# Model Purpose and Complexity

*Slides from*

*Bruce Edmonds*

Centre for Policy Modelling

Manchester Metropolitan University

# A simulation model is not a ‘picture’

- Even if we often think of it in this way…
- …a model is ***not*** a picture of its target.
- Rather it is a tool to help deal with it, e.g.:
  - To understand it
  - To predict it
  - To communicate about it
- Just as machines extend our physical abilities, models extend our mental abilities.

# Different tools for different jobs

- A good tool is well designed for its purpose
- Each model is just such a tool
- However, there are many alternative models for every target so that we do not know what model is good for what purpose and what target
- So to be worth bothering other people about our models, to not waste their time…
- …our models needs to be justified with respect to a stated purpose and target etc.

# There are *no generic models*…

- ..yet and there seem little prospect for them in the foreseeable future.
- A ‘Darwin’ for the social sciences has not arrived with a integrative explanatory theory that connects social phenomena in a way that checks out against evidence (although we are working on this)
- Most work is in a ‘pre-integrative’ stage – specific models  for specific purposes

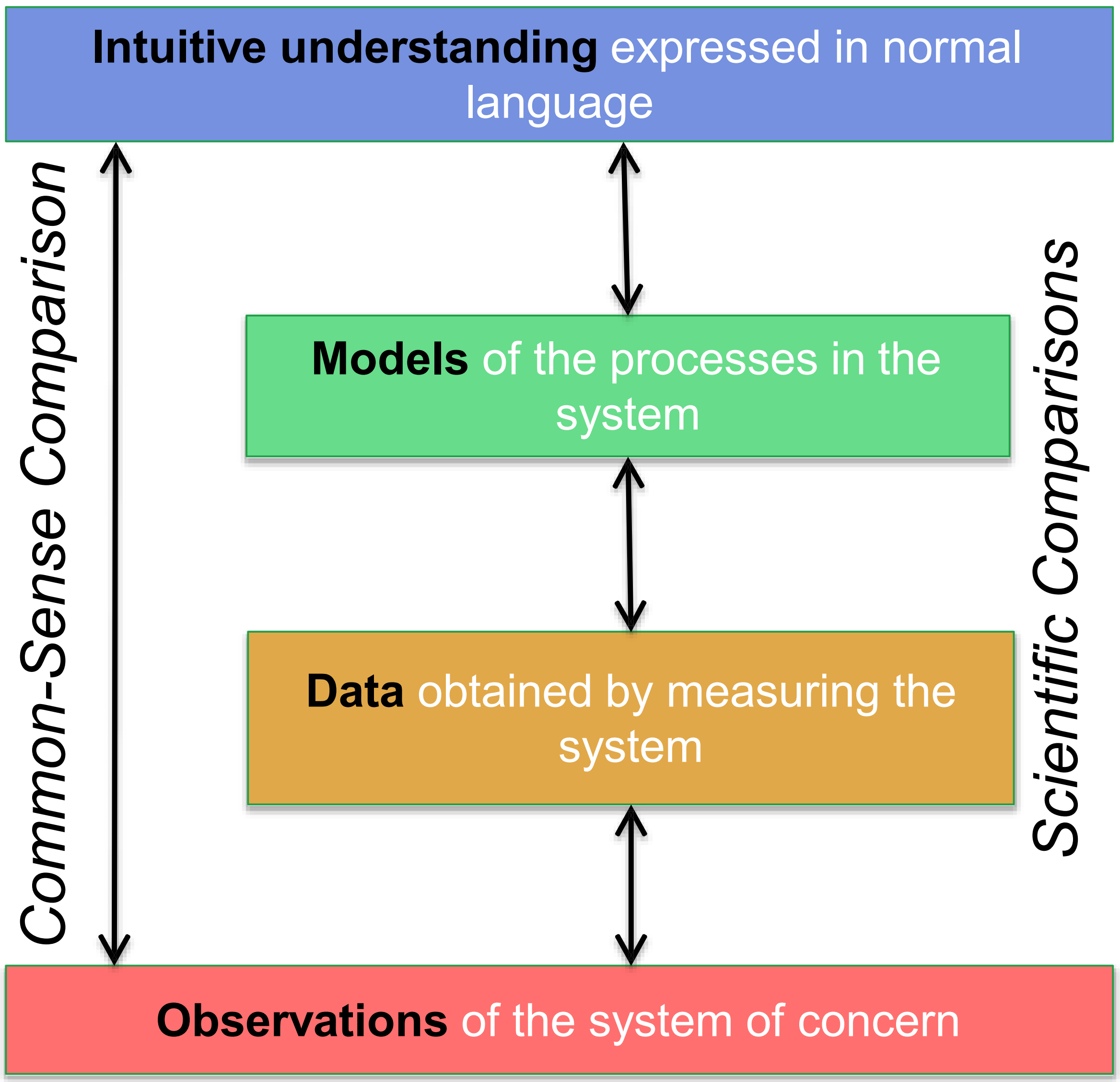
# Simpler is not more general

- Whilst one can add in specific detail to a simpler model to fit what is known about a specific case (making it less general)…
- ...the other way around does not generally work, *making a model simpler usually makes it less general*!
- This is because we do not know which of the processes are essential to a target, and might simplify these away
- Imagine removing acceleration from Newtonian dynamics to just use linear equations – the resulting approximations would only work at a few specific points!

