Building a better metaweb: predicting spatiotemporally explicit plant-pollinator networks

Michael D. Catchen ^{1,2} Paul CaraDonna ^{3,4} Jane E. Ogilvie ³ Francis Banville ³ Dominique Caron ^{1,2} Philippe Desjardins-Proulx ^{5,2} Norma R. Forero-Muñoz ^{5,2} Dominique Gravel ^{6,2} Laura Pollock ^{1,2} Tanya Strydom ^{5,2} Timothée Poisot ^{5,2} Julian Resasco ⁷ Andrew Gonzalez ^{1,2}

Correspondance to:

Colorado Boulder

Michael D. Catchen — michael.catchen@mail.mcgill.ca

This work is released by its authors under a CC-BY 4.0 license

Last revision: May 6, 2022

 ¹ McGill University
² Québec Centre for Biodiversity Sciences
³ Rocky Mountain Biological Laboratory
⁴ Chicago Bontanic Garden
⁵ Université de Montréal
⁶ Université de Sherbrooke
⁷ University of

Purpose: This template provides a series of scripts to render a markdown document into an interactive website and a series of PDFs.

Motivation: It makes collaborating on text with GitHub easier, and means that we never need to think about the output.

Internals: GitHub actions and a series of python scritpts. The markdown is handled with pandoc.

1 Abstract

- 2 Using a data set of [DESCRIBE EACH DATASET IN A NICE WAY], we predict a spatiotemporally explicit
- metaweb of interactions between bumblebees (Bombus) and wildflowers (within find clade). We integrate
- 4 this data with crowdsourced occurrence data and climate data to [best paint the picture of the Colorado
- 5 bumblebee-plant metaweb]. Using temporal climate data, we forecast how the spatiotemporal overlap of
- 6 interacting species will change under proposed climate scenarios. We use this to estimate what
- 7 interactions between bees and plants need the most attention to prevent the spatiotemporal decoupling of
- 8 an interactions from threatening ecosystem functioning or the persistence of a species.

9 Introduction

- 50 Species interactions are important. It is ultimately interactions between individuals of different species
- that drive the structure, dynamics, and persistence of ecosystems, and the abundance and diversity of the
- species within them. Plant-pollinator interactions specifically drive the function and persistence of
- "architecture of biodiversity" (Bascompte & Jordano 2007). However, we are far from a robust
- understanding of plant-pollinator networks. This is because sampling interactions is costly. Interactions
- vary in space and time (Poisot et al. 2015)—particularly relevent in this system (CaraDonna et al. 2014).
- 16 This is why there is interest in using models to predict interactions from sparse data (**Strydom2021?**). In
- this paper, we combine several datasets, each spanning several years, to produce spatially and temporally
- explicit predictions of the bumblebee (genus *Bombus*) and wildflower pollination network across the state
- 19 of Colorado.
- 20 We do this in two parts: (1) metaweb prediction and (2) conditioning our metaweb prediction on
- 21 co-occurrence probability. First, we build a model to predict the metaweb—the network of all
- 22 interactions, aggregated across all times and spatial locations—of *Bombus* and wildflower species across
- ²³ Colorado. (Why do this? The metaweb is more predictable than local interactions.) We do this using
- network embedding (cite?). Network embedding takes each node in the network (either a bumblebee or a
- ²⁵ wildflower) and represents it in a latent *n* dimensional space. Combination of running models on
- Temporal niche (T), Phylogenetic niche (P), Environmental niche (E), and relative abundance in
- 27 community (RA).

- 28 Second, we then use this metaweb to predict the structure of networks at specific locations and times of
- 29 year (Gravel et al. 2019). Finally we suggest a map of sampling priority, which suggests the locations to
- sample that will best improve our understanding of the Colorado *Bombus* pollination metaweb.
- Why is this good for science, what does this contribute to our understanding of plant-pollinator ints,
- networks, Bombus, predictive models, etc., and how can these results be useful.

33 Data

We use three separate field datasets to estimate the Colorado *Bombus* metaweb.

35 Methods

36

[Figure 1 about here.]

37 Metaweb Model

- 38 Phylogeny Construction
- 39 Feature Embedding
- 40 Relative Abundance
- 41 Phylogenetic features
- 42 Environmental niche features
- 43 Temporal niche features
- 44 Metaweb Model Fitting and Validation

[Figure 2 about here.]

