Theory-driven analysis of Ecological data – Day 1

10:30-12:00 **What types of theoretical models in ecology?**

13:45-14:45 **How to build a model?**

14:45-15:45 **How to analyze a model?**

Isabelle Gounand

CESAB, Montpellier,
March 11, 2024

What types of theoretical models in ecology?

Content

1. What system? What question? What hypotheses
2. What model formalism?
    - Deterministic – stochastic processes
    - Time: discrete – continuous
    - Accounting for space?
3. What technical choices?
    - Analytical vs Numerical
    - Agent Based Models vs Equations
4. Some classical models used in ecology

1. What system? What question? What hypotheses?

Global Ecology

Ecology

Population biology

Evolutive biology

Organism biology, physiology, ethology

Cellular and molecular biology

Biochemistry

Biosphere
↑
Biome
↑
Ecoregion

Spatial Homogeneity

Ecosystem
↑
Community
↑
Population
↑
Individual
↑
Cell
Organic Macromolecule
↑
Organic Molecule

Spatial Heterogeneity

Landscape – Metaecosystem
↑
Metacommunity
↑
Metapopulation

What scale?
What process?
[do we want to model]

What can we ignore?

1. What system? What question? What hypotheses?

System + Question → **Scale** → What can we ignore? → What assumptions do we make?
→ **Variable + Processes**

Example: landscape with plants and herbivores

Individuals
Which factors determine individuals development?
Physiology, morphology, behavior etc

Population
How do resources regulate population growth?
Intra-specific competition, pop level rates, etc.

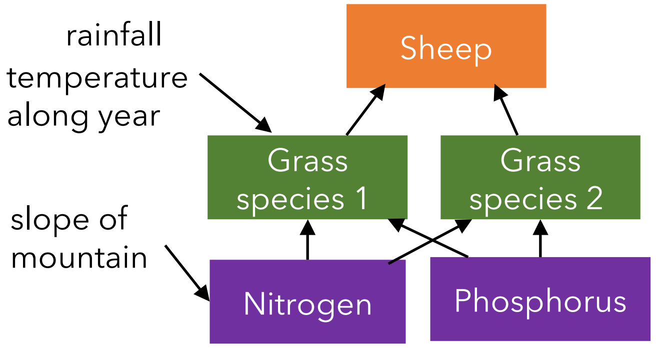
Community
How does grazing impact plant diversity?
Herbivore preferences, plants relative growth, etc.

Ecosystem
Can grazing increase primary production?
Ecosystem fluxes, recycling, etc.

Landscape
Can spatial heterogeneity promote plant diversity?
Spatial connectivity, dispersal rates, etc.

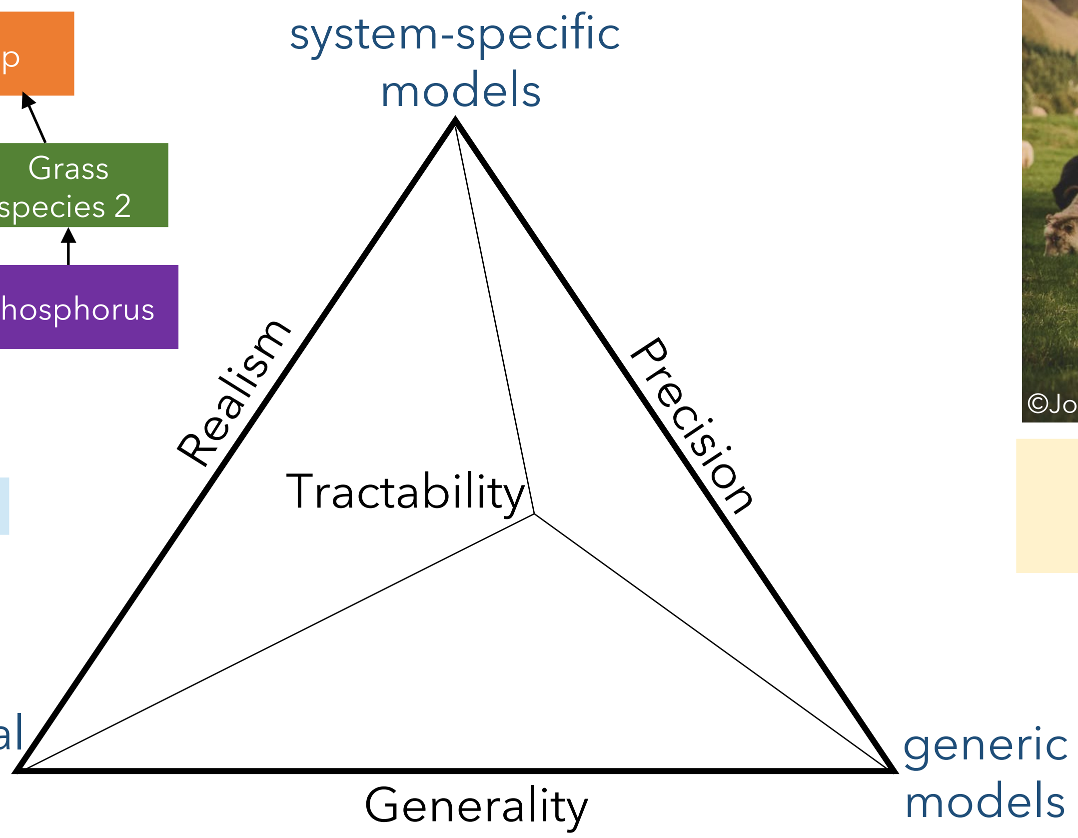
1. What system? What question? What hypotheses?

System + Question → **Scale** → What can we ignore? → What assumptions do we make?
→ **Variable + Processes**

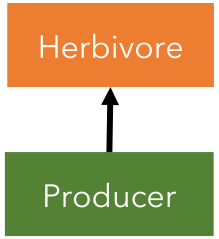


parameter values from specific system

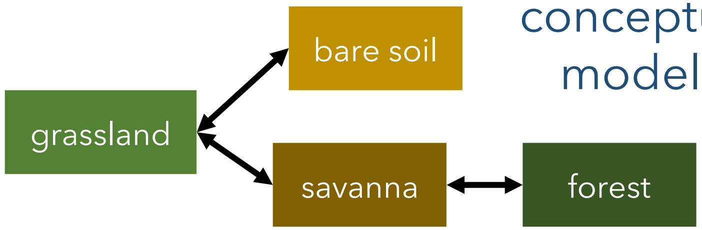
What productivity of grass species 1 this year?



What are the variables and parameters to consider?



How does herbivores regulate plant biomass?



What ecosystem state depending on grazing pressure?

(adapted from Levins 1966, Bullock 2014)

1. What system? What question? What hypotheses?

System + Question → Scale → What can we ignore? → What assumptions do we make?
→ Variable + Processes

This is neither the aim nor relevant to model all details
The upper level is slower than those below and impose constraints
Depending on the system and study question, the dynamics of other levels can be ignored

Physiology question
=> ignore tree dynamics

Long term population dynamics
=> include tree mortality dynamics

Year-scale fish population dynamics
=> ignore human demography

Century-scale fish population dynamics
=> include human demography

1. What system? What question? What hypotheses?

System + Question → Scale → What can we ignore? → **What assumptions do we make?**
→ Variable + Processes

Assumptions → Model → Analysis → Prediction

(simplification of the system False)

(= mathematical formulation of assumptions or series of rules)

(mathematical method)

(truth according to assumptions)

overinterpretation

Types of assumptions

- critical: crucial to test the verbal hypothesis
- exploratory: important to vary and test but not core to the verbal hypothesis
- logistical: those important for tractability

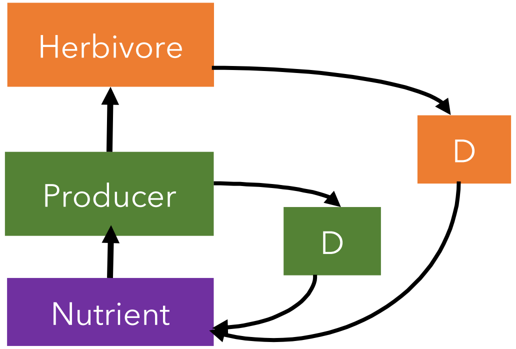
(Servedio et al. 2014)

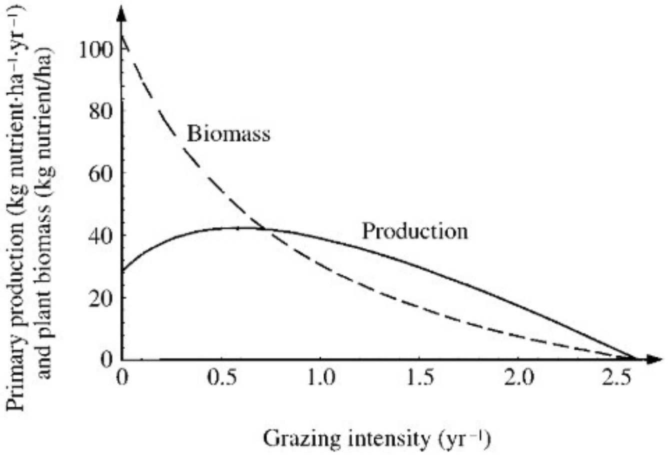
1. What system? What question? What hypotheses?

System + Question → Scale → What can we ignore? → **What assumptions do we make?**
→ Variable + Processes

Question: *Can grazing increase primary production?* (de Mazancourt et al. 1998 Ecology)

Hypothesis: *Herbivory can maximize primary production if herbivore recycling path is faster than plant ones*





Types of assumptions

- critical: crucial to test the verbal hypothesis => 2 paths of recycling
- exploratory: important to vary and test but not core to the verbal hypothesis => functional response (donor vs recipient controlled)
- logistical: those important for tractability => ODE deterministic

(Servedio et al. 2014)

2. What model formalism?

1. Do we need **deterministic or stochastic** dynamics?
2. How do the modelled processes depend on **time** structure?
3. Do we need to consider **space** explicitly?

2. What model formalism? (1) Determinism

What sort of stochasticity counts in ecology?

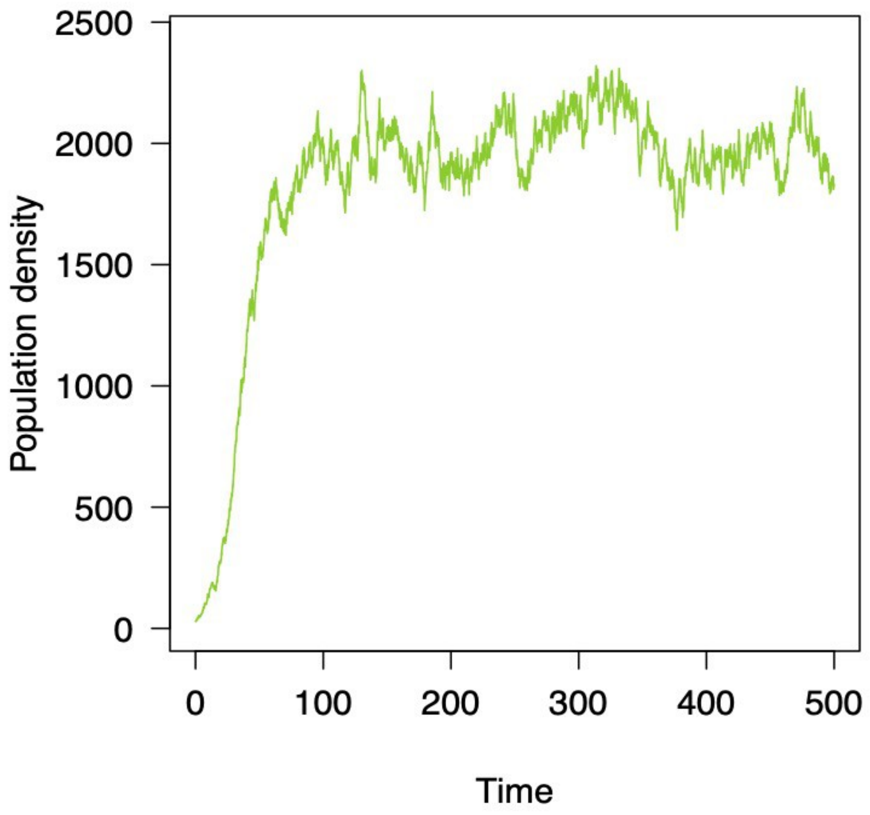
- demographic stochasticity
- environmental stochasticity
- trait variability

When should we account for it?

