Durham University

Simulation

Computational Social Science

Lecture 8

Overview

Models represent key features of a social theory

Simulations are models of a process

- Explicit representation of mechanisms

Two common approaches

- Microsimulation: individuals without interaction
- Agent-based modelling: individuals with interaction

Simulation output combines theory with data

What is a Model?

Durham University

Let’s start with chairs (or close approximations)

Durham University

What makes an object a chair?

Clearly chairs (hopefully)

Why (what do they have in common)?

- Surface to sit on
- Legs
- Back rest

Also decided certain things don't matter

- Style of legs
- Presence of arms
- Construction materials
- Colour
- Typical location (dining room, living room, office)

What makes an object a chair?

Are these all chairs?

- Why or why not?
- Anyone think that one of these is not a chair?

The kneeling chair does not have a back rest

- Maybe back rest not important

What makes an object a chair?

- What about these?
  - Why or why not?
  - Either of these not a chair?
- Starting to use different words
  - ‘chair’ insufficient
  - Subclasses of chairs or not chairs?
- Perhaps inconsistencies?
  - Why is stool different from kneeling chair?

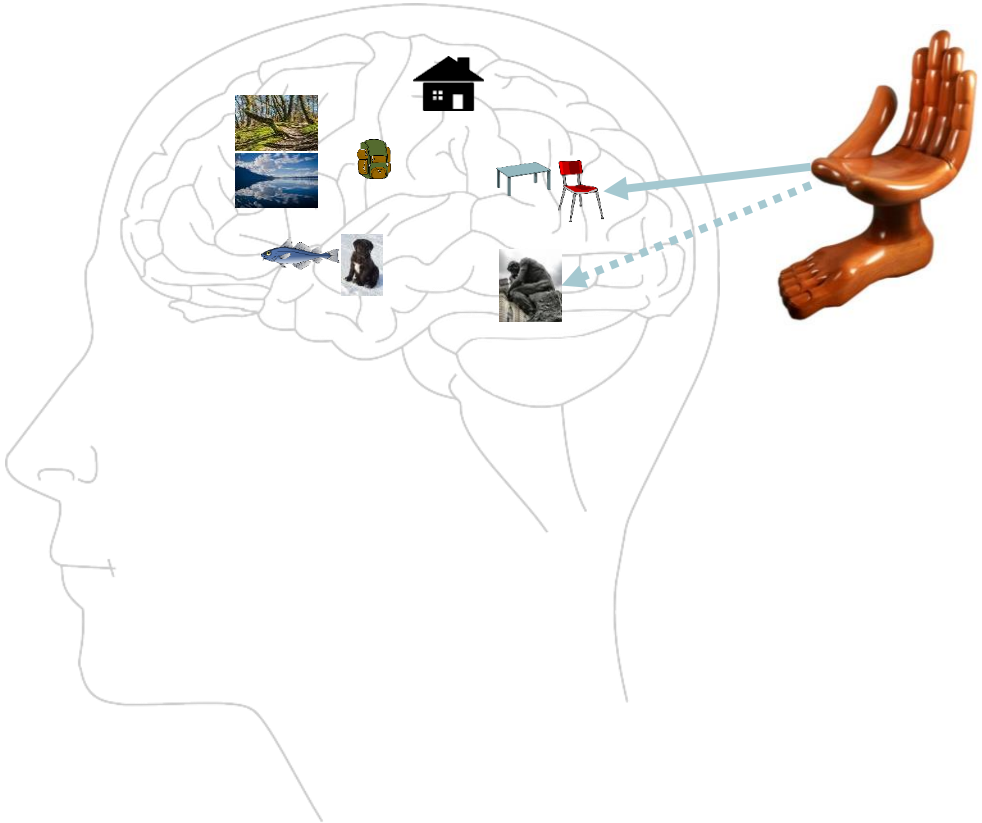
What makes an object a chair?

Table, not a chair

Could be sat on, so why isn't it a chair?

- or at least a stool?

How to decide whether a chair?



Durham University

Why are we talking about chairs?

Definition of a chair is a (text) model of a chair

Setting up a particular way of thinking about what a model is

- Captures the essence of something else

Setting up ideas about what a model can do

- Explicit theory promotes discussion
- Provides rules that classify cases not yet seen

Essence of a chair (approximately)

- Surface to sit on
- Legs
- Intended to be used as a seat

What is a model?

Formal representation…

- the model is constructed using a language, diagram or some other communication medium
- external to the modeller, so allows discussion and interpretation

… of relevant features and relationships…

- simplification that nevertheless captures the important features of the target system

…of some target system

- there is some target object or system that the model is intended to represent

Definition adapted from Greenberger M, MA Crenson & BL Crissey (1976). *Models in the Policy Process*. New York: Russell Sage Foundation.

Relevant features?

Depends on purpose

eg human heart

- Structure
- Surgical practice
- Pumping blood

NOT interchangeable

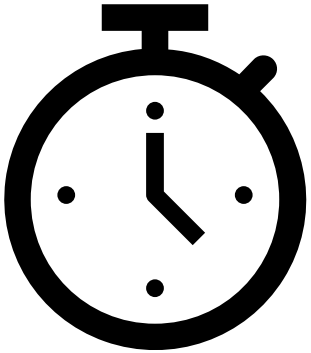
Physical

Behavioural

Functional

What is a simulation?

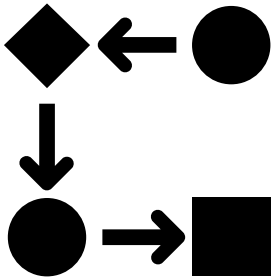
A simulation is a model of a process



What is a process?

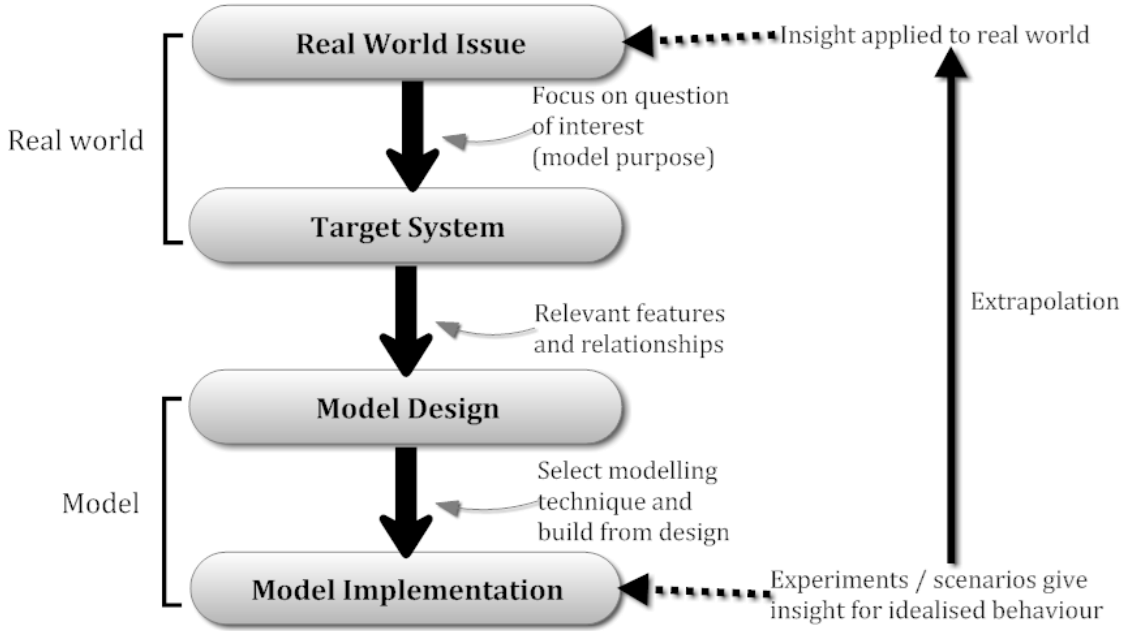
- Occurs through time
- State of world changes

Series of actions / decisions and their consequences



Why is a model useful?

