Simulating social systems with individual-based models: is worth it?

Alison Heppenstall
@ajheppenstall

Nick Malleson
@nickmalleson

Presentation to the Institute of Geography, University of Augsburg

Wed 13th July 2022

MRC/CSO Social and Public Health Sciences Unit

UKRI Medical Research Council

CHIEF SCIENTIST OFFICE

University of Glasgow

erc

European Research Council

Established by the European Commission

University of Glasgow

VIA VERITAS VITA

LEEDS Institute for Data Analytics

Urban Big Data Centre

UNIVERSITY OF LEEDS

UKPRP Prevention Research Partnership

SIPHER Consortium

Systems science In Public Health and Health Economics Research

The Alan Turing Institute

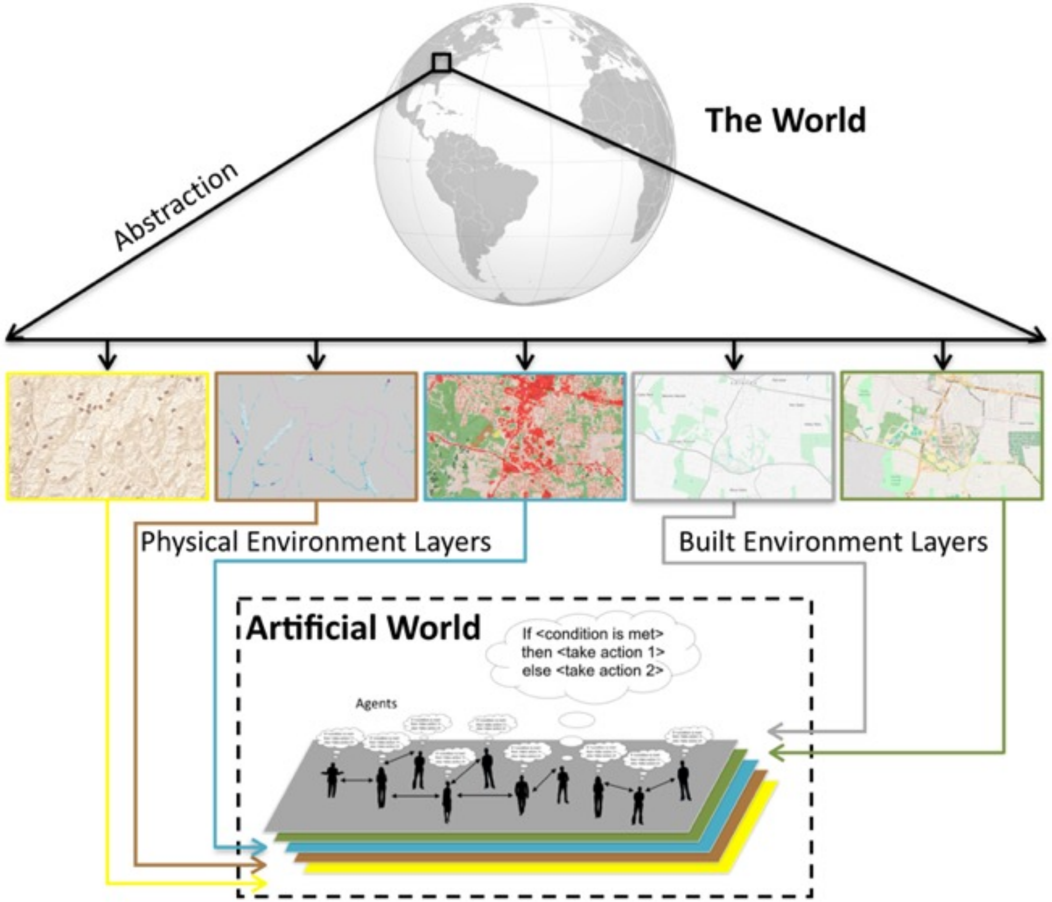
E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

UKRI UK Research and Innovation


LOUIS VUITTON


Why are urban environments difficult to understand and predict?

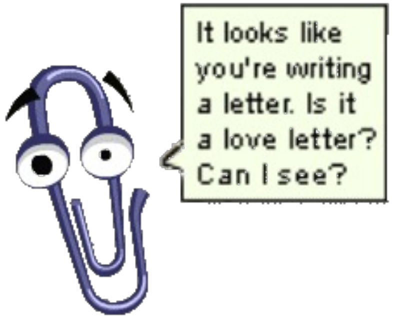
- Geographical space.
    - "…everything is related to everything else, but near things are more related than distant things" (Tobler).
- Human behaviour
- "The appeal is undeniable: it appears obvious that **individual**-level decision-making is the fundamental driver of social systems…" (p.113; O'Sullivan et al, 2012)
- Batty (2014) describes cities as products of networks, comprised of **individual** actors, interconnected at different levels.
- Complexity + systems science



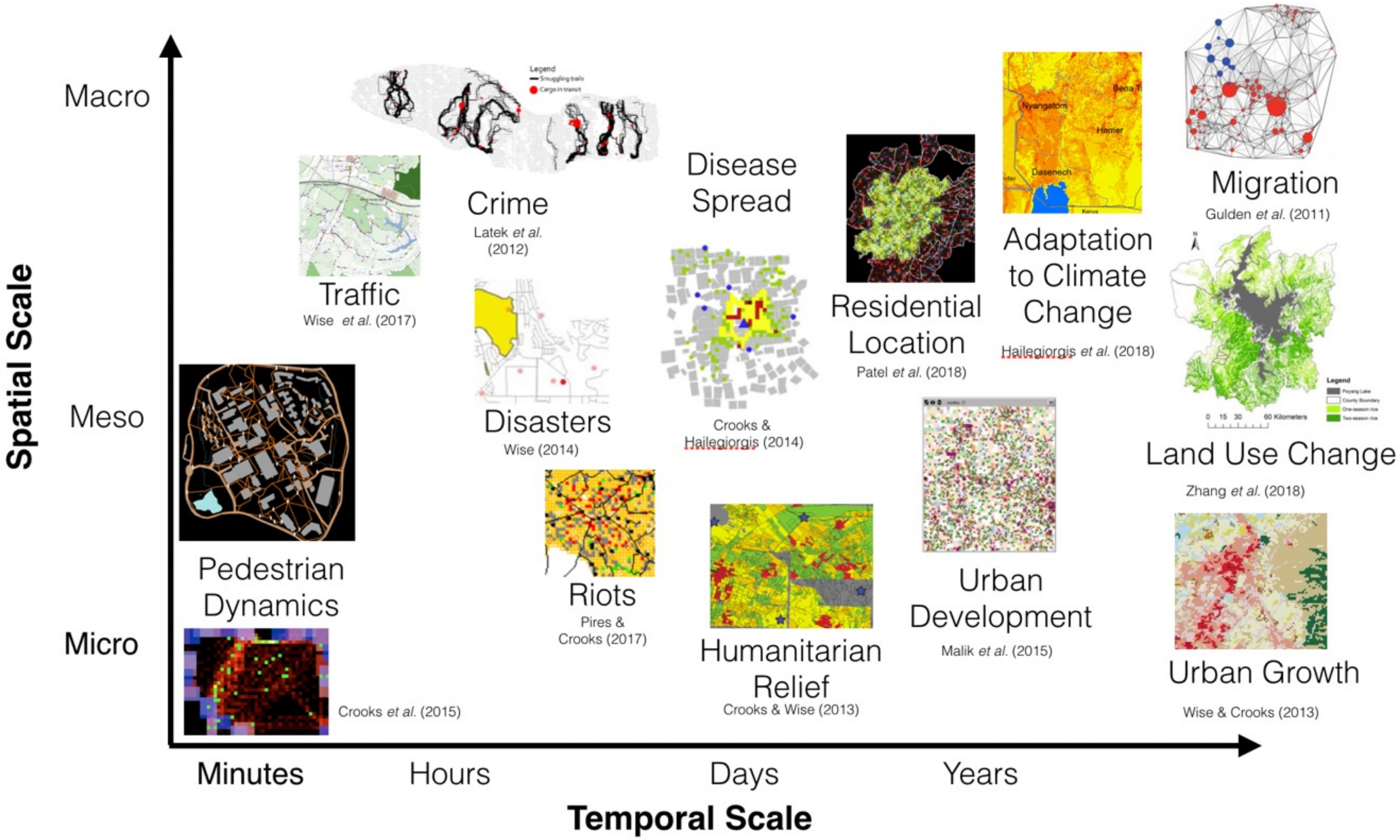
Agent-Based Modelling (ABM)

Agent-Based Modelling (ABM)

- Autonomous, interacting agents
- Represent individuals or groups
- Situated in a virtual environment



Types of problems



Crooks et al. (2019), Agent-based Modelling and Geographical Information Systems: A Practical Primer, Sage, London, UK.

ELSEVIER
Computers, Environment and Urban Systems
Volume 32, Issue 6, November 2008, Pages 417-430

Key challenges in agent-based modelling for geo-spatial simulation

Andrew Crooks, Christian Castle, Michael Batty

Show more

GeoInformatica (2019) 23:169–199
https://doi.org/10.1007/s10707-018-00340-z

CrossMark

Crossing the chasm: a 'tube-map' for agent-based social simulation of policy scenarios in spatially-distributed systems

J. Gareth Polhill[1] • Jiaqi Ge[1] • Matthew P. Hare[1] • Keith B. Matthews[1] • Alessandro Gimona[1] • Douglas Salt[1] • Jagadeesh Yeluripati[1]

Editorial: Meeting Grand Challenges in Agent-Based Models

Li An[1], Volker Grimm[2,3], Billie L. Turner II[4]

[1]Center for Complex Human-Environment Systems, Department of Geography San Diego State University, Storm Hall 303C, San Diego, CA, United States
[2]Department of Ecological Modelling, Helmholtz Centre for Environmental Research - UFZ, Permoserstr. 15, 04318 Leipzig, Germany
[3]University of Potsdam, Potsdam, Germany
[4]School of Geographical Sciences and Urban Planning, College of Liberal Arts and Sciences, Arizona State University, PO Box 875302, Tempe, AZ 85287-5302, United States
Correspondence should be addressed to lan@mail.sdsu.edu

Journal of Artificial Societies and Social Simulation 23(1) 13, 2020
Doi: 10.18564/jasss.4012 Url: http://jasss.soc.surrey.ac.uk/23/1/13.html

Received: 16-03-2019 Accepted: 05-11-2019 Published: 31-01-2020

Methodological Issues of Spatial Agent-Based Models

Steven Manson[1], Li An[2], Keith C. Clarke[3], Alison Heppenstall[4], Jennifer Koch[5], Brittany Krzyzanowski[1], Fraser Morgan[6], David O'Sullivan[7], Bryan C. Runck[8], Eric Shook[1], Leigh Tesfatsion[9]

[1]Department of Geography, Environment, and Society, University of Minnesota, 267, 19th Ave S, Minneapolis, MN 55455, United States
[2]PKU-SDSU Complex Human-Environment Systems Center and Department of Geography, San Diego State University, 5500 Campanile Dr, San Diego, CA 92182, United States
[3]Department of Geography, University of California Santa Barbara, 1720 Ellison Hall, Santa Barbara, CA 93106, United States
[4]School of Geography, University of Leeds, Leeds LS2 9JT, United Kingdom
[5]Department of Geography and Environmental Sustainability, The University of Oklahoma, 100 E Boyd St., Norman, OK 73019, United States
[6]Landscape Policy and Governance Manaaki Whenua — Landcare Research, Private Bag 92170, Auckland Mail Centre, Auckland 1142, New Zealand
[7]School of Geography Environment and Earth Sciences, Victoria University of Wellington, Wellington City 6012, New Zealand

Geographical Analysis (2021) **53**, 76–91

Special Issue

Future Developments in Geographical Agent-Based Models: Challenges and Opportunities

Alison Heppenstall[1,2], Andrew Crooks[3], Nick Malleson[1,2], Ed Manley[1,2], Jiaqi Ge[1], Michael Batty[4]

[1]School of Geography, University of Leeds, Leeds, U.K., [2]Alan Turing Institute, The British Library, London, U.K., [3]Department of Computational and Data Sciences and Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA USA, [4]Centre for Advanced Spatial Analysis (CASA), University College London, London, U.K.

J Land Use Sci. 2016; 11(2): 177–187. PMID: 27158257
Published online 2015 Apr 13. doi: 10.1080/1747423X.2015.1030463

Strategic directions for agent-based modeling: avoiding the YAAWN syndrome

David O'Sullivan,[a,*] Tom Evans,[b] Steven Manson,[c] Sara Metcalf,[d] Arika Ligmann-Zielinska,[e] and Chris Bone[f]

▸ Author information ▸ Copyright and License information Disclaimer

Presentation Overview

- Agent behaviour
- Visualisation
- Data
- Calibration / validation and uncertainty
- Real time ABM
- Computational issues
- Digital Twins

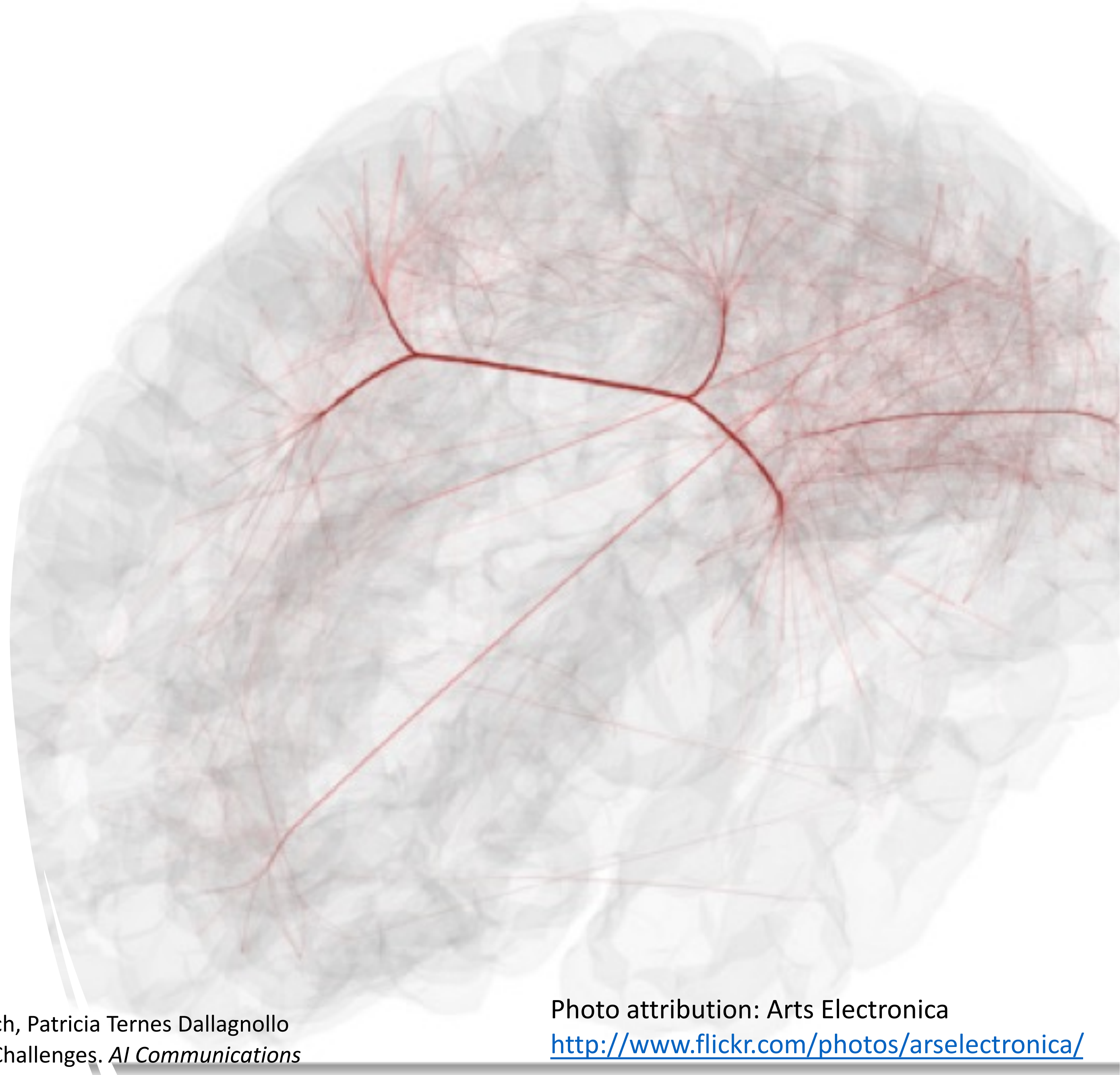


Photo attribution: Arts Electronica
http://www.flickr.com/photos/arselectronica/

Mark Birkin, Daniel Birks, Jiaqi Ge, Alison Heppenstall, Ed Manley, Josie McCulloch, Patricia Ternes Dallagnollo (forthcoming). Agent-Based Modelling for Urban Analytics: State of the Art and Challenges. *AI Communications*

Agent behaviour

Appeal of ABM: Modelling Human behaviour
Why is Behaviour so Tricky?

