



Short Communication

Measuring Rao's Q diversity index from remote sensing: An open source solutionDuccio Rocchini ^{a,*}, Matteo Marcantonio ^{a,b,1}, Carlo Ricotta ^c^a *Fondazione Edmund Mach, Research and Innovation Centre, Department of Biodiversity and Molecular Ecology, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy*^b *Technische Universität Berlin, Department of Landscape Architecture and Environmental Planning, Geoinformation in Environmental Planning Lab, Berlin, Germany*^c *Department of Environmental Biology, University of Rome "La Sapienza", 00185 Rome, Italy*

ARTICLE INFO

Article history:

Received 25 May 2016

Received in revised form 21 July 2016

Accepted 22 July 2016

Keywords:

Biodiversity

Heterogeneity

Landscape metrics

Rao's Q

Remote sensing

Spatial ecology

Shannon's entropy

ABSTRACT

Measuring biodiversity is a key issue in ecology to guarantee effective indicators of ecosystem health at different spatial and time scales. However, estimating biodiversity from field observations might present difficulties related to costs and time needed. Moreover, a continuous data update for biodiversity monitoring purposes might be prohibitive. From this point of view, remote sensing represents a powerful tool since it allows to cover wide areas in a relatively low amount of time. One of the most common indicators of biodiversity is Shannon's entropy H' , which is strictly related to environmental heterogeneity, and thus to species diversity. However, Shannon's entropy might show drawbacks once applied to remote sensing data, since it considers relative abundances but it does not explicitly account for distances among pixels' numerical values. In this paper we propose the use of Rao's Q applied to remotely sensed data, providing a straightforward R-package function to calculate it in 2D systems. We will introduce the theoretical rationale behind Rao's index and then provide applied examples based on the proposed R function.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Measuring biodiversity as an indicator of ecosystem health has been recognized by major initiatives worldwide, including the Group on Earth Observation (GEO BON, <http://www.earthobservations.org/geobon.shtml>) initiative, the International Geosphere Biosphere Programme (IGBP, <http://www.igbp.net/>), the World Climate Research Programme (WCRP, <http://wcrp-climate.org/>), the Committee on Earth Observation Systems (CEOS) Biodiversity task (<http://ceos.org/>), among others.

However, estimating biodiversity from field data presents a number of drawbacks mainly related to time and costs, together with intrinsic difficulties to build standardized procedures for reproducible data gathering (Palmer et al., 2002).

For this purpose, using maps in a GIS environment or heterogeneity-related maps derived from remotely sensed imagery (e.g. Carranza et al., 2007) might help in finding hotspots of diversity over space and track their variation over time (Boyd and Foody, 2011), from local (Feilhauer et al., 2013) to global (Rocchini et al.,

2010) spatial scales. This is true under the light of the Spectral Variation Hypothesis (Palmer et al., 2002) which states that the higher the environmental heterogeneity, the higher will be the species diversity of a certain area. The rationale under the Spectral Variation Hypothesis is that a higher spatial variability (measured by spectral diversity from remotely sensed images) is related to a higher amount of ecological niches for species living therein. Hence, measuring the heterogeneity of a landscape is critical since it is directly related to its diversity (Gillespie et al., 2008; Skidmore et al., 2015). Moreover, landscape diversity is related to the diversity at other ecosystem levels such as species diversity.

It has been demonstrated that the measure being used can lead to very different (and sometimes misleading) results. As an example, one of the mostly used diversity measures of the landscape based on spectral remotely sensed data, i.e. the Shannon's entropy (Shannon, 1948), has a number of implicit drawbacks like: (i) the difficulty to discriminate between differences in richness or relative abundance (Nagendra, 2002) or (ii) the impossibility to consider spectral values as numbers instead of classes (Rocchini and Neteler, 2012b). Concerning the second point, Shannon's entropy accounts for richness and relative abundance of spectral values but it does not explicitly consider the numerical magnitude (values) of pixels.

The aim of this paper is to solve the aforementioned issue, by the application of Rao's Q to remotely sensed data, providing a

* Corresponding author.

E-mail address: duccio.rocchini@fmach.it (D. Rocchini).¹ These authors contributed equally to this article.

straightforward R function to calculate it in 2D systems. We will first introduce the theoretical rationale behind Rao's index and then provide applied examples based on the proposed R function. As far as we know, this is the first attempt to measure Rao's Q in a 2D space applied to remotely sensed data.

2. Theory under the use of the Rao's Q index

Methods for measuring landscape diversity have mostly relied on the classification of remotely sensed image. However, image classification has several drawbacks which should be seriously taken into account, e.g.: (i) the accuracy assessment should be performed in a robust manner, thus requiring time and costs overall when field assessment is involved (Foody, 2002), (ii) it is difficult to build practically sound accuracy assessment protocols (Foody, 2008), (iii) the classification should be performed only by robust algorithms avoiding as much as possible manual digitization (Burnett and Blaschke, 2003), (iv) several issues have to be bypassed when choosing pure training samples in order to avoid mixing effects (Small, 2004). Besides these technical shortcomings, classification is a subjective task in its very nature and it inevitably leads to the degradation of continuous information (Palmer et al., 2002).

