

University of Sheffield

# Re-Implementing and Building upon an Agent-Based Model of Food Web Evolution



Samuel Armstrong

*Supervisor:* Emma Norling

*Module code:* COM3610

This report is submitted in partial fulfilment of the requirements for the degree of Artificial Intelligence and Computer Science by Samuel Armstrong.

May 19, 2021

## Declaration

All sentences or passages quoted in this report from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure in this project and the degree examination as a whole.

Name: Samuel Armstrong

Date: May 19, 2021

## Abstract

The work of Norling (2007) explored the use of an agent-based model to simulate the evolution of species in a food web. The agent-based model implemented in this research attempted to emulate the performance of a corresponding system dynamics model in order to demonstrate the efficacy of agent-based modelling in this field. However, the model did not manage to achieve the diversity or realistic dynamics that the system dynamics approach did. The model that was implemented used a non-spatial approach to model the interactions between agents using a highly abstracted approach that did not accurately correspond to real-world interactions. This leaves many opportunities for this work to be extended to create an agent-based model that is built upon more realistic interactions and behaviours, and yields food webs that more realistically parallel those observed in the real world. This project aims to re-implement the original work and explore potential improvements to the model that could allow it to better simulate the evolution of food webs, with the aim of producing realistically structured food webs and displaying intricate population dynamics.

## COVID-19 Impact Statement

The lockdown imposed because of COVID-19 caused additional challenges for the completion of this project. In the second semester of the project, the university switched to online delivery of all teaching, and university buildings were closed. All project meetings were shifted to email correspondence and video meetings. This project was also somewhat affected by the limited computational hardware available due to working entirely from home.

## Acknowledgements

I owe a great deal of thanks to my supervisor, and inspiration for this project, Dr Emma Norling, whose support, encouragement and feedback throughout this project made my first venture into research both enjoyable and rewarding. I would also like to thank my family and friends for their ceaseless love and support through what has been a tough year for everyone.

Finally, I would like to extend my gratitude to everyone in the Department of Computer Science at the University of Sheffield. This was my final year as an undergraduate student at the university and, although it has not been a normal year, it has been an enjoyable and challenging rollercoaster that I will remember for the rest of my life. This would not have been possible without the hard work that all members of staff have put in to ensure that we can receive the best education possible in this challenging time for the world. I truly appreciate this and cannot thank them enough.

## Contents

Abstract.....	iii
Contents .....	vi
1 Introduction .....	1
1.1 Aims and Objectives.....	2
1.2 Overview of the Report.....	2
1.3 A Precautionary Note.....	2
2 Literature Survey.....	3
2.1 Introduction .....	3
2.2 Background .....	3
2.3 Modelling Options.....	6
2.4 Improvements for the Agent-Based Model .....	7
2.5 Software Choices .....	11
2.6 Evaluating the Performance of Agent-Based Models.....	12
2.7 Summary .....	13
3 Requirements and Analysis.....	14
3.1 Functional Requirements.....	14
3.2 Non-Functional Requirements.....	17
3.3 Model Evaluation .....	18
3.4 Ethical and Legal Considerations .....	19
3.5 Summary .....	19
4 Re-Implementation Model.....	20
4.1 Design.....	20
4.2 Implementation .....	24
4.3 Results and Discussion .....	28
4.4 Findings and Next Steps.....	31
5 Extended Model.....	32
5.1 Design.....	32
5.2 Implementation .....	34
5.3 Results and Discussion .....	36
5.4 Findings and Next Steps.....	40
6 Heterogeneous Behaviour Model.....	41

6.1	Design.....	41
6.2	Implementation and Testing.....	44
6.3	Results and Discussion .....	47
6.4	Findings and Future Steps.....	52
7	Conclusions .....	53
7.1	Findings of the Project .....	53
7.2	Further Work.....	54
8	Bibliography .....	55

# 1 Introduction

The way in which a food web evolves to create such diversity in an ecosystem has long been a key area of research for ecological scientists worldwide. There are many complicated dynamics at play in these systems which combine to create the diverse array of species that co-exist in a food web. Exploring these dynamics, their causes, and the implications this has for the development of the ecosystem is a large-scale area of research and many tools are needed for this research to scale as more advanced technology becomes available.

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. A method of simulation that has become more prevalent in recent years is agent-based modelling – a method of simulation that focusses on individuals in the environment and their interactions, rather than the behaviours of the system as a whole. The principle behind agent-based modelling is that the ‘agent to agent’ and ‘agent to environment’ interactions will lead to emergent behaviour of the system, much in the same way as the interactions between individual animals, and with their environment, in a real-world food web ultimately lead to the evolution of species and the development of that food web.

The paper that will form the basis of this project, ‘Contrasting a System Dynamics Model and an Agent-Based Model of Food Web Evolution’ (Norling, 2007), implemented an agent-based model that attempted to model these food web dynamics, and was contrasted with a more traditional system dynamics model. A simple agent-based model was implemented for this research that had the agents interacting randomly, rather than based upon a notion of spatial location. This model was designed primarily with the purpose of reproducing the results that were found in the system dynamics model – though of course achieving these results through a different means. The results from this model did show some simple food webs emerging, but ‘complex food webs do not arise, even for runs over extended periods’. The paper details a number of possible reasons for this, including the lack of a spatial representation – which unintentionally gave each ‘individual agent a distorted picture of the population’, which potentially caused ‘a lack of stratification of the food web’. 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’.

These statements, along with the fact that only a simple model was implemented in this study, means that there is a lot of potential for the model to be improved. By re-implementing this work and improving/updating the agent-based model, it should be plausible to create a model that can display some of the intricate dynamics of food webs, with some parallels to those observed in real-world food webs. This paper will explore why agent-based modelling is such a useful tool for the purpose of simulating the evolution of food webs, and build upon the baseline of the model developed in Norling (2007). This work will be re-implemented, due to the source code being unavailable (and potentially antiquated due to its age), and it will be extended to try to improve the diversity of the food webs emerging and the realism of population dynamics observed.



## 1.1 Aims and Objectives

The aim of this project is to create an agent-based model of the evolution of a food web that displays similar dynamics to those in a real-world food web, such as predator-prey relationships, and ultimately results in a realistically structured food web. For this to be achieved, a well-researched body of knowledge must be built in the fields of agent-based modelling and the dynamics of an ecosystem.

### 1.1.1 Objectives

- Evaluate why agent-based modelling is the best approach for modelling the evolution of a food web, and understand the potential improvements it offers
- Re-implement the original work (Norling, 2007) using similar mechanics
- Make improvements and updates to the agent-based simulation so that it displays more realistic population dynamics and the food web produced displays diversity and structure closer to that observed in real-world food webs

## 1.2 Overview of the Report

This report assesses the practicalities of using agent-based modelling to model food webs. This is based on the work of Norling (2007) and is seeking to re-implement and extend this work to explore how changes to the model can increase the complexity of emergent food webs. To begin this work, the report reviews literature in relation to:

- The background of simulating the evolution of food webs
- The practicalities of agent-based modelling compared to other models
- The application of agent-based models to food webs and ecosystems
- Ways in which simple agent-based models can be improved to create more realistic simulations
- Software tools and frameworks for creation of agent-based models
- How an implemented model could be evaluated to assess its performance

Further chapters will build on the knowledge obtained from the literature review to analyse, design, and implement three original agent-based models in the chosen software framework to fulfil the requirements laid out. Finally, the results of each model will be evaluated and discussed.

## 1.3 A Precautionary Note

Unfortunately, due to the complexity and spontaneity of real-world systems, making a simulation that replicates the real world with complete accuracy is an impossible goal. The purpose of the simulations that were built in Norling (2007) and will be built for this project are to replicate some of the dynamics found in real-world food webs, rather than to replicate all of the dynamics. Likewise, the models will be simulating a general food web rather than any specific food web. This means that simulations are not directly comparable to real-world data, which can be an issue when it comes to evaluating the performance of the model. This issue is addressed later in the report.

## 2 Literature Survey

### 2.1 Introduction

To be able to make insightful improvements to a simulation of food web evolution that employs an agent-based model, a baseline of academic knowledge must be built in the area of study. The concepts behind food webs and agent-based modelling must be understood; the reasons why an agent-based approach the most suitable for the purposes of this project need to be identified and underpinned; and ways in which the original model can be improved have to be explored. Additionally, the options for software tools and frameworks are assessed to inform the choices made in the project.

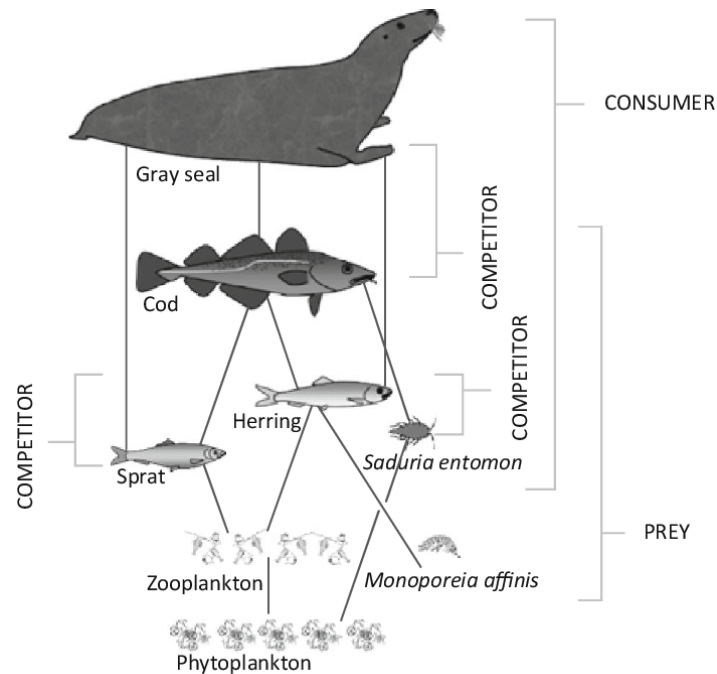
### 2.2 Background

#### *2.2.1 Why do real-world food webs often contain a wide diversity of species?*

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 (Darwin, n.d.). 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 (2016), 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. Figure 2.1 shows an example of a real-world food web.



**Figure 2.1:** Example of a real-world food web (Karlson et al., 2019)

### 2.2.2 What patterns can be 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 (Society, 2021). 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, as can be seen in Figure 2.1. This is an 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' (Morin and Lawler, 1995). These are the sort of characteristics that would ideally be observed in a simulation of the evolution of a food web.

### 2.2.3 Population dynamics

Potts et al. (2020) explains that population dynamics are the reasons behind the changing of the size and structure of a population over time. 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 (Norling, 2007), 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 (2013) 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 (1995) 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 (Burnett, 1997). 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.

Norling (2007) also identifies some population dynamics that are characteristics of real-world food webs. In particular it emphasises the continuous introduction and extinction of species that don't find a niche in the ecosystem, and so are uncompetitive. Conversely, when a new species does find a niche, this can either have little impact on the existing population of species, or it can cause significant changes in the population, such as the extinction of certain 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.

#### *2.2.4 Agent-based modelling*

At its core, agent-based modelling is a method of simulating the dynamics of a system. This method of modelling differs from other methods by producing the system level behaviour through the low-level interactions between 'agents', rather than directly encoding high-level behaviour (Norling, 2007). Agent-based models consist of three core concepts: agents, the relationships between agents, and the environment that the agents persist within (Macal and North, 2010). Agents are individual entities that act autonomously based on their attributes and behaviours – which are set according to the system they are operating in. The process of these agents interacting with each other and with the environment displays the emergent dynamics of the system (Bonabeau, 2002) (Grimm and Railsback, 2005).

## 2.3 Modelling Options

### 2.3.1 *Alternatives to agent-based modelling*

The main alternative to agent-based modelling that is used is system dynamics modelling, which uses mathematical approaches to try to model the high-level dynamics of a system, rather than taking the low-level approach of agent-based modelling. Agent-based modelling is an approach that has been more recently developed, however system dynamics models are still commonplace today and certainly have many practical uses.

Norling (2007) compared the use of an agent-based model with a system dynamics model. The systems dynamics model implemented in this work attempted to model both the population dynamics of the current species in the food web and also the introduction or ‘evolution’ of new species. The entire population is then run through a series of equations to determine how the different species have interacted and what changes to the population (such as the extinction and evolution of species) will result. This system dynamics model produced a realistically structured food web, despite the fact it used no spatial representation, i.e. it had no concept of location. This is one benefit of using system dynamics models, as they can produce reliable results with a global view of the system. Nevertheless, the paper concluded that ‘The 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’, indicating that agent-based modelling can be developed further than system dynamics models. In addition, agent-based models can allow greater insight into the findings of a system dynamics model. In practice, modelling the system-level behaviours can yield useful insights, but being able to explain the phenomena may require agent-based modelling for ‘an understanding at the micro level of why they are so’.

One of the fundamental issues with system dynamics models is that they generally don’t deal well with either environmental or geographical factors. For example, models that incorporate both often suffer from contrasting results and poor modelling performance, as well as being difficult to implement (Sofa, 2011). In contrast, agent-based models are easily able to represent both environmental and geographical factors. An interesting solution to problems such as this is to implement a system dynamics approach in combination with an agent-based modelling approach. A model like this is implemented in Martin and Schlüter (2015) to model social-ecological systems. This paper highlights many benefits that using a combination approach provides, such as the model providing a ‘multi-scopic view on the system’ (Martin and Schlüter, 2015) - essentially where both system-level and micro processes are modelled. However, most of the benefits of this approach come from the context in which it is implemented in this paper. Social-ecological systems combine two different levels of complexity into one system and consequently each level can be modelled using a different method. Therefore, for modelling the system in this project (which only models an ecological system), using a combination approach will not necessarily lead to improved results, and using such an approach would undoubtedly increase the complexity of the simulation – leading to slower run-times and more difficulty in creating a robust solution.

### 2.3.2 *Why is agent-based modelling the most suitable approach?*

There are many benefits that agent-based modelling provides over system dynamics approaches. The main benefit of agent-based modelling is that it finds and deals with dynamics of a system that a mathematical model simply cannot predict or accommodate, meaning that simulating with an agent-based model can reveal emergent phenomena from the interactions of agents (Bonabeau, 2002). Bonabeau states in his paper that the frequent interactions between many agents in an environment can reveal the ‘dynamics out of the reach of purely mathematical models’. Several other benefits stem from this concept. For example, the fact that the system-level behaviours emerge from the interactions of individuals means that the model is highly flexible – it is easy, for example, to increase or decrease the number of agents in the model, or change aspects of the environment they are interacting in.

Agent-based models also ‘provide a natural description of a system’ (Bonabeau, 2002), rather than manufacturing equations that model system-level behaviours. This makes them ideal for the purpose of modelling the evolution of food webs, as the interactions of individual agents should, in principle, provide a far more accurate description of how species evolve than generalised equations. The fact that they can give a more accurate description of the evolution of a food web makes agent-based models potentially more useful for research, as patterns can be observed in the simulation that are more likely to occur in, or give more insight into, real-world food webs. Additionally, this characteristic means that real-world observations can be directly integrated into the model (Grimm, 1999) to make it more realistic, or explore the effects it has on the results. This is a clear advantage over system dynamics approaches, as these tend to be hard wired based on observed trends and patterns. An agent-based model can be quickly and easily modified with different characteristic values to change aspects of the simulation – meaning the system is both far easier to use for people who are not well versed in simulations and can be rapidly changed to explore the impact of different characteristics.

Ultimately, agent-based models are not guaranteed to provide a simulation of a food web that is entirely accurate to real life, but the patterns observed in the results are likely to provide a good insight into the underlying population dynamics. Additionally, the food webs that evolve in the simulation are likely to show similar characteristics to real-world food webs as the simulation can be built on the observed interactions of individuals in the environment.

## 2.4 Improvements for the Agent-Based Model

The agent-based model that was used in the original work (Norling, 2007) was a simple, non-spatial model that was attempting to recreate the ‘system-level’ behaviours observed in a system dynamics model, but emerging from the interactions of agents rather than having the model ‘explicitly encode’ the behaviour. This purpose meant that the agent-based model didn’t fully take advantage of many of the facets of agent-based modelling. This means there are many opportunities for the work to be expanded to enhance the diversity of emergent food webs.

### 2.4.1 Including a spatial representation

Norling (2007) states that possibly the biggest reason for the poor performance of the agent-based model implemented was the lack of a spatial representation. This caused issues with the random sampling giving some agents ‘a distorted picture of the population’. Agent-based models have a surprising variety of potential implementations for a spatially explicit representation, so it is important to understand these before deciding upon the implementation approach.