- Humans, rational animals? Predictable?
- Does the data contain the right processes and drivers?
- Many, many behavioural frameworks - which one?

ELSEVIER
Ecological Economics
Volume 131, January 2017, Pages 21-35

Analysis

A framework for mapping and comparing behavioural theories in models of social-ecological systems

Maja Schlüter[a], Andres Baeza[b, c], Gunnar Dressler[d], Karin Frank[d], Jürgen Groeneveld[d, e], Wander Jager[f], Marco A. Janssen[c], Ryan R.J. McAllister[g], Birgit Müller[d], Kirill Orach[a], Nina Schwarz[h], Nanda Wijermans[a]

Show more

Share Cite

https://doi.org/10.1016/j.ecolecon.2016.08.008

Get rights and content

ELSEVIER
Environmental Modelling & Software
Volume 48, October 2013, Pages 37-48

Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol

Birgit Müller[a], Friedrich Bohn[a], Gunnar Dreßler[a], Jürgen Groeneveld[a, f], Christian Klassert[c], Romina Martin[a], Maja Schlüter[d, e], Jule Schulze[a, b], Hanna Weise[a], Nina Schwarz[b]

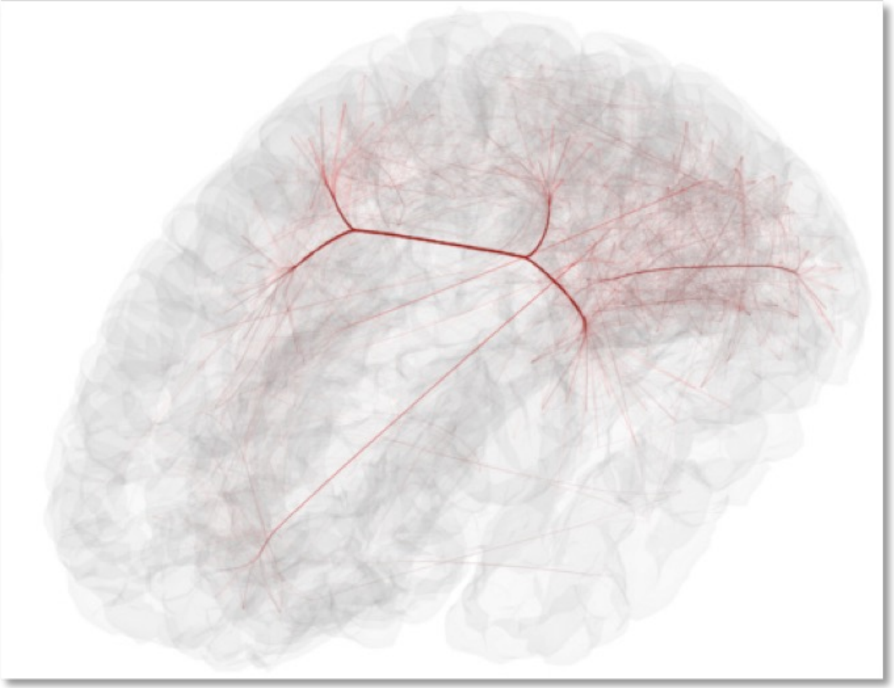
Show more

Share Cite

https://doi.org/10.1016/j.envsoft.2013.06.003

Get rights and content

Appeal of ABM: Modelling Human behaviour



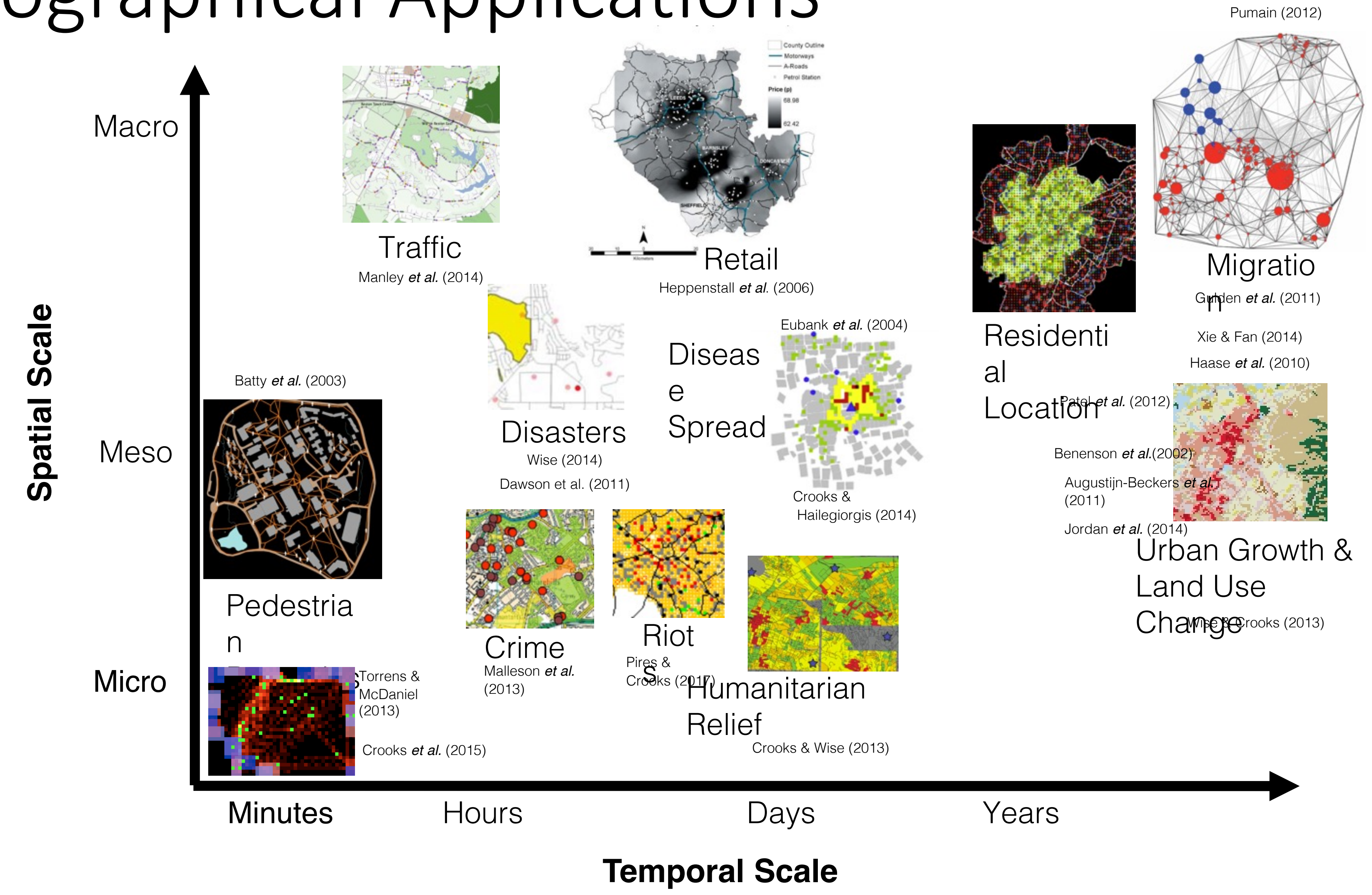
Quiz

- Choose a number between 1 and 4
- Write the number down on a slip of paper
- What percentage of people do you think chose:
  - 1?
  - 2?
  - 3?
  - 4?

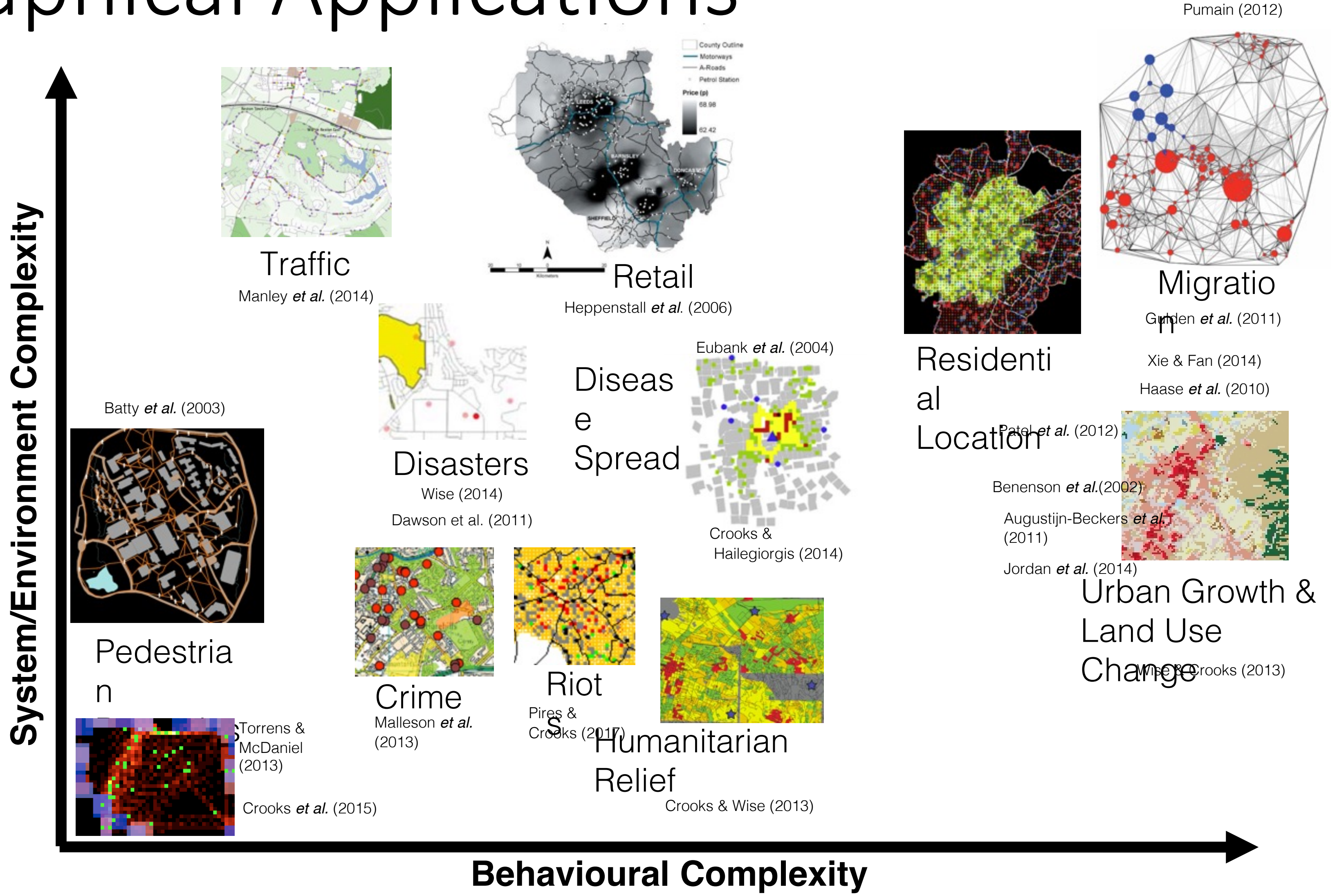
Appeal of ABM: Modelling Human behaviour

"The most common response is **'three'** and there is a secondary effect of this task: people feel a need to explain why they chose whatever answer they did. The second most common answer is **'two'**. Very few people decide to respond with either **'one'** or **'four'**. Sadly, there is not a serious study of this behaviour but undocumented sources suggest that the response statistics are close to 50% for 'three', 30% for 'two' and about 10% for the other two answers." (Kennedy, 2012)

