

The proportion of core species in a community varies with spatial scale and environmental heterogeneity (#30730)

1

First submission

Editor guidance

Please submit by **7 Sep 2018** for the benefit of the authors (and your \$200 publishing discount).



Structure and Criteria

Please read the 'Structure and Criteria' page for general guidance.



Raw data check

Review the raw data. Download from the location [described by the author](#).



Image check

Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the [materials page](#).

8 Figure file(s)



Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. BASIC REPORTING
2. EXPERIMENTAL DESIGN
3. VALIDITY OF THE FINDINGS
4. General comments
5. Confidential notes to the editor






 You can also annotate this PDF and upload it as part of your review

When ready [submit online](#).





Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your [guidance page](#).





BASIC REPORTING

-  Clear, unambiguous, professional English language used throughout.
-  Intro & background to show context. Literature well referenced & relevant.
-  Structure conforms to [PeerJ standards](#), discipline norm, or improved for clarity.
-  Figures are relevant, high quality, well labelled & described.
-  Raw data supplied (see [PeerJ policy](#)).

EXPERIMENTAL DESIGN

-  Original primary research within [Scope of the journal](#).
-  Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
-  Rigorous investigation performed to a high technical & ethical standard.
-  Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

-  Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
-  Data is robust, statistically sound, & controlled.
-  Speculation is welcome, but should be identified as such.
-  Conclusions are well stated, linked to original research question & limited to supporting results.



The best reviewers use these techniques

Tip

Example

Support criticisms with evidence from the text or from other sources

Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.

Give specific suggestions on how to improve the manuscript

Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).

Comment on language and grammar issues

The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 - the current phrasing makes comprehension difficult.

Organize by importance of the issues, and number your points

- 1. Your most important issue*
- 2. The next most important item*
- 3. ...*
- 4. The least important points*

Please provide constructive criticism, and avoid personal opinions

I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC

Comment on strengths (as well as weaknesses) of the manuscript

I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.

The proportion of core species in a community varies with spatial scale and environmental heterogeneity

Molly F Jenkins¹, Ethan P White^{2,3,4}, Allen H Hurlbert^{Corresp. 1,5}

¹ Environment, Ecology, and Energy Program, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States

² Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, Florida, United States

³ Informatics Institute, University of Florida, Gainesville, Florida, United States

⁴ Biodiversity Institute, University of Florida, Gainesville, Florida, United States

⁵ Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States

Corresponding Author: Allen H Hurlbert

Email address: Hurlbert@bio.unc.edu

Ecological communities are composed of a combination of core species that maintain local viable populations and transient species that occur infrequently due to dispersal from surrounding regions. Preliminary work indicates that while core and transient species are both commonly observed in community surveys of a wide range of taxonomic groups, their relative prevalence varies substantially from one community to another depending upon the spatial scale at which the community was characterized and its environmental context. We used a geographically extensive dataset of 968 bird community time series to quantitatively describe how the proportion of core species in a community varies with spatial scale and environmental heterogeneity. We found that the proportion of core species in an assemblage increased with spatial scale in a positive decelerating fashion with a concomitant decrease in the proportion of transient species. Variation in the shape of this scaling relationship between sites was related to regional environmental heterogeneity, with lower proportions of core species at a given scale associated with high environmental heterogeneity. This influence of scale and environmental heterogeneity on the proportion of core species may help resolve discrepancies between studies of biotic interactions, resource availability, and mass effects conducted at different scales, because the importance of these and other ecological processes are expected to differ substantially between core and transient species.

1 **Title: The proportion of core species in a community varies with spatial scale and**
2 **environmental heterogeneity**

3

4

5 Molly F. Jenkins¹, Ethan P. White^{2,3,4}, and Allen H. Hurlbert^{1,5}

6

7 ¹Environment, Ecology, and Energy Program, University of North Carolina, Chapel Hill, NC,
8 27599, USA

9 ²Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL,
10 32611, USA

11 ³Informatics Institute, University of Florida, Gainesville, FL, 32611, USA

12 ⁴Biodiversity Institute, University of Florida, Gainesville, FL, 32611, USA

13 ⁵Department of Biology, University of North Carolina, Chapel Hill, NC, 27599, USA

14

15 Corresponding Author:

16 Allen Hurlbert^{1,5}

17 Email address: Hurlbert@bio.unc.edu

18

19

20

21

22

23 Abstract

24 Ecological communities are composed of a combination of core species that maintain
25 local viable populations and transient species that occur infrequently due to dispersal from
26 surrounding regions. Preliminary work indicates that while core and transient species are both
27 commonly observed in community surveys of a wide range of taxonomic groups, their relative
28 prevalence varies substantially from one community to another depending upon the spatial scale
29 at which the community was characterized and its environmental context. We used a
30 geographically extensive dataset of 968 bird community time series to quantitatively describe
31 how the proportion of core species in a community varies with spatial scale and environmental
32 heterogeneity. We found that the proportion of core species in an assemblage increased with
33 spatial scale in a positive decelerating fashion with a concomitant decrease in the proportion of
34 transient species. Variation in the shape of this scaling relationship between sites was related to
35 regional environmental heterogeneity, with lower proportions of core species at a given scale
36 associated with high environmental heterogeneity. This influence of scale and environmental
37 heterogeneity on the proportion of core species may help resolve discrepancies between studies
38 of biotic interactions, resource availability, and mass effects conducted at different scales,
39 because the importance of these and other ecological processes are expected to differ
40 substantially between core and transient species.

41

42




43

44

45 Introduction

46 Species differ in the temporal persistence with which they occur at any given site. While
47 some species are reliably observed year in and year out, others appear only occasionally (Ulrich
48 and Ollik 2004, Belmaker 2009, Dolan et al. 2009, Gaston et al 2007, Umaña et al. 2017).
49 Indeed, recent work from a broad range of ecological communities has shown that temporal
50 occupancy is typically bimodal, reflecting these two groups which have been referred to as
51 “core” and “transient” species (Coyle et al. 2013, Umaña et al. 2017, Taylor et al. 2018). Core
52 species, in persisting at a site over time, are thought to maintain viable populations through
53 successful reproduction (Coyle et al. 2013, Taylor et al. 2018). In contrast, transients do not
54 persist reliably, and presumably do not maintain viable populations (Magurran and Henderson
55 2003, Umaña et al. 2017). Ecologists have typically ignored this distinction and have assumed
56 that the complete list of species observed over some biological survey constitutes a meaningful
57 “community” of interest for analysis. However, core and transient species interact with their
58 environment in different ways, and in many cases the community of core species may be more
59 relevant for testing theoretical predictions. For example, coexistence theory, niche theory, and
60 other related ideas in ecology are largely predicated upon the occurrence of species that are
61 suited to and influenced by their environments, successfully utilizing those environments for
62 food and reproduction (Umaña et al. 2017). Analyses carried out in communities that support
63 low proportions of core species may poorly align with ecological predictions that are less
64 applicable to transient species. Indeed, previous work has already shown that a wide range of
65 ecological patterns (e.g., species-area relationships, species abundance distributions) differ
66 depending on whether the analysis focuses on core species, transient species, or the entire
67 community (Magurran and Henderson 2003, Taylor et al. 2018). The proportions of core and

68 transient species also vary geographically and therefore influence spatial patterns including
69 species richness gradients (Coyle et al. 2013). Developing general principles regarding the
70 factors that influence the proportion of core species in an assemblage would enable researchers
71 to more effectively compare results between studies and better assess generalities in community
72 ecology.