A spatial representation is where agents in the model are assigned a specific location in an environment (Manson et al., 2020), which contrasts with non-spatial models that use random sampling and give agents no defined locations. One of the key issues with non-spatial models is the fact that they usually assume spatial homogeneity, often leading to a lack of niche spots for new species to fill. Spatial models resolve this issue by allowing different parts of the environment to be change – which allows the effects of various environmental features to have an impact on the agents. In addition, assigning each agent a spatial location prevents the issue of random sampling giving a ‘distorted picture of the population’ (Norling, 2007), leading to more realistic interactions (Gimblett, 2020). This is ideal for modelling food webs, as in the real-world animals have free range in their environment to perform behaviours within, and the environment will inevitably play a role in their interactions. These benefits indicate that there is strong potential for the inclusion of a spatial representation, however there are many options available relating to how the spatial representation is implemented. These need to be researched and evaluated to decide upon an appropriate design choice moving forward.

The paper Baggio et al. (2010) uses a network-based agent-based model that incorporates a form of a spatial representation – different nodes represent different habitats with the links between them representing ways of travelling between habitats. The geographical connectivity of the simulated environment was shown to be one of the most important variables in determining the emergent behaviours of the system. In this simplified spatial model, the spatial connectivity of the different habitats was modified by changing the number of links between them, which was derived from a single variable. For the purpose of making a simulation as accurate as possible, it may be an over-simplified assumption that the geographical connectivity of an environment can boil down to the value of a single variable. In fact, it is all aspects of the environment working in their own way that contribute to the overall effect of connectivity. Nevertheless, the results from this paper show that a spatial representation can be implemented in different ways that have their own benefits. In addition, it shows that it is important to consider the geographical connectivity of the environment, with any spatial representation that is implemented.

Manson et al. (2020) explains that there are many different degrees to which a spatial representation can be included in an agent-based model. Certain simulations don’t require a geographic representation of space, as the simulation is attempting to model dynamics that are not directly linked to the environment in which they occur. Similarly, a spatial model can use a spatial representation that, whilst being important in the dynamics of the simulation, does not map perfectly onto reality – this is the type of representation used in Baggio et al (2010). Finally, spatial agent-based models can use realistic spatial representations that are intended to be accurate to real life, and relevant to the dynamics of the model – this was used in Watkins et al. (2015). This could be the ideal use of a spatial representation for my purposes, as it would allow the inclusion of geographical features, which may allow the evolution of a

more diverse array of species to fill the niche parts of the environment (O'Connell, Streicher, Smith and Fujita, 2017). This often occurs in the real world, with one example being the evolution of species around hydrothermal vents in the ocean (Exploring the evolution of animals at deep-sea hot springs | Natural History Museum, 2020). These geographical structures create unique conditions that species have evolved to survive in, with them not being able to survive in any other environment. Whilst this is an extreme example, it shows that geography offers more niche spots in the environment for species to fill, which is an aspect that can be modelled in a spatial agent-based simulation – especially one that is an accurate representation of real-world environments.

Finally, Norling (2007) states that 'experiments that introduced a form of spatial representation to Model B did not produce significantly different food webs', showing that a spatial representation isn't a singular solution to improving the original agent-based model. Despite this, including a spatial representation does overcome the problem of having the random sampling not accurately representing the population and has had demonstrated success when applied to agent-based models since the original work was completed. Spatial representations also allow further improvements that can help to improve the diversity of the food webs, such as geographical structures. Ultimately, the inclusion of a spatial representation will be an interesting extension to the original work, allow the inclusion of further improvements, and help to make the model more comparable to real-world food webs.

#### *2.4.2 Using more realistic model mechanics*

One thing that is clear from reading the overview of the implementation used in the original work is that the individual agents are a very deep abstraction of the real world. This is understandable, due to the complexity real-world animals exhibit, however there is some potential opportunity to make some minor changes that may improve the performance of the model. These changes are small but relatively fundamental differences in the way in which agents or the environment behave. One example of this is that the original model had agents reproducing by themselves, rather than when they came into contact with another agent of their species. Most animals do not reproduce in this way, so an improvement could be made to the model so that aligns it more closely with reproduction in real world ecosystems.

Norling (2007) also mentions that developing the system to be 'closed', rather than having resources added at each time step, which 'many ecologists consider to be important'. However, the discussion states that this change didn't have much effect on the system behaviour, meaning it may not be the most fruitful extension to pursue for this project.

#### *2.4.3 Heterogeneous behaviour*

One of the biggest potential future developments highlighted by Norling (2007) was the inclusion of 'heterogeneous behaviour'. This is where each agent has a set of unique characteristics that determine their behaviour and abilities, such as speed, size, or average lifespan. These parameters could also be passed from generation to generation, with random mutations and the process of natural selection leading to a natural simulation of the evolution of a food web. The discussion in the paper indicates that this could be implemented in the model to varying degrees, with more advanced models potentially even including 'different behaviours' for 'different species'. In contrast, the original model used a



scoring system of abstract characteristics (which had no defined effects) to determine the outcome of interactions between agents. Similarly, which interactions would take place were simply a result of random sampling. These methods are quite a brittle approach to modelling an ecosystem, as they do not consider the way in which animals roam, hunt, defend themselves, and many other characteristics. By modelling some of these characteristics, it should bring the resulting model far more in line with real ecosystems, and allow greater insight into the emergent behaviours.

Heterogeneous behaviour has the potential to be especially effective alongside a spatial representation, due to the parallels both improvements have with the reality of real-world food webs. This is exemplified in Watkins et al. (2015), which uses both a spatial representation and model parameters for each agent. The parameters used in this model did not mutate and evolve over time, as they were not attempting to model evolution, but some included a degree of stochasticity in each time step – which allowed the model to better resemble the real-world system. This model ultimately achieved results that were generally consistent with the real-world – showing that the use of valued parameters can be used to create an effective agent-based model. This research provides a useful precursor to the inclusion of heterogeneous behaviour for agents, and displays the potential for this extension to improve the performance of the model.

#### *2.4.4 Introducing geographical features*

The inclusion of a spatial representation in the model allows more environmental detail to be encoded within it. One benefit of this is that different geographical features and habitat types can be included, which have features that affect the agents in that area. Previous research has utilised implementations like this, for example R. Vahdati et al. (2019) integrated geographical features into an agent-based model that simulated the ‘Late Pleistocene human dispersals from Africa’. This work shows the usefulness of including such features, as they had a tangible effect on the results of the model – the ‘geographical features prevent populations from dispersing with equal probabilities in all directions and favors certain dispersal paths’. In the context of modelling food webs, the inclusion of geographical features can create more niche spots in the environment which species could potentially adapt to fill – which may allow a more diverse food web to evolve over time (O'Connell, Streicher, Smith and Fujita, 2017).

Watkins et al. (2015) used an agent-based model that used a spatial representation that included geographical features to model the interactions between jaguar populations and their habitats. This model gave each agent unique attributes, including age, gender and energy reserves, which were derived from known jaguar ecology. Each local environment had attributes such as food availability and cell cost, with this information was available to agents located nearby, that differentiated the different geographic regions of the environment. The emergent behaviour of this model showed dynamics that were generally consistent with those observed in natural jaguar populations. These results show that the combination of a spatial representation and geographical features can make a model that produce very realistic findings. This would therefore be an interesting combination of improvements to apply to models in this project.

## 2.5 Software Choices

There are many agent-based modelling software solutions that are widely available and could be used to implement an agent-based model that will simulate the evolution of a food web. These frameworks each have different characteristics that make them more or less suitable for particular problems. Starting from the baseline of all the frameworks detailed in Abar, Theodoropoulos, Lemarinier and O'Hare (2017) the options have been refined down based on what this paper classified as ideal frameworks for 'geospatial simulations'. Those given were Swarm, MASON, Repast Symphony, StarLogo, NetLogo, OBEUS, AgentSheets, and AnyLogic. Manson et al. (2020), also gives six examples of modelling frameworks that are useful for creating geographically explicit agent-based models. The frameworks considered for detailed analysis were recommended by both papers. These were Swarm, MASON, Repast Symphony, and NetLogo.

The ultimate choice of software depends on several factors. Firstly, having a framework specifically designed for the domain of natural/biological sciences may help with the development, as the software will likely include specific features for use in simulating biological systems; although this doesn't necessarily mean a multi-purpose framework shouldn't be used. Additionally, the framework will need at least a reasonable level of modelling strength and scalability to allow the implementation to contain the interactions of many, many agents. Also, having a medium/low development difficulty would help significantly, as the shorter time taken to learn the software package and to develop simulations using it, the more time will be left to explore different improvements and make adjustments.

	Swarm	MASON	Repast Symphony	NetLogo
Language	Java	Java.net	Java	Logo Dialect
Development Difficulty	Hard	Hard	Hard, but flexible	Medium
Modelling strength	Very strong	Medium	Strong	Medium
Scalability	Extreme scale	Large scale	Large scale	Large scale
Domain	Social/Biological sciences, Logistics, Management	Multi-purpose	Multi-purpose	Social/Natural sciences

**Table 2.1:** *Comparison of spatial agent-based modelling software choices (Abar, Theodoropoulos, Lemarinier and O'Hare, 2017) (Manson et al., 2020)*

Overall, each of these four frameworks certainly has the functional capability of modelling biological systems in a practical way for my purposes. Each has their own benefits and drawbacks – in particular Swarm seems suitable for complex models, and NetLogo appears to be the easiest to learn and develop in, without much sacrifice in performance. Ultimately, NetLogo was chosen due to its simplicity of model development, short learning duration and ability to handle relatively complex models. Using NetLogo will allow more time to work on improvements to the simulation. In addition, it is designed specifically for problems in the field of natural sciences, which is the general subject area of this project. NetLogo performs strongly in categories such as modelling strength and scalability, meaning there should not be significant trade-off to using this framework.

## 2.6 Evaluating the Performance of Agent-Based Models

Norling (2007) acknowledged that ‘there are considerable difficulties in getting reliable data on real food webs’. One of the reasons for this that is described in the paper is that it is very difficult to geographically bound most ecosystems. In principle this doesn’t seem like a challenging problem – an ecosystem such as a pond or forest would usually be considered to cover a clear and limited geographic region. The issue arises due to species that are part of both the ecosystem in question and other ecosystems. An example of this are birds that eat fish from a pond, but also eat insects from other ecosystems. This problem ultimately leads to very large and very complex ecosystems from which it is difficult to draw meaningful datasets for model evaluation. This is not the only issue with gathering accurate datasets, as the paper also claims ‘there are tremendous difficulties in gathering data on food webs even for limited geographic domains’.

Acquiring data from third party sources can also be problematic, due to there not being a generalised dataset that applies to all food webs – contrasting with the fact that the agent-based model to be developed will not directly represent a particular ecosystem. Using a dataset for any one given ecosystem would simply not allow useful comparability between this data and the model developed. Interestingly, data such as this has been used for simulation purposes before by Caldarelli, Higgs and McKane (1998), which used averages of the data from many different ecosystems. They used this data despite the fact that by their own admission ‘the problems in getting reliable data on real food webs are readily acknowledged by ecologists’. This means the data is just too unreliable for the purposes of this project and cannot realistically be used for effective evaluation of the models developed.

### 2.6.1 *Evaluating an agent-based model with a dataset*

In contrast to the original work, Watkins et al. (2015) evaluated their agent-based model using a dataset that was obtained using ‘camera trap objects placed in the same locations in the simulated landscape’, which gave a point of comparison between the real-world and the simulation. Whilst this method has clear benefits from an evaluative perspective, it demonstrates the cost and difficulty of obtaining data from ecosystems – just obtaining this data required ‘47 camera traps’ which were run over ‘several months’. In addition, this data only covers the population of jaguars and no other species – which was all that was needed for the study in question, but serves the purpose of showing the difficulty of getting empirical food web data for many different species in an ecosystem. Similarly, ‘Modeling tiger population and territory dynamics using an agent-based approach’ (Carter, Levin, Barlow and Grimm, 2015) used a real-world dataset to evaluate their agent-based model of tiger populations. However, the scope of the work was again limited to modelling a single species, meaning that obtaining meaningful data is far easier than doing the same for an entire ecosystem.

For models where a dataset is available, one method that has been used in the past for evaluating the performance of ecological systems modelled using an agent-based model is pattern-oriented modelling. This requires a pattern to first be observed within the data of a real-world food web, a hypothesis to be made to explain the pattern observed, then predictions to be developed from this hypothesis which can be tested in the agent-based model (Sofa, 2011) (Grimm, 2005). This technique is effective for both testing the performance of an agent-based model and refining the model complexity to find the most

optimal parameters to create the most accurate simulation (Watkins et al., 2015). However, due to the difficulty of obtaining a dataset, and the level of abstraction models of food webs are required to take, this technique is likely to be out of the scope of this project.

### *2.6.2 Evaluating an agent-based model without a dataset*

Due to the issues detailed above, the original model from Norling (2007) did not use a real-world dataset to evaluate its performance. The model was instead evaluated using a qualitative view of how well food webs were developing in the simulation. This was informed by data from the simulation such as the number of species, the total number of agents, and the number of links in the food web. Ultimately, the classification of how well the model was performing was judged based on the complexity of the food webs that were emerging and sustained over a reasonable period of time. This information was gathered from the relationships different species had with each other, and the model was then assessed based on how many 'levels' the food web had at a given time step. This was an effective method of evaluating the model as it allowed different variations of the model to be compared against themselves, and also allowed the model to be compared with other models for similar problems, such as the system dynamics model referred to in the paper.

Similarly to the original work, 'Modeling mountain pine beetle infestation with an agent-based approach at two spatial scales' (Perez and Dragicevic, 2010) did not have a dataset to evaluate the performance of their agent-based model against (although they did have a dataset to inspire the implementation of the model). Instead of evaluating the performance of the model with a dataset, they instead took a qualitative approach to viewing the results. By looking at the results of the model subjectively they were able to see particular dynamics emerging that are known to exist in real-world infestations, which indicate the effectiveness of the model. The arguable disadvantage of this approach is that it requires a level of expert insight into the topic area, so research into the topic area may be necessary.

The solution to the evaluation protocol for the models developed in this project will be discussed in detail in the requirements and analysis stage.

## **2.7 Summary**

To summarise, this literature survey covers the background of simulating the evolution of food webs; the basics of agent-based modelling; alternative modelling options; why agent-based modelling is suitable for a problem such as simulating the evolution of food webs; how the agent-based model from the original work can be improved; the potential software frameworks that the model could be implemented in; and finally, how the model could be evaluated. This provides an academic baseline of knowledge in the topics surrounding the area of study, in preparation for the next stages of the project.

### 3 Requirements and Analysis

The literature survey brought to light many aspects of the project which require review and structure before a solution can be designed and implemented. This section considers the functional and non-functional requirements of the model, as well as evaluation approaches to gauge the performance of the model.

#### 3.1 Functional Requirements

Criticality	Requirement
Must do	Re-implement the work of Norling (2007)
Must do	Include a spatial representation of the agents and environment in the model
Must do	Output graphical and numerical metrics to allow effective evaluation of the model
Should do	Adapt the mechanics of the model so it is more consistent with real-world food webs
Should do	Introduce geographical features to the environment
Could do	Add heterogeneous behaviour to agents
Could do	Modify the mechanics to make it a closed system

**Table 3.1:** *Functional requirements of the project*

##### 3.1.1 *Re-implement the work of Norling (2007)*

The literature survey dissected the benefits of agent-based modelling, and how it has been effectively deployed in the past in related projects. This research discovered many potential changes and extensions that could be added to the original model in order to produce a more complex, diverse, and realistically structured food web. To be able to implement these potential extensions, however, a baseline model for the project is needed. This baseline model will be provided by a re-implementation of the original model that will be developed using NetLogo. Creating an exact re-implementation will not be possible due to several limitations, first of which being that the source code of the original model is not available. Whilst this does allow the freedom of choosing the most appropriate framework, it limits the amount of information known about the implementation. The re-implementation will therefore be based solely on the information detailed in the research paper, with reasonable assumptions being made to fill in for any ambiguity, or anything not directly addressed in the paper. Likewise, the original implementation used Repast Symphony, a Java based modelling framework, but further research has shown NetLogo to be a better choice for reimplementation and model enhancement at the current time. This means there may be some differences in the solution implementation. However, NetLogo is a powerful framework and should allow all aspects of the solution to be included to some extent. The considerations of the research paper will be followed to ensure the re-implemented model is close enough to the original to build upon and draw conclusions from.

