Global Ecology and Biogeography

A
Mac

From zero to infinity: minimum to maximum diversity of the planet by spatio-parametric Rao's quadratic entropy

Journal:	*Global Ecology and Biogeography*
Manuscript ID	GEB-2020-0304.R2
Manuscript Type:	Macroecological Methods
Keywords:	biodiversity, ecological infomatics, modelling, remote sensing, satellite imagery, spatial ecology

From zero to infinity: minimum to maximum diversity of the planet by spatio-parametric Rao's quadratic entropy

January 9, 2021

Abstract

Aim: The majority of work done to gather information on Earth diversity has been carried out by in-situ data, with known issues related to epistemology (e.g., species determination and taxonomy), spatial uncertainty, logistics (time and costs), among others. An alternative way to gather information about spatial ecosystem variability is the use of satellite remote sensing. It works as a powerful tool for attaining rapid and standardized information. Several metrics used to calculate remotely sensed diversity of ecosystems are based on Shannon's Information Theory, namely on the differences in relative abundance of pixel reflectances in a certain area. Additional metrics like the Rao's quadratic entropy allow the use of spectral distance beside abundance, but they are point descriptors of diversity, namely they can account only for a part of the whole diversity continuum. The aim of this paper is thus to generalize the Rao's quadratic entropy by proposing its parameterization for the first time.

Innovation: The parametric Rao's quadratic entropy, coded in R, i) allows to represent the whole continuum of potential diversity indices in one formula, and ii) starting from the Rao's quadratic entropy, allows to explicitly make use of distances among pixel reflectance values, together with relative abundances.

Main conclusions: The proposed unifying measure is an integration between abundance- and distance-based algorithms to map the continuum of diversity given a satellite image at any spatial scale. Being part of the `rasterdiv` R package, the proposed method is expected to ensure high robustness and reproducibility.

Keywords: biodiversity; ecological informatics; modelling; remote sensing; satellite imagery.

1 Introduction

Since Alexander von Humboldt (1769-1859), the spatial component of nature has played a relevant role in natural science. In the development of theoretical and empirical models in ecology, spatial structure represents a key concept to allow scientists to link ecological patterns to the generating processes and to the functional networking among organisms (Borcard & Legendre, 2002). The majority of the work done to gather information about Earth diversity has been carried out by in-situ data, with known issues related to epistemology (e.g., species determination and taxonomy), spatial uncertainty, logistics (time and costs), among others (Rocchini et al., 2011).

Using satellite remote sensing can at least help attaining rapid and standardized information about Earth diversity (Gillespie, 2005; Rocchini et al., 2005). Furthermore, remote sensing can also be used to monitor some ecosystem functions and parameters such as temperatures, precipitation, photosynthesis, vegetation biomass production and precipitation (Schimel et al., 2019; Zellweger et al., 2019) that can be useful to define the different niches of in-situ species, following Goodall (1970) ideas, who envisaged future diversity measures as those based on niche theory (Hutchinson, 1959). The free access to remote sensing data (see Zellweger et al., 2019) has opened new ways to study ecosystem diversity and biodiversity issues (Rocchini et al., 2013). The spectral data related to pixels, as operational geographical units, are descriptions of pieces of land that allow us to define a new kind of Earth "diversity", which may complement in-situ biodiversity measurement (Hernandez-Stefanoni et al., 2012).

Diversity varies with area, thus investigating multiple spatial grains, until wide extents, is important to effectively monitor spatial diversity change in space and time (MacArthur et al., 1966). This is especially true in macroecology, where the primary aim

is to model large-scale spatial patterns to infer the ecological processes which generated them, particularly considering the recent effect of global changes worldwide (Hobohm et al., 2019). In order to determine the horizontal distribution of diversity within a satellite image (i.e. which areas within the image are more diverse than others), diversity indices are usually spatially referenced by calculating the index within a moving window.

Several metrics that measure diversity from satellites rely on the Shannon's theory of entropy (Shannon, 1948), with diversity being measured as $H = -\sum_{i=1}^{N} p_i \log p_i$, where p_i is the proportion of the i-th pixel value (e.g., digital number, DN) found within a moving window containing N pixels. Shannon's H basically summarizes the partition of abundances (*sensu* Whittaker, 1965) by taking into account both relative abundance and richness of DNs (Figure 1).

However, Shannon's entropy is a point descriptor of (remotely sensed) diversity. As such, it shows only one part of the whole potential diversity spectrum at a glance. The use of generalized entropies has been advocated to face such problem. In this case, one single formula represents a parameterized version of a diversity index, thus providing a continuum of potential diversity indices. In the context of the measurement of diversity, the Rényi (1970) parametric entropy

$$H_\alpha = \frac{1}{1-\alpha} \log \sum_{i=1}^{N} p_i^\alpha \tag{1}$$

with $0 \leq \alpha \leq \infty$ represents a powerful tool to account for the continuum of diversity (Figure 1).

One particularly convenient property of H_α is that by varying the parameter α there is a continuum of possible diversity measures, which differ in their sensitivity to rare and

abundant DNs, becoming increasingly dominated by the most common DNs for increasing values of α. Note that for $\alpha \to 1$, H_1 equals the Shannon's entropy. A similar formulation was then proposed by Hill (1973) who expressed parametric diversity as the "numbers equivalent" of Rényi generalized entropy (Appendix S1).

Rényi (and Hill) parametric functions summarize diversity by taking into account the pixel values of a satellite image and their relative abundances. However, they do not allow to explicitly consider the differences among these values. As an example, two arrays of 9 pixels with maximum richness and evenness (i.e. both containing 9 different DNs with relative abundances $p_i = \frac{1}{9}$) but differing in their values will attain the same Shannon diversity irrespective of the values of the DNs in both arrays.

By introducing a distance parameter d_{ij} among each pair of values i and j, Rao's quadratic entropy (Rao , 1982)

$$Q = \sum_{i,j=1}^{N} p_i p_j d_{ij} \tag{2}$$

explicitly considers the differences among the pixel values in the calculation of diversity (Figure 1). Hence, two different pixels with values [2,3] will attain a lower diversity with respect to two pixels with values [0,100]. For instance, to make an ecological parallel, this is somewhat similar to the phylogenetic distance between two species: the values [2,3] would be equivalent to two sister species closely related on the tree of life while [1,100] would be equivalent to two very distant species on the tree of life.

The aim of this paper is thus to propose, for the first time, a parameterization of Rao's quadratic entropy in order to provide a generalized entropy which accounts for both relative abundances and distances among pixel values. The proposed approach is

now part of the `rasterdiv` R package, a package dedicated to diversity measures of spatial matrices, increasing its capability to discern among different diversity measures by a single formula.

2 Spatio-parametric Rao's quadratic entropy

Inter-pixel spectral distances are directly related to landscape heterogeneity and they are capable of describing species habitats, starting with a satellite image (Rocchini et al., 2005). A satellite image can be viewed as a matrix of numbers describing Earth reflectance in different dimensions stored as pixels. A sensor per each light wavelength records the reflectance of a certain object in that wavelength which are stored into numbers in a certain range (e.g., digital numbers in 8 bits, ranging from 0 to 255). In general, the higher the variability in the spectral space defined by the pixel reflectance values, the higher the diversity of the ecosystem under study.

Consider a window of N pixels moving across the whole image to calculate a diversity index. Let i and j be two pixels randomly chosen with repetition within the moving window. Let d_{ij} be a symmetric measure of the (multi)spectral distance between i and j such that $d_{ij} = d_{ji}$ and $d_{ii} = 0$. Rao's Q (Rao , 1982) is defined as:

$$Q = \sum_{i,j=1}^{N} p_i p_j d_{ij} = \sum_{i,j=1}^{N} \frac{1}{N} \times \frac{1}{N} d_{ij} \tag{3}$$

Therefore, Q measures the expected (i.e. mean) distance between two randomly chosen pixels and $\frac{1}{N}$ is the probability to extract each pixel. Note that, unlike H_α or K_α the calculation of Rao's quadratic entropy is not limited to single bands but can

be extended to multispectral systems of any dimension. For the connection between quadratic entropy and variance, see Rocchini et al. (2019).

A more direct approach for developing a parametric version of quadratic entropy stems from the work of Guiasu & Guiasu (2011). Let $\omega_{ij} = \frac{1}{N} \times \frac{1}{N}$ be the combined probability of selecting pixels i and j in this order. Guiasu & Guiasu (2011) noted that Rao's Q can be expressed as a linear function of the combined probabilities of all pairs of pixels:

$$Q = \sum_{i,j=1}^{N} \omega_{ij} d_{ij} = \sum_{i,j=1}^{N} \frac{1}{N} \times \frac{1}{N} d_{ij} = \sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij} \quad (4)$$

In practice, Rao's Q is the arithmetic mean of the distances d_{ij} between all pairs of pixels i and j. Hence, in order to implement a parametric version of Rao's Q, it seems natural to substitute the arithmetic mean in Equation 4 with a generalized mean (Hardy et al., 1952):

$$Q_\alpha = \left(\sum_{i,j=1}^{N} \omega_{ij} d_{ij}^{\alpha} \right)^{\frac{1}{\alpha}} = \left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^{\alpha} \right)^{\frac{1}{\alpha}} \quad (5)$$

This operation connects Q_α with other diversity metrics that are expressed as generalized means, such as Hill's (Hill, 1973) or Jost's (Jost , 2006) numbers (Appendix S1) equivalents (see also Leinster & Cobbold, 2012).

The Rao's Q, viewed as an arithmetic mean, is one of all the possible means in its generalized form Q_α:

$$Q_\alpha = \begin{cases} \alpha \to 0, Q_0 = \sqrt[N^2]{\prod_{i,j=1}^{N} d_{ij}} & \textit{geometric} \\ \alpha = 1, Q_1 = Q = \sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij} & \textit{arithmetic} \\ \alpha = 2, Q_2 = \sqrt{\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^2} & \textit{quadratic} \\ \alpha = 3, Q_3 = \sqrt[3]{\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^3} & \textit{cubic} \\ \alpha \to \infty, Q_{\alpha\to\infty} = \max d_{ij} & \textit{max}_d \end{cases} \quad (6)$$

The mathematical proof that i) for $\alpha \to 0$ Q_0 corresponds to the geometric mean, and ii) for $\alpha \to \infty$ Q_∞ corresponds to the maximum distance between pixel values pairs is provided in Appendix S1.

