Measuring spatial diversity in a free algorithmic environment

Duccio Rocchini[1], Luca Delucchi[1,*],Carlo Ricotta[2], Anne Ghisla[1], Cristina Castellani[1], Roberto Zorer[1], Markus Neteler[1]

[1] IASMA Research and Innovation Centre, Fondazione Edmund Mach, Environment and Natural Resources Area, Via E. Mach 1, 38010 S. Michele all'Adige, TN, Italy. luca.delucchi@iasma.it
[2] Dipartimento di Biologia Vegetale, Universita "La Sapienza" di Roma, piazzale A. Moro 5, 00185 Roma, Italy

1. Introduction

Estimating complexity is complex. This is particularly true when the complexity (i.e. diversity) of the environment at different spatial scales (from regional to global) is taken into account, due to the difficulty of finding adequate proxies of landscape diversity. One of the most powerful tools to be used is spectral diversity, i.e. the diversity of reflectance values recorded by remotely sensed imagery (Palmer et al. 2002). Indices based on compositional and structural diversity are routinely used, but scientists are mainly relying on black box-based (blind) calculations. Bruce Hapke (2005), rearranging the well known Murphy's law, stresses that:

"The equation you need the most contains a typographical error […]"

This can hardly be solved when the code used for building equations is not explicitly available (Neteler et al., 2008).

The aim of this study is to provide a new Open Source script (*r.diversity*) which incorporates the mostly used measures of diversity together with their derivatives and new measures proposed in recent scientific papers.

2. Theory behind the *r.diversity* script

While famous examples of Open Source software exist in other research fields like statistics, GIS benefits from a powerful tool which includes more than 350 modules for for managing and analyzing geographical data, named GRASS GIS, which represents nowadays one of the challenging projects of the Open Source Geospatial Foundation (OSGeo).

Given a remotely sensed image (composed by Digital Number values) or a land use map, calculating diversity by the *r.diversity* script into GRASS GIS relies on the mostly used diversity indices such as:

- the Shannon index H (Shannon, 1948):

$$H = -\sum_{1=1}^{S} p_i \times \ln(p_i) \qquad [1]$$

where S is the total number of different DN values (or classes) and p_i is the proportion of occupancy of each DN value (or class);

- the Simpson index *D* (Simpson, 1949):

$$D=-\sum_{1=1}^{S} p_i^2 \qquad [2]$$

- the Pielou evenness index *J* (Pielou, 1966):

$$J=\frac{H}{\ln(S)} \qquad [3]$$

- the Rényi entropy of order a (Rényi, 1961)

$$H_\alpha=\frac{1}{1-\alpha}\times\ln(\sum_{i=1}^{S} p_i^\alpha) \qquad [4]$$

where $\alpha\geqslant 0$.

All the aforementioned indices are particularly powerful since they account for both richness and relative abundance of e.g. reflectance values or land use classes derived from remotely sensed images.

Refer to Ricotta and Avena (2003) for a complete mathematical dissertation of these indices together with their possible extensions and derivatives.

3. Output of *r.diversity*

Fig.1 represents an example of remotely sensed diversity calculation based on the *r.diversity* script. Once a moving window size is defined (e.g. 3x3 pixels at minimum) *r.diversity* calculates the variability of pixel values into each moving window based on the method parameter (Shannon, Simpson, Pielou, Rényi's entropy) chosen by the analyst, and it attaches to the central point the derived diversity value.

In this example an 8 bit image (256 DN values) is used and two indices (Shannon and Simpson diversity) are presented. Notice that higher values of Shannon and Simpson diversity occur where the input image presents margins among different habitat types (higher habitat heterogeneity).

It is important to note that the used diversity indices work once input images are coded by classes (e.g. DN values at a given radiometric resolution, land use classes, etc.). Hence it is recommended to avoid calculation on continuous values such as NDVI or PCA values. Instead, it is strongly recommended to bin NDVI or PCA continuous values into *N* equal intervals before calculation. Generally the 8-bit integer range (0-255) is used but this range is arbitrary and other ranges could equally be adopted.

4. Advantages and future challenges of the proposed approach

The Open Source nature of the code upon which *r.diversity* is based guarantees that the whole research community working on the diversity theme may benefit from the presented algorithms, on the strength of the notorious “four freedoms” paradigm.

Environmental heterogeneity is considered to be one of the main factors associated with biodiversity given that habitats with highly heterogeneous environments can host more species due to their higher number of available niches. In this view, the use of robust but straightforward tools for measuring remotely sensed variability is challenging since these data provide an inexpensive means of deriving environmental information for large areas in a consistent and regular manner.

We hope that scientists relying on such measures will adopt this Open Source script into GRASS GIS for future development of robust measures of diversity derived from remotely sensed imagery.

