

1 The LAKEOBS_MIX model

This section describes how I developed the ABM model named LAKEOBS_MIX towards the goal of better understanding nonlinear changes in lake ecosystems, using the *NetLogo* programming environment (version 6.1.1). The new lake ecosystem model was written from scratch in *Netlogo* language and designed with the purpose to resemble regime shifts in a MLYB lake observed in chapter 3, with some algorithms based on open-source well-established models (in Appendix B). Netlogo was developed by Wilensky (1999) and written mostly in Scala and partly in Java for agent-based modelling purpose. Owing to being free, performing well, and having rich open sources in its user communities, *NetLogo* is by far the most dominant ABM platform in ecology, social sciences and archaeology (Abar et al., 2017). I have constructed the ABM model based on the conceptual model of feedback loops inferred from palaeoecological data on ecosystem transitions in Lake Taibai by setting up the simplest predation or competition relationships, environment dispersal, and individual life cycles. The aim of this model is to address the broader understanding of the mechanisms by which eutrophication-induced critical transitions have occurred in MLYB lakes over the past century, and to attempt to provide scenario simulations of de-eutrophication.

1.1 Overview

1.1.1 Purpose

This ABM model was constructed primarily to integrate existing modern ecological observations and ecological changes from long timescale palaeoecological reconstructions to

increase understanding of the ecological mechanisms underlying the occurrence of abrupt changes in lake ecosystems, particularly those appearing in the form of critical transitions. Under conditions of known lake ecosystem characteristics and external drivers, I primarily sought to model the mechanisms of formation, hysteresis, and recovery before and after the occurrence of critical transitions in the lake's ecosystem. Based on theories of individual ecology and feedback mechanisms, the emergent power of ABM is used to demonstrate the role of changes in biological functional groups in the formation of critical transitions and to test the performance of multiple recovery pathways to inform the management of some eutrophic water bodies where ecological transitions have already occurred.

1.1.1.1 *Entities, state variables, and scales*

An agent is the basic unit in ABMs and can be a physical or virtual entity that acts, perceives its environment, communicates with others and has skills to achieve its goals (Ferber, 1999). A patch is a specific agent in the NetLogo environment often used to depict the spatial environment. A computational ABM is based on agents and their interactions to simulate system patterns at a higher scale with possible emergent behaviour (Li et al., 2010). The model architecture comprised a grid of 31×31 (961) patches, each representing 1m and in sum 9.61 km² lake area. The model world wrapped both horizontally and vertically so that agents can go through the left end and appear in the right end, resembling the Mercator Projection (cylindrical projection) applied in the common world map. Each grid cell had attributes of 'water?' and 'refuge?' to determine its specific usage. Shades of patch colour indicated lake eutrophication state as defined by nutrient concentration in water.

In the survey of lakes in MLYB (Nanjing Institute of Geographical and Limnological science CAS, 2019), dominant functional groups in ecosystems consist of phytoplankton, zooplankton, zoobenthos, macrophytes, planktivorous fish, herbivorous fish and piscivorous fish. Most of them live in the pelagic zone of lakes and contribute to major lake functions in e.g., water quality and fishery resources. Most debates over the factors regulating trophic interactions in lakes have been focusing on the freshwater pelagic zone of lakes even though many planktivorous fish are also benthivorous and rely to a large extent on benthic food sources (Jeppesen et al., 2003). This is partly because pelagic food webs and the effect on nutrient dynamics is rather straightforward and most nutrients contained in the food can be soon returned to the water through excretion after being taken up by phytoplankton (Scheffer, 1998). Six dominant functional groups in pelagic ecosystems were modelled as agents, phytoplankton, zooplankton, submerged macrophyte, planktivorous fish, herbivorous fish and piscivores (*Figure 1-1*). Agents represented super-individuals of the above groups for the convenience of modelling large populations (Scheffer et al., 1995).

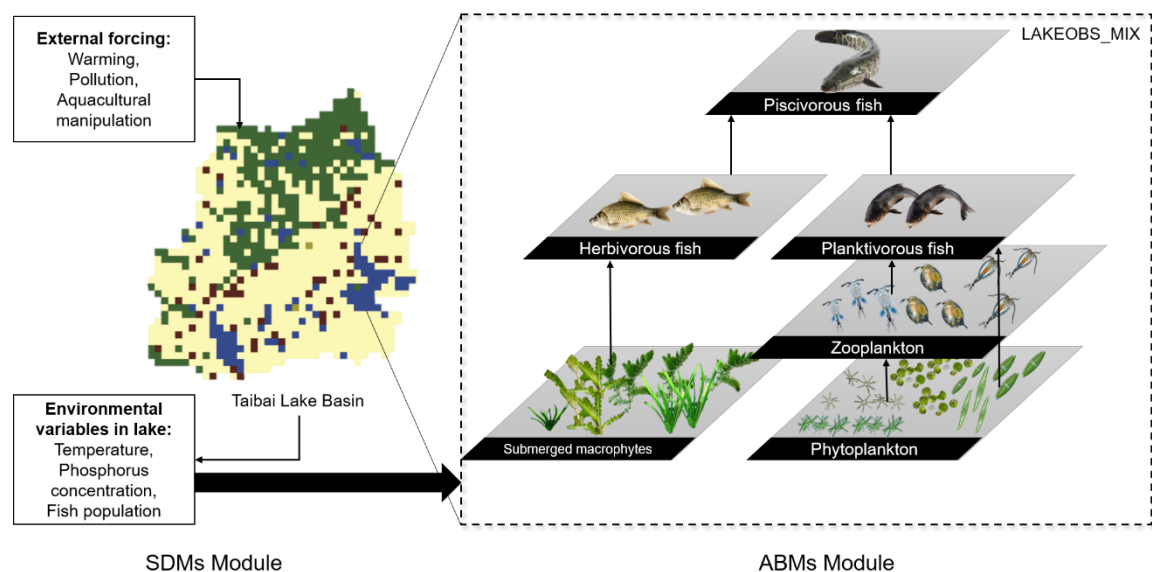


Figure 1-1 the conceptual diagram of the components and interactions in my integrated model LAKEOBS_MIX

All biotic communities live on and interact with the patches close to them in spatial distance. Each community lives with routines of rules and the goal to survive. Primary producers, i.e., phytoplankton and submerged macrophytes, can live alone in suitable conditions and consumers live with (one of) their prey. Each agent has attributes of their identity number, colour, heading, position, breed, energy, biomass, age, and survival energy. Phytoplankton and submerged macrophytes have extra individual variables of light-attenuation and growth rate, facilitating the photosynthesis process. The modulation module for environmental change is implemented with the aid of Netlogo System Dynamics Model (SDM) tool, which reads the driving force change patterns and parameters from the 'Drivers scenarios' section of the ABM, and after judgement and calculation produces new environmental variables influenced by external drivers that are immediately fed into the Ecosystem ABM module.

State variables and parameters

State variables consist of environmental variables, agent-level variables, and aggregation-level variables. Most environmental variables are assigned as 'global' variables, which means the diffusion between patches is zero in time lag and the physical environment is overall homogenous. Individual variables describe the states of each individual and the aggregation variables show higher aggregation-level performance by individuals, e.g., population, biomass, and the reproduction. Environmental and aggregation variables are listed in the following Table. 1-1 and agent-level variables are presented in the 'Submodel' section (1.3).

Table. 1-1 State variables on aggregation level in LAKEOBS_MIX

Code	Abbrev.	Description	Unit
tick	t	time step in model	D
water-temperature	T_w	average temperature of lake water	°C
water-depth	D	average depth of lake water	M
max-pxcor	W	lake width	100m
max-pycor	H	lake height	100m
phytoplankton-amount	N_{phyp}	phytoplankton population	ind.
submerged-macrophyte-amount	N_{sm}	Submerged macrophyte population	ind.
zooplankton-amount	N_z	zooplankton population	ind.
planktivore-fish-amount	N_{pf}	planktivorous fish population	ind.
herbivore-fish-amount	N_{hf}	herbivorous fish population	ind.
piscivore-amount	N_{pisc}	piscivores population	ind.
phytoplankton-concentration	C_{phyp}	phytoplankton biomass, measured as organism mass in volume	mg/m ³
zooplankton-concentration	C_z	zooplankton biomass, measured as organism mass in volume	mg/m ³
sm-concentration	C_{sm}	submerged macrophyte biomass, measured as organism mass in volume	mg/m ³
fish-concentration	C_{fish}	all fish biomass, measured as organism mass in volume	mg/m ³
pf-concentration	C_{pf}	planktivorous fish biomass, measured as organism mass in volume	mg/m ³
hf-concentration	C_{hf}	herbivorous fish biomass, measured as organism mass in volume	mg/m ⁴
pisci-concentration	C_{pisci}	piscivore biomass, measured as organism mass in volume	mg/m ⁵
total-nutrient-concentration	NC_{tot}	total concentration of nutrient in the system	µg/L
global-free-nutrient-concentration	NC_{free}	concentration of nutrient available in water	µg/L
phytoplankton-being-eaten	NBE_{phyp}	Sum of phytoplankton being grazed	ind.
zooplankton-being-eaten	NBE_z	Sum of zooplankton being eaten	ind.
macrophyte-being-eaten	NBE_{sm}	Sum of macrophyte agents being eaten	ind.
planktivore-being-eaten	NBE_{pf}	Sum of planktivorous fish being predated	ind.
herbivore-being-eaten	NBE_{hf}	Sum of herbivorous fish being predated	ind.
piscivore-being-eaten	NBE_{pisci}	Sum of piscivorous fish being predated	ind.
this-tick-being-eaten	NBE	Sum of agents being eaten	ind.
phytoplankton-natural-die	ND_{phyp}	Sum of phytoplankton out of energy	ind.
zooplankton-natural-die	ND_z	Sum of zooplankton out of energy	ind.
macrophyte-natural-die	ND_{sm}	Sum of macrophyte out of energy	ind.
planktivore-natural-die	ND_{pf}	Sum of planktivorous fish out of energy	ind.