2. What model formalism? (1) Determinism

Stochastic models

Randomness of processes is important



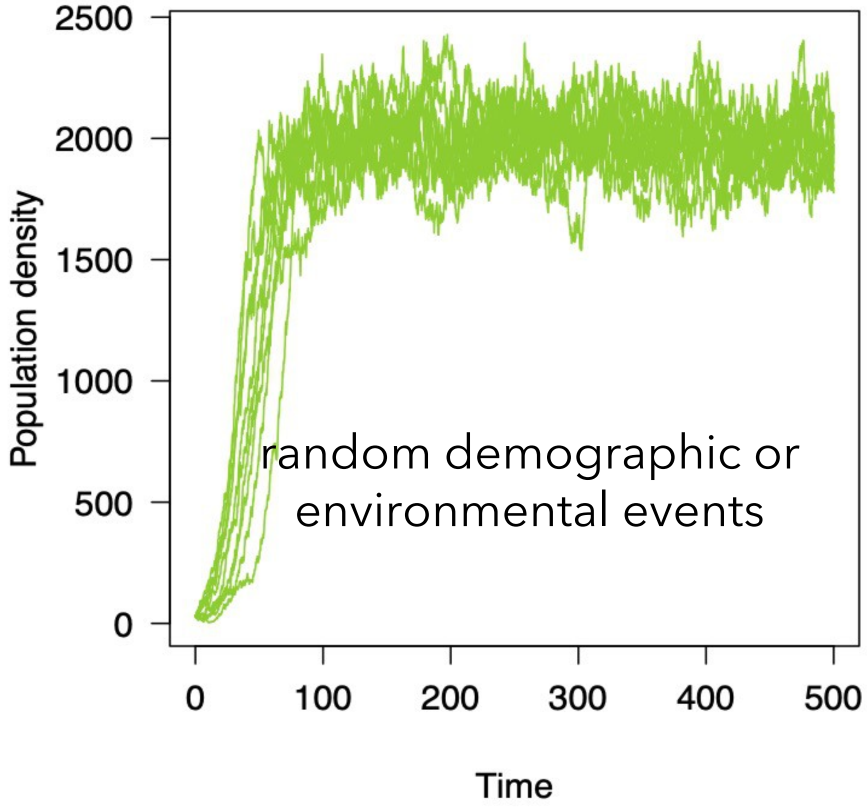
Deterministic models

The noise can be ignored

2. What model formalism? (1) Determinism

Stochastic models

Randomness of processes is important



random demographic or environmental events

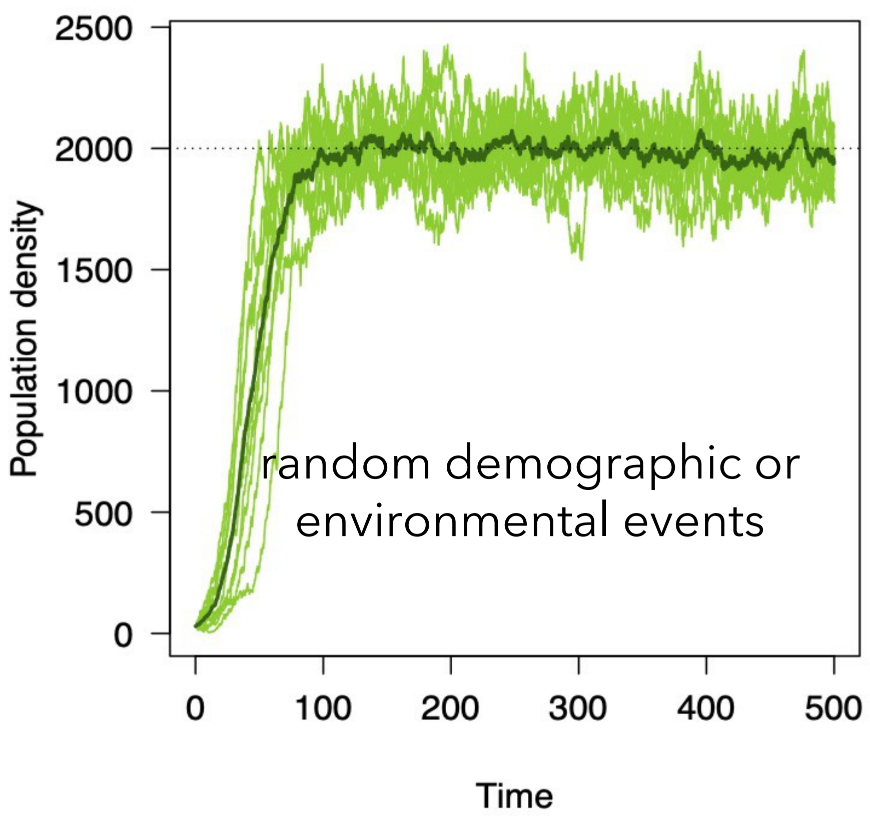
Deterministic models

The noise can be ignored

2. What model formalism? (1) Determinism

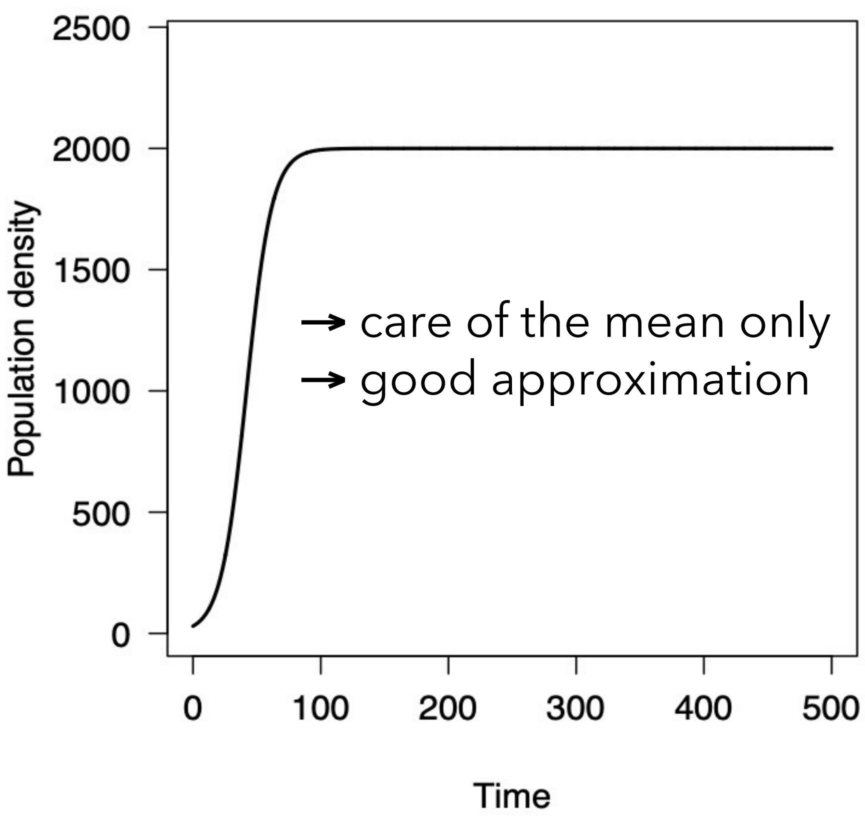
Stochastic models

Randomness of processes is important



Deterministic models

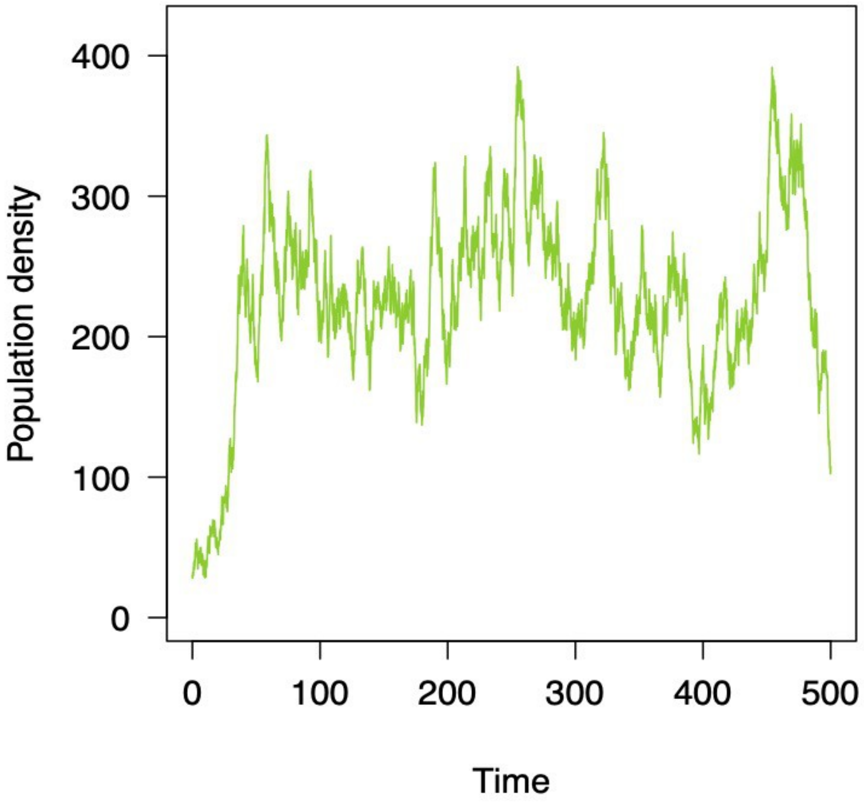
The noise can be ignored



2. What model formalism? (1) Determinism

Stochastic models

Randomness of processes is important



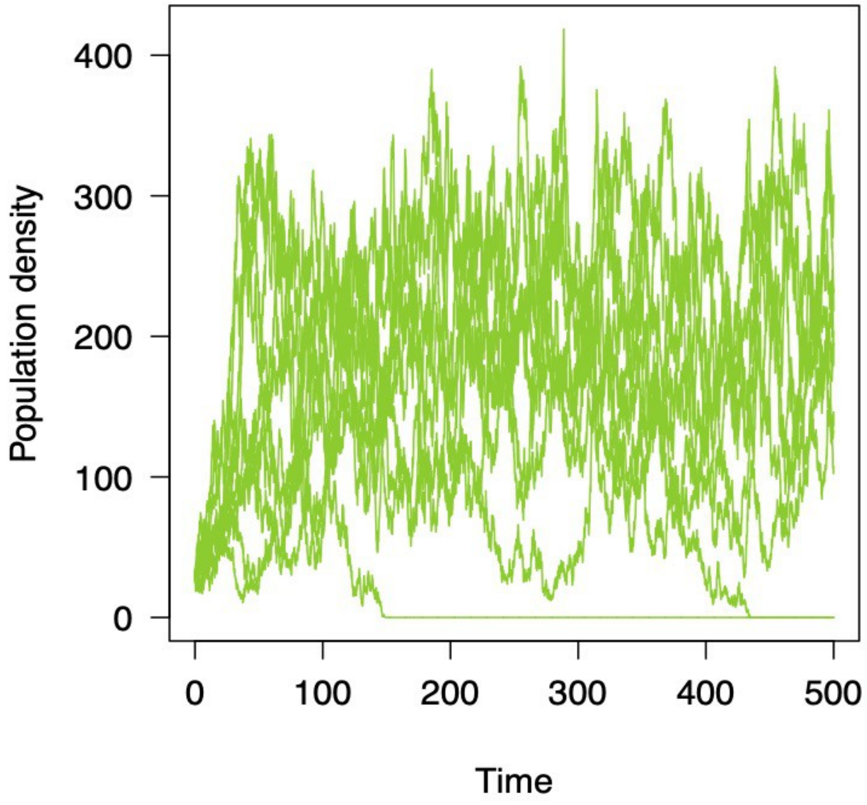
Deterministic models

The noise can be ignored

2. What model formalism? (1) Determinism

Stochastic models

Randomness of processes is important



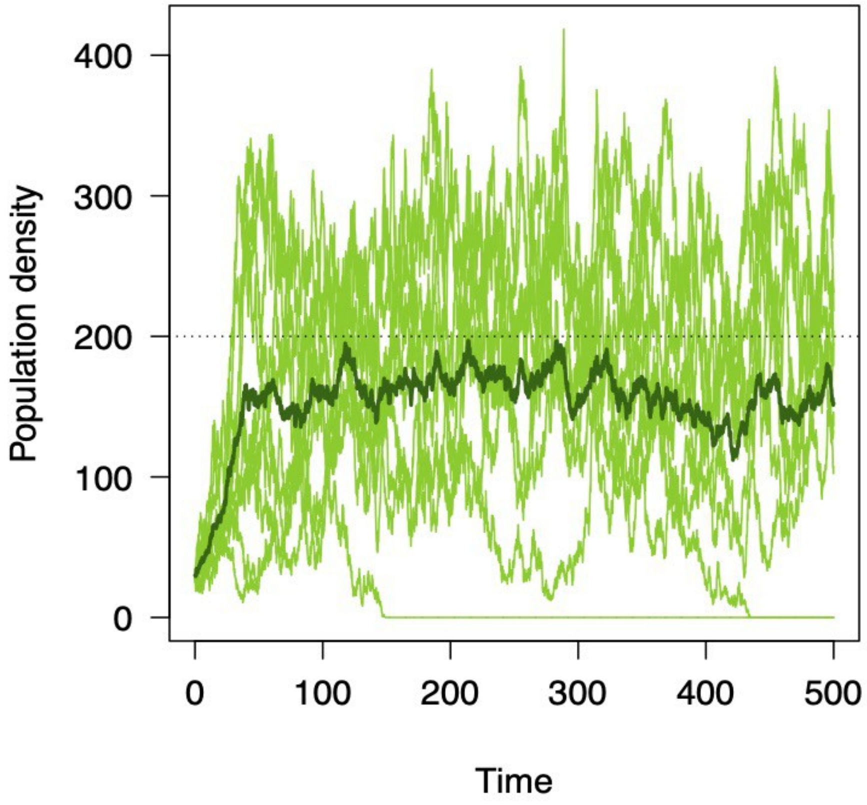
Deterministic models

The noise can be ignored

2. What model formalism? (1) Determinism

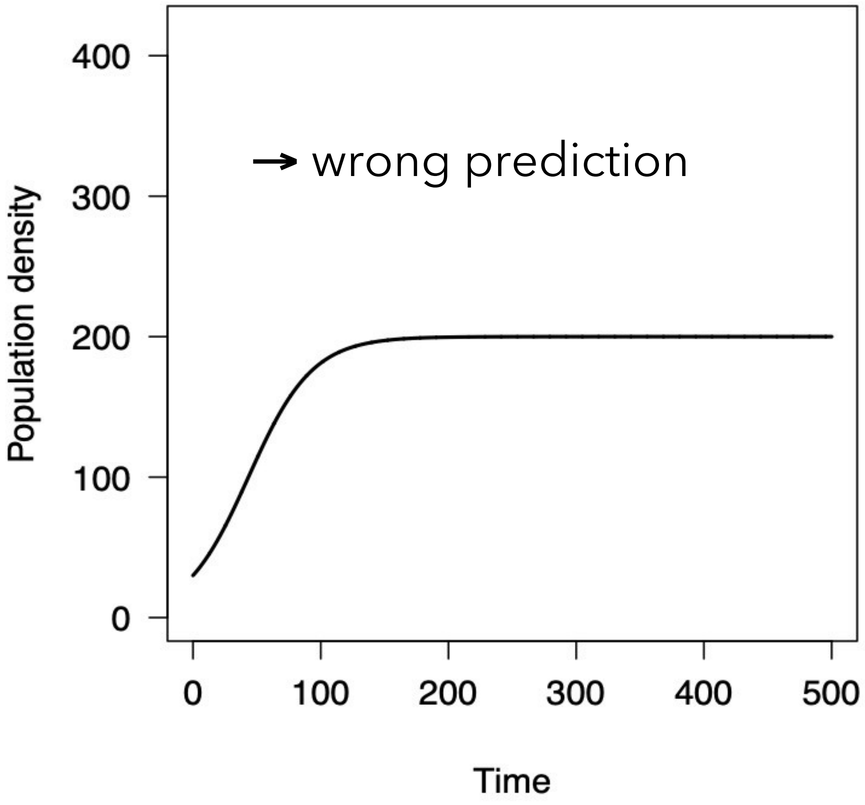
Stochastic models

Randomness of processes is important



Deterministic models

The noise can be ignored



2. What model formalism? (1) Determinism

Stochastic models

Randomness of processes is important

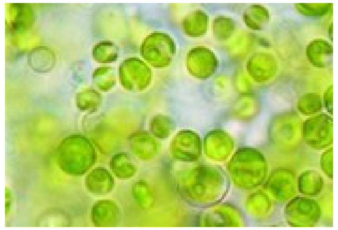
- Neutrality: ecological drift
- Genetic drift (small populations)
- Allee effect
- Demographic stochasticity

