Spatial and temporal variation in species-area relationships in the Fynbos biological hotspot

Lisa L. Manne, Paul H. Williams, Guy F. Midgley, Wilfried Thuiller, Tony Rebelo and Lee Hannah

Species-area relations (SARs) are among the few recognized general patterns of ecology, are empirical relations giving the number of species found within an area of a given size and were initially formulated for island environments. The use of SARs has been extended to mainland environments, and to give baseline estimates of extinction rates attending habitat loss. Using current species distributions based on atlas data, we examined the spatial variation of rates of species accumulation and species-area curves for Proteaceae species for all one-minute by one-minute areas within the Cape Floristic Region, South Africa. We compared SARs for current distributions to those generated from modeled future Protea distributions following climate change. Within one biome and for two different scales, there exists a very large spatial variation in turnover rates for current Proteaceae distributions, and we show that these rates will not remain constant as climate warming progresses. As climate changes in coming years, some areas will gain species due to migration, as other areas lose species, and still other areas maintain current rates of species accumulation/turnover. Both current and future distributions show highly variable rates of species accumulation across the landscape. This means that an average species-area relationship will hide a very large interval of variation among SARs, for both current and future Proteaceae distributions. The naive use of species-area relations to estimate species extinctions following loss of current habitat, or loss of future climatically-suitable area is likely to result in erroneous predictions.

Species-area relations (SARs) are a community-level property that arises from the combination of species’ distributions, and there is an active literature working to understand and reproduce this linkage (He et al. 2002). Species-area relations (SARs) have been used to characterize the number of community members in island, mainland, nested subsets of areas (see review in Begon et al. 1996). SARs are typically fitted as power functions ($S \propto A^z$, but see Tjørve 2003), and others have argued about how a SAR behaves for very small and very large areas (Lomolino 2000, 2001, 2002, Williamson et al. 2001, Williamson 2003). The exponent $z$, a logarithmic rate of species accumulation with increasing area (henceforth “rate of species accumulation”), generally takes values from 0.15 to 0.40 (Hubbell 2001, but see Williamson 2003). Mainland area SARs generally tend to have lower $z$-values than those for oceanic island areas, and accumulation of habitat islands tends to yield the highest $z$-values, which can approach 1.0 for “inter-provincial” areas (Rosenzweig 1995).

SARs have been extended for use in conservation planning, where planners ask how small can an area be, and still support a target number of species (Pimm et al. 1995, Brooks et al. 1997, Thomas et al. 2004)? Whether this use is justified has been questioned (Kinzig et al. 2000), particularly since whether SARs depend on spatial scale (Hubbell 2001, Crawley et al. 2001, He et al. 2003) has not been adequately
examined. In addition, there has been little attention to how much spatial variation in rate of species accumulation exists across the landscape (but see Adler et al. 2003, Adler 2004), or how a SAR might change as climate changes. Understanding these issues is crucial if researchers intend to continue to use SARs as predictive models of extinction.

The Cape Floristic Region (CFR) has an exceptionally rich flora, characterized by island-like patterns of high species: genus ratios (Cowling 1992). We examine 316 species of the well-known Proteaceae (Rebelo 1995). SARs have previously been fitted for the western and eastern CFR (Cowling et al. 2002) using distribution data compiled from several sources. Here, we examine the pattern of SAR variation across the biome and ask, 1) how much variation exists in SARs within the biome? Pursuant to that, 2) does the typically-assumed rate of species accumulation ($z = 0.25$) hold for the majority of the region? 3) Do SARs for current species distributions hold fast under assumptions of climate change? These questions address the consistency and utility of this ecological pattern, lend insight into how the biota of the CFR is spatially structured, and inform how this spatial structure might change under climate change. These questions are pertinent as scientists estimate that this biome’s area will suffer a reduction of 51–65% over the next 50 yr due to climate change (Midgley et al. 2002a).

