Popul Ecol (2011) 53:597–601
DOI 10.1007/s10144-011-0272-7

NOTES AND COMMENTS

Comments on Bartolino et al. (2011): limits of cumulative relative frequency distribution curves for hotspot identification

Luis Cayuela · Lucía Gálvez-Bravo · Luis María Carrascal · Fábio S. de Albuquerque

Received: 20 September 2010 / Accepted: 7 February 2011 / Published online: 26 May 2011

Abstract The recent paper by Bartolino et al. (Popul Ecol 53:351–359, 2011) presents a new method to objectively select hotspots using cumulative relative frequency distribution (CRFD) curves. This method is presented as being independent from the selection of any threshold and, therefore, less arbitrary than traditional approaches. We argue that this method, albeit mathematically sound, is based on likewise arbitrary decisions regarding threshold selection. Specifically, the use of the CRFD curve approach requires the occurrence of two criteria for the method to be applied correctly: the selection of a 45° tangent to the curve, and the need to consider the highest relative value of the study parameter corresponding to a 45° slope tangent to the curve. Using two case studies (dealing with species richness and abundance of a particular species), we demonstrate that these two criteria are really unrelated to the underlying causes that shape the spatial pattern of the phenomena under study, but rather related to sampling design and spatial scale; hence, one could likewise use different but valid criteria. Consequently, the CRFD curve approach is based on the selection of a pre-defined threshold that has little, if any, ecological justification, and that heavily influences the final hotspot selection. Therefore, we conclude that the CRFD curve approach itself is not necessarily better and more objective than any of the global methods typically used for hotspot identification. Indeed, mathematical and/or statistical approaches should not be viewed as a panacea to solve conservation problems, but rather used in combination with biological, practical, economic and social considerations.

Keywords Biodiversity hotspot · Conservation prioritization · Density · High value conservation areas · Species richness

A reply to this "notes and comments" is available at doi: 10.1007/s10144-011-0273-6.

L. Cayuela (✉)
Área de Biodiversidad y Conservación, Universidad Rey Juan Carlos, c/Tulipán s/n, 28933 Móstoles (Madrid), Spain
e-mail: luis.cayuela@urjc.es

L. Gálvez-Bravo
Grupo de Investigación UNGULATA (Ecología, Comportamiento y Conservación de Ungulados), Instituto de Investigación en Recursos Cinegéticos (CSIC-UCLM-JCCM), Ronda de Toledo s/n, 13071 Ciudad Real, Spain

L. M. Carrascal
Dpto. de Biodiversidad y Ecología Evolutiva, Museo Nacional de Ciencias Naturales, MNCN-CSIC, c/José Gutiérrez Abascal 2, 28006 Madrid, Spain

F. S. de Albuquerque
EcoLab, Centro Andaluz de Medio Ambiente, Universidad de Granada-Junta de Andalucía, 18006 Granada, Spain

Introduction

The definition of biodiversity hotspots is clearly a subjective matter with respect to the definition of a cut-off point between areas that contain an extraordinary number of species from those that contain only an ordinary number of species (Fortin and Dale 2005). Defining hotspots is a first step towards understanding processes that favor the presence of especially high species richness in certain areas and a helpful indicator—albeit not the only one—for prioritizing conservation actions at different spatial scales (Myers et al. 2000; Cayuela et al. 2006; Altamirano et al. 2010). Consequently, ecologists and conservationists have been—and still are—interested in developing criteria to discriminate hotspot from non-hotspot locations along a continuum of species richness or any other biological measure (for simplicity, we will refer only to species richness hereafter).

Most often, spatially global methods for hotspot identification are used (Myers et al. 2000; Gjerde et al. 2004; Grand et al. 2004; Orme et al. 2005; Cayuela et al. 2006; Altamirano et al. 2010). These compare the richness value at a given observation with those in the complete dataset, as opposed to spatially local methods, which compare the value at a given observation with locations in the vicinity of that observation (see Nelson and Boots 2008 for a review). Spatially global methods usually rely on the subjective identification of a species richness value above which locations are identified as hotspots. Typically, this threshold ranges between 1 and 5% (Myers et al. 2000; Gjerde et al. 2004; Grand et al. 2004; Orme et al. 2005; Cayuela et al. 2006), but it can be raised up to 25% depending on the specific conservation goals (García 2006). In an attempt to overcome the problem of subjectivity, Bartolino et al. (2011) proposed a new method to objectively select species richness (or abundance of a particular species of interest) hotspots using cumulative relative frequency distribution (CRFD) curves (henceforth referred to as the CRFD curve approach). This method is presented as being independent from the selection of any threshold and, therefore, less arbitrary than traditional approaches. Here, we argue that the method proposed, albeit mathematically sound, is based on likewise arbitrary decisions regarding the selection of a threshold, and that the CRFD curve approach itself is not necessarily better and more objective than any of the global methods typically used for hotspot identification.