Simplification helps us to understand, and to apply our knowledge



A good model combines theory with data (evidence)

Theory sets the model rules

- Model is internally consistent
- Generates stories, sequences of events

Data to calibrate and test the model

- Model is externally coherent, consistent with what is known
- Stories are therefore plausible, given knowledge

Justified stories can be scenarios that are not part of existing evidence base

INTERNATIONAL JOURNAL OF SOCIAL RESEARCH METHODOLOGY
https://doi.org/10.1080/13645579.2022.2137935

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Negotiating a Future that is not like the Past

Corinna Elsenbroich[a] and Jennifer Badham[b]

[a]MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, UK; [b]Department of Sociology, Durham University, Durham, UK

ABSTRACT

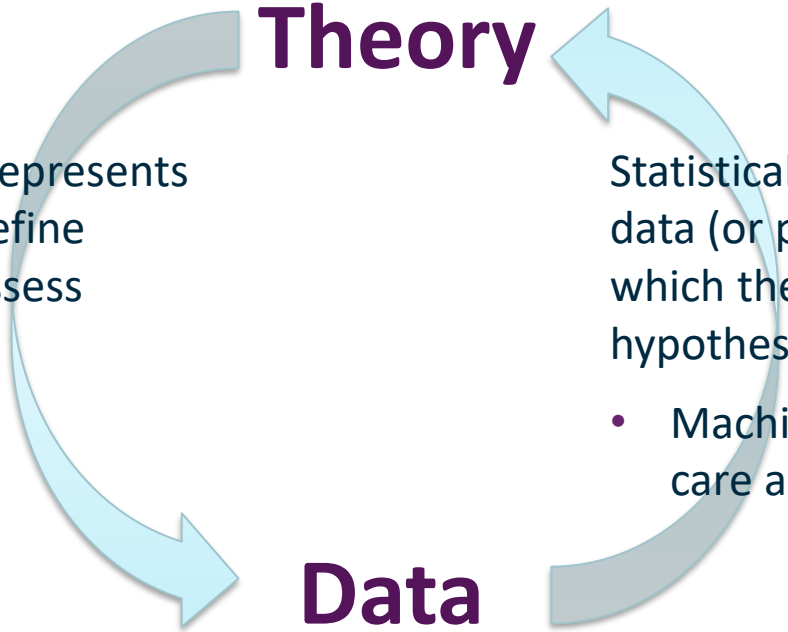
Agent-based models combine data and theory during both development and use of the model. As models have become increasingly data driven, it is easy to start thinking of agent-based modelling as an empirical method, akin to statistical modelling, and reduce the role of theory. We argue that both types of information are important where the past is not a reliable blueprint for the future, which occurs when modelling dynamic complex systems or to explore the implications of change. By balancing theory and data, agent-based modelling is a tool to describe plausible futures, that we call 'justified stories'. We conclude that this balance must be maintained if agent-based models are to serve as a useful decision support tool for policymakers.

KEYWORDS

Prediction; justified stories; modelling change; theory; agent-based modelling

Modelling a process, not modelling a dataset

Meaning of model for purposes of today



Mathematical model represents theory: data used to refine model (calibrate) or assess adequacy of theory

Statistical model represents data (or patterns in the data), which then may generate hypotheses or theory.

- Machine learning doesn't care about theory at all

Representation should match system characteristics

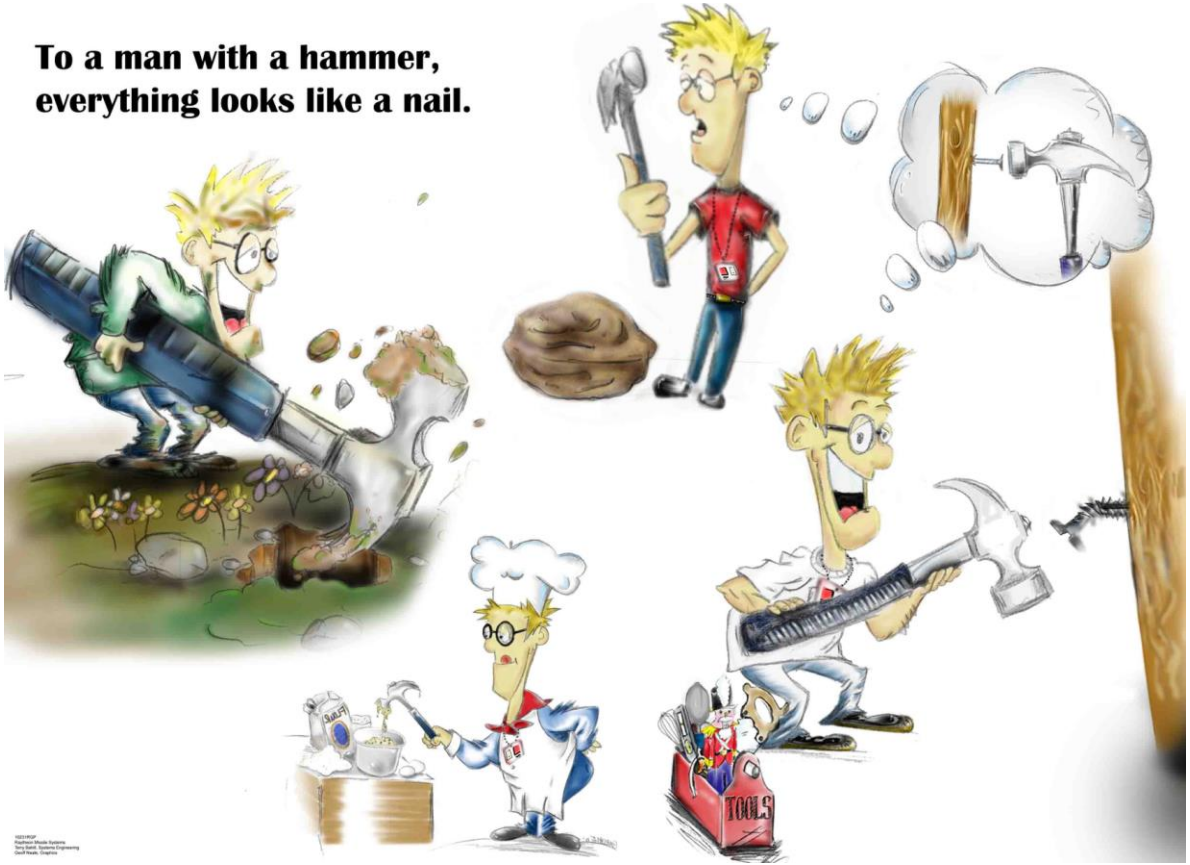


Image source: Terry Bahill
(http://www.sie.arizona.edu/sysengr/)

Durham University

Microsimulation

In the social sciences, a microsimulation model is a computer program that mimics the operation of government programs and demographic processes on individual ("micro") members of a population—people, households, or businesses, for example.

Urban Institute *Microsimulation*
https://www.urban.org/research/data-methods/data-analysis/quantitative-data-analysis/microsimulation

Durham University

Microsimulations replicate detailed individuals

Individuals may be people, households, firms

Large number of attribute

- Often based on Census data, so can include income, age, household structure, education, location, health status

Static microsimulation

- Applies some proposed policy to each individual

Not a simulation in our sense as no process to be modelled

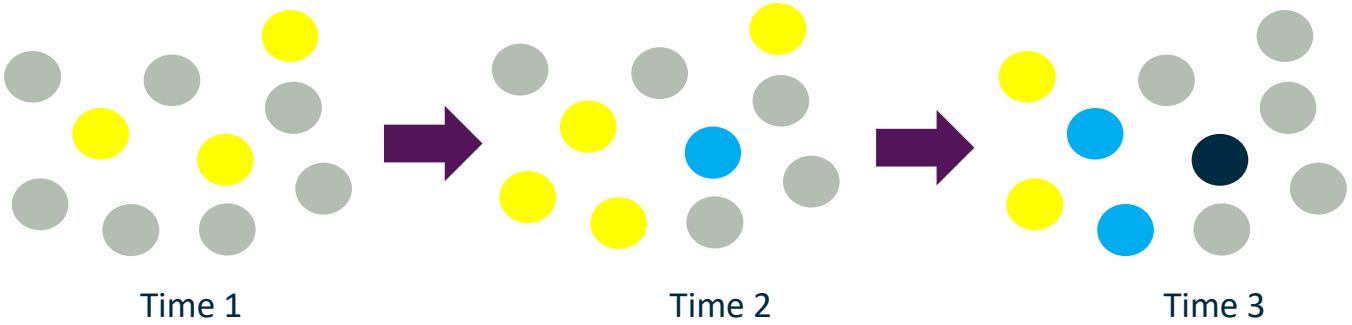
Dynamic microsimulation

- Steps through time and applies probabilities of events occurring based on the detailed profile
- Individuals change states as those probabilities are realised
  - Get married, have children, change jobs, retire, die…

