



### Opening Lecture 2 – September 4<sup>th</sup> 2014

## Land Use Transportation Interaction Models

### **Michael Batty**

m.batty@ucl.ac.uk

@jmichaelbatty

http://www.complexcity.info/
http://www.spatialcomplexity.info/





### **Outline**

- Preliminary Ideas:
  - a. LUTI Models: A Little Bit of History, Graphics Interfaces
  - b. Gravitation and Social Physics
  - c. Spatial Interaction Modelling: Some Basic Ideas
- The Tyndall Model: Integrated Assessment
- SIMULACRA: ARCADIA and SCALE Projects
- Requirements: The Model Design: Models Flows: Physical Movements, Money & the Residential Model
- The Visual Template: The Desktop Model: Running It
- Building a Web-Based Model Interface
- Data Bases: Location, Interactions & Networks
- Current Challenges for Immediate Developments





## **Preliminary Ideas**

a. LUTI Models: A Little Bit of History, Early Graphics

Early models: CATS 1955: The 1960s Models: large scale, aggregate data, policy led, suffered from all the problems of science in public policy that we now know

Largeness, remoteness from users, crude representation, limits on computation, poor links to policy

Lack of understanding of model outcomes

Statics versus dynamics – semi-dynamic models but most operational models predicated in terms of equilibrium and as we have seen, the most developed are structured in terms of a dynamic equilibrium

Disaggregation of sectoral activities





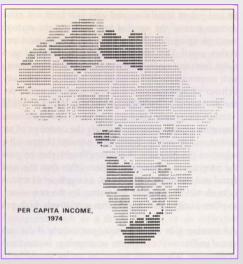
Early attempts at Visualisation: Traffic Flows in CATS, and Schmidt's model of the growth of East Lansing 1967

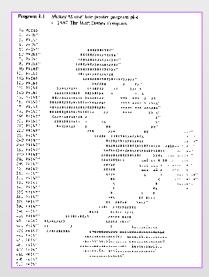
Harvard Lab: SYMAP – Symbol Mapping Systems, 1967-1970

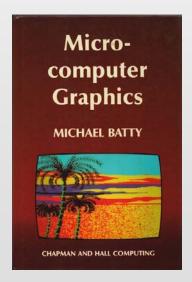
Early cathode ray displays, 1960

Apart from some SYMAP applications, my own attempts began in the early 1980s with the Melbourne Model



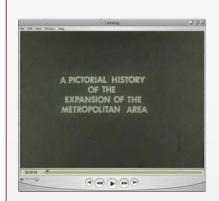
















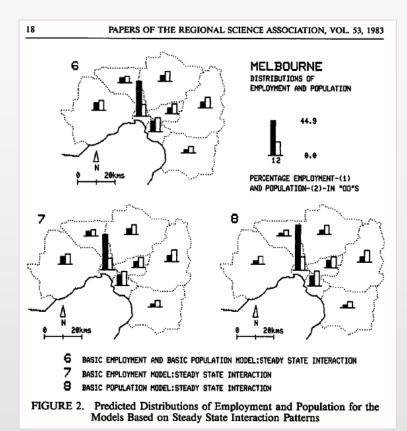


I worked on these models in the 1970s and published a book on them called Urban Modelling – you can get it online and if you go to my blog <a href="www.complexcity.info">www.complexcity.info</a> you can download it Our first foray into these large scale models and their visualisation was for Melbourne in 1982 and then development of a WATFor77 version in 1986. The Melbourne Version is primitive and not very nice to look at because we attached a video to the raster display device









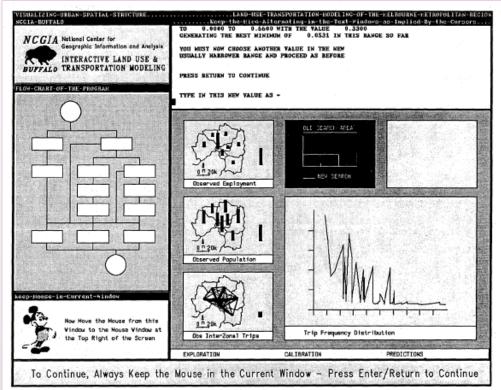


Figure 6. Progress in calibrating the model using dichotomous search.

From a VAX Terminal – A Raster 1982 From a Sun Workstation – Simple Windows - 1991

At the risk of the talk getting longer, I will show the 1986 demo – as some of you may not have been alive in 1986. It still runs under DOS on this laptop





## b. Gravitation and Social Physics

Let me begin to tell you how these models are built with a basic model of interaction- that was first articulated by Newton as his second law of motion – force is proportional to mass times acceleration

In more conventional terms relevant to this treatment, we might write the force between two bodies as

$$F_{12} \sim \frac{M_i M_j}{d_{ii}^2} = G \frac{M_i M_j}{d_{ii}^2}$$

There is a very long history of analogies between force and social interaction going back to Newton himself. There are many references in the books listed at the end of this talk





Let me immediately generalise this and say that we need to define many interactions – we break our system in to areas or points which we define as origins and destinations i and j And then we measure the distance as in von Thunen not as distance per se but as travel cost or rather generalised cost  $c_{ij}$ 

We also define the mass at the origins and destinations as  $O_i$  and  $D_j$  and we then write the conventional spatial interaction of gravity model as

$$T_{ij} \sim \frac{P_i P_j}{c_{ij}^2} = K \frac{P_i P_j}{c_{ij}^{\lambda}}$$

where K is now the 'gravitational' constant and  $\lambda$  a parameter





This is the model that has been used for years but in the late 1960s and 1970s various researchers cast it in a wider framework – deriving the model by setting up a series of constraints on its form which showed how it might be solved and produced various generating mechanisms that could generate consistent models

The *constraints logic* led to **consistent accounting** 

The *generative logic* lead to analogies between **utility and entropy maximising** and opened a door that has not been much exploited to date between entropy, energy, urban forms physical morphology and economic structure

In particular the economic logic rather than the energy entropy logic was called choice theory, specifically discrete choice





Now to introduce all this, we need to define some more terms

We will refer to the size of volume of origins and destinations not as population P but as  $O_i$  and  $D_j$  assuming they are different from one another. We will also assume that the inverse square law on distance or travel cost does not apply and that whenever  $c_{ij}$  appears it will be parameterised with a value that varies which we call  $\lambda$ . Vito used a model like this in his first talk to take account of the weight on a link.

We will assume trips are as we have defined them as  $T_{ij}$  but we can also normalise trips by their total volume T to produce probabilities





$$p_{ij} = \frac{T_{ij}}{T} = \frac{T_{ij}}{\sum_{ij} T_{ij}}$$

Note that we use summation extensively in what follows

We must move quite quickly now so let me introduce the basic constraints on spatial interaction and then state various models

The constraints are usually specified as origin constraints and destination constraints as

$$O_i = \sum_j T_{ij}$$

$$D_j = \sum_i T_{ij}$$

And we can take our basic gravity model and make it subject to either or both of these constraints or not at all





So what we get are four possible models

#### **Unconstrained**

$$T_{ij} = KO_i D_j c_{ij}^{-\lambda}$$

<u>Singly (Origin) Constrained</u> so that the volume of trips at the origins is conserved

$$T_{ij} = A_i O_i D_j c_{ij}^{-\lambda}$$

<u>Singly (Destination) Constrained</u> so that the volume of trips at the destinations is conserved

$$T_{ij} = B_j O_i D_j c_{ij}^{-\lambda}$$

**Doubly Constrained** trip volumes at origin-destinations conserved

$$T_{ij} = A_i B_j O_i D_j c_{ij}^{-\lambda}$$

The first three are location models, the last is a traffic model





Ok, so what are these parameters that enable the constraints to me met — well they can be very easily produced by summing each model over the relevant subscripts — ie origins or destinations and then substituting & rearranging I will do this but I will leave you to work through the algebra if you want to but it isn't necessary as I just want to impress the main ideas. Here are the factors which are sometimes called balancing factors

