

APPLICATION

fuzzySim: applying fuzzy logic to binary similarity indices in ecology

A. Márcia Barbosa*

Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), InBIO Research Network in Biodiversity and Evolutionary Biology, University of Évora, 7004-516 Évora, Portugal

Summary

1. Binary similarity indices are widely used in ecology, for example for detecting associations between species occurrence patterns, comparing regional and temporal species assemblages, and assessing beta diversity patterns, including spatial and temporal species loss and turnover. Such indices have widespread applications in biogeography, global change biology and biodiversity conservation.
2. Similarity indices are commonly calculated upon binary presence/absence (or sometimes modelled suitable/unsuitable) data, which are generally incomplete and more categorical than their underlying natural patterns. Probable false absences are disregarded, amplifying the effects of data deficiencies and the scale dependence of the results.
3. Fuzzy occurrence data, with a degree of uncertainty attributed to localities where presence or absence cannot be safely assigned, could better reflect species distributions, compensating for incomplete knowledge and methodological errors. Similarity indices would therefore also benefit from accommodating such fuzzy data directly.
4. This study proposes fuzzy versions of the binary similarity indices most commonly used in ecology, so that they can be directly applied to continuous (fuzzy) rather than binary occurrence values, thus producing more realistic similarity assessments. Fuzzy occurrence can be obtained with several methods, some of which are also provided. The procedure is robust to data source disparities, gaps or other errors in species occurrence records, even for restricted species for which slight inaccuracies can affect substantial parts of their range.
5. The method is implemented in a free and open-source software package, fuzzySim, which is available for the R statistical software and under implementation for the QGIS geographic information system. It is provided with sample data and an illustrated tutorial suitable for non-experienced users.

Key-words: biogeographic regions, biotic regionalization, beta diversity, chorotypes, fuzzy logic, specific composition, species distributions, vagueness

Introduction

Spatial associations between species distributions provide deep insights into the processes that drive biodiversity patterns. Areas with similar species compositions (biogeographic or biotic regions) (Márquez, Real & Vargas 2001; Holt *et al.* 2013; Olivero, Márquez & Real 2013) or species with similar occurrence patterns (chorotypes) (Baroni Urbani, Ruffo & Vigna Taglianti 1978; Márquez *et al.* 1997; Real, Olivero & Vargas 2008; Olivero, Real & Márquez 2011) serve, for example, as natural units to maximize efficiency in biodiversity management and conservation planning. Beta diversity patterns provide a link between local biodiversity and the broader regional species pool, and are at the core of community ecology (Anderson *et al.* 2011). Like chorotypes and biotic regions, they can be used in tests of hypotheses about the processes driving species distribution and diversity (Baselga & Orme 2012) and can

reveal effects of large-scale historic processes on current biodiversity (Baselga, Gómez-Rodríguez & Lobo 2012).

Identifying chorotypes, biotic regions and beta diversity patterns requires comparing species occurrence patterns via similarity indices. These indices are typically based on the numbers of shared localities among species or the numbers of shared species among localities (Olivero, Real & Márquez 2011; Baselga & Orme 2012; Olivero, Márquez & Real 2013). Such measures are not otherwise spatially explicit, that is, they do not take into account the proximity (whether spatial or environmental) between species occurrence sites. Consequently, the distributions of species recorded at adjacent (even interspersed) and/or highly similar sites are considered as different as those of species occurring at opposite ends of the geographical or environmental space (Barbosa *et al.* 2012) (Fig. 1). This amplifies the effects of false absences and small spatial errors in the georeferencing of species occurrences, and the scale dependence of distributional relationships, precluding the identification of species associations except at broad spatial scales. These problems can particularly affect species with restricted

