INCLUDING SEED DISPERSERS INTO INDIVIDUAL-BASED FOREST MODELS

by ©Fabio Soares Frazao

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

M.Sc. in Environmental Science, Faculty of Science

Memorial University of Newfoundland

April 2018

St. John's Newfoundland and Labrador

Abstract

Tropical forests provide many ecosystem services to human beings, many of which are mediated by animals. Nonetheless, there are few software tools that allow ecologists to explicitly include animals in forest models. I developed packages to implement individual-based models of trees and animals. The animal model focused on arboreal seed dispersers and includes a decision-making algorithm based on the behaviour of primates. I first describe the models independently and provide examples of how the software can be used to address questions of ecological concern, such as the effect of selective logging on carbon stocks, how deforestation and fragmentation affect animal movement and the impacts of defaunation on seed dispersal. Finally, I describe how to integrate seed dispersers into forest models. I found that this integration is not always necessary and provide guidelines on how to decide if seed dispersal should be explicitly modelled.

Acknowledgements

The conclusion of this thesis would not have been possible without the support of many people and organizations. I would like to thank my supervisor Dr. Amy Hurford for her guidance, patience and for giving me the freedom to develop so many valuable skills while working on this project. My supervisory committee Dr. Shawn Leroux and Dr. Yolanda Wiersma for their advice throughout my degree. Dr. Cibele Hummel do Amaral and Dr. Lourdes Peña-Castillo for their helpful examination of a previous version of this thesis. The current and past members of the Theoretical Biology Lab, for providing such a friendly and intellectually stimulating environment. Memorial University, NSERC and AceNet, for providing the necessary infrastructure and financial resources. I would also like to thank my family and friends for always being supportive. Finally, I would like to acknowledge the efforts of the programmers who collectively built the many open-source tools I used.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Tropical Forests as Complex Systems

Conceptual advances in the science of complex systems have equipped ecologists with new tools and approaches to investigate ecosystems. One of these approaches is looking at ecosystems as Complex Adaptive Systems (CAS), in which the properties of a system are considered as emergent from the interactions among its adaptive components (Green et al., 2005). In Ecology, organisms are usually the adaptive components, since they change their behaviours in response to changes in the physical and biotic environments around them. One example of adaptive behaviour is movement: animals decide where and when to move based on the information they have of their surrounding environment and on their needs (Avgar et al., 2013). Each individual is also part of the biotic environment of other organisms, creating a circular relation of causality that results in emergent properties (Grimm and Railsback, 2005), such as the spatiotemporal distribution of trees and processes like seed dispersal (Messier and Puettmann, 2011).

1.1.1 Ecosystem Processes, Functions and Services

The Ecosystem Services Framework (Groot et al., 2012) is an approach for understanding the "ecological characteristics, functions, or processes that directly or indirectly contribute to human well-being" (Costanza (2012), page 27). These functions and processes underpin the goods and services that nature provides to humans (Groot et al., 2010) and upon which human well-being depends (Millennium Ecosystem Assessment, 2005). As we improve our understanding of ecosystem functions, the connections to ecosystem services become clearer (Loreau, 2010). Balvanera et al. (2006) analyzed 446 studies measuring biodiversity and some kind of ecosystem function or biodiversity and an ecosystem service. They concluded that species diversity is positively related to the provision of supporting services

(*e.g.,* production of food or wood), as these are largely based on ecosystem productivity. Regulatory services followed the same trend. For example, they found that increased plant diversity was associated with lower crop damage from pests and reduced abundance of invasive species. Although the the relationship between biodiversity and ecosystem services is well supported, Gamfeldt and Roger (2017) clarifies that this relationship does not change with the number of functions. The results obtained from their simulations contradict the intuitive claim that a greater number of ecosystem functions necessarily requires higher biodiversity. In order to better support natural resources management, they emphasize the need to identify and understand the mechanisms underlying ecosystem services.

The quantification of ecosystem services is often very challenging due to the many nonlinear relationships between dynamic ecosystem properties and the resulting services (Limburg et al., 2002). Modelling is essential for ecosystem services valuation and several approaches have been developed (Kareiva et al., 2011). However, they have rarely included adaptive behaviour as defined by the CAS view. Nelson et al. (2009) divide ecosystem services models into broad or small scale assessments. Broad scale models use a benefit transfer approach: small spatial units are classified according to habitat types and several services are estimated; after establishing the relationships between the services and the habitat types, the models are used to estimate the same services for broader areas based on habitat maps. These models generally assume that every hectare in each habitat type is equivalent and, since the mechanisms generating the services are not modelled, their ability to generate predictions in new scenarios is limited. Troy and Wilson (2006), for example, used GIS techniques to estimate ecosystem services for three study sites in the USA. They mapped land cover, aquatic resources and the estimated economic values for ecosystem services. The results offer the estimated value derived from all services in US dollars by hectare. However, the spatial analyses contribute very little to understanding the mechanisms generating

specific services and how their provision might change over time.

Models categorized as small scale assessment carefully model one single service in small areas, but produce *"ecological production functions"* that attempt to capture how the service depends on several environmental and ecological variables instead of only on habitat type. Models in this category are better suited for estimating services under changing conditions because the independent variables used as input are more closely related to the actual mechanisms than habitat type alone. However, their ability to elucidate the mechanisms underlying observed relationships is still limited, since they are usually aggregated at population or species levels (Grimm et al., 2017). Ricketts et al. (2004) for example, used artificially augmented pollination of coffee plants near and far from forests to demonstrate that wild bees from nearby forest patches provide an economically quantifiable service, but adaptive behaviour was not represented in the model. Nelson et al. (2009) suggested that ecosystem services models must combine "the rigor of the small-scale approach with the breadth of broad-scale assessment" and offer a framework (called "inVest") to do so. Ricketts et al. (2004) used "inVest" to model five ecosystem services under different land-use change scenarios in Sumatra. The model helped to reinforce the conservation value of the Sumatran tiger habitat, but the production functions were based on land cover and related environmental variables, without any representation of the agents involved in the provision of those services. After analyzing different approaches to ecosystem services modelling, Rieb et al. (2017) argue that, in order to improve scientific understanding and decision-making, future models need to advance beyond using land use or land cover categories as a proxy for ecosystem services.

The importance of animals for the functional maintenance of ecosystems is documented (Estes et al., 2011; Redford, 1992; Ripple et al., 2014; Wilmers and Schmitz, 2016) and their explicit consideration is thought to be valuable in the study and quantification of ecosystem services (Filotas et al., 2014; Limburg et al.,

2002) because the resulting models respond better to environmental changes. Railsback and Johnson (2011) for example, explicitly modelled the population densities and foraging behaviour of birds to quantify how forested patches contribute to pest control in coffee farms. Their final model was able to reproduce several observed patterns and they concluded that a useful model for their context required the birds to have some ability to evaluate habitat quality and make informed decisions. Similarly, population models of stream fish also perform better when using an individual-based approach (Railsback et al., 1999).

Grimm et al. (2017) argue that, modelling the adaptive agents that constitute ecosystems and letting system dynamics emerge from individual activities is the most productive way to integrate ecological levels (*i.e.,* individual, population, community, ecosystem) and produce better mechanistic models that are capable of predicting changes in ecosystem services in response to environmental changes. Many of the processes observed at the higher organizational levels depend on the actions taken by individuals. At the same time, individuals are often where ecological theory can be most strongly linked to reality, as researchers can measure and model individual mechanisms in laboratories or using field observations. Levey et al. (2005) provides a good example of how easily observed individual behaviours (bird movement) can be scaled to model landscape-level mechanisms (seed dispersal) even in conditions unknown to the model. In addition, one of the biological theories we have most confidence in is that individuals behave to increase their future fitness (Grimm et al., 2017; Grimm and Railsback, 2005; Grimm et al., 2005; Railsback and Johnson, 2011; Topping et al., 2012). These principles guide individual-based modelling and, despite usually requiring more data, could complement the more aggregated integration suggested by Nelson et al. (2009), resulting in models that are more mechanistic and suitable to address the demand to explain and predict how ecosystem service change in response to climate and land-use change (Grimm et al., 2017).

1.1.2 Individual-Based modelling

Individual-based models (IBMs) are representations of systems of individuals and the environment they live in, in which the system-level properties arise from the decisions of the individuals (often referred to as *agents*) and their interactions with other individuals and the environment (Railsback and Grimm, 2012). Individual-based ecology is the study of ecological systems from a CAS perspective and is built around IBMs (Grimm and Railsback, 2005). In individual based-ecology, the adaptive behaviours of individual organisms are explicitly modelled as decision algorithms, allowing variability among individuals. In IBMs, ecological processes that emerge from individual behaviours, such as seed dispersal, respond well to environmental changes because these affect the actions taken by individuals. Ecologists often summarize the effects of animals and plants on ecosystem functions by grouping species into functional groups (McGill et al., 2006). This strategy is compatible with IBMs, in which the agents can belong to functional instead of species and even represent groups of organisms (*i.e.,* "collectives" such as fish schools or bird flocks).

1.2 Carbon Sequestration as a key process for conservation

Tropical forests store enormous amounts of carbon as tree biomass and in the organic matter of soils and other plants (Betts et al., 2008). As plants grow they convert atmospheric CO_2 into organic materials (*i.e.,* carbon sequestration), which naturally decay later on, returning the carbon dioxide to the atmosphere. The role of forests in the global carbon cycle is determined by how the vegetation and soil stocks change over time. This depends on how fast CO_2 is converted into biomass; how much is naturally released by respiration, decomposition and other processes; how large the stocks are and how much is released by extreme events, natural or otherwise (Daz et al., 2009). Through carbon dioxide absorption and

storage, tropical forests provide an essential service to the maintenance of the atmospheric CO_2 concentrations. Carbon storage is a good indicator of the stature and intactness of tropical forests (Asner et al., 2018). Carbon stocks are frequently used by conservationists as a target for conservation measures as these are positively related with biodiversity and ecosystem functions (Berenguer et al., 2014; Daz et al., 2009). An important example is the conservation framework devised by the United Nations Framework Convention on Climate Change (UNFCCC) as a mechanism to reduce emissions from deforestation and forest degradation (REDD+)(UNFCCC, 2007), which was recently included in the UN agreement in Paris, signed by most member countries (UNFCCC, 2015). REDD+ initiatives are meant to let developed countries financially support developing economies to use their forest resources more sustainably while still growing economically and carbon stocks play a central role in negotiations.

Many of the processes related to the population dynamics of tree species and the maintenance of carbon stocks are mediated by animals (Reiss et al., 2009; Schleuning et al., 2011). Examples include pollination (Kremen et al., 2007), decomposition (Larsen et al., 2005), nutrient cycling (Nichols et al., 2008), seed dispersal (Sekercioglu, 2006) and seed burial (MacMahon et al., 2000; Nichols et al., 2008). Therefore, the inclusion of animals in forest models is relevant to elucidate mechanisms and produce predictive tools that encompass animal-mediated services.

1.2.1 Seed-dispersal as a selected process for Carbon stocks maintenance

Throughout this thesis, I use seed dispersal as an example of animal-mediated process within individual-based forest models. Seed dispersers contribute to drive how aspects of the plant community (*i.e.,* richness, relative abundance and spatial distribution of species) change over time (Levine and Murrell, 2003a). Community

composition, in turn, affects ecosystem properties (Wall et al., 2005). Although plants rely on a variety of dispersal mechanisms, zoochory (*i.e.,* dispersal by animals) is recognized as important (Lambert and Chapman, 2005). With 95% of the tropical seeds being moved by animals (Terborgh et al., 2002), taking their movement into consideration is important to understand spatial patterns in plant communities, but also very difficult (Bialozyt et al., 2014b). Peres et al. (2016) show that dense-wooded tree species in the Amazon forest are replaced by light-wooded species as large seed dispersers (*i.e.,* tapirs and spider monkeys) are removed, resulting in carbon losses that could range between US$ 5.9 trillion and US$ 13.7 trillion in the world's carbon markets. The dispersers model developed on Chapter 3 is loosely based on primates, therefore focusing on primary dispersal (*i.e.,* the displacement of seeds from the parent tree to a surface, usually the ground, Chambers and MacMahon 1994). Everyday, primates move billions of seeds in tropical forests and account for a large proportion of seed dispersal (Lambert and Chapman, 2005). In the Amazon, Howler (*Allouatta spp.*) and Spider (*Ateles spp.*) monkeys are good model organisms to study primary seed dispersal not only because of the significant proportion of seed dispersed by them, but also because they are preferred targets in bushmeat trade and are very sensitive to anthropogenic disturbance (Lambert and Chapman, 2005; Redford, 1992).

Given the role of seed dispersal in the structure of the forest communities (Howe and Miriti, 2000, 2004; Wang and Smith, 2002), it is reasonable to hypothesize that reductions in the dispersers populations will have a negative effect on carbon related services. Redford (1992) highlighted the dangers that defaunation imposes to the provision of essential ecosystem processes. More recently, Bello et al. (2015) studied the effects of removing large seed dispersers on tropical forests carbon storages. They simulated local extinction of trees dispersed by large vertebrates using data from 31 communities in the Brazilian Atlantic Forest and concluded that extirpation of animals may significantly erode carbon stocks. Wilmers and

Schmitz (2016) provide an example of the magnitude that animals can have on the carbon cycle. Their study estimated that gray wolves could increase net ecosystem productivity by 24 to 52 g $C.m^{-2}.yr^{-1}$ in the Isle Royale National park forest. The synthesis presented by Schmitz et al. (2014) highlights that animals affect local and regional carbon cycles through a variety of mechanisms (*e.g.,* by consuming plant biomass, facilitating wildfires and influencing biophysical conditions through destruction of vegetation). They advocate that natural resources management will benefit from improved quantitative understanding of how animals drive carbon sequestration and storage.

1.3 Impacts of anthropogenic disturbances on the carbon cycle

Human-induced habitat changes severely disrupt tropical forests, threatening biodiversity, species interactions and ecosystem functions (Morris, 2010). Zarin et al. (2015) compiled several datasets and estimated the average carbon dioxide emissions from gross tropical deforestation alone to be 2.270 Gt CO_2 y^{-1}. Berenguer et al. (2014) reported the effects of forest degradation (fires, fragmentation and selective logging) on above ground carbon stocks. Berenguer et al. (2014) found that those forests which had experienced both logging and fires, for example, stored 40% less carbon than primary forests.

Land-use change commonly results in fragmentation of natural habitats. Forests are separated into fragments as we build roads, dams, power-transmission lines or open space for farming, mining and urban expansion. Fragmentation affects species richness, community composition and animal-mediated services (Schleuning et al., 2011). The resulting changes in resource distribution and landscape configuration strongly influence the spatial distribution and movement patterns of forest animals (Garcia et al., 2010). The movement patterns of Howler and Spider

monkeys, for example, are well studied (Belle et al., 2013; Kowalewski et al., 2014; Ramos-Fernndez et al., 2004) and known to be related to the spatial configuration of their habitats (Boyer et al., 2006a; Fiore and Suarez, 2007). As a consequence, the seed dispersal functions performed by these animals is also affected by habitat disturbance, creating a feedback mechanism that results in more changes at the tree community level.

Subtler forms of disturbance, like selective logging, can also affect seed dispersers. It is estimated that 30% of all tropical forests are selectively logged (Pfeifer et al., 2015). Compared to clear cut operations (in which the totality of an area is deforested), selective logging is considered a low impact activity, since only a few target trees are removed and most of the forest is left standing (Gibson et al., 2011a). Recent studies, however, have shown that the impacts of selective logging on biodiversity (Burivalova et al., 2014; Solar et al., 2015) and long term consequences for forest structure have been underestimated, even at very low logging intensities (Gatti et al., 2015). In part, the underestimation comes from not taking the fauna and its roles within the forested ecosystems into consideration (Bello et al., 2015). Selective logging can directly affect monkeys by targeting trees that are important for their diets (Felton et al., 2010), but indirect effects can result from structural modification of their habitat (Azevedo-Ramos et al., 2006; Johns, 1986).

Selective-logging still is one of the most adopted forest uses for sustainable management under these programs (Parrotta et al., 2012). In Brazil for example, Reduced Impact Logging (RIL) is guided by the National Environment Council (CONAMA) under the Ministry of Environment. In The Amazon, the directions establish the maximum logging intensity to be 30 m^3/ha and the interval between operations to be between 25 and 35 years (CONAMA-CONSELHO NACIONAL DO MEIO AMBIENTE, 2009). Other countries adopt similar regulations. Based on a meta-analysis of more than 50 publications addressing selective-logging in

tropical forests distributed across 10 countries, Putz et al. (2012) reported that logging cycles vary from 20 to 40 years in most cases.

1.4 Justification

Both carbon sequestration and carbon stocks are results of the biomass gains at the individual tree level. As seed dispersal plays such an important role in the plant community dynamics and dispersers are strongly affected by forest disturbance, it is essential to understand the effects of disturbance on dispersers and the consequences for forest dynamics and the carbon cycle. The system comprised of trees and ecologically linked primary dispersers (such as primates) is a model system to better understand the cascading effects of anthropogenic actions like selective logging on ecosystem services. The use of a CAS perspective will give the opportunity to investigate the relationship between adaptive behaviour (especially movement) and ecosystem services, an approach that is still new to ecosystem services modelling, but has the potential to produce more accurate predictions.

Attempts to model forest ecosystems are not new. In the 1970s, the growing access to computational resources allowed researchers to mathematically represent multiple interacting processes, leading to the first gap models (Botkin et al., 1972) and later to simple agent-based models (Huston et al., 1988). These tools have been used since then to investigate forest succession and long-term ecosystem dynamics (BEFORE (Rademacher et al., 2004), CAIN (Caspersen et al., 2011), LINKAGES (Post and Pastor, 1996), FORMIND (Köhler and Huth, 1998)). However, most of these forest models were developed for temperate forests and do not explicitly include seed dispersers (although most include mathematical functions that allow the simulation of seed dispersal at the landscape level, as a decreasing function of distance). In addition, many are not open-sourced (FORMIND, BEFORE, LINKAGES, CAIN) or well documented, making it hard to add function-

alities. Although FORMIND's source code is not available, the relatively detailed model description and previous use for modelling tropical forests make it the best starting point for this project.

1.5 Objectives

Given the lack of software tools that include seed dispersers in simulations of forest dynamics and limited availability of tropical forest models that are open-source, well-documented and extendable enough to include disperser agents, the objective of this project is to provide software tools able to create individual-based forest models that explicitly include seed dispersers. Following the general introduction in this first Chapter, my thesis is divided into three software and one summary Chapter.

In Chapter 2 I implement an individual-based forest model using FORMIND's model description as a starting point. I indicate the few differences in the model formulation and provide details of my implementation that clarify how the model works. I also describe some additions I made and provide an example to illustrate typical model outputs.

In Chapter 3 I develop an individual-based model to simulate seed dispersers. The model focus on arboreal dispersers such as monkeys and birds. I describe the model formulation, illustrate how parameter values may affect outputs and provide a few examples that illustrate how the model can be used to address questions regarding the effects of environmental disturbances on seed dispersal.

Chapters 2 and 3 illustrate how each model can be used separately, but in Chapter 4 I describe how they can be integrated to produce individual-based forest models that explicitly include dispersers movement and behaviour.

In the 5$^{\text{th}}$ Chapter I give general guidelines on how to decide if seed dispersal should be explicitly or implicitly simulated. The examples given throughout this

thesis were designed to illustrate the software produced and were not intended to test specific hypotheses since they were not calibrated to specific scenarios. Nonetheless, several hypotheses which could be tested with my models arose from this work and I describe them to be addressed in future work. Finally, I provide links to the software packages I developed, including their source code repositories, which also contain documentation, tests and installation instructions.

Specific objectives when developing the software were to:

- Package the two models separately, so that they can be used on their own, in combination with each other or with third-part models;
- Use object-oriented programming, in order to facilitate modification and extension;
- Make the source code available under the GNU General Public License;
- Provide documentation and tests;
- Host the code on GitHub, for transparent development and to encourage contributions from users and developers.

2 Chapter 2: The Trees model

2.1 Model Description

The individual-based forest model structure is largely based on the FORMIND model (Köhler and Huth, 1998), following the description provided by Fischer et al. 2015 (see Rüger et al. (2007) for another version and application of FORMIND). Most of the equations used are based on the formulation detailed in the supplementary material of Fischer et al. (2015). When equations are identical, references are provided in squared brackets, indicating the equation number in that document (*e.g.,* [eq. 15 in Fischer et al. (2015)-S]). A few significant changes were made to the FORMIND model structure and implementation, which are discussed in detail in sections 2.2 and 2.3 and also indicated in the appropriate submodel descriptions in section 2.1.7. The code that implements the model was rewritten in Python 3 as an Object-Oriented, open-source module (See "Objected-Oriented implementation" below). The description in sections 2.1.1-2.1.7 follows the ODD protocol, proposed by Grimm et al. (2010). Although **?** used FORMIND to simulate an African forest, The model is intended represent any tropical forest given appropriate calibration. Although the parameter values used in the examples presented in this thesis are within typical observed ranges (according to Fischer et al. (2015)-S), they were not calibrated to any specific study site. Therefore, these examples are intended to illustrate how the software works and the typical kinds of outputs.

2.1.1 Purpose

The Trees model is an individual-based, spatially explicit model developed to study the long-term impacts of logging and other forms of disturbance on tree abundances and distributions, carbon stocks and carbon sequestration in uneven-aged mixed species tropical forests.

2.1.2 State variables and scales

The model is implemented in three-dimensional space, and is grid and individual-based. Ecological processes are formulated in three organizational scales: trees, patches and landscape (table 2.1). Each tree is represented individually from birth to death. To enable the dynamics of species-rich forests to be simulated more efficiently, species with similar functional characteristics are grouped in plant functional types (PFTs). Groups are based on shade-tolerance and maximum height (however, single species can be represented just as easily). In addition, cohorts are an extra level that aggregate trees of the same age and PFT within a patch. However, this level is used for purely computational reasons and no attributes are accessible (see section 2.2.1 for more details on cohorts and section 2.3 for more details on PFTs).

Patches are squared cells with unique x and y coordinates and area equivalent to the crown area of a mature tree (parameter a in table 2.1). Patches are divided in small vertical layers (Δh in Table 2.1) that might have different light availability due to light diffusion and interception. Light availability is calculated for each layer, as well as the sum of leaf areas. Patches are linked to their neighbours since individual trees in one patch may fall into a neighbouring patch. The landscape is the set of all patches organized in a rectangular grid with periodic boundaries. Forests are simulated in monthly time steps. The total period of time simulated can be decades or centuries, depending on the simulation goals. For the logging example given on section 2.4 the total time was 50 years because that exceeds the time that most sustainable logging operations wait to revisit a site (CONAMA-CONSELHO NACIONAL DO MEIO AMBIENTE, 2009).

Table 2.1: Attributes for each organizational level. Where appropriate, the respective equation numbers are displayed in squared brackets. Abbreviations appear in round brackets. See table 2.2 for units.

Tree	Patch	Landscape
Age	Area (a)	Total carbon in living trees (AGB_{total})[34]
Identification(id)	Cumulative crown area per layer (CCA)	Total carbon in dead trees (S_{dead})[35]
Patch	Position	Total carbon in fast soil (S_{sfast})[36]
Plant functional type (PFT)	Seed bank	Total carbon in slow soil (S_{slow})[37]
Position	Leaf area index per layer (LAI) [8]	Net carbon exchange (NCE)[41]
Diameter at breast height (DBH) [3]		
Height (H) [4]		
Crown length (CL) [5]		
Crown diameter (CD) [6]		
Crown area (CA) [7]		
Gross primary production (GPP) [21]		
Carbon emission through respiration (Cr) [28]		
Above ground biomass (B) [31]		

Table 2.2: Description of Trees model parameters.

Parameter	Description	Unit
PTF parameters		
Dmax	Max DBH increment	cm
h_0, h_1	Height coefficient	-
cl_0	Crown length coefficient	-
cd_0, cd_1, cd_2	Crown diameter coefficient	-
ρ	Wood density	t_{ODM}/m^3
σ	Ratio of total aboveground biomass to stem biomass	-
f_0, f_1	form factor-stem diameter relationship	-
l_0, l_1	LAI-stem diameter relationship	-
I_{seed}	Required radiation for germination	%
N_{seed}	Number of seeds added to seed bank	1/ha.year
a_{fruit}	adjusting factor for number of fruits produced	-
m	Light transmission coefficient	-
R_g	Growth respiration fraction	-
M_{base}	Base mortality probability	-
DBH_{mort}	DBH up until which mortality is increased for small trees	cm
DBH_{fall}	DBH above which a tree can cause damage if it falls	cm
DΔDmax	Diameter at which ΔDmax is observed	cm
ΔDmax	Maximum stem diameter increment	cm
p_{max}	Max leaf photosynthetic rate	$\mu mol_{CO_2}/m^2 s$
dis_{gen}	Proportion of fruits dispersed by the tree's dispersal method	-
α_{fruits}	adjusting factor for fruit production	-
α	Initial slope of light response curve	$\mu mol_{CO_2}/\mu mol_{photon}$
Global Parameters		
Δh	Width of layers of aboveground vertical space discretization	m
I_0	Incoming irradiance on top of canopy	$\mu mol_{photon}/m^2 s$
k	Light extinction coefficient	-
a	Patch area	m^2
x_{max}, y_{max}	Length and width of the grid that represents the landscape	patches
l_{day}	Period of photosynthetic activity in a day	hours
$t_{S_{dead}}$	Fraction of dead wood decomposed by time steps	-
ϕ_{act}	Days of photosynthetic activity in each time step	days

2.1.3 Process overview and scheduling

Within each time-step (1 month) the model proceeds in the following order: re-calculation of light availability within the forest, growth and update of geometrical characteristics. Mortality and recruitment events occur once a year (every 12 time-steps). According to the question of interest, this frequency of these events can be increased, which might capture habitat changes more realistically at the cost of execution time. This version followed the frequency adopted by Fischer et al. (2015) and Rüger et al. (2007). Logging events are executed as scheduled (the default is every 25 years).

Within each patch, all trees compete for light and space following a gap model. Light availability is assumed to be the main driver of tree growth and forest succession. For each patch, the light availability is calculated according to an extinction law that depends on the vertical distribution of the leaf area. Light availability is then determined for each tree. The same equation is used to calculate the light incidence at the soil level (section 2.1.7, submodel III). Annual growth is calculated based on the physiological processes of photosynthesis (section 2.1.7, submodels IV and V) and respiration (section 2.1.7, submodel VI). Growth is expressed as the monthly increment in aboveground biomass (section 2.1.7, submodel VII). Biomass acquisition is translated into DBH (Diameter at Breast Height) increment, which is used to update height, crown area and other geometrical attributes through allometric equations (section 2.1.7, submodel VIII). Tree mortality can occur either through self-thinning in dense patches, stochastic mortality, gap creation by large falling trees or logging (section 2.1.7, submodel II). Recruitment (section 2.1.7, submodel I) occurs when the light intensity at forest floor exceeds a PFT-specific threshold (parameter I_{seed}; see table 2.3). At the beginning of each year, a PFT-specific number of seeds (defined by parameter N_{seed}) is randomly distributed across all the patches. The effective recruitment rates for each PFT

describes the number of seeds which met the minimum light and space criteria to grow over the DBH of 1 cm in each year. At the end of each year mortality events are executed sequentially. If logging operations are scheduled for that year, they occur after natural mortality events. Figure 2.1 illustrates the order in which submodels are executed.

