

Effects of occupancy estimation on abundance-occupancy relationships

Cleber Ten Caten^{a,*}, Lauren A. Holian^a and Tad Dallas^a

^aDepartment of Biological Sciences, University of South Carolina,
Columbia, South Carolina, 29205

*Corresponding author: clebertencaten@gmail.com

Abstract

Abundance-occupancy relationships predict that species that occupy more sites are also more locally abundant, where occupancy is usually estimated following the assumption that species can occupy all sampled sites. Here we use the National Ecological Observatory Network small-mammal data to assess whether this assumption affects abundance-occupancy relationships. We estimated occupancy considering all sampled sites (traditional occupancy) and only the sites found within the species geographic range (spatial occupancy) and realized environmental niche (environmental occupancy). We found that when occupancy was estimated considering only sites possible for the species to colonize (spatial and environmental occupancy) weaker abundance-occupancy relationships were observed. This shows that the assumption that the species can occupy all sampled sites directly affects the assessment of abundance-occupancy relationships. Estimating occupancy considering only sites that are possible for the species to colonize will consequently lead to a more robust assessment of abundance-occupancy relationships.

Running title: Abundance-occupancy in NEON

Funding: This work has been performed with funding to Tad Dallas from the National Science Foundation (NSF-DEB-2017826) *Macrosystems Biology and NEON-Enabled Science* program

Acknowledgements: We would like to thank the NEON data providers.

Data accessibility: *R* code and data to reproduce the analyses are available on figshare at

<https://doi.org/10.6084/m9.figshare.19323644>.

Conflict of interest: The authors have no conflicts of interest to declare.

Keywords: National Ecological Observatory Network, abundance-occupancy, geographic range, environmental niche

Introduction

1 Positive abundance-occupancy relationships – the observation that widely dis-
2 tributed species are also more locally abundant – is a general pattern in ecology
3 [1] that has been described for vertebrates [2, 3, 4], invertebrates [5, 6, 7], plants
4 [8, 9, 10], and bacteria [11]. Resource availability [12, 13], species niche require-
5 ments, and dispersal limitation [5, 14] are among the mechanisms proposed to
6 explain these positive relationships [12]. Although these mechanisms are usually
7 evaluated individually, they can affect species occupancy and abundance simulta-
8 neously [15], with the relative importance of each mechanism being dependent on
9 spatial scale [1]. Moreover, biotic and abiotic factors [16] as well as stochastic dy-
10 namics [17] also affect species abundance and occupancy patterns. This combined
11 effect of different factors affecting species abundance and occupancy might explain
12 why some taxa do not show positive abundance-occupancy relationships [18, 19]
13 as well as why these positive relationships are usually weak [1, 4].

14 Abundance-occupancy relationships can be evaluated at small or large spatial
15 scales [1], where occupancy is usually defined as the number or fraction of sites
16 where a species occurs out of the full set of sampled sites [20, 21]. Thus, the spatial
17 scale sampled in a study can directly impact occupancy estimations. For example,
18 although the major assumption that the species can occupy all sampled sites af-
19 fects the occupancy estimation for all species, species with small geographic ranges
20 are particularly affected by this assumption as they will inherently have exception-
21 ally lower occupancy estimates when large spatial scales are sampled. However,
22 a species occurrence at a site is affected by environmental conditions, dispersal
23 limitation, and biotic interactions [22, 23]. The species environmental niche plays
24 an important role on its ability to occupy sites [24], such that a species can only
25 occupy sites that have environmental conditions that it can tolerate [25, 26]. Con-
26 sequently, species environmental niche breadth is positively related to geographic
27 range size [27, 28] and occupancy [29]. Nevertheless, dispersal limitation [30, 31, 32]
28 and biotic interactions [22, 33] can still prevent a species from occupying environ-
29 mentally suitable sites.

30 Knowledge on the species geographic range can be used to estimate occupancy
31 given that in some cases environmentally suitable sites might be geographically
32 inaccessible for the species occurrence because of dispersal barriers [34] such as
33 mountains and rivers [35, 36]. Similarly, information on the species realized en-
34 vironmental niche (i.e. the set of environmental conditions in which the species
35 was found) can also be used to estimate occupancy as environmental conditions
36 might be unsuitable in parts of the species geographic range [37]. For example,
37 fragmentation processes occurring in parts of the species geographic range could
38 lead to changes in environmental conditions in those locations and render them to
39 be environmentally unsuitable for the species occurrence [38]. Thus, information
40 on both the species realized environmental niche and geographic range can be used
41 to estimate occupancy considering only sites that are possible for the species to
42 colonize, a factor that is often ignored when abundance-occupancy relationships
43 are assessed [10].

