Simulating Very Large Scale Spatial Systems

Modelling Challenges & Software Systems

Michael Batty

m.batty@ucl.ac.uk
@jmichaelbatty

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

My Main Themes

- Scaling Up Models:

  to deal with much larger spatial systems and their environments

- Data Has Changed: A New Take on Big Data

  Big Data: even traditional data is now extensive and we need to deal with ever bigger sets of relations, networks

- Models Have Changed: Extensive Places, Many Interactions

  Models are bigger, more complex, many more types

- Models For 'Everywhere' & 'Everyone'

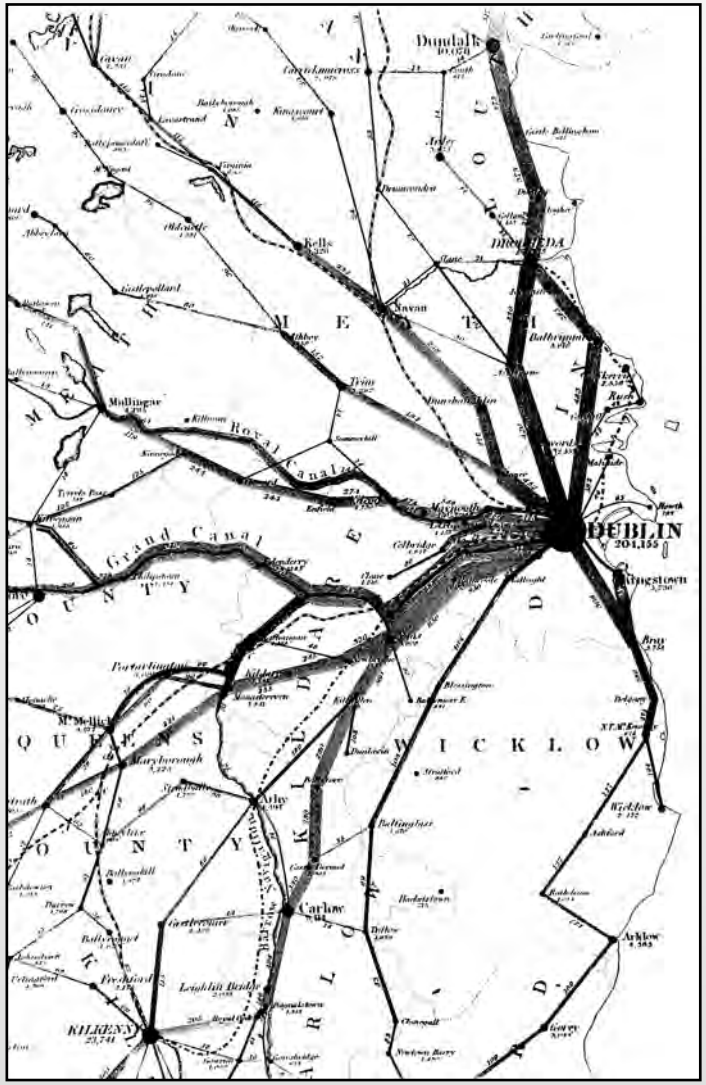
  Our model is a LUTI model for England and Wales

- Let Me Explain the Elements that Go Into this Project

Outline

- Dealing with Networks, Flows and Widespread, Country, even Continent Wide Relationships
- A Web-Based, Fine-Geographic Scale LUTI Model as a Online Tool for Testing Scenarios
- Some Examples: HS2, CrossRail 1 and 2, etc.
- Challenges to Be Addressed by the Model
- The National Infrastructure Models
- The Hardware-Software Facility at Rutherford-Appleton: DAFNI at STFC
- The Software Conundrum

Dealing with Networks, Flows and Widespread, Country, even Continent Wide Relationships



For a long time, since the beginning of the industrial revolution, we have been concerned with the impacts of very large infrastructure projects on economic and urban development.

Here is the classic example: Dublin 1837

For the last 50 years, land use transportation models of various kinds have been built to simulate these impacts.

A key problem is defining the boundary of the spatial system wrt its environment. Inevitably there are massive spillover effects and in a global world where everything is connected to everything else, we need to draw the boundary as wide as possible – everywhere.

Even Lieutenant Harness appreciated this back in Ireland 200 years ago for the infrastructure – the Irish railway – would affect the whole country

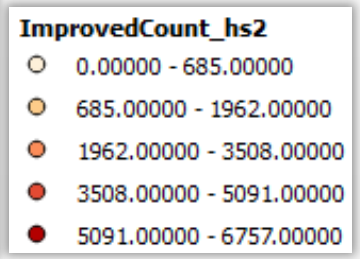
We face similar problems today. In Britain the railway era ended in the 1870s – 150 years – and we now need to renew

High speed trains are just one national project of many that we need to grapple with spatially





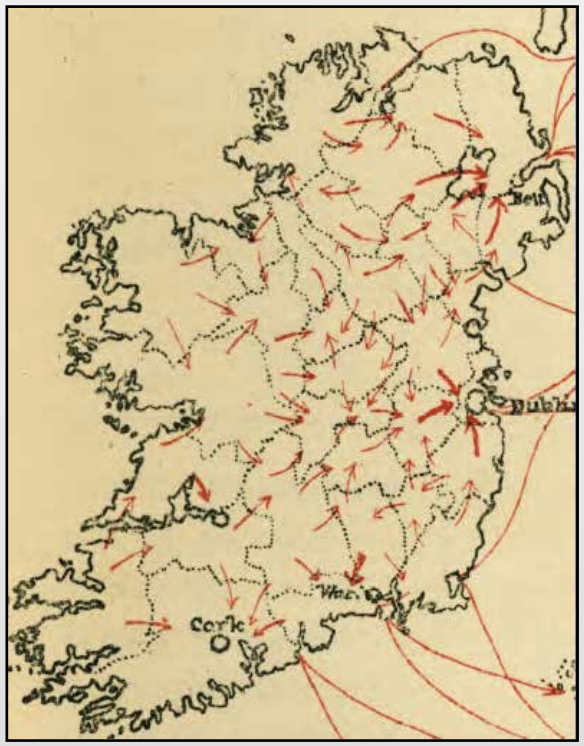
High Speed 2

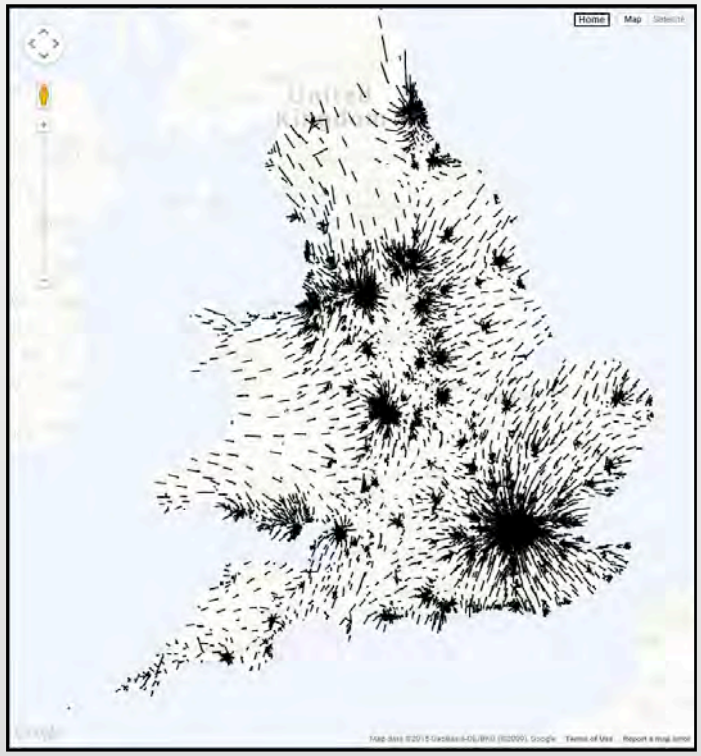




Showing improved journeys greater than 685

One brief digression about interrelationships – networks – which are hard to understand and almost impossible to visualise. 150 years ago, Ravenstein too summarised flows between county districts in Britain using hand drawn crude vector fields – we use these tools extensively now in our new models – this is big data where flows scale as the no. of nodes squared





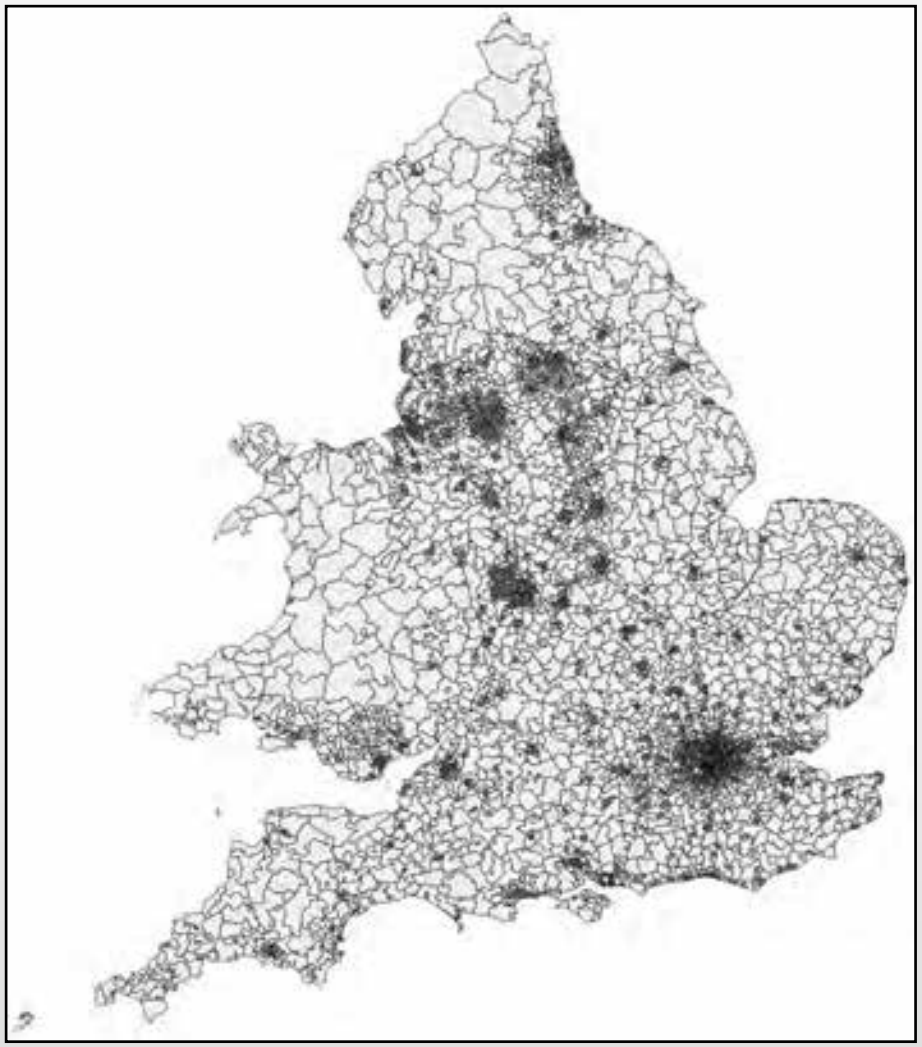
The Web-Based Fine-Geographic Scale LUTI Model as a Online Tool for Testing Scenarios

