Complex Networks

Project: Simulation of a food web under species extinction

Anne Schreiber Brito
Benet Manzanares Salor
Ruben Vera Garcia

Contents

Part 1

Introduction

The objective of the project is to analyze how a food web behaves when one specie is taken out. This is considered as an equivalent of studying the potential consequence of the extinction of a specie. Both simulations and network analysis are leveraged.

In order to conduct the study, a dynamic network will be build, where each node represents a specie and each edge represents the interaction between a pair of creatures. These interactions include prey-predator relationships and competition for food. The Gulf of Cadiz data-set [1] from [2] was employed. For the simulating the ecosystem, we get inspiration from the work *Implicaciones de la similitud estructural en la dinámica de redes ecológicas* by Julia Martinez-Atienza, where she studies the implications of the structural similitude on dynamic ecological networks. In particular, we use the Holling Type I Functional Response [3] model on the same basis defined there.

In the following sections our process description, experiments and results will be described. The final section defines our main learning outcomes, limitations and future work.

Part 2

Process Description

The process of this project has been divided into the finding and selection of the dataset, choosing and understanding the modeling technique and implementation of the simulation and experiments framework.

2.1   Data-set

The data-set we used was taken from a web-page named *Web of Life* that provides a database of ecological networks [2]. The selected data-set [1] contains 44 different species that live at the Gulf of Cadiz. Selection was based on network size (big enough for providing interesting results but not elevating significantly computational cost) and species (wide variety of predators and preys, many of them known). This data-set is given as a predation matrix P where columns represent the relative frequency of prey species in the diet of predator species. In other words, each cell P_{ij} represents the proportion of individuals hunted by specie j that belong to specie i. We decided to not contemplate cannibalism (which was found for 14 species) setting all diagonal of P to zero. Subsequently, we have to adjust the columns for ensuring that all sum 1 as expected (exception is set for those species which are not predator of any other).

In contrast to simpler food webs, predators and preys in this data-set are not disjoint sets. Each specie can be predator of some species and prey of others, relationships can be as predator, prey, or both (bidirectional connections), predators of a specie can have common preys with it (skip-connections in the trophic chain) and even some predators of a specie can be preys of this specie' preys (cycles).

2.2   Modeling

For modeling the behavior of the selected food web, we employ the *Holling Type I Functional Response* [3]. This approach allows modeling mutualistic and competitive relationships between species simultaneously, being suitable for the defined dataset. Additionally, cannibalism is not contemplated. The method defines how the population of a species changes over time using the following formula:

$$\frac{dx_i}{dt} = x_i \left(\alpha_i - \sum_{j}^{S} I_{ij} x_j \right) \tag{1}$$

Where x_i is the amount of population for the specie i, t is time, α_i is the self-growing factor (how the population increases by itself) of specie i, S is the set of species and I is the matrix of interactions.

Each cell I_{ij} represent both the mutualistic and competitive relationship between specie i and specie j. The mutualistic component refers to the predator-prey relationships. Considering the nature of our data-set, we decided to define mutualism with direct proportion to the relationship as prey and indirect to the relationship as predator with the other specie. This represents the fact that having many predators has a negative effect on survival (increasing mutualism) but a larger prey population facilitates survival (decreasing mutualism). To this end, we define the mutualism matrix M employing the predation matrix P through the next formula:

$$M_{ij} = P_{ij} - P_{ji} \tag{2}$$

The competitive component represents how much specie i have to compete with specie j for obtaining food. Subsequently, it is proportional to the ratio between the number of species eaten from both (i and j) and the amount of species eaten by i. Under this premise, we construct each competition component C_{ij} using the matrix of predation P in the following way:

$$C_{ij} = \frac{\sum_{k}^{S} P_{jk} : P_{ik} > 0}{\sum_{l}^{S} P_{il}} \tag{3}$$

For obtaining the I matrix, we perform a weighted sum of the two matrices, allowing to adjust the importance of both factors employing the weights w_m and w_c. Nonetheless, an exception is settled for the diagonal, where values are substituted by the self-limitation d factor. This factor determines how a specie population growing is limited when not mutualism nor competition is found. It can be also interpreted as the ratio of natural deaths of the specie, being the opposite of the self-growing factor. On this basis, the matrix I is defined

by the next formula:

$$I_{ij} = \begin{cases} d, & \text{if } i = j \\ w_m * M_{ij} + w_c * C_{ij}, & \text{otherwise} \end{cases} \tag{4}$$

In some food webs where the sets of predators and preys is completely disjoint each I_{ij} only has a mutualism or competitive component. However, in our dataset species can be preys/predators and competitors of other simultaneously, having both mutualistic and competitive components in sum.

This model requires the manual definition of the following parameters: initial populations x, self-growing factors α, mutualism weight w_m, competition weight w_c and self-limitation factor d. The definition of these parameters is detailed in Section 2.3.2.

