Wednesday, September 28th 2011

Dilemmas of Urban Simulation:

Questions of Scale, Data, Model Design, Theory & Prediction

Michael Batty
The Bartlett, University College London

m.batty@ucl.ac.uk
www.casa.ucl.ac.uk
www.ComplexCity.info

Preamble

I am going to give you a lecture about why we need to re-invigorate GI Science by building models of geospatial systems

This is a hobby horse of mine and it has been this last 40 years since I was a graduate student

I believe that our field is far too wedded to exploring representation, GIS, and far too focussed on spatial analysis than on theories and models for making key predictions about the future.

So this is what I will attempt to portray …….

Outline

- The Geospatial Context
- What Are Models? Models Defined
- Types of Models: Lowry's Famous Paper
- Modelling and Computation: The Virtual Real World
- Principles of Modelling
- Spatial Imperatives – ABM, Integration, Visualisation
- Exemplars: Some of Our Projects in CASA
- Conclusions: Predictions: Informing the Future

The Geospatial Context

What is the Geospatial Context: Mike Goodchild's Definition

In general, geospatial representation and analysis is about informing the world with respect to spatial function

It tends to veer to abstraction as simplification in terms of representation, and then analysis, not modelling per se

Modelling of course shades into representation and analysis but in principle, I believe that modelling is about prediction

In this sense, it is about theory – theories that are computable in that the media for prediction is in the form of a computer model

Here we need to think of models and their media – this is key

Centre for Advanced Spatial Analysis UCL

What Are Models? Models Defined

A model is a simplification of reality: it is an abstraction. It is simpler than the real thing. This is an essential tension. We always lose information, when we build a model and in this sense, it is often *less than the representation* we seek.

We work with models all the time but in the last 50 years they have become central to our vocabulary. If you go back to those times, the term was seldom used: theory was more widely used

Models are abstractions – whatever we do, we abstract and simplify and thus we can define a model of anything. This means that when we have a conversation and we speak of models, we can quite easily talk at cross purposes. This play on the word 'model' is central to this lecture.

Let me show you some definitions to show how broad this area is and to anticipate what we are going to do in this lecture, let me flag up the idea that we can have different degrees of abstraction, and hence different kinds of models of cities.

First what does the new iconography of our times say about the term model: let's Google it

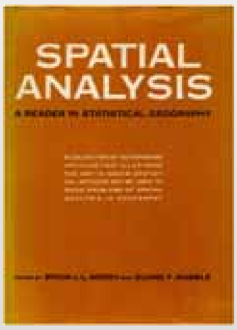
Centre for Advanced Spatial Analysis UCL

Types of Models: Lowry's Famous Paper

In 1965, Ira Lowry wrote a famous paper called "A Short Course in Model Design", in the *Journal of the American Institute of Planners*. There were many such papers written then about models and what they were but this paper was noteworthy because of its clarity and salience.


A SHORT COURSE IN MODEL DESIGN

IRA S. LOWRY



Lowry amongst others produced a simple classification of models, first in terms of their function, and then in terms of their media.

In terms of function he spoke about

- **Descriptive** or representational models, giving *insights*
- **Predictive** or analytic models capable of *forecasting*
- **Prescriptive** or design models capable of *optimisation*

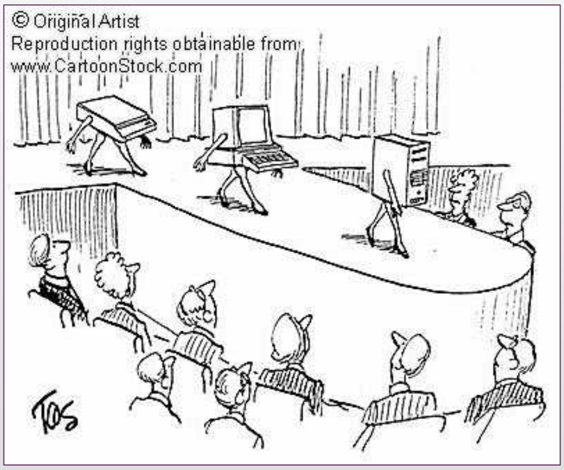
These three kinds of abstraction shade into one another of course but in essence, geospatial analysis has tended to focus on the first of these categories – seldom moving beyond analysis, and certainly not dealing with prediction per se. Insofar as our field deals with prediction, it is to used analysis to inform the decision process which is assumed to be political.

The other classification that Lowry produced is highly relevant to our concerns today and this was in terms of the media used for modelling. He made the distinction between

- **Iconic models** – abstractions that simplified the superficial content of the system, like a toy
- **Analogue models** – abstractions that mirrored the system using some other system – in analogy
- **Symbolic models** – models that abstracted the world into some set of symbols which could be manipulated for functions such as prediction.

In fact iconic and analogue models had less capability for prediction than symbolic

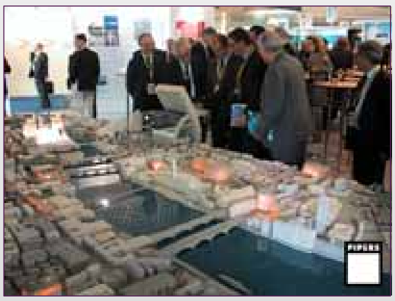
In fact this convergence between all these styles of model and their media is nicely captured by this cartoon



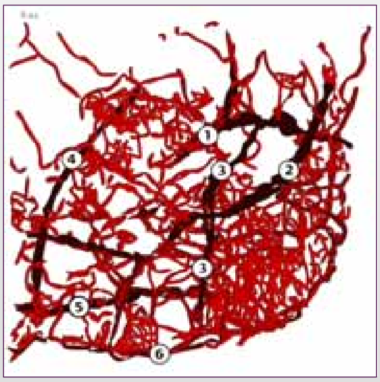
'Computer Models' on the catwalk

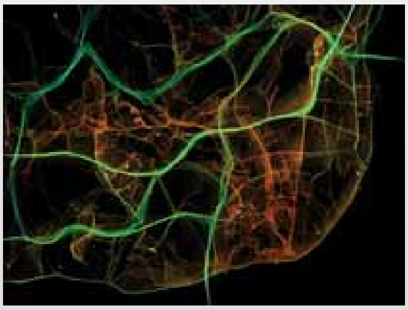
Before I proceed, let me throw out some examples – of each of these styles

Iconic in our field are architects models: here are some large scale ones for central London – it is easy to see that these are eminently useful for participation and dialogue and to pose ideas



Analogue exist much less in our field but good examples are in traffic where the analogy is with fluid flow. In fact here is a rather convoluted example of such flow as visualised using computer graphics – flow of traffic in Lisbon





Centre for Advanced Spatial Analysis

UCL

And last but not least, *Symbolic* or mathematical models which are the essence of what we will develop here. Consider population growth – with a simple differential equation

$$\frac{dP(t)}{dt} = \lambda P(t)$$

which can be integrated or summed to produce a trajectory of change – the classic exponential growth model as

$$P(t) = P(0)e^{\lambda t}$$

Obviously these generate growth profiles of the characteristic form and they are clearly computable. No surprise that the first computer models were symbolic ones.

I will show some of these later but I need to talk about media

Centre for Advanced Spatial Analysis

UCL

Modelling and Computation: The Virtual Real World

The media for our models is computational – numerical and digital – in fact it is entirely possible that we would not be talking about models today – we would not even have this conference – without their being digital computers.

And there are alternative world histories that could have happened without digital computation – unlikely yes but possible.

So almost as soon as computers came on the scene with the Manhattan project and Bletchley Park, with Konrad Zuse in Germany, in the period from 1940-45, they were used to build mathematical models

What happened was that initially models were numerical and digital but complex and computers were part of high science.

Then with the advent of massive transactions processing, information systems came to be developed with mathematical models drawing on such information.