About 3 years ago, prompted by this massive interest in new national infrastructure projects which require scaling up the nation, we decided to extend our standard LUTI model to the whole UK. We have:

- *Extended it from greater London to England and Wales increasing our zones from some 2000 to 7000*
- *Ported it from desktop to web so that anyone can access it anytime from anywhere*
- *Packaged it so that the user – who we assume is 'informed' and a professional policy analyst – can test their own scenarios*
- *Enabled scenarios to be built interactively in the simplest form as new roads, rail, bus and new employment*

Let us look at some pictures of the set up before we explore the model, the software and results

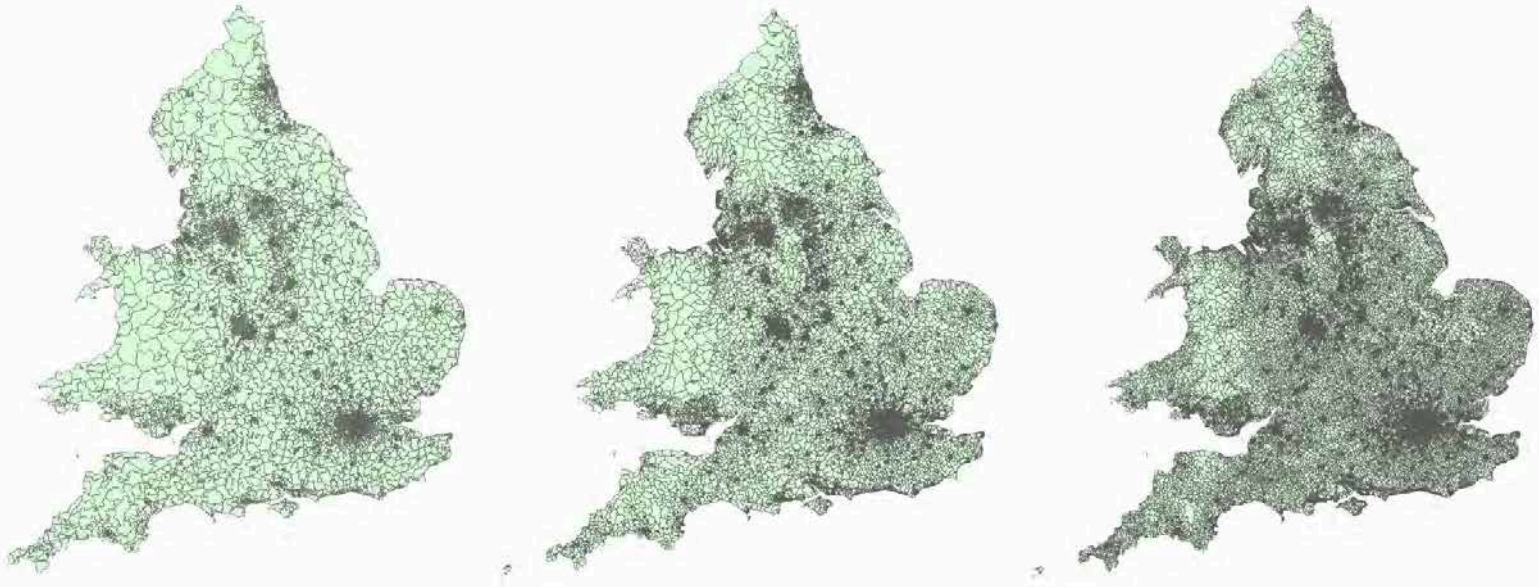
We are using what in the Population Census are called MSOAs of which there are n=7201 in England and Wales.



This generates $n^2=7201^2=51,854,401$, lot of numbers to keep in RAM of which we have several matrices where we also require very fast (<20 seconds model run times)

Scotland not yet added in because of Census issues. It will be.

We could have a finer level of geography – the LSOA although we cannot get flow data. We could go even lower to output area but we need to stick at the higher level – note that in E and W there are some 49 million people so the average size of MSOA is about 7000 persons: bigger than typical TAZ zones



(a) MSOA (A=7201) (b) LSOA (A=34753) (c) OA (A=181408)

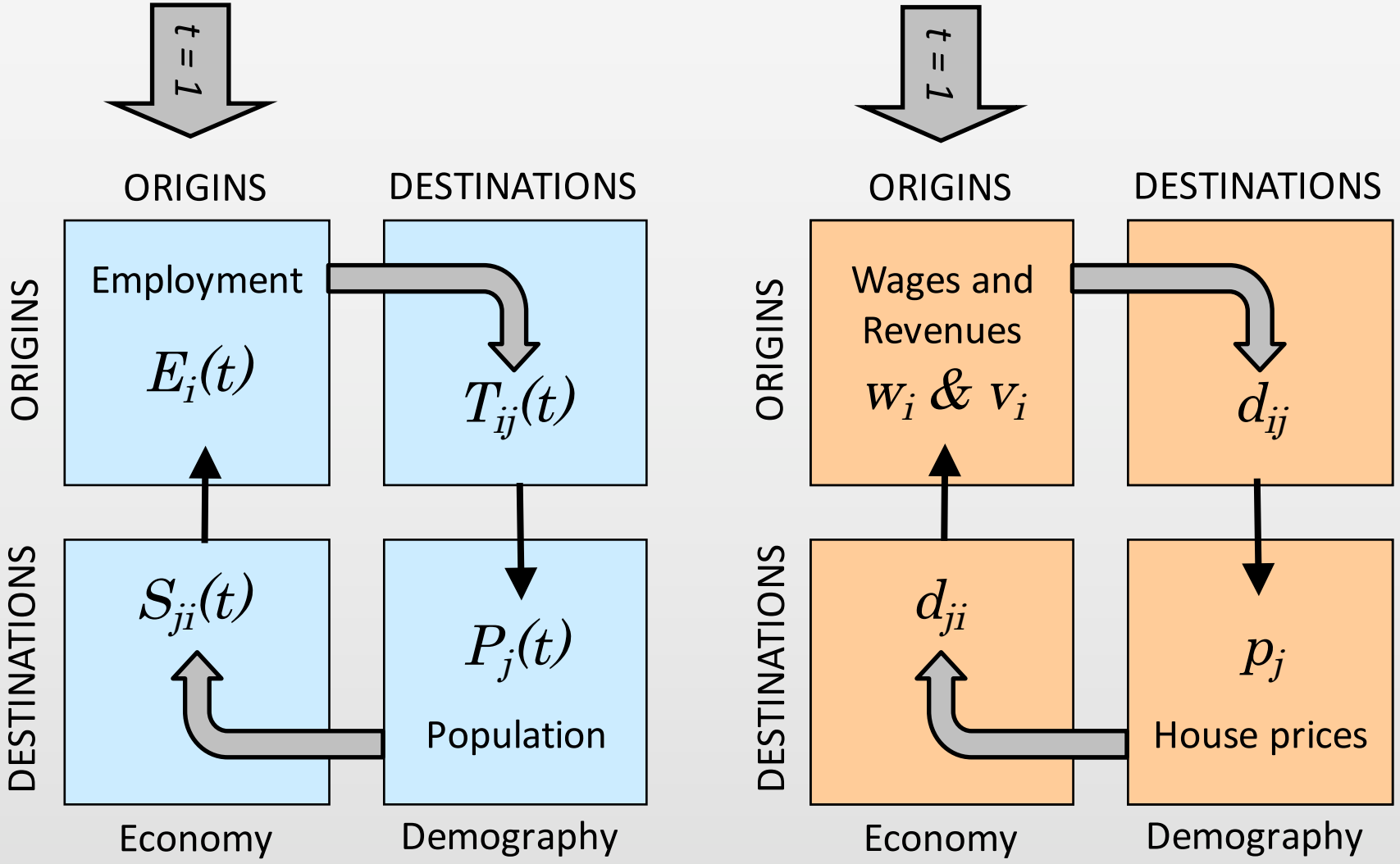
Figure 8.2: ONS Geographies for MSOA, LSOA and OA levels.

I need to do ***two things before I demo the model***: first say something about what the model is and then say something about the software design. I will return to software in the last bit of the talk. Ok the model then

LUTI models essentially conceive of the urban system as being composed of different sectors of employment and population which are linked using flow models. These are very close to discrete choice models in transport and to aggregate input-output models in regional science.

The models are usually disaggregated by mode of transport and have a quasi economic structure in which costs and prices determine who locate where and how the rudimentary market for land is cleared.

This is a typical module linking journey to work and journey to shop – in terms of activities, then prices



Essentially the module I have shown is a couple of nested spatial interaction-discrete choice-like models – one for residential, the other for the retail sector. Note that the model is cross-sectional static

$$T_{ij}^{k} = O_i \frac{D_j^{obs} \exp(-\beta^k d_{ij}^k)}{\sum_k \sum_j D_j^{obs} \exp(-\beta^k d_{ij}^k)} \qquad k = 1,\ k = 2,\ k = 3 \quad road, rail, bus$$

$$P_j = \sum_k \sum_i T_{ij}^k$$

$$S_{j\ell}^{k} = P_j \frac{W_\ell^{obs} \exp(-\lambda^k d_{j\ell}^k)}{\sum_k \sum_\ell W_\ell^{obs} \exp(-\lambda^k d_{j\ell}^k)}$$

