Comparing without gridding species distributions: calculating pairwise intersection, union and similarity between range maps in R

A. Marcia Barbosa*[†][1] and Alba Estrada[2]

[1]CIBIO/InBIO, Universidade de Évora (CIBIO-UE) – Portugal
[2]Dpto. Biología Animal, Universidad de Málaga (UMA) – Espagne

Résumé

The comparison of species' distributions and the definition of typical or shared distribution patterns (chorotypes) have important applications in ecology, biogeography and conservation [1]. They are usually based on indices of pair-wise binary similarity between species presence/absence data on a grid of operative geographic units (OGUs). These OGUs can be administrative (e.g. provinces, counties), natural (e.g. watersheds) or arbitrary (e.g. rectangular grids), all of which can have various shapes and sizes.
A previous study compared the effects of using different OGUs on the definition of chorotypes, in order to identify the type of gridding or territorial partition that produced the least fuzzy and most ecologically interpretable patterns. The results showed that the smallest OGUs produced the best results [2]. In fact, any gridding or partitioning involves some distortion of the data, since it artificially changes the borders of the species' known distribution ranges, usually overestimating them (Figure 1). For example, two species with relatively close distribution areas will be considered to overlap at least partially if both intersect a given grid cell, even if there is really no overlap and those species are physically separated by a geographical barrier that crosses that grid cell. Gridding is especially problematic when the study area includes islands, where, for example, a single grid cell may include parts of different islands or even nearby continents, and a small island can be artificially divided into more than one grid cell (Figure 1). This problem can be minimized by treating islands differently when defining the OGUs, but this can be very laborious, poorly reproducible, involves arbitrary decisions, and does not solve the effect of the grid on the rest of the study area. The most objective solution is not to grid the study area into OGUs, but rather to calculate distribution similarities directly on the distribution range maps.

We propose an adaptation of the binary similarity indices commonly used in the comparison of species distributions, so that they use the size of the intersections and unions between range maps rather than the numbers of shared and non-shared presences on a grid. This allows the analysis of distribution range maps directly without gridding them into OGUs, therefore avoiding the problems associated with islands and barriers and with the choice of the best OGU. It thus eliminates some methodological artefacts in the definition of chorotypes or other analyses that require comparing species distributions based on range maps.

*Intervenant
†Auteur correspondant: barbosa@uevora.pt

This method is implemented within the *fuzzySim* R package [3], namely in functions *pairwiseRangemaps* and *rangemapSim*, which calculate pair-wise similarity matrices directly from range map polygons (e.g., in *ESRI shapefile* format) [4]. For large map sets, this process (namely the map intersections) can be computationally intensive and cumulatively slow. However, there is the option for parallel computing and for calculating the matrix in chunks, cleaning the memory between each chunk and the next, and assembling them all in the end.