

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 27, 2019

¹¹ **Contents**

¹²	Abstract	2
¹³	Introduction	3
¹⁴	Methods	4
¹⁵	1. Data Collection	4
¹⁶	2. Data Manipulation	6
¹⁷	3. SDM – The BIOCLIM method	6
¹⁸	4. LCBD calculation	7
¹⁹	5. Prediction validity	8
²⁰	6. Alternative methods	9
²¹	7. Climate change scenarios & temporal beta diversity	9
²²	Preliminary Results	10
²³	References	17

²⁴ **List of Tables**

²⁵	1 WorldClim 2 climate variables used in the analyses	5
---------------	--	---

²⁶ **List of Figures**

²⁷	1 Single Species Distributions - Raw	13
²⁸	2 Single Species Distributions - SDM	13
²⁹	3 Species Richness - Raw	14
³⁰	4 Species Richness - SDM	14
³¹	5 LCBD values - Raw (transformed)	15
³²	6 LCBD values - SDM	15
³³	7 LCBD-richness relationship - Raw	16
³⁴	8 LCBD-richness relationship - SDM	17

³⁵ **Abstract**

³⁶ Beta diversity is an essential measure to describe the organization of biodiversity in space.
³⁷ The calculation of local contributions to beta diversity (LCBD), specifically, allows for the
³⁸ identification of sites with exceptional diversity within a region of interest, which is useful
³⁹ for both community ecology and conservation purposes. However, beta diversity implies a
⁴⁰ comparison among the sites of a given region, thus, its use is restricted to sites with known
⁴¹ species composition, and to discontinuous spatial scales. We therefore propose a method
⁴² to calculate LCBD indices on continuous scales for a whole region of interest, including
⁴³ unsampled sites. First, species distributions can be predicted on continuous scales using
⁴⁴ species distribution models (SDM). These models, such as the BIOCLIM method, use the
⁴⁵ environmental conditions at sampled sites to predict the presence or absence of each species at
⁴⁶ unsampled locations. Second, LCBD statistics can then be computed on the SDM predictions.
⁴⁷ We therefore show that it is possible to identify beta diversity hotspots on spatially continuous
⁴⁸ and extended scales. Our results confirm that LCBD values are related to species richness,
⁴⁹ and that species-poor sites contribute most to beta diversity.

50 **Introduction**

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

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

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

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

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

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

101 **Methods**

102 **1. Data Collection**

103 We decided to focus our analyses on bird species and collected the data available on eBird
104 for the Warblers family. The complete database contains nearly 600 million observations,
105 and presents two main advantages over other large scale datasets (Johnston et al. 2019): 1)
106 data is structured as checklist and users can explicitly specify their observations as "complete
107 checklists" when all detected species were reported, which allows to infer information on
108 species absences, 2) the dataset is semi-structured and checklists are associated with metadata

describing sampling effort, such as duration of search, distance travelled and number of observers, which can be used as controls in the analyses. We chose to focus specifically on the Warblers family, as it is a diverse group, popular among birders, with over 30 million observations.

We decided to restrict our analyses to North America and collected climate data available in the WorldClim 2 database (Fick and Hijmans 2017). We believe North America represents a suitable scale, large enough to cover a lot of variation in environmental variables and community structure, as well as phenomenons such as species migration. We also expect such extent of the spatial scale to cover for imprecision in estimated species ranges. The WorldClim data consists of spatially interpolated monthly climate data for global areas, available for resolutions from 10 arc-minutes to 30 arc-seconds. The variables used are provided in Table 1, and consists of different measures of temperature and precipitation. We chose to use the coarser 10 arc-minutes resolution in our analyses, again to cover for imprecision, and because we believe it is sufficient for proof of concept.

Table 1: 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

Variable	Description
15	Precipitation Seasonality (Coefficient of Variation)
16	Precipitation of Wettest Quarter
17	Precipitation of Driest Quarter
18	Precipitation of Warmest Quarter
19	Precipitation of Coldest Quarter