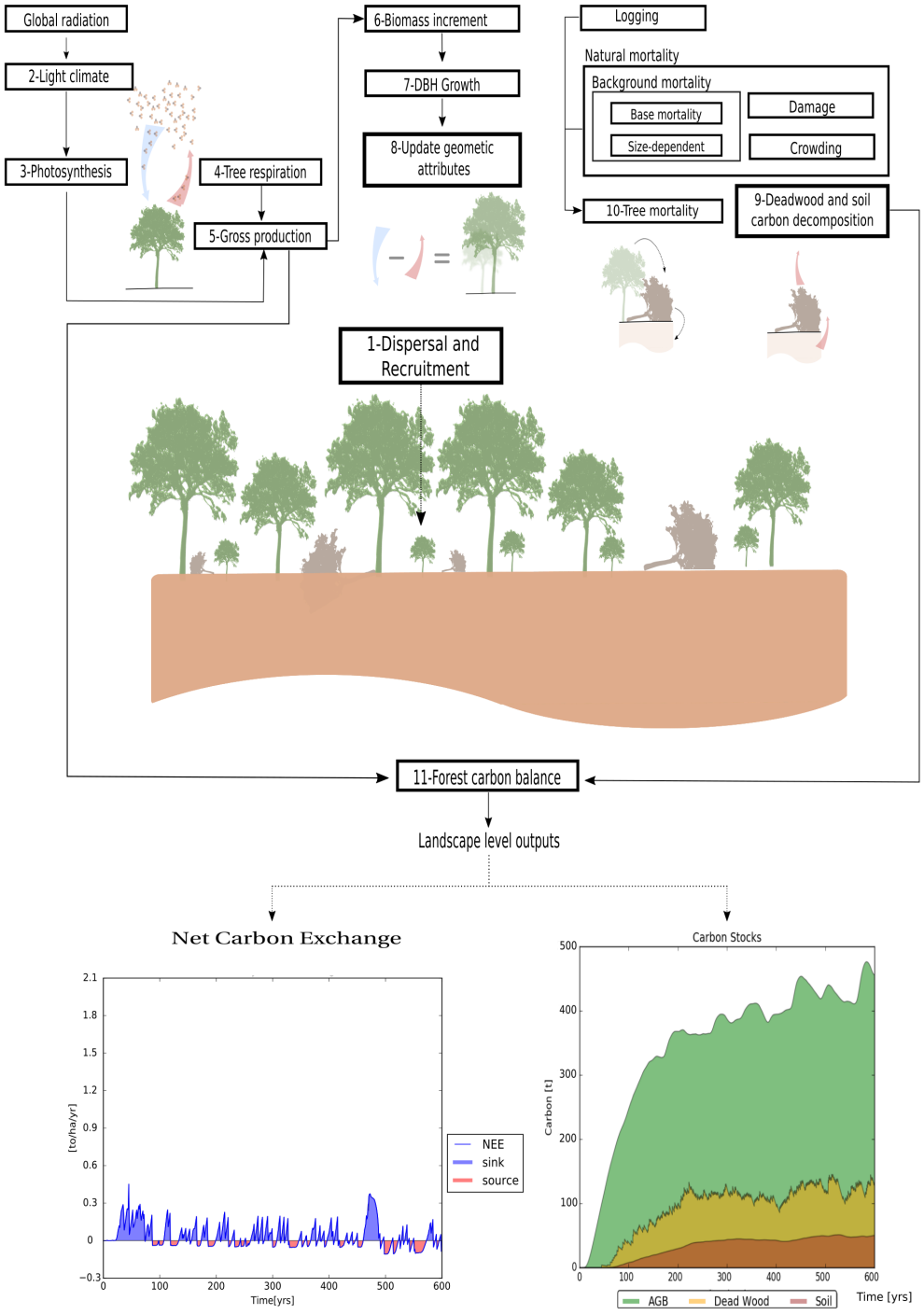

Figure 2.1: Model overview. Boxes correspond to the submodels described in section 2.1.7. Numbers show the order in which events occur. Blue and red arrows represent carbon sequestration and emission respectively. After seed dispersal and recruitment, trees go through the physiological processes leading to growth. Decomposition follows the update of geometric attributes and, after the mortality submodel is executed, landscape level carbon balances are updated. The plots at the bottom display carbon dynamics at the landscape level.

Table 2.3: Parameter values for six Plant Functional Types (PFTs) used in the examples.

	Plant Functional Type					
Parameter	PFT1	PFT2	PFT3	PFT4	PFT5	PFT6
D_{max}	145	120	80	80	47	16
h_0	3.28	4.64	4.82	4.27	4.35	3.0
h_1	0.57	0.41	0.44	0.43	0.34	0.60
cl_0	0.80	0.80	0.30	0.30	0.30	0.30
cd_0	0.60	0.60	0.60	0.60	0.60	0.60
cd_1	0.68	0.68	0.68	0.68	0.68	0.68
cd_2	0	0	0	0	0	0
ρ	0.55	0.54	0.41	0.40	0.52	0.47
σ	0.70	0.70	0.70	0.70	0.70	0.70
f_0	0.77	0.77	0.77	0.77	0.77	0.77
f_1	-0.18	-0.18	-0.18	-0.18	-0.18	-0.18
l_0	2.0	2.0	2.0	2.0	2.0	2.0
l_1	0.10	0.10	0.10	0.10	0.10	0.10
I_{seed}	0.03	0.01	0.05	0.02	0.03	0.02
N_{seed}	30	156	21	300	2	200
m	0.5	0.5	0.5	0.5	0.5	0.5
r_g	0.25	0.25	0.25	0.25	0.25	0.25
M_b	0.015	0.03	0.029	0.04	0.021	0.045
$D\Delta D_{max}$	0.41	0.41	0.28	0.28	0.37	0.28
ΔD_{max}	0.01	0.01	0.01	0.01	0.01	0.01
p_{max}	2.0	3.1	6.8	11	7	12
α	0.36	0.28	0.23	0.20	0.30	0.20

2.1.4 Design concepts

- *Emergence*: Patch level state variables such as light availability at the ground level and vertical layers, as well as landscape level variables such as carbon stocks and net carbon exchange, depend on the abundance and spatial distribution of individual trees.

- *Sensing*: Individuals sense their size (biomass), which affects growth.

- *Interactions*: Trees in the same patch compete for light and space (self-thinning). Trees in different patches may interact when tall trees fall and damage smaller ones.

- *Stochasticity*: All mortality processes are described as probabilities. Seed dispersal can follow a random distribution depending on dispersal mode (uniform or power distribution).

- *Collectives*: Cohorts are collectives of trees with the same age and Functional Type which are located in the same patch. Trees in the same cohort have the same size.

- *Observation*: Observations can be made at the individual (Biomass, Leaf Area and any other individual attribute), population (Abundance, totals and summary statistics for any individual attribute) or landscape levels (Abundance per PFT, totals and summary statistics for any individual attribute, Net Carbon Exchange per year and total carbon stocks in soils, deadwood and living aboveground biomass). Outputs can be recorded hierarchically in a HDF5 database (The HDF Group, 2018). Some plotting functions are also available for the most commonly used data.

2.1.5 Initialization

A simulation can be started with a treeless area, in which case a seed bank is generated by randomly distributing a predefined number of seeds (N_{seed}) across all patches. The total number of seeds is a PFT-specific parameter (N_{seed}) and a uniform distribution is used to determine the position of each seed in the grid. The model can also reconstruct an existing forest from a data file containing information about trees (position in space, id, age, DBH and the PFT to which they belong), the seed bank (how many seed of each PFT are available in the soil) and landscape level carbon stocks. This information can be generated by pausing and saving an ongoing simulation. To test different scenarios on a mature forest, a combination of the two features can be used in order to reduce computation times: starting from a clear area, the model runs for a long period (*e.g.,* 300 years) and

the resulting forest can be saved and used multiple times as the starting point of different scenarios that run for shorter periods of time (*e.g.,* 50 years).

2.1.6 Input

The position and other attributes of all trees can be saved and used as an input, allowing simulations to be paused and continued. Other than that, no other inputs are used. Site conditions are assumed to be homogeneous and there is no inter-annual variability of environmental conditions.

2.1.7 Submodels

The submodels are summarized in figure 2.1.

I. *Recruitment*

Seed dispersal can be implemented by two methods. The first (external seed rain) randomly places seeds across the entire landscape, using the N_{seed} parameter to determine the number for each PFT. The second (implicit seed dispersal) uses a power distribution to disperse the seeds around each tree, based on the fruits available. Each fruit is assumed to contain only one seed. The number of fruits for each tree is calculated by

$$N_{fruits} = DBH \cdot a_{fruits}, \tag{1}$$

where DBH is the Diameter at Breast Height and a_{fruits} is an adjusting factor (*i.e.:* it controls the slope of the linear relationship between DBH and N_{fruits}). Only trees with DBH above 10 cm produce fruits in this version. Chapter 3 describes a third dispersal method (explicit seed dispersal), in which fruits are eaten by explicitly modelled dispersers. In that case, a PFT

specific parameter dis_{gen} determines what proportion of the fruits will be dispersed by the tree's power function (referred to as generic dispersal in Chapter 3) and what proportion of the fruits will remain available for the dispersers to eat.

The probability density distribution for the power function is given by:

$$P(x; a) = ax^{(a-1)}. \tag{2}$$

Each PFT has a light requirement for germination (I_{seed}). At the beginning of each timestep, light incidence at the floor level is calculated (at a patch level). In each patch, all seeds have their PFT-specific light requirement tested against the light incidence. If $I_{floor} > I_{seed}$, then the vertical layer closest to the soil (*i.e.,* the one that the seedling will occupy if it is established) is checked for space. If that layer is not yet fully occupied by seedlings, the current seed is then converted into a small tree with DBH=1cm.

II. *Update of geometric attributes*

A series of allometric equations is used to define the attributes of a tree. Trees are represented as two cylinders (one for the stem and another for the crown) and are addressed by the index i when applicable. See Figure 2.2 (a) for an illustration.

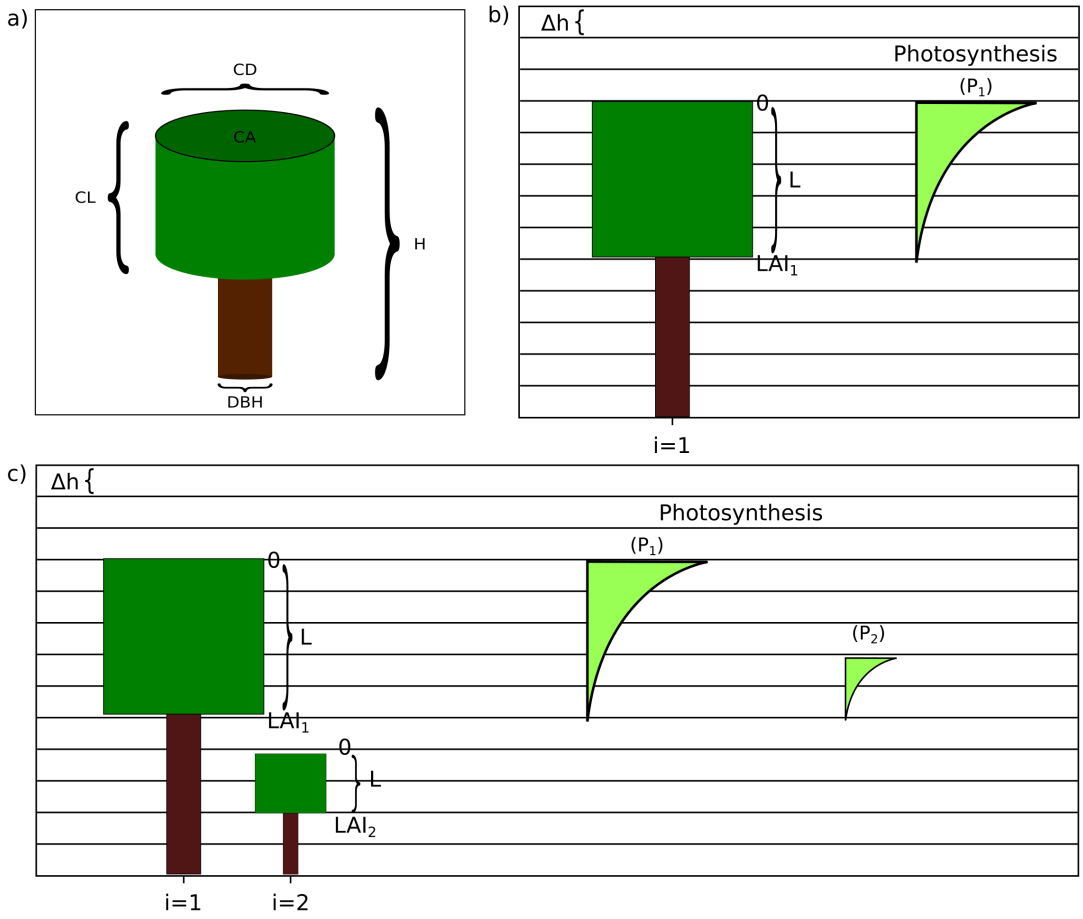

Figure 2.2: Geometrical representation of each tree and vertical stratification. a) Trees are represented as two superimposed cylinders. The crown diameter and length, DBH and tree height are geometrical attributes used in the calculation of biomass. b) Vertical stratification used to calculate photosynthesis for each individual (P_i). Crown Length is divided into layers. Light incidence is higher at the top (layer zero) due to self-shading. c) Taller trees ($i = 1$) limit the photosynthetic activity of smaller individuals ($i = 2$).

When an individual grows (submodel VIII), the biomass gain is translated into DBH increment by the following equation:

$$DBH = \sqrt{\frac{B}{\frac{\pi}{4} \cdot H \cdot f \cdot \frac{\rho}{\sigma}}}, \tag{3}$$

where B is the new biomass, H is the height of the tree, f the form factor that for the idealized cylindrical shape assumed for the stem, ρ the wood density and σ the fraction of stem wood biomass from total tree biomass [derived from eq. 5 in Fischer et al. (2015)-S].

All the other geometrical characteristics are updated based on the new DBH as described below.

Height(H) [eq. 1 in Fischer et al. (2015)-S] is determined by:

$$H = h_0 \cdot DBH^{h_1}, \tag{4}$$

where $h0$ and h_1 are type specific parameters.

Crown length (CL) [eq. 2 in Fischer et al. (2015)-S] is a function of height, with c denoting the type-specific parameter that defines the relationship.

$$CL = c \cdot H. \tag{5}$$

Crown diameter [eq. 3 in Fischer et al. (2015)-S] is proportional to stem diameter, where cd is another type-specific parameter.

$$CD = cd \cdot DBH. \tag{6}$$

Since the crown is represented as a cylinder, crown area (CA) [eq. 4 in Fischer et al. (2015)-S] is

$$CA = \frac{\pi}{4} \cdot CD^2. \tag{7}$$

The leaf area index of a tree (LAI) [eq. 8 in Fischer et al. (2015)-S] represents the total amount of leaf area per unit of crown projection area, which relates to the photosynthetic capacity. It is calculated as a function of DBH,

$$LAI = l_0 \cdot DBH^{l_1}, \tag{8}$$

where l_0 and l_1 are PFT-specific parameters.

III. *Tree mortality* Mortality is based on section E1 in Fischer et al. (2015)-S, but includes additional mortality risks to small trees following Rüger et al. (2007) and excludes deterministic mortality.

(1) Background mortality: each tree has a base mortality (m_{base}) specific to its PFT. The background mortality m_B for big trees is simply m_{base}. Those trees with DBH below the DBH_{mort} threshold have higher background mortality:

$$m_B = \left\{ \begin{array}{ll} m_{max} - m_{max} \cdot \frac{DBH}{DBH_{mort}} + m_{base}, & \text{if } DBH < DBH_{\text{mort}} \\ m_{base}, & \text{else} \end{array} \right\} \tag{9}$$

where m_{max} is the maximum size-dependent mortality of small trees and DBH_{max} is the DBH up to which tree mortality is increased.

(2) Self-thinning: Mortality is increased for trees in dense patches due to competition for space. Crowding based mortality, m_C, is modelled as

$$m_{\text{C}} = 1 - \frac{1}{CCA_{\text{max}}}, \tag{10}$$

where CCA_{max} is the maximum Cumulative Crown Area (which sums the crown areas of all trees with the crowns in that height layer).

(3) Damage mortality

When a large tree dies, it might fall and impact its neighbours. Large trees are those with DBH above the size threshold defined by the parameter DBH_{fall}. Damage is inflicted on the small trees ($DBH < 50$ cm) at least 1 m shorter than the falling tree if they are located on the area where the crowns is going to land. Damage is expressed as the ratio between the crown projection area of the falling tree (see submodel VIII for details on the calculation of CA) and the patch area (parameter a). The direction of the fall is chosen randomly with equal probability. The damage mortality probability is

$$m_{\mathrm{D}} = \frac{CA}{a}. \tag{11}$$

The tree mortality processes are executed sequentially for each tree. A random number between 0 and 1 is drawn from a uniform distribution and the tree dies if the number is lower than m_B, otherwise the process is repeated for m_C and m_D. When a tree dies, its carbon content (44% of the biomass, according to Fischer et al. (2015)) is transferred to a temporary stock S_{mort}, which is subsequently used to update the dead wood stock S_{dead} as described in equation 35.

IV. *Competition for light*

Each patch is vertically divided in small layers of width Δh (see figure 2.2). Light availability is calculated for the individuals within each patch taking

into consideration the canopy layers they occupy. If a patch has only one tree, all the light will be available to that individual. Due to self-shading, higher stratifications of the crown intercept more light, reducing the availability for the layers below. Similarly, with the addition of other trees to the same patch, the tallest ones shade the smallest.

For each patch and height layer, the leaf area accumulated by all individuals is calculated. Each tree contributes parts of its crown leaf area to those height layers which are occupied by its crown. The bottom and top layers [eq. 31 and 32 in Fischer et al. (2015)-S] are defined by:

$$l_{top} = \frac{H}{\Delta h}, \tag{12}$$

and

$$l_{bottom} = \frac{H - CL}{\Delta h}, \tag{13}$$

where H is the tree's height and CL its crown length.

A tree's leaf area index (LAI) [based on eq. 34 in Fischer et al. (2015)-S] contributes equally to each of the patch layers between l_{bottom} and l_{top}:

$$LA_l = \frac{LAI \cdot CA}{l_{top} - l_{bottom}}, \tag{14}$$

where LA_l is the tree's contribution to the leaf area of vertical layer l of that patch and CA is its crown area. The multiplication of LAI by CA gives the leaf area of a single tree in m^2.

To calculate the leaf area per layer (l) at the patch level LAP_l [based on eq. 35 in Fischer et al. (2015)-S], the contribution of each tree's leaf area per layer (LA_l) relative to the patch area (a) is summed,

$$LAP_l = \frac{1}{a} \sum_{\substack{\text{all individuals} \\ \text{w/ crown in } l}} LA_l \tag{15}$$

.

With this information it is possible to calculate the radiation that each tree is able to intercept at the top of its crown (I_{top}) [eq. 36 in Fischer et al. (2015)-S],

$$I_{top} = I_{max} \cdot \exp\left(-k \cdot \sum_{l>l_{top}} LAP_l\right), \tag{16}$$

where $\mathrm{I_{max}}$ is the irradiance above the forest canopy and k the light extinction coefficient. Note that only those layers above l_{top} are used.

V. *Photosynthesis*

Photosynthesis is based on the incoming irradiance at the top of each tree (I_{top}). Following the approach described by Thornley and Johnson (1990) [and eq. 37 in Fischer et al. (2015)-S], the cylindrical crown is considered as a stack of thin disks. The photosynthesis of each disk is modelled by a Michaelis-Menten function describing the relationship between the radiation available at the top of each disk $I(L)$ and its photosynthetic rate $P(I)$:

$$P(I) = \frac{\alpha \cdot I(L) \cdot p_{max}}{\alpha \cdot I(L) + p_{max}}, \tag{17}$$

where α is the initial slope of the light response curve and p_{max} is the maximum rate of photosynthesis. The incident irradiance [eq. 38 in Fischer et al. (2015)-S] on the surface of a leaf is:

$$I(L) = \frac{k}{1-m} \cdot I_{top} \cdot e^{-k \cdot L}, \tag{18}$$

where I_{top} is the irradiance incident on the tree crown, k the light extinction coefficient and m the light transmission coefficient of leaves. Since all leaves are assumed to be distributed homogeneously within the crown, the higher ones will shade the ones below. Thereby, $L = 0$ represents the top of the individual and $L = LAI$ represents the bottom, with LAI being the leaf area index (eq. 8).

The photosynthetic rate per individual (P_i) [eq. 39 in Fischer et al. (2015)-S] is obtained by integrating equation 17 over the individual's leaf area index:

$$P_i = \int_0^{LAI_i} P(I(L))dL. \tag{19}$$

The integration results in [eq. 40 in Fischer et al. (2015)-S],

$$P_i = \frac{p_{max}}{k} \cdot \ln \frac{\alpha \cdot k \cdot I_{top} + p_{max}(1-m)}{\alpha \cdot k \cdot I_{top} e^{-k \cdot LAI_i} + p_{max}(1-m)}. \tag{20}$$

VI. *Gross production*

Equation 20 gives the photosynthetic rate in $\mu mol_{CO_2}/m^2 s$. The Gross Primary Production (GPP) [eq. 42 in Fischer et al. (2015)-S] per tree per year is calculated by multiplying the individual photosynthetic rate P_i by the crown area (CA) and a conversion coefficient ($codm$) that transforms absorbed CO_2 to tons of organic dry mass per year (t_{ODM}/y) :

$$GPP_i = P_i \cdot CA \cdot codm. \tag{21}$$

In the conversion factor [based on eq. 41 in Fischer et al. (2015)-S], 3600 accounts for the conversion from seconds to hours and l_{day} represents the average daylight period. The parameter ϕ_{act} gives the number of days in which there is photosynthetic activity within a time step. For tropical environments, trees are assumed to do photosynthesis every day of the year. The final components $0.63 \cdot 44 \cdot 10^{-12}$ include the molar mass of CO_2, the conversion from grams of CO_2 to grams Organic Dry Matter (44%) and the conversion from *grams* to *tonnes*:

$$codm = 3600 \cdot l_{day} \cdot \phi_{act} \cdot 0.63 \cdot 44 \cdot 10^{-12}. \tag{22}$$

VII. *Respiration*

Respiration is divided into two components: growth and maintenance. Growth respiration is assumed to be a constant fraction of the GPP that is used during the build up of new biomass and is defined by the parameter R_g.

Maintenance respiration (R_m) [eq. 44 in Fischer et al. (2015)-S] is calculated

based on the estimated diameter growth,

$$Rm = \left(GPP - \frac{B_{i,est} - B_i}{1 - Rg} \right), \tag{23}$$

where B_i is the current biomass of tree i (see equation 31) and $B_{i,est}$ is the estimated biomass of the same tree after taking into consideration the diameter growth. The latter is calculated using equation 31, however the DBH used as the input corresponds to the current diameter plus the diameter increment,

$$B_{i,est} = B(DBH_i + g(DBH_i)) \tag{24}$$

The growth function used in the equation above is designed to match empirical measurements [eq. 48 in Fischer et al. (2015)-S]. The diameter increment is given by:

$$g(DBH) = \alpha_0 \cdot DBH \cdot \left(1 - \frac{DBH}{Dmax} \right) \cdot e^{-\alpha 1 \cdot DBH} \tag{25}$$

$$\alpha 0 = \frac{e^{\frac{Dmax - 2 \cdot (D\Delta Dmax \cdot Dmax)}{Dmax - (D\Delta Dmax \cdot Dmax)}} \cdot Dmax \cdot \Delta Dmax}{(Dmax - \Delta Dmax \cdot Dmax) \cdot D_{\Delta Dmax} \cdot Dmax} \tag{26}$$

$$\alpha 1 = \frac{Dmax - 2 \cdot \Delta Dmax \cdot Dmax}{Dmax \cdot D_{\Delta Dmax} \cdot Dmax \cdot (D_{\Delta Dmax} \cdot Dmax)^2} \tag{27}$$

Where Dmax is the maximum diameter a tree can reach, ΔDmax is the maximum increment and $D_{\Delta Dmax}$ is the stem diameter which reaches ΔDmax.

See Köhler and Huth (1998) for more information on how these parameters can be estimated from field measurements.

The equivalent maintenance respiration in tons of organic dry mass is then calculated using the same conversion coefficient used for GPP,

$$Rm = rm \cdot B_i \cdot codm. \tag{28}$$

VIII. *Biomass increment*

The biomass increment [eq. 43 in Fischer et al. (2015)-S] can be calculated by subtracting the yearly respiration value from GPP,

$$B_{inc} = (1 - R_g)(GPP - Rm). \tag{29}$$

The new biomass of a tree is then determined by,

$$B_{\text{new}} = B_{\text{old}} + B_{\text{inc}}. \tag{30}$$

IX. *Growth*

The biomass is updated based on the DBH increment. By rearranging equation 3, the following equation is obtained,

$$B = \frac{\pi}{4} \cdot DBH^2 \cdot H \cdot f \cdot \frac{\rho}{\sigma}, \tag{31}$$

where H is the height of the tree, f a form factor that compensates for the idealized cylindrical shape used to represent the stem, ρ the wood density and σ the fraction of stem wood biomass from total tree biomass.

X. *Carbon balance*

Carbon is stored in four different stocks:

- AGB_{total}: Living trees, which is the sum of the carbon in each living tree.
- S_{dead}: Deadwood, which equals to the total carbon in dead trees.
- S_{fast}: Fast decomposing soil stock, which equals to the amount of carbon resulting from quick decomposition of dead trees.
- S_{slow}: Slow decomposing soil stock, which equals to the amount of carbon resulting from slow decomposition of dead trees.

The dynamics of the total aboveground biomass stock AGB_{total} is determined by the difference between the total carbon captured (see submodel IV-Growth above) as Gross Primary Production (C_{GPP}) [page 28 in Fischer et al. (2015)-S] and total carbon emissions from tree respiration (C_R):

$$C_{GPP} = 0.44 \cdot \sum_{i=1}^{N_t} GPP_i, \tag{32}$$

$$C_R = 0.44 \cdot \sum_{i=1}^{N_t} (R_{m,i} + R_{g,i} \cdot (GPP_i - R_{m,i})), \tag{33}$$

$$AGB_{total} = C_{GPP} - C_R. \tag{34}$$

The following equations [in page 27 of Fischer et al. (2015)-S] calculate the amount of carbon for each of the remaining stocks,

$$S_{dead} = S_{mort} - (t_{S_{dead}-A} + t_{S_{dead}-S_{slow}} + t_{S_{dead}-S_{fast}}) \cdot S_{dead}, \quad (35)$$

$$S_{fast} = t_{S_{dead}-S_{fast}} \cdot S_{dead} - t_{S_{fast}-A} \cdot S_{fast}, \quad (36)$$

$$S_{slow} = t_{S_{dead}-S_{slow}} \cdot S_{dead} - t_{S_{slow}-A} \cdot S_{slow}, \quad (37)$$

Where S_{mort} is the total carbon from trees that died in the past time step. The transition rates [in page 27 of Fischer et al. (2015)-S] depend on how quickly dead wood is decomposed, which is represented by the parameter $t_{S_{dead}}$. It is assumed that 70% of the dead wood emissions go directly to the atmosphere and the other 30% are distributed between the fast and slow decomposing soil stocks. These rates are based on empirical measurements (see Fischer et al. (2015) for details).

$$t_{S_{dead}-A} = 0.7 \cdot t_{S_{dead}}, \quad (38)$$

$$t_{S_{dead}-S_{slow}} = 0.015 \cdot 0.3 \cdot t_{S_{dead}}, \quad (39)$$

$$t_{S_{dead}-S_{fast}} = 0.985 \cdot 0.3 \cdot t_{S_{dead}}. \quad (40)$$

Net carbon exchange [eq. 51 in Fischer et al. (2015)-S] is calculated by sub-

tracting emissions (living trees and stocks) from the total carbon absorption,

$$NCE = C_{GPP} - C_R - t_{S_{dead-A}} \cdot S_{dead} - t_{S_{dead-A}} \cdot S_{fast} - t_{S_{dead-A}} \cdot S_{slow}. \quad (41)$$

XI. *Logging*

The model keeps track of harvestable trees that comply with defined criteria for the logging scenarios (e.g., commercial PFTs, minimum and maximum allowed DBH thresholds for harvesting). Logging events are scheduled according to the frequency determined for the scenario (*e.g.,* every 25 years). Before the logging module is applied, the minimum criterion (i.e., minimum number of trees to be extracted per hectare) is evaluated. If the minimum criterion is met, a logging operation takes place; otherwise logging is omitted. Patches are visited randomly and the largest harvestable tree on the patch is logged, until all patches have been visited at least once or the harvest target has been met. Then, patches are revisited randomly until the harvest target is met.