44 A challenging aspect of estimating species realized environmental niche and ge-
45 ographic range is obtaining enough occurrence points for the species such that its
46 geographic range and realized environmental niche can be confidently estimated.
47 The development of online databases, such as the Global Biodiversity Information
48 Facility (GBIF), where species occurrence points are made publicly available, pro-
49 vide an opportunity to overcome this problem. Here, we use occurrence points
50 obtained from GBIF to estimate the geographic range and realized environmen-
51 tal niche of 122 North American mammal species that have abundance and oc-
52 currence data available in the National Ecological Observatory Network dataset
53 [39, 40]. We use this information on species geographic range and realized envi-
54 ronmental niche to estimate spatial occupancy, and explore how this influences
55 the assessment of abundance-occupancy relationships. Occupancy was estimated
56 as the fraction of all sampled occupied sites, as the fraction of environmentally
57 suitable occupied sites, and as the fraction of occupied sampled sites within the
58 species geographic range. The occurrence of interspecific abundance-occupancy
59 relationships (i.e., the assessment of the scaling between species mean abundance

60 and occupancy) was assessed using these three occupancy estimates. We found
61 that the observed abundance-occupancy relationships became weaker when oc-
62 cupancy estimates were constrained by the species realized environmental niche
63 or geographic range. This occurred because species with small geographic ranges
64 have their occupancy underestimated when it is measured following the traditional
65 approach.

66 Methods

67 **Species abundance and occupancy data** We used the National Ecological
68 Observatory Network (NEON) small mammal data sampled between 2014 and
69 2019. NEON is a continental research platform where occurrence and density data
70 is collected for small mammals in 46 terrestrial sites spread over 20 ecoclimatic
71 domains across the U.S. [39, 40]. Several 10×10 trap grids (plots) are used per site
72 to sample mammals. Each of the 100 traps present in the plots are separated by
73 10m. Although the number of traps is standardized for each plot, there can be 6
74 different types of trap status depending on the sampling outcome. Only traps that
75 had captures or no captures (i.e. trap status 4-6) were used to calculate species
76 abundance. Traps not set, disturbed or with trap door open or closed with feces
77 left behind or with bait missing (i.e. trap status 1-3) were not considered in our
78 analyses. Moreover, individuals recaptured in the same month were not considered
79 when calculating species abundance. Abundance and occurrence data were only
80 obtained for individuals that were identified to the species level ($n = 122$).

81 **Estimating the species realized environmental niche and geographic**
82 **range** Species geographic ranges were estimated with minimum convex poly-
83 gons from occurrence points sampled in the United States obtained from the GBIF
84 database [41]. To estimate the species realized environmental niche we used the 19
85 bioclimatic variables available in the BioClim database [42] at a resolution of 10
86 arc-minutes covering the Americas and performed a Principal Component Anal-
87 ysis (PCA). The first two axes explained more than 80% of the variance in the

88 data, and were selected to estimate the species realized environmental niche. We
89 extracted the environmental values associated with the species occurrence points
90 found in the Americas from the two PCA axes and used minimum convex polygons
91 to estimate the species realized environmental niche (see supplemental material for
92 more details).

93 **Abundance and occupancy estimation** We estimated mean annual abundance
94 as the mean abundance across sampling months and sites, standardizing
95 monthly estimates of abundance based on the number of trapnights. Mean annual
96 occupancy was calculated in three different ways. First, we estimated occupancy
97 using the traditional approach, where occupancy was defined as the number of sites
98 where a species was found divided by the number of total sampled sites, hereafter
99 traditional occupancy. In this case, all sampled sites are used to calculate the
100 species occupancy regardless of whether the sites are suitable for the species oc-
101 currence. An extreme example of a case like this would be estimating occupancy
102 considering sites that do not have the required habitat for the species occurrence.
103 For the second and third cases, we only considered sites found within the species
104 realized environmental niche and known geographic range to estimate occupancy,
105 hereafter environmental and spatial occupancy respectively. Abundance and oc-
106 cupancy estimates were weighted according to the annual number of sites sampled
107 for each species.

108 **Evaluating abundance-occupancy relationships** We assessed the abundance-
109 occupancy relationship using an interspecific approach that evaluates the general-
110 ity of the scaling between species abundance and occupancy across species. Spear-
111 man's rank correlation was used to assess the correlation between the species \log_{10}
112 mean abundance and the three different occupancy metrics estimated.