¹²³ **2. Data Manipulation**

¹²⁴ WorldClim variables and eBird occurrence data are provided in different formats, so they
¹²⁵ require some manipulation to be combined together. WorldClim variables are provided
¹²⁶ in a 2-dimensional grid format, useful for large scale analyses and visualization, where
¹²⁷ each cell or pixel corresponds to the resolution of 10 arc-minutes. Each of the 19 variables
¹²⁸ forms a different grid. On the other hand, eBird records are occurrence-based, so each
¹²⁹ entry in the dataset corresponds to an observation of a single species at a given time and
¹³⁰ location. These entries can easily be matched to the 2D grid format of the WorldClim variables
¹³¹ through their spatial coordinates, which we found more useful for large scale analyses and
¹³² visualization. Hence, for each species, we matched all occurrences in eBird to the grid format
¹³³ of the WorldClim variables, and then created a presence-absence community matrix Y , taking
¹³⁴ all the grid cells as sites. At the 10 arc-minutes resolution, we obtained 39 024 sites with
¹³⁵ occurrences and 62 species. We also applied the Hellinger transformation on the raw presence-
¹³⁶ absence data, although the most appropriate method remains to be determined, especially
¹³⁷ since the data has to be compared with the SDM predictions. All data manipulations and
¹³⁸ further analyses were realized in *Julia v1.2.0* (Bezanson et al. 2017) with the basic structure
¹³⁹ built around the soon-to-be-released `SimpleSDMLayers.jl` package.

¹⁴⁰ **3. SDM – The BIOCLIM method**

¹⁴¹ We used the BIOCLIM method to predict species distributions. BIOCLIM, first introduced by
¹⁴² Nix (1986), is considered as the classic “climate-envelope-model”, and is now available to
¹⁴³ users through the `dismo` package in R (Hijmans et al. 2017). It has long been outperformed
¹⁴⁴ by other methods (Elith et al. 2006), but it is still commonly used for its simplistic approach

145 and ease of understanding, as well as its simple relation to niche theory (Booth et al. 2014;
146 Hijmans et al. 2017). It is also a method designed for presence-only data, which does not
147 require information on absences, nor take them into account if provided (as in our case).
148 Despite that, we chose this method for our preliminary analyses as it was easier to implement
149 and because we believe it to be sufficient for proof-of-concept. We discuss possible alternatives
150 in the “Alternative methods” section below.

151 Briefly, the BIOCLIM method defines species potential range as a multidimensional envi-
152 ronmental hypervolume bounded by the minimum and maximum values of all presences
153 (Franklin 2010). For each species, the algorithm establishes the percentile distribution of
154 the values of each environmental variables at the known locations of occurrences (Hijmans
155 et al. 2017). The environmental variables of all sites are then compared to those percentile
156 distributions and given scores between 0 (1st percentile) and 1 (100th percentile). The median
157 or 50th percentile is considered as the most suitable location, and both tails (e.g. 10th and
158 90th percentile) are not distinguished, the values larger than 0.5 being subtracted from 1.
159 The minimum percentile score across all environmental variables is selected as the prediction
160 value for each site and multiplied by 2 so values are between 0 and 1 (Hijmans et al. 2017).
161 It should be noted that the limiting variable is thus not necessarily the same for all sites.
162 Values of 1 are rare, as it would mean a perfectly median site on all variables, and values
163 of 0 are frequent, since they are assigned whenever an environmental value is outside the
164 range of the observed values (Hijmans et al. 2017). Finally, before calculating richness or beta
165 diversity metrics, we transformed the predictions back to a presence-absence format, where
166 all predictions greater than one are considered as presence. This might tend to overestimate
167 species ranges and create some sort of border effect, but we believe the effects will be mitigated
168 given the spatial extent and coarse scale of our study.

169 4. LCBD calculation

170 We calculated the LCBD statistics through the total variance of the matrix Y for both the raw
171 data and SDM predictions. Legendre and De Cáceres (2013) showed that LCBD coefficients
172 can be calculated directly through the total variance of matrix Y , or through a matrix of
173 dissimilarities among sampling units. We chose the first approach as it also allows to compute
174 species contributions to beta diversity (SCBD), which could also prove useful for conservation

175 purposes, but we did not investigate these for now. Before computing the LCBD statistics, the
176 presence-absence matrix Y had to be transformed in an appropriate way, as mentioned earlier.
177 We chose to apply the Hellinger transformation to the raw data and no transformation on the
178 SDM predictions for now, as the most appropriate one still needs to be determined. We then
179 computed a matrix S of squared deviations from column means and summed all the values
180 of S to obtain the total sum of squares (SS) of the species composition data (Legendre and De
181 Cáceres 2013). LCBD coefficients are then computed as $LCBD_i = SS_i/SS_{Total}$, where SS_i is
182 the sum of squares of a sampling unit i . Finally, since our matrix Y is very large, the LCBD
183 coefficients are very small, so we scaled them to the maximum value observed.