Unconstrained  $K = T/\sum_{i}\sum_{j}O_{i}D_{j}c_{ij}^{-\lambda}$ Origin Constrained  $A_{i} = 1/\sum_{j}D_{j}c_{ij}^{-\lambda}$ Destination Constrained  $B_{j} = 1/\sum_{i}O_{i}c_{ij}^{-\lambda}$ Doubly Constrained  $A_{i} = 1/\sum_{i}B_{j}D_{j}c_{ij}^{-\lambda}$   $B_{j} = 1/\sum_{i}A_{i}O_{i}c_{ij}^{-\lambda}$ 





Let me illustrate in two ways how we can build comprehensive models using this framework

If we say that residential location depends on not only travel cost but also on money available for housing we argue as before that

- The model is singly constrained we know where people work and we want to find out where they live – so origins are workplaces and destinations are housing areas
- The model then lets us predict people in housing
- We argue that people will trade-off money for housing against transport cost

And we then set up the model as follows





Using the trip activity-volume form, we get

$$\sum_{j} T_{ij} = O_{i}$$

$$\sum_{i} \sum_{j} T_{ij} c_{ij} = C$$

$$\sum_{i} \sum_{j} T_{ij} R_{j} = R$$

$$leads to$$

$$T_{ij} = A_{i} O_{i} \exp(\Re R_{i}) \exp(-\lambda c_{ij})$$

Note that we now add a constraint on money available for housing (like rent)  $R_j$ . We can of course find out from this location model how many people live in destination housing zones, so again it is a distribution as well as a location model

$$P_j = \sum_i T_{ij}$$





We can extend this model in lots of ways and we will show some of these later. We also can think about disaggregating the model into different transport modes – let us call each mode k and then set up the model so that we can predict  $T_{ij}^k$  as follows

The model is singly (origin) constrained because we want to predicts how many people travel from work to home. Given we know how many people work at origins, and we want to predict what mode of transport k they travel on. Then

$$\sum_{j} \sum_{k} T_{ij}^{k} = O_{i}$$

$$\sum_{i} \sum_{j} \sum_{k} T_{ij}^{k} F_{j} = F$$

$$\sum_{i} \sum_{j} T_{ij}^{k} c_{ij}^{k} = C^{k}$$





And the model can be specified as

$$T_{ij}^{k} = O_{i} \frac{F_{j} \exp(-\lambda^{k} c_{ij}^{k})}{\sum_{j} \sum_{k} F_{j} \exp(-\lambda^{k} c_{ij}^{k})} = O_{i} \frac{F_{j} \exp(-\lambda^{k} c_{ij}^{k})}{\sum_{j} F_{j} \sum_{k} \exp(-\lambda^{k} c_{ij}^{k})}$$

Note that the mode split is a ratio of the competitive effects of each travel cost, that is

$$\frac{T_{ij}^k}{T_{ij}^\ell} = \frac{\exp(-\lambda^k c_{ij}^k)}{\exp(-\lambda^\ell c_{ij}^\ell)}$$

In short the model is not only distributing trips so that locations compete but also that modes compete BUT modes do not compete per se with locations

Now let us see how we can build this model for real





# The Tyndall Model: Integrated Assessment of Climate Change in the London Region

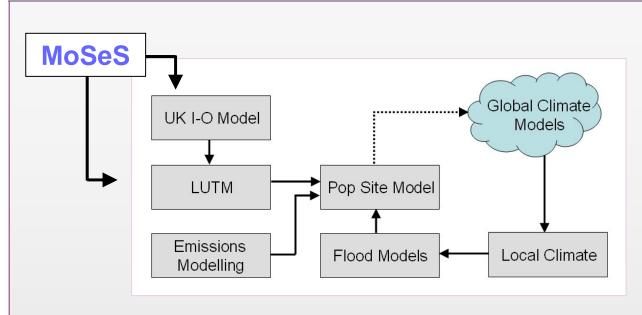
Our first model is a residential location model based on four modes but it is part of a much bigger project involving stringing different models together to look at integrated assessment of sea level rise and its impact on population in the London region. The reason why I have stressed visualisation as a key theme today is that one of our guiding principles in this modelling was to communicate the models as easily and as effectively as possible to our stakeholders

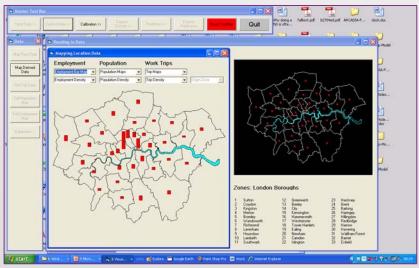
These were largely policy analysts and had some expertise in climate change problems and some knowledge of models of various kinds but not any one particular model

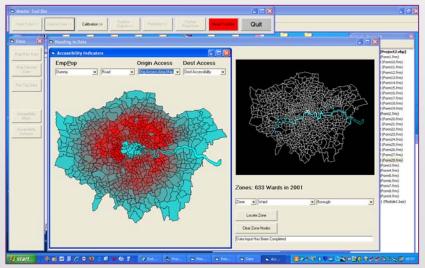
We also needed to communicate these ideas to other modellers, such as the flood modellers, the input-output, people etc.





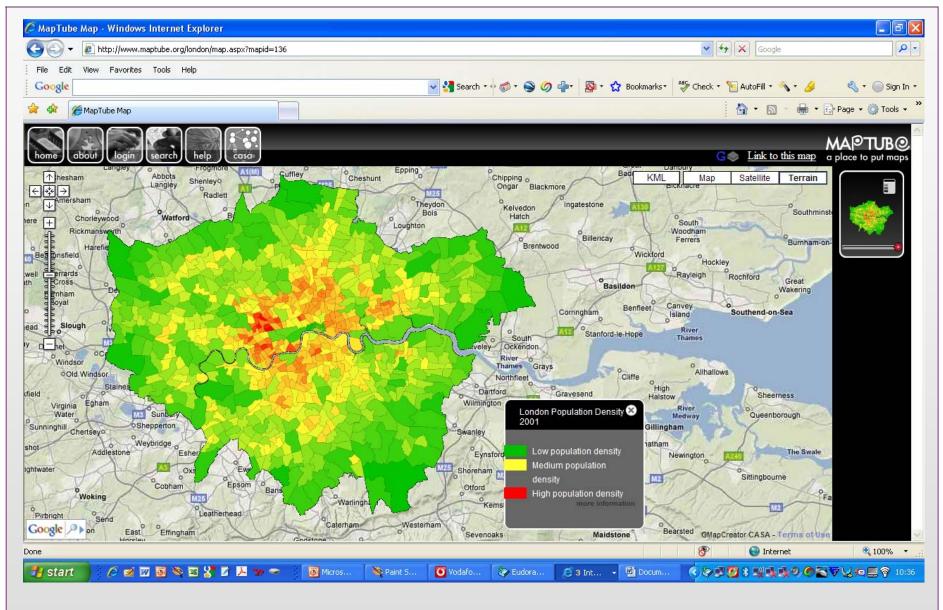














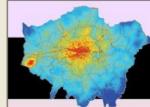




#### The Model Interface

London and the Thames Gateway Land Use Transportation Model





#### Cities Research Programme

## Tyndall°Centre

for Climate Change Research









This program is a rudimentary land-use transportation model built along classical lines which allocates population and employment to small zones of the urban system. It uses spatial interaction principles which bind the population sector (residential or housing) to employment sector (work or industrial and commercial) through the journey to work (work trips) and the demand from services (which loosely translate into trips made to the retail and commercial sector).

The model is being built for Greater London and the Thames Gateway at ward level - 633 in all - so that it can be used in a wider process of integrated assessment focussed on assessing the impact of climate change on small areas in this metropolitan region. In particular rises in sea level and pollution are key issues, and as such the model sits between aggregate assessments of environmental changes associated with global and regional climate change models and environmental input output models, and much more disaggregate models related to the detailed hydrological implication of long term climate change.