→ Questions of conservation
Viability of small populations


Deterministic models

The noise can be ignored

- when processes can be summarised with average parameters: mean growth rate, mass action law
- For large populations

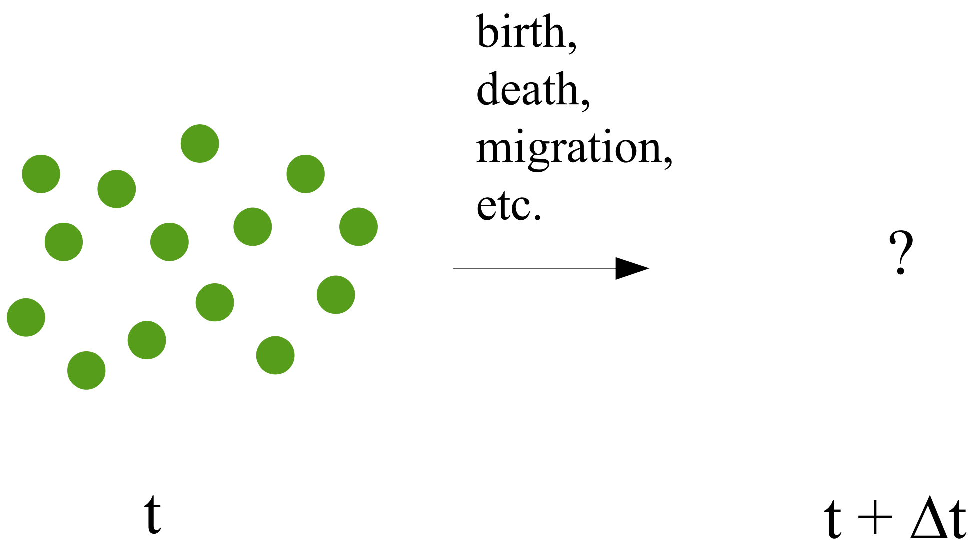



→ NB: Master equations: way to get the deterministic equivalent of stochastic equations

2. What model formalism? (2) Time

How are the modelled processes structured along time?

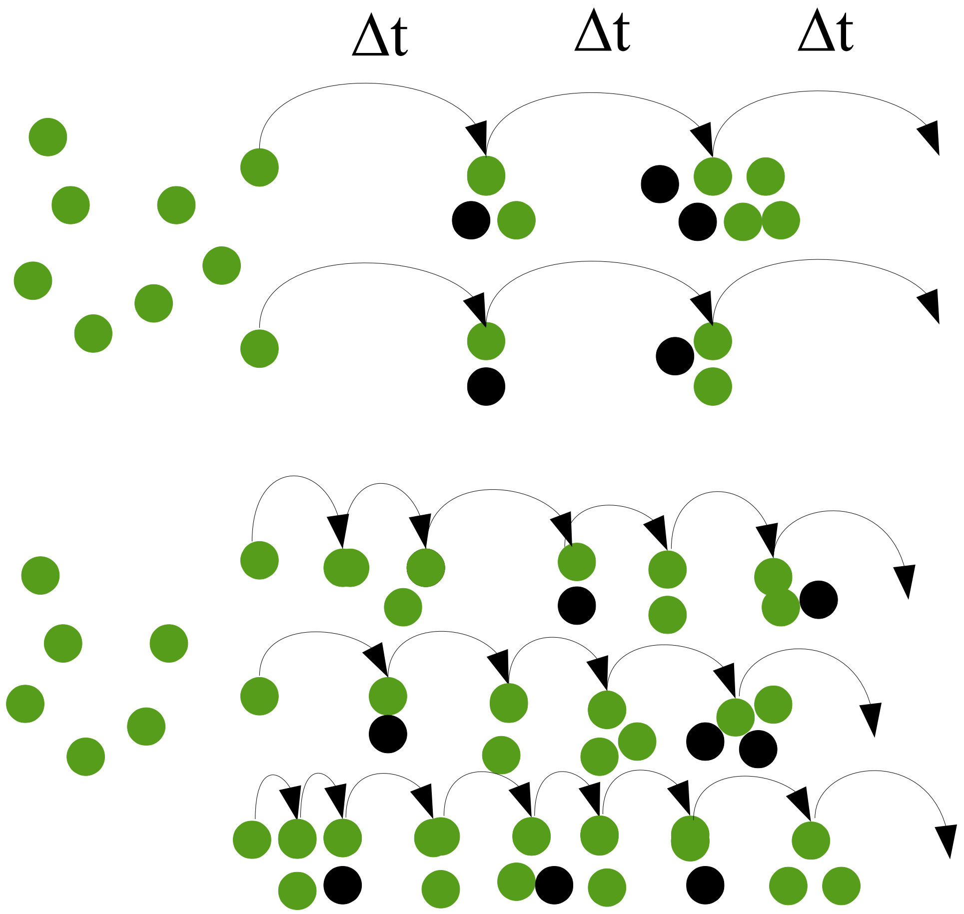
=> When might we want to use discrete or continuous-time formalism?



2. What model formalism? (2) Time

How are the modelled processes structured along time?

=> When might we want to use discrete or continuous-time formalism?



Discrete time:

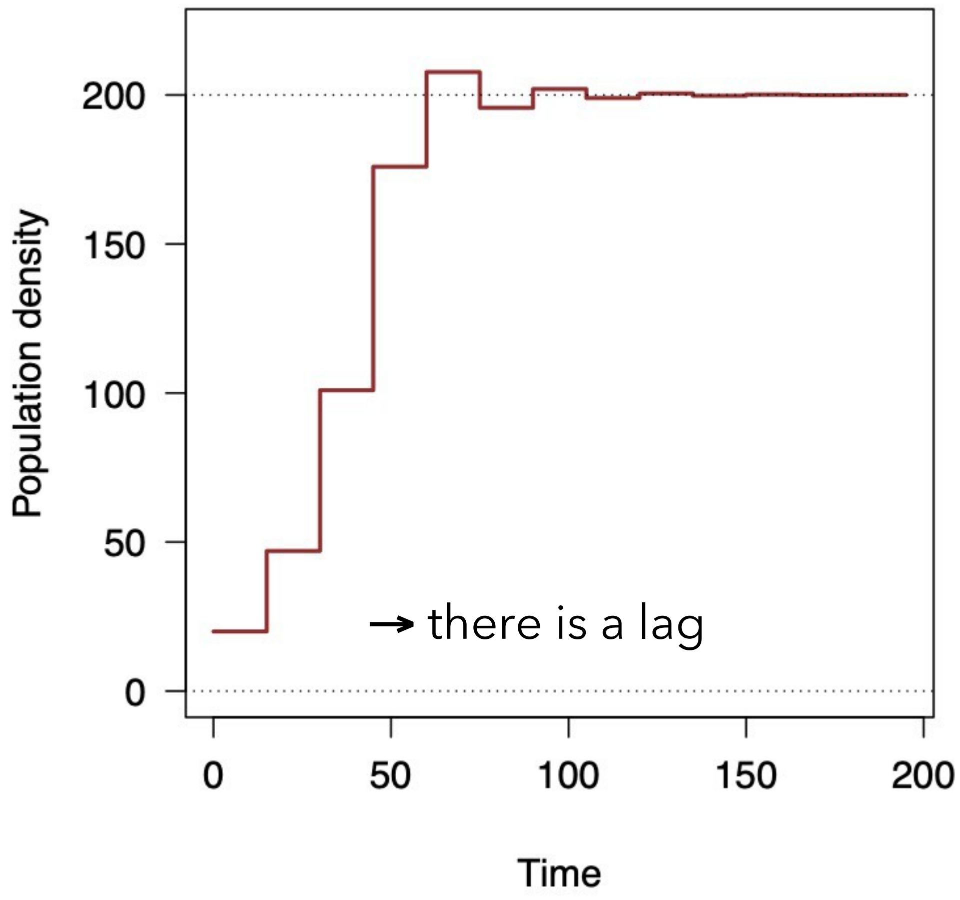
- Fixed generation time
- Synchronization / seasonality
- Sequential processes (e.g.,life-cycle)

Continuous time:

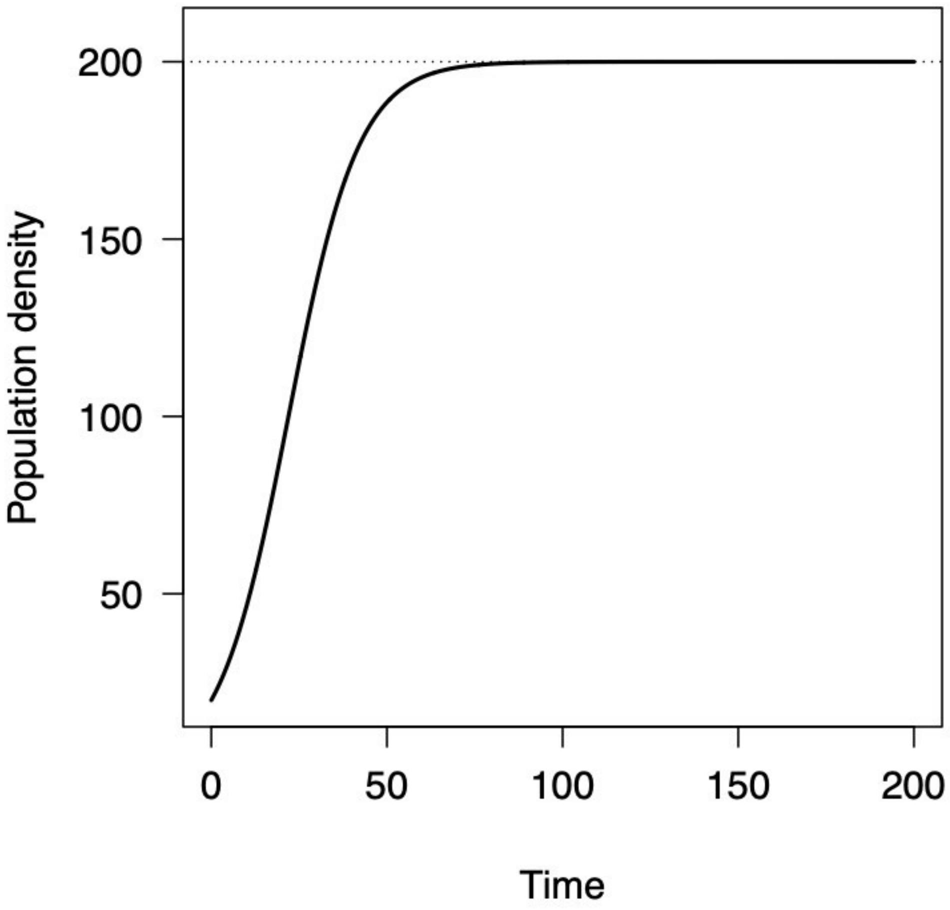
- events happen at any time

2. What model formalism? (2) Time

Discrete time models

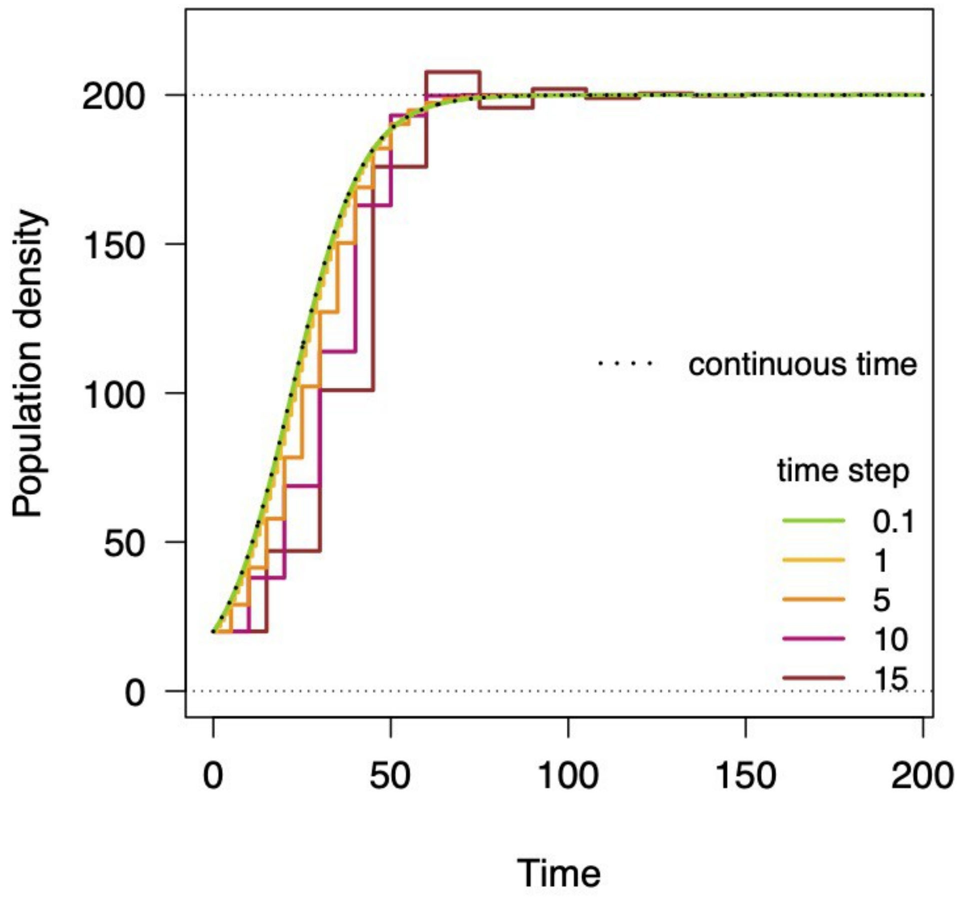


Continuous time models

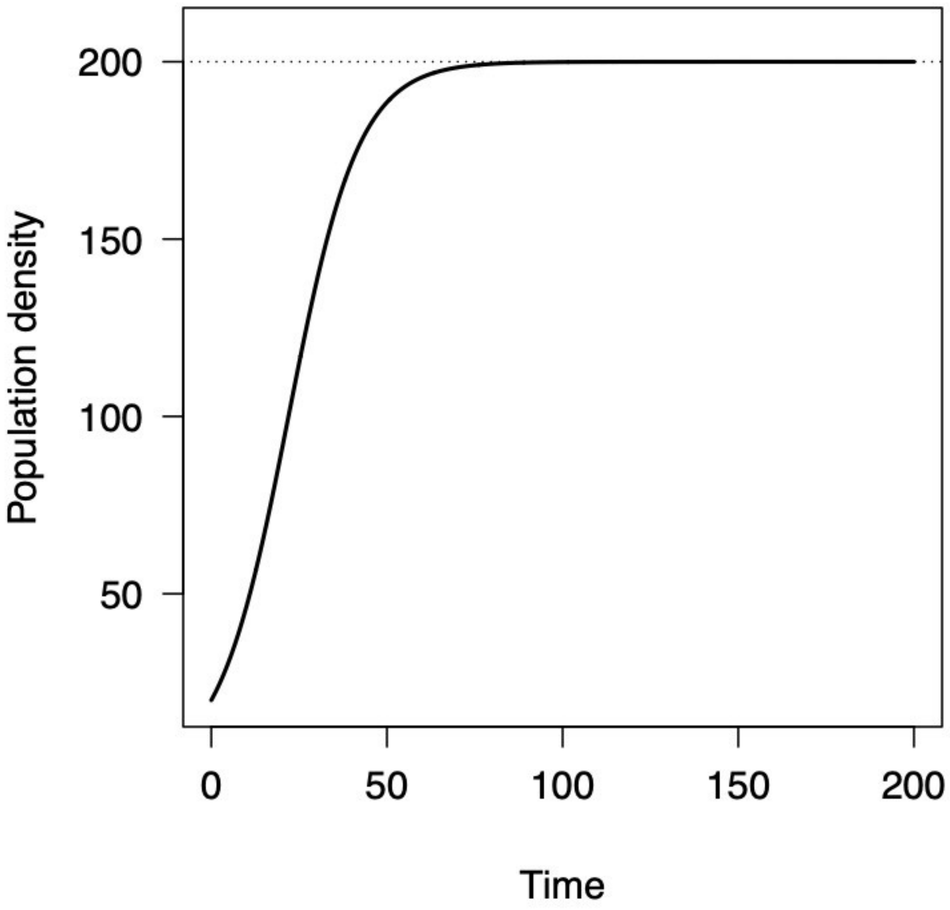


2. What model formalism? (2) Time

Discrete time models



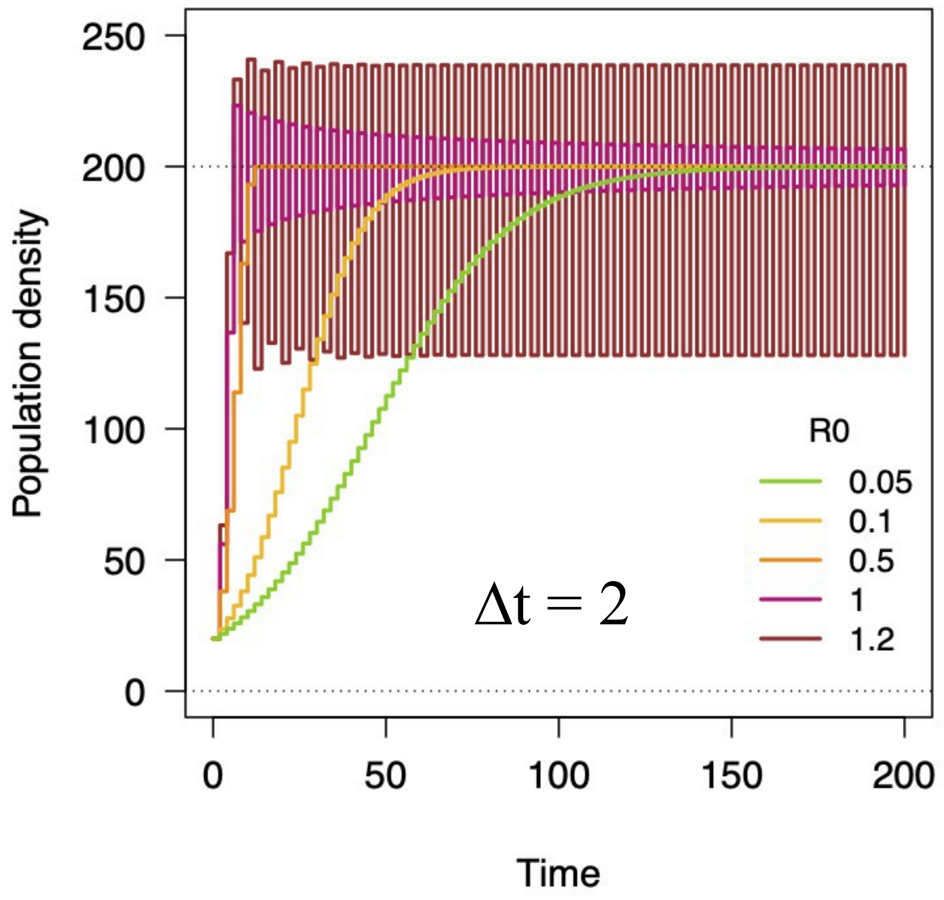
Continuous time models



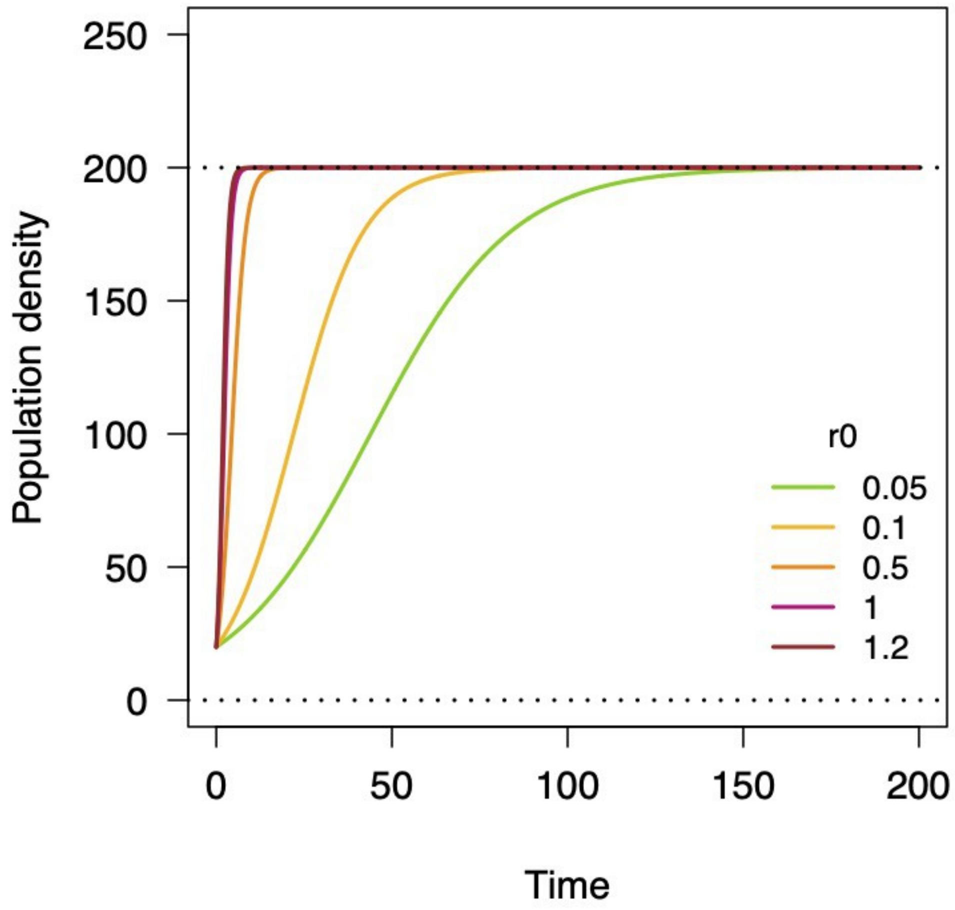
→ small Δt boils down to continuous model

2. What model formalism? (2) Time

Discrete time models



Continuous time models



→ Dynamics can be richer than continuous time model due to the lag

2. What model formalism? (2) Time

Discrete time models

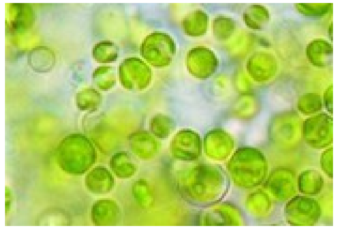
Events are synchronized

- Questions linked to the phenology
- Complex life cycles
- Synchronized generations
- Seasonal dynamics

Continuous time models

Everything can happen at any time

- Process happen continuously
- Generations overlap



→ discrete time models where the time interval is very small boil down to continuous model

→ discrete or continuous time models can be either stochastic or deterministic

2. What model formalism? (3) Space

Is space important to describe your system and answer your question?

Diversity ?

Diversity ~ $f(T, c, e)$ (MacArthur & Wilson 1967 Theory of Island Biogeography)

Space is implicit in colonisation rate c

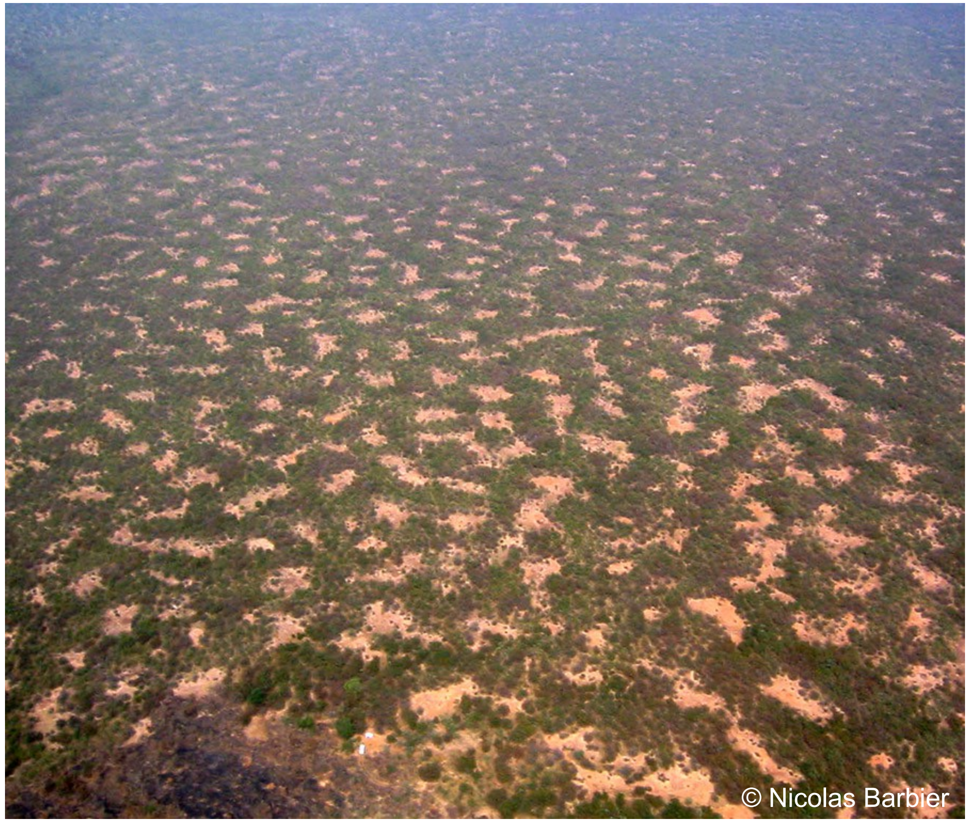
In population models, space is often implicit in the unit, e.g., ind./km^2

2. What model formalism? (3) Space

Is space important to describe your system and answer your question?

Does local dynamics depends on spatial processes and elsewhere local dynamics?