2.3   Implementation

We implemented this project using Python 3 and the NetworkX library. Our code is fully contained in the attached `main.py` file, being possible to configure its behavior modifying the constants of the Settings region. In the following subsections, we define the main tasks of the development.

2.3.1   Interactions matrix

The interactions' matrix I is a fundamental component for our modeling (see Section 2.2). For its construction the mutualism matrix M, competition matrix C, mutualism weight w_m, competition weight w_c and self-limitation factor d are required. Matrices M and C are obtained using the `get_mutualism_and_competition_matrices` method. It first receives the predation matrix P read from the dataset file and adjust it by removing cannibalism and making predation frequencies sum 1. The latter procedure is applied for all species except those that are not predators of any other, to avoid dividing by 0. Afterwards, the equations 2 and 3 are applied for obtaining both matrices. Similarly to the previous case, an exception is defined for Eq.3, avoiding divisions by zero by setting to 0 the competition value for all species that are not predator of any other. Finally, in the `create_iteractions_matrix` method, Eq.4 is applied for obtaining the interactions' matrix I.

2.3.2   Parameters selection

As defined at the end of Section 2.2, the Holling Type I Functional Response modeling technique requires the definition of the initial populations x, self-growing factors α, mutualism weight w_m, competition weight w_c and self-limitation factor d. This multi-parameter selection has critical consequences in the results, so a careful process needs to be carried out for obtaining the most realistic combination.

Our first criteria for this selection was establishing a balanced initial state. In this way, we obtain an ecosystem in equilibrium to which we can apply perturbations (such as removing all the population of a specific specie) and observe how it reacts. In terms of the modeling technique, this is achieved by having $\frac{dx_i}{dt} = 0$ in Eq.1 for all the populations x_i. Since there

are multiple combinations that provide this result, we assume that all the parameters are already defined (with processes detailed later) except the self-growing factors α and adjust these for ensuring the equilibrium. Consequently, α is defined by the following equation:

$$\alpha_i = \sum_{j}^{S} I_{ij} x_j \tag{5}$$

This process is performed in the `balanced_ini_state` method.

For the populations x, we followed the observation that the haunted species usually have more individuals than the predator ones. Under this premise, the population count of a specie was defined directly proportional to the amount of predators. Concretely, for each specie, we establish a minimum population min_x and incremented it proportionally to a predefined value $incr_x$ and predations defined by P:

$$x_i = min_x + incr_x * \sum_{j}^{S} P_{ij} \tag{6}$$

In this way, only have the parameters min_x and $incr_x$ need to be established, rather than the x_i for each specie. We consider that setting a value for $incr_x$ an order of magnitude greater than min_x is a suitable approach. The method `ini_populations` performs this initialization.

Under this framework, min_x and $incr_x$ were arbitrarily defined (min_x=100 and $incr_x$=1000), remaining to establish the values for d, w_m and w_c. Initial experiments proved that results change abruptly for small differences of these parameters. In these parameter selection experiments, a test extinguishing each specie was performed and consequences usually jumped from the extinction of all the other species to the survival of all of them (independently of the initially extinct specie). On this basis, we created the `parameters_search` method that performs a grid search for all the combinations of the parameters d, w_m and w_c in the range [1e-1, 1e-9] jumping by order of magnitude ($9^3 = 729$ combinations). The search criteria relied on the diversity outcome of survivors, having the maximum different amounts of survival species for the different extinct species. This resulted in the parameters d=1e-9, w_m=1e-4 and w_c=1e-9, which obtained 22 different number of survivors for the 44 species.

2.3.3   Network

In the `create_network` method, a network that for representing the food web is defined using the interactions' matrix I as adjacency matrix. This is preferred to the mutualism matrix M since it doesn't clearly provide information about competition factors. We obtain a weighted and directed network where nodes are species and edges represent the weighted sum of mutualism and competitive relationships between the species or self-limitation (as a self-connection). Species populations x and self-growing factors α are established as nodes attributes, so the graph contains in some way all the parameters of the simulation. This allows to apply network-based analysis independently of the modeling simulation.

2.3.4   Simulation

The `simulation` method has been implemented for our experiments. It receives the parameters of the initially stable ecosystem, perturbs it and computes the consequences of applying Eq.1 for a predefined amount of iterations or convergence. Perturbation is performed as the multiplication of a specie population by a predefined *extinction factor* in the range $[0, 1)$. Instead of the NetworkX graph, the NumPy library has leveraged for computing Eq.1 update optimally, defining NumPy arrays for populations x, self-growing factors α and interactions matrix I. The population counts at each iteration are stored, allowing for posterior analysis.

Complementary, we have created `test` method, which calls the `simulation` function for collecting the results of perturbing the populations of each specie. It receives P, M, C, d, w_m and w_c as parameters, creating I, x and α for the simulations. In this way, it can be used for both the `parameters_search` method and the experiments. After each simulation, if the *verbose* setting is True, the `print_simulation_results` and `plot_populations_history` methods are used for providing useful information of the results.