73 The extent to which a species is a core,  regularly occurring member of an assemblage
74 should depend on the spatial scale over which that assemblage is sampled (Figure 1A). Consider
75 two extremes: at the scale of 1 m², no bird species would maintain a viable population and be
76 observed in every sampling period. At the scale of the entire North American continent, nearly
77 all species would be annually present at least somewhere within that extent. Thus, the proportion
 of core species in an assemblage must increase with scale, **but the functional form of this**
79 relationship is less obvious. We expect the shape of the scaling relationship to be a positive
80 decelerating curve (Figure 1C) because as the extent of a region increases, species that are
81 transient at a local scale will shift to become core species, and the proportion of core species will
82 eventually level **off at or below 1.**  This increase will be moderated to some extent by the
83 inclusion of additional transient species from outside the larger regional spatial extent.

84 Another factor that likely impacts the proportion of core species and the shape of the
85 scaling relationship is environmental heterogeneity, which increases the proportion of transient
86 species likely to occur in an assemblage at a given scale via mass effects (Figure 1B; Coyle et al.
87 2013, Taylor et al. 2018). Mass effects are more likely in heterogeneous landscapes—that is,
88 when surrounding areas differ in habitat from the focal assemblage--as species poorly adapted to
89 the local environment arrive via dispersal from adjacent source habitats to which they are better
90 suited (Shmida and Wilson 1985). Environmental heterogeneity may also constrain habitat

91 availability via the partitioning of space by multiple habitat types within the area delimited by the
92 focal assemblage, and the reduction of area per habitat type relative to environmentally
93 homogeneous sites (Allouche et al. 2012). Resources within each habitat may occur at levels
94 below the threshold needed to sustain viable populations (Allouche et al. 2012), constraining the
95 proportion of core species for fine scale sites compared to a homogeneous habitat of the same
96 size. Both effects of environmental heterogeneity on the proportion of core species in an
97 assemblage are expected to be strongest at smaller spatial scales (Figure 1D). At regional scales,
98 most habitat types will have sufficient resources to sustain viable regional populations and an
99 overall larger proportion of core species. Regardless of the specific mechanism, resource-area
100 tradeoffs or mass effects, we expect heterogeneity will contribute to differences in the shape of
101 the overall relationship between the proportion of core species in an assemblage and spatial
102 scale. While we generally expect this relationship to be positive decelerating as described above,
103 effectively smaller habitat patches in heterogeneous environments may result in the proportion of
104 core species increasing slowly at small scales (Figure 1D). While determining the specific
105 mechanisms of heterogeneity influencing assemblages is beyond the scope of this paper,
106 verifying a connection between heterogeneity and community assembly is a critical first step.



107 Here, we make use of a geographically extensive dataset on bird distribution over time
108 which allows us to investigate temporal occupancy, and hence the proportion of core species in
109 an assemblage, over a wide range of spatial scales and environmental contexts. Specifically, we
110 seek to 1) describe the distribution of species' temporal occupancy in ecological assemblages
111 across a gradient of spatial scales, 2) evaluate the relationship between the proportion of core
112 species in a community and the spatial scale at which that community is characterized, and 3)
113 test whether environmental heterogeneity influences that scaling relationship.

114 **Materials & Methods**

115 **Bird data**



116 We used data on the distribution of diurnal land birds (excluding raptors) over time from
117 the North American Breeding Bird Survey (BBS), maintained by the United States Geological
118 Survey (Pardieck 2017). Our data encompassed the 968 BBS routes across the North American
119 continent that were surveyed continuously over the 15 year period from 2000-2014 that had at
120 least 65 neighboring routes within 1,000 km. Each BBS route is a 40 km roadside transect
121 encompassing fifty 3-minute point count stops, each separated by 0.8 km, in which a single
122 observer records all birds detected within 0.4 km. BBS routes were surveyed each year during
123 the breeding season, typically in June.


124 Temporal occupancy, the proportion of years a species was observed over some spatially
125 defined area, was calculated for each species at each site at a range of spatial scales (Figure 2).
126 We defined the proportion of core species in each assemblage as the proportion of species with
127 temporal occupancy greater than two-thirds (i.e. occurring in at least 11 out of the 15 survey
128 years) following Coyle et al. (2013). We also considered alternative thresholds of temporal
129 occupancy for defining core species (0.5 and 0.75) that produced qualitatively similar results
130 (Figures S1, S2). Below the scale of a single BBS route, each route was split into non-
131 overlapping segments of 5, 10 or 25 point count stops (Figure 2), and the proportion of core
132 species was calculated at each spatial scale. To examine spatial scales greater than a single BBS
133 route, for each focal route we sequentially aggregated survey data from an increasing number of
134 nearest neighbor routes, up to a maximum regional scale of the focal route together with its 65
135 nearest neighbors (Figure 2).

136 Our regional scale of 66 neighboring routes was chosen because it was the number of
137 neighbors that fell within a radius of 1,000 km of each focal route even in regions of lower route
138 density in the western US (Figure 2). The entire range of spatial scales we investigated varied
139 from 2.5 km² for a set of 5 point count stops up to 1,659 km² for an area of 66 adjacent BBS
140 routes. Because BBS route density varies across the continent, the spatial extent of the 65 nearest
141 neighbors did vary (Figure 2). However, using a fixed total number of aggregated routes allowed
142 us to keep the total surveyed area characterizing an assemblage constant  and this was the aspect
143 of scale we viewed as most critical for our comparisons. **While regions of the same sampled area
144 but spanning larger extents may encompass a greater range of environmental variation all else
145 equal, we measured this variation directly (see Environmental Data below).** 

146 In addition to spatial scale, we used the total number of individuals observed in the
147 assemblage (community size) as an alternative measure of scale. Community size was found to
148 be a potentially more generalizable measure of scale than area, especially for comparing between
149 taxonomic groups with very different area requirements (Taylor et al. 2018).