113 **Results**

114 **How different are the estimated occupancies?** The mean fraction of occu-
115 pied sites by the species was the lowest for traditional occupancy (mean \pm sd; 0.07 \pm 0.09)

116 followed by environmental occupancy (0.17 ± 0.15) and it was the highest for spa-
117 tial occupancy (0.32 ± 0.27 ; figure 1a-c). Traditional occupancy estimates were
118 lower because it considered all sites when occupancy was estimated whereas spa-
119 tial occupancy was higher than environmental occupancy because it was generally
120 more restrictive in the number of sites a species could potentially occupy (figure
121 1d). Thus, occupancy estimates were higher when fewer sites were considered to
122 estimate it.

123 **How do occupancy estimations affect abundance-occupancy relationships?** We found positive abundance-occupancy relationships using all three oc-
124 cupancy metrics, but, based on the observed Spearman's rank correlation coef-
125 ficient (ρ), the relationship was stronger when using traditional occupancy ($\rho =$
126 0.53 , $p < 0.01$) than when using environmental occupancy ($\rho = 0.39$, $p < 0.01$)
127 or spatial occupancy ($\rho = 0.36$, $p < 0.01$). These differences in the strength of
128 the observed relationship seem to occur because the association between species
129 abundance and occupancy becomes more unclear when occupancy is not estimated
130 traditionally (figure 2a-c). In general, species with small geographic ranges have
131 their occupancies underestimated to a higher degree than species with large ranges
132 (figure 2d), although this underestimation is not dependent on species abundance
133 as there is no relationship between species range size and abundance ($\rho = 0.12$,
134 $p = 0.18$).

136 Discussion

137 Occupancy estimation is a fundamental step for the evaluation of abundance-
138 occupancy relationships, but the assumption that species can occupy all sampled
139 sites is generally overlooked when occupancy is estimated. We show that this
140 assumption directly affects abundance-occupancy relationships, and these rela-
141 tionships become weaker when occupancy is estimated based only on sites possible
142 for the species to colonize. This result is driven mostly by species with small
143 geographic ranges that have their occupancy highly underestimated when occu-

¹⁴⁴ pancy is estimated considering all sampled sites. Thus, removing the unrealistic
¹⁴⁵ assumption that species can occupy all sampled sites [10] has a clear and strong
¹⁴⁶ effect on the assessment of one of the most commonly reported macroecological
¹⁴⁷ relationships.

¹⁴⁸ These effects of occupancy estimation will be more pronounced for smaller-ranged
¹⁴⁹ species, although these effects might be limited when smaller spatial scales are
¹⁵⁰ sampled as most of the species geographic range will be found within the sam-
¹⁵¹ pled area [1]. On the other hand, abundance-occupancy relationships assessed
¹⁵² over broad spatial scales typically consider species with different ecological charac-
¹⁵³teristics. Considering these species differences, especially in terms of geographic
¹⁵⁴ ranges and environmental niche, when estimating occupancy is important as it
¹⁵⁵ can provide a more realistic depiction of abundance-occupancy relationships. For
¹⁵⁶ example, taking these species differences into account will improve our assessment
¹⁵⁷ of the effects of specialist and generalist species on abundance-occupancy relation-
¹⁵⁸ships [5, 43] given that specialist species generally have narrower environmental
¹⁵⁹ niches and smaller geographic ranges than generalist species [44, 45].

¹⁶⁰ The positive relationship between species environmental niche and geographic
¹⁶¹ range size [27, 28] suggests that both factors are intrinsically related and are im-
¹⁶²portant to determine species occurrences. Thus, using knowledge on the species
¹⁶³ geographic range and realized environmental niche provide biological realistic ways
¹⁶⁴ to estimate occupancy given that environmental suitability and geographical ac-
¹⁶⁵cessibility are needed for a species to occur at a location [37, 34]. In general, we
¹⁶⁶ show that species with small geographic ranges are the most affected when occu-
¹⁶⁷pancy is estimated traditionally as several sites that are unsuitable for the species
¹⁶⁸occurrence are considered to estimate their occupancy. This result suggests that
¹⁶⁹attempts to predict species abundance from occupancy patterns [46] should be
¹⁷⁰done carefully as some species occupancy might be underestimated occupancy is
¹⁷¹estimated traditionally.

