

VISUALIZING SPECIES RICHNESS AND SITE SIMILARITY FROM PRESENCE-ABSENCE MATRICES

JORGE SOBERÓN^{1*}, MARLON E. COBOS¹, AND CLAUDIA NUÑEZ-PENICHER¹

¹*Department of Ecology and Evolutionary Biology & Biodiversity Institute, University of Kansas, Lawrence, Kansas, United States.*

Abstract. Species richness and similarity of biotas among distinct sites are important quantities in biogeography. Indices derived from presence-absence matrices are used to represent these quantities in so-called range-diversity plots. The most commonly used range-diversity plot, however, has multiple special cases and its interpretation is cumbersome. Here we present an equivalent formulation that is geometrically simpler and has no special cases. In addition, we introduce a method to identify the statistical significance of the dispersion field, an index that represents how similar species composition is in a cell with respect to the whole area. The new range-diversity plot is a promising tool to explore biodiversity and endemism in a region as the values shown in this plot and their statistical significance can also be represented in geography.

Key words: Biodiversity index, Dispersion field, Range-diversity plot, Endemism, PAM

INTRODUCTION

Representing biodiversity is a complex challenge, mainly because the concept itself is poorly defined (Sarkar 2002). Most often, “biodiversity” is used to mean the list of species in a region. One way of organizing biodiversity data is by using presence-absence matrices (PAMs), where a one represents presence of species j in cell i , and a zero, absence (Fig. 1). PAMs could be regarded as mathematical representations of lists of species per spatial unit. Originally, PAMs were developed for dozens of species in a handful of islands in archipelagos (Connor and Simberloff 1979), which means that, originally, PAMs had a few hundreds of cells. However, it is possible and desirable to expand the concept of a PAM to arbitrary grid partitions of large regions, with extents of 10^6 km² or larger, and grids of resolutions in the order of 1 to a few thousand km². Using these grid partitions allows obtaining more detailed information on the incidence of species in a region, but also helps in visualizing biodiversity patterns via distinct plots of the indices that can be calculated from a PAM (Soberón and Cavner 2015).

PAMs have been analyzed with a variety of techniques (Arita et al. 2008; Christen and Soberón 2009; Ulrich and Gotelli 2012; Soberón and Cavner 2015). One of the methods is the range-diversity plot (Arita et al. 2008; Borregaard and Rahbek 2010; Soberón and Ceballos 2011). This plot displays jointly two indices describing the community composition of every cell in the grid: (1) the number of species, in absolute or relative numbers, and (2) the mean dispersion field (Graves and Rahbek 2005). The dispersion field is simply total range size of all the species occupying a cell. This can be proven to be equivalent to the total amount of overlap of the community of species in a cell, with all the other cells (Soberón & Ceballos, 2011). Instead of total values, in the range-diversity plot, the mean value of the dispersion field is used. The mean is taken with respect of the species existing in each cell. The mean can be taken also with respect to all species in the region, as we shall see later.

The range-diversity plot displays a large amount of data in a compact way that allows interpretations. The mathematical limits of the range-diversity plot are well-defined (Arita et al. 2008; Soberón and Ceballos 2011) and constrains the scatterplot in a very strict way. We refer the reader to these articles for all mathematical details. Moreover, these limits have a straightforward geographic interpretation (Soberón and Ceballos 2011). However, the geometric properties of a range-diversity plot are somewhat cumbersome and subject to many particular cases (i.e., plots are not linear and their shape depends on the specifics

* Corresponding author: Jorge Soberón

Department of Ecology & Evolutionary Biology and Biodiversity Institute, University of Kansas, Lawrence, Kansas 66045, USA. jsoberon@ku.edu

ORCID:

JS: <https://orcid.org/0000-0003-2160-4148>

MEC: <https://orcid.org/0000-0002-2611-1767>

CNP: <https://orcid.org/0000-0001-7442-8593>

of the parameters), making them more difficult to interpret. For instance, the sign of the minimum covariance determines the slope of the lines defining the limits, leading to rather contrasting shapes (Fig. 2); and the relationship between minimum relative richness and minimum covariance determines whether the limits will be truncated or not (not illustrated).