150 **Scaling metrics**

151 We derived a series of metrics characterizing the relationship between the proportion of
152 core species present and scale for each focal route (Figure 1C). We identified the proportion of
153 core species at the smallest scale (p_{min}) and the proportion at the largest scale (p_{max}) for each focal
154 route.  **also identified the slope of the line linking p_{min} and p_{max} for each focal route.** We
155 identified the scale at which the proportion of core species in the community surpassed the
156 threshold of 0.5 for each focal route ($scale_{50}$). Finally, we characterized the degree of curvature
157 in the relationship between the proportion of core species in the community and scale. As a
158 measure of curvature, we estimated the area between the  **erved scaling curve** and the straight

159 line linking p_{min} and p_{max} by summing the differences between the observed values and the values
160 ected from the linear relationship (Figure 1C). Positive values indicate positive decelerating
161 relationships and greater proportions of core species, while negative values indicate positive
162 accelerating relationships and lower proportions of core species relative to a linear relationship.

163 **Environmental data**



164 We acquired raster layers for 0.25 km resolution elevation from Worldclim (Fick &
165 Hijman 2017), and 0.25 km resolution Normalized Difference Vegetation Indices (NDVI) from
166 the NASA GIMMS group (Didan 2015), and calculated mean NDVI and mean elevation for each
167 focal route within a 40 km buffer of the route's starting coordinates. For each environmental
168 variable, we defined regional heterogeneity around each focal route as the variance in mean
169 values across the set of 65 nearest neighbor BBS routes plus the focal route. In order to assess the
170 whether the importance of environmental heterogeneity varied with the spatial scale over which
171 heterogeneity was measured, we also calculated environmental heterogeneity at different scales


 (from 3 to 66 neighboring routes). **We then examined the Pearson's correlation between**
173 heterogeneity and the five scaling metrics describing how the proportion of core species varies
174 across the full range of spatial scales.

175 **Results**




176 At the scale of a single route ($\sim 25 \text{ km}^2$), temporal occupancy was bimodal as expected
177 (Figure 3, dashed line). At larger spatial scales, assemblages were marked by a greater
178 proportion of core species with high temporal occupancy, while at smaller scales, assemblages
179 were characterized by a greater number of transient species and very few core species (Figure 3).
180 The proportion of core species in a community increased on average in a positive decelerating
181 manner with both measures of spatial scale, although there was substantial variability from route



182 to route (Figure 4A). At the largest spatial scales, the proportion of core species exhibited
183 reduced variation, with a mean of 83% and ranging from 75%-90%, while at the smallest spatial
184 scales (2.5 km²) the proportion of core species varied from 11-37%. Using community size in
185 lieu of spatial scale greatly reduced this variation in the proportion of core species at the smallest
186 scale (Figure 4B).

187 Heterogeneity in elevation and heterogeneity in NDVI both  similar effects on the
188 overall shape of the relationship between the proportion of core species and spatial scale,
 although the effects of elevation were stronger for some measures (Figure 5). **Environmentally**
190 **heterogeneous regions had assemblages with a low proportion of core species at both the**
191 **smallest and largest scales, and communities that experienced the greatest increase in the**
192 **proportion of core species between the smallest and largest scales. Assemblages in more**
193 **heterogeneous regions additionally displayed less positive curvature values and a larger spatial**
194 **scale at which the majority of species were identified as core.**

195 The scale at which environmental heterogeneity was measured also affected the strength
196 of the correlation between heterogeneity and scaling curve metrics (Figure 6). Specifically,
 heterogeneity in elevation was **most strongly correlated with all five of the scaling metrics when**
198 **measured at the largest spatial scale**, whereas heterogeneity in NDVI exhibited the strongest
199 correlations with p_{min} , $scale_{50}$, and $slope$ parameters at scales between 15-25 BBS routes (400-
200 600 km²; Figure 6). With the exception of *curvature*, heterogeneity in NDVI was a stronger
201 correlate of our scaling metrics than heterogeneity in elevation at these intermediate scales.

202 **Discussion**


203 Ecologists frequently test hypotheses regarding community assembly and species
204 richness using surveys that reflect a snapshot of a community at a particular point in time.
205 However, it is increasingly recognized that such a snapshot approach fails to differentiate core
206 species from transient species, the former maintaining viable populations and interacting more
207 strongly with their biotic and abiotic environment, and the latter being irregular visitors that are
208 presumably better adapted to other conditions (Magurran and Henderson 2003, White and
209 Hurlbert 2010, Umaña et al 2017). We used a continent-wide dataset on bird assemblages over
210 time to evaluate how the proportion of core species in these assemblages increases with scale and
211 decreases with environmental heterogeneity. Consistent with Coyle et al. (2013), the distribution
212 of temporal occupancy was strongly bimodal at the scale of a single BBS route, reflecting these
213 two distinct groups. However, at scales below the size of a BBS route (<25 km²) few species
214 were present consistently over time, while at scales larger than two aggregated BBS routes (>50
215 km²) most species occurred regularly. The  **smallest scale** assemblages exhibited a fairly wide
 range in the proportion of core species present (11-37%), **at least in part because different sites**
217 differed in the overall number of individuals supported. At the largest spatial scales (1,659 km²
218 of surveyed area distributed across a 1,000 km radius region) there was less variation in the
219 proportion of core species (75-90%). On average, the proportion of core species in a community
220 increased in a positive decelerating manner as a function of spatial scale. As scale increased, so
221 did the probability of including suitable habitat in sufficient quantities to support persistent
222 populations, and species identified as transient at small scales subsequently became core species
 at larger scales. However, **even at the largest scales considered here transient species made up**
224 10% or more of the species observed.

