Spatially continuous identification of beta diversity hotspots using species distribution models

Gabriel Dansereau[1,2,3]

[1]Université de Montréal

[2]BIOS²

[3]Quebec Center for Biodiversity Science

Advisory Committee Meeting
December 5, 2019

Université de Montréal

bios²

CENTRE DE LA SCIENCE DE LA BIODIVERSITÉ DU QUÉBEC

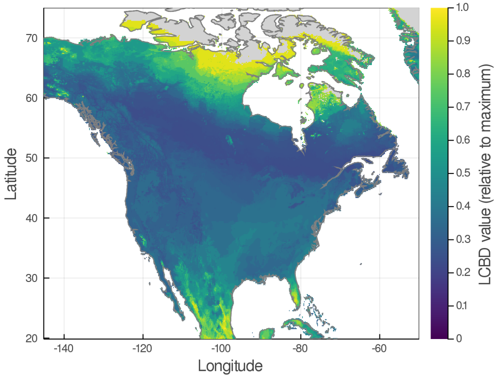
QUEBEC CENTRE FOR BIODIVERSITY SCIENCE

Fonds de recherche Nature et technologies Québec

Objective

Bring together 2 elements:

1. Identification of beta diversity hotspots → LCBD calculation
2. Species distribution modelling on continuous scales → SDMs



Why continuous scales?

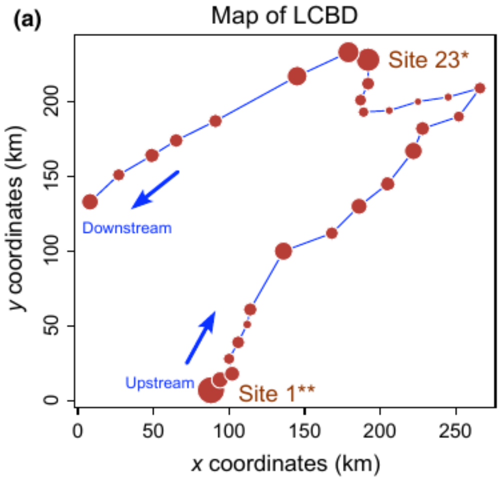


Figure 1: Example of discontinuous LCBD calculation along a river stream (Legendre & De Caceres, 2013)

Why continuous scales?

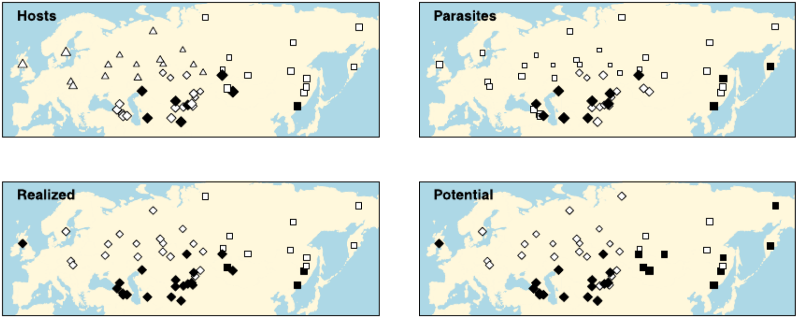


Figure 2: Example of discontinuous LCBD calculation on an extended scale (Poisot et al., 2017)

Why continuous scales?

- Online data on extended scales is increasingly accessible
- Potential for novel ecological insights

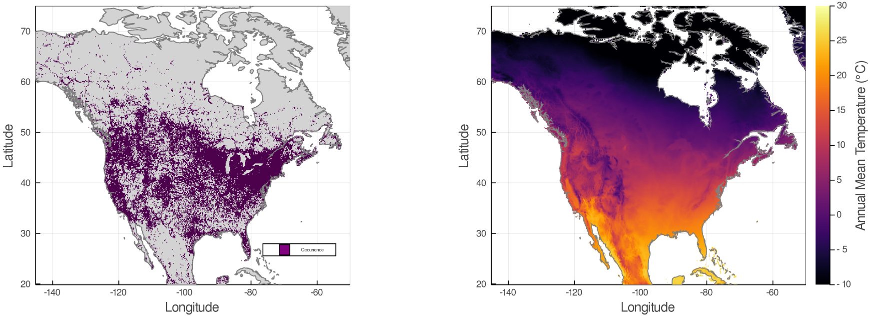
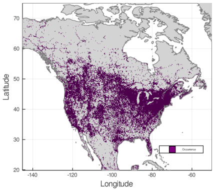
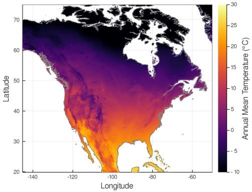


