rasterdiv-An Information Theory tailored R package for measuring ecosystem heterogeneity from space: To the origin and back

Duccio Rocchini, Elisa Thouverai, Matteo Marcantonio, Martina Iannacito, Daniele da Re, Michele Torresani, Giovanni Bacaro, Manuele Bazzichetto, Alessandra Bernardi, Giles M. Foody, et al.

Received: 13 December 2020 | Accepted: 8 February 2021

DOI: 10.1111/2041-210X.13583

APPLICATION

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

rasterdiv—An Information Theory tailored R package for measuring ecosystem heterogeneity from space: To the origin and back

Duccio Rocchini[1,2] | Elisa Thouverai[1] | Matteo Marcantonio[3] | Martina Iannacito[4] | Daniele Da Re[5] | Michele Torresani[6] | Giovanni Bacaro[7] | Manuele Bazzichetto[8] | Alessandra Bernardi[9] | Giles M. Foody[10] | Reinhard Furrer[11,12] | David Kleijn[13] | Stefano Larsen[14,15] | Jonathan Lenoir[16] | Marco Malavasi[2] | Elisa Marchetto[1] | Filippo Messori[1] | Alessandro Montaghi[17] | Vítězslav Moudrý[2] | Babak Naimi[18] | Carlo Ricotta[19] | Micol Rossini[20] | Francesco Santi[1] | Maria J. Santos[21] | Michael E. Schaepman[22] | Fabian D. Schneider[23] | Leila Schuh[11] | Sonia Silvestri[1] | Petra Šímová[2] | Andrew K. Skidmore[24,25] | Clara Tattoni[26] | Enrico Tordoni[7] | Saverio Vicario[27] | Piero Zannini[1] | Martin Wegmann[28]

[1]BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy; [2]Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Praha - Suchdol, Czech Republic; [3]Department of Pathology, Microbiology, and Immunology, School of Veterinary Medicine, University of California, Davis, CA, USA; [4]Inria Bordeaux - Sud-Ouest, Talence, France; [5]Georges Lemaître Center for Earth and Climate Research, Earth and Life Institute, UCLouvain, Louvain-la-Neuve, Belgium; [6]Faculty of Science and Technology, Free University of Bolzano/Bozen, Piazza Universitá/Universitätsplatz 1, Bolzano, Italy; [7]Department of Life Sciences, University of Trieste, Trieste, Italy; [8]EcoBio (Ecosystèmes, Biodiversité, Évolution) - UMR 6553, Université de Rennes, CNRS, Rennes, France; [9]Department of Mathematics, University of Trento, Povo, Italy; [10]School of Geography, University of Nottingham, Nottingham, UK; [11]Department of Mathematics, University of Zurich, Zurich, Switzerland; [12]Department of Computational Science, University of Zurich, Zurich, Switzerland; [13]Plant Ecology and Nature Conservation Group, Wageningen University, Wageningen, The Netherlands; [14]Unit of Computational Biology, Research and Innovation Center, Fondazione Edmund Mach, San Michele all'Adige, Italy; [15]Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy; [16]UR "Ecologie et Dynamique des Systèmes Anthropisés" (EDYSAN, UMR 7058 CNRS-UPJV), Université de Picardie Jules Verne, Amiens, France; [17]DAGRI Department of Agriculture, Food, Environment and Forestry, University of Florence, Firenze, Italy; [18]Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland; [19]Department of Environmental Biology, University of Rome "La Sapienza'", Rome, Italy; [20]Remote Sensing of Environmental Dynamics Laboratory, DISAT, Universitá degli Studi Milano-Bicocca, Milano, Italy; [21]Department of Geography, Earth System Science, University of Zurich, Zurich, Switzerland; [22]Department of Geography, Remote Sensing Laboratories, University of Zurich, Zurich, Switzerland; [23]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA; [24]Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands; [25]Department of Environmental Science, Macquarie University, Sydney, NSW, Australia; [26]Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Firenze, Italy; [27]CNR-IIA C/O Physics Department "M. Merlin" University of Bari, Bari, Italy and [28]Department of Remote Sensing, University of Wuerzburg, Würzburg, Germany

Correspondence
Duccio Rocchini
Email: duccio.rocchini@unibo.it

Funding information
H2020 Project SHOWCASE, Grant/Award Number: 862480; H2020 COST Action, Grant/Award Number: CA17134; Jet Propulsion Laboratory; California Institute of Technology; National Aeronautics

Abstract

1. Ecosystem heterogeneity has been widely recognized as a key ecological indicator of several ecological functions, diversity patterns and change, metapopulation dynamics, population connectivity or gene flow.
2. In this paper, we present a new R package—rasterdiv—to calculate heterogeneity indices based on remotely sensed data. We also provide an ecological application

and Space Administration, Grant/Award Number: 80NM0018D0004; University of Zurich Research Priority Program on Global Change and Biodiversity (URPP GCB); Friuli Venezia Giulia Region, Swiss National Science Foundation, Grant/Award Number: SNSF–175529; Czech University of Life Sciences Prague

Handling Editor: Sarah Goslee

at the landscape scale and demonstrate its power in revealing potentially hidden heterogeneity patterns.

3. The rasterdiv package allows calculating multiple indices, robustly rooted in Information Theory, and based on reproducible open-source algorithms.

KEYWORDS

biodiversity, ecological informatics, modelling, remote sensing, satellite imagery

1 | INTRODUCTION

Ecosystem heterogeneity is related to a number of ecological processes and functions such as species diversity patterns and change (Rocchini et al., 2018), metapopulation dynamics (Fahrig, 2007), population connectivity (Malanson & Cramer, 1999) or gene flow (Lozier et al., 2013). Heterogeneity has been defined in various ways in the scientific literature: (a) as the variation in space and time of qualitative and quantitative descriptors of an environmental variable of interest (Li & Reynolds, 1995); (b) as the horizontal component of habitat variation (August, 1983; Grelle, 2003); (c) as the spatially structured variability of the habitat (Ettema & Wardle, 2002); or (d) as within-habitat variability (Heaney, 2001; Hortal et al., 2009). In this paper, it will be considered as an umbrella concept representing the degree of non-uniformity in land cover, vegetation and physical factors (topography, soil, topoclimate and microclimate; Stein et al., 2014).