Part 3

Experiments

Under the framework previously defined, we have evaluated the effects of the extinction of each specie on the Gulf of Cadiz ecosystem. The selection of parameters from Section 2.3.2 was employed, *extinction factor* was set to zero and number of iterations per simulation is limited to 100, since no important differences has been observed for greater values.

In the following sections, the obtained network, amount of extinctions consequent from initial extinctions and populations evolution throughout the simulations are analyzed.

3.1   Obtained network

Using the methodology defined in Section 2.3.3 (using I as adjacency matrix), a network that represents the food web for the simulation has been obtained. Each node of this network is is labeled with the specie id number (from 0 to 43). The following list depicts the number corresponding each specie:

- 0: Adult hake
- 1: Anchvoy
- 2: Anglerfishes
- 3: Benthic cephalopods
- 4: Benthic invertebrates C

- 5: Benthic invertebrates FDS
- 6: Benthopelagic cephalopods
- 7: Bivalves
- 8: Blue whiting
- 9: Commercial sparids 1

- 10: Commercial sparids 2
- 11: Common octopus
- 12: Crabs
- 13: Deep-sea fishes
- 14: Demersal fishes
- 15: Demersal piscivores
- 16: Detritus
- 17: Discard
- 18: Dolphins
- 19: Flatfishes
- 20: Gelatinous zooplankton
- 21: Horse mackerels
- 22: Import
- 23: Juvenile hake
- 24: Killer whales
- 25: Large benthopelagic fishes
- 26: Loggerhead turtles

- 27: Mackerels
- 28: Macrozooplankton
- 29: Mantis shrimp
- 30: Meso and microzooplankton
- 31: Mullets
- 32: Norway lobster
- 33: Phytoplankton
- 34: Polychaetes
- 35: Rose shrimp
- 36: Sardine
- 37: Seabirds
- 38: Sharks
- 39: Shrimps
- 40: Skates
- 41: Small benthopelagic fishes
- 42: Small demersal fishes
- 43: Suprabenthos

From this list is necessary to explain some species which its name is not sufficiently explanatory.

- Discard: The authors define discard as the bait used by fisherman of the Golf of Cadiz

- Imports are migratory species which are not from the Golf of Cadiz

- Detritus are the dead particulate organic and mineral material, which it is mainly obtained from dissolved organic material

Figure 1 provides a visual representation of the network, showing that there is a strong connectivity. There are only few (such as number 17=Discard, top-right) that have few relationship connections to other species, being usual to have a lot of links to other species. This can be also observed in the network stats shown below the figure, with an average degree of 21.7 that is close to half of the network size (44). In other words, this graph is very different from the well-known scale-free networks, not enabling to observe any pattern even with the reduced amount of nodes. Despite trying, nothing better than the ring representation could be obtained.
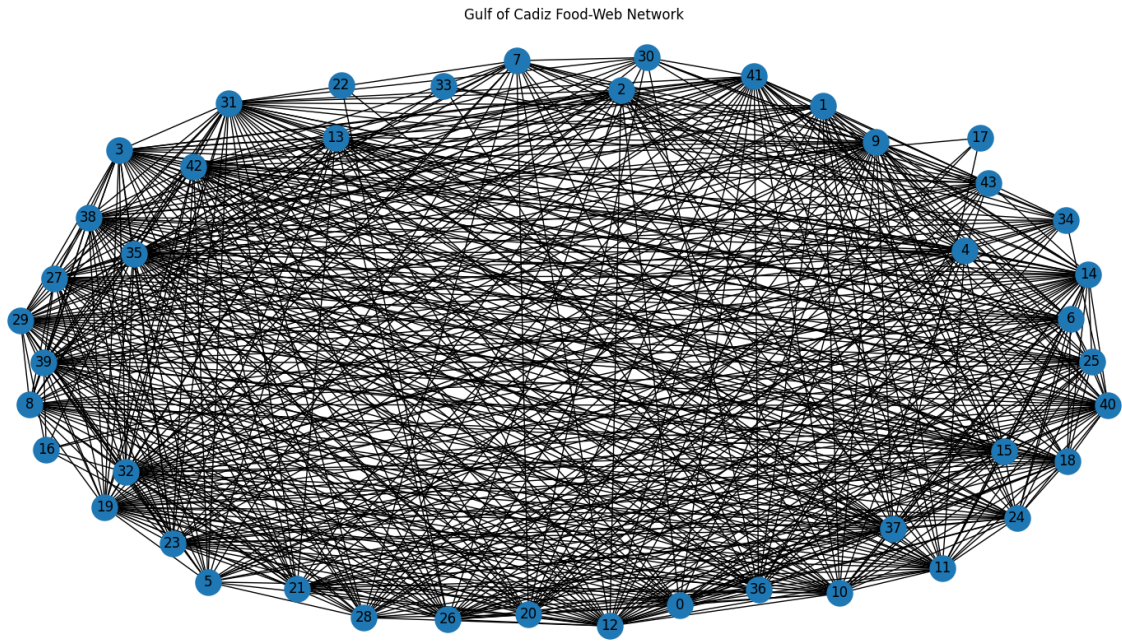
Figure 1: Gulf of Cadiz Food-Web Network