In this contribution, we introduce a formulation of the range-diversity plot, which is equivalent to, but geometrically much simpler than the previous one, and does not have particular cases, facilitating interpretations and further explorations. We also present a way to detect whether observed values of the index related to species ranges (the dispersion field) are statistically significant, allowing for further interpretations in regards to how species are distributed in a region, based on a PAM.

METHODS

Data

To illustrate the application of the new method to produce range-diversity plots, we use a PAM originally composed of $S = 1595$ species of terrestrial vertebrates, on a grid of $N = 711$ cells (resolution of 0.5° , or ~ 55 km at the equator) subdividing Mexico. The original data used to produce the PAM comes from the International Union for the Conservation of Nature (IUCN), which maintains a database of downloadable, machine-readable maps (in shapefile format) for the non-avian terrestrial (increasingly also aquatic) vertebrates of the world (IUCN 2020¹). In this PAM, each cell is represented by the geographic coordinates of its centroid and can be downloaded using the code provided to replicate analysis and plots (data and code are available²). After excluding cells with no species, or species with values of 0 (absence) for all cells, the final number of species and cells was 1573 and 711, respectively (Fig. 1).

Calculations for new plotting method

The basic idea for the range-diversity plot follows from an identity proven in Soberón and Cavner (2015) and based on the variance-covariance matrix of species in cells. This is the covariance in species composition among all cells in the grid. By forming the vector $\bar{\tau}$ of the average covariance of every cell to all other cells, the following identity is proven:

$$S\bar{\tau} = \Phi^* - \alpha / \beta$$

¹ <https://www.iucnredlist.org/resources/spatial-data-download>.

² <https://github.com/jsoberon/PAMs-Mexico>.

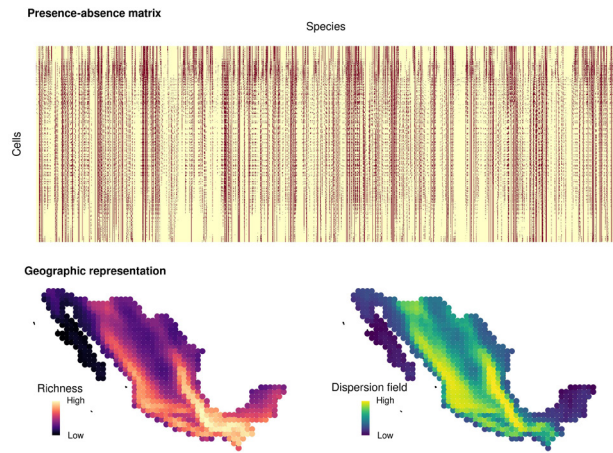


Figure 1. Graphical representation of a presence-absence matrix (top panel, presence = 1 in brown, absence = 0 in yellow), for the PAM (representing 711 cells and 1573 species) of the terrestrial vertebrates of Mexico; and geographic representation of indices of richness and dispersion field derived from it.

Where S is the number of species, Φ^* is the vector of normalized (divided by N) dispersion field values, α is the vector of numbers of species, and β is Whittaker’s (1960) beta “diversity”. A division by N creates a proportion with respect to the size of the region in question. Whittaker’s first equation for beta diversity is simply $S / \bar{\alpha}$, which, by a simple replacement is $\beta = S / \bar{\alpha} = NS / f$ where f is the “fill” of the PAM, or the total number of ones. From the above Arita et al. (2008) obtained the following equation, which is the base of their range-diversity plot:

$$\alpha_i^* = \frac{S\bar{\tau}_i}{\bar{\varphi}_i^* - 1 / \beta}$$

The range-diversity plot represents every cell in the grid, using as y -axis the quantity α_i^* (the asterisk represents that α has been normalized by dividing it by S) and as x -axis the mean dispersion field per cell, or $\bar{\varphi}_i^*$ (with the asterisk representing division by N and to get the mean dispersion field per cell φ_i is divided by α_i). As we shall exemplify below, the above equation implies that the data points in the scatterplot are enclosed within a tent-like “envelope” determined by the minimum and maximum mean covariances, and by the beta diversity (Fig. 2). This envelope is non-linear in shape and has many special cases and one potential discontinuity.

