COM3001/6009
Modelling and Simulation of Natural Systems
Lecture 1: Introduction

Mike Smith* and Luca Manneschi

*m.t.smith@sheffield.ac.uk

First, what is a ‘natural system’?

- One can find various definitions, but “a system that is not manmade” seems to be a good approximation.
- This include those systems which humans can influence (e.g. the spread of infectious diseases).
- We also include ourselves in this (e.g. modelling crowds).
- The methods probably even apply to artificial systems! (e.g. the internet)...

Model of routing on the internet

The problem of oscillations and chaotic behaviour in routing…
Parameter **w** which controls how fast router responds to traffic volume:

- Too slow (small **w**) and the latency will be excessive or inefficient.
- Too fast (big **w**) and the system will oscillate or become chaotic.
- Emergent behaviour…

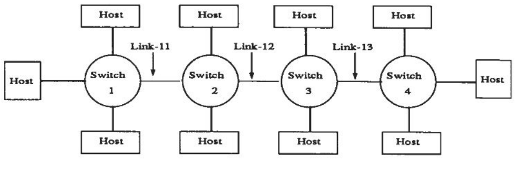


Figure 1. The two network topologies used in simulation tests

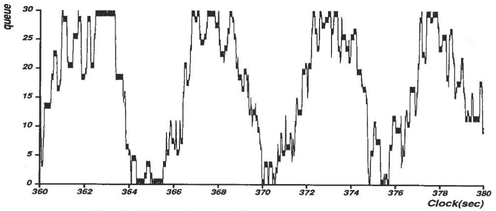
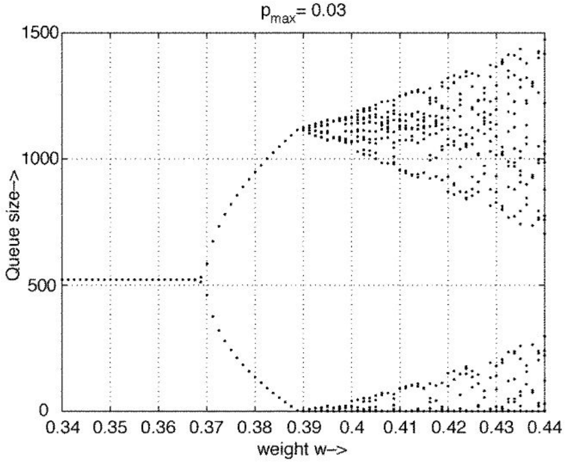


Figure 2. A 20-second trace of packet queue length at Link-2, Switch-1



Compte, Albert, et al. "Cellular and network mechanisms of slow oscillatory activity (< 1 Hz) and wave propagations in a cortical network model." *Journal of neurophysiology* 89.5 (2003): 2707-2725.

Ranjan, Priya, Eyad H. Abed, and Richard J. La. "Nonlinear instabilities in TCP-RED." *IEEE/ACM transactions on networking* 12.6 (2004): 1079-1092.

Models

Need to think what the **purpose** of the model is?

- Estimating a parameter?
- Understanding a system?
- Finding where our current hypotheses are wrong?
- Predicting the future?
    - Making a decision?

We can’t model everything exactly, so need to decide what simplifications are needed.

Typically we try to make our model/simulation as simple as possible while still capturing the features of interest.

Examples of Models

- **Conceptual Models**

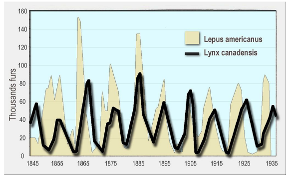
Examples of Models

Earthquake simulator
https://phys.org/news/2017-09-nevada-quake-lab-bridge-mexico.html

- **Physical Models**
  - **Engineering (e.g. earthquake simulation, wind-tunnels)**
  - ***in vivo* (e.g. disease model)**
  - ***in vitro* (e.g. drug effect on particular cell type)**
-

Examples of Models

- **Mathematical or Computational Models**
    - **E.g. Epidemiology (SIR-models)**
    - **Cosmological models**
    - **Ecological models (e.g. Lotka-Volterra)**



Hare and Lynx

wikipedia

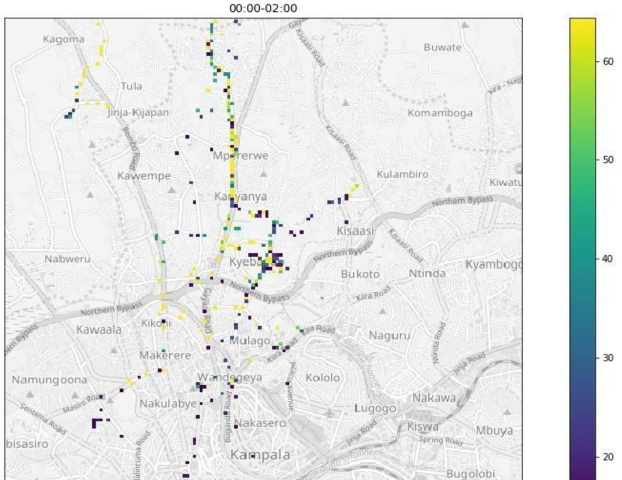
http://www.tng-project.org/ Time evolution of the cosmic magnetic field strength. Simulation: 10 megaparsec wide. Cosmological magnetohydrodynamical simulations of galaxy formation. Helps constrain dark-matter locations, distribution of galaxy types, blackhole sizes and numbers, etc…

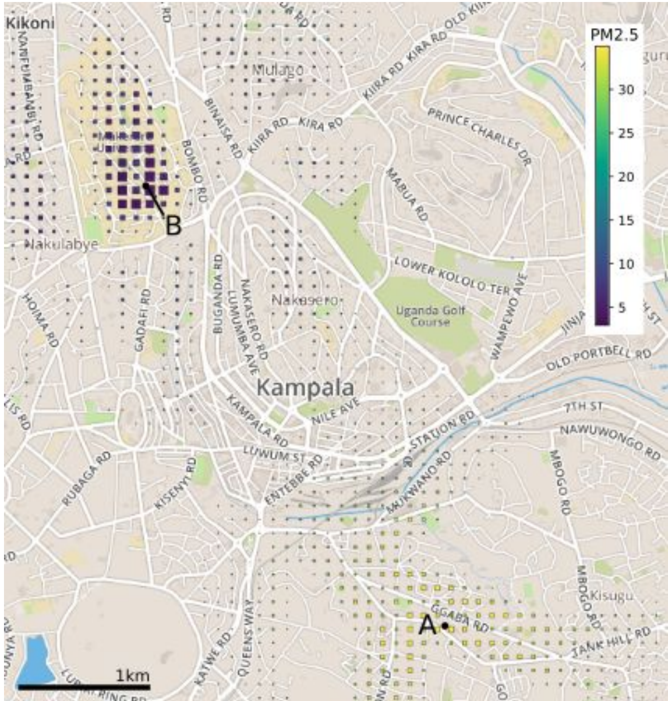
Model: Empirical or Mechanistic Example: Air pollution

- Data: A network of air pollution sensors.
- The questions/aims might be:
    - What is the pollution at location x at time y.
    - Where are the main sources of pollution?

Model: Empirical or Mechanistic
Example: Air pollution

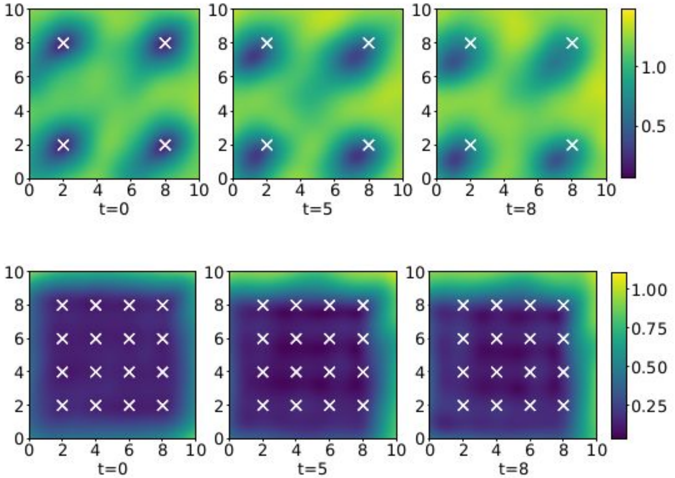
- Empirical approach, just do regression from all the sensors…
  - How to interpolate between sensors?
  - Can’t say where the pollution’s from?
  - Do we just include wind as a regressor?





Model: Empirical or Mechanistic
Example: Air pollution

- Mechanistic approach, need an advection/diffusion model of pollution…
  - Quite computationally intensive
  - Lots of unknowns (e.g. sources of pollution, diffusion constant, etc)
  - But could extrapolate (e.g. to the future)
  - Can say something about sources
- Something in between?
  - Currently working on a way to model using Gaussian processes to allow uncertainty in source term, but still Be quick to calculate.



Adjoint-aided inference of Gaussian process driven differential equations, *in review*.

Model: Empirical or Mechanistic

- ‘Big data’ and machine learning tools allows more empirical (data-driven) approaches. (still just fitting a curve!).
- Advantages of Empirical modelling:
  - Don’t need to know mechanism
  - Simple to use
  - Often scales better with large N
  - Often less computationally intensive

Model: Empirical or Mechanistic

- ‘Big data’ and machine learning tools allows more empirical (data-driven) approaches. (still just fitting a curve!).
- Advantages of Empirical modelling:
    - Don’t need to know mechanism
    - Simple to use
    - Often scales better with large N
    - Often less computationally intensive
- Disadvantages:
    - Needs more data (for same accuracy)
    - Poor extrapolation (probably worse uncertainty quantification)
    - Vulnerable to adversarial attack
    - Can’t be used to estimate physical parameters
- We’ll be focusing on ‘mechanistic’ models in this module. In particular Agent Based Models and Equation Based Models.

Other Dimensions of Modelling

- Descriptive or Predictive
  - maybe we want just a description so we can see what the internal ‘hidden’ units are doing.

Other Dimensions of Modelling

- Descriptive or Predictive
    - maybe we want just a description so we can see what the internal ‘hidden’ units are doing.
- Empirical vs Mechanistic ← we’ve just looked at this.
- Stochastic or deterministic
    - does it do the same each time?
    - Same input -> same output = deterministic

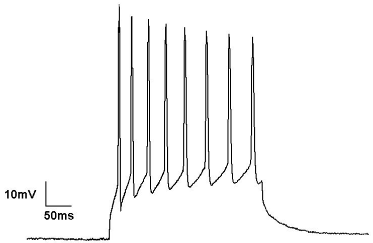
Other Dimensions of Modelling

- Descriptive or Predictive
  - maybe we want just a description so we can see what the internal ‘hidden’ units are doing.
- Empirical vs Mechanistic ← we’ve just looked at this.
- Stochastic or deterministic
  - does it do the same each time?
  - Same input -> same output = deterministic
- Quantitative vs Qualitative
  - Air pollution: we might want to know exact pollution, or we might just want a general idea about where the sources are, or just a sense of what effect does X have?

Other Dimensions of Modelling

- Descriptive or Predictive
  - maybe we want just a description so we can see what the internal ‘hidden’ units are doing.
- Empirical vs Mechanistic ← we’ve just looked at this.
- Stochastic or deterministic
  - does it do the same each time?
  - Same input -> same output = deterministic
- Quantitative vs Qualitative
  - Air pollution: we might want to know exact pollution, or we might just want a general idea about where the sources are, or just a sense of what effect does X have?
- Level of Abstraction
  - My first MSc was modelling single synapses.
  - My second MSc looked at in vivo models of integration of orientation and spatial cues
  - My PhD using fMRI in humans to look at whole brain regions.

The Modelling Process

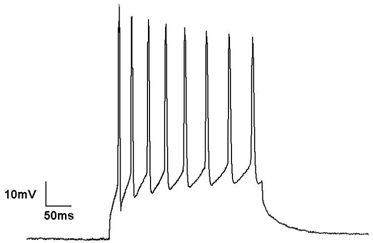


Observations
Microelectrode current-clamp recordings

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

The Modelling Process



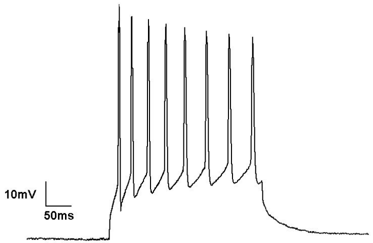
Observations
Microelectrode current-clamp recordings

Hypothesis
Voltage-gated ion channels, for different ions, explain action potentials?

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

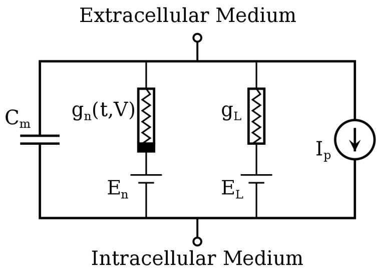
The Modelling Process



Observations
Microelectrode current-clamp recordings

Hypothesis
Voltage-gated ion channels, for different ions, explain action potentials?

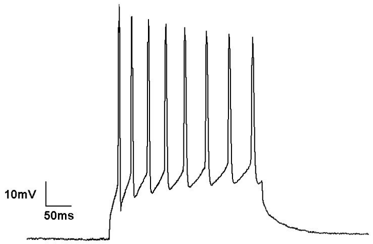
Mechanisms
An electrical circuit representing different components.



1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

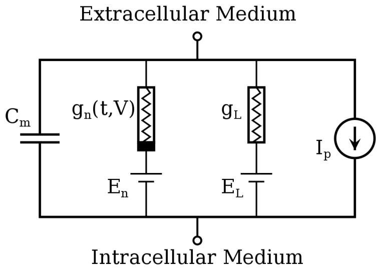
The Modelling Process



Observations
Microelectrode current-clamp recordings

Hypothesis
Voltage-gated ion channels, for different ions, explain action potentials?

Mechanisms
An electrical circuit representing different components.



$$I = C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} + \bar{g}_\mathrm{K} n^4 (V_m - V_K) + \bar{g}_\mathrm{Na} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l),$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m$$

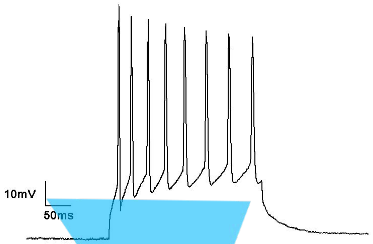
$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h$$

Description
Differential equations describing electrical circuit.

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

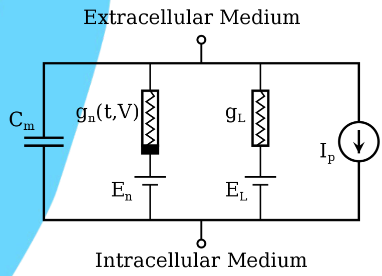
The Modelling Process



Observations
Microelectrode current-clamp recordings

Hypothesis
Voltage-gated ion channels, for different ions, explain action potentials?

Mechanisms
An electrical circuit representing different components.



$$I = C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} + \bar{g}_{\mathrm{K}} n^4 (V_m - V_K) + \bar{g}_{\mathrm{Na}} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l),$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m$$

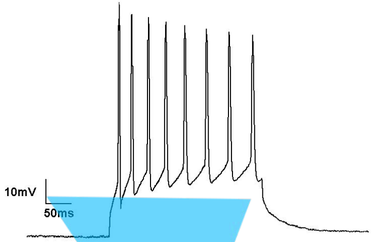
$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h$$

Description
Differential equations describing electrical circuit.

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

The Modelling Process



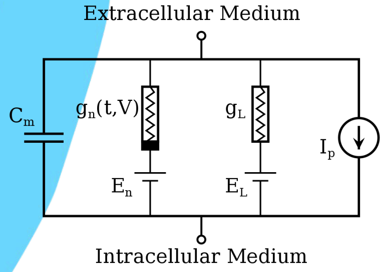
Observations
Microelectrode current-clamp recordings

Hypothesis
Voltage-gated ion channels, for different ions, explain action potentials?

Compare!

Prediction
“Run” the model

Mechanisms
An electrical circuit representing different components.



$$I = C_m \frac{dV_m}{dt} + \bar{g}_K n^4 (V_m - V_K) + \bar{g}_{Na} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l),$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m$$

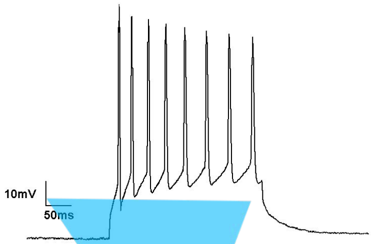
$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h$$

Description
Differential equations describing electrical circuit.

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

The Modelling Process



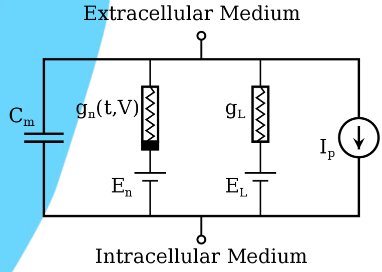
Observations
Microelectrode current-clamp recordings

Hypothesis
Voltage-gated ion channels, for different ions, explain action potentials?

Compare!

Prediction
“Run” the model

Mechanisms
An electrical circuit representing different components.



$$I = C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} + \bar{g}_{\mathrm{K}} n^4 (V_m - V_K) + \bar{g}_{\mathrm{Na}} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l),$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m$$

$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h$$

Description
Differential equations describing electrical circuit.

Understanding
They explained the ionic mechanisms underlying the initiation and propagation of action potentials

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

Detour…

Probably the first fully computational model employed for neuroscience research was the numerical model of the action potential by Hodgkin and Huxley. However, the speed of numerical computation was very slow. Denis Nobel wrote that 'It took Andrew Huxley months with a Brunsviga mechanical calculator[1] to solve for just a few milliseconds of nerve activity.'[35]

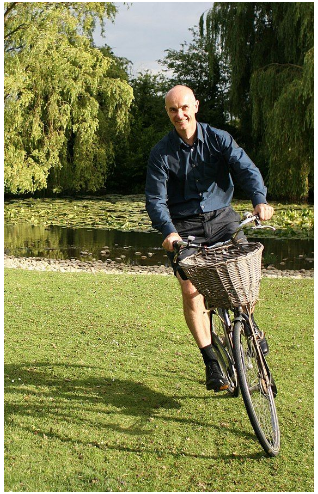
Occam’s Razor

(from wikipedia)... In science, Occam's razor is used as a heuristic to guide scientists in developing theoretical models rather than as an arbiter between published models.

I’d like to add though that Bayesian modelling provides an approach to accounting for model complexity…

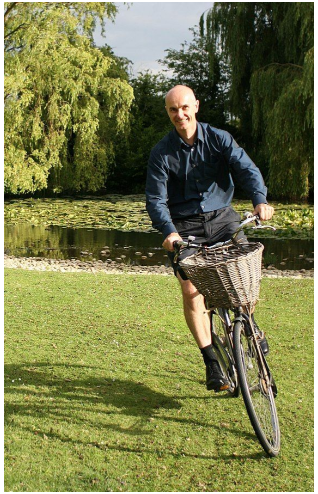
Random Detour: Bayes

From David MacKay…



Random Detour: Bayes

From David MacKay…



Model comparison and Occam's razor

We evaluate the plausibility of two alternative theories $\mathcal{H}_1$ and $\mathcal{H}_2$ in the light of data D as follows: using Bayes' theorem, we relate the plausibility of model $\mathcal{H}_1$ given the data, $P(\mathcal{H}_1 \mid D)$, to the predictions made by the model about the data, $P(D \mid \mathcal{H}_1)$, and the prior plausibility of $\mathcal{H}_1$, $P(\mathcal{H}_1)$. This gives the following probability ratio between theory $\mathcal{H}_1$ and theory $\mathcal{H}_2$:

$$\frac{P(\mathcal{H}_1 \mid D)}{P(\mathcal{H}_2 \mid D)} = \frac{P(\mathcal{H}_1)}{P(\mathcal{H}_2)} \frac{P(D \mid \mathcal{H}_1)}{P(D \mid \mathcal{H}_2)}. \tag{28.1}$$

The first ratio $(P(\mathcal{H}_1)/P(\mathcal{H}_2))$ on the right-hand side measures how much our initial beliefs favoured $\mathcal{H}_1$ over $\mathcal{H}_2$. The second ratio expresses how well the observed data were predicted by $\mathcal{H}_1$, compared to $\mathcal{H}_2$.

the second ratio, the data-dependent factor, embodies Occam's razor *automatically*. Simple models tend to make precise predictions. Complex models, by their nature, are capable of making a greater variety of predictions (figure 28.3). So if $\mathcal{H}_2$ is a more complex model, it must spread its predictive probability $P(D \mid \mathcal{H}_2)$ more thinly over the data space than $\mathcal{H}_1$. Thus, in the case where the data are compatible with both theories, the simpler $\mathcal{H}_1$ will turn out more probable than $\mathcal{H}_2$, without our having to express any subjective dislike for complex models. Our subjective prior just needs to assign equal prior probabilities to the possibilities of simplicity and complexity. Probability theory

This module…

- Deals with **mechanistic** models
  - see Machine Learning modules for empirical ‘data-driven’
  - hopefully there’s some convergence…
- Two types:
  - Agent based (weeks 2-5)
  - Equation based (6-11)

Examples of Agent Based Models…

Dynamics of contagious disease spread

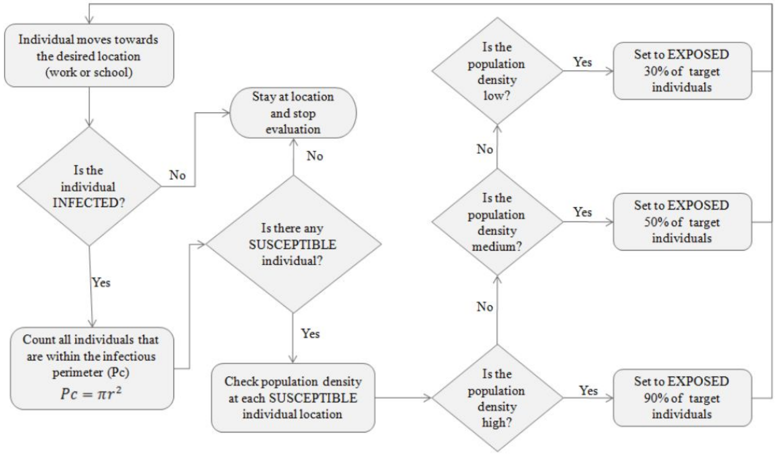


Figure 6
Flow diagram for the infection rules that describe the disease propagation among individuals at physically fixed location.

Perez, Liliana, and Suzana Dragicevic. "An agent-based approach for modeling dynamics of contagious disease spread." *International journal of health geographics* 8.1 (2009): 1-17.

Examples of Agent Based Models…

Microvascular patterning

- Rules at cell level lead to microvasculature patterning.
- Looked at two stimuli (including an angiogenic growth factor added).
- Really ‘cellular automata’ not ABM?

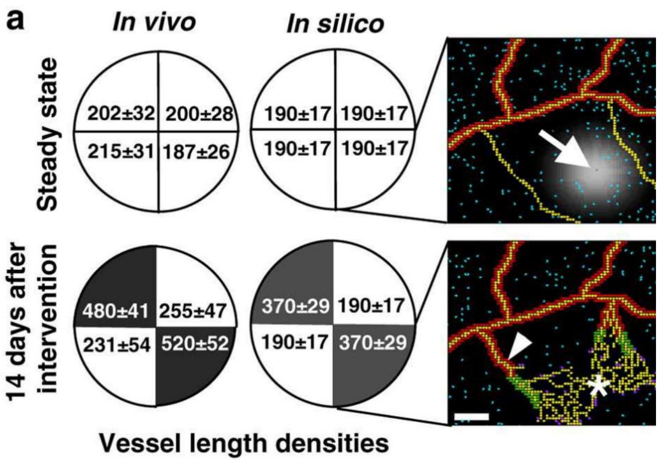
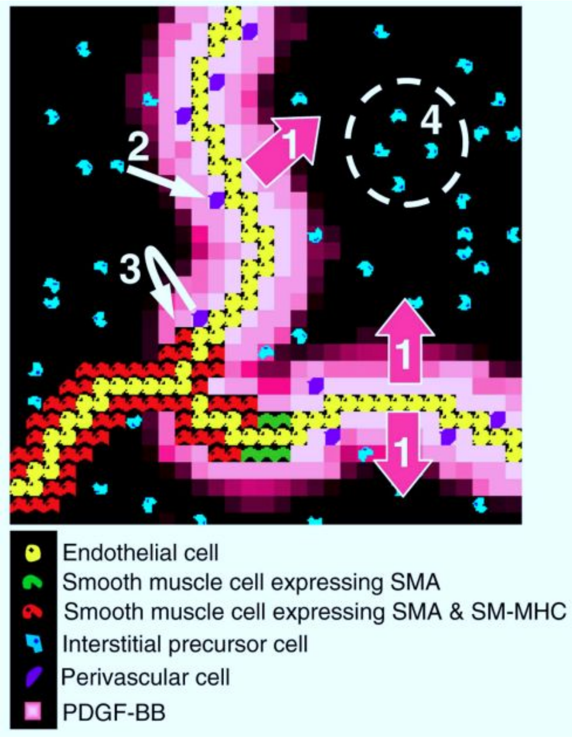


Table 1

Variables and rules for cell behaviors and growth factor diffusion

	Parameter	Definition	Reference
Endothelial cell (EC)	Cell size	100 × 10 μm	(28)
	Proliferation/ migration rate	Negligible	(29, 30, 31)
	Proliferation/ migration rate in response to $VEGF_{164}$	Doubling time = −1329[VEGF] +4828	(32, 33, 34)
Smooth muscle cell (SMC)	Cell size	6 × 88 μm	(35)
	Proliferation/migration rate	Negligible	(36)
	Proliferation/migration rate in response to PDGF-BB	Doubling time = 40+82.8* ($e^{-3.2[PDGFBB]}$) −0.15[PDGFBB]	(37)
	Proliferation/migration rate in response to TGF-	No response	(38, 39)
	Fold change in SM-MHC expression with PDGF-BB	Fold change = −0.0024[PDGFBB]^2 +0.004[PDGFBB]+1.2	(37)
	Fold change in SM-MHC expression with TGF-	Fold change = 0.73*ln([TGFβ])+1.1	(40)
	Cell size	10 × 10 μm	*
	Initial number	0.0035 cells per square micrometer	(42, 20)



Peirce, Shayn M., Eric J. Van Gieson, and Thomas C. Skalak. "Multicellular simulation predicts microvascular patterning and in silico tissue assembly." *The FASEB journal* 18.6 (2004): 731-733.