### *3.1.2 Include a spatial representation of the agents and environment in the model*

A spatial representation is commonly used in many modern-day ecological simulations – exhibited by Watkins et al. (2015) and R. Vahdati et al. (2019). The research into relevant literature revealed that there are many benefits to taking such an approach, particularly in the context of modelling the evolution of a food web. The clearest example of a benefit of including a spatial representation is the fact that, in the wild, animals can only interact with other animals that are in their spatial vicinity – a spatially explicit model captures the consequences of this natural property far better than a model that uses random sampling. This is backed up by Gimblett (2020) which states that ‘to be useful, the simulation requires that realism and attention must be focussed on a thorough examination of the behavioural mechanisms of individuals and to more accurate spatial representations of the environment’. The spatial variety induced by the use of a spatial representation also allows the inclusion of improvements, such as geographical features. With these benefits in mind, and the fact that this approach is endorsed by Norling (2007), the first key extension to the model will be to include a spatial representation, and this is a must-do requirement.

### *3.1.3 Output graphical and numerical metrics to allow effective evaluation of the model*

One of the key aspects of developing extended models is how they will be evaluated. In order for any of the developed models to be effectively evaluated and compared to other models (either the original model or models from related work), data will need to be available to summarise how the model is performing. As stated in the literature survey, acquiring a suitable dataset for the evaluation of this project would be very difficult. Therefore, the information to be drawn from the model needs to be either intuitive enough that logical visual observations can be made, or be directly comparable to relevant work. Some of the metrics that could be displayed from the model include ‘number of species, number of links, links per species, average level ... , maximum level, percentage of basal’ (Norling, 2007), which were also used for evaluating the original model and so would allow a level of comparison. In addition, the ability to pull a food web from the model at a given time step is important for the visual observation of the complexity of the food webs emerging, with complexity being defined as the ‘species richness, connectance, or number of trophic levels’ of a food web (Morin and Lawler, 1995). The precise details of the evaluation protocol to be used are discussed later in this section.

### *3.1.4 Adapt the mechanics of the model so it is more consistent with real-world food webs*

As mentioned in the literature survey, the purpose of this change is to make the model resemble the real world better with its interactions, without changing the fundamental principles of the original model. It should be acknowledged at this point that any model developed will always be a big abstraction from the real world, due to the incredible difficulty in simulating the massive complexity and spontaneity of ecosystems. Regardless, it seems reasonable to hypothesise that integrating more realistic underlying mechanics into the model will yield more realistic food webs. This makes it an interesting extension to explore and evaluate, though it is certainly not guaranteed to improve the performance of the model.

### *3.1.5 Introduce geographical features to the environment*

One of the main benefits of using a full spatial representation in an agent-based model is that it allows environmental and geographical features to be realistically modelled, and have agents interact with them according to their spatial location. Including geographical features provides the benefit of creating more niche spots in the environment that species could potentially evolve to fill – thus increasing the diversity of the resultant food web. It is debatable whether this will be effective using the same behaviour system as the original model, due to the highly abstracted underlying mechanics, but the exploration of the extension both on its own and in tandem with other extensions is worthwhile. Including geographical features has previously been used in R. Vahdati et al. (2019), which, as discussed in the literature survey, found that the inclusion of geographical features had a significant impact of the behaviours observed in the simulation.

### *3.1.6 Add heterogeneous behaviour to agents*

The purpose of this extension is clear: to make the model more similar to real-world ecosystems. The logic behind this is that by making the underlying mechanics of the model more realistic, the emergent behaviours should also be more realistic – thus more complex food webs should emerge. Norling (2007) justifies the validity of this extension by stating ‘it would seem reasonable that different species could and should have different parameters to represent a range of characteristics’. This extension would work particularly well in tandem with a spatial representation and geographical features, as all of these extensions have a similar purpose and should synergise well. This approach has been used before by Watkins et al. (2015), which gave each agent a defined set of characteristics and combined this with a spatially explicit agent-based model and a ‘heterogeneous landscape’ (similar in practice to geographical features). This implementation found ‘behaviour consistent with key characteristics of the dynamics of a natural jaguar population’ – meaning there is literature backed evidence that heterogeneous characteristics can help to create a realistic agent-based model.

One drawback of adding heterogeneous behaviour to agents is that it is unknown how changes in one characteristic affect all other characteristics of animals in the real world, thus there is no clear path of how to map these relationships onto an agent-based model. As an example, increasing the animal size may have knock on consequences to its speed, metabolism (energy consumption), hunting abilities, and many other factors. This means a lot of assumptions would need to be used when developing this model, which could potentially lead to less realistic results. This drawback and the fundamental re-working of the mechanics of the model required for this extension mean it will be a difficult and time-consuming extension to implement, meaning this has been set as a lower priority requirement.

### *3.1.7 Modify the mechanics to make it a closed system*

Construction of a closed system was a potential improvement to the model that was identified in Norling (2007). The reason this change may be useful is that it is referenced as being considered important by ecologists. The paper did some preliminary research into this area and found that there was not a great deal of improvement to the model when this was implemented. As such, this will be

considered as an optional research extension, and implemented if the extended models make sense to be implemented with a form of closed system and the time restrictions allow it.

### 3.2 Non-Functional Requirements

Criticality	Requirement
Must do	Needs to run in reasonable time on a standard computer
Must do	Have easy use controls to change certain model parameters
Should do	Have a visual difference between different species
Should do	Have geographical areas clearly visible in the spatial representation of the model
Could do	Using consistent start and end times for the model

**Table 3.2:** *Non-functional requirements of the project*

#### 3.2.1 *Needs to run in reasonable time on a standard computer*

Due to the limitations brought about by COVID-19, all the simulations will need to be run on my personal computer, which is not very powerful. In addition, for the future development of research, it makes sense to develop a model that can be run on most computers, not just powerful ones. This would make the products of this research accessible to more people. The model also needs to run in reasonable time – an amount of time that can be afforded exclusively for running the models from the tight project schedule.

#### 3.2.2 *Have easy use controls to change certain model parameters*

This requirement would allow parts of the developed model to be quickly tweaked from a dashboard style screen using simple controls. This is a simple but critical part of the project, as would allow someone who understands ecology, but has never used agent-based modelling before, to experiment with the effects of changing key parameters without needing to modify any of the code. This could really help with future research into the topic area, as many ecological experts will not be well versed in programming but will need to be able to modify the parameters. Some parameters that could be controlled in this way are the ecological efficiency value, the initial number of agents, the initial number of species, the amount of energy spent by each agent in each time step, and many more.

#### 3.2.3 *Have a visual difference between different species*

One output difference that could really help with visual evaluation of the model is to give a visual difference to different species in the ecosystem. This would likely come in the form of a different colour for each species. The benefits the visual evaluation of the model, as emergent trends of species would be far easier to identify in the environment, such as if one particular species was booming or another was going extinct. Likewise, this effect also has benefits for ecological research using one of the models developed, as patterns of behaviour can more easily be observed.



### *3.2.4 Have geographical areas clearly visible in the spatial representation of the model*

Having the geographical areas visible in the simulation would allow easier visual evaluation of the model, as it would show whether species were adapting to different niches in the ecosystem. This also allows greater parallels to be drawn between real-world population dynamics and the population dynamics of the model.

### *3.2.5 Using consistent start and end times for the model*

Using consistent start conditions and end times should help the model to show its performance with each update or improvement that is made, and help to standardise the model for testing with different model parameters and features. However, the design and implementation of the solution may prove to be more practical with adaptable start conditions and end times that can be changed to suit observations seen in the results of the model. Therefore, this requirement is marked as desirable, and will be implemented if it becomes practical and sensible to do so.

## **3.3 Model Evaluation**

As discussed in the literature survey, there are some innate difficulties related to the evaluation of the success of an agent-based model in the context of modelling the evolution of a food web. The key problem is the difficulty to obtain a meaningful dataset. A comparison dataset would enable comparison between the model findings and the real-world easier, while allowing the application of techniques like pattern-oriented modelling. Due to the constraints of this project, there will not be the capacity to obtain a dataset for the evaluation of the models developed, and there is no good external dataset available. However, further investigation provided alternative techniques that can be used to draw useful information from the model, which will allow the findings of the model to be compared to either known dynamics of real-world food webs or competing models. Some potential techniques are discussed in this section.

Norling (2007) identified several measures of interest that had been previously employed in the system dynamics model and were used to varying extents in the evaluation of the agent-based model developed. These include 'number of species, number of links, links per species, average level, maximum level, percentage of basal, intermediate and top species, percentage of links between different species types, and ratio of prey to predators'. Using these parameters as part of the evaluation protocol will allow strong points of comparison between the original model and the models developed in this project, and can thus show how well the model is performing.

In addition, the model performance can be evaluated simply on a qualitative view of what the system is achieving and the results it is giving. This approach was taken in Perez and Dragicevic (2010), where the results were analysed and hypotheses for the causation for the phenomena observed were created. For the purposes of this project, this technique can be applied very effectively by drawing up a food web of all the species present at a given time step. This will immediately allow the observation of both how complex and how deep the web is. A deeper and more complex web demonstrates an improvement

over the original model, which could only sustain approximately two levels deep. In addition, the metrics and what is visibly occurring in the environment can be analysed and compared to the known population dynamics and food web patterns that were identified in the literature survey.

One issue when evaluating in this manner is the fact that there will always be some degree of inaccuracy when it comes to simulating something as complicated as a food web. This will make it hard to draw comparisons between the food webs created by the model and real-world food webs, regardless of how closely the model is attempting follow all the factors at play. This is ultimately a limitation of the practice of simulation as a whole – the real world is so complex that nobody could hope to model every aspect at play, instead modelling some of the key factors should be attempted to see how they lead to emergent food webs and population dynamics. Along the same lines, agents will not directly map onto species, so comparison to the real world on a micro-level will not be possible. Again, the goal of these simulations is to capture the emergent dynamics of the system, rather than the individual dynamics of animals.

Ultimately, the evaluation protocol for the agent-based models produced will primarily use a combination of qualitative analysis approaches, with quantitative analysis being used to compare certain model metrics. The effectiveness of each model will be gauged by comparison with the original model and the other models produced in this project. This will be assessed primarily in regard to the food web patterns and population dynamics that are occurring in each simulation, and how these relate to real-world food webs. These will be found using a combination of the approaches that have been discussed in this section. Once these have been observed, the findings will be compared across the different models to assess how realistically each model is simulating the dynamics of a real-world food web. One qualitative measure that will be used to assess the performance of each model is the average number of species it sustains. This can be used to compare the performance of each successive model.

### 3.4 Ethical and Legal Considerations

This project does not have any ethical concerns, as it does not involve any human participants. Any knowledge, insight or information that has been obtained from an external source for this project has been appropriately cited, and any code that was reused in the development of the models for this project has been properly cited and attributed to the producers of said code.

### 3.5 Summary

The requirements and analysis section of this report sets out the necessary requirements for the success of the project. Entering the design phase of the project, some clear objectives and requirements have been specified. The first stage will be to re-implement the model from Norling (2007) using NetLogo. The subsequent stages will apply the various extensions discussed will then be applied, considering the defined priority of each requirement. To evaluate the performance of the model at each stage of development, the evaluation protocol defined in the subsection above will be followed, using a combination of qualitative analysis, and some quantitative analysis, approaches. Fulfilling the objectives laid out in this section aims to allow the simulation of an ecosystem that displays emergent complex food webs and also provides a baseline model for future research into the evolution of food webs.

## 4 Re-Implementation Model

The first step of this project is to re-implement the original model as closely as possible, while incorporating some changes to bring the model up to the current standard of agent-based modelling. This effort will provide a baseline model that can be further built upon and improved in the proceeding models. This section specifies the design requirements for this model based on the original model, the way in which the model has been implemented, and the results were observed.

### 4.1 Design

#### 4.1.1 *The design of the original model*

The model implemented in the original work first defines the method used for calculating the outcome of interactions between agents. For this, a 500x500 matrix of possible features is created that defines all of the possible characteristics for species in the model, 500 in total, and gives each feature a score against each other feature – ‘the value  $M_{xy}$  at any position in the matrix describes how useful feature  $x$  is against feature  $y$ ’. These features are abstractions of species characteristics that are not assigned a real-world context, and are simply numerical values used for calculating which species can consume which other species. For the sake of explanation, they could be considered to be features such as ‘sharp teeth’ or ‘fast runner’. Each species is then given a feature vector of a random sample of ten of these features to define the characteristics for that species.

Before any agent interaction calculations in each time step, the world’s resources need to be distributed out to agents. To calculate how these resources are distributed, the world is assigned a set of features in the same way agent species are. Each agent in the model then has its score calculated against the world, with this value being proportional to the fraction of the world’s resources which that agent receives. This score is calculated in the same way as agent-agent interactions, which is discussed below. Any agents that score zero against the world receive none of its resources, and consequently have to survive by consuming other agents.

Agent interactions are now calculated. Whenever two agents interact in the model, the outcome of the interaction is calculated by first calculating the score value of species  $i$  against species  $j$  with the formula:

$$S_{ij} = \max \left\{ 0, \frac{1}{L} \sum_{\alpha \in i} \sum_{\beta \in j} m_{\alpha\beta} \right\}$$

**Figure 4.1:** *Calculation of the score of agent  $i$  against agent  $j$  (image sourced from Norling, 2007)*

Where  $i$  and  $j$  represent the feature vectors for each of the species, and  $L$  represents the number of features in the feature vectors, in this case 10. This formula finds the scores of each value in the first species’ feature vector and each value in the second species’ feature vector, 100 scores in total, and sums these. This score is then divided by  $L$ , and set to zero if it is a negative number. To determine the

outcome of the interaction, the score of the ‘predator’ species against the ‘prey’ species is taken away from the highest value that any existing species achieves against the ‘prey’ species. If this value is less than the competition parameter, the ‘predator’ species can eat the ‘prey’ species. The competition parameter is a constant used to determine how the resources of the environment are distributed to the different species. This operation is visualised below.

$$S_j^M - S_{i,j} < \delta$$

**Figure 4.2:** Operation to determine if species  $i$  can eat species  $j$  (image sourced from Norling, 2007)

The way in which this model determines which agents interact is through random sampling. At each time step each agent is randomly assigned another agent or the world upon which it will attempt to feed. If the calculations presented above dictate that a given agent can feed on that agent, the ‘prey’ agent dies and is removed from the simulation. Furthermore, the ‘predator’ agent gains the resources currently possessed by the ‘prey’ agent multiplied by the ecological efficiency, a constant between zero and one that reflects the inefficient transfer of energy up the food web. If, after these interactions, an agent has a resource total greater than one, it reproduces with the offspring getting the amount of energy the parent had above one and the agent has its energy reduced down to one. After a fixed number of time steps, or when the population is stable, a mutation occurs which means one of the existing agents has one of its features changed to a different feature, creating a new species. This cycle of agent death, birth, and mutation forms the basis for the simulation.

Additional details given about the model:

- The feature matrix is initiated randomly, but is anti-symmetric: feature  $M_{xy} = -M_{yx}$
- Each feature in the matrix has a score of zero against itself
- If the total population ever becomes zero, including at the very start of the simulation, a new agent is created of a random species
- The resources belonging to the world are added at a constant rate and are distributed to agents in proportion to their score against the world. Agents with a score of zero against the world get no resources from it
- The ecological efficiency used for the model is 0.1
- The competition parameter should be set to  $1 / \text{the number of features each species has}$ , in this case it is  $1/10$

#### 4.1.2 The design of the re-implemented model

This initial model is intended to replicate the performance of the model produced in the original work, though incorporating some changes that bring the model more up to speed with the current advances in agent-based modelling. As such, most of the core underlying mechanics of the new model are the same as the original model. These include the way features are produced and designated to species, using the random sampling of the feature matrix; the ways in which the results of interactions are calculated; the way that agents reproduce; and the way the world’s resources are distributed to the agents.

