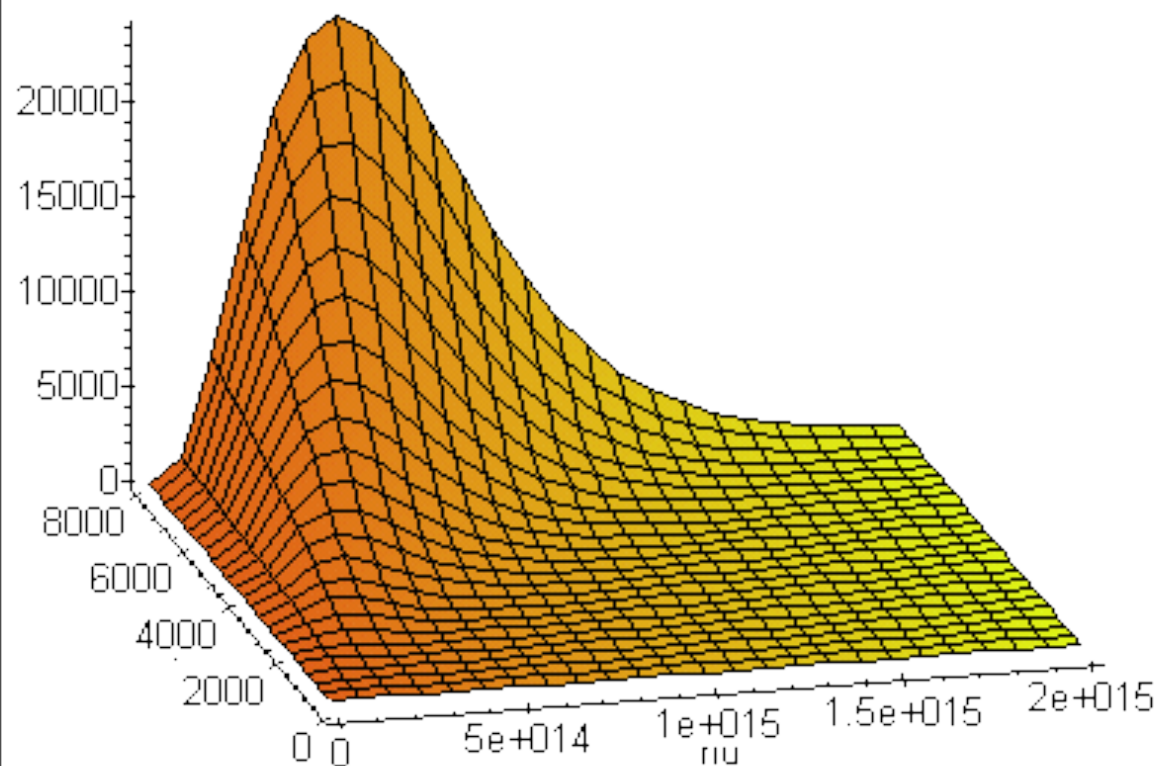
Model building and Validation



Nigel Gilbert

Centre for Research in Social Simulation
University of Surrey

www.simian.ac.uk

Overview

SIMIAN
Simulation Innovation: a Node

- Programming tools
- How an agent-based model is constructed
- Verification and validation
- Validation and model objectives
- Where to learn more
  - Further reading
  - Courses

Tools

SIMIAN
Simulation Innovation: a Node

AgentSheets

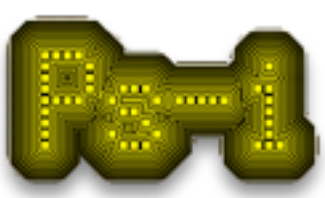
Swarm Development Group
www.swarm.org

Cormas
Natural resources and agent-based simulations
Cirad • Web des SAVOIRS

ASCAPE

M A S O N

RePast



JAS
Java Agent-based Simulation library

www.simian.ac.uk

ESRC National Centre for Research Methods

Tools

SIMIAN
Simulation Innovation: a Node

- Special purpose ‘toolkits’ and ‘packages’
  - Repast, Mason, NetLogo
  - adaptability? complexity?
- Special purpose simulation language
  - Matlab, Mathematica
  - flexibility?
- General purpose programming language
  - C++, Lisp, Smalltalk, Java
  - development tools?