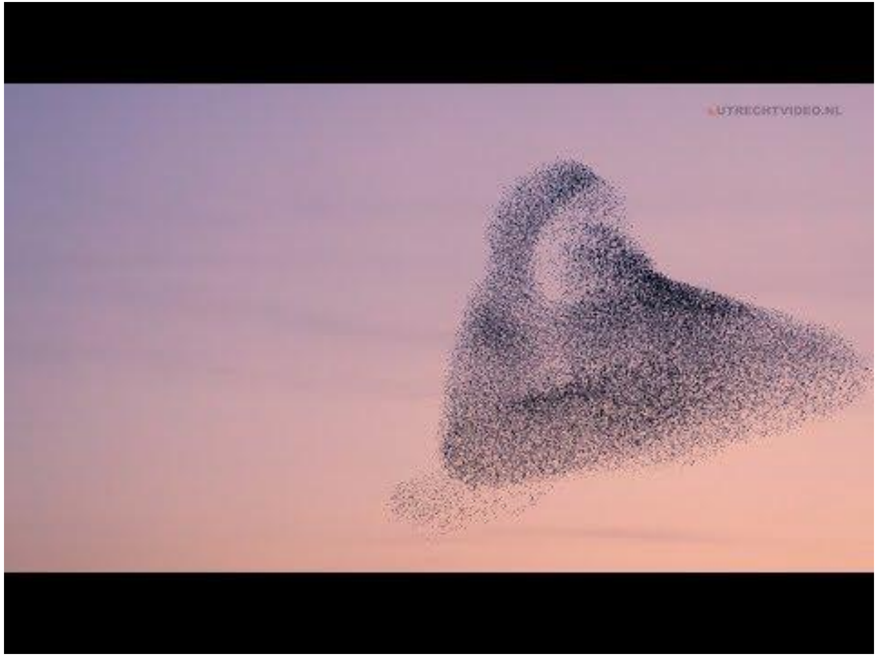
Examples of Agent Based Models…

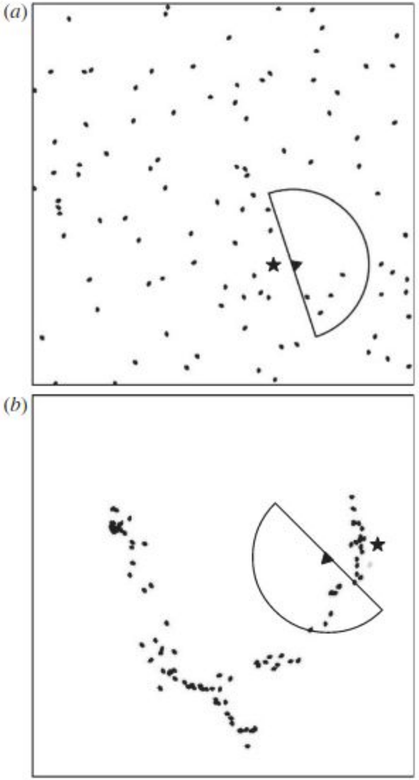
Flocking

Famous example is starling murmuration.

Much debate on **why** they do it. Maybe confuses predators?

Model with evolving agents explores if it helps avoid predation.





Olson, Randal S., et al. "Predator confusion is sufficient to evolve swarming behaviour." *Journal of The Royal Society Interface* 10.85 (2013): 20130305.

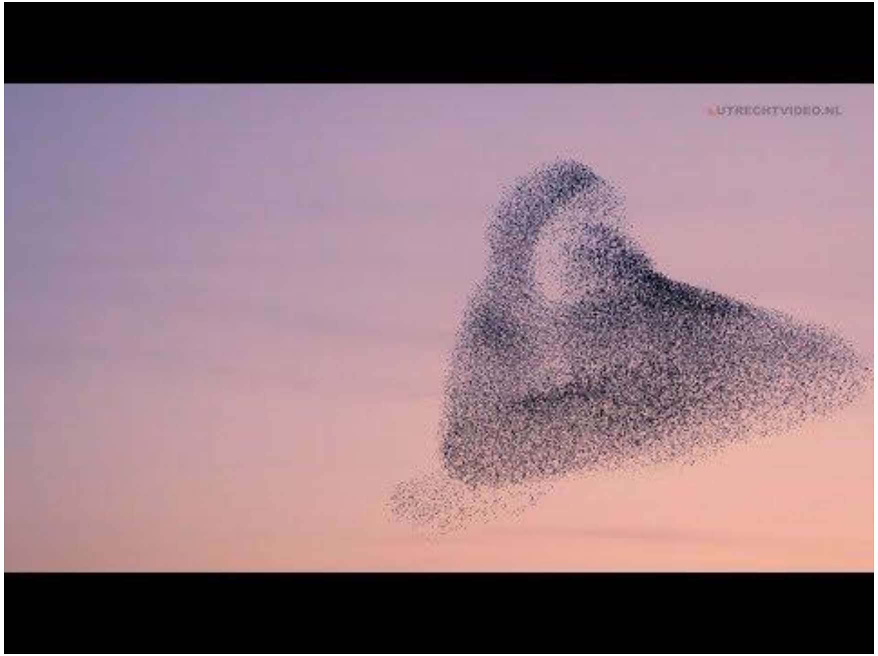
Examples of Agent Based Models…

Flocking

Quick exercise, in pairs or threes, discuss how you might model this as an ABM?

(what rules would you have?)

[5 minutes]

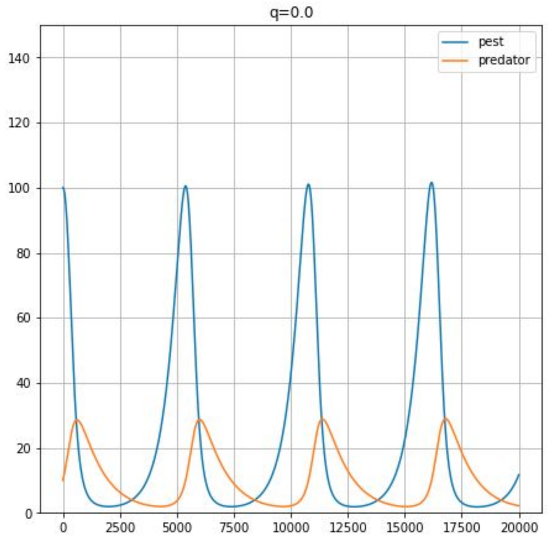


Examples of Equation Based Models…

Predator/Prey Model

Famous example

Lotka-Volterra.



$$\frac{dx}{dt} = \alpha x - \beta xy,$$
$$\frac{dy}{dt} = \delta xy - \gamma y,$$

where

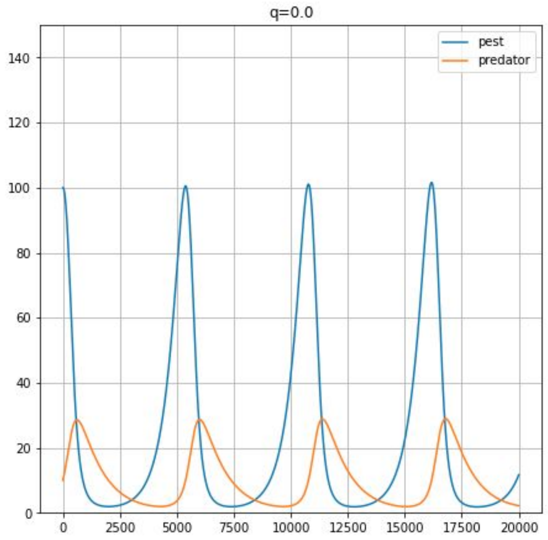
- x is the number of prey (for example, rabbits);
- y is the number of some predator (for example, foxes);
- $\frac{dy}{dt}$ and $\frac{dx}{dt}$ represent the instantaneous growth rates of the two populations;
- t represents time;
- $\alpha, \beta, \gamma, \delta$ are positive real parameters describing the interaction of the two species.

https://en.wikipedia.org/wiki/Lotka%E2%80%93Volterra_equations

Examples of Equation Based Models…

Predator/Prey Model

Famous example
Lotka-Volterra.



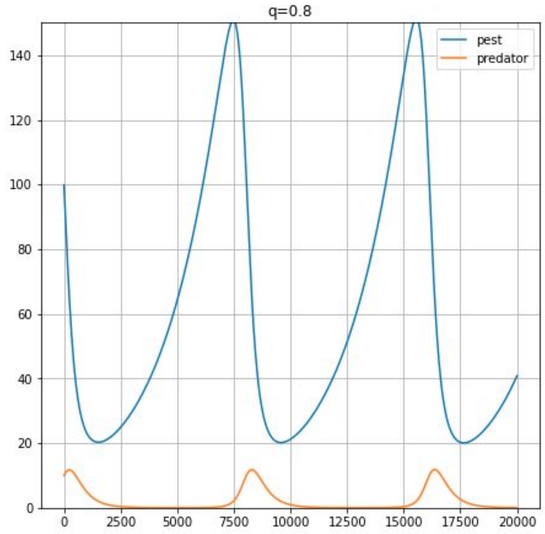

```
H = 100 #prey
P = 10 #pred
r = 1 #rate of growth of prey
c = 0.1 #capture constant
a = 0.2 #rate of prey getting eaten
m = 0.5 #predator mortality rate

for it in range(20000):
    dHdt = H * (r - c * P - q)
    dPdt = P * (a*c*H - m - q)
    H = H + dHdt * 0.002
    P = P + dPdt * 0.002
```

Examples of Equation Based Models…

Predator/Prey Model

Add insecticide…

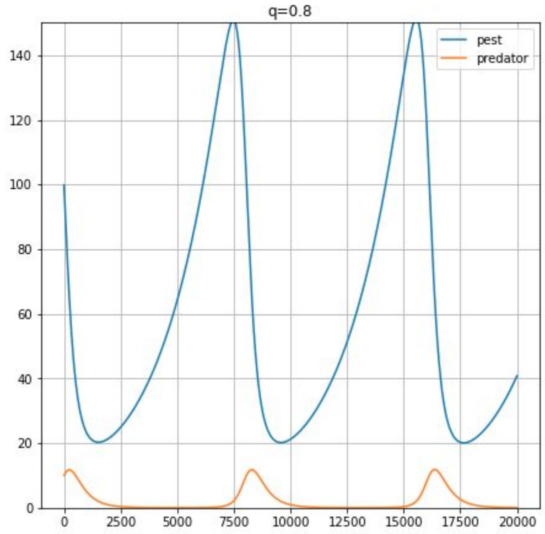


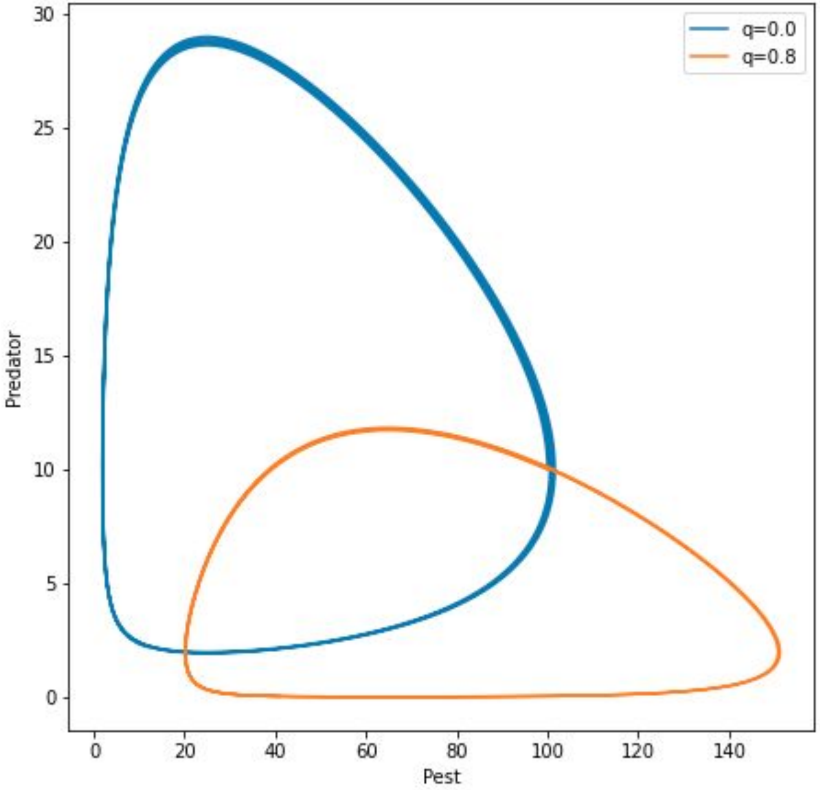

```
H = 100 #prey
P = 10 #pred
r = 1 #rate of growth of prey
c = 0.1 #capture constant
a = 0.2 #rate of prey getting eaten
m = 0.5 #predator mortality rate

for it in range(20000):
    dHdt = H * (r - c * P - q)
    dPdt = P * (a*c*H - m - q)
    H = H + dHdt * 0.002
    P = P + dPdt * 0.002
```

Examples of Equation Based Models…

Predator/Prey Model





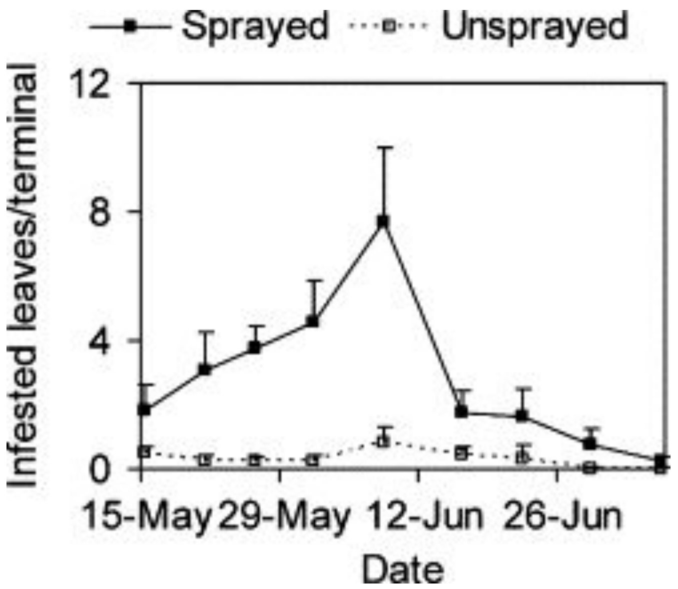
Examples of Equation Based Models…

Predator/Prey Model

Example from apple orchards.

Add pesticide → Should lead to fewer pests?

Often leads to more.



Brown, M. W. "Role of aphid predator guild in controlling spirea aphid populations on apple in West Virginia, USA." *Biological Control* 29.2 (2004): 189-198.

Why do this module?

"There's definitely a shortage of the right people. What we've found is that somebody spot on in terms of the **maths can't do the software**; if they're spot on in terms of the **software, they can't do the maths**."

- Ian Wright, the chief engineer for vehicle dynamics with the Mercedes AMG Petronas Formula One team

Why do this module?

Katy Chapman and team recorded the behaviour of bees entering and leaving nests over the summer to explore the effect of artificial light…

Manually having to label their behaviour.

Potential collaborations with Machine Learning and Agent Based Modelling?

Really benefit from Computer Science | Ecology crossover.

Ben Haden @ben_hadn

By the end of the course, you should be able to…

- **Develop** models that are either:
  - individual (agent) based
  - equations based
- Write and extend code to **simulate and visualise** the dynamics of these models.
- Analyse mathematically the **stability** of simple dynamical systems
- Select the appropriate method for modelling problems.
- Understand how to evaluate and calibration a model.

Prerequisites

You'll need:

- Good A level maths (esp calculus) for the 2nd half of the course
- Some programming skills (ideally in python)
- Imagination!
- Ability to read scientific literature

Assessment

COM3001 [10 credits]

- ABM group project (starts 21st Feb, deadline 25th March) [40%]
- Exam [60%]

COM6009 [15 credits]

- ABM group project (starts 21st Feb, deadline 25th March) [30%]
- Individual assignment (starts 25th March, deadline 13th May) [35%]
- Exam [35%]

Schedule

Wk	Date	Lecture Topic	by	Lab Topic	Assessed Work
1	07 Feb	Intro the module	Mike	Intro to python	
2	14 Feb	Intro to ABM		More python	
3	21 Feb	Implementing ABMs		Intro to group project	21/02: Start of Group Project
4	28 Feb	Validating ABMs		group project	
5	07 Mar	Intro to EBMs & integration		group project	
6	14 Mar	Saturating population model & stablity	Kenneth	Analytic + Numerical Integration	
7	21 Mar	Higher dimensions & numerical methods		Equilibria	25/03: Grp Project deadline & start of Ind. Assignment (COM6009)
8	28 Mar	Higher dims, Eq Stablity, SIR, Lorenz		High dim systems via RK method	
		Easter Break			
9	25 Apr	Comparing ABMs & EBMs. Multiscale	Luca	Pred/Prey vs other methods	
10	02 May	Spiking Neurons		Leaky integrate & fire, networks	
11	09 May	Intro to PDEs		PDEs	
12	16 May	Overflow / Revision		Overflow / Revision	13/05: Deadline for Ind. Assignment

Questions / Help

- Ask at the **lab**
- Ask on the **padlet on blackboard**
- **Email me** m.t.smith@sheffield.ac.uk
- **Office hours [updated]**: I'll be in COM G22 Blue ~~10:30-11:30~~ 1:30-2:30 Monday (after the lecture) for first five weeks (after that Luca will be taking the lead, but feel free to email to organise a meeting with me if you need to).

Group Project

- Groups of about 4 students
- Will discuss and launch in lab on the 21st February (week 3)
- Deadline: 25th March
- Assesment will be by:
    - Group report
    - Peer assessment survey (week 7 or 8)
- Grade is:
    - 40% of COM3001
    - 30% of COM6009
- Next week, you will receive an email asking you to select another student you would like to work with. We’ll then match students into groups. We’ll hopefully have the list of groups by the end of group 2, ready for week 3’s lab.

Feedback

- During lab sessions
- Via padlet & discussion forum
- During meetings (see office hours)
- Assignment marking

Reference Books

The Python Tutorial: https://docs.python.org/3/tutorial/index.html
This is probably written for those who already code a little.

"Differential Equations, dynamic systems and introduction to chaos". - Morris Hirsch, Robert L. Devaney, and Stephen Smale. Chapters 1, 11 and 14.

"Neuronal Dynamics" Werner M. Kistler, Richard Naud, Wulfram Gerstner, Liam Paninski. Chapter 1.

COM3001/6009

Modelling and Simulation of Natural Systems

Lecture 2: Introduction to Agent Based Models

Mike Smith* and Luca Manneschi

*m.t.smith@sheffield.ac.uk

Recap

- A model - method for representing our understanding of a real world system.
- Purpose:
    - Capture knowledge
    - Explore hypotheses
    - Make predictions
- Can be **descriptive** [“empirical” or ML] vs **mechanistic**.

Mathematical Population Models (Predator / Prey models)

Example of a mechanistic model is a population model describing the oscillation in predator and prey over time.

- Uses the Lotka-Volterra Equations which form a pair of Coupled Ordinary Differential Equations (ODEs).

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Rate of change of prey $$\frac{dH}{dt}$$

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Rate of change of prey $$\frac{dH}{dt}$$

Rate of change of predators $$\frac{dP}{dt}$$

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey (depends on current number of prey)

Rate of change of prey
$$\frac{dH}{dt} = \boxed{rH} - cHP$$

Rate of change of predators
$$\frac{dP}{dt} = caHP - mP$$

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Rate of change of prey
$$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators
$$\frac{dP}{dt} = caHP - mP$$

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators (depends on both population of prey AND predators).

Rate of change of prey

$$\frac{dH}{dt} = rH - \boxed{cHP}$$

Rate of change of predators

$$\frac{dP}{dt} = caHP - mP$$

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators

Rate of change of prey
$$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators
$$\frac{dP}{dt} = caHP - mP$$

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators

Rate of change of prey

$$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators

$$\frac{dP}{dt} = \boxed{caHP} - mP$$

Predator births (constrained by prey population AND predator population)

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators

Rate of change of prey $$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators $$\frac{dP}{dt} = caHP - mP$$

Predator births

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators

Rate of change of prey

$$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators

$$\frac{dP}{dt} = caHP - \boxed{mP}$$

Predator births

Predator mortality (due to other causes). Depends on current number of predators.

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators

Rate of change of prey $$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators $$\frac{dP}{dt} = caHP - mP$$

Predator mortality

Predator births

Mathematical Population Models (Predator / Prey models)

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators

Rate of change of prey

$$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators

$$\frac{dP}{dt} = caHP - mP$$

Predator mortality

Predator births

r,c,a and m are model **parameters** that need estimating to fit a real population.

Assumptions…

- Numbers are so large that we can treat them as continuous.
- Homogeneous populations.
- Environment is the same over the whole of space and time.

Assumptions (minimal viable population)

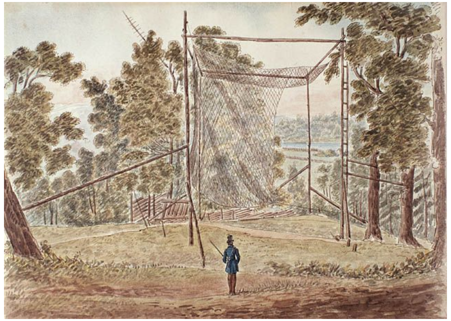
- In conservation ecology it is useful to have estimates of the **Minimal Viable Population**.

  "ecological threshold that specifies the **smallest number of individuals** in a species or population **capable of persisting** at a specific statistical probability level for a predetermined amount of time."

  - https://www.britannica.com/science/minimum-viable-population

Detour…

Minimal Viable Population



- Varies by species. Some examples of why it’s complicated:
- **Passenger Pigeons** (hunting & habitat loss).

Very social, needs large groups.

Detour…

Minimal Viable Population

- Varies by species. Some examples of why it's complicated:
- **Hymenoptera (ants, bees, wasps)...**
- Males (produced with unfertilised egg)
- "We developed an **individual-based** Monte Carlo model to simulate solitary haplodiploid populations" – Zayed and Packer (2005).

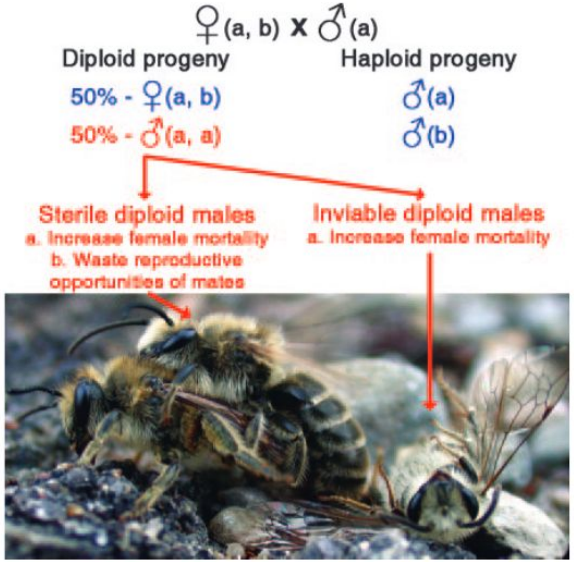


Fig. 1. The cost of sl-CSD. When haplodiploid females mate with males that share a *csd* allele in common (allele *a*), half of their diploid progeny will be homozygous at *csd* and will develop into DMs. Because females fertilize their eggs to produce daughters only, DMP is best viewed as increased female mortality. In some species, DMs have low viability. More often, however, DMs are effectively sterile: they are viable and achieve matings but do not father diploid daughters, thus reducing the reproductive success of their mates.

Zayed, Amro, and Laurence Packer. "Complementary sex determination substantially increases extinction proneness of haplodiploid populations." Proceedings of the National Academy of Sciences 102.30 (2005): 10742-10746.

Detour…

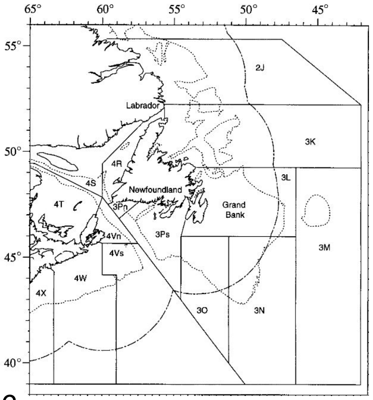
Example models (+ see Allee effect in a sec)

- **North Atlantic Cod** (over fishing).

  "By 1993 six cod populations had collapsed, forcing a belated moratorium on fishing. **Spawning biomass had decreased…by 99% in the case of "northern" cod**, previously the largest cod fishery in the world"

  – wikipedia & Myers et al. (1997).

- Myers et al. (1997) compared research surveys with Virtual Population analysis [VPA] - a type of equation based model, and found **the VPA results were able to detect the declining recruitment before the collapse**, unlike the data from the research surveys.



Myers, Ransom A., Jeffrey A. Hutchings, and Nicholas J. Barrowman. "Why do fish stocks collapse? The example of cod in Atlantic Canada." *Ecological applications* 7.1 (1997): 91-106.

Example models

Why aren’t cod populations returning? (still declining!)