# Using models as an analogy

- Sometimes models are not about the observed world, but related to our ideas/theories.
- In this case a model is used as a kind of complicated analogy – a way of thinking
- It does not relate to data but to a natural language understanding
- The trouble is we apply analogies with great (unconscious) fluidity, inventing a new way of relating the analogy to a situation 'on the fly'
- But this is different from empirical models

# Models stage understanding



# Different model purposes

- There are many different kinds of ways to use a model.
- Each such purpose has its own benefits and dangers...
- ...and needs different things checking for different purposes...
- ...and probably needs to be developed in a different way.

# Different model purposes

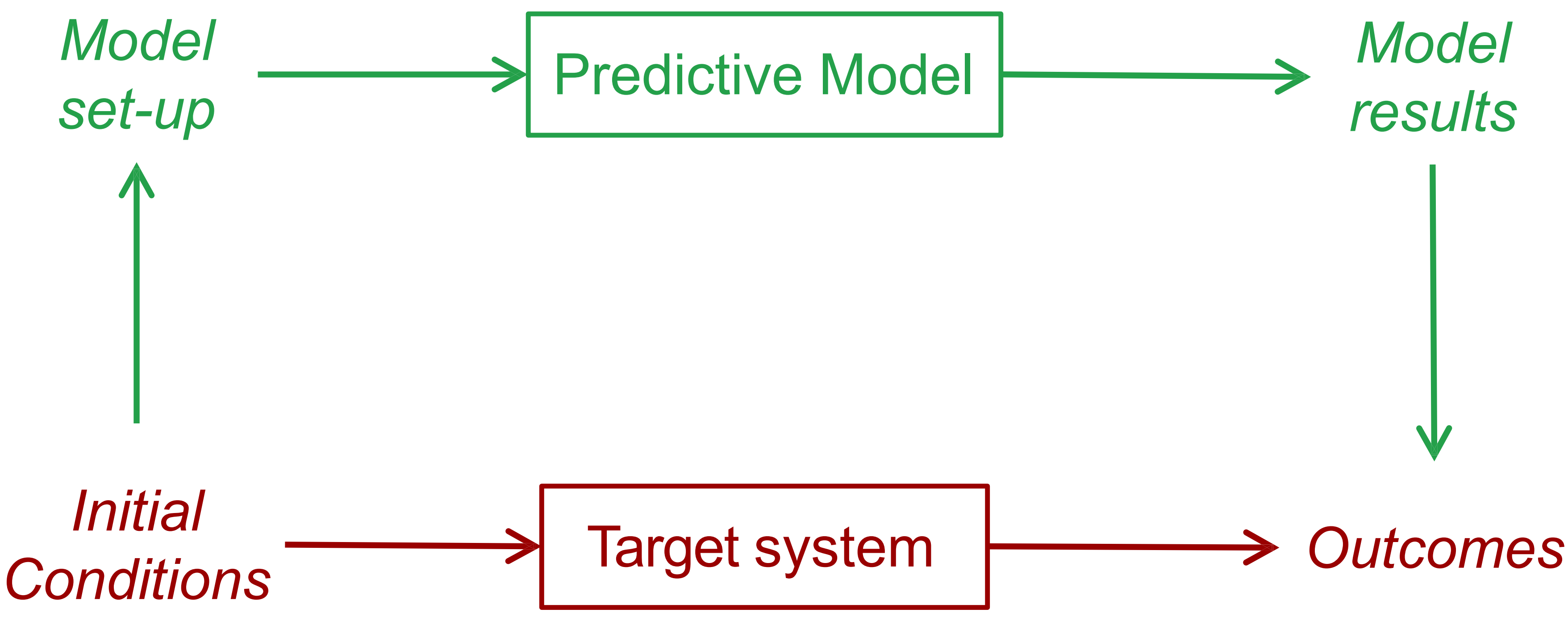
1. Prediction
2. Explanation
3. Theory exposition
4. Illustration
5. Analogy
6. Description

# Purpose 1: **Prediction**

# Motivation

- If you can *reliably* predict something about the world, this is undeniably useful…
- ...even if you do not know *why* your model predicts (e.g. a black-box model)!
- But it has also become the ‘gold standard’ of science…
- ...becuase (unlike many of the other purposes) it is difficult to fudge or fool yourself about – if its wrong this is obvious.

# Predictive modelling



# What it is

*The ability to anticipate unknown data reliably and to a useful degree of accuracy*

- Some idea of the conditions in which it does this well enough have to be understood
- The data it anticipates has to be unknown to the modeller when building the model
- What is a useful degree of accuracy depends on its application
- What is predicted can be: categorical, probability distributions, ranges, negative predictions, etc.