herbivore-natural-die	ND _{hf}	Sum of herbivorous fish out of energy	ind.
piscivore-natural-die	ND _{pisc}	Sum of piscivores out of energy	ind.
this-tick-natural-die	ND	Sum of agents dying out of energy	ind.
this-tick-reproduction	RP	Sum of new-born agents	ind.

In the model, major parameters are listed in the following Table. 1-2. Some parameter values referred to value ranges in the investigation report of lakes in the Eastern region of China, including MLYB lakes (Ministry of Ecology and Environment of PRC, 2020). Modern (2019) mean water temperature in Taibai Lake was 15.47 °C in March and 31.31°C in July. Mean Total Phosphorous (TP) was 69 µg/L in spring and 228 µg/L in summer (Nanjing Institute of Geographical and Limnological science CAS, 2019). The mean depth of Taibai was 3.16 m and the lake area was about 11 km² where the lake depth was deeper than 3.4m. The installation of the system started from “system-composition”, which is a character selector to define which groups involved in the simulation. The selection includes ecologically-possible combinations for convenience of model verification and experimentation, namely "OFF", "only algae", "only macrophyte", "algae + zooplankton", "algae + zooplankton + planktivores", "algae + zooplankton + planktivores + piscivores", "algae + macrophytes ", "algae + zooplankton + macrophytes", "macrophytes + herbivores", "algae + macrophytes + herbivores", "algae + macrophytes + zooplankton + planktivores", "algae + macrophytes + zooplankton + planktivores + herbivores", "algae + macrophytes + zooplankton + planktivores + piscivores", and "algae + macrophytes + zooplankton + planktivores + herbivores + piscivores".

Table. 1-2 Main functional parameters, abbreviations, descriptions and units in the model

Code	Abbrev.	Description	Unit
------	---------	-------------	------

system-composition	SYSC	functional groups that exist in system	chr ¹
ini-water-temperature	T _{ini}	the initial lake water temperature	°C
ini-water-depth	D	the initial water depth	m
Initial-nutrient	NC _{ini}	the initial nutrient concentration	µg/L
photoperiod	PP	portion of daytime	d ⁻¹
phytoplankton-start-amount	N _{p,0}	the initial number of phytoplankton	ind.
macrophyte-start-amount	N _{sm,0}	the initial number of macrophyte	ind.
zooplankton-start-amount	N _{z,0}	the initial number of zooplankton	ind.
planktivore-fish-start-amount	N _{pf,0}	the initial number of planktivorous fish	ind.
herbivore-fish-start-amount	N _{hf,0}	the initial number of herbivorous fish	ind.
piscivore-start-amount	N _{pis,0}	the initial number of piscivores	ind.
zoopk-hunt-rate	HR _z	possibility of zooplankton eating algae	%
pf-hunt-rate	HR _{pf}	possibility of planktivorous fish eating food	%
hf-hunt-rate	HR _{hf}	possibility of herbivorous fish eating macrophytes	%
pisci-hunt-rate	HR _{pisci}	possibility of piscivores eating other fishes	%
phytoplankton-reproduce-rate	RR _p	possibility of phytoplankton reproducing when energy is enough	%
macrophyte-reproduce-rate	RR _{sm}	possibility of macrophyte reproducing when energy is enough	%
zooplankton-reproduce-rate	RR _z	possibility of zooplankton reproducing when energy is enough	%
planktivore-reproduce-rate	RR _{pf}	possibility of planktivores reproducing when energy is enough	%
herbivore-reproduce-rate	RR _{hf}	possibility of herbivores reproducing when energy is enough	%
piscivore-reproduce-rate	RR _{pis}	possibility of piscivores reproducing when energy is enough	%
zooplankton-reproduce-energy	Er _z	threshold energy of zooplankton being mature to reproduce	J

¹ 'chr' is an abbreviation of characters, 'num' of numbers and 'ind.' of individuals in the expressions of units

planktivore-reproduce-energy	Er_{pf}	threshold energy of planktivorous fish being mature to reproduce	J
herbivore-reproduce-energy	Er_{hf}	threshold energy of herbivorous fish being mature to reproduce	J
piscivore-reproduce-energy	Er_{pis}	threshold energy of piscivores being mature to reproduce	J
nutrient-concentration-in-phytoplankton	NC_{phyp}	hypothetical proportion of nutrient absorbed by phytoplankton in weight	g/g
nutrient-concentration-in-zooplankton	NC_z	hypothetical proportion of nutrient absorbed by zooplankton	g/g
nutrient-concentration-in-sm	NC_{sm}	hypothetical proportion of nutrient absorbed by submerged macrophytes	g/g
nutrient-concentration-in-fish	NC_f	hypothetical proportion of nutrient absorbed by fish	g/g

Besides the main body of the model, there is a set of environmental parameters adjusting the changing variables and patterns of the external condition of the ecosystem (Table. 1-3).

Table. 1-3 Parameters defining external forcing, with their units and value ranges

Parameter	Unit	Range
driving-scenario	chr	["Equ", "Manual", "TBNutADD", "TBNutADD_TempADD", "TBComprehens"]
drivers-pattern	chr	["OFF", "temperature", "depth", "nutrient loading", "nutrient loading + temperature", "reclamation", "refuge-on", "Adding herbivores", "Adding planktivores", "adding piscivores", "nutrient loading + reclamation + depth + temperature", "nutrient loading + reclamation + depth + temperature + planktivores"]
temperature-variation-rate	% per year	[-50,50]
depth-variation-rate	% per year	[-50,50]
reclamation-length	% (max-pxcor)	[0,100]
reclamation-width	% (max-pycor)	[0,100]
refuge-length	% (max-pxcor)	[0,100]
refuge-width	% (max-pycor)	[0,100]
nutrient-variation-rate	% per interval	[-50,50]
nutri-adding-interval	num	(input numerical)

herbivore-adding	Ind. per interval	[0,100]
hf-adding-interval	Ind. per interval	[0,100]
planktivore-adding	Ind. per interval	[0,100]
pf-adding-interval	num	(input)
piscivores-adding	num	(input)
pisci-adding-interval	num	(input)

For observation of the changing environment as a result of the system dynamic model (SDM) module in ABM, a list of variables was set up (Table. 1-4).

Table. 1-4 Environmental variables directly modified by parameters of external forces, with units

Variable	Unit
water-area	hectare
water-depth	m
water-temperature	°C
reclamation area	m ²
refuge area	m ²
herbivoresADD	ind./day
planktivoresADD	ind./day
piscivoresADD	ind./day

1.1.1.2 Process overview and scheduling

In this model, the main process sequences are initialisation, individual life cycles, environmental change, and cycles of individual birth and death with environmental change. The initialisation process contains initialisation of the environmental conditions and initialisation of the biota, plus some data cleaning and data logging algorithms in preparation for multiple models runs. In coding order, initialisation runs sequentially through the functions of initialising parameters, setting drivers, creating functional groups, linking the system dynamics model, and preparing to record exported data. Each function is synthesised by a series of instructions that complete the determination of input parameters and the assignment of values to individual variables. The initialisation process can be completed by

clicking the 'setup' button on the Netlogo interactive interface. The initialisation process is followed by the main loop, which consists of updating environmental variables within the system, receiving input on environmental changes, the biome life cycle, and data logging in turn. The main loop process can be executed by clicking the 'go' button on the Netlogo interactive interface. In the default mode the program will automatically cycle through the process, each cycle being marked by a logging time. The environmental change input function is performed by both ABM and SDM, i.e. the parameters are entered into SDM from the ABM interface, processed by SDM and returned to ABM to update the variable assignments from the last cycle. The biome life cycle is the core part of the model and consists of six groups of biomes that are sorted to execute instructions. Finally, the values of the variables at the end of the cycle are recorded and derived, after which the variable 'age' is added to all individuals by one day. The specific algorithms for each procedure are described in the submodels section.