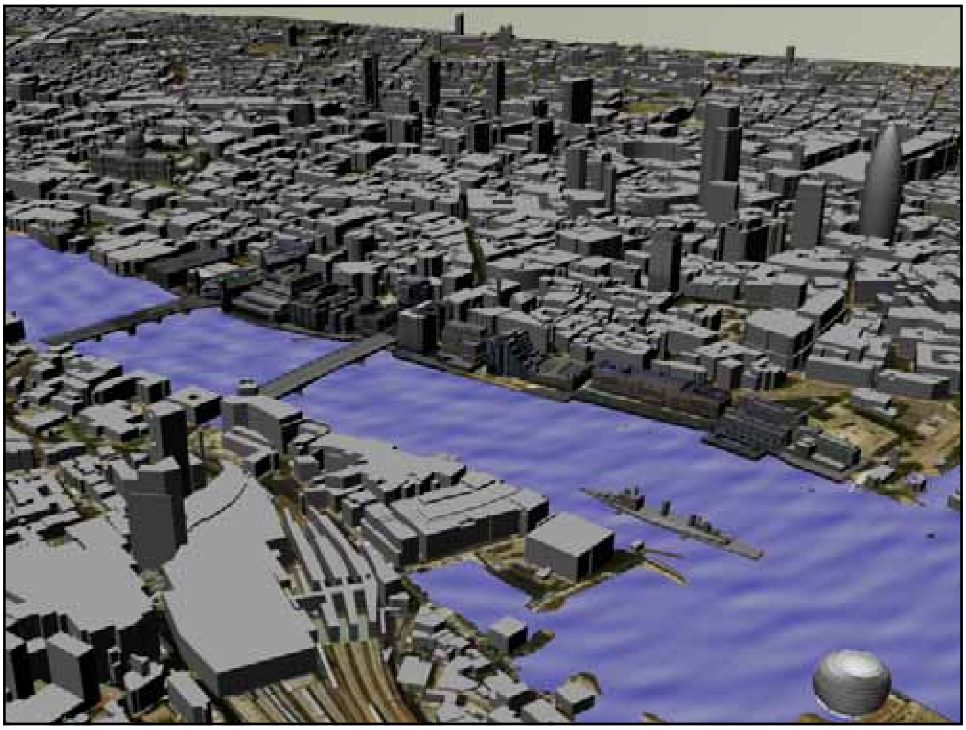
Analogues to an extent were informed by computation but still these were largely physical – egg: wind tunnels and so on – and it was not until the advent of computer graphics that things began to change. This was hastened, indeed massively accelerated by the PC and miniaturisation.

And the idea of developing computer models in analogy to iconic models came onto the agenda. This was perhaps inevitable that the media should spread to all types of models and here are some examples from our own work in London

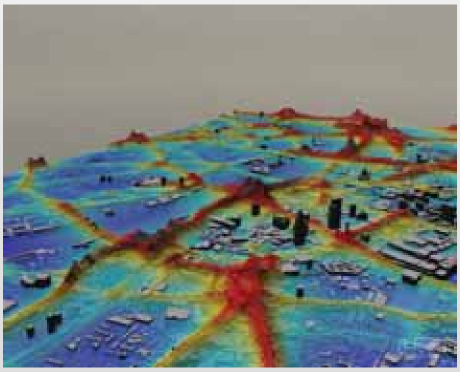
Here we see our 3D Virtual London model which is computable and its being embedded in a virtual world environment for participation and manipulation – it is being used for design or prescription; but it can also be used for simple prediction and this is a surprising development for iconic models

Centre for Advanced Spatial Analysis

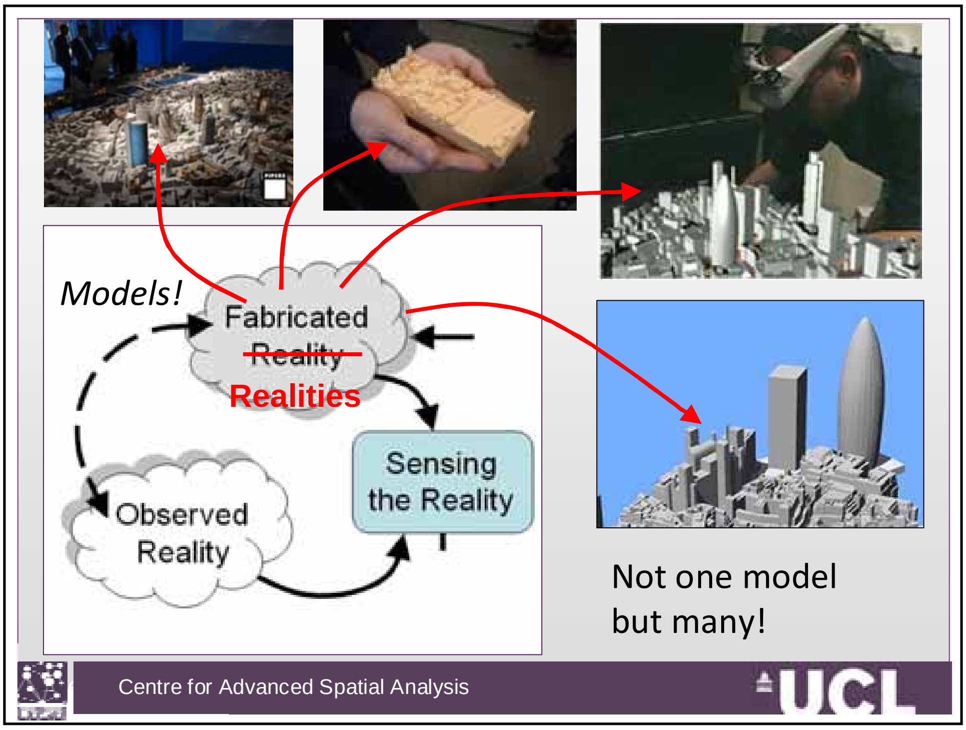
UCL



What perhaps is most surprising is how iconic and symbolic are merging – as for example in our pollution work where the model is being used to represent how pollutants are generated and modelled and flow – in a sense all these models types of coming together: ***a (the) convergence !***



And last but not least we can come full circle and generate analogues and icons from our symbols – here is a graphic of how we can print our model in conventional media from digital



Principles of Modelling

Let me retrack to emphasise three key principles of modelling: these involve simplicity, repeatability, and validity

Simplicity: one of the canons of the scientific method is that models are computable versions of theories – and as such they should be as simple as possible – if you can produce the same quality of prediction with a simpler model, simpler means better in classical science – This is ***Occam's Razor***

Einstein said all models should be simpler but not too simple or something to that effect. This is the principle of *parsimony*

However we are moving very fast beyond age of classical science and these ideas are up for grabs

The second principle is

Repeatability: this is the canon which says that if you have a model, it is only good if you can develop it in one context and show that it works in another. This is the idea of generalisation. Develop the model on one example and see if it works in another, different example. Develop a model for Delft and see if it works for Dordrecht !

This is another of the sacred canons that is under scrutiny. In our world, it is often impossible and in fact this is not just an issue in the human sciences. Any science outside the controlled conditions of the experimental lab or the computer suffers from the lack of repeatability problem

In fact there is a view that models are still useful if they cannot be generalised – this is a quiet heresy that is gaining ground

The third principle is

Falsifiability: this is Karl Popper's canon that says that you cannot prove anything right but you can prove it wrong and good theory, good models, good science withstands the test of time only insofar as it cannot be shown to be wrong

This is a basic conundrum because as the future in knowable, we cannot prove anything to be fundamentally true – we cannot predict the future but we assume we can do.

This is absolutely key to our use of spatial analysis and modelling in forecasting – in decision support

It poses the question that if we cannot predict, what can we do? Simply inform the debate, focus the dialogue – what?

Last but not least a digression in experiment

It is often said that the computer replaces the experimental context in developing theory – but although one can experiment in controlled conditions on a computer, the computer is not the real world

Nor is the experiment you might say but it is much closer to the real world than the computer.