- Minimum degree: 2 by node 22 (Import)

- Maximum degree: 35 by node 32 (Norway lobster)

- Average degree: 21.7

Figure 2 depicts the network obtained using the predations matrix P as adjacency matrix. That would be equivalent to the network provided from the dataset. It can be noted that, compared with the previous network, it has a more much clear structure and a more coherent average degree. This shows that the competitive relationships (the only difference between this network and the previous one) are the responsible for the heavy connectivity observed, significantly chaning network structure.
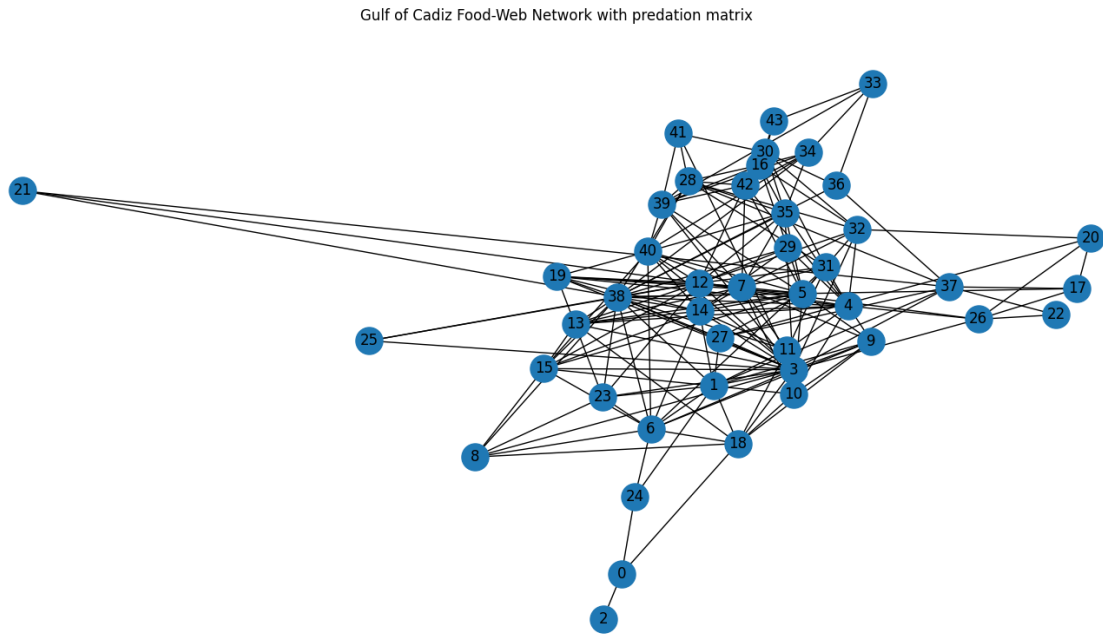
Figure 2: Gulf of Cadiz Food-Web Predator Network

- Minimum degree: 1 by node 2 (Anglerfishes)

- Maximum degree: 23 by node 3 (Benthic cephalopods)

- Average degree: 8.45

3.2   Extinctions dominoes

As a main focus of our project, we have collected the extinctions dominoes. That is, the species that get extinct as a consequence of the initial extinction. In particular, the amount of species that survive are analyzed, since they are considered an importance indicator of the initially extinct specie in the ecosystem (the lower the number, the more important). Table 1 show the amount of species that survive for each initially extinct specie during our simulations.

Initially extinct specie	#Survived	Initially extinct specie	#Survived
Adult hake	40	Import	14
Anchvoy	38	Juvenile hake	23
Anglerfishes	24	Killer whales	18
Benthic cephalopods	17	Large benthopelagic fishes	29
Benthic invertebrates C	38	Loggerhead turtles	34
Benthic invertebrates FDS	37	Mackerels	22
Benthopelagic cephalopods	40	Macrozooplankton	40
Bivalves	40	Mantis shrimp	40
Blue whiting	23	Meso and microzooplankton	41
Commercial sparids 1	40	Mullets	40
Commercial sparids 2	38	Norway lobster	40
Common octopus	14	Phytoplankton	41
Crabs	40	Polychaetes	16
Deep-sea fishes	33	Rose shrimp	40
Demersal fishes	4	Sardine	38
Demersal piscivores	40	Seabirds	30
Detritus	41	Sharks	40
Discard	38	Shrimps	32
Dolphins	40	Skates	9
Flatfishes	28	Small benthopelagic fishes	21
Gelatinous zooplankton	10	Small demersal fishes	38
Horse mackerels	39	Suprabenthos	9

Table 1: Number of survivors after species extinction

As can be observed, Demersal fishes, Skates, Suprabenthos and Gelatinous zooplankton seem the most important species in the ecosystem, since the fewer species survived after their extinction. We don't have a way of validating if these results are correct, since no ground truth has been found. Although, it seems intuitive that Gelatinous zooplankton, which is a first link of the trophic chain, has notorious effects in the ecosystem. However, not reasoning was found to why Demersal fishes, Skates and Suprabenthos are even more critical than Gelatinous zooplankton for the Gulf of Cadiz ecosystem.