225 Much of the variation in the shape of the relationship between the proportion of core
226 species in a community and spatial scale can be explained by the regional environmental
227 heterogeneity surrounding the assemblage. Specifically, landscapes with high environmental
228 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**
230 positively correlated with the proportion of transient species (Coyle et al. 2013, Taylor et al.
231 2018). 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**
233 regional scales (Figure 6). This is likely because variation in elevation encompasses habitat
234 diversity due to the inclusion of different zones of elevation in addition to differences in slope,
235 hydrology, and other topographic features. Variation in NDVI also presumably captures many of
236 these differences, but perhaps less directly as the habitat variation within a given range of NDVI
237 may not be well captured. Ultimately, regional heterogeneity increases the relative proportion of
238 transient species at local scales via the increased likelihood of mass effects by species better
239 adapted to adjacent habitat types (Shmida and Wilson 1985, Coyle et al. 2013, Taylor et al.
240 2018). Landscapes with low environmental heterogeneity should support communities with low
241 temporal turnover (Stegen et al. 2013, Gaston et al. 2007), even at small spatial scales nested
242 within the region as these small scale habitats more closely parallel the resources and
243 composition of the region they occur within. Landscapes with a high degree of environmental
244 heterogeneity are more spatially compartmentalized, effectively decreasing the area and
245 resources available per habitat type to support a viable species population (Allouche et al 2012).
246 Thus, in addition to experiencing greater mass effects, any particular habitat type within a

247 heterogeneous region is less likely to encompass sufficient area and resources necessary to
248 sustain viable populations.

249 These relationships between the proportion of core species and both scale and
250 environmental heterogeneity may help resolve discrepancies between studies regarding the
251 importance of biotic interactions, resource availability, and mass effects for driving community
252 assembly (Henderson and Magurran 2014). Difficulties in synthesizing and generalizing across
253 studies may have arisen from differences in scale and environmental heterogeneity leading to
254 assemblages with different proportions of core species and therefore different apparent
255 mechanisms driving community assembly (e.g., Dorazio et al. 2006, Emerson & Gillespie 2008,
256 Stein et al. 2015). For example, competition and environmental filtering have both been proposed
257 to shape community assembly and influence phylogenetic overdispersion and clustering
258 (Cavender-Bares et al. 2004, Mayfield and Levine 2010). However, the degree of overdispersion
259 or clustering may also be affected by the proportion of core or transient species in a community.
260 Core species are more likely to compete with each other for resources, and would be expected to
261 contribute the most to overdispersion in competition related traits. In addition, core species are
262 expected to be better suited to the local climate or habitat compared to transient species, and so
263 would be expected to exhibit greater clustering of environmental tolerance traits. At small spatial
264 scales, the proportion of transient species will be higher, resulting in a lower likelihood of
265 discerning a nonrandom assembly pattern. The proportion of core species is lowest at small
266 scales, and yet the processes driving core species assembly, like competition, should be most
267 important at these scales where individuals are more likely to interact (Allouche et al. 2012).
268 This may result in seemingly conflicting, or altogether masked, patterns of community assembly
269 in large meta-analyses that include studies conducted at a wide range of scales from disparate

270 taxonomic groups. When testing for aspects of community structure, restricting the analysis to
271 core species should increase the power to detect non-random trait assembly patterns and improve
272 the search for generality.

273 Macroecological analyses of core and transient species use observational time-series to
274 identify these two groups. While this is the only practical way to accomplish this classification at
275 scale (considering thousands of species-site combinations), it can result in two types of
276 classification errors: species may be inferred to be transient when they are core (a false negative),
277 and they may be inferred to be core when they are transient (a false positive). False negatives
278 lead to underestimates of the proportion of core species, and they are expected to occur primarily
279 at intermediate spatial scales. At small scales, few species actually maintain viable populations
280 and nearly all species are truly transient. At large scales, even species that occur at low density
281 will reliably be observed somewhere from year to year, and so nearly all species are y core.



282 The fact that false negatives will be most common at intermediate scales implies that the “true”
283 curve scaling the proportion of core species with area or community size has similar p_{min} , p_{max} ,
284 and slope values to the observed curve. A higher proportion of core species at intermediate scales
285 would reduce the scale at which that proportion exceeded 0.5 ($scale_{50}$) and potentially increase
286 estimates of curvature. False positives are expected to occur primarily at small scales in regions
287 of high environmental heterogeneity. A species that does not sustain a viable population at a
288 local sink site but does in the surrounding region may appear to be a core species at that sink site
289 because neighboring sites support sufficient populations to ensure regular immigration to the
290 sink site. However, the fact that environmental heterogeneity had a negative effect on the
291 observed proportion of core species implies that this bias is minimal. Future research using
292 simulation models to assess misclassification rates for communities across different scales and

293 levels of environmental heterogeneity, and for species with different densities and detection
294 rates, will be necessary for evaluating the extent to which spatial scales and heterogeneity
295 influence classification errors. Alternatively, using stricter thresholds of temporal occupancy for
296 determining the proportion of core species may help reduce the likelihood of false positives
297 (Figures S1, S2).

298 **Conclusions**

299 The distinction between core and transient species is increasingly recognized as being
300 important for properly testing predictions and comparing ecological systems (Magurran and
301 Henderson 2003; Coyle et al. 2013; Supp et al. 2015; Umaña et al. 2017; Taylor et al. 2018),
302 making it critical to understand the factors that influence the relative proportion of these two
303 different groups. Here, we have shown that the proportion of core species in an assemblage is
304 positively associated with spatial scale and negatively associated with environmental
305 heterogeneity. The relative proportion of these two groups of species influences a number of
306 essential patterns in community ecology, including the species-area relationship, species-
307 abundance distribution, temporal turnover, and geographic patterns of biodiversity (Magurran
308 and Henderson 2003; Taylor et al. 2018). All of these patterns are scale-dependent, and
309 investigators have typically assumed an effect of scale itself (Adler et al. 2005, Rahbek 2005,
310 Green and Plotkin 2007). Our results suggest an extra layer of complexity in that scale influences
311 the proportion of core and transient species which may influence ecological patterns independent
312 of scale. **Future work attempting to understand the different ways in which scale influences**
313 **ecological systems should consider this indirect influence of scale.** In general, an understanding
314 of the factors that influence the prevalence of core species is critical for the proper interpretation
315 of synthetic meta-analyses and the evaluation of ecological theory.



316 **Acknowledgements**

317 We are grateful to S. Taylor and J. Coyle for comments that improved an earlier version of this
318 manuscript, and to the thousands of volunteers who annually contribute to data collection for the
319 North American Breeding Bird Survey.