In this sense, computer modelling is not the same as experimentation and indeed there are examples of where computers merge with experiments and are essential to enable the real world to be controlled in the manner it needs to be for experiment. One thing we need to take from this is that the process of validating a computer model is very different from validating a theory through experiment

Spatial Imperatives – ABM, Integration, Visualisation

Ok let me make some points about the great diversity of models in our field and illustrate by way of closure but also focus what I think some of the key issues and dilemmas are.

I want to illustrate three issues – which serve to define these:

- The question of detail – and the development of new models that have too much detail to be parsimonious but detail is needed. ABMs and CA models represent this issue
- The question of visualisation – to cut through complexity by visualisation but to sense the limits of this style of thinking
- The question of model integration – no one model ever fits a single problem context

To do this, I will draw from three of our modelling projects

- The first work by Kiril Stanilov (and a little bit by me) on a CA model of the historical development of west London using Metronamica from RIKS
- The second our integrated model of climate change in London which strings or chains together several aggregate models
- The third our work with a larger but static equilibrium model of the London region which lets us do all kinds of what if style scenarios and visualise these

All these examples illustrate the tension between detail and parsimony, simplicity and complexity and dilemmas posed by informing those who seek to change the future and those who need to predict it

Metronamica for West London: Modelling Urban Development from 1870 to 2011

Centre for Advanced Spatial Analysis UCL

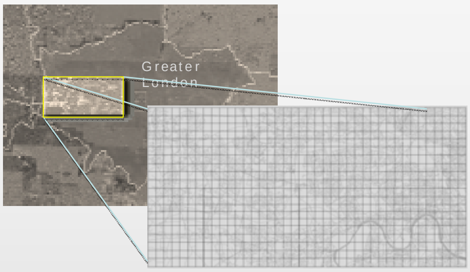
Metronamica for West London: Modelling Urban Development

data

Centre for Advanced Spatial Analysis

UCL

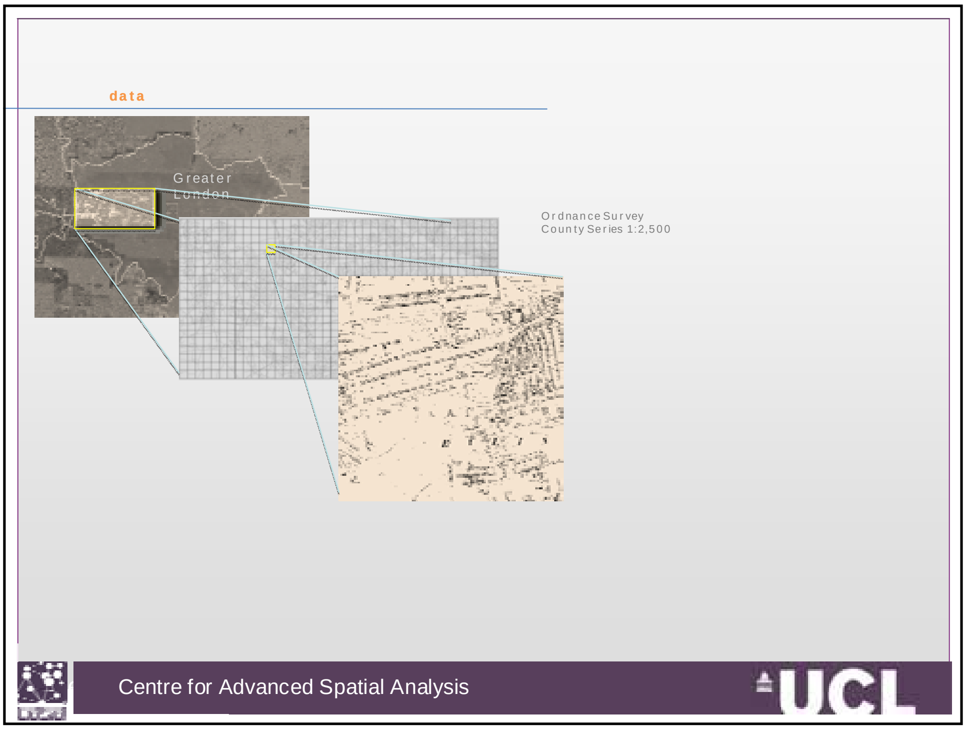
data

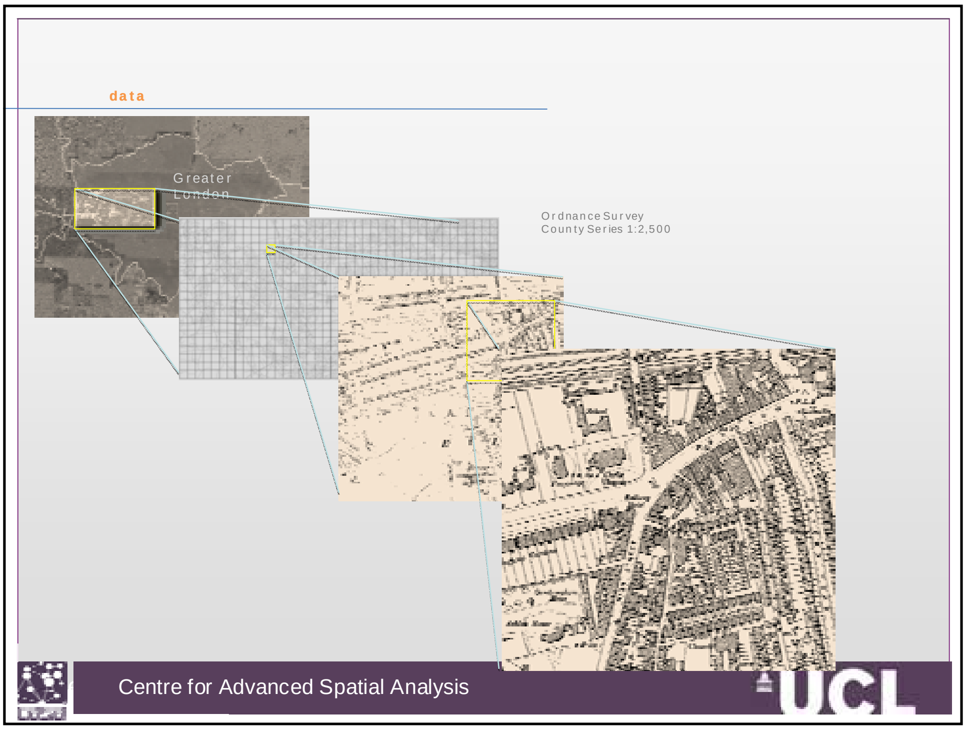


Ordnance Survey
County Series 1:2,500

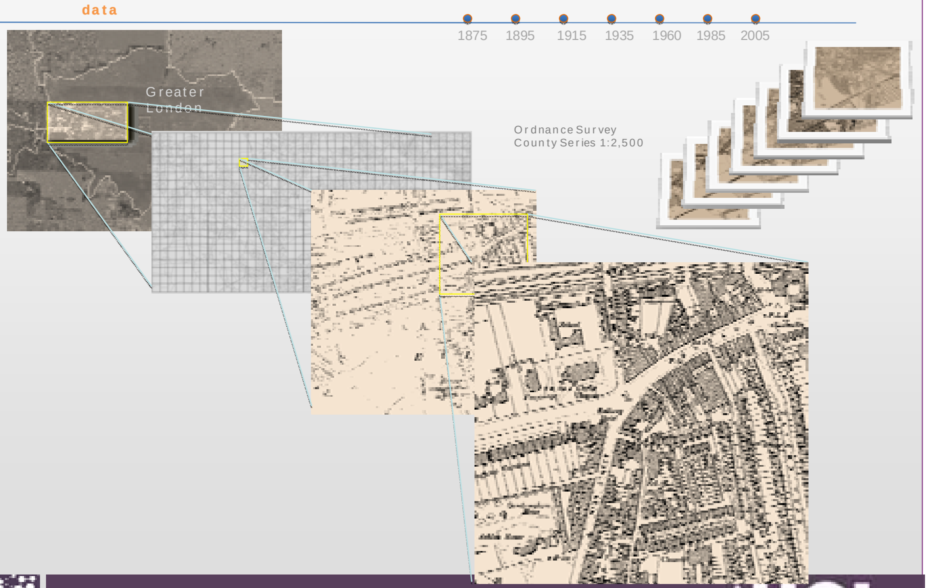
Centre for Advanced Spatial Analysis

UCL


data
Greater
London
Ordnance Survey
County Series 1:2,500
Centre for Advanced Spatial Analysis
UCL