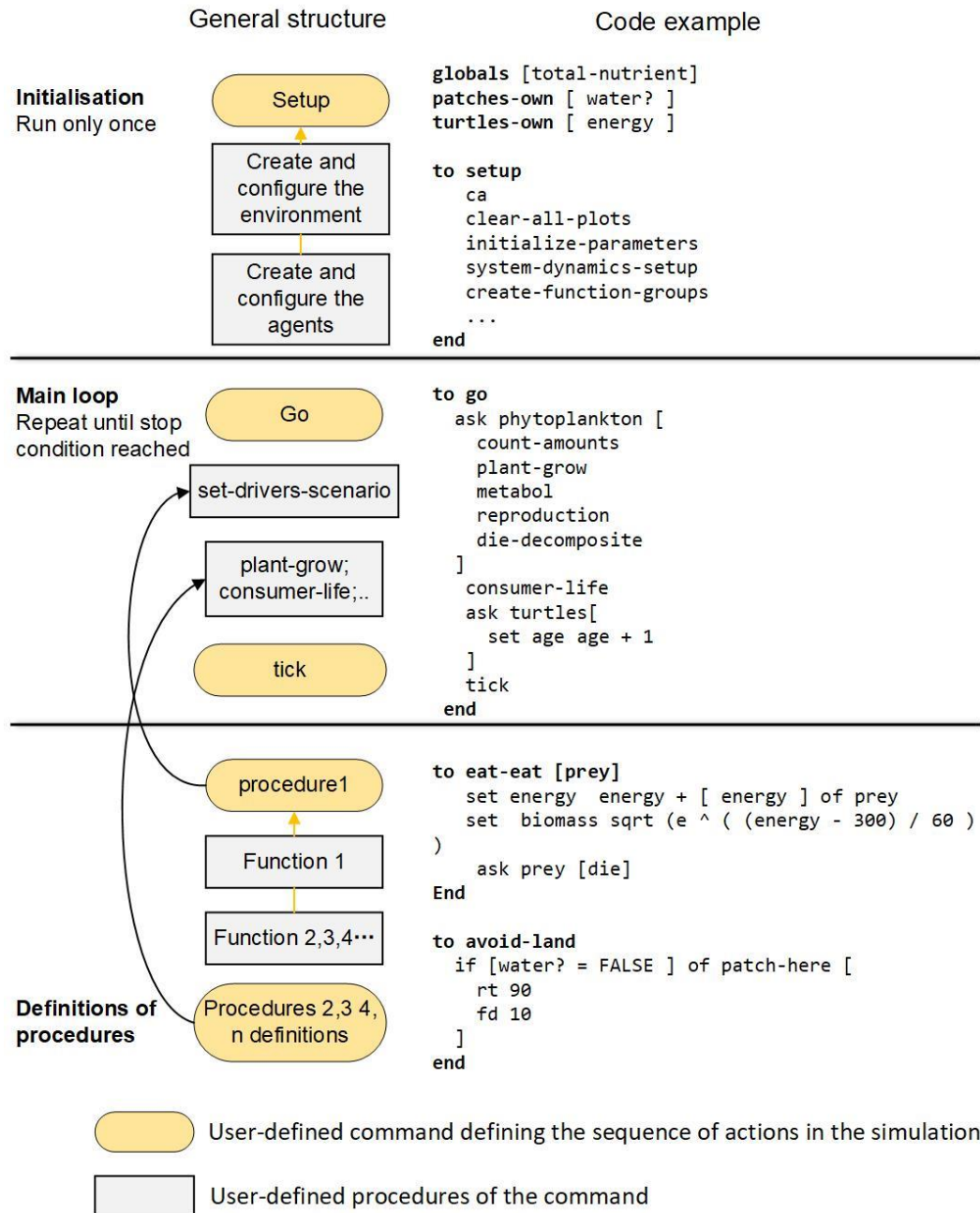


Figure 1-2 Conceptual diagram of the general structure and code examples in compartments of initialisation, main loops and functions in the model. A typical Netlogo model starts the overall initialisation with the command "Setup" **for once** and repeats rules with the continuously executed command "Go". The main execution commands consist of many small procedures, nested within the main execution procedures like "Setup" and "Go". In the code example, the main procedure "Go " includes a procedure "consumer-life" for defining life cycles of animals, which is further composed of detailed behaviours like 'eat prey' and 'keep swimming in water'.

1.2 Design concepts

1.2.1 Basic principles

As the aim of the model is to observe ecosystems undergoing abrupt changes and critical transitions, it is designed with the expectation that a model can develop a state of system equilibrium, as well as population emergent behaviour driven by external forces. It is based on individual biology but can calculate the behaviour of natural systems under uncertainty. The design principles of the ecosystem are therefore as simple and basic as possible. The model has several primary assumptions: 1) total nutrient quality is conserved within individual organisms and within the water column in the absence of external disturbances in the ABM simulation; 2) energy and matter are conserved within living individuals and are imported via sunlight and food; 3) increasing nutrient concentrations within the appropriate range can promote photosynthetic rates in producers; 4) predators search for the nearest food on a spatial scale; and 5) Multiple populations in the absence of disturbance should be able to find a coexisting population dynamic equilibrium.

Once the assumptions of the model are established, individuals are given some basic rules about the life cycle, mainly derived from simplified biological experience and knowledge. After referring to individual ecological knowledge, multiple ecological models and homeostatic models, I established rules of biological behaviour mainly in four procedures: nutrient cycling, photosynthesis, metabolism and mass-energy conversion.

Nutrient cycling

Without external disturbance, the total nutrient concentration is fixed and exists either as an inorganic form in water or as an organic form incorporated in the agents. Parameters defining the ability of agents absorbing nutrient are *nutrient-concentration-in-phytoplankton*, *nutrient-concentration-in-zooplankton*, *nutrient-concentration-in-fish* and *nutrient-concentration-in-sm*. The observation variables are *total-nutrient* (NC_{tot}), *initial-nutrient* and *global-free-nutrient* (NC_{free}), which are related to each other as following *equation 1-1* and diagram (Figure 1-3).

$$NC_{free} = NC_{tot} - C_{phyp} \times NC_p - C_z \times NC_z - C_{sm} \times NC_{sm} - C_{fish} \times NC_{fish} \quad \text{equation 1-1}$$

where NC_p , NC_z , NC_{sm} , and NC_{fish} are nutrient concentration in organisms, and the units are dimensionless.

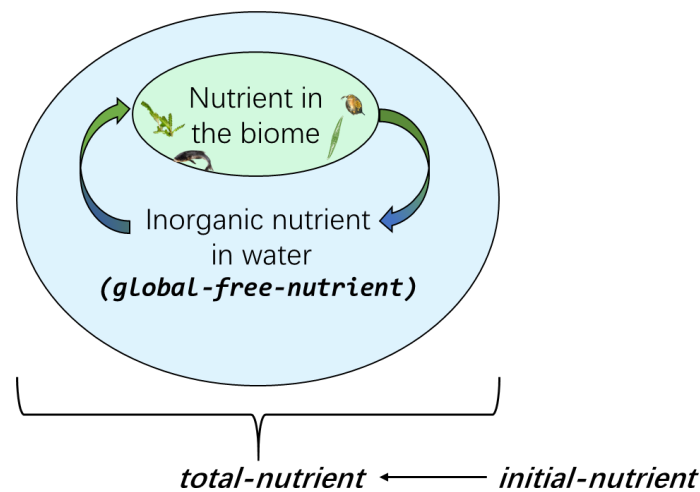


Figure 1-3 A simple diagram showing nutrient cycling and related model variables

Above is a simple diagram showing how the nutrient cycles between the biotic bodies of organisms in the ecosystem and water. The green ovals represent the nutrients in the organisms, possibly for storing energy and participating in vital chemosynthesis, and so

on. Each functional group of organisms has a distinct and fixed nutrient occupancy (C), so that the nutrient content of an individual's body is proportional to its own biomass. In the absence of external inputs, the total nutrient content of the system is fixed, so that any unabsorbed nutrients left in the water are considered to be inorganic nutrients for use (NC_{free}), i.e. the light blue area in the diagram. The nutrients in these two parts dynamically change as organisms grow, reproduce and die, with the total amount remaining constant. The total nutrient content only increases or decreases when there is a nutrient load or reduction in the environment.

