



JISC



Developing e-Infrastructure and Geographical Models

Andy Turner

<http://www.geog.leeds.ac.uk/people/a.turner/>

ACET/PEDAL Seminar, School of Systems Engineering, University of Reading

2009-05-20



JISC



Outline

- Introduction
- Geographical modelling and computational geography
- MoSeS
- GENESIS
- NeISS
- Collaboration
- Conclusion



JISC



Introduction

- Who am I?
- Stan Openshaw
- Who are you?
- My introduction to computational geography and eResearch
- What is computational geography?
- What is eResearch?



JISC



Who am I?

- Background in mathematics and geographical information systems
- Trying to fulfill Stan Openshaw's vision for computational geography
- Researcher
 - Based in the [CCG](#) since 1997
- Programming in Java
 - Since 2000



Stan Openshaw





JISC



Who are you?

- School of Systems Engineering Research
 - Ambient & Pervasive Intelligence Research Group
 - Cybernetic Intelligence Research Group
 - Informatics Research Group
 - Instrumentation & Signal Processing Research Group
 - Interactive Systems Research Group
 - Parallel, Emergent & Distributed Architectures Laboratory
- Centre for Advanced Computing and Emerging Technologies (ACET)
- Anyone who does:
 - Geography/Humanities/Social Science research?
 - Demographic research?
 - Modelling/Simulation work?
- Anyone with an interest in improving physical transport efficiency?
 - Road Safety



JISC



My introduction to computational geography and eResearch

- GIS to GeoComputation
- CCG in crisis
- SIM-UK
 - March 2005
- MoSeS
 - July 2005
- Events and Training
 - Agenda Setting Workshops
 - UK e-Science All Hands Meeting
- e-Social Science and the evolving e-Research community
 - Trying to get others going, making it easy for others to catch up and collaborate
- Grid Computing
 - Crossing organisational boundaries
- 'The Grid'
 - Language, hype and acronym soup



JISC

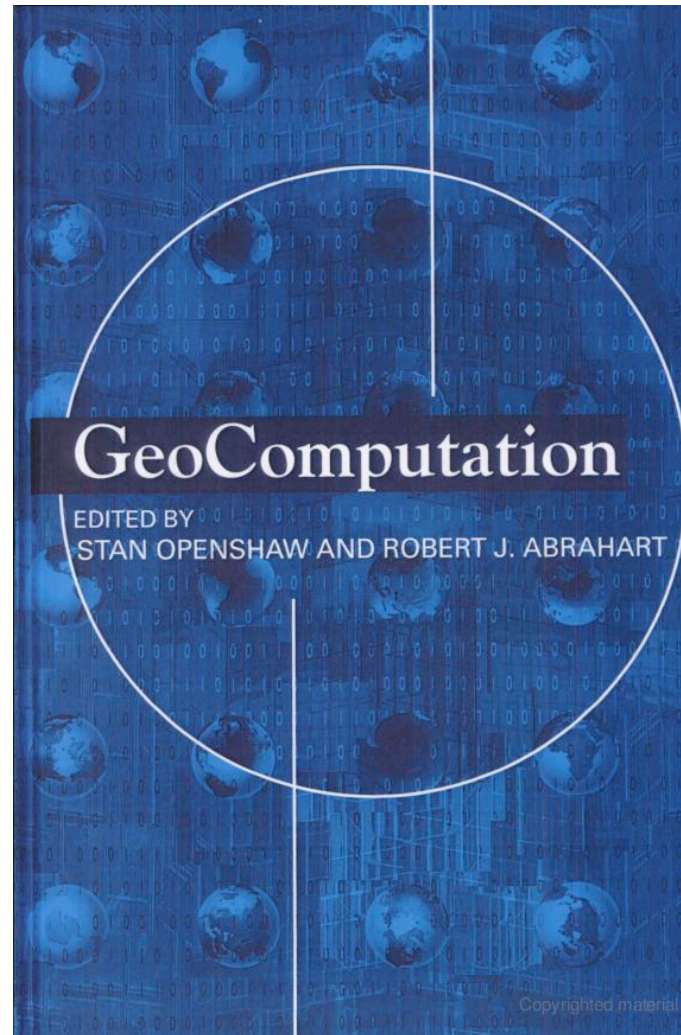


What is computational geography?

- About bringing computers to bear on complex social and environmental problems.
- The application of the computational science paradigm to geographical problems.