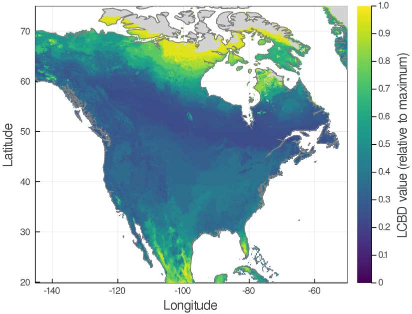
Figure 3: Example of Yellow Warbler occurrence data from eBird (left) and annual mean temperature data from WorldClim 2 (right)

Objective





$\Downarrow$



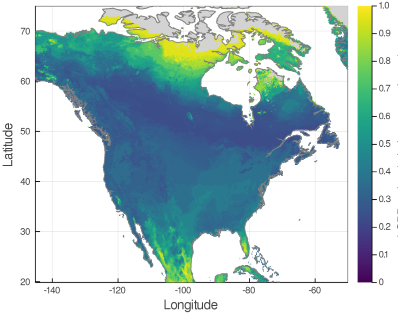
Relevance

Novel ecological insights

- Tool for poorly sampled regions, or with sparse sampling
- Identification of conservation targets

Combination with IPCC climate change scenarios

- Model beta diversity changes
- Identify sites with significant changes



$\Rightarrow$ Insight-oriented approach, exploratory analyses for now

Data - Why eBird & Warblers

According to Johnston et al. (2019):

1. Complete checklists to infer absences
2. Sampling effort metadata to reduce biases

Table 1: Structure of the Warblers (*Parulidae*) occurrence data for North America as checklists in the eBird Dataset

Country	Observations	Checklists	Species	Species per checklist (mean)	Species per checklist (median)	Species per checklist (max)
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

Data - Why WorldClim 2

- Interpolated climate data
- Global range
- Resolutions from 10 arc-minutes to 30 arc-seconds
- High cross-validation coefficients

Variable	Description
1	Annual Mean Temperature
2	Mean Diurnal Range
3	Isothermality
4	Temperature Seasonality
5	Max Temperature of Warmest Month
6	Min Temperature of Coldest Month
7	Temperature Annual Range
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
16	Precipitation of Wettest Quarter
17	Precipitation of Driest Quarter
18	Precipitation of Warmest Quarter
19	Precipitation of Coldest Quarter

Methods - BIOCLIM

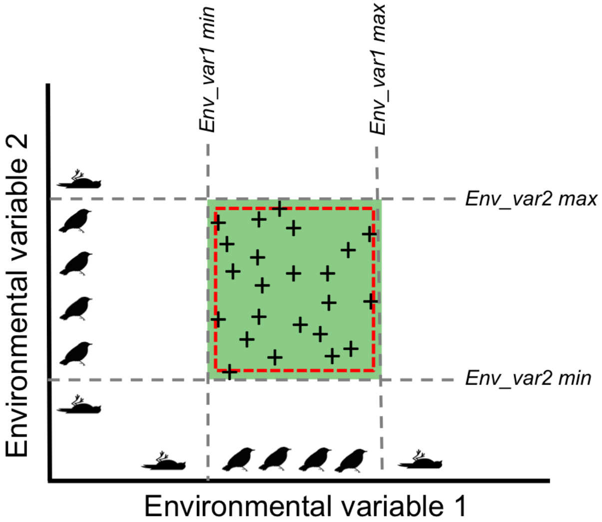


Figure 4: Representation of the climate envelope in the BIOCLIM method[1]