184 **5. Prediction validity**

185 The exact way of testing the validity of the predictions remains to be determined, and will
186 also depend on the exact methods used to make the SDM predictions. A key element to note
187 is that both SDM predictions and LCBD values will have to be validated, so they will likely
188 require different methods. Many metrics are well documented in the literature to test SDM
189 predictions, such as the Kappa index (Franklin 2010), and could be used for the BIOCLIM
190 predictions. Another possible way would be to separate the data into a training and testing
191 dataset, with 70% and 30% of the data for instance, which is a common approach in machine
192 learning techniques. However, this approach reduces the amount of data that can be used
193 in the model, and raises the issue of making sure that the datasets are both random and
194 representative of the data, as well as of the community dynamics. Also, in this framework,
195 the testing data cannot be considered as independent, which prevents using it in certain tests
196 of significance. One interesting approach, suggested by Elith et al. (2006) for SDMs, would be
197 to find independent, well-structured presence-absence datasets for validation, on which beta
198 diversity metrics has been or could be calculated. This validation might not cover the entire
199 extent of the predictions, but it might bring interesting perspectives if combined with other
200 validation methods, mostly because it would bring a closer comparison to the way LCBD
201 metrics are used at the moment.

202 **6. Alternative methods**

203 Other methods could possibly outperform BIOCLIM for the predictions, as shown by Elith
204 et al. (2006). Better predictions will come by two different means: 1) approaches that
205 are better than BIOCLIM to model the relationship between species presence-absence (or
206 even abundance) and environmental variables, and 2) approaches that account for other
207 drivers of species distributions, such as ecological interactions and species migration. The
208 most obvious alternative to BIOCLIM is MAXENT (Phillips, Anderson, and Schapire 2006),
209 another presence-only method that has come to be one of the most widely used methods.
210 Machine learning methods would be also be interesting alternatives that have been proven to
211 outperform BIOCLIM (Franklin 2010). Random Forests, especially are simple methods to put
212 in place, allow for quantification of the variables importance in explaining variation, and offer
213 intrinsic testing metrics. Neural networks could also be an interesting alternative. However,
214 while those methods might return more accurate predictions, they do not implicitly model
215 other drivers of species distribution, among which species interactions and functional niche.
216 Integrating those factors might prove more difficult given our dataset and our focus Warblers
217 species, as no appropriate information on their interaction is available, to our knowledge.
218 Joint species distribution models (JSDMs) might be an interesting way to encompass those, as
219 they attempt to model species co-occurrence, rather than the distribution of single species
220 (Pollock et al. 2014). A different taxonomic group and data datasets could also be used with
221 more details on interactions could also be used, though having a method that can be applied
222 to any taxonomic group would be more interesting. Yet, such an approach might prove to be
223 beyond the scope of the present research.

224 **7. Climate change scenarios & temporal beta diversity**

225 We aim to apply our method to environmental conditions from climate change scenarios, first
226 to model community compositions after climate change on continuous scales through SDMs,
227 and then to identify the sites where the community has changed in the most exceptional ways.
228 This can be done through LCBD values, but also through temporal beta diversity indices
229 (TBI) (Legendre 2019), which allow to study changes in community composition through
230 time from repeated surveys at given sites. Whereas LCBD values essentially measure the
231 contribution to beta diversity of each site compared to all other ones, TBI measure changes in

community composition for a single site between two surveys, and can also be decomposed into species losses and gains. Moreover, TBI can be tested for significance using a permutation test. An approach similar to that of Legendre and Condit (2019) would be most interesting to follow: they first computed LCBD indices and compared the sites that were significant for two surveys 30 years apart, highlighting a swamp region where important changes seemed to have occurred, and then used TBI indices to confirm the sites with significant changes, decompose those into losses and gains and identify the species that had changed the most. Such an approach could be highly informative with our data, although the permutation tests and corrections to apply might cause problems given the number of sites that would be implied in our study. The possibility of using climate change scenarios in the SDMs also needs to be investigated in more details. We did not try to download nor find the appropriate data for now, but we found that the interpolated variables are sometimes different than those used in Worldclim 2.0. The SDM models and predictions might therefore be slightly different than those used for the LCBD calculations, and potentially less reliable. Nonetheless, we believe it will be possible to do some kind of time analysis linking beta diversity, climate change and species distribution modelling, and that it could return highly informative results for conservation purposes.

