21–26.

Diniz-Filho, J. A. F., Soares, T. N., Lima, J. S., Dobrovolski, R., Landeiro, V. L., Telles, M. P. d. C., Rangel, T. F., & Bini, L. M. (2013). Mantel test in population genetics. *Genetics and Molecular Biology*, *36*(4), 475–485. https://doi.org/10.1590/S1415-47572013000400002

Draper, F. C., Baraloto, C., Brodrick, P. G., Phillips, O. L., Martinez, R. V., Honorio Coronado, E. N., Baker, T. R., Zárate Gómez, R., & Amasifuen Guerra, C. A., & Flores, M. (2019). Imaging spectroscopy predicts variable distance decay across contrasting Amazonian tree communities. *Journal of Ecology*, *107*(2), 696–710. https://doi.org/10.1111/1365-2745.13067

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. *Diversity and Distributions*, *13*(3), 252–264. https://doi.org/10.1111/j.1472-4642.2007.00341.x

Fieberg, J. R., Vitense, K., & Johnson, D. H. (2020). Resampling-based methods for biologists. *PeerJ*, *8*, e9089. https://doi.org/10.7717/peerj.9089

Gaston, K. J. (1996). Species-range-size distributions: Patterns, mechanisms and implications. *Trends in Ecology & Evolution*, *11*(5), 197–201. https://doi.org/10.1016/0169-5347(96)10027-6

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

Guillot, G., & Rousset, F. (2013). Dismantling the Mantel tests. *Methods in Ecology and Evolution*, *4*(4), 336–344. https://doi.org/10.1111/2041-210x.12018

Hastie, T. J., & Pregibon, D. (1993). Generalized linear models. In J. M. Chambers & T. J. Hastie (Eds.), *Statistical models in S* (pp. 195–249). Chapman & Hall.

Heinzl, H., Waldhör, T., & Mittlböck, M. (2005). Careful use of pseudoR-squared measures in epidemiological studies. *Statistics in Medicine*, *24*(18), 2867–2872. https://doi.org/10.1002/sim.2168

Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, *54*(2), 187–211. https://doi.org/10.2307/1942661

Legendre, P., Fortin, M., & Borcard, D. (2015). Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, *6*(11), 1239–1247. https://doi.org/10.1111/2041-210X.12425

Legendre, P., & Fortin, M.-J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, *10*(5), 831–844. https://doi.org/10.1111/j.1755-0998.2010.02866.x

Legendre, P., Lapointe, F.-J., & Casgrain, P. (1994). Modeling brain evolution from behavior: A permutational regression approach. *Evolution*, *48*(5), 1487–1499. https://doi.org/10.1111/j.1558-5646.1994.tb02191.x

Lichstein, J. W. (2007). Multiple regression on distance matrices: A multivariate spatial analysis tool. *Plant Ecology*, *188*, 117–131.

Löbl, I., & Smetana, A. (2003). *Catalogue of Palaearctic coleoptera. Vol. 1: Archostemata-Myxophaga-Adephaga*. Apollo Books.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2 Part 1), 209–220.

Marschner, I. C. (2011). glm2: Fitting generalized linear models with convergence problems. R package version 1.2.1. https://cran.r-project.org/package=glm2

Matsuoka, S., Kawaguchi, E., & Osono, T. (2016). Temporal distance decay of similarity of ectomycorrhizal fungal community composition in a subtropical evergreen forest in Japan. *FEMS Microbiology Ecology*, *92*(5), fiw061. https://doi.org/10.1093/femsec/fiw061

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models. Monographs on statistics and applied probability 37* (2nd ed.). Chapman & Hall.