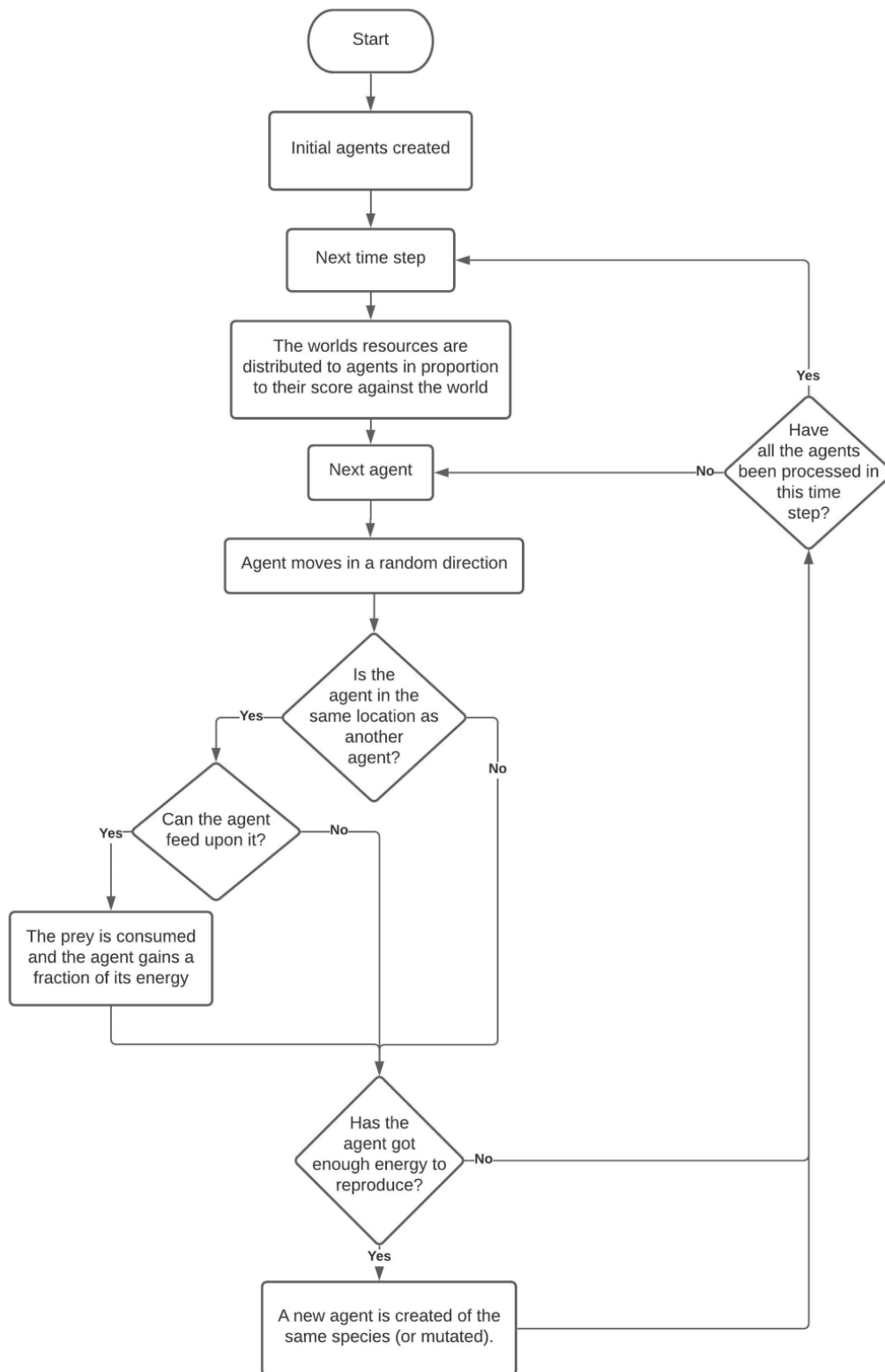
The major change to be included in the new model is the use of a spatial representation. NetLogo is a development framework that is centred around a spatial representation of models, making it more logical to develop all the models using this approach. This decision is justified by the conclusions of the literature survey, which identified the inclusion of a spatial representation as important to allow further extensions to be built upon it. In addition, it found that a spatial representation alone may not have a big impact on the mechanics of the model, which is useful for allowing comparisons between the re-implemented model and the original. For these reasons, the model to be developed here will incorporate a spatial representation rather than random sampling, as was used in the original work. This means that agents won't interact on a purely artificial basis – automatically interacting with one other random agent in each time step. Rather, the interactions will organically emerge from the random movement of the agents through the environment, and agents will only interact when they are at the same geographic location in the same time-step. When agents interact, the score calculations will be performed with each agent playing the 'predator' role once, as the agents do not have pre-defined interaction roles as they would with random sampling.

Another change that will be made is that instead of mutations occurring at the end of a 'evolutionary time step', they will occur on a random basis. This change is due to the ambiguity in the paper – it is not made clear exactly when an 'evolutionary time step' ends, simply saying 'when a fixed number of time steps has passed' or the 'current population is stable'. To clear up this ambiguity, and to make the model more realistic, random mutations will instead be used. This will mean that whenever an agent reproduces, the offspring will have a given chance of mutating. These mutations occur in the same way as the original model, with one of the species features being swapped out.

Overall, there are some differences between the design of the original model and the re-implemented model, but the models remain easily comparable. Any differences in the results obtained from each model will help to highlight the impact of using a spatially explicit model and random mutations, which will be important information for future development in this subject.

#### *4.1.3 Flowchart of the re-implemented model*

The following flowchart and all other model flowcharts were created using Lucidchart:

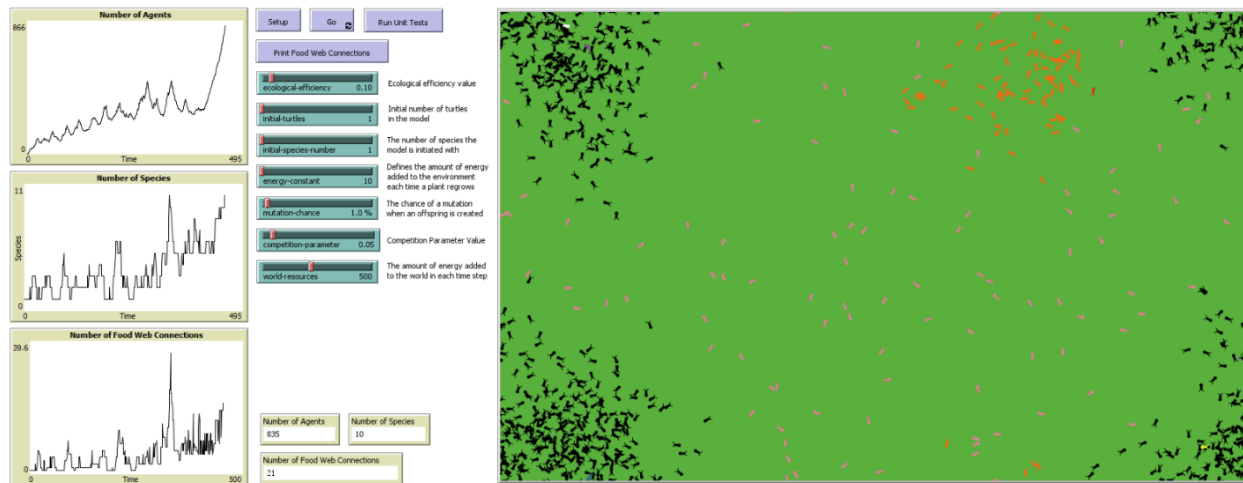


**Figure 4.3:** Flowchart of the mechanics of the re-implementation model

## 4.2 Implementation

### 4.2.1 The model implemented

The design specified above was implemented using NetLogo, as selected in the literature survey as the most appropriate development framework for this context. The interface of the implemented model is shown in Figure 4.4.



**Figure 4.4:** The interface of the re-implementation model

The visualisation and control of the model implemented is shown above, with the simulation in action displayed on the right-hand side of the image. The agents are represented by coloured bug icons overlaid on the world, with each colour depicting a particular species of agents.

The left side of the interface is used for displaying three graphs:

- The number of agents that exist in the model over time
- The number of species that exist in the model over time
- The number of food web connections over time (predator-prey relationships between species)

These are useful visual representations of the activity of the model which can assist with model evaluation. In addition, these graphs are supplemented by corresponding metrics of these values at the current time step.

The interface also includes slider controls to allow some of the model's parameters to be easily modified. The controllable parameters include the ecological efficiency, the initial number of agents, the number of initial species, the mutation chance, the competition parameter, and the number of resources the world distributes in each time step. Buttons are included for setting up and running the model, and also for running the unit tests and printing out an array of the food web connections, which is a prerequisite for part of the model evaluation.

#### *4.2.2 Problems encountered and changes made*

The re-implementation of the original model represented a significant challenge at the outset as the source code of the model was not available and the model needed to be developed in a new framework and language. The model built was based upon the blueprint that was specified in the paper of Norling (2007).

There are a few nuances of the model that were discovered during development that were not discussed in the original work, so appropriate assumptions have been made. The first of these is whether agents should be able to consume agents of their own species. As this was not discussed in the original work, 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 and future models. This essentially means each 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.

One problem that was found with this model, which may also have afflicted the original model, is that if the model is initiated with a single agent that has a score of zero against the world, it will not be able to get any energy from the environment so it can't reproduce, and it can't die as there are no other agents to eat it. This creates an infinite loop of nothing happening in the simulation. This issue could be resolved by implementing a constant 'decay' rate of an agent's energy – a constant value that is subtracted from an agent's energy after each time step. However, the original work did not specify that any solution like this was implemented. Therefore, the results from this model will be produced only from simulations that do not run into this issue, and the issue will be resolved in the extended models.

#### *4.2.3 Evaluation methods*

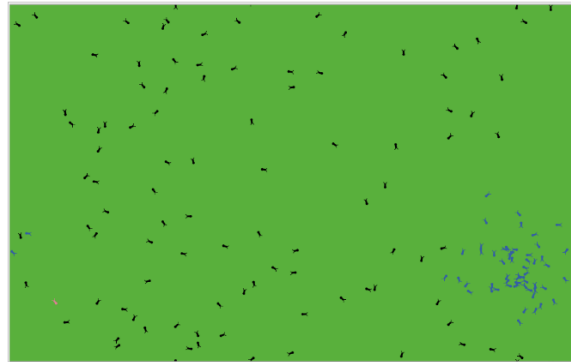
As was decided upon in the analysis stage, the evaluation of the performance of this model will be determined by how well the findings of the model correspond to real-world population dynamics and food webs. The analysis of the findings of this model will be primarily a qualitative view of the results produced. The metrics and techniques that have been used to supplement this analysis are described below.

The first technique is graphically representing the food web that exists in the model at a given time step. Functionality has been included in the interface to print out a list of the relationships between species of agents, and also the world. This information can then be transferred into a graph using GraphViz to display a graph of the food web. This tool is very useful for comparing the diversity and number of connections in the artificial food web to real world food webs. In addition, food web diagrams also formed part of the results of the original work, so this technique should allow effective comparison between the original and re-implemented models.

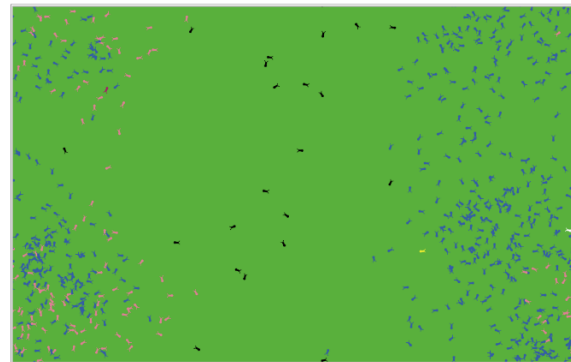


A further important technique for effective qualitative evaluation of the model is observing the changes in the population over time. This can be seen in the in the environment representation on the interface, as the simulation is run. Observation of how species are interacting is simplified by giving each species a unique colour, so evidence of which species are booming or struggling to survive is immediately apparent. Likewise, this can reveal how effective a new species is when introduced and shows the visible changes to the layouts of the population when new highly competitive species are introduced.

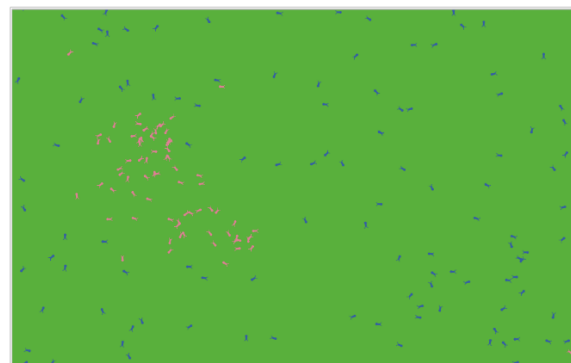
For example, the Figure 4.5 show the introduction of a highly competitive species, shown in blue, that overthrows the previously dominant species, shown in black.



Time step 3712



Time step 3973



Time step 4281

**Figure 4.5:** *The effect a new competitive species can have when introduced in the model (different agent colours correspond to their species)*

There are also several metrics that can be used to assist the analysis of the model. One of the most important of these is the number of agents that exist in the model, which can be graphed over time. This provides insight into the underlying dynamics of the population, such as the introduction of new species leading to population booms or extinctions. Likewise, the number of species and the number of food web connections present in the model indicate how diverse the food webs in the model are, as well as how effectively these are sustained over time.

#### 4.2.4 Model execution speed

This model represents a primary endeavour into the field of research with a simple agent-based model. As such, this model executes relatively quickly, even on limited hardware. This allows the model to be effectively evaluated on its findings of both quite short and very long runs. However, the performance becomes somewhat limited when too many agents exist in the environment, so simulations of the scale used by Norling will not be possible here, though the initial results indicate that this should not be much of an issue.

#### 4.2.5 Testing

Testing the complete functionality of this model is a difficult task, due to the natural stochasticity of many of its functions. Due to this, the model has been tested using a checklist of user acceptance tests to confirm it is working as expected.

Test	Pass/Fail
The simulation can be correctly setup	Pass
The simulation can be started	Pass
The simulation can be stopped	Pass
The number of agents is graphed on the interface	Pass
The number of species is displayed and graphed on the interface	Pass
The number of food web connections is displayed and graphed on the interface	Pass
The food web connections can be printed out	Pass
Model parameters can be changed, and these effect the behaviour of the model	Pass
When the model is run, the agents can be observed interacting with each other	Pass
The model can be reset at any time	Pass

**Table 4.1:** User acceptance tests

As only the first model has been implemented at this stage, not all of the functional and non-functional requirements have been met. Those that have been met are described below.

#### *Re-implement the work of Norling (2007)*

This model closely follows the simulation protocol laid out by Norling (2007), and made assumptions where necessary. This produced a model that successfully re-implemented the work of Norling.

*Include a spatial representation of the agents and environment in the model*

This model has the environment visible on the interface, with the agents and their movement through the environment being visually represented.

*Output graphical and numerical metrics to allow effective evaluation of the model*

The number of agents, number of species, and number of food web connections are all graphically plotted on the interface, so any patterns over time can be observed. In addition, the current values of all of these are also displayed to give an indication of the current state of the simulation.

*Have easy use controls to change certain model parameters*

The interface of the model includes sliders that have the ability to control the ecological efficiency, the initial number of agents, the number of initial species, the mutation chance, the competition parameter, and the number of resources the world distributes in each time step.

*Have a visual difference between different species*

Each agent is set to a particular colour depending on its species, which allows visual observation of how species are interacting in the environment.

#### 4.2.6 Summary

This section has addressed how the designed model has been implemented in NetLogo and the changes that were made to the model during development. In addition, the model has been tested, and meets several of the key requirements that were laid out earlier in this project. This means that this model has been successfully implemented, and can be used to address to successes and pitfalls of this model design.

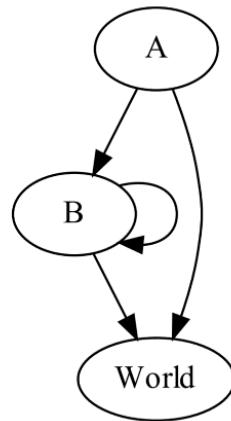
### 4.3 Results and Discussion

#### 4.3.1 Recap of model evaluation and results analysis

As has already been discussed, the main challenge with analysing the results of the model is the lack of a real-world dataset for comparison. To overcome this problem, several evaluation techniques have been identified that can allow comparison between the results of the model and known population and food web dynamics, as well as comparisons between alternative models. These techniques are deployed in this chapter to analyse the results of the re-implementation model in regard to its correspondence with both the original model of Norling (2007) and the evolution of real-world food webs.