However, a much simpler and direct graph comes directly from equation (1). Indeed, element-by-element equation (1) means:

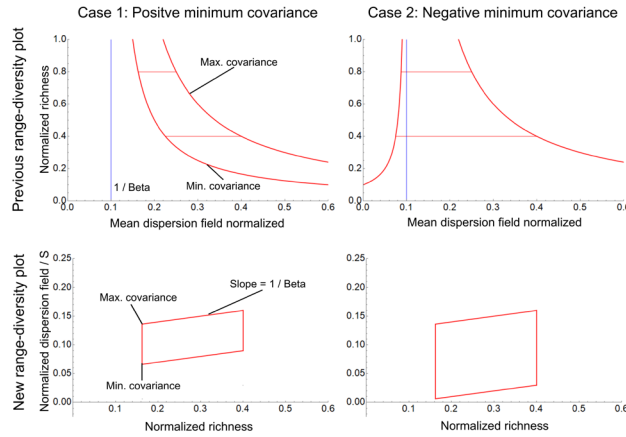


Figure 2. Two special cases of the range-diversity plot. Notice that in the old diagram the change of sign in the minimum covariance produces a very different-looking envelope. In the new diagram, it is the size that changes.

$$S\bar{\tau}_i = \varphi_i^* - \frac{1}{\beta} \alpha_i$$

Or, by an immediate rearrangement:

$$\frac{1}{S} \varphi_i^* = \bar{\tau}_i + \frac{1}{\beta} \alpha_i^*$$

This suggests that φ_i^* / S vs. α_i^* creates a simpler graph, as it is indeed the case. Moreover, the theoretical limits of this new range-diversity plot are straightforward. The limits form a parallelepiped with vertical lines at α_{\min} and α_{\max} , and straight lines with a slope of $1/\beta$ and intercepts at the minimum and maximum values for τ (Fig. 2). Notice the main difference is that instead of plotting the mean normalized dispersion field, with the mean taken by dividing by the local number of species (α_i^*), the mean is taken over all the species in the region. In the discussion, we further expand this point.

Statistical significance of dispersion field

Indices derived from a PAM can be tested for significance. For instance, Schluter (1984) provides a test for the similarity of the distributions of a number of species, using a variance ratio. However, a PAM permits the calculation of a large variety of indices. A more general approach to the problem of statistical significance relies on the randomization of PAMs. PAMs can be randomized under a number of assumptions (Miklós and Podani 2004) and this allows calculation of the distributions of values of the dispersion field, the mean dispersion field, the distribution of richness, overlaps, or many other quantities, from

those that cannot be distinguished from random expectations and others that differ significantly from such a distribution. Randomizing PAMs requires making important decisions (Strona et al. 2018), specifically how the marginal sums of the PAM will be treated. These marginals represent the total number of species per site (richness, the vector α), when summing over columns, or the total number of sites each species uses (incidence, the vector ω) when summing over rows. The question then is whether to leave richnesses, or incidences, or both, constant at the time of randomizing. This is a problem well beyond the purpose of this contribution (see discussion), but we used the approach of randomizing subject to fixed richness and incidence. We used this approach because we assume that both incidence and richness are observed data, deriving from reliable sources of information. To perform the randomization process, we used the function “randomizeMatrix” from the package picante (Kembel et al. 2010) in R 4.0.2 (R Core Team 2020). For the process of randomization, we used the null model “independentswap” which randomizes the matrix with the independent swap algorithm from Gotelli (2000). Considering how the randomization is done, before starting the process, we broke any potential geographic structure remaining in the PAM (where closer rows are also closer cells in the geography) by rearranging randomly the rows (which represent geographic cells) but maintaining the identity of these cells in the results.

After each randomization, the dispersion field is calculated from the resulting PAM and it is normalized. The result of repeating this process a number of iterations (500 in our case) generates a distribution of values of the dispersion field under random expectations. Then, the actual values of any index, for each cell, are compared to the random distribution of values, and the observed values that are as extreme or more extreme than the confidence limits of the null distribution (by default 5%: the 2.5% lower and 97.5% upper limits) are considered statistically significant. Significance is treated differently when the actual value is below the lower limit or when it is above the upper limit as this has implications for interpretation (see discussion).

Calculations of biodiversity indices derived from the PAM, and its randomization, were done using the function “prepare_PAM_CS” from the package biosurvey (Nuñez-Penichet et al. 2020; available³) in R. Geographic representations of the results obtained

³ <https://github.com/claununez/biosurvey>.