320 **References**

- 321 Adler PB., White EP., Lauenroth WK., Kaufman DM., Rassweiler A., Rusak JA. 2005. Evidence
322 for a General Species–Time–Area Relationship. *Ecology* 86:2032–2039. DOI: 10.1890/05-
323 0067.
- 324 Allouche O., Kalyuzhny M., Moreno-Rueda G., Pizarro M., Kadmon R. 2012. Area–
325 heterogeneity tradeoff and the diversity of ecological communities. *Proceedings of the*
326 *National Academy of Sciences* 109:17495–17500. DOI: 10.1073/pnas.1208652109.
- 327 Belmaker J. 2009. Species richness of resident and transient coral-dwelling fish responds
328 differentially to regional diversity. *Global Ecology and Biogeography* 18:426–436. DOI:
329 10.1111/j.1466-8238.2009.00456.x.
- 330 Brown JH., Kodric-Brown A. 1977. Turnover Rates in Insular Biogeography: Effect of
331 Immigration on Extinction. *Ecology* 58:445–449. DOI: 10.2307/1935620.
- 332 Cavender-Bares J., Ackerly DD., Baum DA., Bazzaz FA. 2004. Phylogenetic overdispersion in
333 Floridian oak communities. *The American Naturalist* 163:823–843. DOI: 10.1086/386375.

- 334 Coyle JR., Hurlbert AH., White EP. 2013. Opposing Mechanisms Drive Richness Patterns of
335 Core and Transient Bird Species. *The American Naturalist* 181:E83–E90. DOI:
336 10.1086/669903.
- 337 Didan, K. 2015. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid
338 V006. DOI: 10.5067/MODIS/MOD13Q1.006.
- 339 Dolan JR., Ritchie ME., Tunin-Ley A, Pizay MD 2009. Dynamics of core and occasional species
340 in the marine plankton: tintinnid ciliates in the north-west Mediterranean Sea. *Journal of*
341 *Biogeography* 36:887–895. DOI: 10.1111/j.1365-2699.2008.02046.x.
- 342 Dorazio RM., Royle JA, Söderström B, Glimskär A. 2006. Estimating species richness and
343 accumulation by modeling species occurrence and detectability. *Ecology* 87:842–854. DOI:
344 10.1890/0012-9658(2006)87[842:ESRAAB]2.0.CO;2.
- 345 Emerson BC., Gillespie RG. 2008. Phylogenetic analysis of community assembly and structure
346 over space and time. *Trends in Ecology & Evolution* 23:619–630. DOI:
347 10.1016/j.tree.2008.07.005.
- 348 Fick SE., Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for
349 global land areas: NEW CLIMATE SURFACES FOR GLOBAL LAND AREAS.
350 *International Journal of Climatology* 37:4302–4315. DOI: 10.1002/joc.5086.
- 351 Gaston KJ., Davies RG., Orme CDL., Olson VA., Thomas GH., Ding TS., Rasmussen PC.,
352 Lennon JJ., Bennett PM., Owens IP., Blackburn TM. 2007. Spatial turnover in the global

- 353 avifauna. *Proceedings of the Royal Society B: Biological Sciences* 274:1567–1574. DOI:
354 10.1098/rspb.2007.0236.
- 355 Green JL., Plotkin JB. 2007. A statistical theory for sampling species abundances. *Ecology*
356 *Letters* 10:1037–1045. DOI: 10.1111/j.1461-0248.2007.01101.x.
- 357 Henderson PA., Magurran AE. 2014. Direct evidence that density-dependent regulation
358 underpins the temporal stability of abundant species in a diverse animal community.
359 *Proceedings of the Royal Society B: Biological Sciences* 281. DOI: 10.1098/rspb.2014.1336.
- 360 Hurlbert AH. 2004. Species–energy relationships and habitat complexity in bird communities.
361 *Ecology Letters* 7:714–720. DOI: 10.1111/j.1461-0248.2004.00630.x.
- 362 Lowe JD. 2006. An annotated bibliography of Breeding Bird Census publications. *Bird*
363 *Populations* 7:128–135.
- 364 MacArthur RH. 1957. On the Relative Abundance of Bird Species. *Proceedings of the National*
365 *Academy of Sciences of the United States of America* 43:293–295.
- 366 Magurran AE., Henderson PA. 2003. Explaining the excess of rare species in natural species
367 abundance distributions. *Nature* 422:714–716. DOI: 10.1038/nature01547.
- 368 Magurran AE., McGill BJ. (eds.) 2011. *Biological diversity: frontiers in measurement and*
369 *assessment*. Oxford ; New York: Oxford University Press.

- 370 Pardieck KL., Ziolkowski DJ., Lutmerding M., Campbell K., Hudson M-AR. 2017. North
371 American Breeding Bird Survey Dataset 1966 - 2016, version 2016.0. DOI:
372 10.5066/F7W0944J.
- 373 Rahbek C. 2005. The role of spatial scale and the perception of large-scale species-richness
374 patterns. *Ecology Letters* 8:224–239. DOI: 10.1111/j.1461-0248.2004.00701.x.
- 375 Shmida A., Wilson MV. 1985. Biological Determinants of Species Diversity. *Journal of*
376 *Biogeography* 12:1–20. DOI: 10.2307/2845026.
- 377 Stegen JC., Freestone AL., Crist TO., Anderson MJ., Chase JM., Comita LS., Cornell HV.,
378 Davies KF., Harrison SP., Hurlbert AH., Inouye BD., Kraft NJB., Myers JA., Sanders NJ.,
379 Swenson NG., Vellend M. 2013. Stochastic and deterministic drivers of spatial and temporal
380 turnover in breeding bird communities. *Global Ecology and Biogeography* 22:202–212. DOI:
381 10.1111/j.1466-8238.2012.00780.x.
- 382 Stein A., Gerstner K., Kreft H. 2014. Environmental heterogeneity as a universal driver of
383 species richness across taxa, biomes and spatial scales. *Ecology Letters* 17:866–880. DOI:
384 10.1111/ele.12277.
- 385 Supp SR., Koons DN., Ernest SKM. 2015. Using life history trade-offs to understand core-
386 transient structuring of a small mammal community. *Ecosphere* 6:art187. DOI: 10.1890/ES15-
387 00239.1.
- 388 Taylor SJS., Evans BS., White EP., Hurlbert AH. 2018. The prevalence and impact of transient
389 species in ecological communities. *Ecology* 99:1825–1835. DOI: 10.1002/ecy.2398.

390 Ulrich W., Ollik M. 2004. Frequent and occasional species and the shape of relative-abundance
391 distributions. *Diversity and Distributions* 10:263–269. DOI: 10.1111/j.1366-
392 9516.2004.00082.x.

393 Umaña Maria Natalia, Zhang Caicai, Cao Min, Lin Luxiang, Swenson Nathan G., Suding
394 Katharine 2017. A core-transient framework for trait-based community ecology: an example
395 from a tropical tree seedling community. *Ecology Letters* 20:619–628. DOI:
396 10.1111/ele.12760.