data
Greater
London
Ordnance Survey
County Series 1:2,500
Centre for Advanced Spatial Analysis
UCL

data



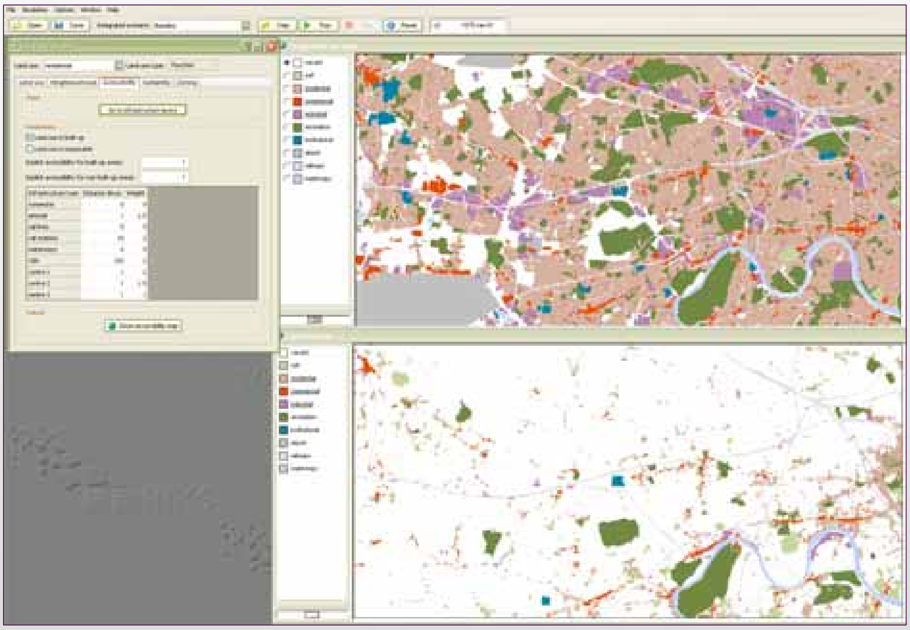
Centre for Advanced Spatial Analysis

UCL

Urban Development from 1870 to 2011

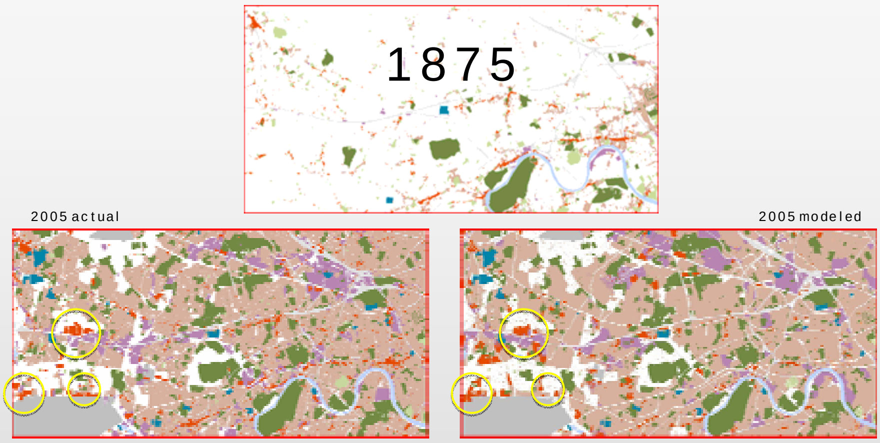
Centre for Advanced Spatial Analysis

UCL



Centre for Advanced Spatial Analysis

UCL



Centre for Advanced Spatial Analysis

UCL

The Need for Integration: Strict Requirements & Motivations for the Models

1. Predictions: Very long time horizons. Very short time horizons – equilibrium models. What If scenarios?
2. Stakeholder Involvement: The need for simple immediate models as well as visual analytics.
3. Complex Problems over Many Scales and Fields: Integrated assessment –visual analytics to communicate
4. Flexibility in Model Design and Extension: The need for developing new but related models quickly
5. The Need to Embrace Organisational Constraints: To build models understood by all members of the Consortia

Grand Challenges and Appropriate Models

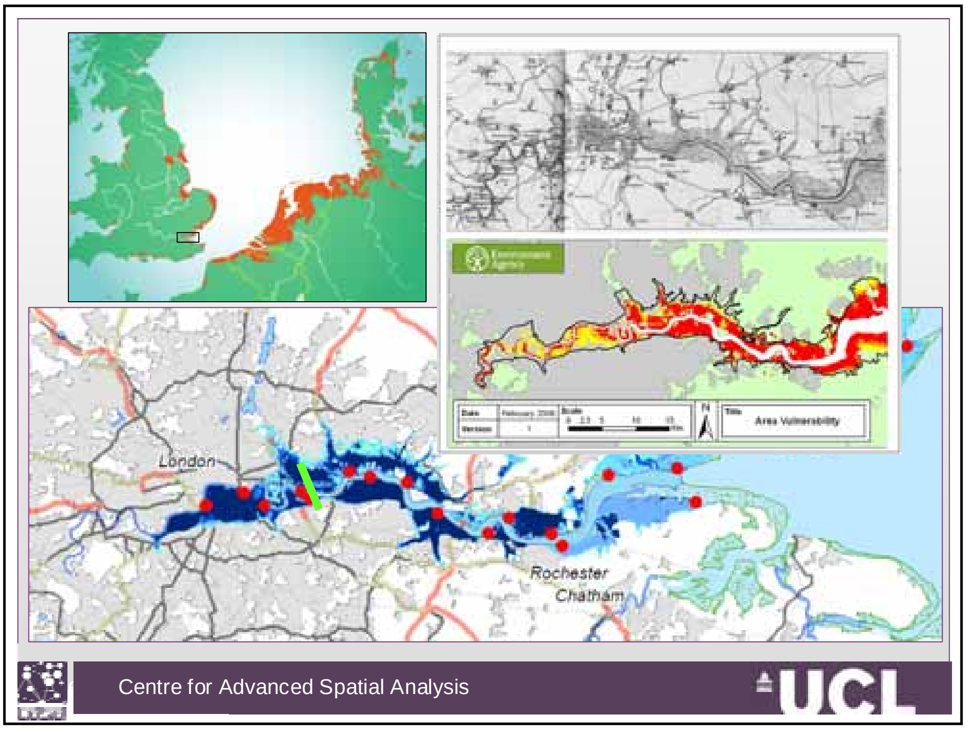
It builds immediacy, accessibility, and visual model operation

