

1

Université de Montréal

2

## **Spatially continuous identification of beta diversity hotspots using species distribution models**

3

4

By

5

**Gabriel Dansereau**

6

20147609

7

Département de sciences biologiques

8

Faculté des arts et des sciences

9

Advisory Committee Meeting

10

November 28, 2019

<sup>11</sup> **Contents**

<sup>12</sup>	Abstract . . . . .	2
<sup>13</sup>	Introduction . . . . .	3
<sup>14</sup>	Methods . . . . .	4
<sup>15</sup>	1. Data Collection . . . . .	4
<sup>16</sup>	2. Data Manipulation . . . . .	6
<sup>17</sup>	3. SDM – The BIOCLIM Method . . . . .	6
<sup>18</sup>	4. LCBD Calculation . . . . .	7
<sup>19</sup>	5. Prediction Validity . . . . .	8
<sup>20</sup>	6. Alternative methods . . . . .	9
<sup>21</sup>	7. Climate Change Scenarios and Temporal Beta Diversity . . . . .	9
<sup>22</sup>	Preliminary Results . . . . .	10
<sup>23</sup>	References . . . . .	19

<sup>24</sup> **List of Tables**

<sup>25</sup>	1 Distribution of the data in the eBird checklists for the countries used in the analyses . . . . .	13
<sup>26</sup>	2 WorldClim 2 climate variables used in the analyses . . . . .	13

<sup>28</sup> **List of Figures**

<sup>29</sup>	1 Single Species Distributions - Raw . . . . .	15
<sup>30</sup>	2 Single Species Distributions - SDM . . . . .	15
<sup>31</sup>	3 Species Richness - Raw . . . . .	16
<sup>32</sup>	4 Species Richness - SDM . . . . .	16
<sup>33</sup>	5 LCBD values - Raw (transformed) . . . . .	17
<sup>34</sup>	6 LCBD values - SDM . . . . .	17
<sup>35</sup>	7 LCBD-richness relationship - Raw . . . . .	18
<sup>36</sup>	8 LCBD-richness relationship - SDM . . . . .	19

<sup>37</sup> **Abstract**

<sup>38</sup> Beta diversity is an essential measure to describe the organization of biodiversity in space.  
<sup>39</sup> The calculation of local contributions to beta diversity (LCBD), specifically, allows for the  
<sup>40</sup> identification of sites with exceptional diversity within a region of interest, which is useful  
<sup>41</sup> for both community ecology and conservation purposes. However, beta diversity implies a  
<sup>42</sup> comparison among the sites of a given region, thus, its use is restricted to sites with known  
<sup>43</sup> species composition, and to discontinuous spatial scales. We therefore propose a method  
<sup>44</sup> to calculate LCBD indices on continuous scales for a whole region of interest, including  
<sup>45</sup> unsampled sites. First, species distributions can be predicted on continuous scales using  
<sup>46</sup> species distribution models (SDM). These models, such as the BIOCLIM method, use the  
<sup>47</sup> environmental conditions at sampled sites to predict the presence or absence of each species at  
<sup>48</sup> unsampled locations. Second, LCBD statistics can then be computed on the SDM predictions.  
<sup>49</sup> We therefore show that it is possible to identify beta diversity hotspots on spatially continuous  
<sup>50</sup> and extended scales. Our results confirm that LCBD values are related to species richness,  
<sup>51</sup> and that species-poor sites contribute most to beta diversity.

52 **Introduction**

53 Beta diversity, defined as the variation in species composition among sites in a geographic  
54 region of interest (Legendre, Borcard, and Peres-Neto 2005), is an essential measure to  
55 describe the organization of biodiversity in space. Total beta diversity within a community  
56 can be partitioned into local contributions to beta diversity (LCBD) (Legendre and De Cáceres  
57 2013), which allows for the identification of sites with exceptional species composition,  
58 hence exceptional biodiversity. Such a method is useful for both community ecology and  
59 conservation biology, as it highlights sites that are most important for their research or  
60 conservation values. However, LCBD calculation methods require complete information  
61 on community composition, such as a community composition matrix  $Y$ , thus they are  
62 inappropriate for partially sampled or unsampled sites. To our knowledge, these methods  
63 have mostly been applied on community data from sampled sites, hence on discontinuous  
64 spatial scales, e.g. at intervals along a river stream (Legendre and De Cáceres 2013). This  
65 raises the following questions: 1) could LCBD indices be extended to continuous spatial  
66 scales, and 2) could this provide novel ecological insights in poorly sampled regions? We  
67 aim to answer these questions by combining the LCBD calculation methods with predictive  
68 biogeography approaches, and suggest that this would allow for the identification of sites  
69 with high conservation value in poorly sampled regions.

70 Species distribution models (SDMs) already allow to make predictions on continuous spatial  
71 scales which could be used to calculate LCBD indices. These methods, also known as  
72 bioclimatic envelope models (Araújo and Peterson 2012), aim to predict species presence or  
73 absence based on observation of occurrences at known locations (Poisot et al. 2019). This way,  
74 they generate novel ecological insights, and represent an approach yet to be applied to LCBD.  
75 We believe that such an approach of generating novel ecological insights for unsampled or  
76 lesser-known locations could be an interesting new perspective in the study. Through them,  
77 we would be able to expand community information already available, and thus work on a  
78 much larger community matrix than in typical LCBD studies.

79 Appropriate data to expand measures of exceptional biodiversity through space is increasingly  
80 available online. For instance, the Worldclim 2.0 database (Fick and Hijmans 2017) provides  
81 interpolated climate data for global land areas at very high spatial resolution, and the eBird

82 platform (Sullivan et al. 2009) provides a growing citizen-contributed database of worldwide  
83 bird observations. Both of these are commonly used in SDMs, and offer relevant information  
84 on extended spatial scales. Hence, we believe that we could use them to predict community  
85 composition and calculate LCBD indices on continuous spatial scales, and that the result  
86 would be representative of the true community structure.