# Examples

- The gas laws (temperature is proportional to pressure at the same volume etc.) predict future measurements on a gas without any indication of why this works
- Nate Silver's team tries to predict the outcome of sports events and elections using computational models. These are usually probabilistic predictions and the wider predicted distribution of outcomes is displayed (http://fivethirtyeight.com)

# Warnings

- There are two different uses of the word ‘predict’: one as above and one to indicate any calculation made using a model.
- This requires repeated attempts at anticipating unknown data (and learning from this)
- because it is impossible to avoid ‘fitting’ known data (e.g. due to publication bias)
- If the outcome is unknown and can be unambiguously checked it could be predictive
- Prediction is ***VERY*** hard in the social sciences, it is rarely done

# Mitigating Measures

- The following are documented:
  - what aspects it predicts
  - roughly when it predicts well
  - what degree of accuracy it predicts with
- That the model has been tested that it predicts on several independent cases
- That the model code is distributed so others can test it and seek to understand how and when it predicts

# Purpose 2: **Explanation**

# Motivation

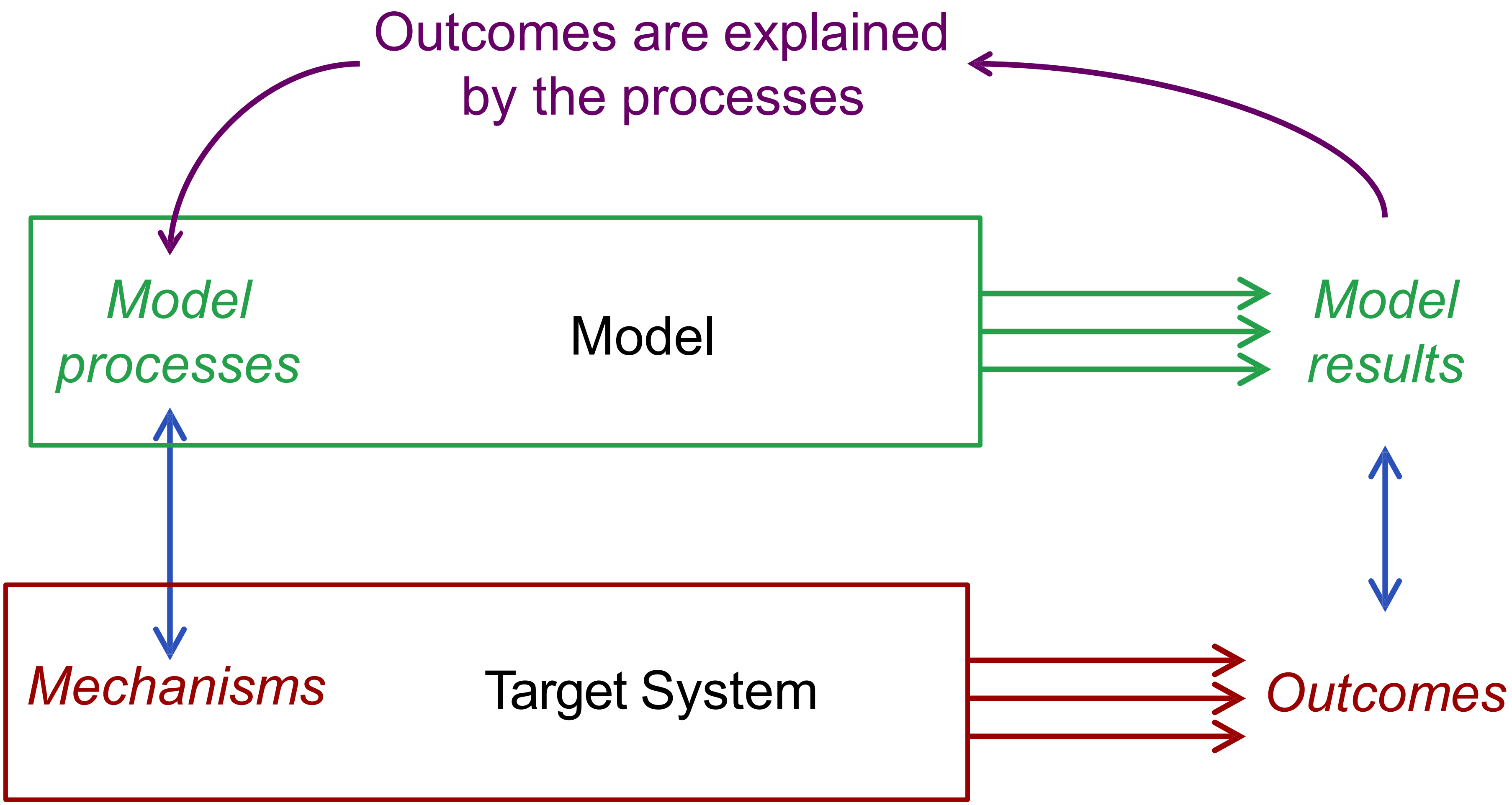
- When one wants to understand *why* or *how* something happens
- One makes a simulation with the mechanisms one wants and then shows that the results fit the observed data
- The intricate workings of the simulation runs support an explanation of the outcomes in terms of those mechanisms

# What it is

*Establishing a possible causal chain from a set-up to its consequences* in terms of the mechanisms of a simulation

- The causation can be deterministic, possibilistic or probabilistic
- The nature of the set-up determines the terms that the explanation is expressed in
- Only some aspects of the results will be relevant to be matched to data

# Explanatory modelling



# Examples

- The model of a gas with atoms randomly bumping around explains what happens in a gas (but does not directly predict the values)
- Lansing & Kramer's (1993) model of water distribution in Bali, explained how the system of water temples acted to enforce social norms and a complicated series of negotiations

# Warnings

- The fit to the target data maybe a very special case which would limit the likelihood of the explanation over other cases
- The process from mechanisms to outcomes might be complex and poorly understood. The explanation should be clearly stated and tested. Assumptions behind this must be tested.
- There might well be more than one possible explanation (and/or model)!

