Effects of Sampling Effort on Estimates of the Mean of Range Size Distributions

Megan Ruffley

The range size distribution (RSD) for a given region describes the frequency of species exhibiting geographic ranges of different sizes. Descriptions of spatial variation in RSD are major elements of current attempts to explain spatial patterns of diversity and identify areas of conservation concern. Thus, it is critical to understand potential bias in estimates of RSD properties. Here, we hypothesize that bias in estimates of the mean of RSD is determined by sampling effort according to the following model of the probability of not discovering a species (P.m): $P.m. = (1-d)^{\sum_{i=1}^{AOO} Ci}$ where d is detectability (the probability of detecting the occurrence of the species in a spatial grid cell i, given the presence of the species and a unit of sampling effort), Ci is sampling effort in grid cell i, AOO is geographic range size measured as area of occupancy, and grid cells i=1 through i=AOO constitute the geographic range of the species. Based on equation 1 we predict that bias in estimates of the mean of RSD for any given region is negatively related to mean sampling effort across the region, and positively related to the spatial aggregation of sampling effort across the region, because in poorly sampled regions narrowly distributed species are more likely to not be discovered than widely distributed species. We tested these predictions in the context of the current belief that mean RSD of Andean plants is smaller than that of Amazonian plants. We simulated the geographic distributions of 100,000 virtual species across the Neotropics by implementing the "stepping stone" model described, and simulated the discovery of each of the virtual species occurring in 100 × 100 km grid cells in the Andes and Amazonia according to sampling effort and equation 1 with detectability = 0.1. For each 100 × 100 km grid cell we calculated bias in the mean of RSD as the difference between true and observed values of the mean of RSD, the latter being the mean AOO of discovered species (excluding undiscovered species). The results supported predictions about the effect of sampling effort on bias in estimates of the mean of RSD. The difference in sampling effort between the Andes and Amazonia can significantly affect bias in estimates of mean RSD. Thus, current descriptions of geographic variation in RSD and the density of narrowly distributed plant species across the Neotropics may be more fiction than substance, and should be regarded as highly tentative at best. Estimates of spatial variation in mean RSD should account for spatial variation in sampling effort.