Simulating Ascendency Statistics from the South Florida Ecosystems Dataset

EN 605.662 Data Visualization - Final Project

Dean Goldman

July 7th, 2022

Abstract

This project is an exploration of a network analysis dataset used for ecological monitoring. The report, titled "Network Analysis of the South Florida Everglades graminoid marshes and comparison with nearby cypress ecosystems" (Ulanowicz et al. 2002) [1] can be described as a study of the carbon cycle, or food chain, as it exists in several ecosystems of South Florida. The goal of this study was to produce dependency relationships between different entities within each ecosystem, and compare aggregated statistics between these ecosystems. These statistics are theoretical indicators of ecosystem resilience and efficiency, and are argued to be critical for land management decision making. The goal of this project will be to investigate several research questions listed below, and communicate the methodologies and analysis for answering them in a visually elegant way.

Introduction

This paper will attempt to address several research questions. Initially, some analysis and visualization will be performed to identify the major sources, sinks, and modes of carbon exchange within these networks. Secondly, an investigation will be made into the author's theory of ascendency and their characterization of each ecosystem within that framework. Given the ascendancy model, certain taxa hold higher importance for ecosystem resiliency, generally because they are widely resourced for food in the community. This paper will experiment with the way these network statistics may change given reasonable error bounds around the data produced. Are there general rules or trends around trophic exchange in these networks that can be discovered?

Background

The entities described in the Trophic Dynamics paper are aggregates of animal species, flora, or organic matter found in the South Florida ecosystem. Some entities correspond to fish, bobcats, or panthers, while others describe smaller scale components to an ecosystem, such as periphyton (water surface microalgae), organic carbon, and certain bacteria. The author's of the paper used a survey of previous studies to aggregate entity weight, population density, and diet information, to produce a directed graph, where each node represents an entity, and its edges represent biomass being transferred from that entity to another, or vice-versa. (one entity eats another, or another eats it). Below are two examples visualized.

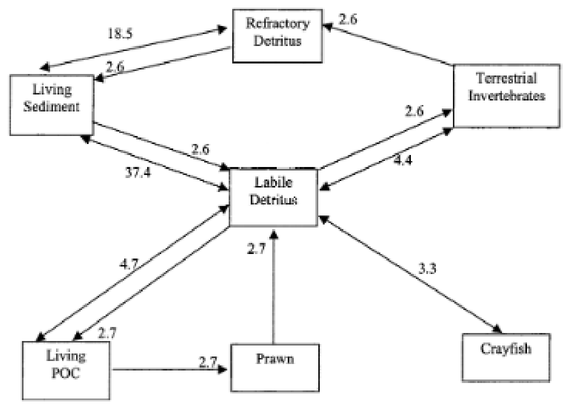


Fig. 1. A graph representation of carbon cycling within the cypress detritus (gCm^{-2} per year). [2]

The authors derive conclusions about the resiliency and efficiency of these different ecosystems using Ulanowicz (1986) theory of ascendancy [3]. The theory of ascendancy is an information theoretic approach to measuring carbon cycling in ecosystems. Ascendancy can be thought of as a spectrum that characterizes ecosystem activity. In low ascendancy systems, mutual information between nodes of carbon exchange is low. This means that for any unit of carbon flowing from one node (a biological taxa), there is uncertainty about where that unit may flow. In other words, multiple taxa may use this first taxa as a resource for food, so it is uncertain where any one unit of carbon may flow. In higher ascendancy systems, carbon exchange is more predictable. Meaning that species are more specialized in their carbon consumption. Low ascendancy systems are said to be more resilient to external forces, e.g. extreme weather, industrial development, because they have more options for carbon cycling, while high ascendancy systems are theorized to be more predictable and efficient. The total potential ascendancy is called the capacity, and the system energy that is spent in redundant cycles is indexed by a quantity called the overhead.

II.7.1 ASCENDANCY: A MEASURE OF ECOSYSTEM PERFORMANCE

In the absence of major perturbations, ecosystems exhibit a propensity towards configurations of ever-greater network ascendancy.