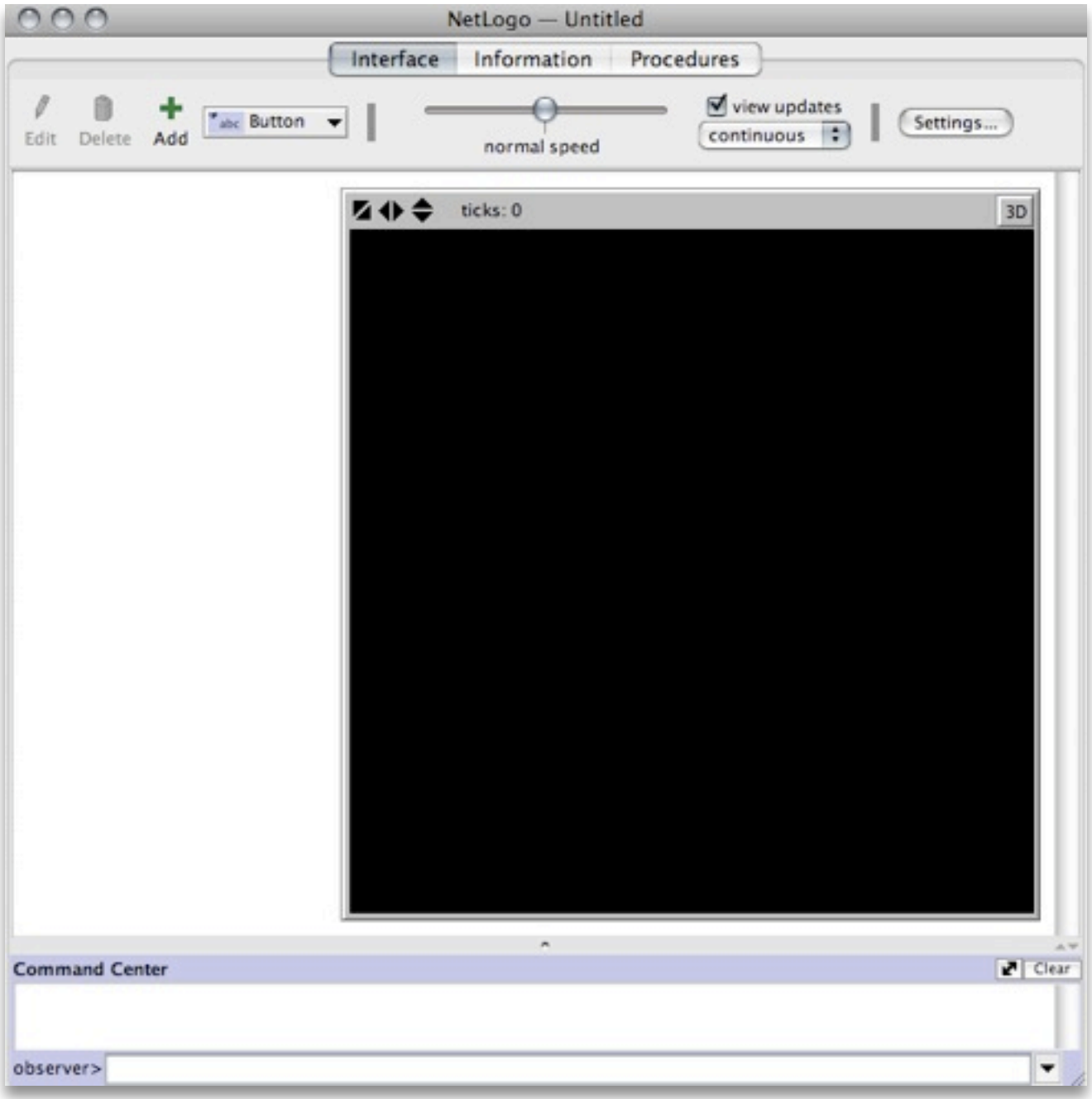
NetLogo

SIMIAN
Simulation Innovation: a Node

- Free
  - but not open source
- Runs on Windows, Mac OS X, Linux
- Own programming language
  - designed for school children
  - distant descendant of Logo
- Good documentation and tutorial
- Very high level language
  - many complicated ideas can be expressed in one or two lines of code
  - built in graphics for plots and controls
- Good user forum
  - provides help if you get stuck