If our answers to the above questions are 1) not a large amount, 2) yes, and 3) yes, then this relationship can be considered a viable conservation tool in the context of ongoing climatic change. We find for this taxon at different spatial scales that different areas within the Fynbos biome show widely varying rates of species accumulation; this variation is so great that a stereotypical SAR with the often-assumed $z$-value of 0.25 holds for only a quarter of the area of the biome. This result leads us to hypothesize that if climate change were to change current biomes into different biomes, then species-area curves in those areas might also change accordingly. We test this idea, and find that future rates of species accumulation, as indexed by $z$-values from SARs fitted from modeled future distributions, will increase in some regions (more turnover), will remain the same in some areas, and will decrease in some areas (due to increased homogenization) as species respond individually to climate warming.

Methods

Current species-area relations in the CFR

We used distribution data for 316 South African Proteaceae species (Rebelo 1995) at two resolutions: a one-minute grid, and a 1/20th degree grid, to construct species-area relationships across the western Cape region. The Proteaceae are particularly apt model organisms for this question, as they are so well-known – all grid cells have been surveyed – and many of these species are restricted to the CFR, an important consideration when constructing species-area curves (Green et al. 2003).

Even though the Proteaceae are well-known, it may be that some individuals were not detected, and thus not represented in the atlas data that we analyzed. Due to sensitivity of SARs to undetected presences, or variation in sampling effort (Cam et al. 2002), we used Proteaceae distributions modeled from relationships between census data and variables critical for plant growth (details of model-fitting below).

Logistic regression has been one of the more popular techniques for modeling species distributions (Manel et al. 1999, Collingham et al. 2000, Bailey et al. 2002, Berg et al. 2004, Eyre et al. 2005, Carter et al. 2006). However, in studies that compare multiple techniques, neural network models and/or generalized additive models (GAMs, Yee et al. 1991) outperform other techniques (Midgley et al. 2003, Segurado et al. 2004, Araujo et al. 2005) or give results that are in line with results from other models (Pearson et al. 2006). A notable exception in this literature is Manel et al. (1999), who found that logistic regression outperformed neural network models for one bird species. Since Thuiller et al. (2003) further showed that GAMs are robust to variation in scale, using GAMs seems reasonable for this analysis.

We fitted GAMs based on five bioclimatic variables critical for plant growth (mean minimum temperature of the coldest month, annual sum of daily temperatures exceeding 18°C, annual potential evaporation, summer soil moisture days, and winter soil moisture days) (Midgley et al. 2002a, 2003) to distribution data from the Protea Atlas Project (PAP; <http://protea.worldonline.co.za>). We used the Agricultural Atlas climate surface dataset (Schulze 1997, Schulze et al. 1999) to represent current climate. We did not include other range determinants as predictors, such as historical factors or biotic interactions, since the current range represents the realized niche of the species, which indirectly includes biotic interactions and historic factors. At the relatively large scale of the Cape Floristic Region, current climate predominates as a range determinant (Thuiller et al. 2004a, c). The Cape flora is relatively new, with high diversification rates in recent years due to high topographic complexity and climatic stability (Cowling et al. 2002, 2005), so that historical factors should be less important for determining range.
In the GAM construction, we randomly partitioned the data into 70% and 30% partitions for model formation and testing, respectively (cf. Fielding et al. 1997, Guisan et al. 2005). This partitioning provided sufficient data to make reliable calibration and enough data to do a proper evaluation. We evaluated each species’ model using the values obtained for the area under the curve (AUC) of a receiver operating characteristic (ROC) plot of sensitivity against (1-specificity) (Swets 1988).

The resulting species distributions are raster distributions in 17,673 one-minute (1.5 × 1.8 km) grid cells covering the ca 47,700 km² Western Cape region. To calculate species-area relationships, we chose a starting cell, noting the number of species in the cell and the area of the cell. Then we examined cells around the starting cell in a radial fashion, accumulating area and species, until we had gathered data on 225 cells (a 15 × 15 grid within the sampling region). We wished to analyze rates of species accumulation with increasing area for relatively small areas within a possibly heterogeneous landscape. We treated edge (of map) cells in the following way. While we did the species accumulations for all cells, when edge cells were within a zone of species accumulation, then species and area calculations were only tallied for land-based grid cells; e.g. consider the case where the second concentric ring of grid cells around a starting grid cell includes 7 sea and 9 land cells. The incremental area for that second concentric ring of cells would thus be 9 cells.