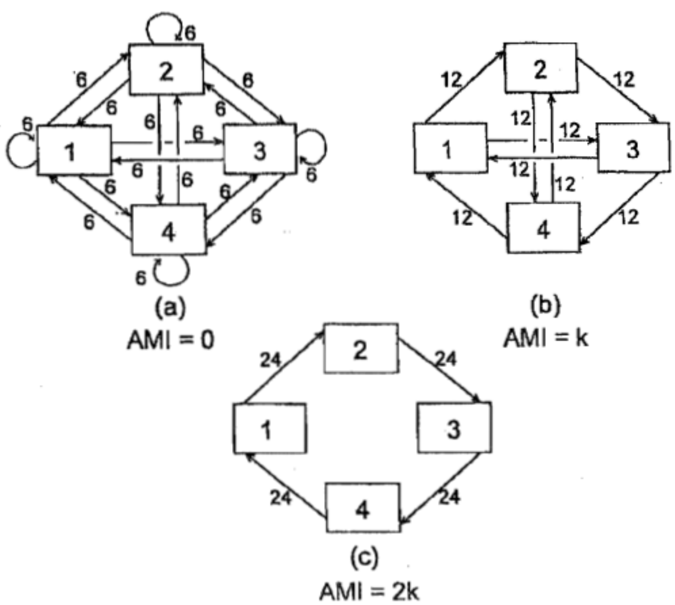


Figure 6: The increase in mutual information as flows become progressively constrained.

The data used in computing the indices of ascendancy were compiled by surveying other ecological studies and conducting fieldwork. Given that these data are static estimates, rather than high precision recordings through time, there is limited room for producing any kind of predictive model that may explain a cause and effect relationship between carbon exchange and network ascendancy. However, it does seem reasonable to produce some general error bounds around these estimates, and conduct experiments in measuring network ascendancy that way. In doing so, it may be possible to identify critical network components and subgraphs that can play a role in land management decision making.

Approach

In order to produce a distribution of experimental outcomes, a Monte Carlo random sampling method was used to generate various perturbations to the data, and compute the resulting ascendancies. The datasets themselves are networks of carbon exchange, which are commonly visualized by a multi-edged directed graph- whose nodes represent taxa and edges represent carbon flow between taxa, or an adjacency matrix, where each row represents carbon outflow for one taxa (what consumes it), and each column represents carbon inflow (what it consumes). Certain general patterns can be gleaned from these networks:

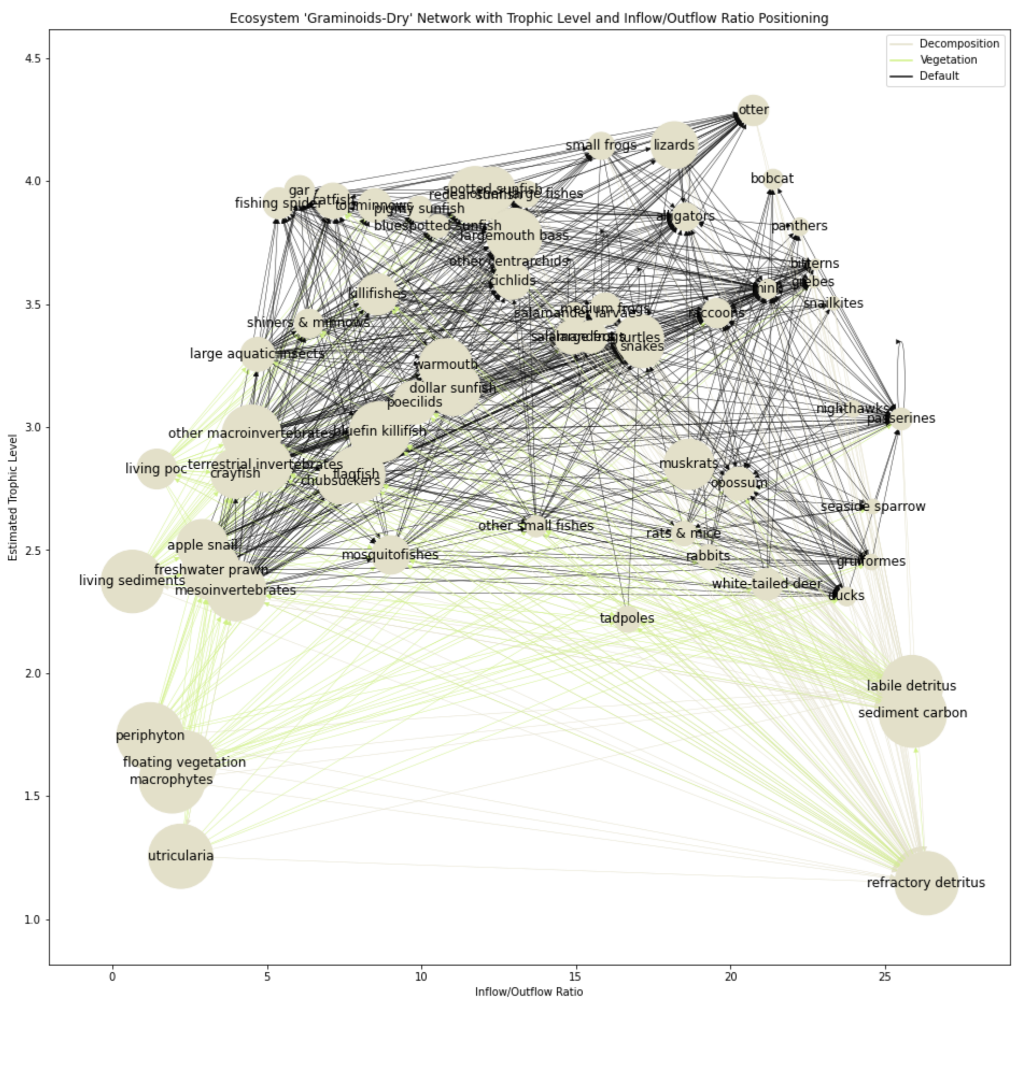


Fig. 3. Graph visualization of the "Cypress-Dry" ecosystem. Node size represents relative biomass, edge color indicates a rough classification of flow type.

Generally, it can be seen that lower trophic level taxa (vegetation, microalgae, periphyton, etc.) take up the majority of biomass in these ecosystems. These are the same nodes that are consumed by relatively the most other taxa as well. It may be determined that these taxa are central to the biology of the ecosystem. In network analysis, this is referred to as 'centrality', and is more a less a ranking of nodes based on their connectedness to other nodes. For this analysis, a ranking of importance will be computed by the PageRank algorithm, often used for ranking results in web search engines [4].