87 The predictive approach we suggest would be especially useful in poorly sampled regions,  
88 or in regions with only sparse sampling. While it doesn't replace a full sampling within the  
89 community, it does provide relevant ecological insights. For instance, the method could help  
90 identify unsampled sites with potential conservation value which should be targeted as soon  
91 as possible in future studies. We also believe that our method could also be combined with  
92 IPCC climate change scenarios, which provide projections for climate variables, in a way that  
93 would allow us to model beta diversity changes with climate change and to identify the sites  
94 where the changes in the community will be most important. Again, this method would be  
95 more relevant as an informative approach to suggest sites to prioritize for future conservation  
96 and more structured research.

97 In this document, we cover in more details the methods that we suggest for this research  
98 project. The preparation part of the project, including data collection and manipulation, has  
99 already been done, and a workflow for the analyses, including code implementation, has been  
100 defined as well. We also detail preliminary analyses and results intended as proof-of-concept  
101 for the approach, which of course needs to be refined. Finally, we discuss methods that we  
102 intend to use in future analyses, and whose feasibility is not as clearly stated.

## 103 **Methods**

### 104 **1. Data Collection**

105 We decided to focus our analyses on bird species and collected the data available on eBird for  
106 the Warblers family. The complete database contains nearly 600 million observations. We  
107 chose to focus specifically on the Warblers family, as it is a diverse group, popular among  
108 birders, with over 30 million observations. Global citizen-contributed databases often present  
109 additional challenges compared to conventional datasets due to their lack of structure, as well  
110 as spatial and taxonomic biases (Johnston et al. 2019). For instance, there was a clear bias in

111 our data towards the United States, where there were much more observations and sampling  
112 events (tbl. 1). However, eBird offers two advantages over other large scale datasets (Johnston  
113 et al. 2019): 1) the data is structured as checklist and users can explicitly specify their  
114 observations as “complete checklists” when all detected species were reported, which allows  
115 to infer information on species absences, and 2) the dataset is semi-structured and checklists  
116 are associated with metadata describing sampling effort, such as duration of search, distance  
117 travelled and number of observers, which can be used as controls in the analyses. Hence,  
118 model performance can be improved by inferring absences and subsampling checklists, while  
119 spatial bias can be compensated by including effort covariates in the model. Therefore, we  
120 believe the dataset can be appropriately used to achieve our objective of expanding measures  
121 of exceptional biodiversity through space.

122 We collected the data available in the WorldClim 2 database (Fick and Hijmans 2017) for  
123 North America, to which we decided to restrict our analyses. The WorldClim data consists of  
124 spatially interpolated monthly climate data for global areas, available for resolutions from 10  
125 arc-minutes to 30 arc-seconds (around 18 km<sup>2</sup> and 1 km<sup>2</sup> at the equator). Since the release of  
126 the first version of the database in 2005 (Hijmans et al. 2005), it became the most common  
127 source of climate data for SDM studies (Booth et al. 2014). The variables we used were  
128 different measures of temperature and precipitation(tbl. 2), and they have been show to have  
129 very high global cross-validation coefficients (> 0.99 and 0.86 respectively) (Fick and Hijmans  
130 2017). We chose to use the coarser 10 arc-minutes resolution in our preliminary analyses,  
131 as we believed it was sufficient for proof of concept of our method. However, Hijmans et  
132 al. (2005) showed high within-grid cell variation in the 10 arc-minutes data, and therefore  
133 recommended the use of the finer resolution, which hid less of the variation known to the  
134 model. Given this, we might reconsider the resolution to use in our final analyses.

135 We chose to restrict our analyses to North America given the high amount of data available in  
136 eBird. We believed it represented a suitable scale for our models, large enough to cover a lot  
137 of variation in environmental variables and community structure, as well as phenomenons  
138 such as species migration. We also expected such extent of the spatial scale to cover for  
139 imprecision in estimated species ranges.

140 **2. Data Manipulation**

141 WorldClim variables and eBird occurrence data are provided in different formats, so they  
142 required some manipulations to be combined together. WorldClim variables are provided in  
143 a 2-dimensional grid format, useful for large scale analyses and visualization, where each  
144 cell or pixel has a size corresponding to the resolution of 10 arc-minutes. Each of the 19  
145 variables forms a different grid. On the other hand, eBird records are occurrence-based, so  
146 each entry in the dataset corresponds to an observation of a single species at a given time  
147 and location. These entries can easily be matched to the 2D grid format of the WorldClim  
148 variables through their spatial coordinates, which we found more useful for large scale  
149 analyses and visualization. Hence, for each species, we matched all occurrences in eBird to  
150 the grid format of the WorldClim variables, and then created a presence-absence community  
151 matrix  $Y$ , taking all the grid cells as sites. At the 10 arc-minutes resolution, we obtained  
152 39 024 sites with occurrences and 62 species in total. All data manipulations and further  
153 analyses were realized in *Julia v1.2.0* (Bezanson et al. 2017) with the basic structure built  
154 around the soon-to-be-released `SimpleSDMLayers.jl` package.

155 **3. SDM – The BIOCLIM Method**

156 We predicted species distributions using the BIOCLIM method (Nix 1986), a climate-envelope  
157 model, considered a classic in the field. The method simply relates a species' distribution  
158 to the ranges of bioclimatic variables at known locations (Booth et al. 2014). It has long  
159 been outperformed by other methods (Elith et al. 2006), but it is still commonly used for  
160 its simplistic approach and ease of understanding, as well as its simple relation to niche  
161 theory (Booth et al. 2014; Hijmans et al. 2017). It is also primarily designed for presence-only  
162 data. Despite that, we chose this method for our preliminary analyses as it was easier to  
163 implement and because we believe it to be sufficient for proof-of-concept. We discuss possible  
164 alternatives in the “Alternative Methods” section below.