A frequency distribution approach for hotspot identification

The methodology proposed by Bartolino et al. (2011) is based on the following steps: (1) plot a CRFD curve of the phenomenon of interest (species richness or abundance of a particular species of interest), where both axes, *x*-variable of interest and *y*-cumulative number of cases, are relativised, ranging from 0 to 1; (2) calculate the tangent to the curve at each point; and (3) select the highest value of frequency distribution of sampling units (i.e., census plots, areas, pixels) corresponding to a 45° slope tangent to the curve. A tangent with >45° slope indicates areas where the relative increase in frequency of sampling units is greater than the relative increase in species richness or abundance, whereas a tangent with <45° slope indicates areas where the relative increase in frequency of sampling units is smaller than the relative increase in species richness or abundance. According to the authors, this threshold will separate areas where the variable considered can be found at lower values (tangent with >45° slope) from those where it is found at higher values (tangent with <45° slope). In their study, Bartolino et al. (2011) also state that application of all global methods, including the CRFD curve approach, can only be effectively applied in the case of spatially autocorrelated and homogeneous processes. We will also discuss the validity of such assumptions in the following section.

Limits to the CRFD curve approach

The geometric approach of Bartolino et al. (2011) offers an easy and practical way to identify hotspots by using geometrical properties of CRFD curves. However, the authors claim that the main strength of this method is its lack of subjectivity as compared to other widely used methods. Our main objection to this claim is that the use of the CRFD curve approach requires the occurrence of two criteria for the method to be applied correctly: the selection of a 45° tangent to the curve, and the need for considering only the 45° tangent occurring for the highest relative value. In our opinion, these criteria pose some problems, since they are not directly related to the phenomenon under study (either the spatial patterns of species richness or abundance or the natural patterns of variation of those parameters according to the shape of their distributions—gaussian, left- or right-skewed) but rather to the sampling process. Although the authors state that "[…] the shape of the CRFD curve varies in relation to how the intensity of the spatial process changes across the area, revealing important aspects of the underlying spatial phenomenon being considered", such aspects are not mentioned at all in their paper. Using a couple of case studies, we will provide further evidence of some rationale problems behind the CRFD curve approach, and will demonstrate that the criteria provided by Bartolino et al. (2011) are arbitrary and their method as subjective as any other global method for identification of hotspots.

To illustrate our point, we applied the same analytical procedure as Bartolino et al. (2011) to two different datasets: one of herb species richness in 10 × 10 km grid cells in the United Kingdom, and another of Canary Islands stonechat *Saxicola dacotiae* abundance (n/km^2) in 2 × 2 km grid cells in Fuerteventura Island, Spain. Distributions maps of terrestrial herbaceous plants for Great Britain were obtained from the New Atlas of the British and Irish Flora (Preston et al. 2002). Plant maps and UTM grid cells of 10 × 10 km each were intersected in ArcGiS 9.3 and herb richness was calculated for each UTM grid cell (further information can be obtained in Albuquerque et al. 2011). Data for the endangered population of Canary Islands stonechat were obtained from the census program described in Seoane et al. (2010a): 1,462 0.5-km line transects during the reproductive seasons in 2005 and 2006, distributed across the whole island, that were assigned to 2 × 2 UTM squares, and estimation of detectability functions to obtain absolute densities. Sample sizes were $n = 2{,}252$ and $n = 286$ for the herb richness and stonechat density datasets, respectively.