The programme enables the user to read in the data and explore it spatially, to calibrate the parameters of the model and explore its outputs spatially and to engage in various predictions ranging from the typical business as usual scenarios to much more radical changes posed limits on spatial behaviour which either result from climate change and, or mandated by government. The predictions and scenarios are intended to go out to 2100 and thus the model is largely designed as a sketch planning tool.

These various stages of the model contained in a master tool bar which is activated when the GO! button is pressed on this screen. The master tool bar enables the users to proceed through the various stages indicated and to display outputs in map and statistical form at any stage.

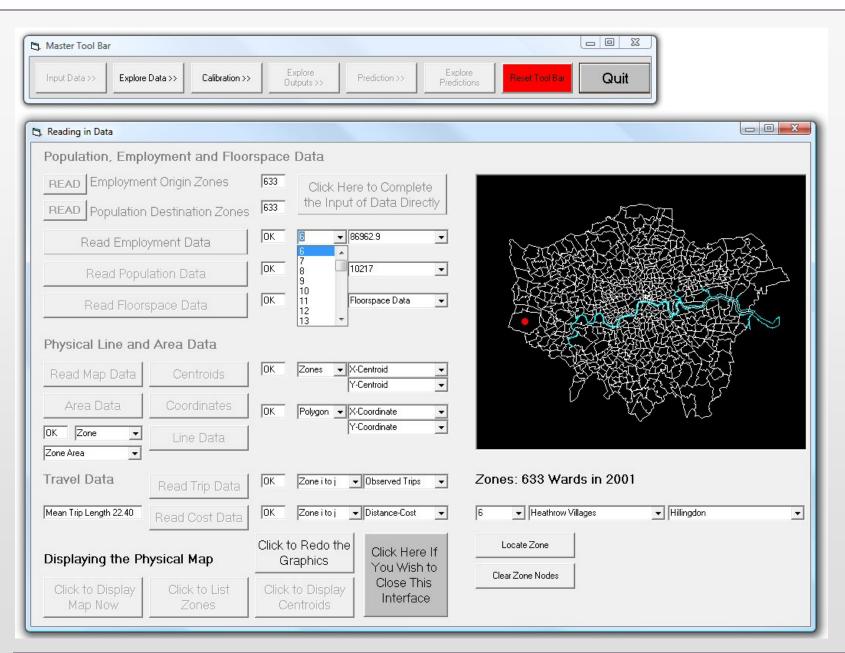




Program Manual

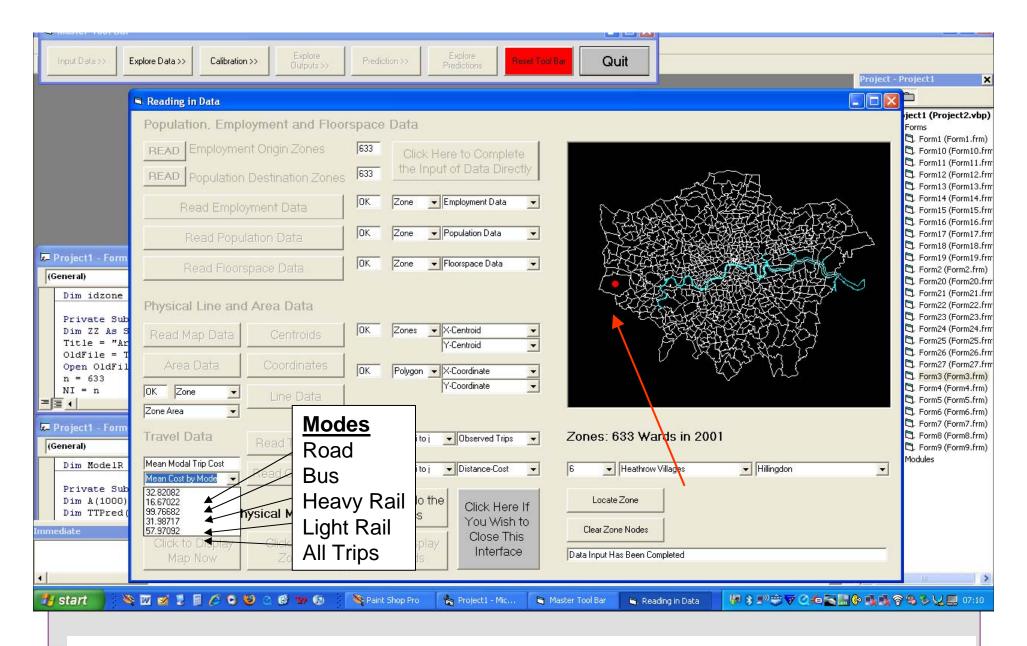








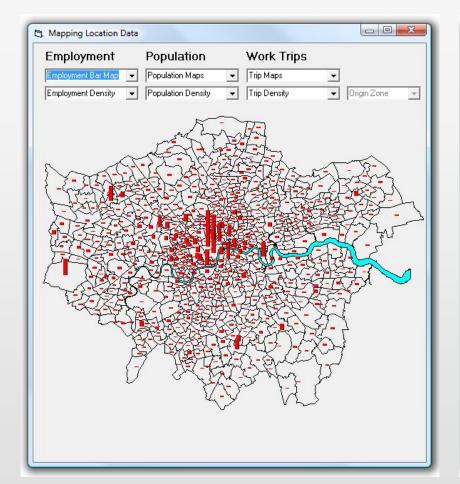


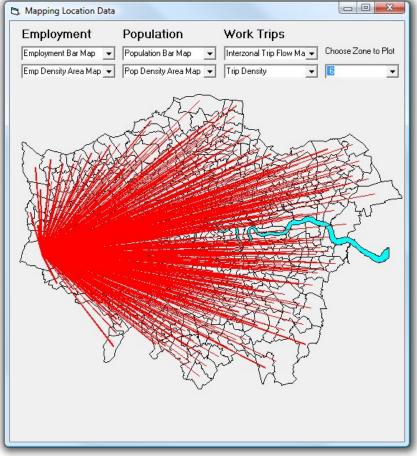


Road: 38%; Bus: 12%: Heavy Rail: 12%: Light Rail 19%; Other (Walk, Bike, Fly): 19%



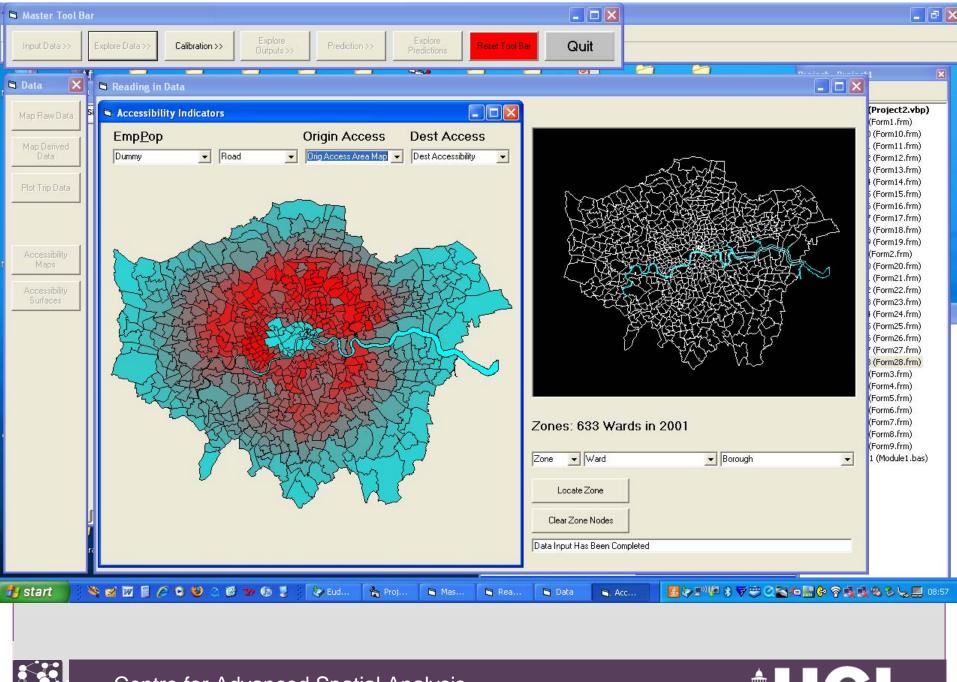










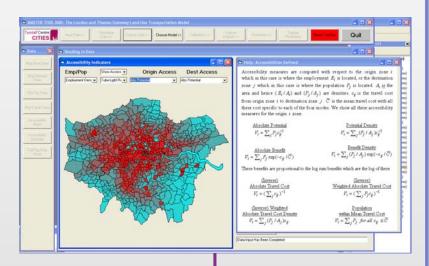






#### Accessibility from the LUTM model

## 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  $(E_i \mid A_i)$  and  $(P_j \mid A_j)$  are densities.  $c_{ij}$  is the travel cost from origin zone i to destination zone j.  $\overline{C}$  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
$$V_i = \sum_{i} P_i c_{ij}^{-1}$$