Moreover, we tried to find if there was correlation between the number of species survived when a specie is initially extinct with the node degree, PageRank, or betweenness value. The metric used to this end was the Pearson correlation, which returns a value near -1 if there is a negative correlation and +1 for a positive correlation. Nevertheless, if the value is near 0, there is no correlation. The results achieved were that node degree with species survived have no correlation (-0.00846794), PageRank with species survived have no correlation (0.24698841) and Betweenness with species survived have no correlation(0.09183374).

3.3　Populations evolutions

In the following, a selection of plots of population counts per iteration during different simulations are depicted. For most of the cases, both plots that including all the species populations and that with only the survival ones are shown. In this way, the detection of differences between extinct populations and survival populations is facilitated.

Figures 3 and 4 show the consequences of the extinction of Crabs. It can be noted that most of the extinctions result after huge peaks in populations, probably due to lack of food for everyone and/or the complete predation of the overpopulated specie of the prey one. In these plots, two clusters of peaks can be observed, probably the latter as a consequence of the extinctions resulting from the former. Other species seem unaffected by the extinction (Import and Discard), maintaining the same population. This may involve the existence of sub-ecosystems (or communities) which can manage to survive if no member is affected.
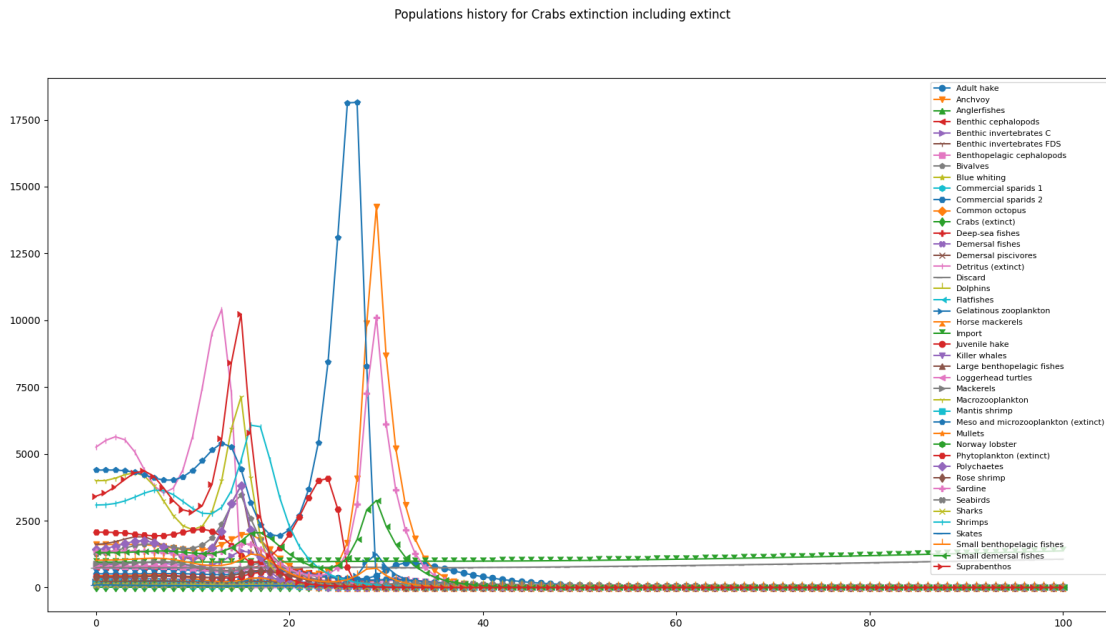