165 The BIOCLIM method defines species potential ranges as a multidimensional environmental  
166 hypervolume bounded by the minimum and maximum values for all occurrences (Franklin  
167 2010). For each species, we established the percentile distribution of each environmental  
168 variable at the known locations of occurrence (Hijmans et al. 2017). All sites were then

169 compared to those percentile distributions and given a score per variable according to  
170 their ranking between 0.0 (1st percentile) and 1.0 (100th percentile). The median or 50th  
171 percentile was considered the most suitable value of the variable, and values larger than  
172 0.5 were subtracted from 1. Therefore, both tails are considered the same. The minimum  
173 percentile score across all environmental variables was then selected as the predicted value  
174 for each site. Values were multiplied by 2 and could therefore be interpreted as probabilities  
175 of species occurrence (Hijmans et al. 2017). Predictions of 1 should be rare by definition,  
176 as they require a perfectly median site on all variables, and values of 0 should be frequent,  
177 since they occur whenever an environmental value is outside the range of the observed ones  
178 (Hijmans et al. 2017).

179 The final step was to convert the probabilities into presence-absence data, so they could be  
180 compared with the raw occurrence data. We transformed the probabilities into 0s and 1s  
181 by converting all values greater than zero to one. Although it might tend to overestimate  
182 species ranges, such a transformation is common in SDMs and can be accounted for during  
183 result validation with specific methods (Franklin 2010). We also considered applying a  
184 threshold, which would be determined by sensitivity analysis. In any case, converting into  
185 presence-absence data allowed easier calculation of the richness and beta diversity metric.

#### 186 4. LCBD Calculation

187 We calculated the LCBD statistics through the total variance of the matrix  $Y$  for both the raw  
188 data and SDM predictions. Legendre and De Cáceres (2013) showed that LCBD coefficients  
189 can be calculated directly through the total variance of matrix  $Y$ , or through a matrix of  
190 dissimilarities among sampling units. We chose the first approach as it also allows to compute  
191 species contributions to beta diversity (SCBD), which could also prove useful for conservation  
192 purposes, but we did not investigate these for now. Before computing the LCBD statistics,  
193 the presence-absence matrix  $Y$  had to be transformed in an appropriate way (Legendre and  
194 De Cáceres 2013). We chose to apply the Hellinger transformation to the raw data and no  
195 transformation on the SDM predictions for now, although we did not investigate these in  
196 detail. The most appropriate transformation still needs to be determined, especially for  
197 the SDM predictions. We then computed a matrix  $S$  of squared deviations from column  
198 means and summed all the values of  $S$  to obtain the total sum of squares ( $SS$ ) of the species

199 composition data (Legendre and De Cáceres 2013). LCBD coefficients are then computed as  
200  $LCBD * i = SS_i / SS * Total$ , where  $SS_i$  is the sum of squares of a sampling unit  $i$ . Finally, since  
201 our matrix  $Y$  is very large, the LCBD coefficients are very small, so we scaled them to the  
202 maximum value observed.

203 **5. Prediction Validity**

204 The exact way of testing the validity of the predictions remains to be determined, and will  
205 also depend on the exact methods used to make the SDM predictions. A key element to note  
206 is that both SDM predictions and LCBD values will have to be validated, hence they might  
207 require different methods. Metrics that measure the accuracy of categorical or probabilistic  
208 predictions in SDMs are well documented, in various forms. Some require absence data to test  
209 against, and can be used on probabilistic predictions directly (area-under-curve, AUC) or after  
210 a conversion of the predictions to binary presence-absence using a certain threshold (Kappa  
211 index, measuring the difference between observed and chance agreement in a confusion  
212 matrix) (Franklin 2010). Other methods are appropriate for presence-only data, such as the  
213 Boyce Index. In any case, measuring prediction error is only one part of validation, and  
214 finding appropriate data for evaluation is almost as essential (Franklin 2010), especially  
215 since we also aim to describe community structure. Separating the data into a training and  
216 testing dataset, with 70% and 30% of the observations for instance, is a possible approach  
217 common in machine learning methods, although all of the available observations might be  
218 needed in some cases (Franklin 2010). An interesting approach, suggested by Elith et al.  
219 (2006) for SDMs, would be to find independent, well-structured presence-absence datasets  
220 for validation, on which both SDM predictions and beta diversity metrics could be tested.  
221 This approach has the advantage that the testing data is truly independent of the training  
222 one, hence it could be used with certain tests of significance. Although it might not cover the  
223 entire extent of the predictions in a single test, this method brings a closer comparison to the  
224 way LCBD metrics are used in most studies, and it would provide interesting perspectives if  
225 combined with other, full-extent validation methods.

226 **6. Alternative methods**

227 Many methods generally outperform BIOCLIM for the predictions, as shown by Elith et al.  
228 (2006). In our case, better predictions will come by two different means: 1) approaches that  
229 are better than BIOCLIM to model the relationship between species presence-absence (or even  
230 abundance) and environmental variables, and 2) approaches that account for other drivers of  
231 species distributions, such as ecological interactions and species migration. Machine learning  
232 methods, especially, would be also be interesting alternatives to consider. MAXENT (Phillips,  
233 Anderson, and Schapire 2006), another presence-only method, has come to be one of the most  
234 widely used methods in SDM studies, often with WorldClim variables (Booth et al. 2014).  
235 Similarly, Random Forests are simple to put in place, take into account both presence and  
236 absence data, allow for quantification of the variables importance in explaining variation, and  
237 offer intrinsic testing metrics (Franklin 2010). However, while those methods might return  
238 more accurate predictions, they do not implicitly model other drivers of species distribution,  
239 among which species interactions and functional niche. Integrating those factors might  
240 prove more difficult given our dataset and our focus Warblers species, as no appropriate  
241 information on their interaction is available, to our knowledge. Joint species distribution  
242 models (JSDMs) might be an interesting way to encompass those, as they attempt to model  
243 species co-occurrence, rather than the distribution of single species (Pollock et al. 2014).  
244 A different taxonomic group and dataset with more details on interactions could also be  
245 used. On the other hand, finding a method that could be applied to any taxonomic group,  
246 especially those well represented in large citizen-contributed datasets, would be most useful  
247 for research and conservation purposes.