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

\_ 🗆 🗆 🗙

$$V_i = \frac{\underline{\text{Absolute Benefit}}}{\sum_j P_j \exp(-c_{ij} / \overline{C})}$$

$$V_i = \sum_{j} (P_j \mid A_j) \exp(-c_{ij} \mid \overline{C})$$

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

$$\frac{\text{(Inverse)}}{\text{Absolute Travel Cost}}$$

$$V_i = \left(\sum_{i} c_{ij}\right)^{-1}$$

$$\frac{\text{(Inverse)}}{\text{Weighted Absolute Travel Cost}}$$

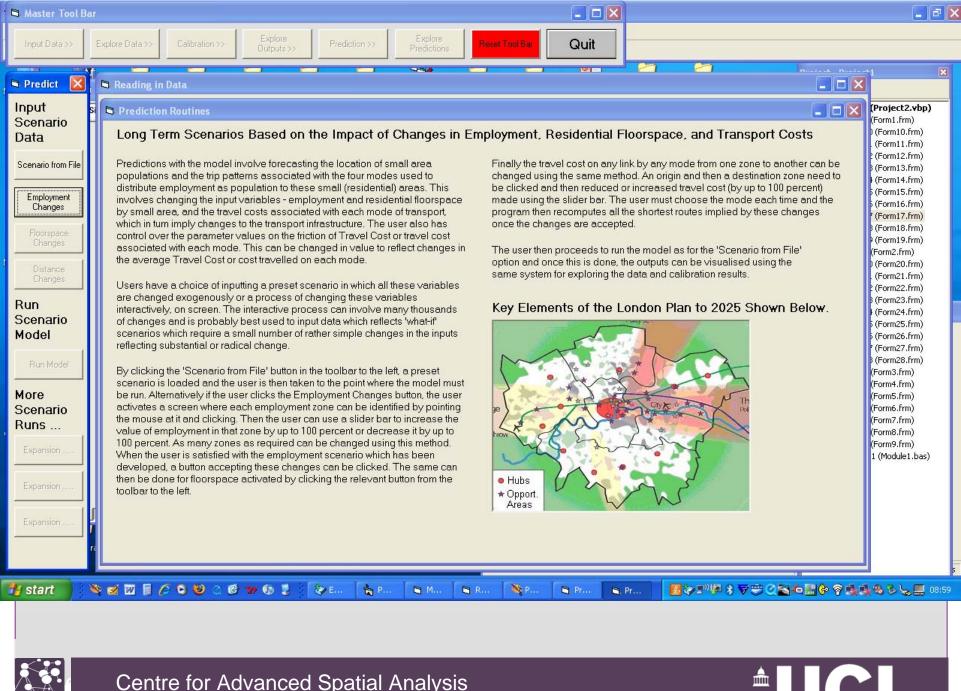
$$V_i = \left(\sum_j P_j c_{ij}\right)^{-1}$$

 $\frac{\text{(Inverse) Weighted}}{\text{Absolute Travel Cost Density}}$   $V_i = \sum_{j} (P_j \mid A_j) c_{jj}$ 

 $\frac{\text{Population}}{\text{within Mean Travel Cost}}$   $V_i = \sum_{i} P_j \text{ for all } c_{ij} \leq \overline{C}$ 

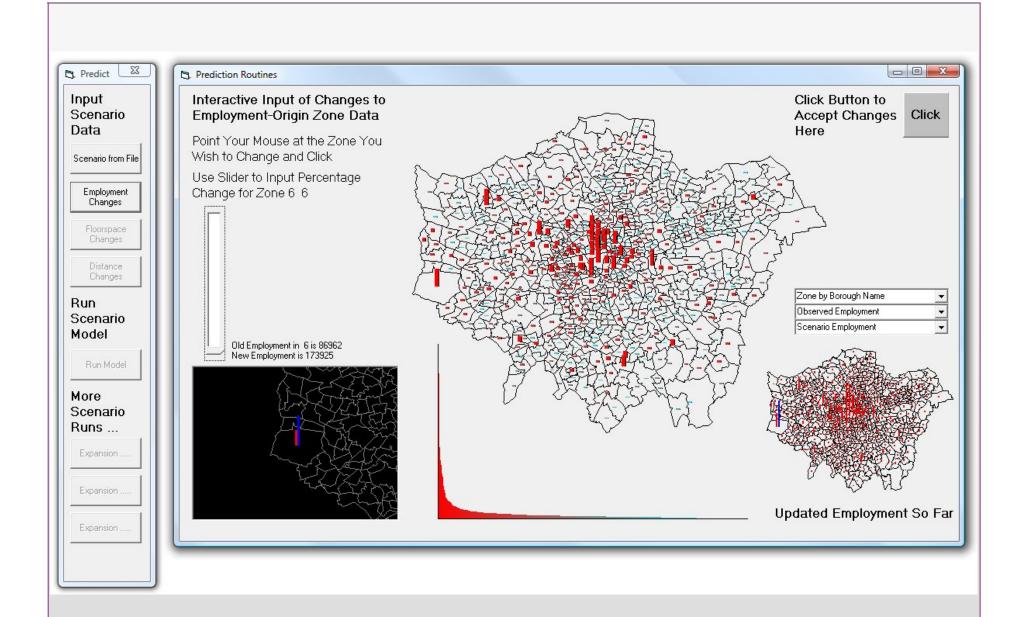






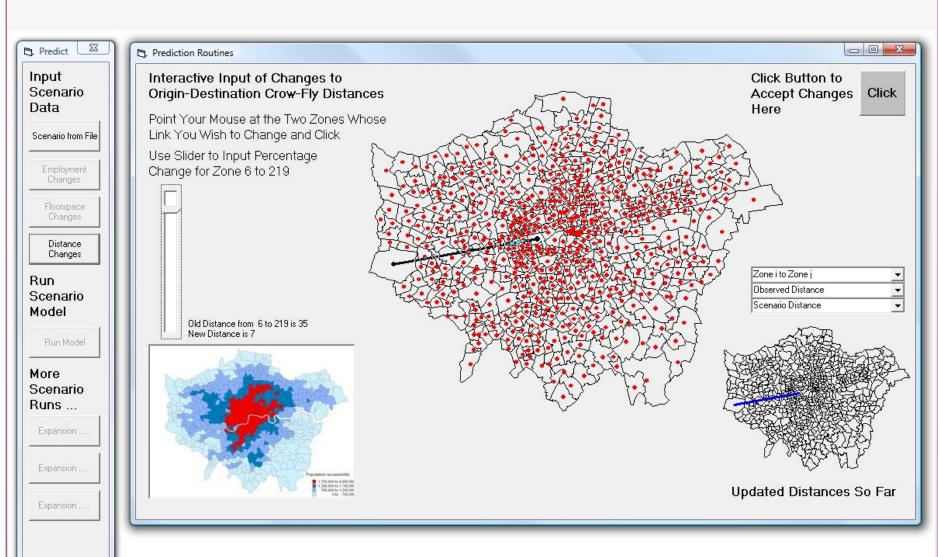










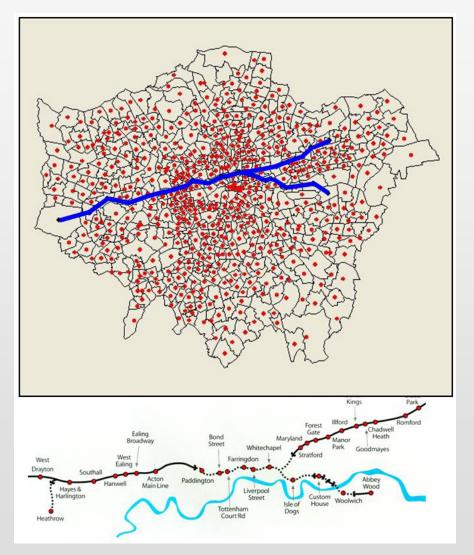


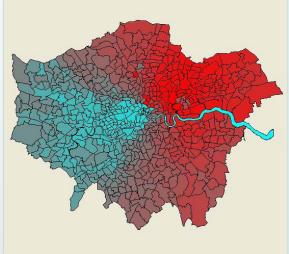
Let us <u>run</u> the model... I need to go to my folder...>>