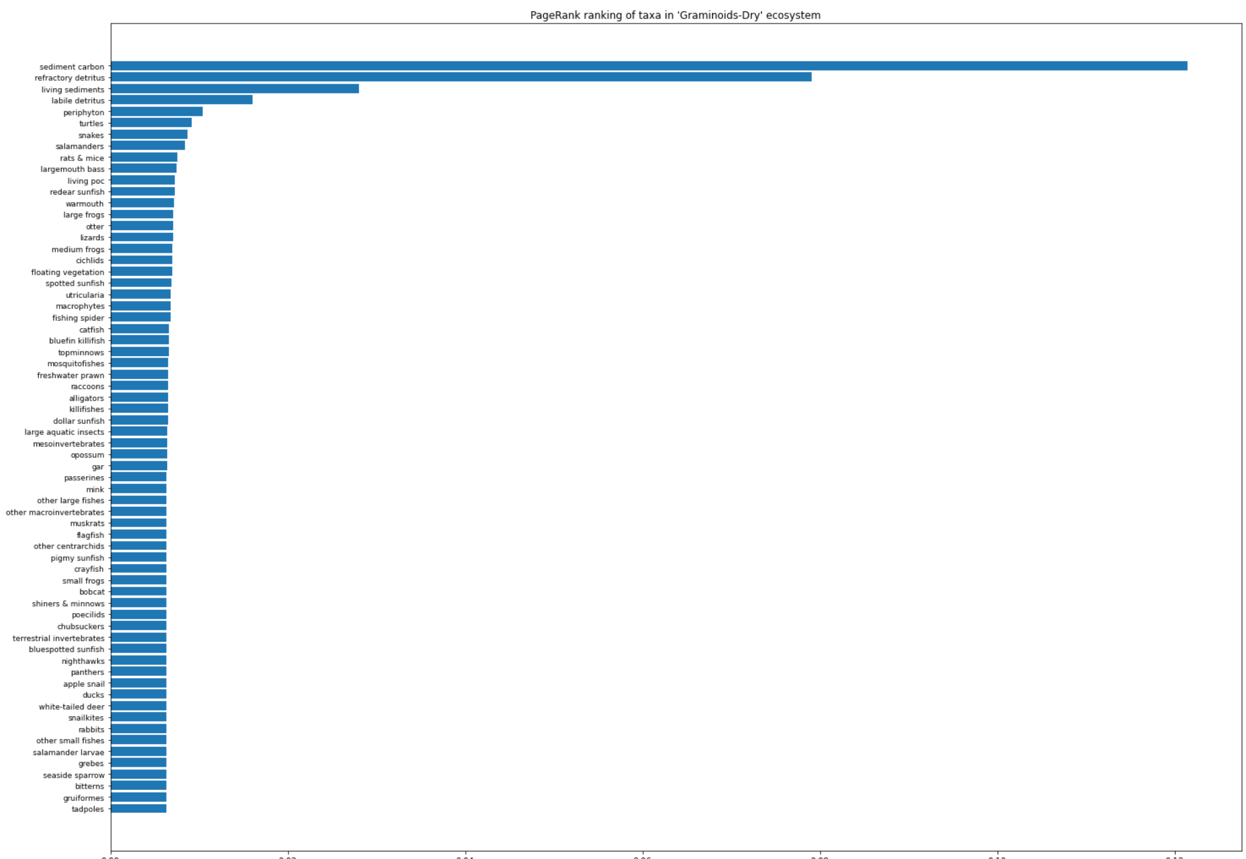


Fig. 4. PageRank ordering of taxa in the Graminoids-Dry ecosystem. Low trophic level organisms play a pivotal role in carbon flow exchange.

Once the important taxa are known, the experiments will focus on perturbing flows of certain taxa- either high ranking, low ranking, or both. These perturbations will then be propagated through the network, and the resulting flow matrix will be used to calculate the ascendancy, capacity, and overhead statistics. The perturbations will involve scaling flow edge weights by some factor drawn from a random uniform distribution, and randomly applying a "dropout" (i.e. scale by 0.0) to some edges.

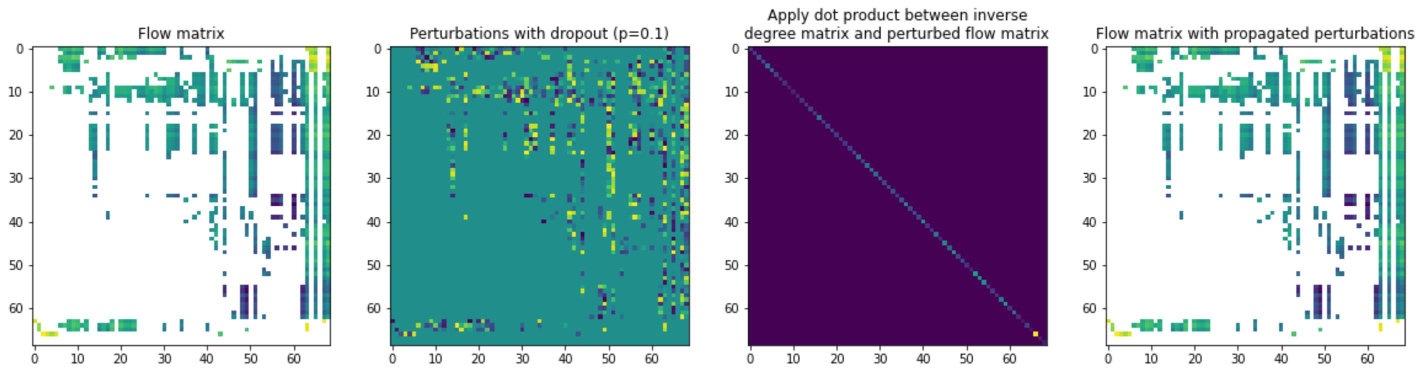


Fig. 5. Step-by-step for perturbing flow matrix. Perturbations are derived from a random uniform distribution, multiplied with the original flow matrix, and propagated throughout the network using the inverse degree matrix to produce a perturbed flow matrix.

The goal of these experiments are to discover potential gradients that may exist when certain flows are dampened or amplified, and thus produce some probabilistic model between distributions of carbon flow and theoretical ascendancy.