Geographical Applications

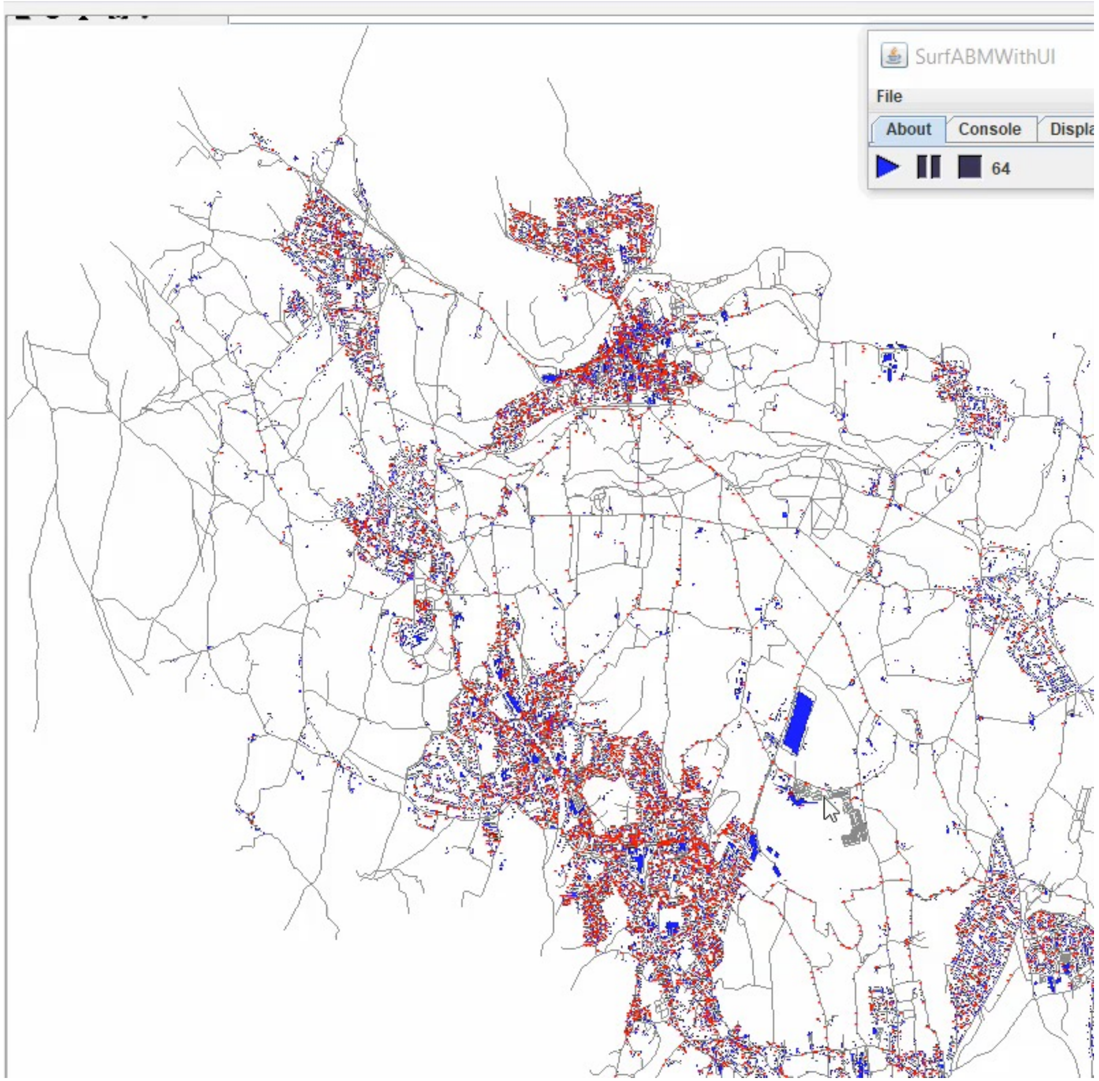


Geographical Applications

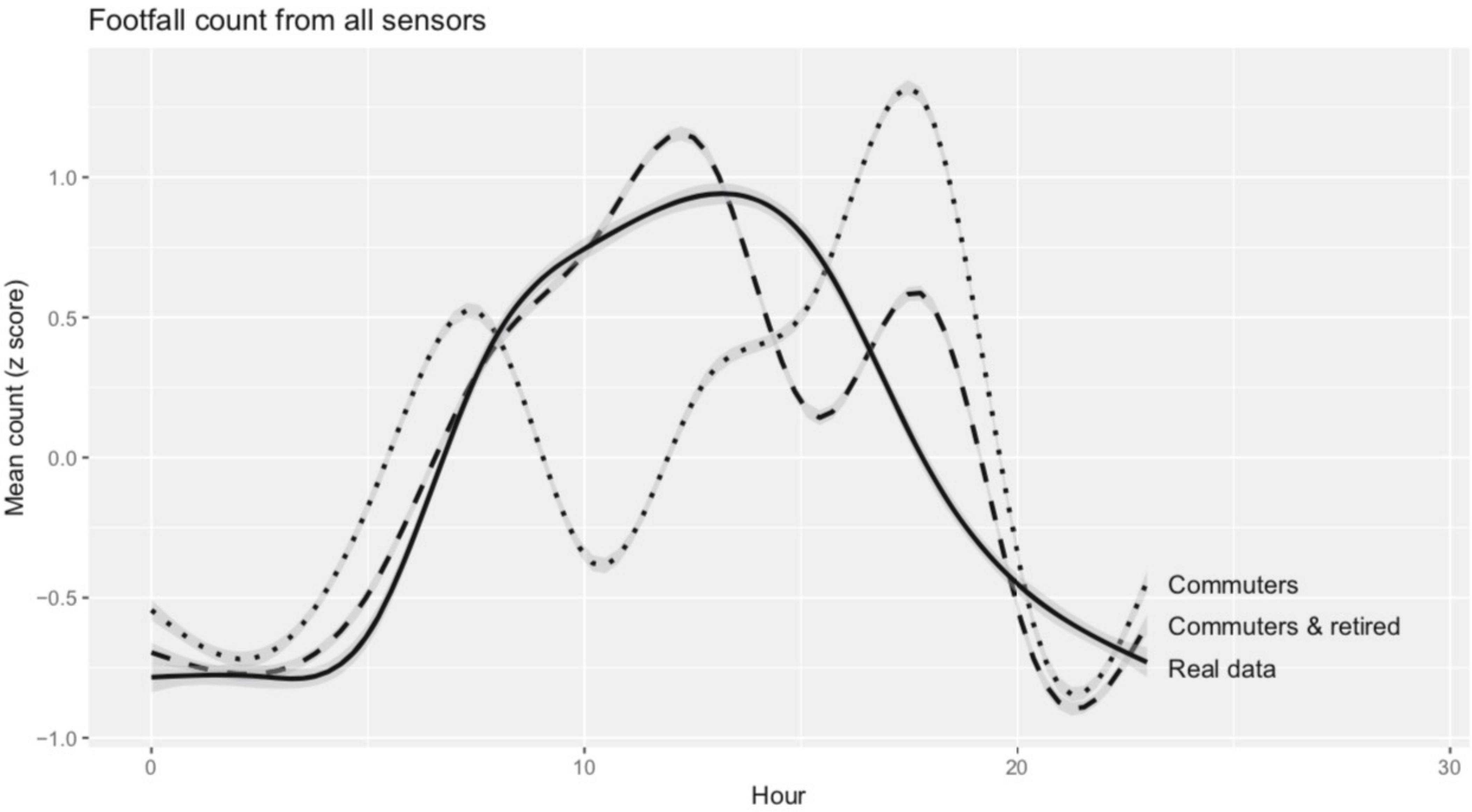


Simulating pedestrian movement

- Can we use data sources to create an accurate picture of how people move around an urban space?
- Use Census to create population
- Use Time/Work survey to put in basic behaviour (commuting)
- Put them in houses and watch them go
- Calibrate against sensor information



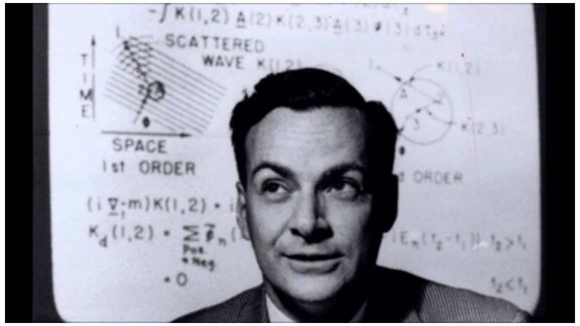
Crols, T. and Malleson, N. (2019) Quantifying the ambient population using hourly population footfall data and an agent-based model of daily mobility. GeoInformatica, 23: 201-220



Crols, T. and Malleson, N. (2019) Quantifying the ambient population using hourly population footfall data and an agent-based model of daily mobility. GeoInformatica, 23: 201-220

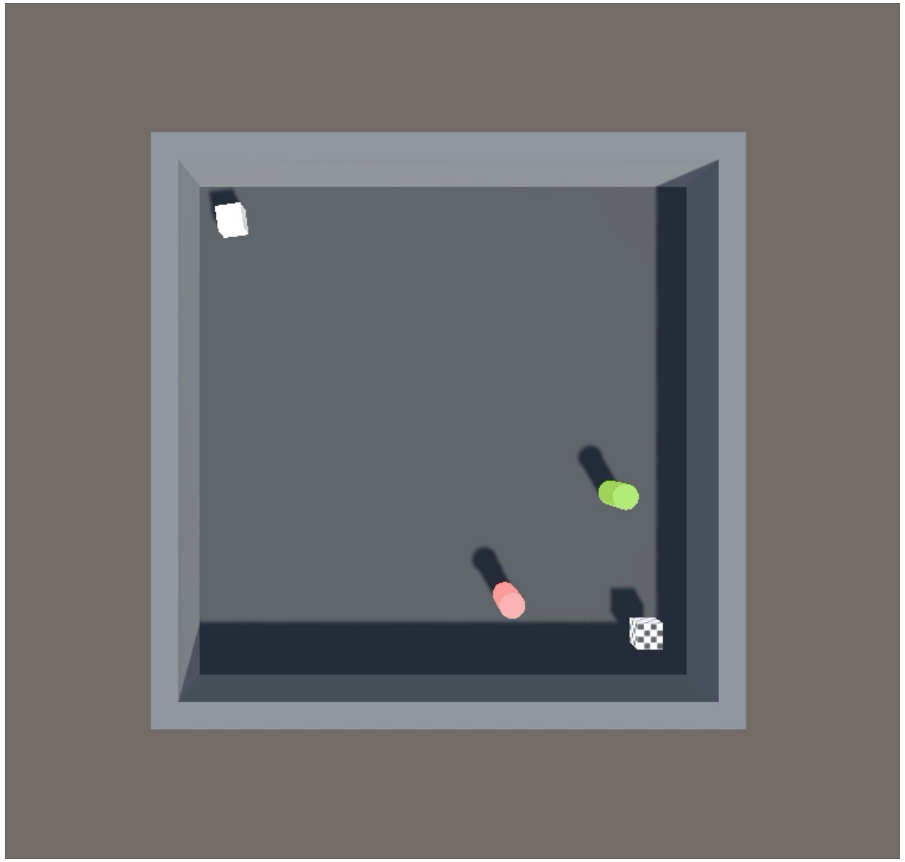
Can we get behaviour right?

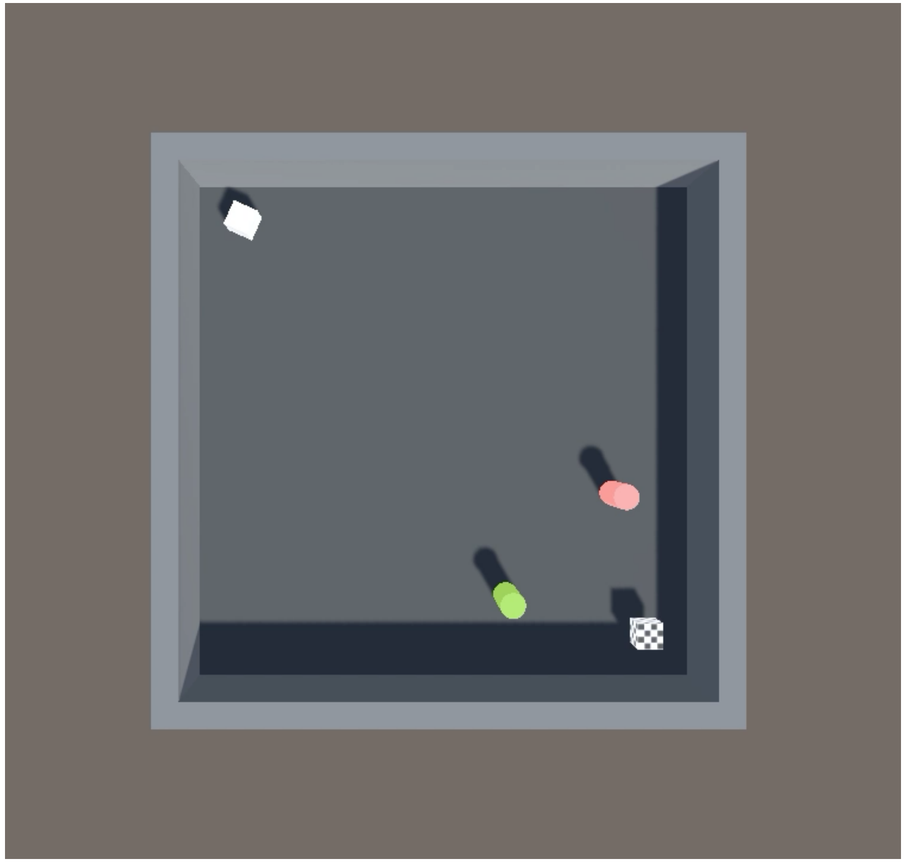
- Behavioural rules often drawn from historical data
- Need rich, individual-level data
  - Contain all events/experiences, results of feedback
  - How extract behavioural rules from qualitative data?
  - Assumptions (rationale, knowledge…)



“Imagine how much harder physics would be if electrons had feelings!”

Acknowledgement: Ed Manley



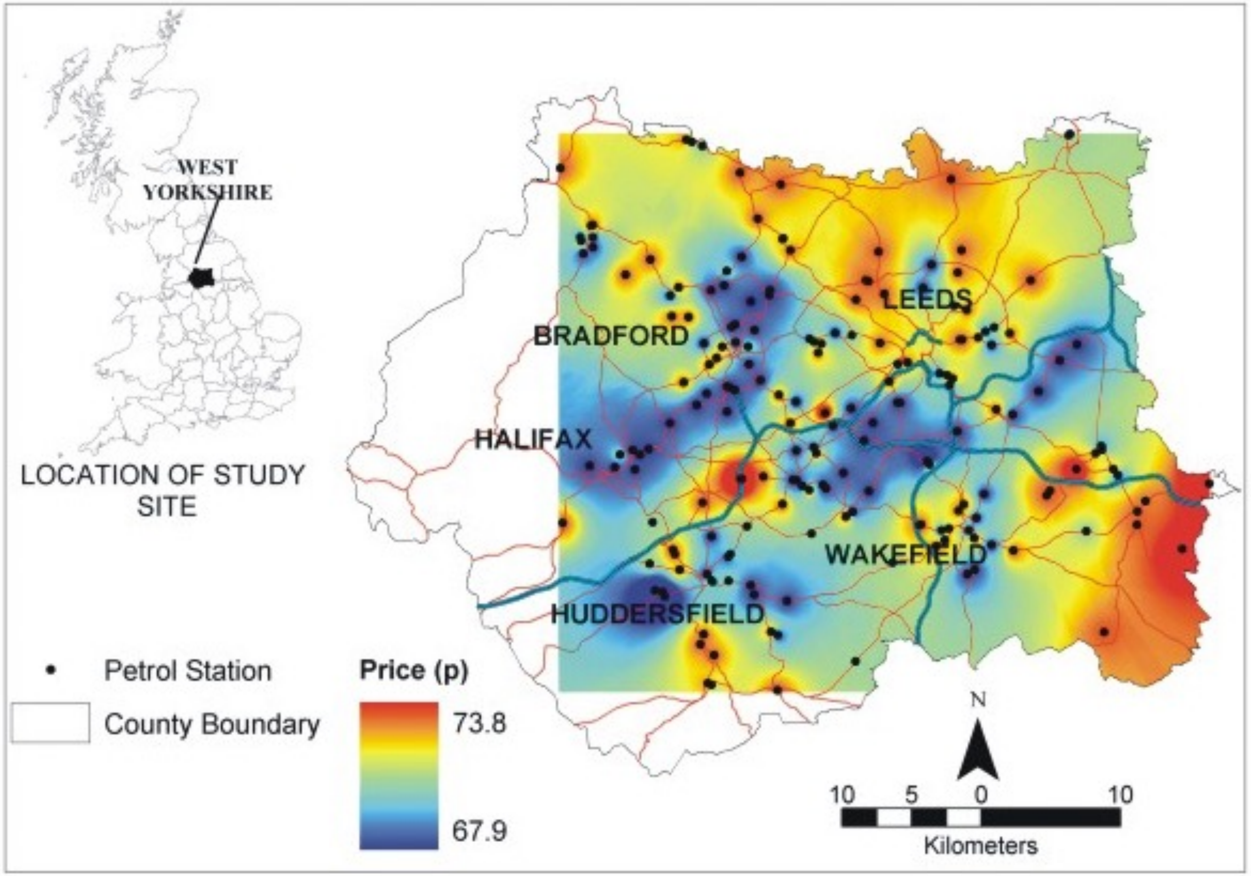


Trained Navigating Agent
Using perspective visual inputs to navigate
Landmarks help guide way to target

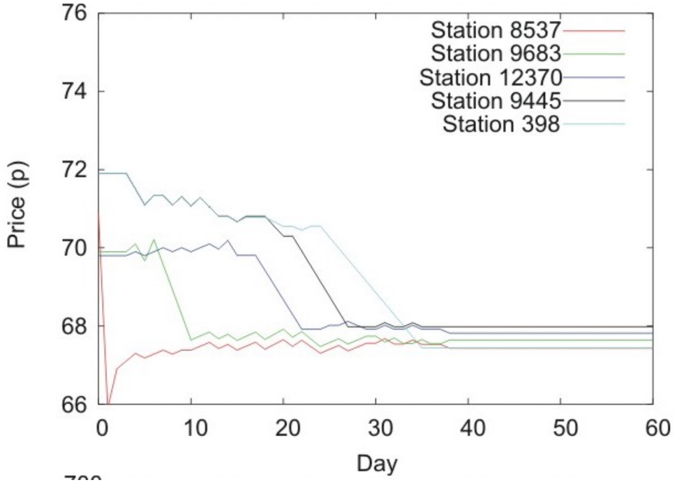
Confused Agent
Landmarks switched
Finding target difficult for agent