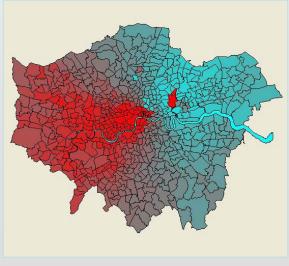




## Testing the impact of Cross Rail-1

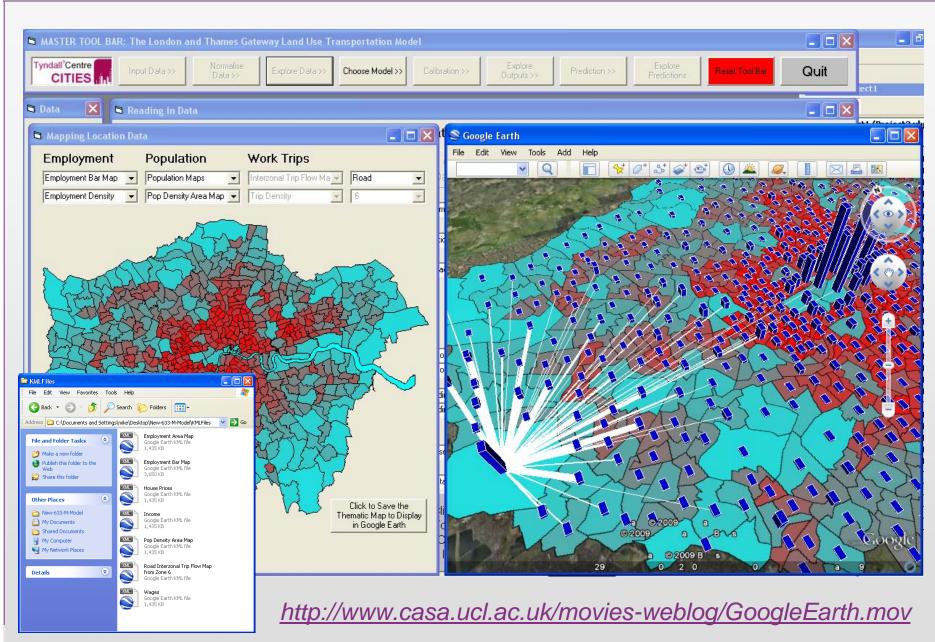








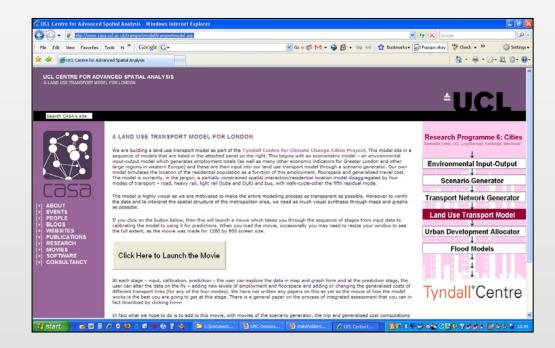








## For a very old movie of all this go to our old web site <a href="http://www.casa.ucl.ac.uk/transportmodel/transportmodel.asp">http://www.casa.ucl.ac.uk/transportmodel/transportmodel.asp</a>



We need to re-track a bit now and for our full model we need to say something more about these kinds of spatial interaction models and how they can be extended





## SIMULACRA: ARCADIA and SCALE Projects

SIMULACRA<sup>1</sup> is a generic set of models that we are building for a series of projects, first the ARCADIA project that is an extension of Tyndall, and then for another EPSRC Project called SCALE which deals with energy change in large cities

We are very keen on building a model framework in which we can develop many different variants, easily and quickly.

I think we are now in a position in this field where we can and should develop lots of variants, which test the robustness of any approach while at the same time, enabling models to be tuned to the problem in hand. But first let me sketch the integrated model – which involves coupling sectors

<sup>1</sup>SIMulation of <u>Urban Landuse</u>, <u>And Commercial and Residential Activities</u>





## Modular Modelling: Coupled Spatial Interaction

Now we have a module for one kind of interaction – consider stringing these together as more than one kind of spatial interaction

Classically we might model flows from home to work and home to shop but there are many more and in this sense, we can use these as building blocks for wider models. This is for next time too

What we will now do is illustrate how we might build such a structure taking a journey to work model from Employment to Population and then to Shopping which we structure as --





First we have the journey from work to home model as

$$T_{ij} = E_i \frac{F_j \exp(-\lambda c_{ij})}{\sum_{j} F_j \exp(-\lambda c_{ij})} , \sum_{j} T_{ij} = E_i$$

$$P_j = \alpha \sum_{i} T_{ij}$$

And then the demand from home to shop

$$S_{jm} = P_j \frac{W_m \exp(-\beta c_{jm})}{\sum_{m} W_m \exp(-\beta c_{jm})}, \quad \sum_{m} S_{jm} = P_j$$

$$S_m = \sum_{j} S_{jm}$$

And there is a potential link back to employment from the retail sector  $E_m = f(S_m)$ 





- Ok so this is the model structure iterative coupling of sectors to take account of simultaneity and we also want to be able to do the following:
- Alter and aggregate the zoning system quickly and easily, on the fly almost
- Alter by adding and deleting different model sectors, so for example running a model based on simply the retailing and other employment sectors without the residential and so on
- Subjecting the model to various kinds of physical constraints, at will and according to external policies
- Extending all sectors to not only predict endogenous activities but to also be subject to exogenous inputs of the same
- To interface the models easily and quickly with other sectoral models, demographic and more established transport models





# Requirements: The Model Design: Models Flows: Physical Movements, Money & the Residential Model

We will now show the current model to present the logic of our framework. Our model has now been scaled up massively to include the outer met area – 1767 zones (33 – 633 – 1767)

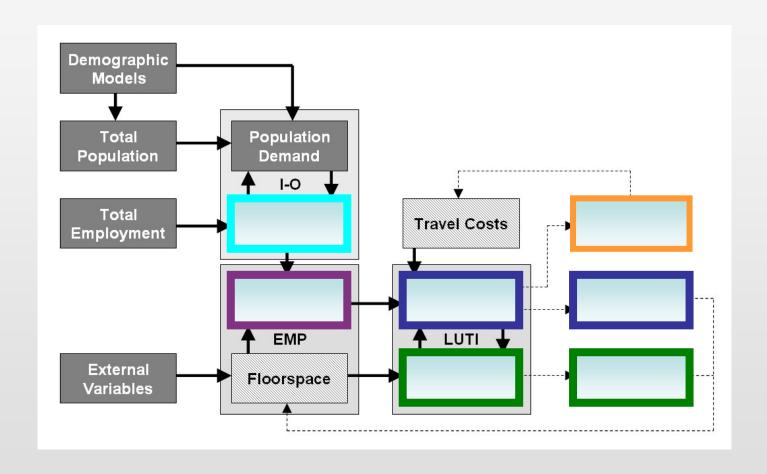
It is now a three sector model, not simply a residential location model as it includes internal employment location, retail location and residential location