$$O_\ell = \sum_k \sum_j S_{j\ell}^k$$

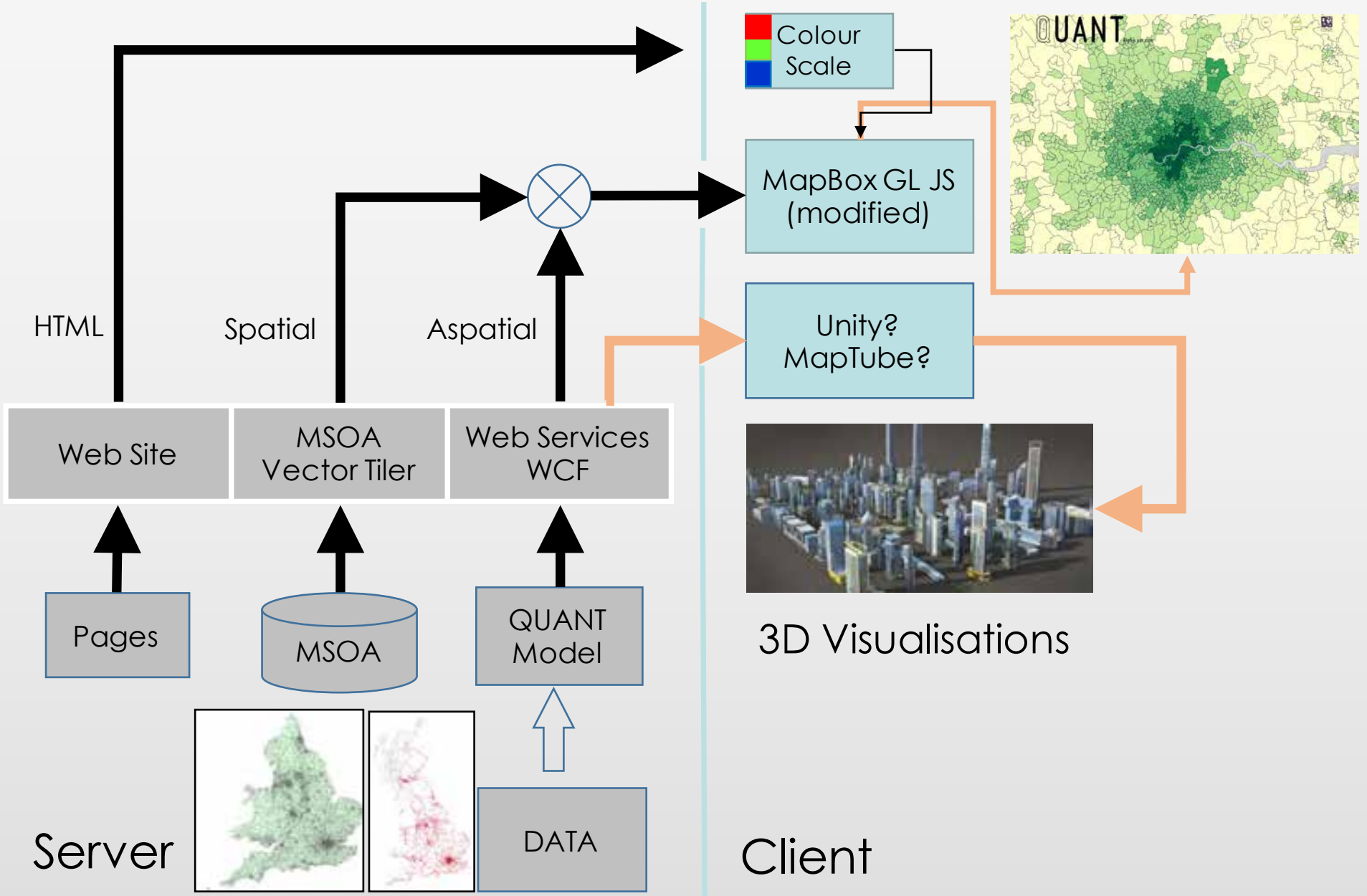
We also embed constraints posed by the supply side reflecting land and transport capacity in the models and this sets up various iterations.

I now need to say something about the software that we have developed – or rather Richard Milton our systems guy has developed.

We have a classic client-server architecture with all (most)computation on the server, all data storage with input and intermediate outputs on the server except for map updates that are a mixture of client and server side functions.

The client is thus of intermediate weight in that the Mapbox technologies used to render the maps and update them are largely on the client side but the zoom and pan capabilities are generated on the server side

Client & Server Architecture



There are many detailed elements of the software but the simulation is mainly written in C# with all web functionality using various kinds of Java script and related technologies from Google etc.

I cant give you chapter and verse on this as I don't do web programming but the whole focus has been so far on building a user interface where the model working on 7201 MSOAs (zones) runs in an average of 15 seconds – anything less is too slow for active interaction with users who are unlikely to tolerate lengthy progress bars.

The map downloads and zooms are quite fast but sometimes dependent on the network and the vector maps can be very slow.

Some Examples: HS2, CrossRail 1 and 2, etc.

The web link is

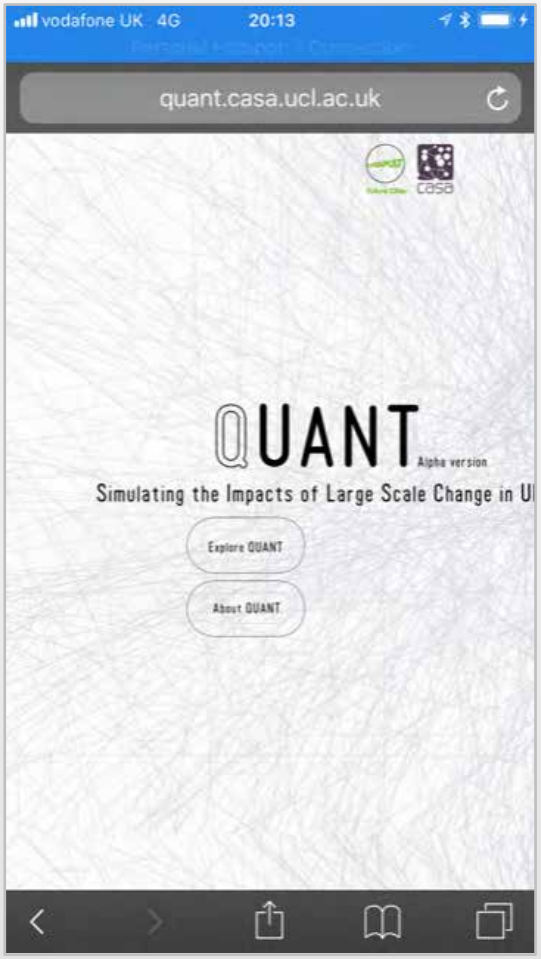
http://quant.casa.ucl.ac.uk/

For those who want to try it out

There is always ***a caveat emptor*** with demo models like this. We are building it as a demo – a prototype and there is much work to do on it before we can really imply it is usable.

It is not supposed to be a flashy app or anything – it is something that local authority planners and consultants might use as a resource along with many other complementary decision support tools

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



m.batty@ucl.ac.uk
richard.milton@ucl.ac.uk

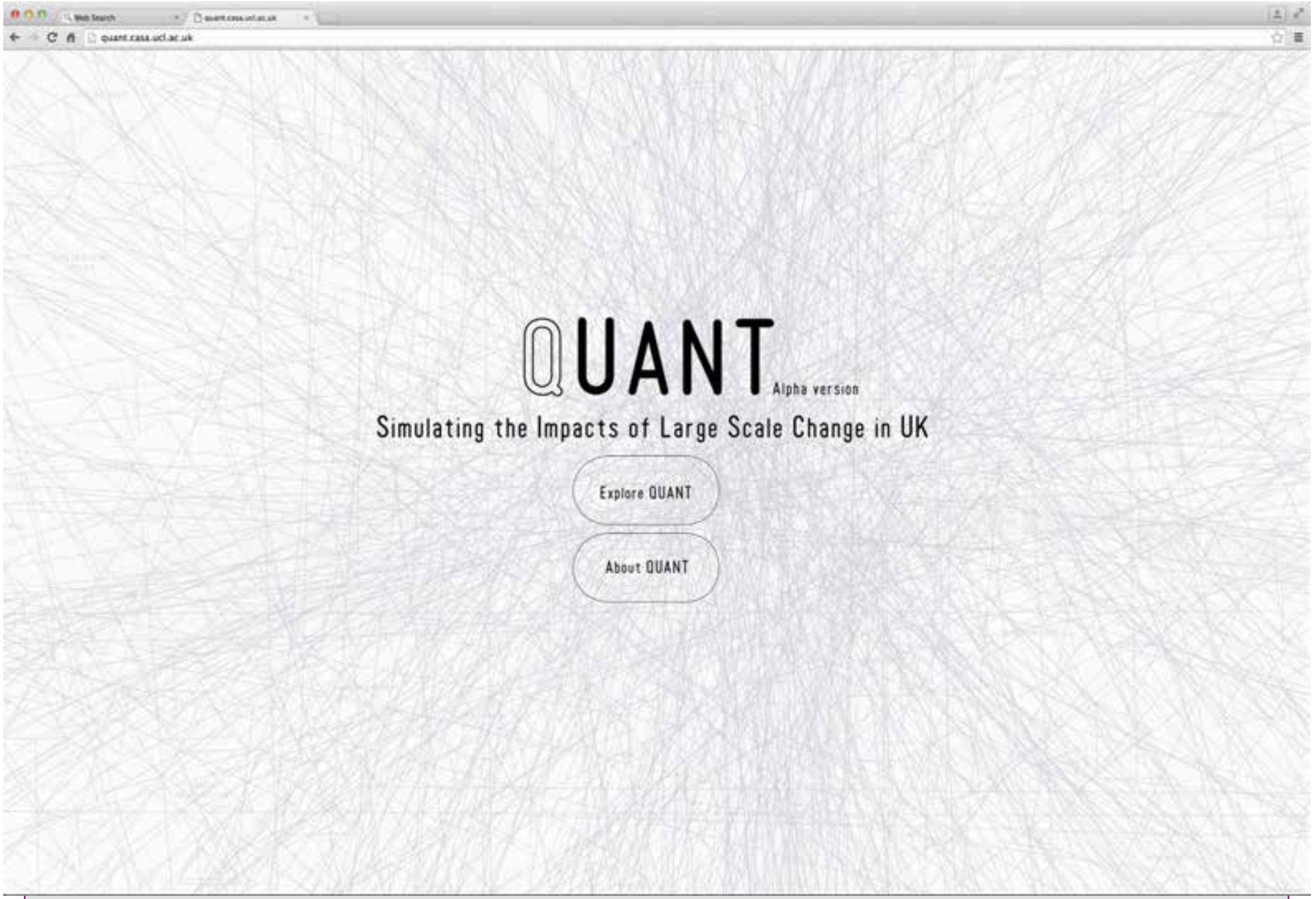
DAFNI
Data & Analytics Facility for National Infrastructure

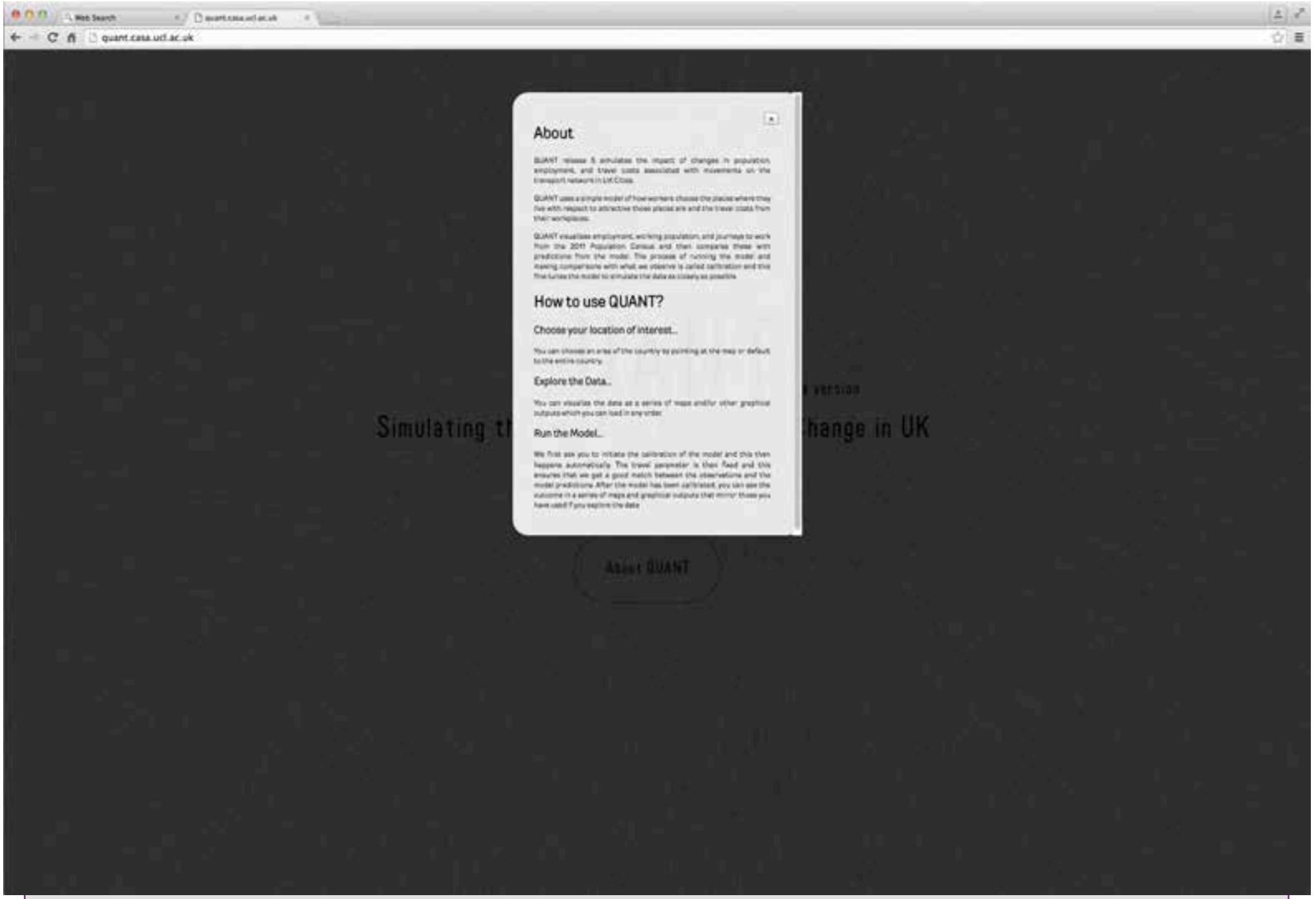
UKCRIC
UK Collaboratorium for Research in Infrastructure & Cities