46 Spatiotemporally Explicit Networks

- Now that we have a metaweb.....
- Figure 3: Maps over time figure and Prob(Connectance) vs. Month figure

49 Sampling Prioiritization

50 Figure 4: Uncertainty and sampling priority map

Discussion

- Bascompte, J. & Jordano, P. (2007). Plant-Animal Mutualistic Networks: The Architecture of Biodiversity.
- Annual Review of Ecology, Evolution, and Systematics, 38, 567–593.
- ⁵⁴ CaraDonna, P.J., Iler, A.M. & Inouye, D.W. (2014). Shifts in flowering phenology reshape a subalpine plant
- community. *Proceedings of the National Academy of Sciences*, 111, 4916–4921.
- Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., et al. (2019). Bringing Elton
- and Grinnell together: A quantitative framework to represent the biogeography of ecological
- interaction networks. *Ecography*, 42, 401–415.
- Poisot, T., Stouffer, D.B. & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary
- through space and time. Oikos, 124, 243–251.

Metaweb Prediction Embedding Models Predicted Metaweb Phylogenetic Embedding Logistic Regression Relative Abundance Embedding Neural Network **₹** Model fitting and validation ADABoost Regression Environment Niche Embedding **Boosted Regression Tree** Random Forest Temporal Niche Embedding Ensemble Model Time of Year **Spatiotemporal Network Prediction** Interaction probability Spatial Co-occurrence probability Temporal Co-occurrence probability

Figure 1: todo

	ROC-AUC									PR-AUC						
T+P+E+R	0.75	0.72	0.84	0.85	0.86	0.87	0.87	T+P+E+R	0.55	0.47	0.67	0.58	0.70	0.73	0.71	
T+E+R	0.76	0.78	0.84	0.83	0.85	0.87	0.86	T+E+R	0.52	0.53	0.67	0.57	0.68	0.71	0.69	
P+E+R	0.75	0.71	0.85	0.84	0.85	0.88	0.87	P+E+R	0.53	0.46	0.69	0.57	0.68	0.74	0.70	
T+P+E	0.73	0.70	0.82	0.84	0.80	0.86	0.86	T+P+E	0.51	0.44	0.65	0.67	0.61	0.71	0.68	
T+P+R	0.75	0.71	0.84	0.83	0.86	0.86	0.87	T+P+R	0.53	0.42	0.68	0.58	0.70	0.72	0.70	
E+R	0.75	0.77	0.85	0.82	0.85	0.87	0.86	E+R	0.49	0.53	0.67	0.59	0.65	0.70	0.68	
P+R	0.75	0.71	0.85	0.84	0.86	0.88	0.87	P+R	0.52	0.46	0.69	0.61	0.69	0.73	0.70	
T+R	0.75	0.77	0.84	0.81	0.85	0.85	0.85	T+R	0.49	0.52	0.68	0.53	0.68	0.69	0.68	
T+P	0.70	0.71	0.81	0.83	0.81	0.85	0.85	T+P	0.47	0.46	0.63	0.64	0.59	0.68	0.66	
T+E	0.65	0.60	0.82	0.83	0.81	0.84	0.84	T+E	0.41	0.32	0.64	0.64	0.63	0.66	0.67	
P+E	0.71	0.70	0.82	0.85	0.79	0.87	0.86	P+E	0.50	0.44	0.62	0.68	0.54	0.70	0.66	
R	0.77	0.77	0.85	0.80	0.84	0.82	0.84	R	0.52	0.52	0.68	0.47	0.67	0.64	0.65	
Е	0.64	0.52	0.79	0.81	0.80	0.81	0.82	Е	0.35	0.23	0.56	0.57	0.56	0.59	0.59	
Р	0.69	0.71	0.80	0.82	0.77	0.84	0.84	Р	0.44	0.44	0.57	0.59	0.52	0.63	0.61	
Т	0.57	0.57	0.80	0.80	0.80	0.82	0.82	Т	0.31	0.31	0.59	0.57	0.61	0.62	0.63	
	Logistic	Neural Network	ADABoost	Decision Tree	Boosted Regression Tree	Random Forest	Ensemble		Logistic	Neural Network	ADABoost	Decision Tree	Boosted Regression Tree	Random Forest	Ensemble	

Figure 2: todo