Landscape heterogeneity across different spatial extents and over different temporal periods can be inferred by applying algorithms based on remote sensing and spatial ecology (Schimel & Schneider, 2019; Skidmore et al., 2011). Remotely sensed spectral heterogeneity measures of a landscape represent a valid alternative to categorical land cover maps, which, especially in the case of non-homogeneous and complex landscape configurations (e.g. mosaic of crops and semi-natural forests), might suffer from oversimplification when investigated through land cover classes (Da Re et al., 2019; Rocchini et al., 2019). Heterogeneous landscapes should present higher spectral heterogeneity values compared to more homogeneous landscapes within the same spatial extent (Rocchini & Ricotta, 2007). It follows that remotely sensed spectral heterogeneity can be profitably used to measure landscape heterogeneity in space and time to convey information on ecosystem processes and functioning (Schneider et al., 2017).

From this point of view, the development of Free and Open-Source algorithms to measure and monitor (i.e. repeated measures over time) landscape or ecosystem heterogeneity from space would allow robust, reproducible and standardized estimates of ecosystem functioning and services (Rocchini & Neteler, 2012). Furthermore, their intrinsic transparency, community-vetoing options, sharing and rapid availability are also valuable additions and reasons to move towards open-source options. Considering the different open-source software options, the R software (R Core Team, 2020) is among the most widely used languages for statistics and modelling and different packages have already been devoted to remote sensing data processing for: (a) raster data management (raster package; Hijmans & van Etten, 2020); (b) remote sensing data analysis (RStoolbox package; Leutner et al., 2019); (c) spectral species diversity mapping (biodivMapR package; Féret & Boissieu, 2020); (d) Sparse Generalized Dissimilarity Modelling based on remote sensing data (sgdm package; Leitão et al., 2017); (e) entropy-based local spatial association (ELSA package; Naimi et al., 2019); or (f) landscape metrics calculation (landscapemetrics package; Hesselbarth et al., 2019), to name just a few. Readers can also refer to https://cran.r-project.org/web/views/Spatial.html for the CRAN Task View on analysis of spatial data.

Nonetheless, no R package currently provides a flow of functions grounded in Information Theory and generalized entropy, incorporating abundance information for each informative value but also the relative numerical distance among said values. In this paper, we introduce the new rasterdiv R package, now available under the Comprehensive R Archive Network (CRAN, https://CRAN.R-project.org/package=rasterdiv), which provides such a set of functions' throughput workflow. The aim of this manuscript is to briefly introduce the theory under the rasterdiv package and to provide an ecological example demonstrating its ability to measure several aspects of landscape or ecosystem heterogeneity.

2 | BRIEF DESCRIPTION OF THE THEORETICAL FRAMEWORK ON INFORMATION-THEORY-BASED METRICS

2.1 | Shannon entropy

Shannon's (1948) theory, profoundly rooted in Eduard Boltzmann's (1872) studies, is a solid basis for calculating landscape heterogeneity and it has been widely used in several ecological applications (refer to Vranken et al., 2014 for a review).

Given a sample area with N pixel values and p_i relative abundances for every $i \in \{1, \ldots, N\}$ the Shannon index is calculated as:

$$H = -\sum_{i=1}^{N} p_i \ln p_i, \quad (1)$$

When applying such indices to remotely sensed data, the image is divided into small chunks of the whole image, commonly defined as 'kernels', 'windows' or 'moving windows' (see Figure 1 in Box 1). These terms will be used throughout this manuscript to relay the local scale of analysis.

In the Shannon index, the relative abundance of pixel values (e.g. reflectance values, vegetation indices) is considered. The higher the richness and turnover, the higher will be the equitability of values and thus the Shannon index.

BOX 1 Description of the moving window approach

Given a raster layer, ecologists usually split it into small chunks, called moving windows. To estimate the value of the previously presented indices, we first select l an odd number, which will correspond to the length of the squared window. With this choice, we have a central entry in the window, that is, the $\left(\frac{l+1}{2}, \frac{l+1}{2}\right)$ entry, which will be placed as a mask, in position (1, 1) over our raster. The index is therefore computed only with the values which the window covers. Notice that with this choice we will have some missing values, which will not contribute to the index computation. The obtained index value is stored in position (1, 1) in the output raster. The following step is moving the window so that its central entry is over the entry (1, 2) of the raster. The index value computed is stored in the corresponding position of the output raster, that is, in (1, 2). We proceed in this way until the last entry of the output raster is filled.

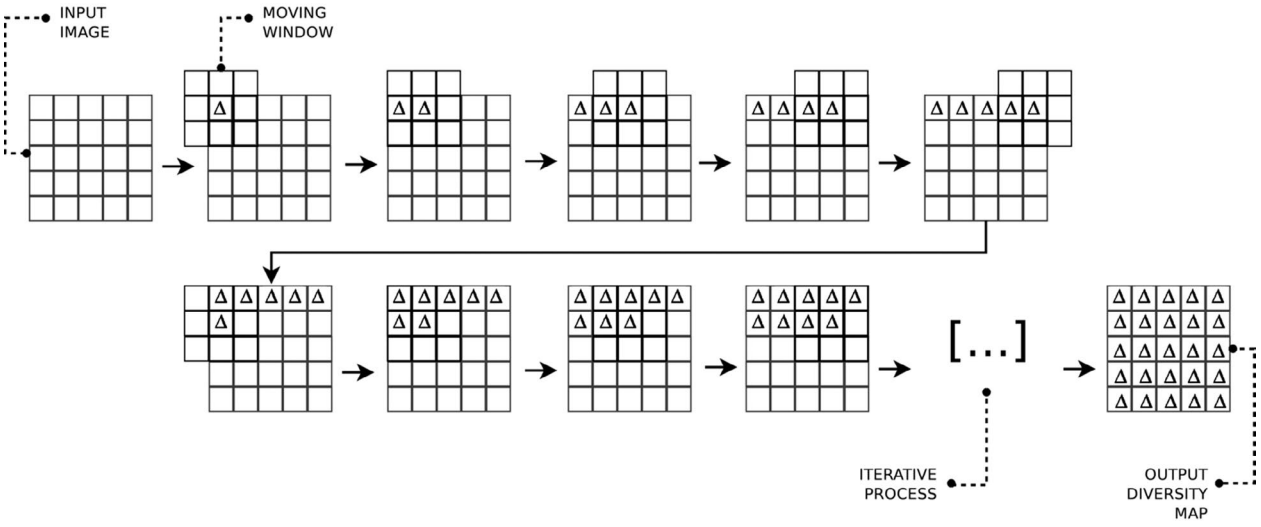

FIGURE 1 The moving window technique for the computation of diversity indices in the rasterdiv package, redrawn from Rocchini et al. (2013)

This technique, visually presented in Figure 1, is extremely popular in ecology since it is quite simple, robust and powerful to simulate a biological boundary. Note that this technique of computing index has a suitable algorithmic structure.