CATAPULT
Future Cities

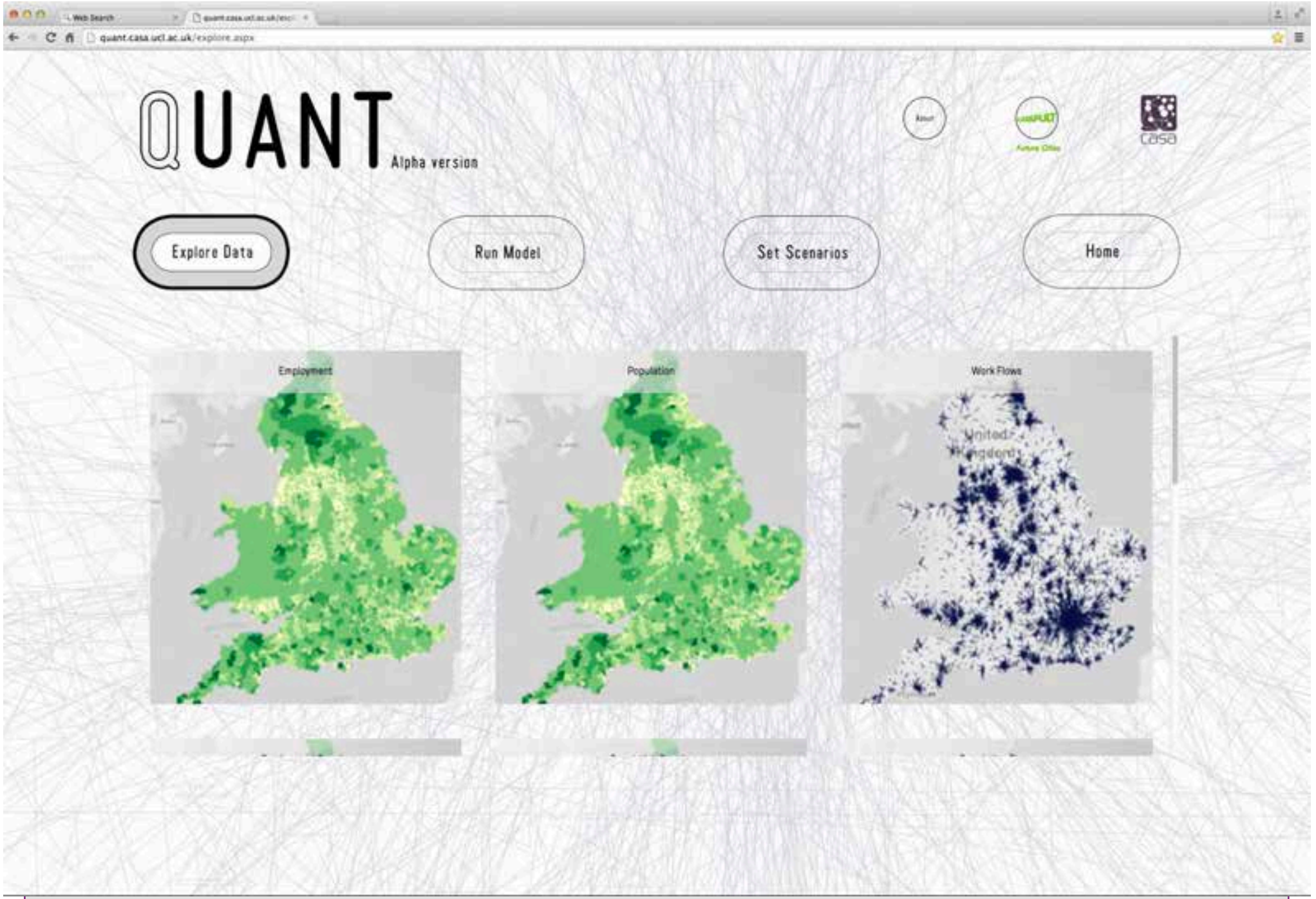
Infrastructure and Projects Authority

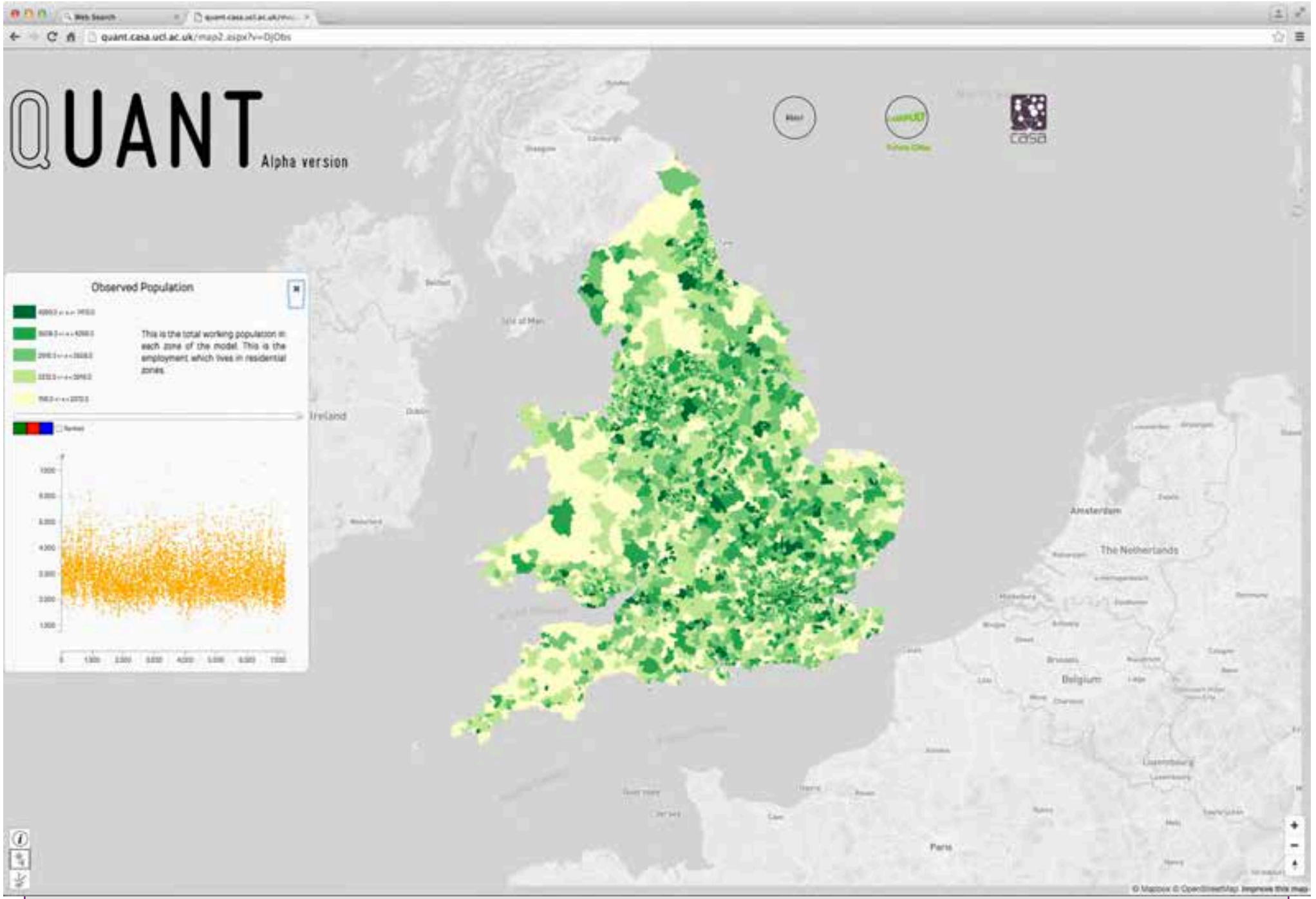
National Infrastructure Commission


quant.casa.ucl.ac.uk
QUANT
Alpha version
Simulating the Impacts of Large Scale Change in UK
Explore QUANT
About QUANT



Simulating Very Large Scale Spatial Systems


quant.casa.ucl.ac.uk/explore.aspx
QUANT
Alpha version
Explore Data
Run Model
Set Scenarios
Home
Employment
Population
Work Flows