Each generalized mean always lies between the smallest and largest of its values. Increasing the parameter α will increase the weight of the highest values of d_{ij}, thus providing a continuum of potential diversity indices (Figure 1).

3 The algorithm

Starting from a satellite image, a spatial moving window might be used to make the calculation on predefined extents of analysis. The grain (*sensu* Dungan et al., 2002) will be the resolution of the image while the extent of analysis will be the size of the moving window (see Figure S1 in Appendix S2). The calculation is based on a distance matrix

of type:

$$M_d = \begin{pmatrix} d_{\lambda_1,\lambda_1} & d_{\lambda_1,\lambda_2} & d_{\lambda_1,\lambda_3} & \cdots & d_{\lambda_1,\lambda_n} \\ d_{\lambda_2,\lambda_1} & d_{\lambda_2,\lambda_2} & d_{\lambda_2,\lambda_3} & \cdots & d_{\lambda_2,\lambda_n} \\ d_{\lambda_3,\lambda_1} & d_{\lambda_3,\lambda_2} & d_{\lambda_3,\lambda_3} & \cdots & d_{\lambda_3,\lambda_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{\lambda_n,\lambda_1} & d_{\lambda_n,\lambda_2} & d_{\lambda_n,\lambda_3} & \cdots & d_{\lambda_n,\lambda_n} \end{pmatrix} \quad (7)$$

among all the potential pairs of pixels inside the moving window. The diagonal terms of the matrix (which equal zero) will have no effect for $\alpha > 0$ (Equation 6), since they would enter the $\sum$ term. On the contrary, for $\alpha \to 0$, they would enter the $\prod$ term by nullifying Q_0.

We coded the proposed parameterization of Rao's quadratic entropy as an R function, implementing the previously developed `rasterdiv` package (Marcantonio et al. (2020), `https://CRAN.R-project.org/package=rasterdiv`). The calculation of different Q_α by automatically changing the range of potential α values is done by the function `paRao`, as:

```
> paRao(x, alpha=c(0:4,Inf), method="classic",
 dist_m="euclidean", window=9, na.tolerance=0.5, simplify=3,
 np=8, cluster.type="SOCK", diag=TRUE)
```

where `x` is the input dataset which can be a `RasterLayer` or a matrix class object, `alpha` is the α parameter of Equation 5, which can be a single value or a vector of integers. In the example above, α is a vector of integers ranging from 0 to 4, plus Inf, which in the R language is a reserved word representing positive infinity ($\alpha \to \infty$). The option `method` decides if paRao is calculated with 1 single layer (`classic`) or with more than one layer (`multidimension`). With `method="multidimension"` then `x` must be a list of

objects. `dist_m` is the type of distance considered in the calculation of the index, and can be set to any distance class implemented in the R package `proxy`, such as `"euclidean"`, `"canberra"` or `"manhattan"`. Moreover, `dist_m` can also be an user-defined matrix of distances. However, if `method` is set to `"classic"` (unidimensional paRao) all distance types reduce to the Euclidean distance. The argument `window` is the side length in cells of the moving window (in this case set to 9), whereas `na.tolerance` is the proportion (0-1) of NA's cell allowed in a moving window: if the proportion of NA's cells in a moving window exceeds `na.tolerance` then the value of the moving window central pixel will be NA. The option `simplify` allows to reduce the number of decimal places to ease the calculation by reducing the number of numerical categories, i.e., if `simplify=3` only the first three digits of data will be considered for the calculation of the index. `np` is the number of parallel processes used in the calculation. If `np`>1 then the `doParallel` package will be called for parallel calculation, and `cluster.type` will indicate the type of cluster to be opened (default is `"SOCK"`, `"MPI"` and `"FORK"` are the alternatives). The `diag` argument refers to the diagonal term of Equation 7. It will have no effect on the function for $\alpha > 0$, while it will nullify the value of Q_α if set to `TRUE`, as previously explained in Equation 7.

In the next section we provide an ecological example of the application of the spatio-paramteric Rao's Q, done at regional scale to allow a comparison with in-situ data on species diversity. Appendix S2 provides an example at worldwide spatial scale.

3.1 Study case: the diversity of vegetation greenness and the ecoregions of California

A comparison between in-situ and remotely sensed diversity at worldwide scale might be difficult due to known biases in e.g. sampling effort, taxonomies, spatial uncertainty (Rocchini et al., 2017). Hence, we decided to calculate the parametric Rao's Q index on a NDVI raster layer of California (USA) to be compared with data in the field on native plant species diversity provided in Thornhill et al. (2017) from Baldwin et al. (2017). We chose California as a case study due to its high ecological diversity as well as to the availability of plant species field-data for this region.

In practice, we aimed at visualizing and describing differences in both diversity and structure of vegetation for the state of California, USA. First, an NDVI raster layer was derived from Copernicus Sentinel-2 data (European Space Agency, reference period: January 2017 to July 2018) and processed through Google Earth Engine to filter out cloud cover, select the greenest pixel of the time series and resample at 100 m pixel resolution. Then, the `paRao` R function was used to derive the Rao's Q index, considering both the original formulation of the Rao's Q ($\alpha = 1$, Equation 6) and the formulation with $\alpha \to \infty$ maximuzing β-diversity (Appendix S2), with a moving window of 9x9 pixels.

A map of plant species richness was derived relying on the potential distribution range of 5,222 native California vascular plants modelled by Thornhill et al. (2017) and used to report alpha-diversity (plant species richness per pixel), whereas a vector map of the ecoregions of California (level III; from the United States Environmental Protection Agency) was used to identify beta-diversity. In Figure 2, we showed NDVI, the Rao's Q indices with $\alpha = 1$ and $\alpha \to \infty$ and native plant species richness, reporting the boundaries of the different ecoregions for California. This comparison revealed macro-ecological and

bio-geographical patterns which can be better interpreted considering the information condensed in the Rao's Q index (Figure 3). We focused on five important ecoregions, which were used in Figure 2 and in the boxplots of Figure 3 to discuss the most evident ecological patterns.

Overall, the Rao's Q index matched native plant species diversity in areas characterized by a lower human impact, such as in the northwest and southeast portions of the state. The "Cascades" (CAS) and "Klamath Mountains" (KM) ecoregions are characterized by highly dissected ridges, foothills, and valleys and host a very diverse flora, rich in endemic and relic species (Griffith et al., 2016). Rao's Q was the highest in these two areas (apart for the agricultural and urbanized areas described later), despite moderate NDVI values (Figures 2 and 3). The low mountains covered by highly productive, rain-drenched evergreen forests composing the "Coast Range" (CR) ecoregion showed lower plant diversity and Rao's Q, despite very high NDVI values. By contrast, the dry and warm "Mohave basin and range " (MBR) is characterized by broad basins and scattered mountains and showed low NDVI, Rao's Q and plant species richness.

On the other hand, Rao's Q performed poorly when compared with native plant species richness in agricultural and developed lands with a high productivity (i.e., high NDVI) and heterogeneity. This is the case of the "California Central Valley" (CCV) region, which is composed of flat, urbanized and intensively farmed plains. The extensive presence of irrigated crops intersected with urbanized areas caused medium to high NDVI values and an apparently high Rao's Q diversity, but a low native species richness, especially in the drier southern portion of the valley.

Passing from the pure Rao's Q index (α=1) to its parameterization with $\alpha \to \infty$ helped to increase the discrimination among areas, due to the fact that when $\alpha \to \infty$

the Rao's Q corresponds to the maximum distance (β-diversity) among pixel values in a site. Very similar gradients of the spatial heterogeneity of California (including BIOMOD variables, NDVI, elevation) as well as environmental DNA (eDNA) data are found in Lin et al. (2020).

4 Discussion

In this paper, we provided a straightforward solution to: i) account for distances in an Information Theory based metric, and ii) provide a generalized formula in order to avoid point description and account for the continuum of diversity. Diversity can be represented by different dimensions (Nakamura et al., 2020). Considering one single metric to account for the whole continuum of diversity metrics might be a powerful addition to the main framework. On the contrary, fragmenting the concept of diversity when trying to capture single aspects of the whole spectrum could be counterproductive.

The proposed unifying measure succeeded to integrate abundance- and distance-based algorithms over a wide variety of diversity metrics. We demonstrated that such integration is not only theoretical but also applicable to real spatial data, considering several dimensions of diversity at the same time. Being part of the `rasterdiv` R package, the proposed method is expected to ensure high robustness and reproducibility.

Remote sensing is obviously not a panacea for all the organismic based diversities like taxonomic-, functional-, genetic-diversity but it can represent an important exploratory tool to detect diversity hotspots and their changes in space and time at the ecosystem level. First of all, it measures heterogeneity of the environment with indirect links to the biodiversity of both plant and animal taxa, but also with potential discrepancies with species diversity, as in the presented case study of the native plant species diversity

of California. This said, depending on the complexity and the resolution at which the proposed parameterized Rao's Q is applied, it might allow finding new insights on the ecological processes acting in a certain ecosystem to shape its diversity. In this paper, the examples provided were based on a single NDVI layer since i) it is a valuable index of vegetation health and ii) it is freely available in the `rasterdiv` package to reproduce the code proposed in this paper (see Appendix S2 for an application at worldwide spatial scale based on the Copernicus Proba-V NDVI freely available in the package). We are aware that NDVI has very limited capacity to track diversity in some habitats like dense forests, because it is saturated at dense vegetation (Mutanga & Skidmore, 2004). From this point of view, imaging spectroscopy offers higher information content, also enabling plant functional trait retrievals (Jetz et al., 2016; Schneider et al., 2019) as well as structural traits by LiDAR data (Schneider et al., 2020). The application of the proposed algorithm to future spaceborne imaging spectroscopy is promising. In other words, the algorithm has been thought to be used with multiple layers, like a whole multispectral image or the most meaningful Principal Components (Peres-Neto et al., 2005), or land use classes probabilities derived from fuzzy set theory (Rocchini & Ricotta, 2007; Feoli, 2018). This is even one of the major advantages of the Rao's Q metric which allows considering both abundance and distance among pixel values, thus being applicable to any continuous raster layer, or to any matrix combination, even in a multiple spectral system.