Results

The experimental results indicate that regardless of perturbation bounds, we can reliably predict a higher or lower ascendency/capacity ratio given that we've either perturbed the highest ranked nodes, or the lowest ranked nodes, respectively. Over all parameters, there exists an interval of roughly 10% where the ascendency/capacity ratio may fall. Interestingly, the scale and sign of perturbations is relatively unimportant, compared to the specific selections of which flow edges are perturb. Perturbing all edges with equal probability maintains an ascendency/capacity ratio very near the original. Over all trials for each parameter setting, there exists a confidence interval of roughly 5%.

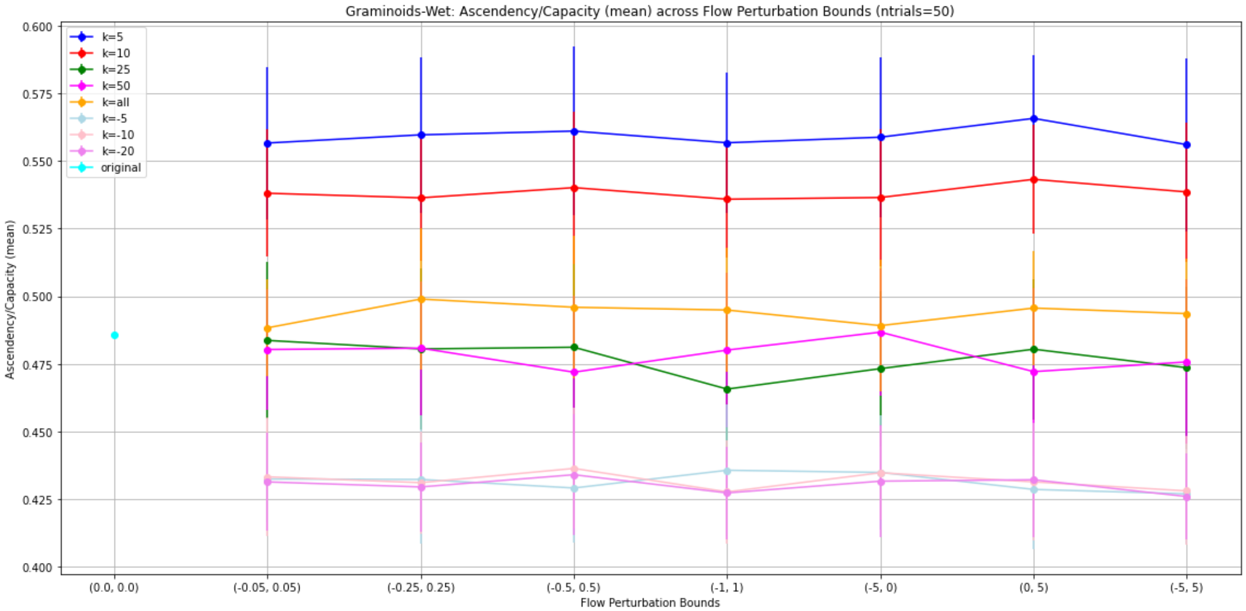


Fig. 6. Experimental results from perturbing ecosystem network flow matrices and computing ascendency/capacity ratios.

Conclusion

These experiments may indicate a possible causal structure in the organization of ecosystems: perturbations along flow channels through highly centralized nodes generally lead to higher ascendency, whereas perturbations along flows between low-centrality nodes may increase overhead. One explanation may be that an increase in energy flow from low-ranking nodes indicates more energy trapped in internal flows, rather than an increase in energy along high-ranking nodes, where that energy is dispersed efficiently across the network. While this may be plausible, it should also be noted that these experimental parameters were chosen with a more or less heuristic approach. It would be conducive to further research to produce more precise distributions of energy flow in these ecosystems. If a time series dataset could be compiled, more precise estimates and predictive modeling could be achieved.

References

[1] Ulanowicz, Robert & Heymans, Johanna & Bondavalli, C. (2002). Network analysis of the South Florida Everglades and graminoid marshes and comparison with nearby cypress ecosystems. Ecological Modeling 149 (2002) 5-23.

[2] Ulanowicz, Robert et al. (2002). Fig. 4. Lindeman Spine for graminoids and cypress habitats. Ecological Modeling 149 (2002) 5-23.

[3] Ulanowicz, R.E. (1986). Theoretical ecology: Ecosystem ascendency. pp. 481-483. In: *1987 Yearbook of Science and Technology* McGraw-Hill, NY. https://people.clas.ufl.edu/ulan/files/McGrwHll.pdf

[4] Brin, Sergey and Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. Computer Science Department, Stanford University. Stanford, CA.