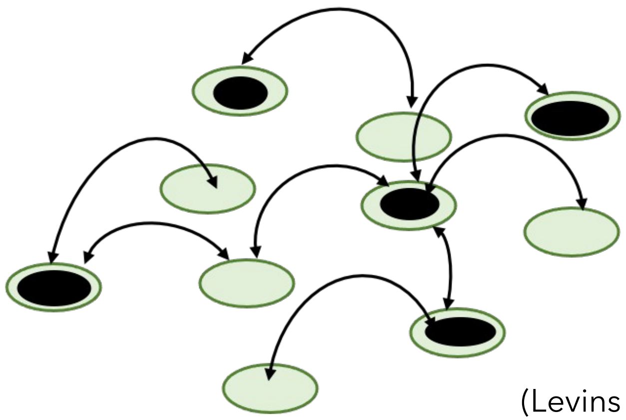
Does spatial patterns emerge from local dynamics?



2. What model formalism? (3) Space

Is space important to describe your system and answer your question?
Does geographical position matter?

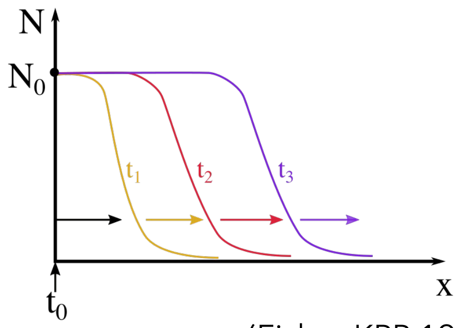
Representation of connectivity only



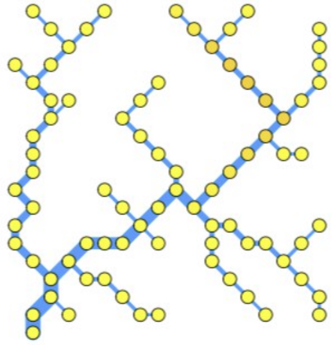
(Levins 1969)

Explicit distances

(Kéfi et al 2007)



(Fisher KPP 1937)



Discrete space

Continuous space

Topology =>
Fragmented landscapes
Connectivity structure effects

Grids =>
Spatial patterns

Continuous space (PDE):
=> Environmental gradient, edge effects, invasion front

3. What technical choices?

1. Analytical vs Numerical
2. Agent Based Models vs Equations

3. What technical choices? (1) Analytical vs simulations

Analytical versus simulation models → Parcimony provides analytical power

When the model has a solution:
We fully know the model behaviour at any time point:

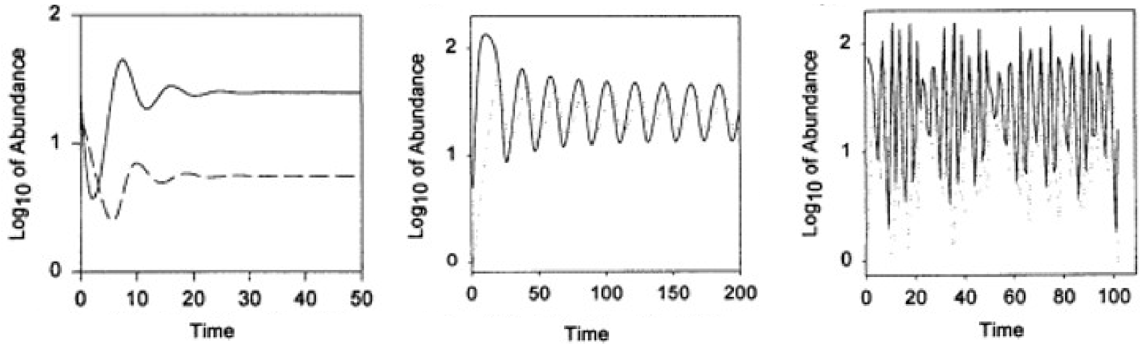
$$\frac{dN}{dt} = f(N) \rightarrow N = f(t)$$

When the model is intractable:
You need to run simulations;

- numerical algorithm (IBM)
- numerical integration (approximation)

Results are specific to the parameter values and to your initial conditions

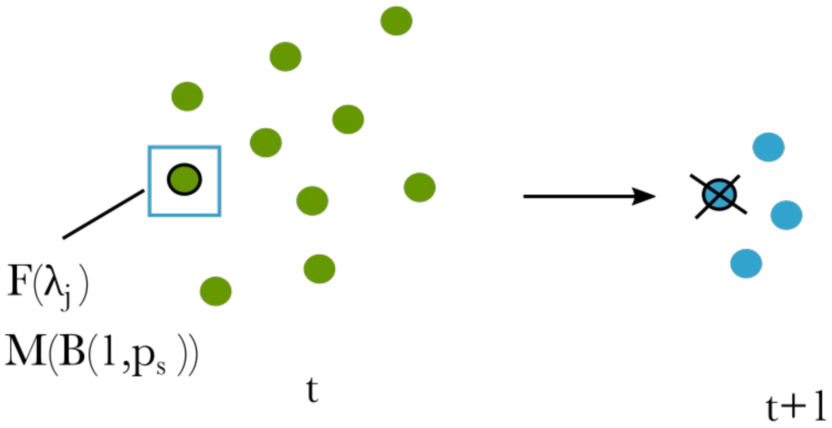
When the model is tractable:
You can describe model behaviour at equilibrium with the ***parameters***.



You can get read of the sensitivity to initial conditions and predict long term dynamics.

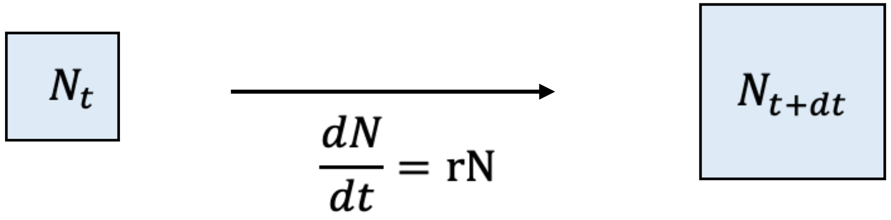
3. What technical choices? (2) rules vs maths

IBM – ABM
Individual-based Modeling



- Variables are individuals or agents
- Processes are formulated as a series of rules involving probabilities, applied to each individual. for example:
  - Number of offsprings from a Poisson distribution
  - Survival probability from a Binomial distribution
  - Other rules for movement, interactions, etc.

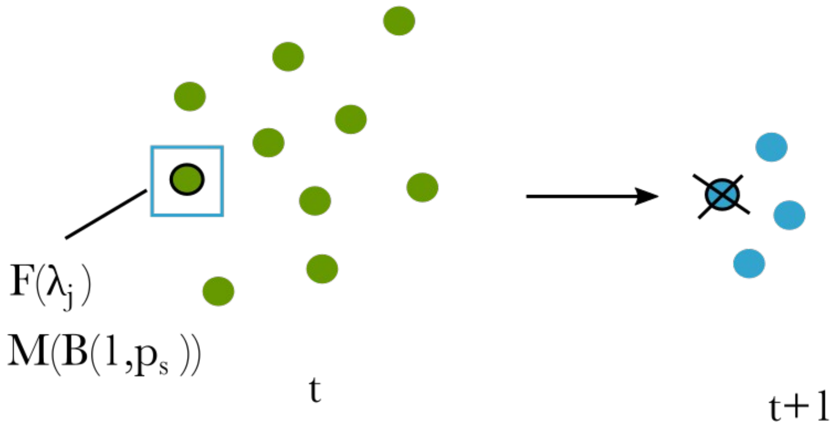
DEs
DE Difference equations
ODE Ordinary Differential Equations
SDE Stochastic Differential Equations
PDE Partial Differential Equations…



- Variables are population densities / biomasses
- We use Math
- Processes are embedded into mean parameters
- What we model is :
  - ODE: the change rate of the variables over time
  - DE: the state of the variable at the next time point
  - PDE: the change over time and space

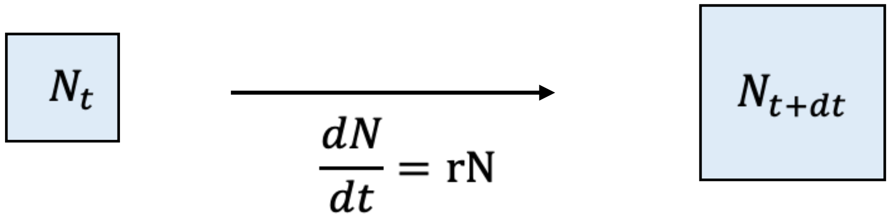
3. What technical choices? (2) rules vs maths

IBM – ABM
Individual-based Modeling



- The dynamics emerge from elemental processes
- Simpler to build from empirical knowledge
- No need for math skills
- High computation consumption
- Rarely tractable

DEs
DE Difference equations
ODE Ordinary Differential Equations
SDE Stochastic Differential Equations
PDE Partial Differential Equations...



- Simplification with math approximations
- Large analysis power for extreme cases
- Fast computation
- Easier to fit to data
- Imposed relations between variables
- Math skill needed

3. What technical choices? (2) rules vs maths

IBM
Individual-based Modeling

In which cases IBMs are particularly relevant?

- when there are not too many parameters
- when stochastic processes are dominant (ecological or genetic drift, Allee effect, demographic stochasticity)
  → **Questions of conservation**
  Viability of small populations

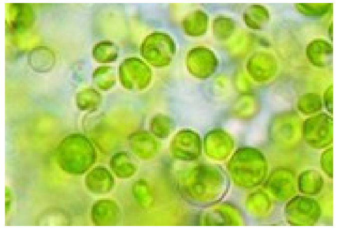


- In domain where processes are hard to synthesise (e.g., behavioral ecology)
- If math skills are lacking ☺

DEs
DE Difference equations
ODE Ordinary Differential Equations
SDE Stochastic Differential Equations
PDE Partial Differential Equations...

In which case DE's are particularly relevant?

- whenever it's possible (analytical power + lower carbon footprint):
  - When processes can be synthesized with average parameters
  - => for relatively large populations




4. Some classical models used in ecology and seen in the next days

- Systems of differential equations without space :

  - Verhulst (logistic growth), Lotka-Volterra predator-prey (Day 2)

  - Lotka-Volterra, food web niche model (Day 4)

- Spatial systems of differential equations: TIB, Levins' occupancy (Day3)
- Spatial IBM: Neutral model of biodiversity from Hubbel (Day 3)

How to build a model?

Content

1. Sketch your system and choose your formalism

2. Identify the assumptions in a classical theoretical model

3. Code the model in R: principle of numerical integration

4. Explore the model

1. Sketch your system

What are your variables?

How are they connected? Which processes do you integrate?

And Choose your formalism

What formalisms in terms of determinism, time, space?

What assumptions on modelled processes?

1. Sketch your system and choose your formalism

When do they coexist?

1. Sketch your system and choose your formalism

Producers-herbivores
in a different ecosystem

When do they coexist?

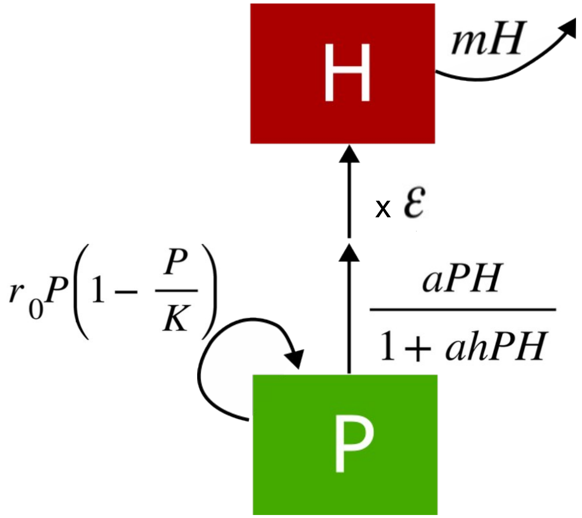
What type of dynamics (stable, oscillating)?