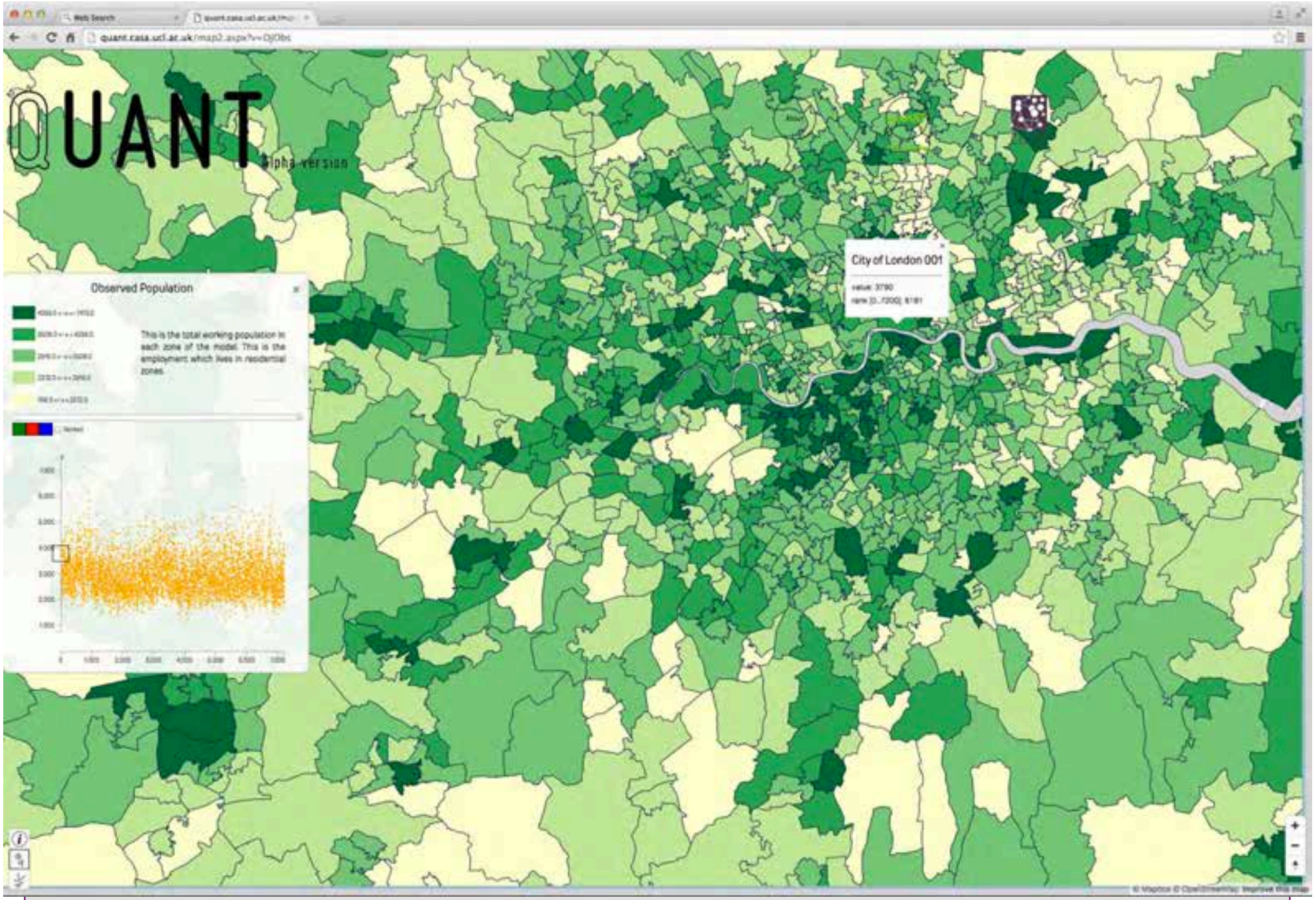


Simulating Very Large Scale Spatial Systems

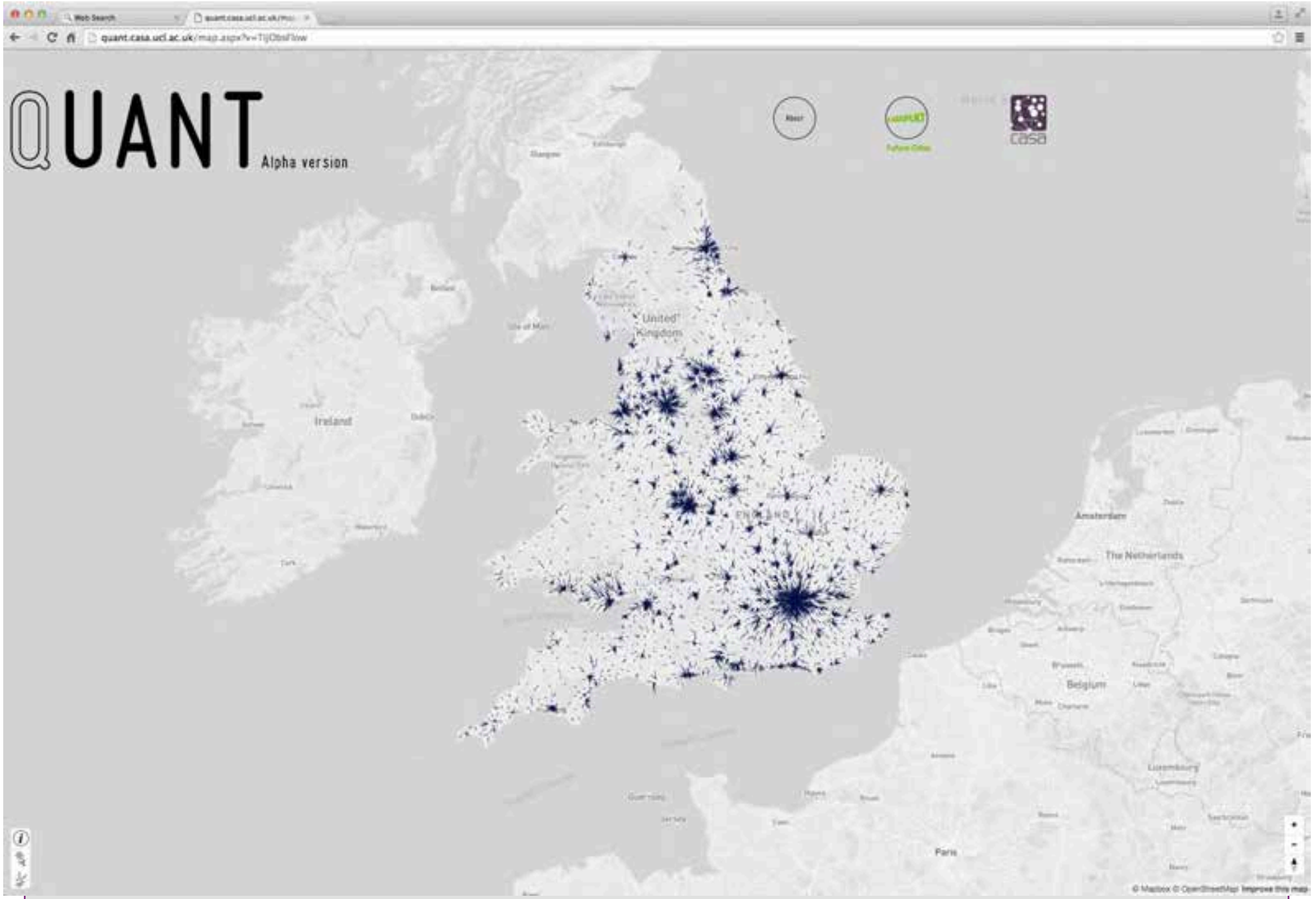

QUANT
Alpha version
Observed Population
This is the total working population in each zone of the model. This is the employment which lives in residential zones.



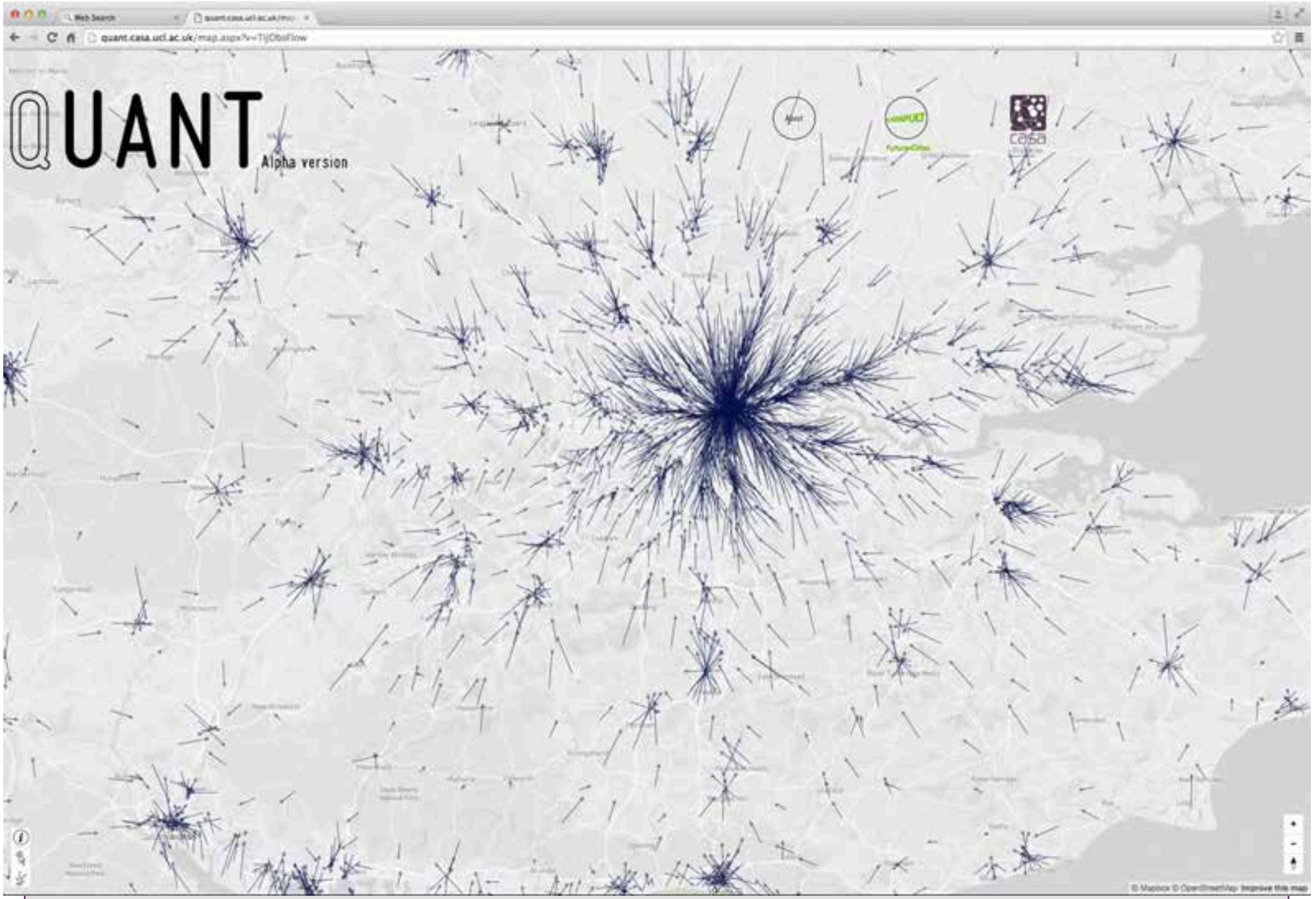
Simulating Very Large Scale Spatial Systems



