Reconsidering an Agent-Based Model of Food Web Evolution

Samuel Armstrong[1,2] and Emma Norling[1,3]

[1] Department of Computer Science,
The University of Sheffield,
United Kingdom
[2] `samuel_e_armstrong@yahoo.co.uk`
[3] `e.j.norling@sheffield.ac.uk`

Abstract. This paper reimplements and extends a prior agent-based model of food web evolution. The earlier work attempted to replicate the results achieved by a system dynamics model of food web evolution, but failed to achieve the diversity or realistic dynamics of the system dynamics approach. This work starts by adding spatial diversity to the model, the lack of this being flagged as a potential problem in the original work. This produced some improvement in the results (with more diverse food webs being produced), but there were still patterns commonly found in the resultant food webs that are uncommon in real-world food webs. To further refine the model, a more complex representation of species traits was added, and methods for classifying species based upon the traits. In particular, an unsupervised learning clustering algorithm has been introduced to classify species in the evolving food web. This has resulted in a model which produces abstract food webs that far more closely mimic the patterns found in real-world food webs.

Keywords: Agent-based simulation · Spatial representation · Food webs

1 Introduction

The way in which a food web evolves to support diversity in an ecosystem has long been a key area of interest for ecological scientists. One method for unlocking some of the complex dynamics behind the evolution of a food web is by simulating it artificially. This has been attempted many times, using many different methodologies and strategies, all with varying degrees of success. The work described in this paper builds upon earlier work by Norling [10], which attempted to create an agent-based model of food web dynamics that mirrored a more traditional system dynamics model. That work did demonstrate the emergence of some simple food webs, but "complex food webs do not arise, even for runs over extended periods." A number of possible reasons were postulated in the original paper, particularly the lack of a spatial representation, which meant there were few advantages to the specialisation of species. Despite these issues, the agent-based model showed "a certain level of correspondence to the system dynamics model," and the paper also acknowledged that an "agent-based model will allow the exploration of a range of variations that would be difficult (if not impossible) to encode in the system dynamics model."

In this project, the original work was replicated, using a modern framework (as agent-based modelling frameworks have evolved considerably in the intervening period), and subsequent iterations looked at the effect of considering spatial diversity and heterogenous behaviour of the agents. The results of these extensions demonstrate a far higher ability to parallel the diversity of species and relationships between species that are observed in real world food webs. Section 2 looks at food webs and their evolution, as well as considering the challenges of modelling these types of systems. Section 3 explains the (reimplementation of) the original model by Norling, and Section 4 considers first the addition of a notion of spatial diversity, and second, the addition of heterogeneous behaviour (dependent on agent characteristics) which evolves with species. The paper concludes in Section 5 with a summary and ideas for future work.

2 Food Webs and their Evolution

In simple terms, food webs are a network of all the predator-prey relationships that exist in a particular ecosystem, and these evolve along with the species within them. According to the Darwinian theory of evolution by natural selection, species will evolve based on a natural bias towards traits that offer them an advantage to survival within their ecosystem [6]. There are a wide variety of traits that are desirable for species in each ecosystem that relate to all aspects of survival. For example, it may benefit a species to have excellent eyesight as this will help them hunt for prey – allowing them to survive; but equally it may benefit them to have thick fur to help them maintain body heat in a cold environment. Each of these factors benefits the survival of the species, though they help the species in very different ways.

As alluded to by Brown, Reilly and Peet [2], there are many external factors that affect which traits are beneficial for a particular species. These include geographic factors, such as different physical features occurring in the environment; biological features, such as competition; and environmental features, such as the availability of resources. Combinations of these factors create a wide array of opportunities in the environment for species to evolve and fill their own niche spot in the food web. This may be especially true with geographical features, as certain characteristics will heavily favour a species if it lives in one environment, but may not be an advantage at all if it lives in another environment. For example, a species could evolve to be very small to easily hide in the cracks of a rock formation where it lives, but if it was living in the nearby desert, its size would not help it hide and would mean that it was prey to a lot of other species.