*Correspondence author. E-mail: barbosa@uevora.pt

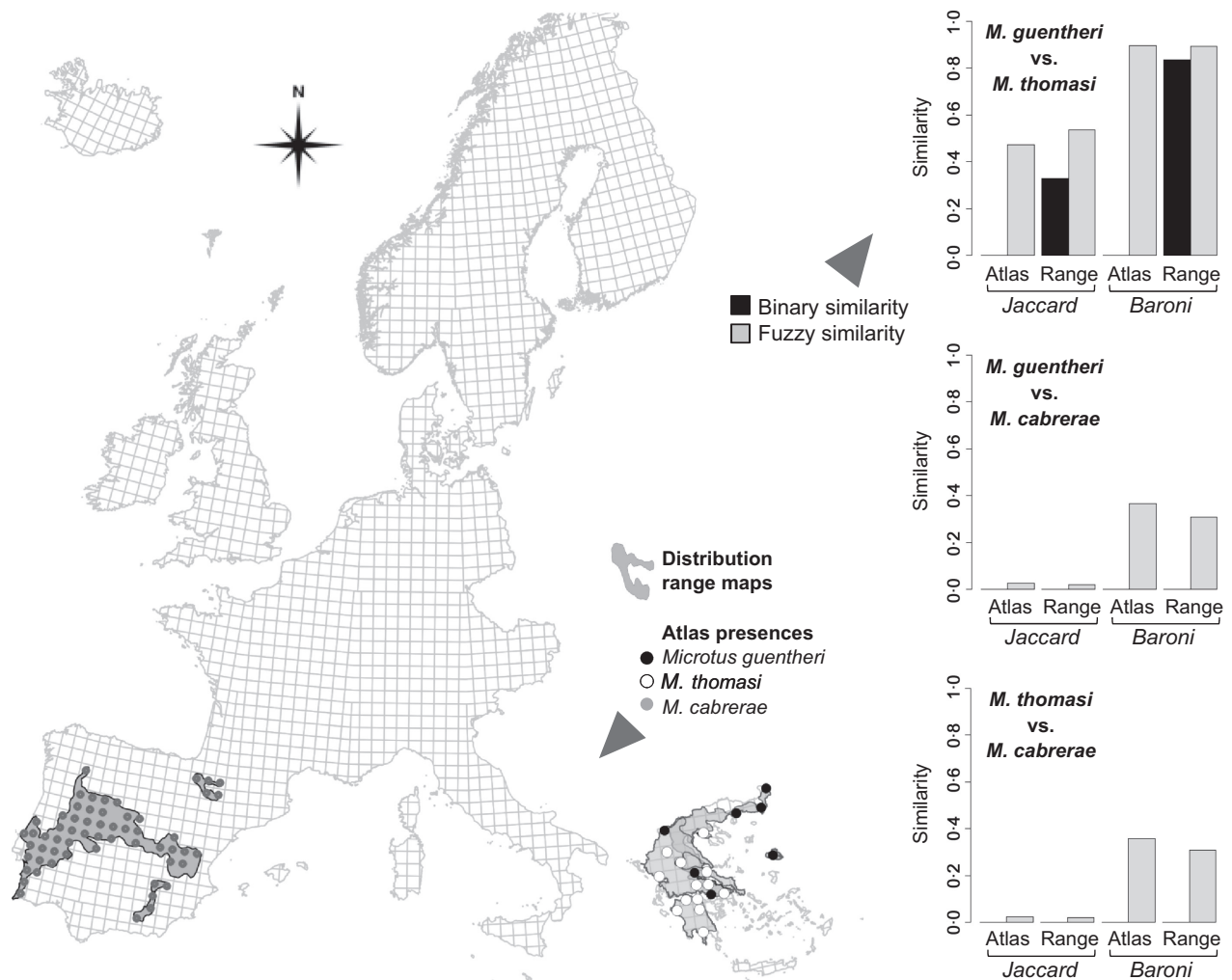


Fig. 1. Left: distributions of three vole (*Microtus*) species in western Europe according to atlas (Mitchell-Jones *et al.* 1999) and range map data (IUCN 2010). Right: pairwise similarities among these distributions using binary (based on presence/absence) and fuzzy (based on distance interpolation of presences) versions of the *Jaccard* and *Baroni* similarity indices. Note that, with atlas data, binary similarity is zero in all cases, while fuzzy similarity detects that *M. guentheri* and *M. thomasi* are more similar to each other (even without any overlapping presences) than to *M. cabrerai*.

distributions, as erroneous or missing distribution records can involve significant parts of their range (Barbosa *et al.* 2012).