2.2 Details of implementation

The model was written in the Python 3 programming language. Six PFTs were used in the example on section 2.4. The model outputs are stored in a HDF5 database. There are plotting functions to generate static plots and also a tool for interactive visualizations (see section 2.3.1). Figure 2.3 contains the UML (Unified Modelling Language) class diagrams illustrating the internal organization of the *trees-ibm* package (see Chapter 5 and table 5.1 for details on the packages.)

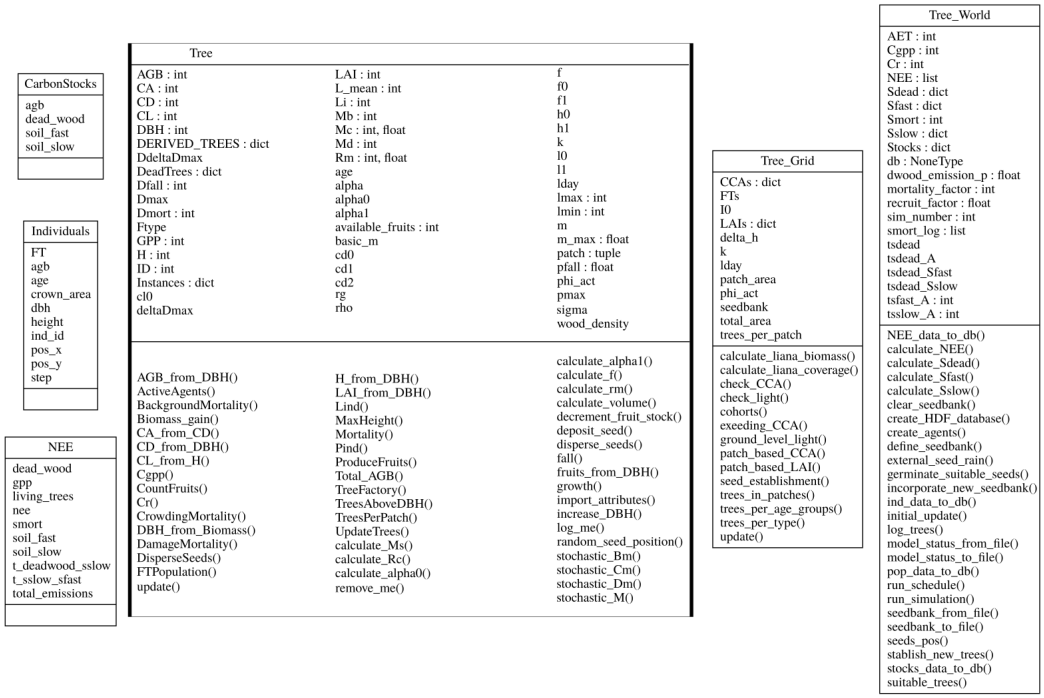

Figure 2.3: Class diagram for the Trees-IBM package. Following the UML (Unified Modelling Language) standards, each class is represented by a block divided into three sections: the class name at the top, followed by the corresponding attributes in the middle and the methods at the bottom.

2.2.1 Cohorts

In order to enhance computational performance, individual trees with the same characteristics are treated as collectives (cohorts). Each cohort contains individuals located in the same patch that have the same type (PFT) and age. Large trees (with DBH $>$ 40 cm) often do not have similar individuals close enough to form a cohort.

In each cohort, one tree is chosen and its biomass calculated. From the biomass all other morphological variables such as DBH, height, crown diameter, crown depth, and stem volume are derived and transferred to the other individuals in the same cohort. Since environmental conditions (*e.g.,* light availability) are defined at the patch level and trees with same type and age will grow at equal rates under

equal conditions, all geometric attributes will change by an equal amount for all trees in a cohort. Therefore these are calculated only once and the values are copied to the remaining members. Execution time was reduced by almost 60% with the adoption of this strategy.

2.2.2 Scaling

One of the drawbacks of spatially explicit models is that the execution time generally increases with the area simulated, since more elements need to be represented. In this model, space is represented by a square grid in which each patch is represented by a cell in the grid. That cell represents an area usually equal to the crown area of a big tree (*e.g.,* 20m x 20m), in order to capture the local competition for light accurately. Figures 2.4 and 2.5 show how execution time and RAM memory use increase as the simulated area increases: both relationships are nearly linear. The simulations were executed on a an iMac with a 3.4 GHz Intel Core i5 cpu and 8 GB of RAM. Future versions of this model will use parallel computing techniques in order to reduce the execution time for larger areas.

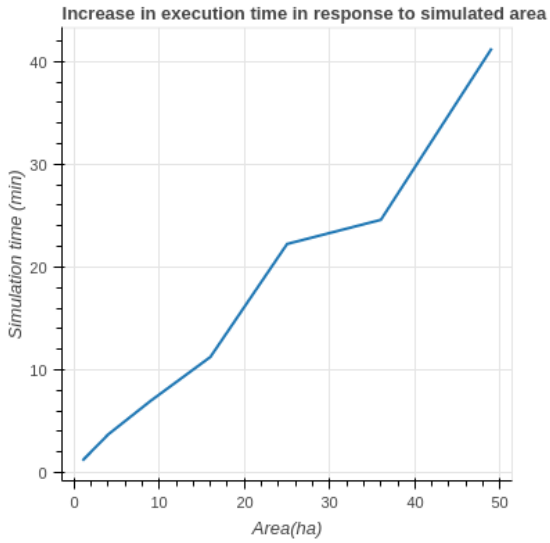

Figure 2.4: Simulation time increases linearly with the area simulated in the Trees model. The x axis goes from 25 cells (a 5 x 5 grid, equivalent to 1 ha) to 1225 cells (a 35 x 35 grid, equivalent to 49ha). Each grid cell is equivalent to 400 mextsuperscript2.

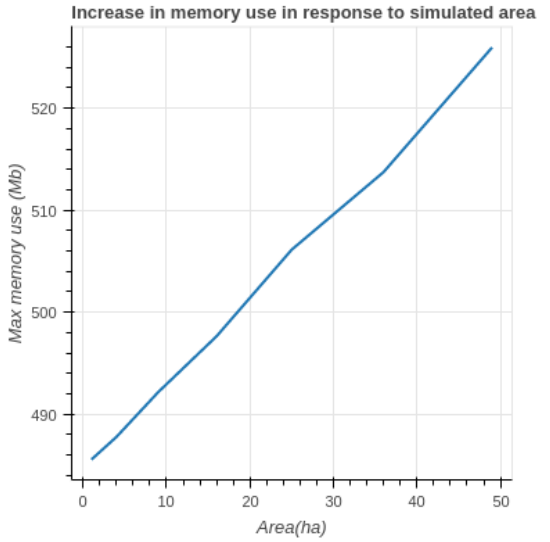

Figure 2.5: Memory use increases linearly with the area simulated in the Trees model. The x axis goes from 25 cells (a 5 x 5 grid, equivalent to 1 ha) to 1225 cells (a 35 x 35 grid, equivalent to 49ha). Each grid cell is equivalent to 400 mextsuperscript2.

2.2.3 Tests

One of the difficulties that arise from the bottom-up approach used in individual-based modelling is that processes that take place at one level of organization interact with other levels (Thiele et al., 2014), making it harder to understand what influences the resulting system-level patterns. This section presents several simplified simulations, breaking down the complexity level present in a typical use of the model (*i.e.,* the example in section 3.4) to clarify how variability in PFT parameters (figure 2.6), stand structure (in terms of PFT and stem size distribution, figure 2.7) and competition (figure 2.8) affect the carbon cycle. At the same time, figure 2.6 demonstrates how the carbon absorbed during an individual's life is slowly but completely released to the atmosphere after death.

Single trees:

Six Plant Functional Types were defined according to the parameter values in Table 2.3. These were based on the PFTs in Fischer et al. (2015). Results from this test simulation are found in figure 2.6. For each type, the simulation started with one single sapling with stem diameter equal to 1 cm and continued for 700 time steps, during which the individual executed the photosynthesis, respiration and growth submodels (section 2.1.7). The regular birth and recruitment submodels described in section 2.1.7 were turned off. At time 700, the tree was killed and its biomass transferred to the deadwood stock. The simulation continued until time 2000. During this second phase, there was no carbon sequestration (since there were no living trees anymore) and the carbon in the deadwood was transferred to the soil stocks and to the atmosphere.

Emissions from the soils further decrease the net carbon balance, which was negative until the end of the simulation. Given enough time, all the carbon once captured by the tree during its growth was released to the atmosphere, resulting in a balance of zero. Finally, these tests illustrate the shape of the carbon absorption

curve that results from the growth model adopted (equation 25): trees grow faster when they are young and the net carbon capture rate declines as they approach the maximum size, which is a largely accepted assumption in plant biology (Carey et al., 2001; Phillips et al., 2008; Weiner and Thomas, 2001). However, Stephenson et al. (2014) questioned this assumption based on a substantial amount of data from many tree species around the globe. Stephenson et al. (2014) argue that large trees capture significantly more carbon than smaller ones, however, there is little support for their alternative assertion in the literature so far and the current version of this model adopts the currently established relationship.

Stands:

Following the more detailed look into how individual trees contribute to the carbon cycle, this section explores the stand level carbon exchanges that emerge from the interaction of multiple individuals. The total sequestration and stocks at a given moment are influenced not only by what kind of trees are in the stand but also their sizes. Trees grow for 600 time steps (50 years). Each stand is comprised of 10 trees and there is no competition among them. In figure 2.7 (a) the carbon sequestration curve follows the same shape seen in figure 2.6 (a), but scaled to the number of trees: the carbon stocks peaks at approximately 500 t as the trees reach their maximum size. In figure 2.7 (b), the effect of heterogeneous sizes is shown to decrease carbon capture rates, which result in lower carbon stocks under this scenario. When half of the stand is substituted by the less productive PFT2, the decrease is even more pronounced (figure 2.7 (c)).

In addition to stand composition, competition also affects total carbon sequestration. Figure 2.8 shows the results of simulations with 3 trees, each belonging to a different PFT. In the absence of competition (a), individuals grow with full resource availability (sunlight). In (b), however, all trees are in the same patch. The tallest PFT2 individual is not affected, while the individual from PFT3 has

its maximum sequestration slightly reduced. The smallest tree, however, not only has its carbon sequestration capacity reduced, but starts to act as a source of carbon. In this case, the tree is intercepting so little solar radiation that its gross primary production is lower than the amount of energy it needs for maintenance, resulting in more emissions through respiration than sequestration. In normal conditions, such a tree would die and would not act as a source of carbon for long. Since in this scenario mortality was turned off in order to highlight the effects of competition on individual carbon exchange, this tree remained in the simulation functioning as a source.

Variation in net carbon exchange of six plant functional types

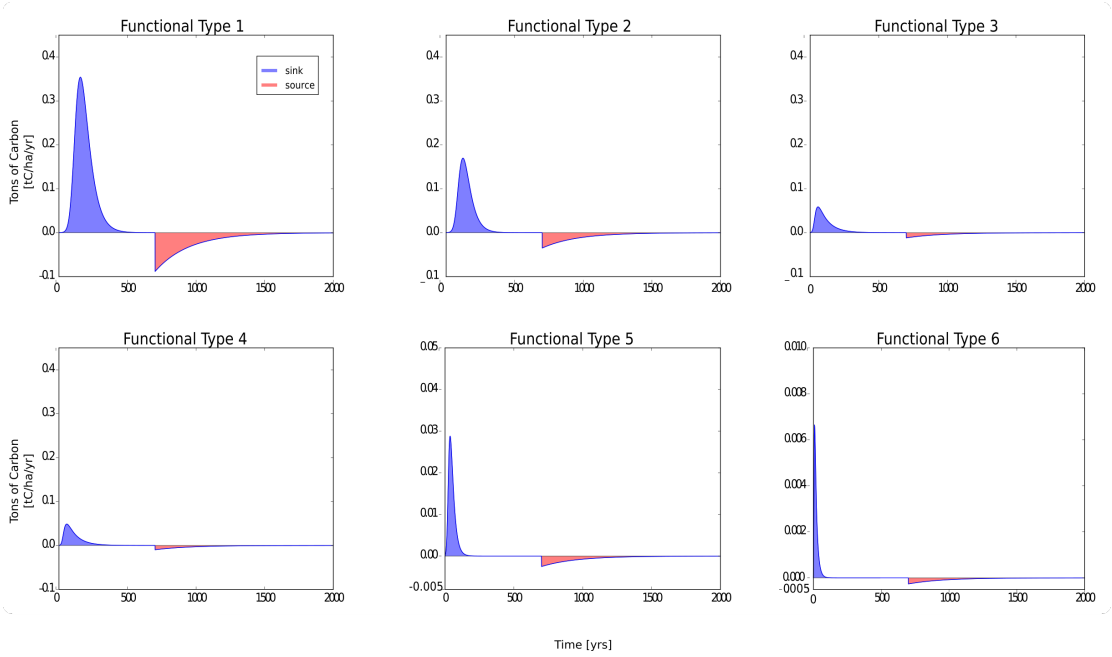

Figure 2.6: Carbon sequestration (blue) for single trees of different functional types and the respective emissions (pink) after death. Each plot show how one individual's sequestration capacity reaches a peak and decreases to zero. When the tree dies ($time = 700$), all carbon stored as biomass is eventually released. The y-axis limit was adjusted in order to increase readability of the PFT5 and 6 plots.

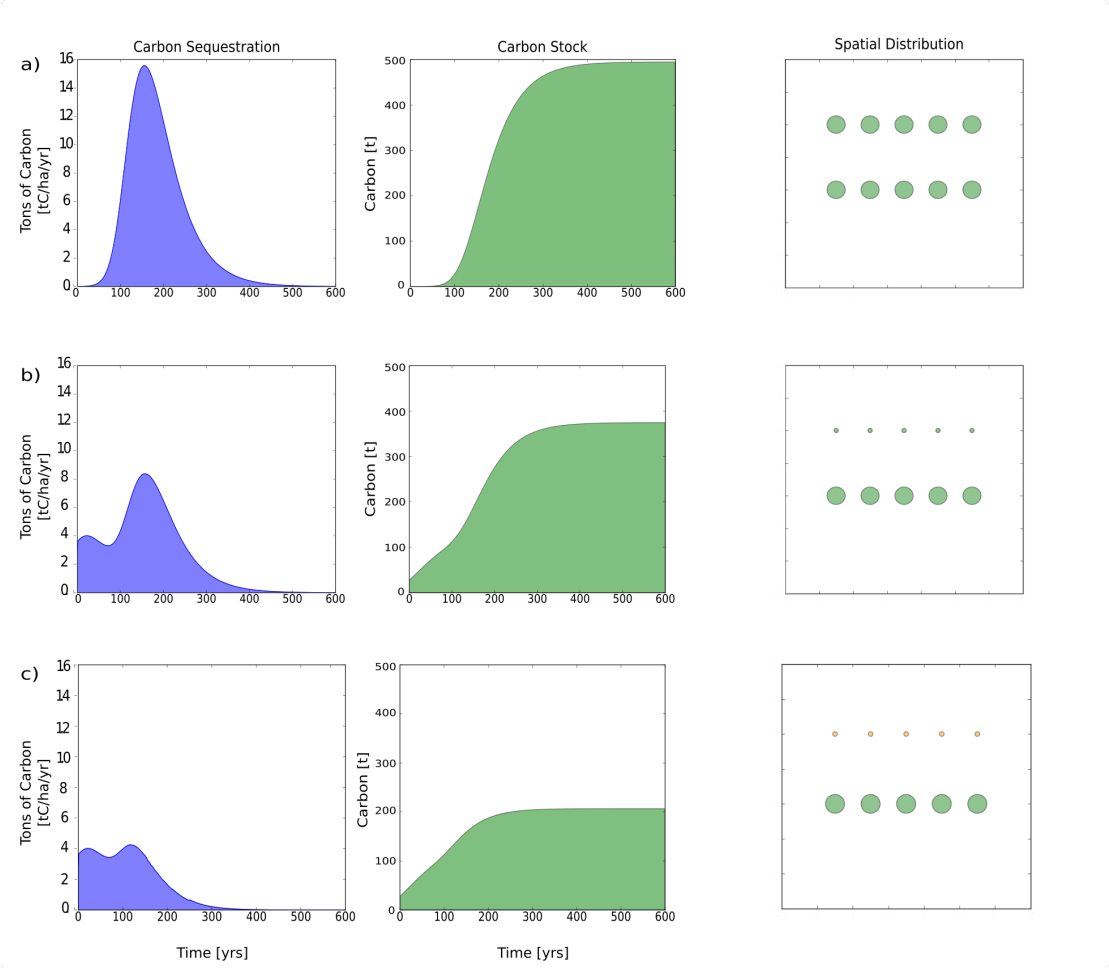

Figure 2.7: Influence of stand composition on carbon exchange and stocks. Stand (a) has 10 trees of PFT1 and same size, (b) has 10 trees of PFT1 and two initial sizes and (c) has 5 trees of PFT1 and 5 of PFT2, with two categories for initial size. Circle sizes indicate initial crown area. Each tree was planted on a different patch to exclude the effect of competition and highlight the influence of size and PFT on carbon sequestration.

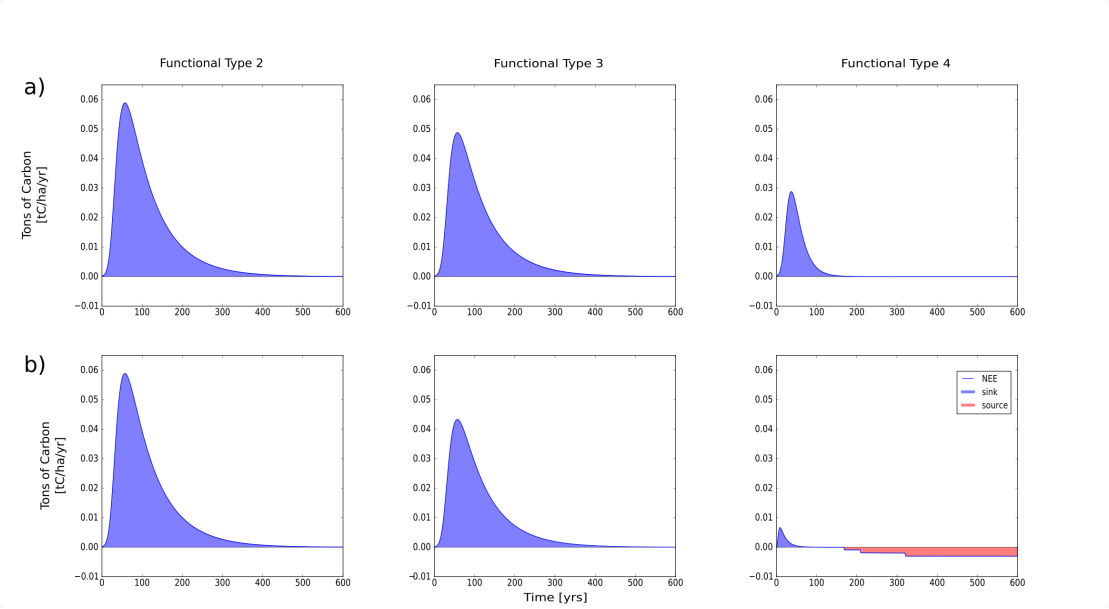

Figure 2.8: Effect of competition on individual carbon exchange. Three seeds of different PFTs were planted in separated patches (a) and in the same patch (b). In the absence of competition (a), each tree grows to its maximum size. When planted in the same patch (b), the tallest individual (PFT2) grows to its maximum size, the PFT3 individual shows a reduced sequestration and the smallest individual (PFT4) starts to act as a carbon source. Mortality was turned of in order to highlight the effects of competition. See table 2.3 for PFT parameter values.

2.3 New features

This section describes the main additions of my implementation in relation to FORMIND.

2.3.1 Object-oriented implementation

Object-oriented programming is a computer programming strategy in which the entities represented in a program are defined as *objects*, data-structures that can hold information in compartments called *attributes* and code organized in *methods*. Objects can exchange information and trigger the execution of specific methods on each other. These features make object-orientation specially suitable for individual-based models, with attributes representing individual characteristics and methods representing their behaviours. The concept of *classes* is also useful

to represent hierarchies common in biological systems. These are blueprints (*e.g.,* species, genus, etc) from which objects are created, and can be analogous to taxonomic or functional groups. In this model, a base *Tree* class is defined, which has all the tree characteristics and processes described in section 2.1.7. Since these are common to every type of tree, each PFT is represented by its own subclass, that inherit all the attributes and methods of the base tree, but declares its own type-specific parameters (*e.g.,* wood density). Each tree is then represented by an object of the corresponding PFT. Two other classes are included in the model: the *TreeWorld* class controls model execution (*e.g.,* by controlling the time) and manages data (*e.g.,* by writing model outputs to the database); the *TreeGrid* class represents space and holds environmental variables (*e.g.,* light).

This approach makes it straightforward to extend or modify the model for purposes other than those addressed in this project. The example below shows how to modify the default *Tree* class to create a new one with a different submodel for maintenance respiration (R_m). If the modeller has multiple PFTs defined as sub-classes of the *Tree* class, all of them will be changed. Selected PFTs can be modified just as easily.

```
# Makes the package containing the Trees model available
import trees_ibm

# Creates a custom tree class that inherits from
# the default Tree class in the package

class MyCustomTree(trees_ibm.Tree):
# At this point, the custom class is identical to the default class

    def __init__(self,position,world,dbh,age=0,id=None):
        # Adds two new parameter
        # These will be used in our respiration submodel
        self.rm_alpha=0.0001
        self.rm_beta=0.85
```

```
        # Passes the required parameters to the parent class
        super().__init__(position=position,world=world,dbh=dbh,id=id)

    #overwrites the original submodel for calculating maintenance respiration

    def calculate_rm(self):
            # Calculates the maintenance respiration using
            # the new parameters
            rm=self.rm_alpha*self.AGB**self.rm_beta
            return rm
```

The rest of the code does not need any modifications. Therefore, if a user would like to compare the effects of two competing submodels on the simulation results, for example, the rest of the code can be kept untouched.

2.3.2 Interactive visualization of model outputs

An interactive visualization tool is included in the package. It reads the simulation results stored in the database generated by the model to display a dashboard in a browser. The spatial distribution of the trees, DBH and age distributions, Net Ecosystem Exchange, total carbon stocks, population sizes for each PFT and the carbon flow between different stocks is shown over time in animated plots. The user can also pause the animation or select any specific time step. In addition, it is possible to zoom on a group of trees and see individual attributes (figure 2.9).

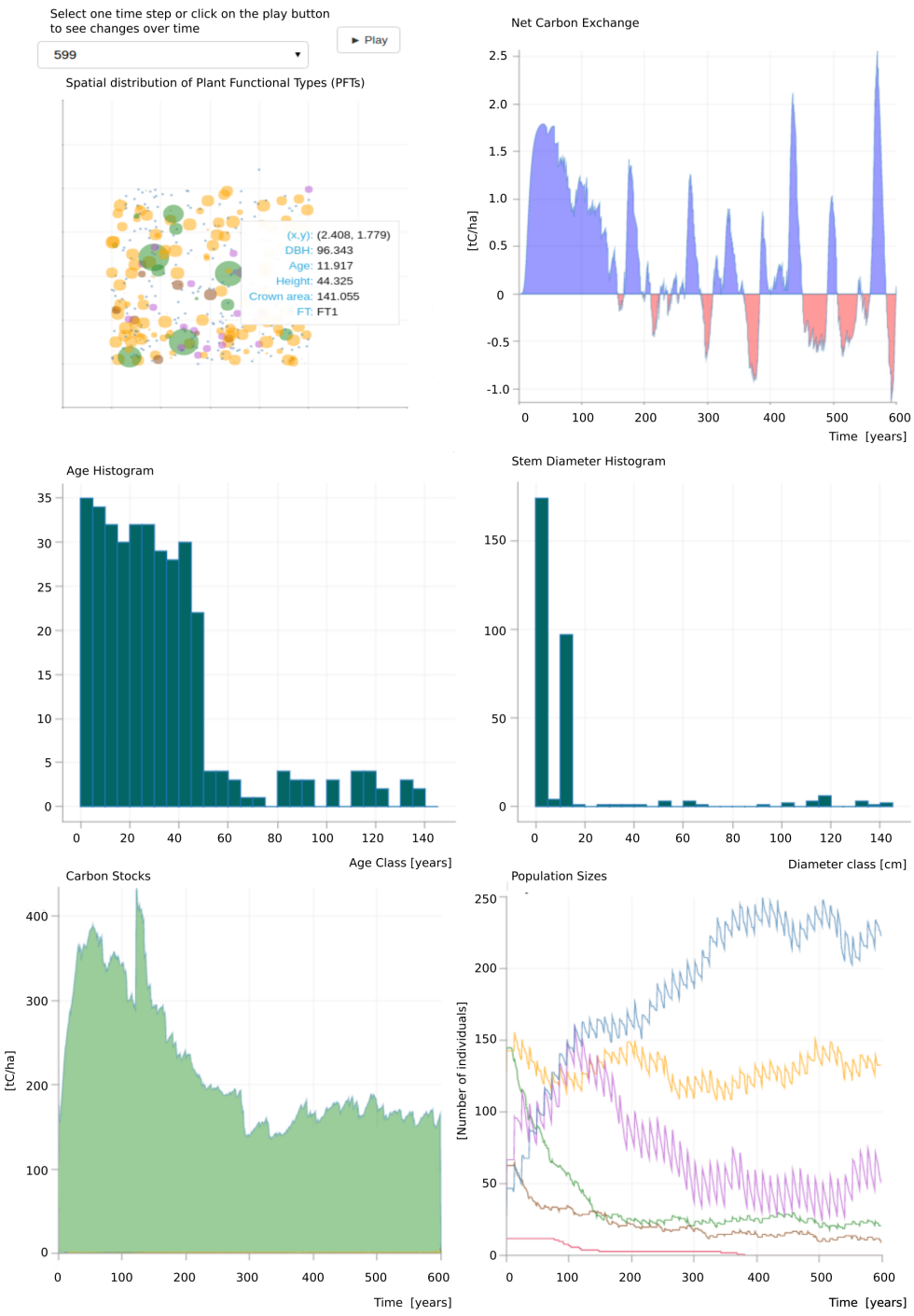

Figure 2.9: Example of output visualization. The top left plot shows the spatial distribution of trees, coloured by PFT. Extra information about one individual can be obtained by clicking on it. The top right plot shows landscape level carbon flow, with sink periods as blue and source periods as pink. The histograms in the middle show age and DBH distribution. The bottom left plot shows landscape level carbon stocks (only for living trees in this example). The bottom right plot shows population sizes for each PFT. Each plot shows a legend when the users positions the mouse over it. Specific time steps can be selected from the drop down box at the top. The play button starts an animation that shows progress over time.

2.3.3 Interactive PFT calibration tool

The use of Plant Functional Types is a flexible strategy that allows the representation of tree communities at different detail resolutions. Virtually, it is possible to create one class for each plant species in a community (or even more, to represent intraspecific patterns), however, in species-rich communities, functional types are often used to group species with similar characteristics (Fischer et al., 2016; Picard and Franc, 2003). The number and characteristics of the PFTs can strongly influence competition and forest dynamics (Köhler and Huth, 1998). It can be challenging to choose how many functional types to use and to estimate the parameter values that best represent field observations, specially when data is scarce or classification schemes for the same study site disagree (Picard et al., 2012). As pointed by Fischer et al. (2016) and Gourlet-Fleury et al. (2005), the issue of choosing a suitable number of PFTs for any specific research question is not resolved. The package developed for this project includes an interactive tool (figure 2.10) that aims to facilitate the determination of type-specific parameters based on commonly measured variables, such as maximum DBH or maximum height. Similarly to the visualization tool described above, it runs on a browser and allows the user to visualize how each parameter affects the selected patterns.