Creating a unique "umbrella" under which all of the potential metrics of diversity can be used is highly beneficial for e.g. monitoring the variation in time of biological systems considering two major axes: i) the α parameter in Equation 5 providing information about the type of diversity at time t_0, ii) the temporal dimension from time t_0 to time t_n given the same α parameter. For the future, exploring such temporal dimension would

allow gathering information of ecosystem changes in different diversity types at a glance.

Moreover, generalized entropy allows us to characterize the dimensionality of diversity (*sensu* Stevens & Tello, 2014) of different habitats/ecosystems. Those areas with a higher diversity dimensionality, namely a higher variability into the diversity spectrum would need a generalized measure to be fully undertaken. On the contrary, ecosystems with a lower dimensionality would have a lower difference among the different diversity measures with a flat curve of the diversity spectrum (Nakamura et al., 2020).

From a functional point of view, when all indices of diversity are highly correlated to each other (low dimensionality), it is expected that the ecological processes underlying diversity are just a few. On the contrary, with a lower correlation among indices (higher dimensionality) there might be a higher number of axes of variation coming out from different processes shaping ecological heterogeneity in space (Stevens & Tello, 2014).

There might be the possibility that a completely random matrix produces a pattern of diversity (Type I error). On the other side, a structured matrix could produce a very low diversity pattern (Type II error, Gotelli (2000)). In both cases, the parametric Rao's Q could allow to determine, thanks to the use of a continuum of diversities, i) why a diversity pattern is still produced even in case of a random matrix, and ii) why a certain landscape shows a very low diversity in a certain point of the whole diversity spectrum. With point descriptors of diversity such inference cannot be done since the investigation is limited to a small window of the entire diversity spectrum, by basically relying on a single final number. In other words, the commonly asked question about what is the index which best describes diversity has no certain answer (Gorelick, 2011). Hence, the use of a trend of diversities will lead to the comprehension of hidden parts of the whole diversity dimensionality.

Furthermore, it is expected that the ecological processes shaping diversity should act at defined spatial scales (Borcard & Legendre, 2002). Hence, different diversity types of the whole dimensionality spectrum are expected to show scale dependent patterns, being apparent only at certain scales and not at some others. The use of a continuum allows measuring the different diversity types altogether in a single step. Changing the extent of analysis by different moving windows would then allow to encompass different spatial structures at different scales.

While geographic gradients of diversity over space are complex to catch in their very nature, biodiversity measurement has mainly relied in the past on few formulas which represented an hegemony (Stevens et al., 2013). In this paper, we demonstrated that diversity is actually multifaceted and should be necessarily approached through a generalized approach. Furthermore, the proposed generalized parametric Rao's Q index might be profitably used to plot multitemporal trends (see e.g. Dornelas et al., 2014) of diversity metrics and discover previously imperceptible differences when making use of single metrics (Figure 4).

5 Conclusion

In order to unfold the dimensionality of diversity methods to directly account for several aspects of diversity at a time are needed. From this point of view, generalized entropy undoubtedly represents a powerful approach for mapping the diversity continuum.

Metrics grounded in Information Theory ensure to make use of relative abundance of pixel values given the same richness in the moving window of analysis. However, distance metrics allow to also account for the relative dispersion in the spectral space of the cloud of pixels in a certain area (Laliberté et al., 2020). The proposed parameterization of the

Rao's Q explicitly considers the dispersion of pixel values in a spectral space (and their relative abundance) by allowing catching the whole dimensionality of diversity.

6 Data availability

The code and the data used in this paper are based on completely Free and Open Source Software, and they are available at the CRAN repository of the R package `rasterdiv`: `https://CRAN.R-project.org/package=rasterdiv`.

References

Baldwin, B. G., Thornhill, A. H., Freyman, W. A., Ackerly, D. D., Kling, M. M., Morueta-Holme, N., Mishler, B. D. (2017). Species richness and endemism in the native flora of California. *American Journal of Botany*, 104, 1-15.

Borcard, D., Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, 153, 51-68.

Dornelas, M., Gotelli, N. J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., Magurran, A. E. (2014). Assemblage time series reveal biodiversity change but not systematic loss. *Science*, 344, 296-299.

Evans, M. R., Grimm, V., Johst, K., Knuuttila, T., de Langhe, R., Lessells, C. M., Merz, M., O'Malley, M. A., Orzack, S. H., Weisberg, M., Wilkinson, D. J., Wolkenhauer, O., Benton, T. G. (2013). Do simple models lead to generality in ecology? *Trends in Ecology & Evolution*, 28, 578-583.

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, 252-264.

Dungan, J. L., Perry, J. N., Dale, M. R. T., Legendre, P., Citron-Pousty, S., Fortin, M. -J., Jakomulska, A., Miriti, M. and Rosenberg, M. S. (2002). A balanced view of scale in spatial statistical analysis. *Ecography*, 25, 626-640.

Gillespie, T. W. (2005). Predicting woody-plant species richness in tropical dry forests: a case study from South Florida, USA. *Ecological Applications*, 15, 27-37.

Gorelick, R. (2011). Do we have a consistent terminology for species diversity? The fallacy of true diversity. *Oecologia*, 167, 885-888.

Goodall, D. W. (1970). Statistical ecology, p. 99-124. In Johnston, R.F., ed. *Annual review of ecology and systematics*, Vol. 1. Annual Reviews, Palo Alto, California, USA.

Null model analysis of species co-occurrence patterns. *Ecology*, 81, 2606-2621.

Griffith, G.E., Omernik, J.M., Smith, D.W., Cook, T.D., Tallyn, E., Moseley, K., Johnson, C.B. (2016). *Ecoregions of California*. USGS, Reston, VA. http://pubs.er.usgs.gov/publication/ofr20161021

Hernandez-Stefanoni, J. L., Gallardo-Cruz, J. A., Meave, J. A., Rocchini, D., Bello-Pineda, J., Lopez-Martinez, J. O. (2012). Modeling alpha- and beta-diversity in a tropical forest from remotely sensed and spatial data. *International Journal of Applied Earth Observation and Geoinformation*, 19, 359-368.

Guiasu, R. C., Guiasu, S. (2011). The weighted quadratic index of biodiversity for pairs of species: a generalization of Rao's index. *Natural Science*, 3, 795-801.

Hardy, G., Littlewood, J. E., Polya, G. (1952). Inequalities. *Cambridge University Press*, Cambridge, UK.

Hill, M. O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54, 427-431.

Hobohm, C., Janisova, M., Steinbauer, M., Landi, S., Field, R., Vanderplank, S., Beierkuhnlein, C., Grytnes, J. -A., Vetaas, R. O., Fidelis, A., de Nascimento, L., Clark, V. P., Fernandez-Palacios, J. M., Franklin, S., Guarino, R., Huang, J., Krestov, P., Ma, K., Onipchenko, V., Palmer, M. W., Fragomeni Simon, M., Stolz, C., Chiarucci, A. (2019). Global endemics-area relationships of vascular plants. *Perspectives in Ecology and Conservation*, 17, 41-49.

Hutchinson, G. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? *American Naturalist*, 93, 145-159.

Feoli, E. (2018). Classification of plant communities and fuzzy diversity of vegetation systems. *Community Ecology*, 19, 186-198.

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., Guralnick, R., Kattge, J., Latimer, A. M., Moorcroft, P., Schaepman, M. E., Schildhauer, M. P., Schneider, F. D., Schrodt, F., Stahl, U., Ustin, S. L. (2016). Monitoring plant functional diversity from space. *Nature Plants*, 2, 16024.

Johnson, P. C. D., Barry, S. J. E., Ferguson, H. M., Müller, P. (2015). Power analysis for generalized linear mixed models in ecology and evolution. *Methods in Ecology and Evolution*, 6, 133-142.

Jost, L. (2006). Entropy and diversity. *Oikos*, 113, 363-375.

Laliberté, E., Schweiger, A. K., Legendre, P. (2019). Partitioning plant spectral diversity into alpha and beta components. *Ecology Letters*, 23, 370-380.

Leinster, T., Cobbold, C. A. (2012). Measuring diversity: the importance of species similarity. *Ecology*, 93, 477-489.

Leitão, P. J., Schwieder, M., Suess, S., Catry, I., Milton, E. J., Moreira, F., Osborne, P. E., Pinto, M. J., van der Linden, S., Hostert, P. (2015), Mapping beta diversity from space: Sparse Generalised Dissimilarity Modelling (SGDM) for analysing high-dimensional data. *Methods in Ecology and Evolution*, 6, 764-771.

Lin, M., Levi Simons, A., Curd, E. E., Harrigan, R. J., Schneider, F. D., Ruiz-Ramos, D. V., Gold, Z., Osborne, M. G., Shirazi, S., Schweizer, T. M., Moore, T. N., Fox, E. A., Turba, R., Garcia-Vedrenne, A. E., Helman, S. K., Rutledge, K., Palacios Mejia, M., Munguia Ramos, M. N., Wetzer, R., Pentcheff, D., McTavish, E. J., Dawson, M. N., Shapiro, B., Wayne, R. K., Meyer, R. S. (2020).A biodiversity composition map of California derived from environmental DNA metabarcoding and Earth observation. bioRxiv 2020.06.19.160374. doi: `https://doi.org/10.1101/2020.06.19.160374`

MacArthur, R. H., Recher, H., Cody, M. (1966). On the relation between habitat selection and species diversity. *American Naturalist*, 100, 319-327.