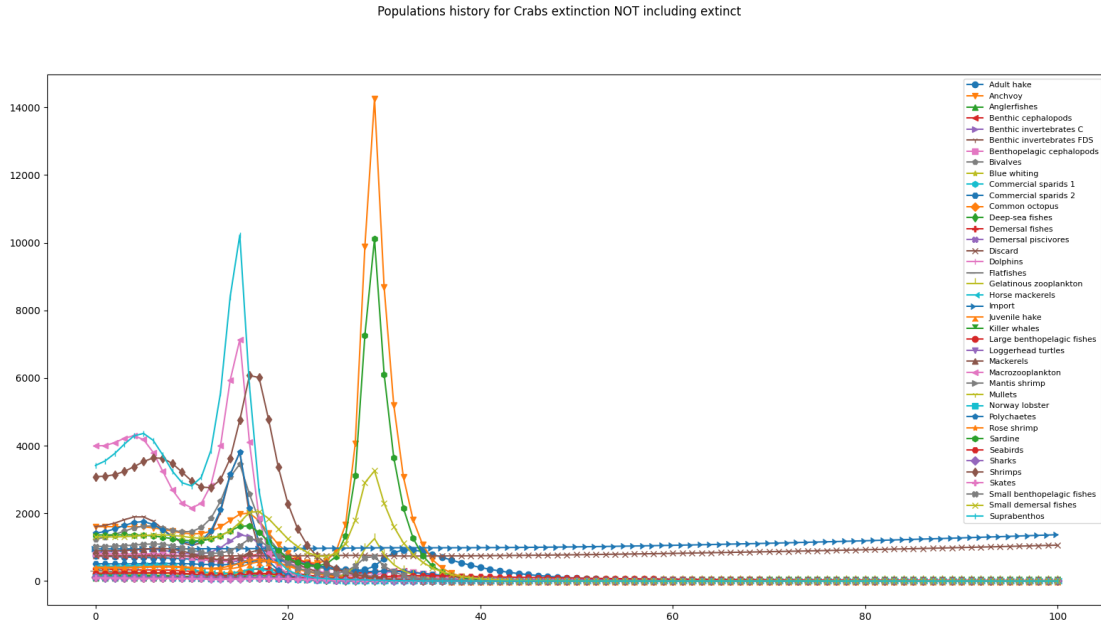Figure 3: Population change after the Crabs first went extinct

Figure 4: Survivor Species after the Crabs first went extinct

In figure 5 we can see that after 75 iterations, a condemning explosion of Dolphins, Commercial sparids 2 and Commercial sparids 1 population is observed. Values of the order of 1e8 and 1e9 are reached, being potentially unrealistic and a sign of the limitations of the model (which does not consider facts such as space and biological restrictions). Parallelly, in figure 6, we can see how Loggerhead turtles populations also increase to 6e6 and afterwards decreases in a logarithm manner to more normal values. This shows that maybe not all explosive growing leads to extinction, but a longer simulation would be required to ensure it.
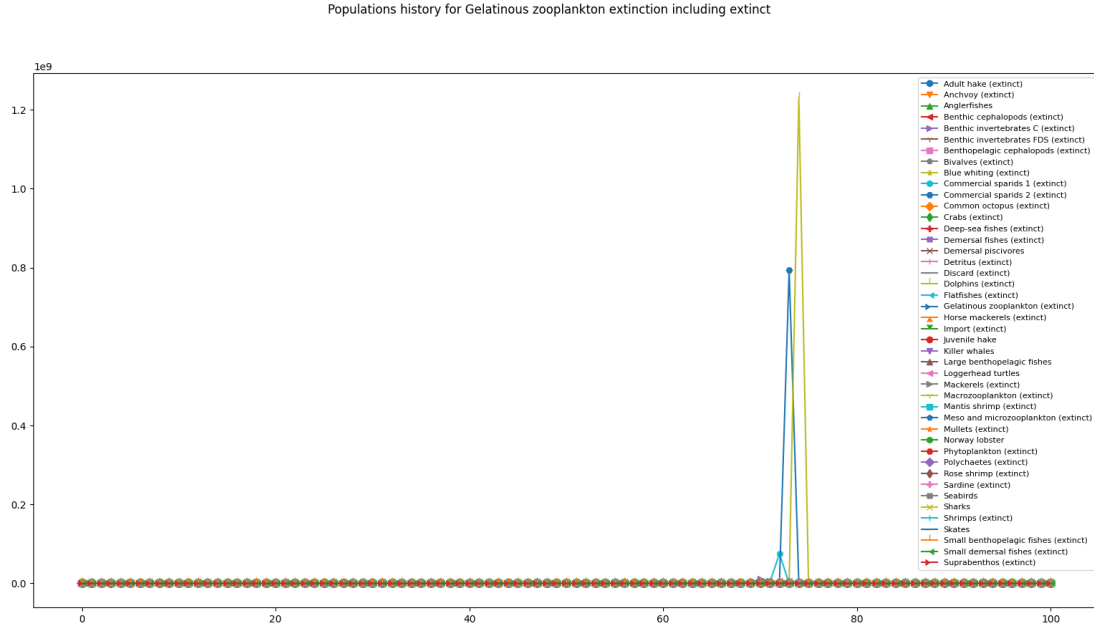
Populations history for Gelatinous zooplankton extinction including extinct