Example of state changes (disease progression)

State transitions for each individual

Hypothetical population over time



Example: cost effectiveness of a health technology
Screening programmes, whether to fund a drug or procedure

Implementation options vary eligible age and how often

QALY is quality adjusted life year

Scenario	Starting age	Frequency	NPV Costs (£m)	NPV QALYs	Cost per QALY
Scenario 1	40	5 years	4.47	1804	2,480
Scenario 2	40	10 years	3.32	1433	2,320
Scenario 3	45	5 years	3.90	1455	2,684
Scenario 4	45	10 years	2.84	1154	2,458
Scenario 5	50	5 years	3.26	1098	2,966
Scenario 6	50	10 years	2.46	912	2,697

Largest benefit

Only slightly more cost-effective for much less benefit

Department of Health (2008), Economic Modelling for Vascular Checks, Table 10
https://webarchive.nationalarchives.gov.uk/ukgwa/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_085917.pdf

Durham University

Projected outcomes rely on detailed assumptions

Of those attending health checks and referred to smoking cessation, only 19% of referred people start the intervention and only 15% of those actually stop smoking

Intervention	Uptake	Compliance	Attribution	RRR of CVD
Smoking	19%	15%	51%	0.36
Anti-Hypertensives	40%	87%	24%	0.24
Exercise	77%	23%	63%	0.14
IFG lifestyle intervention	85%	90%	90%	0.09
Statins prescribing	85%	70%	50%	0.31
Weight management	85%	68%	47%	0.36

Quitting smoking reduces the risk of developing a cardio-vascular disease to 36% of previous level

Department of Health (2008), Economic Modelling for Vascular Checks, Table B1
https://webarchive.nationalarchives.gov.uk/ukgwa/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_085917.pdf

Emphasise distribution impact of policy options (equity)

Microsimulations can calculate the aggregate effect for different subpopulations by summing the effect on individuals within those subpopulation

Example: Simulate a property tax used to fund tax credits for low earners

Decile groups	Static			Labour Supply model			Equilibrium model		
	Couple	Single		Couple	Single		Couple	Single	
		Men	Women		Men	Women		Men	Women
1	7.47	10.49	8.60	8.66	11.43	10.56	6.00	7.97	6.99
2	3.77	4.69	2.70	4.51	4.94	2.78	2.42	2.46	0.67
3	2.10	1.88	2.11	2.58	1.93	1.95	0.89	0.17	0.39
4	1.24	0.82	0.40	1.53	0.81	0.33	0.12	-0.67	-0.72
5	1.00	0.50	0.61	1.22	0.49	0.54	-0.06	-0.61	-0.39
6	0.48	-0.09	-0.34	0.63	-0.10	-0.37	-0.51	-0.86	-1.18
7	0.17	-0.36	-0.54	0.26	-0.37	-0.58	-0.73	-1.12	-1.27
8	-0.01	-0.36	-0.56	0.03	-0.36	-0.53	-0.90	-1.01	-1.12
9	-0.51	-0.77	-0.88	-0.49	-0.76	-0.82	-1.31	-1.31	-1.34
10	-1.16	-1.23	-1.49	-1.16	-1.19	-1.41	-1.86	-1.66	-1.79

Sizeable increase in disposable income

Little to no change in disposable income

Minor decrease in disposable income

Moscarola et al (2020). Shifting taxes from labour to property. A simulation under labour market equilibrium. doi: 10.1016/j.jpolmod.2019.07.002

Durham University

Microsimulations are very detailed

Microsimulations are sensitive to model assumptions

- If assumptions are not grounded in good knowledge then the model will have poor validity
- Assumptions include the initial states: need detailed knowledge of current state of the system being simulated
- Effect of assumptions can be estimated with sensitivity analysis
  - Perturb the values and simulate again

Even good models will be inaccurate in the long-term

- Existing knowledge may not capture all relevant factors
- Systems and conditions may change so the model transitions are no longer appropriate

Each simulated individual has detailed state information

- Demographic
- Health status
- Financial information

Each transition probability incorporates relevant state information

- Is the change in risk when a woman stops smoking at age 40 the same as for a man at age 80?

Microsimulations do not model complexity

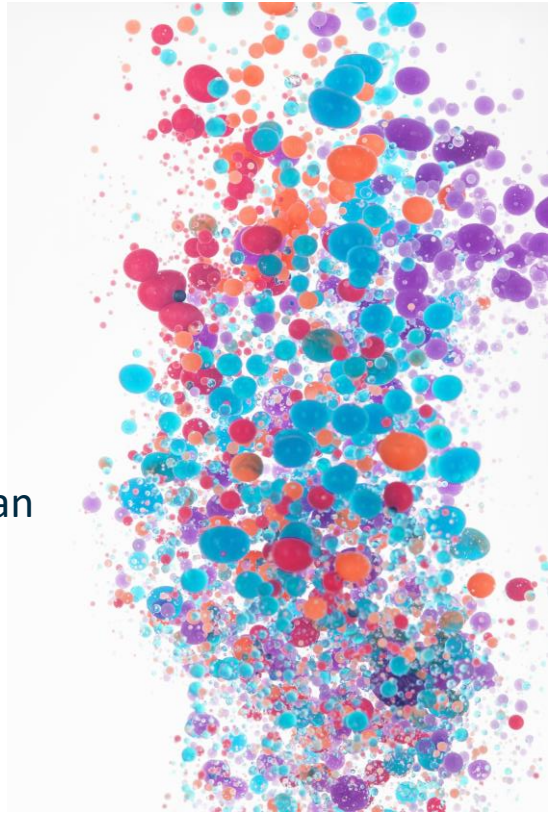
Microsimulation treats cases as independent units

- Cases do not interact with one another in the system
- Cases do not influence the system, but can respond to the system

Projected outcomes are simply aggregated individual probabilistic paths

- More like Weaver's systems of disorganised complexity than organised complexity

They are used to project policy impacts, but… do NOT help to understand social phenomena



Agent-based modelling

Building

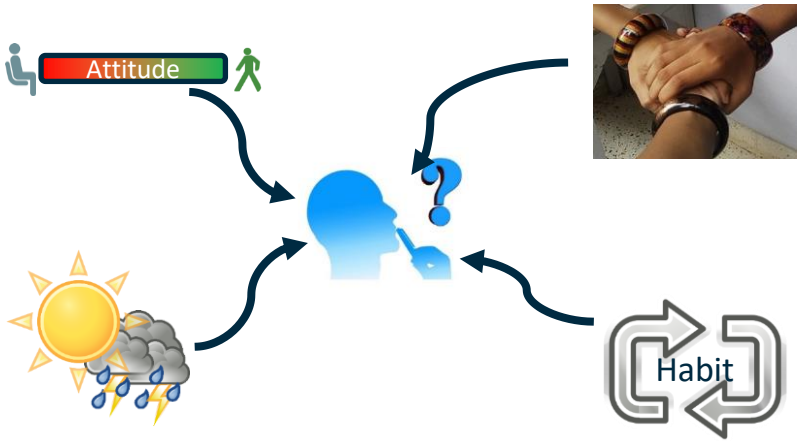
[ABM] … is a computational method that enables a researcher to create, analyse, and experiment with models composed of agents that interact within an environment

Gilbert N (2008). *Agent-Based Models.* Sage Publications.