It is the combination of all the factors above that, over millions of years of evolution, leads to the vast speciation and diversity seen in food webs.

2.1 Patterns Observed in Real-World Food Webs

There are a few characteristic patterns that are often observed in real-world food webs. One such characteristic is the presence of trophic levels. Trophic levels in an ecosystem are groups of species that represent one stage of the transfer of energy up the food chain. As an example, both sharks and orca don't have any natural predators, meaning they are on the highest level of the food web, usually level four or five. Food webs are usually large and complex, depending on the biodiversity of a given ecosystem, and are constantly changed according to the evolution and extinction of species [9]. Food webs also tend to have a feed-forward structure where species in the higher trophic levels feed on species in the lower trophic levels. This is a almost universal principle, with there being few instances where a lower trophic level species can feed on a higher trophic level species, and even the presence of bi-directional connections – where two species can feed on each other – is generally a very rare phenomena. Another characteristic is that omnivores rarely occur as food webs containing them "tend to be less stable than webs without omnivory" [8]. These are the sort of characteristics that would ideally be observed in a simulation of the evolution of a food web.

2.2 Population Dynamics

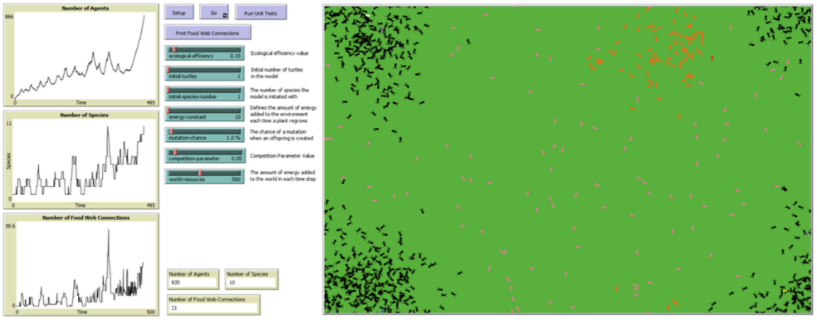
Various studies reinforce the need to understand the changing size and structure of a population over time (recent examples include [1, 5, 7]). The population dynamics are the underlying cause of the system-level effects seen in populations – so in order to accurately depict the diversity found in real-world food webs, it is important to accurately simulate the population dynamics that ultimately lead to this diversity. Traditional methods of simulation often use mathematical models, or a 'system dynamics' approach [10], to attempt to recreate some of these dynamics. These models can, however, over-simplify the population dynamics, which can lead to simulations that do not act in the same manner as real-world food webs. As an example, Rossberg [12] explains that models like this may use an ordinary differential equation to model the exponential growth of a population. However, they do not take into account additional factors, such as the effect of the exponentially growing species on the environment they live in, which would make the real-life effect differ from the simulated effect.

Morin and Lawler [8] identify some of the key population dynamics that have been observed in real-world food webs. The first of these is that food webs of greater complexity tend to have a less stable structure, thus potentially leading to more instances of species extinction or changes in the species that are thriving. Similarly, the introduction or mutation of new species in the environment can have significant impact on the existing population, such as over-predating species or making species uncompetitive, which would likely lead to the extinction of those species. This is one of the key principles of the evolution of food webs, and can be observed in the real world. For example, the introduction of cane toads in Australia has caused a large reduction in the population size of many of the predator species that already existed in the environment, as the toads were better adapted to survival than other species, making the existing species uncompetitive [3]. In addition, if a new species appeared that could take advantage of a niche spot in the environment, such as an inhospitable location, then further species may be able to evolve to feed on this new species.

Overall, population dynamics mean that real-world food webs are an ever-changing network of species that are constantly evolving. When observed over significant time periods, food webs are never completely stable. These are all types of dynamics that should be sought after for a successful simulation of the temporal evolution of food webs.