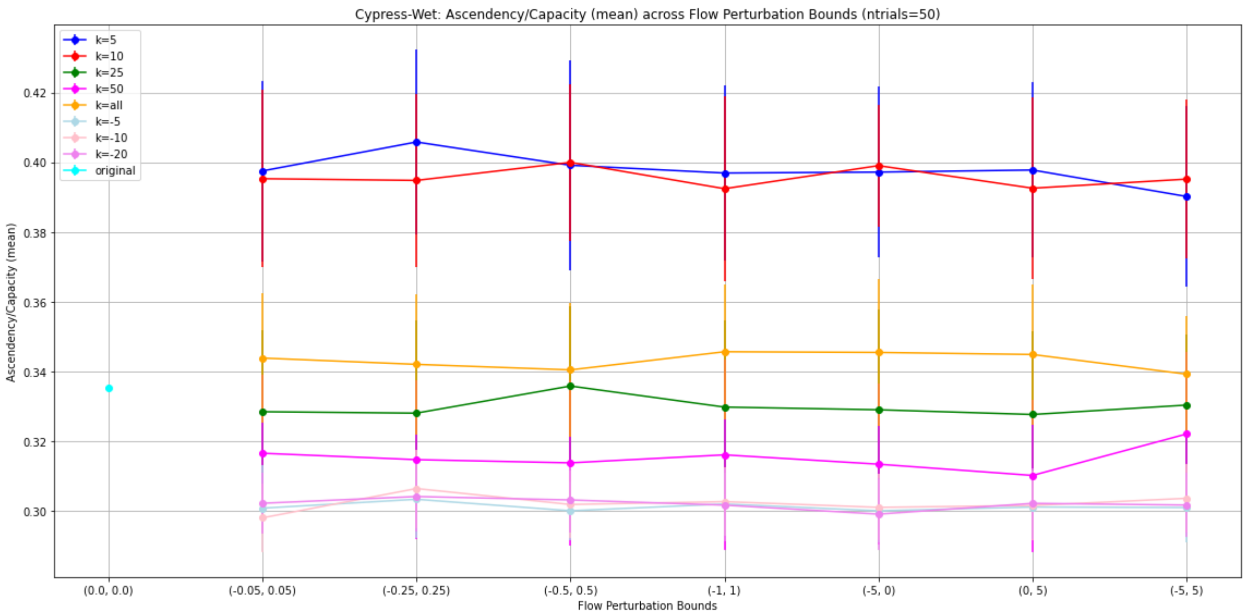
[5] Fath, Brian & Scharler, Ursula & Ulanowicz, Robert & Hannon, Bruce. Ecological network analysis: network construction. Ecological Modeling 208 (2007) 49-55.

[6] Koller, Daphne & Friedman, Nir. 2.2 Graphs - Probabilistic Graphical Models: Principles and Techniques (2009) Massachusetts Institute of Technology. (p.71-75).