Personal and environmental factors influence a decision

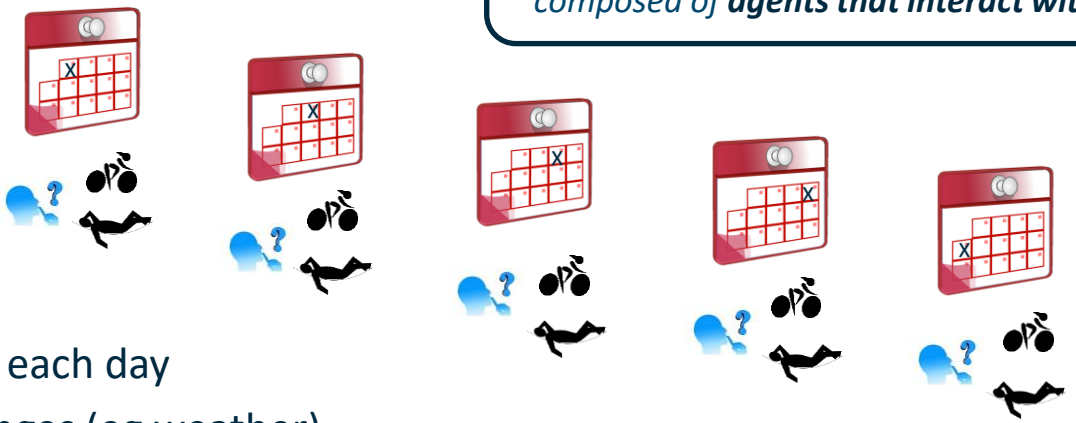
OR

[ABM] … *is a computational method that enables a researcher to create, analyse, and experiment with models composed of* ***agents that interact within an environment***



Process of multiple decisions

[ABM] … *is a computational method that enables a researcher to create, analyse, and experiment with models composed of* ***agents that interact within an environment***



New decision each day

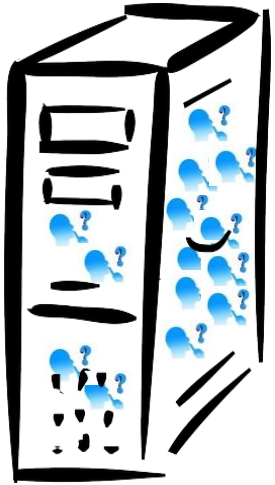
Situation changes (eg weather)

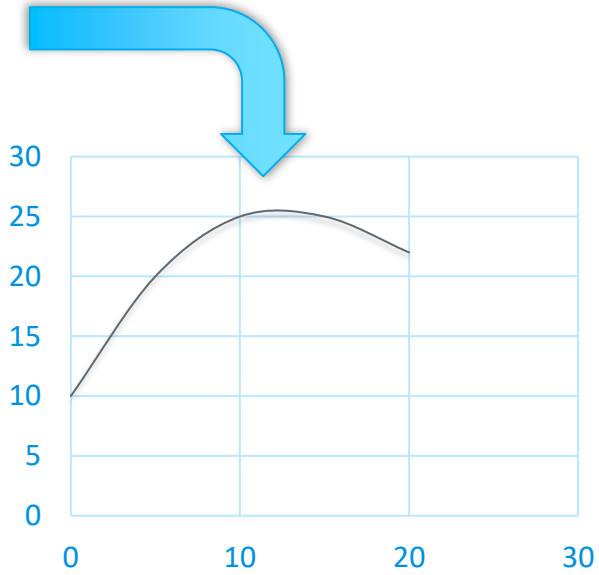
Previous day's decision, and the effect of that decision, is part of the new information

Simulated decisions to see what happens if …

[ABM] … *is a computational method that enables a researcher* ***to create, analyse, and experiment with models*** *composed of agents that interact within an environment*

what if…?





Durham University

ABM as model of process in a complex system

Agent-centric thinking

- I, the agent, have certain characteristics and beliefs of my own as well as information about the world around me, and therefore will decide on some action

Process perspective

- Series of actions and consequences

What features are therefore relevant?

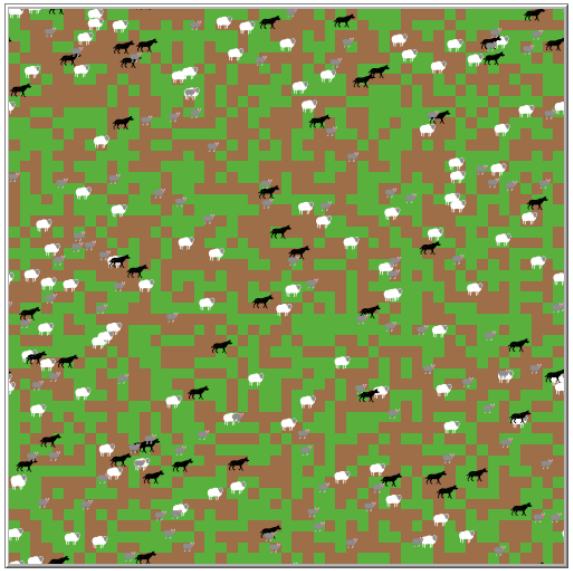
- Personal characteristics that affect agent's choice
- Physical and social environment factors that affect agent's choice
- Rule(s) to combine features and select action
- Impact of that action on self and environment

ABMs represent the mechanisms of a social process

Modelled mechanism

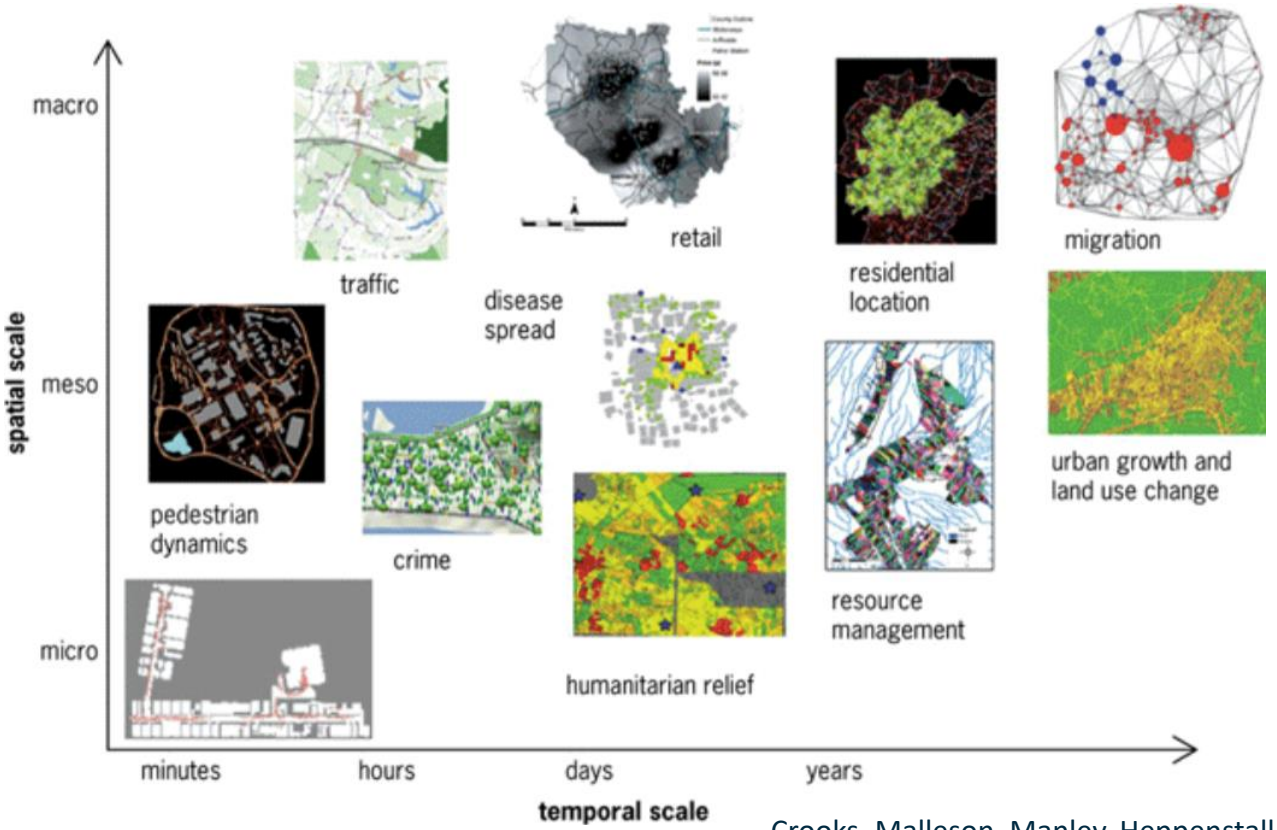
- Environment has patches with or without grass
- Sheep move randomly and eat grass on same patch
- Wolves move randomly and eat sheep on same patch
- Sheep and wolves die if they do not eat regularly
- Sheep and wolves spawn offspring when well fed
- Grass regrows over time