3 Reimplementation of the Previous Agent-Based Model

Rather than try to work with code that was written for a very old version of Repast Simphony [11], the key features of that model (as described in [10]) were reimplemented using NetLogo [14]. This was largely a pragmatic step, as NetLogo has a less steep learning curve than Repast. Fig. 1 shows an example of the interface for this reimplementation.



ecological-efficiency 0.10 Ecological efficiency value
initial-turtles 1 Initial number of turtles in the model
initial-species-number 1 The number of species the model is initiated with
energy-constant 10 Defines the amount of energy added to the environment each time a plant regrows
mutation-chance 1.0 % The chance of a mutation when an offspring is created
competition-parameter 0.05 Competition Parameter Value
world-resources 500 The amount of energy added to the world in each time step

Number of Agents 835
Number of Species 10
Number of Food Web Connections 21

Fig. 1. A snapshot showing a run of the reimplemented model in NetLogo. The graphs display (from top to bottom) the number of agents in the model, the number of species in the model, and the number of food web connections in the model as a function of time. On the right-hand side, different colours represent different species of agents.

In the description of the original model, it was not clear whether agents should be able to consume agents of their own species. In the reimplementation, experiments were done with this both turned on and off, keeping all other factors consistent. It was found that if agents were not allowed to consume agents of their own species, this would lead to one extremely dominant species that destroys any diversity in the ecosystem. In contrast, when they were allowed to consume their own species, it meant that there was a more diverse array of species that were able to co-exist in the environment. Therefore, the decision was made to allow agents to consume their own species for this reimplementation. This essentially means each individual agent is only trying to survive itself and does not consider the health its species as a whole, which is a trait that can be observed in many animal species alive today.

The aim of the original work was to replicate the results from a system dynamics model of food web evolution [4], and as in that system dynamics model, the species within the model do not correspond to those in any particular natural ecosystem, but are stylised species, each defined by a set of features. The features themselves are also abstract - conceptually they could be things such as "tough skin", "fast locomotion" or such; things that could potentially give them an edge over other species. In this reimplementation the same principles have been followed.

3.1 Results from this Model

The food webs produced by the reimplemented model, such as Fig. 2, consistently find that species which solely rely on predation struggle to survive, whereas species that consume both other agents and the world have a much better rate of survival. This leads to food webs that have a somewhat unrealistic structure, as species are well adapted to feed on everything, whereas in real-world food webs, species generally adapt to feed on specific food sources. In addition, having species that are well adapted to consume anything in the environment means that the food web lacks significant species diversity. This is due to most new species not being able to compete with the dominant species, and so quickly dying out, and when a species appears that is competitive it will simply cause the currently dominating species to die out. Fig. 2 shows significantly less diversity than would be found in a real-world food web. This also occurred in the original work, but was even more extreme in the reimplementation.

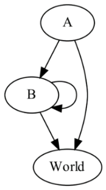


Fig. 2. A typical food web produced by the reimplemented model.

One of the issues with food webs as shown in Fig. 2 is that the higher-level species (A) has a strong advantage over the lower-level one (B): they can feed *both* from the environment and from the other species. This leads to an explosion in the population of A, followed by the extinction of population B, and then a return to lower levels of population A (due to the decrease in food sources). This contrasts with a more classical balance of species that is often considered in agent-based modelling, such as wolf and sheep [13] or foxes and rabbits. In such models, one sees a growth in the predator species as they feast upon their prey, but as the prey becomes scarce, so too do predator numbers fall, due to lack of food.

This reimplemented model generally produces very similar results to those found with the original model, both in terms of the dynamics observed and the food webs produced. As an example, the lack of pure predators found in the food web was also an issue in the original model, and it meant that only simple food webs could be produced. Having this problem appearing in the models shows a good level of correspondence of the reimplemented to the original model, but is an issue that certainly needs to be addressed in future models. The inclusion of a spatial representation in this model did not seem to have any significant impact on the results when compared to the original model. This was somewhat expected, as the literature survey identified a spatial representation as being a tool to unlock new possibilities for improving the agent-based model, but not necessarily being a major improvement on its own. In addition, the use of random rather than fixed mutation did n0t seem to have a significant impact, which was beneficial as the inclusion of the parameter was only due to ambiguity in the original work.