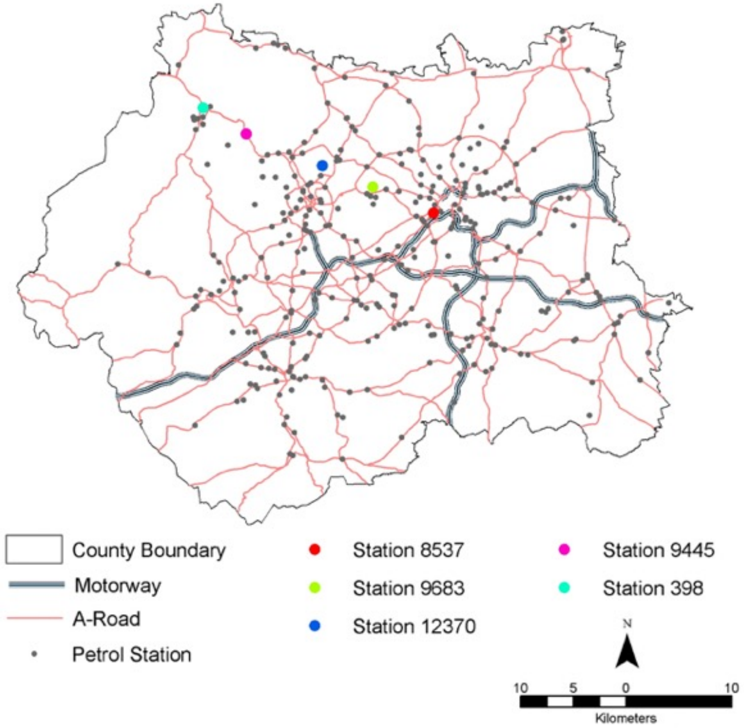
Olmez, S., **Heppenstall, A.,** Birks, D. (2021) Investigating the emergence of complex behaviours in an agent-based model using reinforcement learning. *JASSS*

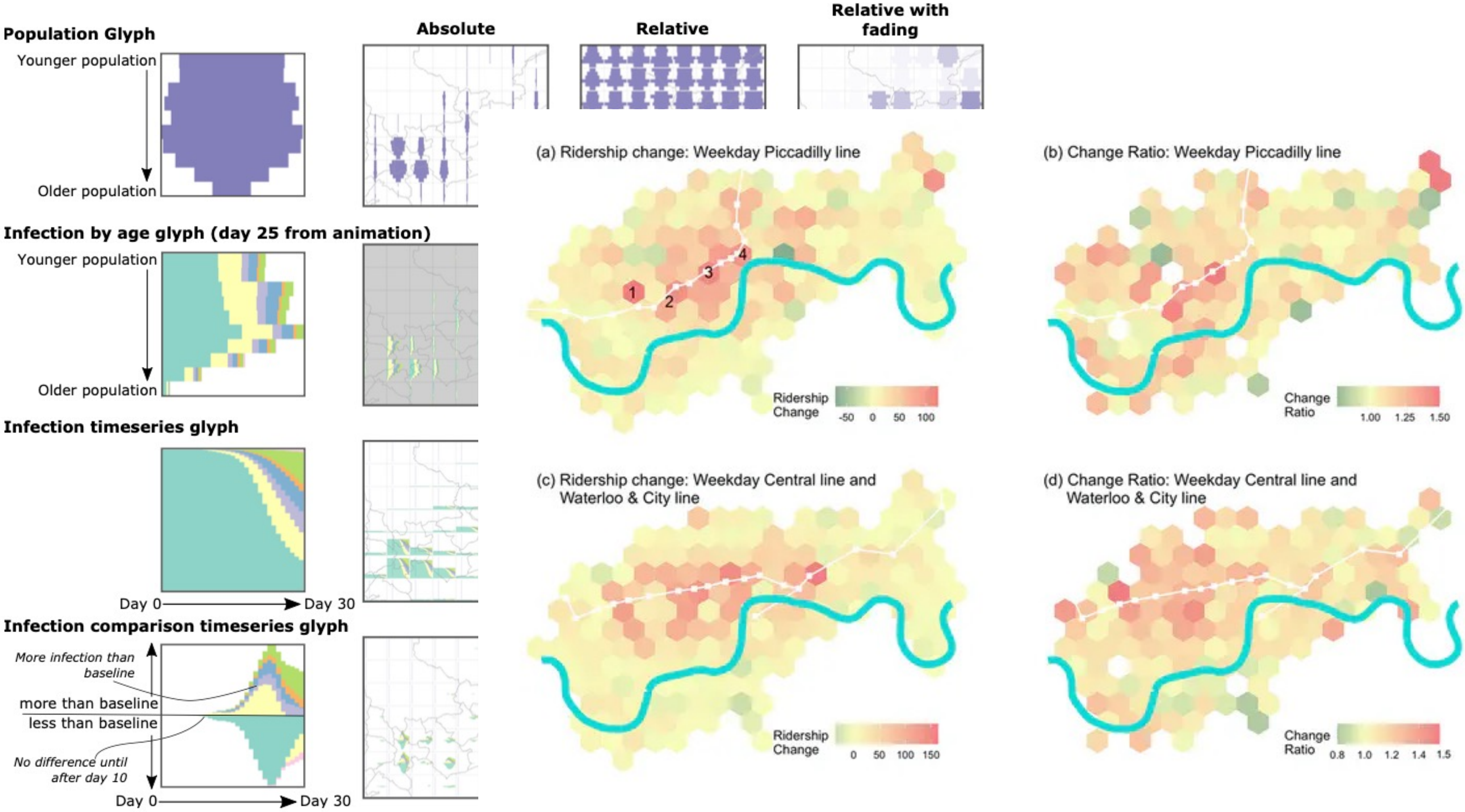
Visualisation



Heppenstall et al (2006); https://www.jasss.org/9/3/2.html







Acknowledgement:
Aidan Slingsby

Yuanxuan Yang et al. (2022).*Journal of Transport Geography*.

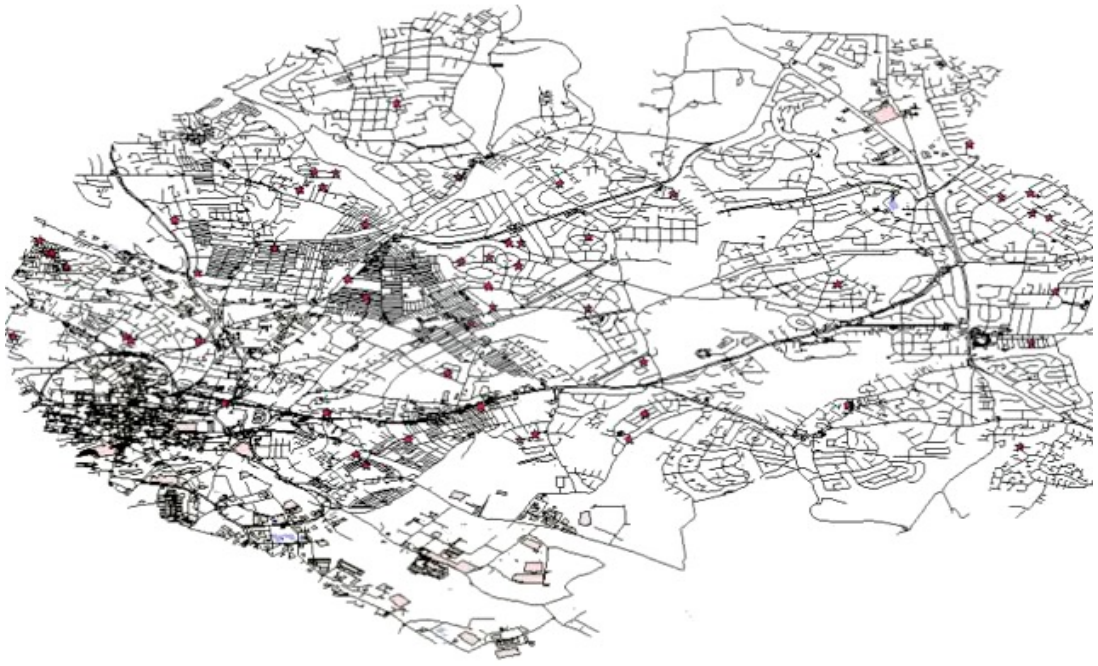
Data

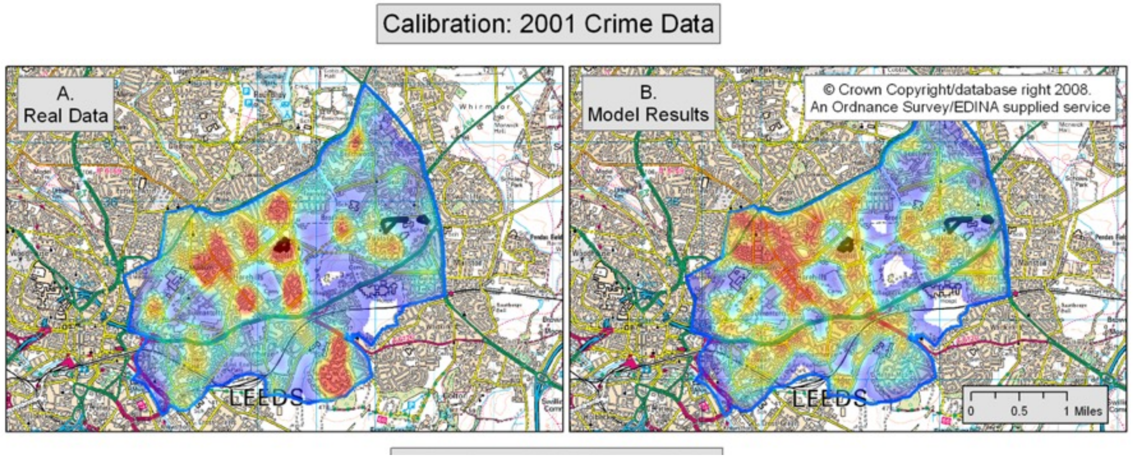
Data

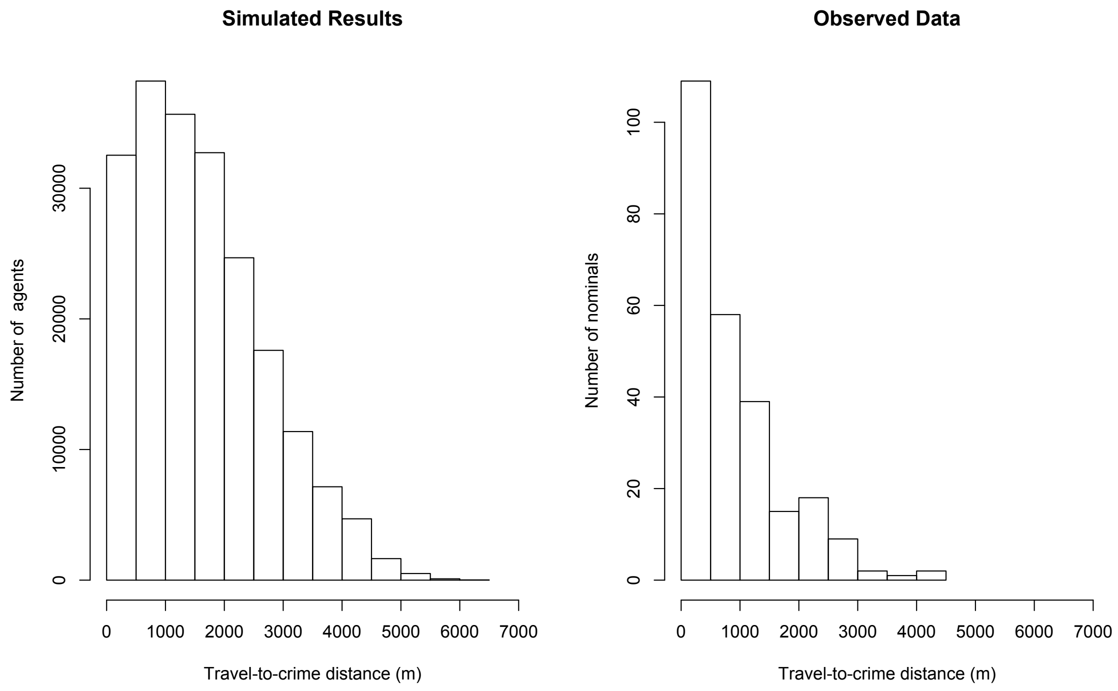
- ABMs are typically very rich (high spatio-temporal resolution)
- But data are often much coarser (usually highly aggregated)
- Difficulties:
    - Identifiability
    - Multi-level validation

Example: right pattern, wrong agents..

- Pattern Oriented Modelling (POM: Grimm et al., 2005)
  - Should evaluate model at multiple scales
  - But not possible with limited data



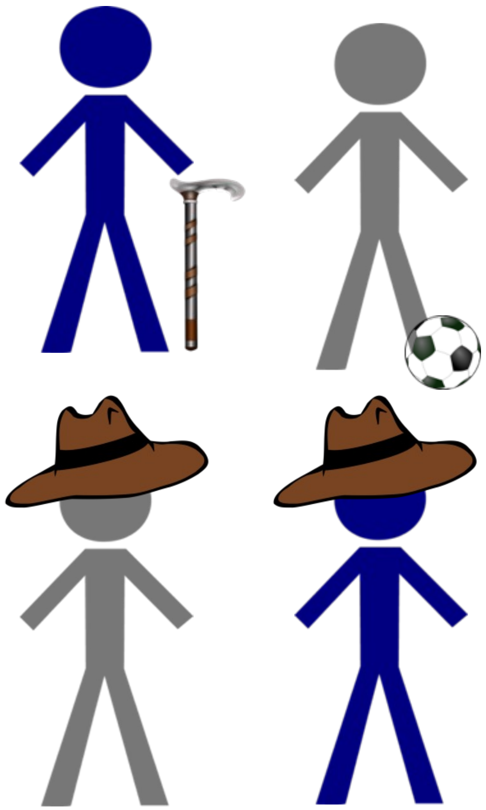




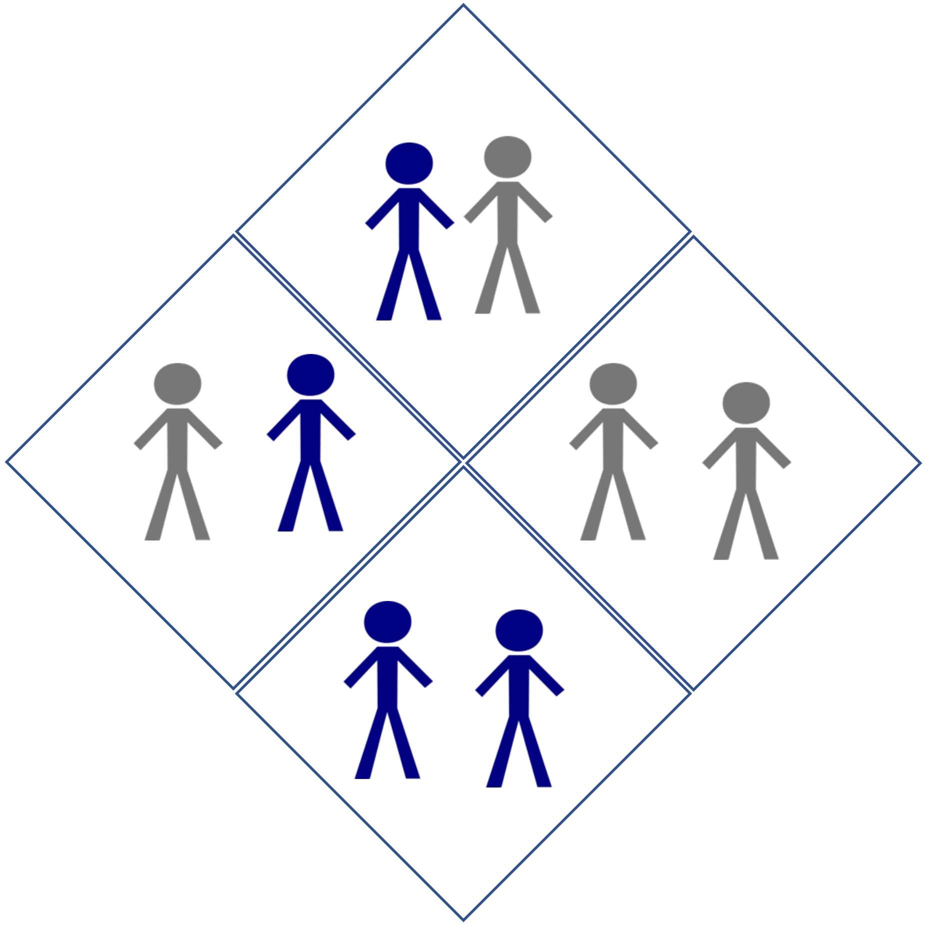
Malleson, N., L. See, A. Evans, and A. Heppenstall (2012). Implementing comprehensive offender behaviour in a realistic agent-based model of burglary. SIMULATION 88(1) 50-71