Power functions are commonly used in fitting SARs (Scheiner 2003, Tjørve 2003), and have been shown to work for even very small scales (Green et al. 2004, Horner-Devine et al. 2004). Power-law SARs are expected to provide a very good fit to data, provided that species abundance distributions are clustered, and that the distribution of abundances conforms to Preston’s lognormal distribution (García Martín et al. 2006). Since for species-area calculations, there need be only one individual of a species present, for that species to be counted, García Martín et al. (2006) extended this rule of thumb to say that power-law formulations will work well if species ranges are clustered, and if the range size distribution conforms to a Preston lognormal, though these authors note that lognormal distributions do not always provide the best fit for many ecosystems (García Martín et al. 2006 and references therein). We checked for spatial autocorrelation (two-tailed test of Moran’s I value, calculated in Idrisi32) to verify that species’ distributions were clustered, then fitted power functions to log-transformed species and area data (Press et al. 1992, Williams 2000) and extracted the fitted z-value.

Models of future, climate, future distributions of Proteaceae species, and future species-area relationships

Using the known climatic associations from the species distribution models (GAMs) above, we modeled expected future distributions for individual Proteaceae species on a one-minute latitude by longitude grid at ten-year time intervals – 2000, 2010, 2020, 2030, 2040, 2050, which we term decadal time slices. Future (2050) climate predictions were produced by perturbing the current climatic data with anomalies derived from climatic simulations produced by the General Circulation Model HadCM2 (<http://cera-www.dkrz.de/IPCC_DDC/IS92a/Hadley-Centre/Readme.hadcm2>), using IS92a emissions assumptions for CO₂ equivalent greenhouse gas concentrations, and excluding sulphate cooling feedback. We assumed a linear change in each climatic variable between 2000 and 2050. This assumption is meant to reduce complexity in the formulation of the models, and in the interpretation of the results. We can imagine situations arising which involve complex temporal shifts in the climate regime, e.g. from any number of possible extreme events. Such complex temporal shifts would introduce additional complexity in interpreting results. In future work, it would be useful and informative to compare this linear assumption to possible outcomes from modeling various types of nonlinear climate events.

In order to predict species’ future ranges with a changing climate, some assumption about species’ dispersal ability must be made. Species’ dispersal ability usually is not explicitly accounted for when projecting future species distributions (but see Iverson et al. 1999, Williams et al. 2005, Broennimann et al. 2006, Midgley et al. 2006). Instead, either dispersal is assumed to be fully effective, so that ranges that have become newly suitable are invariably colonized (“unlimited dispersal” hypothesis), or dispersal is assumed to be zero, so that all individuals of the study are unable to shift to their new ranges (“no dispersal” hypothesis, Thomas et al. 2004, Thuiller et al. 2005). These two extremes encompass the range of possible migration rates, but neither of these approximations is satisfactory, because migration rate depends to a large extent on the capacity of each individual species to migrate, which itself is a composite of individuals’ various abilities. In our study, we follow Midgley et al. (2006) (see this paper for more details) to avoid unreliable future potential distributions.

We calculated range shifts after 2000 using four assumptions about dispersal limitations. First, we estimated migration rate per decadal time slice using the dispersal agent as a surrogate for migration
capability. We assumed dispersal distance to be a maximum of one grid cell per decadal time slice for ant- and rodent-dispersed species; this is likely to be an over-estimate, given the pattern of transport and hoarding of seeds by these species only a few meters from seed sources (Midgley et al. 2002b). We assumed a maximum of three cells per time slice for wind-dispersed species, which is at the high end of the range determined by empirical measurements for seed dispersed by secondary “tumbling” seed dispersal (Bond 1988). The latter rate corresponds to at least 4 km in 10 yr (i.e. per fire event), or an average of 400 m in one year, which may be considered long-distance dispersal (Cain et al. 2000), and compares with high rates inferred from the fossil record (Clark 1998). Second, for every species, if any pixel became suitable under climate change, but was geographically more distant than the maximum dispersal distance of the given species from the source pixel, the probability of dispersal to that pixel was set to 0. Third, we assumed that dispersal events occur on average every decade, depending on decadal fires to provide the only dispersal opportunities to these fire-adapted species—this fire return time is probably just below average for the CFR, but greater than the minimum required for seed set in Proteaceae (Bond et al. 1995). Fourth, we further assumed that bio-climatically unsuitable areas would not persist as viable range between time slices. The result was a set of time-slice models for each species for each of the years 2000, 2010, 2020, 2030, 2040, and 2050 according to all four assumptions.