172 We show that estimating species occupancy considering all sampled sites directly
173 affects the assessment of abundance-occupancy relationships. This assumption
174 ignores the fact that species have different spatial and environmental constraints
175 that can prevent them from occupying a given site. This can particularly affect
176 the assessment of macroecological patterns at large spatial scales where species
177 occurring in an assemblage might show high variation in terms of geographic ranges
178 and environmental niches. This could explain differences in abundance-occupancy
179 relationships observed for different taxa when these relationships are evaluated over
180 broad spatial scales [4]. Thus, a more realistic description of species occupancy
181 patterns will be obtained when species differences are considered during occupancy
182 estimation, and this will also lead to a refined assessment of abundance-occupancy
183 relationships.

184 **References**

185 [1] Gaston KJ. 1996 The multiple forms of the interspecific abundance-
186 distribution relationship. *Oikos* pp. 211–220.

187 [2] Roney NE, Kuparinen A, Hutchings JA. 2015 Comparative analysis of
188 abundance–occupancy relationships for species at risk at both broad taxo-
189 nomic and spatial scales. *Canadian Journal of Zoology* **93**, 7, 515–519.

190 [3] Miranda LE, Killgore KJ. 2019 Abundance–occupancy patterns in a riverine
191 fish assemblage. *Freshwater Biology* **64**, 12, 2221–2233.

192 [4] Ten Caten C, Holian L, Tad D. 2022 Weak but consistent abun-
193 dance–occupancy relationships across taxa, space and time. *Global Ecology
194 and Biogeography* .

195 [5] Verberk WC, Van Der Velde G, Esselink H. 2010 Explaining abundance–
196 occupancy relationships in specialists and generalists: a case study on aquatic
197 macroinvertebrates in standing waters. *Journal of Animal Ecology* **79**, 3, 589–
198 601.

199 [6] Dallas TA, Pöyry J, Leinonen R, Ovaskainen O. 2019 Temporal sampling and
200 abundance measurement influences support for occupancy–abundance rela-
201 tionships. *Journal of Biogeography* **46**, 12, 2839–2849.

202 [7] Marino NA, Céréghino R, Gilbert B, Petermann JS, Srivastava DS, de Om-
203 ena PM, Bautista FO, Guzman LM, Romero GQ, Trzcinski MK, *et al.* 2020
204 Species niches, not traits, determine abundance and occupancy patterns: A
205 multi-site synthesis. *Global Ecology and Biogeography* **29**, 2, 295–308.

206 [8] Riis T, Sand-Jensen K. 2002 Abundance-range size relationships in stream
207 vegetation in denmark. *Plant Ecology* **161**, 2, 175–183.

208 [9] Lovett-Doust J, Hegazy A, Hammouda O, Gomaa N. 2009 Abundance-
209 occupancy relationships and implications for conservation of desert plants
210 in the northwestern red sea region. *Community Ecology* **10**, 1, 91–98.

211 [10] Buckley HL, Freckleton RP. 2010 Understanding the role of species dynamics
212 in abundance–occupancy relationships. *Journal of Ecology* **98**, 3, 645–658.

213 [11] Mateus-Barros E, de Melo ML, Bagatini IL, Caliman A, Sarmento H. 2021
214 Local and geographic factors shape the occupancy-frequency distribution of
215 freshwater bacteria. *Microbial Ecology* **81**, 1, 26–35.

216 [12] Borregaard MK, Rahbek C. 2010 Causality of the relationship between geo-
217 graphic distribution and species abundance. *The Quarterly review of biology*
218 **85**, 1, 3–25.

219 [13] Webb MH, Heinsohn R, Sutherland WJ, Stojanovic D, Terauds A. 2019 An
220 empirical and mechanistic explanation of abundance-occupancy relationships
221 for a critically endangered nomadic migrant. *The American Naturalist* **193**,
222 1, 59–69.

223 [14] Werner EE, Davis CJ, Skelly DK, Relyea RA, Benard MF, McCauley SJ. 2014
224 Cross-scale interactions and the distribution-abundance relationship. *Plos
225 one* **9**, 5, e97387.

226 [15] Gaston KJ, Blackburn TM, Lawton JH. 1997 Interspecific abundance-range
227 size relationships: an appraisal of mechanisms. *Journal of Animal Ecology*
228 pp. 579–601.

229 [16] Holbrook JD, Arkle RS, Rachlow JL, Vierling KT, Pilliod DS, Wiest MM.
230 2016 Occupancy and abundance of predator and prey: Implications of the
231 fire-cheatgrass cycle in sagebrush ecosystems. *Ecosphere* **7**, 6, e01307.

