

Meetings Spark Project

Project Overview

Boolean networks as a way of modeling ecosystems services

Meeting 05/08/2025

Alvaro, Bruno, Carlos, Paulinha at EAWAG

Brainstorming ideas and some key points that were made:

- Usually in Boolean networks there is no population dynamics associated and it would be a great contribution to integrate these two
- Even more if we integrated Stochastic Hybrid Systems (SHS) with Boolean Networks in a community ecology context. In SHS some populations are so rare that their dynamics could be considered stochastic, whereas other populations would follow classic deterministic dynamics
- What Bruno has started involves understanding how costs and benefits of interactions influence the persistence of species within a networks. For that he uses the structure of empirical mutualistic networks and associated costs and benefits for each species. When benefits surpass costs, species are “on” and present on the networks, as the cost of interaction increases, once it surpasses the benefits, species get extinct from the network. He found a sharp transition where above a certain cost, most species go extinct. Costs and benefits are randomly sampled from a distribution. Next steps to explore include:
 - What are the analytical expectation for the boundary conditions (having all species present when $benefits > costs$ and all species extinct when $benefits < costs$ considering normally distributed costs and benefit functions)\$
 - Expectations for random networks, what are the final number of interactions when network is random and cost benefit functions are also random
 - Starting from a random network, what kind of cost and benefit functions would result in networks that resemble the structure of empirical networks? Perhaps here we need to explore different distributions (beyond uniform or normal, for example, beta or lognormal) to give rise to networks with asymmetries like we observe in empirical networks (nestedness patterns for example, which are a common structure found in mutualistic systems)
 - Connection of boolean networks, costs and benefits and game theory – what is the threshold for a Nash equilibrium considering costs and benefits in mutualisms? Is it possible to find an equilibrium even when the cost of the interaction is not zero?
 - Asymmetries in the costs benefit functions (shifting the curves to understand when the system collapses)
- How environmental fluctuations influences the predictability of ecosystem services
 - Fluctuations from the topics to the pole
 - O_2 as a global service *vs* pollination as a local service

Meeting 12/08/2025

Bruno, Carlos & Paulinha @ EAWAG

- Discussion of the probability of all species being present (or extinct) given the distribution of costs and benefits following a normal distribution with means μ_1 and μ_2 and standard deviations σ_1 and σ_2 . The derivation is on the Notes folder
 - Next step: understand the shape of the decay – does it follow the same pattern observed for ecological networks?
- **Environmental function** Within the cost function, there is a physiological cost, which can be associated with an increase in cost given a deviation from an environmental optimum θ . This would become: $\alpha = |\theta - z_i|$, where z_i is the optimum for species i and θ can be either sampled or alternatively it can include environmental fluctuations (akin to the proposal and published paper by Carlos and Paulo). The added environmental fluctuation could be of the form $\theta = \frac{A}{2} (1 + \sin(2\pi\omega t))$
- **Specialization function** considering that benefits cannot accumulate indefinitely, we can incorporate a saturation function into the benefit function. With this we can explore the trade-off between number of partners and costs and benefits of interactions.
 - Consider bottom-up assembly building networks? For now we are considering a “top-down” approach in which the interaction network is given (might be the simplest to keep this and incorporate the saturation when parameterizing the cost-benefit functions). The form of this function is:

$$S = \frac{\left(\sum_j^k A_{ij}B_j\right) B_{max}}{\beta + \sum_j^k A_{ij}B_j}$$

- Connection between network structure and ecosystem services: how are the collapse of networks related to the loss of ecosystem services? Services are lost before the network collapses but what are the “early signs” that can signal such a shift? If we find literature exploring this aspect we can use it to parameterize the ecosystem service in a Boolean way. For example, imagine that one of these signals is the change in one of the properties of the network, as long as that property holds (degree of nestedness, for example), the service is “on”, after a certain threshold, even if you have species and interactions the service is no longer available. In one extreme you can imagine that a single species performs the service. In that case, losing that species would lead to the loss of the service as well. Here we can explore how far a network can be “perturbed” (loss of interaction due to increase in cost) until the service is lost, based on information on properties of the network.

Meeting 14/08/2025

Bruno, Carlos, Paula & Paulo - Zoom

- Aiming for infinity in 5 steps.

For ecosystem services there is probably not much data available but what we can do is build expectations through a theoretical map. How general can we make the rules or does it have to be system specific? How does the topology of the system influences the provision of ecosystem services? Keep in mind potential trade-offs in the maintenance of services and the idea of services and de-services.

1. Theoretical Map

Develop a framework to understand expectations that can generate predictions to be empirically tested. One example could be to explore ecosystem services that rely on species considered the “core” of the network *versus* services that dependent on species that are at the periphery of the network. Additionally we can consider multiple services such as pollination, seed dispersal, food provision, bird watching,...

2. Case studies

- Ecosystem services and indigenous knowledge networks (Cámara-Leret, Fortuna, and Bascompte 2019)
- Bird watching

3. Localities

- Pantanal could be an interesting system to explore (Donatti et al. 2011)

4. Spatial distribution of services (focused on one service)

- Exploring uneven distribution of services across space and time.
- Conflict between micro and macro scales: the services emerges at the micro scale (pollination, carbon storage); process base is local but some have a global component (CO_2 , O_2). For some services we observe “pumps”
- Are there hotspots for different services?