397 White EP., Hurlbert AH. 2010. The Combined Influence of the Local Environment and Regional
398 Enrichment on Bird Species Richness. *The American Naturalist* 175:E35–E43. DOI:
399 10.1086/649578.

Figure 1

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.

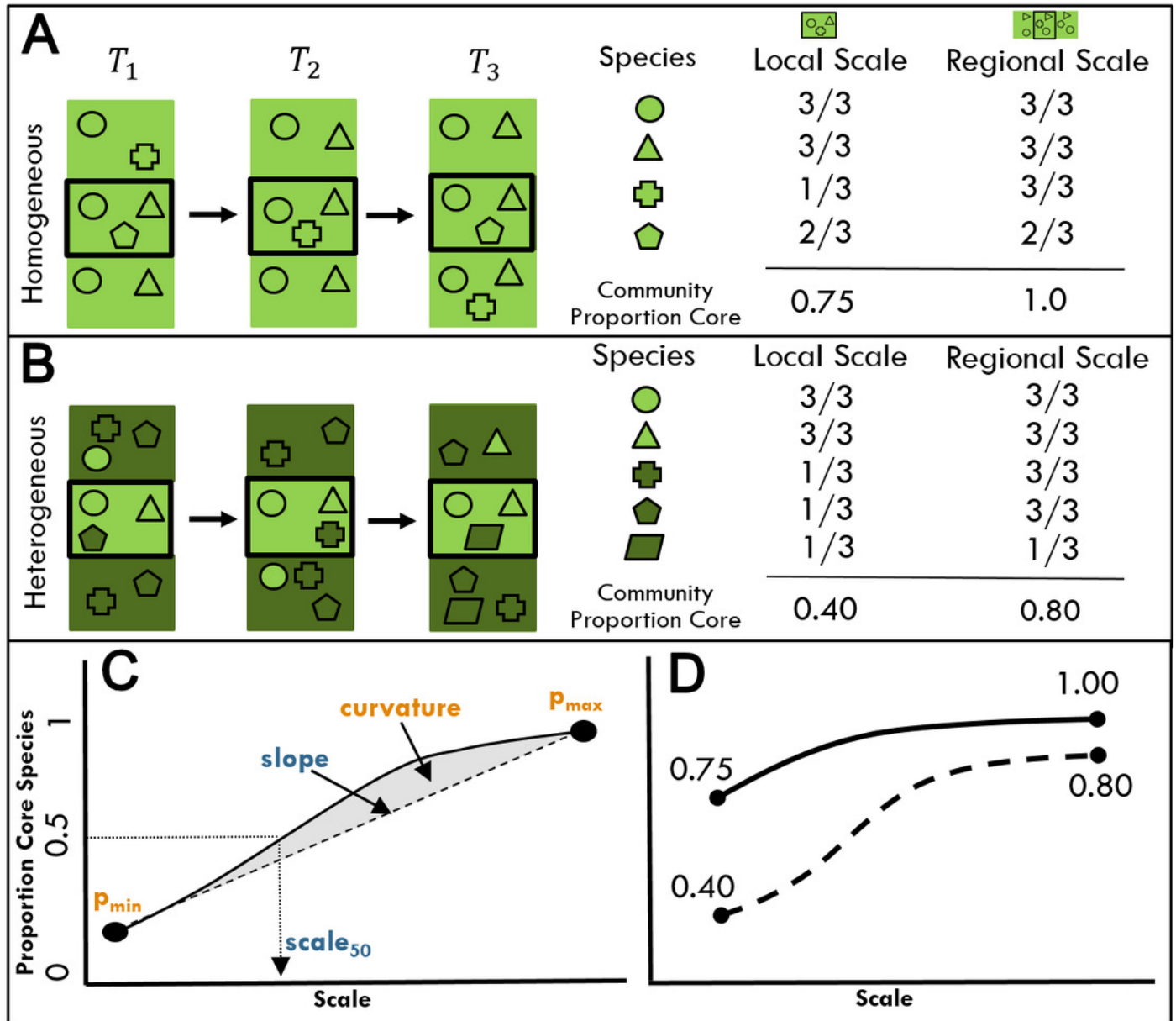


Figure 2(on next page)

Distribution of bird communities and range of spatial scales examined for calculating temporal occupancy and the proportion of core species.

Map of North America shows the 968 Breeding Bird Survey routes used in this study, including two examples of the maximum scale examined: 66 survey routes aggregated together, which span variable extents depending on route density. The inset shows a single survey route made up of 50 point count stops, and the spatial scales examined below the level of a route.