Algorithm 1 Index computation with moving window

Input: a raster $R \in \mathbb{M}^{m \times n}(\mathbb{R})$
Input: an index computation function f
Input: an odd number l
Output: the raster with the index values OR
$m \in \mathbb{M}^{l \times l}(\mathbb{R})$ is the moving window;
for $i = 1, 2, \ldots, m$ **do**
  for $j = 1, 2, \ldots, n$ **do**
    we assign to m the values of R covered by a l side square mask at the step (i, j);
    f computes the index value with values covered by the moving window m;
    $OR_{ij} \leftarrow f(m)$;
  end
end

2.2 | Rényi generalized entropy

Any point descriptor of information heterogeneity, like the previously cited Shannon's *H*, is not able to describe the whole potential spectrum of heterogeneity as it usually measures only one part or component of heterogeneity (e.g. richness, evenness, nestedness, etc.). Hence, no single measure can be used to represent such a wide spectrum (Gorelick, 2011; Nakamura et al., 2020).

Rényi (1970) proposed a method to generalize entropy measurements in just one formula, changing one parameter, called α in its original formulation. Given a sample area with *N* pixel values and p_i relative abundances for every $i \in \{1, \ldots, N\}$, the Rényi entropy index is:

$$H_\alpha = \frac{1}{1-\alpha} \times \ln \sum_{i=1}^{N} p_i^\alpha. \quad (2)$$

Changing the parameter α will lead to different indices starting from the same formula (Hill, 1973). As an example, when $\alpha = 0$, $H_0 = \ln(N)$ where N = richness, namely the maximum possible Shannon index (H_{max}). In practice, with $\alpha = 0$, all the spectral values equally contribute to the index, without making use of their relative abundance. For $\alpha \to 1$, the Rényi entropy index will equal Shannon *H*, according to the l'Hôpital's rule (a mathematical proof is provided in Supporting Information Appendix 1), while for $\alpha = 2$ the Rényi entropy index will equal the $\ln(1/D)$ where *D* is the Simpson's dominance (Simpson, 1949). The theoretical curve relating the Rényi entropy index and α is a negative exponential, that is, it decays until flattening for higher values of α, where the weight of the most abundant spectral values is higher with small differences among the attained heterogeneity maps (Ricotta et al., 2003). Besides the Rényi's generalized entropy, the rasterdiv package includes generalized metrics as the Hill's (1973) numbers.

2.3 | Rao's Q heterogeneity index

The previously described metrics have no dimension. In other words, they do not consider the relative difference among pixel values but just the presence of a different class. For example, having $A = (1, 2, 3, 4, 5, 6, 7, 8, 9)$ and $B = (1, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9)$ as two theoretical arrays containing values that are different from each other, the Shannon index will always be maximum, that is, $H = \ln(9) = 2.197225$ despite the relative numerical distance between pairs of values.

In remote sensing, this is a critical point since contiguous zones of a satellite image might have similar (but not strictly equal) reflectance values. For instance, the variability of a homogeneous surface like a woodland patch or a water area, would be overestimated within the landscape matrix if spectral distances among values are not considered in the calculation.

The Rao's Quadratic heterogeneity measure (hereafter Rao's *Q*; Rao, 1982) can be applied to overcome this issue, considering both relative abundances and spectral distances among pixel values in the calculation.

Given the values of different pixels *i* and *j*, the Rao's *Q* considers their pairwaise distance d_{ij} as:

$$Q = \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij} \times p_i \times p_j. \quad (3)$$

Accordingly, an array with different but spectrally near values will convey a high Shannon's *H* but a low Rao's *Q*. Conversely, an array with different and distant spectral values will convey both a high Shannon's *H* and a high Rao's *Q*.

Given a 3×3 pixels matrix *M*:

$$M = \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_4 & \lambda_5 & \lambda_6 \\ \lambda_7 & \lambda_8 & \lambda_9 \end{pmatrix},$$

where λ is the reflectance value for every pixel in a single 8-bit band (256 possible values), a pairwise distance matrix M_d is derived for all pixel values:

$$M_d = \begin{pmatrix} d_{\lambda_1,\lambda_1} & d_{\lambda_1,\lambda_2} & d_{\lambda_1,\lambda_3} & \cdots & d_{\lambda_1,\lambda_N} \\ d_{\lambda_2,\lambda_1} & d_{\lambda_2,\lambda_2} & d_{\lambda_2,\lambda_3} & \cdots & d_{\lambda_2,\lambda_N} \\ d_{\lambda_3,\lambda_1} & d_{\lambda_3,\lambda_2} & d_{\lambda_3,\lambda_3} & \cdots & d_{\lambda_3,\lambda N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{\lambda_N,\lambda_1} & d_{\lambda N,\lambda_2} & d_{\lambda_N,\lambda_3} & \cdots & d_{\lambda_N,\lambda_N} \end{pmatrix}. \quad (4)$$

Then, according to Equation (3), Rao's *Q* is obtained as the sum of every pairwise distance multiplied by the relative abundances of all the pairs of pixels in the analysed image $d \times (1/N^2)$. Strictly speaking, Rao's *Q* can be defined as the expected difference in reflectance values between two pixels drawn randomly with replacement from the evaluated set of pixels.

It is possible to construct the distance matrix for several dimensions (layers) to consider multiple bands at a time and calculate Rao's *Q* in a multidimensional (multi-layer) system.

To illustrate the benefit of the functions provided in the rasterdiv package, we apply it on an ecological study case in the following section.

3 | APPLICATION: THE ECOSYSTEM HETEROGENEITY OF THE ÖTZI AREA

To show the capabilities of the rasterdiv package, we decided to focus on one of the geologically and biologically most diverse mountain regions worldwide: the Similaun and Ortles glaciers in Italy. This region is not only important for its rich geobiodiversity but also for its fascinating archaeological history, also due to an incredible anthropological discovery of the early nineties: the famous Ötzi Tyrolean iceman (Keller et al., 2012). The study area we are focusing on for applying the rasterdiv package is

included between the Ortles and the Similaun glaciers, in the Alps of northern Italy (Figure 2). Below, we provide a step-by-step tutorial which can be reproduced for any area and by every researcher worldwide. The only required input data are satellite images.