### 4.3.2 Food webs produced by the model

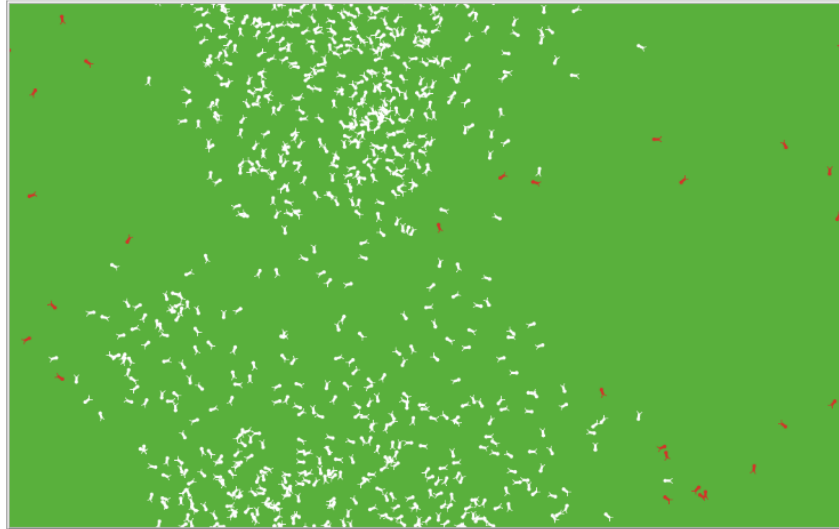
The food webs produced by the re-implementation model, such as Figure 4.6, consistently find that species which solely rely on predation struggle to survive, whereas species that consume both other agents and the world have a much better rate of survival. This leads to food webs that have a somewhat 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. Figure 4.6 shows significantly less diversity than the real-world food web example Figure 2.1, which highlights the unrealistic nature of the results produced by this model.



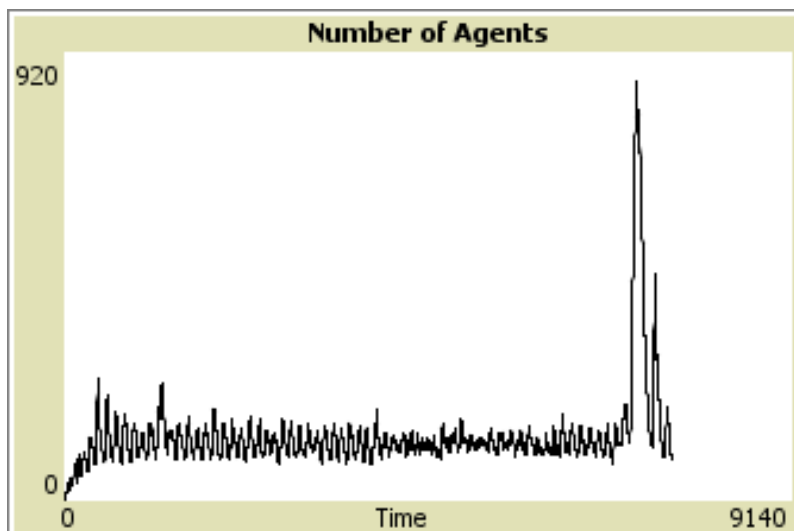
**Figure 4.6:** A food web produced by the re-implementation model

### 4.3.3 Population dynamics observed in the model

Figure 4.6 shows the food web at a time step where species A has been recently introduced into the environment and is the apex predator: no species can feed upon it. In addition, species A can feed on both the world and species B, whereas species B can only feed on the world and itself. These relationships lead to the explosion of the population of species A, denoted as the white agents in Figure 4.7. This is followed quickly by the extinction of species B due to their over predation. This course of events can be observed in Figure 4.8, the number of agents graph, which spikes as the population of species A rapidly grows, then falls back to the standard level once species B is extinct, as there is now less food for species A to consume. Patterns like this have been observed over many runs and show the level of correspondence the re-implemented model has to the original. In addition, these patterns parallel some of the real-world population dynamics that were identified in the literature survey, such as the introduction of new species leading to the extinction of certain existing species. This proves that the model has at least some degree of realism to the results it is producing at this early stage.



**Figure 4.7:** *One species rapidly reproducing in the model*



**Figure 4.8:** *The number of agents existing in the re-implementation model over time*

The pattern that the model is following appear to be closely related to the well-known population dynamics ‘the competitive exclusion principle’ (Population Dynamics | Encyclopedia.com, 2021). This is when two species that occupy the same niche in the environment ‘one will inevitably drive the other to extinction’. This is an interesting dynamic to observe and accentuates one of the key problems with this model – there are no real niche spots in the environment for the species to adapt to fill, meaning they will always fill the same niche.

#### 4.3.4 Comparison with other models

This re-implementation 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 re-implementation 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 didn't 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 didn't 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 a quantifiable point of comparison, the number of species over runs of this model has been calculated. Each run was stopped after 10,000 time-steps has passed.

Run	1	2	3	4	5	Average
Mean Species	1.693	1.795	1.554	1.644	1.427	1.623

**Table 4.2:** *The average number of species in runs of the re-implementation model*

As can be seen in Table 4.2, 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.

#### 4.4 Findings and Next Steps

Overall, there are some promising results that are produced by this model in terms of the population dynamics observed and the correspondence it has to the original model. However, the model struggles to create any significant or sustained diversity in its food webs, and there are also many population and food web dynamics that the model fails to produce, such as species adapting to fill niche spots in the environment.

The similarity in findings between this model and the original does show that the models in their standard configuration has the limitation that they cannot sustain complex food webs over any significant period of time. In addition, the occurrence of the competitive exclusion principle in the model shows that the environment suffers from a lack of niche spots that the species can adapt to fill. The next step in the project is to implement a model that addresses these problems and build on top of the groundwork set out by this re-implementation with improvements that were identified in the requirements and analysis.

## 5 Extended Model

Having produced a baseline model for this project and identified its strengths and weaknesses, a model can now be produced to extend upon this with some of the improvements identified in the requirements and analysis stage. This extended model will use many of the same underlying mechanics as the re-implementation but will attempt to solve some of the issues that were identified in the re-implementation and add additional features seeking to improve the model results.

### 5.1 Design

The first change to the underlying mechanics of this extended model is the introduction of 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. Including a term like this makes logical sense in terms of the energy cost of the activity of agents – something that plainly exists for real world animals. In addition to being a realistic inclusion, this term helps to fight against the overpopulation of the environment, which will speed up the runtime of the model and consequently allow more experimentation with it.

The extended model will also use a more realistic method for distributing the energy from the world than the previous mechanism of automatically distributing energy to the species that score well against the world. This method is to add a plant species that is set in place in the environment and that can be interacted with by agents in the same way they interact with each other. Agent species that score well against the plant species can consume the plants and gain a portion of their energy. The plants are then being removed from that place in the environment for a refractory period whilst they 'regrow'. This method of energy distribution helps to model the impact of scarcity of resources more insightfully, particularly when used in tandem with the energy decay of agents. These act as a natural limit on the number of agents that can exist, in a similar way to how real environments are naturally regulated. This also has the benefit of solving the issue identified in the re-implementation model where the single initial agent could score zero against the world, leaving the program in an infinite loop. Instead, an agent that scores poorly against the plants will still be able to consume them in this model, unless a species that is significantly better adapted exists or is created.

Finally, distinct geographic regions will be included in the environment. This has the benefit of 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. 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. The different environments will have visual differences to assist with the qualitative evaluation of the model, as set out in the non-functional requirements.

## 5.1.1 Flowchart of the extended model

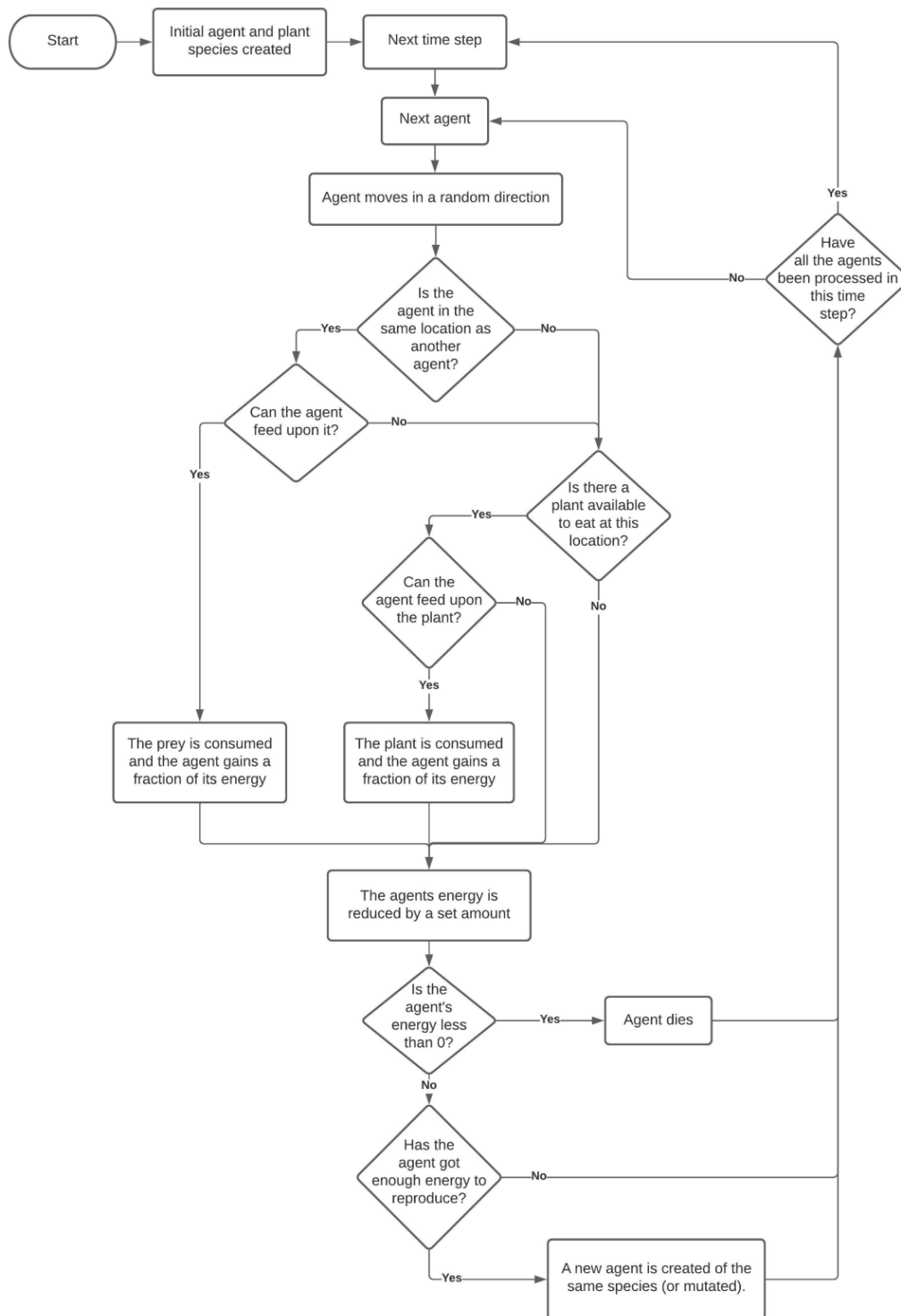


Figure 5.1: Flowchart of the mechanics of the extended model



## 5.2 Implementation

### 5.2.1 The Model Implemented

As is seen in Figure 5.2, the extended model has been successfully implemented in NetLogo following the design protocol laid out in the previous section. The only significant difference in the interface between the extended model and the re-implementation is the simulation environment, which has been segmented into different coloured regions. While these regions are entirely abstract, they can be considered to be mountains, dense forest, field and lake, for the grey, dark green, light green, and blue areas respectively.

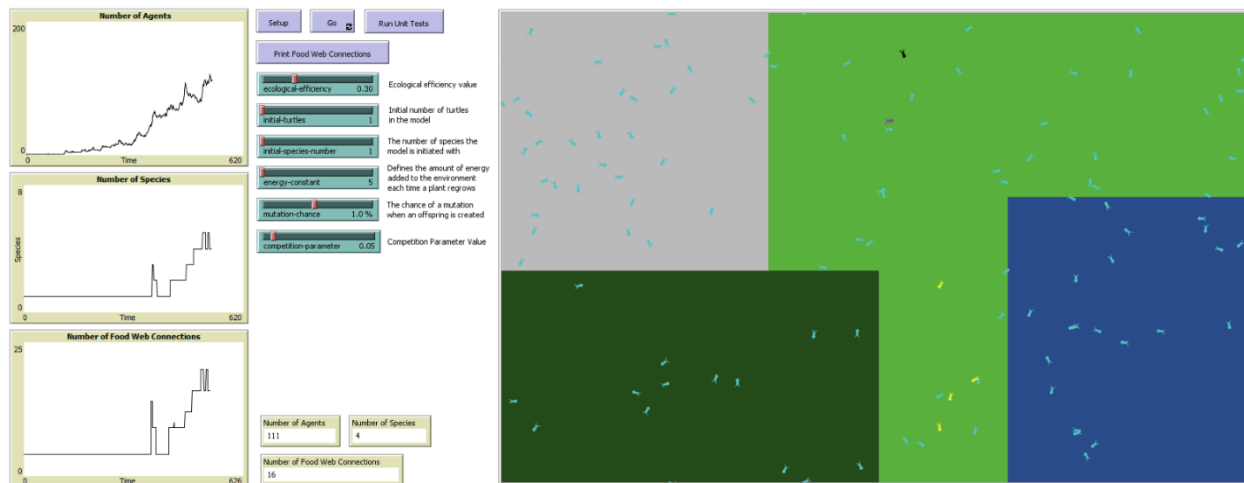


Figure 5.2: The interface of the extended model

### 5.2.2 Problems encountered and changes made

One recurring issue during implementation was that mutated species usually had very similar properties to those of their parent species, which lead to the capabilities of each species being very similar, causing poor diversity in the resulting food web. This was due to the fact that only a single feature was changed during a mutation. To resolve this, the mutation mechanic as changed so that five of the ten features would be modified during a mutation rather than just one.

### 5.2.3 Evaluation methods

This model followed the same evaluation protocol as the re-implemented model, using food webs, graphical data and qualitative evaluation to draw conclusions about the performance of the model. The results produced by this model will be compared and contrasted to those found with the re-implementation model to evaluate the overall success of this model in its goal of producing a realistically structured food web.

### 5.2.4 Model execution speed

This model operates on most of the same basic mechanics as the re-implemented model, meaning it can run in the same order of time as it. This allows effective experimentation with running the model over longer time periods if necessary.

### 5.2.5 Testing

The same user acceptance tests that were used for the re-implemented model were also run with this model and all passed. In addition, the following user acceptance tests were added to the testing protocol of this model.

Test	Pass/Fail
The environment has distinct geographical regions	Pass
The species show some degree of adaptation the certain regions	Pass
The species travel between geographical regions	Pass
The adapted species tend to cluster in the region benefiting them most	Pass
Species with different adaptations develop mostly separate from each other in their region of specialty	Pass

**Table 5.1:** *User acceptance tests*

Similarly, the same functional and non-functional requirements that were met by the re-implemented model have also been met with this extended model. In addition, the following requirements have also been met.

#### *Adapt the mechanics of the model so it is more consistent with real-world food webs*

An energy decay rate ‘metabolism’ has been added to the agents in this model, which better replicates the way in which real-world animals have to expend energy to survive than the previous model. In addition, the use of geographic regions better simulates the diverse habitats that shape some of the fundamental mechanics of real-world food webs. Finally, the introduction of plant-eating and regrowth more accurately simulates the distribution of resources in an environment, as all the energy in real-world food webs originates from basal species, which are some form of plant.

#### *Introduce geographical features to the environment*

In this model different geographic regions are represented both visually and in the underlying mechanics, as each region contains a unique species of plant. This is a highly beneficial extension to the model as it allows species to adapt to eating certain plants, thus encouraging greater speciation and diversity.

#### *Have geographical areas clearly visible in the spatial representation of the model*

The geographical regions that have been included in this model have been clearly displayed on the environment interface in the form of different coloured sections. This allows effective observation of the effect these regions have on the species that appear in the environment, such as species adapting better to a particular habitat.

### 5.2.6 Summary

This section has addressed how the extended model has been implemented in NetLogo and the changes that were made to the model during development. In addition, the model has been tested, and meets additional requirements to the re-implementation model. The extended model has therefore been successfully implemented and can be evaluated regarding the results it produces.

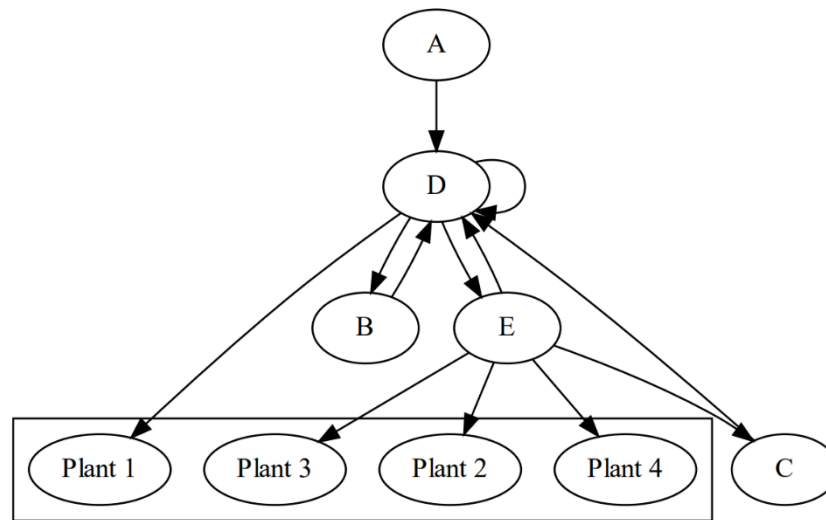
## 5.3 Results and Discussion

### 5.3.1 *Recap of model evaluation and results analysis*

The findings of this model will be analysed by the same protocol followed in the re-implementation model. If the model is performing as intended, the food webs produced should have greater complexity, in particular containing more species and trophic levels. In addition, there should be visible population dynamics occurring, such as species adapting to certain environments and population shifts occurring within these environments.