2.4 Example

To illustrate the use of this model, two scenarios were run. Both simulate 1 ha of forest composed of trees of 6 different PFTs (Table 2.3). To examine the effect of logging on carbon stocks, a harvest operation was executed at time 0 in one of the simulations, removing 30 cubic meters of timber. The simulations ran for 50 years (600 months) and, except for the logging event, were identical. Before these simulations, the model ran for 500 years starting from an empty landscape, in order to provide a mature forest to serve as basis for this example. The same

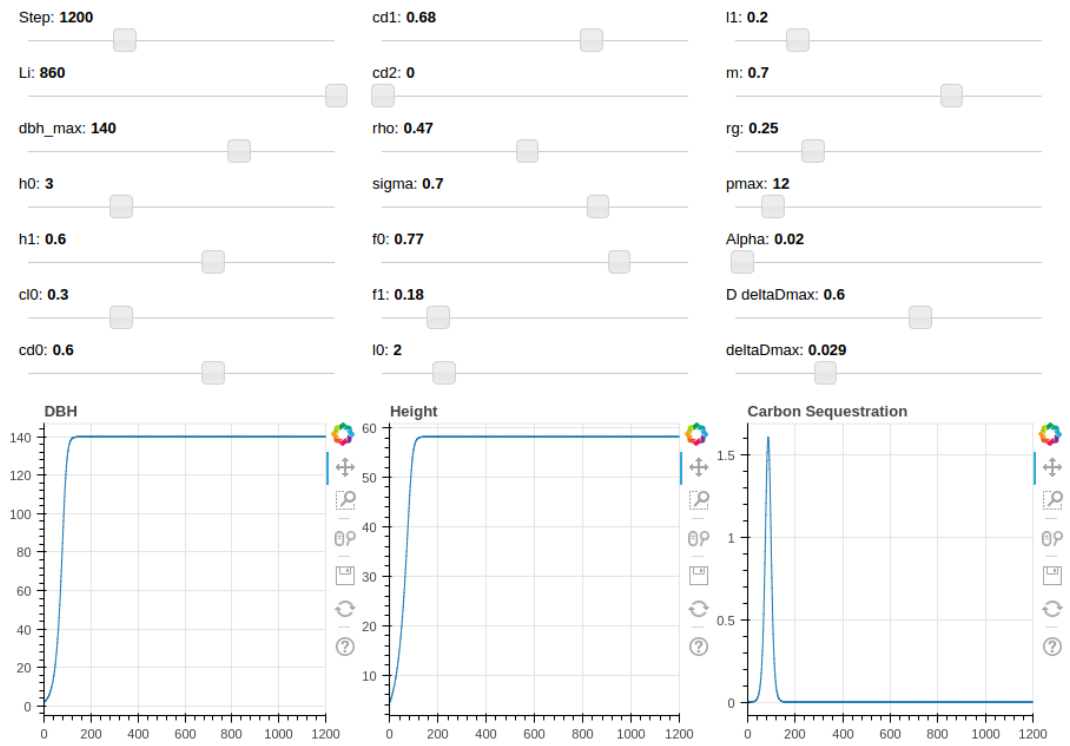

Figure 2.10: Interactive PFT calibration tool. Carbon sequestration, Height and DBH for one tree are updated as parameter values change. This tools is useful to explore the relationship of between PFT level parameters and characteristics of the trees. Parameter values can be exported and used to create PFTs in the model.

random seed was used to guarantee the same outputs from stochastic processes.

The undisturbed scenario presented lower carbon sequestration than the selectively logged forest during certain periods of time. This happens because the gaps open by logging give space for new trees to grow, increasing carbon sequestration. As time passes, succession takes place in the disturbed area and the initial addition to the total absorption fades away. By the end of the simulation (50 years after cut), the value was nearly twice as large for the undisturbed scenario (figure 2.11). Given the natural oscillations in net carbon exchange, the impacts of logging are more evident on the carbon stocks (figure 2.12), where after 50 years, the logged forest stored almost 200 tC less than the undisturbed forest. Finally, figure 2.13 shows that both the age structure and the stem diameter distribution are slightly

different in both scenarios.

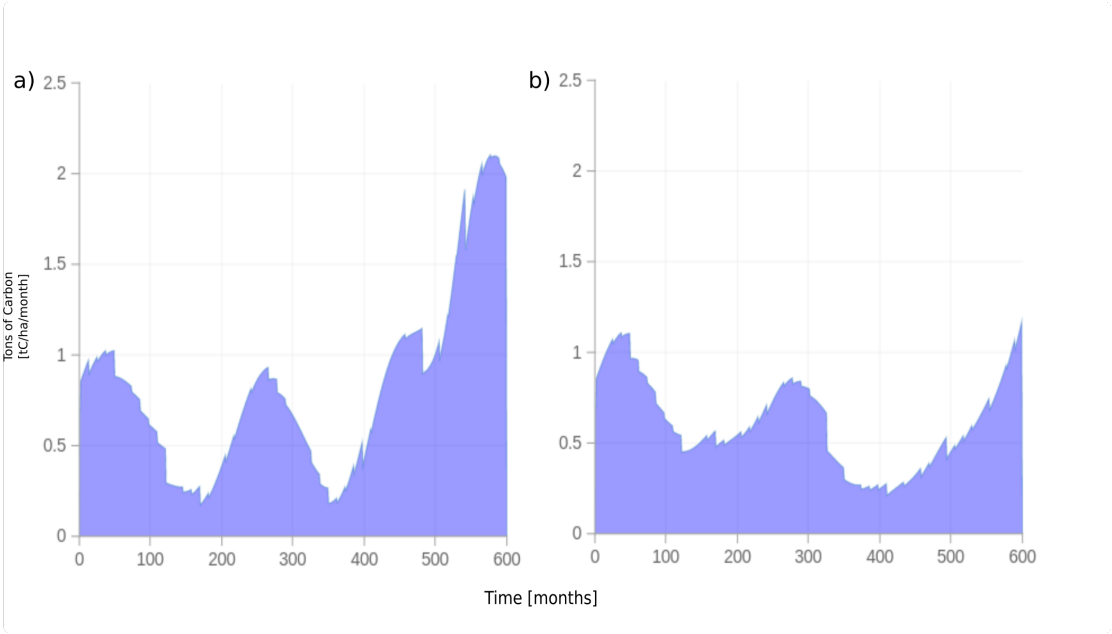

Figure 2.11: Net carbon exchange in undisturbed (a) and selectively logged (b) forest. After 50 years, the amount of carbon absorbed by the logged forest was approximately 50% of the value absorbed by the undisturbed forest.

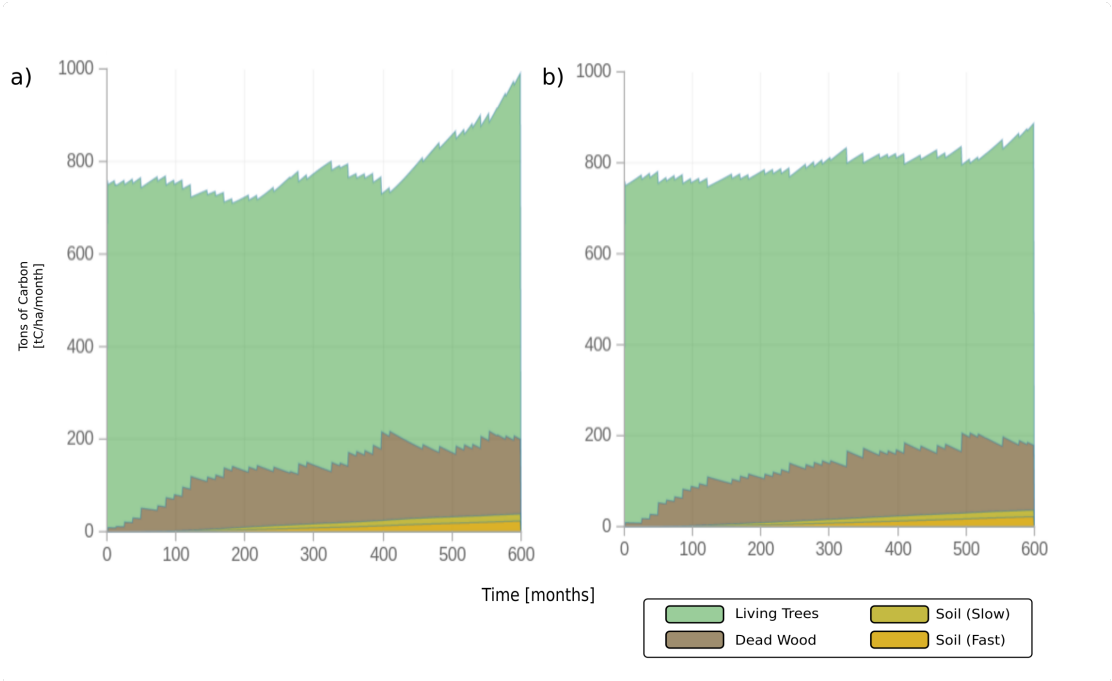

Figure 2.12: Carbon stocks in an undisturbed (a) and selectively logged (b) forest. The undisturbed forest stored nearly 200 tC more than the disturbed forest 50 years after the logging event.

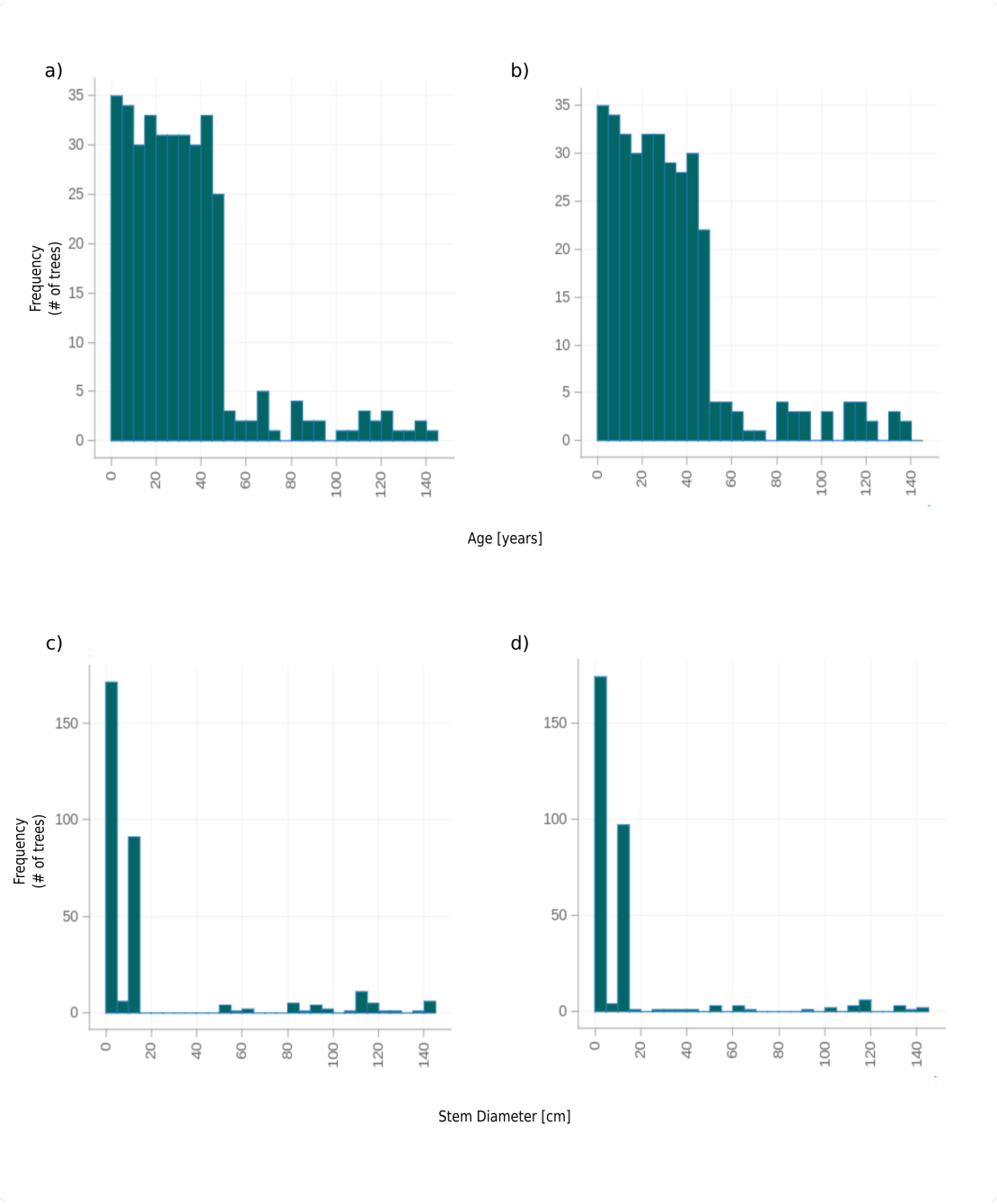

Figure 2.13: Age and stem diameter distributions after 50 years simulation in undisturbed (a and c) and selectively logged forest (b and d).

Although this example used parameter values based on the calibrations performed by Fischer et al. (2015) and Rüger et al. (2007), the purpose here is only to illustrate the model outputs and its possible use for tracking the impacts of logging and other forms of disturbance. As shown in section 2.2.3, carbon exchange varies

with the properties of functional types and calibration is crucial for an accurate representation of specific study sites.

3 Chapter 3: Seed Dispersers model

3.1 Model Description

3.1.1 Purpose

The purpose of the model developed in this chapter is to simulate seed dispersers at an individual level, using a decision algorithm based on energy levels. The model is spatially explicit and the dispersal areas result from the movement patterns of the dispersers. Agents move from tree to tree, mimicking arboreal animals such as birds and monkeys. The model can be coupled to the Trees model described in the previous chapter to link disperser movement and population dynamics to the dynamics of the tree community and the associated carbon stocks. It can also be used independently to model the relationships between animal movement and behaviour and seed dispersal.

3.1.2 Entities, State Variables and scales

The model is comprised of seed dispersers, trees and a landscape divided into patches. The patches and trees are derived from those described in Chapter 2 in order to facilitate integration. Only a subset of the attributes and methods in the Trees model is relevant to the dispersers model and the landscape simply provides a bi-dimensional coordinate system for movement. The patches include *resistance* as an additional attribute that increases the difficulty of moving through areas without trees (see the Foraging and Roaming submodels described in section 3.1.6). The Trees model only needs to provide fruit availability (see eq. 1), position and functional type of trees. The disperser model can obtain the necessary information directly from the Trees model, from files or user-defined functions.

The disperser agents include an attribute *energy* that keeps track of energy levels. During each time step, individuals can gain or lose energy depending on

their activity. Position is also recorded at each time step and is the centre of the current tree. Although the spatial scale is the same as the Trees model, the temporal scale is much smaller, with each time step representing 30 min of the activity.

3.1.3 Process Overview and Scheduling

The model assumes an activity period of 12 hours, which is represented by 24 time steps. Only the portion of the day in which agents are active is simulated. Figure 3.1 illustrates how agents choose their actions. There are three activities a disperser can choose from: feeding, resting and moving. Agents gain energy through the feeding activity and lose energy when moving and, to a lesser extent, resting. Additionally, the activity defecation is selected at the end of each time step. When fruits are ingested, they pass through the disperser's digestive system and, after a period of time t_{gut}, become ready to be expelled. The frequency with which individuals deposit seeds is therefore dependent on gut passage time (t_{gut}) and on the amount of food ingested. If and individual does not ingest any food, it will not defecate. All individuals use the same decision algorithm to decide what action to take (see figure 3.1). All choices are made according to the individual's energy level and how it compares to two thresholds (e_{level1} and e_{level2} on table 3.1). If the energy is below e_{level1}, the agent is hungry and will forage or eat; if the energy is above e_{level2}, agents may choose to rest or roam without targeting feeding trees (see section 3.1.6 for details on the different kinds of movement). This approach is based on the strategy adopted by Bialozyt et al. (2014a). Animals can have several incentives to move, including predator avoidance or the search for food, mates and other resources (Boyer et al., 2006b; Gursky-Doyen and Nekaris, 2007). This model focus on search for food as it does not include predators or explicit reproductive encounters. Energy is only used for the individual's daily activities as described above. In reality, animals allocate their energy to growth,

reproduction and maintenance (van der Vaart et al., 2015), but in this model growth is not simulated and reproduction happens at the population level (see Population growth in section 3.1.6), so the two energy thresholds simplify energy use to maintenance, but still allow individual decision making. This approach has been used by Bialozyt et al. (2014a) to successfully approximate the movement and seed dispersal patterns of two tamarin species (*Saguinus mystax* and *Saguinus nigrifrons*). The model also keeps track of months and years, which is useful to schedule reproduction events. Seed dispersal models like this are normally used to simulate short periods of time (30 days) because that is enough to obtain seed deposition patterns (Bialozyt et al., 2014a; Boyer et al., 2006b; Caughlin et al., 2015). As a consequence, they do not include reproduction. But if long-term impacts are to be modelled, some sort of population dynamics model becomes necessary. This model includes a simple energy-based population growth submodel that allows longer periods of times to be simulated (see section 4.2 for an example and section 5.1 for a list of hypotheses that would require this feature).

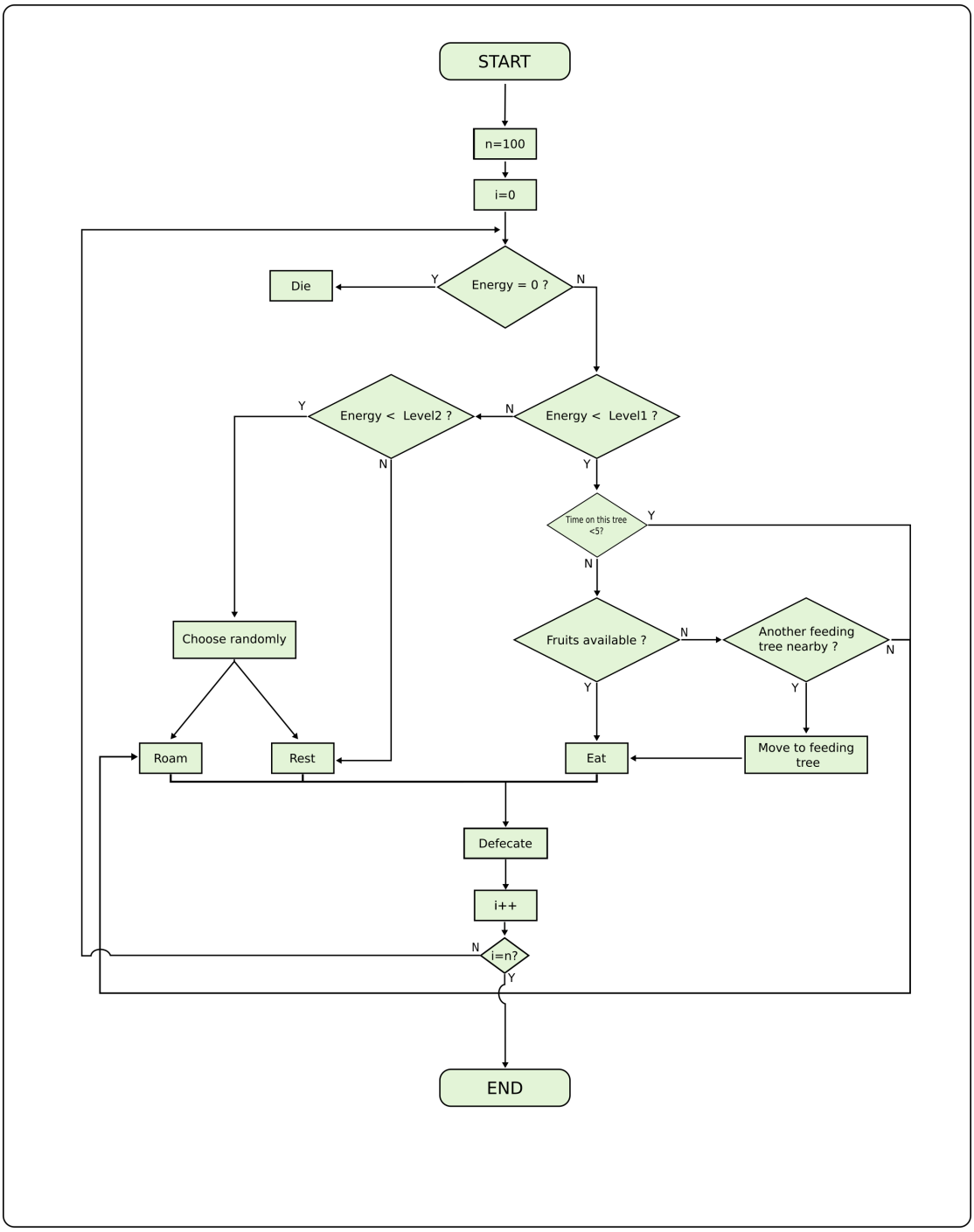

Figure 3.1: Internal decision algorithm used by seed dispersers. Agents choose their actions based on two energy thresholds. The algorithm is repeated until the defined number of time steps (n) is reached. The current iteration is indicated by i.

3.1.4 Design Concepts

- *Basic Principles*: I assumed that the main incentive for dispersers to move is the search for food. Although anti-predation behaviours and mating are important movement incentives (Boyer et al., 2006b; Nathan et al., 2008), these were not included in this model.

- *Emergence*: Since each individual's movement step is determined by the decision algorithm, the home range and average daily path length at the end of the simulation period emerge from the agents behaviour. Seed dispersal distances and kernels reflect movement patterns. Furthermore, the spatial distribution of trees and fruit abundance will also change model outputs.

- *Adaptation*: Agents alter their behaviour in response to their internal energy and the available energy (in form of fruits) in their surroundings. The decision process is guided by their energy levels and two thresholds depicted in figure 3.1. If the energy level falls below 0, the individual dies from starvation.

- *Objectives*: The main objective of the seed dispersers is to maintain their energy level above e_{level1}. Below this threshold, agents feel hungry and search for trees with fruits. In addition to foraging, animals engage in a variety of activities during their daily routine. In order to represent these in the model, agents can rest and move without the goal of finding food if energy levels are above level 2. The actual dispersal of seeds is unintentional and results from foraging behaviour and regular defecation.

- *Sensing*: At any given time step, agents are assumed to know the trees within a radius (parameter r, action radius) of their current position. They also know the number of fruits in each tree and their distance from the current position. In this version of the model agents are assumed to know these

values accurately, which can be interpreted as having a good knowledge of their territory. Alternatively, noise can be added to these values in order to represent an imperfect assessment of the environment. This information is used by submodel *Foraging* to calculate the value for each tree, which will determine the individual's movement. These simplifications abstract the complex cognitive processes that animals use to map their environments (Janson and Byrne, 2007), but similar assumptions have produced realistic movement patterns (Boyer et al., 2006b).

- *Collectives*: Although the agents in this version of the model represent individuals, they can also represent collectives (groups of dispersers), in which case the relevant parameter values should be adjusted (*e.g.,* maximum number of fruits).

- *Stochasticity*: After their energy reaches e_{level2}, the choice of activity is random (Bialozyt et al. (2014a) found that the stochastic selection of activities after satiation better approximated the behaviour of their observed primate species). Stochasticity is also involved in the mortality submodel, the position of new agents resulting from the population growth submodel and the position of dispersed seeds (both those dispersed by the agents and those dispersed by the trees' generic dispersal methods). Details about these submodels can be found in section 3.1.6.

- *Observations*: The location and energy of every disperser is recorded in each time step. Regarding the seeds, their initial and final location are recorded, alongside with the identification number of the parent tree and the mode of dispersal. The trees can have a dispersal method that is independent of the seed dispersers (see submodel *Generic seed dispersal* and *Recruitment* in the Trees model), representing all other dispersal means (*e.g.,* wind, other seed dispersers that are not focal to the simulation and therefore are not

explicitly represented, etc). The model includes functions to calculate four outputs:

- Home range: The home range for each individual is estimated as the convex hull of all the points visited during the simulation. The output is the average of the home ranges of all agents.
- Path length: The euclidean distance between initial and final position is recorded for every movement. This output is the average value for all movement steps performed by all agents during the simulation.
- Seed dispersal distance: The average euclidean distance between the parent tree and the position where each dispersed seed was deposited.
- Dispersal area: The dispersal area (or seed dispersal) for each tree is estimated as the convex hull of all the seeds dispersed during the simulation (calculated with the coordinates of each seed dispersed by an animal). This output is the average area for all trees.

These summary outputs are the main metrics used throughout this chapter. Distances and areas are measured by counting grid cells. For all simulations, each cell has a side of 20 meters (and area of 400 m^2). All data used to compute model outputs are recorded in a database (HDF5) and the seed data can also be saved separately in a file (.json file), which can be used as input for the Trees model described in chapter 2.

3.1.5 Initialization

The distribution of the trees within the landscape is read from a file (alternatively it can be randomly generated at the beginning of a simulation). The total number of dispersers can be defined prior to the simulation to reflect the densities found in the field (Levi and Peres (2013) and Stevenson et al. (2000) report primate

densities around one individual per hectare, which I adopt in the examples). The individuals are assigned an initial tree randomly. When the population growth feature is used, the new individuals will be randomly assigned to available trees. Initial values of the parameters described in table 3.1 are also set at the beginning of the simulation. The model supports the creation of new types of trees and dispersers. All examples in this chapter and the next only used functional types for trees (similar to those described in Chapter 2), but groups with different parameters can also be created to represent multiple disperser functional types.

Table 3.1: Model parameters and descriptions for the Dispersers model. Default values are based on the model described in Bialozyt et al. (2014a)

Parameter	**Description**	**Unit**	**Default value**
r	action radius	m	50
e_{level1}	energy threshold 1	energy units	80
e_{level2}	energy threshold 2	energy units	150
e_{feed}	energy gained per fruit	energy units	8
e_{rest}	energy lost per resting step	energy units	0.7
e_{travel}	energy lost per travel	energy units	1.6
r_p	resistance penalty	-	1.2
w	width of defecation area	grid cells	2
e_r	minimum energy for reproduction	energy units	60
c_1	cost to produce 1 individual	energy units	70
c_2	energy used per individual	energy units/ individual	840
t_{max}	maximum time on the same tree	time steps	5
m_b	background mortality	-	0.03
t_{gut}	gut passage time	time steps	2

3.1.6 Submodels

- *Foraging*

  When energy levels are below e_{level1}, agents will search for food. The equation below is used to calculate the value of all trees within a radius r. This assessment is inspired by the model developed by Boyer et al. (2006b) and reflects an effort to minimize energy waste. When foraging, agents will prefer

the closest tree that has the most fruits. Agents also consider the landscape resistance. Patches without any trees represent a greater energy cost to cross and are avoided (see figure 3.2 for an illustration). Considering these factors, the disperser calculates the value of each tree as

$$value = \frac{n_{fruits}}{distance + resistance \cdot r_p}, \tag{42}$$

where *distance* is the euclidean distance between the tree and the agent and *resistance* is the sum of the resistance value for each patch between the tree and the agent. All patches with at least one tree have resistance equal to 0 while those without any trees have resistance equal to 1. The constant r_p, resistance penalty, adjusts the added cost of empty patches.