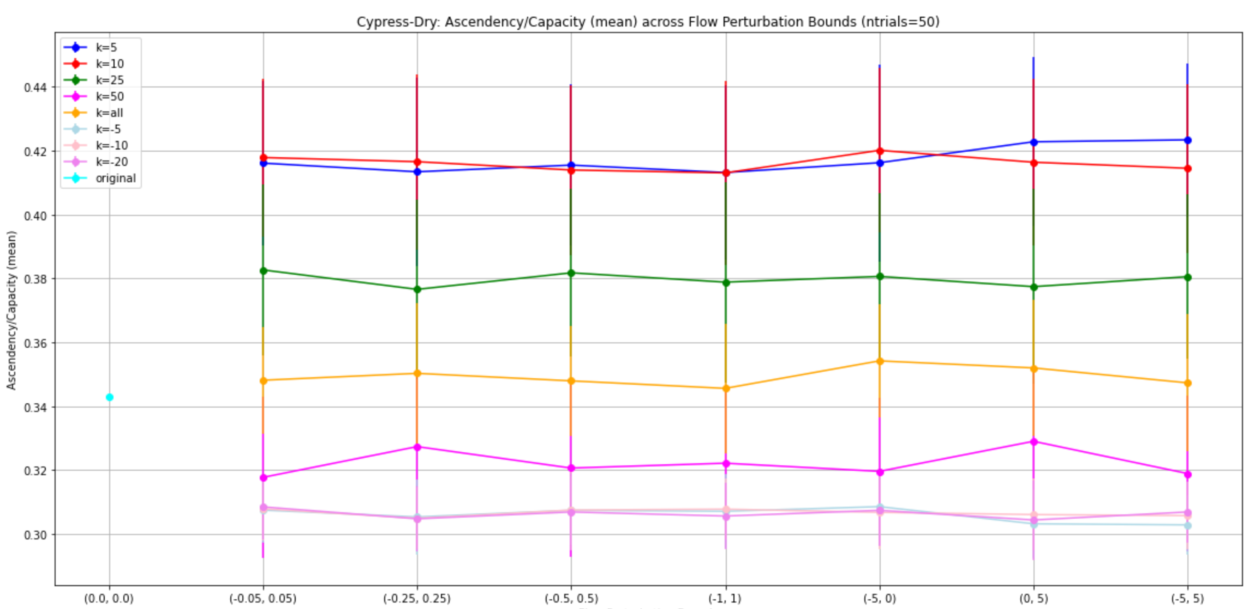
[7] Aric A. Hagberg, Daniel A. Schult and Pieter J. Swart, "Exploring network structure, dynamics, and function using NetworkX", in Proceedings of the 7th Python in Science Conference (SciPy2008), Gäel Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, Aug 2008

[8] Borrett & Lau (2014) enaR: An r package for Ecosystem Network Analysis. Methods in Ecology and Evolution 5: 1206-1213.