Species distributions are inherently complex, and survey methods are always fallible. Species occurrence data therefore have intrinsic, unavoidable errors (Rocchini *et al.* 2011). For example, false absences often arise from insufficient sampling, while range map filling introduces false presences (Barbosa *et al.* 2012). It could therefore be more appropriate to treat distribution data as fuzzy, with localities classified with a degree of uncertainty on whether a species occurs or not, rather than putative presences and absences (Rocchini 2010; Rocchini *et al.* 2011; Duff, Bell & York 2014).

Fuzzy logic has recently been applied to smooth the limits between chorotypes (Olivero, Real & Márquez 2011) and biotic regions (Olivero, Márquez & Real 2013), but these are still primarily defined based on categorical presence/absence data on categorical sampling units. Thus, species recorded at adjacent but not strictly coincident localities (e.g. distribution atlas data for *Microtus thomasi* and *M. guentheri*, Fig. 1) are considered to be as dissimilar from each other as from a species

occurring thousands of kilometres away (e.g. *M. cabrerai*, Fig. 1). However, localities adjacent to (or scattered among) recorded presences, particularly if they are within similar environments and not separated by physical barriers, are often false absences resulting from insufficient sampling or other methodological constraints (e.g., Rocchini *et al.* 2011), so they should be attributed a degree of uncertainty about species presence. Another approach has been to compare species distributions or assemblages based on modelled habitat suitability values (Sillero *et al.* 2009; Albouy *et al.* 2012). However, comparisons are then still based on binary similarity indices after a binarization of suitability, thus discarding relevant quantitative information and introducing abrupt arbitrary thresholds.

Similarity indices that account for fuzziness of location, such as the fuzzy numerical comparison (Visser & de Nijs 2006) or the improved fuzzy kappa (Hagen-Zanker 2009), go beyond site-by-site comparison by giving partial credit to neighbouring sites. They thus introduce tolerance for small spatial differences in species occurrences and have been proposed for improving future analyses (Barbosa & Real 2012; Barbosa *et al.* 2012).

However, these indices are computationally intensive and require the use of complex spatial data structures such as raster maps, rather than simple occurrence data tables. Raster maps also imply that the spatial units under analysis are square pixels, precluding the use of other, sometimes more appropriate divisions such as equal-area units, provinces or watersheds (Carmona *et al.* 1999; Márquez, Real & Vargas 2001). Furthermore, considerable development would be necessary before these indices could be routinely used for such analyses, namely in optimizing their computation for multiple species and in determining their levels of significance (Barbosa *et al.* 2012). Likewise, indices of niche overlap (Warren, Glor & Turelli 2008; Broennimann *et al.* 2012) might be applicable to comparing spatial occurrence patterns, but their use is not established in species distributional and regional clustering, and their significance thresholds entail computer-intensive randomized simulations.

This study proposes an alternative, simple protocol for analysing fuzzy similarity between species occurrence patterns and between regional or temporal species compositions, based on fuzzy versions of the binary similarity indices commonly used in ecology, whose significance thresholds and other properties are therefore already known. The fuzzy indices are directly applicable to fuzzy (continuous) occurrence values, which can be obtained, for example, with distribution modelling, kernel smoothing or spatial interpolation of occurrence localities. To be used within a fuzzy logic framework, fuzzy occurrence must be bounded between 0 and 1 and represent the degree of membership of each locality to the species occurrence area, or of each species to a regional species pool (Zadeh 1965; Olivero, Real & Márquez 2011; Olivero, Márquez & Real 2013).