NetLogo

SIMIAN
Simulation Innovation: a Node



ESRC National Centre for Research Methods

A very simple NetLogo model

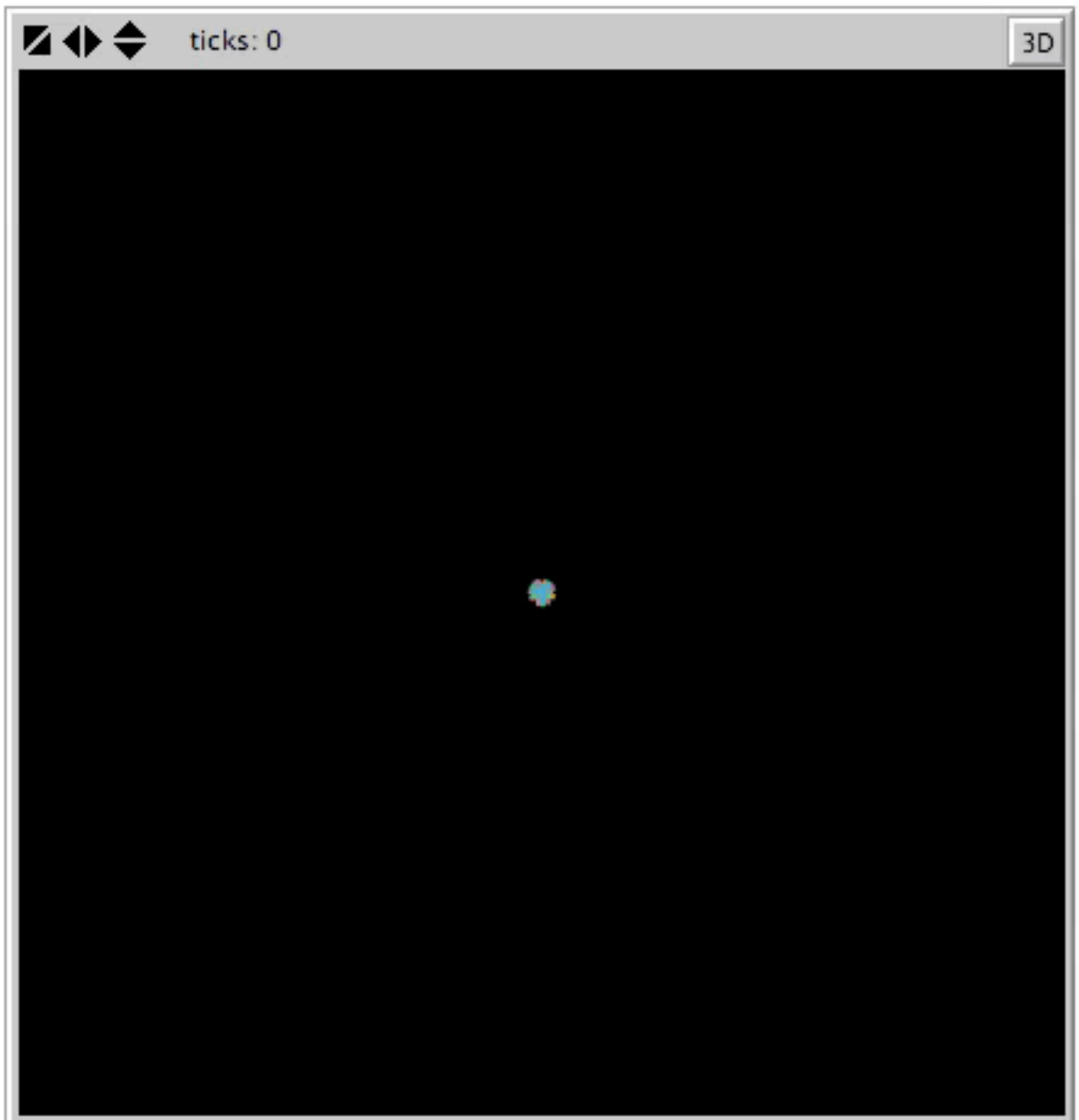
SIMIAN
Simulation Innovation: a Node

```
to setup
  clear-all
  create-turtles 10
end

to go
  ask turtles [
    right (random 360)
    forward 1
    ]
end
```