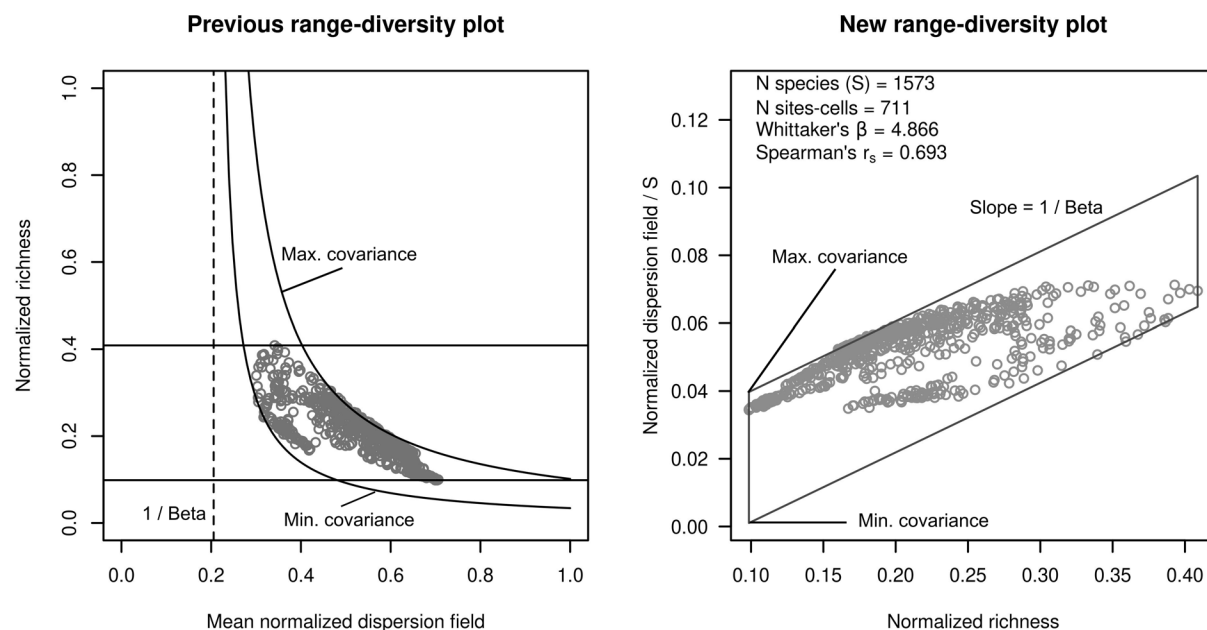


Figure 3. Comparison between previous range-diversity plot and the new plot proposed created with IUCN species ranges for Mexico (summarized in a PAM).

with the new method were done using the package maps (Brownrigg et al. 2018) and other base functions in R.

Exploring the new range-diversity plot

To illustrate the differences between the previous range-diversity plot and the new diagram proposed here, we exemplify two theoretical special cases of these plots. The two cases are: 1) a positive minimum covariance when comparing communities per site (grid cells in our case); and 2) the case with a negative minimum covariance.

Using the PAM based on the terrestrial vertebrate species of Mexico, we produced the original and the new range-diversity plots. Based on the same PAM, we also detected statistically significant values of the dispersion field and identified them in the new range-diversity plot and in a geographic representation for Mexico. As the relationship between points in the range-diversity plot and areas in the geographic region of interest is of high importance, we created visualizations of block-like chunks of points in the diagram and their geographic projections. Blocks in the range-diversity plot were divided using a grid (equal-size cells) of 4 rows and 4 columns using biosurvey in R. Four of the resultant blocks were selected randomly to be used for plots.

RESULTS

In Figure 2 we display the old and the new

range-diversity plots. The plots in the top row in Figure 2 are examples of the original range-diversity diagram proposed by Arita et al. (2008). In this version, the data points are constrained to occur under a tent-like curved region determined by quantities derived from the PAM. The shape and position of the tent-like permitted zone are determined by the beta diversity (Whittaker), the minimum and maximum mean covariance of the community of species in a cell, relative to every other; and indirectly, by the minimum and the maximum number of species in the cells of the grid (Fig 2, upper row). There are several possible particular cases of this diagram, as illustrated in Figure 2 for two possibilities only.