Calculating fuzzy similarity

Jaccard's (1901) index is one of the most widely used similarity indices in ecology, for example for detecting species distributional associations (Real & Vargas 1996; Sillero *et al.* 2009; Barbosa *et al.* 2012), comparing regional species compositions (Chao *et al.* 2004; Anderson, Bolton & Stegen 2009; Engen, Grøtan & Saether 2011), calculating beta diversity patterns (Anderson *et al.* 2011) or assessing temporal species turnover in climate change research (Albouy *et al.* 2012). Sørensen's (1948) index is widely used for assessing compositional similarity between sites or communities and beta diversity patterns in space or time (Anderson *et al.* 2011; Baselga, Gómez-Rodríguez & Lobo 2012; Baselga & Orme 2012; Dobrovolski *et al.* 2012). Simpson's (1960) similarity is the proportion, in the smaller of two samples, of taxa/regions common to both, thereby minimizing the effect of sample size discrepancies. Baroni-Urbani & Buser's (1976) index (hereafter *Baroni* for short) is broadly used in species distributional clustering and biotic regionalization (Márquez *et al.* 1997; Real, Olivero & Vargas 2008; Olivero, Real & Márquez 2011; Moya, Saucède & Manjón-Cabeza 2012; Olivero, Márquez & Real 2013). It accounts for both shared presences and shared absences, but gives greater weight to presences.

All these similarity indices vary between zero (no distributional or compositional overlap) and one (identical distributions or compositions). Tables of significant values per sample size are available (e.g. Baroni-Urbani & Buser 1976; Real 1999), so these indices can be used for identifying significant associations (Olivero, Real & Márquez 2011; Olivero, Márquez & Real 2013). All these indices are calculated on two or more of the following terms: *A* and *B* (the numbers of species/localities in each sample), *C* (the number of shared species/localities) and *D* (the number of species/localities missing from both samples). All these terms have direct correspondence with Boolean logic expressions and can thus be translated into their fuzzy logical equivalents to assess fuzzy binary similarity (Table 1).

Availability and functionality

This methodology is implemented in a free and open-source software package, `fuzzySim`, which works under the R programming environment (R Core Team 2014). The package, including some sample data (Fontaneto *et al.* 2012), is available on the public platform R-Forge (<http://fuzzysim.r-forge.r-project.org>), together with a reference manual and a step-by-step tutorial on its installation and usage. Most functionalities of `fuzzySim` (Fig. 2) are also being implemented as a graphical user interface extension for QGIS (QGIS Development Team 2014), which is also free and open-source.

The `fuzzySim` package allows a variety of methods for converting (multiple) species presence/absence data into continuous, fuzzy surfaces, including inverse distance to presence raised to any power (function `distPres`), trend surface analysis of any given degree (function `multTSA`) and generalized linear models based on presence-absence (function `multGLM`). The former two methods can be useful for purposes other than comparing species distributions and assemblages, for example for defining putative geographical ranges (Takahashi *et al.* 2014) or for delimiting the geographical background in species distribution modelling (Acevedo *et al.* 2012). Besides several methods for selecting predictor variables (including information criteria and false discovery rate), `multGLM` includes an option to convert probability to prevalence-independent favourability values, which have proven appropriate for use within a fuzzy logic framework (Real, Barbosa & Vargas 2006; Acevedo & Real 2012). The package also allows using other continuous distribution data that users can obtain elsewhere, as long as they are bounded between 0 and 1, directly comparable among species

Table 1. Correspondence between the terms in the formulas of binary similarity indices for a given pair of species (sp1 and sp2) and their equivalent expressions in classical and fuzzy set theory (Zadeh 1965)

Term	Boolean logic	Classical sets	Fuzzy sets
A	sp1	sp1	sum(sp1)
B	sp2	sp2	sum(sp2)
C	sp1 AND sp2	sp1 ∩ sp2	sum(minimum(sp1, sp2))
D	NOT sp1 AND NOT sp2	complement (sp1 ∪ sp2)	sum(1 - maximum (sp1, sp2))

