

Simulated neutral metacommunities support the habitat amount hypothesis

F. Laroche^{1*}, M. Balbi¹, T. Grébert², F. Jabot² & F. Archaux¹

¹Irstea, UR EFNO, Domaine des Barres, F-45290 Nogent-sur-Vernisson, France

²Université Clermont Auvergne, Irstea, UR LISC, Centre de Clermont-Ferrand, F-63178 Aubière, France

*corresponding author: fabien.laroche@irstea.fr

Abstract

The Theory of Island Biogeography yielded the important idea that species richness within sites should depend on site connectivity, i.e. its connection with surrounding potential sources of immigrants. However, the effect of connectivity on species presence and community richness in empirical studies is often quite limited, partly because classic indices used to quantify connectivity (e.g. distance to nearest habitat patch) need improvement. Current methodological advances for quantifying connectivity lie along three directions: building better indices that more carefully trace potential fluxes of individual between habitat units, combining such indices and searching for the most appropriate way of describing habitat spatial distribution, using patches or cells as elementary units. Here we assessed the potential of 63 contemporary connectivity indices that explore these different tracks by applying them to virtual habitat maps with contrasted habitat amount and configuration coupled with a neutral metacommunity model. We found that choosing buffers indices (and more generally flux indices computed using habitat cells) is the most fruitful methodological choice to improve the explanation of species richness. Refining the scaling of flux indices could bring an additional improvement although quite limited. Combining indices of different type (connector and flux) also marginally improved the ability to explain species richness. In line with the habitat amount hypothesis, our results suggest that buffer connectivity indices are a simple and robust approach that may prove sufficient for capturing connectivity effects in many contexts, potentially leading to very strong effect sizes upon community richness.

Key-words

Landscape ecology; Structural connectivity; Virtual ecologist; Neutral landscapes; Dispersal; Diversity patterns

Introduction

In community ecology, the Niche Theory [1]–[3] predicts that local environmental conditions and species interactions are the dominant drivers of species composition within sites. It has led to practical tools such as species distribution models, which are remarkably efficient for understanding the distribution of species in space, especially at coarse spatial grain [4]. However, since the Theory of Island Biogeography [5], it is commonly acknowledged that species presence in strongly isolated communities such as islands also depends on their ability to colonize sites through immigration and to avoid random extinction thanks to large population sizes. Area and geographic isolation of sites could thus be additional major drivers of community composition, especially in fragmented habitats where habitat patches behave like islands within an archipelago. TIB principles fostered an important body of empirical research in a wide array of ecosystems about how the surrounding availability and spatial configuration of habitats affect community composition. In particular, the geographic isolation of patches has now been developed into the concept of “patch structural connectivity”, which quantifies the potential flux of immigrants coming into a focal patch from surrounding habitats based on habitat mapping [6]. However, reviews about the effect of the area and the structural connectivity of patches on species presence and community composition in fragmented landscapes suggest that the effects of local environmental conditions within a patch occur more frequently and with stronger magnitudes than the effects of area and connectivity on species presence [7]–[9]. Connectivity and area of patches tend to have weak absolute effect sizes [8] and their respective contributions are sometimes hard to disentangle [10].

Those limited effects are partly due to several limitations of classic patch structural connectivity indices. First, many empirical studies used the distance to the nearest or the few nearest patches as a patch structural connectivity index, whereas [11] and [12] showed that it is a poor predictor of species presence compared to indices that more extensively account for surrounding patches, such as buffers or indices adapted from the incidence function model [13]. Second, patch structural connectivity indices based on the distance between the focal patch and the surrounding habitat (which we call “flux indices” below) do not account for exchanges of individuals among the surrounding patches themselves, which potentially modulate the occupancy of species in the surrounding habitat and thus affect immigration chances within the focal patch. The use of graph theory [14] opens the way for addressing such limitations. In particular, patch-based “connector” indices [15] quantify to what extent a given patch contributes to the connection among the other patches in the landscape. For a given amount of habitat around the focal patch, a lower connector score of the focal patch may therefore indicate that surrounding patches depend less on the focal patch to connect one with another, i.e. that direct fluxes among surrounding patches are stronger. [16] further showed

that connector indices convey complementary information about landscape structure around focal patches compared to flux indices, which suggests that combining them with flux indices to explain species presence and community composition may increase explanatory power without generating confounding effects. Third, describing landscapes as sets of well-delineated patches of habitats embedded in a matrix of non-habitat is often too simplistic, because habitat in real landscape may occur in a more continuous and diffuse manner in space, e.g. following ecological gradients, without forming discrete spatial entities. In particular, [17] challenged the patch-based methodological framework by suggesting that the concept of patch itself, and the subsequent dichotomy between patch area and isolation, may be misleading. Indeed, species presence at some positions in space may rather come from the total amount of habitat in some surrounding area, irrespective of whether this habitat belongs to the focal patch (patch area) or not (structural connectivity). This “habitat amount hypothesis” suggests considering the habitat spatial distribution as a “raster” of cells rather than a “vector map” of patches (using [14] terminology). Consequently, the effect of connectivity on community composition should rather be studied at the finer habitat cell level rather than at the patch level [18].

In a nutshell, studying how structural connectivity affect species presence and community composition within a given location may benefit from methodological advances in several non-excluding directions: (i) optimizing flux connectivity indices, (ii) combining flux and connector connectivity indices and (iii) changing the grain of landscape analysis from habitat patch to habitat cell. Virtual datasets stemming from spatially explicit metacommunity models constitute an ideal context to assess such advances for they offer perfect control of the spatial distribution of habitat and the ecological features of species. By doing so, they eliminate potential observational biases and they allow discussing best strategies as a function of landscape and species features. Such virtual approaches have already been used to study the optimization of indices. For instance, [19] have used neutral metacommunity simulations to show that changing the scaling of patch connectivity indices coming from graph theory (from capturing only the immediate neighborhoods to capturing the whole landscape composition) can change the effect size of connectivity on patch species diversity. They further showed that the higher the dispersal ability of species, the larger the scaling of indices should be. Similarly, a metapopulation simulation study [20] showed that there exists some optimal buffer radius that maximizes the effect size of connectivity upon local species presence. They further suggested that this optimal size, called the “scale of effect” should lie between four and nine times the average dispersal distance of the target species.

Here, we extend these previous studies based on virtual metacommunity models towards exploring the three axes of methodological advances identified earlier. We refine previous results about how optimizing flux connectivity indices modulate the effect size of connectivity

on local species richness by considering a broad set of indices with contrasted scaling and heterogeneous theoretical backgrounds in a single study. We also explore what benefit can be obtained from combining connectivity indices of different kinds compared to using a single connectivity index. At last, our study compare indices based on patches as elementary habitat unit and indices based on habitat cells.

Materials and methods

Landscape generation - We considered binary landscapes made of suitable habitat pixels and inhospitable matrix pixels. We generated virtual landscapes composed of 100×100 cells using a midpoint-displacement algorithm [21] which allowed us covering different levels of habitat quantity and fragmentation. The proportion of habitat cells varied according to three modalities (10%, 20% of 40% of the landscapes). The spatial aggregation of habitat cells varied independently, and was controlled by the Hurst exponent (0.1, 0.5, and 0.9 in increasing order of aggregation; see Fig. S1 for examples). Ten replicates for each of these nine landscape types were generated, resulting in 90 landscapes. Higher values of the Hurst exponent for a given value of habitat proportion increased habitat patch size and decreased the number of patches (Fig. S2). Higher habitat proportion for a constant Hurst coefficient value resulted in higher mean habitat patch size.

Neutral metacommunity simulations - We simulated spatially explicit neutral metacommunities on virtual heterogeneous landscapes. We resorted to using a spatially explicit neutral model of metacommunities, where all species have the same dispersal distance. We used a discrete-time model where the metacommunity changes by steps. All habitat cells were occupied, and community dynamics in each habitat cell followed a zero-sum game, so that habitat cells always harbored 100 individuals at the beginning of a step. One step was made of two consecutive events. Event 1: 10% of individuals die in each cell – they are picked at random. Event 2: dead individuals are replaced by the same number of recruited individuals that are randomly drawn from a multinomial distribution, each species having a weight equal to $0.01 \times \chi_i + \sum_k A_{ik} \exp(-d_{kf} / \lambda_s)$ where χ_i is the relative abundance of species i in the regional pool, A_{ik} is the local abundance of species i in habitat cell k , d_{kf} is the Euclidean distance (in cell unit) between the focal habitat cell f and the source habitat cell k , λ_s represents the average dispersal distance of individuals in the metacommunity and the sum is over all habitat cells k of the landscape. The regional pool was an infinite pool of migrants representing biodiversity at larger spatial scales than the focal landscape, it contained 100 species, the relative abundances of which were sampled once for all at the beginning of the simulation in a Dirichlet distribution with concentration parameters α_i equal to 1 (with i from 1 to 100). Metacommunity were simulated forward in time, with 1000 burn-in steps and 500 steps between each

replicates. Simulation was structured as a torus to remove unwanted border effects in metacommunity dynamics. Metacommunities were simulated with three average dispersal values (“true” simulated average dispersal $\lambda_s = 0.25, 0.5$ and 1 cell). We performed 10 replicates for each dispersal value and in each landscape. This constituted a total of 2700 metacommunity simulations. For each metacommunity simulation, species richness was computed at the cell level with R [22].