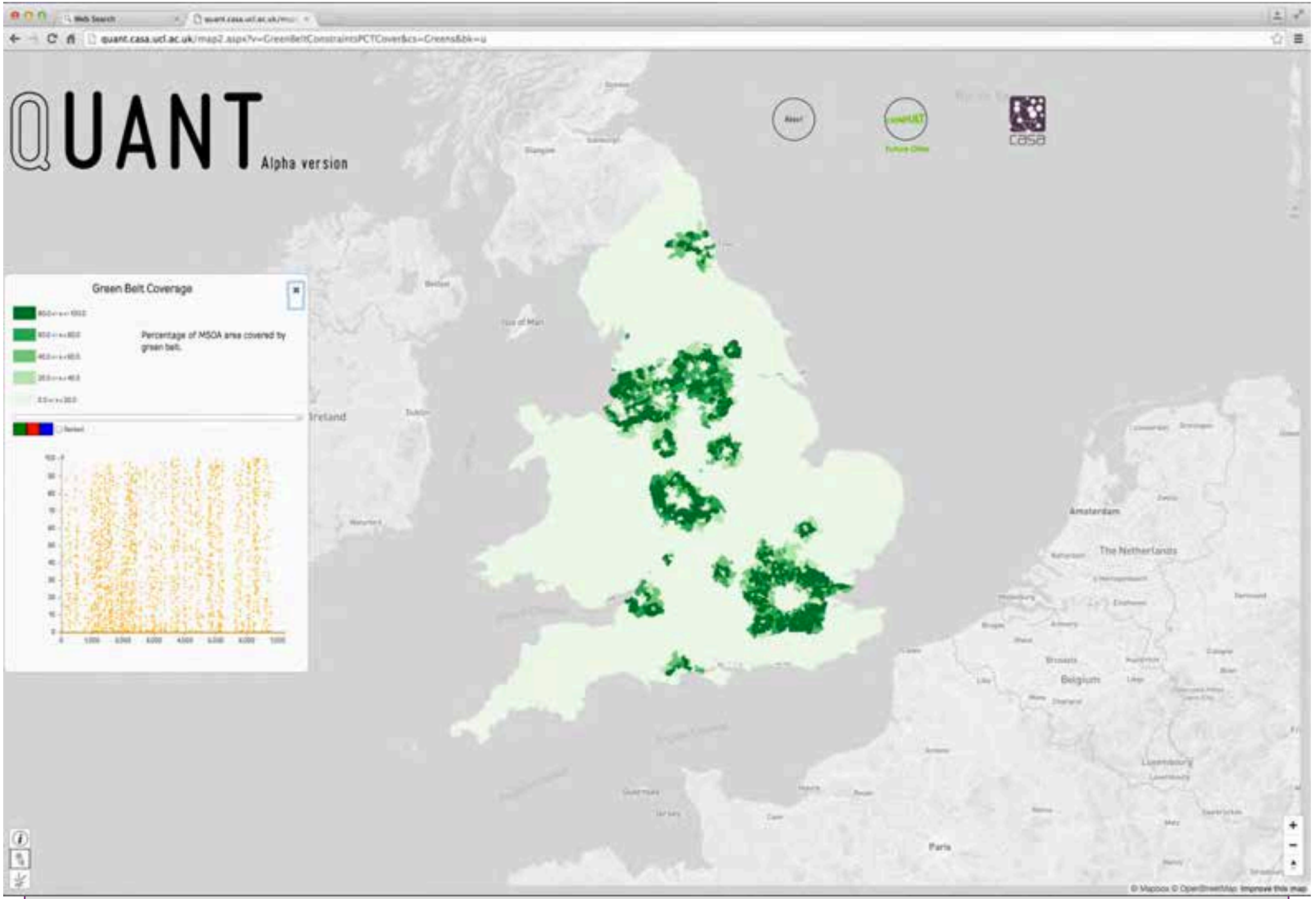
Simulating Very Large Scale Spatial Systems



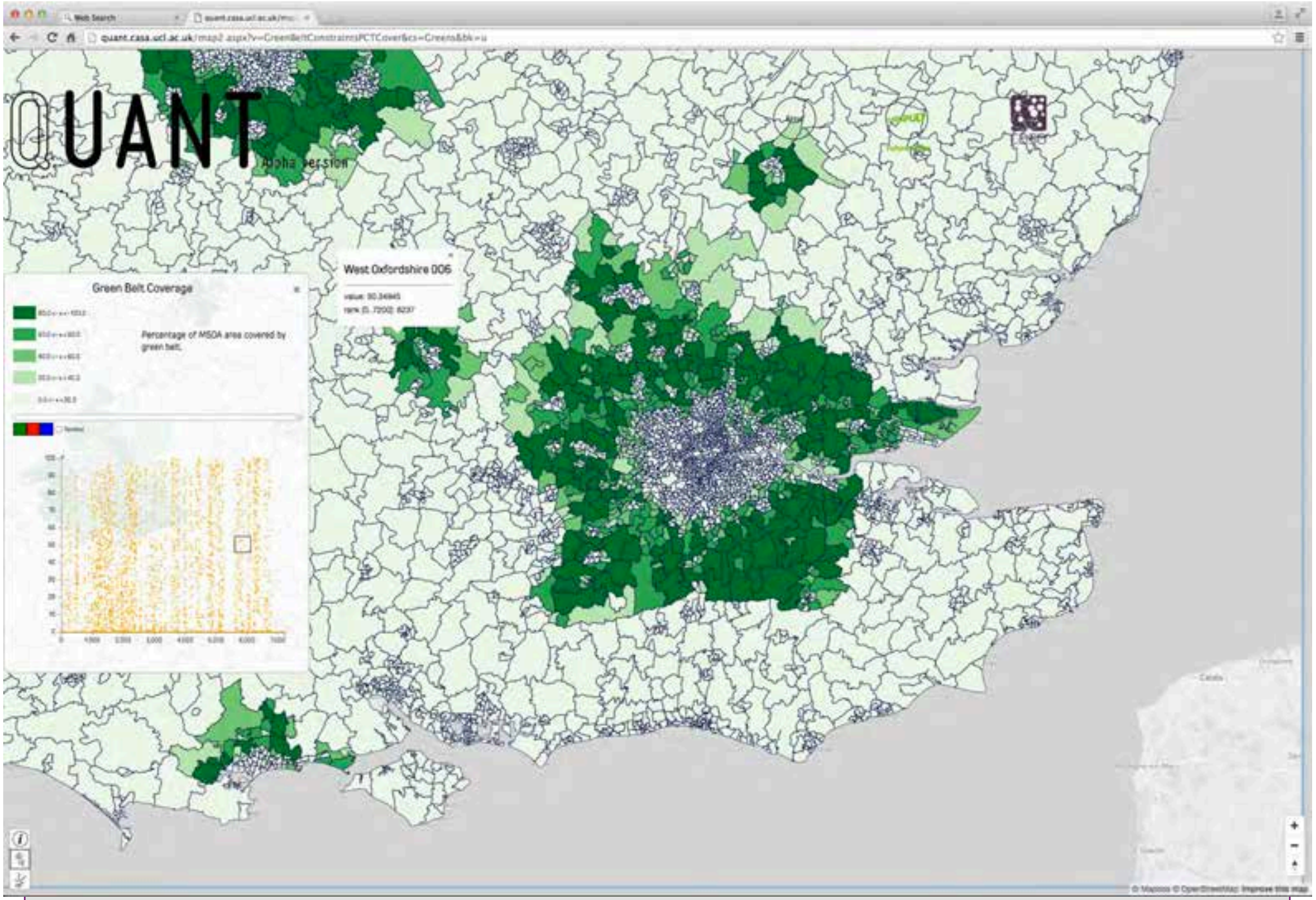
Simulating Very Large Scale Spatial Systems



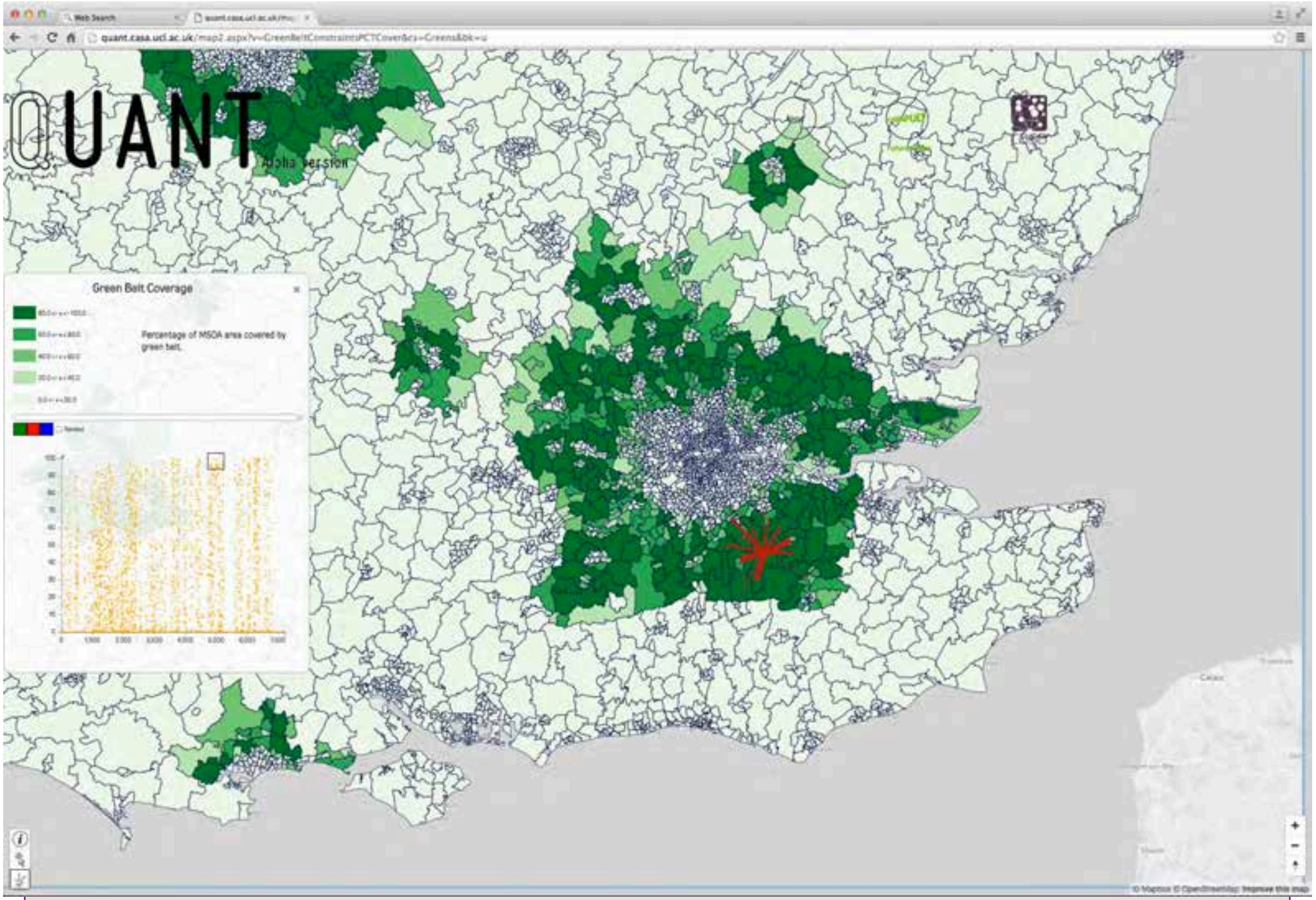
Simulating Very Large Scale Spatial Systems