‘Types’ of Microsimulation (1) **Creating Synthetic Data**

Sample or survey data



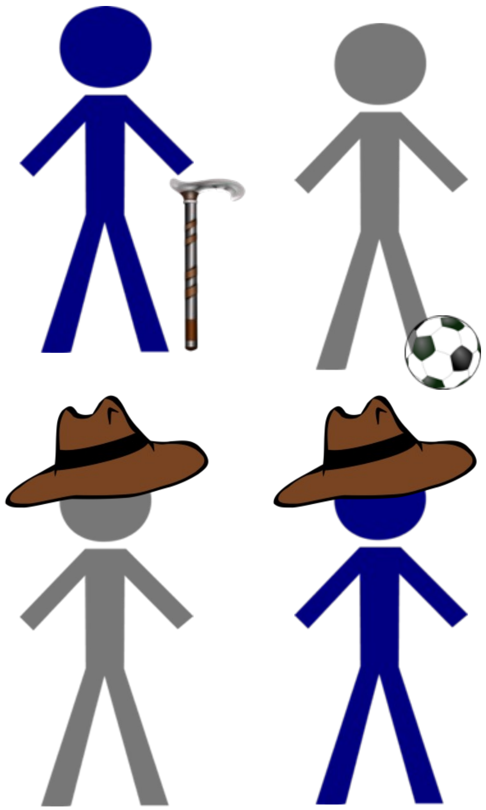
Target or constraining data



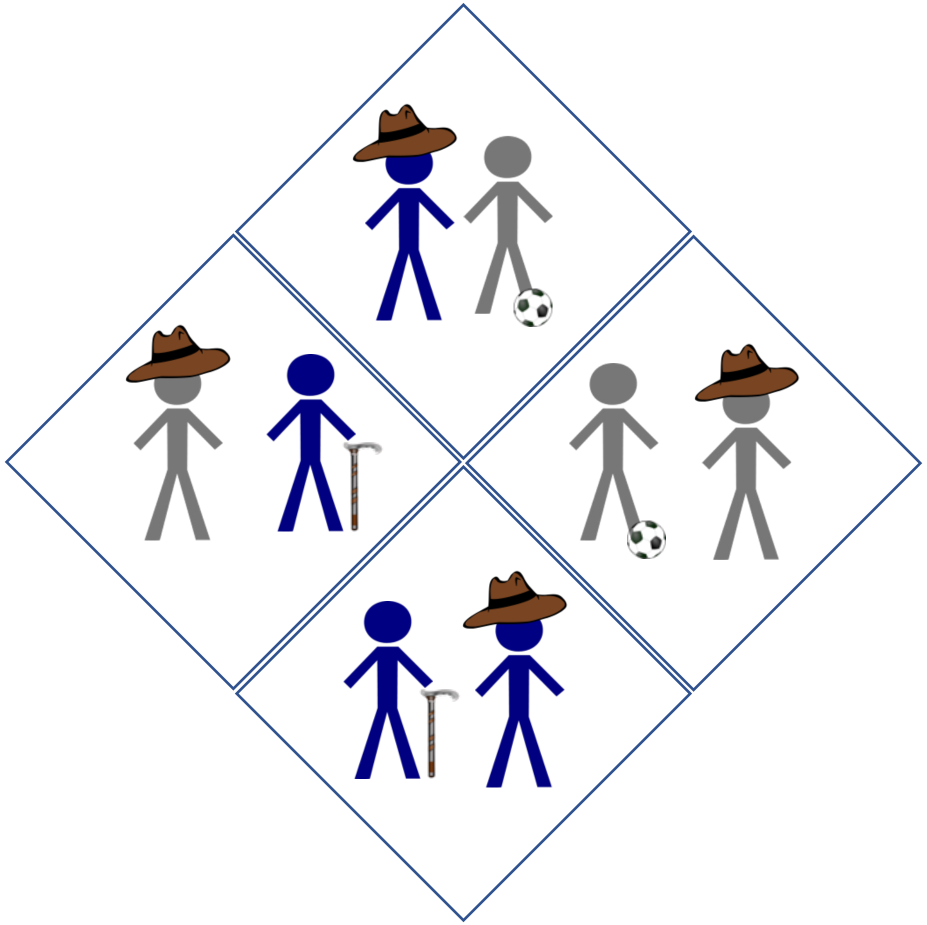
Acknowledgement: Nik Lomax

‘Types’ of Microsimulation (1) **Creating Synthetic Data**

Sample or survey data



Target or constraining data



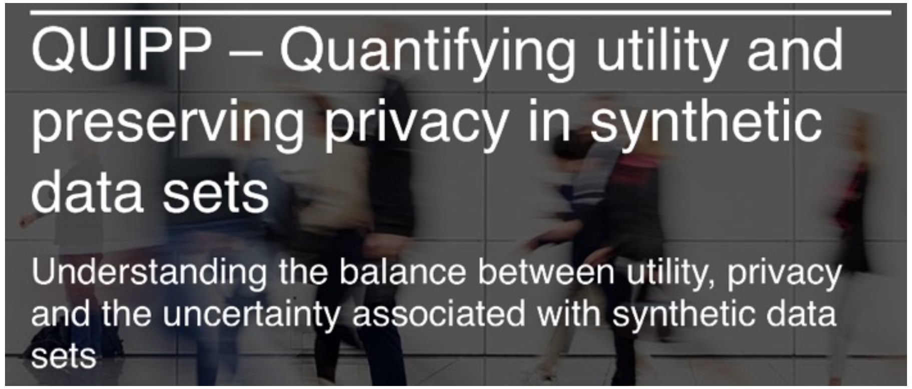
Adding geography as a constraint makes this *spatial* microsimulation

Acknowledgement: Nik Lomax

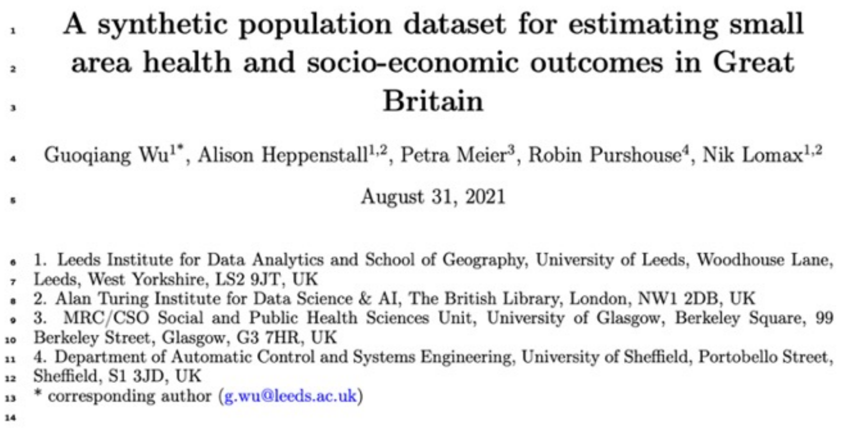
Population synthesis

- Population synthesis
  - Data: census + survey data e.g. UK ESRC Understanding Society
- But what if we can't get hold of the good stuff?

The Alan Turing Institute



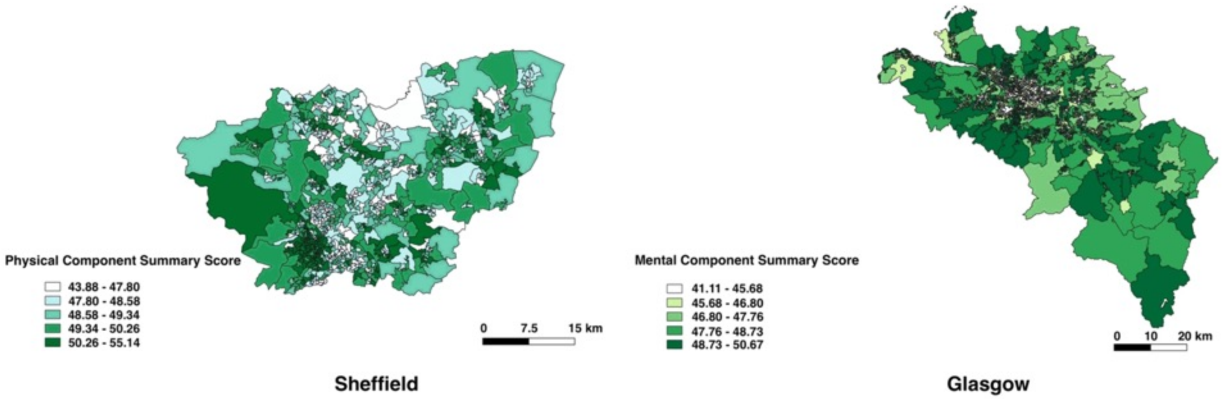
Systems science In Public Health and Health Economics Research


A synthetic population dataset for estimating small area health and socio-economic outcomes in Great Britain

Guoqiang Wu[1*], Alison Heppenstall[1,2], Petra Meier[3], Robin Purshouse[4], Nik Lomax[1,2]

August 31, 2021

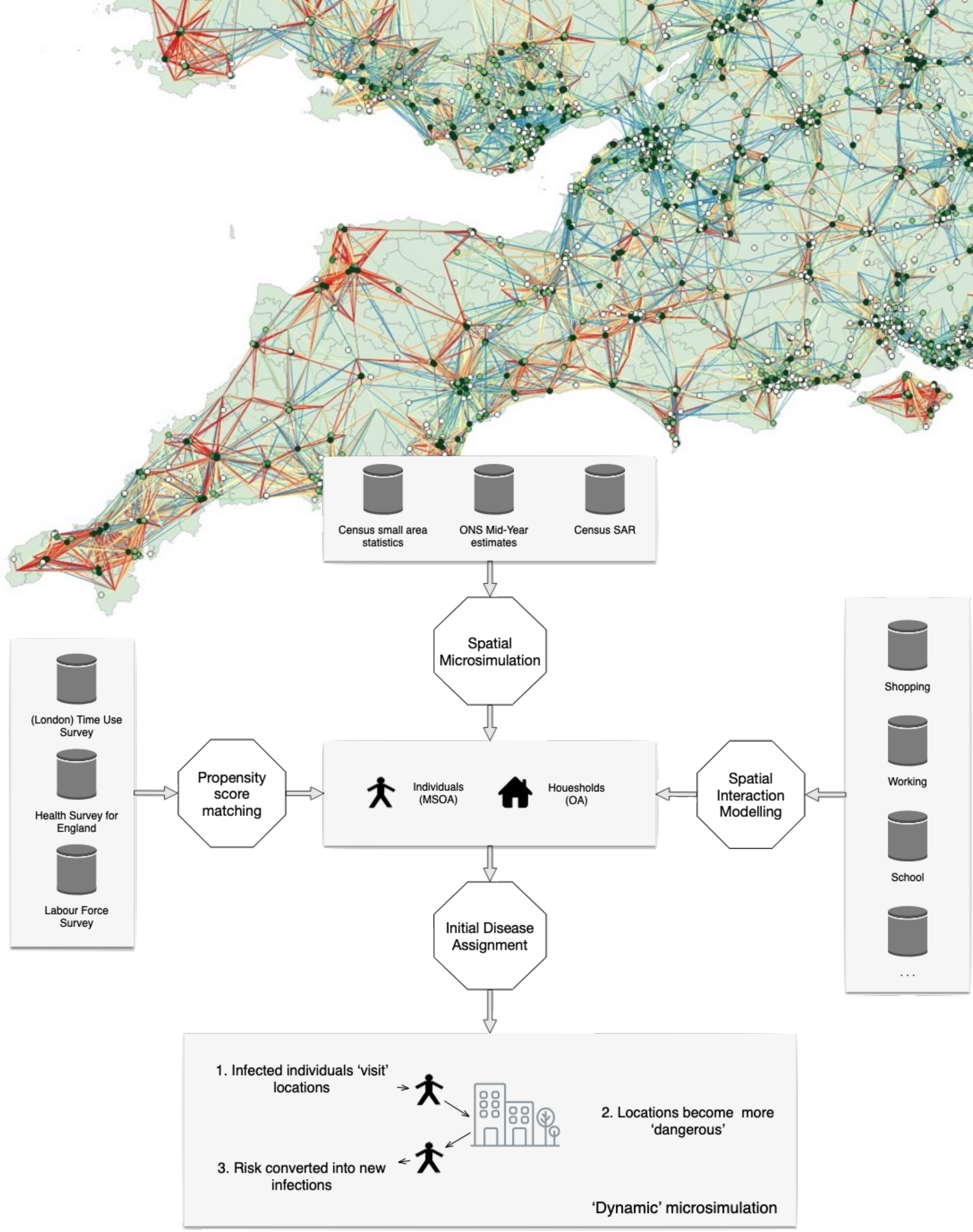
1. Leeds Institute for Data Analytics and School of Geography, University of Leeds, Woodhouse Lane, Leeds, West Yorkshire, LS2 9JT, UK
2. Alan Turing Institute for Data Science & AI, The British Library, London, NW1 2DB, UK
3. MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, Berkeley Square, 99 Berkeley Street, Glasgow, G3 7HR, UK
4. Department of Automatic Control and Systems Engineering, University of Sheffield, Portobello Street, Sheffield, S1 3JD, UK
* corresponding author (g.wu@leeds.ac.uk)



https://www.turing.ac.uk/research/research-projects/quipp-quantifying-utility-and-preserving-privacy-synthetic-data-sets

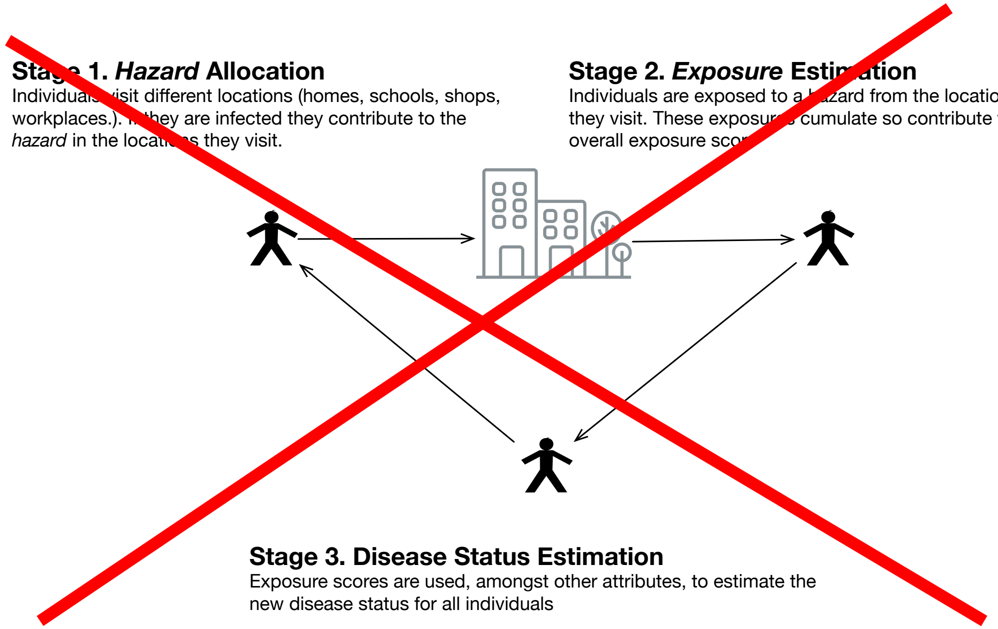
Example: Dynamic Model for Epidemics (DyME)