249 Preliminary Results

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

First, the example of the Yellow Warbler (*Setophaga petechia*), one of the most observed species, showed that the single-species models predicted a broad distribution covering poorly sampled areas, with notable patches of absence across the continent (Fig. 1). Likewise, species richness, defined as the number of species present per site, showed a clear latitude gradient, with the poorest sites to the North and the richest to the South (figs. 3, 4). A form of altitude gradient

262 could also be observed, with the Rockies and other mountains well delimited by their lower
263 values. In both cases, the results make intuitive sense and highlight the models ability to
264 predict species presence despite poor or no sampling. Mexico, for example, has much sparser
265 sampling and fewer observations, but the models predict Yellow Warblers presence in most
266 areas nonetheless, as well as higher species richness than on the highly sampled Atlantic
267 Coast, which make sense for a more southern location. We believe these to be valid insights
268 on poorly sampled locations, although we reckon that intuitive reasoning is not a proof of
269 anything, and that the model might be wrong in important ways. In any case, it highlights
270 the need for an appropriate method of validation, as well as a thoughtful consideration of
271 other factors such as species migration.

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

289 Finally, one disappointing aspect of our method is that the result failed to identify patterns
290 on finer scales. The trends shown by the SDMs for both the species richness and LCBD
291 coefficients were large-scale, latitude-related patterns. Except for mountains, few exceptional

292 sites are actually shown in the middle of the landscape. While it might have been unrealistic
293 to expect such results from a coarse analysis like ours, it would be useful for conservation
294 purposes to be able to identify precise sites within smaller regions. An interesting option
295 would be to train the models and predict species distributions on a large scale, but to compute
296 and scale the LCBD values of finer ones, as it might highlight regional differences in a new
297 way.