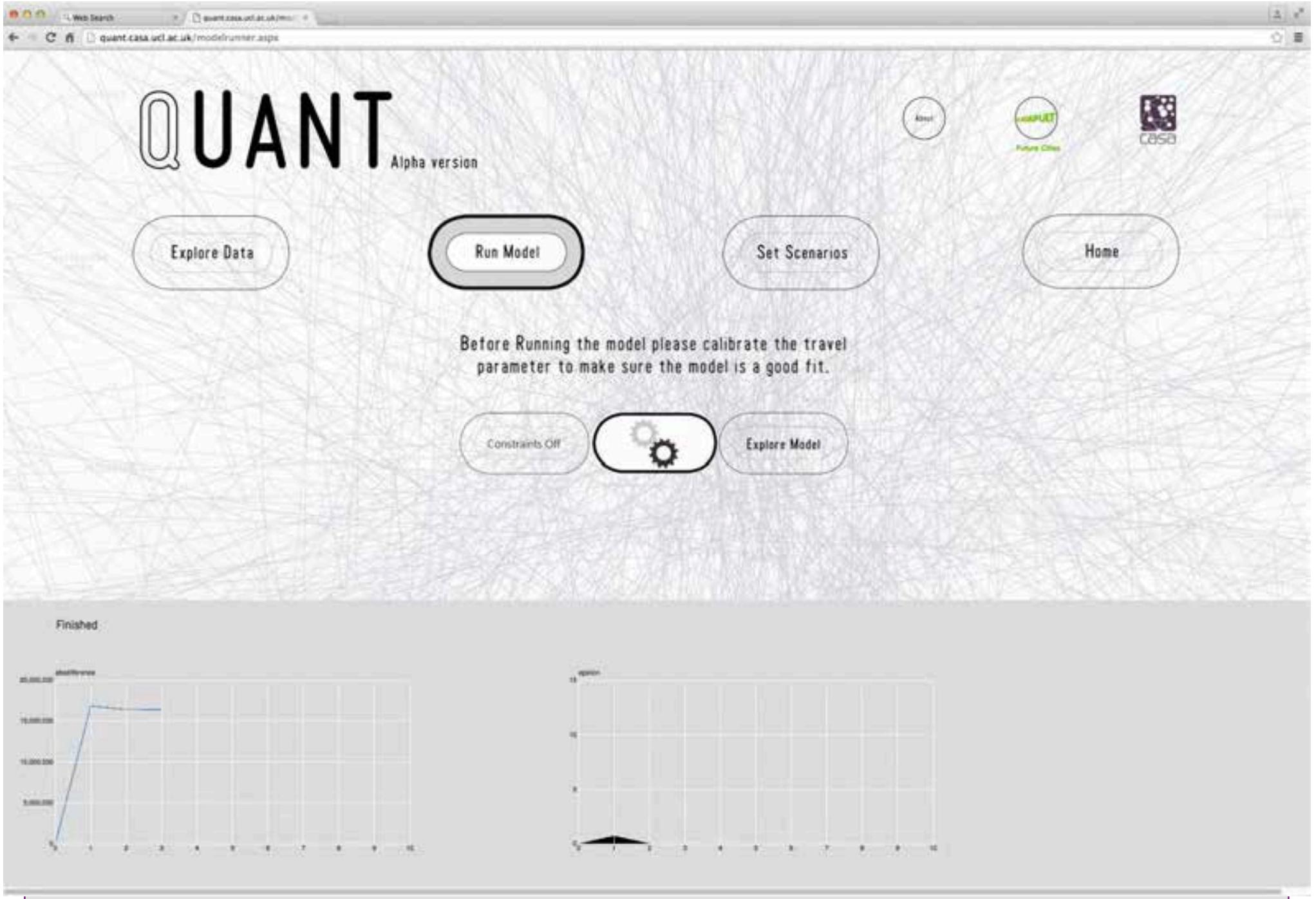
Simulating Very Large Scale Spatial Systems



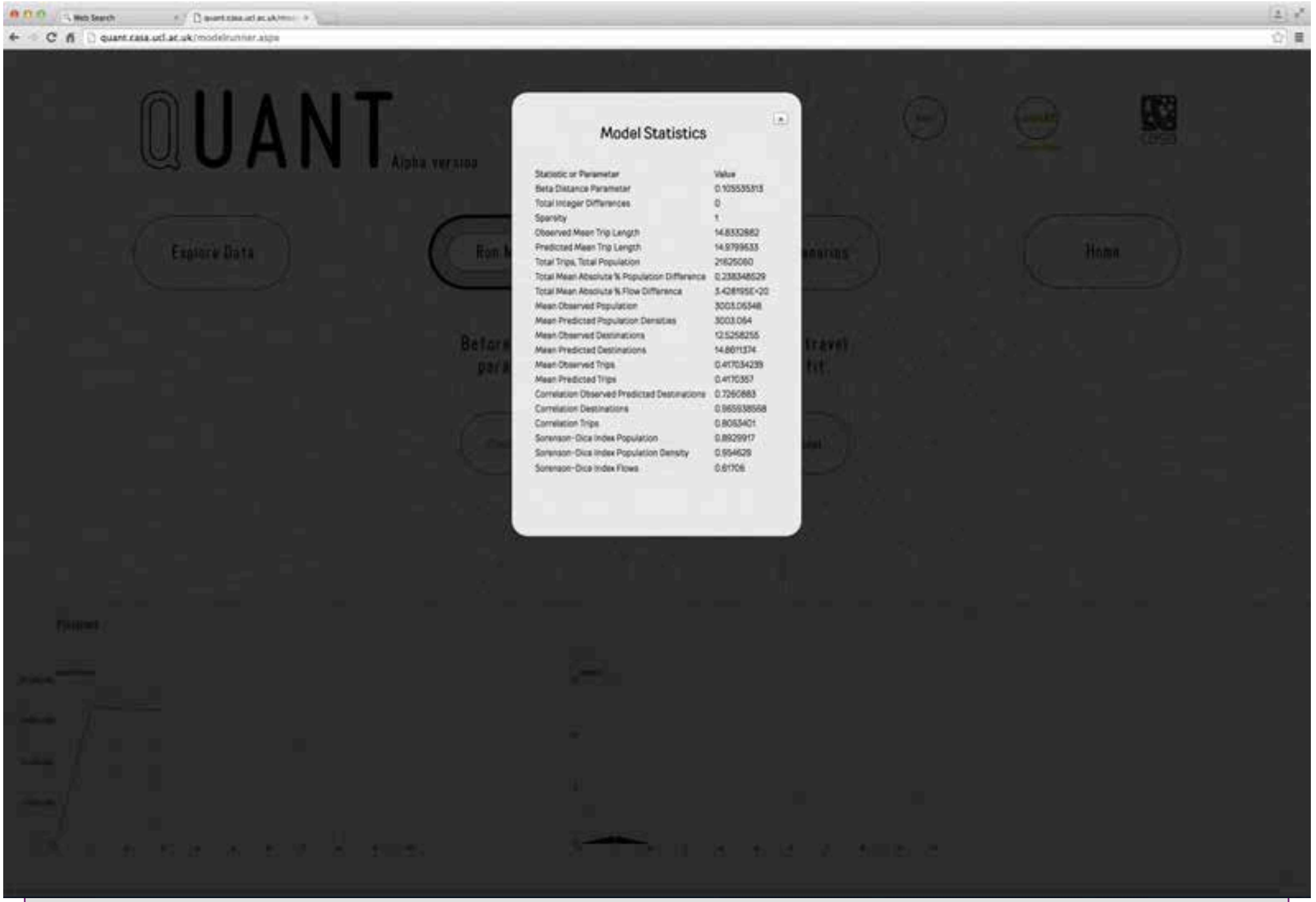
Simulating Very Large Scale Spatial Systems



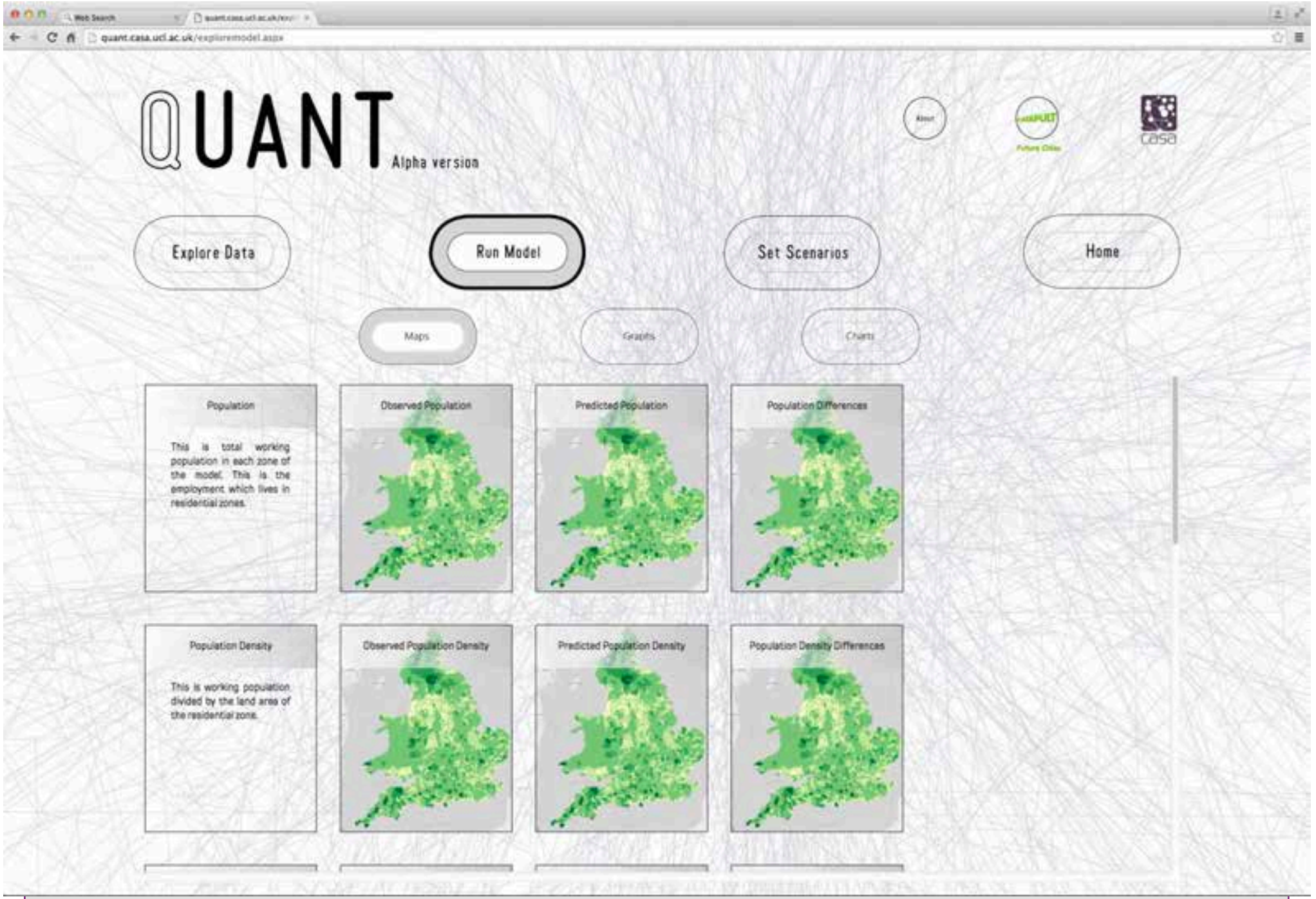
Simulating Very Large Scale Spatial Systems