248 **7. Climate Change Scenarios and Temporal Beta Diversity**

249 We aim to apply our method to environmental conditions from IPCC climate change scenarios,  
250 first to model community compositions after climate change on continuous scales through  
251 SDMs, and then to identify the sites where the community has changed in the most exceptional  
252 ways. This can be seen from the variation in LCBD values, but also through temporal beta  
253 diversity indices (TBI) (Legendre 2019). TBI indices allow to study changes in community  
254 composition through time from repeated surveys at given sites. Whereas LCBD values  
255 essentially measure the contribution to beta diversity of each site compared to all others,

256 TBI measure changes in community composition for a single site between two surveys, and  
257 can also be decomposed into species losses and gains (Legendre 2019). Moreover, TBI can  
258 be tested for significance using a permutation test. An approach similar to that of Legendre  
259 and Condit (2019) would be most interesting to follow. First, they computed LCBD indices  
260 and compared the sites that were significant for two surveys 30 years apart, highlighting  
261 a swamp region where important changes seemed to have occurred. Then, they used TBI  
262 indices to confirm the sites with significant changes, decompose these changes into losses and  
263 gains, and identify the species that had changed the most. Such an approach could be highly  
264 informative with our data, although the permutation tests and corrections to apply might  
265 cause problems given the number of sites implied in our study.

266 The possibility of using climate change scenarios in the SDMs also needs to be assessed.  
267 We did not try to download nor find the appropriate data for now. Interpolated variables  
268 are sometimes different than those used in Worldclim 2.0. Therefore, the SDM models and  
269 predictions might have to be different too, and potentially less reliable. Nonetheless, we  
270 believe it will be possible to do some kind of time analysis linking beta diversity, climate  
271 change and species distribution modelling, and that it could return highly informative results  
272 for conservation purposes.

## 273 Preliminary Results

274 Our preliminary results mainly consisted of comparisons between the raw occurrence data  
275 and the SDM predictions for the four following elements: single-species distribution (figs. 1,  
276 2), species richness (figs. 3, 4), LCBD coefficients (figs. 5, 6), and the relationship between the  
277 species richness and LCBD coefficients (figs. 7, 8). Two main results emerged from them: 1)  
278 the models provided seemingly valid and relevant results for poorly sampled regions, both  
279 expected species-poor and species-rich ones, and 2) there was a strong association between  
280 species richness and LCBD coefficients confirming the relationship shown in other studies.

281 First, the example of the Yellow Warbler (*Setophaga petechia*), one of the most observed species,  
282 showed that the single-species models predicted a broad distribution covering poorly sampled  
283 areas, with notable patches of absence across the continent (figs. 1, 2). Likewise, species  
284 richness, defined as the number of species present per site, showed a clear latitude gradient,

285 with the poorest sites to the North and the richest to the South (figs. 3, 4). A form of altitude  
286 gradient could also be observed, with the Rockies and other mountains well delimited by their  
287 lower values. In both cases, the results make intuitive sense and highlight the models ability  
288 to predict species presence despite poor or no sampling. Mexico, for example, has much  
289 sparser sampling and fewer observations, but the models predict Yellow Warblers presence in  
290 most areas nonetheless, as well as higher species richness than on the highly sampled Atlantic  
291 Coast, which make sense for a more southern location. We believe these to be valid insights  
292 on poorly sampled locations, although we reckon that intuitive reasoning is not a proof of  
293 anything, and that the model might be wrong in important ways. In any case, it highlights  
294 the need for an appropriate method of validation, as well as a thoughtful consideration of  
295 other factors such as species migration.

296 Second, our preliminary LCBD results seemed to confirm the association between species  
297 richness and LCBD coefficients, while also being valid and insightful. They were however  
298 harder to interpret given the use of the Hellinger transformation for the raw occurrence  
299 data only. Raw occurrence data showed a negative relationship between species richness and  
300 LCBD coefficients (fig. 7), as observed previously by Heino and Grönroos (2017), with no  
301 clear geographic pattern (fig. 5). If anything, the highest values seemed to be at the borders  
302 of the most sampled regions, which are about where the sites with the less species observed  
303 are located (fig. 3). On the other hand, SDM predictions showed a quadratic form, with the  
304 LCBD coefficients re-increasing after some threshold (fig. 8). The geographic pattern is also  
305 clearer, with the highest values to the northern and southern extremes (fig. 6). We suggest  
306 that this result makes sense, as LCBD indices should highlight the most exceptional species  
307 composition, and these are both well and continuously represented in a SDM. Although raw  
308 occurrence data results concorded with those of Heino and Grönroos (2017), the “border  
309 effect” and the difference with SDM projections might show the importance of going beyond  
310 occurrence data when using large databases but spatially biased databases such as eBird.  
311 Once again, our results highlight the need for well-thought method and an investigation of  
312 the appropriate transformation to use on the data.