If rates of species accumulation for future species distributions are different to those for current distributions, this result would have serious implications for using these models to predict species losses under climate change projections. Currently-known ecoregions, no matter how defined, are expected to change size and location following climate change (Sala et al. 2000, Ostendorf et al. 2001, Berry et al. 2002, Scott et al. 2002). Consequently, we used projected future distributions to calculate future SARs, and test whether SARs for the Proteaceae of South Africa will vary between 2000 and 2050.

Results

Current rates of species accumulation in the CFR

In order to justify using power-law functions to fit species-area relationships, we needed to verify that the Cape flora species distributions are clustered, and that the species show a lognormal distribution of range sizes. For the Cape flora, 312 of the 316 species show significant autocorrelation at a lag of one spatial step (Moran’s I significantly >0, king’s rule, p < 0.001 for these 312 species). Figure 1 shows the frequency distribution of Moran’s I values for the 316 species. The four non-significantly spatially autocorrelated species are shown with Moran’s I = 0 in this figure. The other bins along the x-axis are 0.20 units in width and are labelled with the midpoint of the bin; for example, 0.50 denotes the bin representing Moran’s I values in the [0.40, 0.60] range, where 0.60 is included in this bin, and 0.40 is included in the previous bin.

We fitted species-area relations within the CFR, with results shown in Fig. 2 (but see the relationship between z and species number in the Discussion), and mapped the z-value onto the grid cell where species accumulation began (Fig. 3). While we calculated SARs and z values for all cells, the results presented in Fig. 2 are only for central areas of the CFR, those areas that are 8 or more cells distant from the edge of the region. For a small number of these central areas in the western CFR (7%, Fig. 2a), the pattern of species increase with increasing area is upward accelerating, resembling exponential growth. For the remaining (central) areas, the typical form of SAR is realized in the data. Some of these areas exhibit step-function behavior (Fig. 2b, 15% of areas), continued slow growth of species number with accumulating area (non-asymptoting, although an asymptotic curve provides a good fit to the data, Fig. 2c,
27% of areas), or typical, asymptotic species-area relations (Fig. 2d, 51% of areas).

The fitted z-values, mapped in Fig. 3a, have a mean of 0.29 and median of 0.25. The z-values range in value from 0.06 to 0.91 (Fig. 3b); the thresholds to the lower and upper 5% tails occur at z \approx 0.116 and z \approx 0.60. We checked whether accumulating species and areas over differing numbers of cells would affect our inference, since, as the area of the accumulation region increases toward the total number of grid cells available (the entire region), the variation in calculated z-value must decrease. We found that SARs calculated for very small species accumulation areas, 49 cells (ca 132 km²), had z-values spanning the range 0.09–1.12 over the CFR; when this species accumulation area increases to larger areas, 225 grid cells (ca 605 km²), z-values span 0.03–0.91. (We note that z-values \( z > 1 \) are biologically not very informative, but can still result from the curve-fitting process.) This change of accumulation area size results in z-values which are statistically indistinguishable (two-tailed paired t-tests); z-values obtained from different sizes of accumulation area correlate very strongly (correlation coefficients of 0.97–0.98).

The results presented in Fig. 2 and 3 are those using the larger area of species accumulation. While the average z-value is near 0.25, this large variation in values belies textbook expectations that one z-value will work for all areas (Rosenzweig 1995). Focusing on those areas that have z-values near 0.25 (Fig. 3c) shows that such z-values are realized for only about a quarter of the land area considered here. These results call into question the utility of one (average) species-area relationship for use in conservation planning, for these taxa in this biome.