Local connectivity indices – For a given habitat cell, we considered 17 types of structural connectivity indices (CIs), summarized in Table 1. Six CIs were computed at the cell grain, i.e. considering a cell as the elementary unit of habitat, while 11 were computed at the patch grain, i.e. considering patches (maximal sets of contiguous habitat cells) as the elementary unit of habitat. In the latter case, all habitat cells within a patch harbored the same value of CI. Six CIs corresponded to measures of “flux” (red in Table 1), i.e. they quantified the intensity of the connection between the focal cell or patch and the surrounding habitat using a weight function decreasing with distance. Note that we considered buffer indices as measures of flux, since they quantify the connection between the focal cell or patch and the surrounding habitat using a step weight function decreasing with distance. Four other CIs were connector indices, i.e. they quantified how the focal cell or patch contributed to create connections among other habitat units (green in Table 1). Two other CIs were related to the area of the focal patch (blue in Table 1). At last, five additional CIs were composite, simultaneously quantifying area, flux and connector features in a single index (white in Table 1).

Buffer indices corresponded to the proportion of habitat cells within circles of different radius ($r_{buf} = 1, 2, 4, 5, 8$ cells) around the focal cell. Graph CIs (dH, dF, dIIC and its components, dPC and its components) were based on nodes corresponding either to cells or to patches, although some of them were computationally too demanding at the cell grain and were thus only considered at the patch grain (see Table S1 for the complete list of indices considered). Pairs of nodes were connected to each other by links. Links’ weights w_{ij} between cells i and j in the network decreased according to the formula $\exp(-d_{ij}/\lambda_c)$, where d_{ij} is the Euclidean distances between cells i and j and λ_c is a scale parameter [14], [23]. λ_c may be interpreted as the hypothesized average dispersal distance of target organisms in the landscape (which may differ from the “true” simulated average dispersal distance due to habitat fragmentation). We considered four scale parameter values ($\lambda_c = 0.25, 0.5, 1$ and 2 cells). Some graph CIs (dH, dIIC and its components) considered a binary graph, where each cell pair was considered either connected (1) or not (0) relatively to a minimal link weight ($w_{min} = 0.005$). All indices with a name starting by “d” were computed using a cell or patch removal procedure, depending on the grain (Table 1). All indices were computed with Conefor 2.7 (command line version for

Linux, furnished by S. Saura, soon publicly available on www.conefor.org; [24]). Altogether, we computed 63 distinct CIs in each sampled cell of each simulation.

Table 1: Connectivity indices (CIs) description at patch (P) and cell grain (C). Red: flux indices; blue: area indices; green: connector indices; white: composite indices.

CI	Definition	Ref.
Buffer	P: undefined; C: $\text{buf}_k = \frac{a}{\pi r^2} \sum_{i=1}^n 1_{d_{ik} \leq r}$	[11]
Harary	P or C: $\text{dH}_k = \frac{100}{H} \left[\sum_{i \neq k}^n \frac{1}{n_{ik}} + \sum_{i \neq k}^{n-1} \sum_{j=i+1}^n \frac{(n_{ij}^{(-k)} - n_{ij})}{n_{ij} n_{ij}^{(-k)}} \right]$	[25]
Integral index of connectivity	P: $\text{dIIC}_k = \frac{100}{IIC} \left[a_k^2 + 2 \sum_{i \neq k}^n \frac{a_k a_i}{1 + n_{ik}} + 2 \sum_{i \neq k}^n \sum_{j=i+1}^n \frac{a_i a_j (n_{ij}^{(-k)} - n_{ij})}{(1 + n_{ij})(1 + n_{ij}^{(-k)})} \right]$ C: $\text{dIIC}_k = \frac{100 a^2}{IIC} \left[1 + 2 \sum_{i \neq k}^n \frac{1}{1 + n_{ik}} + 2 \sum_{i \neq k}^n \sum_{j=i+1}^n \frac{(n_{ij}^{(-k)} - n_{ij})}{(1 + n_{ij})(1 + n_{ij}^{(-k)})} \right]$	[26]
dIIC intra-patch	P: $\text{dIICintra}_k = \frac{100}{IIC} a_k^2$; C: constant	[15]
dIIC connector	P: $\text{dIICconn}_k = \frac{100}{IIC} \left[2 \sum_{i \neq k}^n \sum_{j=i+1}^n \frac{a_i a_j (n_{ij}^{(-k)} - n_{ij})}{(1 + n_{ij})(1 + n_{ij}^{(-k)})} \right]$ C: $\text{dIICconn}_k = \frac{100 a^2}{IIC} \left[2 \sum_{i \neq k}^n \sum_{j=i+1}^n \frac{(n_{ij}^{(-k)} - n_{ij})}{(1 + n_{ij})(1 + n_{ij}^{(-k)})} \right]$	[15]
dIIC flux	P: $\text{dIICflux}_k = \frac{100}{IIC} \left[2 \sum_{i \neq k}^n \frac{a_k a_i}{1 + n_{ik}} \right]$; C: $\text{dIICflux}_k = \frac{100 a^2}{IIC} \left[2 \sum_{i \neq k}^n \frac{1}{1 + n_{ik}} \right]$	[15]
Probability of connectivity	P: $\text{dPC}_k = \frac{100}{PC} \left[a_k^2 + 2 \sum_{i \neq k}^n a_k a_i p_{ik} + 2 \sum_{i \neq k}^n \sum_{j=i+1}^n a_i a_j (p_{ij} - p_{ij}^{(-k)}) \right]$ C: computationally too demanding	[27]
dPC intra-patch	Not reported, because nearly identical to dIICintra_k	[15]
dPC connector	P: $\text{dPCconn}_k = \frac{100}{PC} \left[2 \sum_{i \neq k}^n \sum_{j=i+1}^n a_i a_j (p_{ij} - p_{ij}^{(-k)}) \right]$; C: computationally too demanding	[15]
dPC flux	P: $\text{dPCflux}_k = \frac{100}{PC} \left[2 \sum_{i \neq k}^n a_k a_i p_{ik} \right]$; C: computationally too demanding	[15]
Flux	P or C: $\text{dF}_k = 2 \sum_{i=1}^n w_{ik}$	[14],[21]
Betweenness centrality	P: $\text{BC}_k = \sum_{i \neq k}^n \sum_{j \neq k}^n \frac{nsp_{ij}(k)}{nsp_{ij}}$; C: numerically unstable	[28]
Area	P: $\text{dA}_k = a_k$; C: constant	—

Notations: n : total number of nodes (patches or cells) in a graph; a : area of a cell; a_i : area of patch i ; r : radius of a buffer; n_{ij} : shortest path between nodes i and j in a binary graph; $n_{ij}^{(-k)}$: shortest path between nodes i and j in a binary graph after node k has been removed; $H = \frac{1}{2} \sum_{i=1}^n \sum_{j=1, j \neq i}^n 1/n_{ij}$: Harary index of a graph; $IIC = \sum_{i=1}^n \sum_{j=1}^n a_i a_j / (1 + n_{ij})$: integral index of connectivity of a graph; d_{ij} : Euclidean distance between nodes i and j ; w_{ij} : probability weight of the link between nodes i and j in a weighted graph; p_{ij} : maximal product of links' probability weight across all paths connecting nodes i and j in a weighted graph; $PC = \sum_{i=1}^n \sum_{j=1}^n a_i a_j p_{ij}$: probability of connectivity index of a graph; nsp_{ij} : number of shortest paths between nodes i and j in a binary graph; $nsp_{ij}(k)$: number of shortest paths between nodes i and j in a binary graph that contain node k .

We analyzed correlations among these indices using ascending hierarchical classification. Within each of the 90 simulated landscapes, we computed the values of the 63 indices in all habitat cells, which yielded 63 vectors of length 1000 to 4000 depending on the habitat proportion. We scaled the 63 vectors to mean 0 and variance 1, divided them by the square root of the number of habitat cells in the landscapes and computed pairwise Euclidean distances among them. We thus obtained one 63×63 distance matrix among CIs in each of the 90 landscapes. Note that with our convention, the Euclidean distance between two CIs corresponds to $\sqrt{2 - 2r}$, where r is the Pearson correlation between the two CIs across all habitat cells of the considered landscapes. We then averaged the 90 distance matrices to obtain one single 63×63 distance matrix as a basis for classification. We built two ascending non-supervised classification (*hclust* function of R *base* package), using the *complete* and *single* methods for group merging respectively. A monophyletic group G with common ancestor located at value r has a different meaning depending on the method used. According to the *complete* method, it means that any pair of CIs within G has a correlation above r . According to the *single* method, it means that any CI within G has a correlation with CIs outside G lower than r . We delineated minimum monophyletic groups with common ancestor at a correlation value equal or below 0.7 according to the two classification methods simultaneously. These groups are therefore set of indices such that any two CIs in the set have correlation above 0.7 and any CI in the set has correlation with CIs outside the set below 0.7.

