

THE PROPORTION OF CORE SPECIES IN A COMMUNITY VARIES WITH SPATIAL
SCALE AND ENVIRONMENTAL HETEROGENEITY

Molly F. Jenkins

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in the Curriculum in the Environment, Ecology, and Energy (Ecology).

Chapel Hill
2018

Approved by:

Allen H. Hurlbert

Peter White

James Umbanhowar

© 2018
Molly F. Jenkins
ALL RIGHTS RESERVED

ABSTRACT

Molly F. Jenkins: The Proportion of Core Species in a Community Varies with Spatial Scale and Environmental Heterogeneity
(Under the direction of Allen Hurlbert)

Ecologists define a community as the set of species successfully reproducing and using resources in a shared space. In lieu of a species list, the subset of temporally persistent core species may more appropriately fulfill this definition. Analyses carried out in communities that support low proportions of core species may violate assumptions about the definition of a community, and so poorly align with ecological predictions. We used bird time series data to calculate the proportion of core species across a gradient of scales, to investigate potential generalities in this pattern, and to use these generalities to address discrepancies on drivers of community assembly. We found that the proportion of core species in an assemblage increased in a positive curve with scale and decreased with high environmental heterogeneity (e.g. elevation, vegetation coverage). Communities with high heterogeneity and low proportions of core species were likely dispersal-driven, not resource-limited.

TABLE OF CONTENTS

LIST OF FIGURES	v
INTRODUCTION	1
METHODS	5
Bird data.....	5
Scaling metrics.....	7
Environmental data	7
RESULTS	9
DISCUSSION	11
Considerations.....	17
Conclusions.....	19
APPENDIX A: FIGURES AND FIGURE CAPTIONS.....	20
APPENDIX B: SUPPLEMENTARY MATERIAL	27
REFERENCES.....	29

LIST OF FIGURES

FIGURE 1: Probability density estimate of bird species' temporal occupancy.....	20
FIGURE 2: Methods and Predictions	21
FIGURE 3: Partitioning of boundaries for calculating scale	22
FIGURE 4: Probability densities of temporal occupancy.....	23
FIGURE 5: Proportion of core species present in assemblages as a function of scale.....	24
FIGURE 6: Regional environmental heterogeneity and the occupancy-scale relationship.....	25
FIGURE 7: Environmental heterogeneity across scales and the occupancy-scale relationship...	26
FIGURE 8: Testing core and transient cutoff thresholds: area (km ²).....	27
FIGURE 9: Testing core and transient cutoff thresholds: abundance	28

INTRODUCTION

Species differ in the temporal persistence with which they occur at any given site: while some species are reliably observed year in and year out, others appear only occasionally (Ulrich and Ollik 2004, Belmaker 2009, Dolan et al. 2009, Gaston et al 2007, Umaña et al. 2017). Indeed, recent work has shown that a broad range of ecological communities exhibit a bimodal distribution in temporal occupancy, reflecting these two groups which have been referred to as “core” and “transient” species (Coyle et al. 2013, Umaña et al. 2017, Snell et al. *in review*; Figure 1). Core species, in persisting at a site over time, are thought to maintain viable populations through successful reproduction (MacArthur 1957, Coyle et al. 2013, Snell et al. *in review*). Transients on the other hand do not persist reliably in these locations, and so do not maintain viable populations (Magurran and Henderson 2003, Umaña et al. 2017). Ecologists have typically ignored this distinction and have assumed that the complete list of species observed over some biological survey constitutes a meaningful “community” of interest for analysis. However, the relative representation of core and transient species in an assemblage may vary from assemblage to assemblage, and may in turn impact a wide range of ecological patterns, including species-area relationships, species abundance distributions, temporal turnover, and richness patterns (Magurran and Henderson 2003, Ulrich and Ollik 2004, Belmaker 2009, White and Hurlbert 2010, Coyle et al. 2013, Umaña et al. 2017, Snell et al. *in review*).

For many questions of interest, an ecological community is best defined as the set of species with positive growth rates interacting and inhabiting a shared space (Molles 2013), and so it is the set of core species rather than the list of all species that may be most relevant

(Magurran and Henderson 2003). For example, coexistence theory, niche theory, and other related ideas in ecology are largely predicated upon the occurrence of species that are suited to and influenced by their environments, inhabiting those environments and utilizing those environments for food and reproduction (Umaña et al. 2017). Analyses carried out in communities that support low proportions of core species, and high proportions of transient species may poorly align with ecological predictions that are less applicable to transient species. Distinguishing core from transient species requires detailed knowledge of individual species' niche requirements or long time series that enable the evaluation of temporal occupancy, but many ecological studies are limited by time and funding, and such data are often not feasible (Magurran and Henderson 2003, Umaña et al. 2017). Developing general principles regarding the factors that may influence the proportion of core species in an assemblage would enable researchers to more effectively compare results between studies and better assess generalities in community ecology.

The extent to which a species is a core, regularly occurring member of an assemblage will certainly depend upon the spatial scale over which that assemblage is sampled (Figure 1, Figure 2A). Consider two extremes: at the scale of 1 m², no bird species would maintain a viable population and be observed in every sampling period. At the scale of the entire North American continent, nearly all species would be reliably present at least somewhere within that extent every year. Thus, the proportion of core species in an assemblage must increase with scale, but the functional form of this relationship is less obvious. We expect the shape of the scaling relationship to be a positive decelerating curve (Figure 2C) because the stable proportion of species that a region can biologically support increases as the extent of a region is expanded and because species that are transient in a local context may be core species in the context of the

region; this increase is moderated by the inevitable inclusion of additional transient species over area.

Another factor that likely impacts the proportion of core species in an assemblage is environmental heterogeneity, which increases the proportion of transient species likely to occur via mass effects (Coyle et al. 2013, Snell et al. *in review*). Mass effects are more likely in heterogeneous landscapes, resulting in poorly adapted species dispersing intermittently from adjacent habitats within the region to which they are better suited (Shmida and Wilson 1985). The presence of such transient species which do not maintain viable populations effectively lowers the proportion of core species in such heterogeneous environments (White and Hurlbert 2010, Coyle et al. 2013, Snell et al. *in review*). . However, at larger regional scales these effects should be mitigated by the inclusion of sufficient source habitats that allow for the maintenance of viable and persistent breeding populations for a majority of species.

Environmental heterogeneity may also constrain habitat availability via the partitioning of space by multiple habitat types and the reduction of area per habitat type relative to environmentally homogeneous sites (Allouche et al. 2012). Resources within each habitat may occur at levels below the threshold needed to sustain viable populations (Allouche et al. 2012), constraining the proportion of core species for that scale compared to a homogeneous habitat of the same size. Thus, we expect high environmental heterogeneity to decrease the proportion of core species for local assemblages situated in heterogeneous landscapes (Umaña et al. 2017). While we expect high heterogeneity to have a negative effect on the proportion of core species in the regional assemblage, we predict the negative effect will be larger at the local scales given the prevalence of source-sink dynamics and the area-heterogeneity tradeoff (Figure 2C:D). In contrast, the proportion of transient species should be determined by regional variability in the

number of potential colonists available in the regional pool, which are driven by mass effects and regional environmental heterogeneity (Shmida and Wilson 1984, White and Hurlbert 2010, White et al. 2010, Coyle et al. 2013). We also predict that this relationship will vary at intervening scales such that environmentally heterogeneous sites will be unable to support substantial (>50%) proportions of core species relative to the scale at which homogeneous sites successfully support substantial proportions of core species. Finally, we expect this scaling measurement will contribute to differences in the shape of the overall curve – with large scale thresholds corresponding with accelerating curves, and small scale thresholds typifying a positive decelerating curve.

Here, we make use of a large spatiotemporal dataset on bird distribution and community size which allows us to investigate temporal occupancy, and hence the proportion of core species in an assemblage, over a wide range of spatial scales and environmental contexts. Specifically, we seek to 1) describe the distribution of species' temporal occupancy in ecological assemblages across a gradient of spatial scales, 2) describe the relationship between the proportion of core species in a community and the spatial scale at which that community is characterized, and 3) characterize how environmental heterogeneity influences that scaling relationship.

