MCM Problem 2023 A

Bose Kaikini, Dao Bui, Maxwell Corwin

January 2026

1 Problem Overview

We are tasked with creating a model of a plant community that tracks its growth and adaptation over periods of irregular weather patterns and differing abiotic/biotic factors like habitat loss, pollution, and varying competition and resource usage levels.

2 Literature Insights

The first major hurdle in our plant community modeling process was the trade-off between resolution and generalization. The more generalized the model becomes, and therefore the more applicable, the less new specific information (or resolution) it gives. This is a common problem for ecosystem modeling and left us with a choice between frameworks.

The convention for individual specificity is through the implementation of SDMs, or species diversity models. This allows for very accurate and information-dense modeling about a limited number of plants. Looking at a community, however, requires a much broader lens. It is here that our chosen systems of PFGs comes into play.

Plant functional groups are groups of plant species that are collated based on their demonstration of shared traits. There are a large number of traits chosen based on factors like plant leaf thickness. Importantly, their proportion is similar to the proportion of the individual species and so allows us to use them in our model without losing information.

Next, we had to look at how competition had been implemented in plant models. We saw that for specific species resource density models were the mainstream option. These were able to model resource depletion very effectively and we used a combination of resource depletion and an iterative improvement system for our competition over water supply and a generic resources category.

We were torn between implementing a spatial model for the habitat loss section of the paper, based on what we had seen in the literature. However, we decided that especially to model a variable setting and number of PFGs the best strategy was to use our fixed counter.

We also tried to improve our understanding of our undertaking. Over time, research supported the fact that a higher biodiversity was beneficial for community viability, to an extent. Oversaturation of PFGs, or species in our case, would result in unsustainable competition and radical population drop. Clearly, there was an optimum, and we realized that the literature left gaps for us to generalize this, which was also a deliverable to the problem.

3 Assumptions

- We assume, based on the literature we studied, that the system of PFGs accurately models a community of independent species.
- Next, we also model resources as a single value (outside of water).
- Resource quantity depends only on season. A season of drought ten years into the future will have as much water and our arbitrary resource count as in the present.
- Drought and flood happen in a set proportion at the time (which we arbitrarily biased towards drought because of the setting of the question).
- Micro-evolutions happen on the timeframe we decided.

4 Summary

By employing plant functional groups (pfgs), a way of aggregating plant populations with similar traits into a single entity, we developed a general yet comprehensive model of how different pfgs would interact with each other during irregular weather patterns in arid conditions.

PFGs were randomly generated with several traits (e.g competition, reproduction per water, density). Irregular weather would randomly be selected in the model which would have the possibility of triggering competition through limiting water and resources.

We account for specific survival strategies such as niche partitioning or hibernation through a coefficient. The variation in traits and attributes within the plant functional groups enabled us to assess their performance as the number of PFGs increased or decreased, thereby simulating environments with lower or higher levels of biodiversity.

Through repeated trials, we evaluated the long-term viability associated with different levels of biodiversity and observed the effects of micro-evolution as reflected in changes in relative proportions over time.

5 Model Explanation

This model integrates a mechanistic ecological simulation with data-driven species distribution modeling. The ecological core governs population dynamics through biologically interpretable processes (survival, reproduction, competition, and environmental forcing), while machine-learning components are used to infer species distributions and abundances from environmental co-variates.

5.1 Ecological Simulation Model: Mathematical Formulation

5.1.1 Guiding Equation: Discrete Logistic Growth

The population dynamics are governed by a modified discrete-time logistic growth formulation. Starting from the standard discrete logistic equation,

$$N_{t+1} = N_t + rN_t \left(1 - \frac{N_t}{K}\right), \tag{1}$$

this model separates density-independent survival from density-dependent reproduction and introduces an environmental favorability factor.

For species i at time t:

$$P_{i,t+1} = S_{i,t} + B_{i,t}, \tag{2}$$

where survival and births are defined as

$$P_{i,t+1} = [P_{i,t} \cdot \phi_i \cdot E_{i,t}] + \left[P_{i,t} \cdot (\rho_i - 1) \cdot E_{i,t} \cdot \left(1 - \frac{P_{i,t}}{K_{param}}\right)\right]. \tag{3}$$

Factoring yields

$$P_{i,t+1} = P_{i,t} \cdot E_{i,t} \left[\phi_i + (\rho_i - 1)\left(1 - \frac{P_{i,t}}{K_{param}}\right)\right]. \tag{4}$$

5.1.2 Parameter Definitions

- $P_{i,t}$: Population of species i at time t.

- $E_{i,t}$: Environmental Favorability Factor (0.0–2.0).
- ϕ_i: Base survival rate.

- ρ_i: Base reproduction rate.
- K_{param}: Density regulation parameter (fixed reference capacity).

5.1.3 Species Traits

Each plant functional group is characterized by:

- ϵ_i: Drought resistance.
- ω_i: Water requirement.
- λ_i: Lifespan.
- σ_i: Space requirement.