2. Identify assumptions in theoretical models

The Rosenzweig-MacArthur model (1963)

$$\begin{cases} \dfrac{dP}{dt} = r_0 P\left(1 - \dfrac{P}{K}\right) - \dfrac{aPH}{1 + ahPH} \\ \dfrac{dH}{dt} = \varepsilon \dfrac{aPH}{1 + ahPH} - mH \end{cases}$$

r_0 growth rate
K carrying capacity
a attack rate
h handling time
m mortality rate
ε conversion efficiency



General assumptions from formalism

- Populations are sufficiently large for their biological rates to be approximated with averaged parameters: within a population, all individuals identical
- Generations overlaps in time
- Space is homogeneous

Assumptions from mathematical formulations

- Resources for producers are limited and resource dynamics are much faster than population dynamics
- There is no recycling feedback
- Mass action law: encounter rates are proportional to densities
- Herbivore consumption saturates through time needed to manipulate food
- Herbivores dies without producers (metabolic needs)
- Only a part of herbivore consumption is converted into new biomass

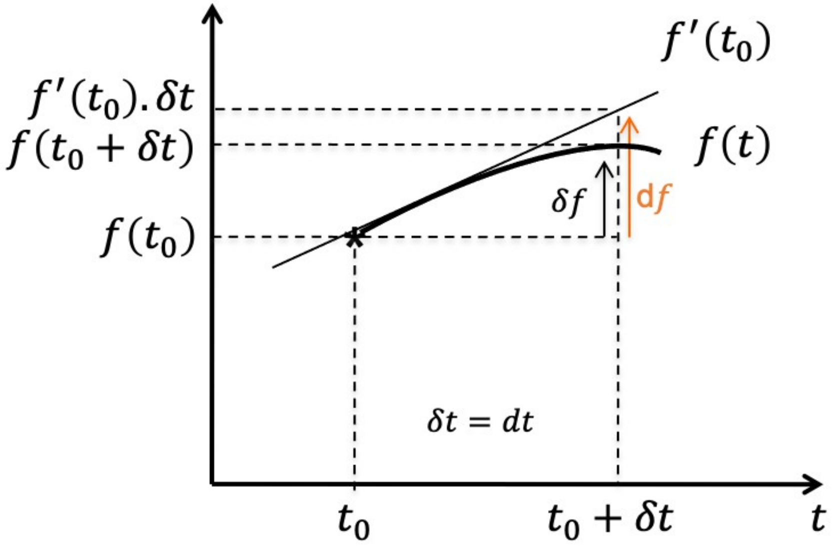
3. Code the model: principle of numerical integration

- Recursive process: approximate the system from the previous time step and the derivative

- Simple example: the Euler method:

$$\tilde{f}(t) = f(t_0 + \delta t) = f(t_0) + f'(t_0).\delta t + \varepsilon$$

- The error depends on time interval and the type of dynamics

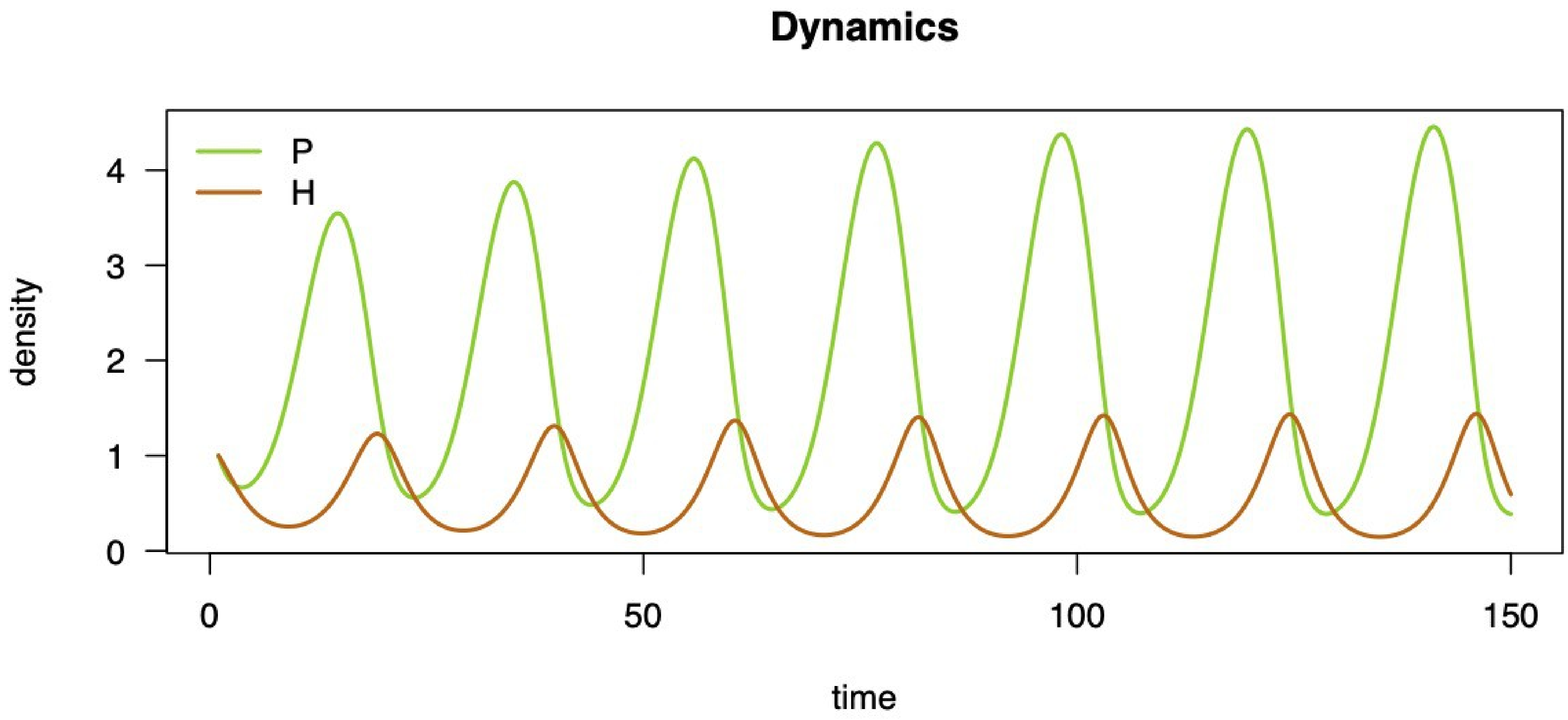


- Mathematicians proposed different algorithms to minimize the error depending on the problem.

- These algorithm are implemented into solvers. Some have adaptive time steps with an error tolerance.

In R we can use the function `ode` of the package `deSolve`

3. Code the model: principle of numerical integration



- In R we can use the function `ode` of the package `deSolve`

4. Explore the model

- Modify the initial conditions. Is the long term result changing?

- Modify the parameters

- **Which strategy to explore the model and answer our question?**

How to analyze a theoretical model?

Content

1. General analysis of the model's behavior
    - Equilibria
    - Graphical representations isoclines, phase plane
    - Local stability analysis (Jacobian matrix)
    - Bifurcation diagrams
    - Dependance to initial conditions

2. Model strategies to answer diverse types of questions
    - Parameter variation
    - Model comparison
    - Experiments with synthetic data

3. Parameter exploration and robustness of conclusions

1. General analysis (1) Equilibria

- Determine the Equilibria (long-term dynamics): solve

$$\begin{cases} \dfrac{dP}{dt} = 0 \\ \dfrac{dH}{dt} = 0 \end{cases}$$

$$P^* = 0,\ H^* = 0$$

$$P^* = K,\ H^* = 0$$

$$\begin{cases} P^* = \dfrac{m}{a(\varepsilon - hm)} \\ H^* = \dfrac{\varepsilon r_0 (aK(\varepsilon - hm) - m)}{a^2 K (\varepsilon - hm)^2} \end{cases}$$

- When tractable, expresses P* and H* with the parameters (symbols) Ụ general expression

- symbolic calculus (Maxima, Mathematica, Matlab)
- In R we can get the numerical calculation of equilibria with the function `stode` of the package `rootSolve` or with the function `searchZeros` of the package `nleqslv`

- Feasibility criteria
- Interpretation on parameters

1. General analysis (2) Graphical representations

A synthetic representation of the system for two variables: Phase plane with null clines to represent the equilibria and the trajectories (see attractors ≠ repellers)

Nullclines for H growth:

$$H = 0$$

$$P = \frac{m}{a(\varepsilon - hm)}$$

Nullclines for P growth:

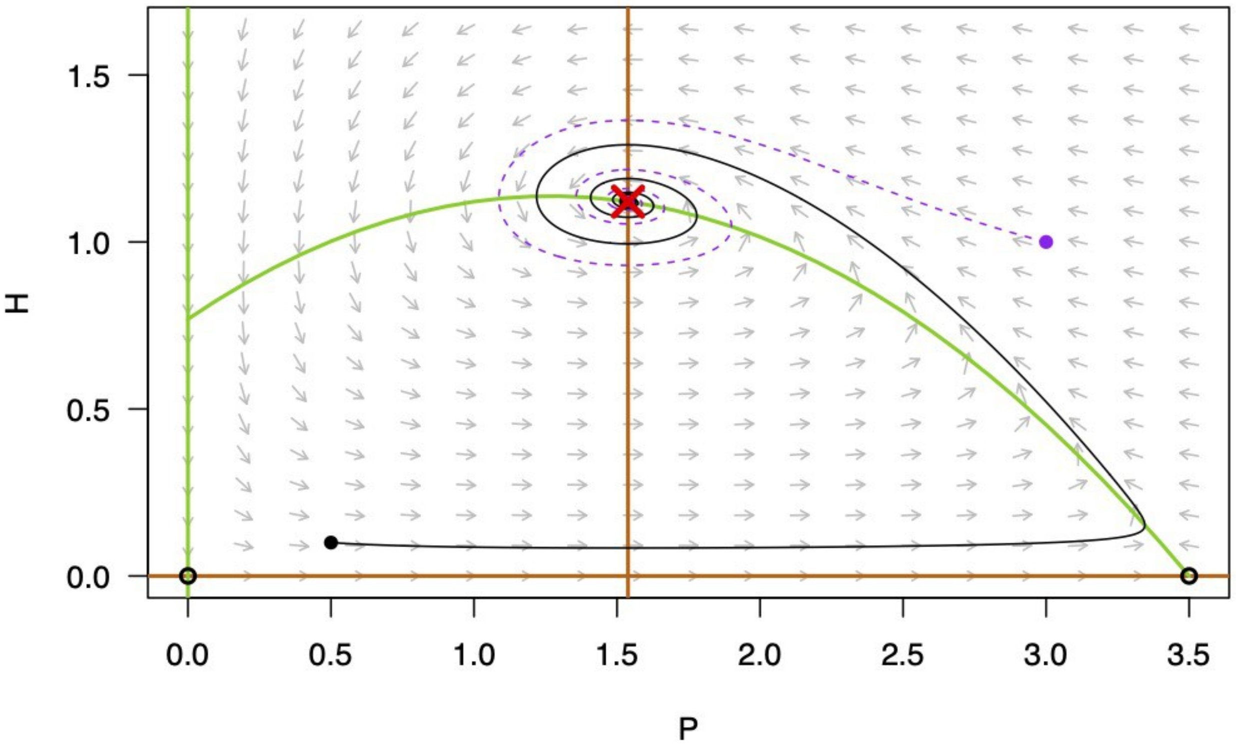
$$P = 0$$

$$H = r_0\left(\frac{1 + ahP}{a}\right)\left(1 - \frac{P}{K}\right)$$

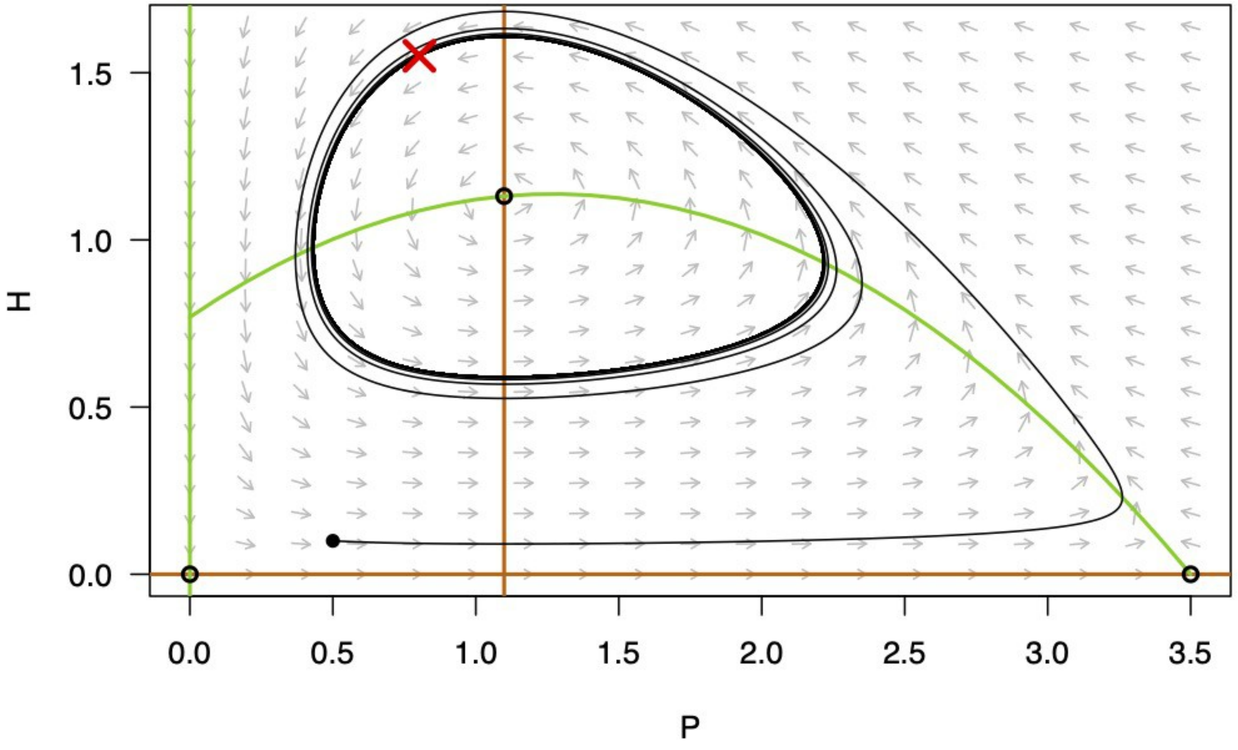
1. General analysis (2) Graphical representations

A synthetic representation of the system for two variables: Phase plane with null clines to represent the equilibria and the trajectories (see attractors ≠ repellers)

Stable equilibrium

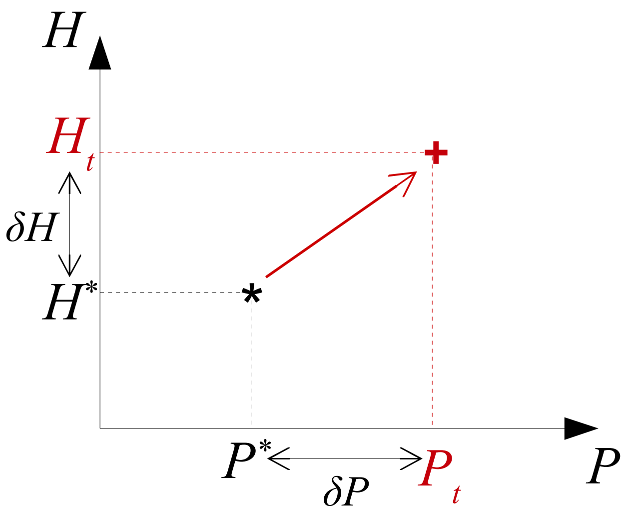


Oscillations



1. General analysis (3) Local stability analysis

- Determine the stability of each equilibrium by analyzing the Jacobian matrix at the equilibrium



$$\begin{pmatrix} \frac{dP_t}{dt} \\ \frac{dH_t}{dt} \end{pmatrix} = J \cdot v_t \quad \text{with} \quad v_t = \begin{pmatrix} \delta P \\ \delta H \end{pmatrix}$$

$$J = \begin{bmatrix} \frac{\partial \frac{dP}{dt}}{\partial P} & \frac{\partial \frac{dP}{dt}}{\partial H} \\ \frac{\partial \frac{dH}{dt}}{\partial P} & \frac{\partial \frac{dH}{dt}}{\partial H} \end{bmatrix}_{P^*,\, H^*}$$