METHODS

Bird data

We used data on the distribution of land bird species over time (excluding owls, kingfishers, and nightjars) from the North American Breeding Bird Survey (BBS), maintained by the United States Geological Survey (Pardieck 2017). Our data encompassed the 968 BBS routes across the North American continent that were surveyed continuously over the 15 year period from 2000-2014 that had at least 65 neighboring routes within 1,000 km. Each BBS route is a 40 km roadside transect encompassing fifty 3-minute point count stops, each separated by 0.8 km, in which a single observer records all birds detected within 0.4 km. Thus, each point count stop surveys an area of 0.5 km² while an entire BBS route surveys an area of 25.1 km². BBS routes were surveyed each year during the breeding season, typically in June.

Temporal occupancy, the proportion of years a species was observed over some spatially defined area, was calculated for each species at each site at a range of spatial scales (Figure 3). We calculated the proportion of core species in each assemblage as the proportion of species with temporal occupancy greater than two-thirds (i.e. occurring in at least 11 out of the 15 survey years) following Coyle et al. (2013). We also considered two alternative thresholds for defining core species that produced qualitatively similar results (Supplemental methods, Appendix). Below the scale of a single BBS route, each route was split into non-overlapping segments of 5, 10 or 25 point count stops, and the proportion of core species was calculated at each spatial scale. To examine spatial scales greater than a single BBS route, for each focal route we

sequentially aggregated survey data from an increasing number of nearest neighbor routes, up to a maximum regional scale of the focal route together with its 65 nearest neighbors (Figure 3).

Our regional scale of 66 neighboring routes was chosen because it was the largest number of neighbors that fell within a radius of 1,000 km of each focal route, even in regions of lower route density in the western US (Figure 3). The entire range of spatial scales we investigated varied from 2.5 km² for a set of 5 point count stops up to 1,659 km² for an area of 66 adjacent BBS routes. Because BBS route density varies across the continent, the spatial extent of the 65 nearest neighbors did vary (Figure 3). However, using a fixed total number of aggregated routes allowed us to keep the total surveyed area characterizing an assemblage constant, and this was the aspect of scale we viewed as most critical for our comparisons. While regions of the same sampled area but spanning larger extents may encompass a greater range of environmental variation all else equal, such regions (with lower BBS route density) tend to occur in the western US which is more environmentally heterogeneous. If this method did affect the measured proportion of core species by introducing extraneous heterogeneity, we would expect the measured proportion to be slightly lower than it is in reality for both homogeneous regions and heterogeneous regions. If this method substantially affected our measured proportion of core species, we would expect to see little to no difference in the proportion of core species at the local scales between environmentally homogeneous and heterogeneous sites.

In addition to spatial scale, we used the total number of individuals observed in the assemblage (community size) as an alternative measure of scale. Community size was found by Snell et al. (*in review*) to be a potentially more generalizable measure of scale than area, especially for comparing between taxonomic groups with very different area requirements.

Scaling metrics

We derived a series of metrics characterizing the relationship between the proportion of core species present and scale for each focal route (Figure 2C). We identified the proportion of core species at the smallest scale (p_{min}) and the proportion at the largest scale (p_{max}) for each focal route. We also identified the slope of the line linking p_{min} and p_{max} for each focal route. We identified the scale at which the proportion of core species in the community surpassed the threshold of 0.5 for each focal route ($scale_{50}$). Finally, we characterized the degree of curvature in the relationship between the proportion of core species in the community and scale. We calculated the area between the curve and a straight line by summing the differences between the observed values and the values expected from a linear relationship between p_{min} and p_{max} (Figure 3), where positive values indicate positive decelerating relationships and greater proportions of core species. Negative values indicate positive accelerating relationships and lower proportions of core species relative to a linear relationship.

Environmental data

We acquired raster layers for elevation from Worldclim (Fick & Hijman, 2017), and Normalized Difference Vegetation Indices (NDVI) from the NASA GIMMS group, and calculated mean NDVI and mean elevation for each focal route within a 40 km buffer. For each environmental variable, we defined regional heterogeneity around each focal route as the variance in mean values across the set of 65 nearest neighbor BBS routes plus the focal route. In order to assess the whether the importance of environmental heterogeneity varied with the spatial scale over which heterogeneity was measured, we calculated environmental heterogeneity at different scales (from 3 to 66 neighboring routes) and then examined the Pearson's correlation

between heterogeneity and the five scaling metrics describing how the proportion of core species varies across the full range of spatial scales.

RESULTS

At the scale of a single route (~25 km²), temporal occupancy was bimodal as expected (Figure 4, dashed line). At larger spatial scales, assemblages were marked by a greater proportion of core species with high temporal occupancy, while at smaller scales, assemblages were characterized by a greater number of transient species and very few core species (Figure 4). The proportion of core species in a community increased on average in a positive decelerating manner with both measures of spatial scale, although there was substantial variability from route to route (Figure 5). At the largest spatial scales, the proportion of core species exhibited reduced variation, with a mean of 83% and ranging from 75%-90%, while at the smallest spatial scales (2.5 km²) the proportion of core species varied from 11-37%. Using community size in lieu of spatial scale greatly reduced this variation in the proportion of core species at the smallest scale (Figure 5).

Heterogeneity in elevation and heterogeneity in NDVI for a region both had similar effects on the overall shape of the relationship between core species and spatial scale, although the effects of elevation were generally stronger (Figure 6). Environmental heterogeneity in a region was associated with communities that were characterized by a low proportion of core species at both the smallest and largest scales, and communities that experienced the greatest increase in the proportion of core species between the smallest and largest scales.

The scale at which environmental heterogeneity was measured also affected the strength of the correlation between heterogeneity and scaling curve metrics (Figure 7). Specifically,

heterogeneity in elevation was most strongly correlated with all five of the scaling metrics when measured at the largest spatial scale, whereas heterogeneity in NDVI was most strongly correlated with the scaling metrics at scales between 15-25 BBS routes (400-600 km²). Both measures of environmental heterogeneity were associated with having a strong negative effect on the minimum and a strong positive effect on the slope, but also with having a strong positive effect on the scale₅₀ parameter at scales below 500 km² (Figure 7).

DISCUSSION

Ecologists frequently test hypotheses regarding community assembly and species richness using surveys which reflect community snapshots. However, it is increasingly recognized that such a snapshot approach fails to differentiate core species from transient species, the former maintaining viable populations and interacting more strongly with their biotic and abiotic environment, and the latter being irregular visitors that are presumably better adapted to other conditions (Magurran and Henderson 2003, White and Hurlbert 2010, Umaña et al 2017). We used a continent-wide dataset on bird assemblages over time to show how the proportion of core species in these assemblages increases with scale and decreases with environmental heterogeneity. Consistent with Coyle et al. (2013), the distribution of temporal occupancy was strongly bimodal at the scale of a single BBS route, reflecting these two distinct groups. However, at scales below the size of a BBS route ($25 < \text{km}^2$) few species were present consistently over time, while at scales larger than two aggregated BBS routes ($>50 \text{ km}^2$) most species occurred regularly. On average, the proportion of core species in a community increased in a positive decelerating manner as a function of spatial scale. Species identified as transient at small scales became core species at larger scales, as the probability of including suitable habitat in sufficient quantities to maintain persistent populations increased. However, for a group like birds with such strong dispersal capabilities, there is always a chance that new transient species will disperse and be observed each year, and so even at the continental scale the proportion of core species never actually reaches 1. The proportion of core species changes in its rate of accumulation over area. Small assemblages ($<5 \text{ km}^2$) exhibited a fairly wide range in the

proportion of core species present (11-37%), at least in part because different sites differed in the overall number of individuals supported. At the largest spatial scales (1,659 km² of surveyed area distributed across a 1,000 km radius region) there was less variation in the proportion of core species (75-90%). This suggests that at local scales, some sites may support source populations of species that are well-suited to the local site, but these sites may also be sinks for other species better suited to adjacent sites. Additionally, sites of the same scale may reflect radically different proportions of core species given how many species are well suited or not, creating variability in the proportion of core species as a function of the relative proportion of species in the pool that relate to the site as a source versus sink site. Viable populations of core species may occur at the scale of the region as they exhibit positive population growth rates for an overall region, but where specifically those populations occur and how prevalent they are at a given location may vary locally with the available resources.

