Thesis proposal

Michael D. Catchen [1,2]

[1] McGill University [2] Québec Centre for Biodiversity Sciences

Correspondance to:

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

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

Last revision: *November 14, 2021*

The proposal for my thesis, *Simulation models for predictive ecology*

Introduction

Within the last several hundred years, human activity has rapidly reshaped Earth's surface. These changes can roughly be divided into two categories: (1) Land-use change, where Earth's surface changes and (2) climate change, words here, as a result of greenhouse gas emissions.

As a result *ecological forecasting*, or building models to estimate how ecological systems will change over time, has as an imperative to mitigate the effect of these changes on Earth's ecosystems, their functioning, and the services they provide to humans (Dietze 2017).

An oft applied definition of the origin of is "the application of the scientific method to natural history." Since its origin ecology has been a descriptive science. This is a natural by-product of the immense variability of Earth's biosphere. emerged to explain particular phenomena at particular scales.

In recent years, there has been an interest in an epistemological shift in ecology. To shift ecology into a predictive science. The justification for this shift is twofold: (1) bogged down philosophy of science, by further rooting our understanding of ecosystem function and dynamics in an ability to predict their structure (Dietze 2017). and (2) the practical need for models for *ecological forecasting*.

Historically the term "theory," as applied in the physical sciences, refers to mathematical models, typically an equation describing how the value of an observable state of the system, represented by a vector of numbers $[x_1, x_2, \ldots, x_n]^T = \vec{x}$ changes as over time.

$\vec{x}(t)$ but instead to define how the state of $\vec{x}$ changes from one time to the next.

Because of its early success in the physical science, the led to framework for bridging theory and data.

A large set of problems in ecology when aiming to confront high-dimensional analytic models with data:

Often assume long-run equilibrium.

Ecological processes vary across more variables than the tools of analytic models are suited for.

As the number of variables in an analytic model increases, so does the ability of the scientist to decern clear relationships between them, and so does overfitting potential.

Curse of dimensionality— Until the 20th century, no theory of the gravitational dynamics of more than 2 bodies. Understanding the gravitational dynamics of more than two planets with any reliability proved difficult. Using the same models (diffeqs), how could we adequately predict ecosystems?

Chaotic dynamics emerge from simple analytic models, and . Whether ecosystems actually exhibit chaotic behavior is a different question.

The term *ecological forecasting* implicitly creates an analogy between predicting how ecosystems will change in the future and weather forecasting. Use of computational methods in NWP. Much as one would not aim to forecast the weather in Quebec by applying Navier-Stokes. NWP has worked because it incorporates information about data and meteorological processes collected at difference scales into models that.

Transition to simulation as the solution: shift toward approach of building models that *generate* data.

(resolving the semantic ambuity of what differentiates “mechanistic” vs “phenomological” models is out of scope for now). —-

Transition to theme of optimization given unknown information. A forecast gives us a range of future values with uncertainty around them. Further a convenient property that a forecasting model’s uncertainty goes up over time (if we assume the underlying process is Markov–this is a strong assumption but oft true of the models we fit to temporal data)

In face of uncertainty, decision making is an optimization problem. We have some goal state for the future, and some estimate of what the state of the world will be given a set of actions. Frame optimization problem mathematically an introduce concept of solution-space and constraint.

Indeed Marx’s most well known quote that “philosophers have hitherto only interpreted the world in various ways; the point is to change it.”

and a necessary step toward establishing a just and sustainable world.

Transition to specifics of this thesis.

[Figure 1 about here.]

CH1 Forecasting the spatial and phenological uncoupling of a plant-pollinator network

This chapter uses several years of data on bee-flower phenology and interactions, combined with spatial records of species occurrence via GBIF, to predict the probability of each realized interaction network as a function of location and time.

Two ways in which this network of interactions can become uncoupled: spatial and temporal. Overlap in ranges and shifts in ranges. Elevational gradient as proxy for range shifts

Data

System description: lots of data on *Bombus* (bumblebees) and wildflowers. Three different sites, (7/7/3) years each, each covering an elevational gradient.

Methods

Split the process into parts.

1) Building an interaction prediction model.
2) Make it spatial based on distributions.
3) Forecast distributions based on CMIP6.

Preliminary Results

Transition to next chapter by discussing uncertainty in interaction prediction across space.

CH2 optimizing sampling of interactions

This chapter discusses the effect of species relative abundance on samples of interaction data, and proposes a method for optimizing spatial sampling of a possible interaction between species as a function of the estimated distribution of both species.