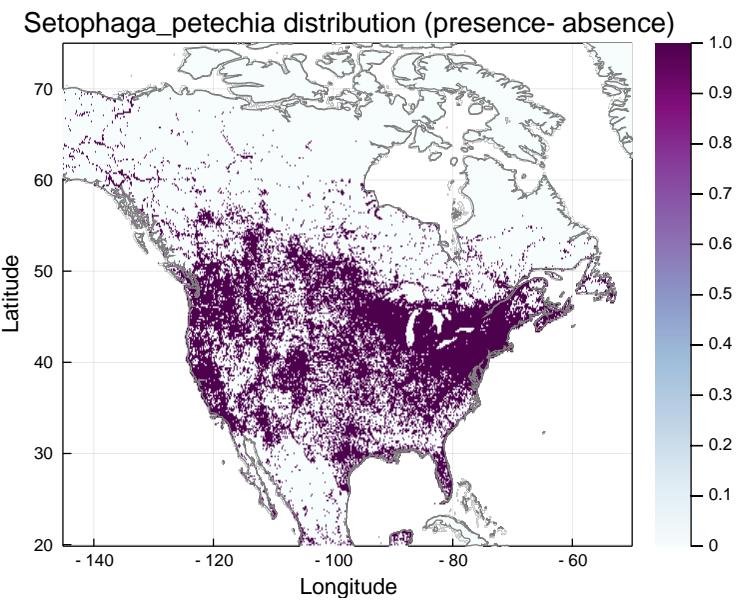


Figure 1: Single Species Distributions - Raw

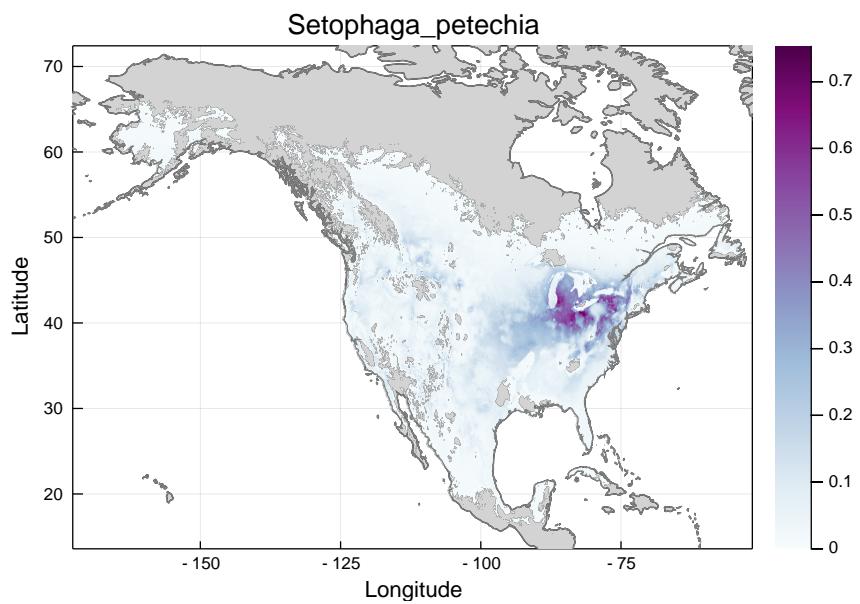


Figure 2: Single Species Distributions - SDM

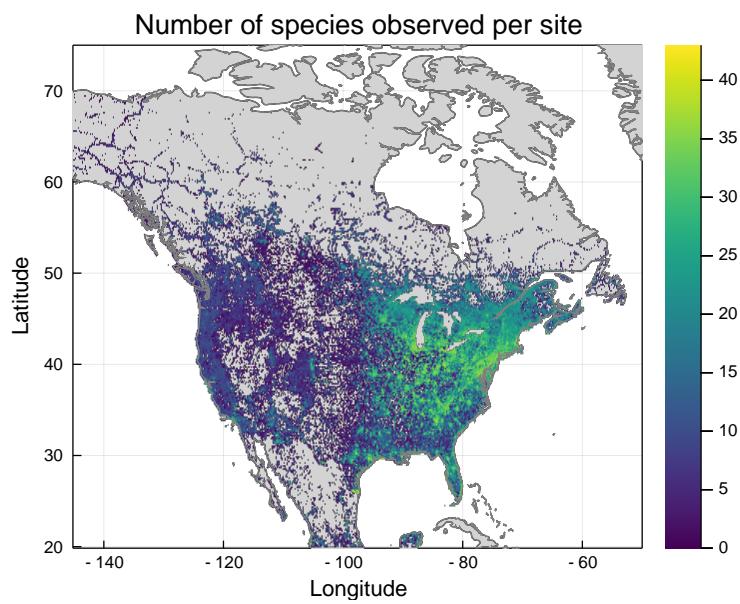


Figure 3: Species Richness - Raw

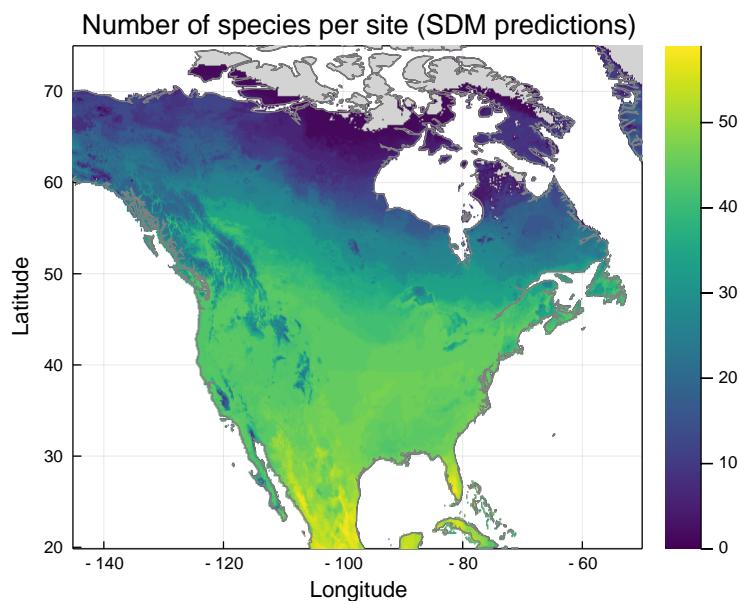


Figure 4: Species Richness - SDM

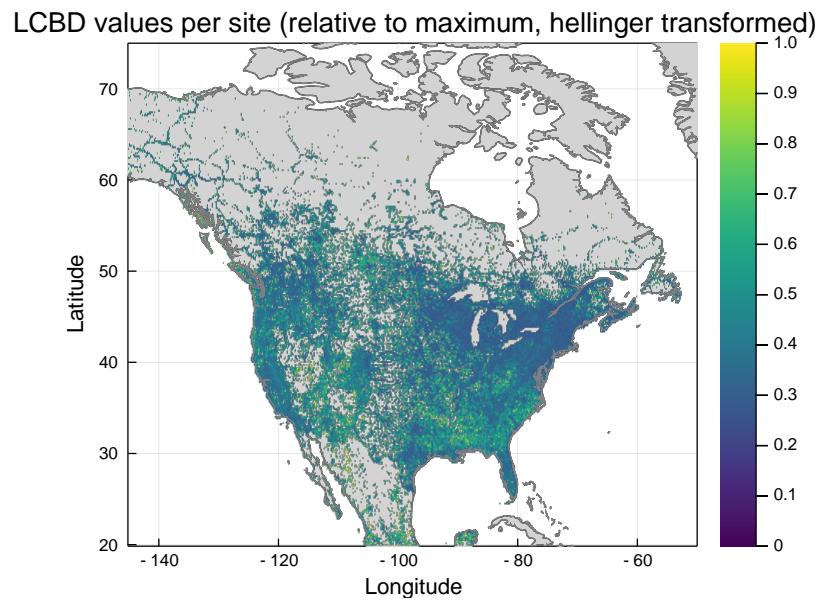


Figure 5: LCBD values - Raw (transformed)