Much of the variation in the shape of the relationship between the proportion of core species in a community and spatial scale can be explained by the regional environmental heterogeneity. Specifically, landscapes with high environmental heterogeneity have proportionally fewer core species, and this effect is strongest at the smallest spatial scales. Consistent with previous findings, we found that environmental heterogeneity was positively correlated with the proportion of transient species (Coyle et al. 2013, Stein et al. 2014, Snell et al. *in review*). This was true whether characterizing heterogeneity based on regional variation in elevation or NDVI, but the effect of elevation was both stronger and more apparent at the regional scales (Figure 7). This is likely because variation in elevation encompasses habitat diversity by virtue of elevational zones in addition to differences in slope, hydrology, shelter, and other topographic features. Variation in NDVI also reflects these differences, but perhaps less

directly as the habitat variation within a given range of NDVI may not be well captured.

Ultimately, regional heterogeneity introduces greater variability from site to site, increasing the relative proportion of transient species at local scales via the increased likelihood that one patch differs in structure from a neighbor patch, and may support very different species as a result (Ricklefs 1987, Coyle et al. 2013, Stegen 2013, Umaña et al. 2017, Snell et al. *in review*).

Landscapes with low environmental heterogeneity should support consistent communities with low turnover (Stegen et al. 2013, Gaston et al. 2007), even at smaller spatial scales nested within the landscape as these landscapes more closely parallel the resources of the region they occur within. This allows competition and resource constraints to be the driving mechanisms to filter out species that are less well suited, permitting only those well-suited species to persist and support viable populations (Coyle et al. 2013). Landscapes with a high degree of environmental heterogeneity are more compartmentalized, effectively decreasing the area and resources available per habitat to support a viable species population (Allouche et al 2012), and instead increasing the number of species in the regional pool that could potentially immigrate into a local habitat via mass effects (Shmida and Wilson 1984, Gaston et al. 2007, White and Hurlbert 2010, Coyle et al. 2013, Stegen et al. 2013). However, our estimate of the relative curvature of the relationship between the proportion of core species and scale was positively related to environmental heterogeneity in the form of NDVI (Figure 6, 7), contrary to our expectations. Some environmental variability in NDVI may actually contribute to higher proportions of core species, particularly at small scales. However, variability in elevation was weakly and negatively related to curvature (Figure 6, 7). More statistically appropriate measures of curvature may yield opposing or stronger relationships, and we plan to address this in future research.

Typically, the scale at which communities surpassed a proportion of 50% core species was an area of 2 BBS routes (about 53 km²) or 1,412 individuals, with a range of 50-200 km² or

400-4000 individuals. Because a substantial proportion of core species are supported by an area the size of a BBS survey, the scale at which BBS surveys are conducted is a reasonable albeit imperfect scale by which to measure a community. The lowest observed proportion of transients at the scale of a BBS route is still about 11% of the community, which may be enough to affect ecological inferences based on commonly examined community ecology patterns (Snell et al. *in review*). To minimize the influence of transient species, assemblages are best surveyed at larger scales (≥ 1400 individuals), but when this is not feasible smaller and environmentally homogeneous sites may be perfectly adequate depending on the ecological question. For example, the Breeding Bird Census (Lowe 2006) is characterized by many sites of variable size (0.1-1 km²) covering fairly uniform habitats - these sites may potentially support less variation in the local-scale proportion of core species when compared to the degree of variation in the proportion of core species at the more size-standardized but environmentally heterogeneous BBS routes. Small heterogeneous sites may be far more problematic for studies that assume communities that are predominantly core species, as these sites are far more subject to both inherent variability in the true proportion of biologically core species and mismatches between the statistical and biological definitions of core species due to detection errors.

This new understanding of how the proportion of core species in an assemblage varies with scale and environmental heterogeneity may help resolve discrepancies between studies with respect to the importance of biotic interactions, resource availability, or mass effects for driving community assembly locally and regionally (Henderson and Magurran 2014). Previous attempts to synthesize and generalize from ecological studies testing concepts such as these may have failed because those studies differed in scale, environmental heterogeneity, or both leading to assemblages with very different proportions of core species and subsequently different apparent

mechanisms driving community assembly (e.g., Dorazio et al. 2006, Emerson & Gillespie 2008, Stein et al. 2015). Such comparisons across scales may pose complications in that certain processes are more or less biologically relevant at different and often taxa-specific scales. For example, competition and environmental filtering are two processes that have been proposed to shape community assembly and result in either phylogenetic overdispersion or clustering (Cavender-Bares et al. 2004). However, the signal strength of dispersal or clustering may be affected by the proportion of core or transient species in a community. Core species are more likely to compete with each other for resources, and are thereby expected to contribute the most to overdispersion in competition related traits. In addition, core species are expected to be better suited to the local climate or habitat compared to transient species, and so they are expected to exhibit a greater degree of clustering of environmental tolerance traits. At small spatial scales, the proportion of transient species will be much higher, and the likelihood of discerning a nonrandom assembly pattern from such a community is low. The proportion of core species is lowest at relatively small scales, and yet processes more likely to drive core species assembly like competition exert greater pressure at smaller scales where resources are more constrained by space (Allouche et al. 2012). This may result in seemingly conflicting patterns of community assembly in large meta-analyses that include studies from disparate taxonomic groups and conducted at a wide range of scales. For example, a meta-analysis on phylogenetic clustering versus overdispersion conducted by Emerson and Gillespie (2008) found an assemblage of Cuban *Anolis* lizards to exhibit unstructured or seemingly random patterns of phylogenetic assembly while an assemblage of dusky salamanders exhibited strong phylogenetic overdispersion. This result may reflect a lack of generality in the degree of overdispersion in community assembly. However, these two studies were carried out at very different spatial scales. The first study may have found unstructured patterns as a result of the small scale of the

study, where we would expect there to be a high relative proportion of transients contributing to noise. However, another single island small-scale *Anolis* clade may contradict this via evident phenotypic and phylogenetic overdispersion, due to a strong signal of competition at the small spatial scale from the proportion of core species present, given the increased likelihood of competitive effects at small scales where resources are constrained. However, at the scale of several islands, *Anolis* communities exhibit phylogenetic clustering and phenotypic overdispersion. This may be due to a high proportion of core species but a low relative competitive signal compared to that expected at the scale of a single island of constrained resources. These potential different outcomes at different scales are still all limited within a single taxonomic group; comparisons of assembly patterns between groups may further strain effective comparisons, as several dozen studies are compared in the synthesis by Emerson & Gillespie (2008). One such study investigated dusky salamanders across the Blue Ridge mountain range, from Southern VA to Northern GA. The salamanders exhibited a strong environmental filtering effect through a combination of overdispersion and conserved evolution. This signal may be due to a combination of the large scale sampled relative to the organism and the high number of core species, as at higher scales the signal strength of competition is generally lessened, and so the effects of environmental filtering may be more apparent. It is possible that within a streambed, while there may be only one or two *Desmognathus* present, those individuals are both reliably persistent in their occupation of the site, as well as the degree to which they are affected by the presence of other amphibians competing for resources in the same streambed. In any assemblage, the presence of transient species likely contributes to greater noise, potentially masking any nonrandomness in assembly patterns. Subsetting for core species appropriate to the surveyed scale and taxa would likely reduce noise and allow for more consistent trait assembly patterns to be evaluated and compared, potentially improving

generality. Additionally, researchers should discern whether one or both measures of scale are more appropriate for validating their particular question of interest. Some ecological questions may be more appropriately addressed using spatial scale as a measure, particularly those that are contained within a single taxonomic group. However, for comparisons across taxonomic groups, community size may be a more appropriate metric as it avoids the complicating and spatially dependent issue of different territory requirements between species and communities. Because different communities have different scale requirements, and because different mechanisms of community assembly operate at different scales, ideally communities should be standardized for comparison using spatial scale or community size using rarefaction-based methods, and that the most ecologically relevant group in the assemblage is analyzed as a subset of the assemblage (Hurlbert 2004, Gotelli and Colwell 2011, Snell et al. *in review*). Comparisons across studies can then be more easily rectified, as the prevalent drivers evident in one study may not contradict, but instead complement the drivers evident in another study.