Population cycles are not directly coded but emerge from conditions of food surplus and scarcity



NetLogo model library
Wolf-Sheep predator prey

Models of all sorts of social processes



Crooks, Malleson, Manley, Heppenstall (2019). Agent-Based Modelling & Geographical Information Systems, Fig 7.1
doi:10.4135/9781529793543

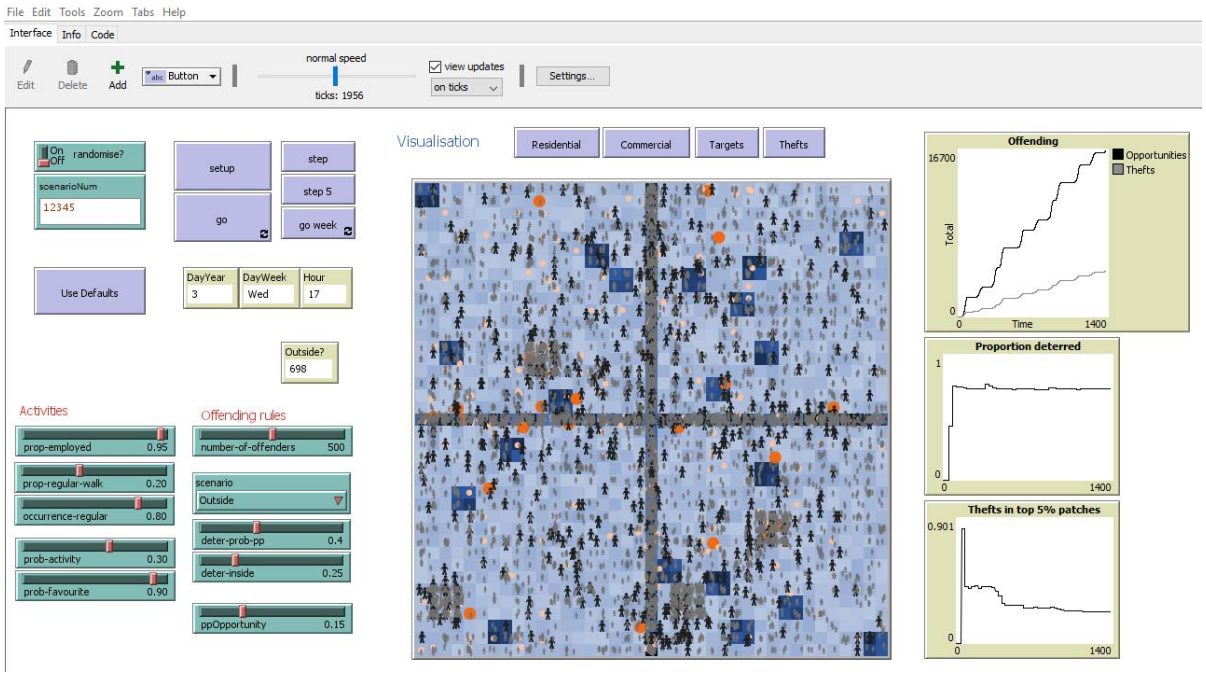
Durham University

Elements of an ABM

What's in the model?

- Agents
- Environment
- Behaviour
- Interactions
- Time

Example model: Guardianship



Groff E, Badham J. 'Examining guardianship against theft' in Gerritsen C, Elffers H (ed) *Agent-based Modelling for Criminological Theory Testing.*

Agents

Concept

The individuals being simulated

- Doing the process

Typically people (in social science)

- Also households or firms

Could be

- Organisations: government, NGOs, countries
- Concepts: eg social practices
- Animals, cells, vehicles

Guardianship model

People

Attributes include:

- Potential offender (or not)
- Work status
- Preferred activities

Environment

Concept

Typically a two-dimensional space

- Can be 3D if altitude is relevant
- Made up of patches or cells in a grid

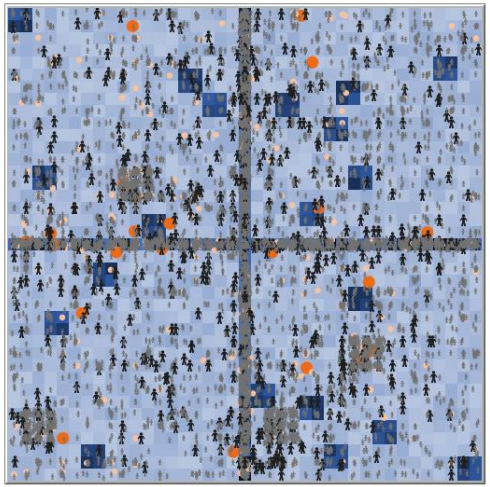
Patches have attributes

- Location: provide distance, direction
- Resources: such as grass

Guardianship model

Grid with residential and business areas

Agents have home and workplace



Time

Concept

Time passes: simulation, not just a model

Discrete timesteps or ticks

- Each tick represents a unit of real world time
- Agents make decisions each tick

Key consideration is duration represented by tick

- Relevant time scale for decisions
- Epidemic may be 1 day
- Housing market may be 3 months

Simulation runs for many ticks

Guardianship model

Time step is two minutes

Schedule of activities

- Commute at similar times each day
- Selection of short activities in the evening
- Selection of long activities on the weekend

Activity examples

25% of commuters use public transport

Short activities: 30% weekday evening, weekend morning

Distance 8-15
32-60 min (return)

Long activities: 30% weekend afternoon

0.5-3.5 hr

Durham University

Behaviour

Concept

Rules that determine agent action

- Each agent makes independent decision (autonomy)
- Action influenced by own situation and environment
- Action impacts on local environment

Guardianship model

Crime requires convergence of

- Motivated offender
  - Agent attribute
- Suitable target
  - Opportunistic, based on number of people on patch
- Absence of guardian

Interactions

Concept

Agent behaviour influenced by situation

Indirect interaction

- Other agents affect the environment
  - Example: resources used so not available

Direct interaction

- Agent decision depends involves others
  - Example: buyer needs a seller
- Agent decision influenced by others
  - Example: social norm to wear masks

Guardianship model

For offenders

- Theft only possible if others present
- If too many others, deterred

Future work

- Agents become familiar with places and each other
- Familiarity leads to confidence to steal or to intervene

Opportunity leads to either theft or deterrence

Offender perspective of guardianship:
deterred by sufficient observable guardians

Opportunity: base + amount
per person present

YES

$$P(deterrence) = 1 - (1 - d)^G$$

d is deterrence per guardian
G is number of guardians

NO

Scenario	Value of G
None	0
Outside	Outside people
Available 0.1	Outside people + 0.1 x Inside
Available 0.25	Outside people + 0.25 x Inside

Durham University

Traps for the novice ABM builder

Arbitrary rather than abstract

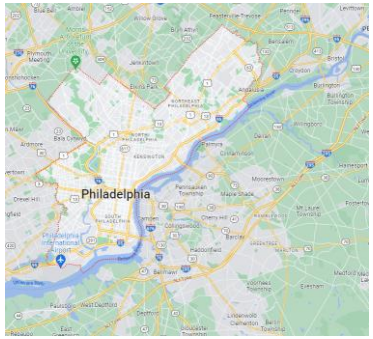
- The simulation rules must be recognisable simplifications of the mechanisms in the real world

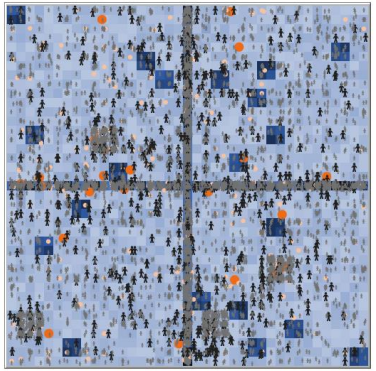
Excessively complicated design

- Don’t include attributes if they don’t change behaviour
- ABMs with too many mechanisms are difficult to interpret as cannot identify the mechanism that generated the result