Figure 5: Population change after the Gelatinous zooplankton first went extinct

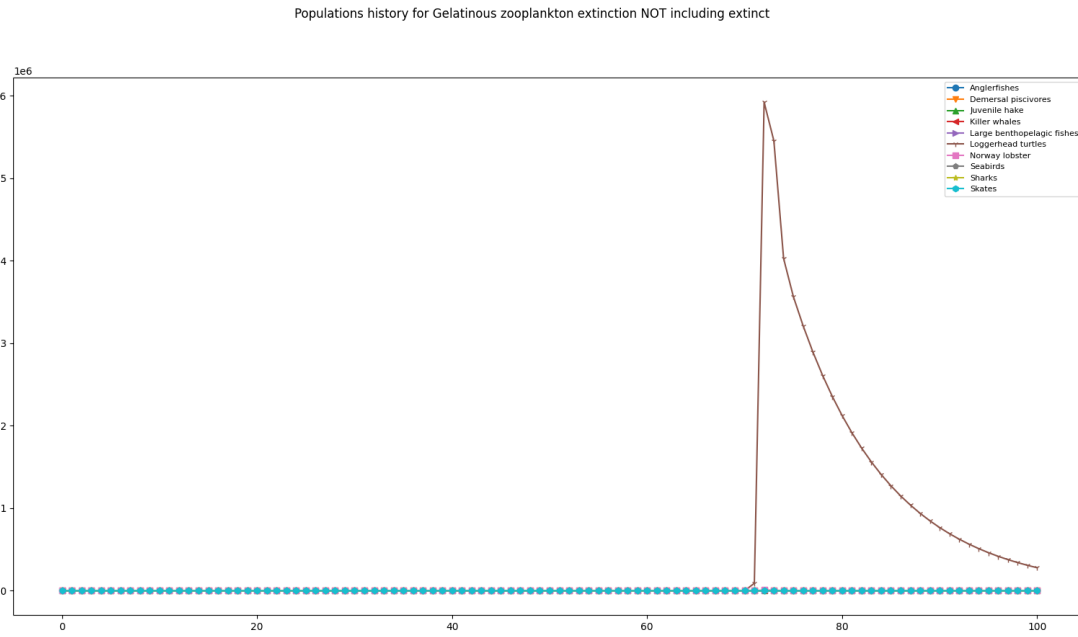Populations history for Gelatinous zooplankton extinction NOT including extinct

Figure 6: Survivor Species after the Gelatinous zooplankton first went extinct

We also plotted for 1000 iterations, and we could see that after the surviving species get stable, some species population rate start to show a periodic waves-like pattern. This can

be seen on figure 7 where the Adult Hake goes extinct. Here, after the species have greater fluctuations in its population between iteration 1 and 100, the species begin to stabilize into constant and lower population rates. If we zoom in figure 7 between iteration 475 and 650, we can see another constant pattern of this wave behavior. In figure 8 we see the zoomed image, where we see that the species Discard, Import, and Benthic invertebrates FDS present the wave pattern. Here we can observe that when the Discard and Import starts to decrease, is when the Benthic invertebrates FDS starts to rise, meaning that when this to spices populations are low, is when the Benthic invertebrates FDS is the highest. This pattern is shown 3 times, the first one between iteration 200 and 400, the next one between 475 and 650 and the last one between iterations 800 and 950.
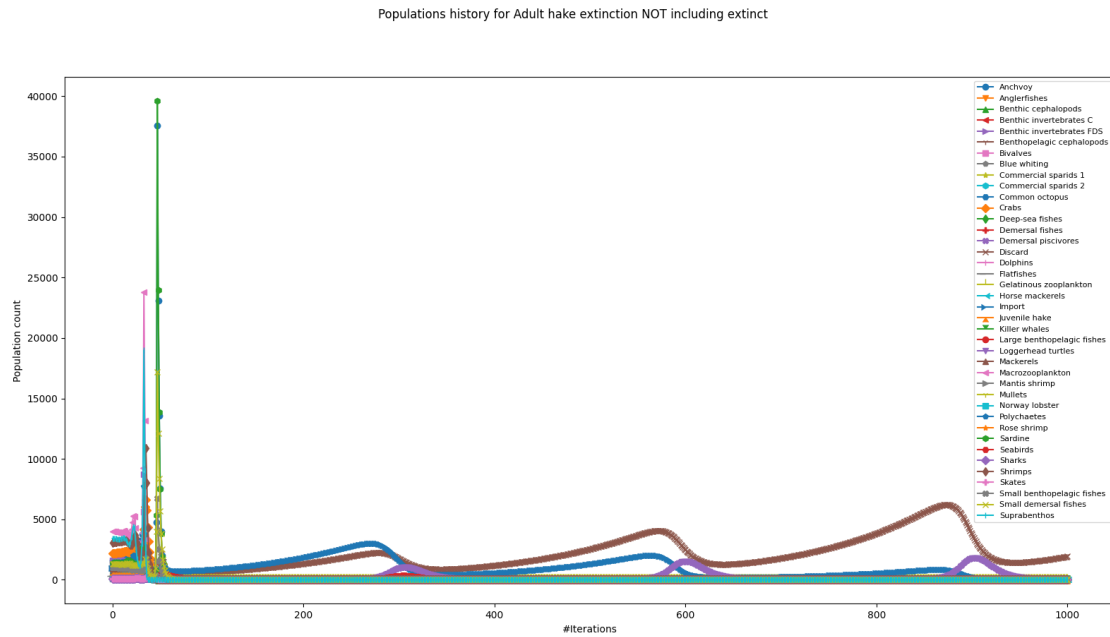