In the bottom row of Figure 2, we present the new range-diversity plot, where the order of the axes is reverted, and the normalizations are not exactly the same. There is still a permitted region, determined by the same values, but the permitted region is now a simple parallelogram, always of the same shape. The area of the parallelogram, though, may change with the values of the parameters (minimum and maximum alpha, beta, and minimum and maximum mean covariances), but the general aspect of the plot is always the same.

The range-diversity plot of the IUCN PAM for 1573 species of terrestrial vertebrates of Mexico is displayed in Figure 3 using the old and the new diagrams. Notice that the points form clusters. As shown below, these clusters have a geographic meaning.

The differences between the old and the new diagrams are: (i) the variables used in the old are the normalized mean (with respect to local richness) dispersion field and the normalized species richness (of every cell), or α_i^* vs. ϕ_i^*/α_i . In the new one, the variables are the normalized dispersion field divided by the total number of species, which can be interpreted as a mean with respect to the number of species in the entire region, and the normalized (with respect to S) species richness. In symbols, we plot ϕ_i^*/S vs. α_i^* . (ii) The limits of the permitted regions are determined by the minimum and maximum mean covariance in both diagrams, in the old one these are hyperbolic lines, whereas in the new plot they are straight lines of slope $1/\beta$.

In Figure 4 we show the new range-diversity plot for the IUCN PAM, as well as a geographic representation of the results obtained. The top-left panel is the new diagram. The top-right panel shows the values of the dispersion field resulted from the randomizations of the PAM. We randomized the PAM, subject to fixed marginals. This randomization changes the values of the dispersion field, but not those of the richness (by construction: richness per cell is kept constant during randomizations). Values above and below the 97.5% and the 2.5% percentiles of the random distribution are illustrated in the bottom-left panel, with the corresponding geographic representation in the bottom-right part of Figure 4. Circles in black are above the 97.5% percentile of the random-

izations, and those closed circles in gray are below the 2.5%. Open circles in gray are inside the 2.5% to 97.5% range of randomizations. Notice that the map shows three clearly differentiated regions: in black are cells with dispersion-fields (remember that the mean dispersion field is identical to mean similarity) significantly larger than random, dark gray cells are those with dispersal fields smaller than random, and light gray cells are those indistinguishable from random. In general, our results show regions of higher than randomly expected similarity in the Nearctic part of Mexico, and lower than randomly expected similarities in the Neotropical part. A transition region where similarities are as expected under a random model is also observed in the transition between Neotropics and Nearctic.

Finally, the range-diversity plot displays clusters or groups of points, which correspond to cells in geography with distinct patterns (some of them clustered but others disjoint; Fig. 5). Exploring this duality helps to understand how richness and similarity of species' composition are distributed in the geography, since recognizing where points in the range-diversity plot are located geographically is not necessarily intuitive.

DISCUSSION

Since its introduction by Arita et al. (2008) the range-diversity plot has been used in a number of analyses (for instance, Borregaard and Rahbek 2010; Villalobos et al. 2013a and b; 2017; Zwiener et al. 2018), but it is a bit cumbersome to read. It has many special cases, contains discontinuities around the value of $1/\beta$, and its "envelope" is non-linear. In this work, we propose a much simpler variant that displays the same information. It can be used, for instance, to obtain richness-rarity quartiles (Villalobos et al. 2013a and b), and has the same properties of relation with maps in geographic space, but it is simpler to interpret and has no differently-shaped special cases (Figs. 2-3). For instance, in the new plot, having negative covariances among-communities composition does not create, as in the previous plot, positive and negative clouds of points; now the entire cloud has a positive slope.