Photosynthesis and plant growth

Growth rates of an agent for? phytoplankton and submerged macrophytes is influenced by local light attenuation (I), and global variables of temperature (T) and available nutrient concentration (N).

$$\text{growth_rate} = f(T) * f(I) * f(N) \quad \text{equation 1-2}$$

The dependence of algal growth on temperature applies Arrhenius function, as it is possible to describe the maximal growth rate as a function of temperature (Goldman & Carpenter, 1974).

$$f(T) = Ae^{-\frac{E_a}{RT}} \quad \text{equation 1-3}$$

In equation 1-3, A is the pre-exponential factor, E_a is the activation energy for the reaction, T is the absolute temperature (in Kelvin scale, °K) and R , which is universal gas constant, takes 8.31 J /mol·K. The parameters A, E_a, m were determined by the general

empirical growth rates of freshwater algae. In culture experiments of freshwater algae, the growth rate of *Chlorella vulgaris*, *Fragilaria crotonensis*, *Staurastrum pingue* and *Synechocystis minima* are 1.30, 0.58, 0.77, and 1.32 day⁻¹ respectively and around 1 in general (Dauta et al., 1990). As a result, an empirical set of parameters for algal maximum growth rate were found as:

$$f(T) = 5.35 \times 10^9 \times e^{-6473 / (T_w + 273.15)} \quad \text{equation 1-4}$$

where T_w is the absolute temperature of the lake water (in Kelvin scale, °K) and the value of $f(T)$ is positively correlated with T_w . This equation was also used to describe the growth of submerged macrophyte as an oversimplification.

When plants overcrowd, they form self-shading and the epiphytic community could shade submerged macrophytes to a large extent too (Sand-Jensen & S ndergaard, 1981). In the model without epiphytes, the sum of phytoplankton and vegetation biomass describes the rough distribution and adjusts the light attenuation available to the growing agent.

$$f(I) = \text{lightAttenuation} = e^{-(\sum B_{phy_3} + \sum B_{sm_3})} \quad \text{equation 1-5}$$

where $\sum B_{phy_3}$ means the sum of phytoplankton's biomass in radius of 3 m from this agent and $\sum B_{sm_3}$ means the sum of submerged macrophyte's biomass in radius of 3 m.

The effect of nutrient limitation on the growth rate of algae and plants is described by the Monod model (Goldman & Carpenter, 1974). The general Monod function, or the Michaelis-Menten equation (equation 1-6), can simulate the grow rate of microbes dependent on limiting factors.

$$\frac{1}{X} \frac{dX}{dt} = \mu = \mu_{max} \frac{S}{K_S + S} \quad \text{equation 1-6}$$

where S is the activator, K_S is the activation coefficient, μ_{max} is maximal production rate. In the LAKEOBS_MIX model, global-free-nutrient NC_{free} indicates the availability of limiting nutrient in water and K_S takes 60.

$$f(N) = 6 * NC_{free} / (NC_{free} + 60) \quad \text{equation 1-7}$$

To prevent any overflow of growth rate, a max-growth-rate $1 \times f(N)$ is set for individuals under optimal conditions.

Maintenance and survival

Individual energy for maintenance is roughly equal to BMR (Basic Metabolism Rate), and can be approximated as

$$BMR = B_0 M^\alpha e^{-E/kT} \quad \text{equation 1-8}$$

where B_0 is a constant, M is body mass, α is a scaling coefficient of around 3/4, E is activation energy in eV of around 0.65 (Sibly et al., 2013), T is body temperature in Kelvins scale and k is Boltzmann's constant (van der Vaart et al., 2016). For modelling purpose, the body temperature of agents was in effect not considered. Thus, the logarithm of equation 1-8 to the base of the mathematical constant e is

$$\ln(BMR) = \ln(B_0) + \alpha \ln(M) - E/kT \quad \text{equation 1-9}$$

Since the term $\ln(B_0) - E/kT$ is seen as a constant here, use a instead and have a simpler equation

$$BMR = e^a \times M^\alpha \quad \text{equation 1-10}$$

The BMR in the model is hence calculated in the exponential term, where α is an individual variable that varies between function groups. The range of α is around the empirical value of 0.75, and body mass takes in the empirical range between 0.1 to 100 g.

Grazing and predation

For consumers, eating food provides the income of energy. In a run of consumers' life cycles, each zooplankton agent first searches for a phytoplankton agent on the same square of one hectare (100m × 100m), and graze all phytoplankton in this area if the random result of 'rolling dice' between 1 to 100 is smaller than the *zoopk – hunt – rate*.

The way zooplankton eats food uses *graze-eat* function

$$\Delta energy_{food} = \sum energy_{prey} \quad \text{equation 1-11}$$

Herbivorous fish and piscivores eat one prey at a time, using *fish-eat-eat* function

$$\Delta energy_{food} = \frac{energy_{prey}}{2} \quad \text{equation 1-12}$$

Predating strategies and moving speed allows fierce fish, i.e., piscivores to search for a fish prey in 4-hectare range and move faster to catch other animals. The energy intake by predation is half the energy of prey as the assimilation rate of predation is lower. For planktivorous fish, each agent can eat two phytoplankton agent and one zooplankton agent in the nearby one hectare.

$$\Delta energy_{food} = energy_{prey} \quad \text{equation 1-13}$$

The parameters in grazing and predation indicate the consumers' ability of energy intake and survival. If the energy input from food exceeds the requirements of maintenance and reproduction, then any excess is stored in the animal's energy reserves without cost of

storing. For modelling purpose, surplus energy from food has no limit of appetite but is constrained by parameters of food-intake probability.

Energy reserve

Energy reserves are stored as carbohydrates and fat in tissues of plants and animals, allowing them to maintain functions during periods of coldness or starvation (Sibly et al., 2013). In the end of the modelling day, the leftover energy reserves are calculated and examined with survival criteria. Agents with remaining energy lower than the least energy required ("survival-energy") die due to exhaustion. If energy reserve of an agent is larger than double the initial energy at age 0, an agent may perhaps reproduce and share a half of energy to its only offspring.

The net energy reserved in a day of a phytoplankton or a submerged macrophyte is the subtraction of BMR from the energy gained in photosynthesis as

$$\Delta plantenergy = energy * growthrate * photoperiod - BMR \quad \text{equation 1-14}$$

where the photoperiod is the proportion of light exposure in a day. For zooplankton and fishes, the net energy gained in a run is

$$\Delta zooenergy = E_{foodgain} - BMR \quad \text{equation 1-15}$$

The biomass changes consequently as

$$M = \sqrt{e^{(energy-300)/60}} \quad \text{equation 1-16}$$

1.2.2 Emergence

All the resultant populations, biomasses, death rates, birth rates, and spatial distributions run with functional groups are emergent phenomena. Besides community-level variables, individual-level emergence includes the life spans, energy reserves and mass and ecosystem-level emergent variables, such as time spent to reach an equilibrium, community composition, and numbers of communities that can coexist.

1.2.3 Adaptation

Phytoplankton and submerged macrophytes adapt to available nutrient concentration in water and the density of surrounding plants (in 9 hectare) in growth rate. Phytoplankton have enzymes to repel each other, so they spread out evenly throughout the water towards areas with lower density of algae when one agent has *light-attenuation* lower than 0.3.

1.2.4 Objectives

The objective of the phytoplankton and submerged macrophyte agents in the model is to maximize the energy from sunlight and reproductive success. The objective of the consumer agents is to maximize their intake of digestible food during the swimming process.

1.2.5 Prediction

In natural ecosystems, fish may make decisions under the principle of reward value maximization as found in training experiments (Torigoe et al., 2021). However, for

simplification, agents have no ability to learn from the previous experiences or predict the result of their behaviour in this model.

1.2.6 Sensing

1.2.6.1 *Environment sensing*

In each run, environmental changes driven by SDM would be sensed firstly by phytoplankton and submerged macrophyte, followed by zooplankton and fish. The agents of primary producers sense the nutrient availability, temperature and surrounding density of algae and plants and change their growth rate accordingly. Animal agents in the model do not sense any environmental variable directly but sense the existence of food and refuge shelter formed by submerged macrophytes around themselves.

1.2.6.2 *Food detection*

Animals can sense their prey nearby and then move to the position and catch the prey. Different 'breeds' have different range of detecting food available. Zooplankton, planktivorous fish and herbivorous fish can only detect food within 1m² while piscivorous fish sense food within nearby 4 hectares.

1.2.6.3 *Obstacle detection*

With a macrophyte agent in the nearby 4-hectare neighbourhood, planktivorous fish and piscivorous fish slow down their swimming speed. When meeting the land due to

reclamation or refuge area preventing fish predation in the front, fish turn 90° right and leave the position.

1.2.7 **Interaction**

Submerged macrophyte agents cannot move once sprouting on a patch, and phytoplankton agents spread only in the overcrowded situation. They have no direct interactions with each other. Animals interact with their prey and avoid obstacles.