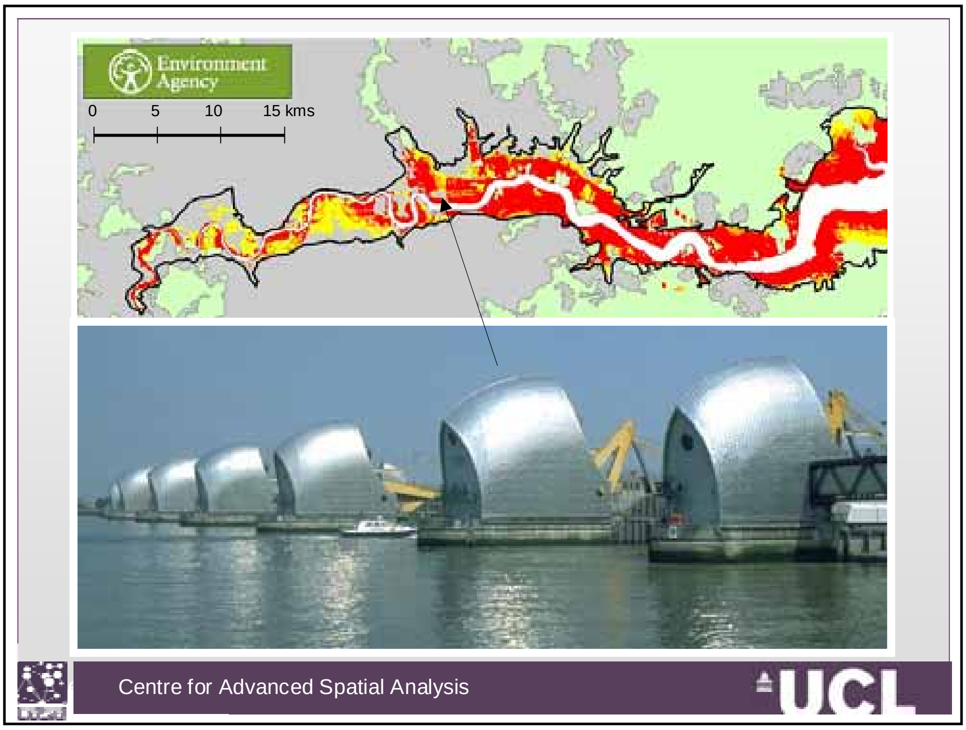
It starts from simple models such as the one I will describe here and it will progress to more complex models

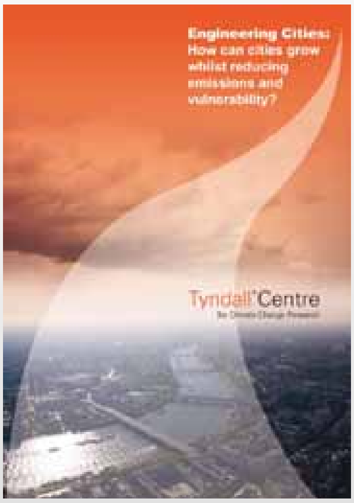
It will deliver models on the desktop but through the web

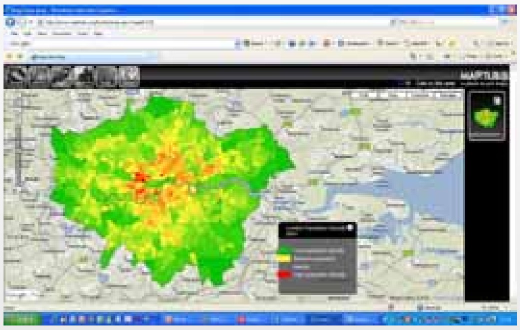
It will build several related models at different levels of disaggregation for the London region, focusing on climate change (Tyndall and ARCADIA), energy change (SCALE) and major policy drivers such as Cross-Rail, the Olympic Games and major retailing developments (GENeSIS). SIMULACRA models: see www.simulacra.blogs.casa.ucl.ac.uk

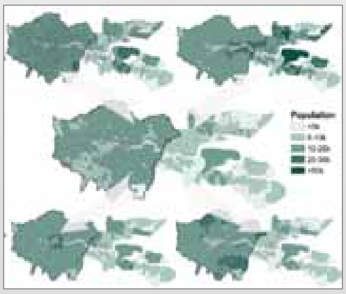
We will illustrate our first model that has been built for Tyndall


London
Rochester
Chatham
Area Vulnerability
Centre for Advanced Spatial Analysis
UCL


Environment Agency
0
5
10
15 kms
Centre for Advanced Spatial Analysis
UCL





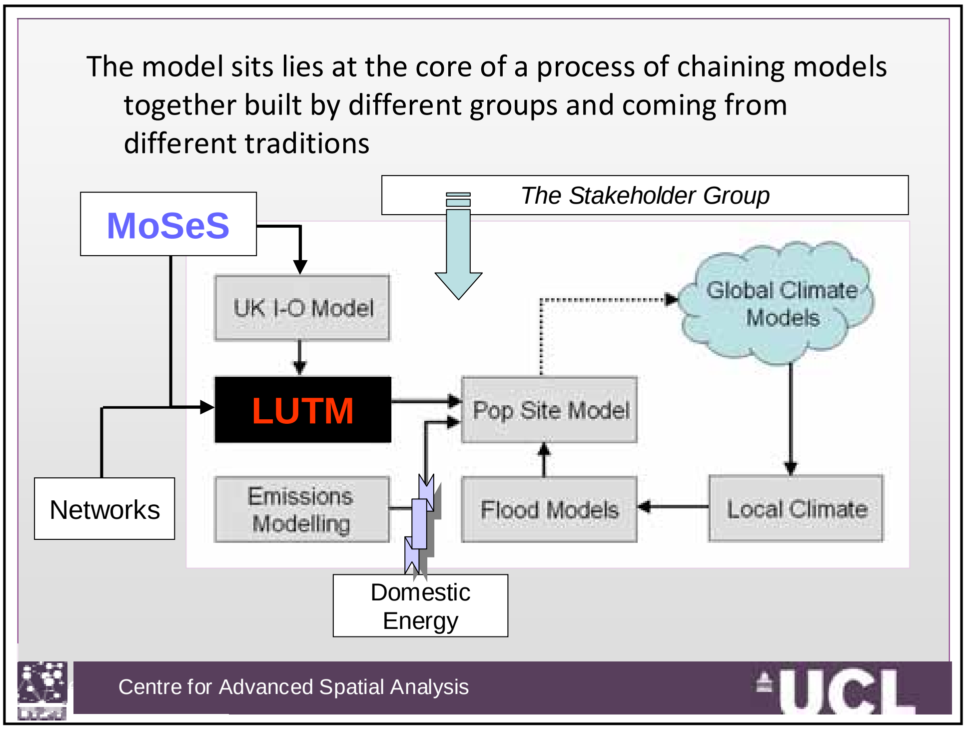


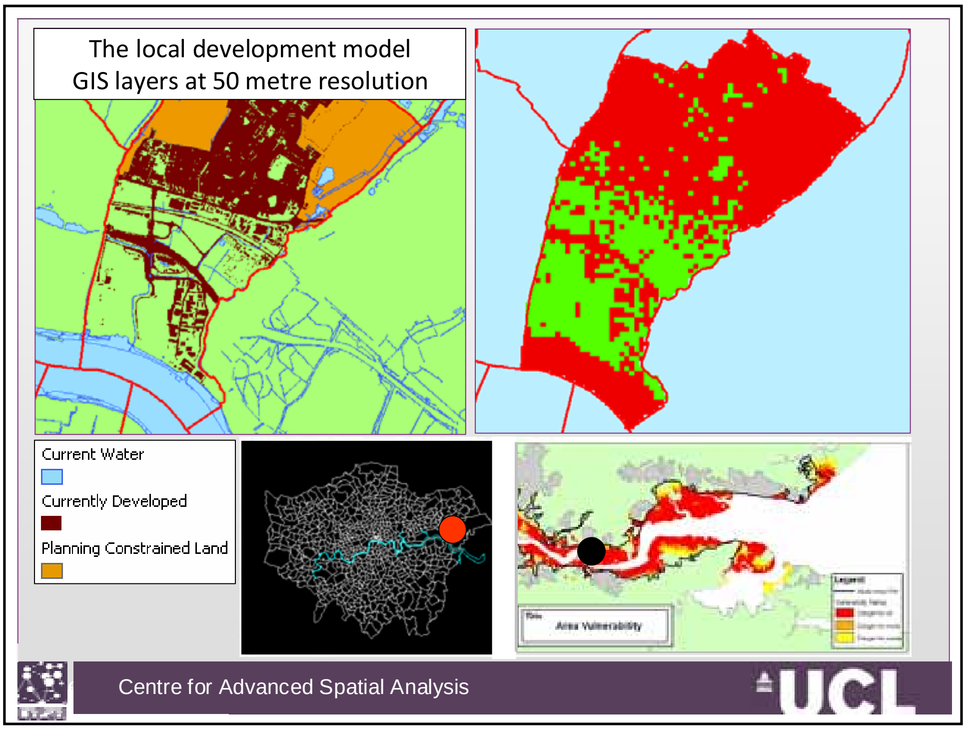
The IPCC forecasts moderated by UKCiP suggest North Sea will rise 2m by 2010, hence serious flooding of Central and East London

Centre for Advanced Spatial Analysis UCL

The Thames Barrier built 1978 to 1984 in operation, likely to be ineffective by 2040? due to new predictions of sea level rise forecast at 1-2 metres by 2100 – somewhat debatable, but ….