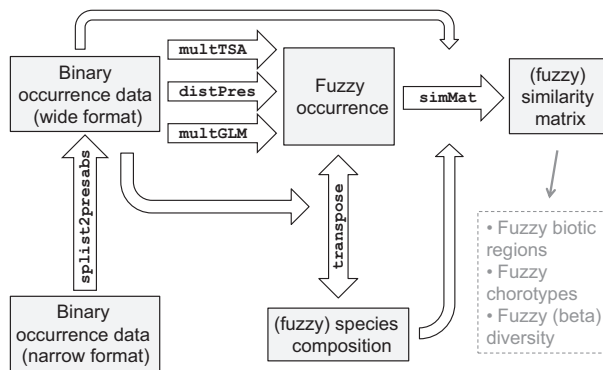


Fig. 2. Basic workflow of the `fuzzySim` package. Text inside arrows represents function names. Text in grey represents operations to do outside `fuzzySim`.

and interpretable as fuzzy membership values (Zadeh 1965; Real, Barbosa & Vargas 2006; Barbosa & Real 2012).

Species occurrence data, whether binary or fuzzy, can be transposed (so that regions go in columns and species in rows) for comparing regional species compositions (Olivero, Márquez & Real 2013). Pairwise similarity matrices between either species distributions or regions' compositions can be calculated with function `simMat` based on fuzzy versions of the *Jaccard*, *Sorensen*, *Simpson* and *Baroni* indices. These metrics were chosen for their simplicity, long-time widespread use and known significance thresholds for meaningful clustering, as well as their comparability to traditional indices. Nonetheless, additional similarity indices can be implemented as necessary, such as fuzzy versions of other measures of agreement between binary variables (e.g. simple matching coefficient, Cohen's kappa, true skill statistic) and of binary correlation indices (Phi, Matthews, Yule), as well as standard correlation coefficients (Pearson, Spearman, Kendall) and measures of niche overlap (Warren, Glor & Turelli 2008; Broennimann *et al.* 2012).

The fuzzy similarity matrices produced with `fuzzySim` can then be plotted, compared, classified, clustered and converted into dendrograms depicting the fuzzy relationships between species distributions or between regional species compositions. R code for all these operations is provided in the tutorials available from the package homepage (<http://fuzzysim.r-forge.r-project.org>). The fuzzy similarity matrices can also be entered in the `RMACOQUI` package (Olivero, Real & Márquez 2011) for a systematic analysis of chorotypes or biotic regions. The fuzzy versions of binary similarity indices can also be integrated within other software packages that currently compute these indices, such as `vegan` (Oksanen *et al.* 2013) or `betapart` (Baselga & Orme 2012).

Example analyses

To allow direct comparison with traditional methods, I used previously analysed data (Barbosa *et al.* 2012) on the occurrence of 156 terrestrial mammal species in Western Europe, based on both a distribution atlas (Mitchell-Jones *et al.* 1999) and a set of range maps (IUCN 2010), under a 50 × 50 km

UTM grid containing 2118 cells (Sastre, Roca & Lobo 2009). These data were previously used for comparing distributional relationships inferred from distribution atlas *versus* range maps, to assess the effects of data type on the results of such analyses. Although there was a good general agreement between the relationships obtained from the two data sources, there were some visible differences, most notably for small-range species (Barbosa *et al.* 2012).

The procedure is here illustrated with two fuzzy versions of species occurrence data: inverse squared distance interpolation of presences (Shepard 1968; Takahashi *et al.* 2014) and environmental favourability for presence (Real, Barbosa & Vargas 2006) based on information criterion selection of WorldClim bioclimatic variables (Hijmans *et al.* 2005). Note that these are simple examples; fuzzy distribution data should be obtained based on knowledge on what best reflects the distributions of the target species. Matrices of fuzzy similarity between species distributions were then calculated with the fuzzy similarity indices. For comparison, binary vs. fuzzy similarity was plotted for two similarity indices, one that considers only shared presences (*Jaccard*) and another that considers also shared absences (*Baroni*).