The results of applying the CRFD curve approach to both datasets are shown in Fig. 1. One can easily realise that a tangent with <45° slope does not necessarily correspond to areas where the variable considered, herb species richness or bird density, respectively, is found at high densities and vice versa, as pointed out by Bartolino et al. (2011). In both datasets, several points with both low and high values fulfill the 45° slope tangent rule (Fig. 1c, f), so clearly they do not necessarily correspond to areas where the variable considered is found at high values. This is probably the reason why Bartolino et al. (2011) suggest that "[…] that the highest x_0 (i.e., relative species richness or abundance) corresponding to a 45° slope tangent to the curve can be used as a global threshold to identify hotspots". In this way, they restrict the 45° slope tangent to the curve rule in order to select only those areas where the variable considered is found at high values. The main conclusion one can extract from this is that the tangent to the CRFD curve may not provide a meaningful criterion related to patterns of species richness or density of a particular taxa, but rather it reflects the average increase in species richness or density as new locations are added to the overall sample. Sites with a tangent of <45° slope are simply those that experience an increase in species richness of at least one species for each new location added to the overall sample. Since this is really unrelated to the underlying causes that shape the spatial pattern of the phenomena under study but rather related to sampling design and spatial scale (Willis and Whittaker 2002; McGill 2010), one could likewise use a different but valid criterion, such as selecting sites with a tangent of <35° slope. Consequently, the CRFD curve approach, as with other global and partially local methods, is based on the selection of a pre-defined threshold that has little, if any, ecological justification, and that heavily influences the final hotspot selection. That threshold is related to the geometry of the CRFD curve itself, derived from the shape of the frequency distribution of absolute values of the ecological parameters under study.

A second source of arbitrariness in the CRFD curve approach relates to the selection of the highest relative value of species richness or abundance corresponding to a 45° slope tangent to the curve, which is proposed as a global threshold to identify hotspots. The application of the CRFD curve approach to our two case studies shows that cells with over 579 herb species (Fig. 1b) and 45.7 birds/km^2 (Fig. 1e) are hotspots in the UK and Fuerteventura Island, respectively. But are there really any biologically meaningful differences between these areas and other areas holding a slightly lower number of species or individuals? It is unlikely. Furthermore, there are several locations that, in spite of harboring a really high number of species, have a tangent with <45° slope, but since these values fall below the highest relative species richness corresponding to a 45° slope tangent to the curve, they are not classified as hotspots (see Fig. 1c). Again, different criteria might be valid here, for example selecting those values that fall below the second or third highest relative species richness or density value corresponding to a 45° slope tangent to the curve. On the other hand, the threshold figure of 45.7 birds/km^2 for the Canary Islands stonechat in Fuerteventura is a very restrictive cut-off point defining population hotspots for the species. This density level only selects 5% of the distribution range of the species in the island (60 km^2), and includes 21.6% of its total population (Seoane et al. 2010b), being of limited value for conservation purposes (both in terms of total population accounted and considering the fragmentation of the selected distribution areas as defined by that threshold, within the whole stonechat range). Therefore, we conclude that there is no biological meaning of this threshold that allows the objective classification of hotspots and non-hotspot areas, and consequently the rationale underpinning the use of this method is limited and only relies on mathematical grounds using the 45° slope as a 'golden standard'.

Finally, it is necessary to clarify that, contrary to what Bartolino et al. (2011) stated in their paper, the use of any global method, including the CRFD curve approach, does not necessarily have to comply with the assumption of spatially autocorrelated surfaces. Meeting the assumption of spatial autocorrelation will enable any global method to incorporate the local structure of the phenomenon of interest, and results will consequently be more similar to those obtained by means of spatially local methods. On the contrary, if spatial autocorrelation is low, global methods will also incorporate the local structure of the phenomenon of interest, but in this case local structure will have little influence on the selection of hotspots. Spatial autocorrelation must be viewed as a characteristic of the spatial pattern of the phenomenon under study and not as a requirement for the performance of the CRFD curve approach (or any other spatially global method). As for the assumption of homogeneous processes, we agree with Bartolino et al. (2011) that identification of hotspots over large areas characterized by non-stationarity will gain from partitioning into smaller regions of investigation if one is