So far, we do not have modal split or any disaggregation of the sectors but we will have five modes and then probably 5 population categories and maybe 5 employment types in terms of occupations – in short the model will ultimately scale up to some 100 times the size of the current model





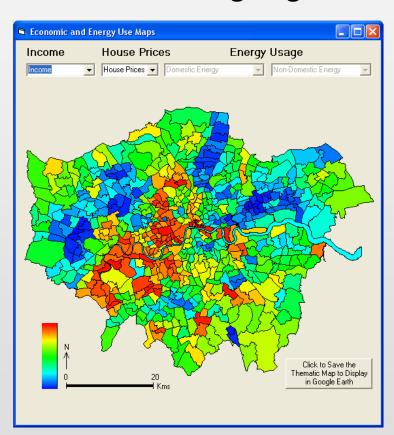
## The Structure of SIMULACRA

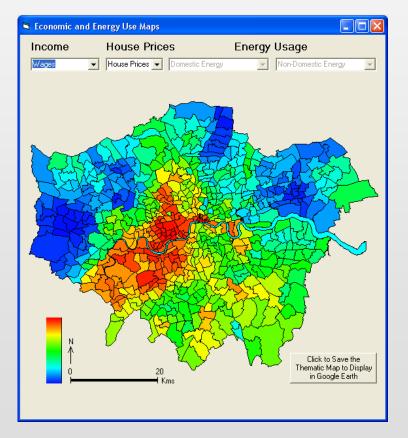






I want to just show very briefly the sort of data that we have in the money sector that is driving this variant of the model and also state the residential location equation so you have some sense of what is going on









#### And the model is formalised as

with travel as a difference or variance  $\sigma^2$  between these two sets of costs. Then, the system must satisfy the constraint

$$\sum_{i} \sum_{j} T_{ij} [(h_i + t_i) - (c_{ij} + \rho_j)]^2 = \sigma^2$$
 (11)

The model that is generated from this constraint and which is the alternative residential location model in the current model variant is

$$T_{ij} = E_i \frac{A_j \exp(-\lambda [(h_i + t_i) - (c_{ij} + \rho_j)]^2}{\sum_i A_j \exp(-\lambda [(h_i + t_i) - (c_{ij} + \rho_j)]^2} , \qquad (12)$$

which is subject to the usual origin constraint, generating population from equation (2) with (12) replacing equation (1).





## The Visual Template: The Desktop Model

Ok – let me quickly tell you our strategy – we are building a fully fledged model using state of the art software and various web–based interfaces which is highly visual and will be as fast as possible

We are also building a mirror model on the desktop which is my contribution to the project and this is a one window minimal model which is for comparative purposes and to enable the bigger model to be tested

This is the model I will now show and then I will sketch the bigger application very briefly which Camilo in our group is developing





#### This is the order in which the operations take place

Sequence of Model Functions

> Activity Totals

Map Graphics Parameter Values

Goodness of Fit Statistics: Deviations & r<sup>2</sup>

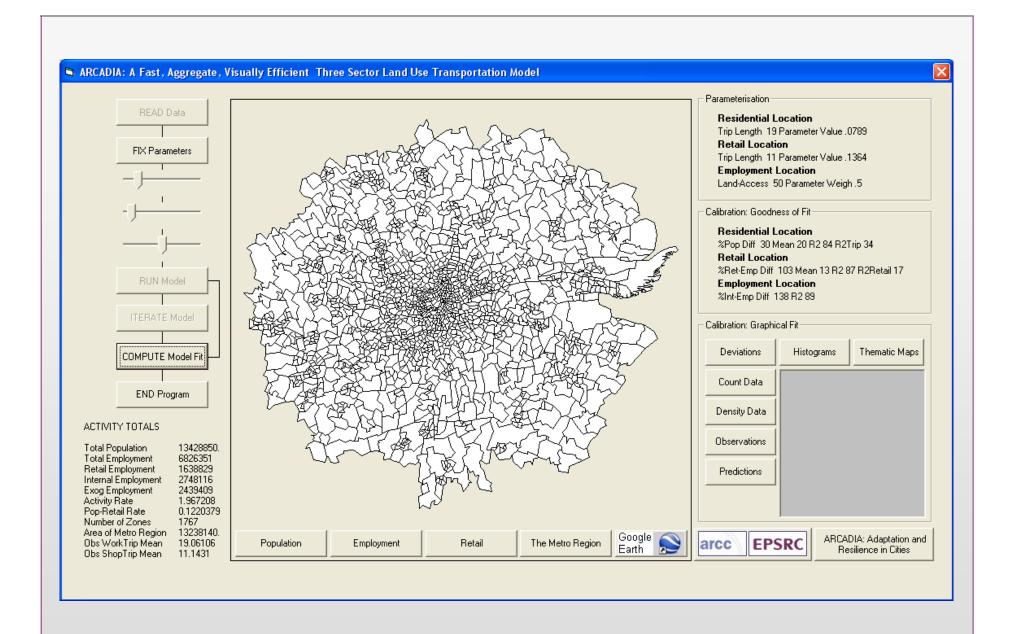
**Graphical Functions** 

Graph Data

Logo



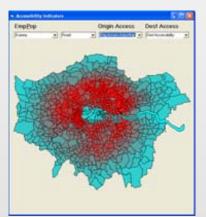


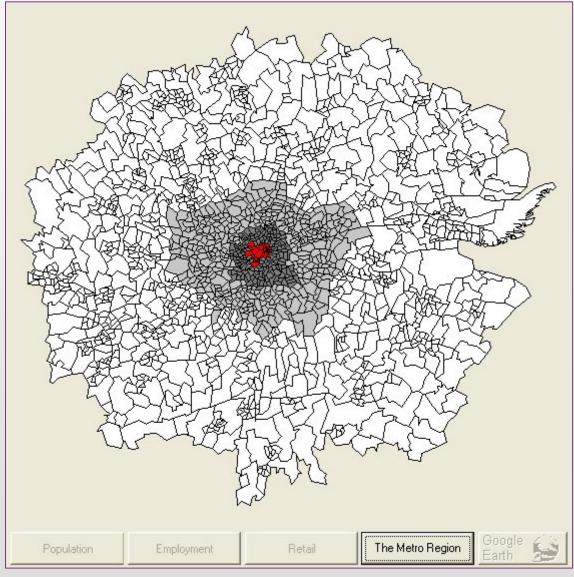








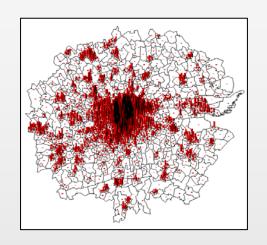


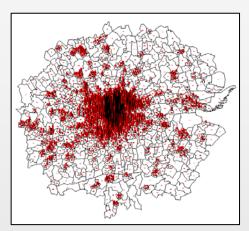


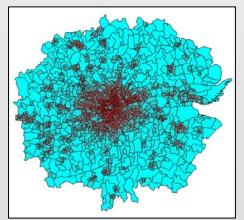


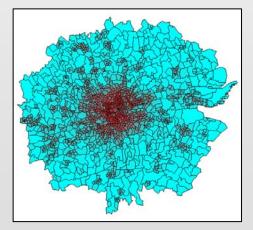


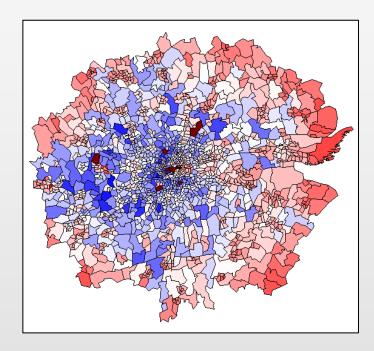
## Here are some sample outputs – I will run the model as speed is important – here goes











RUN





The employment model has been built more recently and it much wider in scale than the current 1767 model – it has 3202 or so zones and this is largely because we need to capture locational change in the wider region.