Considerations

Two types of classification errors are possible when inferring whether species are truly biologically core species indirectly from time series data: a species may be inferred to be transient when it is in fact core (a false negative), and a species may be inferred to be a core species when it is in fact transient (a false positive). False negatives are expected to occur at small scales, when the species in question occurs at low densities and is not detected regularly (Coyle et al. 2013, Henderson and Magurran 2014, Snell et al. *in review*). However, it is unlikely that such false negatives will occur frequently at large scales, as even persistent species with low abundance reliably occur at large scales from year to year. False positives are expected to occur at small scales in regions of high environmental heterogeneity. A species that is in reality

transient at a local sink site but core in the context of the region may be statistically considered as core at that local site because neighboring sites support sufficient viable populations to ensure persistent occurrence in spite of the fact that the species may not be successfully breeding.

Because we expect classification errors to vary as a function of scale, heterogeneity, and the abundance and detectability of species, there are a number of ways that these errors may be addressed by future research. Future research could make use of simulation models to identify false positive and false negative error rates for communities across different scales and relative degrees of environmental heterogeneity, evaluating the specific extent to which small scales and regions of high heterogeneity contribute to classification errors. Additionally, research comparing how temporal occupancy scales for different species with different density and detectability issues could also help identify species or guilds in an assemblage that are more or less susceptible to being improperly classified. For example, some core species may be easy to detect by sight or hearing in spite of their low abundance, while others may be elusive and silent in addition to supporting low abundances, decreasing the likelihood that they will be consistently taken into account as core species and increasing the likelihood that they will be classified as transients. Finally, more strict cutoff thresholds for determining the proportion of core to transient species may bear utility in excluding species that contribute to these errors, as the likelihood of false positives declines with more strict thresholds for qualifying a species as a core species (Appendix B, Figures 8, 9). While a formal investigation into both the likelihood and scaling relationship of these potential classification errors is beyond the scope of this paper, it is nonetheless essential to the utility of the core-scaling relationship.

Conclusions

Community surveys are the main unit of analysis in community ecology, and yet the species lists derived from such surveys generally include a combination of species that are core regularly interacting members of the community and transient species that may play little role in shaping the biotic environment. Here, we've shown that the proportion of core species in an assemblage increases as a function of scale, and decreases as a function of environmental heterogeneity. Since the proportion of core species may influence a number of essential patterns in community ecology, an understanding of the factors that influence the prevalence of core species is critical for the proper interpretation of synthetic meta-analyses and the evaluation of ecological theory. Our findings also have implications for how best to select ecologically relevant scales for improved evaluation of community assemblages.

APPENDIX A: FIGURES AND FIGURE CAPTIONS

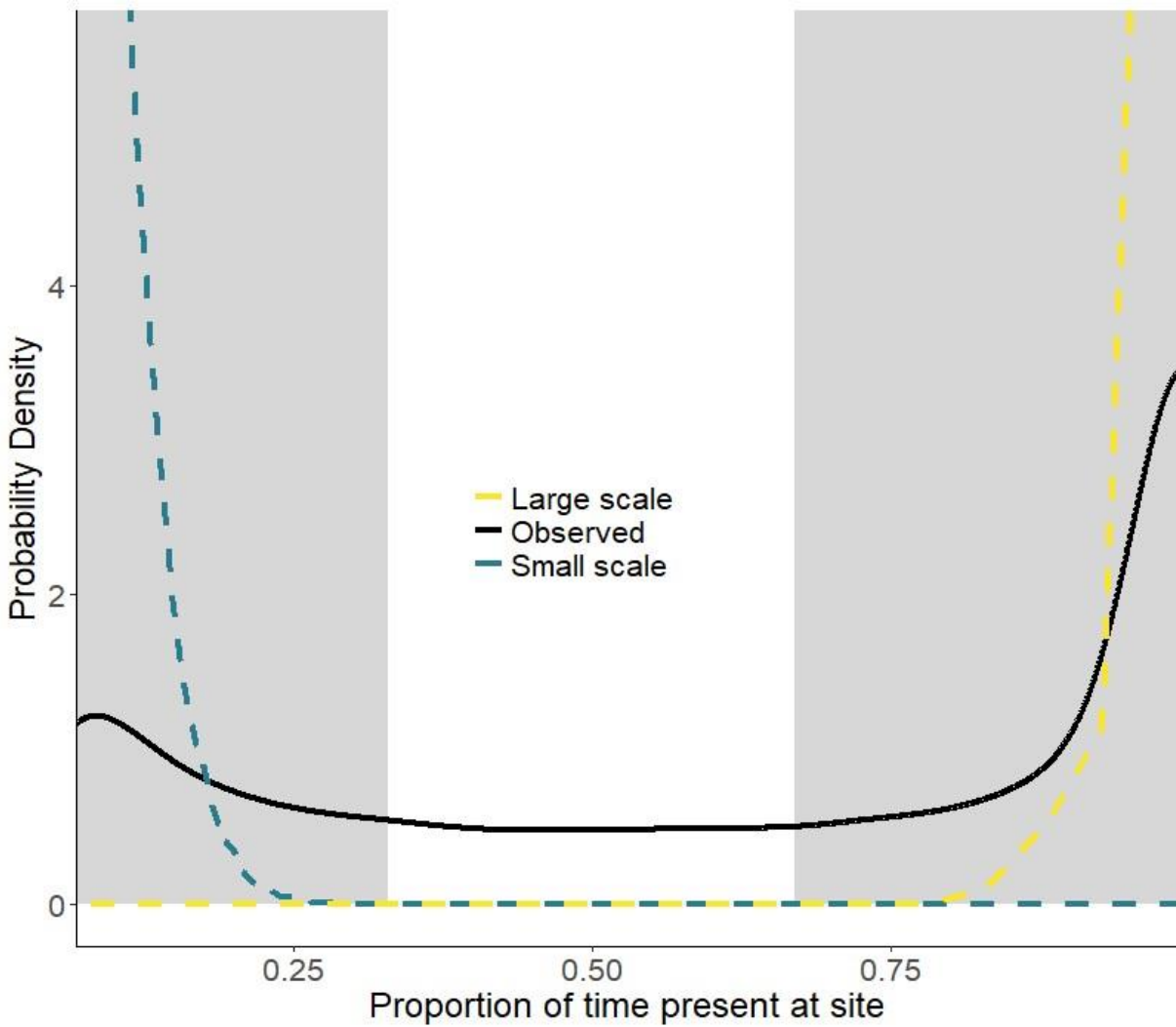


Figure 1. Probability density estimate of bird species' temporal occupancy across 968 North American Breeding Bird Survey routes at the scale of 25 km² (black line), and predicted probability density estimates at fine (blue dashed) and coarse (yellow dashed) spatial scales. The temporal occupancy of a species on a route is the proportion of surveyed years in which the species was recorded present. Grey rectangles indicate temporal occupancy levels used to classify species as core ($>2/3$) or transient ($\leq 1/3$).