As can be seen in Table 1, the model only sustains an average of 1.623 species. This quantitatively backs up the observation of the lack of food web diversity produced by this model and provides a point of comparison for comparing the diversity of the food webs of this model to other models produced in this project.

4 Enhancements to the Model

Replication of the baseline model confirmed the shortcomings observed in [10]; the real aim of this work was to try to improve upon those results. The first step towards this was to introduce a constant decay rate of the energy level of agents. This means that the energy level for each agent is reduced by a constant amount in each time step, resulting in an agent's death if their energy falls to zero. In addition, a different approach to distributing energy from the world to the agents is introduced: a plant species which can be consumed by agents with appropriate characteristics (that is, characteristics determine how successfully the agents can gain energy from plants). The plants are then being removed from that place in the environment for a refractory period whilst they 'regrow'.

4.1 Geographic "niches"

Distinct geographic regions are introduced to the environment, creating additional niche spots in the environment that species can adapt to fill, which is one of the key factors in the speciation of real-world animals. Thus, this change should support a greater diversity of species that are able to survive in the environment. The different environments will be modelled as containing different species of plants, so as to have differentiation between the different areas (and different animals will 'prefer' – that is, consume energy more successfully from

– different plants). To put into context how these species of plants define geographic regions, an example in nature is cacti being in the desert and tall trees existing in the rainforest. Animals have to adapt to consume these in the same way that agents would need to adapt to consume different plants in the model.

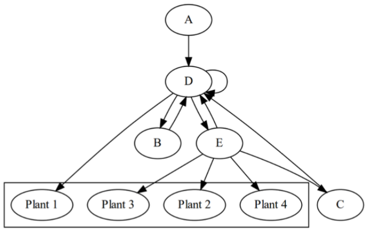


Fig. 3. A typical food web produced in a model with geographic specialisation.

Fig. 3 shows the food web created from one run of the extended model, after around 3000 time steps. The initial observation upon viewing this food web is the additional complexity it contains compared to any food web found with the reimplemented model, sustaining five species on top of the four plant species. This immediately addresses one of the primary issues with the reimplemented model – the lack of diversity found in the food webs produced by the model. In addition, the roles that species are playing in this food web are much better defined and are more comparable to real-world food webs, with one clear basal species and an apex predator. The food web produced is split into four distinct trophic levels, starting with the plants as the first level. Species E is the primary basal species, which is capable of feeding upon most of the plants in the environment. The energy then propagates up the food web to the next trophic level, containing species D. Species D is the sole predator of species E, although it is actually an omnivorous species as it can also consume plant 1. Species A, B and C are all predators that have no capability to eat plants, so solely rely on predation to survive. These form the final trophic level of the food web, with species A being the apex predator as it has no predators.

There are some patterns in this food web that do not correspond to what might be expected in a real-world food web. For example, Fig. 3 shows a 'cycle' of connections where species D can eat E, E can eat C, and C can eat D. Structures like this are highly uncommon in real world food webs. Energy is lost in the transfer from one trophic level to the next, which means that a cycle in a food web should not be able to sustain itself. Another problem with the structure of the food webs produced is the presence of bi-directional connections, two of which can be seen in Fig. 3. While these occasionally do exist in real-world food webs, the frequency with which they occur in this model is unrealistic, once again to the nature of energy transfer between trophic levels.

By including the geographical areas in this model, the additional niche spots mean that this model overcomes the 'competitive exclusion principle' population dynamic that was occurring in the reimplemented model. Instead, the niche spots favour the evolution and survival of new species by some of the same principles that cause diversity in real-world food webs, which were highlighted in the literature survey.