We used Copernicus Sentinel-2 data at a spatial resolution of 10 m (Figure 2) acquired on May 9th 2020. Once the satellite image was downloaded from the Copernicus Open Access Hub (https://scihub.copernicus.eu/), we computed NDVI and, for the purposes of this research, rescaled it to an 8-bit radiometric resolution. Hence, NDVI values were used as the input information (values) to compute the ecosystem heterogeneity indices described in the former section. More specifically, based on the NDVI raster grid used as input object, we ran a set of functions provided in the rasterdiv package and written in Box 2.

We applied the code written in Box 2, using a moving window of 9 × 9 pixels, to both (a) the Similaun glacier upper area, mainly covered by alpine conifer woodlands (dominated by *Picea abies*, *Larix decidua*, *Abies alba*) and rocks, and (b) to the valley bottom (Val Venosta), a human-dominated landscape devoted to agricultural areas and small urban villages.

Concerning the Similaun glacier area, the Shannon index showed medium to high values everywhere, including areas with almost homogeneous rock cover (i.e. alpine habitat) as well as areas with homogeneous tree cover (i.e. coniferous forest habitat) (Figure 3). This is due to the fact that Shannon's *H* does not take into account the distance among pixel values but only the relative abundance of each value within the moving window of 9 × 9 pixels. In this case, NDVI values showed subtle differences among each other, especially in homogeneous areas and even when rescaled at 8 bit (namely 256 possible integer values).

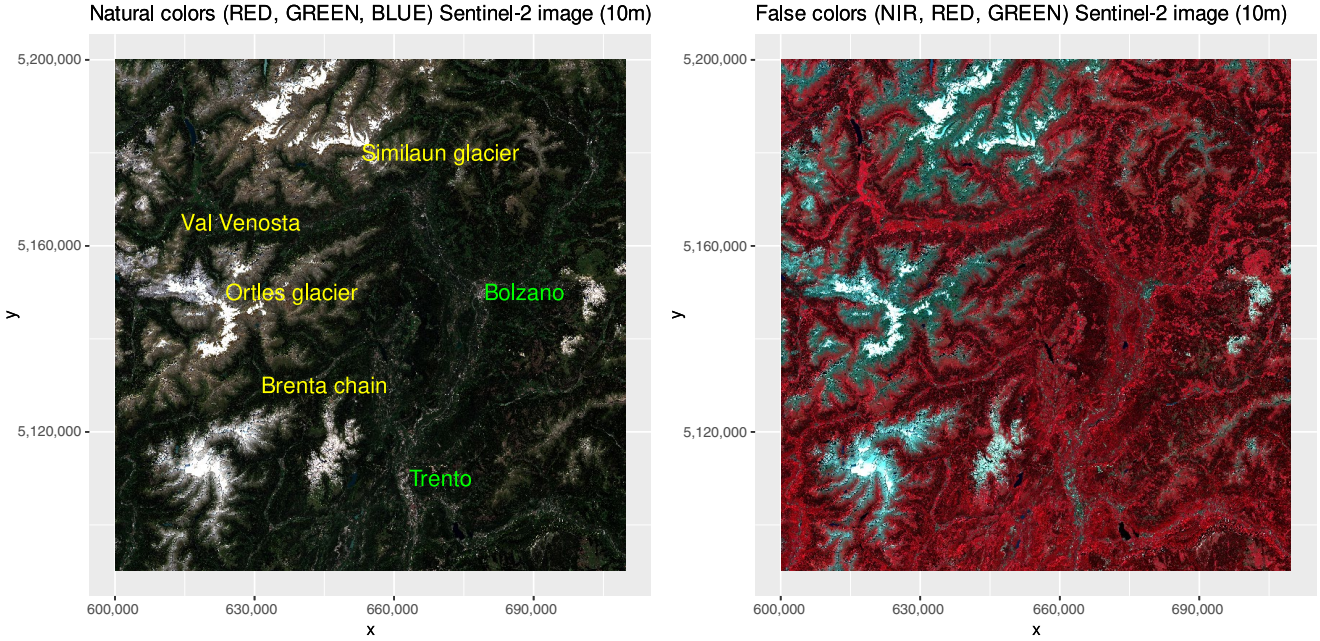

FIGURE 2 The Ötzi area in the northern Italian Alps, used for building an ecological example of the application of the rasterdiv package starting from a Copernicus Sentinel-2 image, represented by an RGB space in natural (red, green and blue) and false (near infrared, red and green) colours. Coordinates are in the UTM (WGS84, zone 32N) reference system

BOX 2 Code used under rasterdiv

```
## Shannon's H
sha <- Shannon(ndvi_8bit, window = 9, na.tolerance = 0.9, np = 3)
# where: ndvi_8bit = input data, window = moving window size, na.tolerance =
proportion of tolerated NA values, np = number of processes (cores)

## Rényi entropy
ren <- Renyi(ndvi_8bit, window = 9, alpha = seq(0,10), na.tolerance = 0.9, np = 3)
# where: alpha = alpha value in the Rényi formula (in this case all integer alpha
values from 0 to 10)

## Rao's Q
rao <- Rao(ndvi_8bit, window = 9, na.tolerance = 0.9, dist_m = 'euclidean', np = 3)
# where: dist_m = distance type used to calculate the pixel values' distance matrix
```