**Future rates of species accumulation in the CFR**

Even though current species distribution patterns are characterized by quite variable SARs, if these current SARs should remain fixed through time, we could predict how many species will remain following climate change. Using the projected future distributions of Proteaceae described above, we calculated future rates of species accumulation (mapped in Fig. 3d, as Fig. 3a). Future rates, like current rates, show a large degree of variation, with a very small proportion of fitted z-values exceeding 1 (Fig. 3e). The future z-values show a mean
of 0.42 and median of 0.33, and many areas within the region do retain similar species accumulation values to those of the year 2000. However, there are two notable differences between Fig. 3a and Fig. 3d: some areas in 2050 (Fig. 3d) are unshaded, reflecting the expectation of complete extinctions in those areas, and some areas in 2050 have quite different rates of species accumulation when compared to 2000 (Fig. 3a). To help visualize this comparison, we overlaid the species accumulation rates of 2000 versus those for 2050 (Fig. 3f), using a three-color scheme (Williams et al. 1998). White areas are where we did not do a comparison, as complete extinction is expected in those areas by 2050. Some areas show where rates of species accumulation will have decreased, e.g. much of the west coast of the Western Cape, while other areas show where rates will have increased by 2050, e.g. in the Agulhas Plain and along the Langeberg Mountains. Still other areas show little change in z-value by 2050. These results are consistent with analyses documenting areas needed to ensure dispersal corridors for the Proteaceae over the next 50 yr (Williams et al. 2005). The differences in turnover rate (as indexed by the fitted z-values) between 2000 and 2050 are statistically significant (two-tailed paired t-test, equal variances, p < 10^-5). Figure 4 shows the proportion of areas experiencing rate changes of different magnitudes.
changes in species accumulation rate, 24% of areas will show an increase in z-value equal in magnitude to the typically-assumed value of 0.25, and a further 29% of areas will show a substantial decrease in z-value.

Discussion

We have shown that rates of species accumulation across this landscape vary between 0.06 and 0.91, with only a quarter of the land area showing a typically-assumed logarithmic rate of species accumulation of 0.25. This result implies that using an average species-area relation to predict the number of extinctions following loss of habitat area may be inappropriate without a consideration of the extent of variation of local species accumulation rates within the broader area. Even though the mean z-value across the CFR is 0.29 (median: 0.25), because of this large spatial variance in z-value, a species-area relationship that would do well at predicting species extinctions following loss of habitat in one area of the CFR may, without consideration of local context, fail miserably in another area of the CFR.

Second, we found that species-area relations will not remain constant under expectations of biotic redistribution in the wake of climatic change; a crucial difference for the whole region is that the mean z-value over the region increases to $z = 0.42$ (with the median z-value increasing to $z = 0.33$). Even disregarding the spatial variation in species accumulation rates documented for current distributions, this latter result implies that the use of species-area relations to calculate species extinctions following loss of climatically-suitable area is likely to result in erroneous predictions.

Naturally, our understanding of these species’ current and (particularly) future geographic ranges may be incorrect in some cases, which would surely impact our results. However, we have worked to assure that our appreciation of current environmental affilia-
tions for these plants, and of how individual species will react under changing climate scenarios, is as accurate as possible. The modeled current and future distributions of Proteaceae species are as realistic as possible (Midgley et al. 2006), including limitations on dispersal imposed by dispersal mode (ant, rodent, wind). Estimates of how far any one plant species may be expected to move within a short time span are conservative, consistent with the idea that only a very few individuals will achieve long distance migration, and that population spread is not well-predicted by the mean dispersal distance (Clark et al. 2003). Species’ individualistic migrations in response to climate change might release other species from competitive exclusion, further complicating our understanding of future distributions, but testing for these eventualities is beyond the scope of this current paper.