Detour…

Other useful models…

Neuenhoff et al. (2019):

- Lots of parameters…

Population dynamics variables			
$\bar{R}^a$	Log mean recruitment (log numbers)		Both
$\varepsilon_a^{ri\,a}$	Recruitment process errors to initialize abundance at age in year 1		Both
$\varepsilon_y^{r\,a}$	Recruitment process errors $y \in \{2, 3, \ldots, T\}$		Both
σ^r	Recruitment process error standard deviation	0.5	Both
$m_k{}^a$	Initial natural mortality for age block k (year^{-1})		Both
μ_k^m	Prior mean for initial natural mortality (year^{-1})	0.5, 0.15	Both
σ_k^m	Prior standard deviation for initial natural mortality	0.1, 0.05	Both
$\varepsilon_{k,y}^{M\,a}$	Natural mortality process errors		Both[b]
σ_k^M	Natural mortality process error standard deviation	0.075, 0.075	Both[b]
q_g	Catchability for fishery $g = 2, 3, 4$		Both
C_y	Catch of Atlantic cod in year y (t)		Both
$N_{a,y}$	Abundance of Atlantic cod at age in year y (numbers)		Both
B_y^S	Spawning biomass in year y (t)		Both
$B_{g,y}^E$	Exploitable biomass for fishery–survey G in year y (t)		Both
$Z_{a,y}$	Instantaneous total mortality rate at age in year y (year^{-1})		Both
F_y	Fully recruited instantaneous fishing mortality in year y (year^{-1})		Both
Fi^a	Fully recruited fishing mortality for initializing abundance at age in year 1		Both
$s_{g,a,y}$	Selectivity at age in year y for fishery g		Both
$s_{50,g,y}{}^a$	Age at 50% selectivity for fishery g in year y		Both
$s_{95,g,y}{}^a$	Age at 95% selectivity for fishery g in year y		Both

Neuenhoff, Rachel D., et al. "Continued decline of a collapsed population of Atlantic cod (Gadus morhua) due to predation-driven Allee effects." *Canadian Journal of Fisheries and Aquatic Sciences* 76.1 (2019): 168-184.

Detour…

Other useful models…

Neuenhoff et al. (2019):

- Lots of parameters…

$w^{B}_{a,y}$	Beginning-of-year mass-at-age (t)		Both
$w^{G}_{g,a,y}$	Mass-at-age for fishery–survey g		Both
$P^{\text{mat}}_{a,y}$	Proportion mature at age in year y		Both
d_g	Survey date (month/12)	0, 0.75, 0.67, 0.75	Both
$M_{a,y}$	Instantaneous natural mortality rate at age in year y (year^{-1})		Both
Mp_y	M due to grey seal predation for age class $k = 2$ ($a \geq 5$)		FR
Mo_y	M due to other causes for age class $k = 2$ ($a \geq 5$)		FR
$m_p{}^a$	Mp_1 (prior mean and variance the same as for m_2)		FR
$m_o{}^a$	Mo_1		FR
μ^{m_o}	Prior mean for m_o (year^{-1})	0.15	FR
σ^{m_o}	Prior standard deviation for m_o	0.03	FR
y_0^{min}	First year for random walk in Mo	7 (1977)	Both
y_0^{max}	Last year for random walk in Mo	32 (2002)	Both

Neuenhoff, Rachel D., et al. "Continued decline of a collapsed population of Atlantic cod (Gadus morhua) due to predation-driven Allee effects." *Canadian Journal of Fisheries and Aquatic Sciences* 76.1 (2019): 168-184.

Detour…

Other useful models…

Neuenhoff et al. (2019):

- Lots of parameters…

Symbol	Description	Value	Model
$\varepsilon_y^{Mo_a}$	Process errors for *Mo* $y \in \{7, 8, \ldots, 32\}$		FR
σ_y^{Mo}	Process error standard deviation for *Mo* $y \in \{7, 8, \ldots, 32\}$	—[c]	FR
Functional response variables			
$\log(q')^a$	$\log_e$ encounter rate between grey seals and Atlantic cod 5 years and older		FR
n_y^{k2}	Per capita consumption of age-block 2 Atlantic cod by grey seals (numbers)		FR
N_y^{k2}	Abundance of age-block 2 Atlantic cod (numbers)		FR
Ho	Handling time of prey other than age-block 2 Atlantic cod		FR
w_y	Average mass of 5+ cod (t)		FR
$C_{\max}$	Maximum annual consumption by grey seals (t)	2.0	FR
$N_{y^*}^{k2}$	Abundance of age-block 2 Atlantic cod in reference year y^* (numbers)	53×10^6	FR
w_{y^*}	Mean mass of age-block 2 Atlantic cod in reference year y^* (t)	0.00092	FR
$P_{y^*}^{\text{diet}}$	Proportional contribution by age-block 2 Atlantic cod by mass to the average grey seal diet in reference year y^* (2010)		FR
Ce_y	Age-block 2 Atlantic cod eaten by seals in year y (numbers)		FR

Neuenhoff, Rachel D., et al. "Continued decline of a collapsed population of Atlantic cod (Gadus morhua) due to predation-driven Allee effects." *Canadian Journal of Fisheries and Aquatic Sciences* 76.1 (2019): 168-184.

Detour…

Other useful models…

Neuenhoff et al. (2019):

- Lots of parameters…

Observation model variables

S_y	Observed grey seal foraging effort (seal-years)		FR
$\hat{S}_y$	Predicted grey seal foraging effort (seal-years)		FR
σ^S	Observation error standard deviation for seal abundance	0.1	FR
$B_y^{S^{RW}}$	Spawning biomass estimates by model RW for year y (t) (for eq. T4.11)		FR
$B_y^{S^{FR}}$	Spawning biomass estimates by model FR for year y (t) (for eq. T4.11)		FR
σ^{B^S}	Observation error standard deviation for SSB (for eq. T4.11)	0.1	FR
$I_{g,y}$	Observed biomass index		Both
$\hat{I}_{g,y}$	Predicted biomass index		Both
$p_{g,a,y}$	Observed proportions at age		Both
$\hat{u}_{g,a,y}$	Predicted proportions at age		Both

Neuenhoff, Rachel D., et al. "Continued decline of a collapsed population of Atlantic cod (Gadus morhua) due to predation-driven Allee effects." *Canadian Journal of Fisheries and Aquatic Sciences* 76.1 (2019): 168-184.

Detour…

Other useful models…

Neuenhoff et al. (2019):

- And equations…

Equation	Formula
Selectivity	
T2.1: selectivity	$s_{g,a,y} = \{1 + \exp[-\log(19)(a - s_{50,g,y})/(s_{95,g,y} - s_{50,g,y})]\}^{-1}$
Mortality	
T2.2[a]: instantaneous rate of natural mortality, $y = 1$	$M_{a,1} = m_i, \quad a \in k_i$
T2.3[a]: instantaneous rate of natural mortality, $y > 1$	$M_{a,y} = M_{a,y-1}\exp(\varepsilon^{M}_{k_i,y}), \quad a \in k_i$
T2.4: instantaneous rate of fishing mortality	$F_{a,y} = s_{1,a,y}F_y$
T2.5: instantaneous rate of total mortality	$Z_{a,y} = M_{a,y} + F_{a,a}$
State dynamics	
T2.6: abundance $a = 1$, $y = 1$	$N_{1,1} = \exp(\bar{R} + \varepsilon^{ri}_1)$
T2.7: abundance at $a \in \{2, 3, \ldots, A-1\}$, $y = 1$	$N_{a,1} = \exp\left[\bar{R} + \varepsilon^{r1}_a - \sum_{a=1}^{a-1}(s_{1,a,1}Fi + M_{a,1})\right]$
T2.8: abundance at $a = A$, $y = 1$	$N_{A,1} = \dfrac{\exp\left[\bar{R} + \varepsilon^{r1}_A - \sum_{a=1}^{A-1}(s_{1,a,1}Fi + M_{a,1})\right]}{1 - \exp[-(s_{1,A,1}Fi + M_{A,1})]}$
T2.9: abundance $a = 1$, $y > 1$	$N_{1,y} = \exp(\bar{R} + \varepsilon^{r}_y)$
T2.10: abundance $a \in \{2, 3, \ldots, A-1\}$, $y > 1$	$N_{a,y} = N_{a-1,y-1}\exp(-Z_{a-1,y-1})$
T2.11: abundance $a = A$, $y > 1$	$N_{A,y} = N_{A-1,y-1}\exp(-Z_{A-1,y-1}) + N_{A,y-1}\exp(-Z_{A,y-1})$

Neuenhoff, Rachel D., et al. "Continued decline of a collapsed population of Atlantic cod (Gadus morhua) due to predation-driven Allee effects." *Canadian Journal of Fisheries and Aquatic Sciences* 76.1 (2019): 168-184.

Detour…

Other useful models…

Neuenhoff et al. (2019):

- Failed recovery is due to severe **increases in the natural mortality** of adult Atlantic cod.
- Examined the role of **predation by grey seals**… incorporating grey seal predation in the population model
- Estimated predation mortality of adult Atlantic cod increased sharply during the cod collapse **and has continued to increase**
- While predation by grey seals appeared to play a minor role in the collapse of Atlantic cod, **we found it to be the main factor preventing recovery**

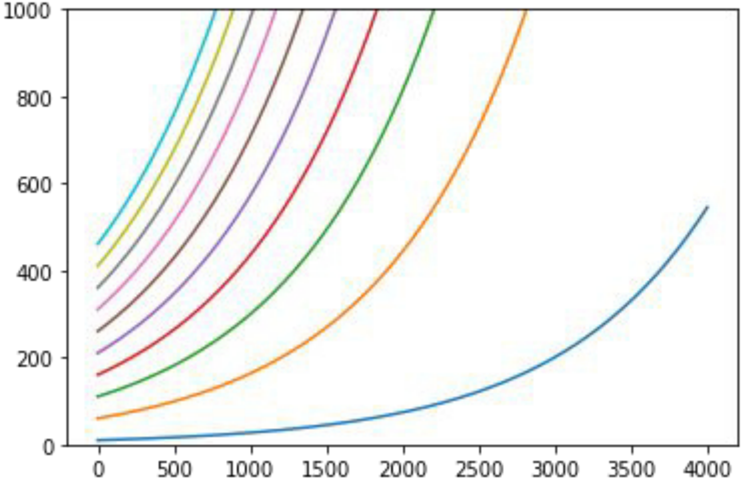
Neuenhoff, Rachel D., et al. "Continued decline of a collapsed population of Atlantic cod (Gadus morhua) due to predation-driven Allee effects." *Canadian Journal of Fisheries and Aquatic Sciences* 76.1 (2019): 168-184.

Allee Effect

The **Allee** effect: “Population growth rate DECREASES as the population size decreases” (we might expect the converse as there would be more resources to go around).

Carrying Capacity term

- A different model of population dynamics incorporates a CARRYING CAPACITY term: (combines prey size, habitat size, nesting sites, etc)

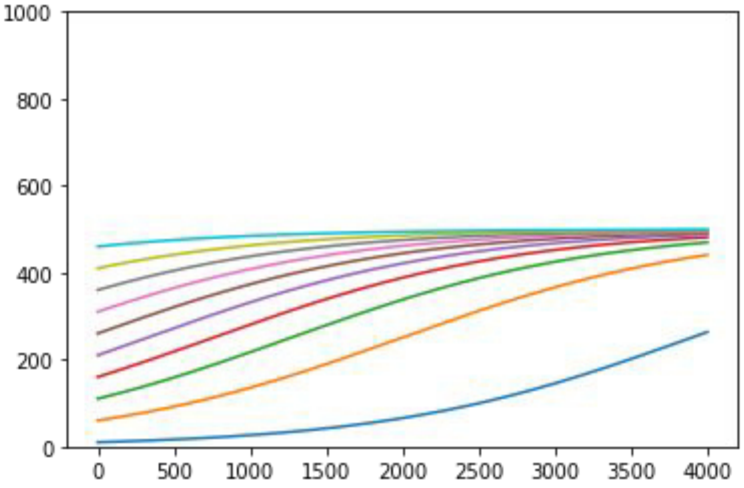


$$\frac{dH}{dt} = rH$$

```
r = 0.1
a = 100
k = 500
```

Carrying Capacity term

- A different model of population dynamics incorporates a CARRYING CAPACITY term: (combines prey size, habitat size, nesting sites, etc)

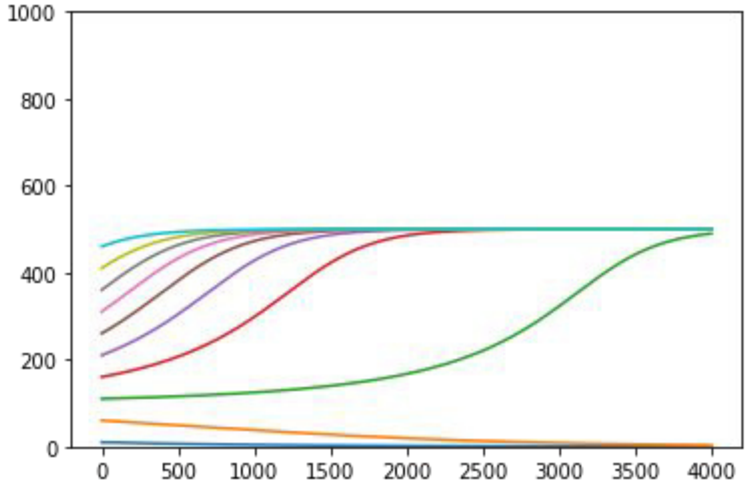


$$\frac{dH}{dt} = rH \times (1 - \frac{H}{k})$$

```
r = 0.1
a = 100
k = 500
```

The Allee effect

- Population growth rate DECREASES as the population size decreases (we might expect the converse as there would be more resources to go around).

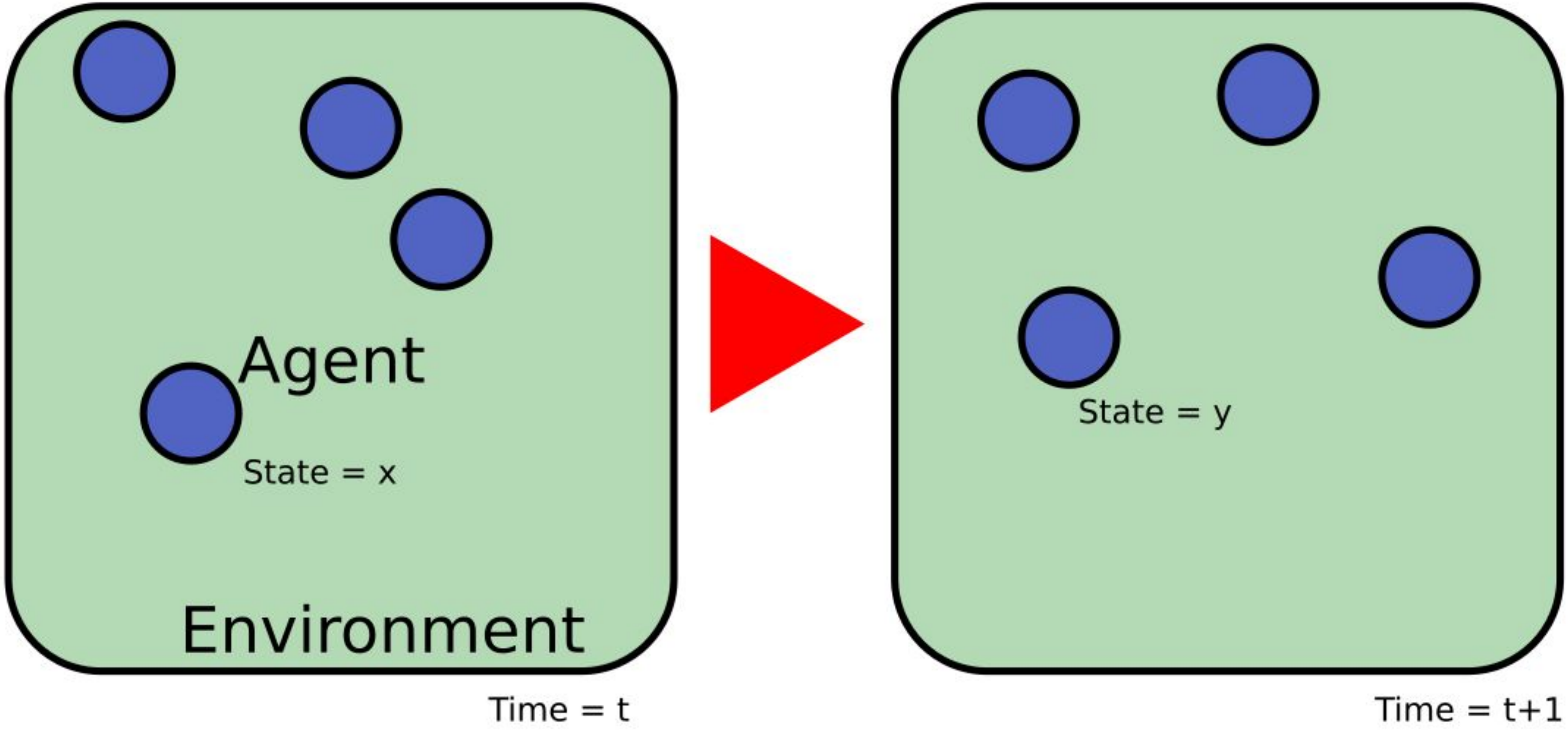


$$\frac{dH}{dt} = rH \times (1 - \frac{H}{k}) \times (\frac{H}{a} - 1)$$

Assumptions…

- Numbers are so large that we can treat them as continuous.
  - If small - they could go locally extinct?
- Homogeneous populations.
  - Might there be variation in a population (health, age, skill, size?)
- Environment is the same over the whole of space and time.
  - Prey Refugees?

Agent-based model

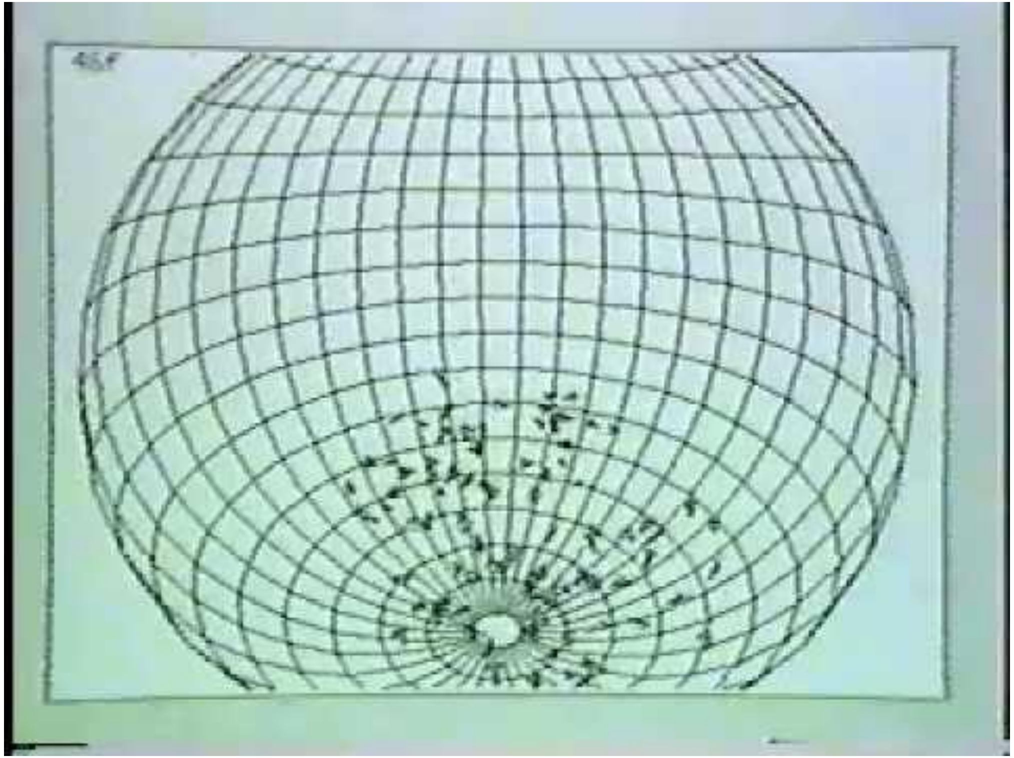


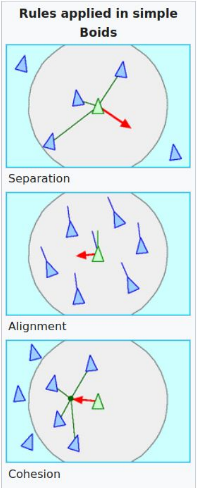
Boids

- Craig Reynolds (1986)

Rules:

- Maintain **separation**.
- **Alignment**: Move in direction aligned with neighbours
- **Cohesion**: Move towards average position.
- Obstacle avoidance.





Emergent Behaviour

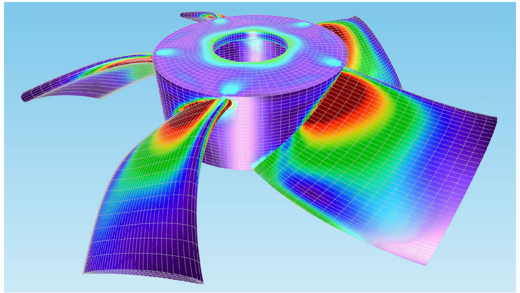
Complex global behaviour that is determined/caused by **local interactions**.

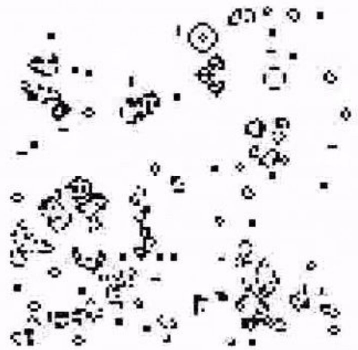
- Flocking
- Microvasculature (and almost everything else in multicellular organisms!)
- Insect foraging
- Galaxy formation
- Crowds/traffic
- Financial markets
- Internet

System level behaviour not always an intuitive result (see pesticide example).

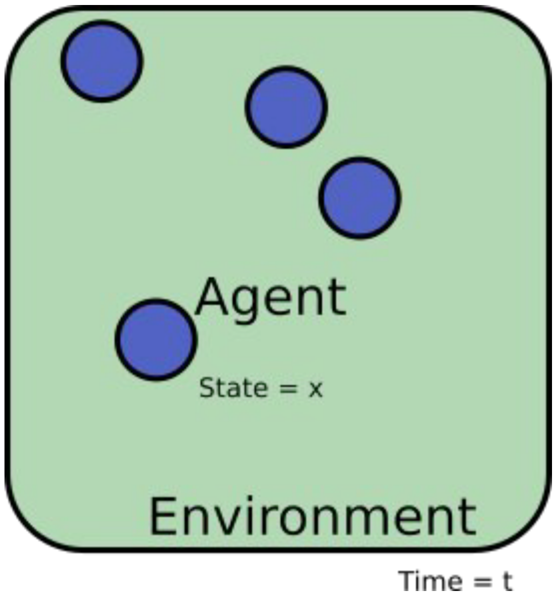
Cellular Automata

- Another similar rule-based modelling approach
- Rules apply to SQUARES not specific INDIVIDUALS
- E.g. might be useful for modelling a fluid, etc, forest fires, etc.

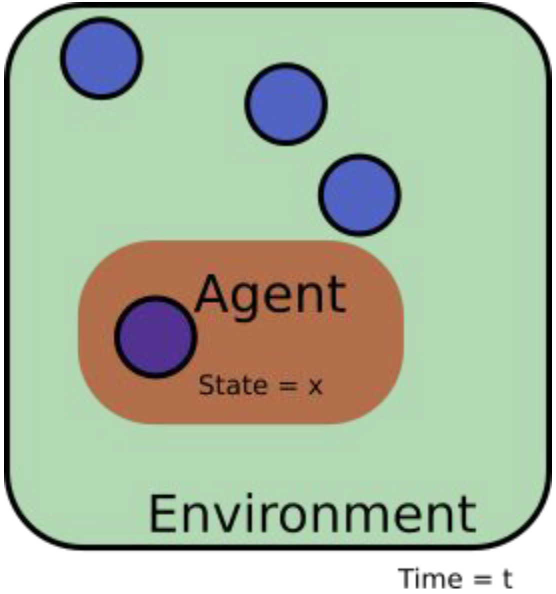




Components of an Agent Based Model

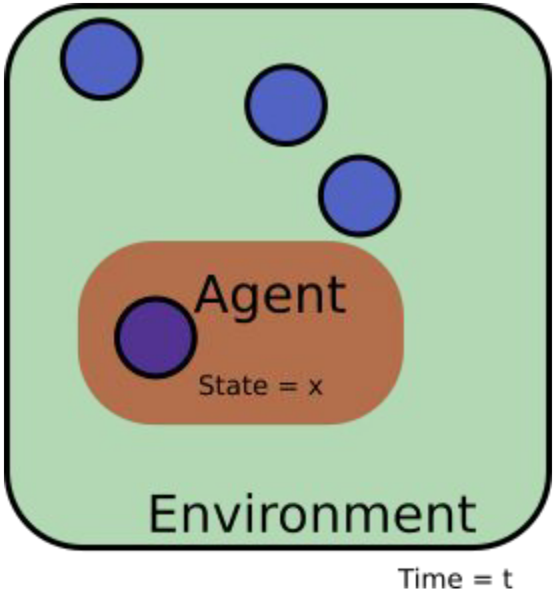


Components of an Agent Based Model: **The Agent**



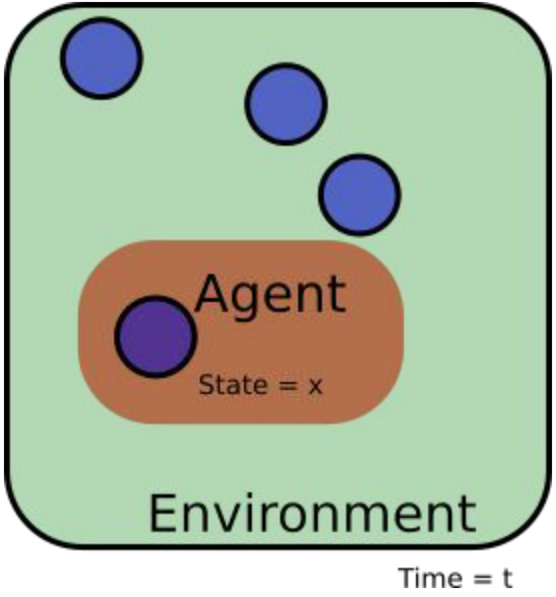
Components of an Agent Based Model: The Agent

- Agents can be of different types (e.g. a fox).
- Data associated:
    - Location (x,y)
    - Age (years)
    - Weight
    - Hunting skill



Components of an Agent Based Model: The Agent

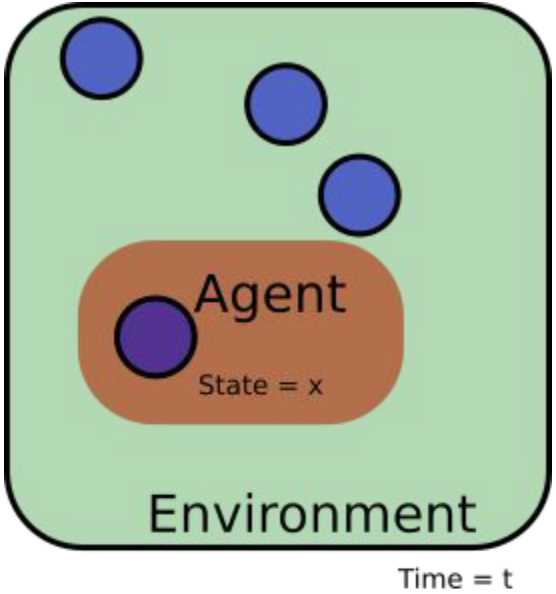
- Agents can be of different types (e.g. a fox).
- Data associated:
    - Location (x,y)
    - Age (years)
    - Weight
    - Hunting skill
- Will exclude:
    - Location of each hair
    - Full genome
    - Brain structure



Components of an Agent Based Model: The Agent

- Agents can be of different types (e.g. a fox).
- Data associated:
  - **Location (x,y)**
  - **Age (years)**
  - **Weight**
  - **Hunting skill**

THE STATE



Components of an Agent Based Model: **Rules**

- Rules describe how the state is updated
  (this includes interactions)

- E.g.