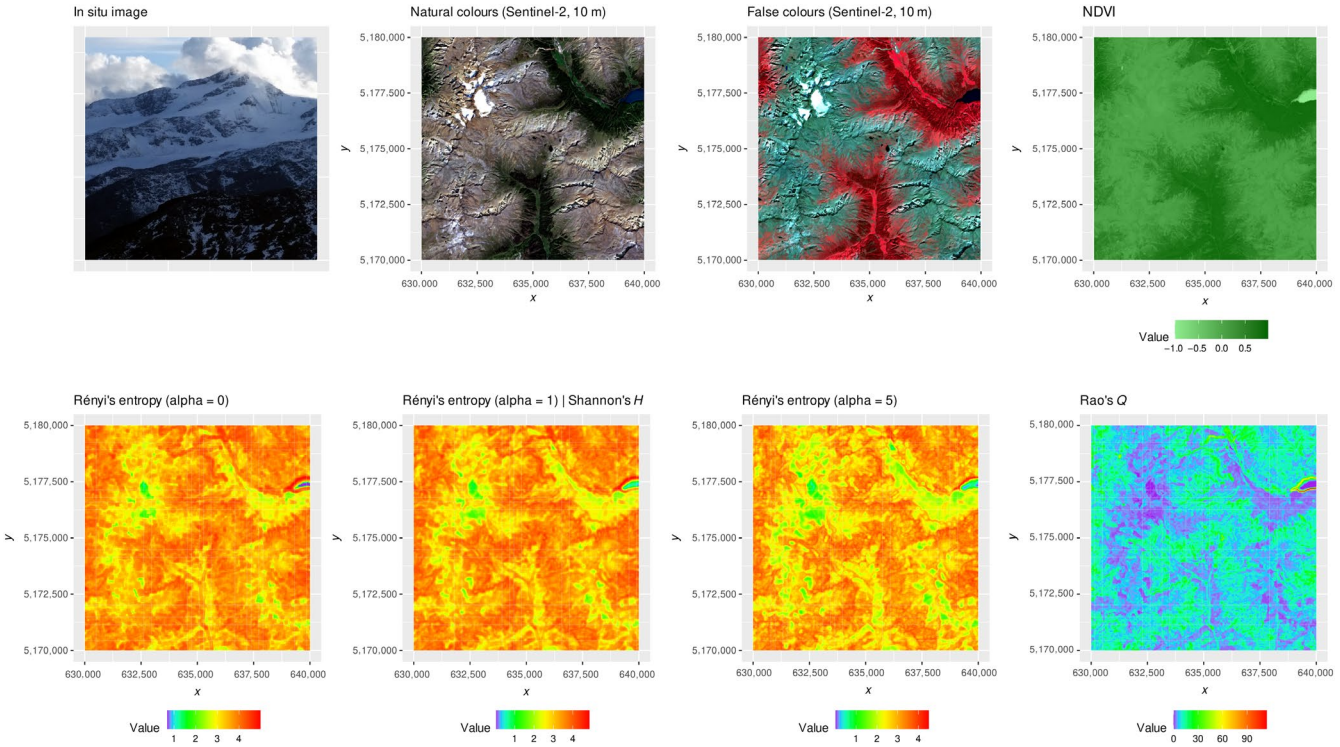

FIGURE 3 The first row of the figure represents the area under study, namely the Similaun glacier as a subset of the study area in Figure 2. NDVI is shown in the natural range but it was rescaled to 8-bit before heterogeneity computation. Then, different metrics were applied: the Rényi's entropy index, the Shannon's *H* (corresponding to the Rényi's entropy with $\alpha = 1$) and the Rao's *Q*. Coordinates are in the UTM (WGS84, zone 32N) reference system

Low heterogeneity values were found in areas with snow cover; in that case, the values of the neighbouring pixels are the same and they lead to a low Shannon's *H* for the focal pixel centered on the moving window. The saturation of high values of heterogeneity was apparent when considering Rényi's entropy at low alpha values (Equation 2). This is related to the aforementioned (see Section 2.2) negative exponential curve relating the value of this index with respect to alpha, shown in Ricotta et al. (2003). The result is a map with saturated values of heterogeneity. This effect is softening when increasing the alpha parameter. The rasterdiv package allows accounting for several indices at a time to avoid heterogeneity saturation effects, for instance by considering distance among pixel values besides relative abundance of each value. More specifically, running the Rao's *Q* function coded in rasterdiv allows to circumvent this issue of saturated values of heterogeneity (refer to the bottom line of Figure 3). In the Similaun glacier area, the homogeneous cover of spruce forests is better reflected by the Rao's *Q* index as it better contrasts against the geological heterogeneity of the upper alpine belt. Maximum Rao's *Q* values were found at the interface between water (i.e. alpine lakes) and the surrounding vegetation and rocks, representing an interesting ecotone area (see the upper right corner of the last panel in Figure 3).

The rasterdiv package, thanks to a combination of functions rooted in Information Theory, thus helps to reveal hidden spatial patterns of heterogeneity, and allows measuring ecosystem heterogeneity related to both biotic and abiotic components. By doing so, the rasterdiv package better discriminates among ecosystem features and functions constituting the landscape.

In the valley area (Val Venosta), this phenomenon was even more pronounced (Figure 4) due to the higher spatial heterogeneity of the landscape matrix under study with patches of small agricultural fields mixed with water bodies (e.g. the Adige river in the middle of the valley), resulting in very high values for both Rényi ($\alpha = 0$) and Shannon indices. Low to medium values on the north- and south-facing slopes indicate grasslands and broadleaf forests. Again, the Rao's *Q* index helped to better discriminate among these areas and among different land uses. Indeed, by relying on the relative numerical distance among the 8-bit-NDVI pixel values, it can differentiate between areas with low to medium heterogeneity values (blue and light blue colours in Figure 4), that is, grasslands and broadleaf forests, and areas with medium to high heterogeneity values, that is, upper mountain rocks at the interface with the treeline and with alpine lakes, as well as riparian vegetation besides the Adige river in the lower part of the valley.

With this study case focusing on a very heterogeneous area in the Alps, we illustrated two main critical components of the rasterdiv package: (a) how it allows users to measure multiple indices simultaneously and (b) how it can help detecting otherwise hidden patterns of heterogeneity in the landscape.

For the sake of clarity and to avoid a catalogue-like article, we did not present all the indices that can be calculated by the rasterdiv package. Instead, we decided to showcase a few, to illustrate to