4.2 Adding Heterogeneous Behaviour

Heterogeneous behaviour is where each species has a set of unique characteristics that determine their behaviour and abilities, each of which has a value. These values are on a continuous scale and represent traits that affect multiple aspects of their behaviour, rather than just their interactions, as was the case with the previous models. For example, an agent might have the ability to run faster, but this could come at the cost of a higher metabolism. In this version of the model, each agent has four features which are capable of mutation: predatory ability, plant eating ability, defence, and speed. To ensure that these values do not simply grow unconstrained, a fifth characteristic, metabolism, has been added which derives from these:

$$Metabolism = PredatoryAbility + PlantEatingAbility + Defence + \frac{Speed}{5} \quad (1)$$

This metabolism characteristic provides a trade-off for having better values for other characteristics, and thus prevents the values of these characteristics growing exponentially for the whole population. In this calculation the speed is divided by five as it is initialised as a value that is on average five times bigger than any of the other values, which are all initialised within the same range. In addition to agents having characteristics, plants will also have a set of characteristics which don't evolve over time, but do define different 'species' of plant according to the area in which they preside.

As in the previous models, an agent can eat another agent or plant when they are in the same geographic location. Whether or not a given agent can feed on another agent is defined by the following interaction rules:

$$x = \begin{cases} Eat, & \text{if } PredatoryAbility \times random(0,2) > Defence \\ Ignore, & \text{otherwise} \end{cases} \quad (2)$$

The random value is designed to add a level of stochasticity to interactions, which should aid the evolution process. In addition to these rules, in this model, an agent cannot eat another agent that is of the same species as itself, and an agent cannot eat another agent that has a higher predatory ability score than itself.

In this model species reproduce in the same way as they did in the first two models – they produce an offspring once they are above a critical energy threshold. When an offspring is created, each of the feature values of the parent are randomly modified by a small amount, meaning the offspring is slightly mutated. This mechanic replaces the mutation mechanic of the first two models and makes this process more organic, as the mutated agents are able to find a niche over time by mutating further rather than immediately either finding a niche or not finding one – and either thriving or going extinct as a result.

An additional modification in this version was that instead of each area having a distinct plant species, each area has unique plant defence and plant energy values. The purpose of this is to allow species to adapt to particular regions, for example to adapt to an area that has a higher plant defence, but also has less competition. Watkins et al. (2015) used a similar method for differentiating between different geographic regions with great success. However, evaluation of this model determined that this method of differentiating the different environments was not as effective as the methods used in the extended model, which is perhaps due to the regions not being expressed by enough unique characteristics. Future work in this area could experiment with alternative methods or characteristic to describe the different regions.

The most challenging task with this variation was how to identify species within the model. In the previous models, there are clear boundaries which mean each agent has a clearly-defined species. In this variant, each agent has a unique feature vector of values rather than a species. The initial approach to this problem assigned each agent an individual class for each of each of its features. These classes were determined by how the agent's value for a given feature compared to the maximum value that any existing agent had for that feature. This feature would then be assigned a value from one to four depending on how its value compared. This was run for all the features of each agent and the array of classes assigned to an agent would determine its species – this is referred to as *linear species classification*. As can be seen in Fig. 4 (left example), this approach produces a diverse food web, but in fact *too* diverse – indicating that the classification mechanism might be dividing the agents into too many species.

To address this issue, a form of unsupervised learning clustering algorithm was introduced to identify species. To the best of our knowledge, unsupervised learning has never been used in this context before, so represents a novel approach to species classification in models of food web evolution. Many options for algorithms were considered and experimented with, but ultimately the DBSCAN clustering algorithm was applied, as it both runs reasonably quickly and, most importantly, does not require a predetermined number of clusters. The flexibility to produce an undefined number of clusters was a key attribute required for the problem, as the number of species that exists at any given time is not known. The DBSCAN algorithm works by starting with a random initial data point and looking for any other data points that are within a defined distance, epsilon, of it. If it finds any data points, these are added to the cluster and the algorithm will be run again to try to find any more data points are within the scope of any point in the cluster. Once no more data points can be found, if the size of the cluster exceeds a given threshold, min_samples, then the cluster will be finalised and the algorithm will randomly select a data point that is not in a cluster to continue the process. If the threshold is not met the data points found will not be assigned to a cluster.