  Rule: **MOVE**

  This might move the agent towards prey.

Components of an Agent Based Model: Rules

- Rules describe how the state is updated (this includes interactions)
- E.g.

  Rule: **MOVE**

  This might move the agent towards prey.

Evolutionary Models: Rules can change over time (from first lecture the model looking at **why** starlings flock into murmurations).

Components of an Agent Based Model: Rules

- Rules describe how the state is updated (this includes interactions)

- E.g.

  Rule: **MOVE**

  This might move the agent towards prey.

Evolutionary Models: Rules can change over time (from first lecture the model looking at **why** starlings flock into murmurations).

What rules?

- Need to be clearly defined and justified!
- Ideally sensitivity analysis should happen.

Components of an Agent Based Model: Rules

- Rules describe how the state is updated (this includes interactions)

- More concrete…

  Rule: **MOVE**

  ```
  if (self.food=0) then
          self.position = self.position + random
  ```

Probabilistic (stochastic)

Components of an Agent Based Model: Rules

- Rules describe how the state is updated (this includes interactions)

- More concrete…

  Rule: **MOVE**

  if (self.food=0) then
  food_location = environment.nearbyfood(self.position)
  move self.location towards food_location

Deterministic

Components of an Agent Based Model: **Environment**

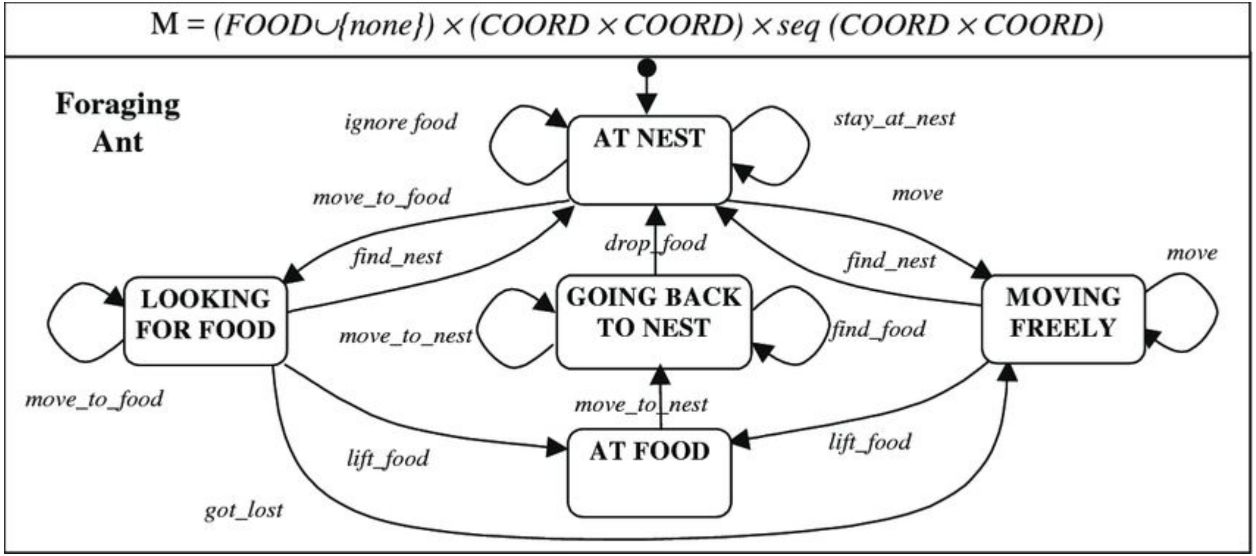
- Typically a constrained area or volume
- Can often be modified by agents
- Presumably will alter agent behaviour or state (otherwise it's a bit pointless)
- Often agents can detect/interact with their **local** environment.
    - E.g. large gravitational model - only local gravitational gradients matter.
    - E.g. Modelling rabbit populations - only the grass local to the current location of the rabbit will be eaten.
    - E.g. Internet routing - only directly connected routers will provide information/data.
    - Etc…
- Sometimes a grid to record local features (e.g. local food)

Components of an Agent Based Model: **Environment**

- Typically a constrained area or volume
- Can often be modified by agents
- Presumably will alter agent behaviour or state (otherwise it's a bit pointless)
- Often agents can detect/interact with their **local** environment.
    - E.g. large gravitational model - only local gravitational gradients matter.
    - E.g. Modelling rabbit populations - only the grass local to the current location of the rabbit will be eaten.
    - E.g. Internet routing - only directly connected routers will provide information/data.
    - Etc…
- Sometimes a grid to record local features (e.g. local food)

Formal approach: X-Machines

- Written as a finite state machine.
  - Input & output streams
  - State (memory)



Kefalas, P., Holcombe, M., Eleftherakis, G., & Gheorghe, M. (2003). A formal method for the development of agent-based systems. In *Intelligent agent software engineering* (pp. 68-98). IGI Global.

Designing an Agent Based Model (conceptual ideas)

- What would we write on paper!

Need to define:

Agents:

- What the agents are? [categories?]
- What do we want to capture/model?

Rules/Behaviours

Environment

Designing an Agent Based Model (conceptual ideas)

- What would we write on paper!

Need to define:

Agents:

- What the agents are? [categories?]
- What do we want to capture/model?

Rules/Behaviours

Environment

What times/lengthscales

- (step size + total length)

Interactions?

- Agent - Agent
- Agent - Environment

Designing an Agent Based Model (conceptual ideas)

- What would we write on paper!

Need to define:

Agents:

- What the agents are? [categories?]
- What do we want to capture/model?

Rules/Behaviours

Environment

What times/lengthscales

- (step size + total length)

Interactions?

- Agent - Agent
- Agent - Environment

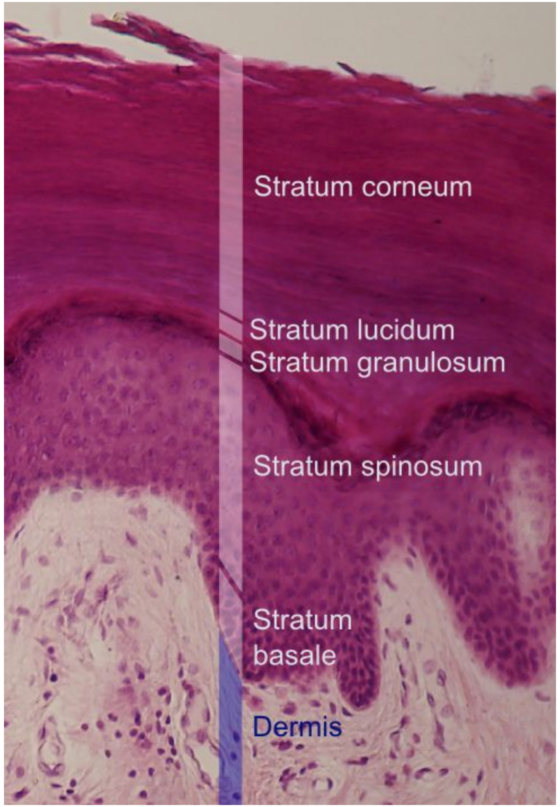
How?

- Talk to experts.
- What do we want to answer (this will drive most decisions).
- What data is there to constrain and validate the model?

Designing an Agent Based Model (conceptual ideas)

Example:

- **Skin**
- Chemical and physical interactions of cells.
- A sub-set of cells can perform cell division.
- Cell differentiation.



By Mikael Häggström, based on work by Wbensmith - File:WVSOM Meissner's corpuslce.JPG by Wbensmith.Layers were drawn according to image at Home Page of Deborah S. Dempsey, Department of Biological SciencesNorthern Kentucky University > V. SKIN > 2 LAYERS ([1]), CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10759398

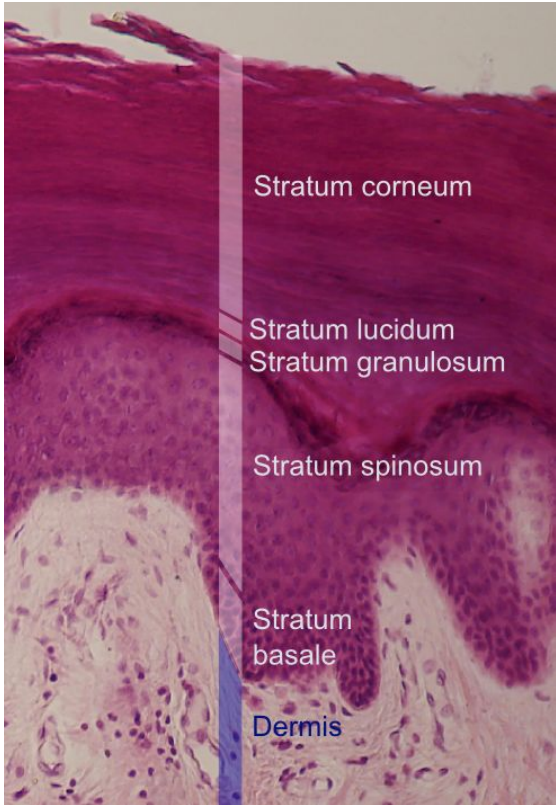
Designing an Agent Based Model (conceptual ideas)

Example:

- **Skin**
- Chemical and physical interactions of cells.
- A sub-set of cells can perform cell division.
- Cell differentiation.

What is the question?

- How does individual cell behaviour affect tissue turnover (& potential abnormalities).



Designing an Agent Based Model (conceptual ideas)

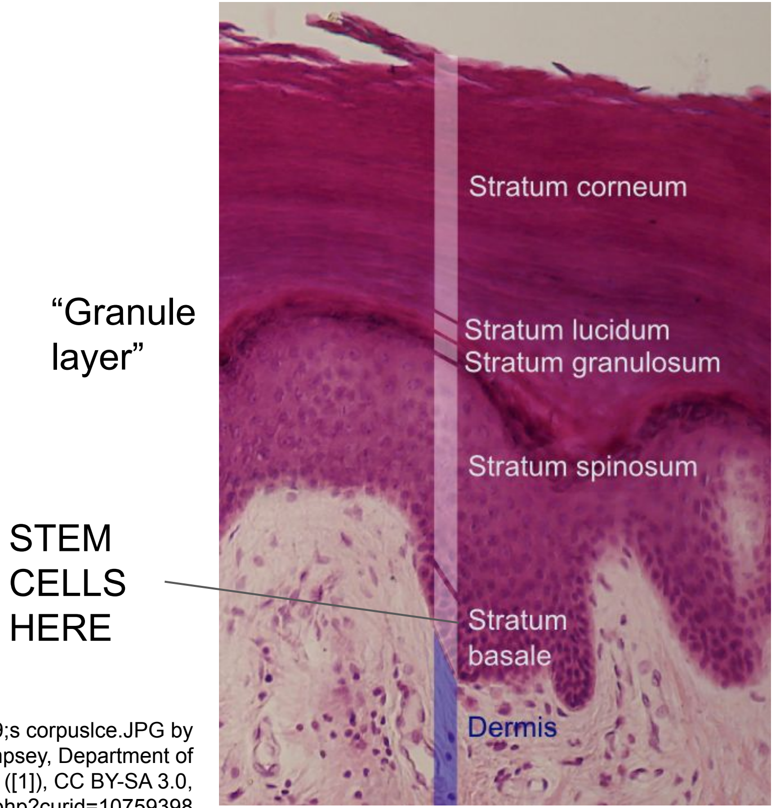
Example:

- **Skin**
- Chemical and physical interactions of cells.
- A sub-set of cells can perform cell division.
- Cell differentiation.

What is the question?

- How does individual cell behaviour affect tissue turnover (& potential abnormalities).

[cell lineage, cells migrate up from stratum basale, differentiating as they go + can only divide limited times].



Designing an Agent Based Model (conceptual ideas)

What are the **agents**?

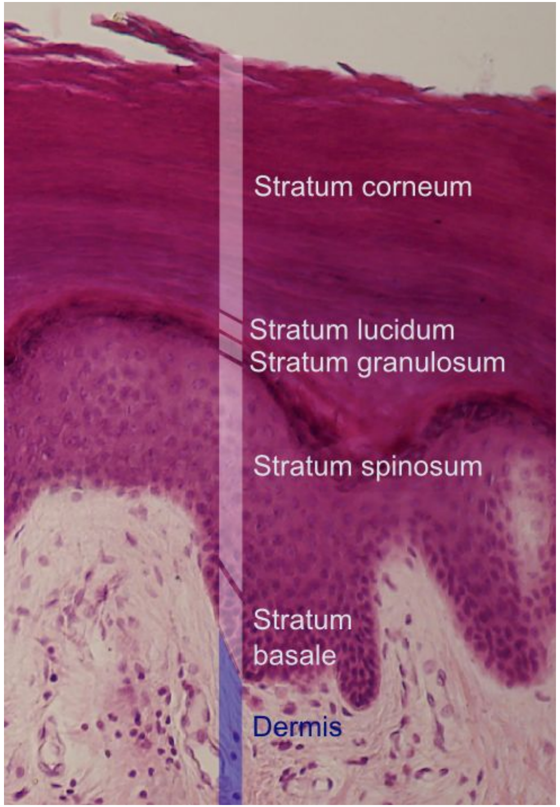
Cells

Sub classes?

E.g. cells in different layers.

Parameters/state?

- How long since last division
- How many divisions
- Location (x,y…z?)
- Protein expression?

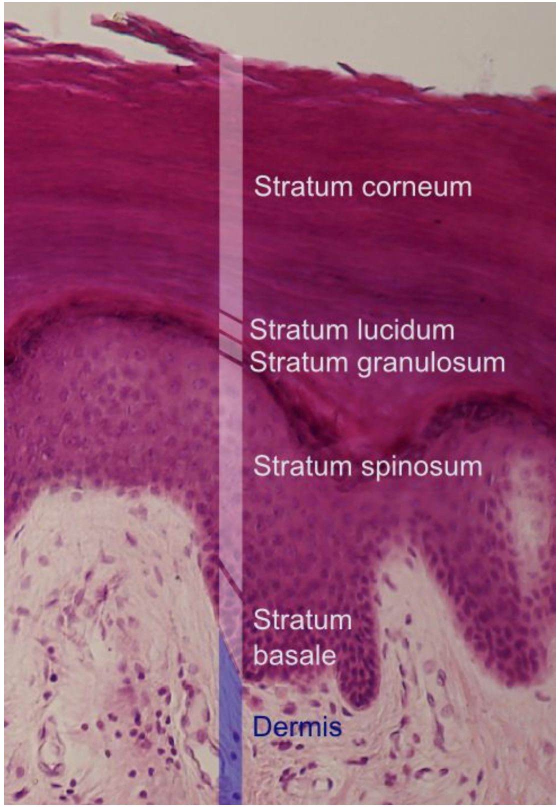


By Mikael Häggström, based on work by Wbensmith - File:WVSOM Meissner's corpuslce.JPG by Wbensmith.Layers were drawn according to image at Home Page of Deborah S. Dempsey, Department of Biological SciencesNorthern Kentucky University > V. SKIN > 2 LAYERS ([1]), CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10759398

Designing an Agent Based Model (conceptual ideas)

What are the **rules**?

Cell division
Cell differentiation
Upward movement (stratification)
Loss of cells from top of tissue (+cell death).

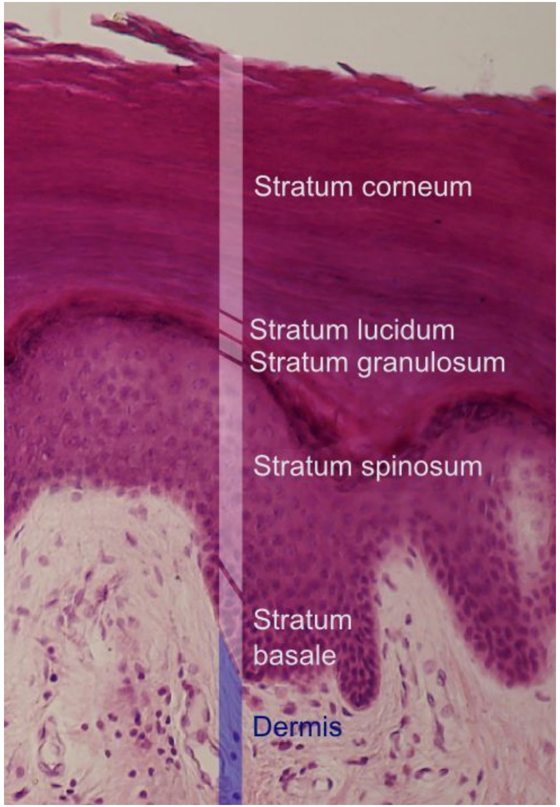


Designing an Agent Based Model (conceptual ideas)

What is the environment?

Simply the dimensions of space.

(do we need distributions of chemical signals, nutrients, oxygen? – would need associated rules).

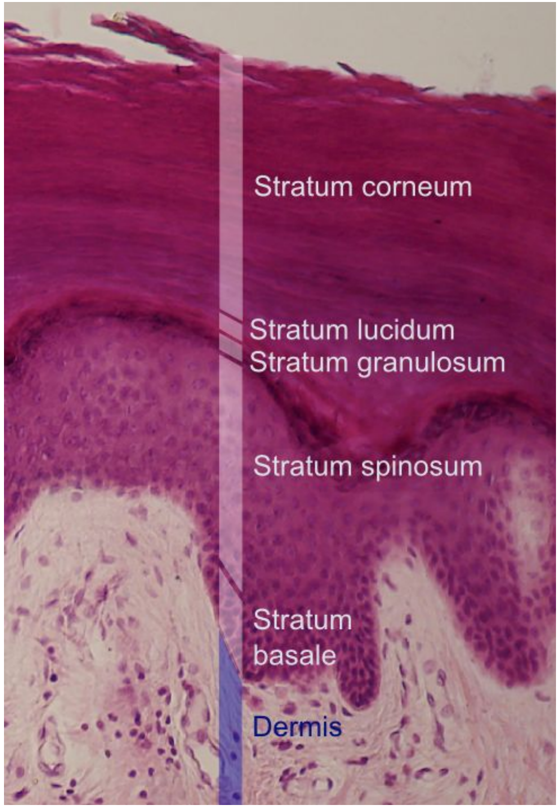


By Mikael Häggström, based on work by Wbensmith - File:WVSOM Meissner's corpuslce.JPG by Wbensmith.Layers were drawn according to image at Home Page of Deborah S. Dempsey, Department of Biological SciencesNorthern Kentucky University > V. SKIN > 2 LAYERS ([1]), CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10759398

Designing an Agent Based Model (conceptual ideas)

Time & Length scales:

Spatial scale:
well, cells are about 10um. What scale is the emergent behaviour? Maybe 1mm? So whole model 1mm across. Agents about 10um.

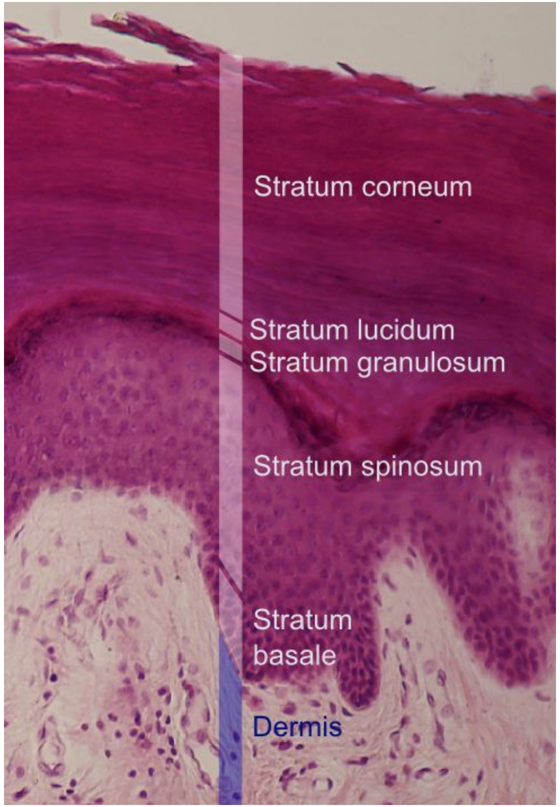


Designing an Agent Based Model (conceptual ideas)

Time & Length scales:

Spatial scale:
well, cells are about 10um. What scale is the emergent behaviour? Maybe 1mm? So whole model 1mm across. Agents about 10um.

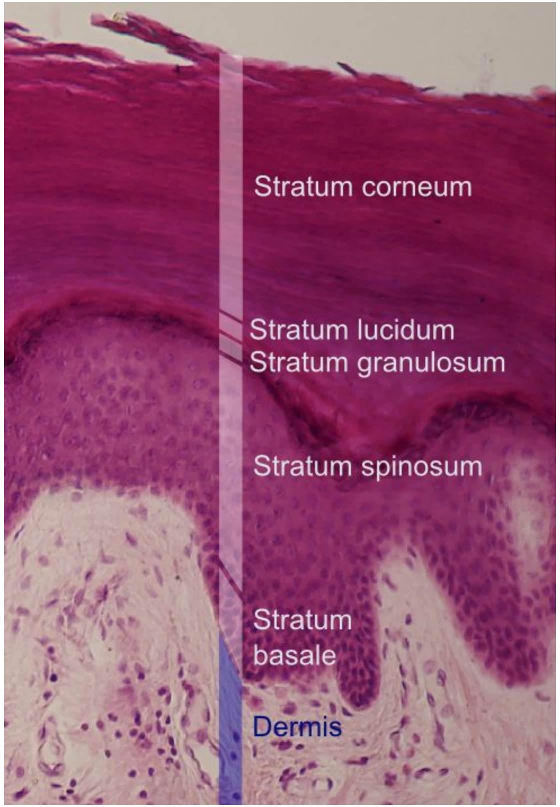
Time steps/total time:
Cell division happens every few hours/days. Total turnover happens over about 28 days. So time step = 1 hour. Total time = 1 month. So will need about 700 iterations.



Designing an Agent Based Model (conceptual ideas)

What types of information will agents exchange between themselves and the environment.

- Local cell density
  - (will need to ‘know’ location of other agents).
  - Could have a **global message list** that cells send their locations to.
- How do cells bond to each other?



Designing an Agent Based Model (conceptual ideas)

Summary. Need to define the:

- Agents
- Rules
- Environment
- Spatial/temporal steps/scales.

Designing an Agent Based Model (conceptual ideas)

Summary. Need to define the:

- Agents
- Rules
- Environment
- Spatial/temporal steps/scales.

Next lecture: How to convert this **conceptual model** into a **simulation**.

COM3001/6009

Modelling and Simulation of Natural Systems

Lecture 3: Implementation of Agent Based Models

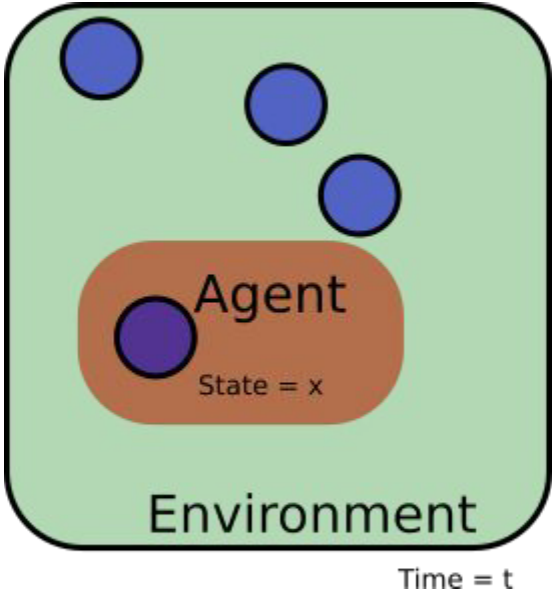
Mike Smith* and Luca Manneschi

*m.t.smith@sheffield.ac.uk

Recap

Identify:

- Agents
- Rule Sets
- Environment
- Spatial/temporal scale & units

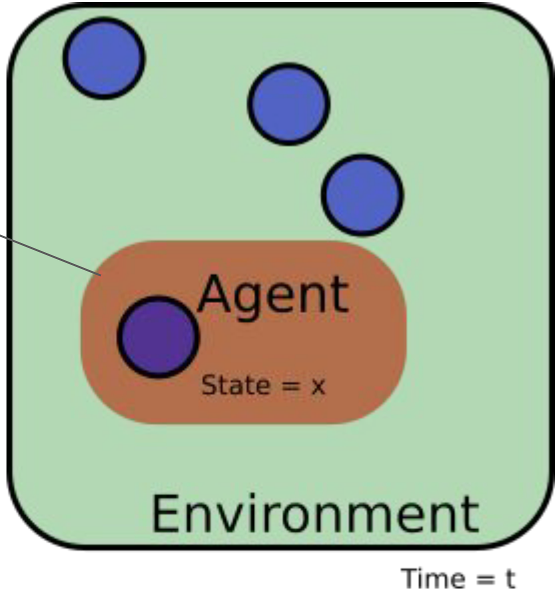


Recap

Identify:

- Agents
- Rule Sets
- Environment
- Spatial/temporal scale & units

Agent simplified into a state - which is updated by model rules.



Agent Implementation

Example: Want to model effect of distribution of forage (flowers) on the expected foraging behaviour of bumblebees & success of cuckoo bees.