### 5.3.2 *Food webs produced by the model*

Figure 5.3 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 re-implemented model, sustaining five species on top of the four plant species. This immediately addresses one of the primary issues with the re-implemented 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.

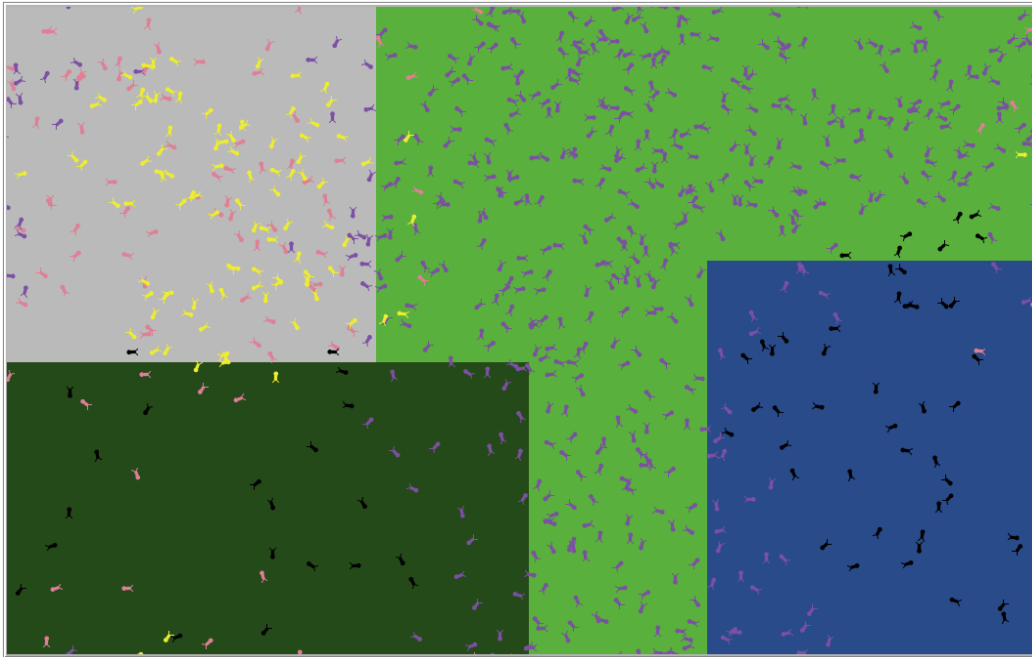


**Figure 5.3:** *A food web produced by the extended model*

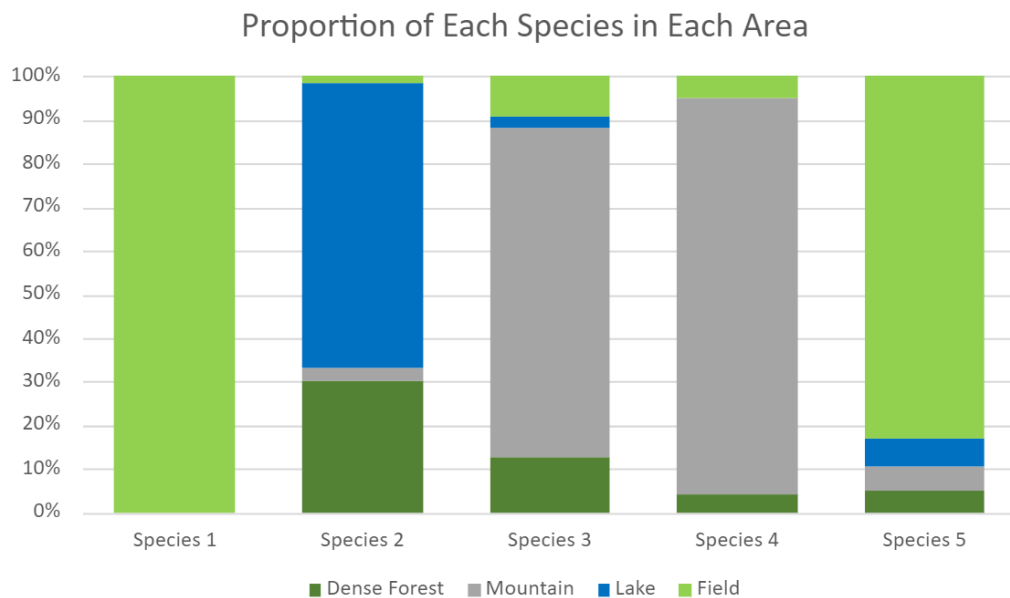
However, there are some patterns in this food web that do not accurately line up with what might be expected in a real-world food web. For example, Figure 5.3 shows a ‘cycle’ of connections where species D can eat E, E can eat C, and C can eat D. Structures like this are highly uncommon in real world food webs, and logically do not make sense. As was found in the literature survey, energy is lost in the transfer for one trophic level to the next, which means that a cycle in a food web should, in all likelihood, 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 Figure 5.3. 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.

### 5.3.3 Population dynamics observed in the model

Figure 5.4 shows the state the model was in when the snapshot of this food web was taken. It is visible that certain species are more common in certain areas, such as the purple species being primarily in the field area or the yellow species being primarily in the mountain area. This is strong evidence for individual species being best adapted to one or more of the environments, and this shows that the intention of including geographic regions is being fulfilled. This links back to the real-world population dynamics of ecosystems, where species often evolve to fill a niche in the food web, leading to greater speciation and diversity of species. This change is therefore likely to be one of the primary reasons for the more complex food web observed in Figure 5.3. This effect is also visualised in a more quantitative manner in Figure 5.5, which represents the proportion of each species in each geographic region in the form of a bar chart. This further reconciles the qualitative judgement about species being better adapted to certain areas with data that has been extracted directly from the model. Once again, the area names used in Figure 5.5 (such as ‘Dense Forest’) are conceptualisations – they are included for the benefit of visualisation and do not have any true meaning.



**Figure 5.4:** Adaptation of species to geographic locations (the colour of each agent denotes its species)

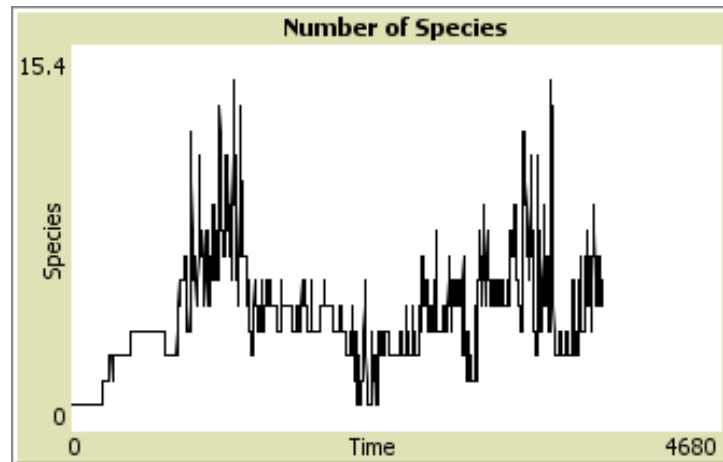


**Figure 5.5:** The proportion of each species that persists in each environment

An interesting observation from Figure 5.4 is that both the yellow species and the pink species appear to be well adapted to the mountain area. This fact links back to the known population dynamic identified in the literature survey where one species adapting well to a niche spot in the environment can help additional species to survive within it.

### 5.3.4 Comparison with other models

In contrast to the re-implemented model, the extended model is capable of sustaining a reasonably complex food web over a significant period of time, as can be seen in Figure 5.6. In particular, Figure 5.6 shows that between approximately iterations 1000 to 2000, a food web is sustained that almost always contains a minimum of four species, with the average number of species being closer to five or six. This once again provides empirical evidence for the success of the extended model, as it sustains greater complexity in the food webs than can be found with the re-implemented model, such as Figure 4.6.



**Figure 5.6:** *The number of species found in the extended model over time*

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 re-implementation 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. This is in contrast to the re-implementation model and the original model, both of which did not encourage species diversity.

The number of species over runs of this extended model has been calculated in the same manner as the re-implementation model, with each run being stopped after 500 time-steps – as it runs more slowly than the re-implementation model.

Run	1	2	3	4	5	Average
Mean Species	5.691	7.365	4.397	4.084	5.093	5.326

**Table 5.2:** *The average number of species in runs of the extended model*

The average number of species that is sustained in the model is 5.326. This shows significant improvement over the re-implementation model average of 1.623 species, quantitatively proving that the improvements introduced in this model successfully increased the diversity of the emergent food webs.

## 5.4 Findings and Next Steps

The overall performance of this model showed some significant improvements over the re-implemented model. The first major evidence of this was the more complex and hierarchical structure of the food webs produced by the model. In addition, these food webs are sustained over significant periods of time, which was one of the biggest issues of the re-implemented model, which was unable to sustain any complex food web over a significant period of time. The benefits of adding geographical features to this model can also be seen, with species visibly more common in certain areas – showing they have found a niche in the environment. Likewise, the population dynamics observed in the model effectively parallel some known population dynamics of real-world food webs, such as species adapting to certain environments allowing other species being able to survive in that environment. This model shows the true value in including a spatial representation in the model, because the inclusion of geographic areas would not have been possible without one.

Despite all of these promising results, there are still some issues that could be resolved to further enhance the model. One such problem is the nature of the food webs that are produced by the model – and how realistically structured they are. The extended model pushed the boundaries of the mechanics used in the original model, so in order to develop a model that can further improve upon these results, a more fundamental reworking of the model is necessary. The original plan in this project was to culminate the research with the extended model, however the results produced by it did not fulfil everything the project aimed to achieve at the outset. Therefore, a model will next be developed that rethinks the model mechanics by including heterogeneous behaviour for agents, and addresses the problems with the extended model that have been identified through the findings.

## 6 Heterogeneous Behaviour Model

Norling (2007) discussed the potential of creating a model that integrated heterogeneous behaviour, as this provides a more realistic method for simulating the behaviour of species and the way in which they evolve. As was discussed in the literature survey, 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.

### 6.1 Design

#### 6.1.1 *Agent characteristics*

For the purposes of this model, each agent has four features which are capable of mutation: predatory ability, plant eating ability, defence, and speed. Agents also have a fifth characteristic, metabolism, though this doesn't mutate or evolve directly and is instead determined by the calculation below.

$$Metabolism = PredatoryAbility + PlantEatingAbility + Defence + \frac{Speed}{5}$$

**Equation 6.1:** *Metabolism calculation for a given agent*

Having 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 reside. These are used to define the geographic regions in the environment, which are visually represented in this model the same as the previous models. These characteristics are plant defence and plant energy.

The predatory ability and defence characteristics will be used for calculating the result of agent-agent interactions, and the plant eating ability of an agent and the plant defence of a plant will be used for calculating the outcome of agent-environment interactions. Conversely, the speed and metabolism metrics determine agent behaviours rather than interaction rules, with speed defining how quickly an agent can traverse the environment, and metabolism defining the rate at which an agent uses its energy stores.

#### 6.1.2 *Interaction rules*

In the same way as 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 rule; if the predatory ability of the predator agent multiplied by a random floating-point number between zero and two is greater than the defence of the prey agent, then the predator



can eat the prey. Otherwise, it cannot. A similar operation determines whether an agent can eat a given plant. These are visualised in Equation 6.2 and 6.3 respectively.

$$x = \begin{cases} \text{PredatoryAbility} * \text{random}(0, 2) > \text{Defence} & x = \text{Eat} \\ \text{Otherwise,} & x = \text{Ignore} \end{cases}$$

**Equation 6.2:** Operation to determine whether a given agent can eat another agent

$$x = \begin{cases} \text{PlantEatingAbility} * \text{random}(0, 2) > \text{PlantDefence} & x = \text{Eat} \\ \text{Otherwise,} & x = \text{Ignore} \end{cases}$$

**Equation 6.3:** Operation to determine whether a given agent can eat a given plant

The random values applied in these equations are designed to add a level of stochasticity to interactions, which should aid the evolution process. In addition to these rules, 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. These rules were intended to give the interaction rules more correspondence to the real-world, as in the majority of real-world food webs there is no cannibalism and higher trophic level predators cannot be eaten by lower trophic level predators. These additional restraints on the agent interactions should therefore allow a more realistically structured food web to emerge from the model.

### 6.1.3 Mutation

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.

## 6.1.4 Flowchart of the heterogeneous behaviour model:

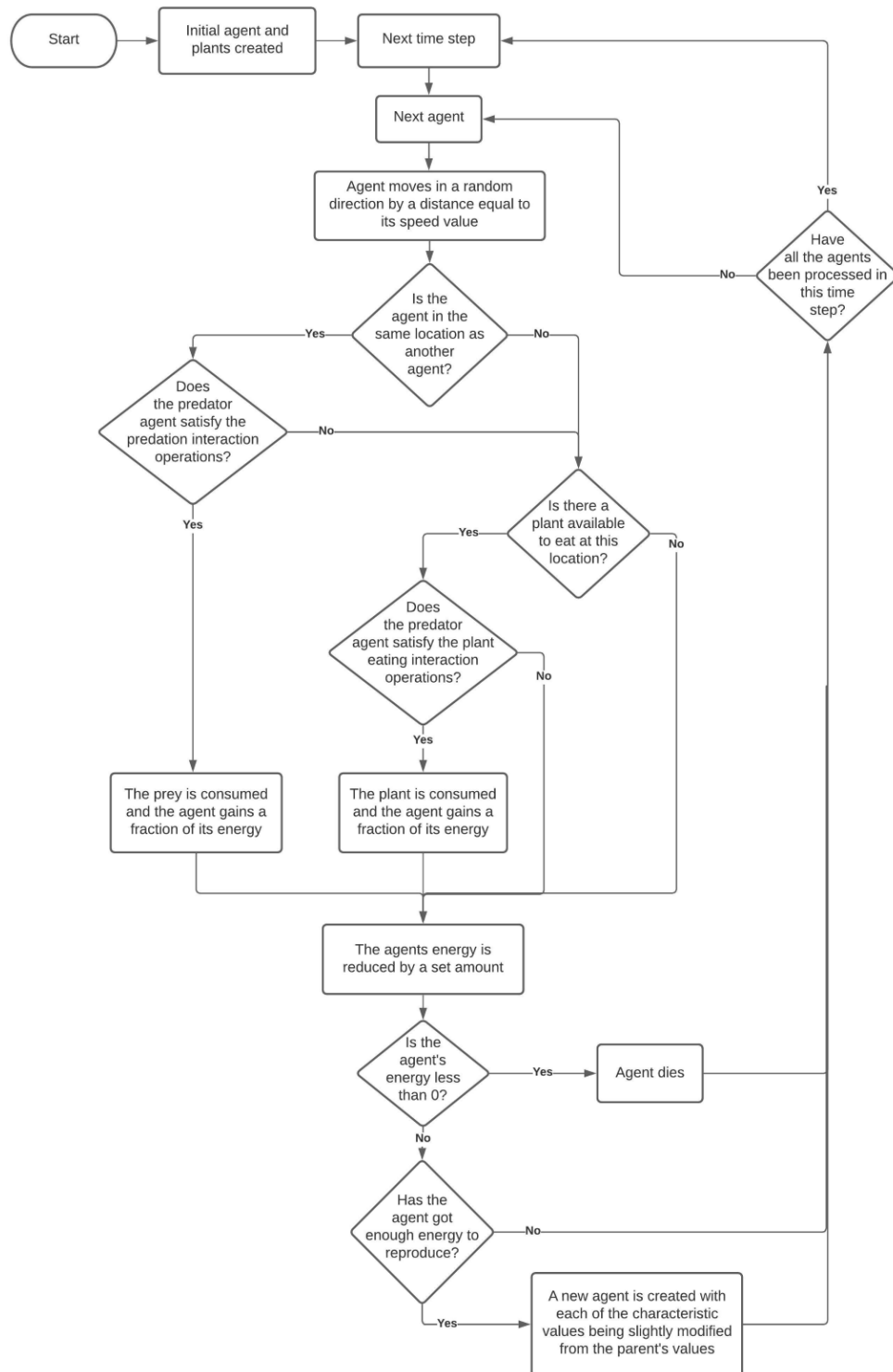
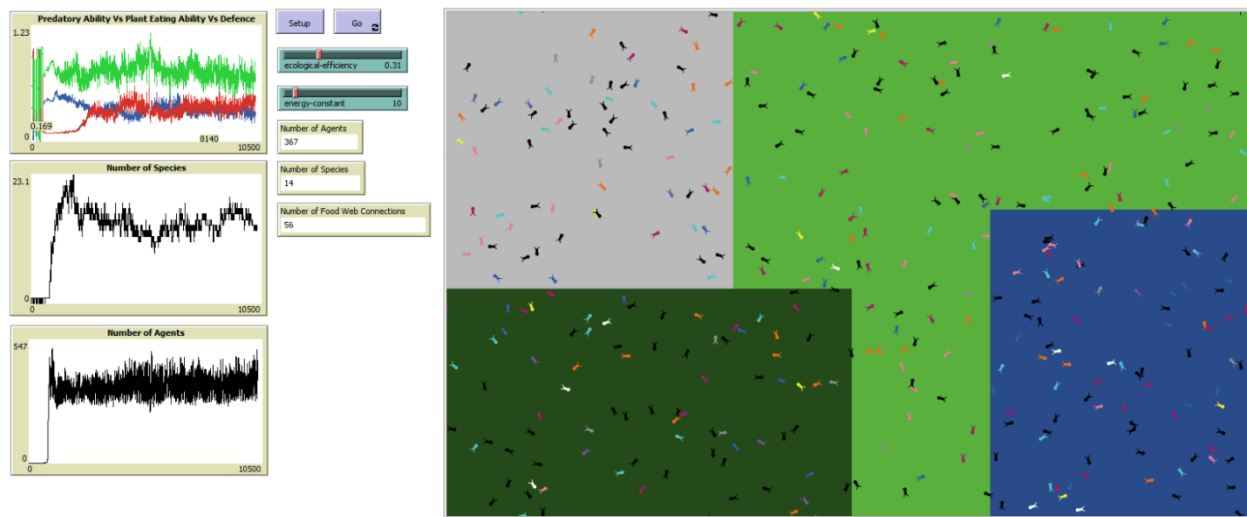