Marcantonio, M., Iannacito, M., Thouverai, E., Da Re, D., Tattoni, C., Bacaro, G., Vicario, S., Rocchini, D. (2020). *rasterdiv: Diversity Indices for Numerical Matrices*. R package version 0.2-0. `https://CRAN.R-project.org/package=rasterdiv`

Mutanga, O., Skidmore, A. K. (2004). Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International Journal of Remote Sensing*, 25, 3999-4014.

Nakamura, G., Gonçalves, L. O., Duarte, L. d. S. (2020). Revisiting the dimensionality of biological diversity. *Ecography*, 43, 539-548.

Patil, G. P., Taillie, C. (1982). Diversity as a concept and its measurement. *Journal of the American Statistical Association*, 77, 548-561.

Peres-Neto, P. R., Jackson, D. A., Somers, K. M. (2005). How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, 49, 974-997.

Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, 21, 24-43.

Rényi, A. (1970). *Probability Theory*. North Holland Publishing Company, Amsterdam.

Rocchini, D., Andreini Butini, S., Chiarucci, A. (2005). Maximizing plant species inventory efficiency by means of remotely sensed spectral distances. *Global Ecology and Biogeography*, 14, 431-437.

Rocchini, D., Delucchi, L., Bacaro, G., Cavallini, P., Feilhauer, H., Foody, G. M., He, K. S., Nagendra, H., Porta, C., Ricotta, C., Schmidtlein, S., Spano, L. D., Wegmann, M., Neteler, M. (2013). Calculating landscape diversity with information-theory based indices: A GRASS GIS solution. *Ecological Informatics*, 17, 82-93.

Rocchini, D., Garzon-Lopez, C. X., Marcantonio, M., Amici, V., Bacaro, G., Bastin, L., Brummitt, N., Chiarucci, A., Foody, G. M., Hauffe, H. C., He, K. S., Ricotta, C., Rizzoli, A., Rosá, R. (2017). Anticipating species distributions: handling sampling effort bias under a Bayesian framework. *Science of the Total Environment*, 584-585, 282-290.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J. M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G., Chiarucci, A. (2011). Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. *Progress in Physical Geography*, 35, 211-226.

Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N., Feilhauer, H., Féret, J. -B., Foody, G. M., Gavish, Y., Godinho, S., Kunin, W. E., Lausch, A., Leitão, P. J., Marcantonio, M., Neteler, M., Ricotta, C., Schmidtlein, S., Vihervaara, P., Wegmann, M., Nagendra, H. (2018). Measuring β-diversity by remote sensing: a challenge for biodiversity monitoring. *Methods in Ecology and Evolution*, 9, 1787-1798.

Rocchini, D., Marcantonio, M., Da Re, D., Chirici, G., Galluzzi, M., Lenoir, J., Ricotta, C., Torresani, M., Ziv, G. (2019). Time-lapsing biodiversity: an open source method for measuring diversity changes by remote sensing. *Remote Sensing of Environment*, 231, 111192.

Rocchini, D., Ricotta, C. (2007). Are landscapes as crisp as we may think? *Ecological Modelling*, 204, 535-539.

Schimel, D., Schneider, F. D., JPL Carbon and Ecosystem Participants (2019). Flux towers in the sky: global ecology from space. *New Phytologist*, 224, 570-584.

Schneider, F. D., Ferraz, A., Schimel, D. (2019). Watching Earth's Interconnected Systems at Work. *Eos*, 100.

Schneider, F. D., Ferraz, A., Hancock, S., Duncanson, L. I., Dubayah, R. O., Pavlick, R. P., Schimel, D. S. (2020). Towards mapping the diversity of canopy structure from space with GEDI. *Environmental Research Letters*, 15, 115006.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423, 623-656.

Stevens, R. D., Tello, J. S. (2014). On the measurement of dimensionality of biodiversity. *Global Ecology and Biogeography*, 23, 1115-1125.

Stevens, R. D., Tello, J. S., Gavilanez, M. M. (2013). Stronger tests of mechanisms underlying geographic gradients of biodiversity: insights from the dimensionality of biodiversity. *PLOS ONE*, 8, e56853.

Thornhill, A. H., Baldwin, B. G., Freyman, W. A., Nosratinia, S., Kling, M. M., Morueta-Holme, N., Madsen, T. P., Ackerly, D. D., Mishler, B. D. (2017). Spatial phylogenetics of the native California flora. *BMC Biology*, 15, 96.

Whittaker, R. H. (1965). Dominance and diversity in land plant communities. *Science*, 147, 250-260.

Zellweger, F., De Frenne, P., Lenoir, J., Rocchini, D., Coomes, D. (2019). Advances in microclimate ecology arising from remote sensing. *Trends in Ecology & Evolution*, 34, 327-341.

Figures

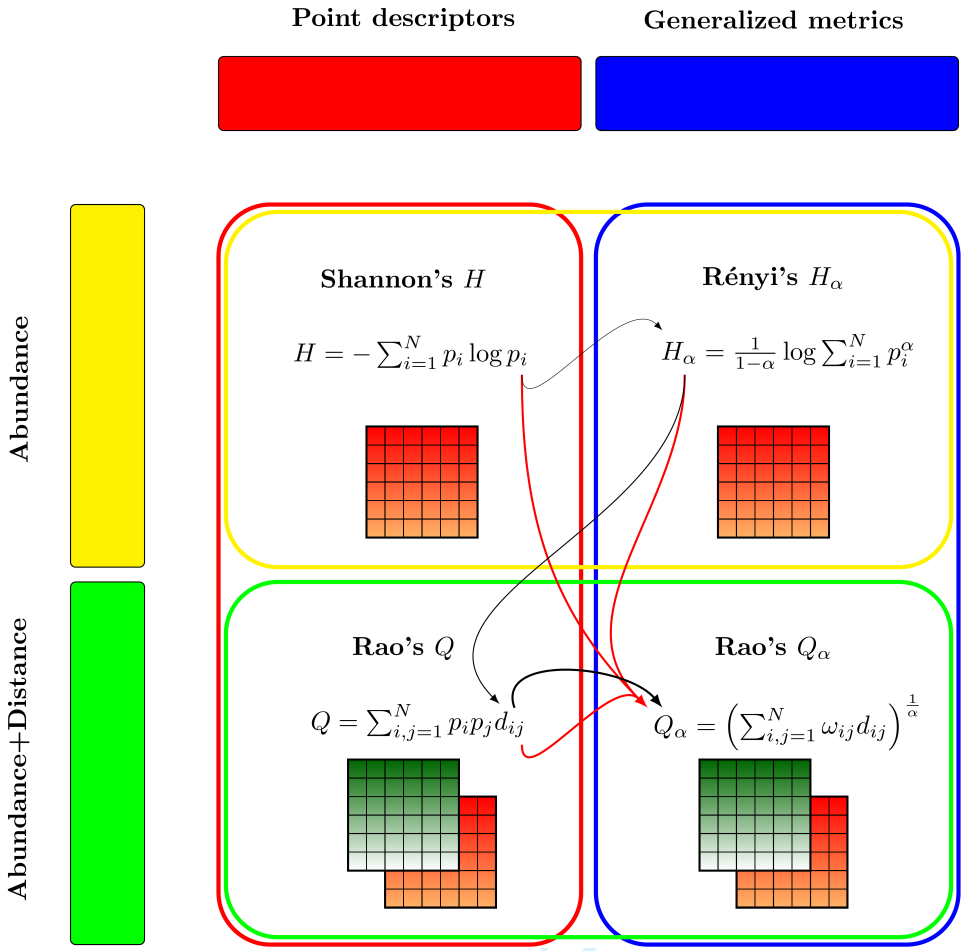


Figure 1: Grounding theory of this paper. Diversity measures can encompass abundance-based as well as abundance-distance-based metrics (yellow and green boxes, respectively). Abundance-distance-based metrics allow multiple layers to be used. The black lines represent the theoretical flow of this paper, with the thickness representing the complexity of each index, starting from Shannon's Information Theory (point descriptor) to Rényi's H_α (generalized entropy), which do not make use of distance. Distance enters the Rao's Q formula, but this is still a point descriptor of diversity. Finally, parametric Rao's Q_α comprises the use of distances and the generalized entropy concept. The red arrows represent the properties of the Rao's Q_α: i) it is grounded in Information Theory starting from Shannon's H, ii) it is a generalized entropy like the Rényi H_α, and iii) it makes use of distances like the Rao's Q.