Buff-tailed bumblebee
Bombus terrestris (Bill Temples, BBCT)

To represent the agent:

```python
class Bumblebee:
  """
  Base class that can be used for different bumblebees
  """
  def __init__(self,position,food,nest,species):
    self.position = position
    self.food = food
    self.nest = nest
    self.species = species
```

Agent Implementation

Example: Want to model effect of distribution of forage (flowers) on the expected foraging behaviour of bumblebees & success of cuckoo bees.

Buff-tailed bumblebee *Bombus terrestris* (Bill Temples, BBCT)

To represent the agent:

```
class Bumblebee:
  """
  Base class that can be used for different bumblebees
  """
  def __init__(self,position,food,nest,species):
    self.position = position
    self.food = food
    self.nest = nest
    self.species = species
```

State typically refers to fields that can change

Agent Implementation

If there are multiple classes, then object orientation is a useful approach.

Each caste could have its own class, for example, all inherited from the base class.

Updates to the state can made in class methods.

```python
class Worker(Bumblebee):
  """
  Workers visit forage to collect pollen & nectar.
  They learn where good flower patches are. The ids of
  the flower patches it knows are stored in self.flower_patches.
  """
  def __init__(self,position,food,nest,species,flower_patches):
    super().__init__(position,food,nest,species)
    self.flower_patches = flower_patches

class Queen(Bumblebee):
  """
  Queens, once the nest has active workers, will stop foraging
  and remain in the nest. So we don't need to worry about its position.
  """
  def __init__(self,food,nest,species):
    self.food = food
    self.nest = nest
    self.species = species

class Male(Bumblebee):
  """
  Males leave the nest, and don't return (usually), so we don't need
  to store which nest they're from.
  """
  def __init__(self,position,food,species):
    self.position = position
    self.food = food
    self.species = species
```

Agent Implementation

```
bee1 = Bumblebee([2,3],1.0,3,'Buff-tailed bumblebee')
bee2 = Bumblebee([4,5],1.0,4,'Early bumblebee')

bee3 = Worker([1,5],1.0,4,'Early bumblebee',[5,2,3])
bee4 = Male([1,5],1.0,'Early bumblebee')
bee5 = Queen(1.0,4,'Forest cuckoo bumblebee') #in the early bumblebee's nest
```

Environment Implementation

- Geographically constrained (some feature might vary).
- Can be modified by agents
    - (e.g. a flower, when visited will have its level of nectar and pollen reduced).
- Design decision: Do we want flowers to be agents?

```python
import numpy as np
class Environment:
  shape = 'square'
  units = 'm'
  tilesize = 10
  def __init__(self):
    self.flower_density = np.zeros([100,100])
    self.nectar_levels = np.zeros([100,100])  #will be changed by agents
```

Environment Implementation

- Geographically constrained (some feature might vary).
- Can be modified by agents
  - (e.g. a flower, when visited will have its level of nectar and pollen reduced).
- Design decision: Do we want flowers to be agents?
- Do we want a more complex environment?
  - Types of forage?
  - Other landmarks?
  - Good/poor nest habitat?
  - Elevation?
  - Wind?

Rule Implementation

- Makes sense to have rules defined in methods (if using OOP)
- Typically modifies states (of agent and environment).

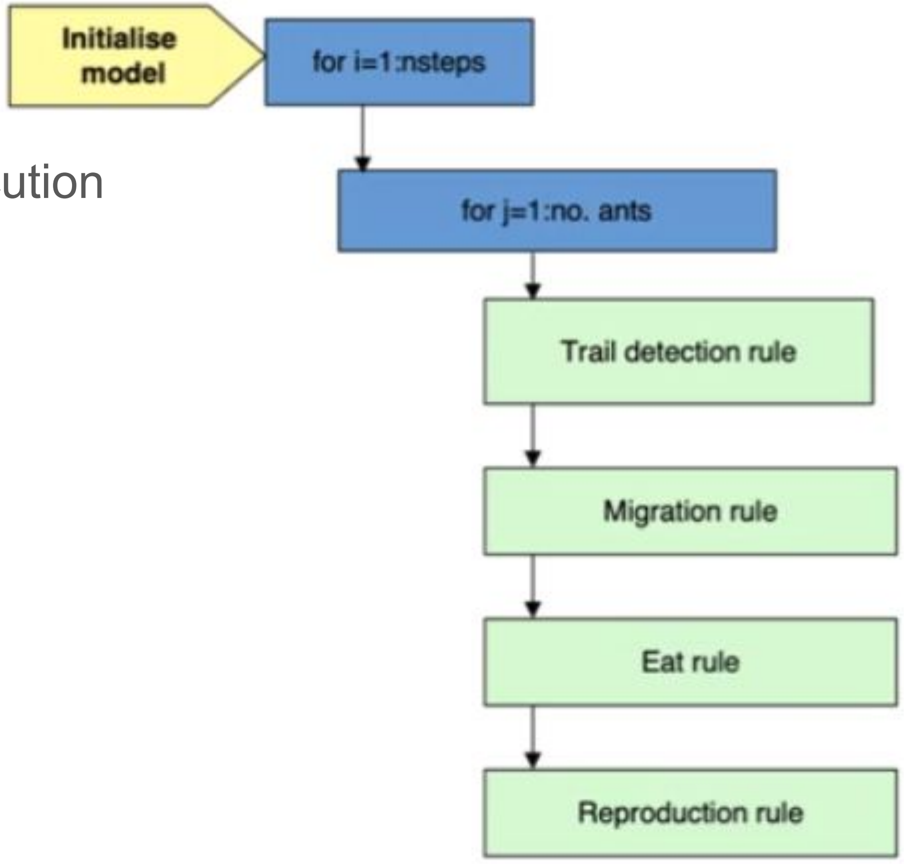
Probably want to have a 'behaviour' variable that says if the bee is flying or feeding?

```
class Worker(Bumblebee):
  """
  Workers visit forage to collect pollen & nectar.
  They learn where good flower patches are. The ids of
  the flower patches it knows are stored in self.flower_patches.
  """
  def __init__(self,position,food,nest,species,flower_patches):
    super().__init__(position,food,nest,species)
    self.flower_patches = flower_patches

  def feed(self,env):
    if env.nectar_levels[self.position[0],self.position[1]]>0:
      self.food+=0.1 #feeding
```

Execution

Need to think about the ORDER of execution



Ant ABM example (from Dawn's lecture)

Model Initialisation Phase

From the ecolab3 example you'll use in the lab:

```
#create default environment
env = Environment(shape=[40,40],growrate=50,maxgrass=5,startgrass=1)

#create agents (rabbits and foxes)
Nrabbits = 40
Nfoxes = 10
agents = []
for i in range(Nrabbits):
    r = Rabbit(position=env.get_random_location(),speed=1)
    agents.append(r)
for i in range(Nfoxes):
    f = Fox(position=env.get_random_location(),speed=5)
    agents.append(f)
```

Execution Phase

Need to think about the ORDER of execution

- Example code from ecolab3
  - Think about what the consequences are of the order…

An agent could be dead and still the breed method will be called…

```
for it in range(Niterations):
    #for each agent, apply rules (move, eat, breed)
    for agent in agents:
        agent.move(env)
        agent.eat(env,agents)
        a = agent.breed()
        if a is not None: agents.append(a)

    #removed dead agents
    agents = [a for a in agents if not a.die()]

    #grow more grass
    env.grow()
```

Activity

In pairs/tables think about the following questions to be answered using an ABM:

- Bumblebee (a haplodiploid genus) what is the minimal viable population?
- North Atlantic cod & predation by seals / fishing (what will the population do next?)
- Internet ‘oscillations’ due to queuing rules (what effect do different rules have?)

Define:

- agent & its variables/parameters/states
- rules
- environment
- length/time scales [resolution & extent]

6 minutes (2 minutes each)

Communication and Timing

Agent Based Models are often interested in the effect of **interactions**.

A method implementing an agent's rule will need to know:

- Where other agents are
- Their state or type (e.g. is it a predator?)

When/how information is exchanged?

- Can lead to unwanted artefacts
- Computational and memory issues

Communication and Timing

- Suppose Agent A (a fox) attacks and kills Agent B (a rabbit). How does Agent B “know” that it has been eaten?

Direct Communication: During Agent A’s “feed” method, Agent B is updated

-
-
-

Option 2 **Indirect Communication**: Agent A sends a message that Agent B will later read.

-
-

Communication and Timing

- Suppose Agent A (a fox) attacks and kills Agent B (a rabbit). How does Agent B “know” that it has been eaten?

Direct Communication: During Agent A’s “feed” method, Agent B is updated

- Less computation needed & no need to manage messages
- Harder to trace interactions
- Synchronous model update isn’t possible

Option 2 **Indirect Communication**: Agent A sends a message that Agent B will later read.

-
-

Communication and Timing

- Suppose Agent A (a fox) attacks and kills Agent B (a rabbit). How does Agent B “know” that it has been eaten?

Direct Communication: During Agent A’s “feed” method, Agent B is updated

- Less computation needed & no need to manage messages
- Harder to trace interactions
- Synchronous model update isn’t possible

Option 2 **Indirect Communication**: Agent A sends a message that Agent B will later read.

- Easier to trace interactions
- Needs extra memory and code

Communication and Timing

- Suppose Agent A (a fox) attacks and kills Agent B (a rabbit). How does Agent B “know” that it has been eaten?

Direct Communication: Du

- Less computation need
- Harder to trace interacti
- Synchronous model upd

Ecolab3 uses **direct communication**. However, Dawn used indirect communication in her version of the module.

Option 2 **Indirect Communication**: Agent A sends a message that Agent B will later read.

- Easier to trace interactions
- Needs extra memory and code

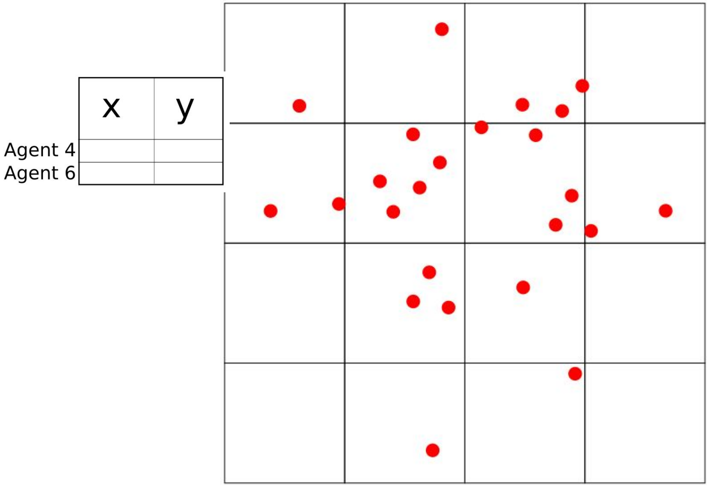
Messages (for Indirect Communication)

- **Global** Message Lists
    - E.g. a table of the locations of all agents
    - Each agent will write its location to this list
    - This is **in addition** to their local representation [it might not match, as this might be written to at the start of the loop]

	x	y
Agent 1		
Agent 2		
Agent 3		
Agent 4		
Agent 5		
Agent 6		
Agent 7		
Agent 8		

Messages (for Indirect Communication)

- **Global** Message Lists
  - E.g. a table of the locations of all agents
  - Each agent will write its location to this list
  - This is **in addition** to their local representation [it might not match, as this might be written to at the start of the loop]
- **Local** Message Lists
  - Might divide into a grid
  - Means only nearby agents will be ‘visible’.
  - Might have computational advantages! (parallelism)
  - FLAME:
    - FLAME is a **flexible and generic agent-based modelling platform**



Messages (for Indirect Communication)

- **Directed** Messages
  - Rather than a list of messages, we might want to share direct messages with other agents.
  - This could have an NxN array (or a single ‘in-box’) list for each agent

Might use a mix of approaches depending on rule.

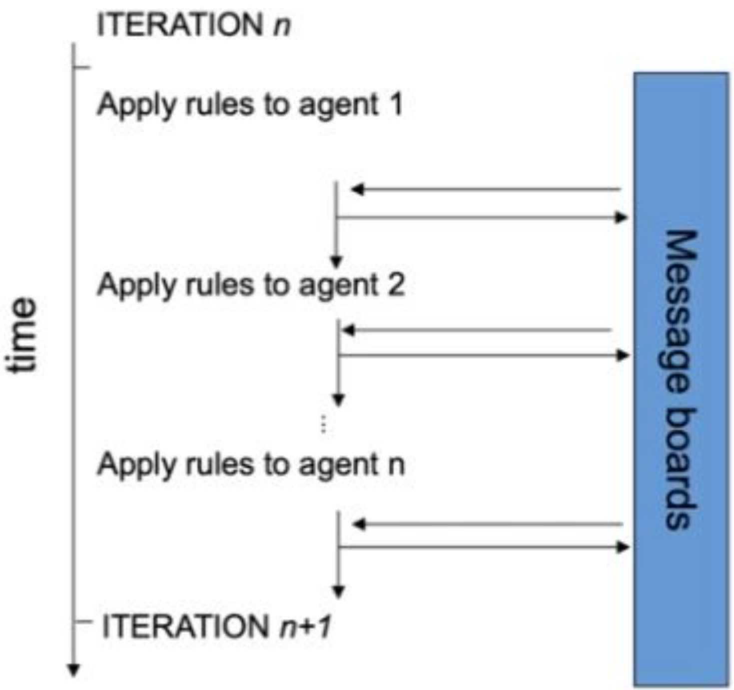
Timing and Synchronisation

- Real World: Time is continuous
- Model: Discrete time steps
    - Can lead to artefacts

Asynchronous Update

From Dawn’s slides:

- Here we apply rules AND write messages immediately.
- Communication occurs immediately after each agent is updated.
-
- Advantage: Agents act on current info & can be implemented with direct communication.
- Disadvantage: First agent can get an advantage.

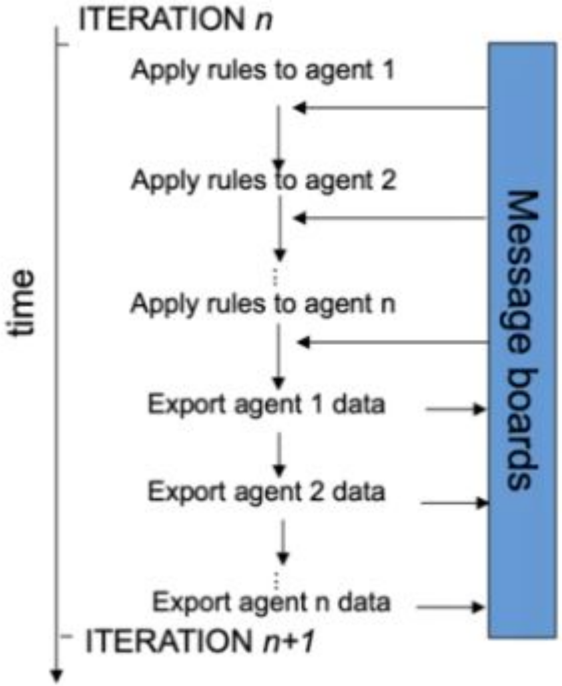


Synchronous Update

From Dawn's slides:

- Communication only occurs at the end of each loop.

- Advantage: Ordering doesn't give any advantages.
- Disadvantage: Might act on incorrect information.



Example of asynchronous issue

- If we have three agents:
  - Fox A is first in the list.
  - Fox B is second in the list (right next to the rabbit).
  - Rabbit C is third in the list.
- We loop over the agents.
  - Fox A catches and eats Rabbit C.
  - Fox B, although was closer, had its ‘hunt’ method run after Fox A’s.
  - Rabbit C could have escaped but its ‘move’ method was run last.
- Can be solved partly by randomising the order of agents.
- (not always necessary!)

Fox A

Fox B

Rabbit C

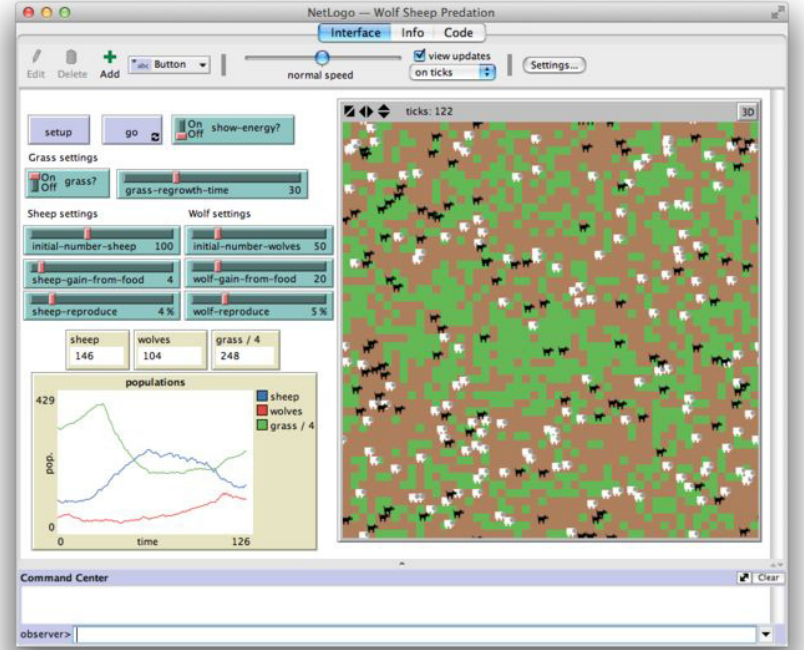
ABM Tools

Will be asking you to code your ABM directly (using python) not using a tool.
Because:

- Often tools are quite restrictive
- Need to learn something that’s not very relevant.
- Also by implementing it yourself you can learn more.

ABM Tools

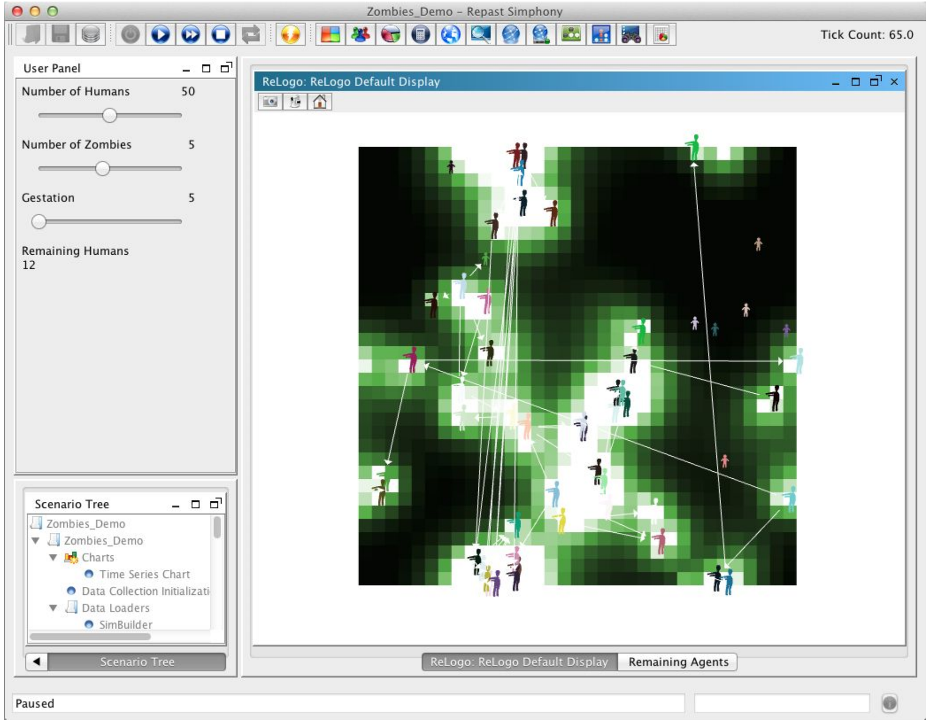
NetLogo



ABM Tools

NetLogo

Repast

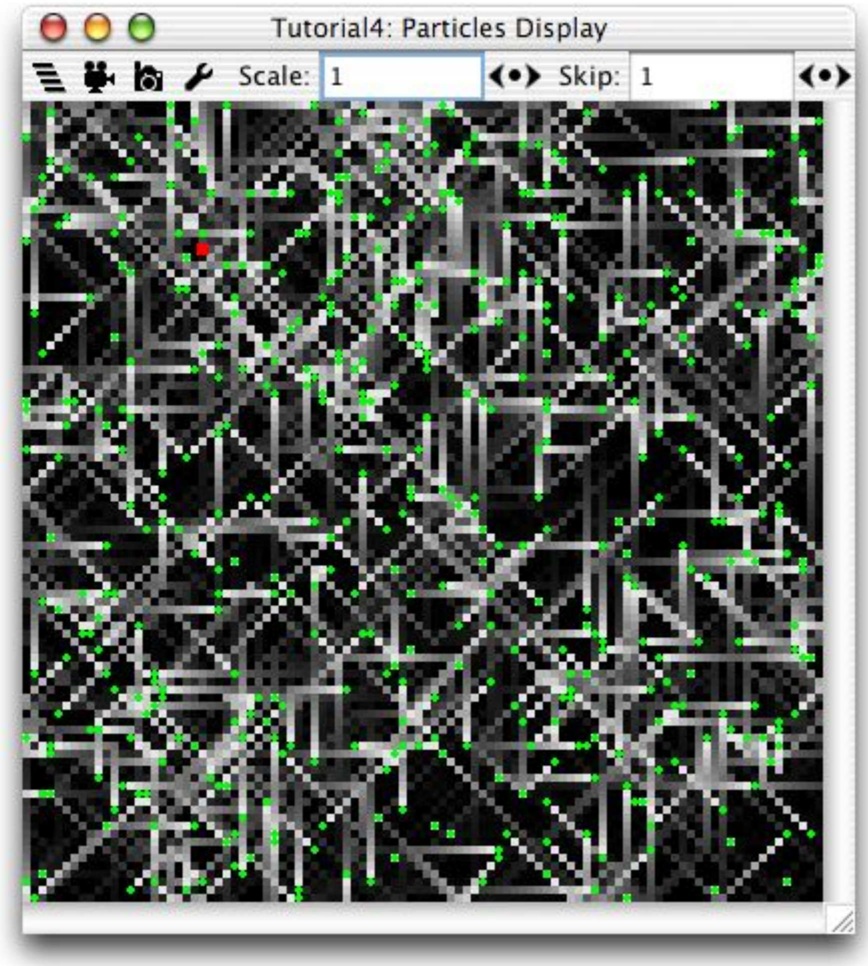


ABM Tools

NetLogo

Repast

MASON



ABM Tools

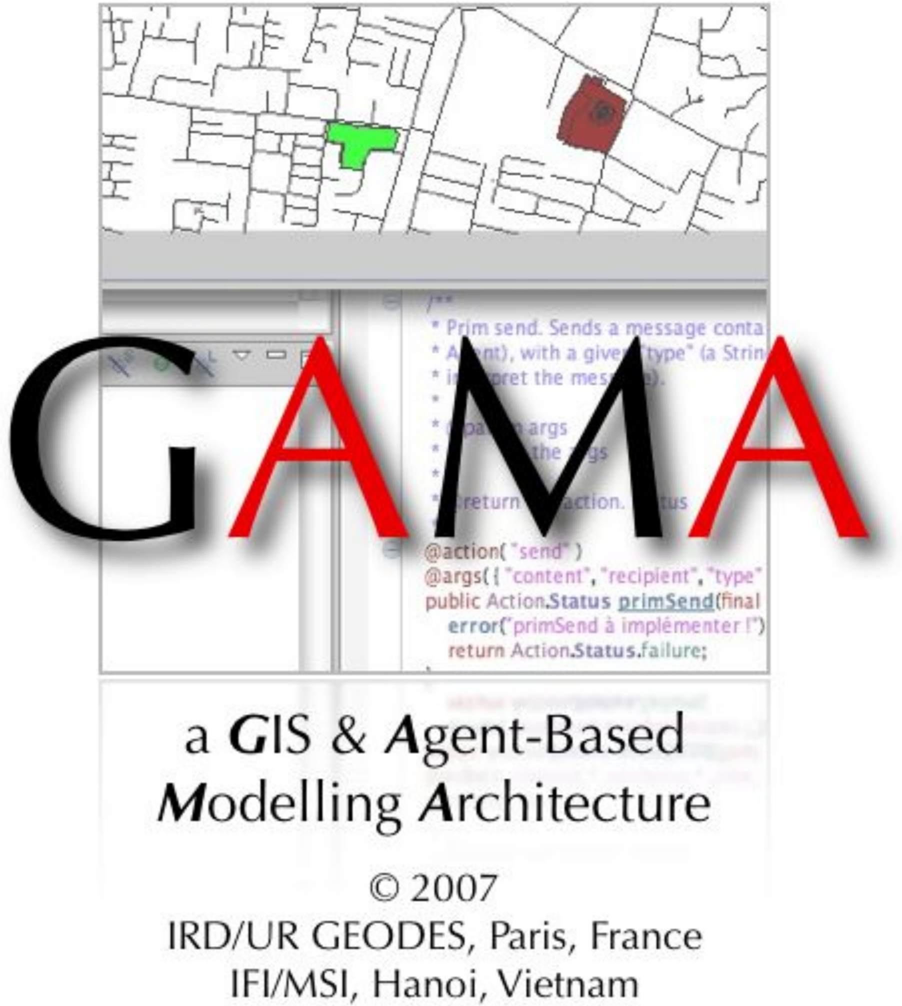
NetLogo

Repast

MASON

GAMA

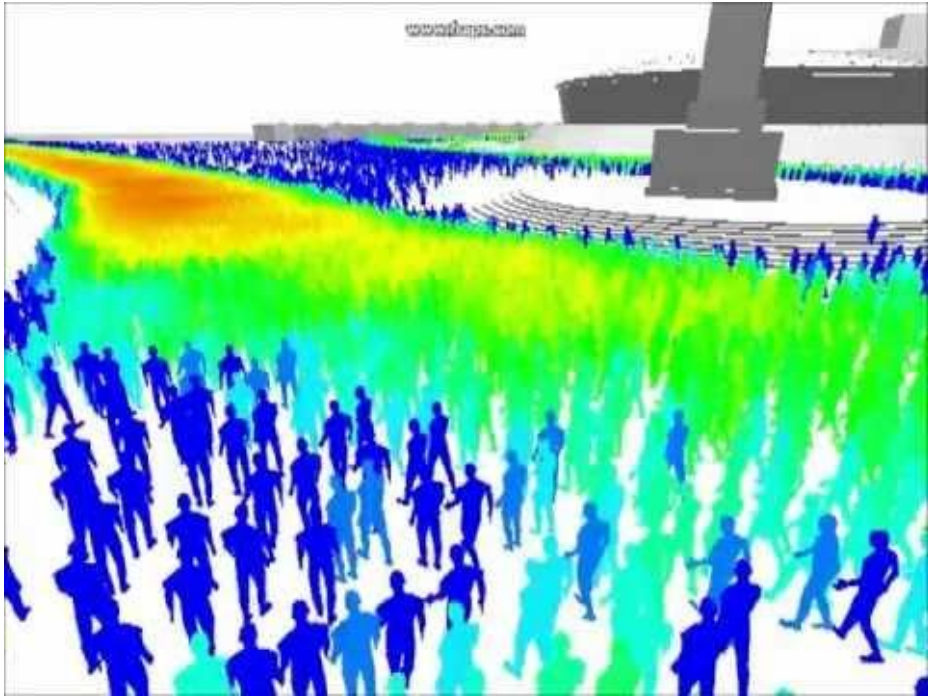
FLAME



ABM Tools

- FLAME / FLAME GPU:
- Flexible Large Scale Agent Modelling Environment.
- Model defined with an XML representation (quite verbose).
- Can run on GPUs.