# Mitigating Measures

- Ensure the built-in mechanisms are plausible and at the right "level“
  - Using ABM makes this easier
- Be clear which aspects of the output are considered significant and which artifacts of the simulation
- Probe the simulation to find when the explanation works (noise, assumptions etc)
- Do classical experiments to show your explanation works for your code

# Purpose 3: **Theory Exposition**

# Motivation

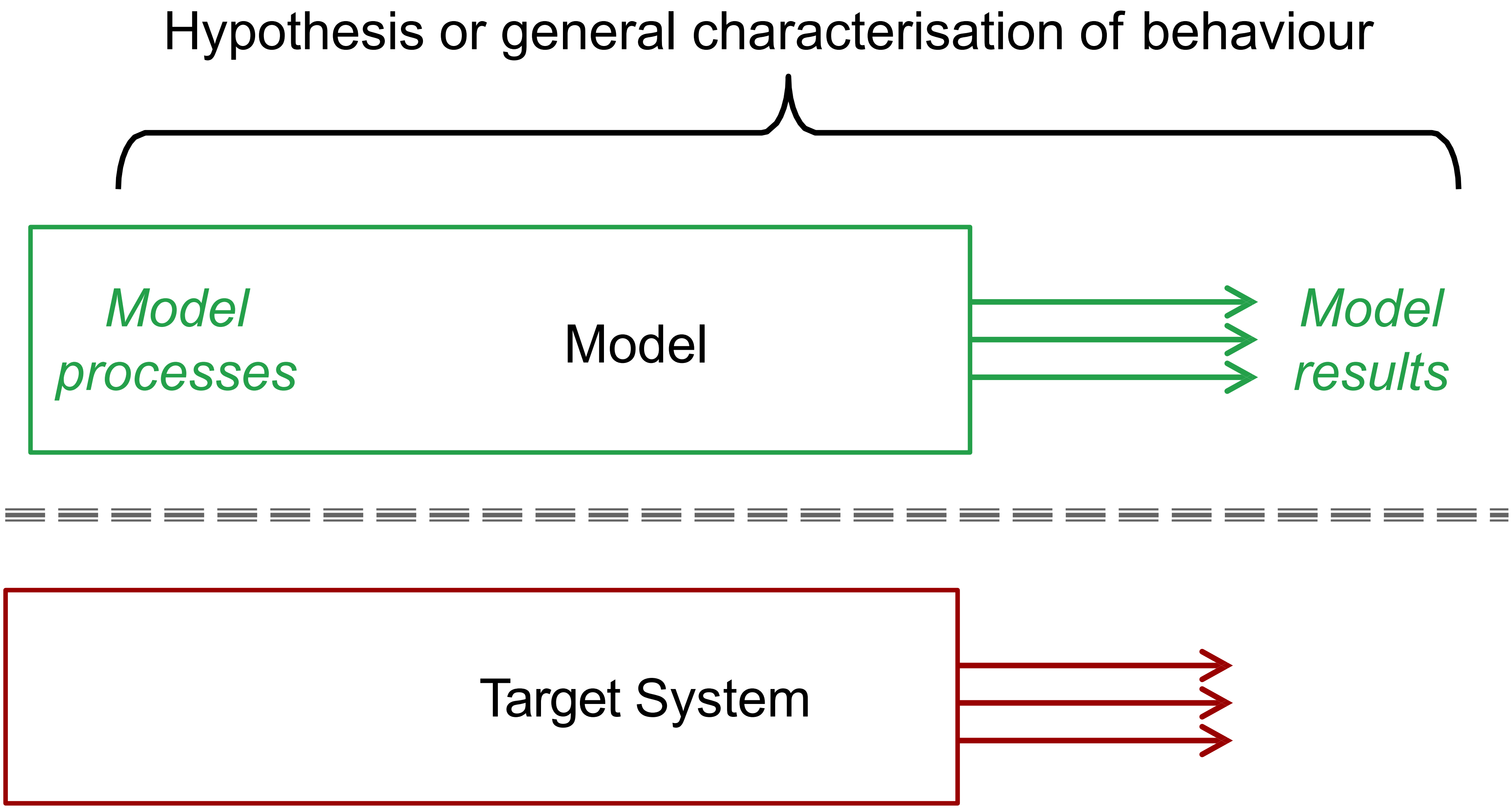
- If one has a system of equations, sometimes one can analytically solve the equations to get a general solution
- When this is not possible (almost all complicated systems) we can calculate specific examples – to simulate it!
- We aim to sufficiently explore the whole space of behaviour to understand a particular set of abstract mechanisms
  - E.g. using value trees to connect values to actions

# What it is

Discovering then *establishing (or refuting) hypotheses* about the *general behaviour* of a set of mechanisms

- The hypotheses may need to be discovered
- But crucially showing the hypotheses hold (or are refuted) by the set of experiments
- the hypotheses need to be quite general for the exercise to be useful to others
- Does not say anything about the observed world!