JISC





JISC



GeoComputation

- Essentially the follow-on revolution from Geographic Information Science
- It comes about once we have a critical mass of data that is geographically referenced
- Exploits developments in IT and new data gathering and earth observing technologies, and takes the notion of GIS beyond data and towards its analysis, modelling, and use in problem solving
- The book
 - provides pointers on how to harness these technologies in tandem and in the context of multiple different subjects and problem areas
 - seeks to establish the principles and set the foundations for subsequent growth.



JISC



Geographical Modelling

- Geography as a subject is concerned with describing the world.
- Geographical modelling and simulation is about generalising the processes and having a dynamic update across space and over time
- Involves defining classes of things and instantiating objects in a system that are enacted to interact



JISC



Modelling is generic and focussed in some context

- Modelling is pretty much the same no matter what you are modelling
- The focus of a model is bounded in terms of spatial and temporal extent and in the resolutions of these dimensions
- Most things are physically constrained and constrained on networks, but in order to generate a model we need to generalise and understand things on an object level
- Modellers seek patterns and generality
- This is done as much by a geographer in developing a geographical model as:
 - a biologist in developing a model of how a body works
 - a chemist in developing a model of a reaction of substance
 - a physicist in developing a model of the fabric of reality
 - at the extremes of scale
 - the context in which all physical models operate



JISC



A geographical focus

- So geographical modelling is a focus on modelling everything
 - Focuses on here and now at human scales
 - Down to the spatial resolution of individual beings
 - Higher resolution in some cases
 - Up to the spatial scale of their maximum range
 - Is there a minimum temporal resolution?
 - I am currently thinking about modelling things from one second to the next
 - » Fractions of a second can be important
 - » Often geographical modelling has considered changes on a higher aggregate year to year basis
- Raster versus Vector
 - Typically thought of spatially
 - Temporally this is more like
 - fixed time step versus event based



JISC



Complex Systems

- Human systems have become increasingly complex
- We live in a very complex world full of complex subsystems and feedback
- Most systems when looked at in detail are complex!



JISC



eResearch

- e-Science
- e-Social Science
- Discipline agnostic
- An open and collaborative research activity using advances in information and communication technology
 - Advances in computer science drive this
 - By definition it is collaborative
- The approach provides us with a means to achieve a critical mass of interest (resource) to begin to address some of the really big and complex problems



JISC



eResearch MO

1. Be open about what you are planning to do
2. Put the information about this online
3. Search for others doing anything similar and explore collaboration opportunities
4. Do some work
5. Publish and reflect openly
6. Repeat 1-6



JISC



eResearch projects/ collaborations

- MoSeS
- NCeSS e-Infrastructure for the Social Sciences Project
- GENESIS
- NeISS/e-ISS



JISC



MoSeS Rationale

- The idea is to provide planners, policy makers and the public with a tool to help them analyse the potential impacts and the likely effect of planning and policy changes.
- Example Application:
 - There may be a housing policy to do with joint ownership, taxation and planning restriction legislation that can be developed to alleviate problems to do with lack of affordable housing and workers without precipitating a crash in the housing market and economy as a whole
 - A balanced policy may be easier to develop by running a large number of simulations within a system like SimCity for real to understand the sensitivities involved

MoSeS Vision

- Suppose that computational power and data storage were not an issue what would you build?
 - SimCity
 - <http://en.wikipedia.org/wiki/SimCity>
 - For real on a national scale





JISC



Initial Tasks

- Develop methods to generate individual human population data for the UK from 2001 UK human population census data
- Develop a Toy Model
 - Dynamic agent based microsimulation modelling toolkit and apply it to simulate change in the UK
- Develop applications for
 - Health
 - Business
 - Transport



JISC



Challenges

- Grid enabling the data and tools
- Visualisation
 - Google Earth
 - Computer Games
- Collaboration
- Retaining a problem focus
- Design and Development



JISC



Generic MoSeS Approach

- MoSeS to date has approached Modelling and Simulation from a specific angle
 - Geographic
 - Demographic
 - Contemporary
 - About the UK
 - Targeted towards supporting a developing set of applications
- It is not a requirement to make it clear what steps can be followed by other Social Scientists wanting to Model and Simulate something different
 - However, the generic work of MoSeS should be relevant and we are working towards this



JISC



MoSeS First Steps

- The development of a national demographic model
- The development of 3 applications
 - Health care
 - Transport
 - Business
- The development of a portal interface to support the development and resulting applications by providing access to the data, models and simulations and presenting information to users (application developers) in a secure way



JISC



My MoSeS Checklist

- Outputs to be made as openly available as possible
- Use appropriate standards
- Automate with free and open source software.
- Results to be replicable
- Be open and up front about what we were trying to do and how
- Adopt best practice and learn from others in NCeSS and think about what else they and what formed to be called the e-Research community wanted.