The mean dispersion field is simply the sum of Jaccard similarities among regions (Soberón and Ceballos 2011). When taking the mean relative to the S species in the region, rather than to α_i , the local species number, one gets a smaller measure of similarity,

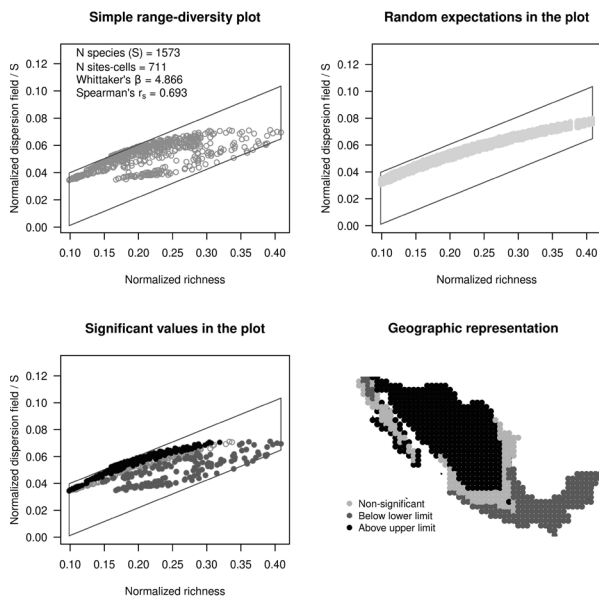


Figure 4. Plotting options for the new range-diversity plot and representation of results in the geography. The non-statistically significant cells map into a region of Mexico that represents the transition between tropical and temperate conditions.

since $S >$ any local α . This idea needs to be kept in mind when interpreting the new plot.

We also propose an application of randomization of a PAM to obtain intervals of confidence for one statistic derived from the PAM, the normalized dispersion field. Because our randomization process starts by reordering arbitrarily the rows in the matrix (cells or sites), spatial autocorrelation is, in some sense, broken, preventing artifacts due to it. Although we do not prove it here, several metrics of a PAM are invariant to randomization with fixed marginals (obviously the mean richness and Whittaker's beta, but also the mean dispersion field, the mean diversity field, and others). Fortunately, the dispersion field is sensitive to such type of randomization, and since, as it has been proven in Soberón and Ceballos (2011), the dispersion field of a cell is mathematically equivalent to the total number of shared species with all the others cells, identifying those cells that are statistically significant is a useful result (Figs. 4-5).

The randomization of the entire PAM allows testing the significance of any of the parameters derived from it. In this sense, our approach is more general than those that rely on the significance of specific indices. As said before, Schluter (1984) presented a variance test to compare the composition of two regions. This has been applied, for instance, by Villalobos et al. (2014). Our approach would permit to test directly whether composition, or nestedness, or any parameter sensitive to the randomization of a PAM has an extreme value, with respect to the null model of fixed marginals.

The null models obtained by fixing either of the marginals of a PAM have biological interpretations. In an interesting paper, Gibert and Escarguel (2019) suggest that fixing the species numbers marginal is equivalent to hypothesizing that the dominant community assembly processes are "niche" based, and fixing the species ranges marginal is a hypothesis about the dominance of dispersal factors. Although we do not explore these ideas here, we notice that the range-diversity plots are ideally suited to do this exploration.

Statistically significant values of the dispersion field may have a geographic meaning (Table 1). Indeed, cells with observed values below the lower confidence limits of random expectations are those with smaller similarity to others, or in other words, inhabited with species of more restricted ranges. This can be regarded as cells with a species' composition "more endemic." Those above the upper limits of

random expectations indicate cells that share more species with others than expected randomly. This is, cells with a species' composition shared widely with others, or less "endemic." That the region indistinguishable from random appears to separate geographical regions of higher than random similarity from those of lower than random similarity is quite interesting. In other explorations (not presented), we also found that geographic regions with dispersion field values above or below random expectations are bordered or separated by the areas with values indistinguishable from random. We believe that our method would be useful to suggest transitional zones in terms of levels of endemism, but this requires further empirical exploration.

For a given cell, the value of its dispersion field is positively correlated with the number of species in the cell, since it is calculated as the sum of the ranges of all the species in the cell (thus more species, more terms in the sum). Therefore for a given matrix,