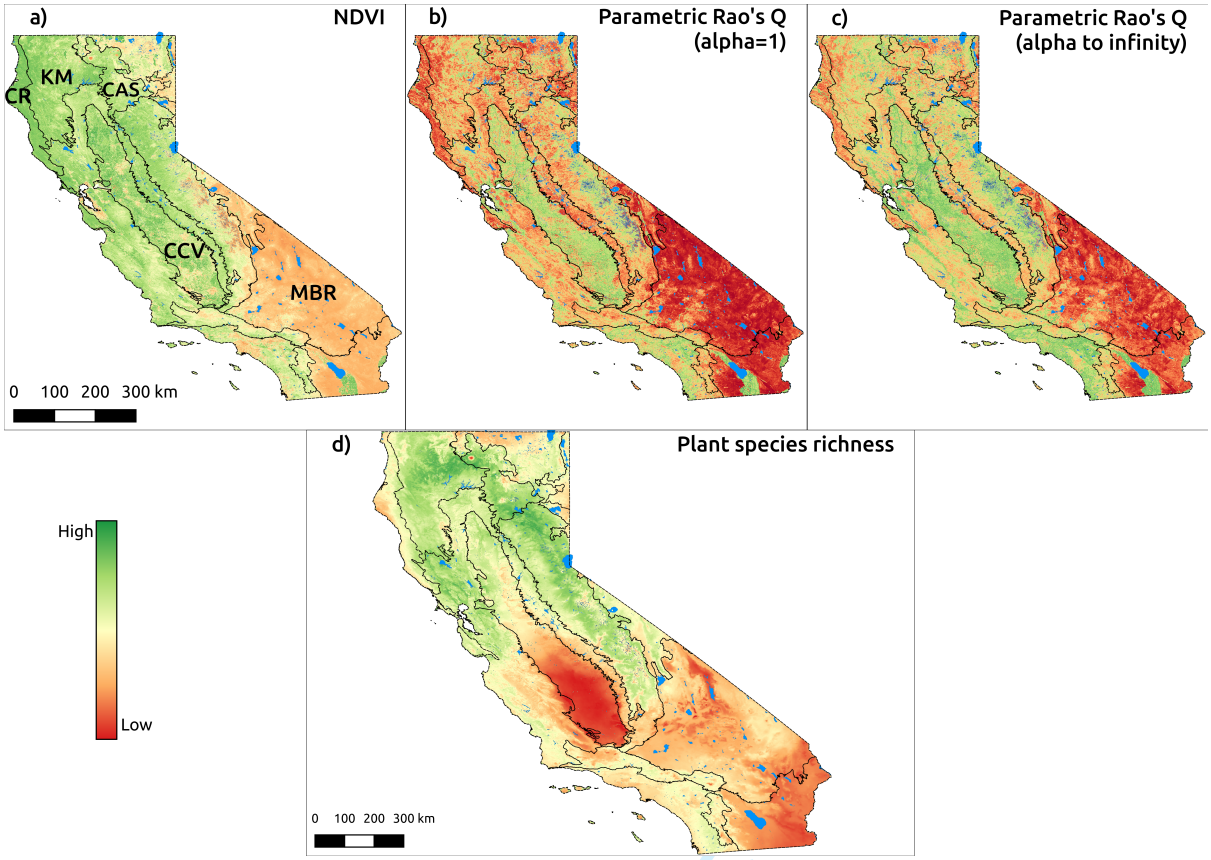


Figure 2: Maps showing NDVI, the Rao's Q Index, native plant species richness for the ecoregions of California. The NDVI values shown in the top-left box (100 m resolution) were derived from the ESA Copernicus Sentinel-2 dataset, processed with Google Earth Engine and range between -0.26 (red) and 0.99 (dark green). The Rao's Q indices in the top-right boxes were calculated from the NDVI map using $\alpha = 1$ and $\alpha \to \infty$ and a moving window of 9x9 pixels. Rao's Q higher values represent pixels whose surrounding NDVI values are more "diverse" than pixels with lower Rao's Q values. The bottom map, reporting the pixels' potential native plant species richness (resolution: 810 m), was derived summing the binary potential distribution ranges of 5,222 native plant species modelled by Thornhill et al. (2017). Species richness ranges between 134 (red) to 1029 (green) species per pixel (1 km^2). Five ecoregions were labelled in the NDVI map and discussed in the main text: Coast range (CR), Klamath Mountains (KM), Cascades (CAS), California Central Valley (CCV) and Mohave basin and range (MBR). A mask with water bodies was added to the maps. Refer to the main text for additional information.

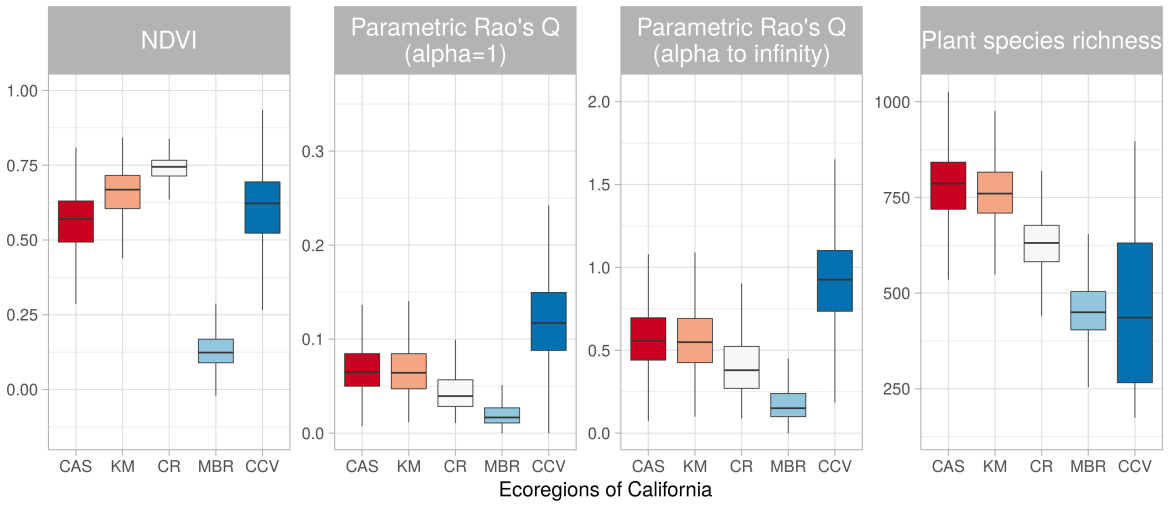


Figure 3: Boxplot showing the values distribution of NDVI, the parametric Rao's Q (with $\alpha = 1$ and $\alpha \to \infty$) and plant species richness. The Rao's Q index matched native plant species diversity in those ecoregions characterized by a lower human impact, such as Cascades (CAS), Klamath Mountains (KM), Coast range (CR), Mohave basin and range (MBR), California Central Valley (CCV). On the other hand, Rao's Q performed poorly when compared with native plant species richness in agricultural and developed lands with a high productivity (i.e., high NDVI) and heterogeneity. See also Figure 2 and the main text for additional information.

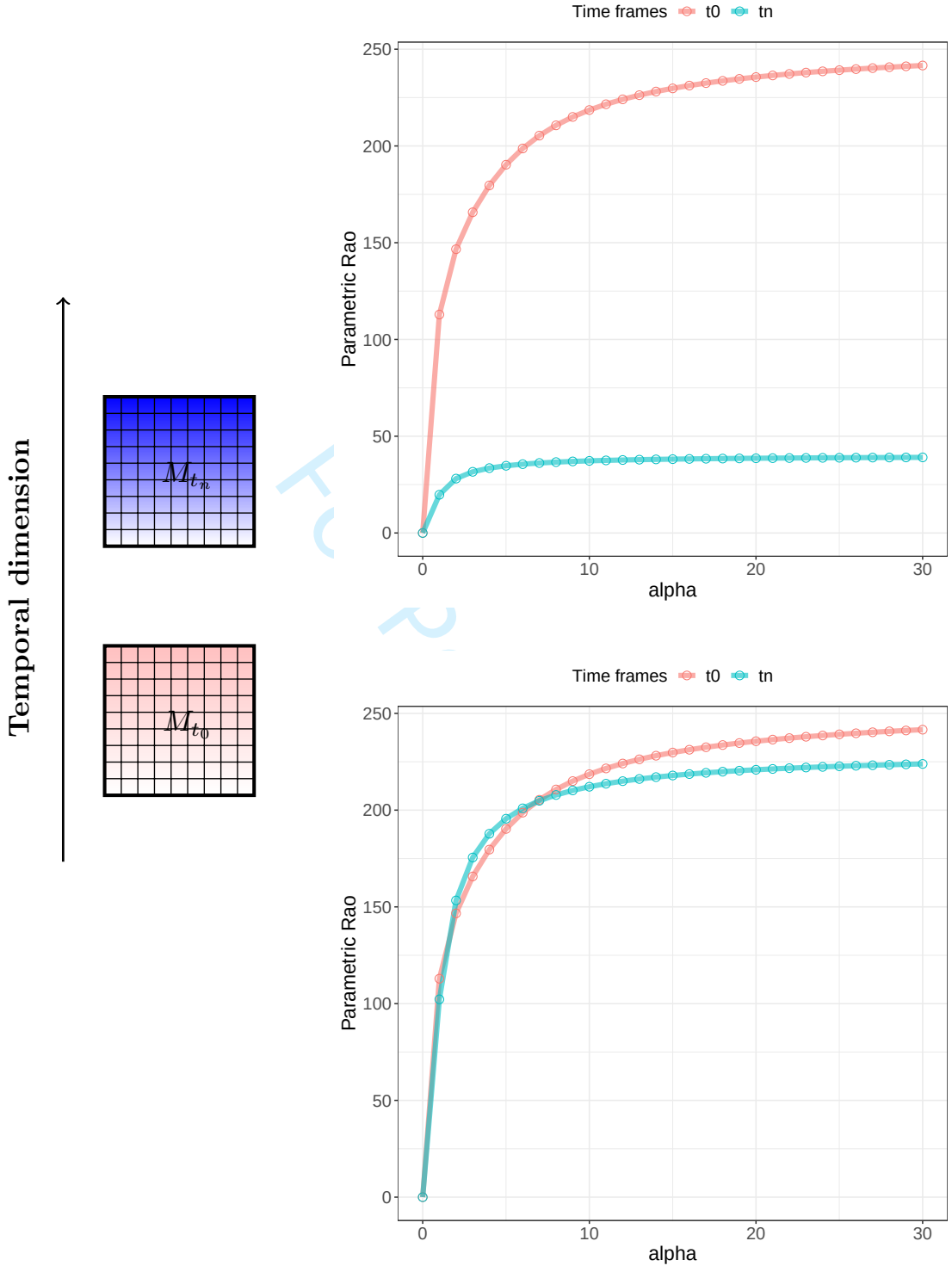


Figure 4: A theoretical example of the power of using generalized entropy for monitoring purposes. Given a landscape at times t_0 (pink) and t_n (blue), calculating generalized entropy will allow the formation of a graph showing the continuum of Rao's Q values observed over a range of values for α. The same landscape in different times might show an abrupt change (e.g., a catastrophic event) with an apparent diversity decrease (top). In this case, point descriptors (e.g., single α values) of diversity may be sufficient to describe this pattern. When the change in diversity is subtle (bottom), using a point descriptor might fail to detect it but it becomes manifest in the continuum of diversities based on generalized entropy. The complete code for reproducing this theoretical example is available in Appendix S3.