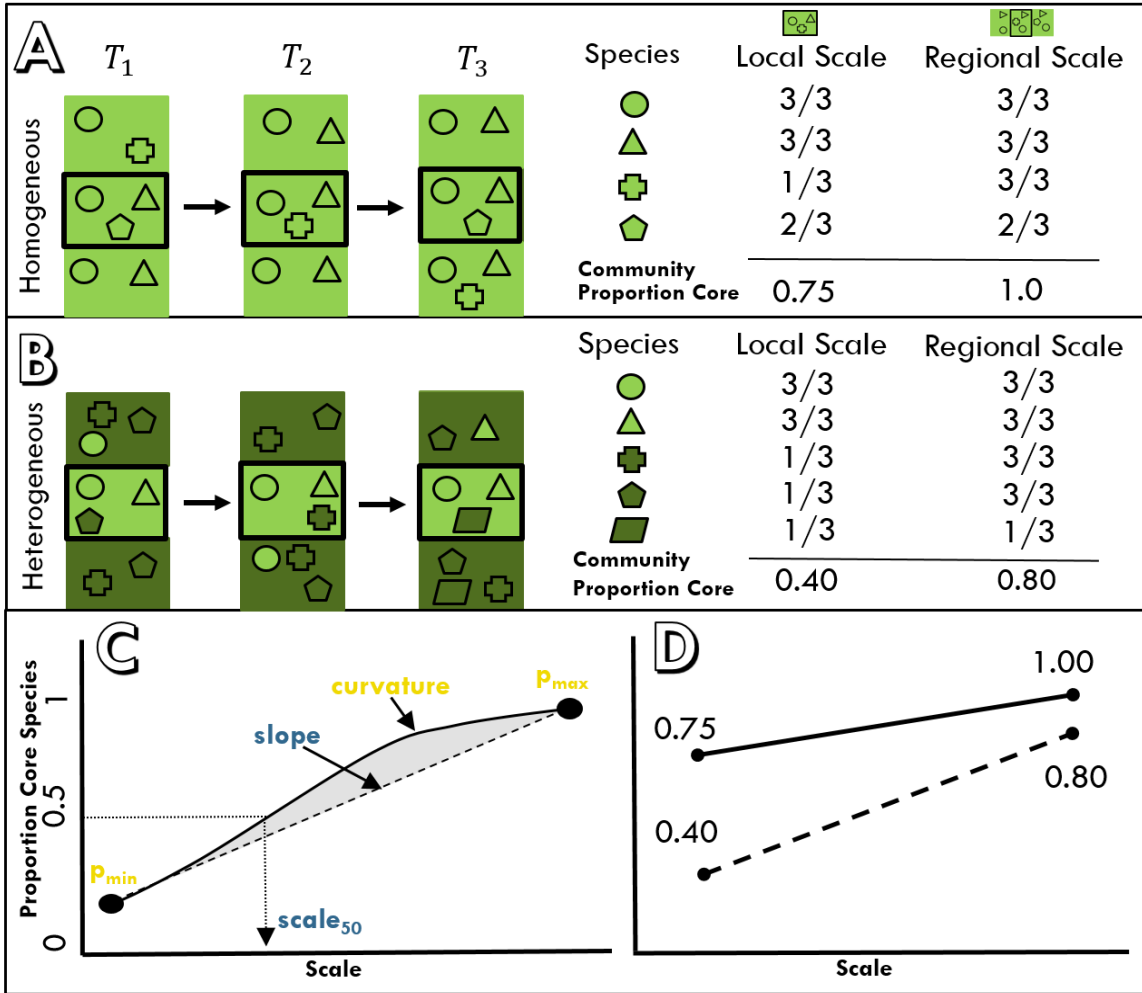


Figure 2. The proportion of core species in a community is expected to vary with scale and environmental heterogeneity. (A, B) Species (symbols) are distributed across an environmentally homogeneous (A) or heterogeneous (B) landscape over three time periods (T1, T2, T3). The temporal occupancy of each species as well as the proportion of core species in the assemblage that occur in 2/3 or more time periods is assessed at both the local (central black boxes) and regional (rectangles) scales. The color of species symbols indicates habitat affinities for landscapes of the same color. (C) A generalized scaling relationship for the proportion of core species in a community. We consider the following parameters from this curve: 1) p_{min} , proportion of core species at the minimum spatial scale, 2) $scale_{50}$, the spatial scale at which the community first exceeds 50% core species, 3) p_{max} , proportion of core species at the maximum spatial scale, 4) slope, the slope of the line linking the minimum and maximum values, and 5) curvature, calculated as the area between the scaling curve and the straight line connecting min and max values. Parameters in yellow are expected to be negatively related to environmental heterogeneity, while parameters in blue are expected to be positively related to environmental heterogeneity. (D) The proportion of core species in (A) and (B) at local versus regional scales for landscapes of high and low environmental heterogeneity.

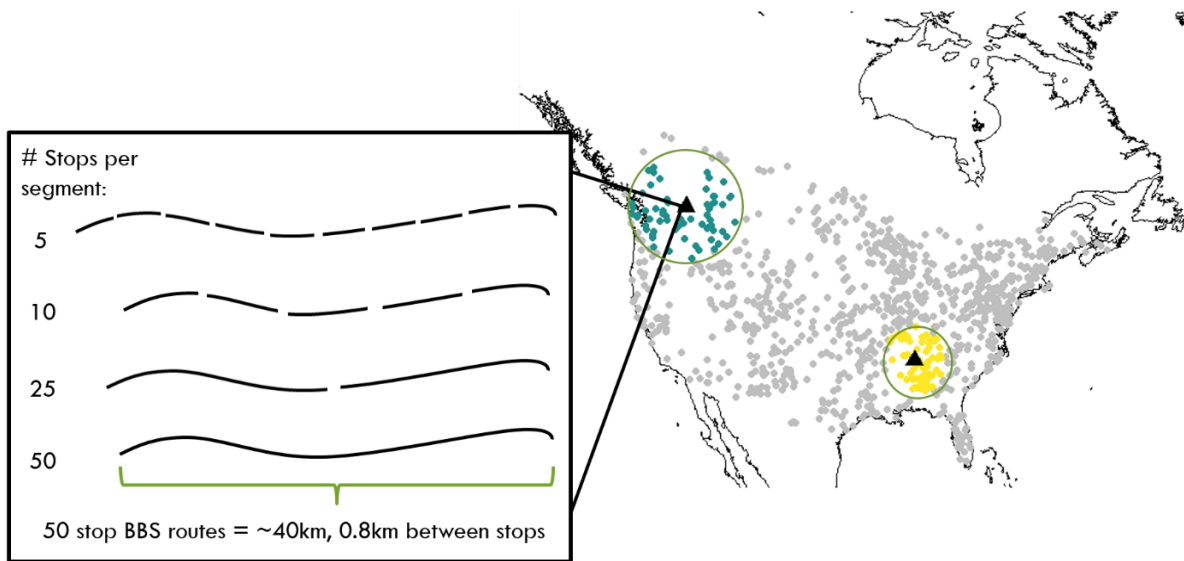


Figure 3. Aggregation and partitioning of spatial boundaries of ecological communities for calculating temporal occupancy and proportion of core species. On the right - a map of North America highlights all of our BBS route sites in grey ($n = 968$), while also highlighting the maximum (66) routes aggregated total to encompass the regional scale. On the left - a breakdown of how we segmented routes for the smallest and most local scales, based on the number of point counts conducted within each segment.