232 [17] Chase JM, Myers JA. 2011 Disentangling the importance of ecological niches
233 from stochastic processes across scales. *Philosophical transactions of the Royal
234 Society B: Biological sciences* **366**, 1576, 2351–2363.

235 [18] Thompson K, Hodgson JG, Gaston KJ. 1998 Abundance–range size relation-
236 ships in the herbaceous flora of central england. *Journal of ecology* **86**, 3,
237 439–448.

238 [19] Komonen A, Päivinen J, Kotiaho JS. 2009 Missing the rarest: is the positive
239 interspecific abundance–distribution relationship a truly general macroecological
240 pattern? *Biology Letters* **5**, 4, 492–494.

241 [20] Hartley S. 1998 A positive relationship between local abundance and regional
242 occupancy is almost inevitable (but not all positive relationships are the
243 same). *Journal of Animal Ecology* pp. 992–994.

244 [21] Wilson PD. 2008 The pervasive influence of sampling and methodological arte-
245 facts on a macroecological pattern: the abundance–occupancy relationship.
246 *Global Ecology and Biogeography* **17**, 4, 457–464.

247 [22] Pulliam HR. 2000 On the relationship between niche and distribution. *Ecology*
248 *letters* **3**, 4, 349–361.

249 [23] Sexton JP, McIntyre PJ, Angert AL, Rice KJ. 2009 Evolution and ecology of
250 species range limits. *Annu. Rev. Ecol. Evol. Syst.* **40**, 415–436.

251 [24] Leibold MA. 1995 The niche concept revisited: mechanistic models and com-
252 munity context. *Ecology* **76**, 5, 1371–1382.

253 [25] Kraft NJ, Adler PB, Godoy O, James EC, Fuller S, Levine JM. 2015 Commu-
254 nity assembly, coexistence and the environmental filtering metaphor. *Func-*
255 *tional ecology* **29**, 5, 592–599.

256 [26] Cadotte MW, Tucker CM. 2017 Should environmental filtering be abandoned?
257 *Trends in ecology & evolution* **32**, 6, 429–437.

258 [27] Slatyer RA, Hirst M, Sexton JP. 2013 Niche breadth predicts geographical
259 range size: a general ecological pattern. *Ecology letters* **16**, 8, 1104–1114.

260 [28] Kambach S, Lenoir J, Decocq G, Welk E, Seidler G, Dullinger S, Gégout JC,
261 Guisan A, Pauli H, Svenning JC, *et al.* 2019 Of niches and distributions: range
262 size increases with niche breadth both globally and regionally but regional
263 estimates poorly relate to global estimates. *Ecography* **42**, 3, 467–477.

264 [29] Heino J, Tolonen KT. 2018 Ecological niche features override biological traits
265 and taxonomic relatedness as predictors of occupancy and abundance in lake
266 littoral macroinvertebrates. *Ecography* **41**, 12, 2092–2103.

267 [30] Ehrlén J, Eriksson O. 2000 Dispersal limitation and patch occupancy in forest
268 herbs. *Ecology* **81**, 6, 1667–1674.

269 [31] Ozinga WA, Schaminée JH, Bekker RM, Bonn S, Poschlod P, Tackenberg O,
270 Bakker J, Groenendaal JMv. 2005 Predictability of plant species composition
271 from environmental conditions is constrained by dispersal limitation. *Oikos*
272 **108**, 3, 555–561.

273 [32] Pinto SM, MacDougall AS. 2010 Dispersal limitation and environmental
274 structure interact to restrict the occupation of optimal habitat. *The American
275 Naturalist* **175**, 6, 675–686.

276 [33] Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, Dor-
277 mann CF, Forchhammer MC, Grytnes JA, Guisan A, *et al.* 2013 The role of
278 biotic interactions in shaping distributions and realised assemblages of species:
279 implications for species distribution modelling. *Biological reviews* **88**, 1, 15–
280 30.

281 [34] Soberón J, Peterson AT. 2005 Interpretation of models of fundamental eco-
282 logical niches and species' distributional areas .

283 [35] Von Oheimb PV, Albrecht C, Riedel F, Bössneck U, Zhang H, Wilke T. 2013
284 Testing the role of the himalaya mountains as a dispersal barrier in freshwater
285 gastropods (*gyraulus* spp.). *Biological Journal of the Linnean Society* **109**, 3,
286 526–534.