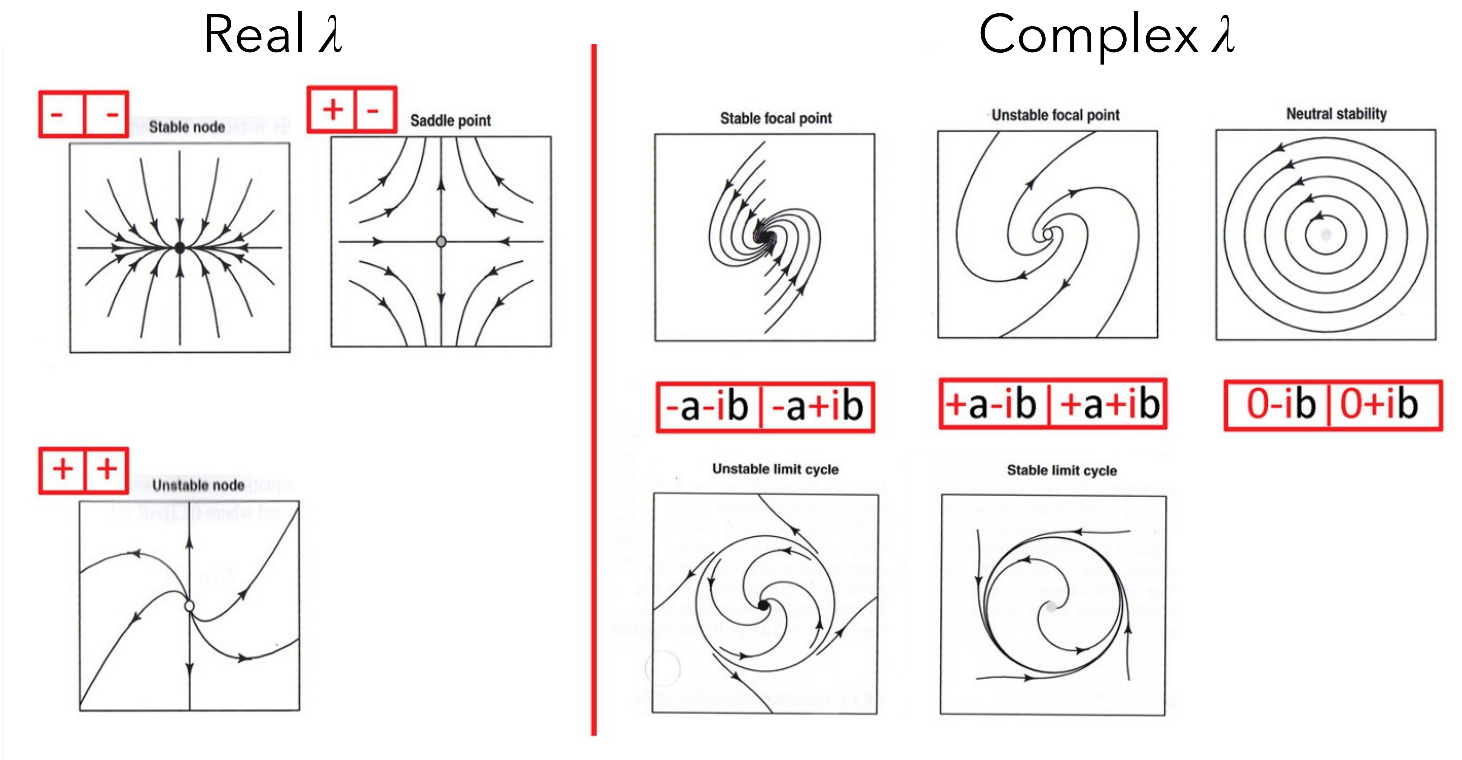
$$v_t = C_1 U_1 e^{\lambda_1 t} + C_2 U_2 e^{\lambda_2 t}$$

with λ the eigenvalues of J

if negative $\lambda \rightarrow$ $\lim_{t \to \infty} v_t = \vec{0}$ Stable equilibrium

1. General analysis (3) Local stability analysis

- Determine the stability of each equilibrium by analyzing the Jacobian matrix at the equilibrium



$$J = \begin{bmatrix} \frac{\partial \frac{dP}{dt}}{\partial P} & \frac{\partial \frac{dP}{dt}}{\partial H} \\ \frac{\partial \frac{dH}{dt}}{\partial P} & \frac{\partial \frac{dH}{dt}}{\partial H} \end{bmatrix}_{P^*,\, H^*}$$

with λ the eigenvalues of J

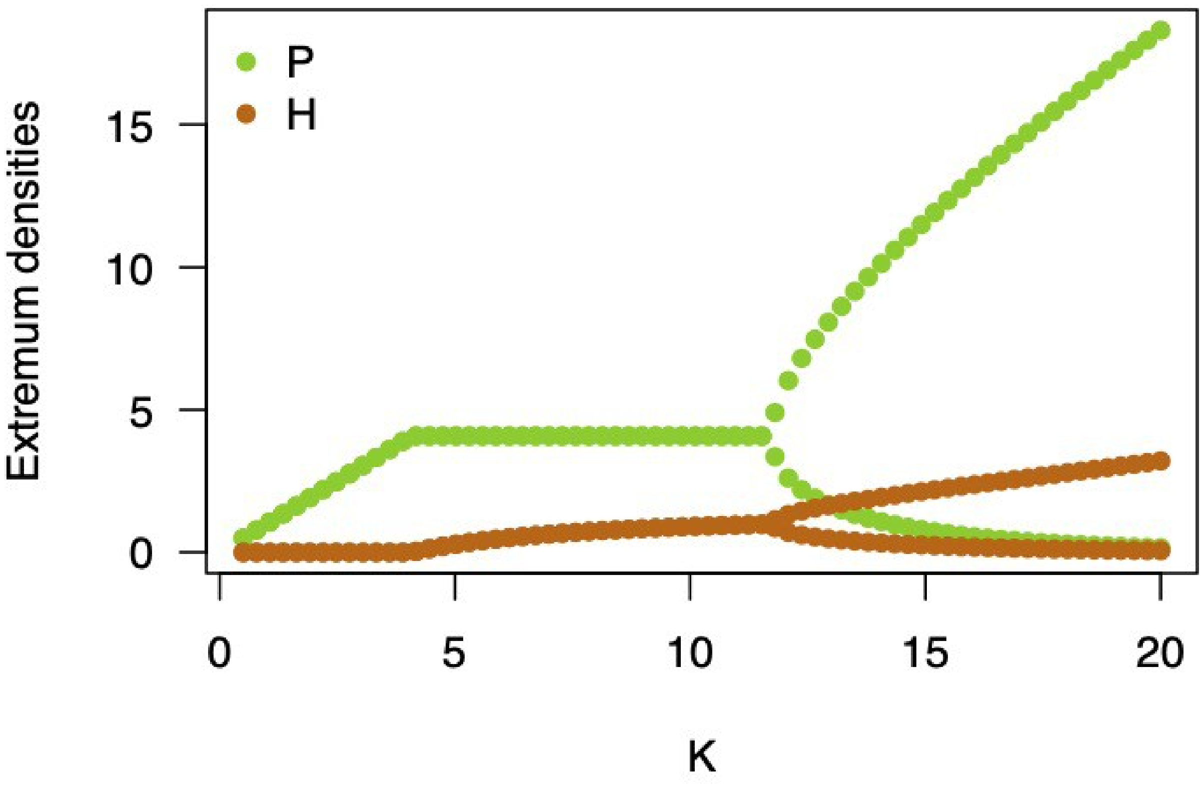
if negative $\lambda \rightarrow$ $\lim_{t \to \infty} v_t = \vec{0}$ Stable equilibrium

J from the function `fully.jacobian` (`rootSolve`) λ from the function `eigen`

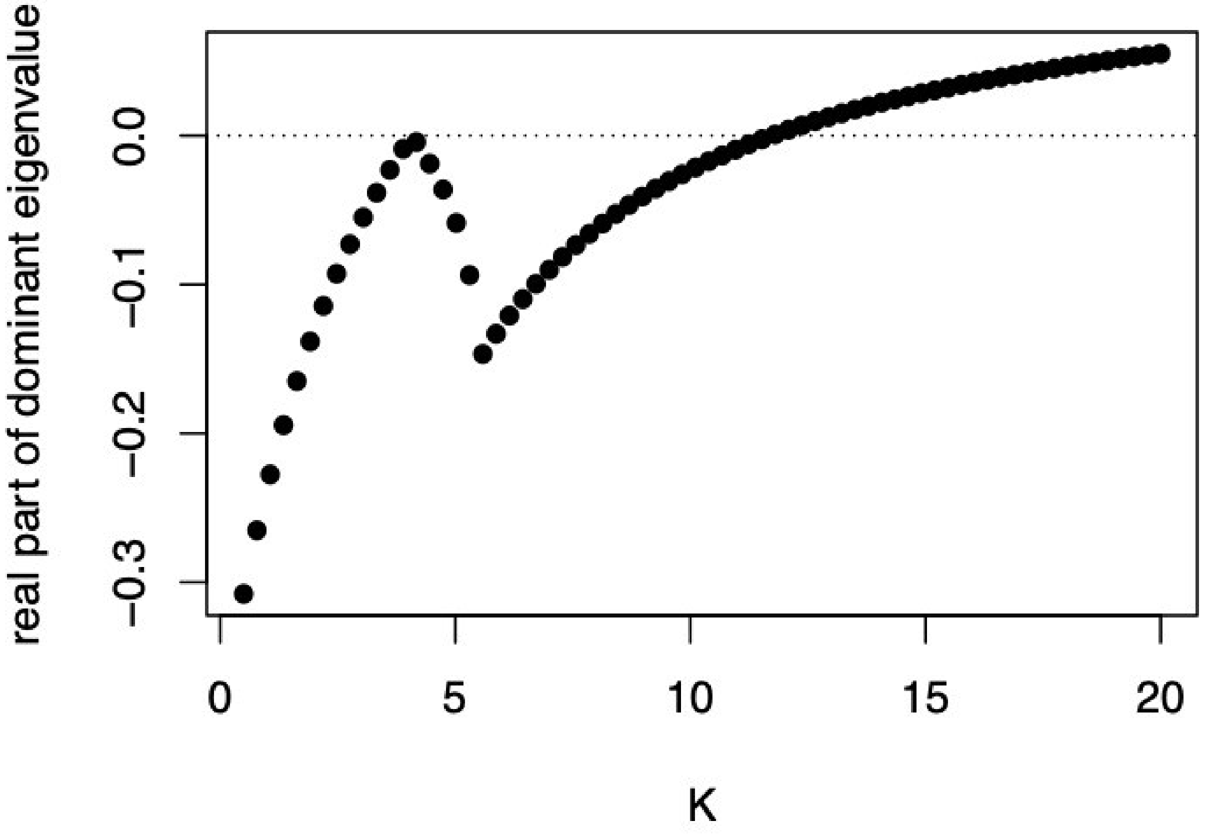
1. General analysis (4) Bifurcation diagrams

- How long-term (asymptotic) state of the system varies with one parameter

Variables

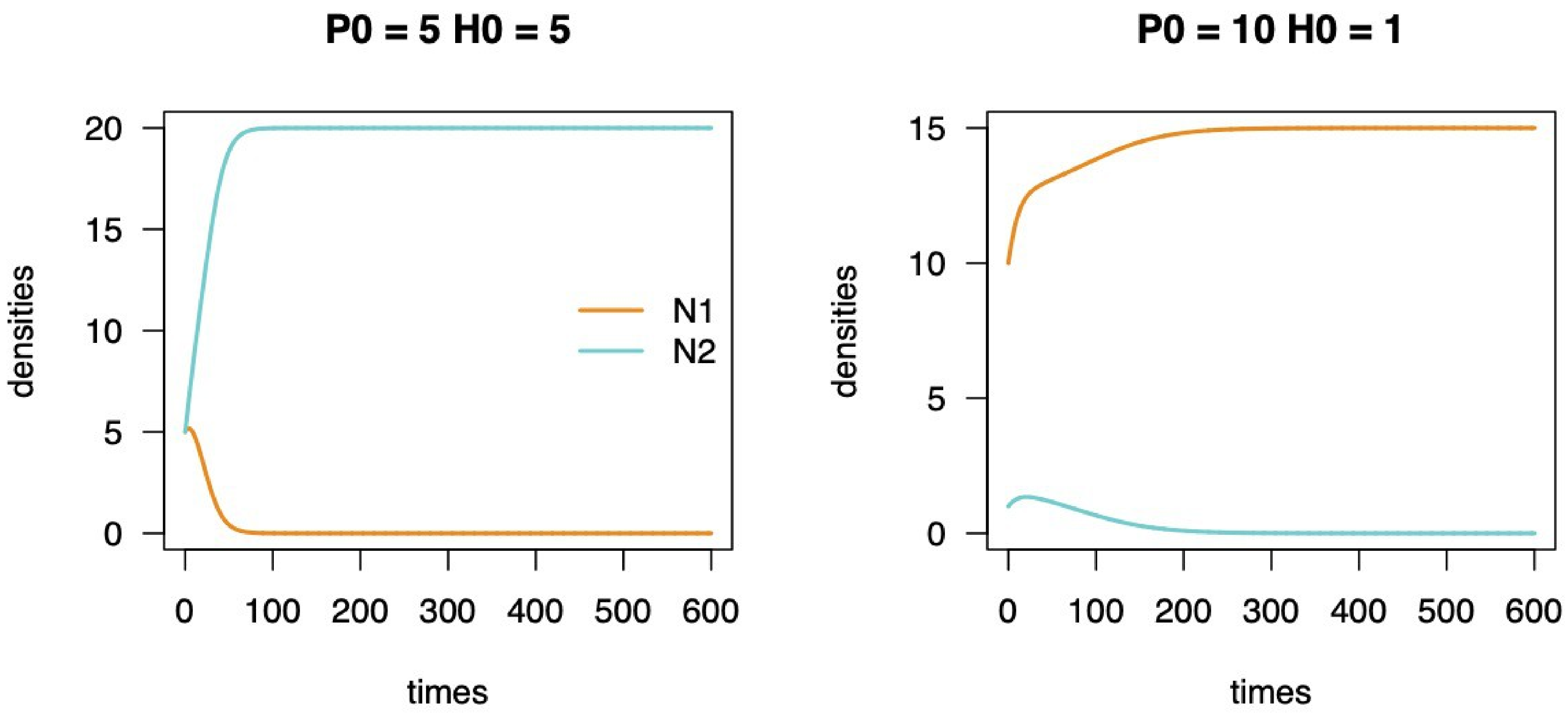


Stability



1. General analysis (5) Dependance to initial conditions

- we can observe several equilibrium points for the same parameters (historical effects)
- Example of Lotka-Volterra competition only initial densities differing:



- Screen series of initial densities to find all the equilibria using `searchZeros` in `nleqslv`

3. What model analysis strategy?

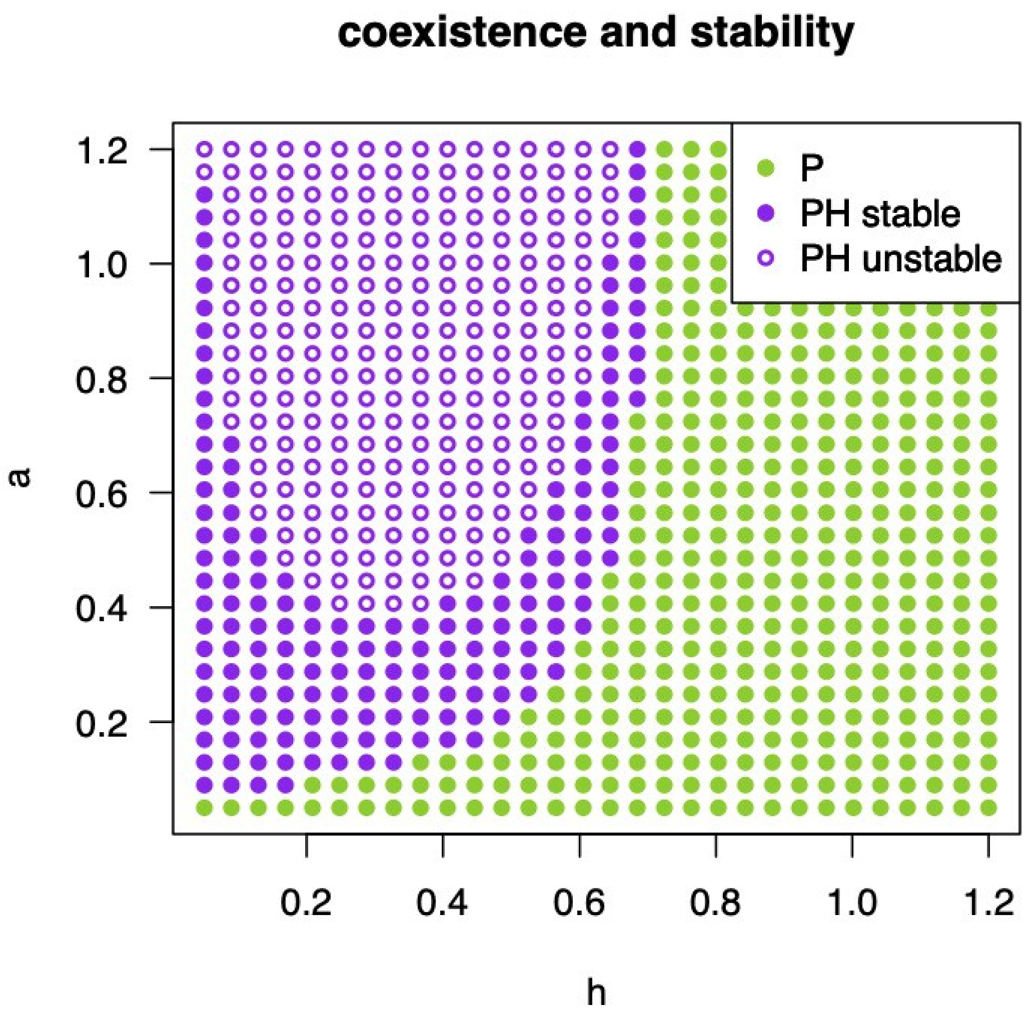
1. Parameter exploration
2. model comparison
3. *in sillico* experiment on synthetic data

2. Analysis strategy (1) Parameter variation

- Generalisation of bifurcation diagrams with 2-D parameter space exploration.
- The aim is to identify all the possible behaviors of the model within 'reasonable' parameter ranges

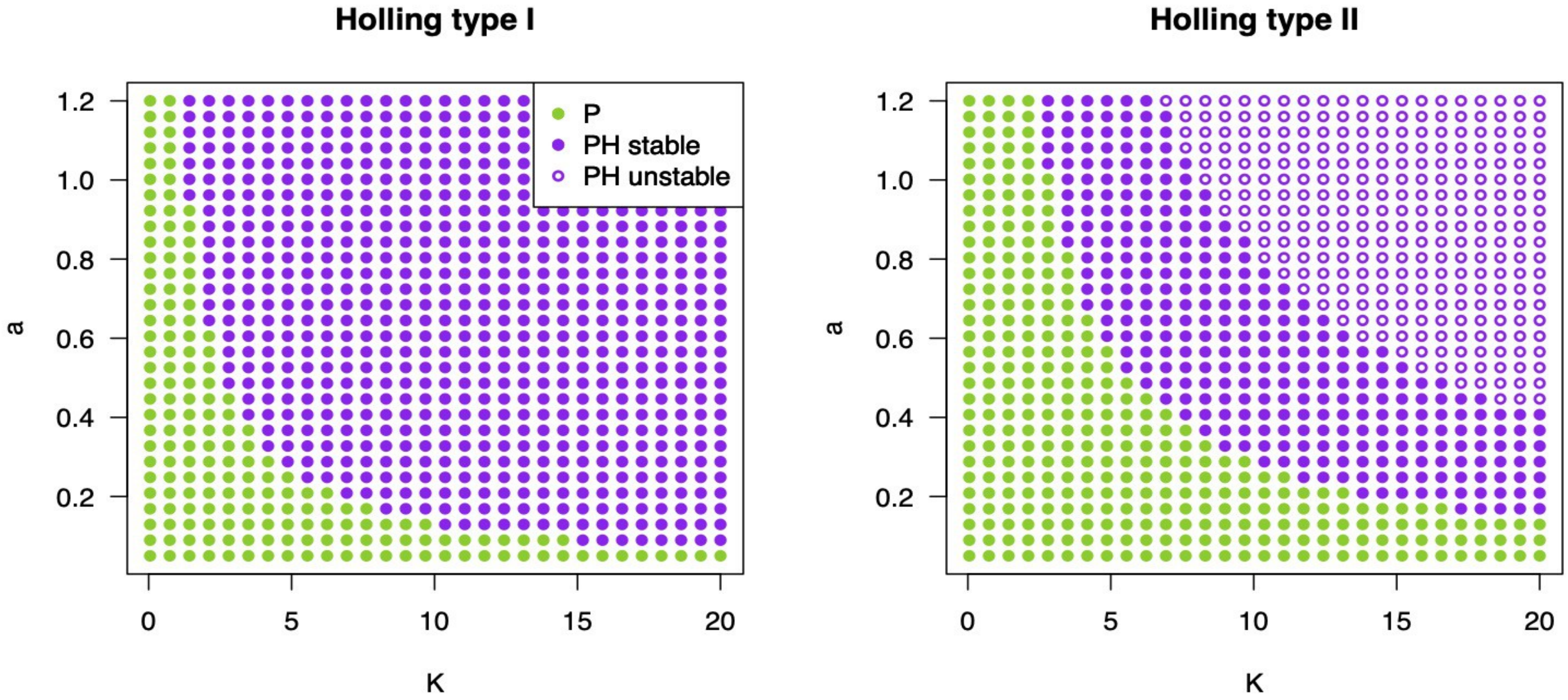
Here we vary h the handling time and a the grazing rate

- h should be sufficiently small, for H to persist
- Increasing a allows to compensate high h
- Increasing a destabilizes the system



2. Analysis strategy (2) Model comparison

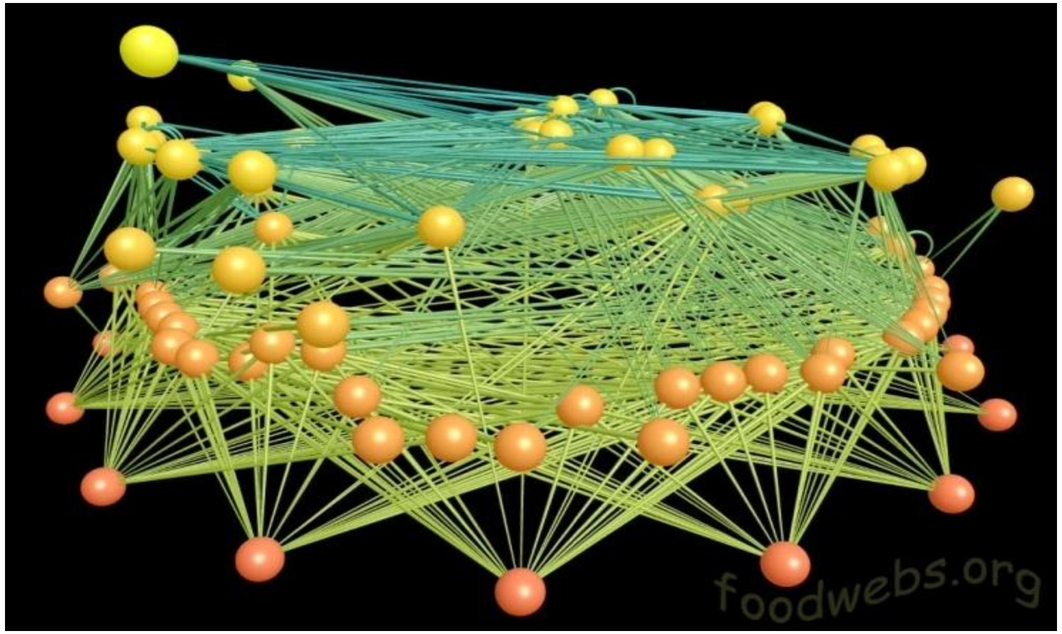
Here we compare models with different functional responses for the herbivore



In our system a type I (linear) increases persistence and stability compared to a type II (saturating) functional response because it creates a lag between P and H growths.

2. Analysis strategy (3) Experiments with synthetic data

- Complex system experiments not feasible in nature Ụ create realistic virtual data, for example food webs having the same general properties as natural food webs, to do perturbation experiments and observe how this would modify food we structure.



+

3. Parameter exploration and robustness of conclusions

- Tractability: We know everything about the model
- Parameterisation: fix some of them for which we have empirical knowledge

or restrain their range to reduce the parameter space to explore

- Sensitivity analysis:
  - Check the sensitivity of the results to variation in parameters ± 10%
  - Simulate many parameter sets to include variations
  - Many algorithms exist to explore efficiently large parameter spaces

❖ Some useful references

Breckling, B., Jopp, F., & Reuter, H. (2011). Historical background of ecological modelling and its importance for modern ecology. In Modelling Complex Ecological Dynamics: An Introduction into Ecological Modelling for Students, Teachers & Scientists (pp. 29-40). Berlin, Heidelberg: Springer Berlin Heidelberg.

Caswell, H. (1988). Theory and models in ecology: a different perspective. Ecological modelling, 43(1-2), 33-44.

Grainger, T. N., Senthilnathan, A., Ke, P. J., Barbour, M. A., Jones, N. T., DeLong, J. P., ... & Germain, R. M. (2022). An empiricist's guide to using ecological theory. The American Naturalist, 199(1), 1-20.

Levins, R. (1966). The strategy of model building in population biology. American scientist, 54(4), 421-431.

Otto, S. P., & Day, T. (2007). A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press

Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C. L., Goldberg, E. E., Stern, C. A., ... & Yeh, D. J. (2014). Not just a theory—the utility of mathematical models in evolutionary biology. PLoS biology, 12(12), e1002017.

❖ Other quoted references

Fisher, R. A. (1937). The wave of advance of advantageous genes. Annals of eugenics, 7(4), 355-369.

Kéfi, S., Rietkerk, M., Alados, C. L., Pueyo, Y., Papanastasis, V. P., ElAich, A., & De Ruiter, P. C. (2007). Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. Nature, 449(7159), 213-217.

Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the ESA, 15(3), 237-240.

MacArthur, R. H., & Wilson, E. O. (2001). The theory of island biogeography (Vol. 1). Princeton university press.

Rosenzweig, M. L., & MacArthur, R. H. (1963). Graphical representation and stability conditions of predator-prey interaction. The American Naturalist, XCVII(895), 209–223.