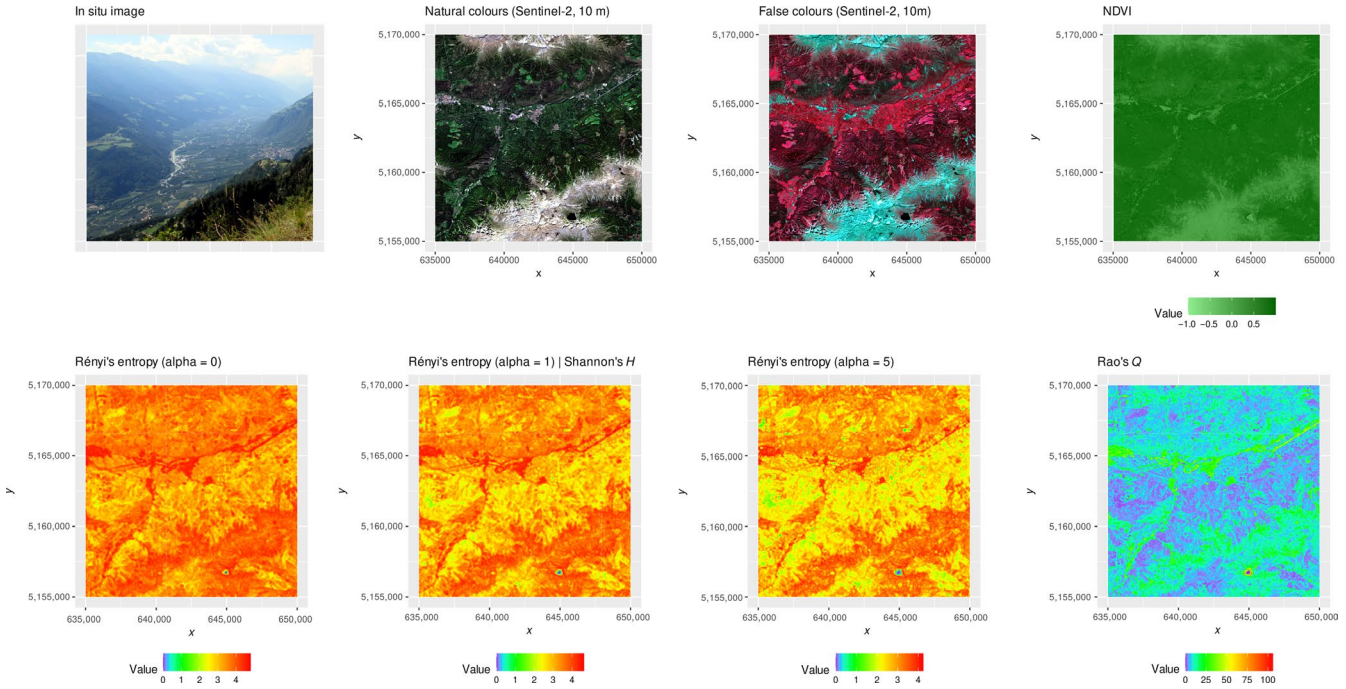

FIGURE 4 The first row of the figure represents the area under study, namely the Val Venosta, a subset at a lower elevation with respect to the Similaun glacier of Figure 3, and thus with higher human impact. NDVI is shown in the natural range but it was rescaled to 8-bit before heterogeneity computation. Then, different metrics were applied: the Rényi's entropy, the Shannon's *H* (corresponding to the Rényi's entropy with $\alpha = 1$) and the Rao's *Q*. Coordinates are in the UTM (WGS84, zone 32N) reference system

ecologists the high potential of the rasterdiv package for applications in landscape ecology, macroecology, biogeography and the analysis of spatiotemporal dynamics in general. We refer to the manual of the package (https://CRAN.R-project.org/package=rasterdiv) and its vignettes to find additional metrics and examples related to the package.

4 | DISCUSSION

In this paper, we provide an ecological overview of new R package rasterdiv. The overarching rationale for proposing this new R package is to present a set of methods and ready-to-use functions for calculating landscape heterogeneity metrics from space (e.g. satellite images), as well as from airborne or ground-based devices, to monitor and analyse, among other things, biodiversity change, habitat fragmentation and land use and cover changes.

The importance of computing continuous spectral heterogeneity measures from satellite-borne or airborne sensors to better discriminate among the various components in the landscape has been highlighted in several studies (Doxa & Prastacos, 2020; Godinho et al., 2018; Karlson et al., 2015; Ribeiro et al., 2019). Nevertheless, caution is recommended when making use of continuous remotely sensed data, and the radiometry of pixel values should be carefully considered before applying such measures. For instance, relying on float (decimal) precision data such as the NDVI (which ranges from −1 to 1) may lead to a high neighbouring heterogeneity which could actually be the effect of data binning rather than the effect of an ecological underlying pattern. In general, an 8-bit image (and therefore composed of 256 integer values/classes) is preferable when applying spectral heterogeneity measures. In this paper, we used an 8-bit NDVI layer rescaled from Copernicus data. However, one might even rely on a multispectral system reduced to a single 8-bit layer through by means of the first component of a principal component analysis or any multidimensionality reduction technique (Féret & Boissieu, 2020).

Most metrics based on Information Theory can accommodate only one layer at a time when relying on indices using abundance information only (in the shown example, the Rényi entropy index and Shannon's *H*). However, the rasterdiv package includes accounting for pixel values distances, such as the Rao's *Q* index, which can integrate multiple layers of ecological information such as multiple bands of satellite images or physical and biotic data (e.g. vegetation cover, soil pH, topography).

In general, remotely sensed data are simplifications of more complex systems depending on the radiometric and spectral properties of one or more images. From an ecological point of view, the spectral space of an image might be associated with the Hutchinson's hypervolume which orders geometrically the variables shaping species' ecological niches (Blonder, 2018; Hutchinson, 1959). Hence, calculating heterogeneity in such a space could provide important information about species niches variability or at least on the landscape variability shaping species distribution (Rocchini et al., 2018).

Future local and global changes are expected to impact ecosystem heterogeneity. Since remotely sensed data nowadays allow us to rely on relatively long and standardized time series,

applying different measures of heterogeneity to multi-temporal stacks would enhance the power to estimate and potentially forecast ecosystem heterogeneity shifts in space and time. This will be an invaluable tool to allow targeted and efficient monitoring and planning practices. For instance, due to the unprecedented rate of climate change, the adaptation of species to climate change is a benchmark in ecology (Stein et al., 2014). The rasterdiv package might also be particularly useful when aiming at calculating climate-related heterogeneity, which is likely to shape ecosystem heterogeneity patterns that species have adapted to. This could be directly done running the functions on remotely sensed climate data (Metz et al., 2014; Senner et al., 2018; Zellweger et al., 2019), which are expected to drive several ecological functions at different spatial scales.

5 | CONCLUSION

Measuring heterogeneity from space to understand ecological patterns and processes acting across the landscape and over different time periods is crucial to guide effective management practices, especially in the Anthropocene epoch, in which human intervention is leading to rapid environmental changes (Randin et al., 2020).