Lack of coding skills

- ABMs require specialist skills to produce
- Experience in computer programming is required





```
to setup-patches
  ; default values for patches
  ask patches
  [ set residential? false
    set commercial? false
    set station? false
    set park? false
    set numVisitsOffenders 0
    set numVisitsGuardians 0
  ]
  ; transport station and/or park on the block?
  ask n-of round (count patches / 10) patches [ set station? true ]
  ask n-of (count patches / 10) patches [ set park? true ]
  ; main cross streets are commercial and residential
  ask patches with [pxcor = 0 or pycor = 0]
  [ set commercial? true
    set residential? true
  ]
  ; commercial patches, some 3x3 patches - places where they work, shop etc
  ask n-of 4 patches
  [ set commercial? true
    ask neighbors [ set commercial? true ]
  ]
  ask patches with [commercial?]  ; high commercial density 450 to 550 (arbitrary
  [ set commercialDensity 500 + random 50 * one-of [-1 1]
  ]
  ; residential patches have popDensity (population density)
  ask patches with [not commercial?] [ set residential? true ]
  ask patches with [residential?] [ set popDensity 100 + random 20 * one-of [-1 1]
  let high-residential patch-set n-of 20 patches with [residential?]
```

Agent-based modelling

Using

Durham University

Process of ABM in research is messy

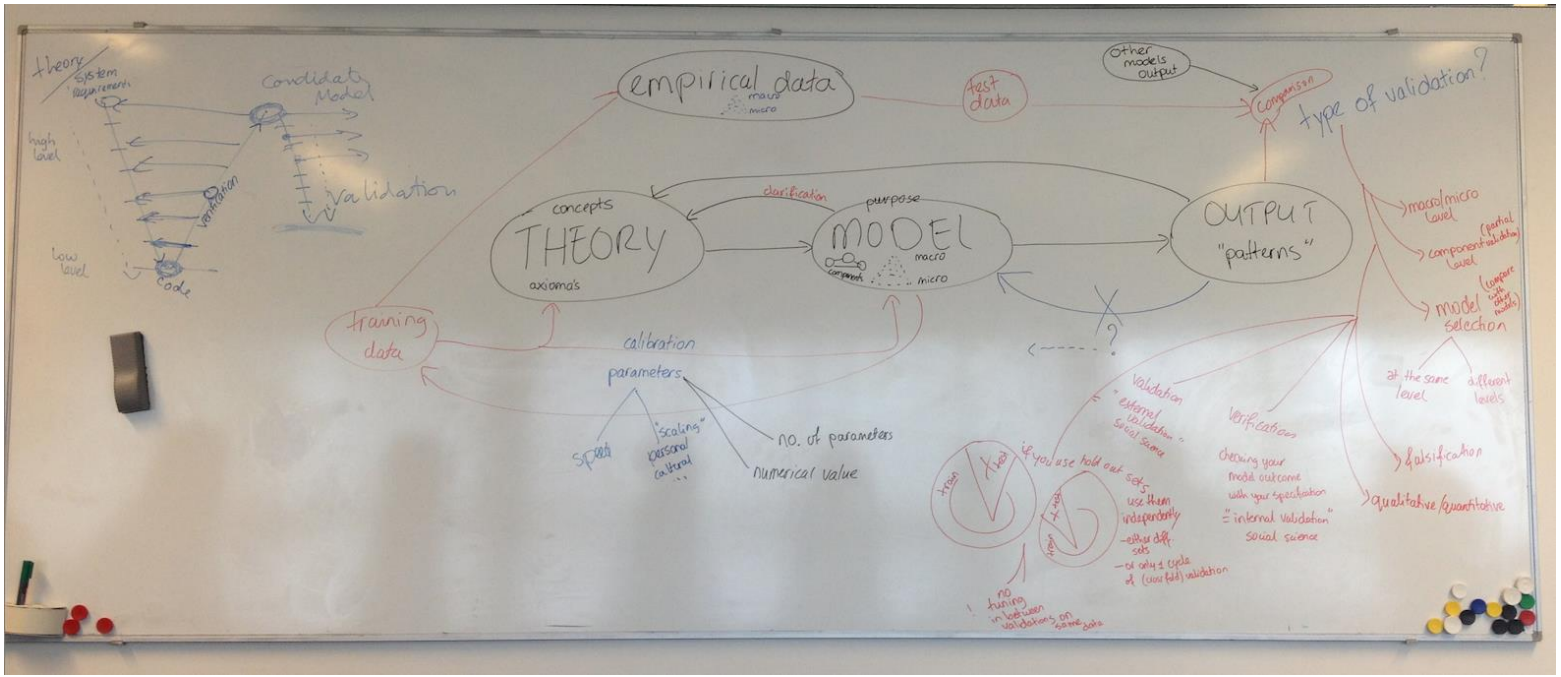


Photo courtesy of Dr Corinna Elsenbroich

Parameterisation and calibration

Scenario defined by a particular set of parameter values

Parameterisation: select values for parameters

- Some parameters may have target system equivalents (eg disease state duration), so value taken from real world
- Parameter sweeping explores model behaviour

Calibration: select parameter values so as to generate most realistic model output

- Matching patterns, not necessarily some measure
- Part of validity assessment

Verification and validation

Verification: Is the model ‘right’?

- Is it doing what we say/think it is doing?
- Syntax errors are easy to find, throw error messages
- Logic errors are difficult, particular where complexity
  - Critical to build and test incrementally

Validation: Is it the right model?

- Replication: model output matches historical data
- Prediction: model output matches target system data collected later
- Structural: model reproduces target system behaviour using the same mechanisms as occur in the target system

Scenarios are experiments with the ABM

Experiments for various purpose

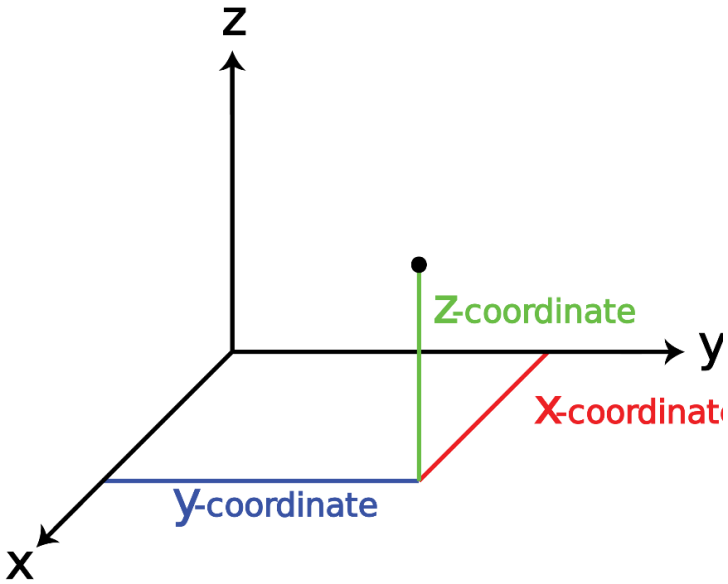
- Research question
  - what relationship is hypothesised?
  - implications of the implemented theory (process)?
- Understand the limitations of the model
  - sensitivity and uncertainty analysis
- Effect of different scenarios

Common: effect of inputs on output(s)

- Input is parameter value (number) or category (scenario)
- Output is pattern or value

Some parameters are control (fixed) variables

Curse of dimensionality



Each parameter

- multiple settings
  - each category
  - numeric value ranges
- multiplies the number of simulations
  - eg If 3 inputs have 4 possible values each, 4x4x4 = 64 combinations
- gets out of control VERY easily

Parameter space

- term used to denote the combination of all input variable values
- one dimension for each variable

How many simulations?