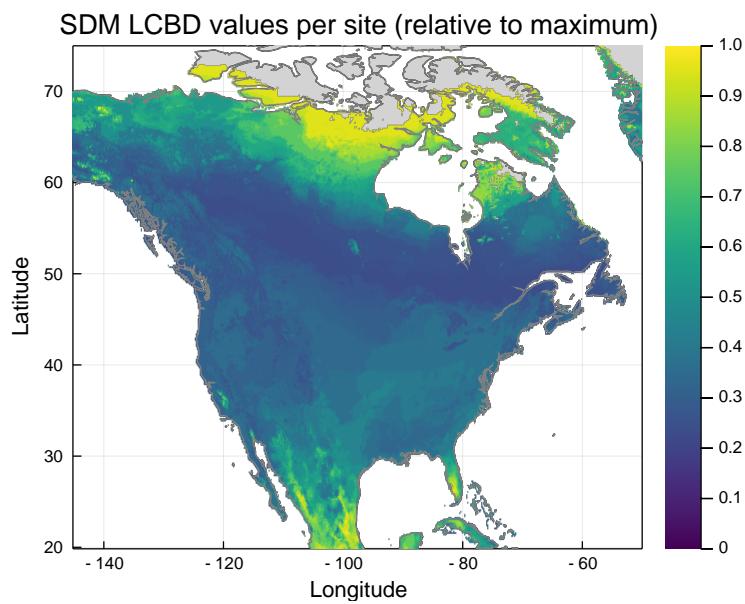


Figure 6: LCBD values - SDM

Relationship between LCBD (hellinger transformed) and species richness

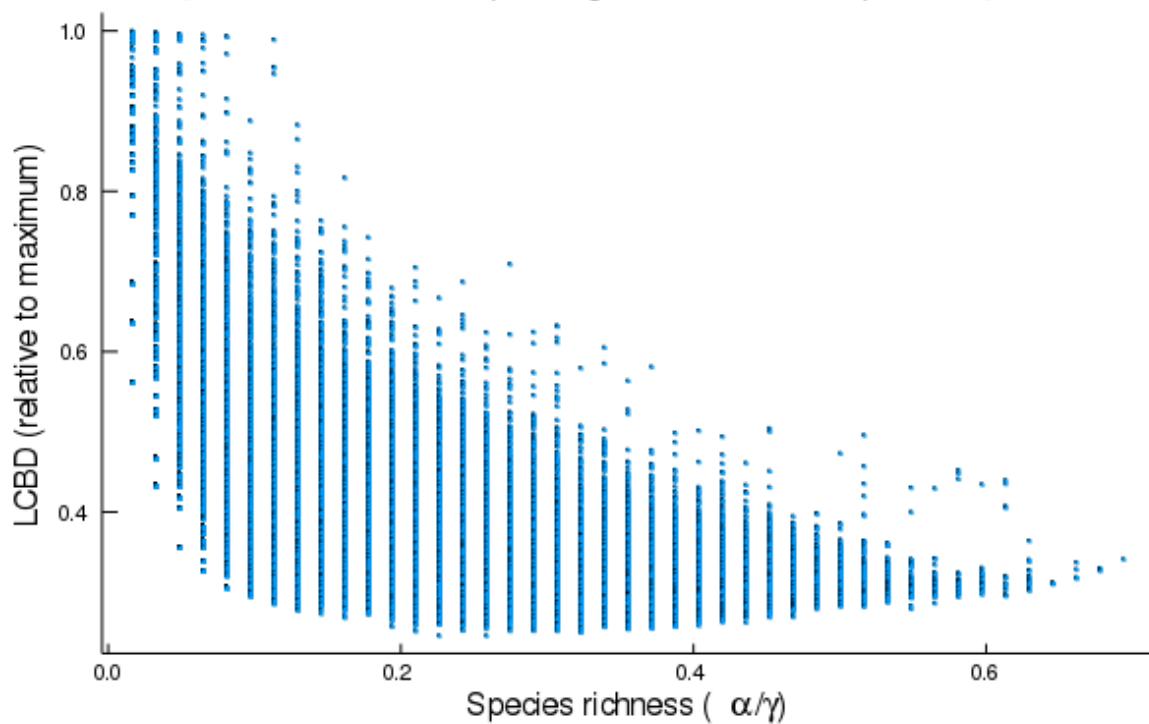


Figure 7: LCBD-richness relationship - Raw

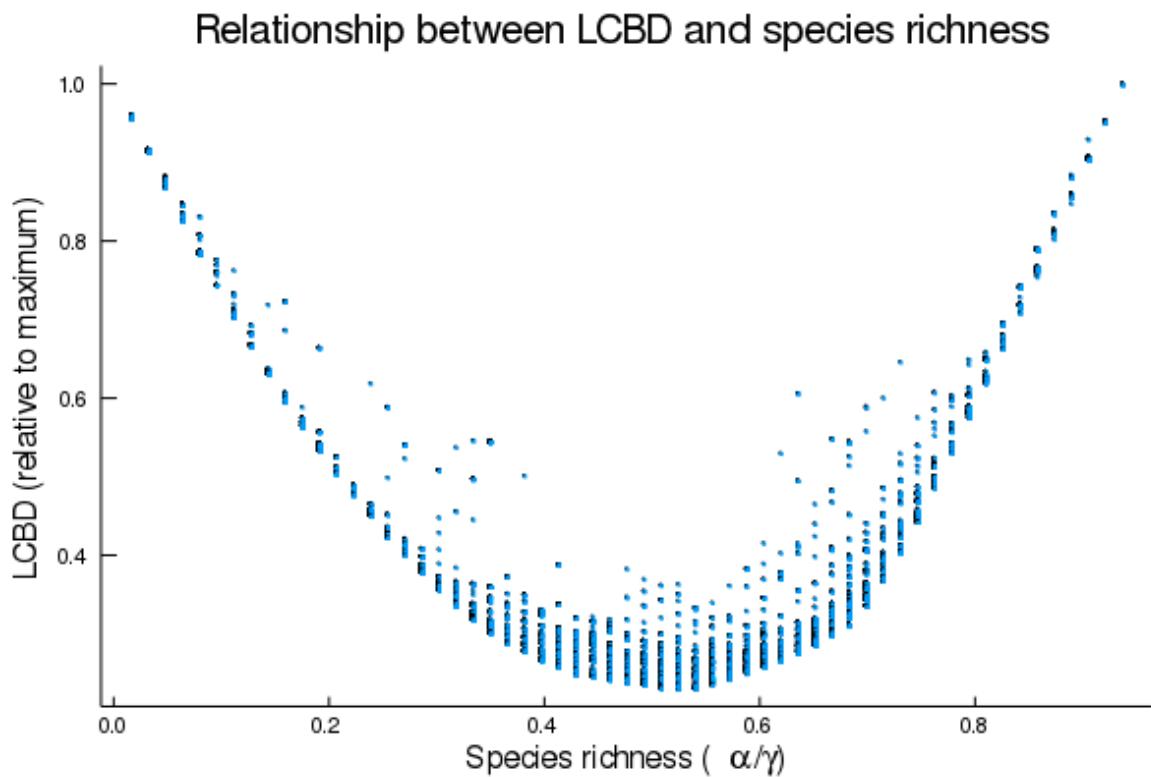


Figure 8: LCBD-richness relationship - SDM

298 References

- 299 Araújo, Miguel B., and A. Townsend Peterson. 2012. "Uses and Misuses of Bioclimatic
300 Envelope Modeling." *Ecology* 93 (7): 1527–39. <https://doi.org/10.1890/11-1930.1>.
- 301 Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B. Shah. 2017. "Julia: A Fresh
302 Approach to Numerical Computing." *SIAM Review* 59 (1): 65–98. <https://doi.org/10.1137/141000671>.
- 303
- 304 Booth, Trevor H., Henry A. Nix, John R. Busby, and Michael F. Hutchinson. 2014. "BIOCLIM:
305 The First Species Distribution Modelling Package, Its Early Applications and Relevance to
306 Most Current MaxEnt Studies." *Diversity and Distributions* 20 (1): 1–9. <https://doi.org/10.1111/ddi.12144>.
- 307
- 308 Elith, Jane, Catherine H. Graham, Robert P. Anderson, Miroslav Dudík, Simon Ferrier, Antoine
309 Guisan, Robert J. Hijmans, et al. 2006. "Novel Methods Improve Prediction of Species'
310 Distributions from Occurrence Data." *Ecography* 29 (2): 129–51. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>.
- 311