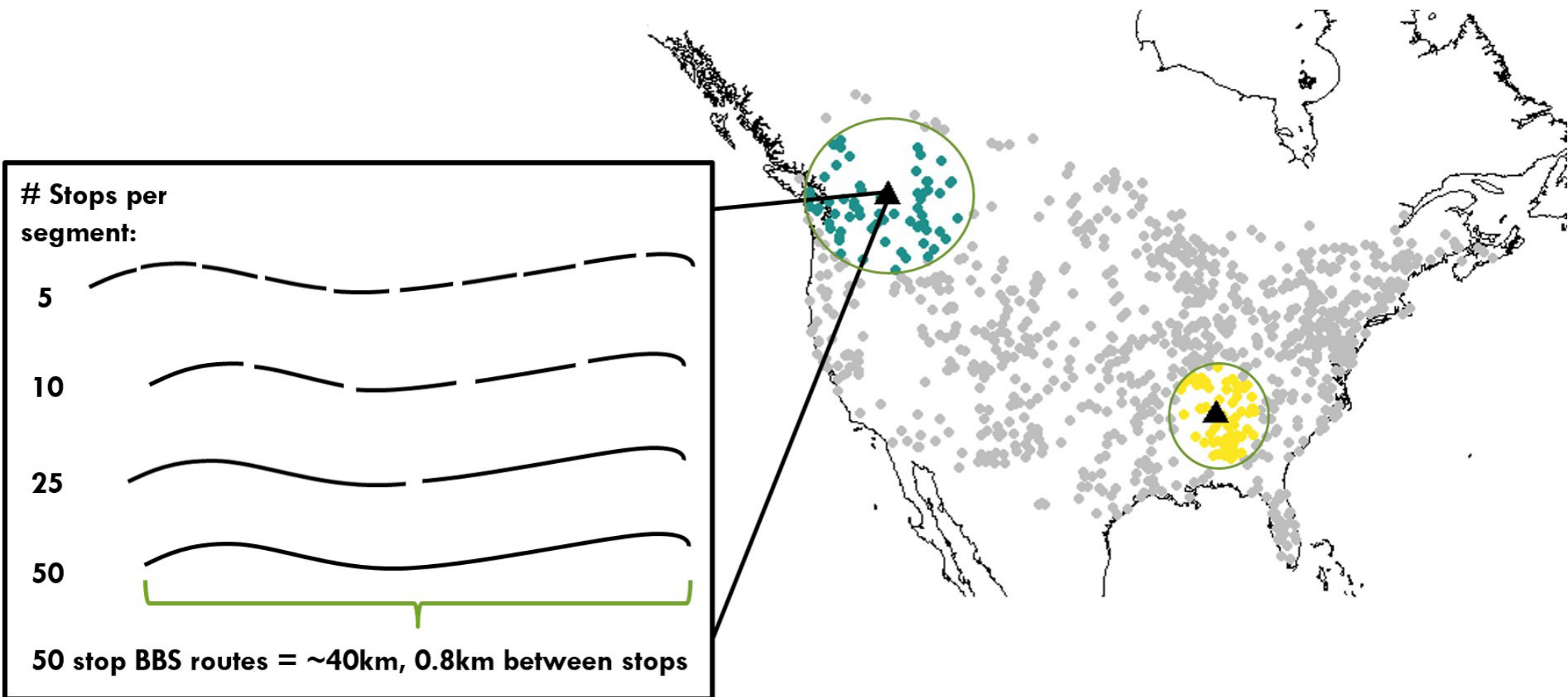


Figure 3

Average probability densities of temporal occupancy for the bird species present at a



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.

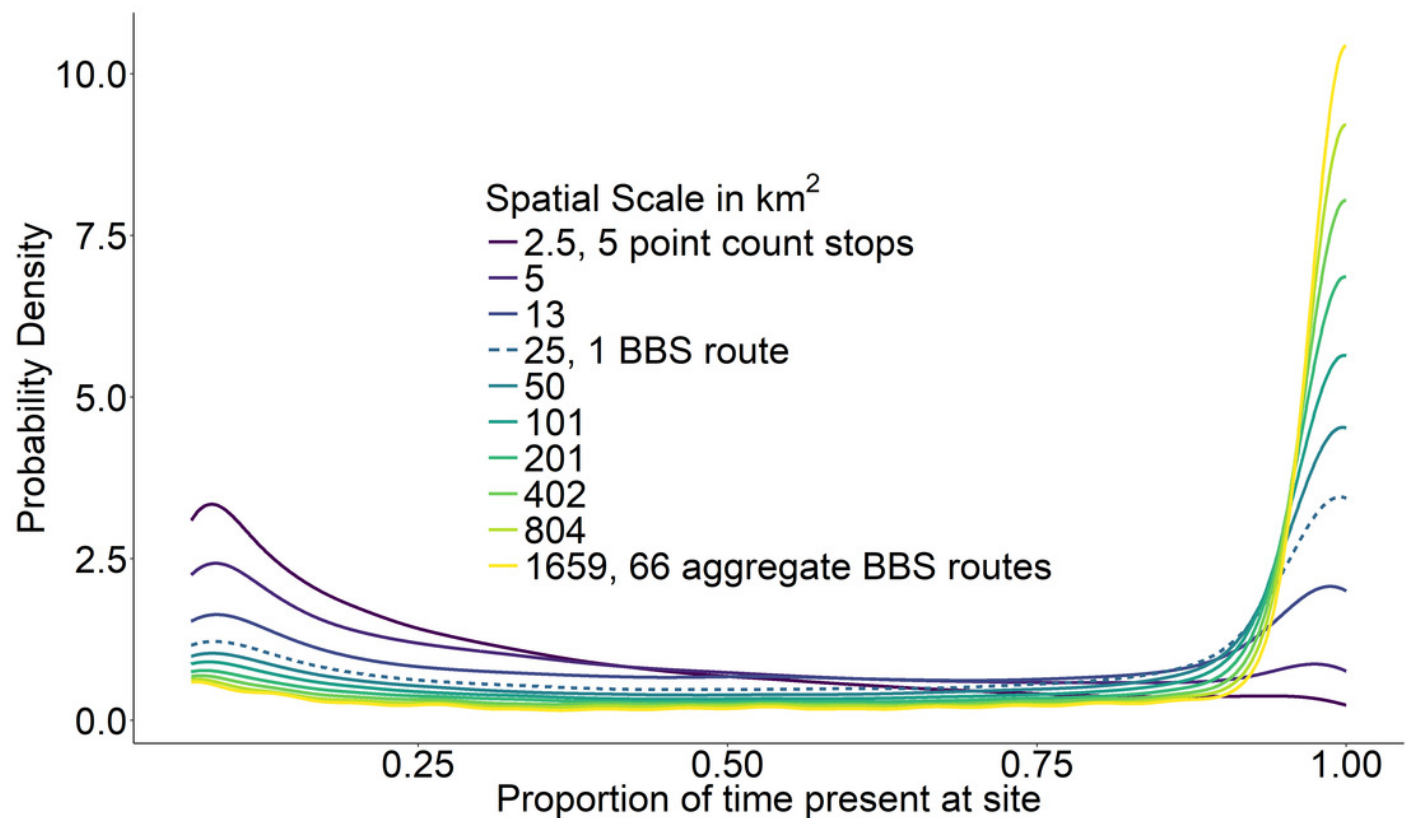


Figure 4

Proportion of core species present in assemblages as a function of (a) scale as measured by area and (b) scale as measured by community size.

Each line represents a single focal BBS route; we examined 968 routes total. Average across all BBS routes indicated by the bold black line. Highlighted routes exemplify low environmental heterogeneity (purple, Illinois, route 54) and high environmental heterogeneity (orange, Utah, route 169).