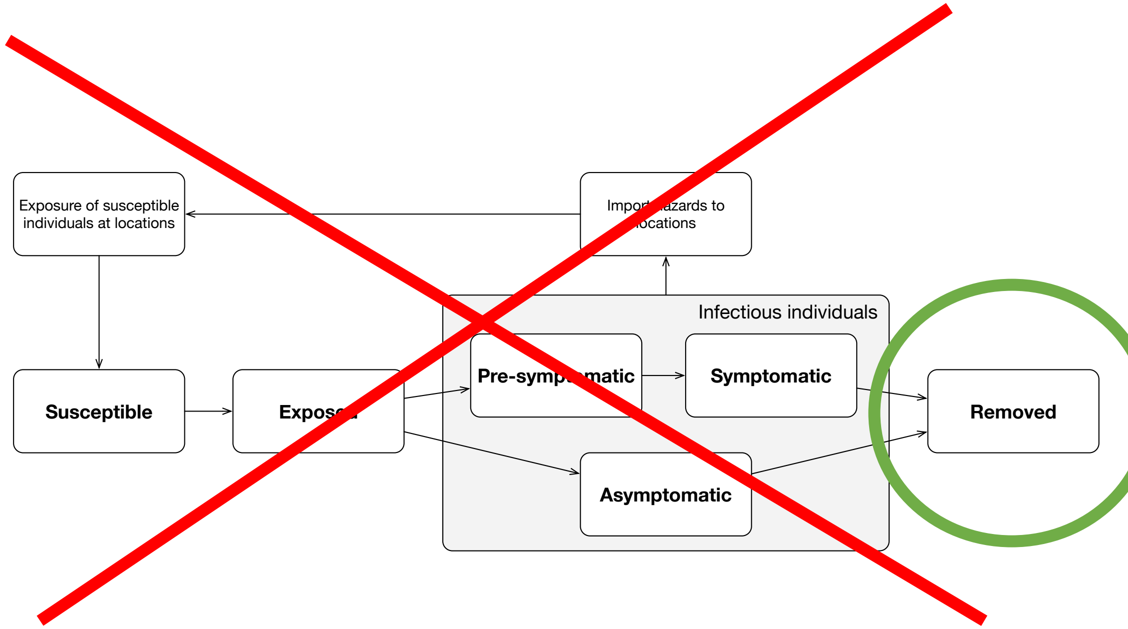
- Spooner et al. (2021)
- Part of the Royal Society Rapid Assistance in Modelling the Pandemic (RAMP) call
- COVID transmission model including dynamic spatial microsimulation, spatial interaction model, data linkage, ….
- Represents all individuals in a study area with activities: *home, shopping, working, schooling*
- Daily timestep



Dynamic Model for Epidemics (DyME)

- Incredible detailed model!
- Only data available for validation: COVID cases and hospital deaths
- Only quantify a tiny part of the transmission dynamics
- Modelling was the easy part …
- No ‘solution’, but better use of the available data might help ….





Calibration / validation and uncertainty

Calibration, Validation, and Uncertainty

- Essential, particularly for policy implications
- Draw on **Uncertainty Quantification** for more **efficient calibration** and for quantifying **understanding uncertainty:**
- History Matching to quantify uncertainties and rule out implausible parameter ranges
- Approximate Bayesian Computation to find suitable parameter distributions

JASSS
THE JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION
JASSS is an interdisciplinary journal for the exploration and understanding of social processes by means of computer simulation

ENHANCED BY Google

Homepage Journal information Journal statistics Journal Content Contact us

Home > 25 (2), 1

Calibrating Agent-Based Models Using Uncertainty Quantification Methods

Josie McCulloch[a], Jiaqi Ge[a], Jonathan A. Ward[b], Alison Heppenstall[c], J. Gareth Polhill[d] and Nick Malleson[e]
[a]University of Leeds, United Kingdom; [b]School of Mathematics, University of Leeds, United Kingdom; [c]University of Glasgow, United Kingdom; [d]The James Hutton Institute, United Kingdom; [e]School of Geography, University of Leeds, United Kingdom

Other articles by these authors

Journal of Artificial Societies and Social Simulation **25** (2) 1
<https://www.jasss.org/25/2/1.html>
DOI: 10.18564/jasss.4791 Save citation...

Received: 19-May-2021 Accepted: 29-Jan-2022 Published: 31-Mar-2022

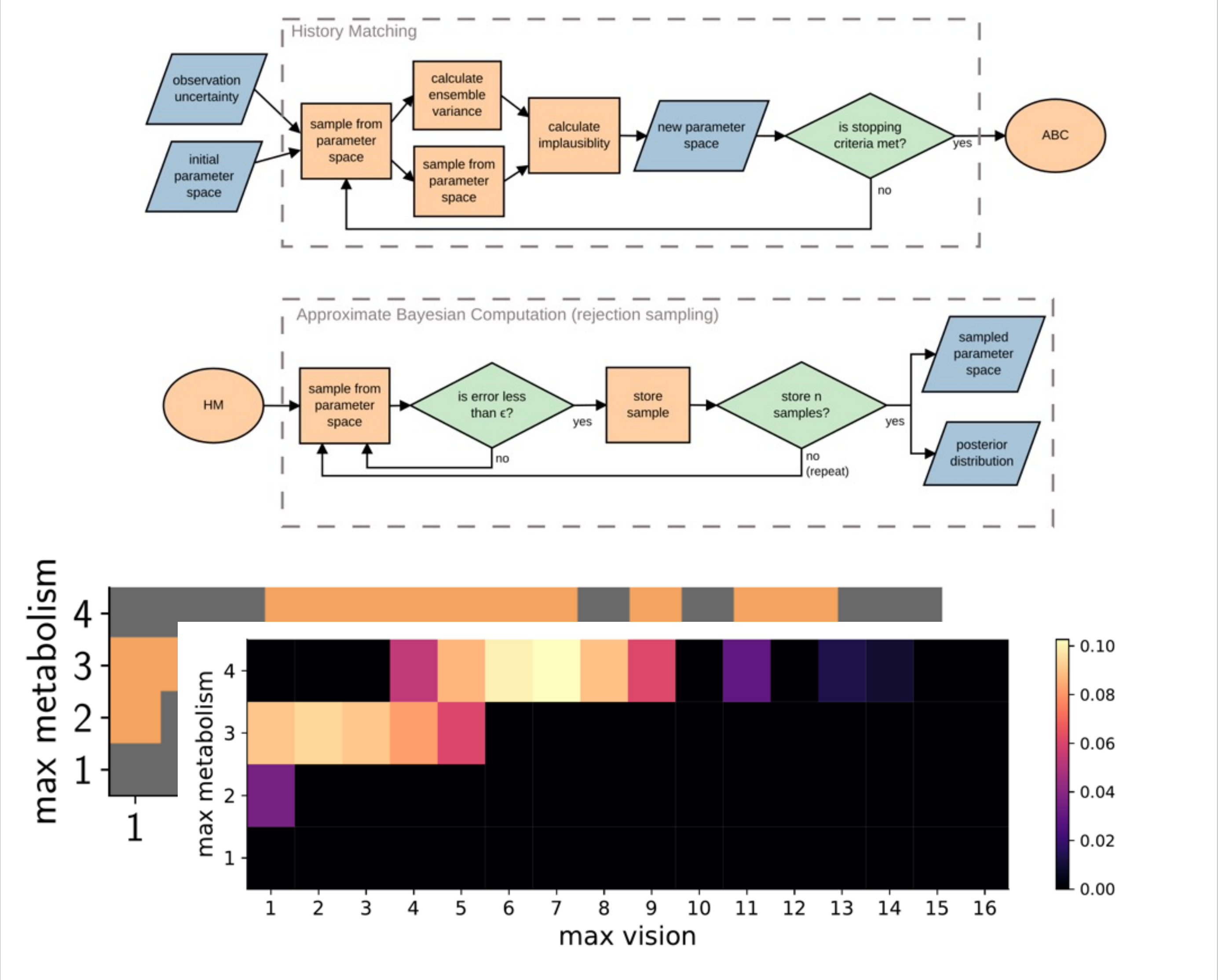
Abstract

Agent-based models (ABMs) can be found across a number of diverse application areas ranging from simulating consumer behaviour to infectious disease modelling. Part of their popularity is due to their ability to simulate individual behaviours and decisions over space and time. However, whilst there are plentiful examples within the academic literature, these models are only beginning to make an impact within policy areas. Whilst frameworks such as NetLogo make the creation of ABMs relatively easy, a number of key methodological issues, including the quantification of uncertainty, remain. In this paper we draw on state-of-the-art approaches from the fields of uncertainty quantification and model optimisation to describe a novel framework for the calibration of ABMs using History Matching and Approximate Bayesian Computation. The utility of the framework is demonstrated on three example models of increasing complexity: (i) Sugarscape to illustrate the approach on a toy example; (ii) a model of the movement of birds to explore the efficacy of our framework and compare it to alternative calibration approaches and; (iii) the RISC model of farmer decision making to demonstrate its value in a real application. The results highlight the efficiency and accuracy with which this approach can be used to calibrate ABMs. This method can readily be applied to local or national-scale ABMs, such as those linked to the creation or tailoring of key policy decisions.

Abstract
Introduction
Background
Uncertainty and agent based models
Calibration of agent-b models
Approximate Bayesian Computation (ABC)
History Matching
Methods
History Matching (HM
Approximate Bayesian Computation (ABC)
A framework for robus validation: SugarScape example
Define the parameter space to be explored
Quantify all uncertaint the model and observ
Run HM on the param space
Run ABC, using the HM results as a uniform pr
A Step-by-Step Example: SugarSc
Define the parameter space to be explored
Quantify all uncertaint in the model and observation
Model discrepancy
Ensemble variance
Observation uncertain
Run HM on the param space
Run ABC, using the HM results as a uniform p
Experiments and

Calibration, Validation, and Uncertainty

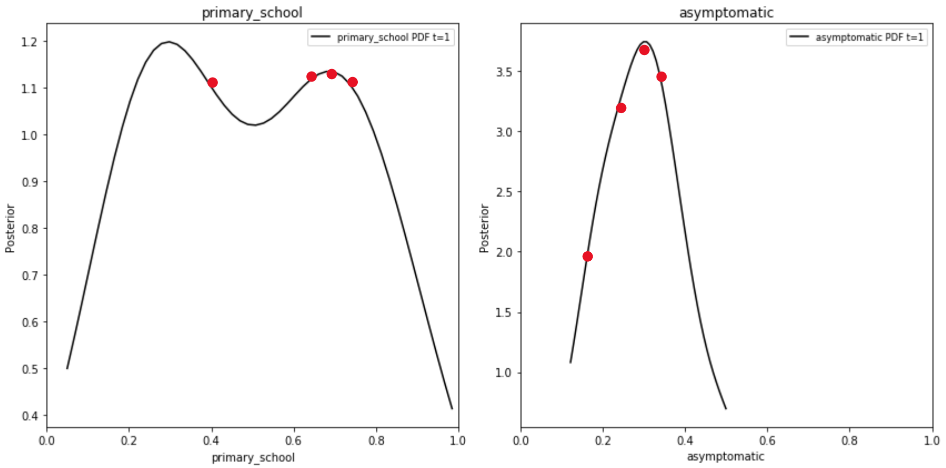
1. Define parameter space to be explored
2. Quantify uncertainties:
   - Model discrepancy (how well the model outcomes reflect the data)
   - Ensemble variance (how much the model varies with the same parameter values)
   - Observation uncertainty
3. Run History Matching to identify implausible parameter regions
4. Run Approximate Bayesian Computation, using uniform priors from HM



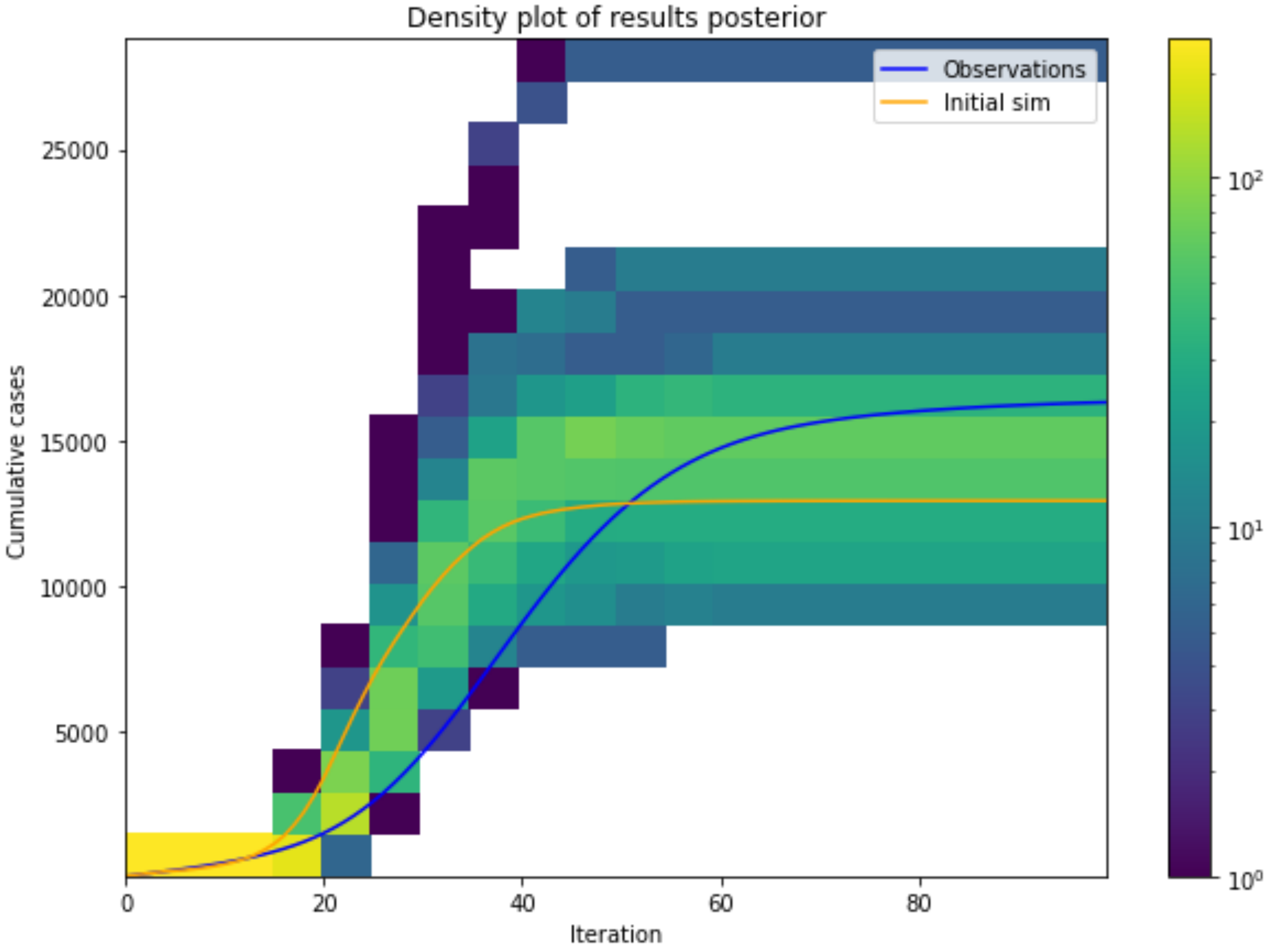
Source: McCulloch (2022): https://www.jasss.org/25/2/1.html