1.2.8 **Stochasticity**

Multiple elements of stochasticity play a key role in simulating emergent scenarios in the model. The stochasticity exists in: the initial spatial distribution of agents, heading directions in movement, chances to encounter an agent to interact, decisions to eat food or not in a particular day, decisions to reproduce or not, position selection for offspring, availability of food in the surrounding, and obstacles in the surrounding. At the beginning of each run, a certain number of agents in six breeds are created and randomly distributed in water area. When animals meet their food, there is a fixed probability for them to eat and obtain energy using a dice-rolling algorithm. The stochastic distribution of agents in setup stage may decide the final pattern emergence, influenced by stochasticity in movement, reproduction and eating processes.

1.2.9 Collectives

For modelling purpose, the agents' sizes in this model are not realistic and act as "super-individuals", e.g., an agent of phytoplankton represents an algal patch of one hectare and one zooplankton agent can eat this patch up in a day. No collectives are represented in the model.

1.2.10 Observation

Since the emergence of ecosystem state shifts or critical transitions are of interest, the data produced by the model are mainly on a community-level and ecosystem-level. Spatially explicit patch data and agents' attributes of age, energy, mass and growth rate are also aggregated on each day and examined for model verification. In the model setup stage, if 'logData' is selected as 'On', a list of environmental conditions and community descriptions will be collected for every day. The list includes *drivers-pattern*, *total-nutrient*, *global-free-nutrient*, *water-temperature*, *water depth*, reclamation area, lake area, refuge area, population of six groups, biomass of six functional groups with fishes aggregated, number of phytoplankton being eaten, number of zooplankton being eaten, number of submerged macrophyte being eaten, and additive number of herbivorous fish, planktivorous fish and piscivores.

The areas for parameter setting and variable observation were arranged in categories on the model interface (Figure 1-4). The green slider, the green input window, the selector and the switches on the left side of the figure are used to adjust the values and

options of the parameters. The parameters are basically divided into three blocks, which belong to the categories "initial condition", "Biological communities" and "Drivers scenarios". The purple button is the button for executing the command, which can be used for a single or continuous execution of the function.

The observation area includes three kinds of data: text boxes, displays of individual variable values, time series plots, and an option to enable the 'logData' function to export pre-defined variable values for each day of a single simulation directly to an external directory. The text box (on the top left of Figure 1-4) contains a description of the initial basic physical conditions of the lake system and the random seed used in the current simulation (e.g. 1496922334). Single variable values are displayed for relatively important and low-variability variables such as the total-nutrient level of the system, the components of the system and the population of each functional group. Individual variables can give an accurate sense of the quantity. When dynamic changes in the time series are significant, I created three main blocks of graphs. The purpose of designing and classifying these graphs is to facilitate a fast examination of the current state of various dimensions of the ecosystem.

At the top are the more macroscopic graphs of the environment and trophic levels, which include a graph of nutrient changes over time, a graph of total energy changes in the trophic levels, and the current water depth, water temperature, reclamation area, refuge area, and total nutrient concentration. In the middle are observations of population size and biomass, including current population size and population and biomass trends over time for each of the six functional groups. There is also a graphical window in the middle to allow visual observation of the spatial position of individuals in real time in the model system.

Light green dots represent phytoplankton, green plant shapes represent submerged plants, blue worm shapes represent zooplankton, grey fish represent planktivorous fish, blue fish represent herbivorous fish and black fish represent piscivorous fish. The graph window is surrounded by a chronological display (need to preset the starting age), the surface area of the lake (default is 961 hectare) and the number of components in the whole ecosystem (between 0 and 6). The bottom section is about individual properties and interaction changes. The graphs comprise statistics on the number of mortalities (the number of mortalities due to predation and the number of natural mortalities due to energy depletion are counted separately), statistics on the total number of births and statistics on the number of external additions (currently there are only three additions: planktivorous, herbivorous and piscivorous fish species). These observations help the model user to understand the dynamics of the simulated lake ecosystem and the way it responds to environmental change, from the community level to the individual level.

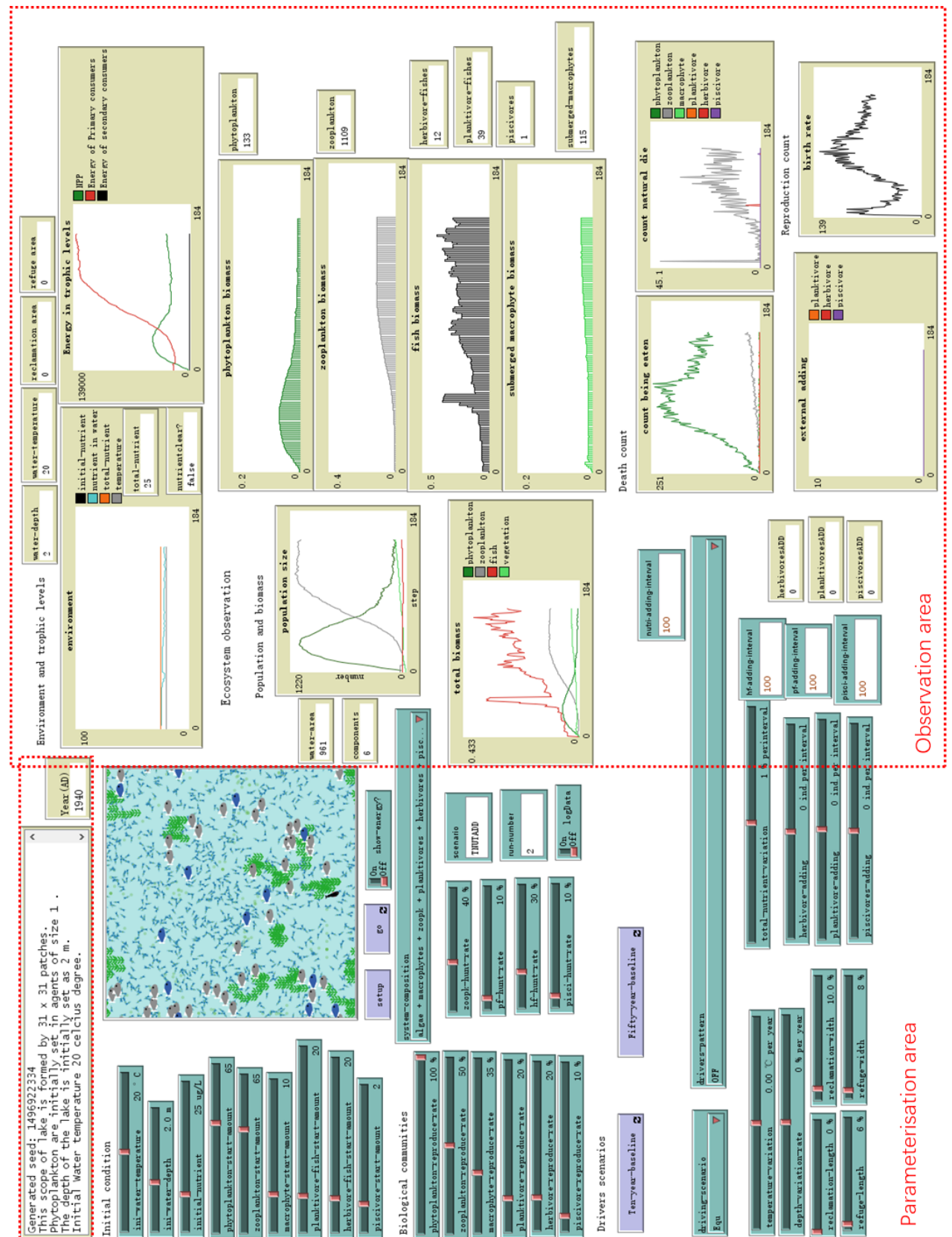


Figure 1-4 The main interface of my ABM model in Netlogo environment in Windows 10 system. In the red dashed boxes are entities for observation, including generated random seed, the current time (Year), and biological and environmental variables.

1.2.11 Initialisation

All parameters are required to be set up before each simulation begins and in the first day in the model, all agents live together (as in a new aquarium) with certain numbers of functional groups with a set environmental condition. In the first 1000 days, the groups cope with the new environment and each other, and gradually become a balanced ecosystem with comparatively stable system composition without external disturbance.

The model does not use input from external sources such as data files. Empirical and theoretical parameters use approximate ranges from literature if possible. After taking a suite of initial values as parameters, no more input is needed. To manipulate the external condition, parameters related to drivers can be set as a following-up input.

1.3 Submodels

1.3.1.1 Environment-change submodel

The environment is the background to the biological activity in the model and is updated first each day. Once the initialisation parameters for the environment are set, the simulation does not require any input unless scenarios for environmental change are set. The input variables to the Environmental-change submodel consist mainly of two selective variables, which determine the names of the parameters to be manipulated, and 14 quantitative variables, which describe the magnitude of each of these variables. As summarised from empirical studies in MLYB, causes for abrupt ecological shifts during 1950s – 1980s may include climate change (as moderating factor), land reclamation, agricultural

intensification and industrialization (Guo et al., 2019; Zhang et al., 2018). In addition, the aquacultural manipulation, i.e. introducing large amount of juvenile planktivorous fish, is evaluated as a significant factor. The overall relationships of these external factors are shown as the coloured blocks in Figure 1-5 in a simplified network resembling realistic environmental impacts.