Centre for Advanced Spatial Analysis UCL

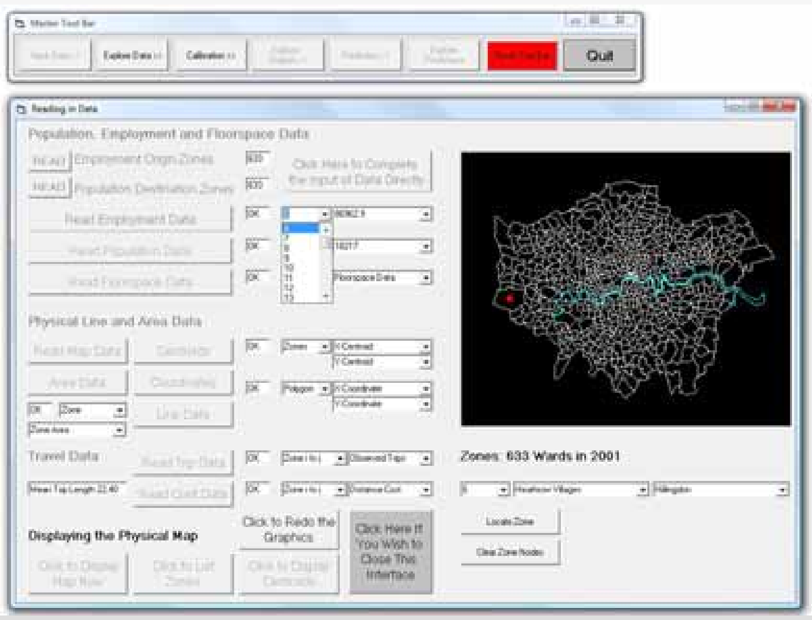

The model sits lies at the core of a process of chaining models together built by different groups and coming from different traditions
MoSeS
The Stakeholder Group
UK I-O Model
Global Climate Models
LUTM
Pop Site Model
Networks
Emissions Modelling
Flood Models
Local Climate
Domestic Energy
Centre for Advanced Spatial Analysis
UCL


The local development model
GIS layers at 50 metre resolution
Current Water
Currently Developed
Planning Constrained Land
Centre for Advanced Spatial Analysis
UCL

Visual Analytics and Modelling Processes

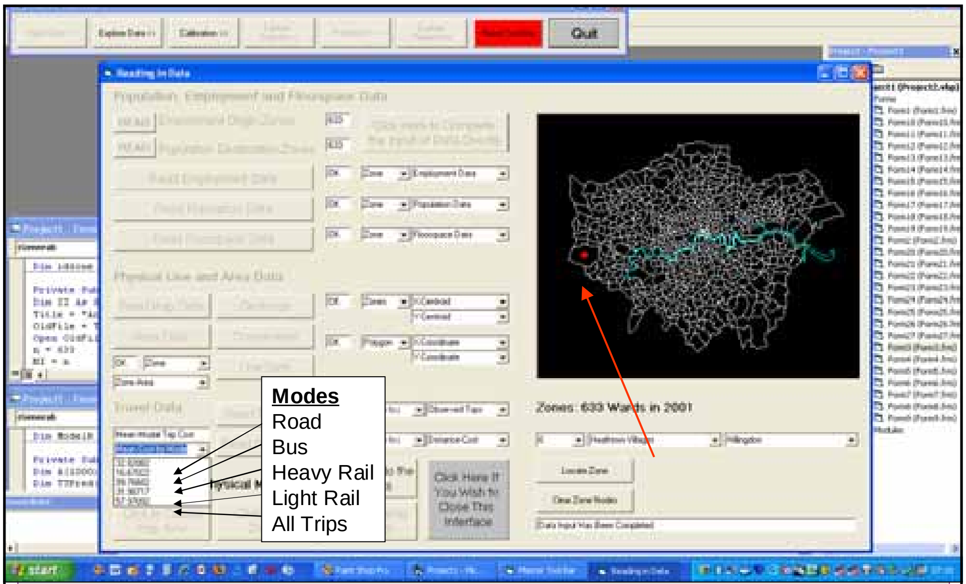
Centre for Advanced Spatial Analysis

UCL



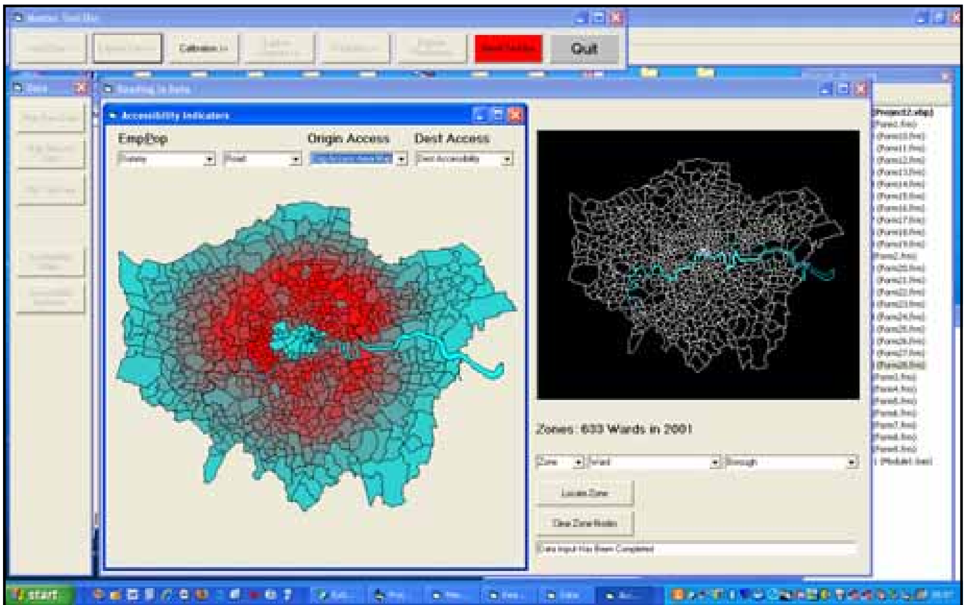
Centre for Advanced Spatial Analysis

UCL

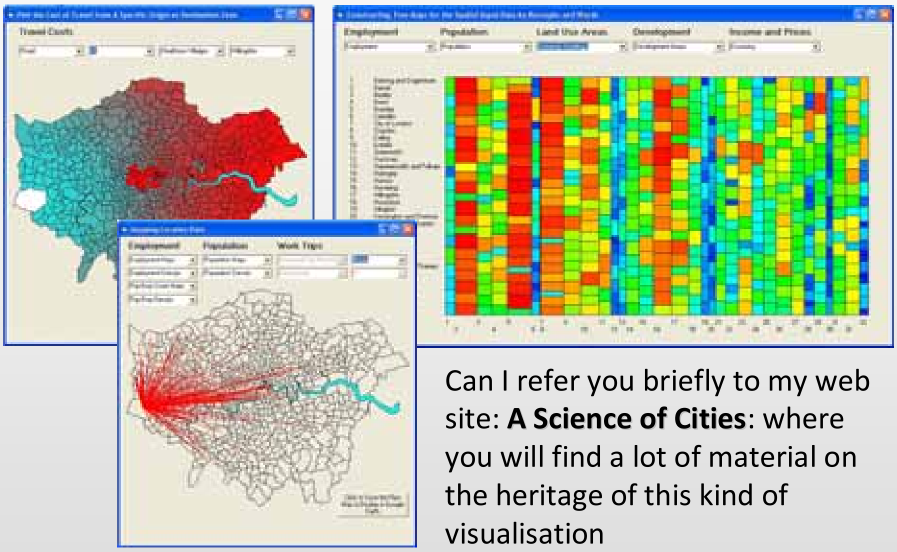


Road: 38%; Bus: 12%: Heavy Rail: 12%: Light Rail 19%; Other (Walk, Bike, Fly): 19%