313 Finally, one disappointing aspect of our method is that the result failed to identify patterns  
314 on finer scales. The trends shown by the SDMs for both the species richness and LCBD

315 coefficients were large-scale, latitude-related patterns. Except for mountains, few exceptional  
316 sites are actually shown in the middle of the landscape. While it might have been unrealistic  
317 to expect such results from a coarse analysis like ours, it would be useful for conservation  
318 purposes to be able to identify precise sites within smaller regions. This might be achieved  
319 by using a finer resolution, which we should probably reconsider, or by using a different  
320 technique, such as training the models and predicting species distributions on large scales,  
321 but computing and scaling LCBD values on finer local ones, which might highlight regional  
322 differences in a new way.

Table 1: Distribution of the data in the eBird checklists for the countries used in the analyses

Country	Observations	Checklists	Species	Species per	Species per	Species per
				checklist	checklist	checklist
	(mean)	(median)	(maximum)			
US	19 206 453	7 840 526	56	2.450	2.0	34
CA	3 360 650	1 115 625	45	3.012	2.0	31
MX	407 227	147 599	61	2.759	2.0	21
Total	22 974 330	9 103 750	63	2.523	2.0	34

Table 2: WorldClim 2 climate variables used in the analyses

Variable	Description
1	Annual Mean Temperature
2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
3	Isothermality (BIO2/BIO7) (* 100)
4	Temperature Seasonality (standard deviation *100)
5	Max Temperature of Warmest Month
6	Min Temperature of Coldest Month
7	Temperature Annual Range (BIO5-BIO6)
8	Mean Temperature of Wettest Quarter
9	Mean Temperature of Driest Quarter
10	Mean Temperature of Warmest Quarter
11	Mean Temperature of Coldest Quarter
12	Annual Precipitation
13	Precipitation of Wettest Month
14	Precipitation of Driest Month
15	Precipitation Seasonality (Coefficient of Variation)
16	Precipitation of Wettest Quarter
17	Precipitation of Driest Quarter
18	Precipitation of Warmest Quarter