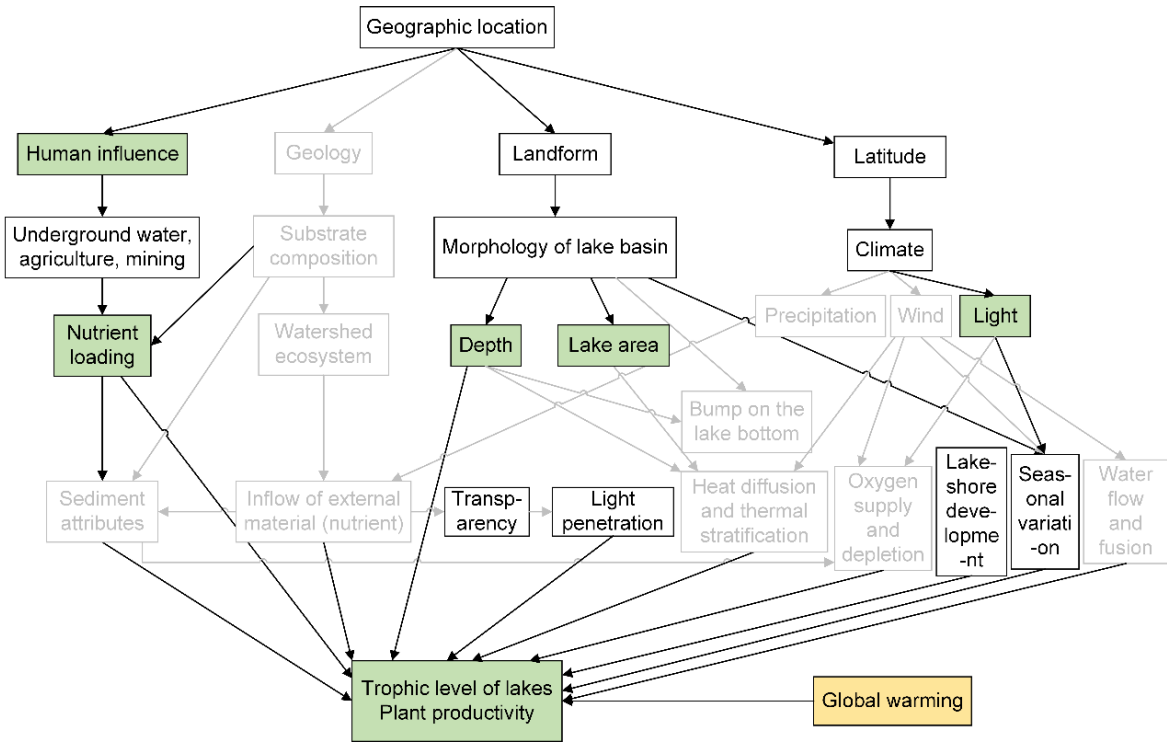


Figure 1-5 External factors considered in the model based on empirical study

The parameters that control the pattern of environmental change in the model are shown in the table below. The “drivers-pattern” contains items to manipulate in the environmental submodel and can be modified according to experiments. After defining what drivers are, model user can quantitatively modify the changing rate of drivers.

Table. 1-5 Variables, value ranges and units in the environmental submodel

Variable	Values/ Ranges	Unit
driving-scenario	"Equ" "Manual" "TBNutADD"	<chr>

	"TBNutADD_TempADD" "TBComprehens" "restoration-1" ...	
drivers-pattern	"modifying piscivores" "nutrient loading + temperature + planktivores + herbivores" "nutrient loading + reclamation + depth + temperature" "nutrient loading + reclamation + depth + temperature + planktivores" "nutrient loading + reclamation + depth + temperature + planktivores + herbivores" ...	<chr>
Temperature-variation	[0, 10]	°C per year
Depth-variation-rate	[-50, 50]	% per year
Reclamation-length	[0, 100]	% max-pxcor
Reclamation-width	[0, 100]	% max-pycor
Refuge-length	[0, 100]	% max-pxcor
Refuge-width	[0, 100]	% max-pycor
Total-nutrient-variation	[-50, 50]	% initial-nutrient per interval
Nutri-adding-interval	[1, ∞]	day
Herbivore-adding	[-∞, ∞]	ind. per interval
Hf-adding-interval	[1, ∞]	day
Planktivore-adding	[-∞, ∞]	ind. per interval
Pf-adding-interval	[1, ∞]	day
Piscovores-adding	[-∞, ∞]	ind. per interval
Pisci-adding-interval	[1, ∞]	day

In SDM, each variable changes separately according to the parameter that is relevant to it. After Δd , i.e. one day, the increment of water temperature is calculated as:

$$\Delta temp = \text{temperature_variation} * \alpha_t * \text{temp} / 3650$$

where $\alpha_t = \begin{cases} 1, & \text{member? "temperature" drivers_pattern} \\ 0, & \text{!member? "temperature" drivers_pattern} \end{cases}$ and temp is the

water temperature before the day. The increment of water depth is

$$\Delta depth = depth_variation_rate * \alpha_d * depth / 36500$$

where $\alpha_d = \begin{cases} 1, & \text{member? "depth" drivers_pattern} \\ 0, & \text{!member? "depth" drivers_pattern} \end{cases}$ and depth is the water

depth before the change. For nutrient load into the lake, the increment of nutrient load is

$$\Delta nutri = total_nutrient_variation * \alpha_n * \beta_n * initial_nutrient / 1000$$

where $\alpha_n = \begin{cases} 1, & \text{member? "nutrient" drivers_pattern} \\ 0, & \text{!member? "nutrient" drivers_pattern} \end{cases}$, $\beta_n =$

$\begin{cases} 1, & \text{Day mod nutri_adding_interval} = 0 \\ 0, & \text{Day mod nutri_adding_interval} \neq 0 \end{cases}$, and Day means the simulation date ("ticks")

in the model.

The above increases (and decreases) in temperature, depth and nutrient can be achieved by adjusting the "temperature_variation", "depth_variation_rate", and "total_nutrient_variation" to positive (or negative) values. Fish placement and removal are implemented in a similar way and with different statements. In the case of "planktivores" (or "herbivores" or "piscivores") in the drivers-pattern In the case of "planktivores" (or "herbivores" or "piscivores"), to add a specific number of fish to the system, set the number of additions to a positive number, and to reduce the number of specific fish from the system, set the number of additions to a negative number. Perform a "pf-adding-interval" (or "hf-adding-interval" or "pisci-adding-interval ") into the system for the number of "planktivore-adding" attributes is initially set for that species of fish. This combination allows flexible customisation of the environmental-changing pattern, for example to simulate a gradual adding pattern set the adding interval to very small with a small number of fish fry at

a time, or to simulate a large number of fry adding pattern set the adding interval to larger with a larger number of fry at a time. To reduce a fish from the system to simulate a fishing scenario, set the number of additions to negative. In the case of planktivorous fish, when the procedure identifies a negative planktivore-adding, a planktivorous fish with an abs (planktivore-adding) number of existing fish is randomly selected to end its life cycle and its nutrient concentration is also subtracted from the system.

To facilitate the simulation of specific scenarios, the driving-scenario parameter selector sets up scenarios of environmental change with different parameter settings, including no change ("Equ"), eutrophic change ("TBNutADD", "TBNutADD_ TempADD", and 'TBComprehens') and de-eutrophic changes ("restoration-*n*"). The 'Equ' scenario is used to test the steady-state equilibrium of a lake ecosystem in the absence of disturbance. The 'Manual' scenario can also be selected for manual parameter adjustment.

1.3.1.2 Primary-producer submodel

The primary producers in the model, i.e. submerged plants and phytoplankton, have essentially similar life processes. In the case of phytoplankton (Figure 4-6), for example, each individual undergoes several processes each day: growth, reproduction, dispersal and death. The process of growth involves photosynthesis and respiration, and the energy accumulated is the remainder of the solar energy assimilated in photosynthesis minus respiration (see section 4.3.2.1 for basic principles). Due to differences in reproduction rates in reality, the reproduction rate of submerged plants is generally set to less than the reproduction rate of phytoplankton (e.g. 30%), and the reproduction rate of phytoplankton is set to 100% by

default. The process of reproduction is that the energy of the parent is halved and the energy of the offspring will be half of the energy of the parent before reproduction, which is basically the same as the initial energy value of the individual. At the end, the remaining energy of all individuals will be checked. If there is an individual that falls below the minimum threshold required for survival, then that individual dies naturally because it cannot sustain itself, and after death its energy disappears and the nutrients return to the water.