# Modelling to understand Theory



# Examples

- Many economic models are explorations of sets of abstract mechanisms
- Deffuant, G., et al. (2002) How can extremism prevail? jasss.soc.surrey.ac.uk/5/4/1.html
- Edmonds & Hales (2003) Replication… jasss.soc.surrey.ac.uk/6/4/11.html

# Warnings & Mitigation

- A bug in the code is fatal to this purpose
- A general idea of the outcome behaviour so the exploration needs to be extensive
- Clarity about what is claimed, the model description etc. is very important
- It it tempting to use the model as a way of thinking about the world, but this is dangerous!

# Purpose 4: **Illustration**

# Motivation & What it is

- An idea is new but complex and one wants to simply illustrate it
- This is a way of communicating through a single (but maybe complex) example

A behaviour or system is illustrated precisely using a simulation

- It might be a very special case, no generality is established
- It might be used as a counter-example

# Examples

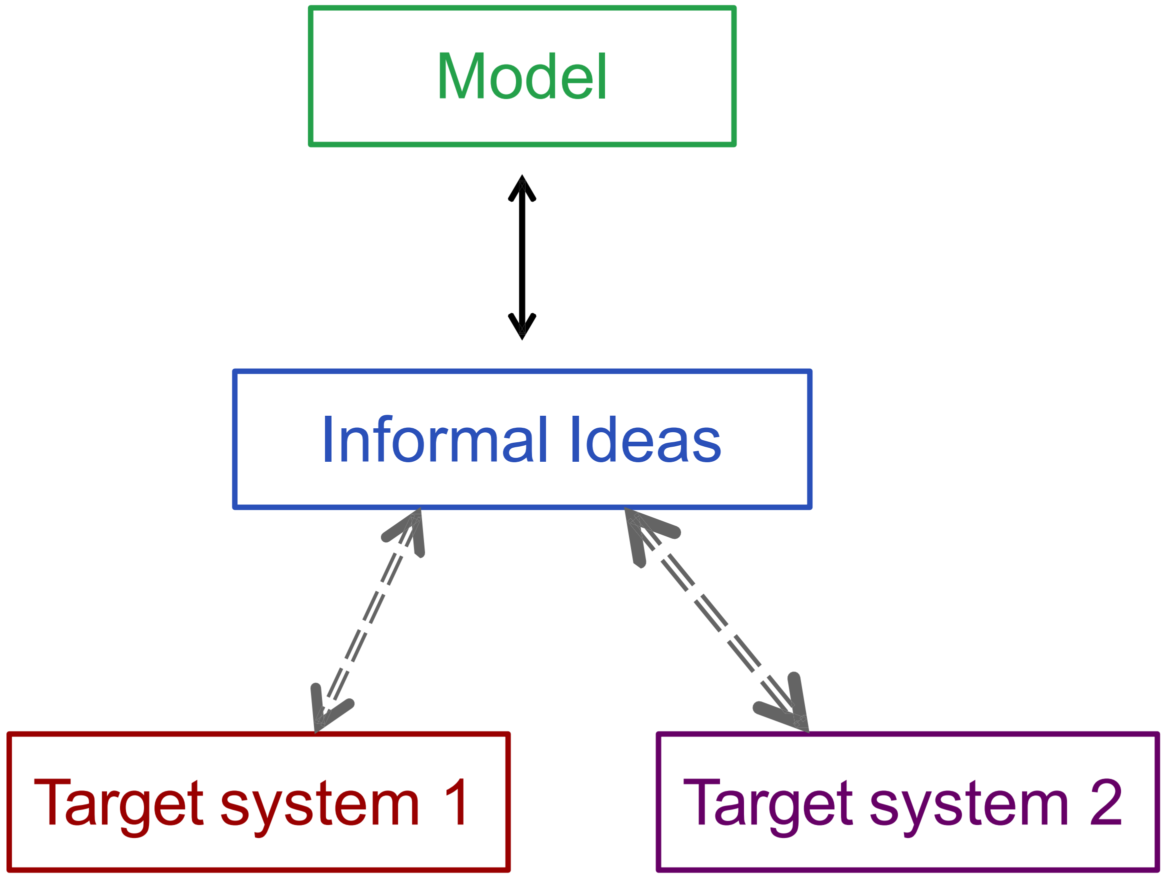
- Sakoda/Schelling’s 2D Model of segregation which showed that a high level of racial intollerance was not necessary to explain patterns of segregation
- Riolo et al. (2001) Evolution of cooperation without reciprocity, Nature 414:441-443.
- Baum, E. (1996) Toward a model of mind as a laissez-faire economy of idiots.

# Purpose 5: **Analogy**

# Motivation & What it is

- Provides a ‘way of thinking about’ stuff
- The model is not (directly) about anything observed, but about ideas (which, in turn, may or may not relate to something observed)
- It can suggest new insights or new future directions for research
- We need analogies to help us think about what to do (e.g. what and how to model)

# An illustration of analogical use of a model



# Examples

- Axelrod's Evolution of Cooperation models (1984 etc.)
- Hammond & Axelrod (2006) The Evolution of Ethnocentrism. Journal of Conflict Research
- Many economic models which show the efficiency of markets
- Many ecological models showing how systems reach an equilibrium

# Warnings

- When one has played with a model the whole world looks like that model
- But this does not make this true!
- Such models can be very influential but (as with the economic models of risk about lending) can be very misleading