Statistical analysis - We considered communities in habitat cells away from each other's for a minimal distance of 12 cells, to reduce spatial auto-correlation. We also reduced potential landscape border effect (that could decouple landscape indices and actual migrants received) by excluding cells near landscape borders (to a distance inferior or equal to eight cells, equivalent to the longest buffer radius). Each landscape counted in average 25 sampled cells (CI-95% = [23, 27]). For each CI, each landscape and each metacommunity replicate, we computed the Pearson correlation coefficient r between CIs and species richness in sampled cells, which yielded $2700 \times 63 = 170100$ r values. We analyzed how CIs type, scaling of CIs, habitat proportion, habitat aggregation and community dispersal affected the value of r using linear models (*lm* function in R *base* package), considering all the covariates as categorical variables.

We tested whether combining a connector index (dIICconnector) to a cell-based dF index improved the model of species richness. In each simulation, we first selected the dF index that had the most significant effect upon species richness in a linear model. Then we tested whether some dIICconnector at the cell grain could have a significant additional effect using a likelihood ratio test. If so, we measured the relative increase of the model R^2 induced by adding the selected dIIC connector index. We also analyzed how habitat proportion, habitat aggregation

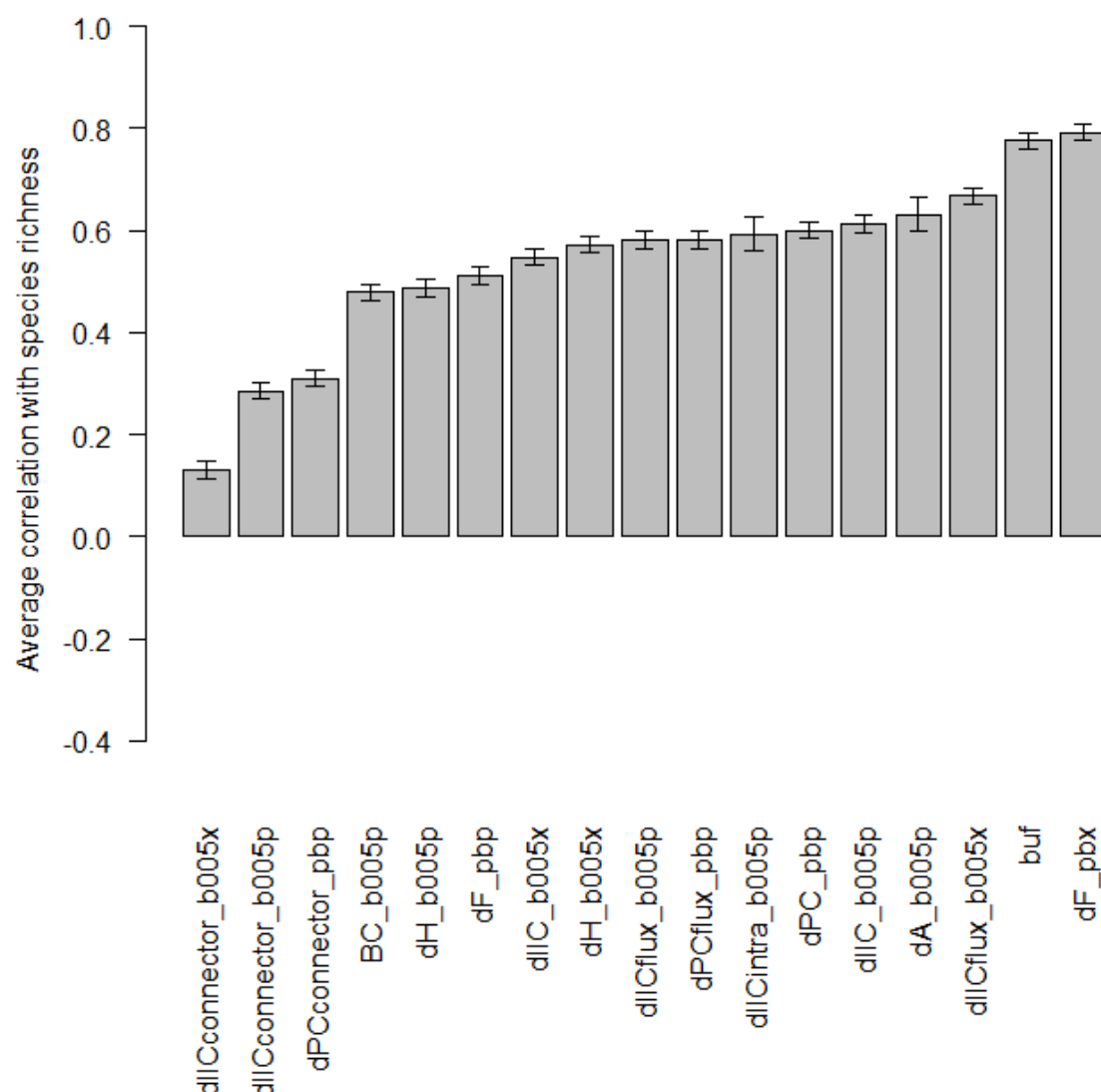
and community dispersal modulated the potential gain of combining dF and dIIConnector indices.

Results

Correlation among CIs - Our preliminary analysis of pairwise correlations among CIs revealed that average pairwise correlation were globally high with many extreme values towards 1 (median = 0.49; CI-95% = [-0.0353, 0.9998]). Classification of indices based on the *complete* method (searching for correlation within groups above 0.7; Fig. 1) clearly distinguished clusters corresponding to patch and cell-based indices respectively. Among patch-based indices, those that quantified the connector fraction of connectivity formed one cluster (plus two isolated singletons), while the rest of patch based indices formed another cluster. By contrast, a more complex structure emerged among cell-based indices where scale, edges handling, and focal component of connectivity (flux versus connector fraction) all mattered in the delineation of clusters. Classification based on the *single* method (searching for correlation among groups below 0.7; Fig. S3) was much simpler: nearly all the indices fell into the same cluster, irrespective of grain, type and scale, suggesting that most of the clusters evidenced in the first classification partially overlap one with another and cannot be treated as independent groups without risk. Only cell-based binary connector CIs stood out as having profiles distinct from the vast majority of indices. The combination of our two classification hence suggested that the only combination of CIs showing a robust pattern of complementarity across landscapes involves cell-based binary connector indices (potentially several scale at a time) and one index chosen among all the others.

Correlation between CIs and species richness - Averaging over all landscape configuration, all community dispersal levels, all scaling parameter values and all replicates, every CI showed a significantly positive correlation with species richness (Fig. 2). Connector indices at patch and cell grain showed a significantly lower correlation than other CIs. Buffers and dF at cell grain clearly stepped out as yielding the highest correlation values on average (average correlation with species richness above 0.7; Fig. 2).

Figure 2: Average correlation between local connectivity indices (CIs) and species richness. For each CI, we reported the average correlation across all landscape types, all landscape replicates and all community replicates (grey bars). Error bars correspond to estimated 95% confidence interval of the mean, computed as 1.96 times the standard error of the mean.



The effect of dF (computed at cell grain) and buffer indices on species richness depended on a triple interaction among community dispersal, habitat proportion and habitat aggregation and on a double interaction between index scaling and community dispersal. However, for the sake of interpretation, we studied how the effect of dF and buffer on species richness varied according to habitat proportion, to habitat aggregation, to community dispersal and to the interaction between index scaling and community dispersal separately. The effects of buffers and dF at cell grain on species richness were maximized for intermediary community dispersal level and increased with the habitat proportion in the landscape (Fig. 3). In addition, the effect of habitat aggregation on the correlation with species richness was unimodal for buffers and decreasing for dF. The scale parameter values that maximized cell-based dF and buffer

correlations with species richness increased with community dispersal level (Table 2). The optimal radius (r_{buf}) for buffers was about 8 times the community dispersal level (λ_s), while the optimal scale parameter of dF at cell grain (λ_c) was about 2 times the community dispersal level (λ_s).

Figure 3: Community dispersal and habitat proportion effects on the correlation between buffers (resp. dF at cell grain) and species richness. Panels A and B (resp. C and D) correspond to buffer indices (resp. dF at cell grain). Panels A and C (resp. B and D) show the effect of community dispersal (resp. habitat proportion). Grey bars indicate the average correlation across all the simulation corresponding to the considered dispersal or landscape parameter. The error bars correspond to 1.96 times the standard error of the mean.