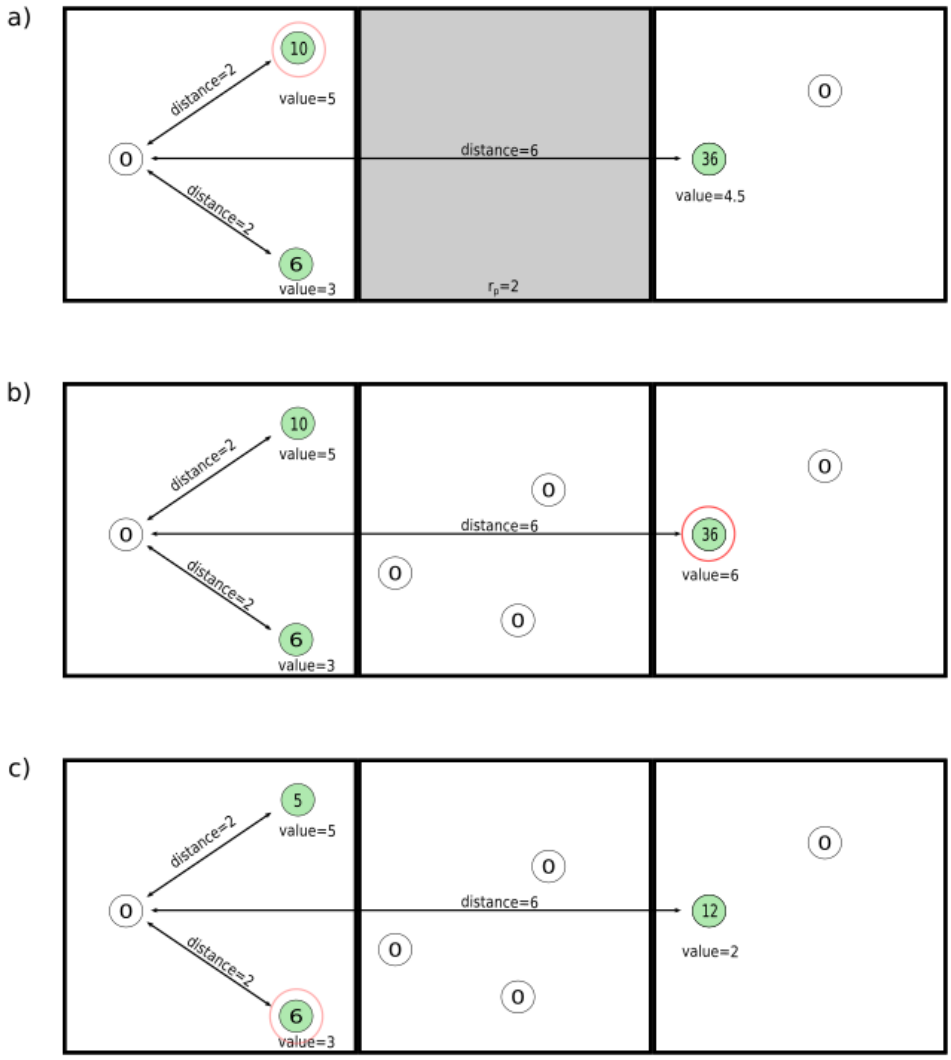

Figure 3.2: Tree values as perceived by a foraging disperser. Circles indicate trees and number within show the fruit availability. A foraging agent will select a different tree (indicated by the red circle) depending context. In a) the empty middle patch adds a resistance cost to the furthest tree, reducing its value. If the patches were not isolated (b), that tree would be selected because trees that are closer have less fruits available (0, 6 and 10). The depletion of resources (c) in one tree (*i.e.,* by competing agents) may also alter the destination choice.

- *Roaming*

  If individuals are not hungry, they might still move without the goal of obtaining food. In this travel mode, the fruit availability is not considered, so individuals may use non-fruiting trees as well. Distance and landscape

resistance are still considered:

$$value_{roaming} = \frac{1}{distance + resistance \cdot r_p}. \tag{43}$$

- *Defecation*

  Defecation happens at the end of each time step and the seeds that have passed through the digestion process (if any) are expelled. Since a time step represents a period of 30 min, seeds are deposited between the current position and the disperser's position at the beginning of that time step. The area of possible deposition is represented by a rectangle of length equal to this distance and with a width defined by parameter w (see figure 3.3). Given the 30 min temporal resolution in this model, this approach relaxes the assumption that dispersers move linearly from one tree to the next and considers that individuals might have moved within the rectangular area in the last 30 min.

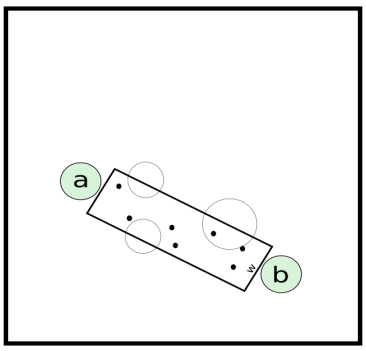

Figure 3.3: Seed deposition between two trees. Points represent location of seed within the possible rectangular area between tree a and tree b. Empty circles represent other trees.

- *Generic seed dispersal* In order to account for seeds that are dispersed by means other than the modelled dispersers (*e.g.,* wind), each tree type can

have a generic dispersal method defined as a power distribution (eq. 2). A proportion gen_{dis} of each tree's seeds are dispersed according to that distribution, leaving the rest available for the disperser agents to consume. A value of 1 indicates that the tree is exclusively dispersed by other means.

- *Population growth*

  Reproduction is not explicitly modelled as a behaviour of the agents. Instead, I use a simple population growth model based on the classic Verhulst logistic model. The maximum population size (*i.e.,* the carrying capacity K) is limited by the amount of fruits available in the landscape (eq. 1 describes how fruits are produced in the Trees model.) and potential growth depends on the total energy reserves in the disperser population (E_r). Only the energy of individuals above a threshold is counted towards the total,

$$E_r = \sum_{i=1}^{N} \left\{ \begin{array}{ll} e_{d_i} - e_r, & \text{if } e_{d_i} > e_r \\ 0, & \text{else} \end{array} \right\}, \tag{44}$$

  where N is the number of dispersers in the population, e_{d_i} is the energy level of the i$^{\text{th}}$ disperser and e_r is the minimum energy required for reproduction. I assume that population level reproduction is related to E_r, with the number of births given by

$$\Delta N = \frac{E_r}{c_1} \cdot \Big(\frac{K - N}{K}\Big), \tag{45}$$

  where c_1 is the energy cost to produce a new individual. The carrying capacity K calculated as,

$$K = \frac{n_{fruitsL} \cdot e_{feed}}{c_2}. \tag{46}$$

In the equation above, $n_{fruitsL}$ is the number of fruits available in the whole landscape, e_{feed} is the energy gained by each fruit consumed and c_2 is the average energy used by a disperser between reproduction periods. Parameters c_1 a c_2 relate to the average energy used by an individual throughout the year. In the examples presented in this and the following Chapter, the average energy was around 70, which is the value used for c_1, while c_2 was set to 840 (12*70). However, these parameters can be calibrated using census data for specific species/study areas. The total energy cost used for reproduction is equally divided and subtracted from the energetic reserves of all individuals who contributed to the total energy E_r.

- *Mortality*

  Agents can die of starvation (if their energy levels falls below 0) and are also subject to stochastic mortality with probability mb. Once a year, a random number is drawn from a uniform distribution for each agent. If that number is $\leq \mathrm{m}_b$, the agent is removed from the population.

3.2 Details of implementation

The Dispersers model was also written in the Python 3 programming language, taking advantage of object-oriented programming (see figure 3.4 for the class diagram). The individuals are represents by instances of the Dispersers class. Each activity is defined as one method and the decision algorithm illustrated in figure 3.1 is implemented by the *schedule* method, which defines the order and conditions for the activities to be executed. The current decision algorithm is designed

to capture the behaviour of arboreal primary-dispersers with a level abstraction enough for it to be used to represent a variety of species. However,the code design makes it easy to modify the decision algorithm and make it more specific to a given species or group of species. Simple modifications (*e.g.:* changing the order of activities) can be done by creating a subclass of the Dispersers class and modifying only a few lines of code in the *schedule* method. More complex modifications can be done by adding new activities (*i.e.:* new methods in the subclass) or modifying the existing ones.

In addition to the Dispersers class, a SimpleTree class is also included. This is a simplified version of the the trees described in Chapter 2 and were included to allow the Dispersers model to be used independently of the Trees model. If the user does not need to simulate the population dynamics of trees to address the question, only the Dispersers package needs to be installed. However, the code was designed to allow the integration with the Trees model (more details on Chapter 4).

Figure 3.4 contains the UML (Unified Modelling Language) class diagrams illustrating the internal organization of the *dispersers* package (see Chapter 5 and table 5.1 for details on the packages.)

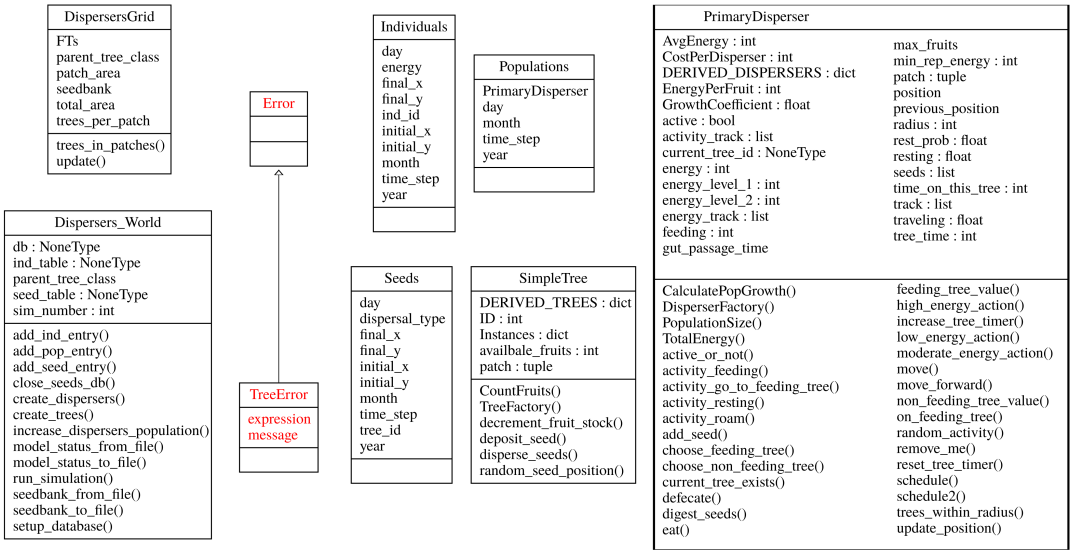

Figure 3.4: Class diagram for the Dispersers package. Following the UML (Unified Modelling Language) standards, each class is represented by a block divided into three sections: the class name at the top, followed by the corresponding attributes in the middle and the methods at the bottom.

3.3 Influence of main parameters

This section explores how model outputs (described in the *Observations* section) are affected by the main individual parameters.

As the amount of energy lost during resting periods gets closer to zero (fig. 3.5), the path length and dispersal distance slightly increase. Home range is not affected, while dispersal area decreases for values close to zero. All outputs were affected positively (with two to four-fold increments) as the radius action parameter, r, increased. This can be explained by the way the radius limits the movement possibilities, as only trees within that distance are perceived and evaluated by the dispersers. This was the only parameter to affect the home range, which showed large variations but no increases or decreases for other parameters. Given enough time and energy, agents are likely to explore the totality of their territory since resources were uniformly distributed in these simulations and there

was no shortage of fruits. The average path length and dispersal distance increase approximately 1.5 times as the energy gained from each fruit (e_{feed}) varied from 0 to 10. This relationship followed a saturation curve and fits the expectation that agents need to spend less time feeding and have more time free to roam their territories. Similarly, figure 3.6 shows that these values also increased as the costs of travelling (e_{travel}) decreased although the change in magnitude was less and the relationship more linear. As the energy costs associated with travelling and resting approached zero, the dispersal area was reduced. The dispersal distance was slightly reduced as the threshold value e_{level1} increased, but this parameter did not affect other outputs.

Figure 3.7 illustrates how the logistic population growth submodel presented in equations 44 to 46 responds to resource availability. The carrying capacity is set by the total energy available in the landscape. For this experiment, an initial population of 10 dispersers was placed in a 100ha landscape. For each scenario, only a percentage (20-100%) of the trees had fruits. For the purposes of this experiment, fruit production was fixed to 15 fruits per month throughout the year because the goal here was simply to demonstrate how the population growth submodel works. In experiments where a complete forest model is used to simulate the trees (like in the example presented in Chapter 4), the number of fruits produced per tree is calculated using equation 1. Therefore, heterogeneous fruit distributions might emerge, according to the aspects of the tree community (*e.g.,* abundance and spatial distribution of each PFT, tree size and time of the year). Consequently, the total available energy in the landscape and the carrying capacity fluctuate according to the tree community dynamics.

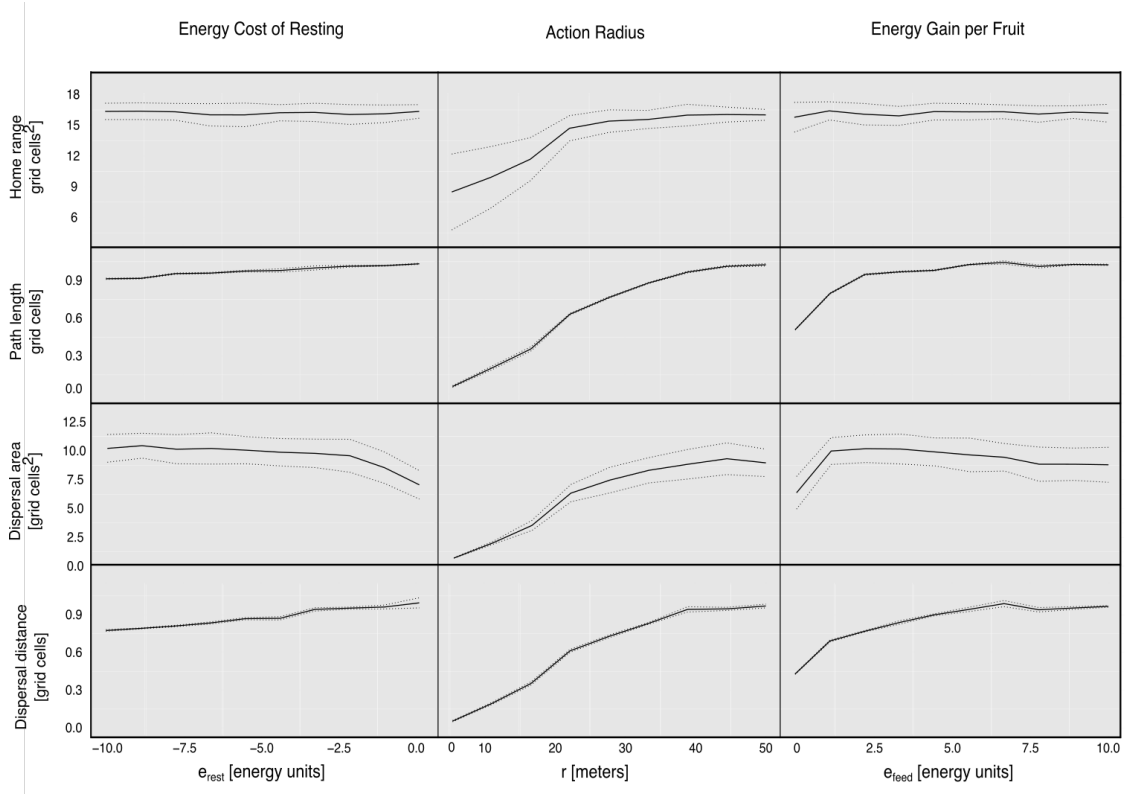

Figure 3.5: Effects of energy costs associated with resting, action radius and energy gains per fruit on animal movement and seed dispersal metrics. Bigger action radius result in higher seed dispersal and animal movement metrics. Higher energy gains per fruit result in similar effects. Changes in the amount of energy lost when resting only affect the metrics slightly, decreasing dispersal area as the cost of resting approaches zero and increasing all others by small quantities.

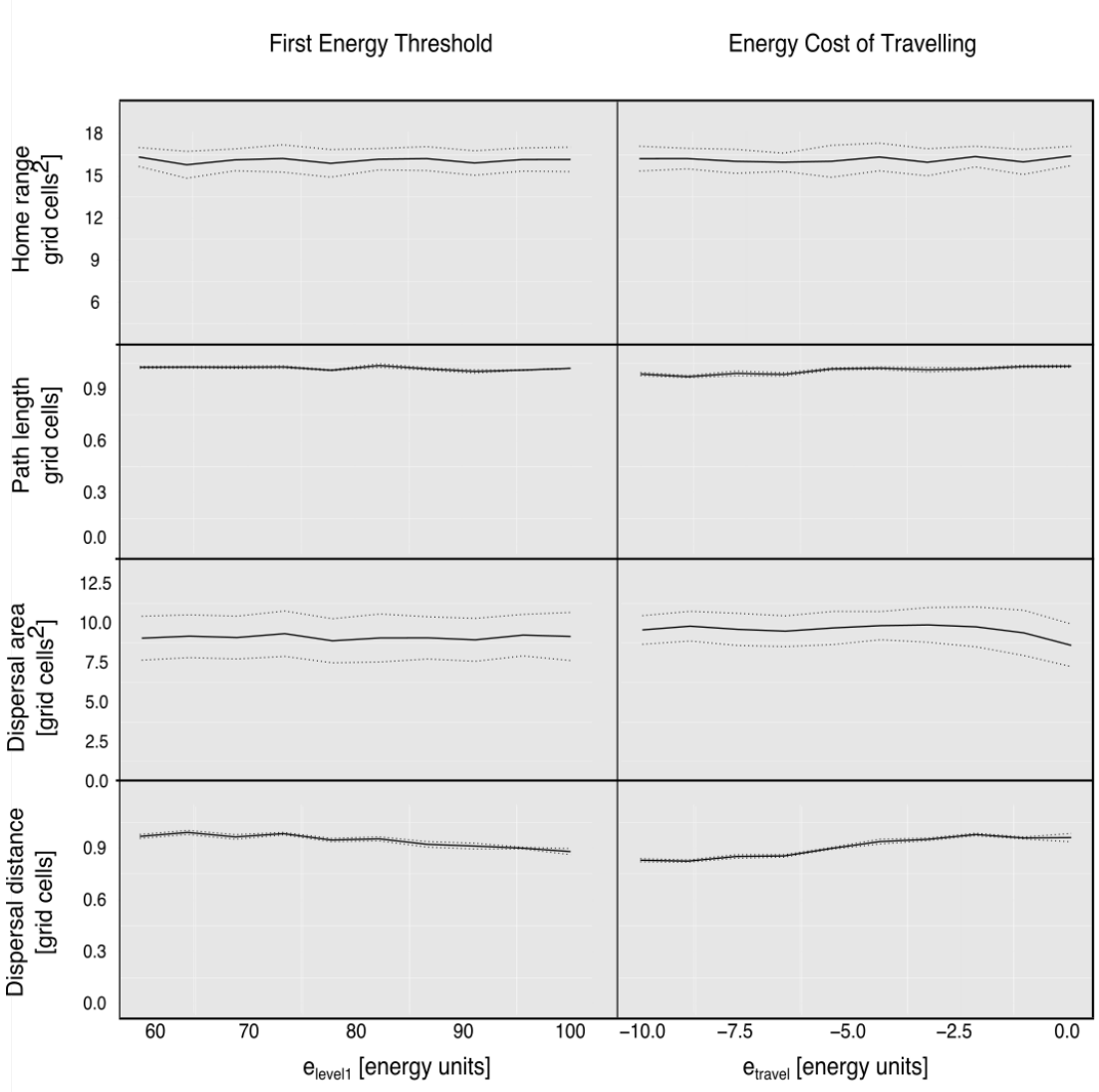

Figure 3.6: Effects of main parameter values on animal movement and seed dispersal metrics. Increasing the first energy threshold and decreasing the cost of travelling only affect seed dispersal and movement metrics slightly.

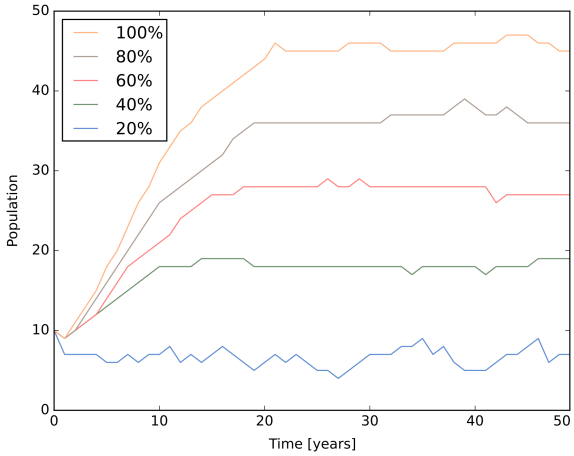

Figure 3.7: Resource availability limits population size.

3.4 Examples

3.4.1 Effects of defaunation on seed dispersal

The decline of animal populations due to human activities is a widespread and yet cryptic form of environmental degradation (Galetti and Dirzo, 2013). Different from fragmentation, deforestation and other consequences of land-use change, defaunation (the human-driven extinction or extirpation of large and medium vertebrates) is less evident from direct observations and often considered an invisible threat (Phillips, 1996). The loss of such animals affects the ecosystem functions they perform, such as seed dispersal (Vidal et al., 2013b). One application of the dispersal model is to investigate the effects of defaunation on seed dispersal. The simulation described below used a landscape of 50 ha in which trees were uniformly distributed. Half of the trees had fruits (5 fruits each) and at the beginning of each simulation the corresponding number of dispersers was randomly distributed among the all trees. The simulation period was 10 days.

All model outputs drop drastically at low densities (approximately 15 dispersers). Home range increased 60% until this threshold, with path length and dispersal distance showing the same pattern with less evident increases (fig. 3.8).

This can be explained by scramble competition, which increases the per capita resource availability and allows individuals to reach their satiation threshold e_{level1} quicker, giving them more time to roam their territories. Although these summary metrics only show a significant drop after a threshold, the total number of dispersed seeds declines linearly with the population (fig 3.9). This is likely to have strong negative effects on the establishment rates of those trees that rely on zoochory, favouring wind-dispersed species.

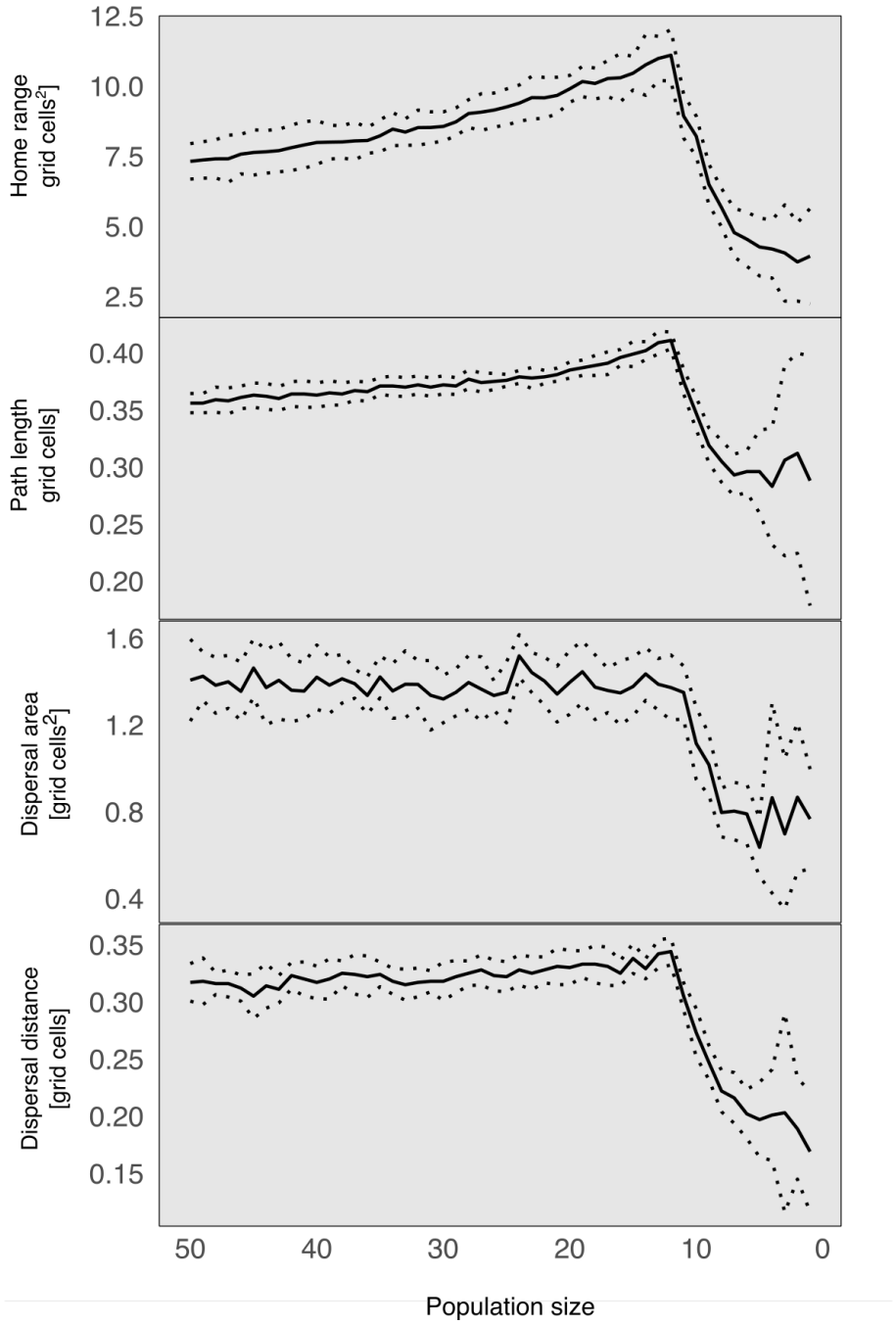

Figure 3.8: Effects of disperser densities on animal movement and seed dispersal metrics. Low disperser densities decrease the seed dispersal and animal movement metrics. As population densities are reduced, home range and path length slightly increase until the number of dispersers reaches 12 individuals. At this point, home range, path length, dispersal distance and dispersal area all drop significantly.

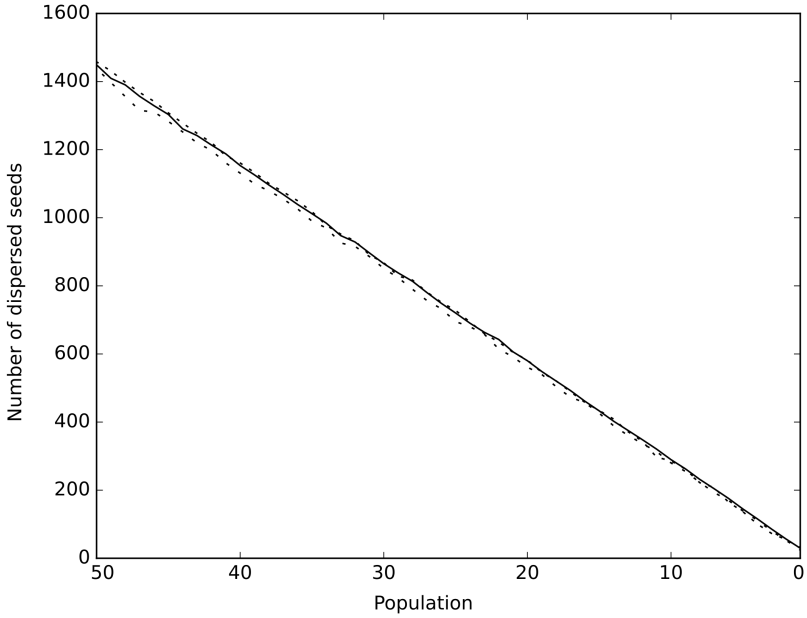

Figure 3.9: Linear effect of population size on the number of dispersed seeds. The deforestation gradient goes from 100 ha of forest to 10 ha. For the isolation gradient, each of the two fragments has an area of 8 ha and the total area is kept at 16 ha (two fragments) for all levels of isolation. The total area in the fragmentation gradient is also 16 ha and kept constant for all levels of fragmentation (number of fragments). The distance between adjacent fragments is kept constant (100 m).

3.4.2 Effects of degradation and fragmentation on seed dispersal

Habitat degradation and fragmentation impose limitations on seed dispersal and can also affect post-dispersal demographic processes, such as germination and seedling establishment. This section explores the effects of deforestation and fragmentation on the model outputs through three gradients. The first gradient (fig 3.10 a) is simply the reduction of the forested area, representing deforestation. The other two gradients account for two aspects of fragmentation: isolation, for which the distance between two fragments of equal area is gradually increased (fig 3.10 b) and the number of fragments, which kept the total area and distance between any two adjacent fragments constant. The initial forested area in gradient **a** was 100 ha, while the total forested area in **b** and **c** was 16 ha. The distance

between adjacent fragments in the fragmentation gradient was 100 m.