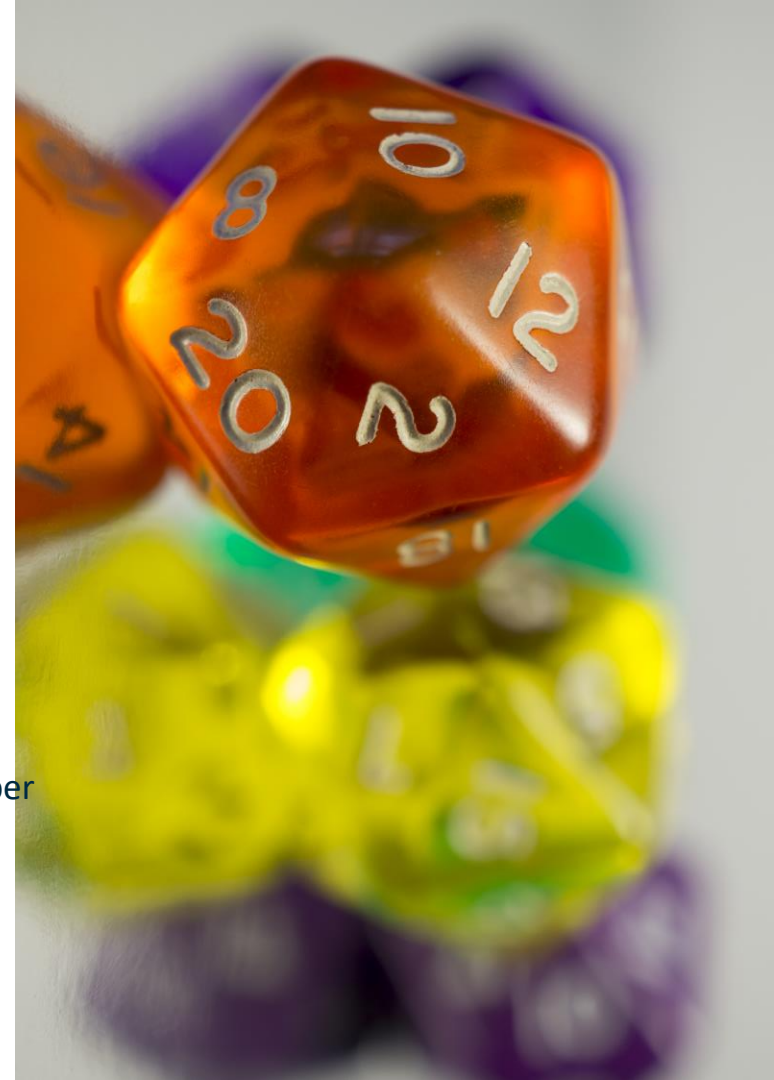
Most ABMs include random elements

- Starting configuration
- Order in which agents make their decisions
  - and affect the environment of other agents
- Consequence of agent-agent interaction
  - such as whether information is transmitted
- Movement direction and/or distance

This means need multiple simulations for each scenario

No good answer for how many

- Statistical significance is meaningless, arbitrarily increase number
- May do preliminary experiments to assess variability



Durham University

ABMs used to test or generate theories

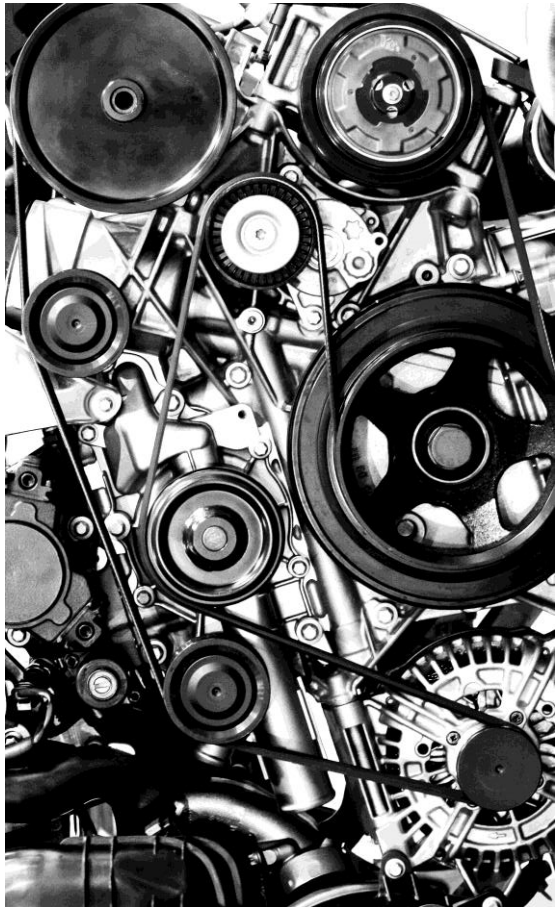
Test whether the purported mechanisms are able to generate the social phenomenon of interest

- Not proof that the theory is correct, just that it is a plausible explanation

Rule-based simulations allow us to see mechanisms and their effects in action and by changing parameters allow us to explore how they work in different ways

- Exploring ideas is a function of ABMs but such exploration should be systematic and theoretically informed

Unexpected outcomes may stimulate further hypotheses



Scenarios

Concept

These are the experiments: what if…

Policy scenarios

- Effect of different implementations

Theory scenarios

- Under what parameters is the mechanism a plausible generator of phenomenon?

Sensitivity scenarios

- Effect of unknowns in data or theory

Guardianship model

Theory scenarios

Effect of different levels of intervention

- Which agents are considered for deterrence

Future work

- Effect of familiarity

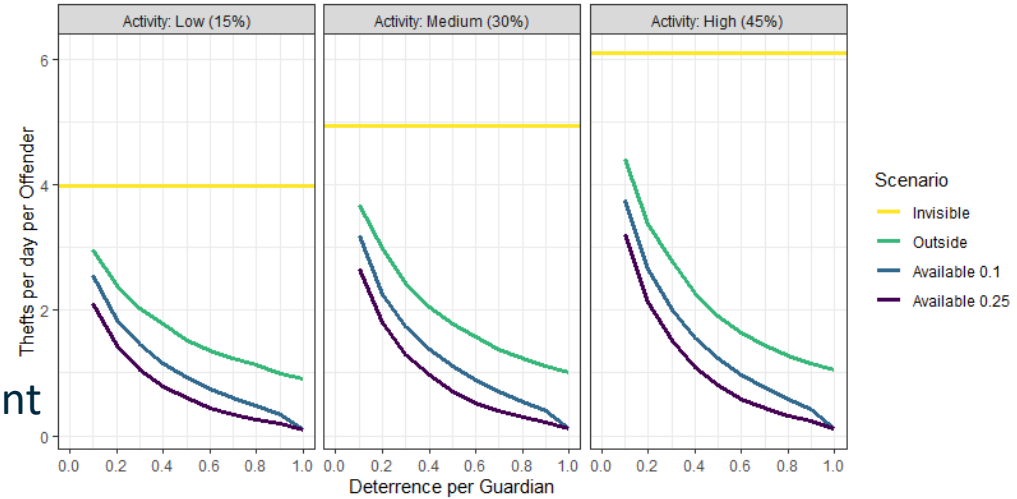
Results (guardianship model): Theft levels

Increased vigilance reduces thefts

- No surprise there

Greater activity increases thefts

- Less obvious than it may appear
- Opportunities increase faster than guardianship as other people present



Durham University

Results (guardianship model): Theft distribution

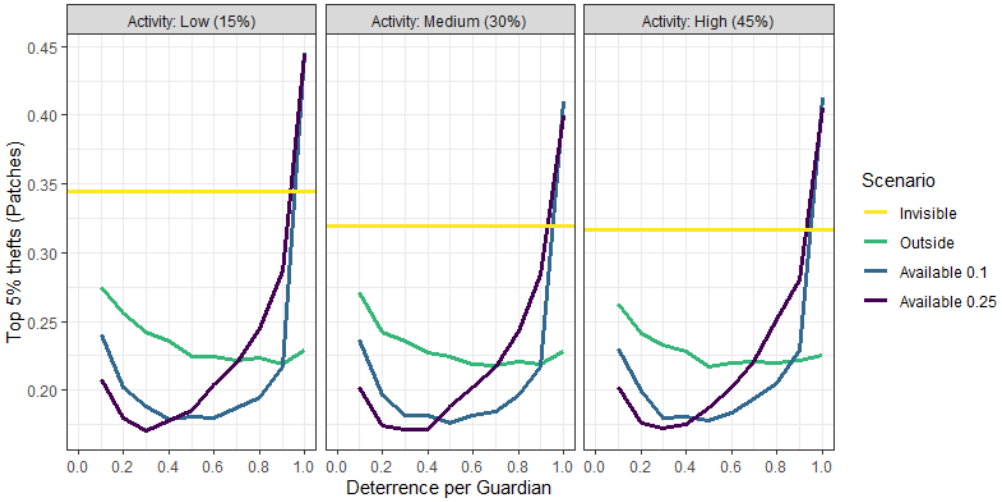
At low deterrence per guardian

- Thefts more widespread as deterrence increases
- Thefts more spatially concentrated as guardians added