This feature of the simulation was implemented in Python using NetLogo's 'py' extension. It used Scikit-Learn's 'cluster' module (Pedregosa et al., 2011), which was passed the values for each parameter of each agent that were standardised with a zero mean and unit variance using Scikit-Learn's 'StandardScaler' class. The value of epsilon used was 0.5 and the min_samples value used was 5. Python then returned the labels found by the algorithm to NetLogo, where they are used for plotting the number of species, colouring the agents according to their species, and finding the food web connections between the species.

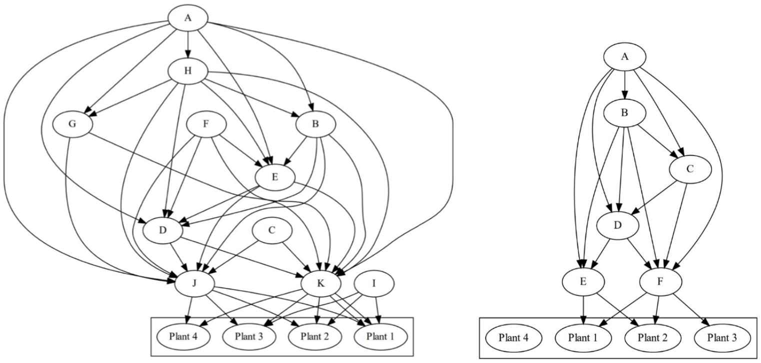


Fig. 4. Typical food webs produced in a model with heterogeneous behaviours. In the left-most example, species were identified using linear species classification, while in the right example, species were identified using unsupervised learning.

Fig. 4 shows food webs that were created from typical runs of the simulation, with one using linear species classification on the left, and one using the unsupervised learning approach on the right. The first observable difference is their complexity. In the version with linear species classification (left), there are eleven species of agents in total, with a great deal of food web connections between the species. The food web also does not contain the bi-lateral connections or cycles that were found in the food webs of the extended model, meaning this food web has a more natural feedforward structure, with better defined trophic levels. There are around five trophic levels that can be seen in this food web, which lines up well with the number of trophic levels found in real-world food webs, per the literature survey. Likewise, this food web does not contain any omnivores, which the literature survey identified as a relatively rare occurrence, meaning this food web is certainly overall the most realistic that any model has produced during this project. This certainly shows initial promise to the model, but equally these results could be due to the linear species classification approach having too many possible categories for each characteristic.

Table 1. The average number of species evolved in runs of the different variants of the model (averaged over 500 time steps)

Run	1	2	3	4	5	Average
Replicated model	1.693	1.795	1.554	1.644	1.427	1.623
Geographic niches	5.691	7.365	4.397	4.084	5.093	5.326
Heterogenous behaviours	7.808	9.746	10.591	9.914	6.849	8.982

Once again, with this unsupervised learning classification method the average number of species were considered, again over 500 steps, as shown in Table **??**. The average number of species the is sustained in the heterogeneous behaviour model is significantly better than both the extended model and the reimplemented model – 8.982 compared to 5.326 and 1.623 respectively.

5 Discussion and Future Work

This final model achieved some very interesting results and produced the most realistic food webs of any model developed in this project. It introduced the novel approach of using unsupervised clustering to determine the different species in an agent-set, with the results achieved corresponding well to real world food-webs and producing more realistic species than a linear species classification approach.