# Purpose 6: **Description**

# Motivation & What it is

- Much science involves a lot of description (e.g. Darwin drawing Finches)
- Simulations can also be used in this way

This is an attempt to partially represent what is important of a specific observed case

- It does abstract (as all modelling does) but cautiously, retaining as much relevant detail as possible
- Later we might go back to the description and learn something new from it

# Examples

- Scott Moss's (1998) Model of handling crises in a water pumping station. JASSS 1(4):1, jasss.soc.surrey.ac.uk/1/4/1.html
- Richard Taylors (2003) thesis. http://cfpm.org/cpmrep137.html
- Sukaina Bharwani's (2004) thesis. Translating interviews with farmers into a simulation http://goo.gl/MzJJR7

# Warnings & Mitagation

- These models should have a good evidential base – qualitative or quantitative
- Might use expert or stakeholder opinion
- Assumptions and their basis should be very carefully documented
- They maybe complex but have NO level of generality – they are a particular case
- Need later further work for generalisation or analysis of complicated simulations

# Summary of Modelling Purposes

| ***Modelling Purpose*** | **Essential features** | **Particular risks** (apart from that of lacking the essential features) |
|---|---|---|
| *Prediction* | Anticipates *unknown* data | Conditions of application unclear |
| *Explanation* | Uses *plausible* mechanisms to match outcome data in a *well-defined* manner | Model is brittle, so minor changes in the set-up result in bad fit to explained data |
| *Theory* | Systematically maps out or establishes the consequences of some mechanisms | Bugs in the code; inadequate coverage of possibilities |
| *Analogy* | Provides new insights | Danger of thinking it is general and true |
| *Description* | Relates directly to evidence for a small set of cases | Unclear documentation; over generalisation from cases described |
| *Illustration* | Shows an idea clearly | Over interpretation to make theoretical or empirical claims |

# Some common confusions

- Firstly in many publications researchers do not make their model purpose clear
- So the model is hard to judge properly
- Some have simply not thought about it! Some common confusions:
- Theory → Analogy
- Illustration → Explanation
- Description → Explanation
- Explanation → Prediction

# Pragmatics of Model Development

# Exploratory & Consolidation

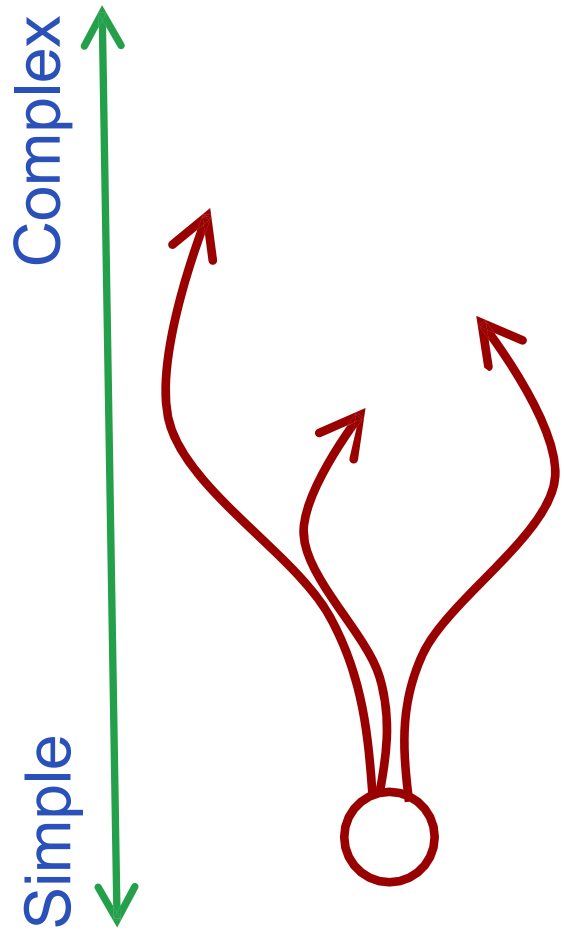
- It is common that one does not know clearly what one wants to program
- In this case you might ‘play’ with models, exploring the possibilities, getting an idea of what works, what is possible etc.
- But this does not make a model suitable for communicating with others
- Rather you then need to re-do the model properly and justify what it does rigourously

# How complicated should you make your model?

- Simpler models are easier for us humans to deal with. Easier to: program, check, understand, communicate, analyse etc.
- But if you leave out crucial processes your model may simply not be adequate to predict, explain etc. your target
- In general, ***we have no idea what is necessary and what is not*** – this is a major area of contention and confusion!
- So what to do? What strategy to follow?