Submerged plants differ from phytoplankton in that phytoplankton can move their offspring at birth to neighbouring areas with the lowest density of surrounding phytoplankton when the individual's perceived light attenuation is less than 0.3 (figures 4-7). Submerged plants are limited by their root systems to randomly sprouting new plants in the vicinity of the parent plant.

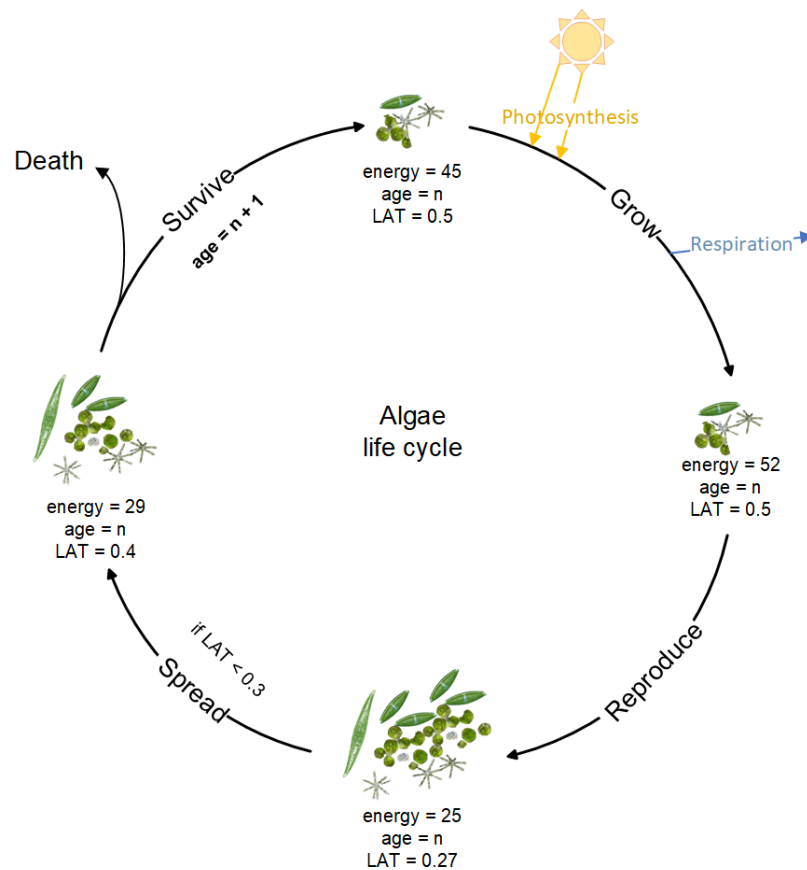


Figure 1-6 Daily life cycle of phytoplankton submodel in the model, developed and adapted from Bronmark & Hansson (2005).

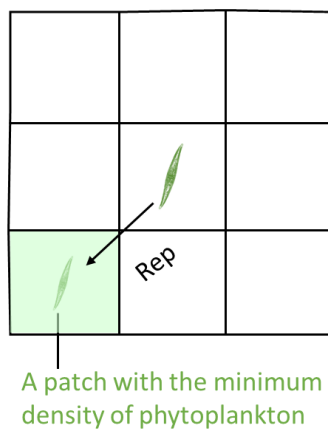


Figure 1-7 Spatial selection of a phytoplankton offspring in reproduction procedure. When the original phytoplankton in the middle has to choose one neighbouring patch for its offspring, it would go through all the surrounding 8 patches and choose a patch with the minimum phytoplankton on it to lower the competition intensity.

1.3.1.3 Zooplankton agent

As primary consumers, the main source of food for zooplankton is phytoplankton and the main energy dissipation is metabolic, with the added risk of being grazed by planktivorous fish. Zooplankton have a certain chance of grazing an individual phytoplankton from within 1 hectare of their location per day, determined by the zooplankton-hunt-rate parameter. The zooplankton move at a speed of 20m/day, with the direction of movement randomly adjusted within 45° of the current direction. When? BUT CHECK MEANING The energy threshold required for reproduction is reached and an offspring can be hatched, the rate of reproductive behaviour is determined by the zooplankton-reproduce-rate.

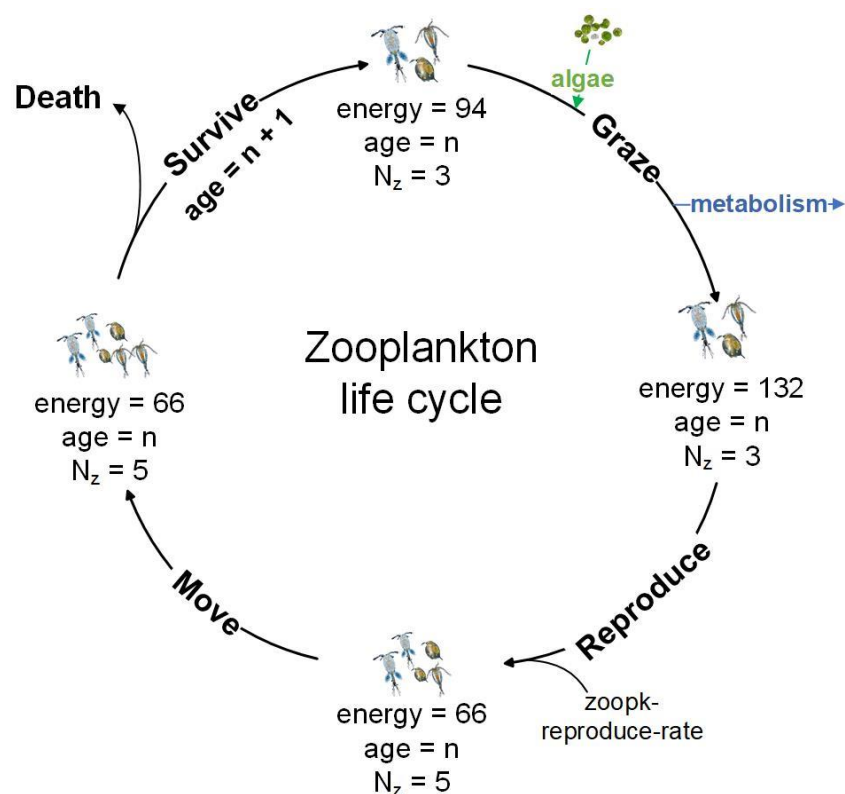


Figure 1-8 A diagram showing the life cycle of zooplankton agents. The diagram shows three agents aged n as the initial state, with an average individual energy reserve of 94 J. After reproduction the number of zooplankton increased to 5, and the average energy per individual decreased while the total energy was almost the same. The age of each

zooplankton increased by 1 after a full day of life cycle. The diagram was developed and adapted from Bronmark & Hansson (2005).

1.3.1.4 Fish submodel

The life cycle of fish agents is similar to that of zooplankton in that they obtain energy through predation and assimilation, metabolic processes consume energy and may be preyed upon by natural predators. The mass of an individual fish-agent is much greater than that of a phytoplankton-agent and therefore the metabolic rate is higher. In the default initial settings, the mass of a phytoplankton-agent is 1 g, the mass of a zooplankton agent is 1.5 g, the mass of a submerged plant-agent is 3 g, while the mass of planktivorous and herbivorous fish-agent is 20 g and the mass of piscivorous fish agent is 40 g. Based on the initial mass settings, the daily metabolic energy of a zooplankton-agent is 0.387 J, while the daily metabolic energy of a piscivorous fish-agent is 1.148 J. The metabolic energy of planktivorous fish-agent is 1.148 J per day and the metabolic energy of piscivorous fish-agent is 2.828 J per day. The metabolic energy consumption increases as the body weight of the individual increases.

The daily diet of planktivorous fish-agents consists of up to two phytoplankton and one zooplankton within a patch (i.e. 1 hectare). The daily diet of herbivorous fish-agent consists of one individual submerged plant within a patch. Piscivorous fish-agent feed daily on other fish-agent present within 200 m of their patch and can only feed on one fish-agent a day, with a limited chance of not feeding within the pisci-hunt-rate parameter. When

submerged vegetation is present within the patch, the movement speed of planktivorous and herbivorous fish-agent decreases to 20 m/day and that of piscivorous fish-agent to 50 m/day. Both are limited by energy thresholds and probability parameters.