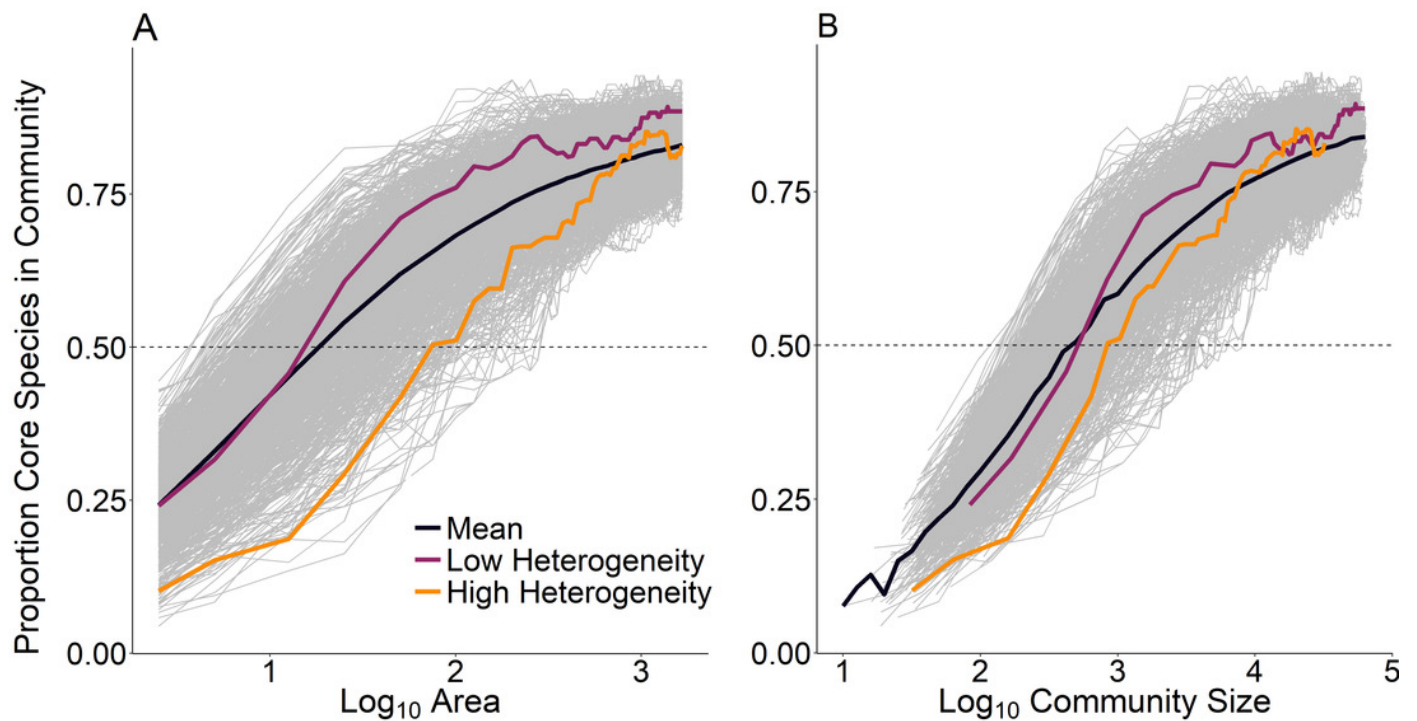


Figure 5

Correlation between two measures of regional environmental heterogeneity

and five parameters describing how the proportion of core species increases with scale.

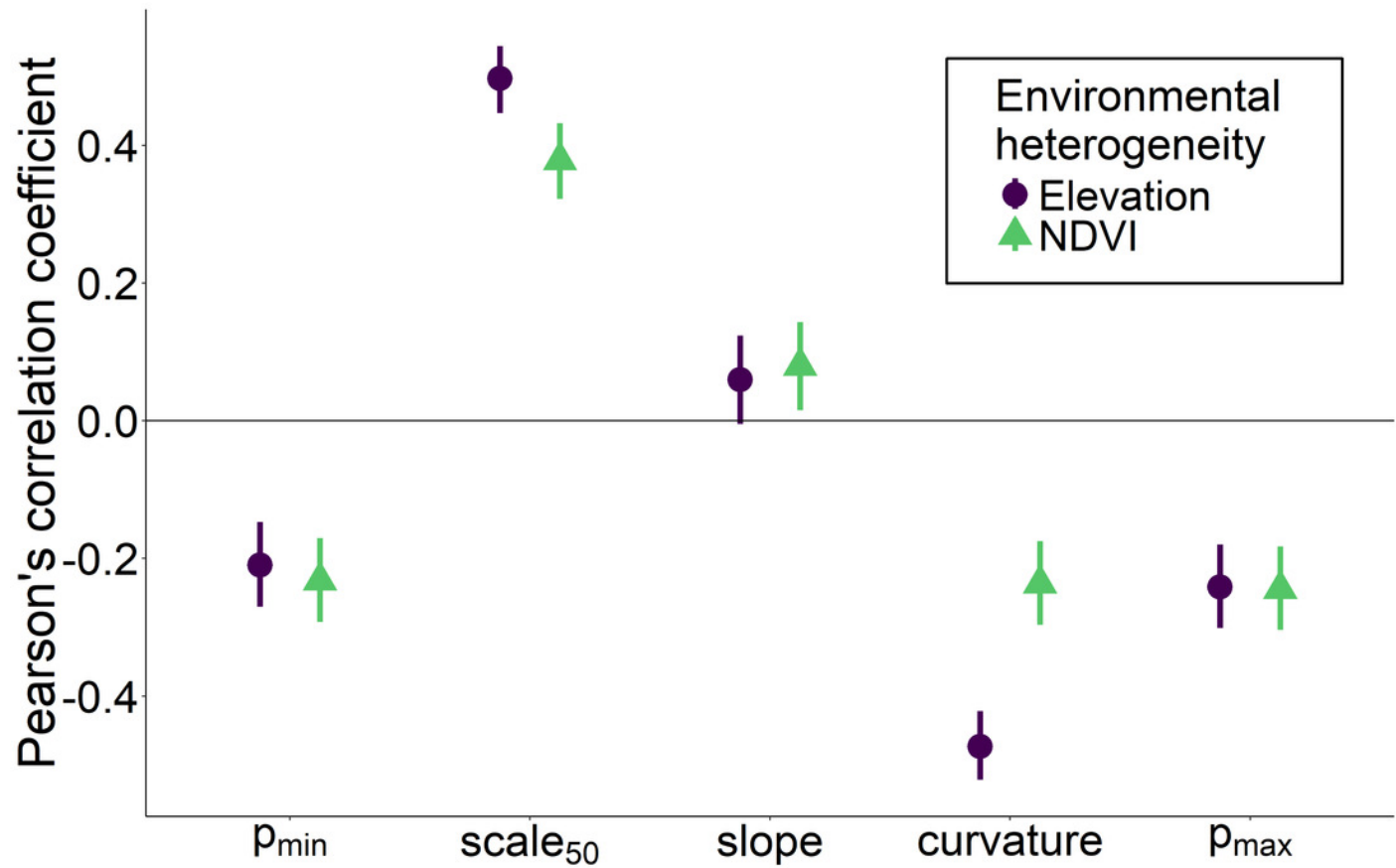


Figure 6

Correlation between two measures of environmental heterogeneity and five parameters describing how the proportion of core species increases with scale as a function of the spatial scale over which environmental heterogeneity was characterized.