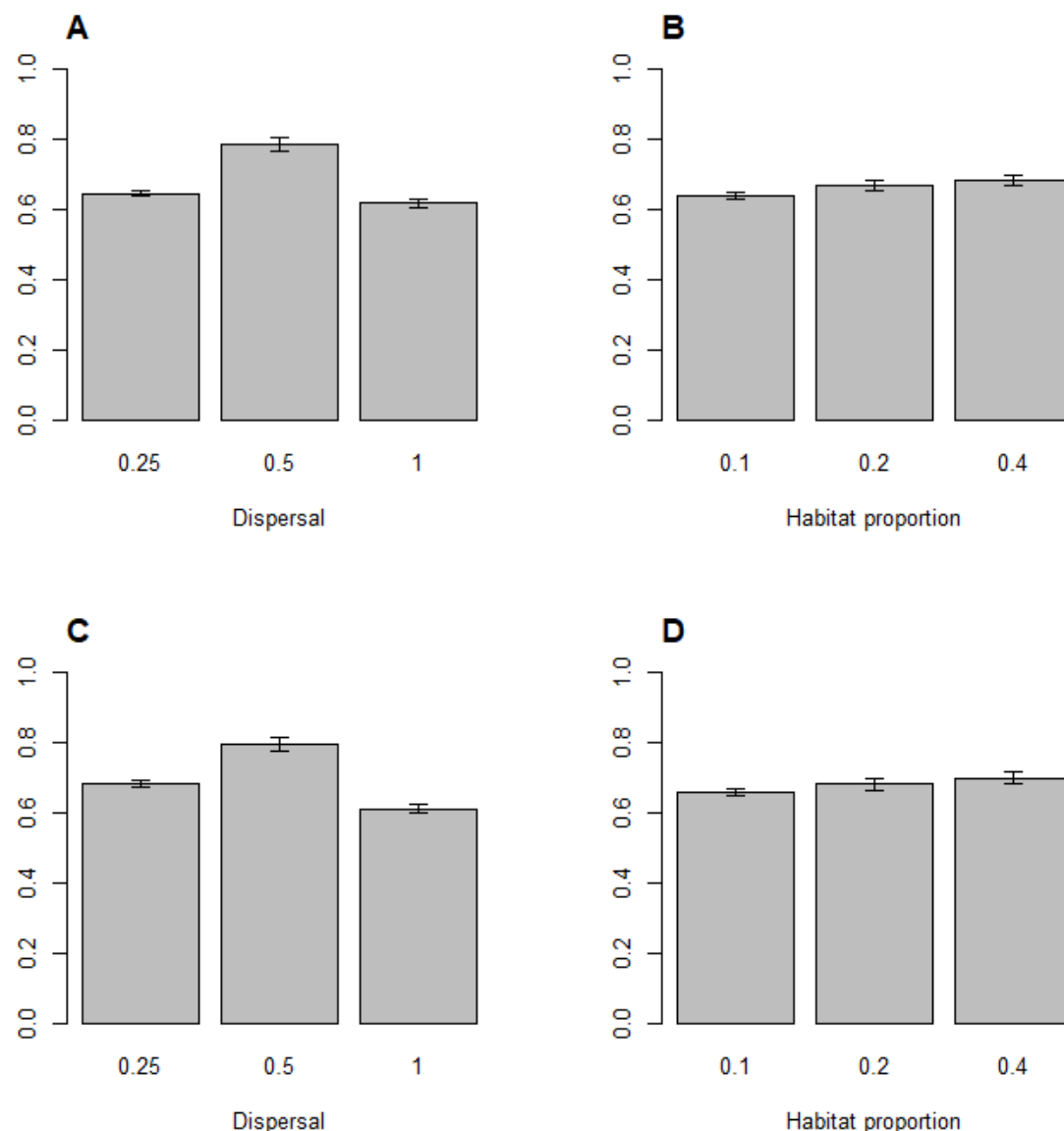
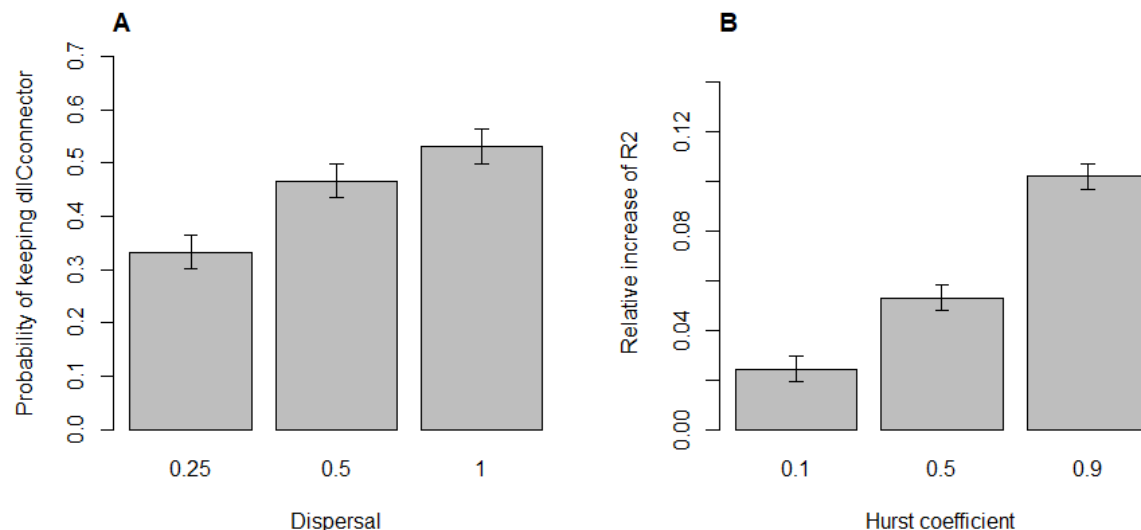


Table 2: Effect of buffer radius (r_{buf}) and dF at cell grain scale parameter (λ_c) on the correlation with species richness as a function of community dispersal (λ_s). Figures in the table correspond to average correlation value across simulation for a given CI scale parameter (or radius) and a given community dispersal level. For each value of community dispersal (table rows), and each CI type: (i) we indicated in bold the maximal correlation value obtained across scales; (ii) we indicated in normal black font the correlation values that did not differ significantly from the maximal value (checking that their estimated 95% confidence intervals, obtained as 1.96 times the standard error of the mean, overlapped with the highest correlation value); (iii) we indicated in light grey font the correlation values that were significantly lower than the maximal correlation value.

		Buffer radius (r_{buf})					dF scale parameter (λ_c)			
		1	2	4	6	8	0.25	0.5	1	2
Community dispersal	0.25	0.68	0.69	0.66	0.63	0.60	0.69	0.70	0.69	0.66
	0.5	0.70	0.81	0.84	0.83	0.80	0.74	0.80	0.84	0.85
	1	0.55	0.60	0.63	0.66	0.66	0.57	0.60	0.63	0.66

Combining different CIs to predict species richness - We explored to what extent complementing dF computed at cell grain – that showed among the strongest correlation with species richness – with a dIICconnector index allowed explaining an even higher proportion of species richness variance across sampled sites. dIICconnector had a significant additional effect on species richness in about 44% of the simulations. As an element of comparison, adding a second dF index with different scaling instead of a dIICconnector index significantly improved the effect on species richness in only 9% of the simulations. Including dIICconnector as a complementary index more often resulted in a significantly improved model fit to species richness with increasing community dispersal levels (Fig. 4). However, the relative R^2 increase when dIICconnector significantly improved the model fit was quite variable, but generally minor, with a skewed distribution (between 0% and 122%, with a median at 2%). In particular, higher relative increase of R^2 was observed for habitats with higher aggregation (Fig. 4).

Figure 4: Effect of adding dIIConnector at cell grain to dF at cell grain on model's fit to species richness. Across simulations, the additional contribution of dIIConnector could be significant or not, which is assessed as the proportion of simulations where dIIConnector complementary effect is detected as significant using a likelihood ratio test with level 0.05. Panel A shows this proportion as a function of community dispersal used in simulation. When significant, dIIConnector complementary effect could be quantified as a relative increase in model's R^2 compared to the model with dF only. Panel B shows the relative increase of R^2 induced by dIIConnector when significant, as a function of landscape Hurst coefficient in simulation (Hurst coefficient was the main driver of this quantity).