- 312 Fick, Stephen E., and Robert J. Hijmans. 2017. "WorldClim 2: New 1-Km Spatial Resolution
313 Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37 (12):
314 4302–15. <https://doi.org/10.1002/joc.5086>.
- 315 Franklin, Janet. 2010. "Moving Beyond Static Species Distribution Models in Support of
316 Conservation Biogeography: Moving Beyond Static Species Distribution Models." *Diversity*
317 and *Distributions* 16 (3): 321–30. <https://doi.org/10.1111/j.1472-4642.2010.00641.x>.
- 318 Heino, Jani, and Mira Grönroos. 2017. "Exploring Species and Site Contributions to Beta
319 Diversity in Stream Insect Assemblages." *Oecologia* 183 (1): 151–60. <https://doi.org/10.1007/s00442-016-3754-7>.
- 321 Hijmans, Robert J., Steven Phillips, John Leathwick, and Jane Elith. 2017. *Dismo: Species*
322 *Distribution Modeling*. <https://CRAN.R-project.org/package=dismo>.
- 323 Johnston, A., W. M. Hochachka, M. E. Strimas-Mackey, V. Ruiz Gutierrez, O. J. Robinson,
324 E. T. Miller, T. Auer, S. T. Kelling, and D. Fink. 2019. "Best Practices for Making Reli-
325 able Inferences from Citizen Science Data: Case Study Using eBird to Estimate Species
326 Distributions." *bioRxiv*, March, 574392. <https://doi.org/10.1101/574392>.
- 327 Legendre, Pierre. 2019. "A Temporal Beta-Diversity Index to Identify Sites That Have
328 Changed in Exceptional Ways in Space-Time Surveys." *Ecology and Evolution* 9 (6): 3500–
329 3514. <https://doi.org/10.1002/ece3.4984>.
- 330 Legendre, Pierre, Daniel Borcard, and Pedro R. Peres-Neto. 2005. "Analyzing Beta Diver-
331 sity: Partitioning the Spatial Variation of Community Composition Data." *Ecological*
332 *Monographs* 75 (4): 435–50. <https://doi.org/10.1890/05-0549>.
- 333 Legendre, Pierre, and Richard Condit. 2019. "Spatial and Temporal Analysis of Beta Diversity
334 in the Barro Colorado Island Forest Dynamics Plot, Panama." *Forest Ecosystems* 6 (1): 7.
335 <https://doi.org/10.1186/s40663-019-0164-4>.
- 336 Legendre, Pierre, and Miquel De Cáceres. 2013. "Beta Diversity as the Variance of Community
337 Data: Dissimilarity Coefficients and Partitioning." *Ecology Letters* 16 (8): 951–63. [//doi.org/10.1111/ele.12141](https://doi.org/10.1111/ele.12141).
- 339 Nix, Henry A. 1986. "A Biogeographic Analysis of Australian Elapid Snakes." *Atlas of Elapid*

340 *Snakes of Australia* 7: 4–15.

341 Phillips, Steven J., Robert P. Anderson, and Robert E. Schapire. 2006. “Maximum Entropy
342 Modeling of Species Geographic Distributions.” *Ecological Modelling* 190 (3): 231–59.
343 <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.

344 Poisot, Timothée, Richard LaBrie, Erin Larson, Anastasia Rahlin, and Benno I. Simmons. 2019.
345 “Data-Based, Synthesis-Driven: Setting the Agenda for Computational Ecology.” *Ideas in
346 Ecology and Evolution* 12 (July). <https://doi.org/10.24908/iee.2019.12.2.e>.

347 Pollock, Laura J., Reid Tingley, William K. Morris, Nick Golding, Robert B. O’Hara, Kirsten
348 M. Parris, Peter A. Vesk, and Michael A. McCarthy. 2014. “Understanding Co-Occurrence
349 by Modelling Species Simultaneously with a Joint Species Distribution Model (JSDM).”
350 *Methods in Ecology and Evolution* 5 (5): 397–406. <https://doi.org/10.1111/2041-210X.12180>.

352 Sullivan, Brian L., Christopher L. Wood, Marshall J. Iliff, Rick E. Bonney, Daniel Fink, and
353 Steve Kelling. 2009. “eBird: A Citizen-Based Bird Observation Network in the Biological
354 Sciences.” *Biological Conservation* 142 (10): 2282–92. <https://doi.org/10.1016/j.biocon.2009.05.006>.