Opposite patterns at high deterrence per guardian

Activity level has minimal effect

Not at all obvious



Durham University

ABMs as tools for discussion

Particularly in participatory modelling settings

- Stakeholders co-design the model

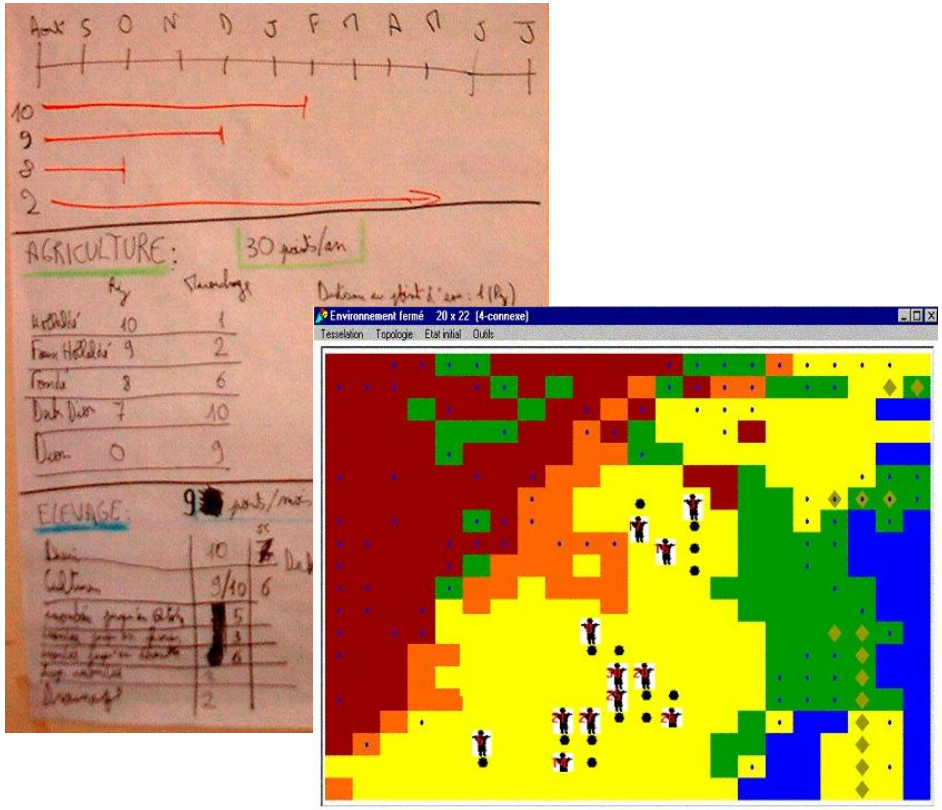
Design process elicits knowledge and competing perceptions

Model provides shared understanding

Playing model moves from entrenched positions

- Example: G'nith land use

Can criticise model instead of each other



D'Aquino, Le Page, Bousquet, Bah (2003). Using Self-designed Role-playing Games and a Multi-agent System to Empower a Local Decision-making Process for Land Use Management. JASSS 6(3) 5

Simulation

(wrap up)

Durham University

Two common simulation methods in social research

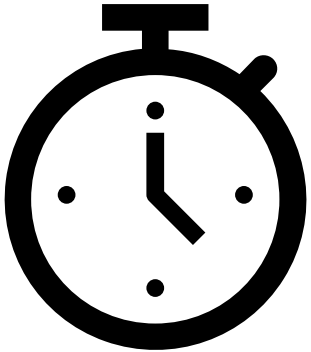
Microsimulation and Agent-Based Modelling both simulate individuals stepping through time

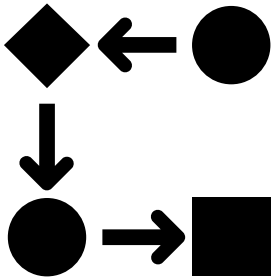
Microsimulation focus is state transitions

- Detailed state descriptions
- Individual transitions are independent

ABM focus is theoretical mechanisms of change

- Relatively simple states
- Agents interact with each other and environment





Characteristics of ABMs

Simulation

- Models a series of actions and their effects

Actions are taken by autonomous individuals (agents)

- What are my choices?
- What do I need to know?
- What do I do?

Heterogeneity

- Same agent, different situation → potentially different action
- Different agents, same situation → potentially different actions

Interaction

- Agents can perceive other agents
- Agent actions affect the situation of others and their future self

Fundamentals of Agent-Based Modelling

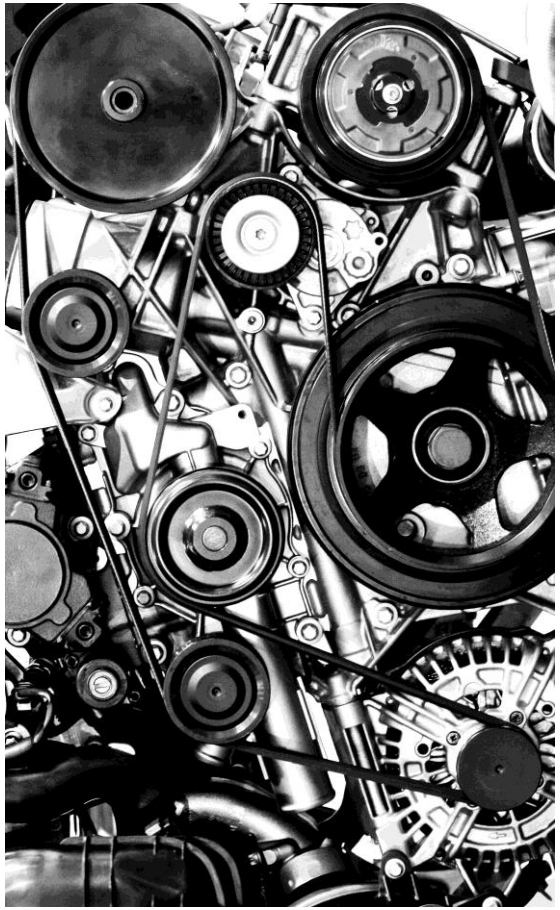
ABMs focus on the macro-consequence of micro-actions

You need to understand the system that you wish to model

- What process are you modelling?
- How do people make decisions that enact the process of interest?
- You need to have answers, or at least a plan to get them

ABMs represent the target system but are not the system itself

- Understand whether mechanisms are plausible explanation
- Cannot tell us what happened in any specific case; they cannot replace empirical study

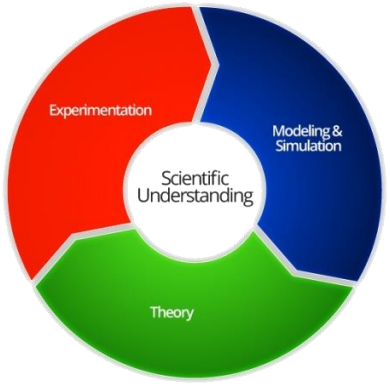


Types of research questions

Simulations can be used to extrapolate plausible future outcomes of complex systems…

…to test theories and hypotheses derived from empirical data

…or to produce theories and hypotheses for empirical examination



Justified Stories: formalising “what if?”

Stories

- Sequence of events (dynamic)
- Coherent
- Internally consistent

Justified

- Represents knowledge about the world
- Multiple sources: data, theory, expert opinion
- Externally consistent
  - plausible, given knowledge

Models assist insight

Tool for thinking: Model captures some theory of process and allows us to examine the implications of that theory

Next session: Agent-Based Modelling Workshop

Work with a simple NetLogo model

- Conformity with majority vote

Exercises

- Understand the existing model
- Experiment with the model
- Extend the model

NetLogo: Specialist ABM software

- Download from https://ccl.northwestern.edu/netlogo

Start model is on Blackboard Ultra