[1] https://support.bccvl.org.au/support/solutions/articles/6000083201-bioclim

Preliminary Results

Single species example - Raw data

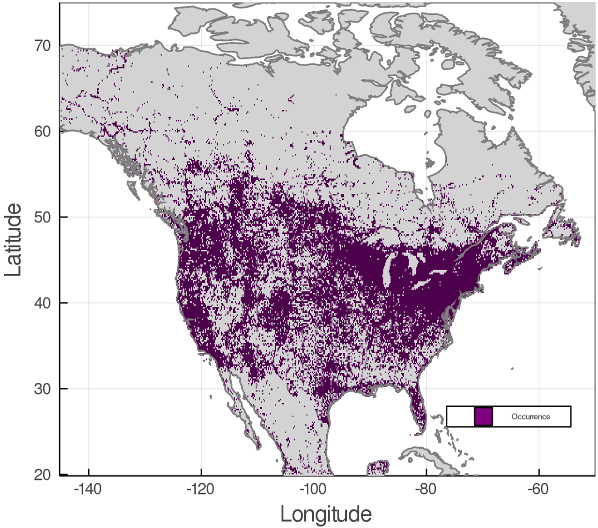


Figure 5: Distribution of the Yellow Warbler based on the raw occurrence data after transformation into presence-absence per site

Single species example - SDM with threshold

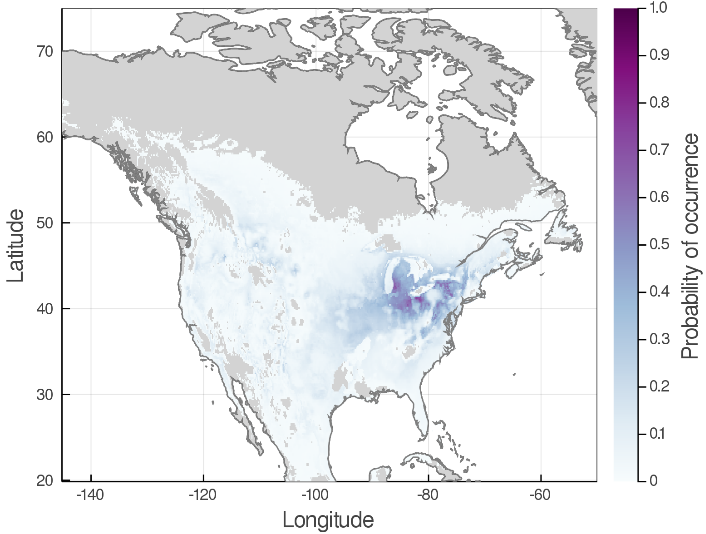


Figure 6: Distribution of the Yellow Warbler based on the SDM predictions with a threshold of 5%

Single species example - SDM without threshold

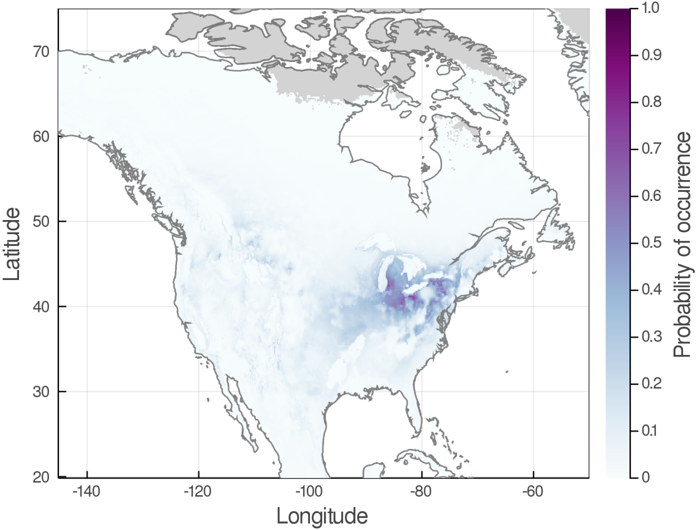


Figure 7: Distribution of the Yellow Warbler based on the SDM predictions without a threshold