Centre for Advanced Spatial Analysis UCL



Centre for Advanced Spatial Analysis UCL

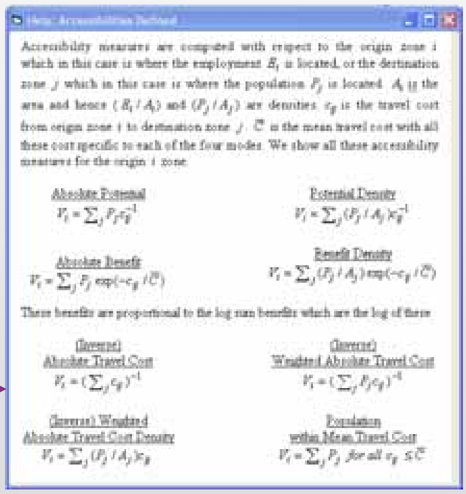


Can I refer you briefly to my web site: **A Science of Cities**: where you will find a lot of material on the heritage of this kind of visualisation

http://www.complexcity.info/media/movies/urban-models-luti/

Accessibility from the LUTM model

Many different accessibility measures, 8 in all



Accessibility measures are computed with respect to the origin zone i which in this case is where the employment E_i is located, or the destination zone j which in this case is where the population P_j is located. A_i is the area and hence (E_i / A_i) and (P_j / A_j) are densities. c_{ij} is the travel cost from origin zone i to destination zone j. $\bar{C}$ is the mean travel cost with all these cost specific to each of the four modes. We show all these accessibility measures for the origin i zone.

Absolute Potential
$$V_i = \sum_j P_j c_{ij}^{-1}$$

Potential Density
$$V_i = \sum_j (P_j / A_j) c_{ij}^{-1}$$

Absolute Benefit
$$V_i = \sum_j P_j \exp(-c_{ij} / \bar{C})$$

Benefit Density
$$V_i = \sum_j (P_j / A_j) \exp(-c_{ij} / \bar{C})$$

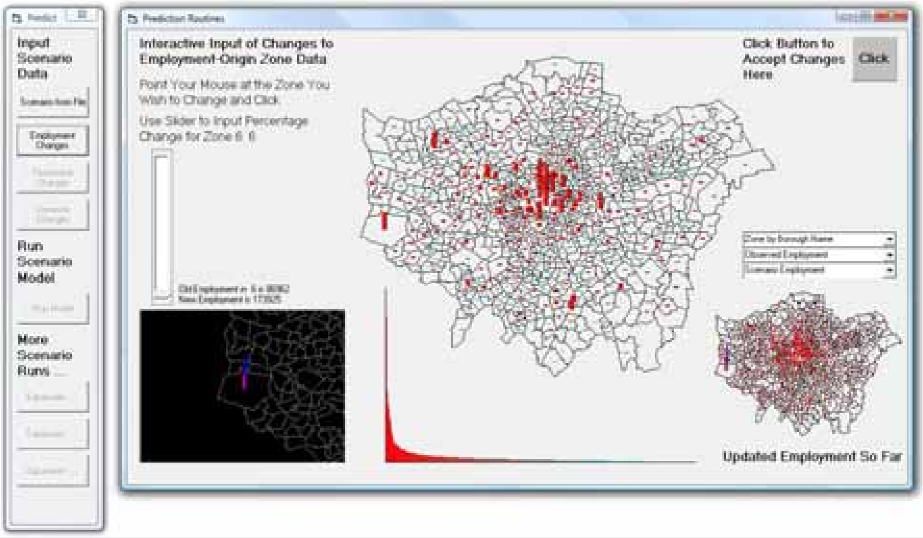
These benefits are proportional to the log sum benefits which are the log of these

(Inverse) Absolute Travel Cost
$$V_i = (\sum_j c_{ij})^{-1}$$

(Inverse) Weighted Absolute Travel Cost
$$V_i = (\sum_j P_j c_{ij})^{-1}$$

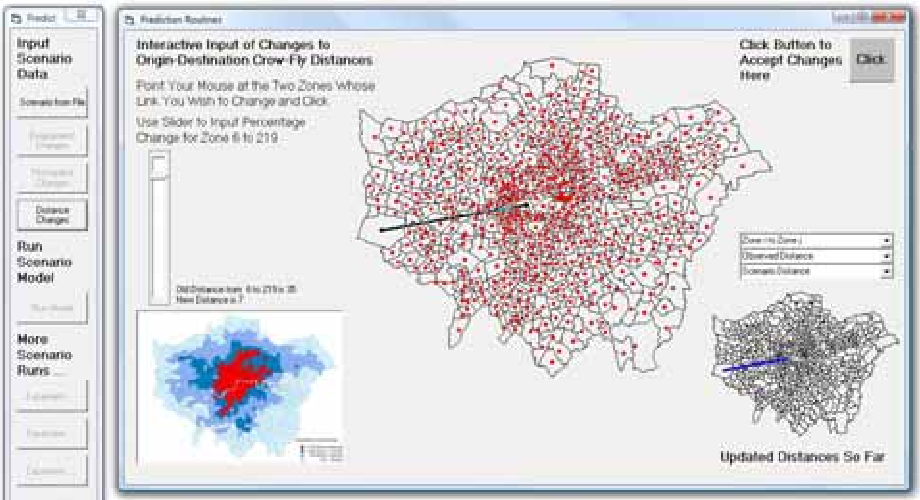
(Inverse) Weighted Absolute Travel Cost Density
$$V_i = \sum_j (P_j / A_j) c_{ij}$$

Population within Mean Travel Cost
$$V_i = \sum_j P_j \text{ for all } c_{ij} \le \bar{C}$$



Centre for Advanced Spatial Analysis

UCL



I would like to show you it running – but no time…..

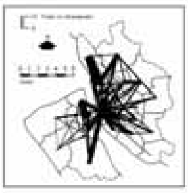
Centre for Advanced Spatial Analysis

UCL

For several movies of all our models go to www.complexcity.info/media/movies Drill down to find

Early Computer Movies: 1967 to 1990

Here we show how crude movies of urban development were fashioned from SYMAP plots stitched together using crude in-betweening. And then we move onto pixel arrays on PCs and VDUs – visual display units – attached to mini and even main frame computers, which ultimately turned into work stations by the early 1990s.



Urban Models (LUTI)

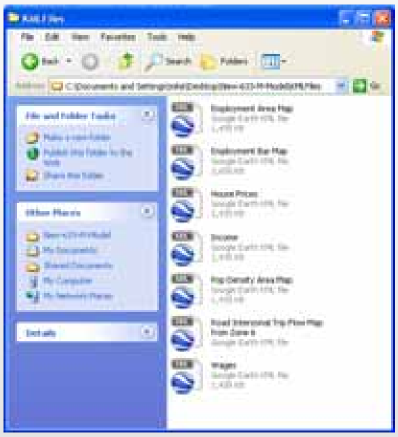
These are land use transportation interaction models that in our context evolved from early visual representations developed on workstations and then PCs. The Tyndall model and its successor ARCADIA are the key examples we show here. We have linked them to 3D media (Google Earth) loaded on the fly as these simulations proceed.