5. Generalizations

Based on the information we gain from the combination of the previous steps, can we make generalizations about the ecosystem services and ecological networks?

- Current understanding of ecosystem service debt is mainly based on species loss and not necessarily on loss of interactions
- Static value of nature
- **Potential extensions** as a first approximation we are removing species that have a negative net benefit but another idea would be to remove interactions (instead of species); turning off interactions that are negative.

Meeting 18/08/2025

Bruno, Carlos, Paulinha @ Eawag

- Proposal of an Ecosystem Service matrix, in which services are organized in rows and columns represent species that contribute to the service. Therefore, the ecosystem services provided by species follows:

$$E = \mathbf{P}\vec{x}$$

where, \mathbf{P} is a matrix of $k \times s$ dimension, in which k is the ecosystem services vector and s is the species vector. The entries P_{ij} depicts the effect of each species j on ecosystem service i . The vector \vec{x} is related to the state of species. For instance, present or absent, abundance and so on. For our purpose:

$$E_{ij} = \begin{cases} P_{ij}, & \text{if } x_i = 1 \\ 0, & \text{if } x_i = 0 \end{cases}$$

- Based on this we can explore the structure of this matrix by:
 - Assuming services are correlated (positively or negatively)
 - Services are randomly structured
 - The role of the species in the adjacency matrix determines the services it proves (core vs periphery)
 -
- **Keeping in mind** Currently our assumption is of a Boolean function operating in an static environment; fluctuations can arise via parameter α regarding costs of interactions, influencing the cost-benefit function.

- **Next steps** Challenges on allowing species to come back to the network, there is a problem of the arbitrary choice of species order when estimating CB function; another challenge is to establish when to stop simulations if species keep changing. Regarding the second challenge we discussed including species “life cycles”, which arise naturally when thinking about the fluctuating environments in the α parameter (e.g., follow 5 complete cycles) and another way would be to establish an extinction threshold in which we create a vector v and species can persist in the network even when they are not interacting/have negative benefits for a certain number of Boolean iterations (which would be akin to a “cycle”)

Meeting 18/08/2025

Bruno, Carlos, Paulinha @ Eawag

- Discussion around the chose of beta distribution as the distribution to set costs and benefits parameters. This distribution was chosen for some reasons:
 - The values are all positive. As do not exist negative cost or negative benefit, it fit to our data;
 - The values are defined on the interval [0,1]. It could be used as the percentage of costs and benefits;
 - As the values are bounded between 0 and 1, we know the maximum value of benefits for one species, that is $B \leq k_i$, where k is the degree of i ;
 - The shape of the curve assume different forms as a function of the parameters. In nature we really do not know the shape of the distributions of costs and benefits;
 - It have a solution for our case, which the other option of distribution (log normal) do not.

Meeting 22/08/2025

Bruno, Carlos, Paulinha @ Eawag

- The discussion move around the theoretical expectations of beta distribution. We can obtain it by the PDF of the difference of two *beta* distributions, $\theta_d = \theta_1 - \theta_2$ and can be derived using the Appell Hypergeometric function ($\$F_1$).
- During the discussion, we figure out that to find the theoretical expectations of species extinction, we need to identify the expectations for all species or to individual species. If we want the expectations for all species if is dependent of the number of species at the network. Therefore, the expectation for all species being extinct or be maintained at the networks follows:

$$E(x) \sim \text{beta}(\alpha, \beta)^N$$

where, N is the number of species.

- During the conversation, to reach the theoretical expectations we identified that we need a little change at the model, as it could be impossible to find the expectations. In the original version of the model, we are summing rows of costs and benefits and subtracted than. Now, we subtracted the benefits by the ecological costs. It allows us to find the theoretical expectations and make the model more elegant, since we can now segregate between a biotic and an abiotic component. The biotic ones comprehends the benefits and the ecological costs. The abiotic component comprehends the physiological costs and the alpha value, which is related with environmental fluctuations. Therefore,

$$w_i = \sum_j (B_{ij} - C_{ij}^e) - C^p \alpha$$

where w is the fitness of species i . C^e is the ecological cost and C^p is the physiological cost.