ESRC National Centre for Research Methods

SIMIAN
Simulation Innovation: a Node



Setup

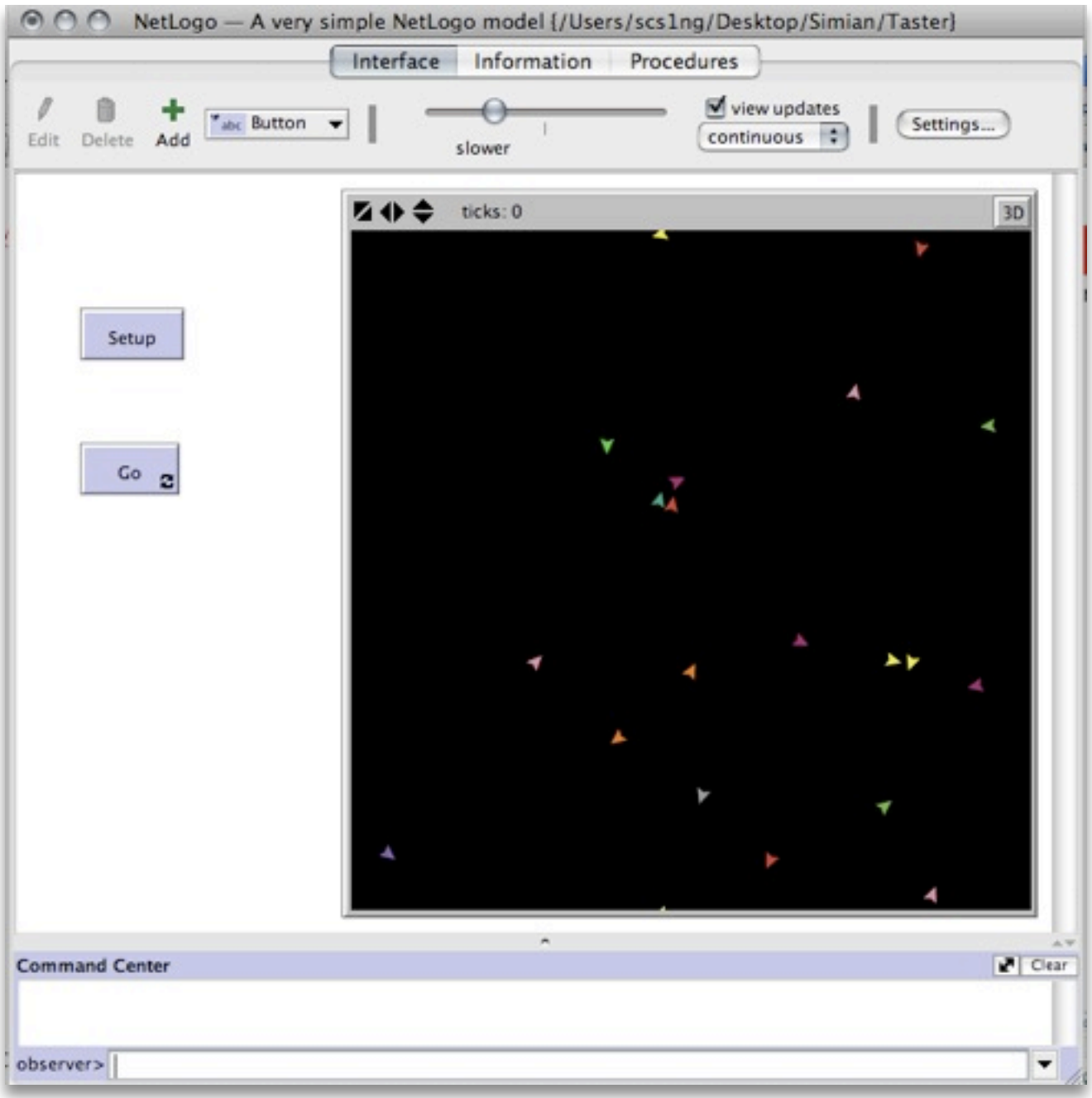
Go

www.simian.ac.uk

ESRC National Centre for Research Methods

A very simple NetLogo model

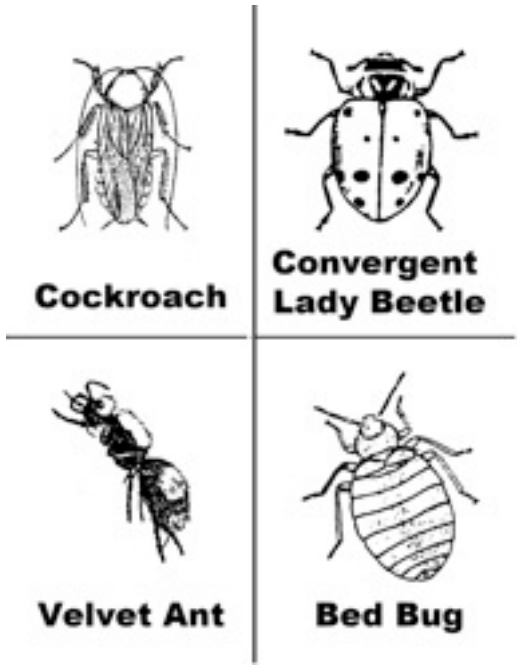
SIMIAN
Simulation Innovation: a Node



ESRC National Centre for Research Methods

Verification and validation

SIMIAN
Simulation Innovation: a Node



- Verification
  - **Is the model right?**
  - Getting rid of bugs
- Validation
  - **Is it the right model?**
  - Checking whether the model is a *good* model of something
  - 'Good' depends on one's objectives

www.simian.ac.uk

ESRC National Centre for Research Methods

Validation

SIMIAN
Simulation Innovation: a Node

- Is the model a good model?
- It depends on the modeller's objectives