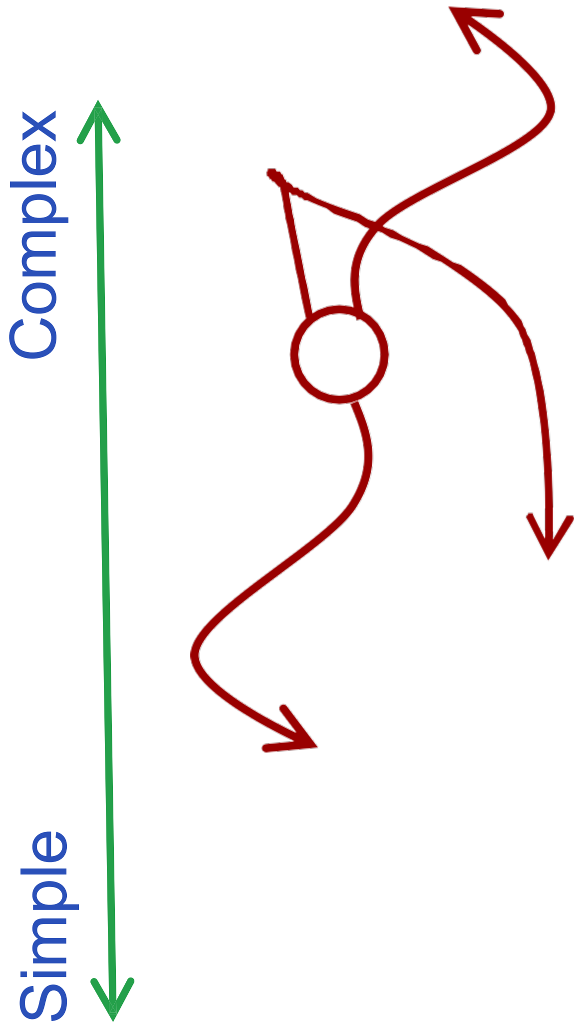
# KISS – keep it simple stupid!

- An engineering approach
- Start with a simple model and add features only when the simpler is shown not to work
- Trouble is it maybe a combination of mechanisms that are required – trying them one at a time might not work



# KIDS – keep it descriptive stupid!

- Start with the available evidence of what is important
- Then explore variations from there, maybe showing some are not required
- Trouble is this makes for complicated models so care is needed in development and further work needed to understand your model



# Which is better?

- There is no reason to suppose that simple models will be adequate for socio-ecological phenomena
- Limitations on: human understanding, time, resources, data are inevitable and should simply be honestly declared
- For each feature the decision as to whether to include it *or not include* it is important and needs justified
- This depends on the model purpose!

# Features and Purpose

| ***Modelling Purpose*** | **What you need in it** | **Evidence needs** |
|---|---|---|
| *Prediction* | Aspects/parameters that cause the significant outcomes in a reliable manner | Lots of numerical data to adjust the model, lots of new cases to check it works |
| *Explanation* | The mechanisms and structures you want the explanation in terms of | Enough evidence and data at a variety of levels, qualitative and quantitative for validation |
| *Theoretical exposition* | Not too many parameters, simplified mechanisms | None |
| *Description* | Whatever seems relevant about a particular case and that you have evidence for | Rich evidence and data at a variety of levels, qualitative and quantitative |
| *Analogy* | Just enough for the analogy to work | Terms in model need real referents, but no other |
| *Illustration* | Just enough to show the example dynamics | Terms in model need real referents, particular case needs to be plausible |

# Using models together

- Part of the answer seems to be that we have to use 'clusters' of models together

For example:

- A simpler model of a more complex model to understand it (model chains, meta-modelling)
- Different (but related) models for different aspects and different purposes
- Simpler models for understanding, but more complex for actual use but checked against the simpler

# Many models but distinct purposes

- Use different kinds of model together, but keep them distinct!
- Each model for a single purpose
- Think how the models relate to each other
- If a model has two purposes, it is better to think of them as different models (e.g. modelA-theory and modelA-analogy)
- In presentations/publications you need to justify a model for each use seperately

# Conclusion: Reflective Modelling

# Think!

- Think about what you are trying to do, and ***why*** as you develop or use your simulations
- Keep the models and purposes distinct but think how they relate
- Analogies are needed to think, including what we are going to model and why…
- But the serious models themselves are the business, they either demonstrate or not that they are adequate to their purpose

# Check!

- There are many problems when finding interesting results
- Simulations need a lot of work before it  is worth wasting others' time with them
- Lots of checking of the code, the results etc. are necessary (jasss.soc.surrey.ac.uk/12/1/1.html)
- Lots of documentation of your code, your results, your experiments, what fails etc. are necessary

# Share!

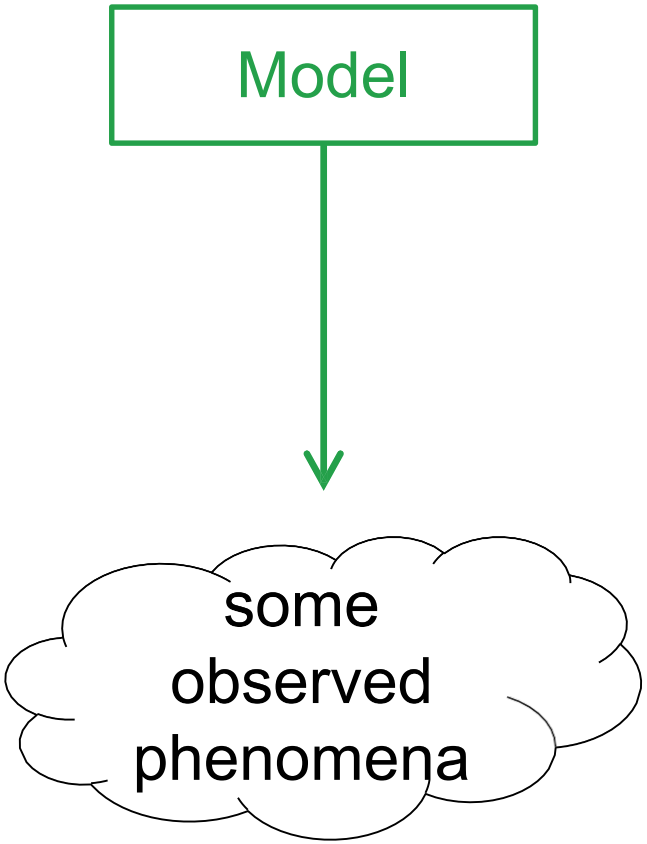
- You are not alone!
- The power of formal models (like computer simulations) is they can be unambiguously shared with others who can then run, check, inspect, change etc.
- Share your models and ideas at an early stage (e.g. at OpenABM.org), be honest and open – this is the way to learn
- But distinguish between developing work and mature work so others can understand how seriously to take it