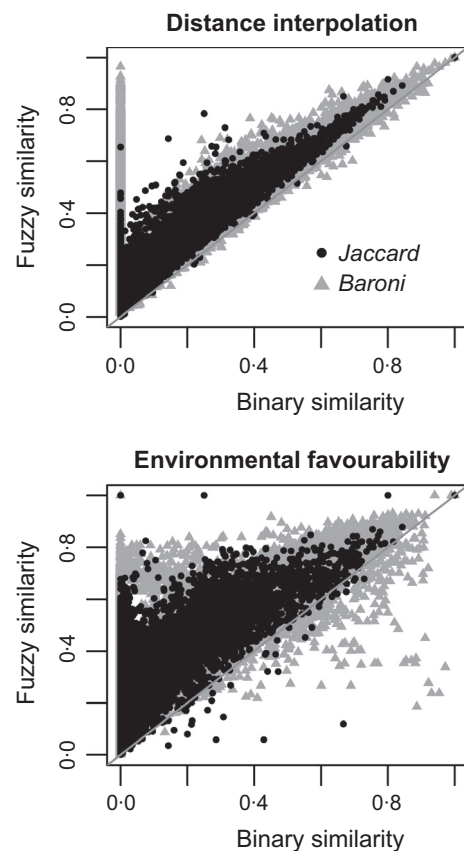


Fig. 3. Pairwise distributional similarities among 156 species in the European mammal atlas (Mitchell-Jones *et al.* 1999) when using binary vs. fuzzy versions of the *Jaccard* and *Baroni* similarity indices. Fuzzy distribution data were obtained with inverse squared distance interpolation of presences (top) and with environmental favourability based on informative bioclimatic variables (bottom).

As expected, fuzzy similarity was generally higher than binary similarity (Fig. 3), as similarity among species distributions is more than the strict coincidence of their recorded localities. Also, fuzzy similarity matrices obtained from atlas and from range map data were more similar (Mantel rank correlation tests, $\rho = 0.98$ for Jaccard and 0.97 for Baroni) than with binary similarity (Barbosa *et al.* 2012), indicating that fuzzy similarity minimizes the effects of data type on the results of chorological analyses. More notably, the fuzzy approach solved the problem posed by small-range species, for which slight differences between data sets could mean their coincidence or not in substantial parts of their distribution areas. Chorotypes, biogeographic regions and beta diversity patterns defined with fuzzy similarity indices are thus more likely to be robust to disparities, errors or gaps in species occurrence data, even for narrowly distributed species, which usually face higher conservation concern and where slight inaccuracies can affect substantial parts of their recorded range.

Acknowledgements

This work was supported by Fundação para a Ciência e a Tecnologia through 'FCT Investigator' contract IF/00266/2013 and exploratory project CP1168/CT0001. The IUCN and Societas Europaea Mammalogica (particularly A.J. Mitchell-Jones) kindly provided the distribution data for analysis. Raimundo Real, Jesús Olivero and Ana Luz Márquez provided valuable early discussions on chorotypes and fuzzy logic. Alba Estrada, Raquel Garcia, Rich Grenyer and Duccio Rocchini provided valuable feedback.

Data accessibility

Mammal atlas maps are available from the European Mammal Society (<http://www.european-mammals.org>). Range maps are available at the IUCN website (<http://www.iucnredlist.org>). Data used in the `fuzzysim` tutorial are included in the supplementary material of Fontaneto *et al.* (2012) and in the `r` package (<http://fuzzysim.r-forge.r-project.org>).