Discussion

Only connector indices at cell grain can be combined with other CIs for analyzing species richness - Our simulation allowed us to explore virtual landscapes that covered a wide variety of spatial distribution of habitats (Figs. S1, S2). Over this range of scenarios, we could identify strongly correlated CIs that may capture similar information about the spatial configuration of the surrounding habitat. [29] already performed classification of CIs according to their similarity in patch prioritization. They identified three main groups of CIs: those that quantify immigrant fluxes within the focal patch, those that quantify the role of the focal patch as a connector, and patch area. Here we extended and refined this analysis in several aspects : (i) we considered a broader set of CIs and include the new axis of landscape grain (patch versus cell); (ii) we considered distinct scale parameter values (or radius) for each indices, with no a priori hypothesis about whether indices of same type but different scaling should be more similar one with another that indices with different types but similar scaling; (iii) our clusters were directly defined in terms of Pearson correlation thresholds, which contributed to characterize which CIs can be used as complementary covariates in statistical modeling of metacommunity biodiversity; (iv) we used a methodology where clusters stems from the classification itself with clear thresholds, rather than a posteriori graphical interpretation of NMDS, which made our

approach more replicable. Our classification of CIs partially echoed the results of [29] for patch based indices: when making groups of CIs with tight pairwise correlation within groups (complete methods), we found that connector indices (red in Fig. 1) were clearly separated from indices based on patch area and those based on flux coming within patches. However, we did not retrieve the separation proposed by the authors between patch area and flux indices, which all lied in the same cluster (purple in Fig. 1, meaning correlation above 0.7). Looking more closely at the internal structure of this cluster, we could retrieve the separation between area and flux indices, but it occurred at higher correlation thresholds. The typology of [29] should therefore be simplified, flux and area indices being in practice too tightly connected to make efficient complementary indices. CI classification was more complex when moving to a cell-based analysis of the landscape, scaling and type showing an interacting effect on the correlation among indices (Fig. 1).

A striking feature of our classification of indices computed at the cell grain is the tight correlation between buffers and dF indices at cell grain. In our study, a buffer index resembled a dF index when the buffer radius was about 4 times the dF scale parameter value. [11] had already evidenced that correlations between IFM index (a generalization of the dF index; [13]) and buffers could reach 0.9 in a real landscape (their study did not focus on how the scaling of both indices could affect the correlation). However, their finding was based on IFM index computed at patch grain. Surprisingly, we did not find a similar strong relationship between buffers and dF at patch grain. This apparent paradox probably comes from the fact that the real landscape considered by [11] is highly fragmented, making patches and cells similar spatial entities. Our virtual study thus suggests that the pattern evidenced by [11] is restricted to highly fragmented contexts, while buffer – dF correlation at cell grain should apply to a broader variety of landscapes. Such a similarity between buffers and dF at cell grain was quite expected since both indices share the same general structure: a weighted sum of surrounding habitat cells contribution where weights decreases with distance following a particular kernel. dIICflux at cell grain also has a similar structure, considering topological rather than Euclidean distance, and this index was also tightly correlated with buffers and dF. Regarding the shape of the kernel, buffers are based on a step function while dF indices are based on an decreasing exponential kernel. We therefore interpret our results as the fact that changing the decreasing function used as a kernel may little affect the local connectivity as long as scaling is adjusted. This may explain why [30] found that: (i) switching from buffer to continuously decreasing kernel little affected AIC or pseudo R^2 of models used to predict species abundances; (ii) neither continuously decreasing nor step function was uniformly better to explain species abundance across four case studies; (iii) different continuous shapes of kernel had quite indiscernible predictive performance.

In a nutshell, our classification analysis based on two complementary similarity metrics (Fig. 1 and Fig. S3) revealed that there were only two groups of CIs that could be considered as relatively independent: connector indices at cell grain on the one hand, and a vast group gathering all the other CIs on the other hand. In the context of our study, it is therefore of little interest to combine several flux indices to predict species richness. By contrast, our analysis of CIs correlation suggests that combining a connector index at cell grain with some other CI (e.g. a flux index) may provide complementary insights, since both aspects of local connectivity are quite uncorrelated. This means in particular that considering separately connector $dIIC_{connector}$ and flux fraction $dIIC_{flux}$ of the integral index of connectivity $dIIC$ unravel two independent types of pattern that cannot be disentangled when using $dIIC$ index, as already emphasized by [15].

Cell-based flux indices and buffers are the best CIs for explaining species richness - All CIs except connectors reached good levels of correlation of 0.5 or more with species richness. Best indices were buffers and dF at cell grain, that reached very strong correlation (average above 0.7) outperforming patch-based indices. Finding dF and buffers as the best-performing indices brings support to the habitat amount hypothesis [17]. It demonstrates that quantifying the amount of habitat around the community of interest through a continuous (dF) or discontinuous kernel (buffer) is an efficient approach to capture immigration effects upon community species richness under our metacommunity model assumptions. Optimizing buffer (resp. dF) radius (resp. scaling parameter) could further improve the correlation, sometimes by more than 0.1. For buffers, we found that the optimal radius was about 8 times the community dispersal level implemented in simulations, which was similar to the previous findings of [20] for predicting species occurrence. Since we mentioned earlier that dF and buffers had very similar profiles when the scale parameter of dF was approximately 1/4 the radius of the buffer, it came as no surprise that, for cell-based dF , we found the best correlation with species richness when the scale parameter was equal to about twice the community dispersal level. The effect of varying buffer or dF scaling parameter upon the correlation with species richness (Table 2) was however quite moderate compared to variation among types of CIs or grain of analysis (Fig. 2), and all the tested scale parameter values yielded high correlations. This tended to suggest that finding the scaling parameter that maximizes the correlation is probably not an accurate method to obtain estimate of species dispersal level. It also means, as a corollary, that buffer or dF indices are quite robust to inaccurate knowledge of dispersal distances of species. This matches the observation that, in empirical systems, buffer radii showing correlation with species richness similar to that obtained from the “scale of effect” radius can spread over a large array of distances, sometimes covering several orders of magnitude (e.g. [31]).

Combining cell-based flux and connector indices can improve the explanation of species richness - Adding a connector index to a flux index often lead to a significant but moderate improvement of the R^2 of species richness models in our virtual study. The relative increase of R^2 could reach 10% in landscape with rather continuous habitat, which roughly corresponds to a 30% increase in R . Although limited, this effect is higher than the effect of varying scale parameters in flux indices (buffer or dF). The contribution of connector indices to species richness explanation might come from the fact that they implicitly capture habitat amount at very large scales, beyond the reach of flux indices considered in our study. It could also be due to some other phenomenon. For instance, while buffers and dF indices would quantify the quantity of migrants coming into the focal cell, connector indices may be modulating the species diversity of these migrants.

How robust are our results with respect to assumptions underpinning virtual experiments? In our simulations, best CIs reached very strong correlation (average above 0.7). Such degrees of correlation never occur in empirical studies, because many other processes than immigration and drift shape community composition and potentially break down this ideal relationship. Placing ourselves in an idealized virtual framework allowed us revealing performance contrasts and possible combinations among CIs, but our conclusions are obviously strongly dependent on the basic assumption that species considered are similar both in terms of habitat preference and in dispersal abilities. In that respect, our results should be seen as setting a reference describing how indices should perform if communities were neutral. Discrepancies with field observations would then convey information about non-neutral processes at work [32].

While using a neutral model is undoubtedly an advantage for our theoretical exploration of CI properties, we must ask to what extent the ranking of CIs obtained here depends on the other instrumental choices we made in our model beyond the neutrality assumption. Part of the relative success of cell-based dF and buffers compared to patch-based indices probably stems from the fact that we did not include different resistance values to habitat and matrix cells. Strong resistance of matrix should give more biological meaning to patch frontiers, as habitat continuity become a central feature for individuals dispersal. In addition, when heterogeneous resistance occurs, landscape connectivity including displacement costs (e.g. least cost path, circuit theory) can be markedly different from prediction based on Euclidean distance only [33], and may better capture the movement of organisms in real case study [34], [35]. This probably also applies to patch or cell connectivity and may partially explain why we do not observe in empirical systems as strong effects of Euclidean indices as in our simulations. However, we advocate that most of our results here should at least qualitatively apply to scenarios with

heterogeneous resistance. Indeed, since our study is purely virtual, we could as well consider that distances among cells in the model are not Euclidean but ecological distances. This would have amounted to saying that landscapes considered in our study are “distorted” maps. Based on this mind experiment, we still expect a strong effect of dF in landscape with heterogeneous resistance, provided that it is computed with ecological distances rather than Euclidean distances. Buffers in terms of ecological distances (which are not circles anymore in the Euclidean space) should also have strong effects on species richness. Importantly, all these conclusions would still depend on the neutrality assumption, a central hypothesis of our study, which would imply that resistance of the landscape should be similar for all species.

The grid cell resolution used to depict CI at the cell grain exactly matches the grid used to simulate the metacommunity processes of demographic stochasticity and dispersal. One may question whether this does not provide an advantage to indices computed at the cell grain. We believe that it is true, and that cell-based indices may show decreased performance if the size of a cell is very different from the size of a community, which we define as the maximum size where internal distance is not a strong limiting factor for interactions between individuals. We advocate that the grid resolution used for landscape analysis should be based on such ecological consideration, and that cells should match the typical size of communities.