The fitted SARs for the current distributions depend upon modeled species ranges. This aspect of the analysis might engender two criticisms. First, we have only explored one species distribution modelling method, which could potentially be biasing. However, in a comparison of several species distribution models for four species, Pearson et al. (2006) demonstrated that GAMs provides consensual predictions and projections compared to other techniques. The use of additional modeling techniques might therefore, change the results slightly, but not our inference. Second, the modeled ranges may be more dependent on aspects of habitat, rather than climatic factors. This type of bias would result in the species accumulation process accumulating the underlying habitats, so that the fitted z-values represent rate of accumulation of habitats. This phenomenon is not likely to be driving our results, as the majority of the Proteaceae species considered here have distributions that span multiple habitats, as defined by Cowling et al. (2001).

Scale, or resolution of the analysis may be sources of bias for the species-area calculations. We checked for this effect and found that the resolution did not affect our inference. Performing analyses for the same species and study region, but at a more coarse ($0.05^\circ$) resolution, does not change our quantitative results. Further, SARs for current distributions of sub-Saharan terrestrial birds (Manne unpubl.), with data at a $1^\circ$ resolution, yield the same widely-varying rates of species accumulation ($z$ spanning 0.13 to 1.0+).
The magnitude of the fitted z-value is likely to be very heavily-influenced by the number of species in the starting cell. For example, when starting with a small number of species, the rate of accumulation may be very rapid; however if beginning with a large number of species, species accumulation rate necessarily must be lower. The rate of Proteaceae species accumulation in an area is significantly-, but not well-predicted by the number of species in the starting area. Figure 5 shows a plot of z-value versus species richness, for which a linear regression model explains only 23% of the variance in the data. Consequently, the wide variation in species accumulation rates is not predominantly an effect of the number of species in the starting area.

Some may think it inappropriate or speculative to calculate SARs from future (expected) species ranges. Our calculations here are meant to highlight the community-level consequences of these future species distributions. If it is too speculative to calculate SARs from expected future ranges, then it must follow that the prediction of future ranges is also inappropriate.

It should be borne in mind that while the results here show that there is cause for concern over changing patterns of beta diversity over time and expected locations of extinctions, there are locations that are expected to become more diverse. Increased homogenization is not a blanket outcome for the Cape Floristic Region. However, these different outcomes point out the importance of not relying too heavily upon a simple rule of thumb to predict numbers of species remaining following some amount of climatic change.

It may be that ecologists have attempted to push the predictive power of the species-area relationship farther than is reasonable. The original species-area relationship was conceptualized from relatively small, well-defined, island areas that likely consisted of more homogenous environments than those found in a mainland area. We suspect that the wide variation in species accumulation rates, found elsewhere with simulated distributions (Sizling et al. 2004), may relate directly to habitat heterogeneity (Scheiner et al. 2000, Proches et al. 2003, Thuiller et al. 2006). An average SAR assumed, or formulated, for an entire biome has no way to capture such dependence, and thus should fail to capture smaller-scale variation across the biome.

Conclusion

Climate change will likely result in Proteaceae extinctions in some areas, but other areas will maintain their current rate of turnover, and still other areas could realize increased numbers of species and rates of species accumulation. As well, the number of extinctions in some areas, incurred as a result of changing climate, may be only modest. In terms of predicting these patterns, we have found that an average SAR for the Fynbos biological hotspot masks a very large interval of variation among SARs, for current Proteaceae distributions. Further, SARs as indices of turnover patterns do not hold fast under assumptions of biotic redistribution as climate changes. This variability of SARs in space and time calls into question, as others have done (Thuiller et al. 2004b), whether SARs can be applied for conservation if the possibility of large spatial variation in SARs is not considered (Thomas et al. 2004).

Acknowledgements – The authors would like to thank Conservation International (all) and the Natural Sciences & Engineering Research Council of Canada for funding (LLM), J. Brown, J. Harte, J. Kerr, S. Pimm, D. Srivastava, E. Weiher, and three anonymous reviewers for comments on draft versions of the manuscript, and R. Cowling for helpful conversations during manuscript preparation. Distribution data were obtained (December 2002) from the Protea Atlas Project, South African National Biodiversity Inst.

References

Rosenzweig, M. L. 1995. Species diversity in space and time. – Cambridge Univ. Press.