287 [36] Pirani RM, Werneck FP, Thomaz AT, Kenney ML, Sturaro MJ, Ávila-Pires
288 TC, Peloso PL, Rodrigues MT, Knowles LL. 2019 Testing main amazonian
289 rivers as barriers across time and space within widespread taxa. *Journal of
290 Biogeography* **46**, 11, 2444–2456.

291 [37] Gaston KJ. 1991 How large is a species' geographic range? *Oikos* pp. 434–438.

292 [38] Gaston KJ, Fuller RA. 2009 The sizes of species' geographic ranges. *Journal*
293 *of applied ecology* **46**, 1, 1–9.

294 [39] Kao RH, Gibson CM, Gallery RE, Meier CL, Barnett DT, Docherty KM,
295 Blevins KK, Travers PD, Azuaje E, Springer YP, *et al.* 2012 Neon terres-
296 trial field observations: designing continental-scale, standardized sampling.
297 *Ecosphere* **3**, 12, 1–17.

298 [40] Thorpe AS, Barnett DT, Elmendorf SC, Hinckley ELS, Hoekman D, Jones
299 KD, LeVan KE, Meier CL, Stanish LF, Thibault KM. 2016 Introduction to the
300 sampling designs of the national ecological observatory network terrestrial
301 observation system. *Ecosphere* **7**, 12, e01627.

302 [41] Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, Ram K.
303 2021 *rgbif: Interface to the Global Biodiversity Information Facility API*. R
304 package version 3.5.2.

305 [42] Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005 Very high
306 resolution interpolated climate surfaces for global land areas. *International*
307 *Journal of Climatology: A Journal of the Royal Meteorological Society* **25**, 15,
308 1965–1978.

309 [43] Sudta C, Salcido DM, Forister ML, Walla T, Villamarín-Cortez S, Dyer LA.
310 2020 Jack-of-all-trades paradigm meets long-term data: generalist herbivores
311 are more widespread and locally less abundant. *Ecology Letters* **00**, 1–10.

312 [44] Clavel J, Julliard R, Devictor V. 2011 Worldwide decline of specialist species:
313 toward a global functional homogenization? *Frontiers in Ecology and the*
314 *Environment* **9**, 4, 222–228.

315 [45] Boulangeat I, Lavergne S, Van Es J, Garraud L, Thuiller W. 2012 Niche
316 breadth, rarity and ecological characteristics within a regional flora spanning
317 large environmental gradients. *Journal of Biogeography* **39**, 1, 204–214.

³¹⁸ [46] Hanski I. 1982 Dynamics of regional distribution: the core and satellite species
³¹⁹ hypothesis. *Oikos* pp. 210–221.