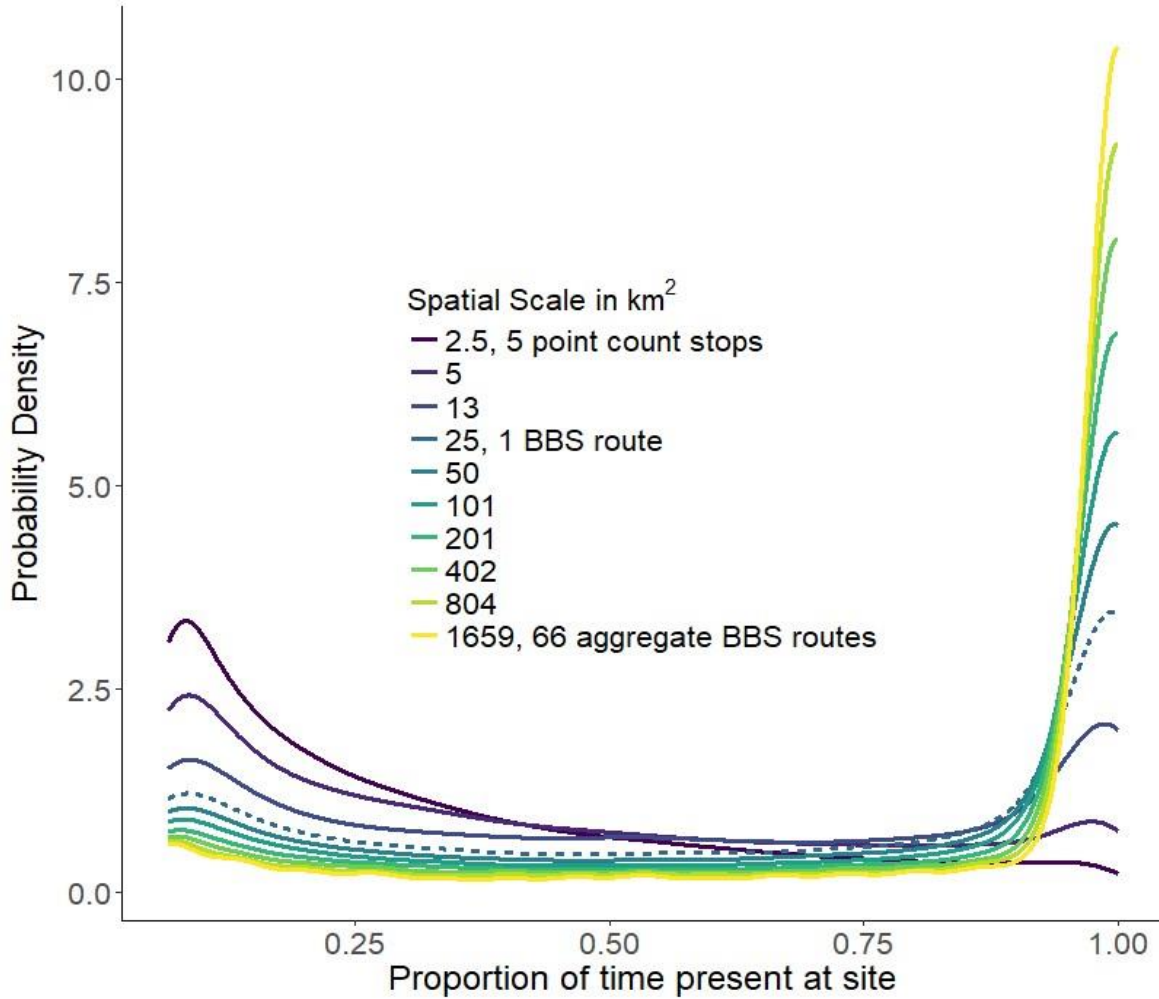


Figure 4. Average probability densities of temporal occupancy for the bird species present at a site, calculated over ten spatial scales from small (dark) to large (light). Each curve represents the average probability density across 968 BBS routes at a particular scale. BBS route scale highlighted with dashed line.

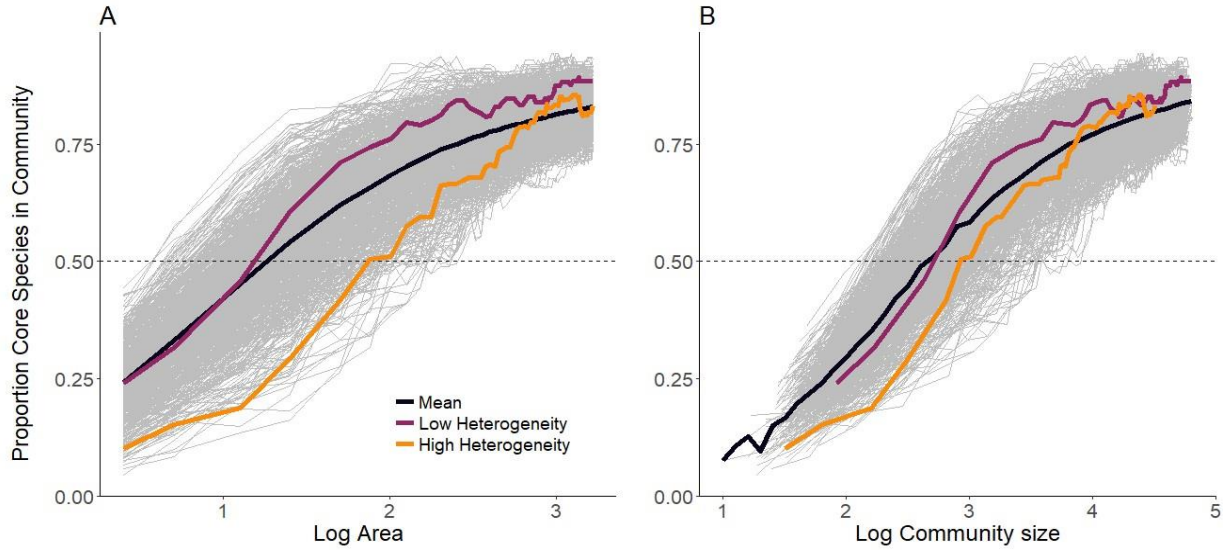


Figure 5. Proportion of core species present in assemblages as a function of (a) spatial scale in km^2 and (b) scale as measured by community size. Each line represents a single focal BBS route; we examined 968 routes total. Overall moving window average across all BBS routes indicated by the bold black line. Additional highlighted routes exemplify low environmental heterogeneity (purple, Illinois) and high environmental heterogeneity (orange, Utah).

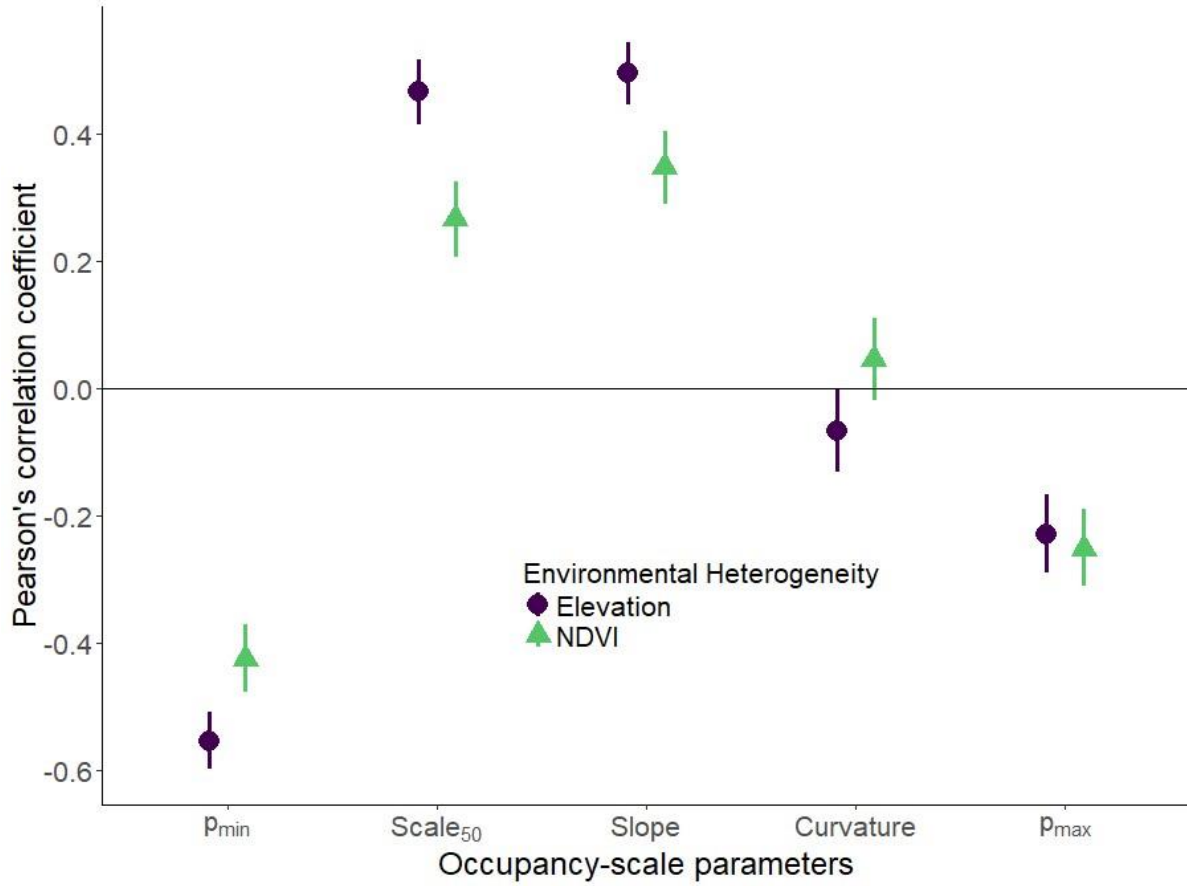


Figure 6. Relative importance of environmental heterogeneity at the regional scale (66 BBS routes) to the occupancy-scale relationship as characterized by five parameters. The strongest effects are highlighted by the “minimum” parameter (p_{min}) and the “slope” parameter, with elevation playing the stronger role compared to NDVI.