Millar, R. B., Anderson, M. J., & Tolimieri, N. (2011). Much ado about nothings: Using zero similarity points in distance-decay curves. *Ecology*, *92*(9), 1717–1722. https://doi.org/10.1890/11-0029.1

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.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. *Journal of Biogeography*, *26*(4), 867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x

O'Hara, R., & Kotze, J. (2010). Do not log-transform count data. *Nature Precedings*. https://www.nature.com/articles/npre.2010.4136.1#citeas

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, P., Solymos, G. L., Stevens, M. H. H., Szoecs, E. & Wagner, H. (2019). Vegan: Community ecology package. R package version 2.5-4. https://CRAN.R-project.org/package=vegan

Peguero, G., Ferrín, M., Sardans, J., Verbruggen, E., Ramírez-Rojas, I., Van Langenhove, L., Verryckt, L. T., Murienne, J., Iribar, A., Zinger, L., Grau, O., Orivel, L., Stahl, C., Courtois, E. A., Asensio, D., Gargallo-Garriga, A., Llusià, J., Margalef, O., Ogaya, R., … Peñuelas, J. (2021). Decay of similarity across tropical forest communities: Integrating spatial distance with soil nutrients. *Ecology*, e03599. https://doi.org/10.1002/ecy.3599

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ribeiro Jr, P. J., Diggle, P. J., Christensen, O., Schlather, M., Bivand, R. & Ripley, B. (2020). geoR: Analysis of geostatistical data. R package version 1.7-5.2.2. https://CRAN.R-project.org/package=geoR

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. *Journal of Biogeography*, *42*(11), 2101–2111. https://doi.org/10.1111/jbi.12577

Smouse, P. E., Long, J. C., & Sokal, R. R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, *35*(4), 627. https://doi.org/10.2307/2413122

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

Tornero, I., Boix, D., Bagella, S., Pinto-Cruz, C., Caria, M. C., Belo, A., Lumbreras, A., Sala, J., Compte, J., & Gascón, S. (2018). Dispersal mode and spatial extent influence distance-decay patterns in pond metacommunities. *PLoS ONE*, *13*(8), e0203119. https://doi.org/10.1371/journal.pone.0203119

Tuomisto, H., Ruokolainen, K., & Yli-Halla, M. (2003). Dispersal, environment, and floristic variation of Western Amazonian forests. *Science*, *299*(5604), 241–244. https://doi.org/10.1126/science.1078037

Yang, Q., Weigelt, P., Fristoe, T. S., Zhang, Z., Kreft, H., Stein, A., van Seebens, H., Dawson, W., Essl, F., König, C., Lenzner, B., Pergl, J., Pouteau, R., Pyšek, P., Winter, M., Ebel, A. L., Fuentes, N., Giehl, E. L. H., Kartesz, J., … van Kleunen, M. (2021). The global loss of floristic uniqueness. *Nature Communications*, *12*(1), 7290. https://doi.org/10.1038/s41467-021-27603-y

Zeller, K. A., Creech, T. G., Millette, K. L., Crowhurst, R. S., Long, R. A., Wagner, H. H., Balkenhol, N., & Landguth, E. L. (2016). Using simulations to evaluate Mantel-based methods for assessing landscape resistance to gene flow. *Ecology and Evolution*, *6*(12), 4115–4128. https://doi.org/10.1002/ece3.2154

Zhang, W., Lei, M., Li, Y., Wang, P., Wang, C., Gao, Y., Wu, H., Xu, C., Niu, L., Wang, L., & Zhang, H. (2019). Determination of vertical and horizontal assemblage drivers of bacterial community in a heavily polluted urban river. *Water Research*, *161*, 98–107. https://doi.org/10.1016/j.watres.2019.05.107

BIOSKETCH

The Biogeography Lab (SM-S, RM-D, CG-R and AB, at BiBiCI, University of Santiago de Compostela) studies macroecological patterns to contribute to the understanding of the processes driving biodiversity.

Authors' contributions: All authors conceived the ideas and the simulation design. S.M.-S wrote the code, analysed the data and led the writing of the manuscript. All authors revised and critically contributed to the drafts and gave final approval for publication.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: 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. *Journal of Biogeography*, 49, 968–978. https://doi.org/10.1111/jbi.14351