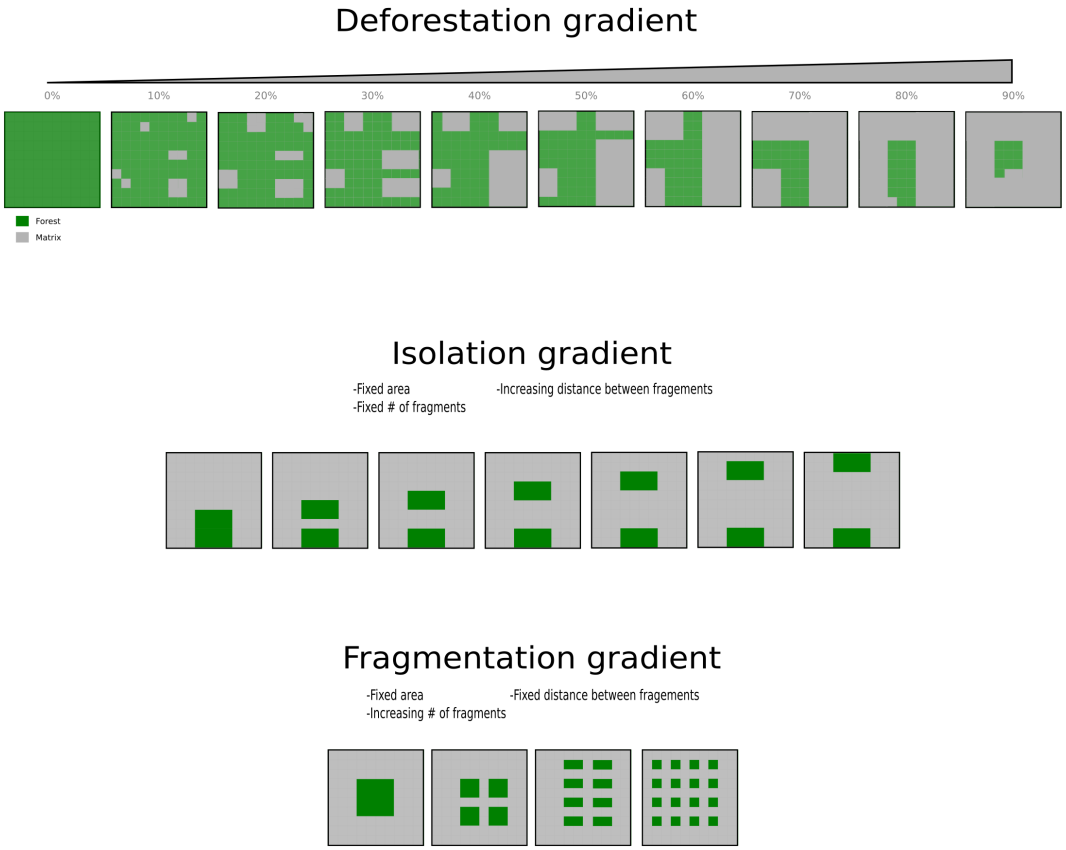

Figure 3.10: Illustration of the degradation, isolation and fragmentation gradients used for simulations.

All outputs declined with increasing deforestation, isolation and number of fragments (figure 3.11). Isolation caused a gradual decrease, while both deforestation and the number of fragments resulted in abrupt decreases early in the gradient. The home range was the least affected output in all cases. These results indicate the potentially disrupting effects that even low intensity disturbances can have on dispersers movement and seed dispersal.

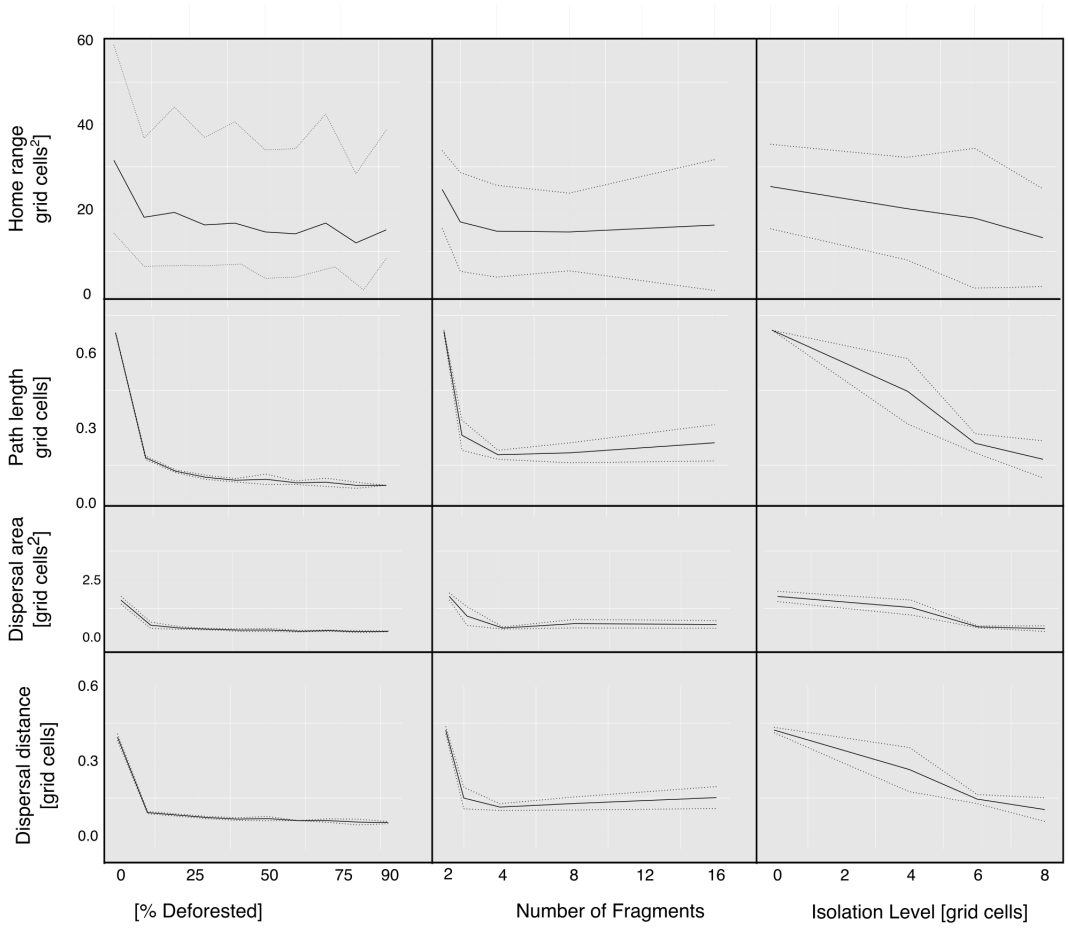

Figure 3.11: Negative effects of degradation, number of fragments and isolation on seed dispersal and animal movement metrics. Forest degradation and number of fragments drastically decrease outputs even at low levels of disturbance. Fragment isolation reduces all metrics more linearly.

4 Chapter 4: Integrated Model

This chapter discusses a procedure for explicitly including seed disperser movement into the individual-based forest model described in Chapter 2. Seed dispersal is usually simulated by dispersal kernels, mathematical functions that describe the probability of a seed being deposited at a particular distance from the parent tree (Morales and Carlo, 2006). The characteristics of these functions (for example scale and shape) can affect ecological patterns such as population dynamics, carrying capacity and community composition (Levin et al., 2003; Levine and Murrell,

2003b). For animal-dispersed plants (*i.e.,* most woody species), kernels are functions of seed traits and frugivore behaviours that determine movement patterns (Nathan and Muller-Landau, 2000), both of which are highly variable and difficult to quantify (Howe and Smallwood, 1982; Patterson et al., 2008). Most simulation studies use fixed kernels, without including adaptive behaviour and assuming that dispersal is uniform throughout the simulation, but in changing environments the kernel characteristics are likely to be variable. Levey et al. (2005) illustrates that algorithms that capture small-scale movement well enough can be used to scale long distance seed dispersal with a high degree of accuracy. The study used easily observed behaviours at the scale of 20 m to estimate the seed rain at the landscape scale, accurately predicting the effect of corridors on seed dispersal. By integrating the two models previously discussed in chapters 2 and 3, seed dispersal can be simulated explicitly, incorporating the adaptive characteristics of individual-based models into the seed dispersal process.

4.1 Model integration

One of the biggest difficulties in explicitly including animals into an individual-based forest model such as the one described on Chapter 2 is the difference in scales. Forest gap models usually operate on yearly time steps and divide the space into patches that can accommodate several trees, depending on their size, while animals must be modelled at much finer scales in order to capture changes in the agent states. The Trees model described in this thesis can operate at time steps of one year or one month. Finer time steps are possible, but come at the expense of computational time. Since the dispersers model operates at 30 min time steps, the integration needs to deal with processes occurring with different frequencies.

The integration uses multiple inheritance, a feature in many object-oriented

programming languages that allows a new object to inherit characteristics from multiple parents. Both models were implemented as independent packages and include a *World* class (among others), which coordinates the agents and controls the execution of events during a simulation. By importing these classes, a new *Integrated World* is created, inheriting from the *Tree World* and *Dispersers World*. The code below demonstrates how a new class implementing the integration represented in figure 4.1 is created. The events executed by the *Tree World* or *Dispersers World* can be rearranged in a variety of ways. Figure 4.1 illustrates the arrangement used in this chapter's demonstration. Different arrangements can be easily obtained by modifying the order or location of method calls.

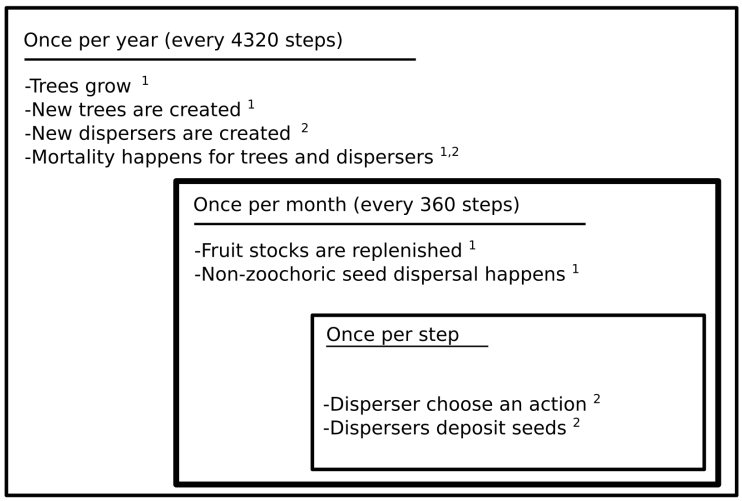

Figure 4.1: Frequency with which events are executed in the integrated model used in this chapter. 1) indicates methods inherited from *Tree World* and 2) indicates methods inherited from *Dispersers World*.

```
import trees_ibm
import dispersers

class Integrated_World(dispersers.world.Dispersers_World,trees_ibm.world.Tree_World):
    def __init__(self,topology, parent_tree_class):
        super().__init__(topology,parent_tree_class)

    def model_status_from_file(self,input_file):
        trees_ibm.world.Tree_World.model_status_from_file(self,input_file)
```

```
def
     run_trees_simulation(self,logging_settings,dispersal_settings,produce_fruits,h5file):

    trees_ibm.world.Tree_World.run_simulation(self,n=1,logging_settings=logging_settings,
    dispersal_settings=dispersal_settings, produce_fruits=produce_fruits,
    increment_time=False,h5file=h5file)

def run_integrated_simulation(self,n,logging_settings,dispersal_settings,h5file):

    for step in range(n):
        dispersers_ids=list(dispersers.PrimaryDisperser.Instances.keys())

        for ind in dispersers_ids:
            d=dispersers.PrimaryDisperser.Instances.get(ind)
            d.schedule()
            self.add_ind_entry(ind_id=d.id,
            initial_x=d.previous_position[0],
            initial_y=d.previous_position[1],
            final_x=d.position[0],
            final_y=d.position[1],
            energy=d.energy)

        if self.day==1 and self.step==1:#every month
            self.parent_tree_class.ProduceFruits()
            self.parent_tree_class.DisperseSeeds()



            if world.month==1: #every_year
                self.run_trees_simulation(logging_settings=logging_settings,
                dispersal_settings=dispersal_settings,produce_fruits=False,
                h5file=h5file)
                self.increase_dispersers_population()
                self.add_pop_entry()

        self.increment_time()
```

4.2 Example: Selective Logging

In order to demonstrate how the integrated model is applied, I use a selective logging case study similar to the one presented in Section 2.4. This example investigates how logging intensity and frequency affect carbon stocks and seed dispersal during a 55 year simulation. This time span is long enough to encompass the commonly adopted 30 years logging cycle CONAMA-CONSELHO NACIONAL DO MEIO AMBIENTE (2009), while also accommodating cycles that are shorter and longer than the usual. Starting from a 16 ha clear area, the Trees model was executed on its own for 300 years in order to produce a mature forest. This initial simulation uses the same 6 PFTs as Chapter 2 and seed dispersal was modelled using the external seed rain mode described in Section 2.1.7-I (Recruitment). The resulting mature forest is used as the starting point for the selective logging scenarios, which use the integrated model and explore a factorial design with 3 logging intensities (30, 60 and 120 m^3/ha), 3 logging frequencies (every 10, 25 and 50 years) and a control, with logging intensity set to 0. The first logging operation is executed at the beginning of year 6, making the frequencies equivalent to a total of 5, 2 and 1 logging events, respectively. As described in submodel 2.1.7-XI, trees are logged randomly within the landscape until the volume quota is met (assuming there are enough trees that fit the diameter limits). In order to obtain the mean and variance for output values, each scenario was executed 30 times using the parameter values listed on table 4.1.

Table 4.1: Integrated model parameters. The parameters for the integrated model are the same as those used for the Trees and the Dispersers models described in chapters 2 and 3, respectively. Complete descriptions and units are found in tables 2.2 and 3.1.

	Plant Functional Type					
Parameters PFT1	PFT2	PFT3	PFT4	PFT5	PFT6	
D_{max}	145	58	58	44	16	16
h_0	3.3	4.6	4.8	4.3	4.3	3.0
h_1	0.60	0.4	0.4	0.4	0.3	0.60
cl_0	0.80	0.80	0.30	0.30	0.30	0.30
cd_0	0.60	0.60	0.60	0.60	0.60	0.60
cd_1	0.68	0.68	0.68	0.68	0.68	0.68
cd_2	0	0	0	0	0	0
ρ	0.55	0.55	0.41	0.40	0.52	0.47
σ	0.70	0.70	0.70	0.70	0.70	0.70
f_0	0.77	0.77	0.77	0.77	0.77	0.77
f_1	-0.18	-0.18	-0.18	-0.18	-0.18	-0.18
l_0	2.0	2.0	2.0	2.0	2.0	2.0
l_1	0.10	0.10	0.10	0.10	0.10	0.10
I_{seed}	0.03	0.01	0.05	0.02	0.03	0.02
N_{seed}	20	15	21	50	2	200
m	0.5	0.5	0.5	0.5	0.5	0.5
r_g	0.25	0.25	0.25	0.25	0.25	0.25
M_b	0.015	0.03	0.03	0.04	0.021	0.045
$D\Delta D_{max}$	0.33	0.34	0.23	0.60	0.33	0.60
ΔD_{max}	0.012	0.012	0.019	0.029	0.011	0.029
p_{max}	2.0	3.1	6.8	11	7	12
α	0.36	0.28	0.23	0.20	0.30	0.20
α_{fruits}	0.7	0.7	0.7	0.7	0.7	0.7
gen_{dis}	0.0	1.0	0.2	0.2	0.2	0.2
A_{disp}	0.03	0.2	0.2	0.2	0.2	0.2
N_{fruit}	30	30	30	30	30	30

	Dispersers parameters
r	50
e_{level1}	80
e_{level2}	150
e_{feed}	8
e_{rest}	0.7
e_{travel}	1.6
r_p	1.2
w	2
e_r	60
c_1	70
c_2	840
t_{max}	5

Figures 4.2-c and d show that selective logging reduces carbon stocks and the dispersers populations even at mild intensities. Scenarios with the highest logging intensity and frequency resulted in carbon stocks 3.5 times lower than the control with dispersers populations 2.5 times smaller. Intermediate intensities and frequencies also resulted in reductions in carbon stocks and disperser population sizes, except for the scenarios with lowest logging intensity (30 m^3/ha), which remained close to the control (15% lower) at low and medium frequencies (50 and 25 years). It can take decades for carbon stocks to recover from a single logging event (Figure 4.2b). It took 35.5 years (sd=11.49) on average for the carbon stocks to recover from an operation that removed 30 m^3 of timber per hectare. The required time increased to 41.0 (sd=11.75) and 54.3 (sd=8.81) years as intensities increased to 60 and 120 m^3/ha.

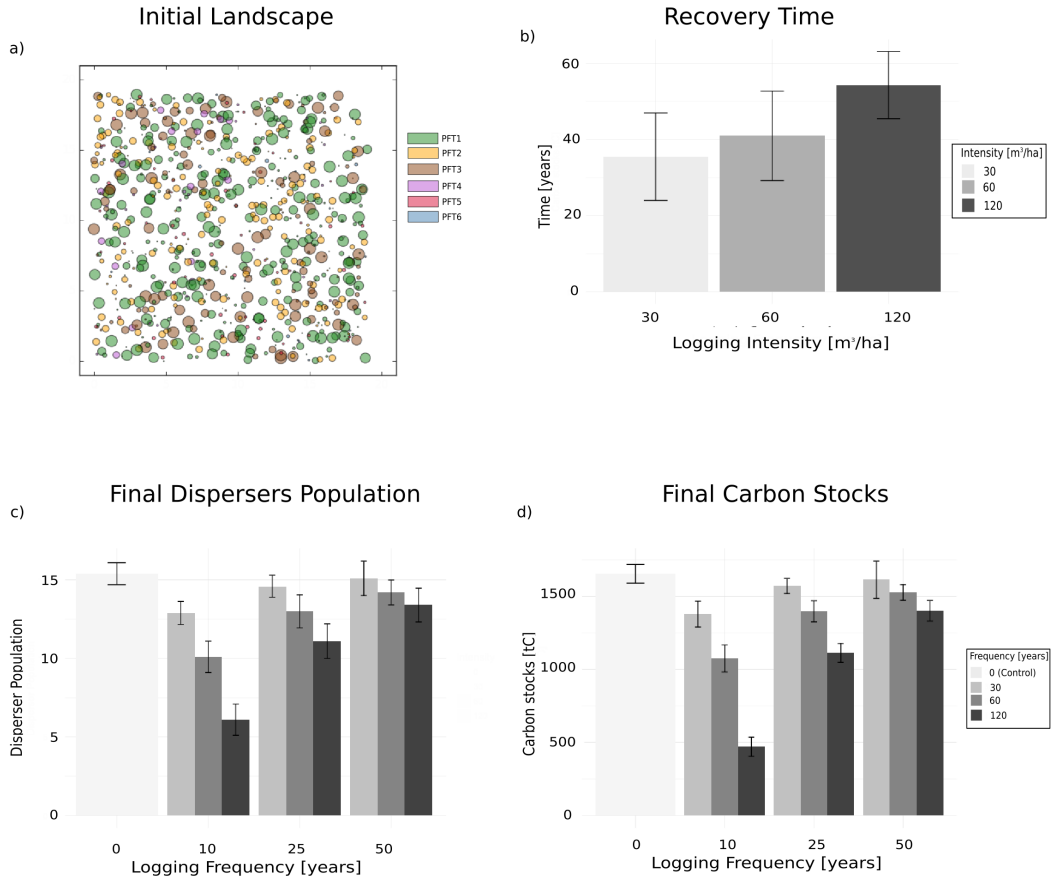

Figure 4.2: Increasing logging frequency and intensity decreases carbon stocks and dispersers populations. a) the spatial configuration of the initial 16 ha forest plot, b) time taken for carbon stocks in the living trees to recover after a single logging event. c) the effect of frequency and intensity on the population size of seed dispersers. d) the effect of frequency and intensity on carbon stocks. Carbon stocks are shown in term of aboveground biomass of living trees. The model was run 30 times for each treatment. Error bars represent standard deviation.

Selective logging affects only a subset of the measured seed dispersal characteristics (figure 4.3). The number of dispersed seeds decreases as logging intensity and frequency increase. Average path length and home range were slightly higher at the 120 m^3/ha intensity, which was expected since dispersers need to travel further if trees are more sparsely distributed. The small effect suggests that even the most severe of the scenarios did not disturb the habitat structure enough to affect movement patterns (see *Model improvements* below). The seed displace-

ment metrics (average dispersal distance and dispersal area) are not affected by any of the scenarios, indicating that some aspects seed dispersal services can still be maintained at these levels of selective-logging.

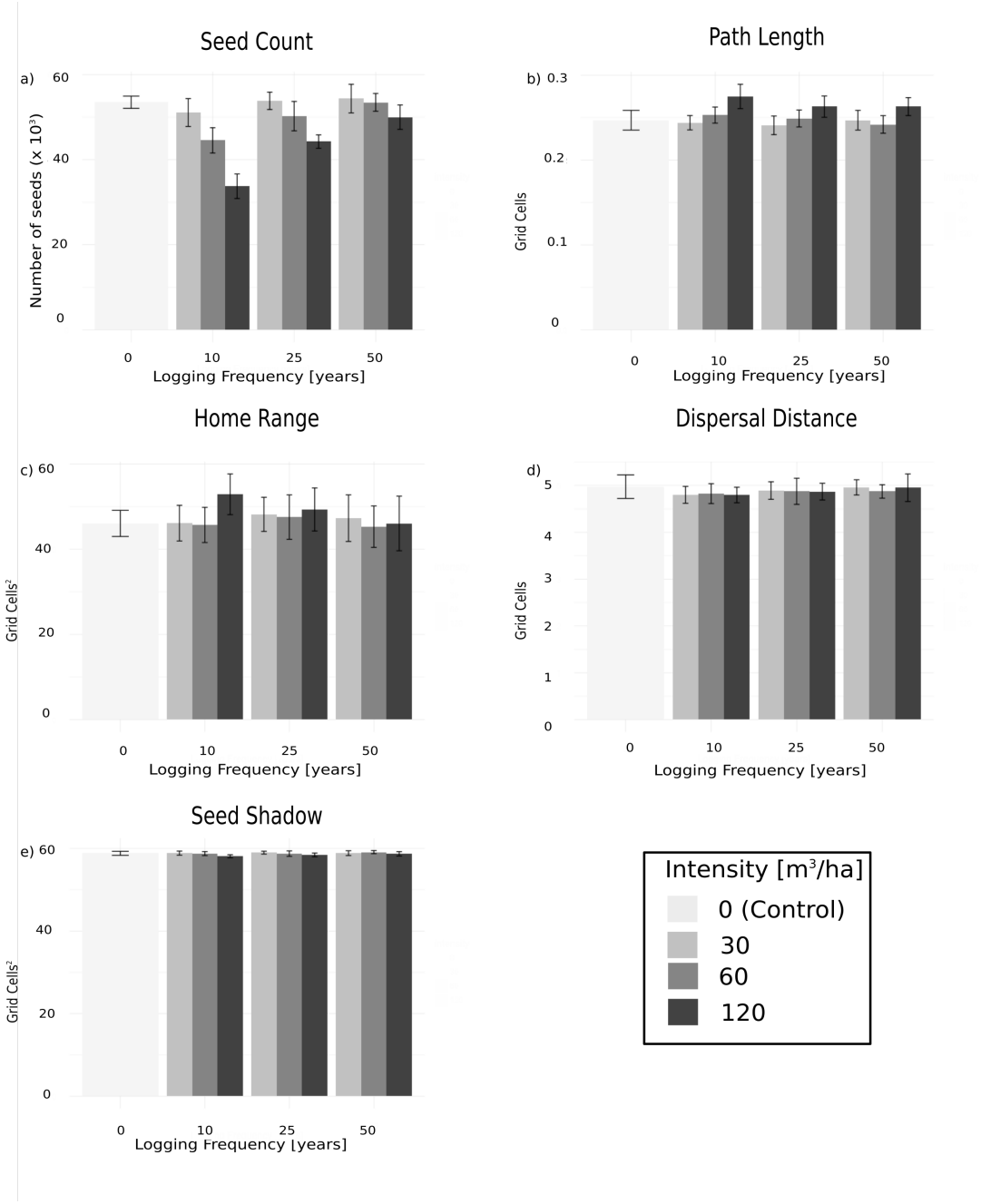

Figure 4.3: The long term effects of logging on seed dispersal are weak. The number of seeds dispersed (a) in the final year decreases with more frequent and intense logging events. Disperser movement is also affected to a lesser extent (b and c), particularly by the strongest disturbance scenario (120 m^3/ha every 10 years). Average dispersal distance (d) and area (e) are not affected.The model was run 30 times for each treatment. Error bars represent standard deviation.

The impacts of logging on the same dispersal metrics could not be detected immediately after the cut (fig. 4.4). This indicates that the changes detected at the

end of the simulated period are cumulative and that short-term monitoring might not be suitable for properly detecting the effects of selective-logging (Sebbenn et al., 2008).

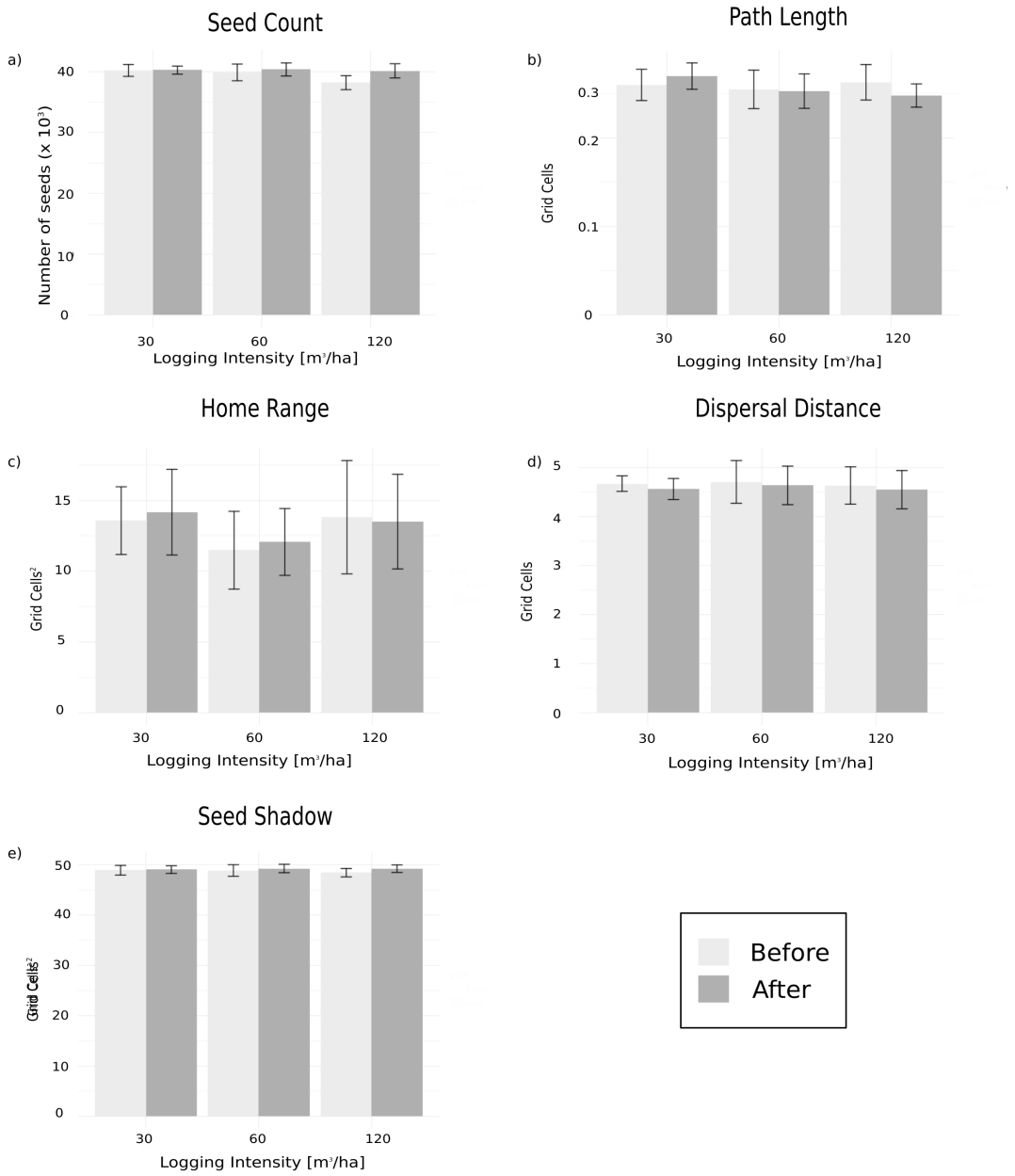

Figure 4.4: Immediate impacts of logging are not detectable. Dispersal metrics calculated immediately before and 1 year after logging do not show differences. The model was run 30 times for each treatment. Error bars represent standard deviation.