320 Figures

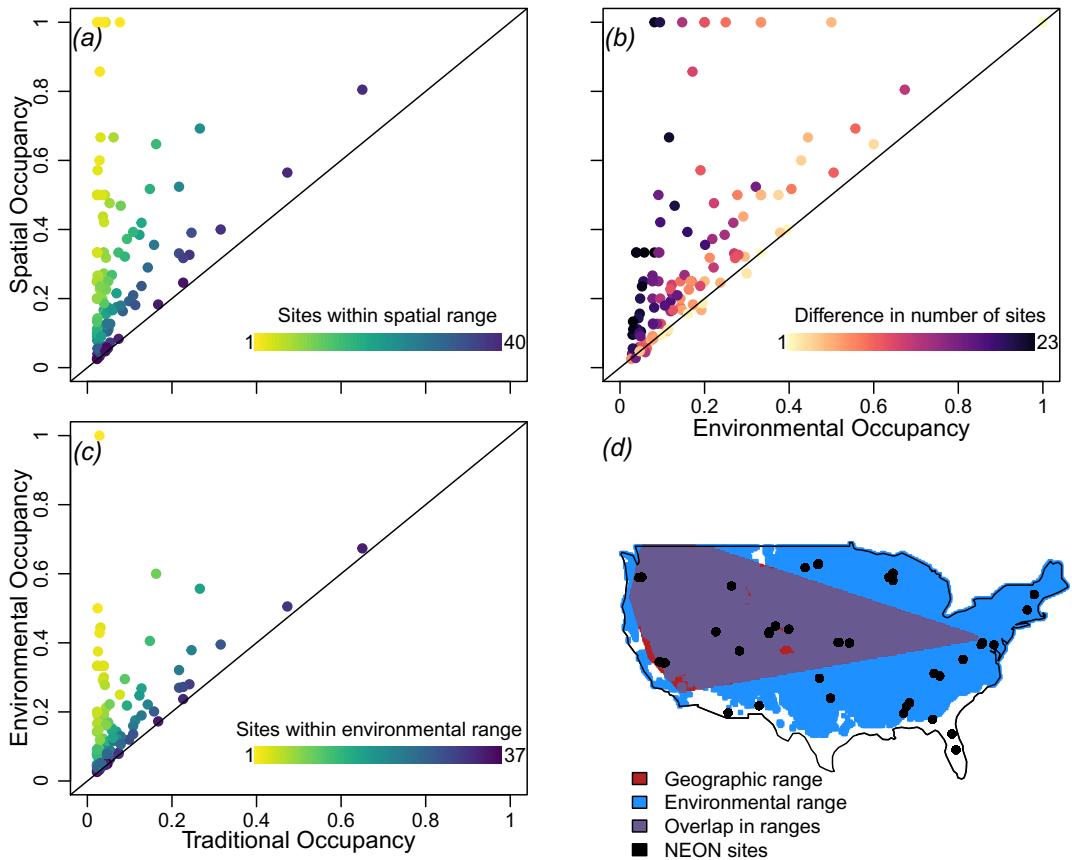


Figure 1: Comparison between traditional and spatial (a), environmental and spatial (b), and traditional and environmental (c) occupancy estimations. Points closer to the identity line represent species that have more similar occupancy estimates in the compared approaches. Legends represent the number of sites within species spatial range (a) and environmental niche (c), and the difference in the number of sites within the species spatial range and environmental niche (b). In panel (d) we show areas suitable for *Ochotona princeps* occurrence based on its geographic range (in red) and realized environmental niche (in blue). Areas in purple represent locations that are suitable for the species occurrence based on both the species geographic range and realized environmental niche and black points are the sampled NEON sites.

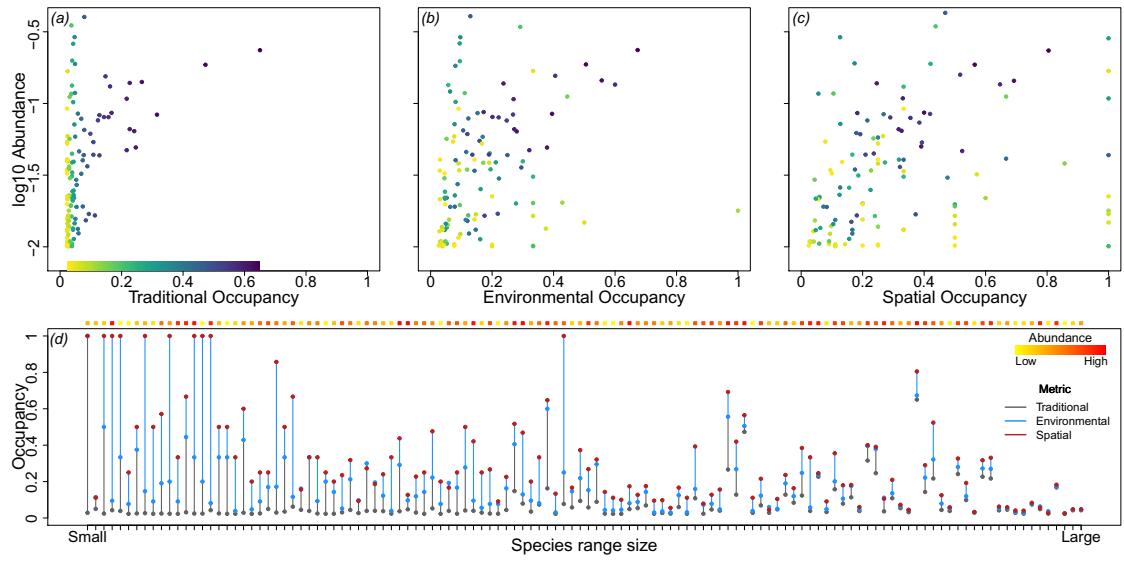


Figure 2: Relationships between \log_{10} species abundance and traditional (a), environmental (b) and spatial (c) occupancies. Points are colored based on the traditional occupancy estimation in the three panels. Panel (d) shows how species with small geographic ranges have their occupancies more strongly underestimated (i.e. traditional occupancy is substantially smaller than environmental or spatial occupancies) than species with larger geographic ranges. These differences in occupancy estimates were independent of abundance given the lack of a clear relationship between species abundance (shown in the squared symbols) and differences in the species occupancy estimates.