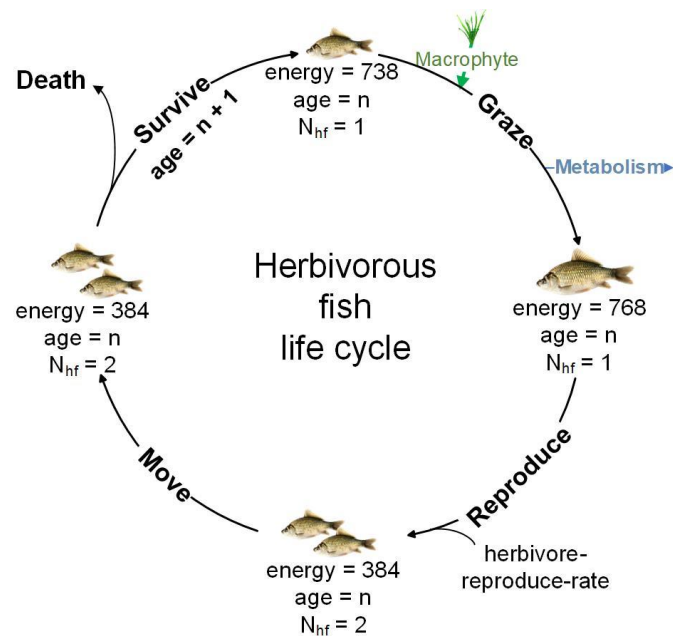


Figure 1-9 The life cycle of an herbivorous fish-agent in one day. This herbivorous fish-agent first gains a net increase in energy of 30 J by grazing on vegetation, so that its energy reserves exceed the threshold for reproduction (e.g. 760 J) and a probabilistic decision is made to reproduce, at which point the fish-agent completes its 'split' into two small fish-agent, each with about half the energy of the previous one. This offspring of the herbivorous fish-agent will then move once away from the mother. If it is incapable of reproducing and its energy reserve falls below the basic value, the fish-agent dies. Fish-agent with energy savings above the basic value increase in age by one day and wait for other individuals to interact. The diagram was developed and adapted from Bronmark & Hansson (2005).

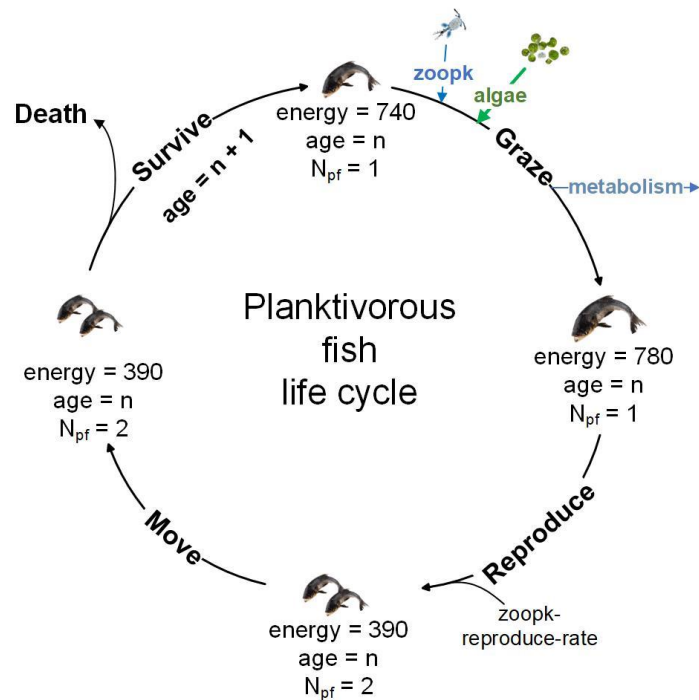


Figure 1-10 The life cycle of an herbivorous fish-agent in one day. The rules a planktivorous fish-agent obey are the same to herbivorous fish-agent in Figure 1-9 except the diet. Planktivorous fish-agent eats 2 phytoplankton agents and 1 zooplankton on spot at maximum. The diagram was developed and adapted from Bronmark & Hansson (2005).

2 REFERENCE

Abar, S., Theodoropoulos, G. K., Lemarinier, P., & O'Hare, G. M. P. (2017). Agent Based

Modelling and Simulation tools: A review of the state-of-art software. *Computer*

Science Review, 24, 13–33. <https://doi.org/10.1016/j.cosrev.2017.03.001>

Bronmark, C., & Hansson, L.-A. (2005). *The Biology of lakes and ponds - Second Edition*.

Dauta, A., Devaux, J., Piquemal, F., & Boumnich, L. (1990). Growth rate of four freshwater algae in relation to light and temperature. *Hydrobiologia*, 207(1), 221–226.

<https://doi.org/10.1007/BF00041459>

Ferber, J. (1999). *Multi-agent systems - an introduction to distributed artificial intelligence*.

Goldman, J. C., & Carpenter, E. J. (1974). A kinetic approach to the effect of temperature on

algal growth. *Limnology and Oceanography*, 19(5), 756–766.

<https://doi.org/10.4319/lo.1974.19.5.0756>

Guo, C., Zhu, G., Qin, B., Zhang, Y., Zhu, M., Xu, H., Chen, Y., & Paerl, H. W. (2019). Climate

exerts a greater modulating effect on the phytoplankton community after 2007 in

eutrophic Lake Taihu, China: Evidence from 25 years of recordings. *Ecological*

Indicators, 105(January), 82–91. <https://doi.org/10.1016/j.ecolind.2019.05.034>

Jeppesen, E., Jensen, J. P., Jensen, C., Faafeng, B., Hessen, D. O., Sondergaard, M., Lauridsen,

T., Brettum, P., & Christoffersen, K. (2003). The impact of nutrient state and lake depth

on top-down control in the pelagic zone of lakes: A study of 466 lakes from the

temperate zone to the arctic. *Ecosystems*, 6(4), 313–325.

<https://doi.org/10.1007/PL00021503>

Li, H., Mynett, A., Penning, E., & Qi, H. (2010). Revealing spatial pattern dynamics in aquatic

ecosystem modelling with Multi-Agent Systems in Lake Veluwe. *Ecological Informatics*,

5(2), 97–107. <https://doi.org/10.1016/j.ecoinf.2009.08.003>

Ministry of Ecology and Environment of PRC. (2020). *Lake Nutrient Standards-Central and*

Eastern Lakes (Total phosphorus, total nitrogen and Chl a) 2020.

<https://doi.org/000014672/2020-01882>

Nanjing Institute of Geographical and Limnological science CAS. (2019). *Survey Report of*

Chinese Lakes [中国湖泊调查报告]. 科学出版社.

<https://books.google.co.uk/books?id=fca2zQEACAAJ>

Sand-Jensen, K., & Søndergaard, M. (1981). Phytoplankton and Epiphyte Development and

Their Shading Effect on Submerged Macrophytes in Lakes of Different Nutrient Status.

Internationale Revue Der Gesamten Hydrobiologie Und Hydrographie, 66(4), 529–552.

<https://doi.org/10.1002/iroh.19810660406>

Scheffer, M. (1998). Ecology of Shallow Lakes. In *Population and Community Biology Series*

(1998th ed., Vol. 68, Issue 1). Chapman and Hall. [https://doi.org/10.1016/S0304-](https://doi.org/10.1016/S0304-3770(00)00103-0)

[3770\(00\)00103-0](https://doi.org/10.1016/S0304-3770(00)00103-0)

Scheffer, M., Baveco, J. M., Deangelis, D. L., Rose, K. A., & Vannes, E. H. (1995). Super-

Individuals a Simple Solution for Modeling Large Populations on an Individual Basis.

Ecological Modelling, 80(2–3), 161–170. [https://doi.org/10.1016/0304-3800\(94\)00055-](https://doi.org/10.1016/0304-3800(94)00055-M)

[M](https://doi.org/10.1016/0304-3800(94)00055-M)

Sibly, R. M., Grimm, V., Martin, B. T., Johnston, A. S. A., Topping, C. J., Calow, P., Nabe-

nielsen, J., Thorbek, P., & Deangelis, D. L. (2013). Representing the acquisition and use

of energy by individuals in agent-based models of animal populations. *Methods in*

Ecology and Evolution, 4, 151–161. <https://doi.org/10.1111/2041-210x.12002>

Torigoe, M., Islam, T., Kakinuma, H., Fung, C. C. A., Isomura, T., Shimazaki, H., Aoki, T., Fukai,

T., & Okamoto, H. (2021). Zebrafish capable of generating future state prediction error

show improved active avoidance behavior in virtual reality. *Nature Communications*, 12(1), 1–21. <https://doi.org/10.1038/s41467-021-26010-7>

van der Vaart, E., Johnston, A. S. A., & Sibly, R. M. (2016). Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. *Ecological Modelling*. <https://doi.org/10.1016/j.ecolmodel.2015.08.012>

Wilensky, U. (1999). *Netlogo*. Center for Connected Learning and Computer-Based Modeling, Northwestern University. <http://ccl.northwestern.edu/netlogo/>

Zhang, K., Yang, X., Xu, M., Lin, Q., Kattel, G., & Shen, J. (2018). Confronting challenges of managing degraded lake ecosystems in the Anthropocene, exemplified from the Yangtze River Basin in China. *Anthropocene*, 24, 30–39. <https://doi.org/10.1016/j.ancene.2018.11.001>