4.3 Model improvements

Although the parameter values used in table 4.1 produced results that were biologically sensible (*e.g.:* carbon stocks and net emissions per ha compared to the results found by Fischer et al. (2015); Rüger et al. (2007)), there are many other parameter combinations that may also produce acceptable outputs. For example, section 3.2 demonstrates how some of the dispersers parameters can influence movement metrics. Depending on those values (*i.e.,* depending on the type of dispersers modelled), the same levels of disturbance might have bigger or smaller effects on seed dispersal services. Figure 4.5 depicts modelling as an iterative process, with data being used at different stages in order to refine and improve models. Future versions of the models described here will benefit from field data for calibration and validation. The code structure adopted facilitates adjustments to specific cases, for example, not only different parameter values, but also other minimum modifications, such as a different number of PFTs or different disperser decision algorithms, can be incorporated with few lines of code.

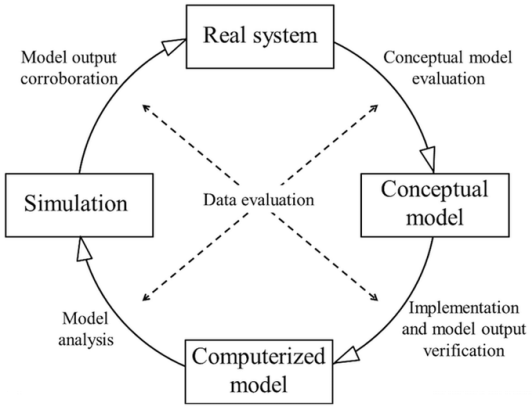

Figure 4.5: The modelling cycle (adapted from Schulze et al. (2017).

Future work will also focus on performance improvements. The 300 simulations used in this example ran in parallel using Cedar, one of Compute Canada's high performance clusters and were completed in approximately 20 hours. On a modern desktop computer (an iMac with a 3.4 GHz Intel Core i5 cpu and 8 GB of RAM),

one single integrated simulation took on average 5 hours to run. Considering that hundreds of simulations are usually required, better performance would allow users without access to supercomputers or cloud infrastructure to use the models more easily.

Figures 4.6 and 4.7 show the execution time profiles for the Trees model and the integrated model respectively. Faster methods to calculate the environmental variables should be the focus of future optimization for the Trees model. Specifically, the functions that calculate cumulative crown and leaf area for each patch could be rewritten with more efficient algorithms and/or in a compiled language, such as C or Fortran. For the disperser model, the decision algorithm takes most of the execution time and should be the focus of optimization efforts. Figure 4.7 also shows that, in the integrated model, only approximately 3 % of the time is spent on the trees (10 minutes in a 5 hour simulation). In situations where implicit seed dispersal can be used, only the Trees model is necessary, which results in lower simulation times. See section 5.4 for guidelines on how to decide if seed dispersal can be simulated implicitly (*i.e.,* without explicitly modelling the disperser agents).

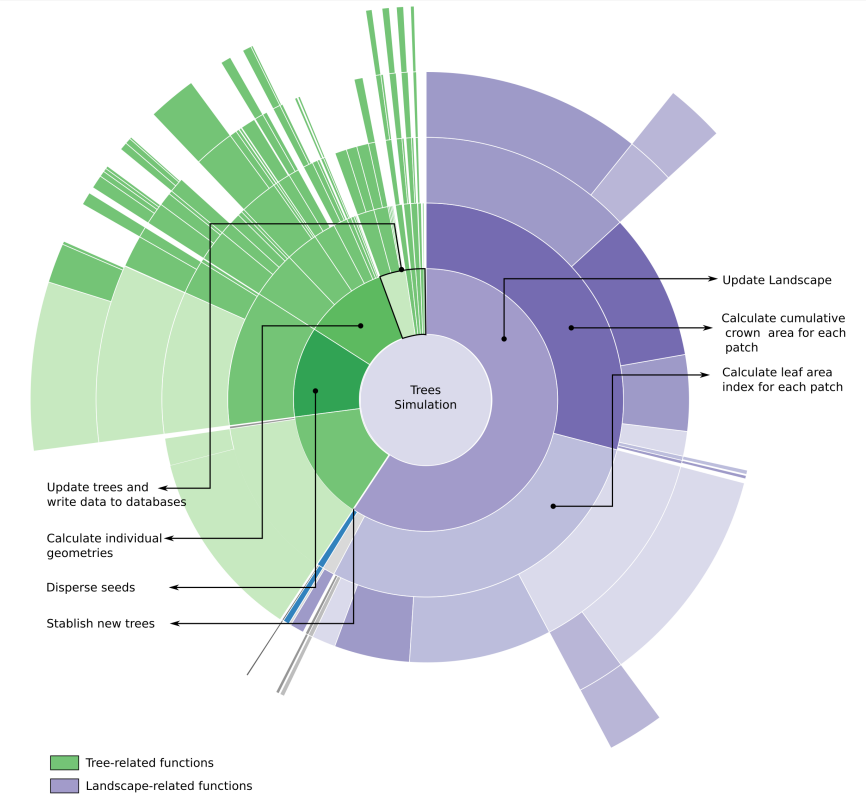

Figure 4.6: Trees model execution time profile. Each arc represents a function or method executed during a simulation. The time spent in each function is denoted by the angular width of the arc. The circle in the middle corresponds to a root function, then the functions those functions call, and so on.

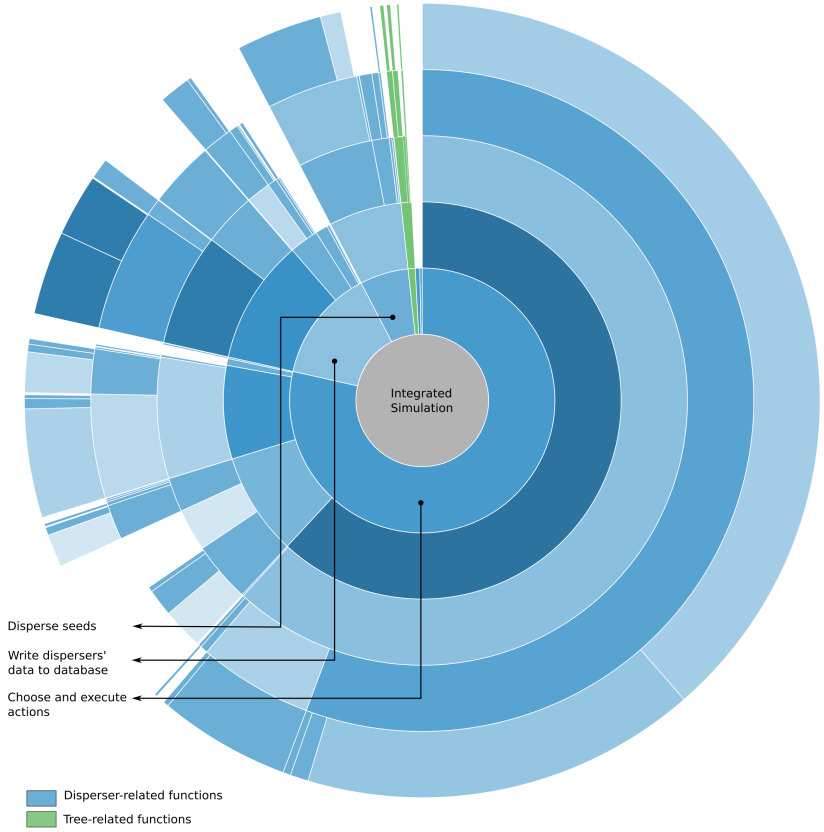

Figure 4.7: Integrated model execution profile. Most of the time is spent simulating the dispersers. Only 3% of execution time is spent on the trees.

5 Summary and conclusions

The key contributions of this project were:

- The software package that implement the models discussed Chapters 2 and 3;
- The procedure for including animals in forest models described and exemplified in Chapter 4;
- Hypotheses to be addressed in future work.

5.1 Future Directions

The inclusion of seed dispersers into individual-based forest models opens possibilities to explore a variety of real-world scenarios using the complex adaptive systems approach. This section describes some of the questions that have been recently raised in the literature, starting with concerns regarding selective-logging. For each of these topics I suggest a few hypotheses that could be tested using the integrated model and the general guidelines described in Section 5.2-Recommendations. The current version of the software packages would be sufficient to assemble the integrated model for some of the hypotheses, while others would require minimum additions to the code base (indicated when applicable). Appropriate parameter estimations would be required for all the cases.

In the study case presented in section 4.2, the seed dispersal collapse that is expected as a consequence of population declines (Pérez-Méndez et al., 2016) did not happen. Although the total number of seeds dispersed significantly decreased with logging intensity and frequency, dispersal length and dispersal area did not differ significantly from the control. Those simulations assumed that all dispersers in the landscape had the same characteristics and equivalent dispersal functions. But trees rely on a variety of disperser species, with different movement patterns and preferences (Bascompte and Jordano, 2007; González-Castro et al., 2015).

Habitat disturbance also affects frugivores in different ways, with large-bodied species being specially affected (Dirzo et al., 2014; Hansen and Galetti, 2009). The remaining species tend to be smaller and, although their abundance might increase (Peres and Dolman, 2000), the seed dispersal effectiveness is not always compensated (González-Castro et al., 2015; Wotton and Kelly, 2011) since large species are responsible for the largest fraction of dispersed seeds (Jordano et al., 2007; Vidal et al., 2013a). As such, splitting seed dispersers into groups (Zamora, 2000) that reflect functional differences related to body size might better capture the impacts of logging.

Hypotheses:

- Selective-logging results in density compensation by small-bodied frugivores.
- Population declines of large-bodied frugivores reduces dispersal-length and dispersal area in logged forests.

  Required modifications: Add the ability to create functional groups based on body size. Parameters values related to the seed dispersers could be obtained for each functional group.

Selective-logging is often followed by increased hunting pressure, driven by the easier access that roads provide to sites that used to be isolated (Jansen and Zuidema, 2001).

Hypothesis:

- Hunting following selective logging events significantly reduces average length of seed dispersal and seed shadow (dispersal area).

  Required modifications: Disperser mortality (parameter m_b) would be increased after a cut.

Fragmentation is also known to have strong negative effects on seed dispersal (Bacles et al., 2006). Results from Chapter 3 (section 3.3.2) indicate that the size

of fragments and distance between fragments affect both disperser movement and seed dispersal metrics. The integrated model can evaluate how long it takes to restore the different scenarios presented in figure 3.10 to a reference state. The different points in the deforestation, isolation and number of fragments gradient could be defined as initial states for the model. Simulations would then track the time required for each of these initial scenarios to reach a state in which the landscape is completely covered by forest again. In addition to forest coverage, community composition (in terms of abundances by PFT) and carbon stocks can also be analyzed, as studies suggest that these aspects might be different even when forest coverage has been restored (Bello et al., 2015; Gibson et al., 2011b).

Hypotheses:

- Restoration time increases as fragment isolation increases.
- Restoration time increases as fragment size decreases.
- Restoration time increases as the total abundance of dispersers decreases.

A meta-analysis done by Gibson et al. (2011b) suggests that once tropical forests have been sufficiently disturbed, there is no evidence that they can be restored to sustain the same biodiversity. Studies indicate that defaunation can lead to changes in the tree community composition, reducing the abundance or persistence of bigger, animal-dispersed trees and increasing the populations of smaller, wind-dispersed species (Peres et al., 2016). The integrated model could be used to simulate landscapes with different levels of defaunation over time.

Hypothesis:

- Tree composition in highly defaunated forests becomes dominated by wind-dispersed species over time
- Carbon stocks in highly defaunated forests decrease over time.

5.2 Recommendations

5.2.1 When to use the integrated model

The results presented in the case study show that the simulated levels of selective logging did not affect seed dispersal metrics and had little impact average movement metrics. The dispersal kernel changed very little throughout time.It is important to reiterate that all simulations presented in this thesis have the purpose of illustrating how the software works and possible outputs. Since these simulations were not calibrated to reflect any particular site, I do not claim that the results obtained should be generalized to real-world logging scenarios without carefully calibrating the model to the corresponding forests. Nonetheless,in situations where the disturbance does not significantly affect animal movement, it is advantageous to approximate the seed dispersal kernel generated by the dispersers as a simple mathematical function and use the dispersal method included in the Trees model. After using dispersal distance data (*i.e.,* the euclidean distance from each seed's deposition site to the correspondent parent tree) to estimate a probability density function, the parameter α on equation 2 can be estimated. This is a way to implicitly model seed dispersal and assumes that the fitted function sufficiently approximates dispersal at any time during the simulation. This strategy significantly reduces the simulation times because the dispersers model is only used once. After it is verified that the dispersal metrics (*i.e.,* dispersal distance and seed shadow) have small variance and that disturbance does not affect movement (figure 5.1-a), the deposition patterns can be used to estimate the kernel function (figure 5.1-d). Subsequently, only the trees model is used. In cases where animal movement and seed dispersal are affected by the disturbance, but the variance is still small (figure 5.1-b), the same strategy can be applied to the period prior to the disturbance, but inclusion of seed dispersers is recommended after that point. Finally, if movement and dispersal metrics are highly variable (figure 5.1-c), the

dispersal kernel will also change frequently; in this case, the integrated model is recommended for the entire simulation.

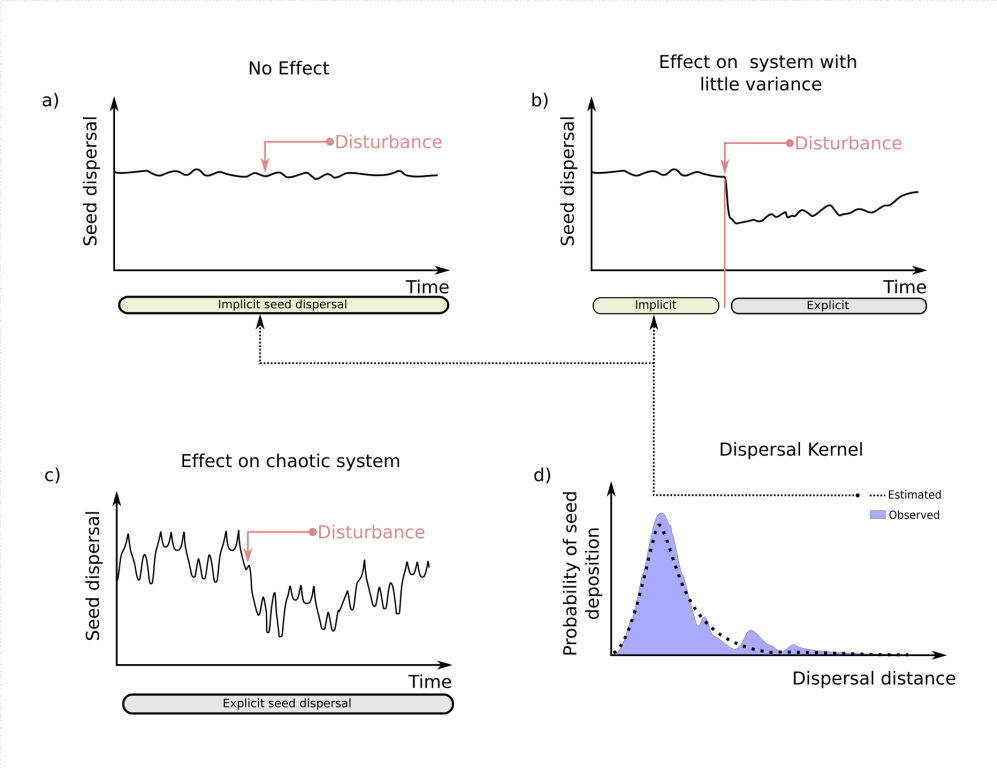

Figure 5.1: Guidelines on when to model seed dispersal implicitly (using only the Trees model) or explicitly (with the integrated model). Implicit dispersal is used when there is no effect of disturbance on the dispersal metrics (a). For scenarios where there is an effect of disturbance on dispersal metrics, implicit dispersal is used before the change if the variance is small (b), otherwise it is recommended to model dispersal explicitly (c). The dispersers model is used to estimate the dispersal function for the implicit model implementation.

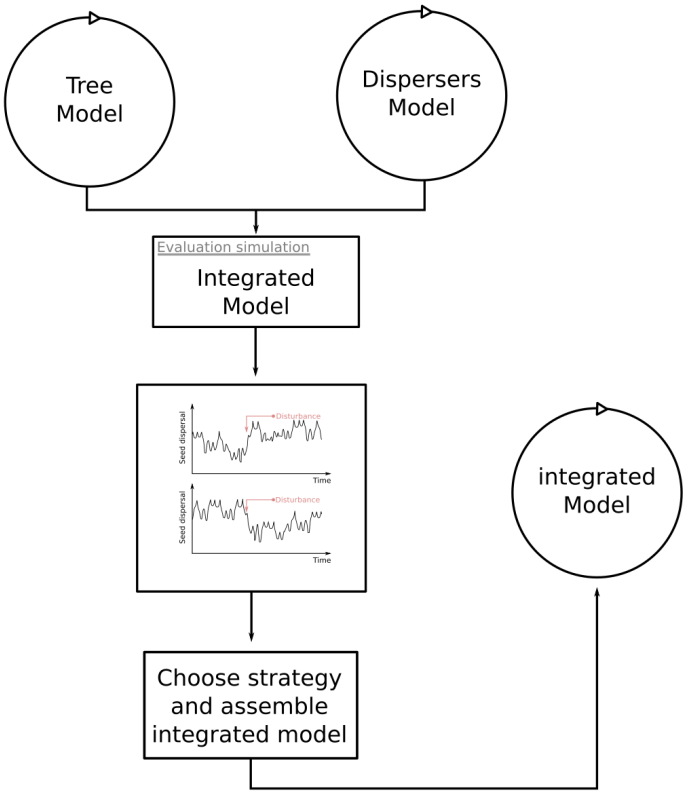

Figure 5.2: General guidelines on using the integrated model. The circles represent the model cycle described by figure 4.5. After adjusting the trees and dispersers models to the study system, and evaluation simulation is executed using the integrated model. The outputs are used to decide whether or not explicit seed dispersal is necessary and the final version of the model is assembled accordingly.

5.2.2 Model calibration

The trade-off between model simplicity and realism is a classical problem in ecological modelling. Individual-based models such as the Trees and Dispersers models presented in Chapters 2 and 3 offer flexibility to be adapted to different environments and contexts, but the large number of parameters often makes calibration more difficult than in simpler models. However, most parameters can be obtained from standard field measurements or estimated from well-known patterns.

In the Trees model, parameters related to morphometric relationships are es-

timated from standard measurements in forest inventories, such as DBH, tree height and wood density. Growth-related parameters require repeated diameter measurements or tree ring data, which can be rare for some sites (Brienen et al., 2017). The most difficult parameters to estimate are those related to mortality (*i.e.,* M_b) and birth rates (*i.e.,* I_{seed}), as well as seed-dispersal (*i.e.,* N_{seeds}, α_{fruits}). These are measurements rarely included in typical forest inventories because they require long-term censuses. However, recent advances in statistical methods are improving the calibration quality for uncertain parameters (Fischer et al., 2016). Specifically, Tolson and Shoemaker (2007) developed a collection of stochastic search algorithms that significantly decrease the computation time required to estimate parameter values. By automatically running the model thousands of times and comparing system-level patterns that are available from inventories and literature (*e.g.,* relative abundances, average basal area, number of trees per hectare), these methods provide good estimates for those parameters that are hard to measure directly (see Lehmann and Huth (2015) for an example applied to forest models). Furthermore, Bayesian methods have been used in combination with Markov chain Monte Carlo techniques to assess parameter uncertainty (van der Vaart et al., 2016).

Most parameters in the Dispersers model are directly measured by field observations of individual organisms, such as average time spent on each tree (t_{max}), the maximum number of fruits a disperser can hold in it's stomach (f_{max}) or average distance travelled per day (which is related to action radius r) (Stevenson et al., 2000). Parameters related to energy levels are hard to measure directly and techniques similar to the ones described for forest models can be applied. Bialozyt et al. (2014a) found the energy gain from each ingested fruit to be one of the most important parameters in their individual-based primate model and I observed the same in mine (parameter e_{feed}). In the population growth submodel, the energy cost to produce one individual (c_1), energy used per individual (c_2)

and background mortality (m_b) are also critical for population dynamics and can be calibrated using census data as reference. When population data is missing, typical density values can be obtained from the literature (*e.g.,* Levi and Peres (2013) and Stevenson et al. (2000)) and used as approximations. One promising approach is the inclusion of energy budgets from the Dynamic Energy Budget framework (Sousa et al., 2010) combined with Approximate Bayesian Computations to calibrate energy-related parameters (van der Vaart et al., 2016).

5.3 Summary of contributions

Throughout this thesis, I developed software tools to implement individual-based models of trees and animals that are able to explicitly include animal behaviour into forest models. Chapter 2 was dedicated to the Trees model. I described the underlying assumptions and equations, which were based on an existing model (FORMIND). My goal was to provide a simple implementation that could serve as a base for modellers wishing to develop individual-based forest models. In addition to the core processes related to tree recruitment, growth and competition, the current version includes a logging submodel, which was used to exemplify how the model works, what are the typical outputs and how it can be applied to Natural resources management (*i.e.,* selective-logging). The open-source code and object oriented implementation are flexible enough to add new processes or modify the current algorithms in order to address the specific needs of modellers. I also developed interactive tools to help users explore model outputs and understand the tree parameters.

Differently from forest models, the availability of software that explicitly models seed dispersers is reduced to source files published as appendices of a few studies (*e.g.,* Bialozyt et al. 2014a; Boyer et al. 2006b). In Chapter 3, I described a model that simulates arboreal seed disperser. It individually represents disperser agents

and their movements, using a decision algorithm inspired by primate behaviour. The included examples demonstrate how the model can be used to investigate the effects of anthropogenic disturbance (defaunation, deforestation and fragmentation) on seed dispersal. Although the decision algorithm is intended to represent a general arboreal disperser (such as a monkey or bird), the software can be modified to approximate the behaviour of a particular species more accurately.

The Trees model can be used to simulate tree communities without animals (like the example in section 2.4, which uses implicit seed dispersal) and the Dispersers model can be used without simulating the tree community processes (*i.e.,* recruitment, growth, competition, carbon storage, population sizes). However, one of the main novelties of this project was the explicit simulation of animals into forest models, discussed in Chapter 4. One of the challenges of integrating these two kinds of models is that they operate in different time and spatial scales. Section 4.1 provides an algorithm and code to perform the integration of the Trees and the Dispersers model, but the same strategy could be used with other animal models (*e.g.,* a model that explicitly simulates herbivores).

The possibility to explicitly include animals into forest models expands the tool set available for ecologists to investigate many ecological and natural resources management questions related to emergent properties of ecosystems (see the hypotheses in section 5.1), including the provision of ecosystem services, such as climate change mitigation through carbon sequestration. However, the explicit simulation of animals is not always necessary. In section 4.2, for example, I applied the integrated model to selective logging scenarios and found that logging activities did not impact seed dispersal distance and the seed shadow. Given the higher computational costs of the integrated model compared to the Trees model alone, once the modeller verifies that the animal-mediated process of interest (*e.g.,* seed dispersal) is not affected by the disturbance in question (*e.g.,* selective logging), I recommend the use of a simpler, implicit representation of the process (*e.g.,*

implicit seed dispersal as described in section 2.1.7-I, as opposed to the explicit seed dispersal used in section 4.2). Nonetheless, whether or not animal-mediated processes should be implicitly or explicitly modelled highly depends on the questions and systems being investigated and I recommend the procedure described in section 5.2.1 to be used with appropriate calibrations.

The code that is used as a base to implement each of the models discussed in Chapters 2 and 3 is organized as independent Python packages (*trees-ibm* and *dispersers*, respectively). They can be used independently or together to implement the individual-based models described in this document or variations of them. In addition, a third package (*py-ibm*) contains general code that is common to many IBMs and should facilitate the development of IBMs in Python. All packages are available (see table 5.1 for links) under the GNU General Public License v3.0 from PyPI, the official Python repository (https://pypi.python.org/pypi) and can be installed using the python package manager (pip). Figure 5.3 contains the UML (Unified Modelling Language) class diagrams illustrating the internal organization of the *py-ibm* package. Diagrams for the *trees-ibm* and *dispersers* can be found in figures 2.3 and 3.4, respectively. The GitHub repositories include example scripts, documentation, detailed installation instructions and requirements, as well as guidelines for submitting contributions and reporting bugs through GitHub's issues and pull requests systems. All code may be freely used, modified and redistributed.

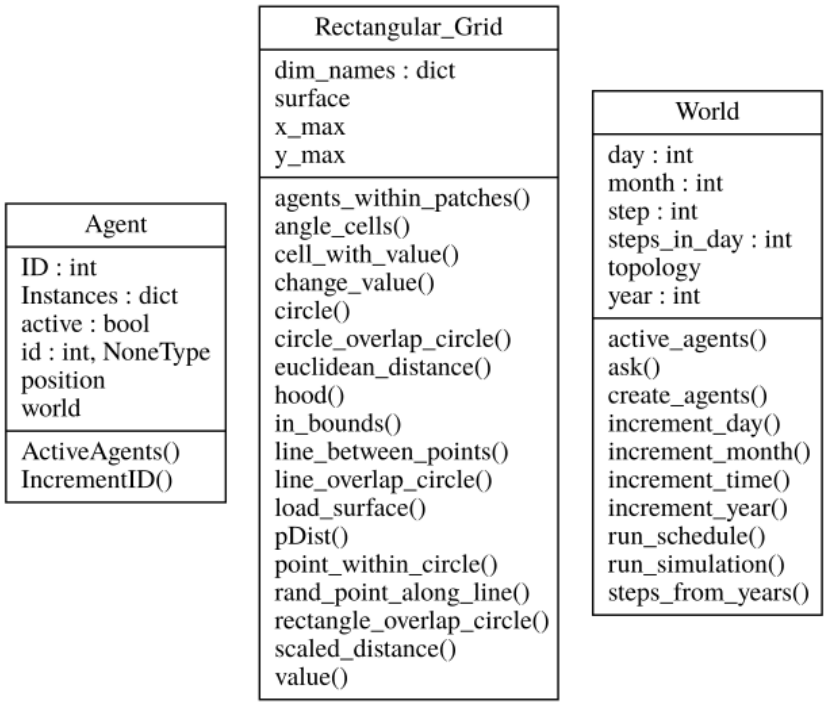

Figure 5.3: Class diagram for the Py-IBM package. Following the UML (Unified Modelling Language) standards, each class is represented by a block divided into three sections: the class name at the top, followed by the corresponding attributes in the middle and the methods at the bottom.

Table 5.1: Python packages produced for this thesis. Packages can be downloaded as compressed files, accessed on GitHub or installed via pip (Python package manager).