Conclusion

We had identified three non-excluding methodological axes for improving the potential of CIs for explaining species presence and community richness in space: optimizing indices, combining indices and changing the grain of habitat description. Our study suggests that the most promising direction is to select appropriate indices capturing a concept of flux (buffers, dF, dIICflux), and applying them at cell grain. In line with the habitat amount hypothesis, simple buffers are probably the best tool to capture the effect of immigration upon community species richness, although more complex notion of distance should probably be used in real context, accounting for matrix resistance and dispersal barriers. These indices can reach very strong levels of correlation in a neutral metacommunity, suggesting that poorer performance in real landscape stems from the preponderance of non-neutral processes such as heterogeneity in dispersal and contrasted niches and habitats of species. By contrast combining CIs of different types showed a quite limited potential for improving the prediction of species richness. Buffers might therefore be a sufficient tool to capture landscape effects on community composition within a locality, while more complex indices may have more decisive contribution when studying landscape-scale connectivity.

Acknowledgments

We thank Santiago Saura for sharing CONEFOR code. This work was part of the Patrames project funded by the convention 2016 n°2101870336 of the French Ministry of Environment (Water and Biodiversity Direction).

References

- [1] C. S. Elton, *Animal ecology*. University of Chicago Press, 2001.
- [2] J. Grinnell, « The niche-relationships of the California Thrasher », *The Auk*, vol. 34, n° 4, p. 427-433, 1917.
- [3] G. E. Hutchinson, « Concluding remarks », *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 22, p. 415-427, 1957.
- [4] W. Thuiller, L. J. Pollock, M. Gueguen, et T. Münkemüller, « From species distributions to meta-communities », *Ecology Letters*, vol. 18, n° 12, p. 1321-1328, déc. 2015.
- [5] R. H. MacArthur et E. O. Wilson, *The theory of island biogeography*, Princeton University Press. Princeton, NJ, 1967.
- [6] L. Tischendorf et L. Fahrig, « On the use of connectivity measures in spatial ecology. A reply », *Oikos*, vol. 95, n° 1, p. 152-155, oct. 2001.
- [7] M. J. Mazerolle et M.-A. Villard, « Patch characteristics and landscape context as predictors of species presence and abundance: A review1 », *Écoscience*, vol. 6, n° 1, p. 117-124, janv. 1999.
- [8] L. R. Prugh, K. E. Hodges, A. R. E. Sinclair, et J. S. Brashares, « Effect of habitat area and isolation on fragmented animal populations », *Proc Natl Acad Sci USA*, vol. 105, n° 52, p. 20770, déc. 2008.
- [9] D. H. Thornton, L. C. Branch, et M. E. Sunkist, « The influence of landscape, patch, and within-patch factors on species presence and abundance: a review of focal patch studies », *Landscape Ecology*, vol. 26, n° 1, p. 7-18, janv. 2011.
- [10] S. A. Cushman et K. McGarigal, « Hierarchical analysis of forest bird species-environment relationships in the oregon coast range », *Ecological Applications*, vol. 14, n° 4, p. 1090-1105, août 2004.
- [11] A. Moilanen et M. Nieminen, « Simple connectivity measures in spatial ecology », *Ecology*, vol. 83, n° 4, p. 1131-1145, avr. 2002.
- [12] M. V. Vieira, M. Almeida-Gomes, A. C. Delciellos, R. Cerqueira, et R. Crouzeilles, « Fair tests of the habitat amount hypothesis require appropriate metrics of patch isolation: An example with small mammals in the Brazilian Atlantic Forest », *Biological Conservation*, vol. 226, p. 264-270, oct. 2018.
- [13] I. Hanski, « A practical model of metapopulation dynamics », *Journal of Animal Ecology*, vol. 63, n° 1, p. 151-162, 1994.
- [14] D. Urban et T. Keitt, « Landscape connectivity: a graph-theoretic perspective », *Ecology*, vol. 82, n° 5, p. 1205-1218, mai 2001.
- [15] S. Saura et L. Rubio, « A common currency for the different ways in which patches and links can contribute to habitat availability and connectivity in the landscape », *Ecography*, vol. 33, n° 3, p. 523-537, 2010.
- [16] G. Baranyi, S. Saura, J. Podani, et F. Jordán, « Contribution of habitat patches to network connectivity: Redundancy and uniqueness of topological indices », *Ecological Indicators*, vol. 11, n° 5, p. 1301-1310, sept. 2011.
- [17] L. Fahrig, « Rethinking patch size and isolation effects: the habitat amount hypothesis », *Journal of Biogeography*, vol. 40, n° 9, p. 1649-1663, mai 2013.
- [18] L. Tischendorf et L. Fahrig, « On the usage and measurement of landscape connectivity », *Oikos*, vol. 90, n° 1, p. 7-19, juill. 2000.
- [19] E. P. Economo et T. H. Keitt, « Network isolation and local diversity in neutral metacommunities », *Oikos*, vol. 119, n° 8, p. 1355-1363, 2010.

- [20] H. B. Jackson et L. Fahrig, « What size is a biologically relevant landscape? », *Landscape Ecology*, vol. 27, n° 7, p. 929-941, août 2012.
- [21] T. R. Etherington, E. P. Holland, et D. O'Sullivan, « NLMpy: a python software package for the creation of neutral landscape models within a general numerical framework », *Methods in Ecology and Evolution*, vol. 6, n° 2, p. 164-168, févr. 2015.
- [22] R Core Team, *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- [23] A. G. Bunn, D. L. Urban, et T. H. Keitt, « Landscape connectivity: A conservation application of graph theory », *Journal of Environmental Management*, vol. 59, n° 4, p. 265-278, août 2000.
- [24] S. Saura et J. Torné, « Conefor Sensinode 2.2: A software package for quantifying the importance of habitat patches for landscape connectivity », *Environmental Modelling & Software*, vol. 24, n° 1, p. 135-139, janv. 2009.
- [25] C. Ricotta, A. Stanisci, G. C. Avena, et C. Blasi, « Quantifying the network connectivity of landscape mosaics: a graph-theoretical approach », *Community Ecology*, vol. 1, n° 1, p. 89-94, 2000.
- [26] L. Pascual-Hortal et S. Saura, « Comparison and development of new graph-based landscape connectivity indices: towards the prioritization of habitat patches and corridors for conservation », *Landscape Ecol*, vol. 21, n° 7, p. 959-967, oct. 2006.
- [27] S. Saura et L. Pascual-Hortal, « A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study », *Landscape and Urban Planning*, vol. 83, n° 2, p. 91-103, nov. 2007.
- [28] L. C. Freeman, « A Set of Measures of Centrality Based on Betweenness », *Sociometry*, vol. 40, n° 1, p. 35-41, 1977.
- [29] G. Baranyi, S. Saura, J. Podani, et F. Jordán, « Contribution of habitat patches to network connectivity: Redundancy and uniqueness of topological indices », *Ecological Indicators*, vol. 11, n° 5, p. 1301-1310, sept. 2011.
- [30] P. Miguët, L. Fahrig, et C. Lavigne, « How to quantify a distance-dependent landscape effect on a biological response », *Methods in Ecology and Evolution*, vol. 8, n° 12, p. 1717-1724, juin 2017.
- [31] K.-O. Bergman, N. Jansson, K. Claesson, M. W. Palmer, et P. Milberg, « How much and at what scale? Multiscale analyses as decision support for conservation of saproxylic oak beetles », *Forest Ecology and Management*, vol. 265, p. 133-141, 2012.
- [32] D. Alonso, R. S. Etienne, et A. J. McKane, « The merits of neutral theory », *Trends in Ecology & Evolution*, vol. 21, n° 8, p. 451-457, août 2006.
- [33] C. E. Simpkins, T. E. Dennis, T. R. Etherington, et G. L. W. Perry, « Assessing the performance of common landscape connectivity metrics using a virtual ecologist approach », *Ecological Modelling*, vol. 367, p. 13-23, janv. 2018.
- [34] S. L. Emel et A. Storfer, « Landscape genetics and genetic structure of the southern torrent salamander, *Rhyacotriton variegatus* », *Conservation Genetics*, vol. 16, n° 1, p. 209-221, févr. 2015.
- [35] S. Vuilleumier et P. Fontanillas, « Landscape structure affects dispersal in the greater white-toothed shrew: Inference between genetic and simulated ecological distances », *Ecological Modelling*, vol. 201, n° 3, p. 369-376, mars 2007.

Supplementary figures

Figure S1: Four examples of extreme landscapes in terms of aggregation (hurst) and habitat proportion (prop) in our study. Habitat is pictured in green, matrix in brown.