---

Variable	Description
19	Precipitation of Coldest Quarter

---

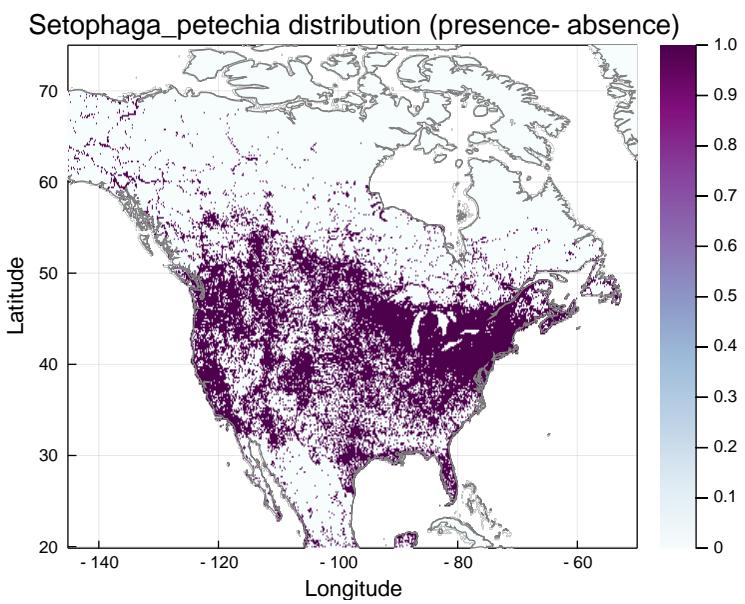


Figure 1: Single Species Distributions - Raw

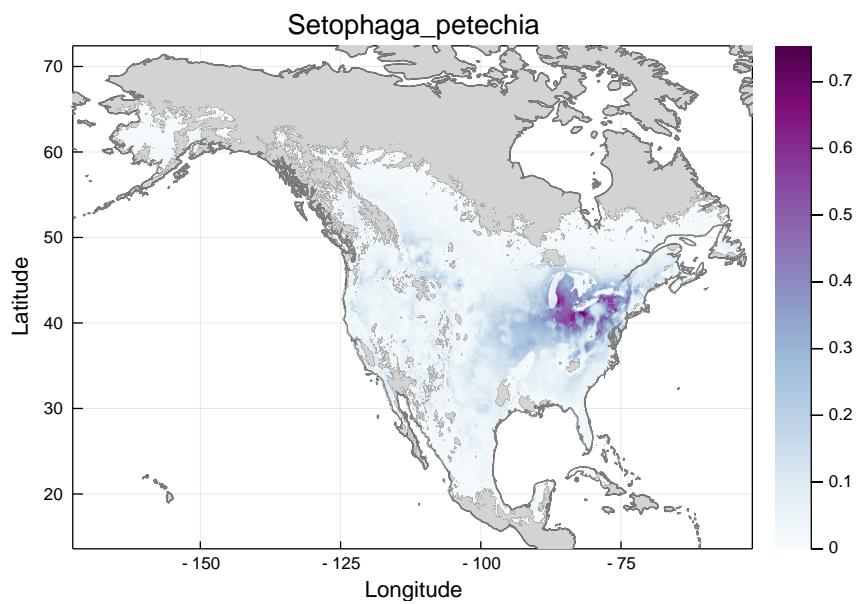


Figure 2: Single Species Distributions - SDM

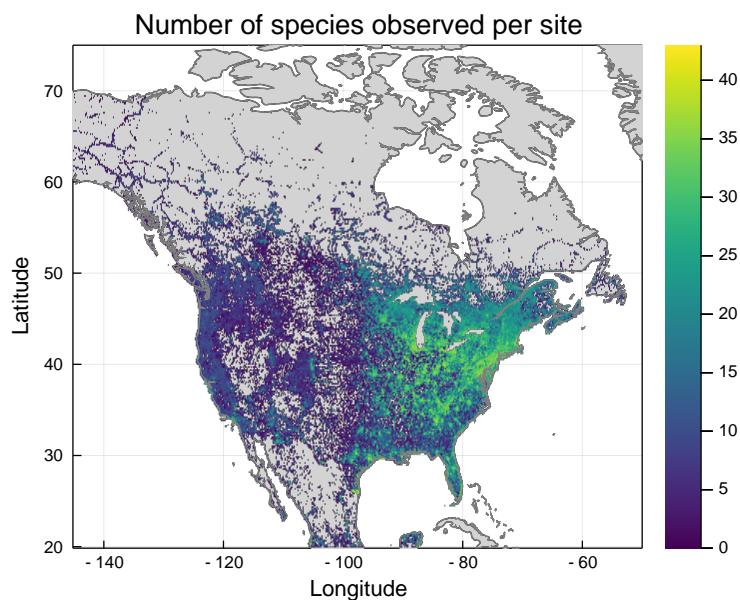


Figure 3: Species Richness - Raw

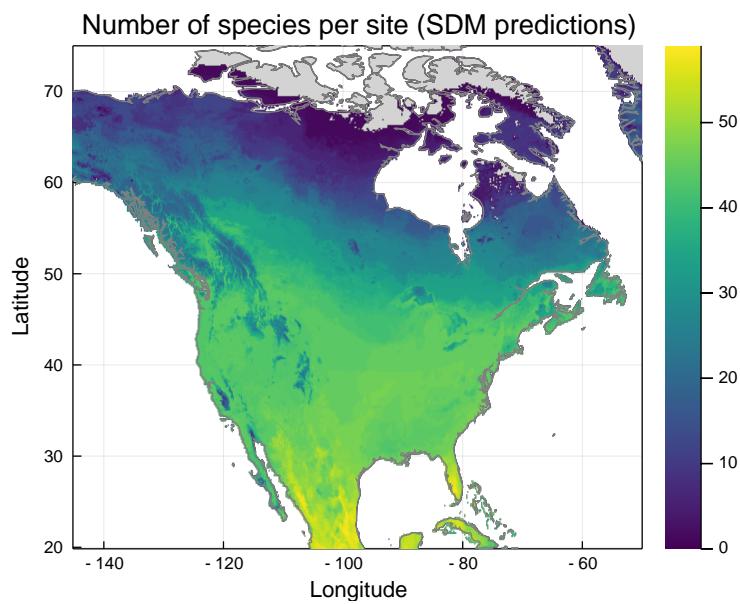


Figure 4: Species Richness - SDM

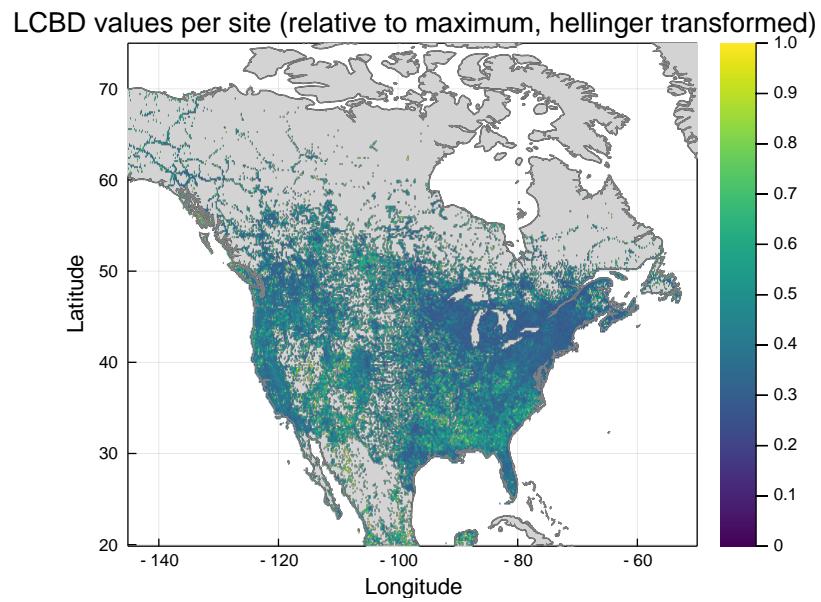


Figure 5: LCBD values - Raw (transformed)