Py-IBM	
File	https://pypi.python.org/pypi/py_ibm/0.1.1
GitHub	http://github.com/fsfrazao/Py_IBM
pip command	pip install py-ibm
Trees-IBM	
File	https://pypi.python.org/pypi/trees_ibm/0.2.2
GitHub	http://github.com/fsfrazao/Trees
pip command	pip install trees-ibm
Dispersers	
File	https://pypi.python.org/pypi/dispersers/0.2.2
GitHub	http://github.com/fsfrazao/Dispersers
pip command	pip install dispersers

6 Bibliography

Bibliography

Asner, G. P., Brodrick, P. G., Philipson, C., Vaughn, N. R., Martin, R. E., Knapp, D. E., Heckler, J., Evans, L. J., Jucker, T., Goossens, B., et al. (2018). Mapped aboveground carbon stocks to advance forest conservation and recovery in malaysian borneo. *Biological Conservation*, 217:289–310.

Avgar, T., Deardon, R., and Fryxell, J. M. (2013). An empirically parameterized individual based model of animal movement, perception, and memory. *Ecological Modelling*, 251:158–172.

Azevedo-Ramos, C., de Carvalho Jr., O., and do Amaral, B. D. (2006). Short-term effects of reduced-impact logging on eastern amazon fauna. *Forest Ecology and Management*, 232(13):26–35.

Bacles, C. F., Lowe, A. J., and Ennos, R. A. (2006). Effective seed dispersal across a fragmented landscape. *Science*, 311(5761):628–628.

Balvanera, P., Pfisterer, A. B., Buchmann, N., He, J.-S., Nakashizuka, T., Raffaelli, D., and Schmid, B. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters*, 9(10):1146–1156.

Bascompte, J. and Jordano, P. (2007). Plant-animal mutualistic networks: the architecture of biodiversity. *Annu. Rev. Ecol. Evol. Syst.*, 38:567–593.

Belle, S. V., Estrada, A., and Garber, P. A. (2013). Collective group movement and leadership in wild black howler monkeys (alouatta pigra). *Behavioral Ecology and Sociobiology*, 67(1):31–41.

Bello, C., Galetti, M., Pizo, M. A., Magnago, L. F. S., Rocha, M. F., Lima, R. A.,

Peres, C. A., Ovaskainen, O., and Jordano, P. (2015). Defaunation affects carbon storage in tropical forests. *Science Advances*, 1(11):e1501105.

Berenguer, E., Ferreira, J., Gardner, T. A., Arago, L. E. O. C., Camargo, P. B. D., Cerri, C. E., Durigan, M., Oliveira, R. C. D., Vieira, I. C. G., and Barlow, J. (2014). A largescale field assessment of carbon stocks in humanmodified tropical forests. *Global Change Biology*, 20(12):3713–3726.

Betts, R. A., Malhi, Y., and Roberts, J. T. (2008). The future of the amazon: new perspectives from climate, ecosystem and social sciences. *Philosophical transactions of the Royal Society of London.Series B, Biological sciences*, 363(1498):1729–1735. LR: 20140904; JID: 7503623; RF: 30; OID: NLM: PMC2367686; ppublish.

Bialozyt, R., Flinkerbusch, S., Niggemann, M., and Heymann, E. W. (2014a). Predicting the seed shadows of a neotropical tree species dispersed by primates using an agent-based model with internal decision making for movements. *Ecological Modelling*, 278:74–84.

Bialozyt, R., Flinkerbusch, S., Niggemann, M., and Heymann, E. W. (2014b). Predicting the seed shadows of a neotropical tree species dispersed by primates using an agent-based model with internal decision making for movements. *Ecological Modelling*, 278:74–84.

Botkin, D. B., Janak, J. F., and Wallis, J. R. (1972). Some ecological consequences of a computer model of forest growth. *The Journal of Ecology*, pages 849–872.

Boyer, D., Ramos-Fernandez, G., Miramontes, O., Mateos, J. L., Cocho, G., Larralde, H., Ramos, H., and Rojas, F. (2006a). Scale-free foraging by primates emerges from their interaction with a complex environment. *Proceedings.Biological sciences / The Royal Society*, 273(1595):1743–1750. LR: 20140909; JID: 101245157; OID: NLM: PMC1634795; ppublish.

Boyer, D., Ramos-Fernández, G., Miramontes, O., Mateos, J. L., Cocho, G., Larralde, H., Ramos, H., and Rojas, F. (2006b). Scale-free foraging by primates emerges from their interaction with a complex environment. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1595):1743–1750.

Brienen, R. J., Gloor, M., and Ziv, G. (2017). Tree demography dominates long-term growth trends inferred from tree rings. *Global change biology*, 23(2):474–484.

Burivalova, Z., aan Hakk ekerciolu, and Koh, L. P. (2014). Thresholds of logging intensity to maintain tropical forest biodiversity. *Current biology*, 24(16):1893–1898.

Carey, E. V., Sala, A., Keane, R., and Callaway, R. M. (2001). Are old forests underestimated as global carbon sinks? *Global Change Biology*, 7(4):339–344.

Caspersen, J. P., Vanderwel, M. C., Cole, W. G., and Purves, D. W. (2011). How stand productivity results from size- and competition-dependent growth and mortality. *PloS one*, 6(12):e28660.

Caughlin, T. T., Ferguson, J. M., Lichstein, J. W., Zuidema, P. A., Bunyavejchewin, S., and Levey, D. J. (2015). Loss of animal seed dispersal increases extinction risk in a tropical tree species due to pervasive negative density dependence across life stages. In *Proc. R. Soc. B*, volume 282, page 20142095. The Royal Society.

Chambers, J. C. and MacMahon, J. A. (1994). A day in the life of a seed: movements and fates of seeds and their implications for natural and managed systems. *Annual Review of Ecology and Systematics*, pages 263–292.

CONAMA-CONSELHO NACIONAL DO MEIO AMBIENTE (2009). RESOLUCAO No 406-Estabelece parametros tecnicos a serem ado-

tados na elaboracao, apresentacao, avaliacao tecnica e execucao de Plano de Manejo Florestal Sustentavel- (PMFS) com fins madeireiros, para florestas nativas e suas formas de sucessao no bioma Amazonia. http://www.mma.gov.br/port/conama/legiabre.cfm?codlegi=597. (In Portuguese, accessed on: 13/apr/2018).

Costanza, R. (2012). Ecosystem health and ecological engineering. *Ecological Engineering*, 45:24–29.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., and Collen, B. (2014). Defaunation in the anthropocene. *science*, 345(6195):401–406.

Daz, S., Hector, A., and Wardle, D. A. (2009). Biodiversity in forest carbon sequestration initiatives: not just a side benefit. *Current Opinion in Environmental Sustainability*, 1(1):55–60.

Estes, J. A., Terborgh, J., Brashares, J. S., Power, M. E., Berger, J., Bond, W. J., Carpenter, S. R., Essington, T. E., Holt, R. D., Jackson, J. B., et al. (2011). Trophic downgrading of planet earth. *science*, 333(6040):301–306.

Felton, A., Felton, A., Foley, W., and Lindenmayer, D. (2010). The role of timber tree species in the nutritional ecology of spider monkeys in a certified logging concession, bolivia. *Forest Ecology and Management*, 259(8):1642–1649.

Filotas, E., Parrott, L., Burton, P. J., Chazdon, R. L., Coates, K. D., Coll, L., Haeussler, S., Martin, K., Nocentini, S., Puettmann, K. J., et al. (2014). Viewing forests through the lens of complex systems science. *Ecosphere*, 5(1):1–23.

Fiore, A. D. and Suarez, S. A. (2007). Route-based travel and shared routes in sympatric spider and woolly monkeys: cognitive and evolutionary implications. *Animal cognition*, 10(3):317–329.

Fischer, R., Bohn, F., de Paula, M. D., Dislich, C., Groeneveld, J., Gutiérrez, A. G., Kazmierczak, M., Knapp, N., Lehmann, S., Paulick, S., et al. (2016). Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests. *Ecological modelling*, 326:124–133.

Fischer, R., Ensslin, A., Rutten, G., Fischer, M., Costa, D. S., Kleyer, M., Hemp, A., Paulick, S., and Huth, A. (2015). Simulating carbon stocks and fluxes of an african tropical montane forest with an individual-based forest model. *PLoS One*, 10(4):e0123300.

Galetti, M. and Dirzo, R. (2013). Ecological and evolutionary consequences of living in a defaunated world. *Biological Conservation*, 163:1–6.

Gamfeldt, L. and Roger, F. (2017). Revisiting the biodiversity–ecosystem multifunctionality relationship. *Nature ecology & evolution*, 1(7):0168.

Garcia, D., Zamora, R., and Amico, G. C. (2010). Birds as suppliers of seed dispersal in temperate ecosystems: conservation guidelines from real-world landscapes. *Conservation Biology*, 24(4):1070–1079.

Gatti, R. C., Castaldi, S., Lindsell, J., Coomes, D., Marchetti, M., Maesano, M., Paola, A. D., Paparella, F., and Valentini, R. (2015). The impact of selective logging and clearcutting on forest structure, tree diversity and above-ground biomass of african tropical forests. *Ecological Research*, 30(1):119–132.

Gibson, L., Lee, T. M., Koh, L. P., Brook, B. W., Gardner, T. A., Barlow, J., Peres, C. A., Bradshaw, C. J., Laurance, W. F., and Lovejoy, T. E. (2011a). Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature*, 478(7369):378–381.

Gibson, L., Lee, T. M., Koh, L. P., Brook, B. W., Gardner, T. A., Barlow, J., Peres, C. A., Bradshaw, C. J., Laurance, W. F., Lovejoy, T. E., et al. (2011b).

Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature*, 478(7369):378–381.

González-Castro, A., Calviño-Cancela, M., and Nogales, M. (2015). Comparing seed dispersal effectiveness by frugivores at the community level. *Ecology*, 96(3):808–818.

Gourlet-Fleury, S., Blanc, L., Picard, N., Sist, P., Dick, J., Nasi, R., Swaine, M. D., and Forni, E. (2005). Grouping species for predicting mixed tropical forest dynamics: looking for a strategy. *Annals of forest science*, 62(8):785–796.

Green, J. L., Hastings, A., Arzberger, P., Ayala, F. J., Cottingham, K. L., Cuddington, K., Davis, F., Dunne, J. A., Fortin, M.-J., and Gerber, L. (2005). Complexity in ecology and conservation: mathematical, statistical, and computational challenges. *Bioscience*, 55(6):501–510.

Grimm, V., Ayllón, D., and Railsback, S. F. (2017). Next-generation individual-based models integrate biodiversity and ecosystems: yes we can, and yes we must. *Ecosystems*, 20(2):229–236.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., and Railsback, S. F. (2010). The odd protocol: a review and first update. *Ecological modelling*, 221(23):2760–2768.

Grimm, V. and Railsback, S. F. (2005). *Individual-based modeling and ecology*. Princeton university press.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H. H., Weiner, J., Wiegand, T., and DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science (New York, N.Y.)*, 310(5750):987–991. LR: 20070319; JID: 0404511; RF: 27; ppublish.

Groot, R. D., Brander, L., Ploeg, S. V. D., Costanza, R., Bernard, F., Braat, L., Christie, M., Crossman, N., Ghermandi, A., and Hein, L. (2012). Global estimates of the value of ecosystems and their services in monetary units. *Ecosystem services*, 1(1):50–61.

Groot, R. S. D., Alkemade, R., Braat, L., Hein, L., and Willemen, L. (2010). Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. *Ecological Complexity*, 7(3):260–272.

Gursky-Doyen, S. and Nekaris, K. A.-I. (2007). *Primate anti-predator strategies*. Springer Science & Business Media.

Hansen, D. M. and Galetti, M. (2009). The forgotten megafauna. *Science*, 324(5923):42–43.

Howe, H. F. and Miriti, M. N. (2000). No question: seed dispersal matters. *Trends in Ecology and Evolution*, 15(11):434–436.

Howe, H. F. and Miriti, M. N. (2004). When seed dispersal matters. *Bioscience*, 54(7):651–660.

Howe, H. F. and Smallwood, J. (1982). Ecology of seed dispersal. *Annual review of ecology and systematics*, 13(1):201–228.

Huston, M., DeAngelis, D., and Post, W. (1988). New computer models unify ecological theory. *BioScience*, 38(10):682–691.

Jansen, P. A. and Zuidema, P. A. (2001). Logging, seed dispersal by vertebrates, and natural regeneration of tropical timber trees. *The cutting edge: conserving wildlife in logged tropical forests. Columbia University Press, New York*, pages 35–60.

Janson, C. H. and Byrne, R. (2007). What wild primates know about resources: opening up the black box. *Animal cognition*, 10(3):357–367.

Johns, A. D. (1986). Effects of selective logging on the behavioral ecology of west malaysian primates. *Ecology*, pages 684–694.

Jordano, P., Garcia, C., Godoy, J., and García-Castaño, J. L. (2007). Differential contribution of frugivores to complex seed dispersal patterns. *Proceedings of the National Academy of Sciences*, 104(9):3278–3282.

Kareiva, P., Tallis, H., Ricketts, T. H., Daily, G. C., and Polasky, S. (2011). *Natural capital: theory and practice of mapping ecosystem services*. Oxford University Press.

Köhler, P. and Huth, A. (1998). The effects of tree species grouping in tropical rainforest modelling: Simulations with the individual-based model formind. *Ecological Modelling*, 109(3):301–321.

Kowalewski, M. M., Garber, P. A., Corts-Ortiz, L., Urbani, B., and Youlatos, D. (2014). *Howler Monkeys: Behavior, Ecology, and Conservation*. Springer.

Kremen, C., Williams, N. M., Aizen, M. A., Gemmill-Herren, B., LeBuhn, G., Minckley, R., Packer, L., Potts, S. G., Steffan-Dewenter, I., Vázquez, D. P., et al. (2007). Pollination and other ecosystem services produced by mobile organisms: a conceptual framework for the effects of land-use change. *Ecology letters*, 10(4):299–314.

Lambert, J. E. and Chapman, C. A. (2005). The fate of primate-dispersed seeds: Deposition pattern, dispersal distance and implications for conservation. *Seed fate: Predation, dispersal and seedling establishment*, pages 137–150.

Larsen, T. H., Williams, N. M., and Kremen, C. (2005). Extinction order and

altered community structure rapidly disrupt ecosystem functioning. *Ecology letters*, 8(5):538–547.

Lehmann, S. and Huth, A. (2015). Fast calibration of a dynamic vegetation model with minimum observation data. *Ecological modelling*, 301:98–105.

Levey, D. J., Bolker, B. M., Tewksbury, J. J., Sargent, S., and Haddad, N. M. (2005). Effects of landscape corridors on seed dispersal by birds. *Science*, 309(5731):146–148.

Levi, T. and Peres, C. A. (2013). Dispersal vacuum in the seedling recruitment of a primate-dispersed amazonian tree. *Biological Conservation*, 163(Supplement C):99 – 106. Special Issue: Defaunations impact in terrestrial tropical ecosystems.

Levin, S. A., Muller-Landau, . H. C., Nathan, . R., and Chave, . J. (2003). The ecology and evolution of seed dispersal: a theoretical perspective. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):575–604.

Levine, J. M. and Murrell, D. J. (2003a). The community-level consequences of seed dispersal patterns. *Annual Review of Ecology, Evolution, and Systematics*, 34:549–574.

Levine, J. M. and Murrell, D. J. (2003b). The community-level consequences of seed dispersal patterns. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):549–574.

Limburg, K. E., O'Neill, R. V., Costanza, R., and Farber, S. (2002). Complex systems and valuation. *Ecological Economics*, 41(3):409–420.

Loreau, M. (2010). Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philosophical transactions of the Royal Society of London.Series*

B, Biological sciences, 365(1537):49–60. LR: 20141204; JID: 7503623; RF: 91; OID: NLM: PMC2842700; ppublish.

MacMahon, J. A., Mull, J. F., and Crist, T. O. (2000). Harvester ants (pogonomyrmex spp.): their community and ecosystem influences. *Annual Review of Ecology and Systematics*, 31(1):265–291.

McGill, B. J., Enquist, B. J., Weiher, E., and Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in ecology & evolution*, 21(4):178–185.

Messier, C. and Puettmann, K. J. (2011). Forests as complex adaptive systems: implications for forest management and modelling. *Italian Journal of Forest and Mountain Environments*, 66(3):249–258.

Millennium Ecosystem Assessment, M. (2005). *Ecosystems and human well-being*, volume 5. Island Press Washington, DC.

Morales, J. M. and Carlo, T. A. (2006). The effects of plant distribution and frugivore density on the scale and shape of dispersal kernels. *Ecology*, 87(6):1489–1496.

Morris, R. J. (2010). Anthropogenic impacts on tropical forest biodiversity: a network structure and ecosystem functioning perspective. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1558):3709–3718.

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., and Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, 105(49):19052–19059.

Nathan, R. and Muller-Landau, H. C. (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in ecology & evolution*, 15(7):278–285.

Nelson, E., Mendoza, G., Regetz, J., Polasky, S., Tallis, H., Cameron, D., Chan, K. M., Daily, G. C., Goldstein, J., and Kareiva, P. M. (2009). Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment*, 7(1):4–11.

Nichols, E., Spector, S., Louzada, J., Larsen, T., Amezquita, S., Favila, M., and Network, T. S. R. (2008). Ecological functions and ecosystem services provided by scarabaeinae dung beetles. *Biological Conservation*, 141(6):1461–1474.

Parrotta, J. A., Wildburger, C., and Mansourian, S. (2012). *Understanding relationships between biodiversity, carbon, forests and people: the key to achieving REDD objectives*. International Union of Forest Research Organizations (IUFRO).

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. (2008). State–space models of individual animal movement. *Trends in ecology & evolution*, 23(2):87–94.

Peres, C. A. and Dolman, P. M. (2000). Density compensation in neotropical primate communities: evidence from 56 hunted and nonhunted amazonian forests of varying productivity. *Oecologia*, 122(2):175–189.

Peres, C. A., Emilio, T., Schietti, J., Desmoulière, S. J., and Levi, T. (2016). Dispersal limitation induces long-term biomass collapse in overhunted amazonian forests. *Proceedings of the National Academy of Sciences*, 113(4):892–897.

Pérez-Méndez, N., Jordano, P., García, C., and Valido, A. (2016). The signatures

of anthropocene defaunation: cascading effects of the seed dispersal collapse. *Scientific reports*, 6.

Pfeifer, M., Lefebvre, V., Turner, E., Cusack, J., Khoo, M., Chey, V. K., Peni, M., and Ewers, R. M. (2015). Deadwood biomass: an underestimated carbon stock in degraded tropical forests? *Environmental Research Letters*, 10(4):044019.

Phillips, N. G., Buckley, T. N., and Tissue, D. T. (2008). Capacity of old trees to respond to environmental change. *Journal of Integrative Plant Biology*, 50(11):1355–1364.

Phillips, O. L. (1996). The changing ecology of tropical forests. *Biodiversity Conservation*, pages 291–311.

Picard, N. and Franc, A. (2003). Are ecological groups of species optimal for forest dynamics modelling? *Ecological Modelling*, 163(3):175–186.

Picard, N., Köhler, P., Mortier, F., and Gourlet-Fleury, S. (2012). A comparison of five classifications of species into functional groups in tropical forests of french guiana. *Ecological Complexity*, 11:75–83.

Post, W. M. and Pastor, J. (1996). Linkagesan individual-based forest ecosystem model. *Climatic Change*, 34(2):253–261.

Putz, F. E., Zuidema, P. A., Synnott, T., Peña-Claros, M., Pinard, M. A., Sheil, D., Vanclay, J. K., Sist, P., Gourlet-Fleury, S., Griscom, B., et al. (2012). Sustaining conservation values in selectively logged tropical forests: the attained and the attainable. *Conservation Letters*, 5(4):296–303.

Rademacher, C., Neuert, C., Grundmann, V., Wissel, C., and Grimm, V. (2004). Reconstructing spatiotemporal dynamics of central european natural beech forests: the rule-based forest model before. *Forest Ecology and Management*, 194(1):349–368.

Railsback, S. F. and Grimm, V. (2012). *Agent-based and individual-based modeling: a practical introduction*. Princeton university press.

Railsback, S. F. and Johnson, M. D. (2011). Pattern-oriented modeling of bird foraging and pest control in coffee farms. *Ecological Modelling*, 222(18):3305–3319.

Railsback, S. F., Lamberson, R. H., Harvey, B. C., and Duffy, W. E. (1999). Movement rules for individual-based models of stream fish. *Ecological Modelling*, 123(2):73–89.

Ramos-Fernndez, G., Mateos, J. L., Miramontes, O., Cocho, G., Larralde, H., and Ayala-Orozco, B. (2004). Lvy walk patterns in the foraging movements of spider monkeys (ateles geoffroyi). *Behavioral Ecology and Sociobiology*, 55(3):223–230.

Redford, K. H. (1992). The empty forest. *Bioscience*, pages 412–422.

Reiss, J., Bridle, J. R., Montoya, J. M., and Woodward, G. (2009). Emerging horizons in biodiversity and ecosystem functioning research. *Trends in ecology & evolution*, 24(9):505–514.

Ricketts, T. H., Daily, G. C., Ehrlich, P. R., and Michener, C. D. (2004). Economic value of tropical forest to coffee production. *Proceedings of the National Academy of Sciences of the United States of America*, 101(34):12579–12582. LR: 20140608; JID: 7505876; 0 (Coffee); OID: NLM: PMC515099; 2004/08/11 [aheadofprint]; ppublish.

Rieb, J. T., Chaplin-Kramer, R., Daily, G. C., Armsworth, P. R., Böhning-Gaese, K., Bonn, A., Cumming, G. S., Eigenbrod, F., Grimm, V., Jackson, B. M., et al. (2017). When, where, and how nature matters for ecosystem services: Challenges for the next generation of ecosystem service models. *BioScience*, 67(9):820–833.

Ripple, W. J., Estes, J. A., Beschta, R. L., Wilmers, C. C., Ritchie, E. G., Hebblewhite, M., Berger, J., Elmhagen, B., Letnic, M., Nelson, M. P., et al. (2014). Status and ecological effects of the worlds largest carnivores. *Science*, 343(6167):1241484.

Rüger, N., Gutiérrez, A. G., Kissling, W. D., Armesto, J. J., and Huth, A. (2007). Ecological impacts of different harvesting scenarios for temperate evergreen rain forest in southern chilea simulation experiment. *Forest Ecology and Management*, 252(1):52–66.

Schleuning, M., Farwig, N., Peters, M. K., Bergsdorf, T., Bleher, B., Brandl, R., Dalitz, H., Fischer, G., Freund, W., Gikungu, M. W., et al. (2011). Forest fragmentation and selective logging have inconsistent effects on multiple animal-mediated ecosystem processes in a tropical forest. *PLoS One*, 6(11):e27785.

Schmitz, O. J., Raymond, P. A., Estes, J. A., Kurz, W. A., Holtgrieve, G. W., Ritchie, M. E., Schindler, D. E., Spivak, A. C., Wilson, R. W., Bradford, M. A., et al. (2014). Animating the carbon cycle. *Ecosystems*, 17(2):344–359.

Schulze, J., Müller, B., Groeneveld, J., and Grimm, V. (2017). Agent-based modelling of social-ecological systems: Achievements, challenges, and a way forward. *Journal of Artificial Societies & Social Simulation*, 20(2).

Sebbenn, A. M., Degen, B., Azevedo, V. C., Silva, M. B., de Lacerda, A. E., Ciampi, A. Y., Kanashiro, M., Carneiro, F. d. S., Thompson, I., and Loveless, M. D. (2008). Modelling the long-term impacts of selective logging on genetic diversity and demographic structure of four tropical tree species in the amazon forest. *Forest ecology and management*, 254(2):335–349.

Sekercioglu, C. H. (2006). Increasing awareness of avian ecological function. *Trends in ecology & evolution*, 21(8):464–471.

Solar, R. R. d. C., Barlow, J., Ferreira, J., Berenguer, E., Lees, A. C., Thomson, J. R., Louzada, J., Maus, M., Moura, N. G., and Oliveira, V. H. (2015). How pervasive is biotic homogenization in humanmodified tropical forest landscapes? *Ecology Letters*, 18(10):1108–1118.

Sousa, T., Domingos, T., Poggiale, J.-C., and Kooijman, S. (2010). Dynamic energy budget theory restores coherence in biology.

Stephenson, N. L., Das, A., Condit, R., Russo, S., Baker, P., Beckman, N., Coomes, D., Lines, E., Morris, W., Rüger, N., et al. (2014). Rate of tree carbon accumulation increases continuously with tree size. *Nature*, 507(7490):90–93.

Stevenson, P. R. et al. (2000). Seed dispersal by woolly monkeys(lagothrix lagothricha) at tinigua national park, colombia: Dispersal distance, germination rates, and dispersal quantity. *American Journal of Primatology*, 50(4):275–289.

Terborgh, J., Pitman, N., Silman, M., Schichter, H., and Nez, P. (2002). Maintenance of tree diversity in tropical forests. *Seed dispersal and frugivory: ecology, evolution and conservation.CABI Publishing, Wallingford*, pages 1–17.

The HDF Group (1997-2018). Hierarchical data format version 5. http://www.hdfgroup.org/HDF5.

Thiele, J. C., Kurth, W., and Grimm, V. (2014). Facilitating parameter estimation and sensitivity analysis of agent-based models: A cookbook using netlogo and r. *Journal of Artificial Societies and Social Simulation*, 17(3):11.

Thornley, J. and Johnson, I. (1990). *Plant and crop modelling: a mathematical approach to plant and crop physiology*. Oxford University Press.

Tolson, B. A. and Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1).

Topping, C. J., Dalkvist, T., and Grimm, V. (2012). Post-hoc pattern-oriented testing and tuning of an existing large model: lessons from the field vole. *PL o S One*, 7(9).

Troy, A. and Wilson, M. A. (2006). Mapping ecosystem services: practical challenges and opportunities in linking gis and value transfer. *Ecological economics*, 60(2):435–449.

UNFCCC (2007). Report of the conference of the parties on its thirteenth session, held in bali from 3 to 15 december 2007. In *United Nations Framework Convention on Climate Change (UN FCCC)*.

UNFCCC, C. (2015). Adoption of the paris agreement. proposal by the president. Technical report, United Nations Office at Geneva.

van der Vaart, E., Johnston, A. S., and Sibly, R. M. (2015). Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. *Ecological Modelling*.

van der Vaart, E., Johnston, A. S., and Sibly, R. M. (2016). Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. *Ecological modelling*, 326:113–123.

Vidal, M. M., Pires, M. M., and Guimarães, P. R. (2013a). Large vertebrates as the missing components of seed-dispersal networks. *Biological Conservation*, 163:42–48.

Vidal, M. M., Pires, M. M., and Guimares, P. R. (2013b). Large vertebrates as the missing components of seed-dispersal networks. *Biological Conservation*, 163(Supplement C):42 – 48. Special Issue: Defaunations impact in terrestrial tropical ecosystems.

Wall, S. B. V., Forget, P.-M., Lambert, J. E., and Hulme, P. E. (2005). Seed fate pathways: filling the gap between parent and offspring. *Seed fate: Predation, dispersal and seedling establishment*, pages 1–8.

Wang, B. C. and Smith, T. B. (2002). Closing the seed dispersal loop. *Trends in Ecology and Evolution*, 17(8):379–386.

Weiner, J. and Thomas, S. C. (2001). The nature of tree growth and the" age-related decline in forest productivity". *Oikos*, 94(2):374–376.

Wilmers, C. C. and Schmitz, O. J. (2016). Effects of gray wolf-induced trophic cascades on ecosystem carbon cycling. *Ecosphere*, 7(10).

Wotton, D. M. and Kelly, D. (2011). Frugivore loss limits recruitment of large-seeded trees. *Proceedings of the Royal Society of London B: Biological Sciences*, 278(1723):3345–3354.

Zamora, R. (2000). Functional equivalence in plant-animal interactions: ecological and evolutionary consequences. *Oikos*, 88(2):442–447.

Zarin, D. J., Harris, N. L., Baccini, A., Aksenov, D., Hansen, M. C., Ramos, C. A., Azevedo, T., Margono, B. A., Alencar, A. C., and Gabris, C. (2015). Can carbon emissions from tropical deforestation drop by 50years? *Global Change Biology*.