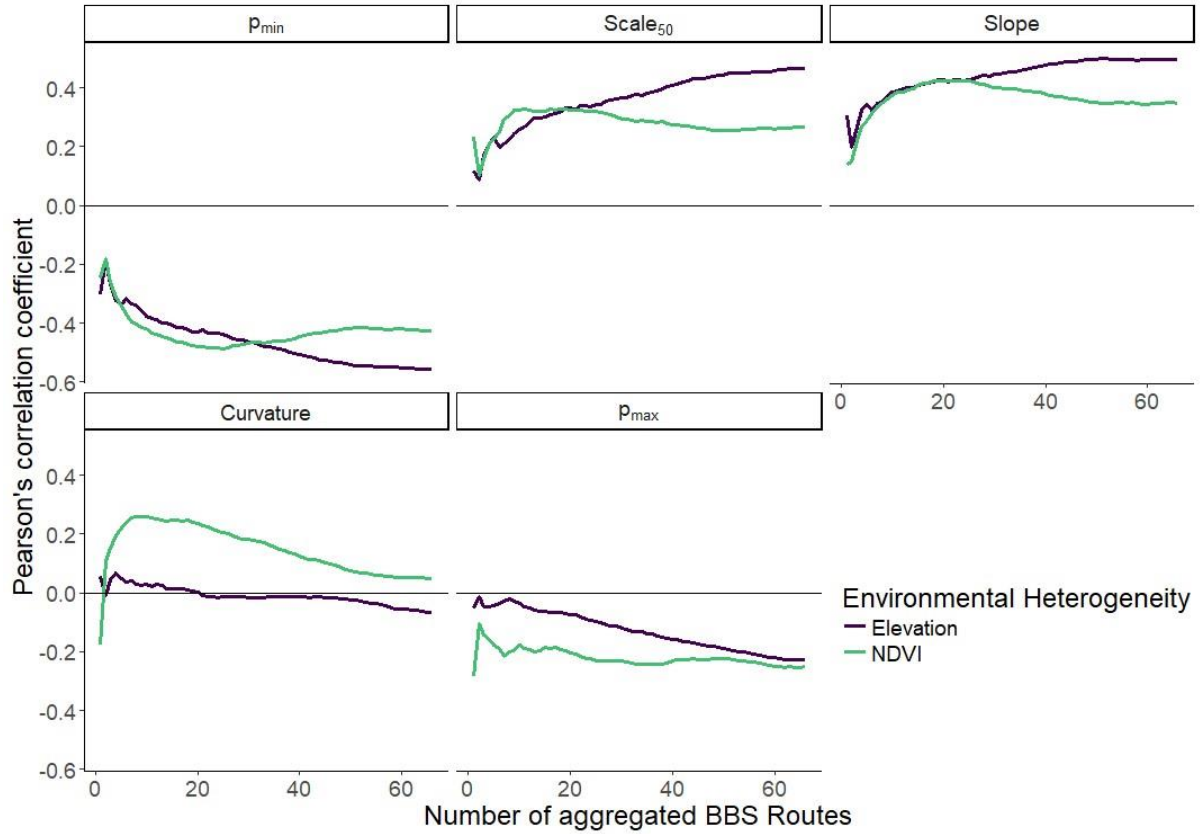


Figure 7. The correlation between environmental heterogeneity (elevational heterogeneity in purple, NDVI heterogeneity in green) and five metrics describing the shape of the proportion of core species-scale relationship as a function of the scale (i.e., number of nearest neighbor BBS routes) over which environmental heterogeneity was measured.

APPENDIX B: SUPPLEMENTARY MATERIAL

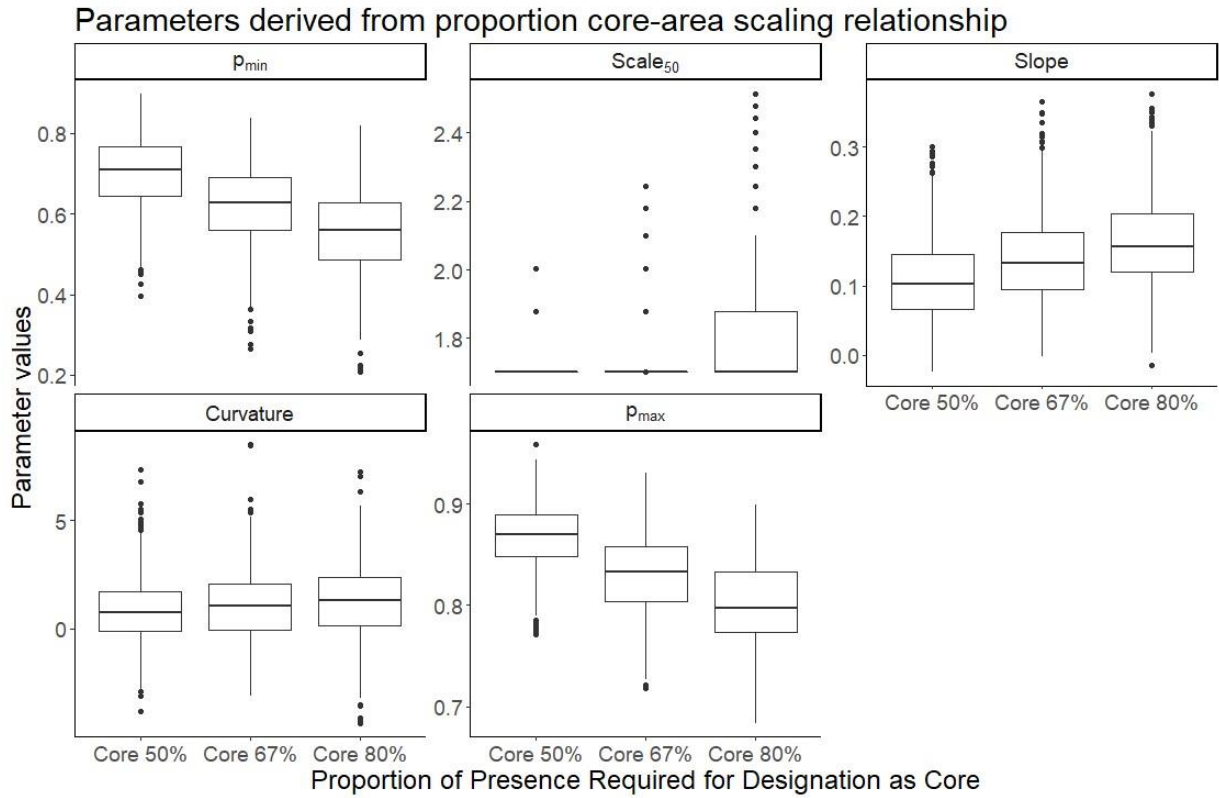


Figure 8. The variation in parameter estimates for the core abundance relationship across different core and transient cutoff thresholds. While there was not an appreciable difference between cutoffs, slight differences may indicate the resilience of higher cutoff thresholds against classification errors.