ABC for Uncertain Predictions

- ABC estimates a posterior

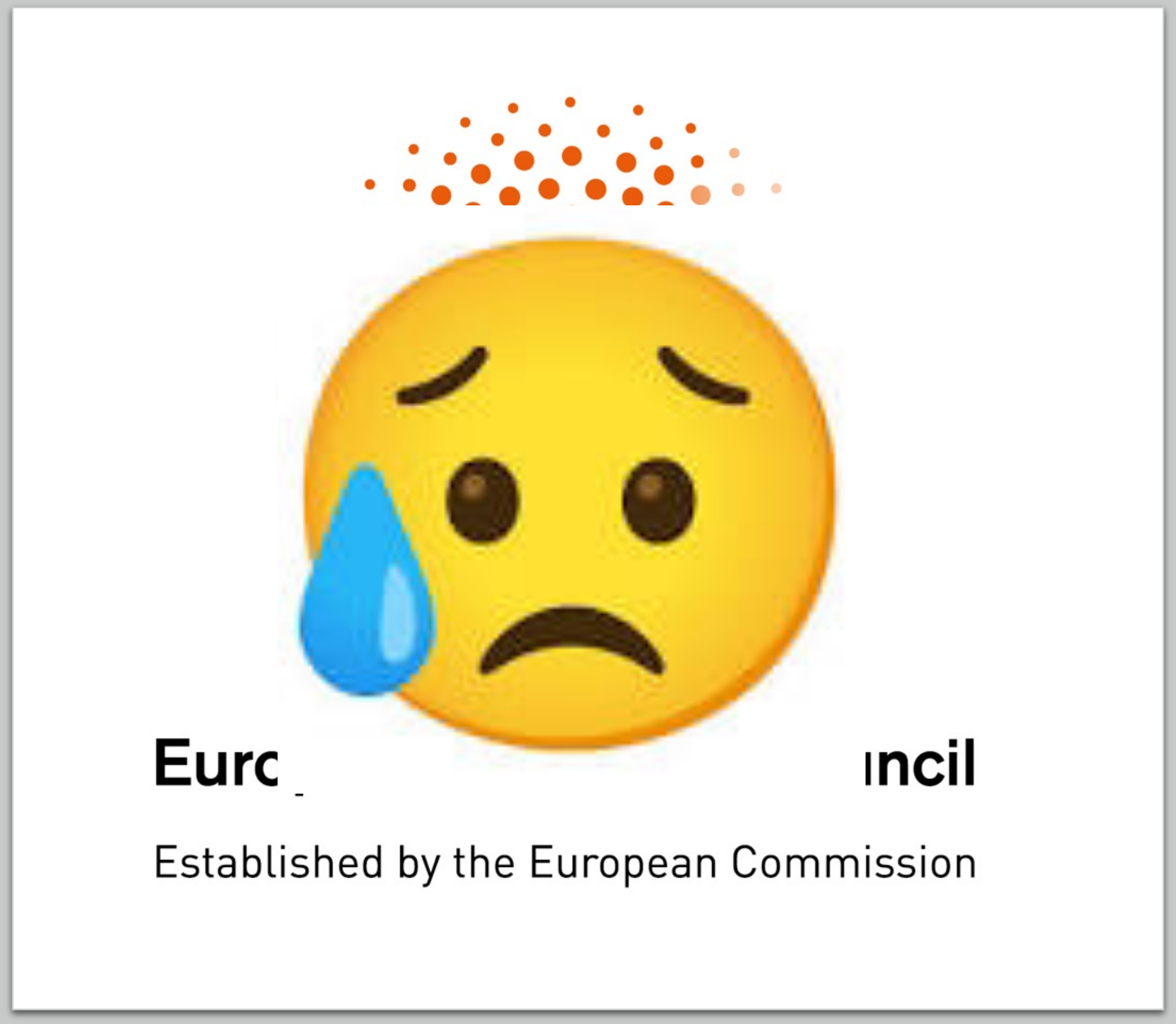


- Can sample from the posterior to make predictions



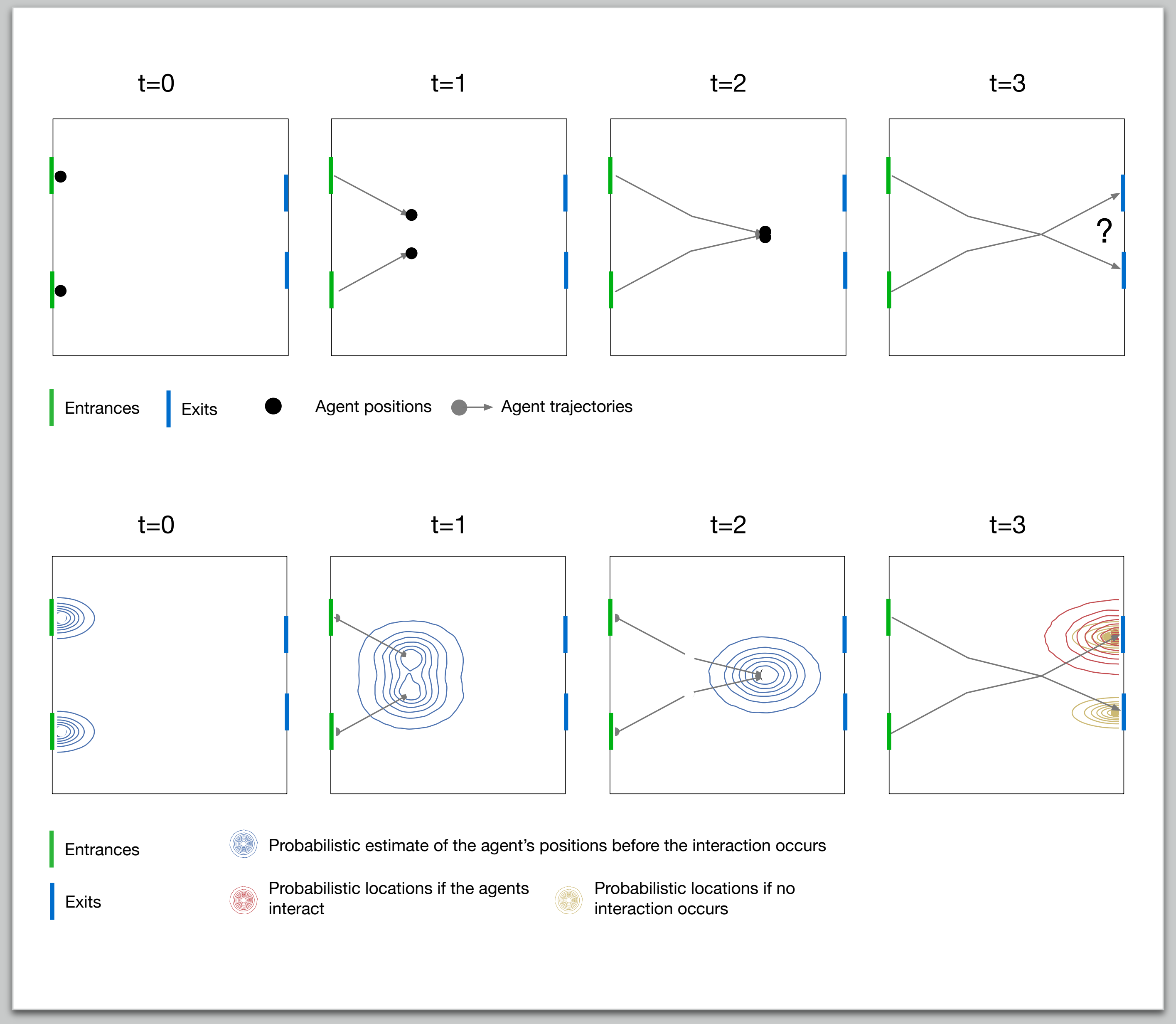
(Towards) Probabilistic Agent-Based Modelling

- Rather than running an ABM thousands or millions of times to explore its uncertainties, can we treat agents as fundamentally probabilistic?
- Instead of representing agents as points, represent them as probability distributions.
- Loads of questions about how this would work and what would happen, but might be worth exploring…



(Towards) Probabilistic Agent-Based Modelling

- Rather than running an ABM thousands or millions of times to explore its uncertainties, can we treat agents as fundamentally probabilistic?
- Instead of representing agents as points, represent them as probability distributions.
- Loads of questions about how this would work and what would happen, but might be worth exploring…



Real-Time Model Updating

Data Assimilation for Agent-Based Models

erc
European Research Council
Established by the European Commission

- We know that models will diverge from reality due to uncertainty in:
  - Parameters
  - Input data
  - Model structure, etc.

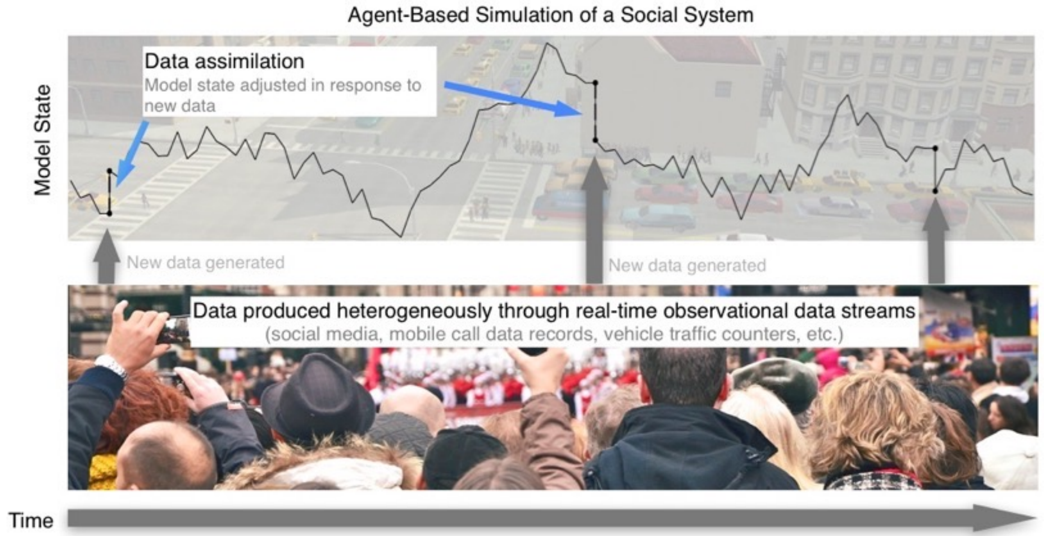
- Need a way to update the model state in response to new data

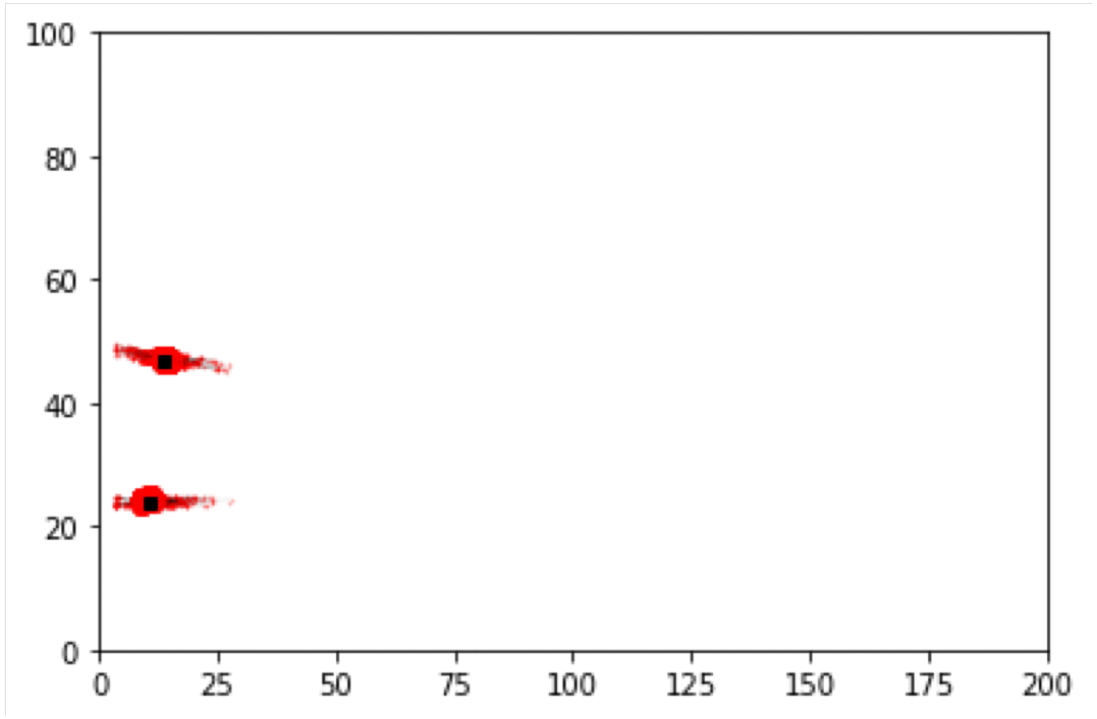
Data Assimilation

- Used in meteorology and hydrology to constrain models to reality.
- Assumptions:
  - Data have relatively low uncertainty, but are sparse
  - Models are detailed, but uncertain
- Try to improve estimates of the true system state by combining:
  - Noisy, real-world observations
  - Model estimates of the system state
- Should be more accurate than data / observations in isolation.
- https://dust.leeds.ac.uk/

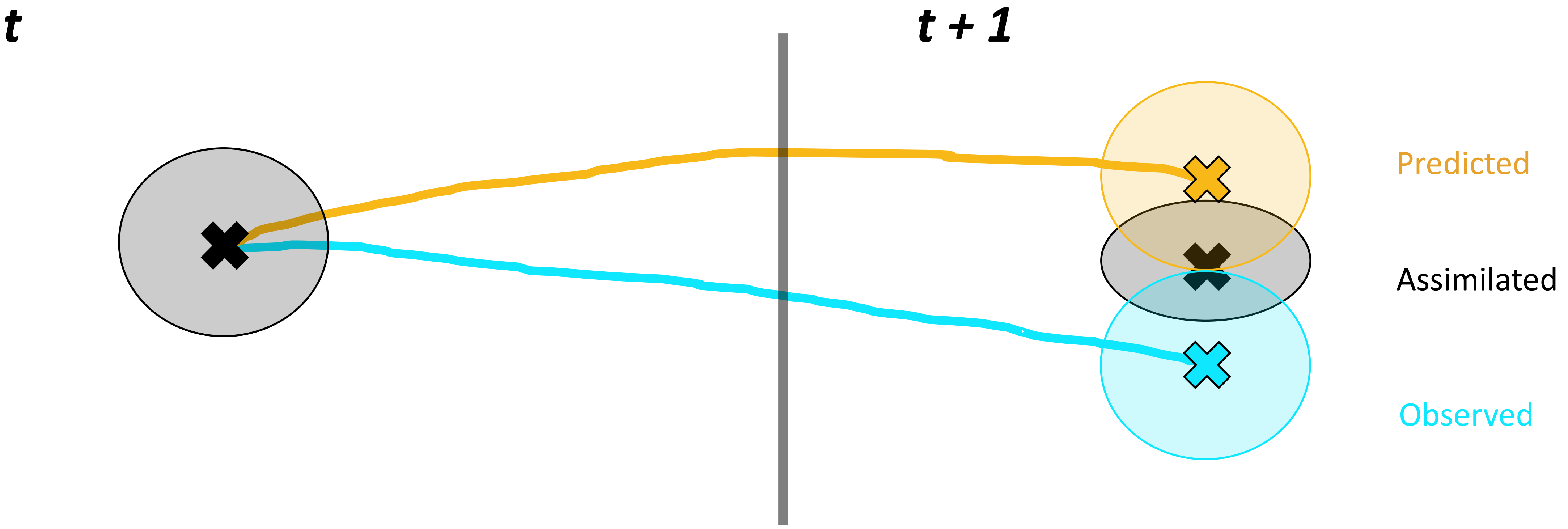
Real time digital twins?