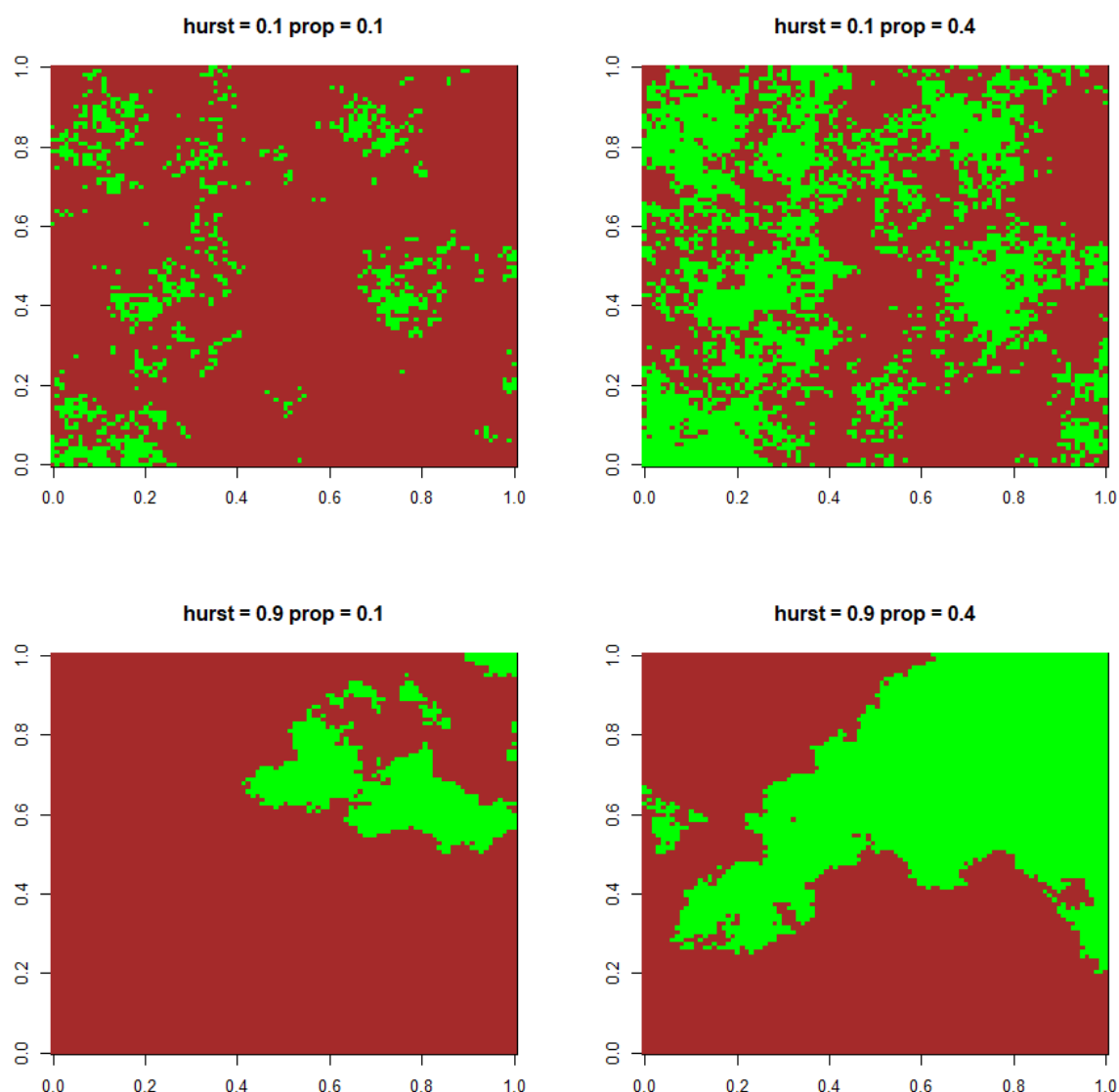


Figure S2: Average size and number of patches in the virtual landscapes used in our study. Colors correspond to distinct combinations of Hurst exponent and habitat proportion. Ellipses correspond to 95%-CI of a fitted bivariate Student distribution.

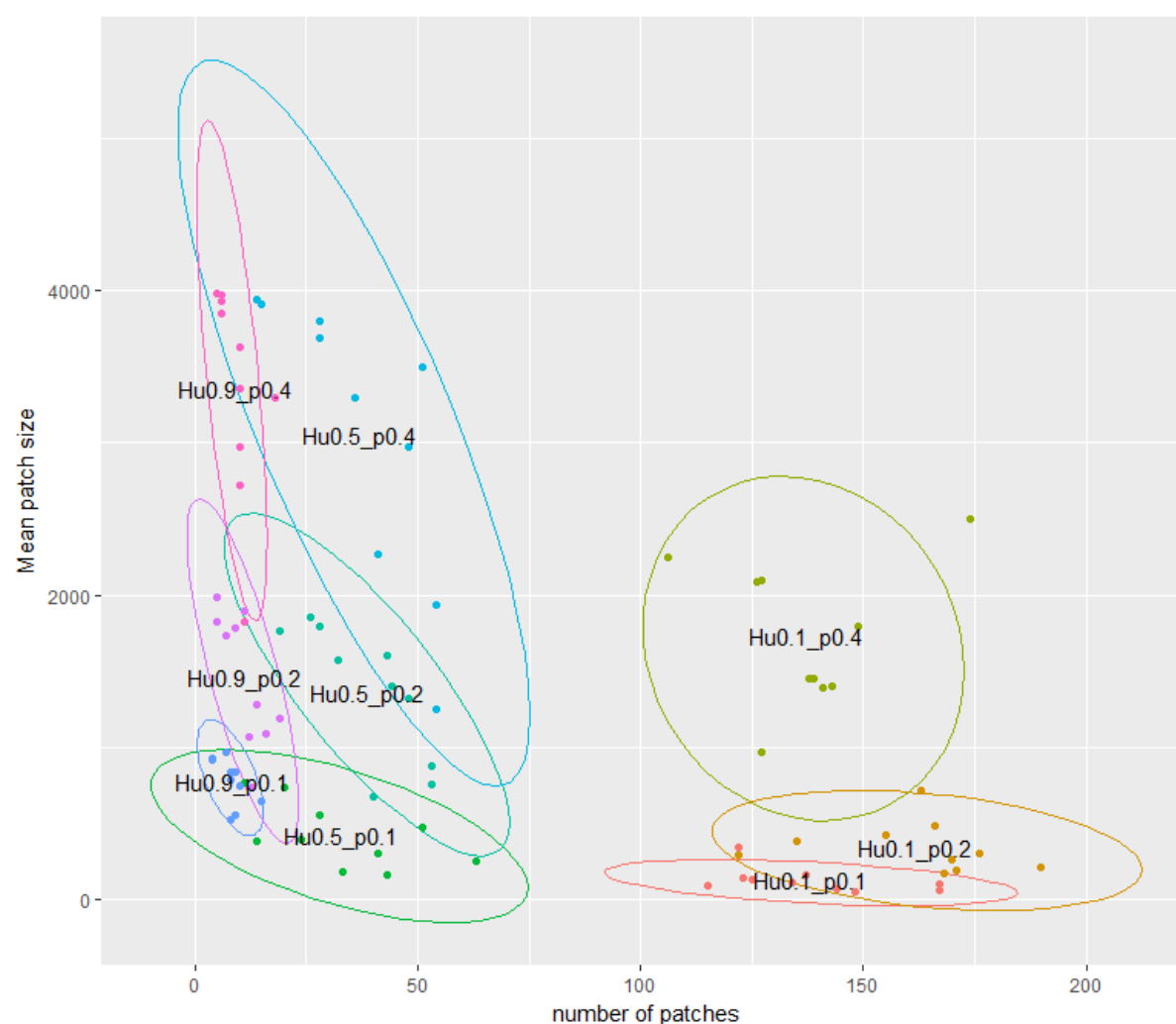
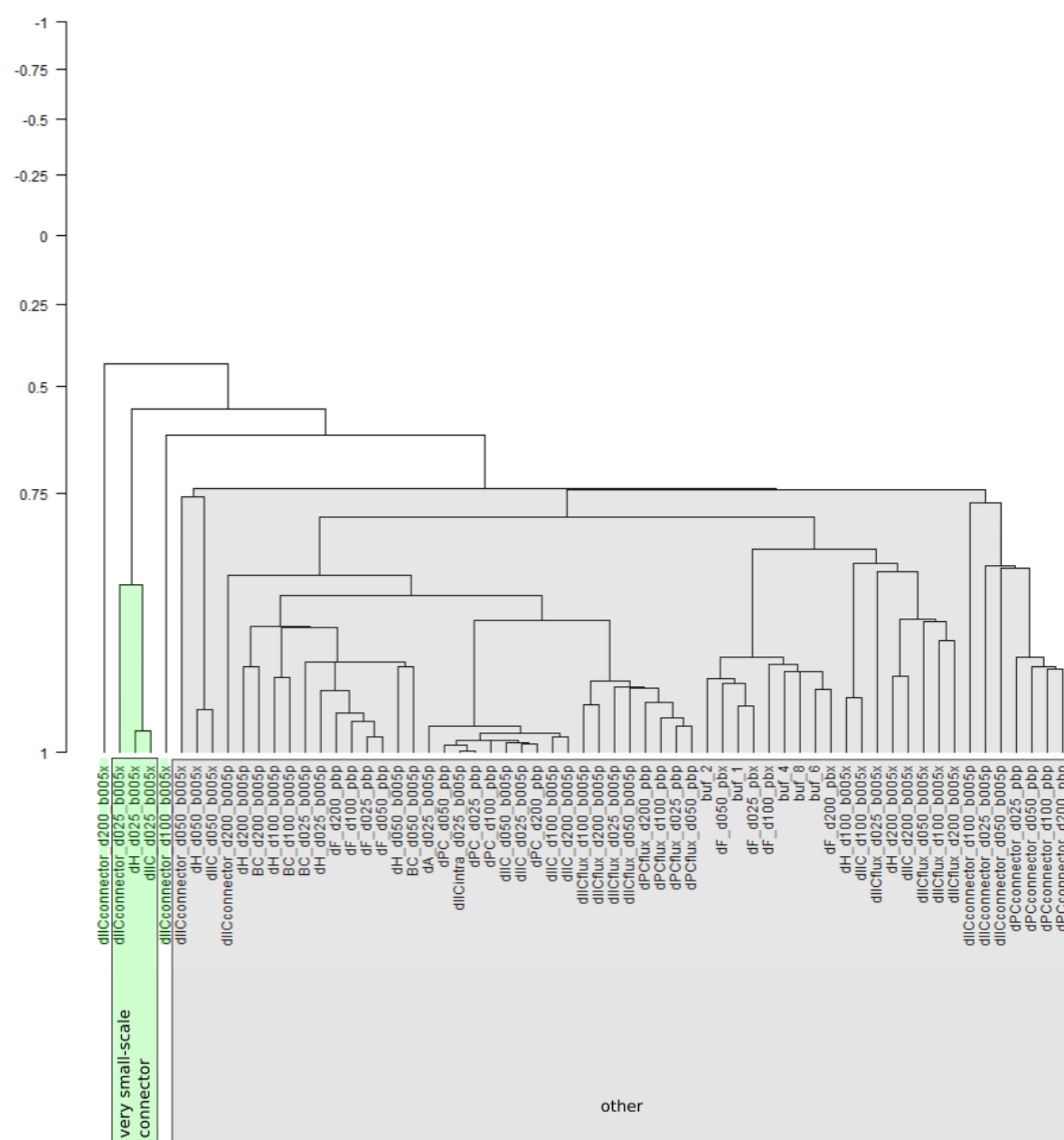