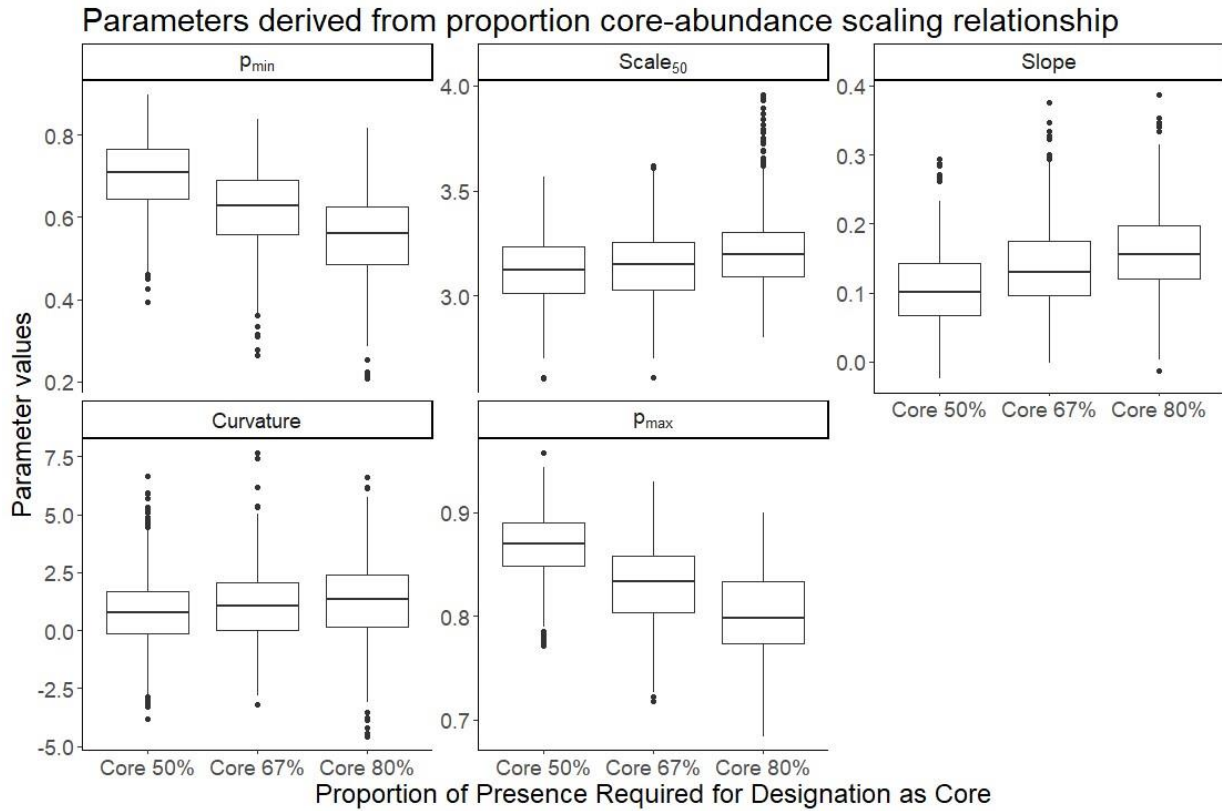


Figure 9. The variation in parameter estimates for the core abundance relationship across different core and transient cutoff thresholds. While there was not an appreciable difference between cutoffs, slight differences may indicate the resilience of higher cutoff thresholds against classification errors.

REFERENCES

- Allouche, O., M. Kalyuzhny, G. Moreno-Rueda, M. Pizarro, and R. Kadmon. 2012. Area–heterogeneity tradeoff and the diversity of ecological communities. *Proceedings of the National Academy of Sciences* 109:17495–17500.
- Belmaker, J. 2009. Species richness of resident and transient coral-dwelling fish responds differentially to regional diversity. *Global Ecology and Biogeography* 18:426–436.
- Brown, J. H., and A. Kodric-Brown. 1977. Turnover Rates in Insular Biogeography: Effect of Immigration on Extinction. *Ecology* 58:445–449.
- Cavender-Bares, J., D. D. Ackerly, D. A. Baum, and F. A. Bazzaz. 2004. Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist* 163:823–843.
- Coyle, J. R., A. H. Hurlbert, and E. P. White. 2013. Opposing Mechanisms Drive Richness Patterns of Core and Transient Bird Species. *The American Naturalist* 181:E83–E90.
- Dolan John R., Ritchie Mark E., Tunin-Ley Alina, and Pizay Marie-Dominique. 2009. Dynamics of core and occasional species in the marine plankton: tintinnid ciliates in the north-west Mediterranean Sea. *Journal of Biogeography* 36:887–895.
- Dorazio Robert M., Royle J. Andrew, Söderström Bo, and Glimskär Anders. 2006. Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology* 87:842–854.
- Emerson, B. C., and R. G. Gillespie. 2008. Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution* 23:619–630.
- Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas: NEW CLIMATE SURFACES FOR GLOBAL LAND AREAS. *International Journal of Climatology* 37:4302–4315.
- Gaston, K. J., R. G. Davies, C. D. L. Orme, V. A. Olson, G. H. Thomas, T.-S. Ding, P. C. Rasmussen, et al. 2007. Spatial turnover in the global avifauna. *Proceedings of the Royal Society B: Biological Sciences* 274:1567–1574.
- Gotelli, N. J., & Colwell, R. K. 2011. Estimating species richness. *Biological diversity: frontiers in measurement and assessment*, 12, 39-54.
- Henderson, P. A., and A. E. Magurran. 2014. Direct evidence that density-dependent regulation underpins the temporal stability of abundant species in a diverse animal community. *Proceedings of the Royal Society B: Biological Sciences* 281.

- Hurlbert, Allen H. 2004. Species–energy relationships and habitat complexity in bird communities. *Ecology Letters* 7:714–720.
- Lowe, J. D. 2006. An annotated bibliography of Breeding Bird Census publications. *Bird Populations* 7:128–135.
- MacArthur, R. H. 1957. On the Relative Abundance of Bird Species. *Proceedings of the National Academy of Sciences of the United States of America* 43:293–295.
- Magurran, A. E., and P. A. Henderson. 2003. Explaining the excess of rare species in natural species abundance distributions. *Nature* 422:714–716.
- Molles, M. C. 2013. *Ecology: concepts and applications* (6th ed.). McGraw-Hill, New York.
- Pardieck, K. L., D. J. Ziolkowski, M. Lutmerding, K. Campbell, and M.-A. R. Hudson. 2017. North American Breeding Bird Survey Dataset 1966 - 2016, version 2016.0. U.S. Geological Survey.
- Shmida, A., and M. V. Wilson. 1985. Biological Determinants of Species Diversity. *Journal of Biogeography* 12:1–20.
- Snell, S., B. Evans, E. White, and A. Hurlbert. 2017. The prevalence and impact of transient species in ecological communities. *bioRxiv* 163816.
- Stegen, J. C., A. L. Freestone, T. O. Crist, M. J. Anderson, J. M. Chase, L. S. Comita, H. V. Cornell, et al. 2013. Stochastic and deterministic drivers of spatial and temporal turnover in breeding bird communities. *Global Ecology and Biogeography* 22:202–212.
- Stein, A., K. Gerstner, and H. Kreft. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* 17:866–880.
- Stein, A., and H. Kreft. 2015. Terminology and quantification of environmental heterogeneity in species-richness research. *Biological Reviews* 90:815–836.
- Ulrich, W., and M. Ollik. 2004. Frequent and occasional species and the shape of relative abundance distributions. *Diversity and Distributions* 10:263–269.
- Umaña Maria Natalia, Zhang Caicai, Cao Min, Lin Luxiang, Swenson Nathan G., and Suding Katharine. 2017. A core-transient framework for trait-based community ecology: an example from a tropical tree seedling community. *Ecology Letters* 20:619–628.
- White, E. P., and A. H. Hurlbert. 2010. The Combined Influence of the Local Environment and Regional Enrichment on Bird Species Richness. *The American Naturalist* 175:E35–E43.