Figure 6.1: Flowchart of the mechanics of the heterogeneous behaviour model

## 6.2 Implementation and Testing

### 6.2.1 Model implemented

The interface of the heterogeneous behaviour model, Figure 6.2, appears similar to those of the previous models. The average predatory ability, plant-eating ability, and defence over all the agents is plotted on the graph in the top left corner to reveal some of the evolutionary dynamics of the food web that are occurring. As these values are an average over all the agents, rather than over a single species, this information's practicality may be somewhat limited, but it should still give some idea of the traits that offer better survivability.



**Figure 6.2:** *The interface of the heterogeneous behaviour model*

### 6.2.2 Problems encountered and changes made

One change made during the implementation was that the mechanics of the geographic regions were implemented differently. Instead of each area having a distinct plant species, in this model 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.

### 6.2.3 Evaluation methods

The biggest issue that arose from taking this new approach to the simulation was how this model would

be evaluated. Unlike the previous models, there are no specified boundaries in this model that define different species. Instead each agent has a unique feature vector of values. Distinguishing different species in this multi-dimensional feature space poses a challenging problem that needs to be resolved to allow this model to be effectively evaluated and compared to the previous models.

The initial solution used for 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 was not a perfect method for differentiating between species, due to the hard boundaries that differentiate species, but it should give a good enough indication of how species have evolved to allow effective evaluation of the model. This approach to classifying the different species in the environment will be referred to as linear species classification.

A more experimental solution that had the potential to provide more accurate species differentiation was a form of unsupervised learning clustering algorithm. According to my research, unsupervised learning has never been used in this context before, so represents a novel approach to species classification in models of food web evolution. It will therefore be interesting to consider the results produced by this model in light of this and evaluate whether it is an effective technique for this purpose. 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.

In order to implement this algorithm in NetLogo, the simulation was linked to Python using NetLogo's 'py' extension. In Python, the algorithm was implemented using Scikit-Learn's 'cluster' module (Pedregosa et al., 2011), which was passed the values for each parameter of each agent that were standardized 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. One downside of this approach was the difficulty in picking suitable values for epsilon and min\_samples. For the purposes of this model the values were picked based on their perceived good performance, but for future studies using this approach a more evidence-based approach should be taken for deciding these values. A further shortcoming of using unsupervised clustering was that it significantly increased the runtime of the model. Whilst the model still ran quickly enough to be feasible, it performed noticeably slower than

the linear species classification approach – meaning that the testing of the model over extended periods of time and with different model parameters was somewhat limited.

Overall, neither of these methods provide a perfectly optimal solution to this problem, but they should each provide a reasonable indication of the species that exist in the simulation. The results from each method will be compared to draw further conclusions about the behaviour of the model. The problem of differentiating between species remains an ongoing area of debate and research in not only the field of agent-based modelling of ecological systems, but also ecology as a whole.

Aside from the problem of determining the boundaries between species, this model will be evaluated in the same manner as the previous models. This will involve constructing the food webs found in the model using GraphViz, monitoring the number of species present in the environment, and observing the way species are behaving in the environment. In addition, the average values for the traits over the whole population will be monitored. This may not help with observing the way that specific species are adapting, but provides insight into the features that are being favoured, and the relationships the features have with each other.

#### 6.2.4 Model execution speed

The execution speed of this model, especially when using unsupervised clustering, is significantly slower than the previous two models. Whilst it is still feasible to use this model over shorter runs, this limits the ability of this model to be used to produce findings over very long runs. This shouldn't have an overly adverse effect on this model's comparability with other models, but it is nevertheless worth noting.

#### 6.2.5 Testing

This model was run through a set of user acceptance tests to assess its performance.

Test	Pass/Fail
The simulation can be correctly setup	Pass
The simulation can be started and stopped	Pass
The number of agents is graphed on the interface	Pass
The number of species is displayed and graphed on the interface	Pass
The number of food web connections is graphed on the interface	Pass
The food web connections can be printed out	Pass
Model parameters can be changed, and these affect the behaviour of the model	Pass
When the model is run, the agents can be observed interacting with each other	Pass
The model can be reset at any time	Pass
The environment has distinct geographical regions	Pass
The species show some degree of adaptation the certain regions	Fail <sup>1</sup>

**Table 6.1: User acceptance tests**

---

<sup>1</sup> Small amounts of adaptation can sometimes be observed, but this wasn't considered to be significant enough for the acceptance test to pass

All the tests passed, except the species in the model showing clear adaptation to particular environments. The cause of this was identified as probably a result of the different geographic regions not having enough distinct characteristics. As the model provided interesting results other than this, which is discussed later, the model was not changed to attempt to add more distinct features to the different regions. However, finding an effective solution to this problem could be developed upon for future research in this subject area.

The functional and non-functional requirements of the model have been met in this model in a similar way to those discussed with the previous models. In addition, the following requirements have been fulfilled by this heterogeneous behaviour model.

#### *Add heterogeneous behaviour to agents*

This model allows each agent to have a value for each of a set of four characteristics: predatory ability, plant eating ability, defence, and speed. Each of these can mutate when an offspring is produced, which allows these characteristics to evolve over time and form different species. This works in a similar way to the evolution of animals in the real-world, though simplified to allow the characteristics and evolution to be computationally modelled.

#### *Adapt the mechanics of the model so it is more consistent with real-world food webs*

The fundamental mechanics changes that were introduced in this model all had the main intention of giving the model more parallels to the real-world, as this should provide the most realistic results from the simulation. For example, the interaction rules state that an agent with a higher predatory ability cannot eat an agent with a lower predatory ability. This was intended to replicate observed trends in real-world food webs, such as sharks being able to eat tuna, also a predator, though tuna cannot eat sharks.

### **6.2.6 Summary**

This section has addressed how the heterogeneous behaviour model has been implemented in NetLogo and the changes that were made to the model during development. In addition, the model has been tested, and meets all the 'must do' and 'should do' requirements of the project and many of the 'could do' requirements. The heterogeneous model has therefore been successfully implemented and can be evaluated regarding its results and findings.

## **6.3 Results and Discussion**

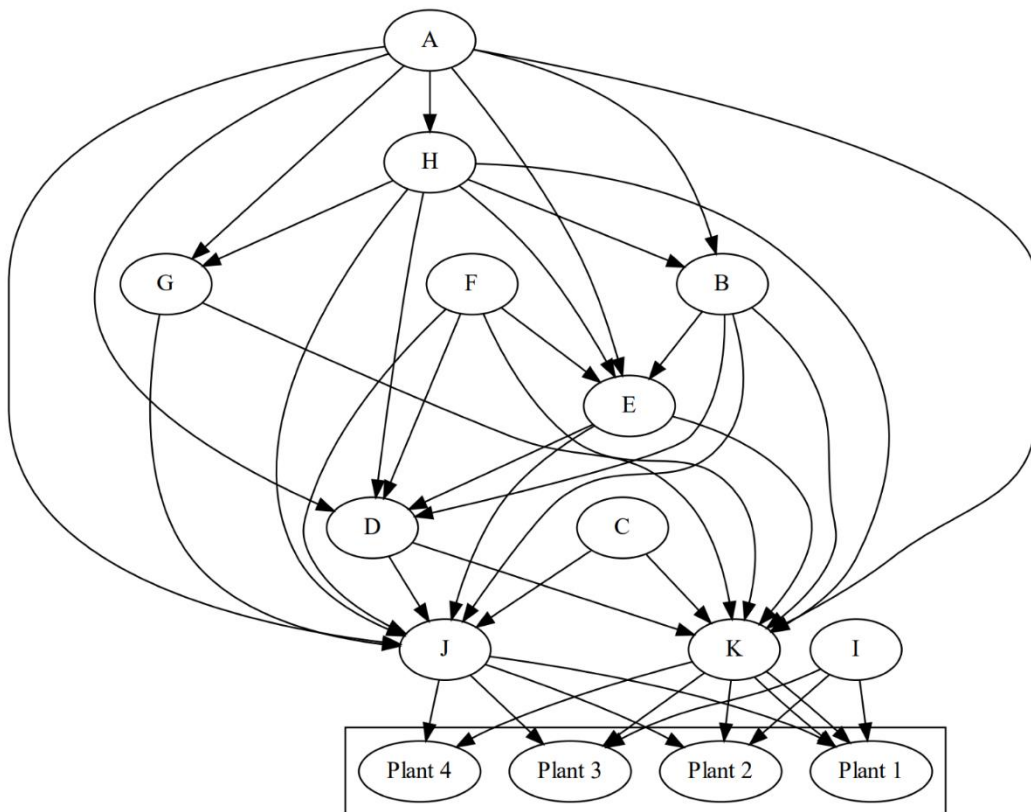
### **6.3.1 Recap of model evaluation and results analysis**

The findings of this model will be analysed and evaluated based on the same protocol as the previous models; namely the depth and structure of the food webs it produces, and the population dynamics

observed in the model. Unfortunately, the method for differentiating between species that is used in this model is less sure than the previous models, so the findings produced by this model have the caveat. Nevertheless, the evaluation techniques used provide points of comparison between all of the models produced during this project, and will allow evaluation of how effective each model has been.

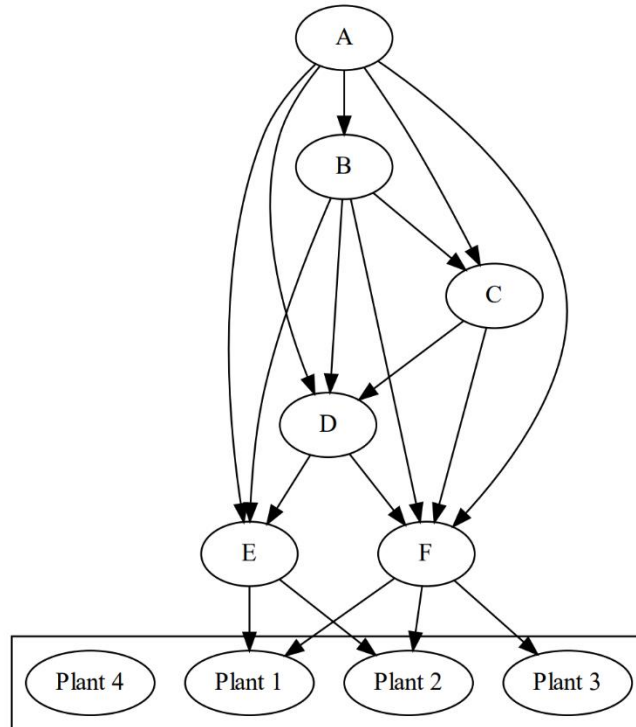
### 6.3.2 Food webs produced by the model

The initial results produced with the model were produced using linear species classification, rather than the unsupervised learning approach, which is explored later. Figure 6.3 shows a food web that was created from one such run of the simulation as described. The first observable difference of this food web to any that have previously been found is its complexity. 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.



**Figure 6.3:** A food web formed by the heterogeneous behaviour model using linear species classification

This food web does, however, seem slightly excessive – indicating that the linear species classification might be dividing the agents into too many species. At this point, the linear species classification was replaced with unsupervised learning clustering, both as this is a more advanced technique for determining different species, and as it provides a second set of results from the model, which should help to reconcile the results found. Figure 6.4 shows a food web that was found using this new species classification approach.



**Figure 6.4:** A food web produced by the heterogeneous behaviour model using unsupervised clustering

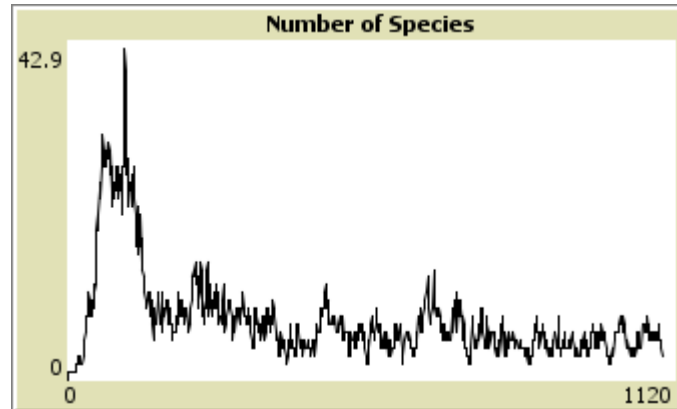
Using this approach consistently produced a food web that is more modest and constrained in the number of species and number of connections than Figure 6.3, but equally it maintains most of the key benefits identified in that food web. Figure 6.4 shows the food web has a clear hierarchical feed-forward structure and multiple trophic levels, again without and bi-directional connections or cycles. Visual observation of this food web in comparison to the real world food web in Figure 2.1 shows that these have a very similar structure and diversity. This shows the effectiveness of the new model mechanics used – and shows that the unsupervised clustering more realistically defines the species in the ecosystem than the linear classification approach.