Appendix S1 - Mathematical dissertation on the proposed algorithms

From zero to infinity: minimum to maximum diversity of the planet by spatio-parametric Rao's quadratic entropy

January 9, 2021

1 Hill's numbers and generalized entropy

Hill (1973) expressed parametric diversity as the "numbers equivalent" of Rényi's generalized entropy, as:

$$K_\alpha = \frac{1}{\left(\sum_{i=1}^{N} p_i \times p_i^{\alpha-1}\right)^{\frac{1}{\alpha-1}}} \tag{1}$$

where the numbers equivalent K_α is the theoretical number of equally-abundant DNs (i.e. all those with $p_i = \frac{1}{K_\alpha}$) that are needed in order that its diversity be H_α (**?**).

Hill's K_α has the form of the reciprocal of a generalized mean of order $\alpha - 1$. Jost (2006) further showed that, like for H_α, the numbers equivalents of all parametric and non-parametric measures of diversity that can be expressed as monotonic functions of $\sum p_i^\alpha$ have the form of the reciprocal of a generalized mean of order $\alpha - 1$ (for details, Jost , 2006).

2 Mathematical proof: for $\alpha \to 0$ Q_0 is the geometric mean among the generalized means, for $\alpha \to \infty$ Q_∞ is the maximum distance between pixel values pairs

We want to compute

$$\lim_{\alpha \to 0} Q_\alpha \quad \text{where} \quad Q_\alpha = \left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^\alpha\right)^{\frac{1}{\alpha}}. \tag{2}$$

By $\exp(\log(x)) = x$ we can rewrite Q_α as

$$Q_\alpha = \left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^\alpha\right)^{\frac{1}{\alpha}} = \exp\left(\log\left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^\alpha\right)^{\frac{1}{\alpha}}\right) = \exp\left(\frac{1}{\alpha}\log\left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^\alpha\right)\right)$$

reminding that if $N > 1$, there is at least one distance $d_{ij} > 0$. We use this last expression to calculate (2). We use the following two well known results.

Theorem 1 (De l'Hôpital)**.** *Let $f_1, g_1 : (a, b) \mapsto \mathbb{R}$ be two functions such that*

- $\lim_{x\to a} f_1(x) = \lim_{x\to a} g_1(x) = 0$
- *f_1 and g_1 are differentiable in (a, b) with $g_1'(x) \neq 0$ for every $x \in (a, b)$*
- *the limit $\lim_{x\to a} \frac{f_1'(x)}{g_1'(x)} = L$ with $L \in \mathbb{R}$*

then

$$\lim_{x\to a} \frac{f_1(x)}{g_1(x)} = L.$$

Theorem 2 (Limit composition)**.** *Let $f_2 : (a, b) \mapsto \mathbb{R}$ and let $g_2 : (c, d) \mapsto \mathbb{R}$ be two functions such that the image set of g_2 is contained in the domain of f_2, i.e. $\mathcal{I}mg(g_2) \subseteq (a, b)$. Let $x_0 \in (c, d)$, if it holds that*

- $\lim_{x\to x_0} g_2(x) = y_0$ *with* $g_2(x) \neq y_0$ *definitely for* $x \to x_0$
- $\lim_{y\to y_0} f_2(y) = l$

with $a, b, c, d, x_0, y_0, l \in \mathbb{R} \cup \pm\infty$ then

$$\lim_{x\to x_0} (f_2 \circ g_2)(x) = l.$$

We apply Theorem (2) to calculate the limit (2) with $f_2(x) = \exp(x)$ and

$$g_2(\alpha) = \frac{1}{\alpha}\log\left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^\alpha\right).$$

(all assumptions of the theorem hold). Setting $x_0 = 0$, we have to compute

$$\lim_{\alpha \to 0} g_2(\alpha). \tag{3}$$

which will be accomplished using Theorem (1) by setting $f_1 : (0, +\infty) \mapsto \mathbb{R}$

$$f_1(\alpha) = \log\left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^{\alpha}\right)$$

and $g_2 : (0, +\infty) \mapsto \mathbb{R}$, $g_2(\alpha) = \alpha$. Then we have

$$f_1(0) = \lim_{\alpha \to 0} f_1(\alpha) = \log(\frac{1}{N^2} \sum_{i,j=1}^{N} 1) = \log(1) = 0$$

as the limit exists and

$$g_1(0) = \lim_{\alpha \to 0} g_1(\alpha) = 0.$$

Both functions f_1 and g_1 are differentiable. Lastly we observe that $g_1'(\alpha) \equiv 1$. Since all the assumptions of Theorem 1 hold then

$$\begin{aligned} \lim_{\alpha \to 0} \frac{f_1(\alpha)}{g_1(\alpha)} = \lim_{\alpha \to 0} \frac{f_1'(\alpha)}{g_1'(\alpha)} &= \lim_{\alpha \to 0} \frac{(\frac{1}{N^2} \sum_{i,j=1}^{N} d_{ij}^{\alpha})^{-1} (\frac{1}{N^2} \sum_{i,j=1}^{N} d_{ij}^{\alpha} \log d_{ij})}{1} \\ &= \frac{1}{N^2} \sum_{i,j=1}^{N} \log d_{ij} = \sum_{i,j=1}^{N} \log(d_{ij})^{\frac{1}{N^2}} = \prod_{i,j=1}^{N} \log(d_{ij}^{\frac{1}{N^2}}) \end{aligned} \tag{4}$$

By Equation (4) we have the expression of Equation 3. Let

$$y_0 = \prod_{i,j=1}^{N} \log(d_{ij}^{\frac{1}{N^2}})$$

and we conclude by observing

$$\lim_{y \to y_0} \exp(y) = \exp\left(\prod_{i,j=1}^{N} \log(d_{ij}^{\frac{1}{N^2}})\right) = \prod_{i,j=1}^{N} \exp(\log(d_{ij}^{\frac{1}{N^2}})) = \prod_{i,j=1}^{N} d_{ij}^{\frac{1}{N^2}} = \sqrt[N^2]{\prod_{i,j=1}^{N} d_{ij}}.$$

Now we want to compute

$$\lim_{\alpha \to +\infty} Q_\alpha \quad \text{where} \quad Q_\alpha = \left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^{\alpha}\right)^{\frac{1}{\alpha}}$$

We define $d = \max\{d_{ij} | i, j \in \{1, \dots, N\}\}$ and we rewrite Q_α as

$$Q_\alpha = \left(\sum_{i,j=1}^{N} \frac{1}{N^2} d_{ij}^\alpha \right)^{\frac{1}{\alpha}} = \left(\sum_{i,j=1}^{N} \frac{1}{N^2} d^\alpha \left(\frac{d_{ij}}{d} \right)^\alpha \right)^{\frac{1}{\alpha}} = d \left(\sum_{i,j=1}^{N} \frac{1}{N^2} \left(\frac{d_{ij}}{d} \right)^\alpha \right)^{\frac{1}{\alpha}}$$

Next we observe that

$$\frac{d_{ij}}{d} \leq 1$$

by construction and there exist a pair $(\bar{i}, \bar{j})$ such that $\frac{d_{\bar{i},\bar{j}}}{d} = 1$. Therefore it follows that

$$\sum_{i,j=1}^{N} \frac{1}{N^2} \left(\frac{d_{ij}}{d} \right)^\alpha = \frac{1}{N^2} \sum_{i,j=1}^{N} \left(\frac{d_{ij}}{d} \right)^\alpha = \frac{1}{N^2} \left(1 + \sum_{\substack{i,j=1 \\ (i,j) \neq (\bar{i},\bar{j})}}^{N} \left(\frac{d_{ij}}{d} \right)^\alpha \right) \leq 1$$

for every $\alpha > 1$. And the limit in (4) is

$$\lim_{\alpha \to +\infty} d \left(\sum_{i,j=1}^{N} \frac{1}{N^2} \left(\frac{d_{ij}}{d} \right)^\alpha \right)^{\frac{1}{\alpha}} = d = \max_{i,j} d_{ij}.$$

References

Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. Ecology, 54: 427-431.

Jost, L. (2006). Entropy and diversity. Oikos, 113: 363-375.

Appendix S2 - The parametric Rao's Q variation over the planet

From zero to infinity: minimum to maximum diversity of the planet by spatio-parametric Rao's quadratic entropy

December 19, 2020

We applied the algorithm to a Copernicus Proba-V NDVI (Normalized Difference Vegetation Index) long term average image (June 21st 1999-2017) at 5 km grain, also provided in the `rasterdiv` package as a free `Rasterlayer` dataset which can be loaded by the function `data()` (Figure 1). The parametric Rao's Q algorithm can also be applied to multispectral data; in such a case distances are calculated in the multisystem created by the values of the pixels in each axis/band. The moving window passing throughout the whole image will return M_{Q_α} matrices/layers where α is the value chosen in the R function `paRao`.

With $\alpha \to 0$ the $\prod$ in Equation 6 (in the main text of the manuscript) leads to zeroes throughout the whole map (Figure 2). Increasing α will increase the weight of higher distances among different values until reaching the maximum distance value for $\alpha \to \infty$. In this case the maximum turnover is reached and areas with maximum β-diversity will be apparent. In this case, a multitemporal set is used (long term average NDVI from June 21st 1999-2017). Hence, areas with the highest spatial and temporal turnover are enhanced, namely major mountain ridges. We expect that using single frame images would lead to the enhancement of the spatial component of diversity.

Since the whole process is based on distances in a spectral space between pairs of pixels in terms of their "spectral characters" or in the "spectral space", it is important to notice some cornerstone aspects on the use of

distances from satellite images, especially when comparing different images or the same image in different times. In satellite images, the measure of distances could be impacted by: i) the use of different sensors with different radiometric resolutions, as an example an 8-bit ($2^8 = 256$ values) with respect to a 16-bit ($2^{16} = 65536$ values) image, or ii) the radiometric calibration which has been performed, e.g. with a non-linear transform. Therefore, care should be taken when making use of distances in remote sensing data, explicitly taking into account how the vector of proportions between pixels belonging to some defined classes (e.g., digital numbers, DNs) was obtained.

The complete code of the function can be directly seen in R by typing the `paRao` function name. Moreover, a complete R coding session, to perform the above described analysis is provided at the end of this Appendix.