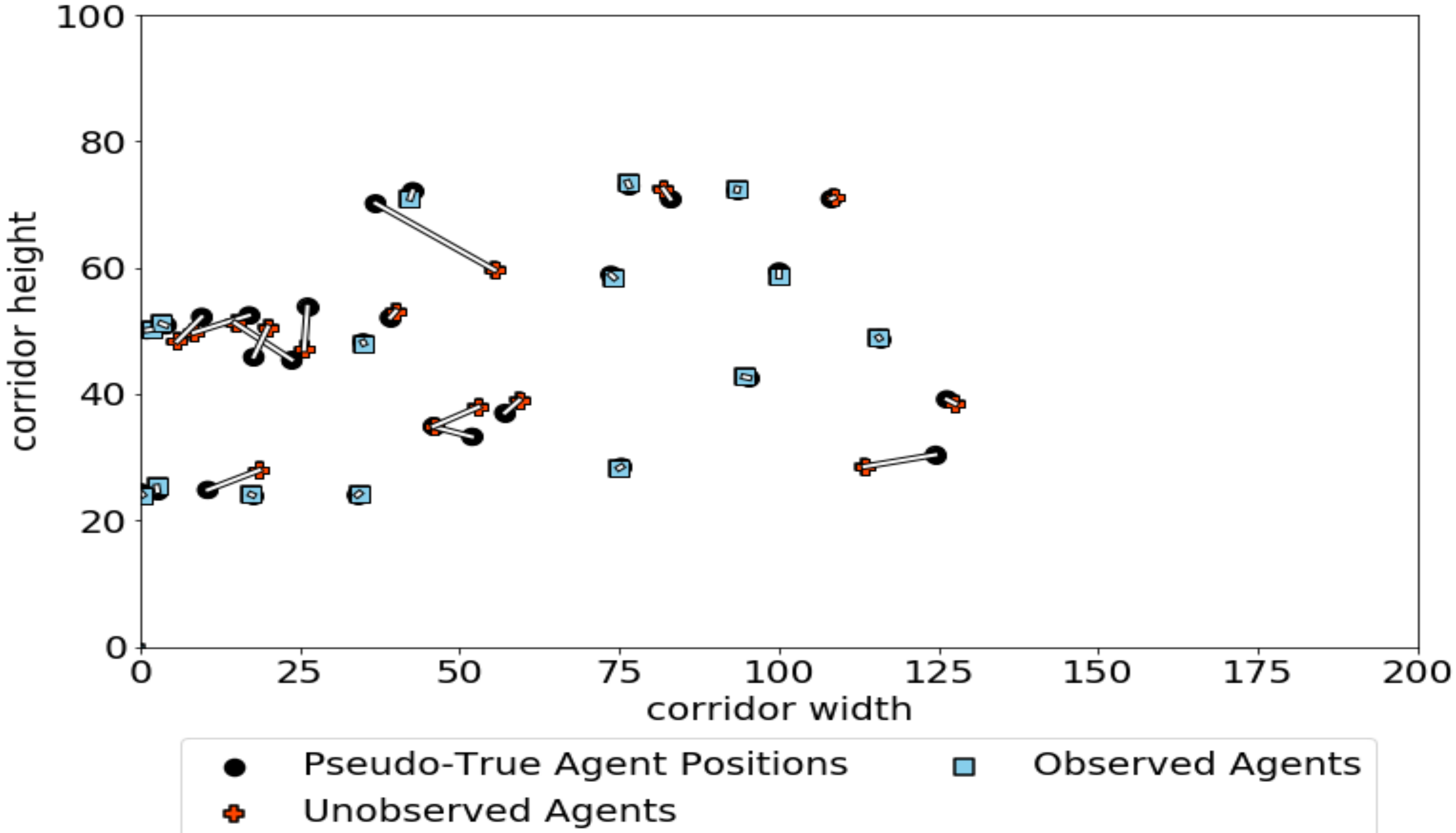
- Ultimately work towards models of cities (using anonymous data from (traffic/pedestrian counters, social media, mobile phones, etc.)
- But this is much too hard!
- At the moment we're modelling crowds
- Methods:
  - Ensemble Kalman filter (Suchak)
  - Unscented Kalman filter ([Clay et al., 2021]())
  - Particle Filter ([Malleson et al., 2022]())
  - Quantum field theory ([Tang, 2019]())
  - Agent Based MCMC ([Tang and Malleson, 2022]())
- https://urban-analytics.github.io/dust/publications.html





Data Assimilation (DA)





100
80
60
40
20
0
0
25
50
75
100
125
150
175
200
corridor height
corridor width
Pseudo-True Agent Positions
Observed Agents
Unobserved Agents

Challenges

- Categorical parameters
- Model discrepancy (the ’Serge’ effect)
- Computational issues

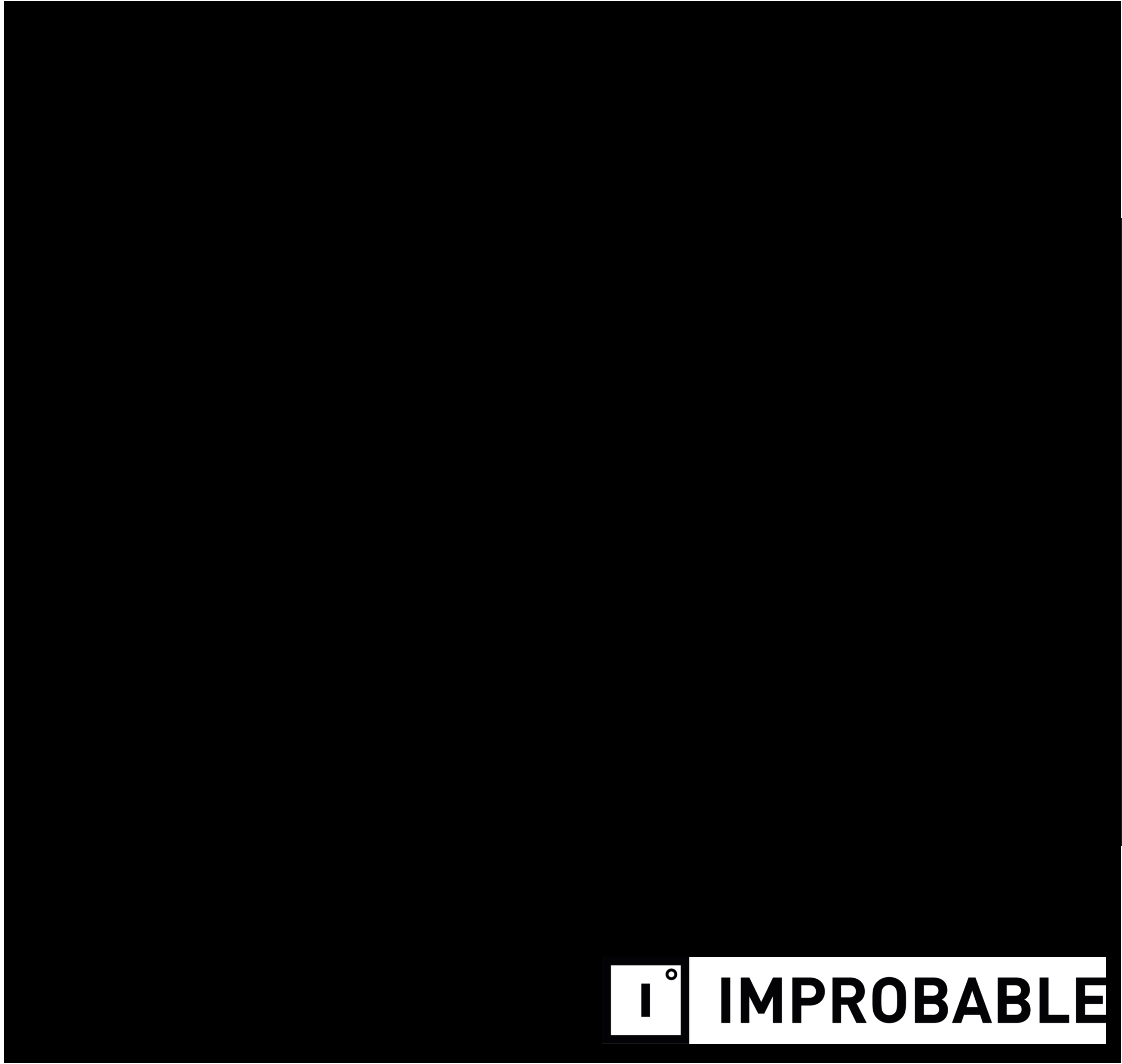
Computational Issues

Computational Issues

- ABMs are typically computationally expensive
- This prevents the use of more advanced methods (need 1000s+ model runs)
- Big computers can help
- But maybe if modelers were better at programming …

DyME Python and OpenCL

- Dynamic Model for Epidemics (DyME) applied to Devon
  - ~800,000 individuals
- Lots of interactions
- Python/R implementation:
  - 2 hours
- Rewritten by Improbable using OpenCL (python and C):
  - 5 seconds!
  - Opens new & exciting opportunities for model inference etc.

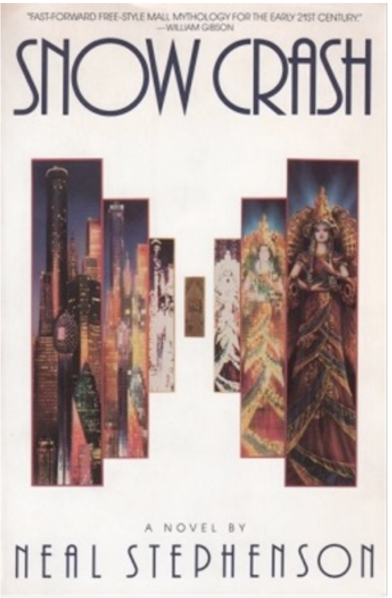


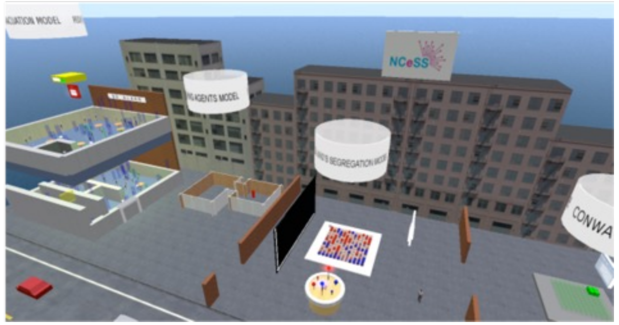
Digital Twins

Digital Twins

- Significant interest from government (and industry / academia) in digital twins
- Pieces coming together (SIPHER, DyME, QUANT, GALLANT…)
- Problems:
  - Data (and multi-level validation)
  - Compute
  - Sharing and linking models
  - Ensuring they are equitable

Summary





- Understanding how cities function now and in the future is crucial.
- How we go about simulating these systems is gaining a lot of publicity at present
- Ensuring our models are transparent and robust and the outputs can be understood by policymakers is of upmost importance
- Spatial ABMs have some unique challenges: how do we identify and simulate key processes?
- Need to look at methods in other disciplines, what utility can they have for our models?

Summary

- Collaborative model building
  - Get out from our silos and share code
- Data
  - POM, identify trends, create representative individuals. Need lots of data.
- Behavioural models
  - Move from data driven rules to reinforcement learning?
- Data Assimilation and Managing Uncertainty
  - Need more research into methods for quantifying uncertainty (quantum field theory?) and move us towards real-time simulation of cities
- ABM has a lot to offer, but we need to ensure that we work together, share best practice and produce robust and rigorously evaluated models.

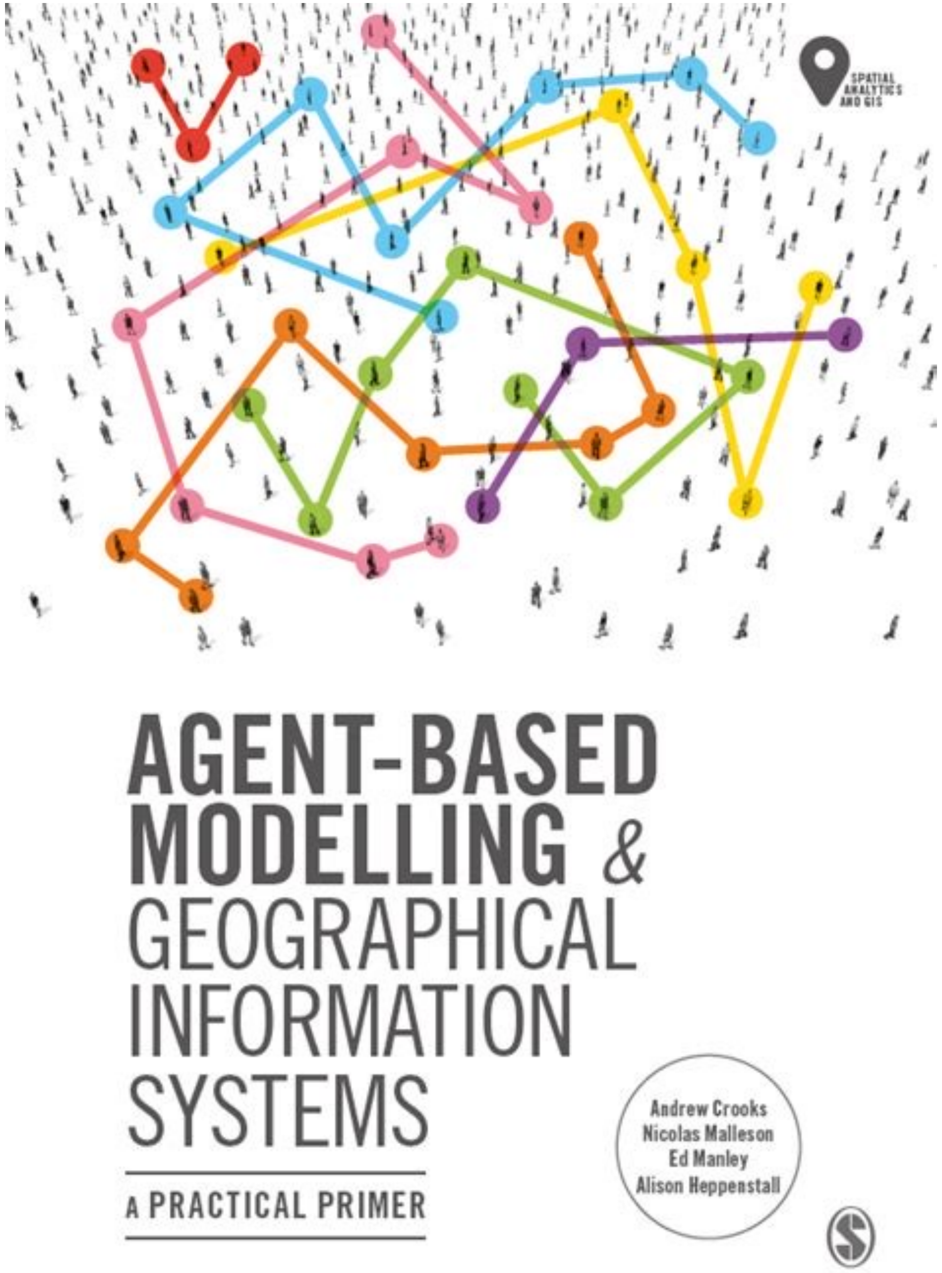
Future?

- Role of Machine Learning
  - Extract behaviour/patterns from data?
  - More ‘intelligent’ decision-making
  - Real-time simulations
- Platforms…
  - Unity: gaming platform
  - Virtual Reality: Behavioural dynamics of human and non-human agents in environments.

Agent-based Modelling and Geographical Information Systems: A Practical Primer

n.s.malleson@leeds.ac.uk @nickmalleson

Alison.heppenstall@glasgow.ac.uk @ajheppenstall



Save the date: GIScience 2023

- We are delighted to announce that the **2023 GIScience conference** will be held at the **University of Leeds, UK**, from **Wed 13th - Friday 15th September 2023**

Simulating social systems with individual-based models: **is it worth it?**

Alison Heppenstall
@ajheppenstall

Nick Malleson
@nickmalleson

Presentation to the Institute of Geography, University of Augsburg
Wed 13th July 2022