The model is econometric based on predicting floorspace which in turn predicts employment. The two typical equations

The models that have been fitted are variants of the generic structure that follows:

$$\begin{split} F_{j}^{z} &= a_{0} + \sum_{\substack{\ell=1,2,3\\\ell\neq z}} a_{k} F_{j}^{\ell} + a_{4} B_{j} + a_{5} D_{j} + a_{6} A_{j}^{z} + a_{7} T_{j} + a_{8} G_{j} + \\ &+ a_{9} H_{j} + a_{10} Q_{j} + a_{11} M_{j} + \sum_{\ell=1,2,3} s_{\ell} \delta_{j}^{\ell} + \sum_{\ell=1,2} f_{\ell} \delta_{j}^{\ell} \end{split} \tag{20}$$

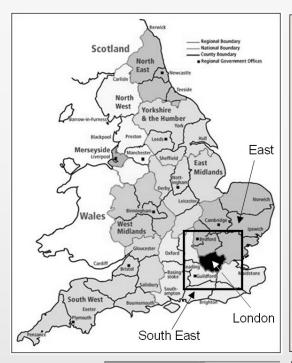
similar to that for floorspace and we state it as

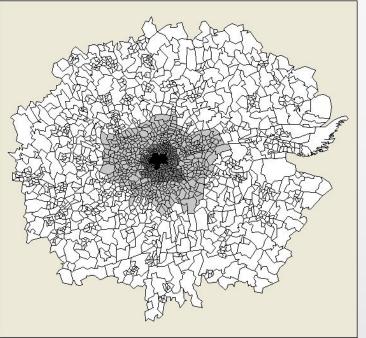
$$E_{jk} = b_0 + \sum_{z=1,2,3} b_k F_j^z + b_4 A_j^{PT} + b_5 A_j^R + b_6 M_j + b_7 J_j + b_8 H_j + b_9 h_j + b_{10} Q_j$$

$$\tag{21}$$



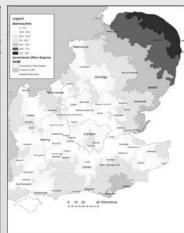


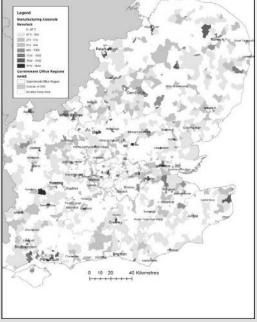








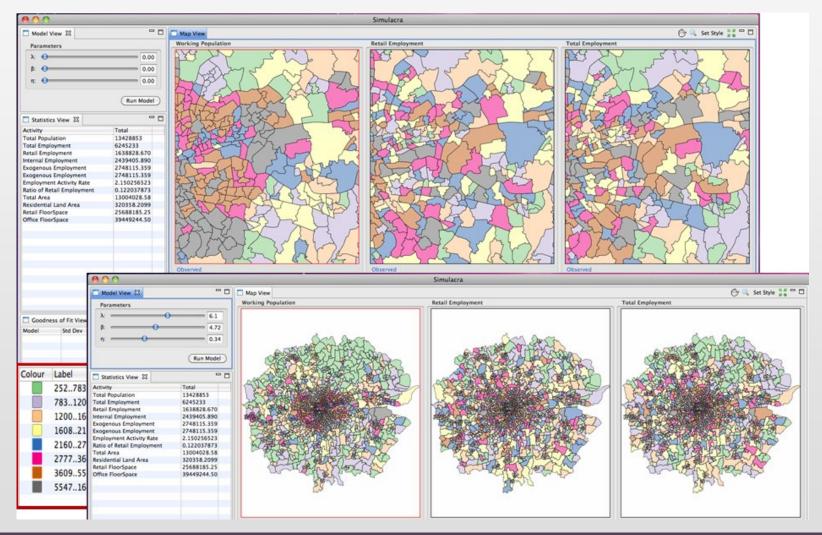






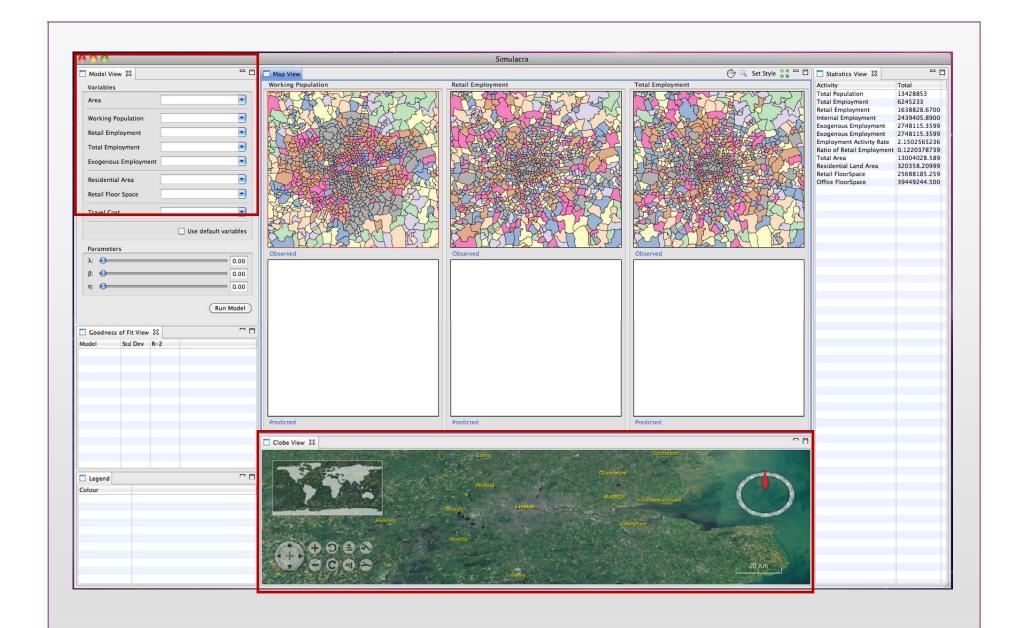


And here are some screen shots from the desktop version of the web based application – on Mac and PC





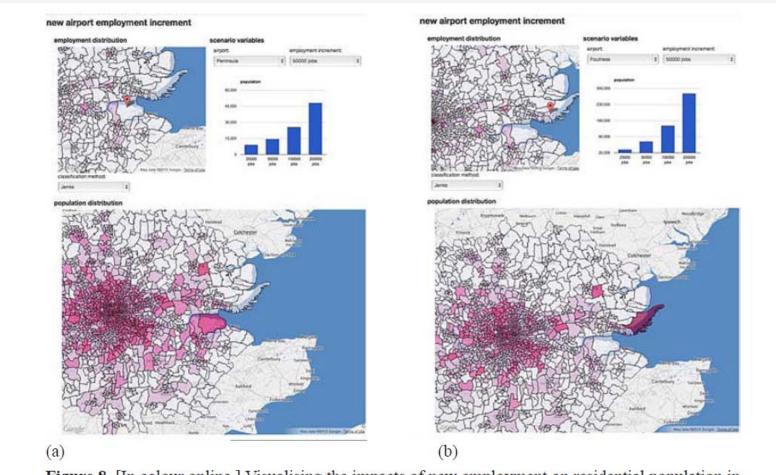








## And here are some examples of testing the impact of new airports on existing urban structure in the London region



**Figure 8.** [In colour online.] Visualising the impacts of new employment on residential population in the vicinity of two proposed sites for a new London airport: (a) Isle of Grain, (b) Maplin Sands.





### Data Bases: Location, Interactions & Networks

We have a big problem in getting the networks sorted out for the aggregate model as these networks are at a very fine scale

We need them to be at a coarser scale for the model as we need to do all the assignment and capacity checking at the level of the model

This is a long standing issue, we know, but we cannot afford to move down the local fine scale level to do the assignment of trips to the network because this would simply destroy our basic principle of accessibility of the model to users and also the speed requirements we need

We will show some of the detail we have by way of illustrating our work in progress.