Figure S3: Classification of connectivity indices (CIs) across all landscapes using the “single” method for group merging. We delineated minimum monophyletic groups with common ancestor at a correlation value equal or below 0.7, meaning two indices belonging to distinct groups has correlation below 0.7 across our simulated landscapes. These groups are presented using colors and frames. Indices in green are connector indices. The grey frame encompasses a mixture of connector, area and flux indices. CIs without frame constitute monophyletic groups on their own. The “very small-scale connector” group is identical to that of Fig. 1 in main text and indices within it are computed at cell grain. See Table S1 for the precise meaning of indices names.



Supplementary tables

Table S1: List of all patch structural connectivity indices (CIs) considered in our study.

CIs labels are made of three parts separated by underscores “_”. The first part of the name indicates the type of the index (“buf”, “dA”, “dF”, “dH”, “BC”, “dPC”, “dPCflux”, “dPCconnector”, “dIIC”, “dIICintra”, “dIICflux”, “dIICconnector”; see Table 1 for description). The second part of the name indicates the scale parameter of the index (“d025”, “d050”, “d100”, “d200” corresponding to $\lambda_c = 0.25, 0.5, 1, 2$ cells respectively, and “1”, “2”, “4”, “6”, “8” corresponding to buffer radius r_{buf} in cells). The last part refers to the grain used to compute the index (“ppb” and “b005p” corresponding to patch grain, (“pbx” and “b005x” corresponding to cell grain)

CI labels	Edges handling	CI type	Habitat grain	CI scaling /radius
buf_1	-	buffer	cell	1
buf_2	-	buffer	cell	2
buf_4	-	buffer	cell	4
buf_6	-	buffer	cell	6
buf_8	-	buffer	cell	8
dA_d025_b005p	-	area	patch	-
dIICintra_d025_b005p	-	internal fraction of the integral index of connectivity	patch	-
BC_d025_b005p	binary	betweenness connectivity	patch	0,25
BC_d050_b005p	binary	betweenness connectivity	patch	0,5
BC_d100_b005p	binary	betweenness connectivity	patch	1
BC_d200_b005p	binary	betweenness connectivity	patch	2
dH_d025_b005p	binary	Harary index	patch	0,25
dH_d050_b005p	binary	Harary index	patch	0,5
dH_d100_b005p	binary	Harary index	patch	1
dH_d200_b005p	binary	Harary index	patch	2
dH_d025_b005x	binary	Harary index	cell	0,25
dH_d050_b005x	binary	Harary index	cell	0,5
dH_d100_b005x	binary	Harary index	cell	1
dH_d200_b005x	binary	Harary index	cell	2
dIIC_d025_b005p	binary	integral index of connectivity	patch	0,25
dIIC_d050_b005p	binary	integral index of connectivity	patch	0,5
dIIC_d100_b005p	binary	integral index of connectivity	patch	1
dIIC_d200_b005p	binary	integral index of connectivity	patch	2
dIIC_d025_b005x	binary	integral index of connectivity	cell	0,25
dIIC_d050_b005x	binary	integral index of connectivity	cell	0,5
dIIC_d100_b005x	binary	integral index of connectivity	cell	1
dIIC_d200_b005x	binary	integral index of connectivity	cell	2
dIICconnector_d025_b005p	binary	connector fraction of the integral index of connectivity	patch	0,25

dIIConnector_d050_b005p	binary	connector fraction of the integral index of connectivity	patch	0,5
dIIConnector_d100_b005p	binary	connector fraction of the integral index of connectivity	patch	1
dIIConnector_d200_b005p	binary	connector fraction of the integral index of connectivity	patch	2
dIIConnector_d025_b005x	binary	connector fraction of the integral index of connectivity	cell	0,25
dIIConnector_d050_b005x	binary	connector fraction of the integral index of connectivity	cell	0,5
dIIConnector_d100_b005x	binary	connector fraction of the integral index of connectivity	cell	1
dIIConnector_d200_b005x	binary	connector fraction of the integral index of connectivity	cell	2
dIIcflux_d025_b005p	binary	flux fraction of the integral index of connectivity	patch	0,25
dIIcflux_d050_b005p	binary	flux fraction of the integral index of connectivity	patch	0,5
dIIcflux_d100_b005p	binary	flux fraction of the integral index of connectivity	patch	1
dIIcflux_d200_b005p	binary	flux fraction of the integral index of connectivity	patch	2
dIIcflux_d025_b005x	binary	flux fraction of the integral index of connectivity	cell	0,25
dIIcflux_d050_b005x	binary	flux fraction of the integral index of connectivity	cell	0,5
dIIcflux_d100_b005x	binary	flux fraction of the integral index of connectivity	cell	1
dIIcflux_d200_b005x	binary	flux fraction of the integral index of connectivity	cell	2
dF_d025_pbp	continuous weights	flux index	patch	0,25
dF_d050_pbp	continuous weights	flux index	patch	0,5
dF_d100_pbp	continuous weights	flux index	patch	1
dF_d200_pbp	continuous weights	flux index	patch	2
dF_d025_pbx	continuous weights	flux index	cell	0,25
dF_d050_pbx	continuous weights	flux index	cell	0,5
dF_d100_pbx	continuous weights	flux index	cell	1

dF_d200_pbx	continuous weights	flux index	cell	2
dPC_d025_pbp	continuous weights	probability of connectivity index	patch	0,25
dPC_d050_pbp	continuous weights	probability of connectivity index	patch	0,5
dPC_d100_pbp	continuous weights	probability of connectivity index	patch	1
dPC_d200_pbp	continuous weights	probability of connectivity index	patch	2
dPCflux_d025_pbp	continuous weights	flux fraction of probability of connectivity index	patch	0,25
dPCflux_d050_pbp	continuous weights	flux fraction of probability of connectivity index	patch	0,5
dPCflux_d100_pbp	continuous weights	flux fraction of probability of connectivity index	patch	1
dPCflux_d200_pbp	continuous weights	flux fraction of probability of connectivity index	patch	2
dPCconnector_d025_pbp	continuous weights	connector fraction of probability of connectivity index	patch	0,25
dPCconnector_d050_pbp	continuous weights	connector fraction of probability of connectivity index	patch	0,5
dPCconnector_d100_pbp	continuous weights	connector fraction of probability of connectivity index	patch	1
dPCconnector_d200_pbp	continuous weights	connector fraction of probability of connectivity index	patch	2