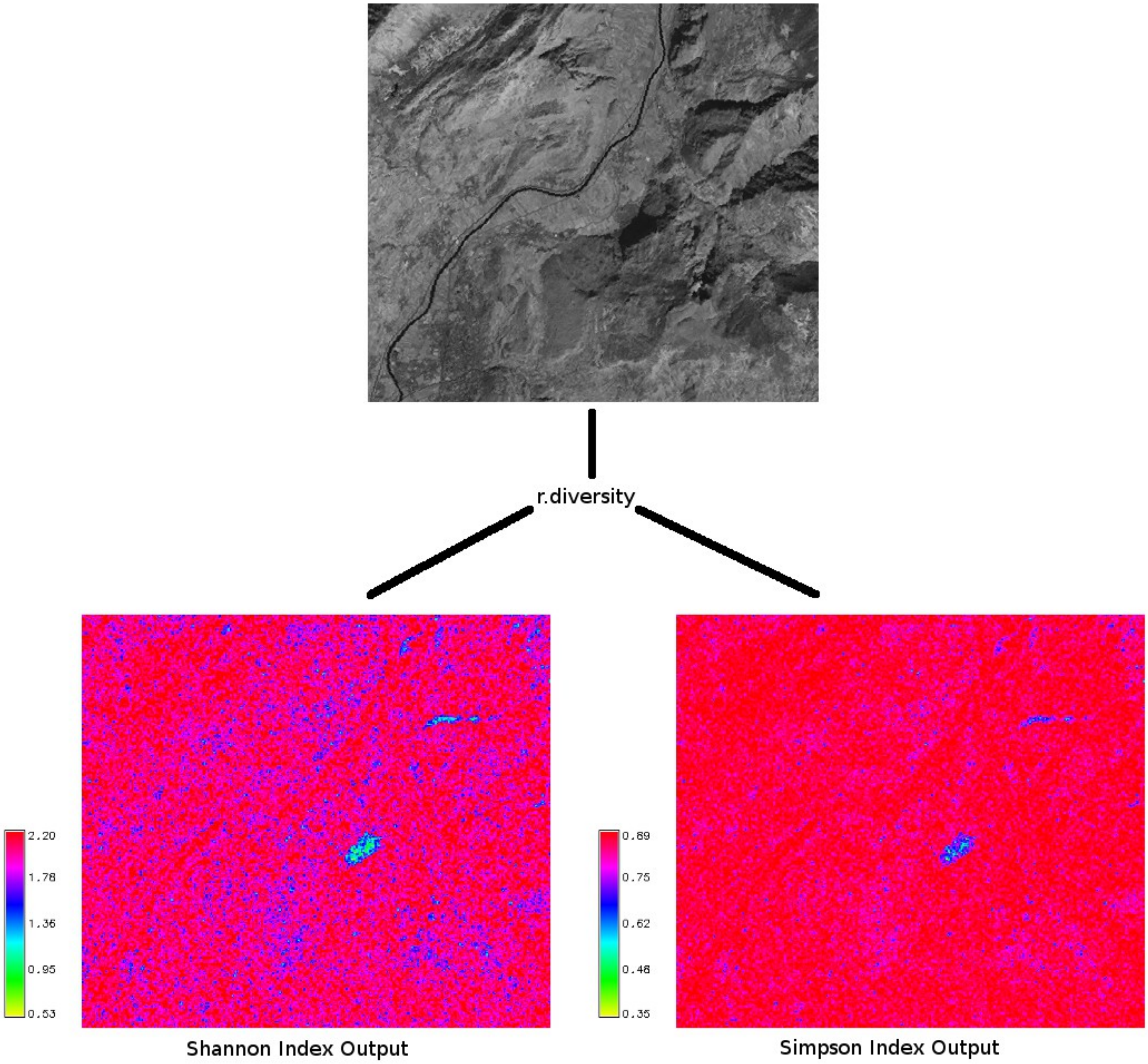

Figure 1: Example of the *r.diversity* script procedure. Once the moving window size is defined, *r.diversity* calculates the variability of pixel values into the spatial window based on the method parameter chosen by the analyst. In this example, an 8-bit Landsat image is used. The Shannon and Simpson indices are presented as an output example.

References

Hapke, B., 2005. 'Theory of reflectance and emittance spectroscopy'. Cambridge University Press, UK.

Neteler, M., Beaudette, D.E., Cavallini, P., Lami, L., Cepicky, J., 2008. 'GRASS GIS'. In: G.B. Hall (Ed), 'Open Source Approaches to Spatial Data Handling', pp. 171-199, Springer, New York.

Palmer, M.W., Earls, P., Hoagland, B.W., White, P.S., Wohlgemuth, T., 2002. 'Quantitative tools for perfecting species lists'. *Environmetrics* 13, 121-137.

Pielou, E.C., 1966. 'Shannon's formula as a measure of species diversity: its use and misuse'. *Am. Nat.* 100, 463–465.

Rényi, A., 1961. 'On measures of entropy and information'. In: Neyman, G. (Ed.), Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1. University of California Press, Berkeley, CA, pp. 547–561.

Ricotta, C., Avena, G.C., 2003. 'On the relationship between Pielou's evenness and landscape dominance within the context of Hill's diversity profiles'. *Ecological Indicators* 2, 361-365.

Shannon, C., 1948. 'A mathematical theory of communication'. *Bell Syst. Tech. J.* 27, 379–423.

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