The proposed rasterdiv package is a powerful tool for monitoring spatial and temporal variation of ecosystems' properties (Rocchini et al., 2018), given the intrinsic relationship (sensu Laliberté et al., 2019) between the spatial variation of ecosystems and that of the spectral signal from pixel values (Rocchini et al., 2019). No single measure can provide a full description of all the different aspects of ecosystem heterogeneity. That is why the rasterdiv package offers multiple approaches to disentangle the complexity of ecosystem heterogeneity in space and time through calculations deeply rooted in Information Theory and based on reproducible Free and Open-Source algorithms.

ACKNOWLEDGEMENTS

We are grateful to the handling Editor and two anonymous reviewers who helped us improving a previous version of this manuscript with their precious suggestions. D.R. and D.K. were partially supported by the H2020 Project SHOWCASE (Grant agreement No 862480). D.R. was also partially supported by the H2020 COST Action CA17134 'Optical synergies for spatiotemporal sensing of scalable ecophysiological traits (SENSECO)'. The research carried out at the Jet Propulsion Laboratory, California Institute of Technology, was under a contract with the National Aeronautics and Space Administration (80NM0018D0004). Government sponsorship is acknowledged. This work was supported by Friuli Venezia Giulia Region Operative Program, European Social Fund—2014/2020 Program, Specific Action n. 53/16: Integrative professional training courses within degree programs. RF was partially supported by SNSF-175529. We thank the Czech University of Life Sciences Prague. Faculty of Environmental Sciences for supporting their research group Spatial Sciences in Ecology and Environment. The contribution of R.F. M.J.S., M.E.S. and L.S. is supported by the University of Zurich Research Priority Program on Global Change and Biodiversity (URPP GCB).

AUTHORS' CONTRIBUTIONS

D.R., M.I., E.T., M.Mar., D.D.R., G.B., E.M., C.T., S.V. and C.R. contributed to the development of the algorithms and the coding of the rasterdiv package. D.R., M.T., M.B., A.B., G.M.F., R.F., D.K., S.L., J.L., M.Mal., F.M., A.M., V.M., B.N., M.R., F.S., M.J.S., M.E.S., F.D.S., L.S., S.S., P.S., A.K.S., E.T., P.Z. and M.W. contributed to the conceptual development of the theoretical background of the rasterdiv package. All authors contributed to the writing of the manuscript.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13583.

DATA AVAILABILITY STATEMENT

Free data for applying the proposed functions are available directly into the rasterdiv package (https://CRAN.R-project.org/package=rasterdiv). The Copernicus Sentinel-2 image used in this paper can be freely downloaded from: https://scihub.copernicus.eu/, image tile reference: \S2A_MSIL2A_20200905T101031_N0214_R022_T32TPS_20200905T130252.

ORCID

Duccio Rocchini https://orcid.org/0000-0003-0087-0594
Matteo Marcantonio https://orcid.org/0000-0003-3896-2355
Daniele Da Re https://orcid.org/0000-0002-3398-9295
Giovanni Bacaro https://orcid.org/0000-0003-0946-4496
Manuele Bazzichetto https://orcid.org/0000-0002-9874-5064
Alessandra Bernardi https://orcid.org/0000-0002-2970-2730
Giles M. Foody https://orcid.org/0000-0001-6464-3054
Reinhard Furrer https://orcid.org/0000-0002-6319-2332
David Kleijn https://orcid.org/0000-0003-2500-7164
Stefano Larsen https://orcid.org/0000-0002-6774-1407
Jonathan Lenoir https://orcid.org/0000-0003-0638-9582
Marco Malavasi https://orcid.org/0000-0002-9639-1784
Vítězslav Moudrý https://orcid.org/0000-0002-3194-451X
Babak Naimi https://orcid.org/0000-0001-5431-2729
Carlo Ricotta https://orcid.org/0000-0003-0818-3959
Micol Rossini https://orcid.org/0000-0002-6052-3140
Maria J. Santos https://orcid.org/0000-0002-6558-7477
Michael E. Schaepman https://orcid.org/0000-0002-9627-9565
Fabian D. Schneider https://orcid.org/0000-0003-1791-2041
Sonia Silvestri https://orcid.org/0000-0003-3302-4335
Petra Šímová https://orcid.org/0000-0003-2480-1171
Andrew K. Skidmore https://orcid.org/0000-0002-7446-8429
Clara Tattoni https://orcid.org/0000-0003-1555-5669
Enrico Tordoni https://orcid.org/0000-0002-9722-6692
Piero Zannini https://orcid.org/0000-0003-2466-4402
Martin Wegmann https://orcid.org/0000-0003-0335-9601

REFERENCES

August, P. V. (1983). The role of habitat complexity and heterogeneity in structuring tropical mammal communities. *Ecology*, *64*, 1495–1507.

Blonder, B. (2018). Hypervolume concepts in niche- and trait-based ecology. *Ecography*, *41*, 1441–1455. https://doi.org/10.1111/ecog.03187

Boltzmann, L. E. (1872). Weitere studien über das waärmegleichgewicht unter gasmolekälen. *S K Akad Wiss Wein*, *66*, 275–370.

Da Re, D., De Clercq, E. M., Tordoni, E., Madder, M., Rousseau, R., & Vanwambeke, S. O. (2019). Looking for ticks from space: Using remotely sensed spectral diversity to assess *Amblyomma* and *Hyalomma* tick abundance. *Remote Sensing*, *11*, 770.

Doxa, A., & Prastacos, P. (2020). Using Rao's quadratic entropy to define environmental heterogeneity priority areas in the European Mediterranean biome. *Biological Conservation*, *241*, 108366.

Ettema, C. H., & Wardle, D. A. (2002). Spatial soil ecology. *Trends in Ecology & Evolution*, *17*, 177–183.

Fahrig, L. (2007). Landscape heterogeneity and metapopulation dynamics. In J. Wu & R. Hobbs (Eds.), *Key topics in landscape ecology* (pp. 78–91). Cambridge University Press.

Féret, J.-B., & de Boissieu, F. (2020). biodivMapR: An R package for α- and β-diversity mapping using remotely sensed images. *Methods in Ecology and Evolution*, *11*, 64–70.

Godinho, S., Guiomar, N., & Gil, A. (2018). Estimating tree canopy cover percentage in a mediterranean silvopastoral systems using Sentinel-2A imagery and the stochastic gradient boosting algorithm. *International Journal of Remote Sensing*, *39*, 4640–4662.