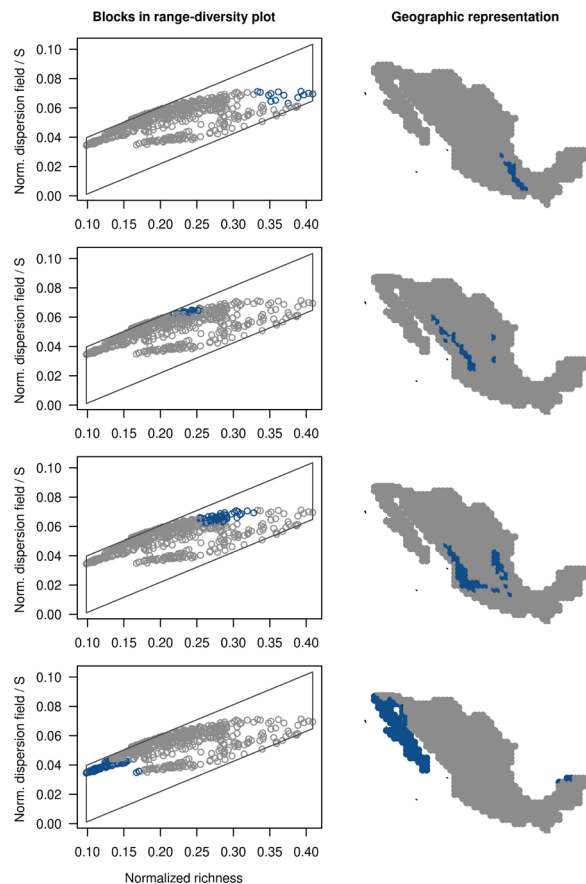


Figure 5. Regions in the new range-diversity plot and their representations in the geography. There is a consistent relationship between clusters of cells in the diagram and geographic regions. Blocks of points highlighted in the diagram were selected randomly.

Table 1. Interpretation of dispersion field and significance values in the range-diversity plot.

| Place in the null distribution | Low | Intermediate | High |
|---------------------------------------|--------------------------------|--------------------------------|--------------------------------|
| Below 2.5% CI of random expectations | High endemism | High endemism | High endemism |
| | Low richness | Intermediate richness | High richness |
| Within 95% of random expectations | Cannot say much about endemism | Cannot say much about endemism | Cannot say much about endemism |
| Above 97.5% CI of random expectations | Low endemism | Low endemism | Low endemism |
| | Low richness | Intermediate richness | High richness |

it is possible to have a large value of the dispersion field (relative to all in the matrix) that nevertheless is significantly low, for cells (rows in the matrix) with many species; and the opposite, relatively low values of the dispersion field can be significantly high, in cells with a small number of species. In other words, high values of the dispersion field can be identified to be below random expectations (therefore highly endemic) because the high value of this index is directly related to a high species richness in that particular cell, and vice versa (Table 1, Fig. 4). This last point is of special interest as it suggests that interpretations of low or high values of the dispersion should be more appropriate if accompanied by considerations of species richness patterns. We anticipate that the plot we proposed here, and the geographic projection of its results can aid in promoting better interpretations of patterns of endemism in a region.

We showed that, for our data, the partition of dispersion fields in the upper, middle, and lower quantile ranges has a clear geographic interpretation. This is a fact first noticed by Soberón & Ceballos (2011) and used by Villalobos et al. (2013a). The geographic pattern of significant values is dependent on the region of interest, as species composition changes and such changes have different patterns in distinct areas of the planet. Further explorations using the range-diversity plot and its projections will provide more information on how the patterns of endemism are rescued from PAMs in distinct regions of the world (Soberón and Ceballos 2011), or in different parts of a phylogeny (Villalobos et al. 2017). For instance, the geographic projection of points with high values of both indices in the range-diversity plot (Fig. 5) corresponds with areas that have been described to harbor a large number of species, which ranges are among the most restricted of Mexican avifauna (Villalobos et al. 2013b).

With the formulation of the new range-diversity plots, we also introduce software to prepare such fig-

ures and to randomize PAMs. This implementation is parallelized, and depending on the features of the computer allows randomization of large matrices, with millions of elements (i.e., thousands of cells and species). We anticipate that the software provided will help in the exploration of the effects of randomizing PAMs and the implications of the results that can be obtained for different regions of the planet. As in any other analysis, the quality of input data is one of the most important factors to consider. Researchers should acknowledge limitations deriving from data quality and determine appropriate features of the PAMs to be created. Selecting appropriate cell sizes in the PAM, for instance, can help to avoid over-interpretations of resulting range-diversity plots.

ACKNOWLEDGMENTS

The authors would like to thank the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), especially Raul Sierra, for facilitating access to data. Fabricio Villalobos and an anonymous reviewer gave useful suggestions. The authors declare that no competing interests exist.