  - Formalising a theory
    - Usually an abstract model
  - Developing middle range theory
    - Model of a class of phenomena
  - Modelling a specific situation
    - Facsimile models

Abstract models

SIMIAN
Simulation Innovation: a Node

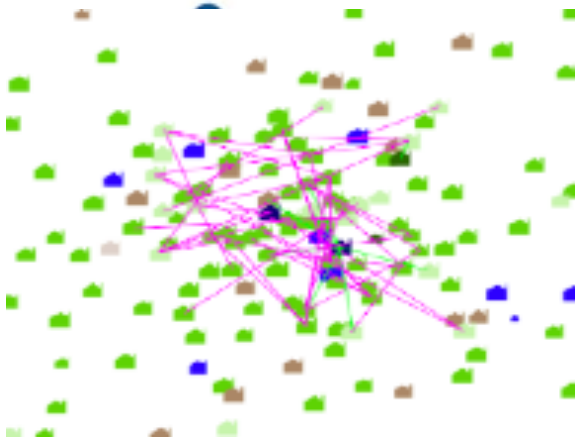
- Aim: demonstrate some (probably emergent) social process or mechanism
- No corresponding specific empirical case
- Example:

  Models of opinion dynamics

- Validation criterion:

  Does it generate more specific ('middle range') theories that can be tested empirically?

Middle range models



- Aim: describe the general characteristics of a particular social phenomenon
- Should be applicable to many specific cases
- Example:

  models of innovation networks, supply chains
- Validation criterion:

  Qualitative resemblance

  Similar dynamics

  'History friendly' models

Facsimile models

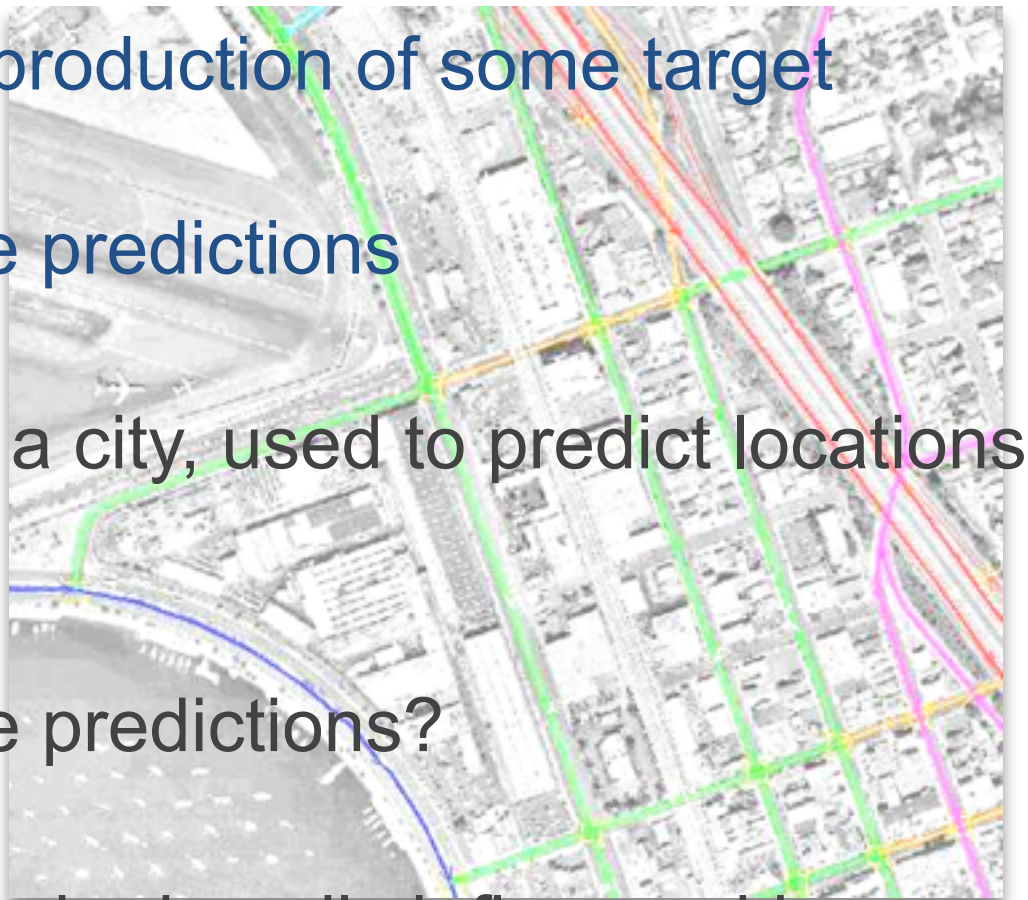
SIMIAN
Simulation Innovation: a Node

- Aim: provide an exact reproduction of some target phenomenon
- Often intended to provide predictions
- Example:

  a model of the traffic in a city, used to predict locations of potential jams
- Validation criterion

  Does it lead to accurate predictions?
- Problem:

  behaviour of model may be heavily influenced by random events (simulated using a random number generator)

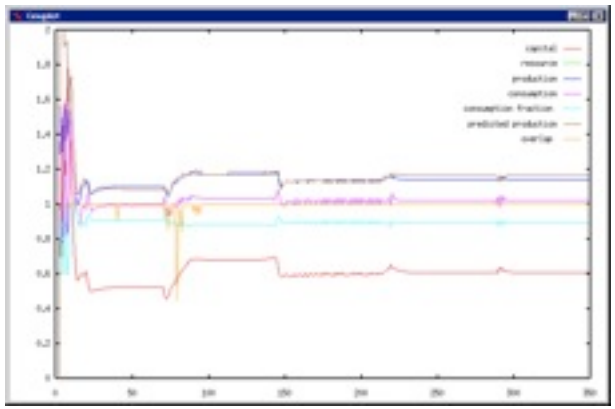


Comparing outputs with observed data

SIMIAN
Simulation Innovation: a Node

- ‘Observations’ may be impossible
  - too abstract
  - inaccessible (e.g. social complexity in 20,000BC)
- Differences may be due to any or all of:
  - bad model
  - bad data
  - model is an abstraction of the target
  - ‘random’ variations, but sampling distributions are unknown

Comparing model output and data



- Basic idea:
  - Use the same statistics as one would use to compare a statistical model with data
    - E.g. R^2, the regression coefficient
    - Spearman's rho for ordinal data
    - Measure of association such as phi for categorical data
  - Caution!
    - It is unlikely that the model output will be normally distributed
    - When comparing time series, consider auto-correlation
      - i.e. value at time *t*+1 is not independent of value at time *t*