Species richness - Raw data

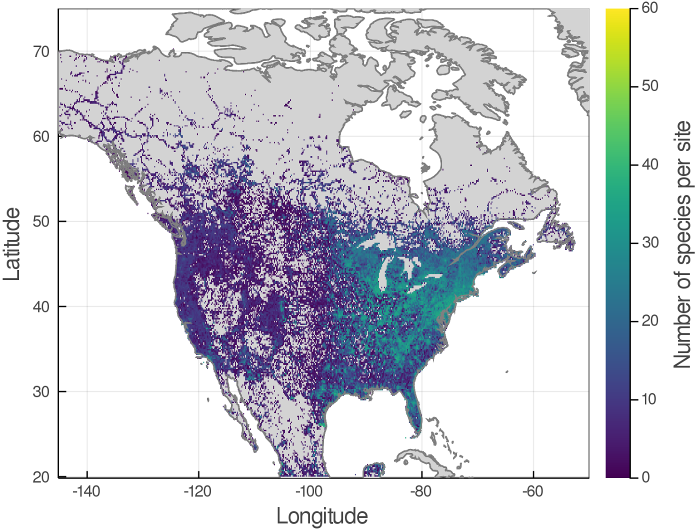


Figure 8: Species richness based on the raw occurrence data, defined as the number of species present per site

Species richness - SDM without threshold

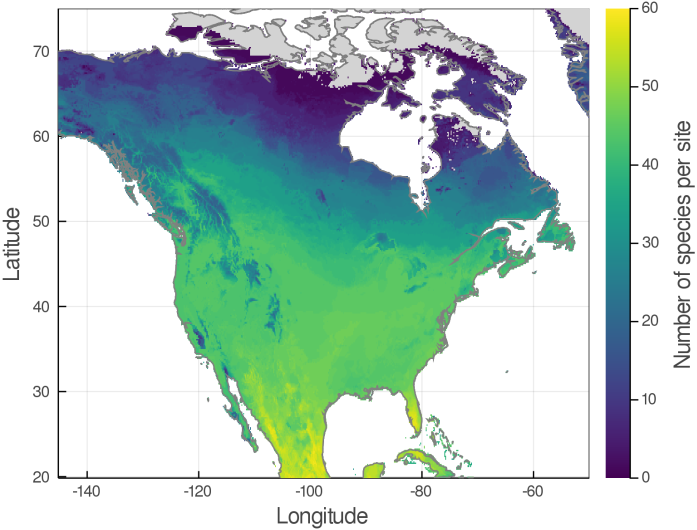


Figure 9: Species richness based on the SDM predictions without threshold, and defined as the number of species present per site

LCBD - Raw data (with Hellinger transformation)

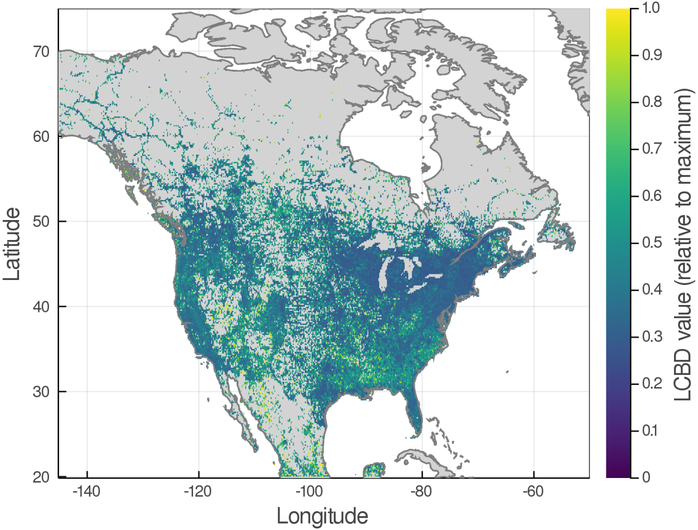


Figure 10: LCBD values relative to maximum value based on the raw data after Hellinger transformation

LCBD - SDM without threshold (no transformation)