(developed at Sheffield)



Summary

- How/when agents communicate information
    - Direct or Indirect?
    - Use messages?
        - Global, Local or Directed?
- Timing
    - Be careful not to cause artefacts!
- Update:
    - Asynchronous or Synchronous?

(no general right/wrong way - but need to have a reason for a given scenario).

Next: Parameterisation and Validation in ABMs.

COM3001/6009
Modelling and Simulation of Natural Systems
Lecture 4: Building & Parameterising ABMs

Mike Smith* and Luca Manneschi

*m.t.smith@sheffield.ac.uk

Introduction

- Problem Modelling Biological (vs Physical) systems
- Sources of error
- How to relate agent rules to empirically observed data
- Introduce CoSMoS
- Model Validation / Parameter Sensitivity

Problem Modelling Biological (vs Physical) systems

Pendulum:

Can model very well (simple harmonic motion)

$$\frac{d^2x}{dt^2} = -w^2x$$

Has a closed form, exact solution, e.g.

$$C\cos(wt + \theta)$$

Problem Modelling Biological (vs Physical) systems

Advection/Diffusion of (e.g.) pollution over space:

$$\frac{\partial u}{\partial t} + p_1 \nabla u - \nabla.(p_2 \nabla u) = f$$

f = forcing function (might vary over space & time)
u = concentration of pollution

Problem Modelling Biological (vs Physical) systems

- As systems get more complicated, we stop being able to write down a list of physical laws that we can solve: There are no “fundamental” theoretical laws for biology.

Problem Modelling Biological (vs Physical) systems

- As systems get more complicated, we stop being able to write down a list of physical laws that we can solve: There are no “fundamental” theoretical laws for biology.
- Behaviour in particular needs something agent based.

Problem Modelling Biological (vs Physical) systems

- As systems get more complicated, we stop being able to write down a list of physical laws that we can solve: There are no “fundamental” theoretical laws for biology.
- Behaviour in particular needs something agent based.
- More difficult to observe and measure, e.g.:
    - Hard to record from many neurons simultaneously
    - Hard to track biological targets
    - Often need to run in the lab (or in vitro) which might not be representative.

Sources of Error

- Abstraction Error - How much our model is fundamentally wrong.
- Epistemic Error - Uncertainty in our model parameters.
  - I like to split this again into uncertainty in our model's hyperparameters and those parameters which describe the current state.
- Aleatoric Error - in some sense this is unknowable errors/noise.

Can't eliminate these sources, but it is important to estimate the scale of these errors (or their consequence on the features of interest).

On board: predicting air pollution

Model Parameters

The equation based model for predator/prey population dynamics had four parameters - these were all ‘rates’.

Let H be the number of prey, and P the number of predators.

Births of prey

Prey being eaten by predators

Rate of change of prey

$$\frac{dH}{dt} = rH - cHP$$

Rate of change of predators

$$\frac{dP}{dt} = caHP - mP$$

Predator births

Predator mortality

Conversely, Agent Based Model parameters usually parameterise behaviour/rules.

How to choose parameters?

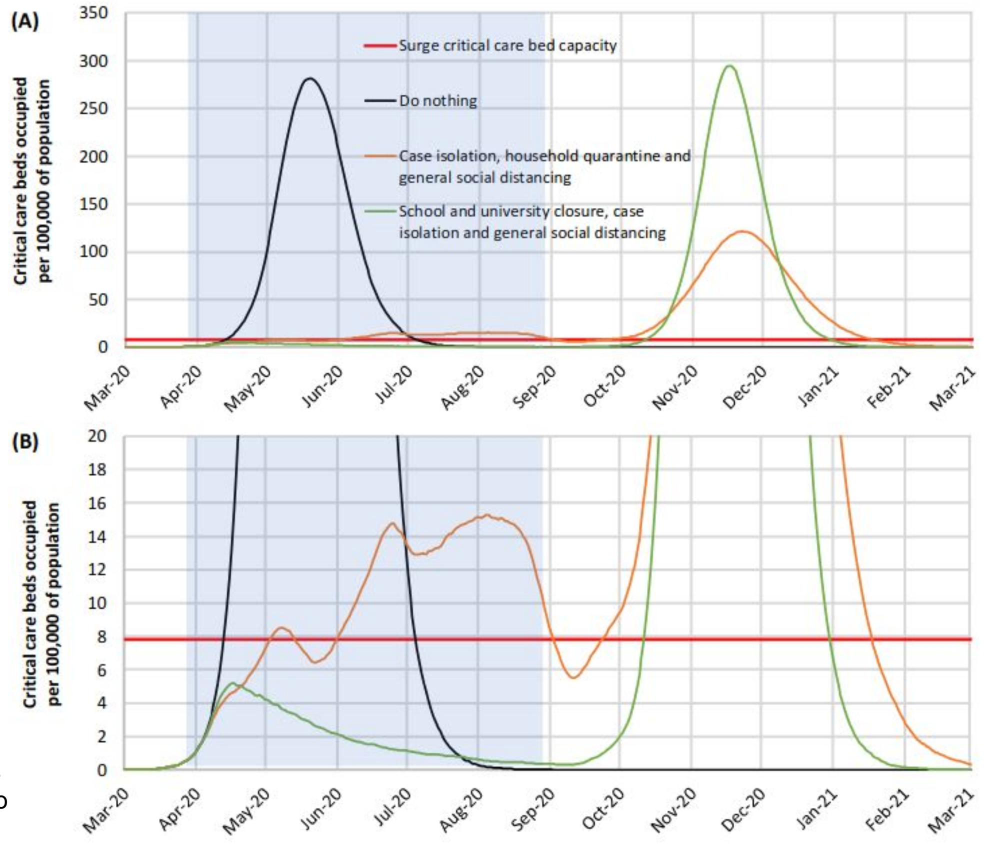
- COVID19 example…
- If we think back to early March 2020
- Ferguson's team's modelling - **used an ABM**.
    - Arguably the conclusions of the report were quite obvious from earlier: Maybe could have just looked at Italy / Wuhan, but still a lot of uncertainty.
        - (separate discussion about using a cost matrix, and how to reason under uncertainty).
- Purpose: They wanted to consider different measures to see what effect they would have on critical care.

The moment I realised exactly how serious things were was when I read the report by Ferguson et al. 2020 (~14-16 March). A week later lockdown.
Ferguson, Neil M., et al. "**Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand**." (2020).
[cited 3495 times]

How to choose parameters?

In spite of these criticisms they seemed to be roughly correct:

- First wave peak was covid occupied 5.8 critical care beds / 100,000; peaking mid-April).
- “social distancing measures might have to be imposed for 18 months or more, at least intermittently, until a vaccine is developed and tested”*



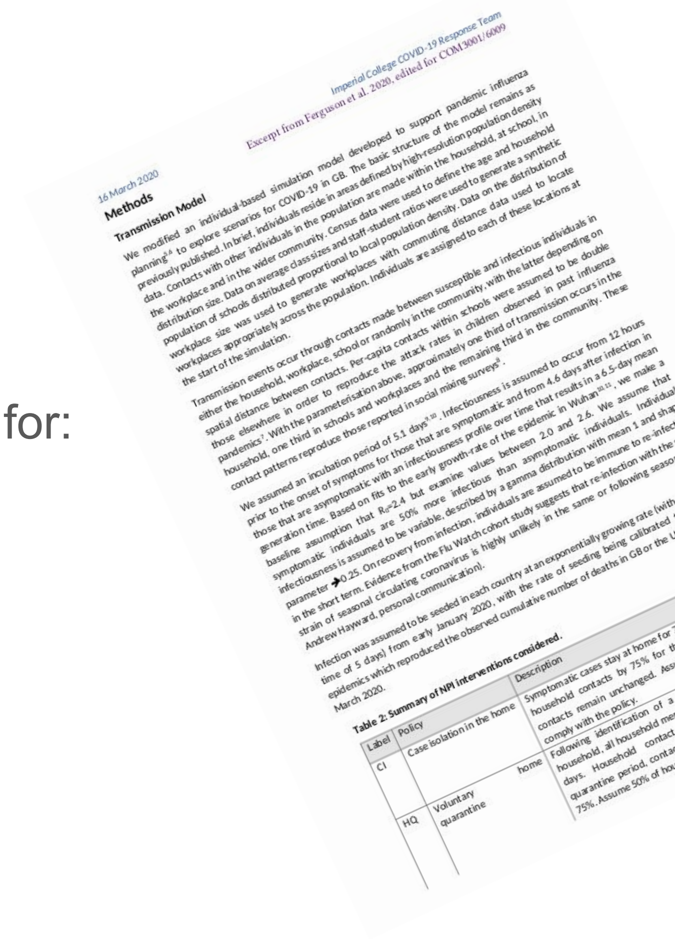
*From a press discussion/critque of the report and its wider policy implicatications https://www.nytimes.com/2020/03/17/world/europe/coronavirus-imperial-college-johnson.html

How to choose parameters? **Activity**

In groups of 2-4:

Skim through this excerpt of the methods section, look for:

- How they've based parameters on literature, etc.



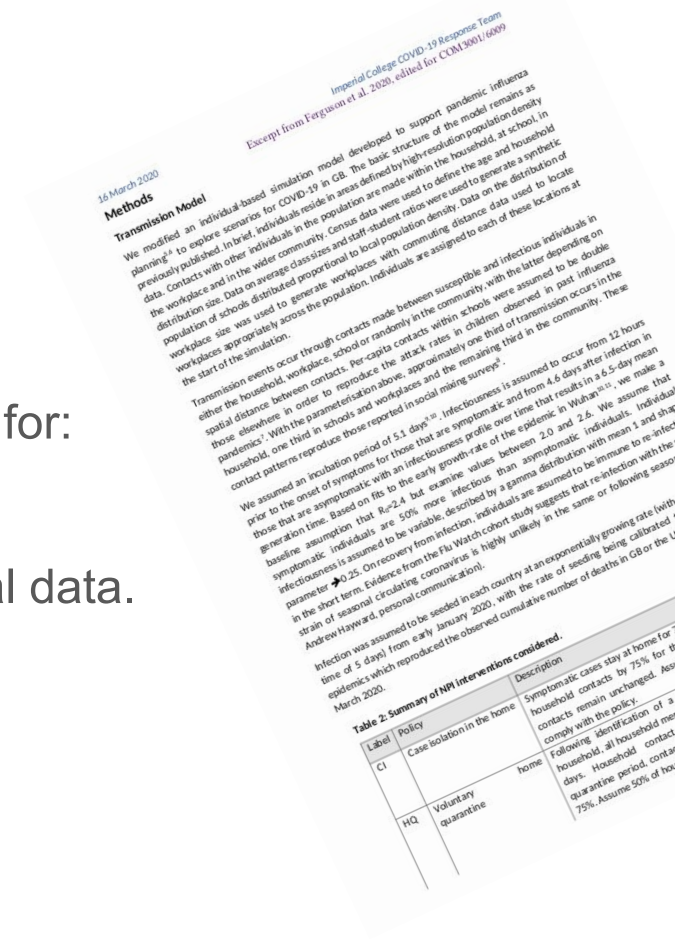
5 - 10 minutes

How to choose parameters? **Activity**

In groups of 2-4:

Skim through this excerpt of the methods section, look for:

- How they’ve based parameters on literature, etc.
- Validations of parts of their model against empirical data.



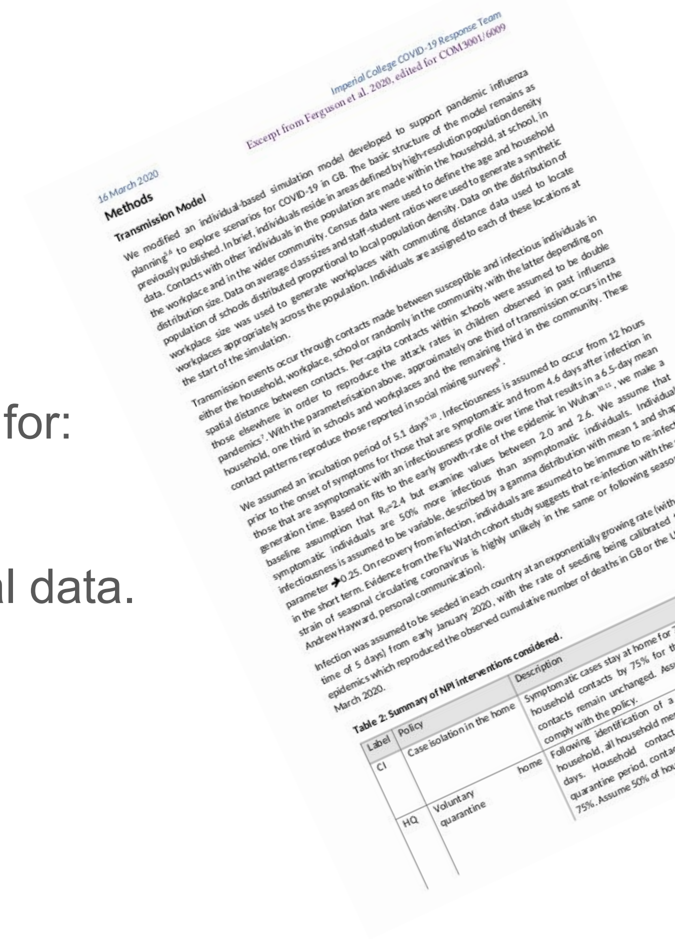
5 - 10 minutes

How to choose parameters? **Activity**

In groups of 2-4:

Skim through this excerpt of the methods section, look for:

- How they’ve based parameters on literature, etc.
- Validations of parts of their model against empirical data.
- Where guesses have been made.


Imperial College COVID-19 Response Team
Excerpt from Ferguson et al. 2020, edited for COM3001/6009

16 March 2020

Methods

Transmission Model

We modified an individual-based simulation model developed to support pandemic influenza planning[3,4] to explore scenarios for COVID-19 in GB. The basic structure of the model remains as previously published. In brief, individuals reside in areas defined by high-resolution population density data. Contacts with other individuals in the population are made within the household, at school, in the workplace and in the wider community. Census data were used to define the age and household distribution size. Data on average class sizes and staff-student ratios were used to generate a synthetic population of schools distributed proportional to local population density. Data on the distribution of workplace size was used to generate workplaces with commuting distance data used to locate workplaces appropriately across the population. Individuals are assigned to each of these locations at the start of the simulation.

Transmission events occur through contacts made between susceptible and infectious individuals in either the household, workplace, school or randomly in the community, with the latter depending on spatial distance between contacts. Per-capita contacts within schools were assumed to be double those elsewhere in order to reproduce the attack rates in children observed in past influenza pandemics[7]. With the parameterisation above, approximately one third of transmission occurs in the household, one third in schools and workplaces and the remaining third in the community. These contact patterns reproduce those reported in social mixing surveys[8].

We assumed an incubation period of 5.1 days[9,10]. Infectiousness is assumed to occur from 12 hours prior to the onset of symptoms for those that are symptomatic and from 4.6 days after infection in those that are asymptomatic with an infectiousness profile over time that results in a 6.5-day mean generation time. Based on fits to the early growth-rate of the epidemic in Wuhan[10,11], we make a baseline assumption that $R_0=2.4$ but examine values between 2.0 and 2.6. We assume that symptomatic individuals are 50% more infectious than asymptomatic individuals. Individual infectiousness is assumed to be variable, described by a gamma distribution with mean 1 and shap parameter ➔0.25. On recovery from infection, individuals are assumed to be immune to re-infect in the short term. Evidence from the Flu Watch cohort study suggests that re-infection with the strain of seasonal circulating coronavirus is highly unlikely in the same or following seaso Andrew Hayward, personal communication).

Infection was assumed to be seeded in each country at an exponentially growing rate (with time of 5 days) from early January 2020, with the rate of seeding being calibrated epidemics which reproduced the observed cumulative number of deaths in GB or the U March 2020.

Table 2: Summary of NPI interventions considered.

Label	Policy	Description
CI	Case isolation in the home	Symptomatic cases stay at home for household contacts by 75% for th contacts remain unchanged. Ass comply with the policy.
HQ	Voluntary home quarantine	Following identification of a household, all household me days. Household contac quarantine period, conta 75%. Assume 50% of hou

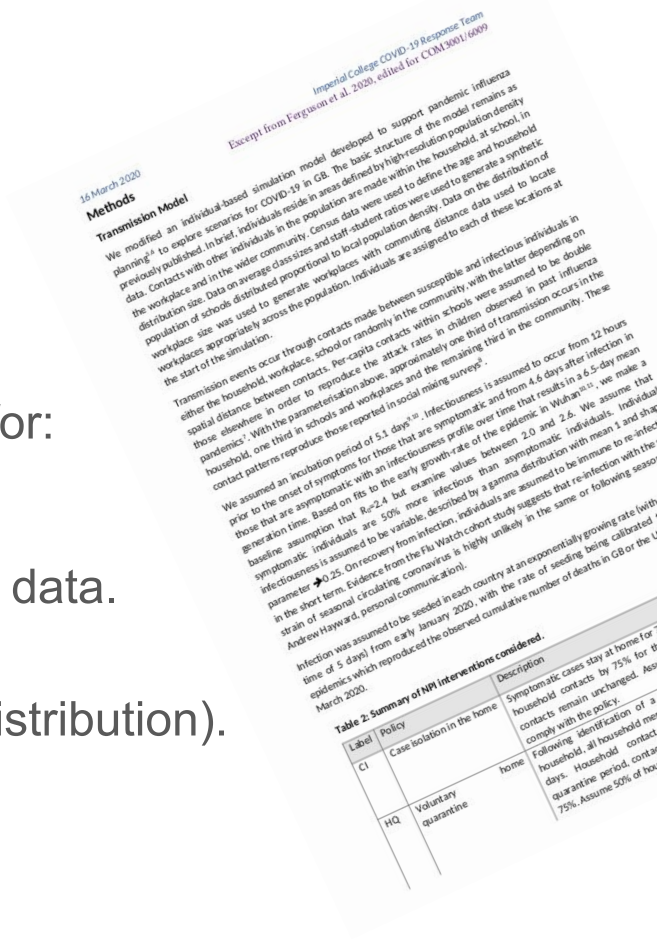
5 - 10 minutes

How to choose parameters? **Activity**

In groups of 2-4:

Skim through this excerpt of the methods section, look for:

- How they've based parameters on literature, etc.
- Validations of parts of their model against empirical data.
- Where guesses have been made.
- Where the variation of agents is modelled (with a distribution).



Imperial College COVID-19 Response Team

Excerpt from Ferguson et al. 2020, edited for COM3001/6009

16 March 2020

Methods

Transmission Model

We modified an individual-based simulation model developed to support pandemic influenza planning[3,4] to explore scenarios for COVID-19 in GB. The basic structure of the model remains as previously published. In brief, individuals reside in areas defined by high-resolution population density data. Contacts with other individuals in the population are made within the household, at school, in the workplace and in the wider community. Census data were used to define the age and household distribution size. Data on average class sizes and staff-student ratios were used to generate a synthetic population of schools distributed proportional to local population density. Data on the distribution of workplace size was used to generate workplaces with commuting distance data used to locate workplaces appropriately across the population. Individuals are assigned to each of these locations at the start of the simulation.

Transmission events occur through contacts made between susceptible and infectious individuals in either the household, workplace, school or randomly in the community, with the latter depending on spatial distance between contacts. Per-capita contacts within schools were assumed to be double those elsewhere in order to reproduce the attack rates in children observed in past influenza pandemics[7]. With the parameterisation above, approximately one third of transmission occurs in the household, one third in schools and workplaces and the remaining third in the community. These contact patterns reproduce those reported in social mixing surveys[8].

We assumed an incubation period of 5.1 days[9,10]. Infectiousness is assumed to occur from 12 hours prior to the onset of symptoms for those that are symptomatic and from 4.6 days after infection in those that are asymptomatic with an infectiousness profile over time that results in a 6.5-day mean generation time. Based on fits to the early growth-rate of the epidemic in Wuhan[10,11], we make a baseline assumption that R_0=2.4 but examine values between 2.0 and 2.6. We assume that symptomatic individuals are 50% more infectious than asymptomatic individuals. Individual infectiousness is assumed to be variable, described by a gamma distribution with mean 1 and shap parameter → 0.25. On recovery from infection, individuals are assumed to be immune to re-infect in the short term. Evidence from the Flu Watch cohort study suggests that re-infection with the strain of seasonal circulating coronavirus is highly unlikely in the same or following seaso (Andrew Hayward, personal communication).

Infection was assumed to be seeded in each country at an exponentially growing rate (wit time of 5 days) from early January 2020, with the rate of seeding being calibrated epidemics which reproduced the observed cumulative number of deaths in GB or the U March 2020.

Table 2: Summary of NPI interventions considered.

Label	Policy	Description
CI	Case isolation in the home	Symptomatic cases stay at home for household contacts by 75% for th contacts remain unchanged. Ass comply with the policy.
HQ	Voluntary home quarantine	Following identification of a household, all household me days. Household contac quarantine period, conta 75%. Assume 50% of hou

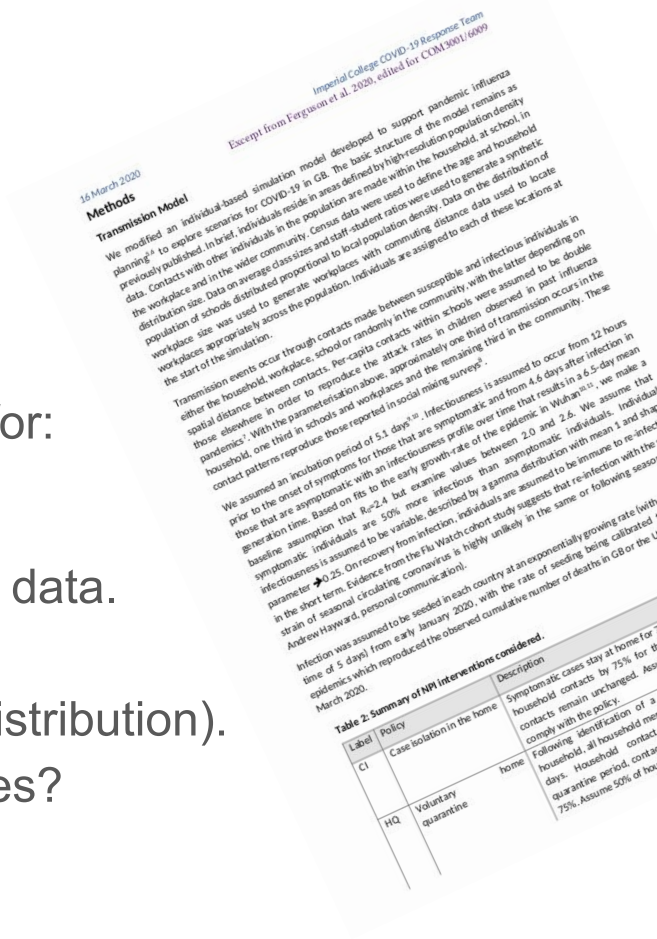
5 - 10 minutes

How to choose parameters? **Activity**

In groups of 2-4:

Skim through this excerpt of the methods section, look for:

- How they've based parameters on literature, etc.
- Validations of parts of their model against empirical data.
- Where guesses have been made.
- Where the variation of agents is modelled (with a distribution).
- Where do they consider a range of parameter values?



Imperial College COVID-19 Response Team
Excerpt from Ferguson et al. 2020, edited for COM3001/6009

16 March 2020

Methods

Transmission Model

We modified an individual-based simulation model developed to support pandemic influenza planning[5,6] to explore scenarios for COVID-19 in GB. The basic structure of the model remains as previously published. In brief, individuals reside in areas defined by high-resolution population density data. Contacts with other individuals in the population are made within the household, at school, in the workplace and in the wider community. Census data were used to define the age and household distribution size. Data on average class sizes and staff-student ratios were used to generate a synthetic population of schools distributed proportional to local population density. Data on the distribution of workplace size was used to generate workplaces with commuting distance data used to locate workplaces appropriately across the population. Individuals are assigned to each of these locations at the start of the simulation.

Transmission events occur through contacts made between susceptible and infectious individuals in either the household, workplace, school or randomly in the community, with the latter depending on spatial distance between contacts. Per-capita contacts within schools were assumed to be double those elsewhere in order to reproduce the attack rates in children observed in past influenza pandemics[7]. With the parameterisation above, approximately one third of transmission occurs in the household, one third in schools and workplaces and the remaining third in the community. These contact patterns reproduce those reported in social mixing surveys[8].

We assumed an incubation period of 5.1 days[9,10]. Infectiousness is assumed to occur from 12 hours prior to the onset of symptoms for those that are symptomatic and from 4.6 days after infection in those that are asymptomatic with an infectiousness profile over time that results in a 6.5-day mean generation time. Based on fits to the early growth-rate of the epidemic in Wuhan[10,11], we make a baseline assumption that R_0=2.4 but examine values between 2.0 and 2.6. We assume that
symptomatic individuals are 50% more infectious than asymptomatic individuals. Individual
infectiousness is assumed to be variable, described by a gamma distribution with mean 1 and shap
parameter → 0.25. On recovery from infection, individuals are assumed to be immune to re-infec
in the short term. Evidence from the Flu Watch cohort study suggests that re-infection with the
strain of seasonal circulating coronavirus is highly unlikely in the same or following seaso
Andrew Hayward, personal communication).