Life-history strategy is determined by lifespan:

- **Annuals** ($\lambda_i \leq 2$): $\rho = 2.5$, $\phi = 0.05$.
- **Perennials** ($\lambda_i > 2$): $\rho = 1.3$, $\phi = 0.75$.

5.1.4 Environmental Logic

The environmental factor is defined as

$$E_{i,t} = F_{res}(t) \cdot F_{weather}(i,t), \tag{5}$$

linking shared resource competition and stochastic weather forcing.

Resource demand is computed as

$$D_{water} = \sum_i P_{i,t} \cdot \omega_i \cdot (1 - 0.05\epsilon_i), \quad D_{space} = \sum_i P_{i,t} \cdot \sigma_i. \tag{6}$$

Availability ratios are

$$A_{water} = \frac{Water_{supply}(t)}{D_{water}}, \quad A_{space} = \frac{Space_{supply}(t)}{D_{space}}, \tag{7}$$

with co-limitation captured by the geometric mean:

$$F_{res} = clamp\left(\sqrt{A_{water} \cdot A_{space}}, 0.2, 2.0\right). \tag{8}$$

Weather states introduce supply shocks and species-specific physiological stress. For example:

- **Drought**: $F_{weather} = 0.5 + \epsilon_i/20$.
- **Flood**: $F_{weather} = 0.6 + 0.3\,(1 - \sigma_i/10)$.

5.1.5 Equilibrium Behavior

Under constant conditions ($E_{i,t} = 1$), equilibrium occurs when $P_{t+1} = P_t$, yielding

$$N^* = K \left(1 - \frac{1-\phi}{\rho - 1}\right), \tag{9}$$

showing that long-term population size is determined by life-history strategy.

5.2 Data-Driven Species Distribution Modeling

To complement the mechanistic simulation, species distributions and abundances are estimated using machine-learning techniques, primarily random forest models. Random forests are well suited for species distribution modeling because they are:

1. **Non-parametric**, allowing flexible relationships between predictors and responses.
2. Capable of capturing **non-linear effects and hierarchical interactions** among environmental variables.
3. Effective with **presence–absence as well as abundance data**.

5.2.1 Random Forest Modeling Approaches

Maps of 28 tree species are generated using three complementary random forest-based approaches:

- **RF_Bin**: Independent binary classification models predicting presence/absence for each species (28 models).
- **RF_Abund**: Independent regression models predicting continuous basal area (m^2 ha^{-1}) for each species (28 models).
- **RF_NN**: A multivariate random forest nearest-neighbor imputation model that predicts basal area for all species simultaneously using a single model.

5.2.2 Nearest-Neighbor Imputation

The **RF_NN** approach relies on random forest proximity measures to identify ecologically similar sites in predictor space. Species abundances at unsampled locations are then imputed from the most similar observed sites, preserving realistic covariance structures among species.

5.3 Model Integration

The ecological simulation provides dynamically evolving environmental states and resource constraints, while the random forest models translate environmental conditions into spatially explicit estimates of species presence and abundance. Together, this hybrid framework combines process-based interpretability with data-driven predictive power.

6 Results

6.1 Baseline

We tested our model on various numbers of PFGs. It is important to note that if the number of PFGs looks lower than the expected number of species, this is because the PFGs are fundamentally archetypes of species and represent a much broader number. Below are the results for 3, 5, and 10 PFGS:

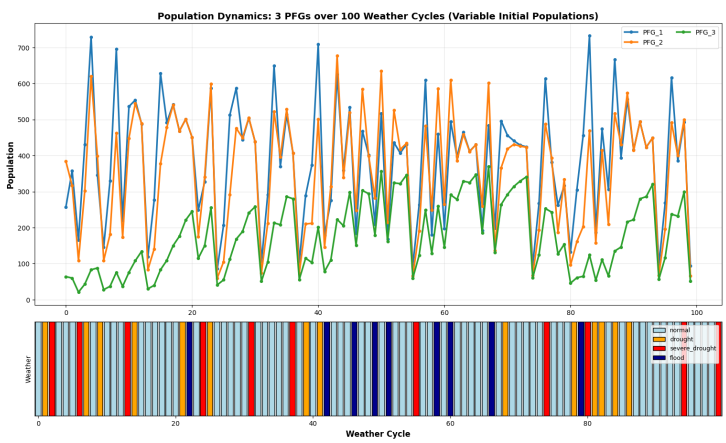

Figure 1: Plant Functional Groups (n=3)

What does this tell us? There were several insights to take away from these graphs. First, the periodicity. Like with animal population trends where this feature is clearly apparent, the populations consistently vary in the same proportions. This is indicative of the fact that resources are shared among the populations of PFGs, so the most competitive will naturally have a greater share. Some PFGs performed better in lower resource environments, probably due to better drought resistance traits.