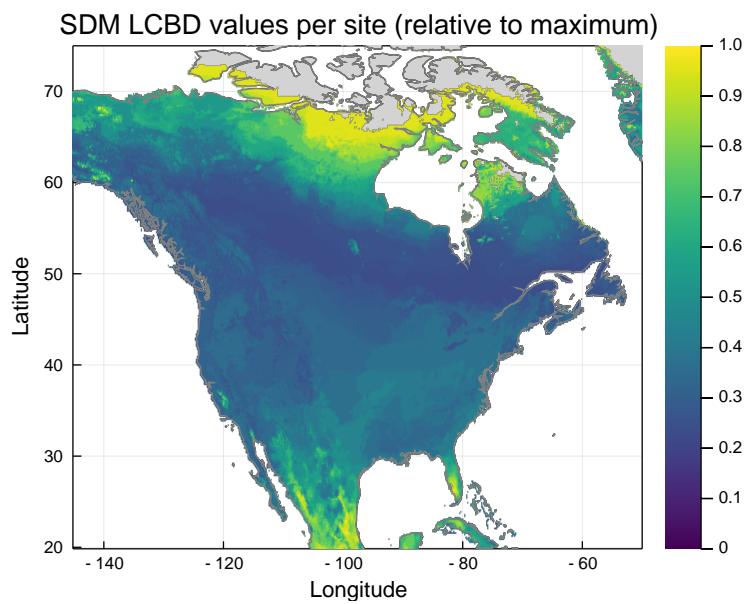


Figure 6: LCBD values - SDM

## Relationship between LCBD (hellinger transformed) and species richness

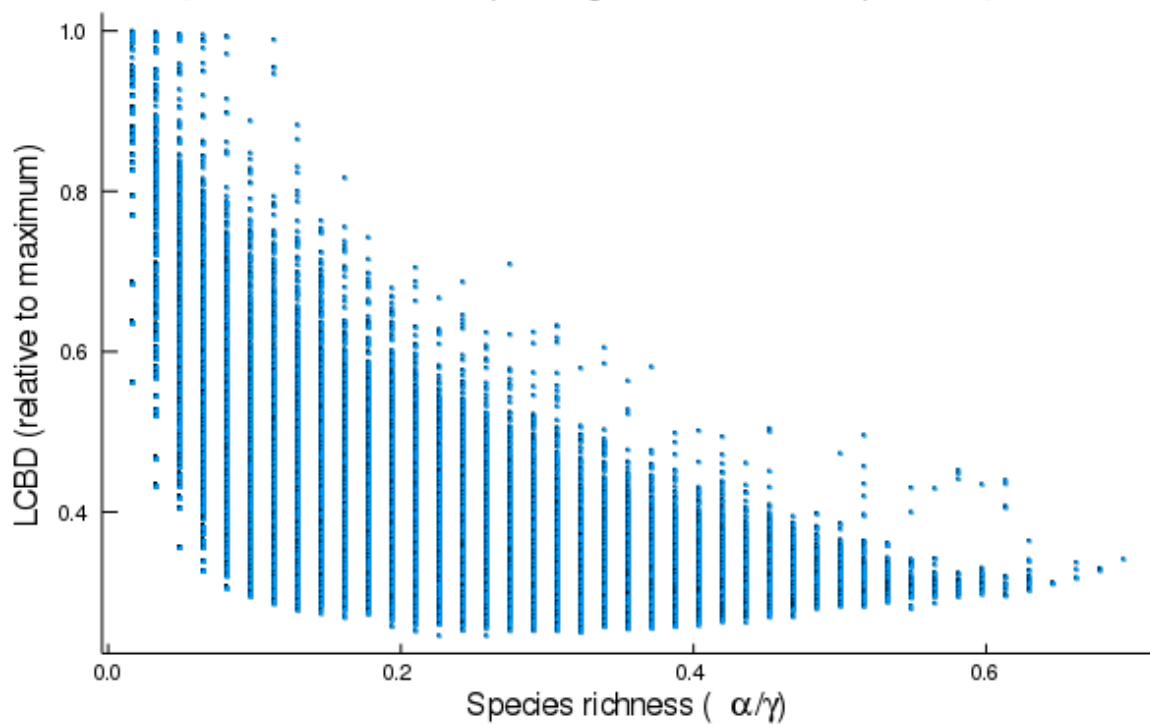


Figure 7: LCBD-richness relationship - Raw

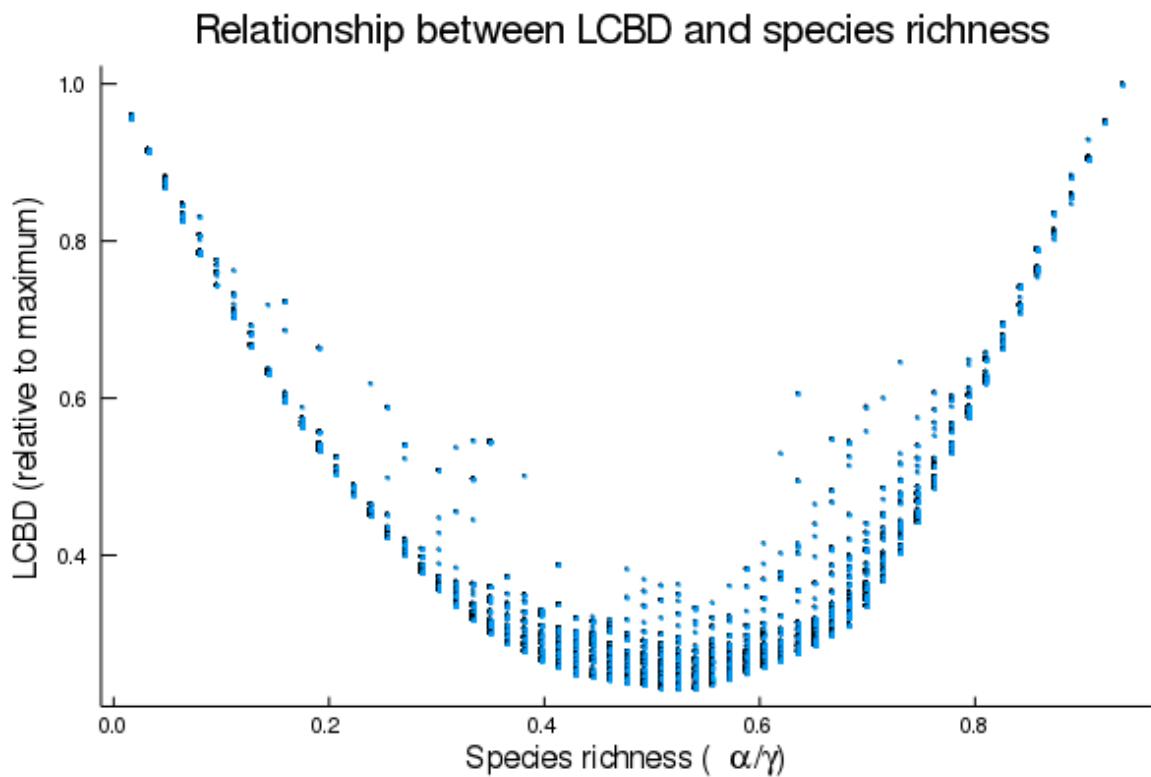


Figure 8: LCBD-richness relationship - SDM

**323 References**

- 324 Araújo, Miguel B., and A. Townsend Peterson. 2012. “Uses and Misuses of Bioclimatic  
325 Envelope Modeling.” *Ecology* 93 (7): 1527–39. <https://doi.org/10.1890/11-1930.1>.
- 326 Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B. Shah. 2017. “Julia: A Fresh  
327 Approach to Numerical Computing.” *SIAM Review* 59 (1): 65–98. <https://doi.org/10.1137/141000671>.
- 328
- 329 Booth, Trevor H., Henry A. Nix, John R. Busby, and Michael F. Hutchinson. 2014. “BIOCLIM:  
330 The First Species Distribution Modelling Package, Its Early Applications and Relevance to  
331 Most Current MaxEnt Studies.” *Diversity and Distributions* 20 (1): 1–9. <https://doi.org/10.1111/ddi.12144>.
- 332
- 333 Elith, Jane, Catherine H. Graham, Robert P. Anderson, Miroslav Dudík, Simon Ferrier, Antoine  
334 Guisan, Robert J. Hijmans, et al. 2006. “Novel Methods Improve Prediction of Species’  
335 Distributions from Occurrence Data.” *Ecography* 29 (2): 129–51. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>.
- 336

- 337 Fick, Stephen E., and Robert J. Hijmans. 2017. "WorldClim 2: New 1-Km Spatial Resolution  
338 Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37 (12):  
339 4302–15. <https://doi.org/10.1002/joc.5086>.
- 340 Franklin, Janet. 2010. "Moving Beyond Static Species Distribution Models in Support of  
341 Conservation Biogeography: Moving Beyond Static Species Distribution Models." *Diversity*  
342 and *Distributions* 16 (3): 321–30. <https://doi.org/10.1111/j.1472-4642.2010.00641.x>.
- 343 Heino, Jani, and Mira Grönroos. 2017. "Exploring Species and Site Contributions to Beta  
344 Diversity in Stream Insect Assemblages." *Oecologia* 183 (1): 151–60. <https://doi.org/10.1007/s00442-016-3754-7>.
- 345
- 346 Hijmans, Robert J., Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis. 2005.  
347 "Very High Resolution Interpolated Climate Surfaces for Global Land Areas." *International*  
348 *Journal of Climatology* 25 (15): 1965–78. <https://doi.org/10.1002/joc.1276>.
- 349
- 350 Hijmans, Robert J., Steven Phillips, John Leathwick, and Jane Elith. 2017. *Dismo: Species*  
*Distribution Modeling*. <https://CRAN.R-project.org/package=dismo>.
- 351
- 352 Johnston, A., W. M. Hochachka, M. E. Strimas-Mackey, V. Ruiz Gutierrez, O. J. Robinson,  