### 6.3.3 Population dynamics observed in the model

One interesting observation from this model is that as the simulation is run, there is an initial peak in the number of species which then levels off to a more standard level. This represents an interesting population dynamic emerging from the interaction of agents. During the start of the simulation, new species regularly evolve, but are not generally well adapted to the environment that they find

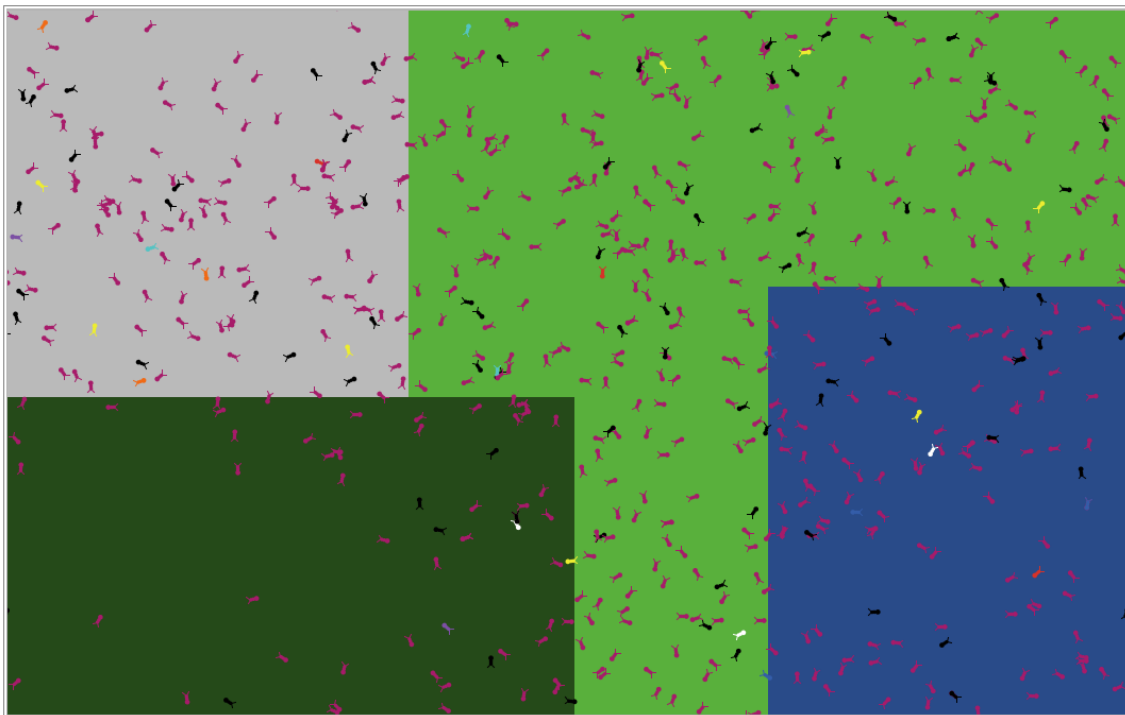


themselves in. They therefore often do not find a niche in the environment and consequently die out quickly, leading to large species churn. As time passes species get better adapted to the environment and start to successfully find niche spots to fill, meaning that the number of species reduces to only those that have found a niche and are well adapted for survival. This trend can be seen in Figure 6.5.



**Figure 6.5:** *The number of species found in a run of the heterogeneous behaviour model over time*

As was discussed in the previous section, one problem with this model is that species do not visibly adapt to different geographic regions in the same way they did in the extended model, as can be seen in Figure 6.6. This is most likely due to there being fewer differences between the regions in this model – the only differences are the plant defence and plant energy. In contrast, in the extended model each area had a unique plant species that only agent species that scored well against them could eat. Once again, this is an improvement that could be investigated further in future research into this topic.



**Figure 6.6:** *An example of the environment found when running the heterogeneous behaviour model*

### 6.3.4 Comparison with other models

Looking at the first and most visual point of comparison, the food webs, Figures 6.3 and 6.4 show a significant increase in diversity over Figure 5.3, which was produced by the extended model. The food webs produced with this model also display a clearly hierarchical structure with multiple trophic levels, whereas the extended model produced this to some extent – but was limited by inconsistencies with real-world food webs, such as bi-directional connections and connection cycles. These findings show clear improvements of this model over both the extended model and the re-implementation model.

In contrast, the lack of species adaptation that is observed in Figure 6.6 shows a backward step in the findings of this model, as Figure 5.4 of the extended model displayed clear adaptation of species to certain environments. This lack of adaptation could be due to either the species differentiation technique being applied or the fact that the environments are not different enough to induce significant speciation by location. Either way, this is an issue that should be addressed in future work.

The heterogeneous behaviour model and the extended model show differences in their emergent population dynamics, but both have some positive aspects in this regard. Whereas the extended model shows intricate dynamics when it comes to the adaptation of species to niche spots in the environment, the heterogeneous behaviour model shows general trends in the population of there being initially many poorly adapted species that reduce in number as they become better adapted to the environment. Both of these population dynamics trends appear to be realistic representations of real-world dynamics, though the dynamics of the extended model may be more insightful for both micro and macro level dynamics of food web evolution.

The number of species over runs of this extended model has been calculated in the same manner as the previous models, with each run being stopped after 500 time-steps – as this model runs quite slowly. The average number of species sustained in each model is shown in Table 6.3 to show the progression of the results throughout this project.

Run	1	2	3	4	5	Average
Mean Species	7.808	9.746	10.591	9.914	6.849	8.982

**Table 6.2:** The average number of species in runs of the extended model

Model	Re-Implementation	Extended	Heterogeneous Behaviour
Mean Species	1.623	5.326	8.982

**Table 6.3:** The average number of species produced by each model

The average number of species the is sustained in the heterogeneous behaviour model is significantly better than both the extended model and the re-implementation model – 8.982 compared to 5.326 and 1.623 respectively. This quantitatively backups up the finding that this model produces the most diverse and complex food webs of the models produced. In addition, Table 6.3 shows how the success of the results produced has progressed over the course of the project, culminating in the highly diverse food webs produced by the heterogeneous behaviour model with unsupervised clustering.

## 6.4 Findings and Future Steps

Overall, this model achieved some very interesting results and ultimately 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 certainly producing more realistic species than a linear species classification approach. However, it is debatable whether a clustering approach is optimal for this application – or whether an alternative approach, such as a tree structure that defines species based on an agent's parent, would be superior. There are also some downsides to this model, such as it showing less visible adaptation of species to survive in certain geographic regions than was seen in the extended model.

Ultimately, this model proved that heterogeneous behaviour with unsupervised clustering is an interesting and fruitful concept for the purpose of simulating the evolution of a food web. However, this incarnation is not perfect and there are many potential future developments that could build upon its promising results. One such improvement could be to add some additional differences to the geographic regions to allow greater adaptation to certain environments. Likewise, more agent characteristics could be added to the model to allow species to evolve to take advantage of niche spots in the environment in different ways. To further increase the realism of the model and how well it parallels the real-world, multi-agent reproduction could be added to the model to observe whether this has any significant impact on how the model performs. For future development of this work, it may be more optimal to implement the model in a more powerful alternative modelling framework such as Repast Symphony, as the heterogeneous behaviour model using unsupervised clustering runs slowly in NetLogo.

## 7 Conclusions

This project was undertaken because agent-based modelling has a great amount of potential for research in the field of ecology, but it had yet to be applied effectively to modelling the evolution of food webs. Norling (2007) attempted to achieve this, but the model produced in the paper failed to realise the full potential of agent-based modelling, with the model being outperformed by a system dynamics model. Using Norling's paper as a basis, this project aimed to re-implement this original model, and make improvements and updates to the agent-based simulation so that it displays more realistic population dynamics and produces food webs that display diversity and structure closer to those observed in real-world food webs. To achieve the aims of this project, a re-implementation model was produced first, followed by two succeeding models that implemented various improvements to improve the correspondence of the model with real-world food webs.

There were several changes made to the original model over the course of this project, but the key ones were the inclusion of a spatial representation, geographical features, and heterogeneous behaviour for agents. One of these changes was added to each of the models developed, with the research of this project culminating with the heterogeneous behaviour model that included all of these improvements, alongside many smaller changes to mechanics.

### 7.1 Findings of the Project

This project produced some interesting findings that relate to a number of key areas that have been discussed over the course of this paper. The main finding is that agent-based modelling is an effective technique for modelling the evolution of food webs. Unlike the original agent-based model produced in Norling (2007), the extended model and heterogeneous behaviour model produced in this project were able to parallel system dynamics models in terms of the diversity of species sustained. The food webs produced by the heterogeneous behaviour model also successfully replicate many patterns that can be observed in real-world food webs. They sustained an average of over 8 species with at least 4 trophic levels, proving the viability of this modelling technique for ecological research and scientific applications. In addition, this project proved that agent-based models of food web evolution can yield interesting insights into some of the underlying population dynamics of the ecosystem, such as how species are adapting to habitat features, which is something that system dynamics approaches cannot achieve. This further realises the potential for the application of agent-based modelling in this field.

The project also explored the novel application of unsupervised clustering to the agent-set of the heterogeneous behaviour model, to determine the different species from the set of continuous data. This approach worked very effectively, assisting the production of realistically structured and diverse food webs in the model, such as Figure 6.4. This finding and proof of concept within this model may have applications across the field of computational ecology, and has the potential to be used in other ecological models in the future.

Another finding of this project was that implementing a spatially explicit agent-based model is an important prerequisite to achieving diverse food webs. However, it does not make any significant impact if other improvements, such as geographical features, are implemented alongside it. This leads

on to a more far-reaching finding of this project, which is that more realistic model mechanics always tended to produce more diverse and realistically structured food webs. Examples of this were inclusions such as geographic features, energy decay of agents, and continuous features in agents, all of which improved the structure of the food webs produced by their respective models. This finding sets the stage for potential future improvements to the model – as it is likely that the models can be improved further by adapting the mechanics to better represent what is occurring in real-world ecosystems. The heterogeneous behaviour model is the most advanced and realistic model that was produced in this project, so this is likely to be the model that would form the basis for future development.

## 7.2 Further Work

The heterogeneous behaviour model represents the first steps into a new style of agent-based modelling of the evolution of food webs. It showed some very promising results, but there are still many potential opportunities for it to be further improved, related to more realistic model mechanics or otherwise. Some such potential future improvements include:

- *Multi-agent reproduction* – currently all the models produced have a simple system of reproduction so when any agent surpasses a threshold of energy, an offspring agent is produced that has that same or slightly mutated characteristics from the parent. Including multi-agent reproduction 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. Most real-world animals also reproduce in this manner, so this change would bring the mechanics more in line with real-world ecosystems.
- *Additional environmental diversity* – this would allow more species to adapt to niche spots in the environment, leading to greater diversity of species. This is another feature that many real-world ecosystems have that leads to their wide diversity, as was explained in section 2.2.1.
- *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.
- *New approach for defining the species* – this could use a different clustering algorithm or some form of tree structure that defines species based on the parents of agents. The unsupervised clustering approach worked well in this project, but it is possible that an alternative approach would be more suitable in this context. Accurately defining the species is important for evaluating the results of the model, so this is a potentially key area that could be improved.
- *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. However, this would require more powerful hardware and potentially a more powerful modelling framework than was used for this project.

## 8 Bibliography

Abar, S., Theodoropoulos, G., Lemarinier, P. and O'Hare, G., 2017. Agent Based Modelling and Simulation tools: A review of the state-of-art software. *Computer Science Review*, 24, pp.13-33.

Baggio, J., Salau, K., Janssen, M., Schoon, M. and Bodin, Ö., 2010. Landscape connectivity and predator–prey population dynamics. *Landscape Ecology*, 26(1), pp.33-45.

Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(Supplement 3), pp.7280-7287.

Brown, R., Reilly, L. and Peet, R., 2016. Species Richness: Small Scale.

Burnett, S., 1997. Colonizing Cane Toads cause population declines in native predators: reliable anecdotal information and management implications. *Pacific Conservation Biology* 3, 65-72.

Caldarelli, G., Higgs, P. and McKane, A., 1998. Modelling Coevolution in Multispecies Communities. *Journal of Theoretical Biology*, 193(2), pp.345-358.

Carter, N., Levin, S., Barlow, A. and Grimm, V., 2015. Modeling tiger population and territory dynamics using an agent-based approach. *Ecological Modelling*, 312, pp.347-362.

Darwin, C., n.d. The Origin Of Species By Means Of Natural Selection.

Encyclopedia.com. 2021. *Population Dynamics* | *Encyclopedia.com*. [online] Available at: <<https://www.encyclopedia.com/earth-and-environment/ecology-and-environmentalism/environmental-studies/population-dynamics>> [Accessed 3 May 2021].

Gimblett, H. Randal. Integrating Geographic Information Systems and Agent-based Modeling Techniques for Simulating Social and Ecological Processes / [electronic Resource]. New York, New York: Oxford UP, 2020. Santa Fe Institute Studies in the Sciences of Complexity. Web.

GraphViz [online] Available at: <[www.graphviz.org](http://www.graphviz.org)> [Accessed 27 March 2021].

Grimm, V., 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future?. *Ecological Modelling*, 115(2-3), pp.129-148.

Grimm, V., 2005. Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science*, 310(5750), pp.987-991.

Grimm, V. and Railsback, S., 2005. *Individual-Based Modeling And Ecology*. Princeton University Press.

Karlson, A., Gorokhova, E., Gårdmark, A., Pekcan-Hekim, Z., Casini, M., Albertsson, J., Sundelin, B., Karlsson, O. and Bergström, L., 2019. Linking consumer physiological status to food-web structure and prey food value in the Baltic Sea. *Ambio*, 49(2), pp.391-406.

Lucidchart. 2021. *Online Diagram Software & Visual Solution | Lucidchart*. [online] Available at: <<https://www.lucidchart.com/>>

Macal, C. and North, M., 2010. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 4(3), pp.151-162.

Manson, S., An, L., Clarke, K., Heppenstall, A., Koch, J., Krzyzanowski, B., Morgan, F., O'Sullivan, D., Runck, B., Shook, E. and Tesfatsion, L., 2020. Methodological Issues of Spatial Agent-Based Models. *Journal of Artificial Societies and Social Simulation*, 23(1).

Martin, R. and Schlüter, M., 2015. Combining system dynamics and agent-based modeling to analyze social-ecological interactions—an example from modeling restoration of a shallow lake. *Frontiers in Environmental Science*, 3.

Morin, P. and Lawler, S., 1995. Food Web Architecture and Population Dynamics: Theory and Empirical Evidence. *Annual Review of Ecology and Systematics*, 26(1), pp.505-529.

Nhm.ac.uk. 2020. *Exploring The Evolution Of Animals At Deep-Sea Hot Springs | Natural History Museum*. [online] Available at: <<https://www.nhm.ac.uk/our-science/our-work/origins-evolution-and-futures/exploring-the-evolution-of-animals-at-deep-sea-hot-springs.html>> [Accessed 19 November 2020].

Norling, E., 2007. Contrasting a System Dynamics Model and an Agent-Based Model of Food Web Evolution. *Multi-Agent-Based Simulation VII*, pp.57-68.

O'Connell, K., Streicher, J., Smith, E. and Fujita, M., 2017. Geographical features are the predominant driver of molecular diversification in widely distributed North American whipsnakes. *Molecular Ecology*, 26(20), pp.5729-5751.

Pedregosa et al., 2011. Scikit-learn: Machine Learning in Python, *JMLR* 12, pp. 2825-2830

Perez, L. and Dragicevic, S., 2010. Modeling mountain pine beetle infestation with an agent-based approach at two spatial scales. *Environmental Modelling & Software*, 25(2), pp.223-236.

Potts, R., Jouval, F., Bigot, L., Bureau, S., Quod, J., Penin, L., Adjeroud, M., Dematteis, B., Ferrucci, M., Coulleri, J., Berzins, L., Dawson, R., Morrissey, C., Clark, R., Silk, M., Hodgson, D., Campetella, G., Chelli, S., Simonetti, E., Damiani, C., Bartha, S., Wellstein, C., Giorgini, D., Puletti, N., Mucina, L., Cervellini, M., Canullo, R., Belsare, A., Vanak, A., Potts, R., Kueppers, L., Armarego-Marriott, T., Ferraz, G., Rutz, C., Loretto, M., Bates, A., Davidson, S., Duarte, C., Jetz, W., Johnson, M., Kato, A., Kays, R., Mueller, T., Primack, R., Ropert-Coudert, Y., Tucker, M., Wikelski, M., Cagnacci, F. and Armarego-Marriott, T., 2020. *Population Dynamics - Latest Research And News | Nature*. [online] Nature.com. Available at: <<https://www.nature.com/subjects/population-dynamics>> [Accessed 1 December 2020].

Rossberg, A., 2013. Food Webs And Biodiversity : Foundations, Models, Data.

R. Vahdati, A., Weissmann, J., Timmermann, A., Ponce de León, M. and Zollikofer, C., 2019. Drivers of Late Pleistocene human survival and dispersal: an agent-based modeling and machine learning approach. *Quaternary Science Reviews*, 221, p.105867.

Society, N., 2021. *Food Web*. [online] National Geographic Society. Available at: <<https://www.nationalgeographic.org/encyclopedia/food-web/>> [Accessed 3 March 2021].

Sofo, A., 2011. *Biodiversity*. Rijeka: InTech.

Van Rossum, G. & Drake, F.L., 2009. Python 3, *CreateSpace*.

Watkins, A., Noble, J., Foster, R., Harmsen, B. and Doncaster, C., 2015. A spatially explicit agent-based model of the interactions between jaguar populations and their habitats. *Ecological Modelling*, 306, pp.268-277.

Wilensky, U., 1999. NetLogo [online] Available at: <<https://ccl.northwestern.edu/netlogo/>> [Accessed 20 March 2021].

Wilson, 1998. Resolving Discrepancies between Deterministic Population Models and Individual-Based Simulations. *The American Naturalist*, 151(2), p.116.