The heterogeneous behaviour model represents the first steps into a new style of agent-based modelling of the evolution of food webs, showing that heterogeneous behaviour with unsupervised clustering is an interesting and fruitful concept for the purpose of simulating the evolution of a food web. Some such potential future improvements include:

- Multi-agent reproduction – currently all the models presented here have a simple system of reproduction, such that when any agent surpasses an energy threshold, an offspring agent is produced that has that same or slightly mutated characteristics from the parent. Including multi-agent reproduction (i.e. two agents to be together and offspring to have characteristics draw from both parents) would potentially help species adapt more effectively to local regions, as the individuals in that species would have to be in the same spatial location to reproduce.
- Additional wider environmental diversity – currently environmental niches are determined solely by the plant species that exist there. Varying terrain or water availability could allow more species to adapt to niche spots in the environment, leading to greater diversity of species.
- More descriptive features for agents – this could come in the form of adding additional features to agents, such as size or strength, or having the existing features more intricately describe agent behaviour (such as making the existing feature vector multi-dimensional with multiple vector values representing features like the predatory ability).

- Exploring further approaches for defining the species – such a different clustering algorithm or some form of tree structure that defines species based on the parents of agents.
- Increasing the model scale – including more environments and many more agents could help to yield more insightful results as these conditions better parallel those found in the real-world.

References

1. Bestley, S., Ropert-Coudert, Y., Bengtson Nash, S., Brooks, C.M., Cotté, C., Dewar, M., Friedlaender, A.S., Jackson, J.A., Labrousse, S., Lowther, A.D., et al.: Marine ecosystem assessment for the southern ocean: birds and marine mammals in a changing climate. Frontiers in Ecology and Evolution **8**, 338 (2020)
2. Brown, R.L., Jacobs, L.A., Peet, R.K.: Species richness: small scale. eLS (2007)
3. Burnett, S.: Colonizing cane toads cause population declines in native predators: reliable anecdotal information and management implications. Pacific Conservation Biology **3**(1), 65–72 (1997)
4. Caldarelli, G., Higgs, P.G., McKane, A.J.: Modelling coevolution in multispecies communities. Journal of theoretical biology **193**(2), 345–358 (1998)
5. Campetella, G., Chelli, S., Simonetti, E., Damiani, C., Bartha, S., Wellstein, C., Giorgini, D., Puletti, N., Mucina, L., Cervellini, M., et al.: Plant functional traits are correlated with species persistence in the herb layer of old-growth beech forests. Scientific reports **10**(1), 1–13 (2020)
6. Darwin, C., Bynum, W.F.: The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life. Penguin Harmondsworth (2009)
7. Jouval, F., Bigot, L., Bureau, S., Quod, J.P., Penin, L., Adjeroud, M.: Diversity, structure and demography of coral assemblages on underwater lava flows of different ages at reunion island and implications for ecological succession hypotheses. Scientific reports **10**(1), 1–13 (2020)
8. Morin, P.J., Lawler, S.P.: Food web architecture and population dynamics: theory and empirical evidence. Annual Review of Ecology and Systematics **26**(1), 505–529 (1995)
9. National Geographic Society: Food web. Available at: https://www.nationalgeographic.org/encyclopedia/food-web/ [Accessed Jan 2022]
10. Norling, E.: Contrasting a system dynamics model and an agent-based model of food web evolution. In: International Workshop on Multi-Agent Systems and Agent-Based Simulation. pp. 57–68. Springer (2006)
11. North, M.J., Collier, N.T., Ozik, J., Tatara, E.R., Macal, C.M., Bragen, M., Sydelko, P.: Complex adaptive systems modeling with repast simphony. Complex adaptive systems modeling **1**(1), 1–26 (2013)
12. Rossberg, A.G.: Food webs and biodiversity: foundations, models, data. John Wiley & Sons (2013)
13. Tatara, E., North, M., Howe, T., Collier, N., Vos, J., et al.: An indroduction to repast simphony modeling using a simple predator-prey example. In: Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects. Citeseer (2006)
14. Tisue, S., Wilensky, U.: Netlogo: A simple environment for modeling complexity. In: International conference on complex systems. vol. 21, pp. 16–21. Boston, MA (2004)