Infection was assumed to be seeded in each country at an exponentially growing rate (with
time of 5 days) from early January 2020, with the rate of seeding being calibrated
epidemics which reproduced the observed cumulative number of deaths in GB or the U
March 2020.

Table 2: Summary of NPI interventions considered.

Label	Policy	Description
CI	Case isolation in the home	Symptomatic cases stay at home for household contacts by 75% for th contacts remain unchanged. Ass comply with the policy.
HQ	Voluntary home quarantine	Following identification of a household, all household me days. Household contac quarantine period, contac 75%. Assume 50% of hou

5 - 10 minutes

How to choose parameters?

Transmission Model

We modified an individual-based simulation model developed to support pandemic influenza planning[5,6] to explore scenarios for COVID-19 in GB. The basic structure of the model remains as previously published. In brief, individuals reside in areas defined by high-resolution population density data. Contacts with other individuals in the population are made within the household, at school, in the workplace and in the wider community. Census data were used to define the age and household distribution size. Data on average class sizes and staff-student ratios were used to generate a synthetic population of schools distributed proportional to local population density. Data on the distribution of workplace size was used to generate workplaces with commuting distance data used to locate workplaces appropriately across the population. Individuals are assigned to each of these locations at the start of the simulation.

How to choose parameters?

- Ideally we base our parameters on empirical data (e.g. the census, experiments, or calculated from other sources).

Transmission Model

We modified an individual-based simulation model developed to support pandemic influenza planning[5,6] to explore scenarios for COVID-19 in GB. The basic structure of the model remains as previously published. In brief, individuals reside in areas defined by high-resolution population density data. Contacts with other individuals in the population are made within the household, at school, in the workplace and in the wider community. Census data were used to define the age and household distribution size. Data on average class sizes and staff-student ratios were used to generate a synthetic population of schools distributed proportional to local population density. Data on the distribution of workplace size was used to generate workplaces with commuting distance data used to locate workplaces appropriately across the population. Individuals are assigned to each of these locations at the start of the simulation.

How to choose parameters?

Transmission events occur through contacts made between susceptible and infectious individuals in either the household, workplace, school or randomly in the community, with the latter depending on spatial distance between contacts. Per-capita contacts within schools were assumed to be double those elsewhere in order to reproduce the attack rates in children observed in past influenza pandemics[7]. With the parameterisation above, approximately one third of transmission occurs in the household, one third in schools and workplaces and the remaining third in the community. These contact patterns reproduce those reported in social mixing surveys[8].

How to choose parameters?

- Validation (by comparing outputs to some observations)

Transmission events occur through contacts made between susceptible and infectious individuals in either the household, workplace, school or randomly in the community, with the latter depending on spatial distance between contacts. Per-capita contacts within schools were assumed to be double those elsewhere in order to reproduce the attack rates in children observed in past influenza pandemics[7]. With the parameterisation above, approximately one third of transmission occurs in the household, one third in schools and workplaces and the remaining third in the community. These contact patterns reproduce those reported in social mixing surveys[8].

How to choose parameters?

We assumed an incubation period of 5.1 days[9,10]. Infectiousness is assumed to occur from 12 hours prior to the onset of symptoms for those that are symptomatic and from 4.6 days after infection in those that are asymptomatic with an infectiousness profile over time that results in a 6.5-day mean generation time. Based on fits to the early growth-rate of the epidemic in Wuhan[10,11], we make a baseline assumption that R_0=2.4 but examine values between 2.0 and 2.6. We assume that symptomatic individuals are 50% more infectious than asymptomatic individuals. Individual infectiousness is assumed to be variable, described by a gamma distribution with mean 1 and shape parameter α=0.25. On recovery from infection, individuals are assumed to be immune to re-infection in the short term. Evidence from the Flu Watch cohort study suggests that re-infection with the same strain of seasonal circulating coronavirus is highly unlikely in the same or following season (Prof Andrew Hayward, personal communication).

How to choose parameters?

- Often individuals vary. Need to model the **distribution** of parameters.

We assumed an incubation period of 5.1 days[9,10]. Infectiousness is assumed to occur from 12 hours prior to the onset of symptoms for those that are symptomatic and from 4.6 days after infection in those that are asymptomatic with an infectiousness profile over time that results in a 6.5-day mean generation time. Based on fits to the early growth-rate of the epidemic in Wuhan[10,11], we make a baseline assumption that R_0=2.4 but examine values between 2.0 and 2.6. We assume that symptomatic individuals are 50% more infectious than asymptomatic individuals. Individual infectiousness is assumed to be variable, described by a gamma distribution with mean 1 and shape parameter α=0.25. On recovery from infection, individuals are assumed to be immune to re-infection in the short term. Evidence from the Flu Watch cohort study suggests that re-infection with the same strain of seasonal circulating coronavirus is highly unlikely in the same or following season (Prof Andrew Hayward, personal communication).

How to choose parameters?

Disease Progression and Healthcare Demand

Analyses of data from China as well as data from those returning on repatriation flights suggest that 40-50% of infections were not identified as cases[12]. This may include asymptomatic infections, mild disease and a level of under-ascertainment. We therefore assume that two-thirds of cases are sufficiently symptomatic to self-isolate (if required by policy) within 1 day of symptom onset, and a mean delay from onset of symptoms to hospitalisation of 5 days. The age-stratified proportion of infections that require hospitalisation and the infection fatality ratio (IFR) were obtained from an analysis of a subset of cases from China[12]. These estimates were corrected for non-uniform attack rates by age and when applied to the GB population result in an IFR of 0.9% with 4.4% of infections hospitalised (Table 1). We assume that 30% of those that are hospitalised will require critical care (invasive mechanical ventilation or ECMO) based on early reports from COVID-19 cases in the UK, China and Italy (Professor Nicholas Hart, personal communication). Based on expert clinical opinion, we assume that 50% of those in critical care will die and an age-dependent proportion of those that do not require critical care die (calculated to match the overall IFR). We calculate bed demand numbers assuming a total duration of stay in hospital of 8 days if critical care is not required and 16 days (with 10 days in ICU) if critical care is required. With 30% of hospitalised cases requiring critical

How to choose parameters?

Label	Policy	Description
CI	Case isolation in the home	Symptomatic cases stay at home for 7 days, reducing non-household contacts by 75% for this period. Household contacts remain unchanged. Assume 70% of household comply with the policy.
HQ	Voluntary home quarantine	Following identification of a symptomatic case in the household, all household members remain at home for 14 days. Household contact rates double during this quarantine period, contacts in the community reduce by 75%. Assume 50% of household comply with the policy.
SDO	Social distancing of those over 70 years of age	Reduce contacts by 50% in workplaces, increase household contacts by 25% and reduce other contacts by 75%. Assume 75% compliance with policy.

How to choose parameters?

- We might not be able to find such data: **It is ok to make a 'guess' about these parameters - but need to document.** E.g. (from Ferguson *et al.*)

Label	Policy	Description
CI	Case isolation in the home	Symptomatic cases stay at home for 7 days, reducing non-household contacts by 75% for this period. Household contacts remain unchanged. Assume 70% of household comply with the policy.
HQ	Voluntary home quarantine	Following identification of a symptomatic case in the household, all household members remain at home for 14 days. Household contact rates double during this quarantine period, contacts in the community reduce by 75%. Assume 50% of household comply with the policy.
SDO	Social distancing of those over 70 years of age	Reduce contacts by 50% in workplaces, increase household contacts by 25% and reduce other contacts by 75%. Assume 75% compliance with policy.

How to choose parameters?

- Ideally we base our parameters on empirical data (e.g. the census, experiments, etc).
- We might not be able to find such data: **It is ok to make a ‘guess’ about these parameters - but need to document.**
- Ideally need to try out different values of these to see how sensitive the model is to them (and report that range in output).
- Some parameters might matter more than others.

Summary:

- Need to be clear about where parameter estimates are from. Make it easy to update in light of new evidence. Quoting: The report’s “conclusion has only been reached in the last few days, with the refinement of estimates of likely ICU demand due to COVID-19 **based on experience in Italy and the UK**.”

Modelling Process

Assumptions,
Rules &
Parameters

Natural System

Modelling Process

Assumptions,
Rules &
Parameters

Simulation

Predictions

Natural System

Modelling Process

Assumptions,
Rules &
Parameters

Simulation

Predictions

Compare to observations
(validation)

Natural System

Modelling Process

Assumptions,
Rules &
Parameters

Simulation

Predictions

Compare to observations
(validation)

Natural System

Modelling Process

Group Project:

- You’ll go through this process for your group project.
- But the validation step, you will be mainly doing just once.
- In research, would be working with a domain expert.
- Model would be refined iteratively. (careful with test data)

Assumptions, Rules & Parameters

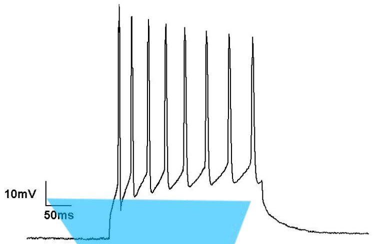
Simulation

Predictions

Compare to observations (validation)

Natural System

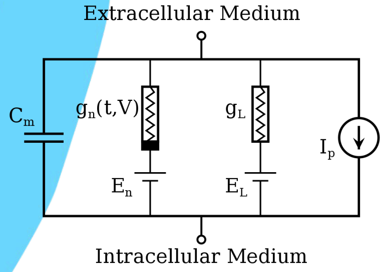
The Modelling Process



Observations
Microelectrode current-clamp recordings

Hypothesis
Voltage-gated ion channels, for different ions, explain action potentials?

Mechanisms
An electrical circuit representing different components.



$$I = C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} + \bar{g}_{\mathrm{K}} n^4 (V_m - V_K) + \bar{g}_{\mathrm{Na}} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l),$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m$$

$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h$$

Description
Differential equations describing electrical circuit.

Prediction
“Run” the model

Compare!

Understanding
They explained the ionic mechanisms underlying the initiation and propagation of action potentials

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

CosMos Process

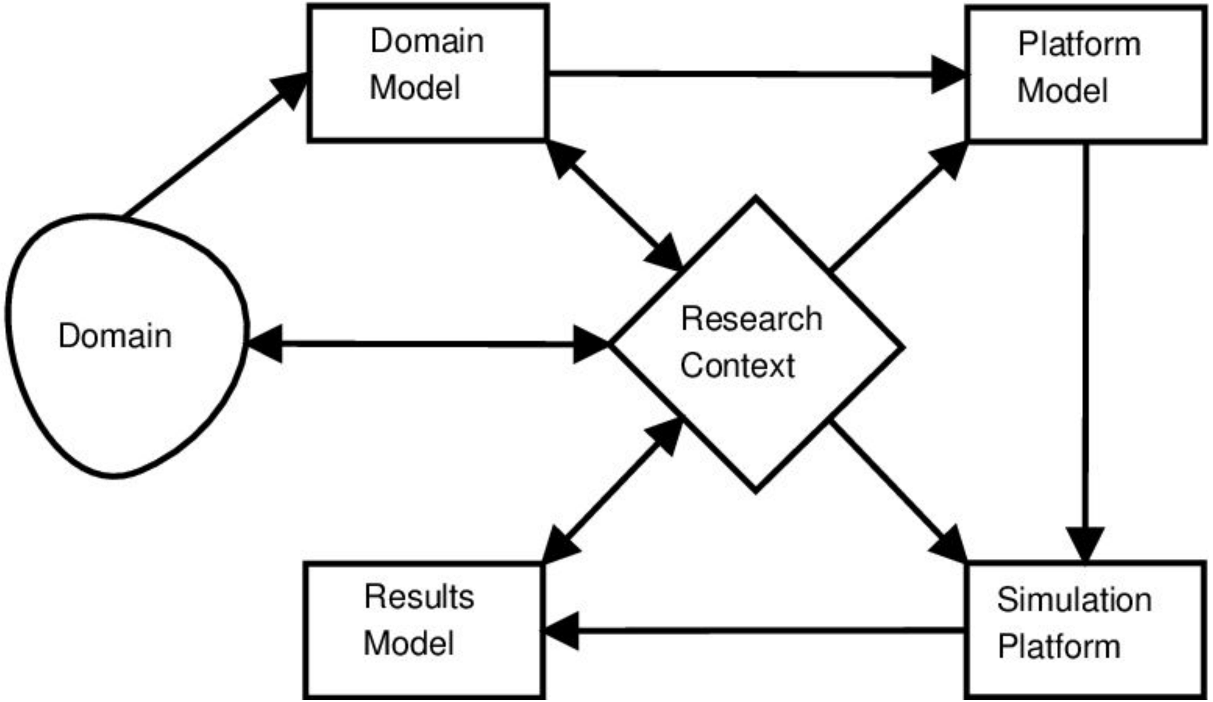


Figure from Garnett, Philip, et al. "Using the CoSMoS process to enhance an executable model of auxin transport canalisation." *CoSMoS* 2010 (2010): 9-32.

CosMos developed by Complex Systems Group (University of York)

CosMos Process

Domain expert makes observations/experiments

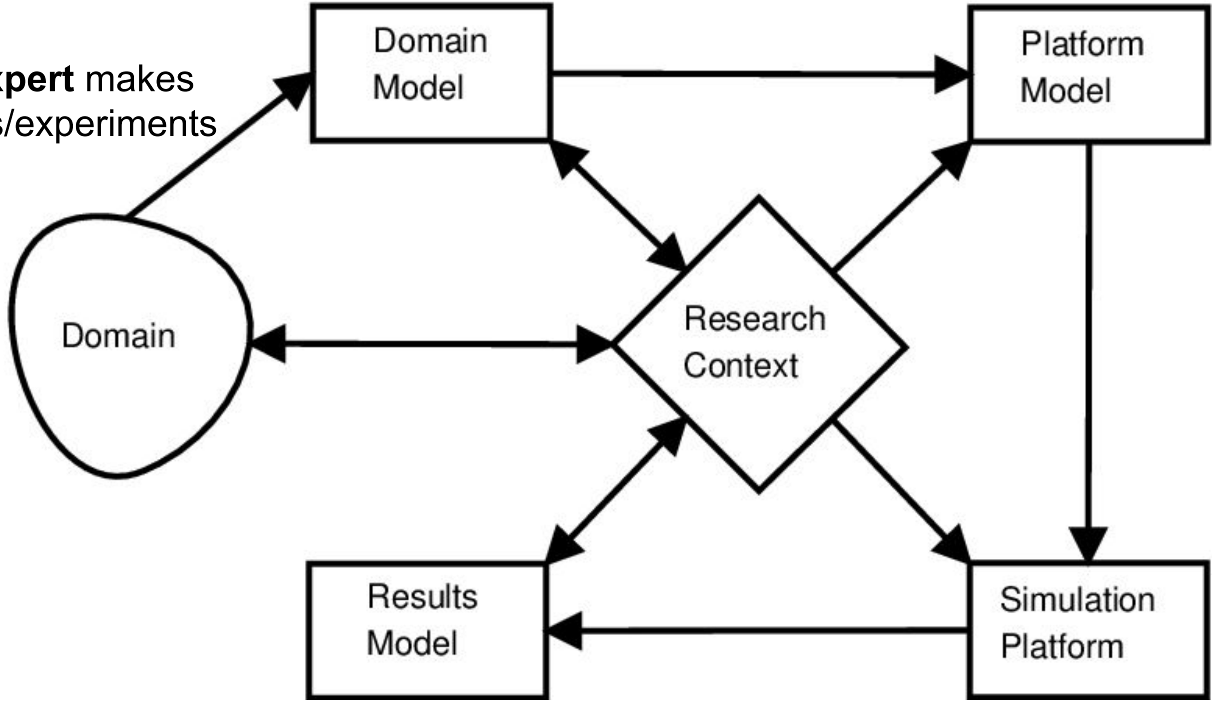


Figure from Garnett, Philip, et al. "Using the CoSMoS process to enhance an executable model of auxin transport canalisation." *CoSMoS* 2010 (2010): 9-32.

CosMos developed by Complex Systems Group (University of York)

CosMos Process

Domain expert makes observations/experiments

Convert conceptual model to a more quantitative model

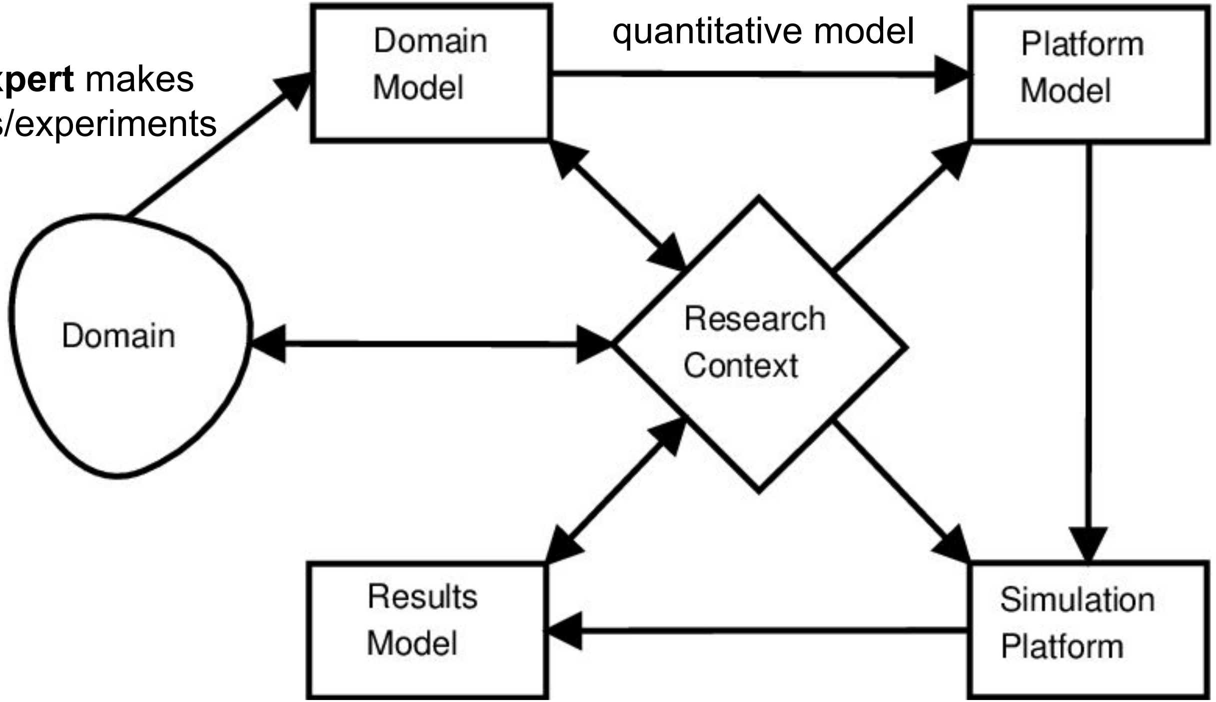


Figure from Garnett, Philip, et al. "Using the CoSMoS process to enhance an executable model of auxin transport canalisation." *CoSMoS* 2010 (2010): 9-32.

CosMos developed by Complex Systems Group (University of York)

CosMos Process

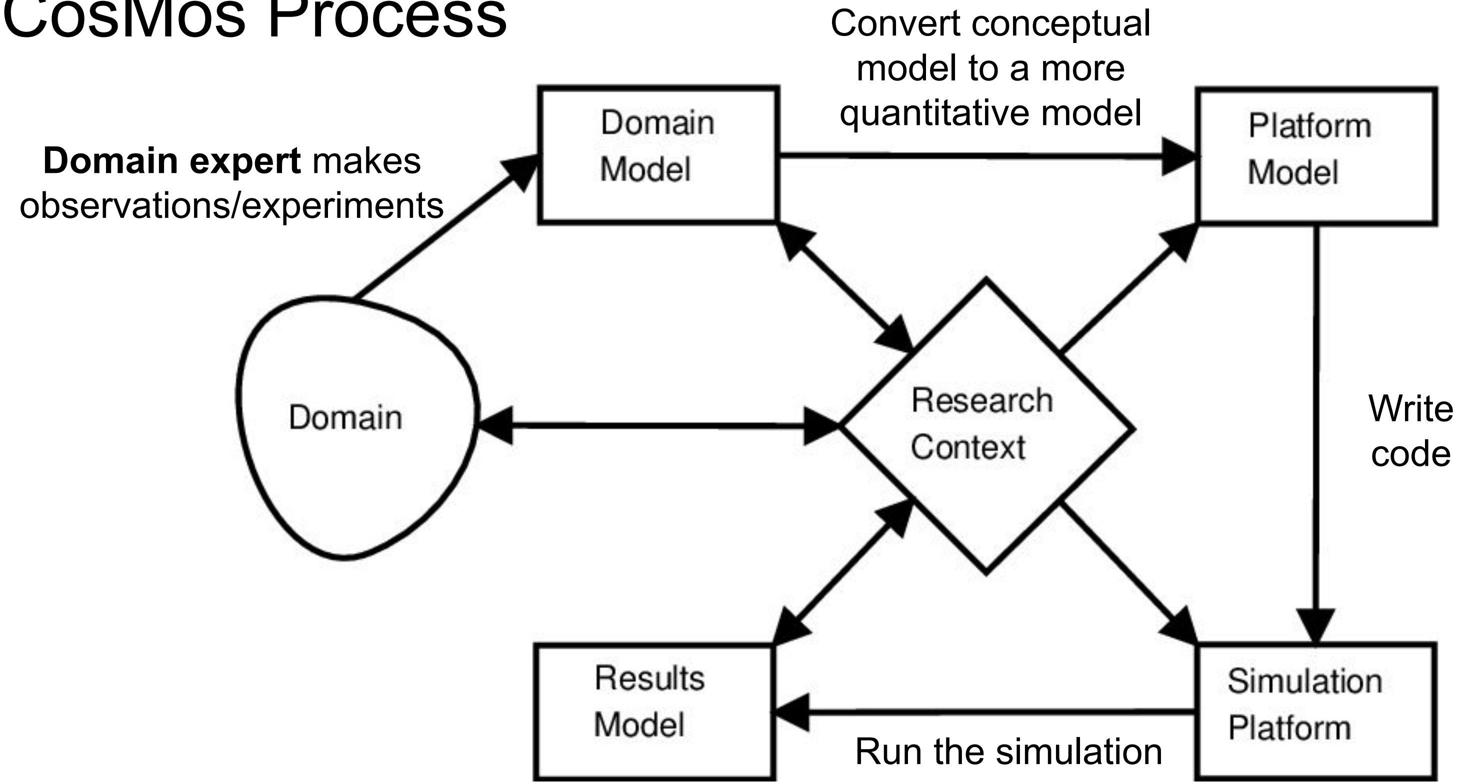


Figure from Garnett, Philip, et al. "Using the CoSMoS process to enhance an executable model of auxin transport canalisation." *CoSMoS* 2010 (2010): 9-32.

CosMos developed by Complex Systems Group (University of York)

CosMos Process

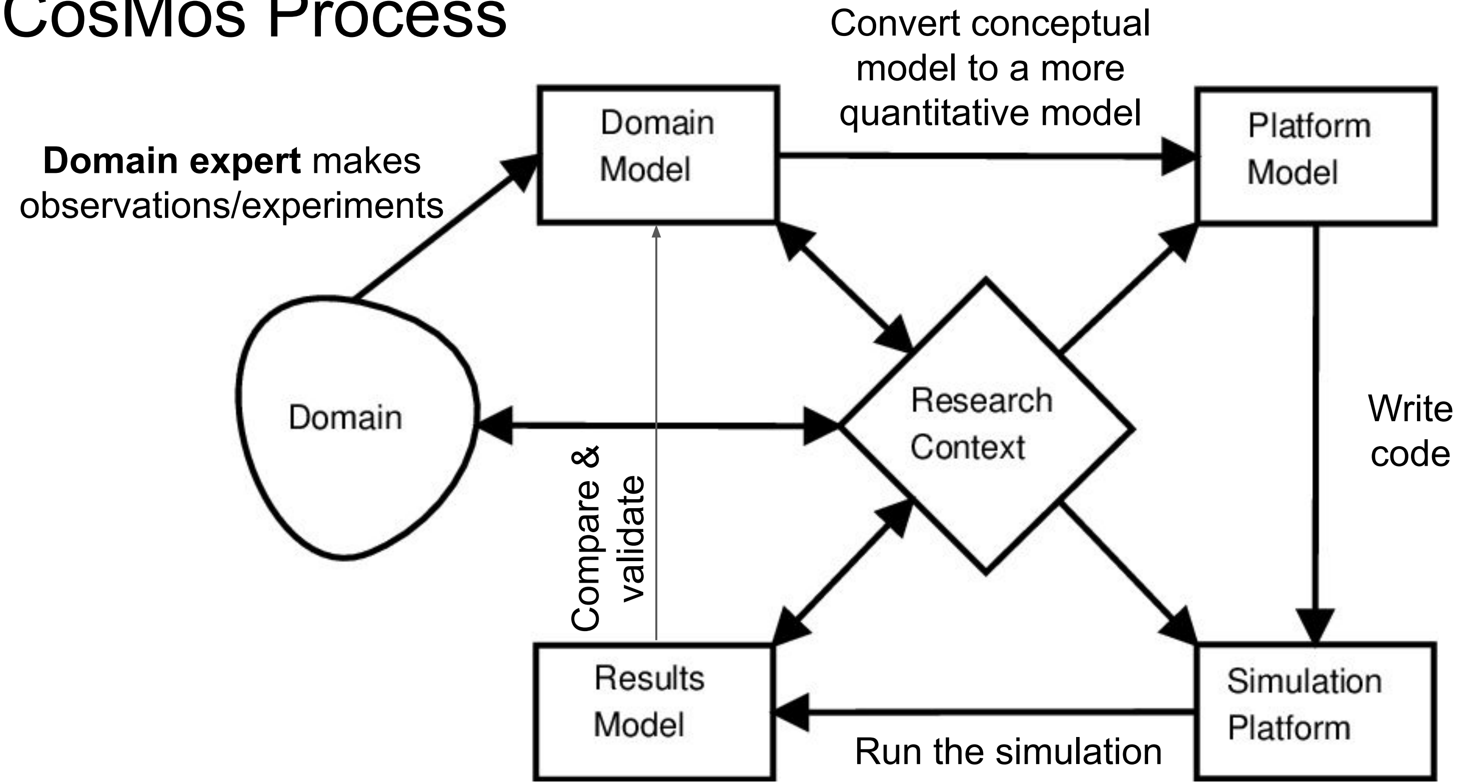


Figure from Garnett, Philip, et al. "Using the CoSMoS process to enhance an executable model of auxin transport canalisation." *CoSMoS* 2010 (2010): 9-32.