References

- Acevedo, P. & Real, R. (2012) Favourability: concept, distinctive characteristics and potential usefulness. *Naturwissenschaften*, **99**, 515–522.
- Acevedo, P., Jiménez-Valverde, A., Lobo, J.M. & Real, R. (2012) Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, **39**, 1383–1390.
- Albouy, C., Guilhaumon, F., Araújo, M.B., Mouillot, D. & Leprieux, F. (2012) Combining projected changes in species richness and composition reveals climate change impacts on coastal Mediterranean fish assemblages. *Global Change Biology*, **18**, 2995–3003.
- Anderson, R.J., Bolton, J.J. & Stegenga, H. (2009) Using the biogeographical distribution and diversity of seaweed species to test the efficacy of marine protected areas in the warm-temperate Agulhas Marine Province, South Africa. *Diversity and Distributions*, **15**, 1017–1027.
- Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L. *et al.* (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, **14**, 19–28.
- Barbosa, A.M. & Real, R. (2012) Applying fuzzy logic to comparative distribution modelling: a case study with two sympatric amphibians. *The Scientific World Journal*, **2012**, Article ID 428206.
- Barbosa, A.M., Estrada, A., Márquez, A.L., Purvis, A. & Orme, C.D.L. (2012) Atlas versus range maps: robustness of chorological relationships to distribution data types in European mammals. *Journal of Biogeography*, **39**, 1391–1400.
- Baroni Urbani, C., Ruffo, S. & Vigna Taglianti, A. (1978) Materiali per una biogeografia italiana fondata su alcuni generi di coleotteri Cicindelidi, Carabidi e Crisomelidi. *Memorie della Società Entomologica Italiana*, **56**, 35–92.
- Baroni-Urbani, C. & Buser, M.W. (1976) Similarity of binary data. *Systematic Zoology*, **25**, 251.
- Baselga, A., Gómez-Rodríguez, C. & Lobo, J.M. (2012) Historical legacies in world amphibian diversity revealed by the turnover and nestedness components of Beta diversity. *PLoS ONE*, **7**, e32341.
- Baselga, A. & Orme, C.D.L. (2012) betapart: an R package for the study of beta diversity. *Methods in Ecology and Evolution*, **3**, 808–812.
- Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G. *et al.* (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.
- Carmona, J.A., Doadrio, I., Márquez, A.L., Real, R., Huguely, B. & Vargas, J.M. (1999) Distribution patterns of indigenous freshwater fishes in the Tagus River basin, Spain. *Environmental Biology of Fishes*, **54**, 371–387.
- Chao, A., Chazdon, R.L., Colwell, R.K. & Shen, T.-J. (2004) A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters*, **8**, 148–159.
- Dobrovolski, R., Melo, A.S., Cassemiro, F.A.S. & Diniz-Filho, J.A.F. (2012) Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, **21**, 191–197.
- Duff, T.J., Bell, T.L. & York, A. (2014) Recognising fuzzy vegetation pattern: the spatial prediction of floristically defined fuzzy communities using species distribution modelling methods. *Journal of Vegetation Science*, **25**, 323–337.
- Engen, S., Grotan, V. & Saether, B.-E. (2011) Estimating similarity of communities: a parametric approach to spatio-temporal analysis of species diversity. *Ecography*, **34**, 220–231.
- Fontaneto, D., Barbosa, A.M., Segers, H. & Pautasso, M. (2012) The 'rotiferologist' effect and other global correlates of species richness in monogonont rotifers. *Ecography*, **35**, 174–182.
- Hagen-Zanker, A. (2009) An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation. *International Journal of Geographical Information Science*, **23**, 61–73.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Holt, B.G., Lessard, J.-P., Borregaard, M.K., Fritz, S.A., Araújo, M.B., Dimitrov, D. *et al.* (2013) An update of Wallace's zoogeographic regions of the world. *Science*, **339**, 74–78.
- IUCN. (2010) IUCN Red List of Threatened Species, version 2010.4. Retrieved from <http://www.iucnredlist.org/technical-documents/spatial-data>
- Jaccard, P. (1901) Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin del la Société Vaudoise des Sciences Naturelles*, **37**, 547–579.
- Márquez, A.L., Real, R. & Vargas, J.M. (2001) Methods for comparison of biotic regionalizations: the case of pteridophytes in the Iberian Peninsula. *Ecography*, **24**, 659–670.
- Márquez, A.L., Real, R., Vargas, J.M. & Salvo, A.E. (1997) On identifying common distribution patterns and their causal factors: a probabilistic method applied to pteridophytes in the Iberian Peninsula. *Journal of Biogeography*, **24**, 613–631.
- Mitchell-Jones, A.J., Amori, G., Bogdanowicz, W., Krystufek, B., Reijnders, P.J.H., Spitzenberger, F. *et al.* (1999). *The Atlas of European Mammals*. Academic Press, London.
- Moya, F., Saucède, T. & Manjón-Cabeza, M.E. (2012) Environmental control on the structure of echinoid assemblages in the Bellingshausen Sea (Antarctica). *Polar Biology*, **35**, 1343–1357.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B. *et al.* (2013). *vegan: Community Ecology Package*.
- Olivero, J., Márquez, A.L. & Real, R. (2013) Integrating fuzzy logic and statistics to improve the reliable delimitation of biogeographic regions and transition zones. *Systematic Biology*, **62**, 1–21.
- Olivero, J., Real, R. & Márquez, A.L. (2011) Fuzzy chorotypes as a conceptual tool to improve insight into biogeographic patterns. *Systematic Biology*, **60**, 645–660.
- QGIS Development Team. (2014). QGIS Geographic Information System. OSGeo - Open Source Geospatial Foundation. Available at: <http://www.qgis.org>.
- R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.r-project.org>.
- Real, R. (1999) Tables of significant values of Jaccard's index of similarity. *Miscellanea Zoologica*, **22**, 29–40.
- Real, R., Barbosa, A.M. & Vargas, J.M. (2006) Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, **13**, 237–245.