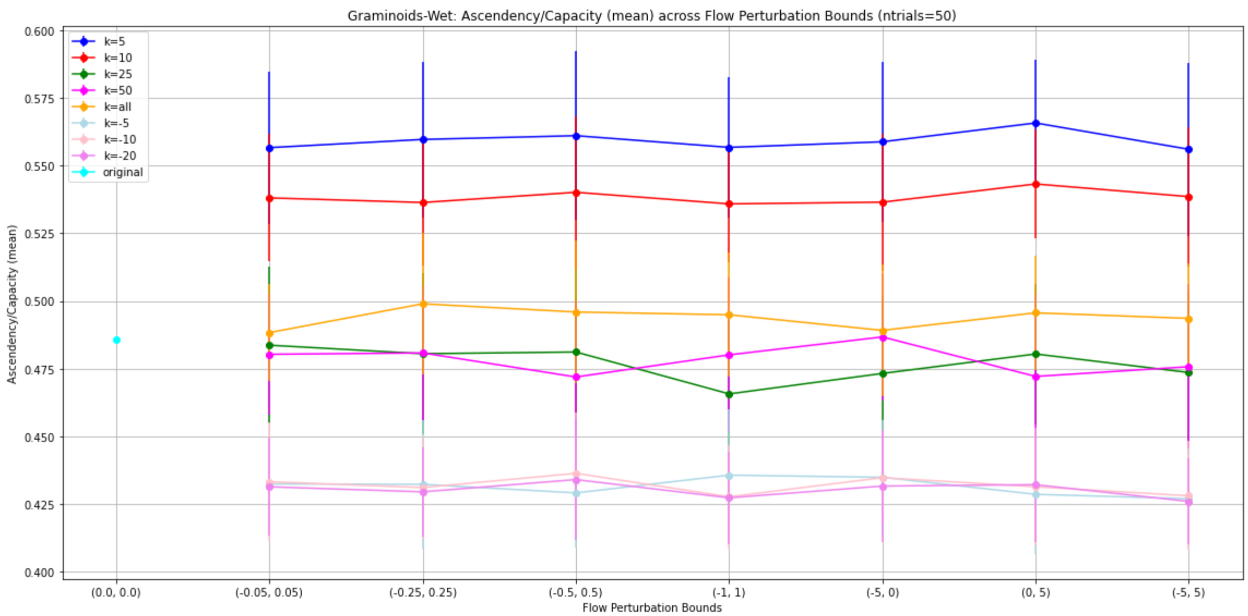
Appendix

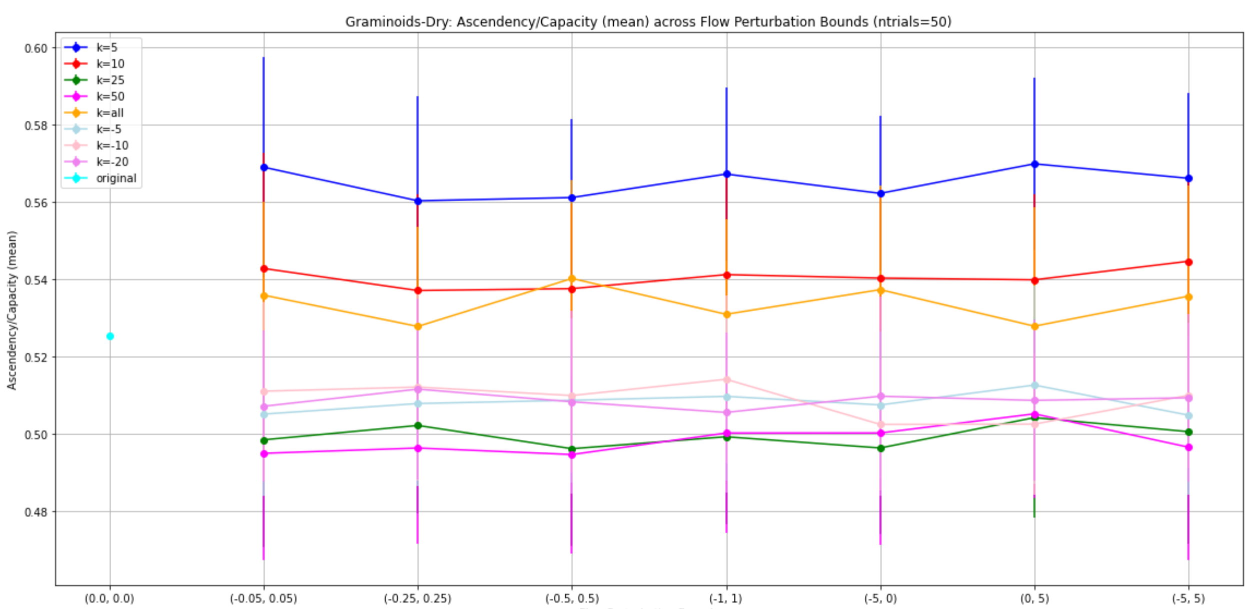


Cypress-Wet: Ascendency/Capacity (mean) across Flow Perturbation Bounds (ntrials=50)
k=5
k=10
k=25
k=50
k=all
k=-5
k=-10
k=-20
original
Ascendency/Capacity (mean)
0.42
0.40
0.38
0.36
0.34
0.32
0.30
(0.0, 0.0)
(-0.05, 0.05)
(-0.25, 0.25)
(-0.5, 0.5)
(-1, 1)
(-5, 0)
(0, 5)
(-5, 5)
Flow Perturbation Bounds



Cypress-Dry: Ascendency/Capacity (mean) across Flow Perturbation Bounds (ntrials=50)
k=5
k=10
k=25
k=50
k=all
k=-5
k=-10
k=-20
original
Ascendency/Capacity (mean)
0.44
0.42
0.40
0.38
0.36
0.34
0.32
0.30
(0.0, 0.0)
(-0.05, 0.05)
(-0.25, 0.25)
(-0.5, 0.5)
(-1, 1)
(-5, 0)
(0, 5)
(-5, 5)
Flow Perturbation Bounds


Graminoids-Wet: Ascendency/Capacity (mean) across Flow Perturbation Bounds (ntrials=50)
k=5
k=10
k=25
k=50
k=all
k=-5
k=-10
k=-20
original
Ascendency/Capacity (mean)
Flow Perturbation Bounds
(0.0, 0.0)
(-0.05, 0.05)
(-0.25, 0.25)
(-0.5, 0.5)
(-1, 1)
(-5, 0)
(0, 5)
(-5, 5)


Graminoids-Dry: Ascendency/Capacity (mean) across Flow Perturbation Bounds (ntrials=50)
k=5
k=10
k=25
k=50
k=all
k=-5
k=-10
k=-20
original
Ascendency/Capacity (mean)
Flow Perturbation Bounds
(0.0, 0.0)
(-0.05, 0.05)
(-0.25, 0.25)
(-0.5, 0.5)
(-1, 1)
(-5, 0)
(0, 5)
(-5, 5)