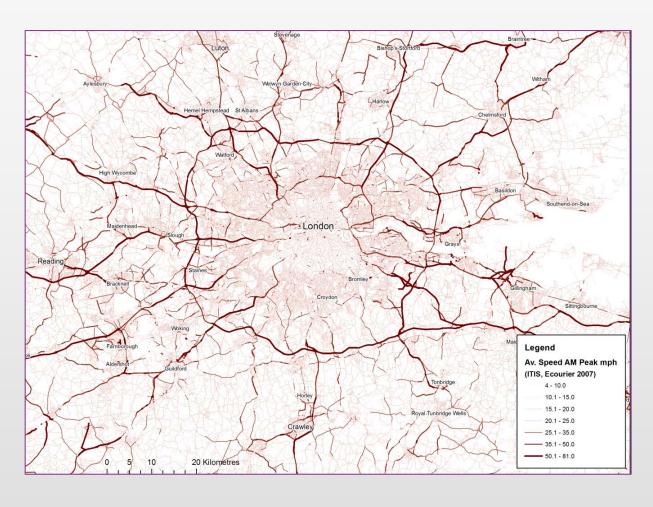




#### **Road Costs**

Used GPS data for realistic road speeds across the South East. Sourced from ITIS and Ecourier.

Future improvements with dynamic consideration of congestion.





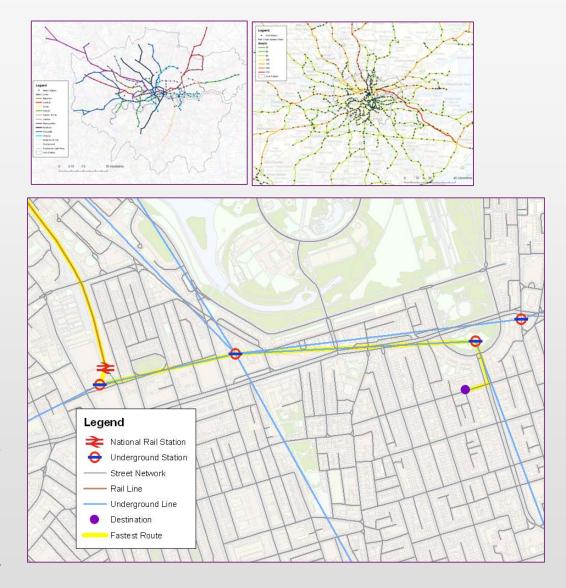


#### **Public Transport Costs**

Based on network geometry and timetabled services. Initially using model presented by Duncan last term. Allows multimodal PT trips.

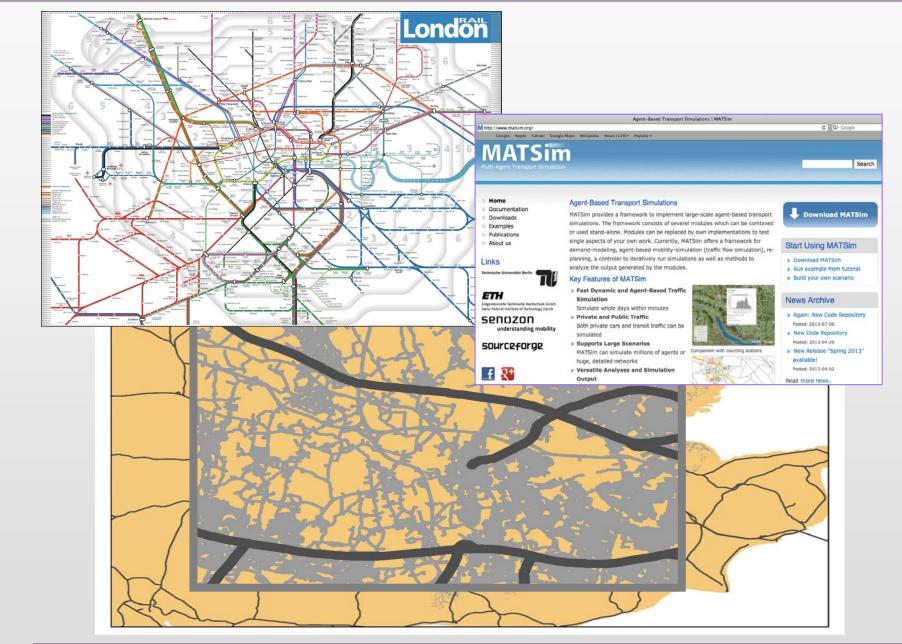
#### **TransXchange**

Full UK PT timetable available in XML format. Could be used to automate process of generating PT networks.





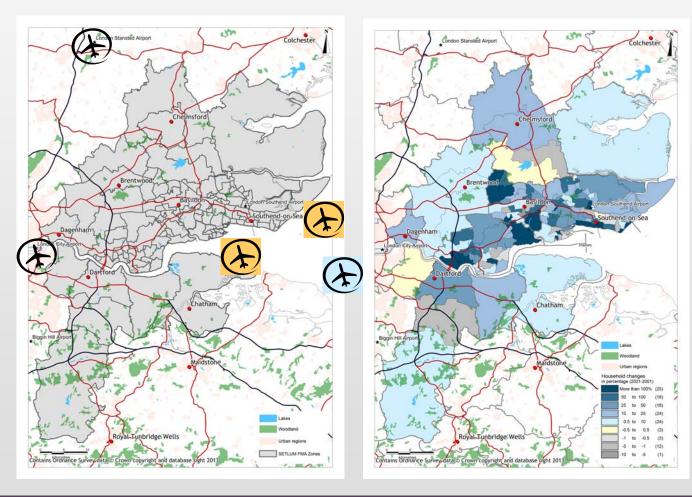








And we are linking all this to Melanie Bosredon's Work is adding energy issues and data to the The DSC SETLUM (South Essex Land Use Model)







### Key Challenges for Immediate Developments

- Speed of Models
- Quick and Effective Visualisation
- Running the Model with Users/Stakeholders
- Building a Residential Model Based on the Housing Market Cost, prices, travel and energy costs etc - The Wegener Principle
- Moving to a Semi-Dynamic Model with Inertia and Internal Migration
- A Local UK Dimension: Thinking of the Modelling Strategy as being Informed by National Data Bases such as Neighbourhood Statistics





## Thanks

http://www.spatialcomplexity.info/

http://www.complexcity.info/

http://www.mechanicty.info/

http://blogs.casa.ucl.ac.uk/

http://www.casa.ucl.ac.uk/

#### **Acknowledgements**

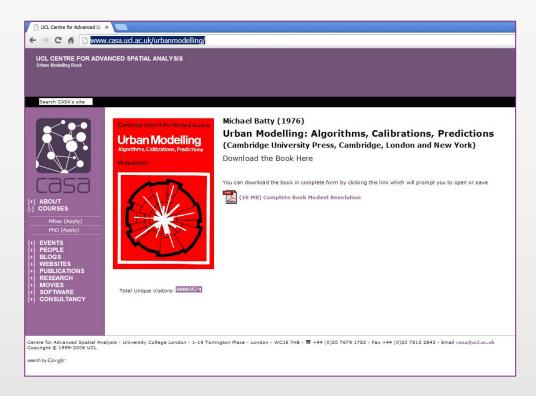
Melanie Bosredon, Richard Milton, Pete Ferguson, Joan Serras, Camilo-Vargas-Ruiz We are also working on these models with your own group here which leads the EU INSIGHT Project and EUNOIA which is led by NOMMON from Madrid











## http://www.casa.ucl.ac.uk/urbanmodelling/

Environment and Planning B 2013 40 987 - 1002

SIMULACRA: fast land-use-transportation models for the rapid assessment of urban futures

Michael Batty, Camilo Vargas, Duncan Smith, Joan Serras, Jon Reades, Anders Johansson

*Environment and Planning A* 2013 **45** 532 – 552

<u>Visually-Driven Urban Simulation: exploring fast and slow change in residential location</u>
Michael **Batty** 