- Real, R., Olivero, J. & Vargas, J.M. (2008) Using chorotypes to deconstruct biogeographical and biodiversity patterns: the case of breeding waterbirds in Europe. *Global Ecology and Biogeography*, **17**, 735–746.
- Real, R. & Vargas, J.M. (1996) The probabilistic basis of Jaccard's index of similarity. *Systematic Biology*, **45**, 380–385.
- Rocchini, D. (2010) While Boolean sets non-gently rip: a theoretical framework on fuzzy sets for mapping landscape patterns. *Ecological Complexity*, **7**, 125–129.
- 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.
- Sastre, P., Roca, P. & Lobo, J.M. (2009) A Geoplatform for improving accessibility to environmental cartography. *Journal of Biogeography*, **36**, 568.
- Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM national conference* - pp. 517–524. ACM Press, New York, NY, USA.
- Sillero, N., Brito, J.C., Skidmore, A.K. & Toxopeus, A.G. (2009) Biogeographical patterns derived from remote sensing variables: the amphibians and reptiles of the Iberian Peninsula. *Amphibia-Reptilia*, **30**, 185–206.
- Simpson, G.G. (1960) Notes on the measurement of faunal resemblance. *American Journal of Science*, **258A**, 300–311.
- Sørensen, T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biologiske Skrifter/Kongelige Danske Videnskabernes Selskab*, **5**, 1–34.
- Takahashi, M.K., Eastman, J.M., Griffin, D.A., Baumsteiger, J., Parris, M.J. & Storfer, A. (2014) A stable niche assumption-free test of ecological divergence. *Molecular Phylogenetics and Evolution*, **76**, 211–226.
- Visser, H. & de Nijs, T. (2006) The map comparison kit. *Environmental Modelling & Software*, **21**, 346–358.
- Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.
- Zadeh, L.A. (1965) Fuzzy sets. *Information and Control*, **8**, 338–353.

Received 5 February 2015; accepted 10 March 2015

Handling Editor: Robert B. O'Hara