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

Grelle, C. E. V. (2003). Forest structure and vertical stratification of small mammals in a secondary Atlantic forest, southeastern Brazil. *Studies on Neotropical Fauna and Environment*, *38*, 81–85.

Heaney, L. R. (2001). Small mammal diversity along elevational gradients in the Philippines: An assessment of patterns and hypotheses. *Global Ecology and Biogeography*, *10*, 15–39.

Hesselbarth, M. H. K., Sciaini, M., With, K. A., Wiegand, K., & Nowosad, J. (2019). *landscapemetrics*: An open-source R tool to calculate landscape metrics. *Ecography*, *42*, 1648–1657.

Hijmans, R. J., & van Etten, J. (2020). *raster: Geographic analysis and modeling with raster data*. R package version 3.0-12. Retrieved from http://CRAN.R-project.org/package=raster

Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology*, *54*, 427–431. https://doi.org/10.2307/1934352

Hortal, J., Triantis, K. A., Meiri, S., Thébault, E., & Sfenthourakis, S. (2009). Island species richness increases with habitat diversity. *The American Naturalist*, *174*, E205–E217. https://doi.org/10.1086/645085

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

Karlson, M., Ostwald, M., Reese, H., Sanou, J., Tankoano, B., & Mattsson, E. (2015). Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and Random Forest. *Remote Sensing*, *7*, 10017–10041.

Keller, A., Graefen, A., Ball, M., Matzas, M., Boisguerin, V., Maixner, F., Leidinger, P., Backes, C., Khairat, R., Forster, M., Stade, B., Franke, A., Mayer, J., Spangler, J., McLaughlin, S., Shah, M., Lee, C., Harkins, T. T., Sartori, A., … Zink, A. (2012). New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nature Communications*, *3*, 698.

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

Leitão, P. J., Schwieder, M., & Senf, C. (2017). sgdm: An R package for performing Sparse Generalized Dissimilarity Modelling with tools for gdm. *ISPRS International Journal of Geo-Information*, *6*, 23.

Leutner, B., Horning, N., Schwalb-Willmann, J., & Hijmans, R. J. (2019). *RStoolbox: Tools for remote sensing data analysis*. R package version 0.2.6. Retrieved from http://CRAN.R-project.org/package=RStoolbox

Li, H., & Reynolds, J. (1995). On definition and quantification of heterogeneity. *Oikos*, *73*, 280–284.

Lozier, J. D., Strange, J. P., & Koch, J. B. (2013). Landscape heterogeneity predicts gene flow in a widespread polymorphic bumble bee, *Bombus bifarius* (Hymenoptera: Apidae). *Conservation Genetics*, *14*, 1099–1110. https://doi.org/10.1007/s10592-013-0498-3

Malanson, G. P., & Cramer, B. E. (1999). Landscape heterogeneity, connectivity, and critical landscapes for conservation. *Diversity and Distributions*, *5*, 27–39.

Metz, M., Rocchini, D., & Neteler, M. (2014). Surface temperatures at the continental scale: Tracking changes with remote sensing at unprecedented detail. *Remote Sensing*, *6*, 3822–3840.

Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., Toxopeus, A. G., & Alibakhshi, S. (2019). ELSA: Entropy-based local indicator of spatial association. *Spatial Statistics*, *29*, 66–88.

Nakamura, G., Goncalves, L. O., & Duarte, L. D. S. (2020). Revisiting the dimensionality of biological diversity. *Ecography*, *43*(4), 539–548. https://doi.org/10.1111/ecog.04574

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Randin, C. F., Ashcroft, M., Bolliger, J., Cavender-Bares, J., Coops, N., Dullinger, S., Dirnböck, T., Eckert, S., Ellis, E., Giuliani, G., Guisan, A., Jetz, W., Joost, S., Karger, D., Lembrechts, J., Lenoir, J., Luoto, M., Morin, X., Price, B., … Payne, D. (2020). Monitoring biodiversity in the Anthropocene using remote sensing in species distribution models. *Remote Sensing of Environment*, *239*, 111626.

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.

Ribeiro, I., Proenca, V., Serra, P., Palma, J., Domingo-Marimon, C., Pons, X., & Domingos, T. (2019). Remotely sensed indicators and open-access biodiversity data to assess bird diversity patterns in mediterranean rural landscapes. *Scientific Reports*, *9*, 1–13.

Ricotta, C., Corona, P., Marchetti, M., Chirici, G., & Innamorati, S. (2003). LaDy: Software for assessing local landscape diversity profiles of raster land cover maps using geographic windows. *Environmental Modelling & Software*, *18*, 373–378.

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., 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., … 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., & Neteler, M. (2012). Let the four freedoms paradigm apply to ecology. *Trends in Ecology & Evolution*, *27*, 310–311.

Rocchini, D., & Ricotta, C. (2007). Are landscapes as crisp as we may think? *Ecological Modelling*, *204*, 535–539. https://doi.org/10.1016/j.ecolmodel.2006.12.028

Schimel, D., & Schneider, F. D. (2019). Flux towers in the sky: Global ecology from space. *New Phytologist*, *224*, 570–584. https://doi.org/10.1111/nph.15934

Schneider, F. D., Morsdorf, F., Schmid, B., Petchey, O. L., Hueni, A., Schimel, D. S., & Schaepman, M. E. (2017). Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nature Communications*, *8*, 1441.

Senner, N. R., Stager, M., & Cheviron, Z. A. (2018). Spatial and temporal heterogeneity in climate change limits species' dispersal capabilities and adaptive potential. *Ecography*, *41*, 1428–1440.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(379–423), 623–656. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*, 688.

Skidmore, A. K., Franklin, J., Dawsonc, T. P., & Pilesjo, P. (2011). Geospatial tools address emerging issues in spatial ecology: A review and commentary on the Special Issue. *International Journal of Geographical Information Science*, *25*, 337–365.

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, *17*(7), 866–880.

Vranken, I., Baudry, J., Aubinet, M., Visser, M., & Bogaert, J. (2014). A review on the use of entropy in landscape ecology: Heterogeneity, unpredictability, scale dependence and their links with thermodynamics. *Landscape Ecology*, *30*, 51–65.

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.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Rocchini D, Thouverai E, Marcantonio M, et al. rasterdiv—An Information Theory tailored R package for measuring ecosystem heterogeneity from space: To the origin and back. *Methods Ecol Evol*. 2021;12:1093–1102. https://doi.org/10.1111/2041-210X.13583