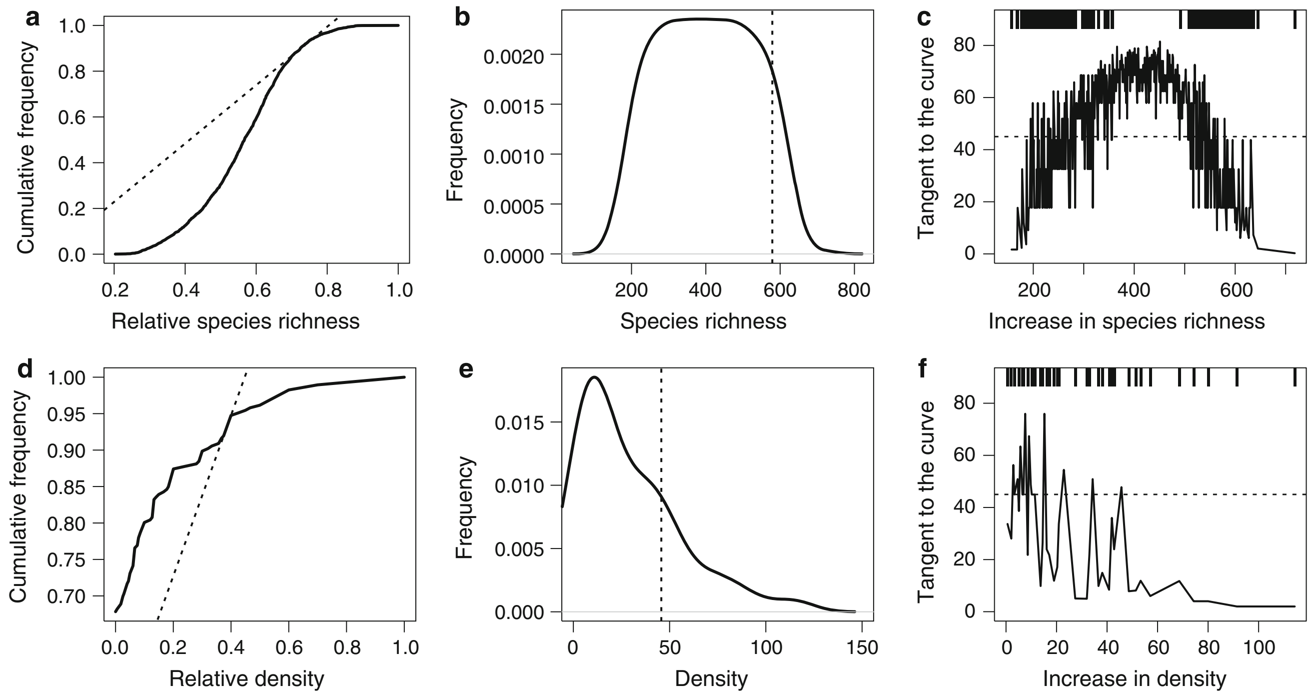

Fig. 1 **a**, **d** Cumulative relative frequency distribution curve of herb species richness and Canary Islands stonechat *Saxicola dacotiae* abundance (n/km^2) showing the 45° slope tangent to the curve (*dotted line*) at the highest species richness value in 10 × 10 km grid cells in the UK, or the highest density in 2 × 2 km grid cells in Fuerteventura Island, respectively. **b**, **e** Frequency distribution of herb species richness and Canary Islands stonechat abundance in 10 × 10 km grid cells in the UK and 2 × 2 km grid cells in Fuerteventura Island, Spain, respectively; the *dotted line* separates the proportion of observations/locations above the identified threshold. **c**, **f** Slope of tangent to the curve $f(x)$ at a point $(x_0, f(x_0))$, in degrees, corresponding to the derivative of $f(x)$ in x_0 (sensu Bartolino et al. 2011), for herb species richness and Canary Islands stonechat in 10 × 10 km grid cells in the UK and 2 × 2 km grid cells in Fuerteventura Island, Spain, respectively. *Small vertical lines* in the *upper part* of the graphs indicate tangents to the curve with <45° slope

interested in not losing locally relevant hotspots. Nonetheless, locally important hotspots can be irrelevant for the goals of a particular study (e.g., Myers et al. 2000). In such cases, non-stationarity will not be a problem for the application of any spatially global method.

Further prospects in the use of methods for hotspot identification

In most cases, species richness and abundance patterns change across a continuum, and almost any criterion to harden this continuum into classes (hotspot vs. non hotspot) will fall into some degree of subjectivity. In their study, Bartolino et al. (2011) compared the results obtained from the CRFD curve approach with those obtained using other spatially local and global methods, and found that the main hotspots were detected by all the techniques they considered. They observed, however, wide variations in the number and size of the identified hotspots, particularly when comparing global methods using datasets with low spatial autocorrelation with spatially local methods. Since the definition of thresholds to identify hotspots from non-hotspots areas based on the CRFD curve approach may be as arbitrary as any other commonly used global threshold (e.g., considering the 5, 10, or 15% highest values as hotspots), this method should be used with caution. Alternatively, we suggest the use of the CRFD curve approach in combination with existing spatially local and global methods to build a consensual output. The robustness of the CRFD curve approach to variations in the distribution shapes of the ecological parameters employed to define hotspots should also be investigated, because this method relies on the relative value of species richness or abundance compared to the highest recorded figures, and very different shapes of parameter distributions can be obtained with the same maximum values. In other words, how does the definition of hotspot change according to the kurtosis and skewness of the studied parameters? How is the shape of these distributions affected by both the total number of hotspots and their spatial distribution (e.g., interconnection among them, fragmentation of the study region due to hotspot selection)?