Figures

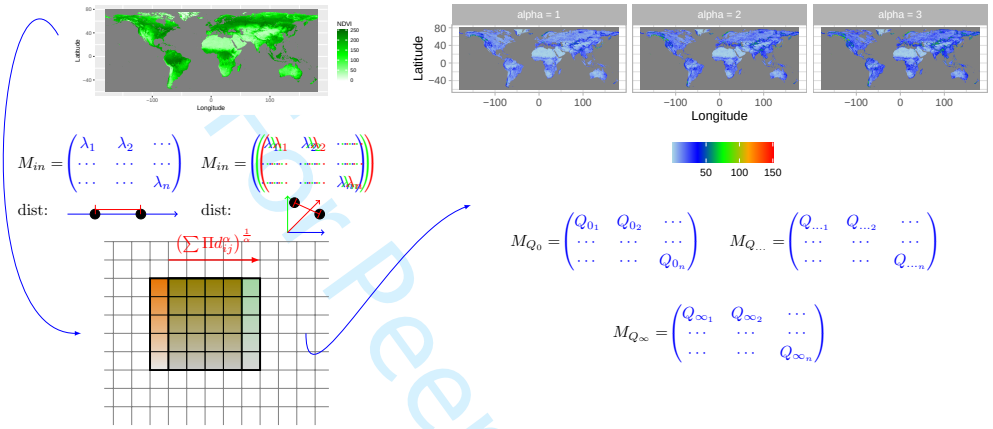


Figure 1: Starting from Copernicus Proba-V NDVI (Normalized Difference Vegetation Index) long term average image (June 21st 1999-2017) at 5km grain, parametric Rao's Q is calculated in a moving window. In this paper NDVI was used as a single layer to calculate distances on one axis, but several layers can be used as well. In this example, three layers (blue, green and red matrices) are shown to calculate distances. The algorithm is based on a moving window passing throughout the whole image, calculating the Rao's Q_α and saving the output in the central pixel. In this example a moving window of 5x5 pixels is passing (red arrow) from one position (orange) to the other (green). The output is a stack of layers each of which represents a different mean of the whole generalized mean spectrum of Equation 5 (in the main text of the paper).

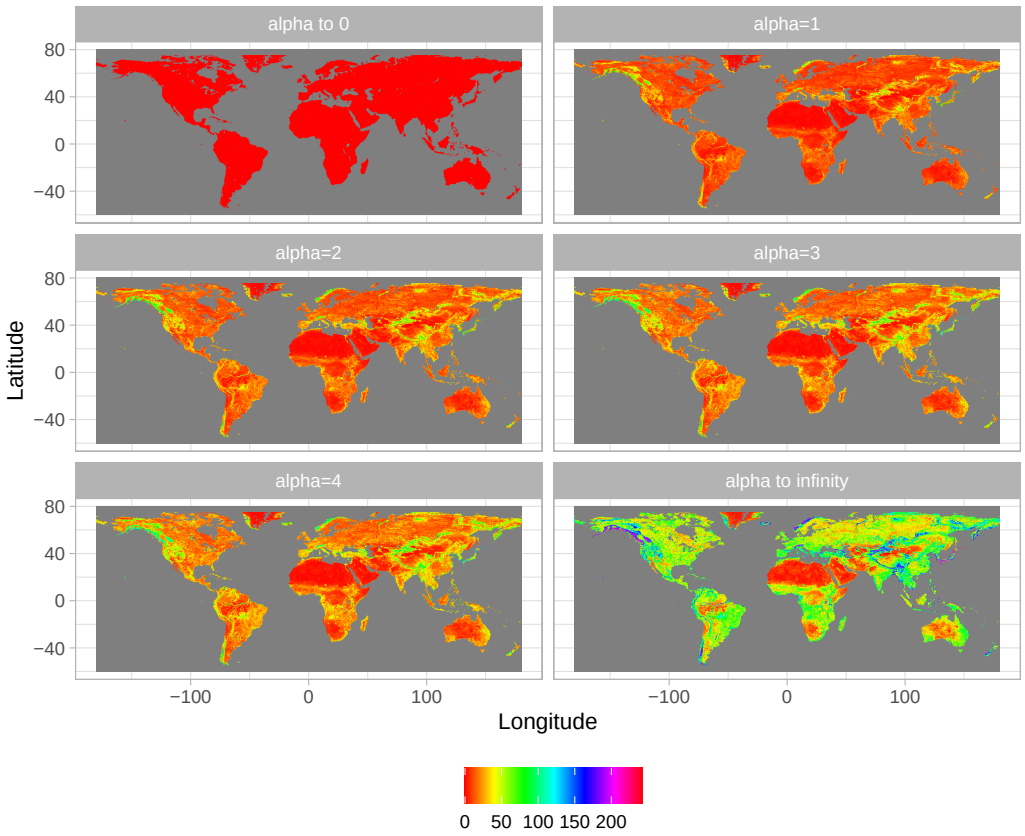


Figure 2: Output of the application of the algorithm shown in Figure 1, achieved by applying different α values: from 0 to 4 until $\alpha \to \infty$. The higher the value of the parameter α, the higher the weight of highest distances among pixel values, until reaching the maximum potential β-diversity (maximum distance) at $\alpha \to \infty$.

Code

0.1 paRao function

```
function (x, dist_m = "euclidean", window = 9, alpha = 1,
    method = "classic",
     rasterOut = TRUE, lambda = 0, na.tolerance = 0, rescale =
     FALSE,
     diag = TRUE, simplify = 3, np = 1, cluster.type = "SOCK",
     debugging = FALSE)
{
     is.wholenumber <- function(x, tol = .Machine$double.eps
    ^0.5) abs(x -
         round(x)) < tol
     if (!(is(x, "matrix") | is(x, "SpatialGridDataFrame") |
    is(x,
         "RasterLayer") | is(x, "list"))) {
         stop("\nNot a valid x object.")
     }
     if (is(x, "SpatialGridDataFrame")) {
         x <- raster(x)
     }
     else if (is(x, "matrix") | is(x, "RasterLayer")) {
         rasterm <- x
     }
     else if (is(x, "list")) {
         rasterm <- x[[1]]
     }
     if (na.tolerance > 1 | na.tolerance < 0) {
         stop("na.tolerance must be in the [0-1] interval.
    Exiting...")
     }
     if (any(!is.numeric(alpha))) {
         stop("alpha must be a numeric vector. Exiting...")
     }
     if (any(alpha < 0)) {
         stop("alphas must be only positive numbers. Exiting
    ...")
     }
     if (method == "classic" & is(x, "RasterLayer")) {
         isfloat <- FALSE
         if (!is.wholenumber(rasterm@data@min) | !is.
    wholenumber(rasterm@data@max) |
             is.infinite(rasterm@data@min) | !is.wholenumber(
    median(getValues(rasterm),
             na.rm = T))) {
```

```
            message("Input data are float numbers. Converting x data in an integer matrix...")
            isfloat <- TRUE
            mfactor <- 100^simplify
            rasterm <- getValues(rasterm) * mfactor
            rasterm <- as.integer(rasterm)
            rasterm <- matrix(rasterm, nrow(x), ncol(x), byrow = TRUE)
            gc()
        }
        else {
            rasterm <- matrix(getValues(rasterm), ncol = ncol(x),
                nrow = nrow(x), byrow = TRUE)
        }
        message("Matrix check OK: \nParametric Rao output matrix will be returned")
    }
    else if (method == "classic" & (is(x, "matrix") | is(x, "list"))) {
        isfloat <- FALSE
        if (!is.integer(rasterm)) {
            message("Input data are float numbers. Converting x in an integer matrix...")
            isfloat <- TRUE
            mfactor <- 100^simplify
            rasterm <- as.integer(rasterm * mfactor)
            rasterm <- matrix(rasterm, nrow(x), ncol(x), byrow = TRUE)
            gc()
        }
        else {
            rasterm <- as.matrix(rasterm)
        }
        message("Matrix check OK: \nParametric Rao output matrix will be returned")
    }
    else ("The class of x is not recognized. Exiting...")
    if (window%%2 == 1) {
        w <- (window - 1)/2
    }
    else {
        stop("The size of the moving window must be an odd number. Exiting...")
    }
    if (np == 1) {
        if (method == "classic") {
            out <- lapply(alpha, paRaoS, rasterm = rasterm, w = w,
```

```
                dist_m = dist_m, na.tolerance = na.tolerance,
                diag = diag, debugging = debugging, isfloat =
 isfloat,
                mfactor = mfactor)
        }
        else if (method == "multidimension") {
            out <- lapply(alpha, mpaRaoS, x = x, rasterm =
rasterm,
                w = w, dist_m = dist_m, na.tolerance = na.
tolerance,
                rescale = rescale, lambda = lambda, diag =
diag,
                debugging = debugging)
        }
        if (rasterOut == T & class(x) == "RasterLayer") {
            outR <- lapply(out, raster, template = x)
            return(outR)
        }
        else {
            return(out)
        }
    }
    else if (np > 1) {
        if (method == "multidimension") {
            stop("Multidimensional paRao not yet implemented,
 set 'np=1'. Exiting...")
        }
        else {
            message("\n###################### Starting
parallel calculation ########################")
            if (debugging) {
                cat("#check: Before parallel function.")
            }
            if (cluster.type == "SOCK" || cluster.type == "
FORK") {
                cls <- makeCluster(np, type = cluster.type,
outfile = "",
                  useXDR = FALSE, methods = FALSE, output = "
")
            }
            else if (cluster.type == "MPI") {
                cls <- makeCluster(np, outfile = "", useXDR =
 FALSE,
                  methods = FALSE, output = "")
            }
            else {
                message("Wrong definition for 'cluster.type'.
 Exiting...")
            }
```

```
            doParallel::registerDoParallel(cls)
            on.exit(stopCluster(cls))
            gc()
            out <- lapply(alpha, paRaoP, rasterm = rasterm, w = w,
                dist_m = dist_m, na.tolerance = na.tolerance,
                diag = diag, debugging = debugging, isfloat = isfloat,
                mfactor = mfactor)
            if (rasterOut == T & class(x) == "RasterLayer") {
                outR <- lapply(out, raster, template = x)
                return(outR)
            }
            else {
                return(out)
            }
        }
    }
}
```