CosMos developed by Complex Systems Group (University of York)

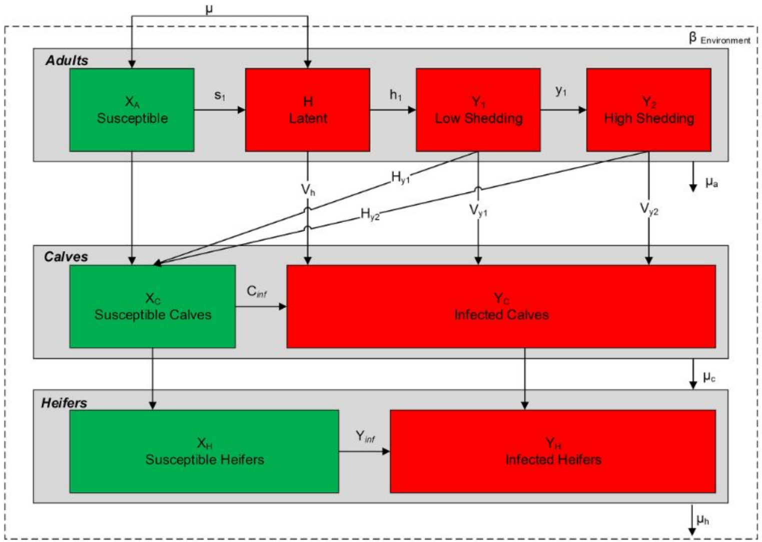
Validation and Sensitivity

We compare the predictions of emergent feature with empirical data.

- We might compare some feature that isn't necessarily the one we want to learn about (in the Ferguson 2020 paper they looked at proportion of contacts in different contexts).
- Types of comparison:
    - Qualitative (e.g. do we see oscillations in predator/prey model?)
    - Quantitative (e.g. do absolute/relative values match real system?)
    - Statistical test (e.g. is there a significant difference between the true data and simulated).
- Can't 'prove' our model is correct…
- Is our model useful or "good enough"?

Validation

Example Al-Mamun et al. (2018) used an ABM for TB in cattle.



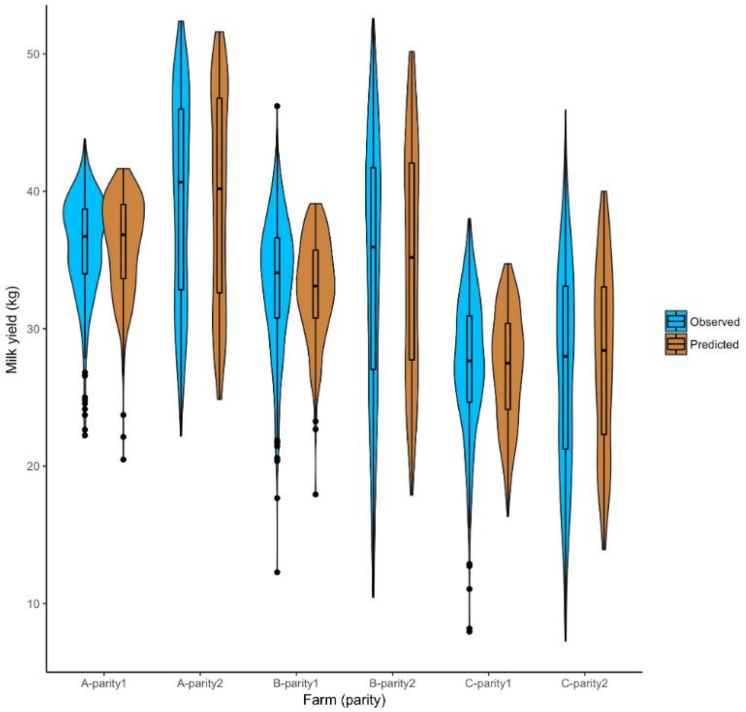
Al-Mamun, Mohammad A., et al. "A data-driven individual-based model of infectious disease in livestock operation: A validation study for paratuberculosis." *Plos one* 13.12 (2018): e0203177.

Validation

Example Al-Mamun et al. (2018) used an ABM for TB in cattle.

- They looked at the **distribution** of individuals & compared to the distribution in the real data.
- Both **quantitative** and **qualitative** validation.
- They didn't compare using a statistical test.

In this case there was careful fitting to data: So generalisation might be a problem. Ideally test on another dataset.

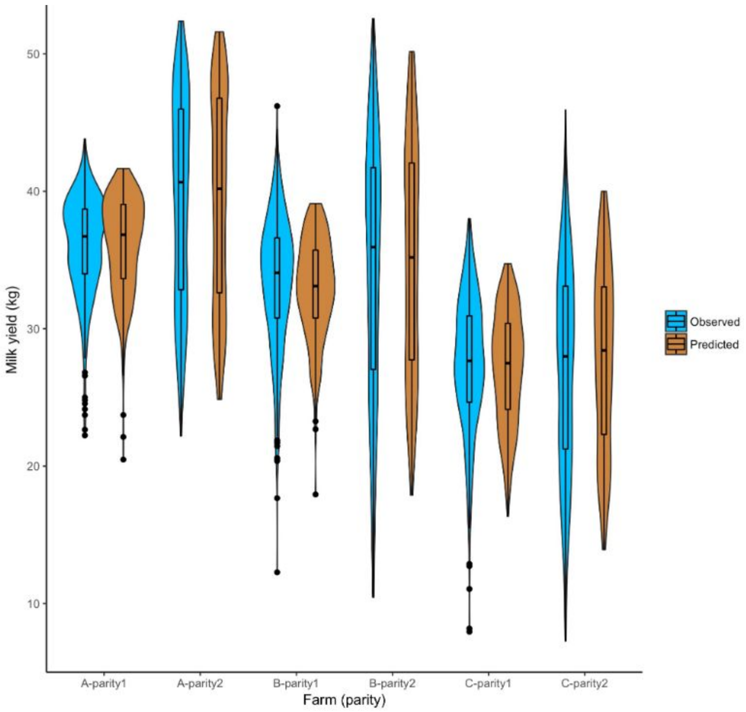


Al-Mamun, Mohammad A., et al. "A data-driven individual-based model of infectious disease in livestock operation: A validation study for paratuberculosis." *Plos one* 13.12 (2018): e0203177.

Validation

Ideally you might be able to find some data or statistic online you can compare your model output to.

Might have to compare just relative change, or even just qualitative structures.



Al-Mamun, Mohammad A., et al. "A data-driven individual-based model of infectious disease in livestock operation: A validation study for paratuberculosis." *Plos one* 13.12 (2018): e0203177.

Quantitative Validation

- If you have a **stochastic model**, you'll need to run it many times to get a distribution.
- You will probably also want to **sample from the priors of the parameters**.

Either sample randomly, pseudorandomly, latin hypercube… etc…

Summary:

- You'll get distributions of the observed features.
- Might want to check a true measurement lies within our 95% CI, etc.

Explained on whiteboard.

Quantitative Validation

- If you have a **stochastic model**, you'll need to run it many times to get a distribution.
- You will probably also want to **sample from the priors of the parameters**.

$$p(\overset{\text{observation}}{y} \mid \overset{\text{model}}{M}) = \int p(y, \theta \mid M)\, d\theta$$

$$= \int p(y \mid \theta, M)\, p(\theta \mid M)\, d\theta$$

$$= E_{p(\theta|M)}\, p(y \mid \theta, M)$$

$$\approx \frac{1}{N} \sum_{i=1}^{N} \underbrace{y_i^{(\theta)}}_{\text{sampled from model with parameter } \theta} \quad \theta \sim p(\theta \mid M)$$

Abductive Validation

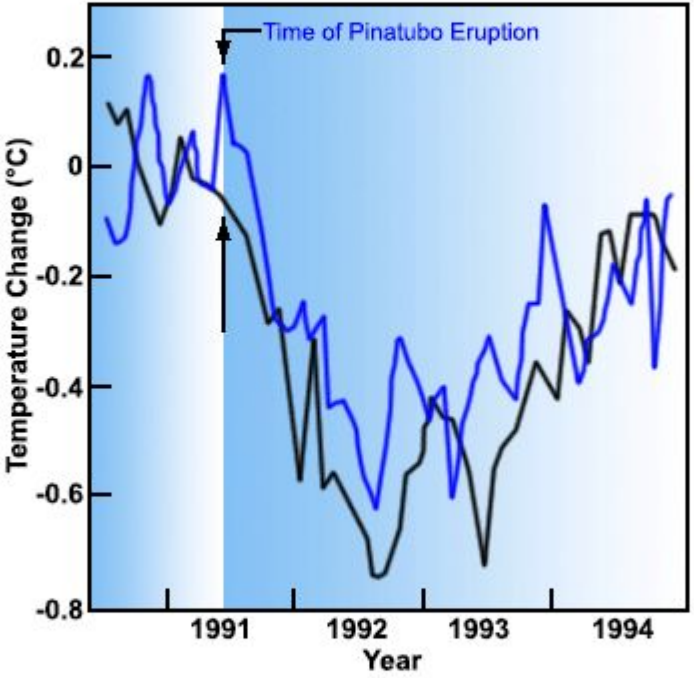
- If you can experiment on the real world system…
- Perturb real world & the model, and see if the perturbation causes similar changes to both.
- Examples (from an EBM)...
- Mount Pinatubo. 20 megatonnes of SO_2 in one day.

Mount Pinatubo in the Philippines on June 15, 1991

Abductive Validation

- If you can experiment on the real world system…
- Perturb real world & the model, and see if the perturbation causes similar changes to both.
- Examples (from an EBM)...
- Mount Pinatubo. 20 megatonnes of SO2 in one day.

Used to check Hansen’s climate model (black), real data (blue).



From https://www.e-education.psu.edu/meteo469/?q=book/export/html/141

Cooling effect appears slightly over estimated, but extended el Nino warmed Earth a little more than predicted.

Sensitivity Analysis

How much effect do our parameters have on the system?

In the simple predator-prey model, **how does the prey growth rate effect prey maximum**?

1) Vary our parameter by a small amount [keeping all other parameters fixed] & see what happens to the emergent feature of interest.

“Local [one-at-a-time] sensitivity analysis”

- Change the parameter by only a small amount (e.g. < 10% of mean)

Might want to look at the change in all parameters (global sensitivity analysis)

Open lab4 colab.

Sensitivity Analysis

Summary

- Approximates the gradient (partial derivative).
- Gives us a sense of the epistemic uncertainty.

Model Calibration

- Adjust parameters to make the feature of interest match the observed data.

$$\frac{\partial f}{\partial x} \approx \frac{\Delta f}{\Delta x}$$

change in feature

change in parameter

Partial derivative.
(there are other parameters).

Reminder! Sources of Error

- Abstraction Error - How much our model is fundamentally wrong.
- Epistemic Error - Uncertainty in our model parameters.
- Aleatoric Error - in some sense this is unknowable errors/noise.

Summary

- Explain why you’ve made the modelling decisions (including parameters).
- Quantify uncertainty.
- Compare to real data.
- Iterate.

Important final warning:

- Just because the model agrees with some data, doesn’t mean it has correctly captured the underlying system in a valid way.

COM3001/6009

Modelling and Simulation of Natural Systems

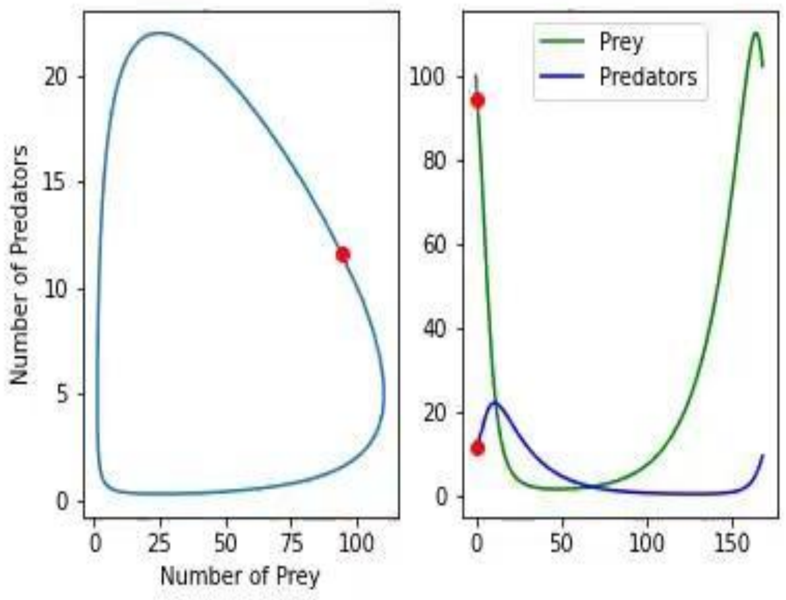
Lecture 5: Introduction to Equation Based Models

Mike Smith* and Luca Manneschi

*m.t.smith@sheffield.ac.uk

Dynamical System / EBM

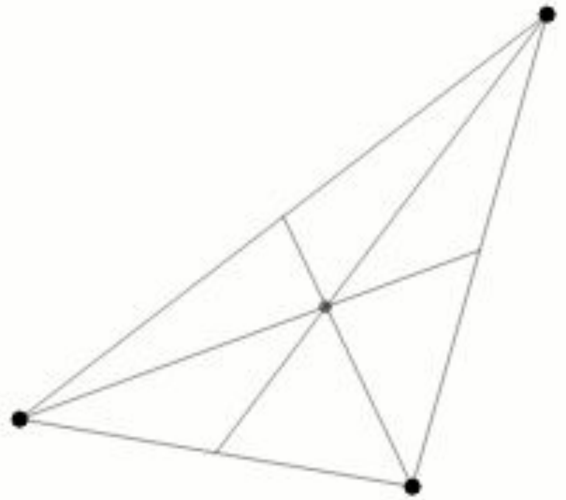
- Set of differential equations
- At a point in time the system has a particular **state**.
- s(t) = [x(t) y(t)]
- Only need the current state to determine next step (markov property).
- The Lotka-Volterra example is deterministic.
- But could also imagine stochastic systems.



Another example…

- Newton's universal model of gravity.
- The 3 body problem - no closed form solution.
- Solve by numerically integrating over the differential equations.
- Force of gravity between a pair of masses:

$$F = G\frac{m_1 m_2}{r^2},$$



https://en.wikipedia.org/wiki/Three-body_problem

https://javalab.org/en/three_body_problem_en/

A 3rd Example from Lecture 1: Hodgkin-Huxley Model

The current across the membrane of a neuron can be modelled during an action potential with a set of differential equations

$$I = C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} + \bar{g}_{\mathrm{K}} n^4 (V_m - V_K) + \bar{g}_{\mathrm{Na}} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l),$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m$$

$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h$$

Hodgkin–Huxley 1963 Nobel Prize in Physiology or Medicine

1. https://en.wikipedia.org/wiki/Electrophysiology A whole-cell current-clamp recording of Substantia Nigra Pars Reticulata neuron. A small amount of negative current is tonically injected to pause tonic firing, and then approximately 200pA of positive current is injected

Discrete and Continuous Time

In **discrete time**, we define a **difference equation**.

$N_{i+1} - N_i \equiv \Delta N_i = f(N_i)$

For example if something is growing exponentially, then $f(N_i) = N_i$.

To illustrate, compute a few steps…

- $N_0 = 1$, $N_1 - N_0 = 1$ so $N_1 = 2$
- $N_1 = 2$, $N_2 - N_1 = 2$ so $N_2 = 4$
- $N_2 = 4$, $N_3 - N_2 = 4$ so $N_3 = 8$… etc

This discrete step event could equate to an annual “mating season” in a population.

Discrete and Continuous Time

If changes are (approximately) continuous (e.g. biomass etc):

In **continuous time**, we define a **differential equation**.

$$\frac{dN(t)}{dt} = \lim_{\Delta t \to 0} \frac{N(t + \Delta t) - N(t)}{\Delta t}$$

Difference equation → Differential equation

$$N_{i+1} - N_i = f(N_i)$$

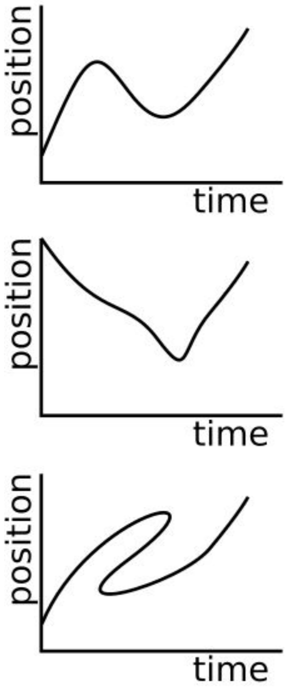
Here we've assumed a time interval…

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

Functions

E.g. position of rabbit over time.

Which one can't be a function?



Functions

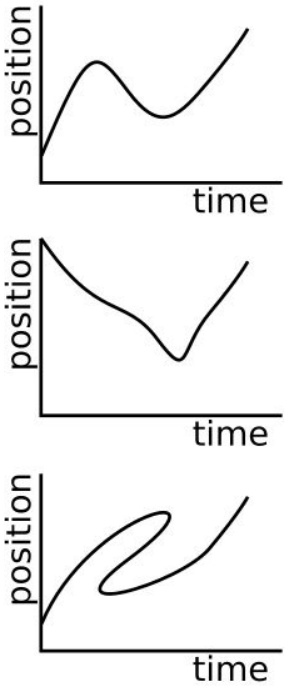
E.g. position of rabbit over time.

Which one can’t be a function?

The last one… can only map to one value.

From wikipedia:

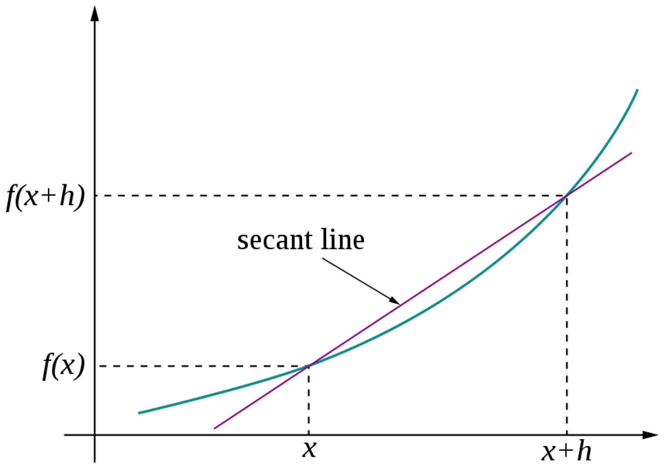
a function from a set X to a set Y assigns to each element of X exactly one element of Y



Derivatives

E.g. position of rabbit over time.

If the rabbit travels from the 2m mark at time 3s, to the 8m mark at time 6s, how fast is it going on average?



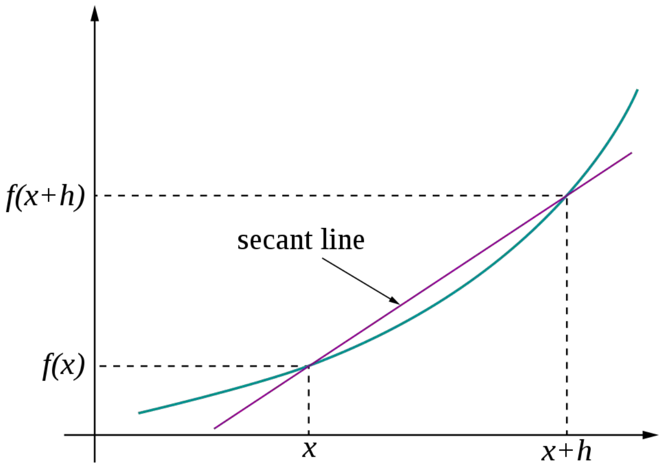
https://en.wikipedia.org/wiki/Derivative

Derivatives

E.g. position of rabbit over time.

If the rabbit travels from the 2m mark at time 3s, to the 8m mark at time 6s, how fast is it going on average?

(8-2)/(6-3) = 6/3 = 2m/s



https://en.wikipedia.org/wiki/Derivative

Derivatives

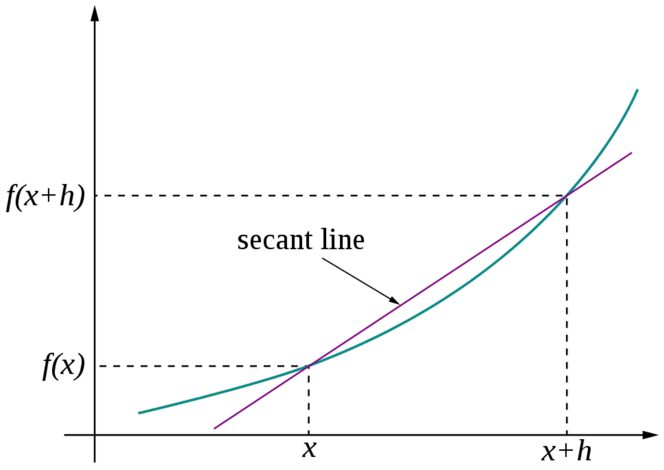
E.g. position of rabbit over time.

If the rabbit travels from the 2m mark at time 3s, to the 8m mark at time 6s, how fast is it going on average?

(8-2)/(6-3) = 6/3 = 2m/s

This is just the average over that time.

If we want a more precise estimate of its speed at time 3s, we would need to measure immediately before and after 3s.



https://en.wikipedia.org/wiki/Derivative

Draw limits on board.

Derivatives

Derive derivative for simple $x = ct^2$ case (on board)

The derivative is a function of time too.

Might be worth before next week revising your differentiation if you're a bit rusty.

This **basic calculus refresher** might be useful (spend ~1 hour doing the exercises) http://web.thu.edu.tw/wenwei/www/Courses/calculus/calculus.pdf

Derivatives

Derive derivative for simple x = ct^2 case (on board)

If you have a vector function of multiple input dimensions, then you can differentiate with respect to each, to give the jacobian.

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial \mathbf{f}}{\partial x_1} & \cdots & \dfrac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla^{\mathrm{T}} f_1 \\ \vdots \\ \nabla^{\mathrm{T}} f_m \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}$$

From wikipedia

Numerical Solutions to differential equations

Assume we’re given some data about disease cases:

Date	Cases
5th	4
7th	15
10th	120
12th	507

Numerical Solutions to differential equations

Assume we're given some data about disease cases:

Date	Cases
5th	4
7th	15
10th	120
12th	507

Which model might fit this?

- $dN/dt = c$
- $dN/dt = cN^2$
- $dN/dt = cN$

We think that the number of cases is growing exponentially. This differential equation is of the last form.

Numerical Solutions to differential equations

Assume we're given some data about disease cases:

Date	Cases
5th	4
7th	15
10th	120
12th	507

Which model might fit this?

- $dN/dt = c$
- $dN/dt = cN_2$
- $dN/dt = cN$

What are the units of c?

We think that the number of cases is growing exponentially. This differential equation is of the last form.

Numerical Solutions to differential equations

- We want to know $N(t_*)$...
- We can do this by starting at a known start point, e.g. $t=0$, with known $N(0)$.
- We compute dN/dt, compute estimate of $N(t+\Delta t)$ [see board], repeat…

This is Euler's method:

- Simply: $N(t+\Delta t) \approx \Delta t \, dN/dt + N(t)$
- The Euler method less accurate than other higher-order techniques such as Runge-Kutta methods, etc.

$$N(t + \text{delta } t) \sim (dN/dt) * \text{delta } t + N(t)$$

Numerical Solutions to differential equations

we have a model, that the number of cases grows with $dN/dt = N$.

Solving analytically...

$\int \frac{1}{N} dN = \int dt$

$\ln N + C = t$

$N = Ae^t$

Let's say that at time zero, $N = 4$, so $A = 4$.

We can compute analytically N at any time.

Often we don't have an analytical solution, so we have to solve it numerically. Euler's method is a simple way of doing this, although there are far more accurate approaches.

We've already explained this in the lecture, so we just write the method:

$N(t + \Delta t) \approx N(t) + \Delta \frac{dN}{dt}$

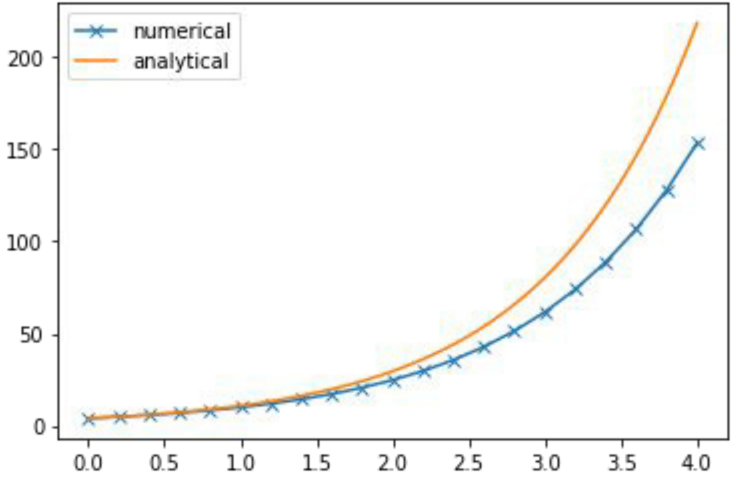
We know $\frac{dN}{dt}$. We actually defined it as equal to N. [just check: $\frac{dN}{dt} = \frac{d}{dt} Ae^t = Ae^t = N$].

Numerical Solutions to differential equations

```
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline

N = 4
timestep = 0.2
record = []
record.append(N)
for it in range(20):
  N = N + timestep * N
  record.append(N)

plt.plot(np.arange(21)*timestep,record,'-x',label='numerical')
plt.plot(np.linspace(0,4,100),4*np.exp(np.linspace(0,4,100)),label='analytical')
plt.legend()
```



Runge Kutta uses gradient at start, end and mid-point etc…

Summary

- Dynamical System: Next state just depends on current state at time t.
- Usually written as a set of coupled ODEs.
- Can be solved numerically, but be careful with time step.

- Revise differentiation for next week's lecture.