353 E. T. Miller, T. Auer, S. T. Kelling, and D. Fink. 2019. "Best Practices for Making Reli-  
354 able Inferences from Citizen Science Data: Case Study Using eBird to Estimate Species  
355 Distributions." *bioRxiv*, March, 574392. <https://doi.org/10.1101/574392>.
- 356
- 357 Legendre, Pierre. 2019. "A Temporal Beta-Diversity Index to Identify Sites That Have  
358 Changed in Exceptional Ways in Space-Time Surveys." *Ecology and Evolution* 9 (6): 3500–  
359 3514. <https://doi.org/10.1002/ece3.4984>.
- 360
- 361 Legendre, Pierre, Daniel Borcard, and Pedro R. Peres-Neto. 2005. "Analyzing Beta Diver-  
362 sity: Partitioning the Spatial Variation of Community Composition Data." *Ecological*  
363 *Monographs* 75 (4): 435–50. <https://doi.org/10.1890/05-0549>.
- 364
- Legendre, Pierre, and Richard Condit. 2019. "Spatial and Temporal Analysis of Beta Diversity  
in the Barro Colorado Island Forest Dynamics Plot, Panama." *Forest Ecosystems* 6 (1): 7.  
<https://doi.org/10.1186/s40663-019-0164-4>.
- Legendre, Pierre, and Miquel De Cáceres. 2013. "Beta Diversity as the Variance of Community

- 365 Data: Dissimilarity Coefficients and Partitioning." *Ecology Letters* 16 (8): 951–63. <https://doi.org/10.1111/ele.12141>.
- 366
- 367 Nix, Henry A. 1986. "A Biogeographic Analysis of Australian Elapid Snakes." *Atlas of Elapid Snakes of Australia* 7: 4–15.
- 368
- 369 Phillips, Steven J., Robert P. Anderson, and Robert E. Schapire. 2006. "Maximum Entropy Modeling of Species Geographic Distributions." *Ecological Modelling* 190 (3): 231–59. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- 370
- 371
- 372 Poisot, Timothée, Richard LaBrie, Erin Larson, Anastasia Rahlin, and Benno I. Simmons. 2019. "Data-Based, Synthesis-Driven: Setting the Agenda for Computational Ecology." *Ideas in Ecology and Evolution* 12 (July). <https://doi.org/10.24908/iee.2019.12.2.e>.
- 373
- 374
- 375 Pollock, Laura J., Reid Tingley, William K. Morris, Nick Golding, Robert B. O'Hara, Kirsten M. Parris, Peter A. Vesk, and Michael A. McCarthy. 2014. "Understanding Co-Occurrence by Modelling Species Simultaneously with a Joint Species Distribution Model (JSDM)." *Methods in Ecology and Evolution* 5 (5): 397–406. <https://doi.org/10.1111/2041-210X.12180>.
- 376
- 377
- 378
- 379
- 380 Sullivan, Brian L., Christopher L. Wood, Marshall J. Iliff, Rick E. Bonney, Daniel Fink, and Steve Kelling. 2009. "eBird: A Citizen-Based Bird Observation Network in the Biological Sciences." *Biological Conservation* 142 (10): 2282–92. <https://doi.org/10.1016/j.biocon.2009.05.006>.
- 381
- 382
- 383