# Beware!

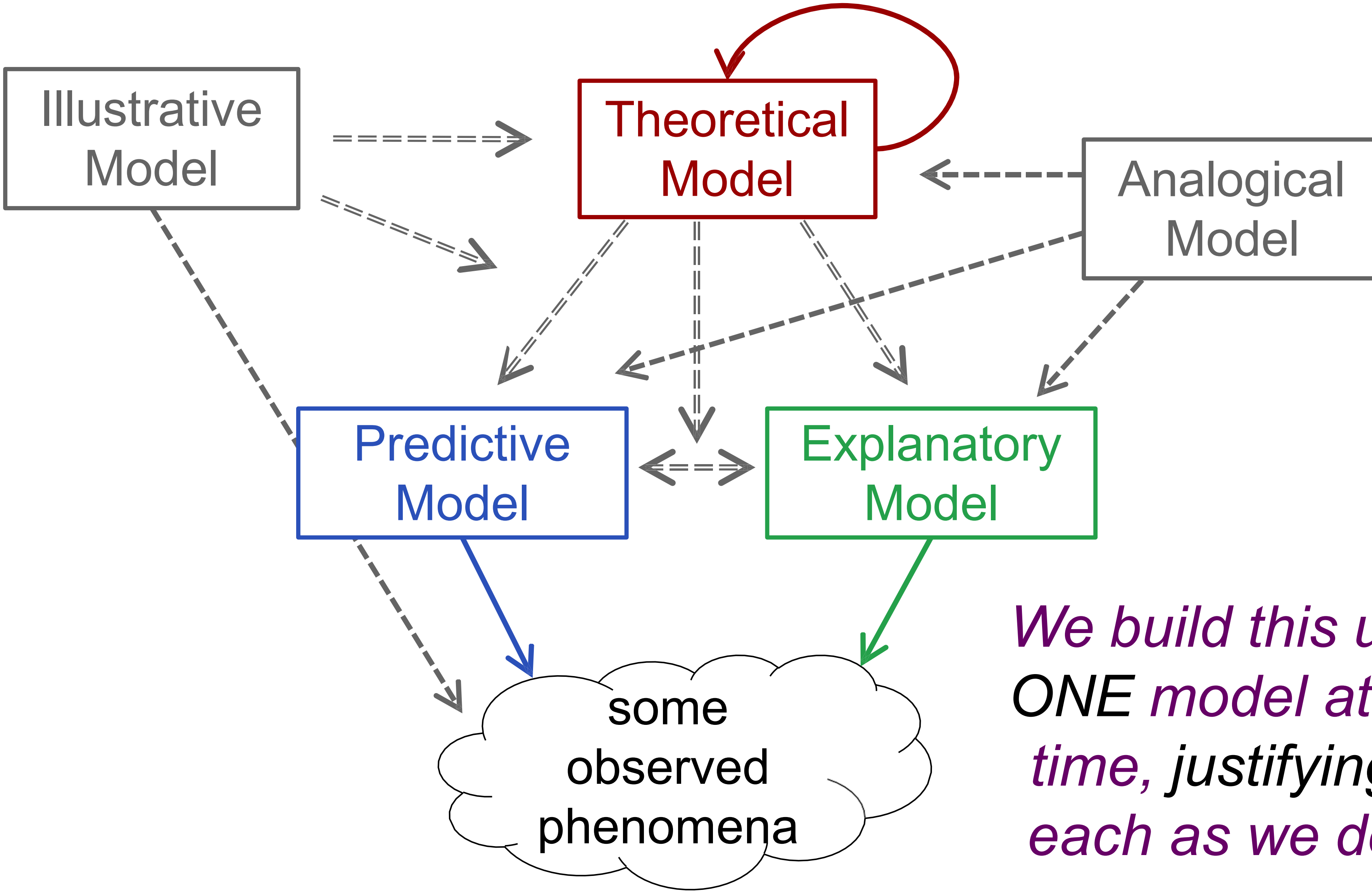
- We, as humans, are very, VERY, ***VERY*** good at deceiving ourselves
- Simulation models make this worse!
- We use models as a way of thinking about things and then it is hard to think about them in other ways (including new ways)
- Any way to shake us up and reconsider this is good – empirical data is a good way of doing this, critiquing models can also help
- You should be throwing away ***most*** of your simulations, not elaborating bad models

# The key idea of this talk

*Simplistic picture of a universal modelling approach*



# The key idea of this talk




*We build this up* ONE *model at a time,* justifying *each as we do*