Figure 7: Species survivors after 1000 iterations when Adult Hake goes extinct
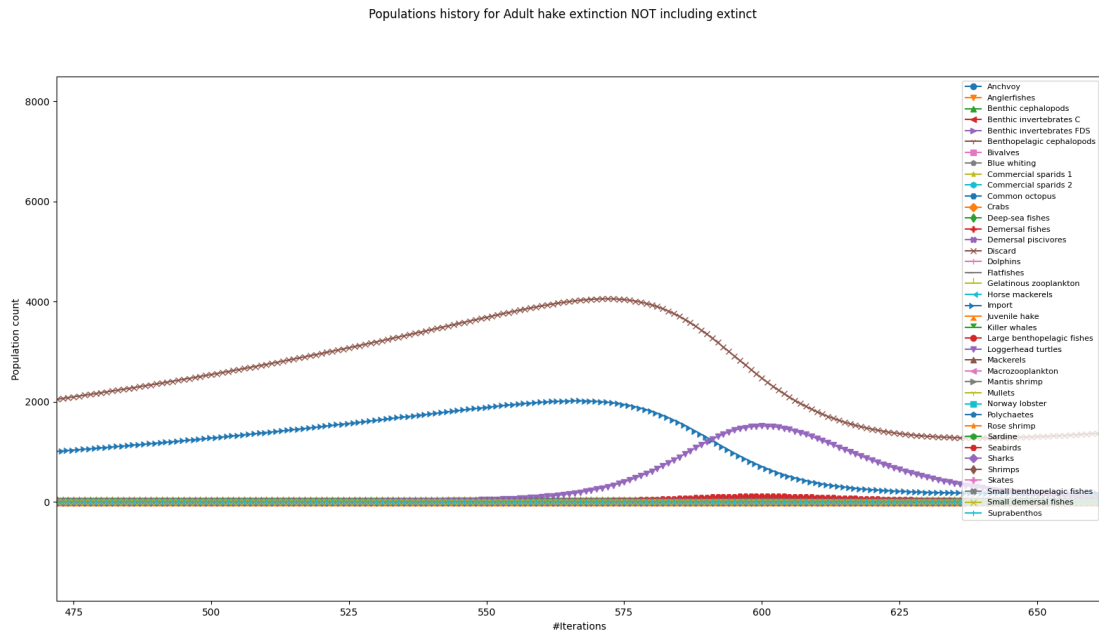
Figure 8: Species survivors after 1000 iterations when Adult Hake goes extinct - Zoom between 475 and 650 iterations

Part 4

Conclusions

We were capable of obtaining coherent outcomes of species population behaviors after species extinction. Considering the common knowledge of which specie preys on another, some results can be predicted, such as the catastrophic consequences of the Gelatinous zooplankton extinction. Nevertheless, the extinction of other species such as the Demersal fishes unpredictably resulted in even more extinctions. The number of extinctions resulting from each initially extinct specie is considered an important finding of our work, being a possible indicator of the specie importance in the Gulf of Cadiz ecosystem. Attempts on predicting this effect using network descriptors have failed (with close-to-zero correlations) so, until better network-based methods are defined, simulations seem the better approach for this task. Analysis of populations histories throughout simulations have allowed the discovery of interesting patterns, such as the potential relationship between populations explosive growing and extinction and the periodic repetition of wave-like patterns.

However, we have to consider that the selected modeling technique has limitations. First, it ignores many important factors, such as other important species (i.e. human fisherman), space limitations and biological restrictions. This can be observed during some simulations, where the population of species such as Dolphins reach values in the order of 1e9. Other problem is that the predation relative frequencies keep the same independently of the preys populations. In the case that these relative frequencies are due to limitations in predator success this is correct but, if it is only due to predator preferences, it should change depending on the amount of individual of each prey specie. Modeling this requires additional knowledge from the species and significantly increments computational cost, but it's an interesting future work that would lead to more realistic results. In any case, since we don't have a way of validating how close to reality the simulation results are (no ground truth has been found nor easy to obtain), no strong conclusions should be extracted from them.

As future research direction, the proposed or an improved simulation framework can be employed for modeling the extinction of multiple species at the initial state or exploring the consequences of a partial extinction rather than a complete one.

References

[1] Torres MA., Coll M., Heymans JJ., Christensen V,. Sobrino I. (2013) Food-web structure of and fishing impacts on the Gulf of Cadiz ecosystem (South-western Spain). Ecological Modelling 265 (2013) 26–44

[2] Orterga R., Fortuna MA., Bascompte J. (ND) Structure and dynamics of ecological networks. Source: `https://www.web-of-life.es/map.php`

[3] Holling, Crawford S. "The components of predation as revealed by a study of small-mammal predation of the European Pine Sawfly1." The Canadian Entomologist 91.5 (1959): 293-320.