Methods

- the missing link paper, turn this into optimizing with two different SDMs

- relative abundance and its effect on false negative

- non-independent associations in samples
- simulate species distribution and efficacy of detection given a set of observation points where the dist from observation site decays.
- optimize set of repeated sampling locations L for a *known* distribution D.
- address SDM not being the territory

Results

CH3 optimizing corridor placement

This chapter proposes an algorithm for optimizing (corridorplacement/restoration effort) given a raster where each cell indicates land-cover. The optimization method uses the result of a simulated process (specifically occupancy dynamics in the landscape) and uses simulated annealing to estimate the global optimum of the targetstate (specfically mean-time-to-extinction for the occupancy dynamics example).

Methods

- land cover -> resistance -> extinction time
- simulated annealing to optimize landscape optimization

CH4 a software note on the resulting packages.

(MetacommunityDynamics.jl: a virtual laboratory for community ecology): a collection of modules in the Julia language for different aspects of metacommunity ecology, including most of the code used for the preceding chapters.

- TK conceptual figure with interfaces between what I'm writing / have contributed to and linked with other libraries
- `Observatories.jl`, `Corridors.jl`, `MCD.jl`

References

Dietze, M.C. (2017). Prediction in ecology: A first-principles framework. *Ecological Applications*, 27, 2048–2060.

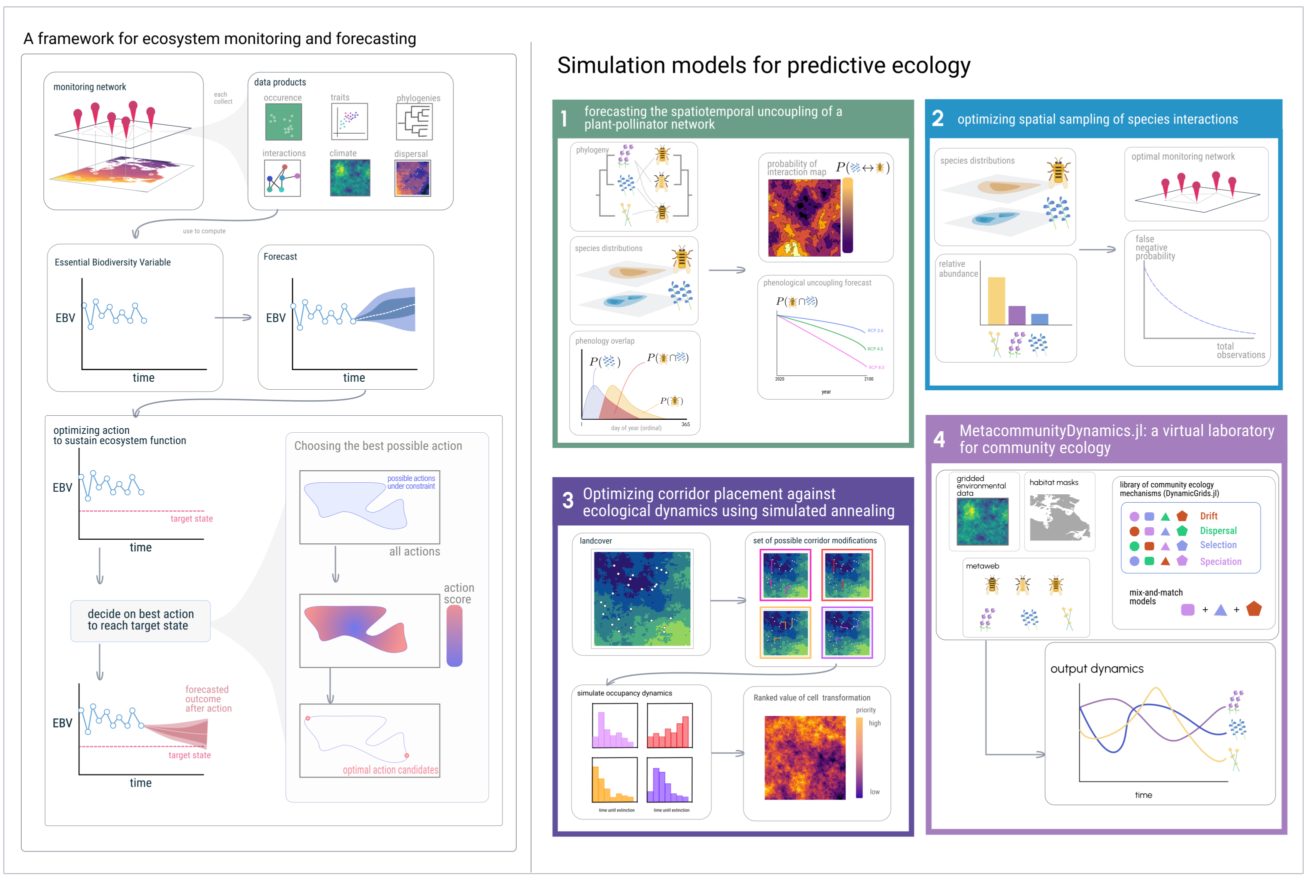

Figure 1: thesis concept