0.2 Application of the paRao function to a synthetic set

```
# install standalone rastediv
install.packages('rasterdiv_0.2-0.tar.gz', repos = NULL, type = "source")

library(raster)
library(rasterdiv)

# generate matrix
synth <- raster(ncol = 8, nrow = 8, xmn = 1, xmx = 6, ymn = 1, ymx = 6)
values(synth) <- rpois(ncell(synth), lambda=3)

# paRao function, using the code in the manuscript
synth.parao <- paRao(synth, alpha = c(0:4,30^9), dist_m = "euclidean", window = 9, na.tolerance = 0.5, simplify = 3, diag = T, rasterOut = T)
```

0.3 Application of the paRao function to the 8bit copNDVI dataset

```
library(rasterdiv)

st <- paRao(copNDVI, alpha = c(0:4,Inf),
```

```
 dist_m = "euclidean", window = 9, na.tolerance = 0.5,
   simplify = 3, diag = TRUE, rasterOut = TRUE)
```

0.4 Output plot

```
library(raster)
library(ggplot2)
library(rasterVis)
library(RColorBrewer)

var.labs=c("layer.1" = "alpha to 0", "layer.2" = "alpha=1", "
   layer.3" = "alpha=2", "layer.4" = "alpha=3", "layer.5" = "
   alpha=4", "layer.6" = "alpha to infinity")

gplot(st, maxpixels=500000) +
  geom_raster(aes(fill = value),  color = "black") +
  labs(x="Longitude",y="Latitude", fill="")+
  scale_fill_gradientn(colors=rainbow(100)) +
  coord_equal()+
  theme_light()+
  facet_wrap(~ variable, ncol = 2,  labeller = labeller(
   variable = var.labs))+
  theme(legend.position = "bottom") +
  NULL
```

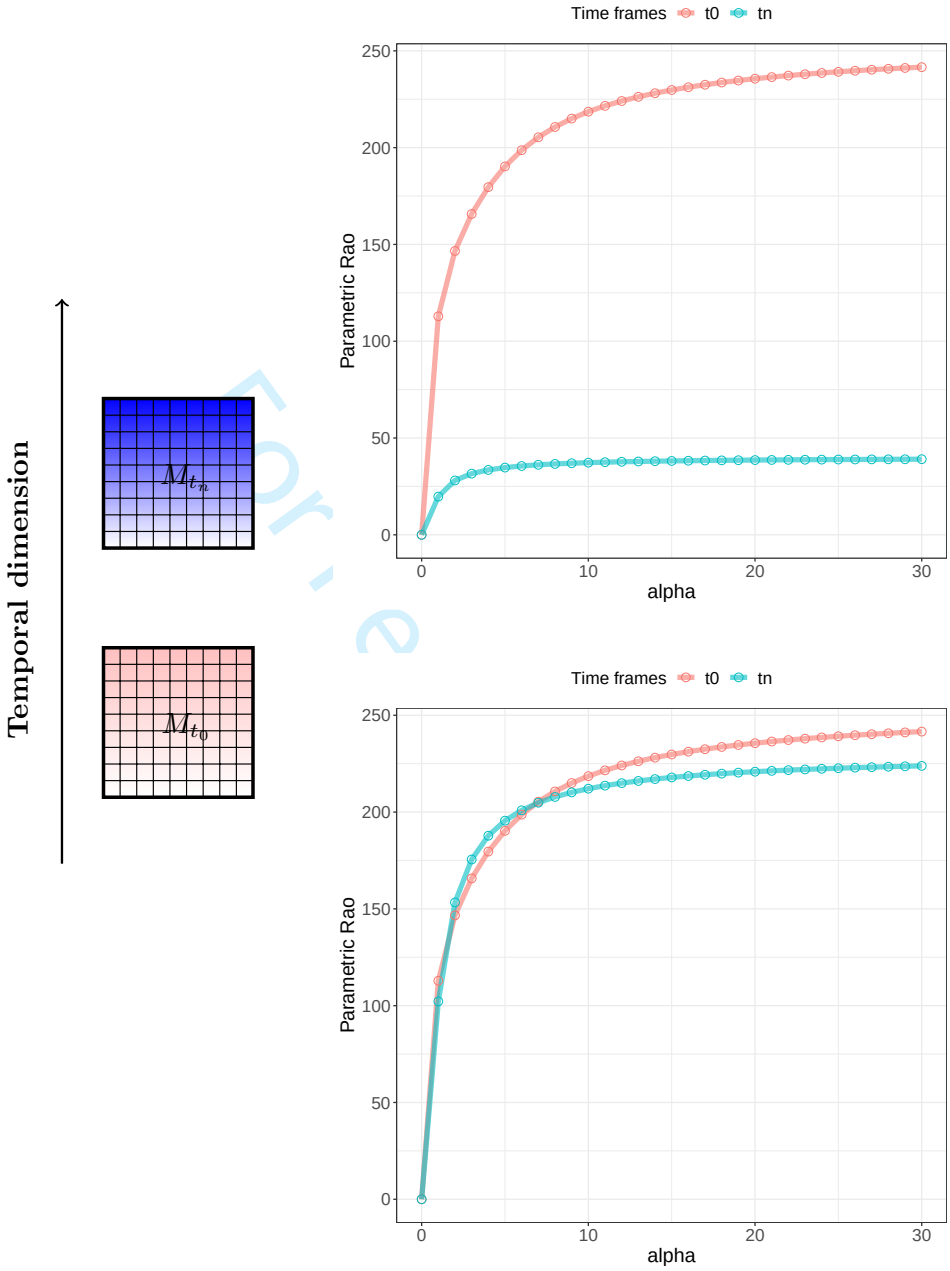
Appendix S3 - Code for Figure 4

From zero to infinity: minimum to maximum diversity of the planet by spatio-parametric Rao's quadratic entropy

January 9, 2021

```
library(ggplot2)
library(rasterdiv)
x1 <- matrix(c(255, 128, 1, 255, 128, 1, 255, 128, 1),ncol=3)
x2 <- matrix(c(10, 10, 10, 10, 50, 50, 50, 50, 50),ncol=3)
p1 <- paRao(x1,window=3,np=1,na.tolerance=0.1,dist_m="
    euclidean",alpha=2)
p2 <- paRao(x2,window=3,np=1,na.tolerance=0.1,dist_m="
    euclidean",alpha=2)
alphas <- seq(0,30,1)
out1 <- paRao(x1,window=3,np=1,na.tolerance=0.1,dist_m="
    euclidean",alpha=alphas)
out2 <- paRao(x2,window=3,np=1,na.tolerance=0.1,dist_m="
    euclidean",alpha=alphas)
r1 <- sapply(out1, function(y) {y[2,2]})
r2 <- sapply(out2, function(y) {y[2,2]})
ggp <- rbind.data.frame(
cbind.data.frame(raop=r1,alphas,"Time frames"=rep("t0",length
    (alphas))),
cbind.data.frame(raop=r2,alphas,"Time frames"=rep("tn",length
    (alphas))))

pdf("landscapes.pdf")
ggplot(ggp, aes(x=alphas, y=raop,col=`Time frames`)) +
     geom_line(size=2,alpha=0.6) +
     geom_point(cex=3,pch=21) +
     theme_bw() +
     xlab("alpha") +
     ylab("Parametric Rao") +
     theme(axis.text.x = element_text(size=14), axis.text.y =
    element_text(size=14)) +
     theme(axis.title.x = element_text(size=16), axis.title.y
     = element_text(size=16))+
     theme(legend.position="top",legend.title=element_text(
    size=14),legend.text=element_text(size=14))
dev.off()



########################################

#### Second graph

library(raster)
library(rasterdiv)
library(ggplot2)

x1 <- matrix(c(255, 128, 1, 255, 128, 1, 255, 128, 1),ncol=3)
x2 <- matrix(c(10, 10, 10, 10, 50, 50, 50, 50, 50),ncol=3)
```

```
x3 <- matrix(c(rep(20,3),rep(250,6)),ncol=3)
alphas <- seq(0,30,1)
out1 <- paRao(x1,window=3,np=1,na.tolerance=0.1,dist_m="
   euclidean",alpha=alphas)
out2 <- paRao(x2,window=3,np=1,na.tolerance=0.1,dist_m="
   euclidean",alpha=alphas)
out3 <- paRao(x3,window=3,np=1,na.tolerance=0.1,dist_m="
   euclidean",alpha=alphas)
r1 <- sapply(out1, function(y) {y[2,2]})
r2 <- sapply(out2, function(y) {y[2,2]})
r3 <- sapply(out3, function(y) {y[2,2]})
ggp <- rbind.data.frame(
cbind.data.frame(raop=r1,alphas,"Time frames"=rep("t0",length
   (alphas))),
cbind.data.frame(raop=r3,alphas,"Time frames"=rep("tn",length
   (alphas))))

pdf("landscapes2.pdf")
ggplot(ggp, aes(x=alphas, y=raop,col=`Time frames`)) +
    geom_line(size=2,alpha=0.6) +
    geom_point(cex=3,pch=21) +
    theme_bw() +
    xlab("alpha") +
    ylab("Parametric Rao") +
    theme(axis.text.x = element_text(size=14), axis.text.y =
   element_text(size=14)) +
    theme(axis.title.x = element_text(size=16), axis.title.y
    = element_text(size=16))+
    theme(legend.position="top",legend.title=element_text(
   size=14),legend.text=element_text(size=14))
    ggsave("~/paRao_comparison1.png",dpi=600,scale=0.5,width
   =10,height=10)
dev.off()
```


Temporal dimension
M_{t_n}
M_{t_0}
Time frames t0 tn
Parametric Rao
alpha
0
50
100
150
200
250
0
10
20
30
Time frames t0 tn
Parametric Rao
alpha
0
50
100
150
200
250
0
10
20
30

88