DATA ACCESSIBILITY STATEMENT

All data and code used in this research are openly available at <https://github.com/jsoberon/PAMs-Mexico>

REFERENCES

- Arita, H. T., J. A. Christen, P. Rodríguez, and J. Soberón. 2008. Species diversity and distribution in presence-absence matrices: mathematical relationships and biological implications. *American Naturalist* 172:519–532.
- Borregaard, M. K., and C. Rahbek. 2010. Dispersion fields, diversity fields and null models: uniting range sizes and species richness. *Ecography* 33:402–407.
- Brownrigg, R., T. P. Minka, A. Deckmyn, R. A. Becker, and A. R. Wilks. 2018. maps: Draw geographical

- maps. R package. <https://cran.r-project.org/web/packages/maps/index.html>
- Christen, J. A., and J. Soberón. 2009. Anidamiento y los análisis Rq y Qr en PAM's. *Micelánea Matemática* 49:51–61.
- Connor, E. F., and D. Simberloff. 1979. The assembly of species communities: chance or competition? *Ecology* 60:1132–1140.
- Gibert, C., and G. Escarguel. 2019. PER-SIMPER—A new tool for inferring community assembly processes from taxon occurrences. *Global Ecology and Biogeography* 28:374–385.
- Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. *Ecology* 81:2606–2621.
- Graves, G. R., and C. Rahbek. 2005. Source pool geometry and the assembly of continental avifaunas. *PNAS* 102:7871–7876.
- IUCN. 2020. The IUCN Red List of Threatened Species. Version 2020-6.2.
- Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463–1464.
- Miklós, I., and J. Podani. 2004. Randomization of presence-absence matrices: comments and new algorithms. *Ecology* 85:86–92.
- Núñez-Penichet, C., M. E. Cobos, A. T. Peterson, J. Soberón, N. Barve, V. Barve, and T. Gueta. 2020. biosurvey: tools for biological survey planning. R package. <https://github.com/claununez/biosurvey>.
- R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Sarkar, S. 2002. Defining “biodiversity”; assessing biodiversity. *Monist* 85:131–155.
- Schluter, D. 1984. A variance test for detecting species associations, with some example applications. *Ecology* 65:998–1005.
- Soberón, J., and J. Cavner. 2015. Indices of biodiversity pattern based on presence-absence matrices: a GIS implementation. *Biodiversity Informatics* 10:22–34.
- Soberón, J., and G. Ceballos. 2011. Species richness and range size of the terrestrial mammals of the world: biological signal within mathematical constraints. *PLoS ONE* 6:e19359.
- Strona, G., W. Ulrich, and N. J. Gotelli. 2018. Bi-dimensional null model analysis of presence-absence binary matrices. *Ecology* 99:103–115.
- Ulrich, W., and N. J. Gotelli. 2012. A null model algorithm for presence-absence matrices based on proportional resampling. *Ecological Modelling* 244:20–27.
- Villalobos, F., R. Dobrovolski, D. B. Provete, and S. F. Gouveia. 2013a. Is rich and rare the common share? Describing biodiversity patterns to inform conservation practices for South American anurans. *PLoS ONE* 8:e56073.
- Villalobos, F., A. Lira-Noriega, J. Soberón, and H. T. Arita. 2013b. Range-diversity plots for conservation assessments: using richness and rarity in priority setting. *Biological Conservation* 158:313–320.
- Villalobos, F., A. Lira-Noriega, J. Soberón, and H. T. Arita. 2014. Co-diversity and co-distribution in phyllostomid bats: evaluating the relative roles of climate and niche conservatism. *Basic and Applied Ecology* 15:85–91.
- Villalobos, F., M. Á. Olalla-Tárraga, M. V. Cianciaruso, T. F. Rangel, and J. A. F. Diniz-Filho. 2017. Global patterns of mammalian co-occurrence: phylogenetic and body size structure within species ranges. *Journal of Biogeography* 44:136–146.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279–338.
- Zwiener, V. P., A. Lira-Noriega, C. J. Grady, A. A. Padiál, and J. R. S. Vitule. 2018. Climate change as a driver of biotic homogenization of woody plants in the Atlantic Forest. *Global Ecology and Biogeography* 27:298–309.