321 **Supplemental Material**

322 Effects of occupancy estimation on abundance-occupancy relationships

323

324 Cleber Ten Caten^{a,*}, Lauren A. Holian^a and Tad Dallas^a

325

326 ^a Department of Biological Sciences, University of South Carolina, Columbia,
327 South Carolina, 29205 * clebertencaten@gmail.com

328

329 **Methods**

330 We obtained occurrence points from the Global Biodiversity Information Facility
331 (GBIF) for the 122 species used in our study. These occurrence points were subse-
332 quently used to estimate the species realized environmental niche and geographic
333 range. To estimate the species geographic range we used the species occurrence
334 points found in the United States and minimum convex polygons. The minum
335 convex polygon is defined as the smallest polygon in which no internal angle ex-
336 ceeds 180 degrees and it contains all the points used to build it. Figure S1a) shows
337 the occurrence points of *Sorex fumeus* in the geographic space and its geographic
338 range (black line) obtained from minimum convex polygon.

339 To estimate the species realize environmental niche we used the 19 bioclimatic
340 variables available in the BioClim database. These variables constitute different
341 facets of temperature and precipitation patterns. More especifically, these bio-
342 climatic variables represent annual trends (e.g. mean annual temperature and
343 precipitation), seasonality (e.g. mean diurnal range in temperature) and extreme
344 environmental factors (e.g. precipitation of the driest and coldest quarter) associ-
345 ated with temperature and precipitation. The bioclimatic variables were obtained
346 at a resolution of 10 arc-minutes (i.e. $\approx 18 \text{ km}^2$) covering the Americas. We
347 selected the resolution of 10 arc-minutes because it can be used to obtain a fine
348 information of the species realized environmental niche. A Principal Component
349 Analysis (PCA) was performed on these variables and the first two axes explained

350 80% of the variance of the data and were used to estimate the realized environmen-
351 tal niche of the species. To achieve this goal, we used the species occurrence points
352 and extracted the environmental values associated with these points and used min-
353 imum convex polygons to estimate the species realized environmental niche. Fig-
354 ure S1b) shows the occurrence points of *Sorex fumeus* in the environmental space
355 and its realized environmental niche (black line) obtained from minimum convex
356 polygon.

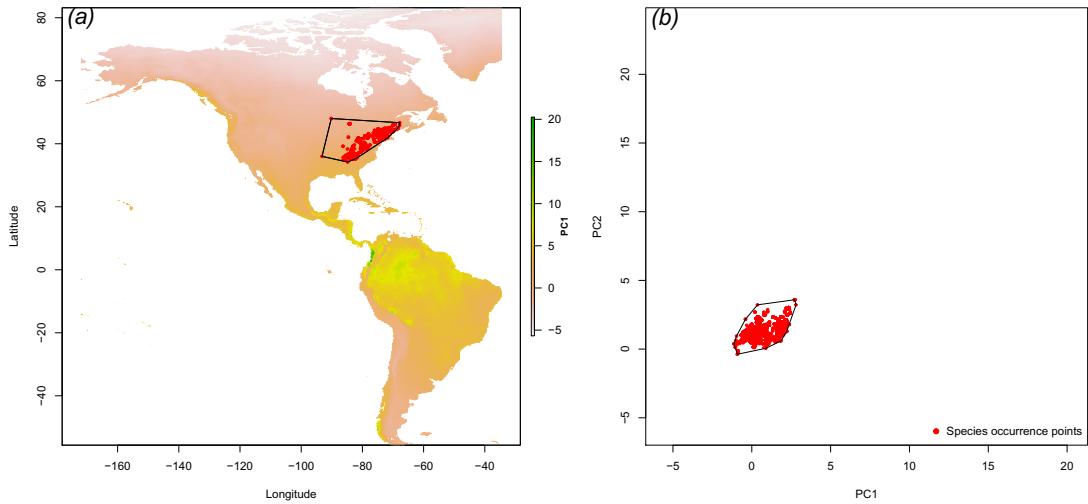


Figure S1: Geographic range of *Sorex fumeus* obtained from minimum convex polygons represented by the black line surrounding the species occurrence points (red dots) plotted in the geographic space *a*. The color in the map represents the environmental values of the first axis of the PCA (PC1) in the geographic space. Panel *b* shows the *Sorex fumeus* realized environmental niche obtained from the minimum convex polygons represented by the black line surrounding the species occurrence points (red dots) plotted in the environmental space. Here, the environmental space is represented by the first and second PCA axes, PC1 and PC2 respectively.