The next feature to take away is the sustained diversity. Ecological resistance in our simulations was measured considering potential micro-evolutions over time. We see from our graphs that these micro-evolutions certainly had an impact. We can also roughly see the timeline it takes for these changes to

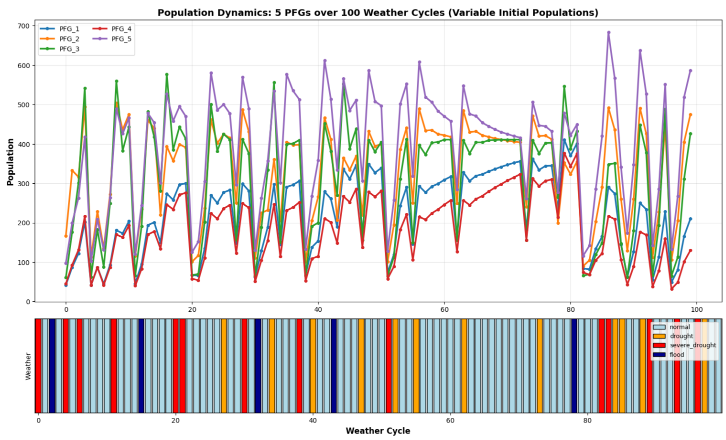

Figure 2: Plant Functional Groups (n=5)

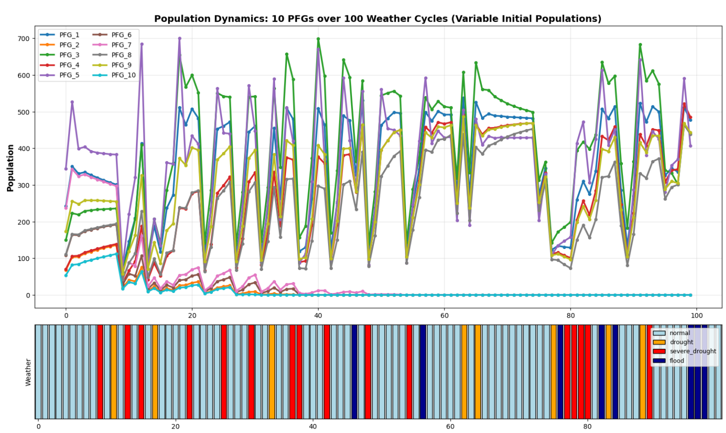

Figure 3: Plant Functional Groups (n=10)

occur, around some tens of weather cycles in our simulation.

Finally, we can also see that the number of PFGs that allows for the optimal quantity of sustained evolution. At 5 PFGs, none of the PFG populations drop off and the proportions allow for the ideal amount sharing of resources and diversity.

7 Model Optimization

First, let's note that we only considered three different numbers of PFGs in our baseline. Clearly, an insight from this limited set will not be as substantive as a more broad search.

To accomplish this, we must employ global optimization techniques on our objective function for the PFG count. We can do this through a gradient search algorithm or by finding the partial derivatives of our objective function and equating them to 0. We choose gradient search.

8 Strengths and Limitations

8.1 Baseline

Strengths

- Models many different archetypes of species.
- Accurately predicts which traits allow for optimal chances of survivability.
- Accounts for different weather cycles and ecological constraints like habitat loss.

Limitations

- Used a limited number of traits instead of the full range typical PFG research would involve.
- Periodicity is a worrying indicator of cycles that may not reflect real world phenomena.
- The evolution mechanic is limited in the sense that it cannot develop new traits and has a fixed, deterministic formula for what traits change.

9 Conclusion

9.1 Baseline

We present a first attempt at building a plant community under irregular weather patterns and varying ecological pressures. After an initial literature review, we chose to represent species using plant functional groups (PFGs) in order to generalize the model. This approach allowed us to focus on their behavior within the community without using specific species data.

Our model is guided through a logistic growth equation with simplified representations of competition, random weather, and micro-evolution. While a lot of these components are assumed and abstracted, the simulation produces several expected ecological behaviors. In specific, our results suggest that neither very low nor very high bioversity leads to the most stable outcomes, and that the middle ground may allow for greater long-term resilience.

At the same time, this work has its limitations. The restricted set of traits and fixed treatment of resources limits the realism of the model.

Overall, this model serves as our initial foundation, and it demonstrates how PFGs can be used as a method for modeling populations.

10 Works Cited

Isbell, Forest, et al. "Benefits of Biodiversity Exceed Those of Individual Species." Nature Communications, vol. 8, no. 14389, 2017, pp. 1-7, https://doi.org/10.1038/ncomms14389.

Lozano-Parra, Javier, et al. "A Novel Framework to Generate Plant Functional Groups for Ecological Modelling." Ecological Indicators, vol. 166, no. 112217, Sept. 2024, https://doi.org/10.1016/j.ecolind.2024.112217.

Nowak, David J. "Institutionalizing Urban Forestry as a Biotechnology to Improve Environmental Quality and Human Health in the Post-2015 World." USDA Forest Service, Jan. 2015, https://research.fs.usda.gov/treesearch/47499.

Wang, Jing, et al. "Effects of Plant Diversity on the Stability of Plant Communities in Arid and Semi-Arid Regions." Sustainability, vol. 18, no. 2, 2024, p. 641, https://doi.org/10.3390/su18020641.

CITE TEXTBOOKS (1 and 2)