JISC



Reflection on MoSeS

- Never-ending story...
- Too early to judge?
- There are many positives:
 - I have learned a great deal over the last 3 going on 4 years and found a community of collaborators that I am happy and excited to work with.
 - I have developed a lot of structured information about me and my research interests.
 - I have participated in lots of surveys.



JISC



- All MoSeS code we developed is open source and Java based
- We used great tools developed by others
 - Netbeans/Eclipse
 - MPJ Express
 - Gridsphere/Sakai
- We did use a variety of computational resources
 - We got more results thanks to the UK NGS
- MoSeS was highly ambitious and we didn't get as far as hoped



JISC



NCeSS e-Infrastructure for the Social Sciences Project

- <http://url.ca/4jx7>
- ESRC funded
- MoSeS node responsible for a small part of this work
 - Workpackage 4.1 focuses on the provision of a Geographic Linkage Service and a mapping service
- My role is to manage this
- Paul Townend and Junaid Arshad from the [Distributed Systems and Services Group at Leeds](#) have been doing the work
 - Junaid has been focussing on security which is a major issue in social science generally
 - It is security of the data that is key
 - Not exposing the data such that it can be accessed by non-registered users and that when it is accessed by registered users they are made aware/reminded of the data usage restrictions
 - The hope is that Use Case based security can be provided to Virtual Organisations
 - The data may never have to leave the system in a potentially disclosive form, but that the outputs can be made available for reports that can be publicly consumed and used in administrative decision making.



JISC



GENESIS

- <http://portal.ncess.ac.uk/access/wiki/site/GENESIS/home.html>
- Second phase research node of NCeSS
 - Began in October and ongoing for 3 years subject to review
 - Focus on generative geographical simulation
- Combines the modelling strengths of MoSeS with the visualisation strength of GeoVUE
- My work on GENESIS
 - Developing "agent based models" of two types:
 1. based on publicly available data that can be freely distributed,
 2. that take account of more restricted data, but which might be more realistic/detailed.
 - I am developing models that operate
 - with a daily time step and run for years
 - on a second time step and run for a day
 - The computational requirements for these models are considerable and I am investigating ways to scale and partition the task
 - To begin with I am considering something city scale with approximately 1 Million individuals
 - I would like to do something on a national scale or even global scale



JISC



NeISS

- <http://portal.ncess.ac.uk/access/wiki/site/e-iss/home.html>
- JISC funded
 - Part of the Information Environment 2009-2011 Programme
 - 3 years
 - Started in April
- Developing a national e-Infrastructure for social simulation
- Key components of the work:
 - orchestrate the data and computation
 - develop modelling and analysis tools
 - close the publication cycle by encapsulating the process in workflows that can be enacted and re-enacted
- To be done as open eResearch using the NCeSS Sakai Portal to help manage and organise information about the project



Collaboration

- Let us think collectively about what we do and how we might collaborate
- Assessing the performance of MPI jobs when run in a VM either locally or in the cloud.
 - [Ruben Santiago Montero](#) is interested to explore a collaboration around this
- Funding opportunities
 - EPSRC
 - Complexity
 - <http://www.epsrc.ac.uk/CallsForProposals/complexsciencecall.htm>
 - Software sustainability
 - <http://www.epsrc.ac.uk/CallsForProposals/SoftwareSustainabiity.htm>
- I want to get to grips with g-eclipse
- Topics of personal interest
 - Road safety research
 - Use mobile technology road sensors sensors in vehicles and information about the road conditions and environment to develop a warning system for vehicle and drivers
 - Systems to help people with Cystic Fibrosis
 - Developing a virtual Stan and systems to communicate with the main man



JISC



Thanks and Acknowledgements

- eResearch community
 - NCeSS
- ACET/MPJ Express developers
 - Aamir Shafi
 - Bryan Carpenter
 - Mark Baker
- University of Leeds
 - School of Geography
 - Centre for Computational Geography
- EC/ESRC/JISC
- Organisers
- You