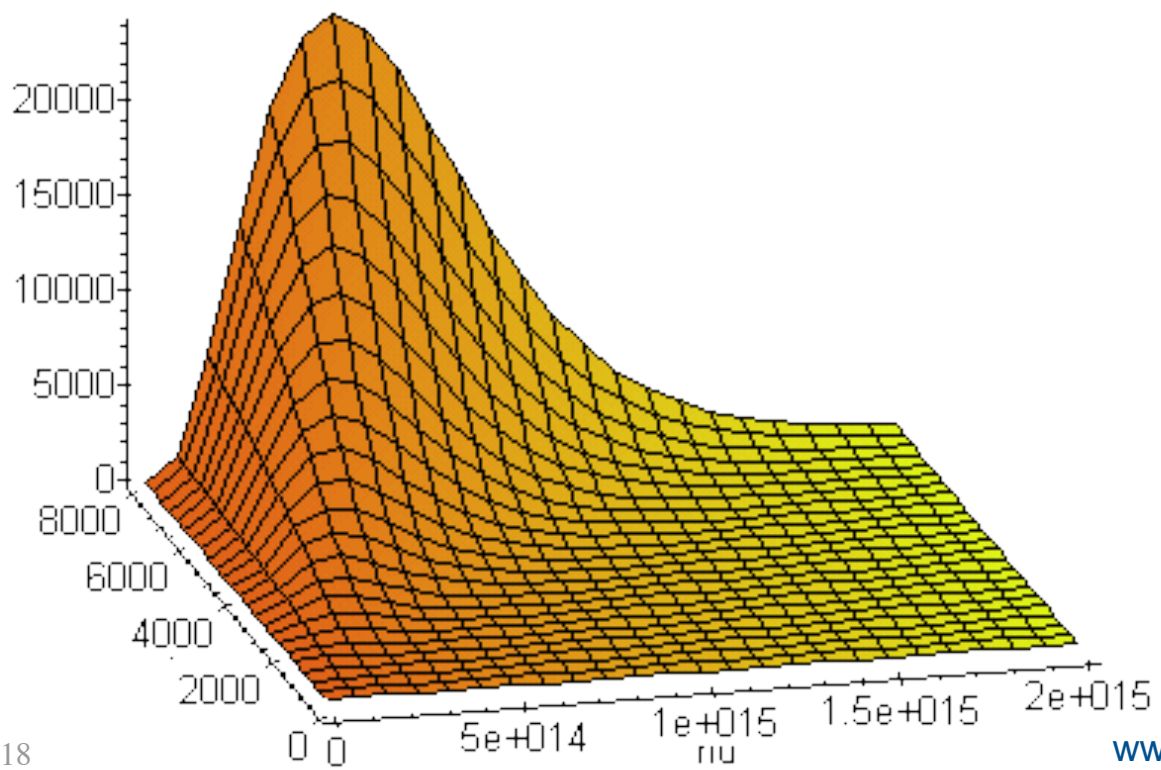
Validity

- Other related questions
  - sensitivity to values of the input parameters
    - do small changes in the values of the inputs result in large changes in the outputs?
    - if so, are you sure that the values used are the right ones?
  - repeatability
    - is the output similar on every run?
    - if not, are you sure that the runs are typical?
  - simplicity
    - could the model be simplified without affecting its validity?
    - if so, why use a more complicated model than necessary?

Validity

SIMIAN
Simulation Innovation: a Node

- Sensitivity analysis
  - repeatedly run the model with small variations in input parameters and observe outputs
  - but the space of possible input values is exceedingly large



www.simian.ac.uk

ESRC National Centre for Research Methods

Questions?