Finally, it is the role of scientists to inform policy makers engaged in conservation about which areas to prioritize according to different criteria. One such criterion is species richness, but there can be many others, such as species rarity, vulnerability (Rey Benayas and de la Montaña 2003) or complementarity (Justus and Sarkar 2002; Cayuela et al. 2006). Identifying hotspots from non-hotspot areas is one way to prioritize high-value conservation areas. Scientists can achieve this through an intellectual exercise and as a way of highlighting the need to strengthen the conservation of natural areas (e.g., Myers et al. 2000), but this must be taken as just an indication of what needs to be and what can be preserved given political, economic and social considerations. Therefore, whether the criteria used to identify hotspots are arbitrary or not is not really relevant, because hotspot selection must be understood in the particular context of each case study. We go further and believe that any criterion will be to some extent arbitrary, even those that claim to be objective, such as the CRFD curve approach. Mathematical and/or statistical approaches should not be viewed as a panacea to solve conservation problems for two main reasons: currently, they all involve inherently subjective decisions, and they miss biological, practical, economic and social considerations. However, once the most appropriate tool for hotspot definition is selected, given the specific characteristics of the case study at hand and justifying its use and threshold selection with sound biological knowledge, hotspot identification can be a useful tool to be included in conservation prioritization processes.

References

Albuquerque FS, Castro-Díez P, Rodríguez MA, Cayuela L (2011) Assessing the influence of environmental and human factors on native an exotic species richness. Acta Oecol 37:51–57

Altamirano A, Field R, Cayuela L, Aplin P, Lara A, Rey-Benayas JM (2010) Woody species diversity in temperate Andean forests: the need for new conservation strategies. Biol Conserv 143:2080–2091

Bartolino V, Maiorano L, Colloca F (2011) A frequency distribution approach to hotspot identification. Popul Ecol 53:351–359

Cayuela L, Rey Benayas JM, Justel A, Salas-Rey J (2006) Modelling tree diversity in a highly fragmented tropical montane landscape. Global Ecol Biogeogr 15:602–613

Fortin MJ, Dale MRT (2005) Spatial analysis: a guide for ecologists. Cambridge University Press, Cambridge

García A (2006) Using ecological niche modeling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. Biol Conserv 130:25–46

Gjerde I, Rolstad J, Blom HH, Storaunet KO (2004) Fine-scale diversity and rarity hotspots in northern forests. Conserv Biol 18:1032–1042

Grand J, Buonaccorsi J, Cushman SA, Griffin CR, Neel MC (2004) A multiscale landscape approach to predicting bird and moth rarity hotspots in a threatened pitch pine-scrub oak community. Conserv Biol 18:1063–1077

Justus J, Sarkar S (2002) The principle of complementarity in the design of reserve networks to conserve biodiversity: a preliminary history. J Biosci 27:421–435

McGill BJ (2010) Matters of scale. Science 328:575–576

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403:853–858

Nelson TA, Boots B (2008) Detecting spatial hot spots in landscape ecology. Ecography 31:556–566

Orme CDL, Davies RG, Burgess M, Eigenbrod F, Pickup N, Olson VA, Webster AJ, Ding TS, Rasmussen PC, Ridgely RS, Statterfield AJ, Bennett PM, Blackburn TM, Gaston KJ, Owen IPF (2005) Global hotspots of species richness are not congruent with endemism or threat. Nature 436:1016–1019

Preston CD, Pearman DA, Dines TD (2002) New atlas of the British and Irish flora: an atlas of the vascular plants of Britain, Ireland, the Isle of Man and the Islands. Oxford University Press, Oxford

Rey Benayas JM, de la Montaña E (2003) Identifying area of high-value vertebrate diversity for strengthening conservation. Biol Conserv 114:357–370

Seoane J, Kouri A, Illera JC, Palomino D, Alonso CL, Carrascal LM (2010a) New data on the population, distribution and habitat preferences of the Canary Islands stonechat *Saxicola dacotiae*. Ardeola 57:387–405

Seoane J, Kouri A, Illera JC, Palomino D, Alonso CL, Carrascal LM (2010b) La tarabilla canaria en España. Población reproductora en 2005–2006 y método de censo. Monografía SEO/BirdLife, Madrid (in Spanish)

Willis KJ, Whittaker RJ (2002) Species diversity—scale matters. Science 295:1245–1248