And now a little bit more about the model extending the model visually to disseminate to stakeholders

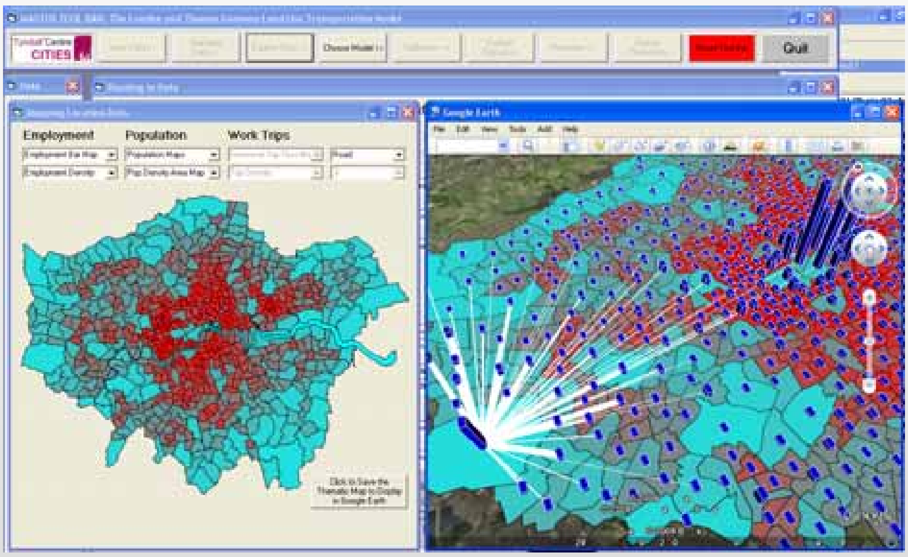
Extending the Software

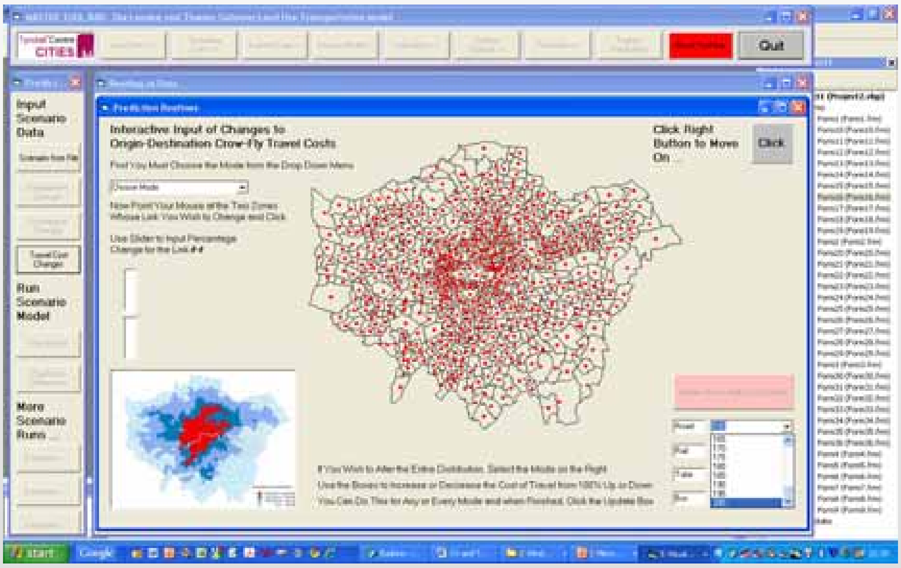
- Currently we do not have good zoom, pan, overlay facilities in the model due to difficulties of such programming in VB. I suspect these could be developed but we also need to share the data and the predictions and a quick possibility is to use a non-proprietary open map visualisation system to link on the fly to the model: this should be web-based
- The best way forward at present is to generate KML files in the program and then feed them to Google Earth where we have overlay, 3D, and external data facilities. You have seen this.

- In this way, we can extend massively our ability to visualise as well as providing a storage facility for the model input and output data
- What is impressive about this is that the speed of doing all this is not slower than the interactive program in VB



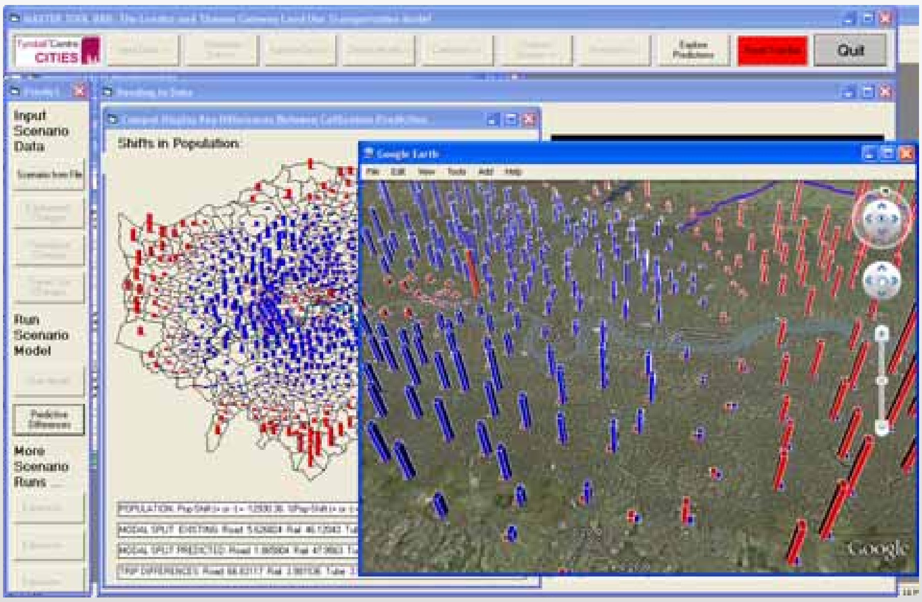
It is based on pragmatic use of available software but it generates familiarity. It also lets us store data in a convenient fashion and add other external data as KML files – particularly physical data which is hard to store in the customised software





Centre for Advanced Spatial Analysis

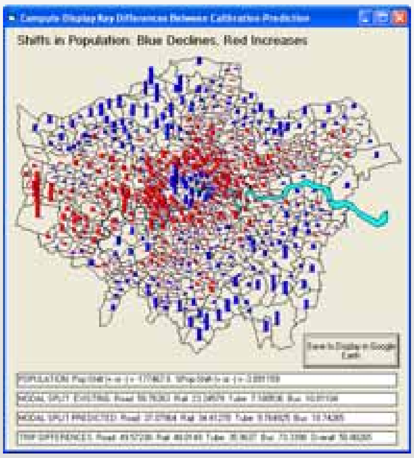
UCL



Centre for Advanced Spatial Analysis

UCL

- The following figures show what happens if gas costs rise by 100% i.e. double



Mode	Observed	Percent Shift
Road	39%	-50%
Rail	12%	+48%
Tube	33%	+36%
Bus	16%	+73%
Population Shift		4%

Conclusions: Predictions: Informing the Future

I know this is a lot of material to digest but I hope it has given you a sense of where I think we stand in geospatial modelling.

I believe many of the things that we have held sacred in science about repeatability, generalisation, simplicity and so on are under enormous scrutiny.

We stand at the focal point in this debate and need to engage in it. I have not spoken at all about how new data is enriching our world and changing what we are interested in and what we consider important

This too will change our science. We stand on a threshold of enormous change in how we develop our field.

If there is time,

I will answer any

Questions

www.casa.ucl.ac.uk

www.ComplexCity.info/tweetpad/