SIMIAN
Simulation Innovation: a Node

.... and then some suggestions for further reading

ESRC National Centre for Research Methods

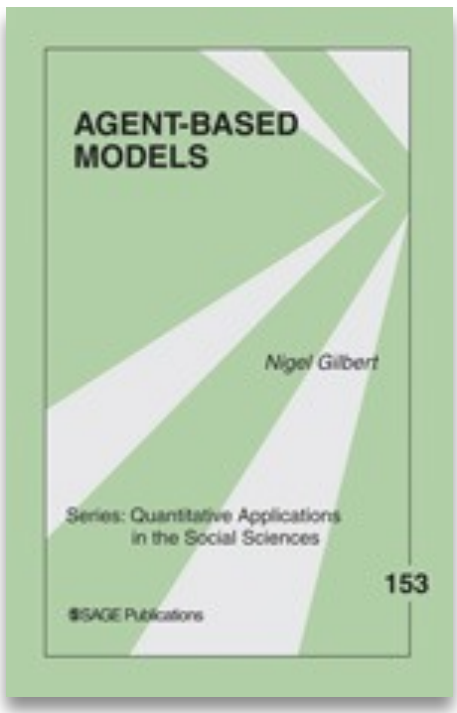
To learn more

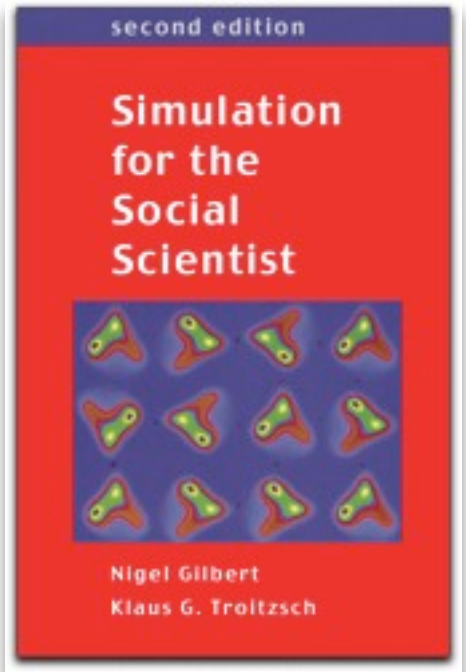
SIMIAN
Simulation Innovation: a Node

http://www.soc.surrey.ac.uk/JASSS/

JASSS

Journal of Artificial Societies and Social Simulation





essa

European Social Simulation Association
http://www.essa.eu.org

www.simian.ac.uk

ESRC National Centre for Research Methods

Email list

SIMIAN
Simulation Innovation: a Node

- http://www.jiscmail.ac.uk

Bibliography

SIMIAN
Simulation Innovation: a Node

- Ahrweiler, P., Pyka, A., & Gilbert, N. (2004). Simulating knowledge dynamics in innovation networks (skin). In R. Leombruni & M. Richiardi (Eds.), *Industry and labor dynamics: The agent-based computational economics approach*. Singapore: World Scientific Press.
- Chattoe, E., Saam, N. J., and Möhring, M. 2000. Sensitivity Analysis in the Social Sciences: Problems and Prospects. In *Tools and Techniques For Social Science Simulation*. R. Suleiman, K. G. Troitzsch, and G. N. Gilbert, Eds. Physica-Verlag, 243-273.
- Deffuant, G., Amblard, F., & Weisbuch, G. (2002). How can extremism prevail? A study based on the relative agreement interaction model. *Journal of Artificial Societies and Social Simulation*, 5(4).
- Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, 4(5), 41-60.
- Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies : Social science from the bottom up*. Washington, D.C. ; Cambridge, Mass. ; London: Brookings Institution Press : MIT Press.
- Gilbert, Nigel (2007) *Agent Based Models*. Quantitative Applications in the Social Sciences 153. Sage.

Bibliography

SIMIAN
Simulation Innovation: a Node

- Gilbert, N., & Abbott, A. (Eds.). (2005). Special issue: Social science computation. *American Journal of Sociology*. (Vol. 110 (4)). Chicago: The University of Chicago Press.
- Gilbert, N., & Terna, P. (2000). How to build and use agent-based models in social science. *Mind and Society*, 1(1), 57 - 72.
- Macy, M., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, 28, 143-166.
- Ramanath, A. M., & Gilbert, N. (2004). The design of participatory agent-based social simulations. *Journal of Artificial Societies and Social Simulation*, 7(4).
- Tesfatsion, l., & Judd, K. (2006). *Handbook of computational economics* (Vol. 2): North-Holland.
- Wilensky, U. (1999). *NetLogo*. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/