Rocchini et al. (2010) summarize several approaches to measure ecosystem diversity from remotely sensed images, mainly based on the continuous variability of pixel values (e.g. original digital numbers of a satellite image). Such approaches show their full power when relying on Free and Open Source algorithms. Open Source algorithms allow indeed robustness and reproducibility thanks to the public availability of the used code (Rocchini and Neteler, 2012a).

Among the most spread diversity indexes used in ecology there is the Shannon entropy index (H' , Shannon, 1948). This index can be easily applied to remote sensed data. Given a certain number of reflectance values in a remotely sensed image, also referred to as digital numbers, H' can be calculated as $H' = -\sum p_i \times \log(p_i)$. In this particular application H' takes into account the relative proportion p of each reflectance value i . Generalizing, H' considers the equitability of the system. Furthermore, when transformed to the Pielou evenness index J (Pielou, 1969), calculated as $J = H'/H_{max}$, it shows the maximum possible diversity within the same number of reflectance values. Quoting Ricotta and Avena (2003), who provided an elegant mathematical dissertation about the Pielou index applied to both species and landscape classes:

“The normalization of H' with respect to maximum entropy ($J = H'/H_{max}$) is termed ‘evenness’ because it measures deviation from an even distribution of individuals amongst the N species”.

Translating the sentence from species to spectral diversity, it turns out to be:

“The normalization of H' with respect to maximum entropy ($J = H'/H_{max}$) is termed ‘evenness’ because it measures deviation from an even distribution of individual pixels amongst the N reflectance values.”

However, Shannon and Pielou indices only rely on the relative abundance of reflectance, not considering the numerical value of reflectance *per se*. Facing the problem from a mathematical point of view, let M be an image of 3×3 pixels (indicated by c to avoid confusion with p , used in this manuscript to indicate the proportion of area of each category):

$$M = \begin{pmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{pmatrix} \quad (1)$$

Let i and j be two different pixel values, e.g. two Digital Numbers (DNs) of a 8 bit image with $i \neq j$, as:

$$M = \begin{pmatrix} i & i & j \\ i & j & j \\ j & j & j \end{pmatrix} \quad (2)$$

In this case $H = -\sum p \times \ln p = -(3/9 \times \ln(3/9) + 6/9 \times \ln(6/9)) = 0.637$.

Shannon entropy does not take into account the value of i and j but just the proportion of i and j values. Therefore, it does not discriminate among different contexts such as, (a) $i = 1$ and $j = 200$ or (b) $i = 201$ and $j = 200$.

On the contrary, Rao's Q index does take into account i and j value by considering their pairwise distance d_{ij} :

$$Q = \sum \sum d_{ij} \times p_i \times p_j \quad (3)$$

As an example, in case of (a) $Q = 88.444$, while in case of (b) $Q = 0.444$.

As a consequence, deriving Rao's Q involves calculating a distance matrix M_d for all the pixel values:

$$M_d = \begin{pmatrix} d_{1,1} & d_{1,2} & d_{1,3} & \dots & d_{1,n} \\ d_{2,1} & d_{2,2} & d_{2,3} & \dots & d_{2,n} \\ d_{3,1} & d_{3,2} & d_{3,3} & \dots & d_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n,1} & d_{n,2} & d_{n,3} & \dots & d_{n,n} \end{pmatrix} \quad (4)$$

or more simply $\begin{matrix} \boxed{1} & \boxed{2} & \boxed{3} & \dots & \boxed{N} \\ \hline 1 & 2 & 3 & \dots & N \end{matrix}$, when N pixels are considered (see also Rocchini (2007) on distance matrices in a spectral space). Thus, Rao's Q is related to the sum of all the pixel values pairwise distances, each of which is multiplied by the relative abundance of each pair of pixels in the analysed image $d \times (1/N^2)$. In other words, Rao's Q is the expected difference in reflectance values between two pixels drawn randomly with replacement from the considered evaluated pixels set. The distance matrix can be built in several dimensions (layers), thus allowing to consider more than one band at a time. As a consequence Rao's Q can be calculated in a multidimensional (multi-layers) system.

In remote sensing applications the derivation of synthetic indexes of any sort (i.e., diversity) is often performed considering small chunks of the whole image per time, commonly defined as ‘windows’ or ‘moving windows’. From now on, we will use this terminology to indicate the local space of analysis.

3. Coding Rao's Q in R

The function `spectralrao()` to derive Rao's Q, written in the R statistical language (R Core Team, 2016), is reported in Appendix 1 and stored in the GitHub repository <https://github.com/mattmar/spectralrao>. The function accepts matrix, RasterLayer or Spatial-GridDataFrame object as input (or a list of them). It can be ran with two different settings, using (i) a single matrix (`mode="classic"`) or (ii) more matrices (`mode="multidimensional"`) as input. Distance can be calculated relying on Euclidean, Manhattan and Canberra distances by the `distance_m` parameter. Appendix 2 provides a complete description of such distances, with their advantages and disadvantages, together with proper reference to previous ecological papers using them. Further, a user-defined distance matrix can be also provided through the function argument `distance_m`.

In this manner, it is possible to obtain H' as output, together with Rao's Q, setting the option `shannon=TRUE`. On the other hand, if `mode="multidimensional"`, a list of matrices must be provided as input. The overall distance matrix is thus calculated in a multi- or

Download English Version:

<https://daneshyari.com/en/article/6292744>

Download Persian Version:

<https://daneshyari.com/article/6292744>

[Daneshyari.com](https://daneshyari.com)