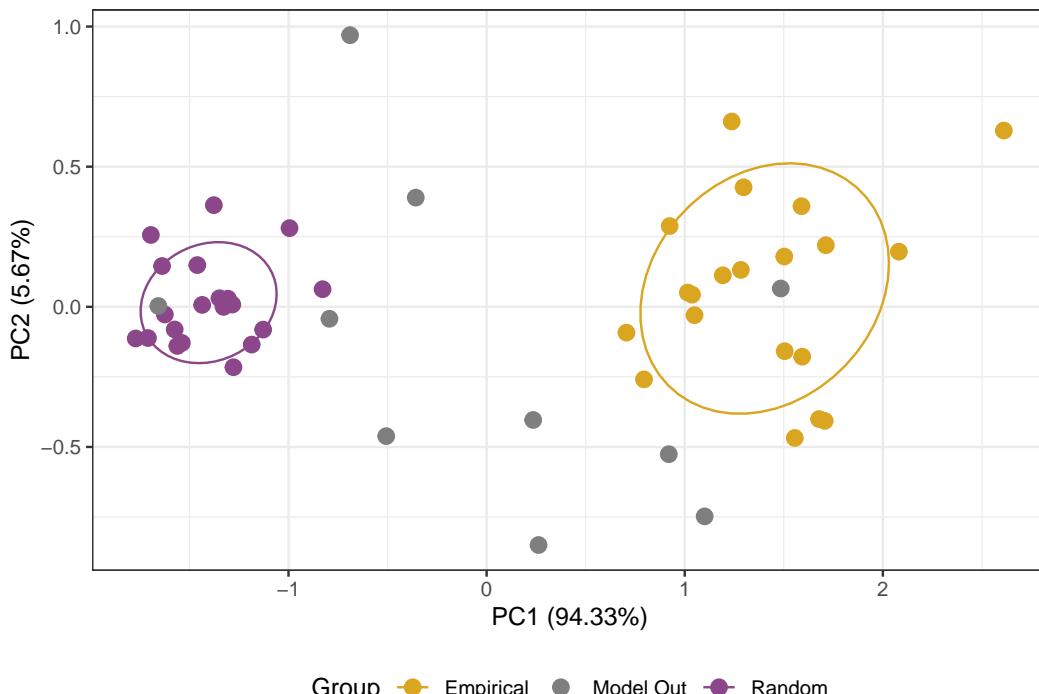
- We also pass through an idea of adaptive landscape driven by mutualistic interactions and environmental fluctuations where mutualistic interactions favor species to disperse to less favorable environments. The presence of species in less favorable environments depend on the correlations of presence of partner species and how distant a species are from their environmental optima.

Meeting 08/12/2025

Bruno, Carlos, Paulinha and Marcelo @ Puxadinho IB - USP

Robustness Challenge Discussion * Recap of the model considering the single framework and the two routes being taken by Paulinha and Bruno, identifying the main challenge which is **How robust is our framework to parameter choice**. We need to address this issue and be confident of the robustness of our model/parameter choices.

- Paulinha is leading the route considering how processes encoded in our model framework lead to the emergence of empirical patterns, whereas Bruninho is leading the route that considers empirical patterns as a starting point and explores the consequences of our framework to these patterns.
- How will the inference be done? – in Paulinha's case, where we start from a random expectation with all possible interactions and use to model to “prune” interactions as a framework to understand processes shaping ecological systems. One route for inference would be the reproducibility of networks properties. In this case, we could do a PCA of the network metrics (initial ones and the output ones comparing with the empirical distribution) and see how parameter combinations (i.e, symmetry and shape of cost benefit functions) lead to properties that are similar to the ones observed in empirical systems. An expectation would be a PCA where empirical networks cluster together with respect to their properties, whereas random (initial) networks would cluster separately and the networks emerging from model output would cluster together with the empirical ones. Alternatively, we might not see a signal in there (see below an example of a PCA with simulated data to show how this analysis would look like)



- The real challenge here is to explore the parameter space of the cost/benefit functions, it is huge! One possible solution is to derive the Markinho and analytical approximations. Ideally we can develop a generalized cost function allowing us to incorporate different types of noise

- Developing a neutral Markov Chain/Process that would allow to deal with exploration of model parameters and sensitivity to initial conditions and choice of parameter values; we also discussed writing a descent gradient function (it could start linear and go to non-linear as well)
- Thinking about the route Bruninho is taking, studying the structure of the Ecosystem Service Network (Services by Species, which are influenced by the cost benefit functions that are a result of the adjacency matrix, defining species interactions), each of these networks can be seen as a layer (Adjacency matrix A , Cost and Benefit matrices C and B , and Ecosystem Service matrix E) and the transition matrices connecting the layers can be described as a Markov Process. But importantly considering the connection between layers and the framework: **What is the general equation in matrix form?** Writing this down will help us understand the robustness of the framework, will also allow us to quantify the number of absorbing states with the Markov Chain. One simplification we could do is to start with a single service, derive the predictions, compare with the results of the model (if we're doing things right, they should match)
- To move forward we talked about two main things that need to be done (early January):
 1. A Road Map of the project (common aspects between both routes and specific goals/questions for each one);
 2. README files for each route with input/output of the algorithm and a detailed explanation of the algorithm – this will facilitate the development of Markinho

References

- Cámará-Leret, Rodrigo, Miguel A Fortuna, and Jordi Bascompte. 2019. “Indigenous Knowledge Networks in the Face of Global Change.” *Proceedings of the National Academy of Sciences* 116 (20): 9913–18.
- Donatti, Camila I, Paulo R Guimarães, Mauro Galetti, Marco Aurélio Pizo, Flávia MD Marquitti, and Rodolfo Dirzo. 2011. “Analysis of a Hyper-Diverse Seed Dispersal Network: Modularity and Underlying Mechanisms.” *Ecology Letters* 14 (8): 773–81.