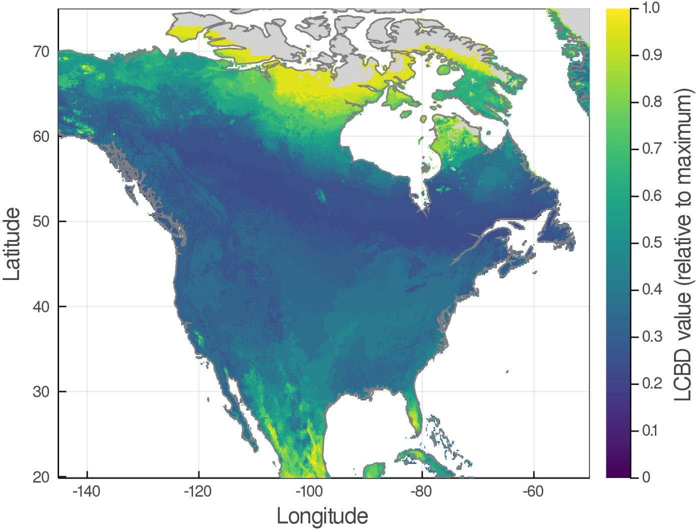


Figure 11: LCBD values relation to maximum value based on the SDM predictions without threshold or transformation

LCBD-richness relationship

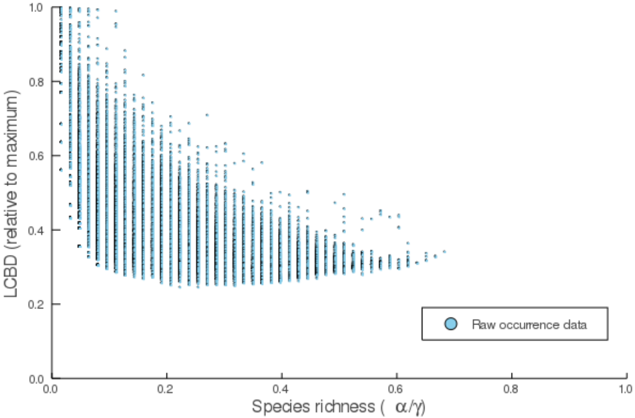


Figure 12: Relationship between the relative LCBD values and species richness, defined as the number of species (α) divided by the total number of species (γ)

LCBD-richness relationship

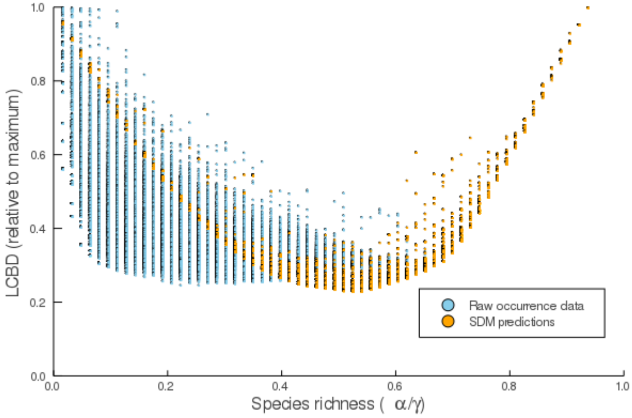


Figure 13: Relationship between the relative LCBD values and species richness, defined as the number of species (α) divided by the total number of species (γ)

Elements to discuss

1. Validating results
   - Method for prediction error
   - Data for evaluation
2. Improving SDM predictions
   - Better modelling of presence-absence through environmental conditions
   - Integrating other drivers of species distributions
3. Improving LCBD calculation understanding values
   - Data transformation
   - Relationship interpretation
4. Determining scale & resolution to focus on

References

Johnston, A., W. M. Hochachka, M. E. Strimas-Mackey, V. Ruiz Gutierrez, O. J. Robinson, E. T. Miller, T. Auer, S. T. Kelling, and D. Fink. 2019. "Best Practices for Making Reliable Inferences from Citizen Science Data: Case Study Using eBird to Estimate Species Distributions." bioRxiv, March, 574392. https://doi.org/10.1101/574392.

Legendre, Pierre, and Miquel De Cáceres. 2013. "Beta Diversity as the Variance of Community Data: Dissimilarity Coefficients and Partitioning." Ecology Letters 16 (8): 951–63. https://doi.org/10.1111/ele.12141.

Poisot, Timothée, Cynthia Guéveneux-Julien, Marie-Josée Fortin, Dominique Gravel, and Pierre Legendre. 2017. "Hosts, Parasites and Their Interactions Respond to Different Climatic Variables." Global Ecology and Biogeography 26 (8): 942–51. https://doi.org/10.1111/geb.12602.