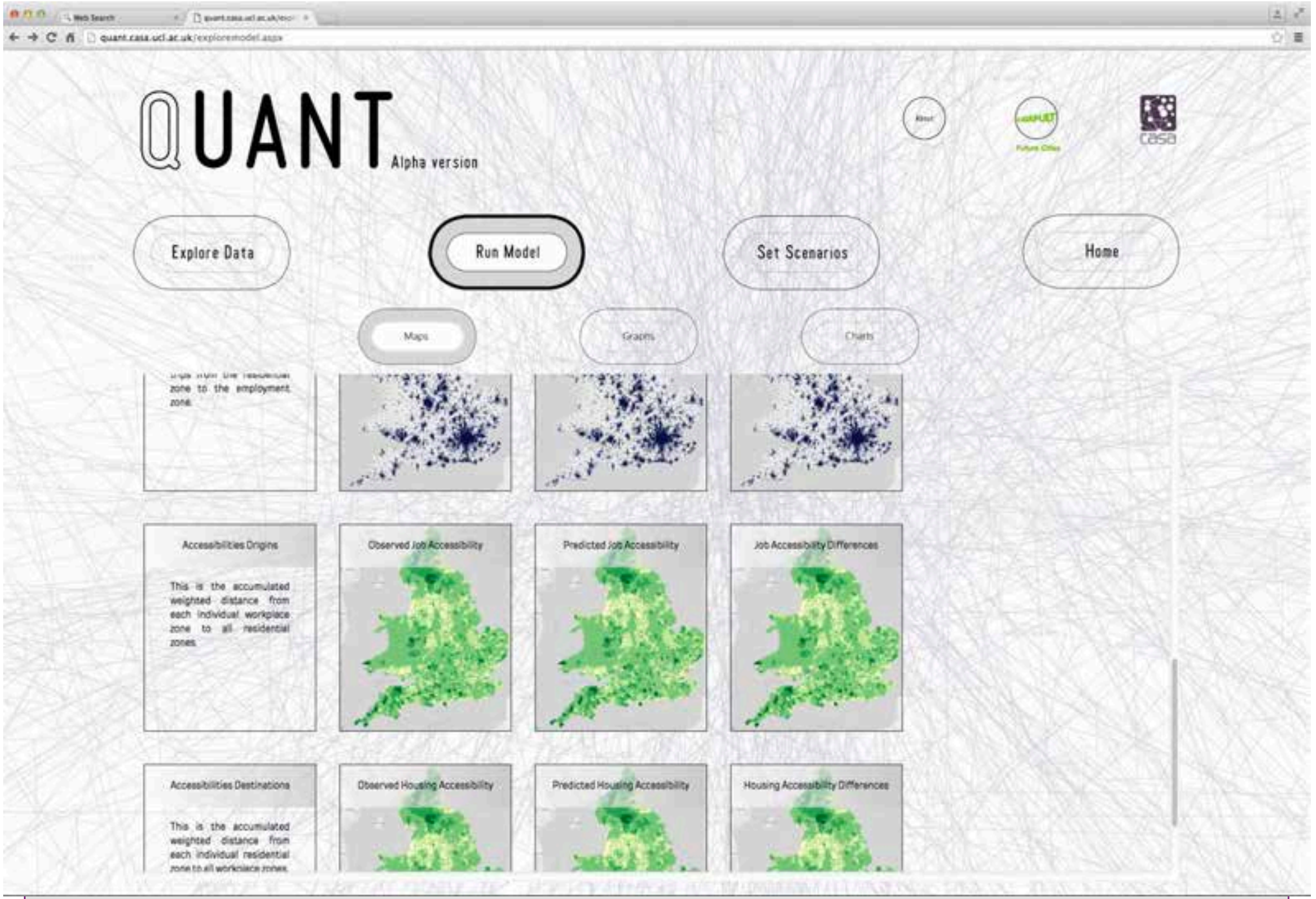


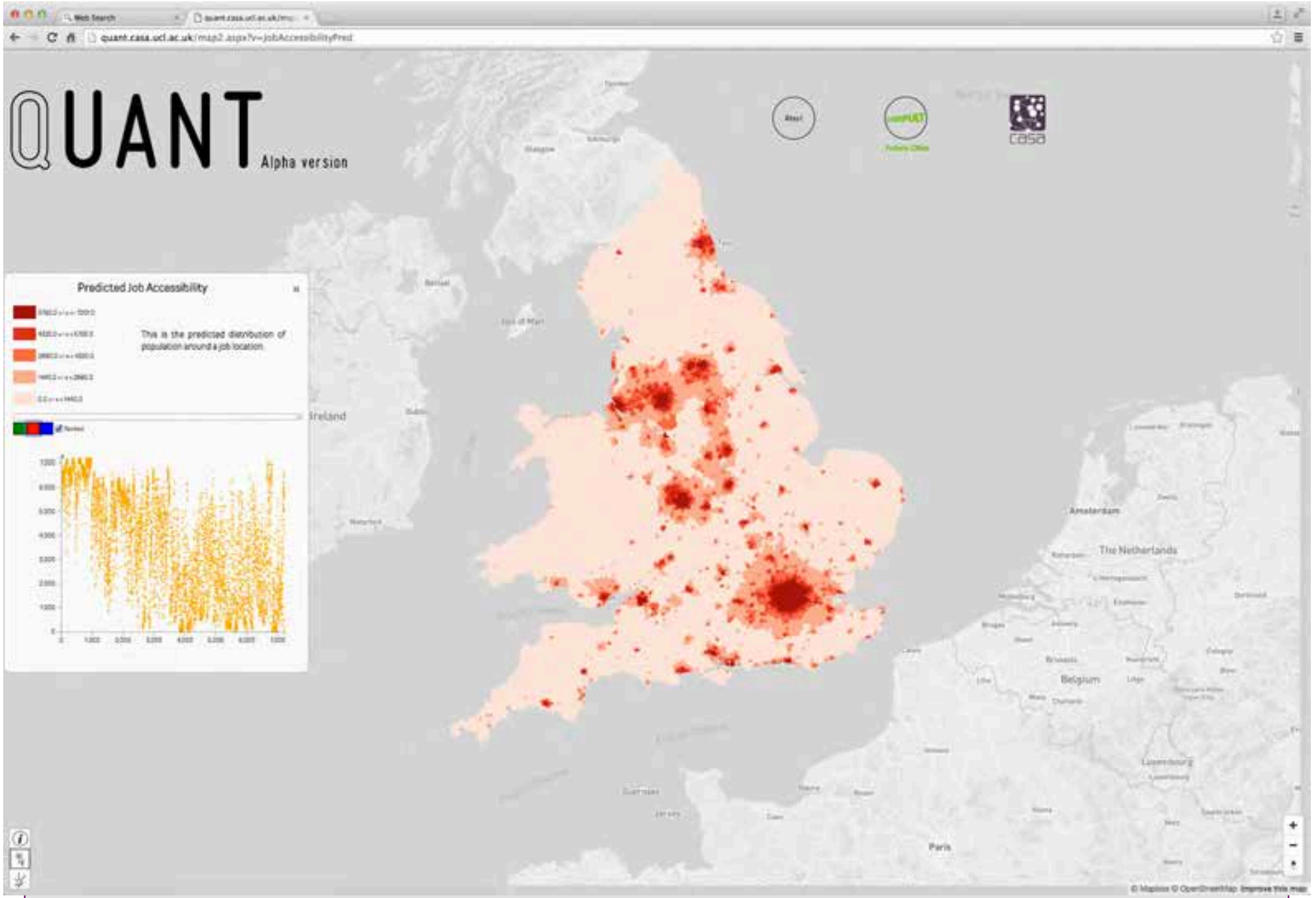
Simulating Very Large Scale Spatial Systems



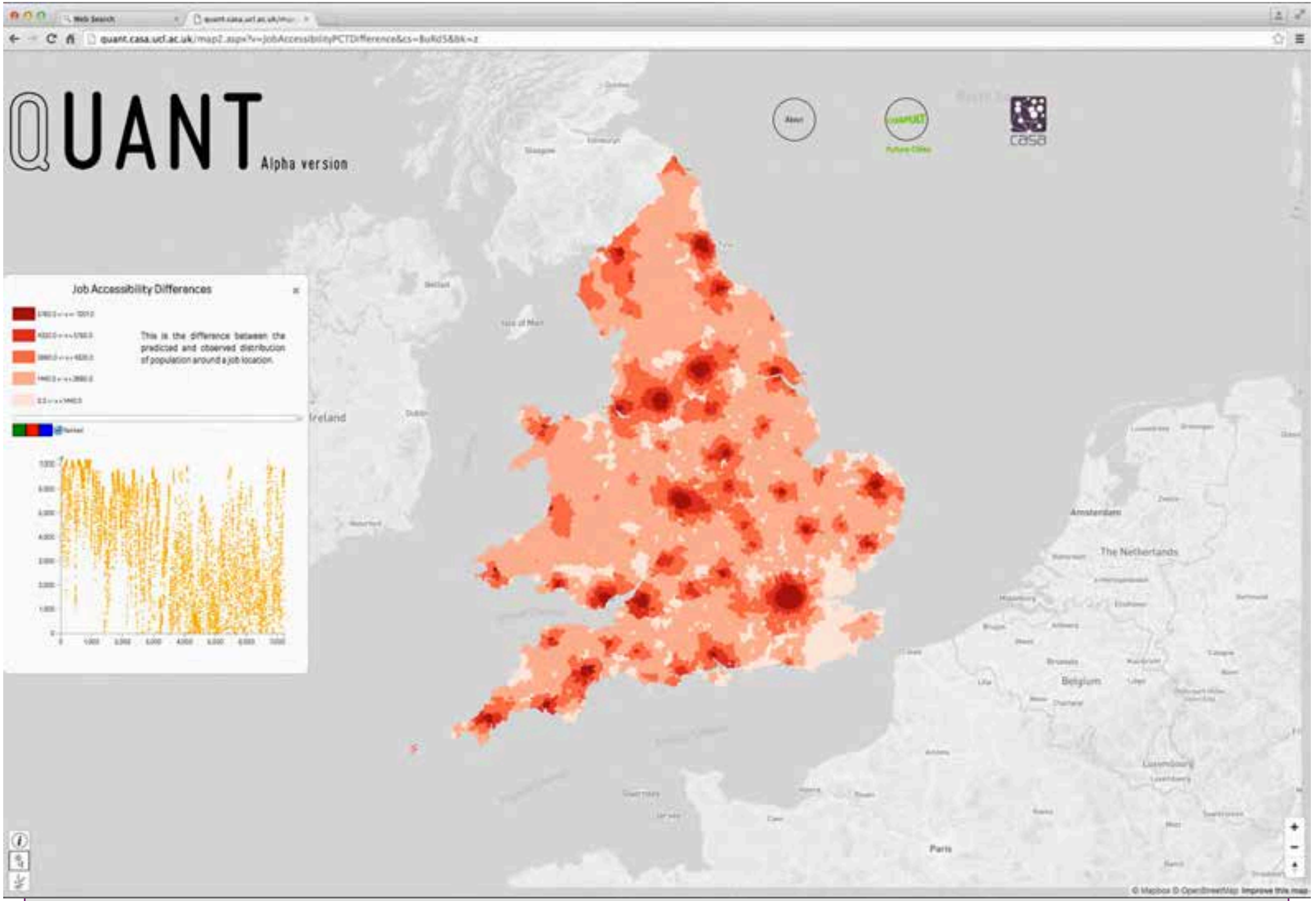
Simulating Very Large Scale Spatial Systems

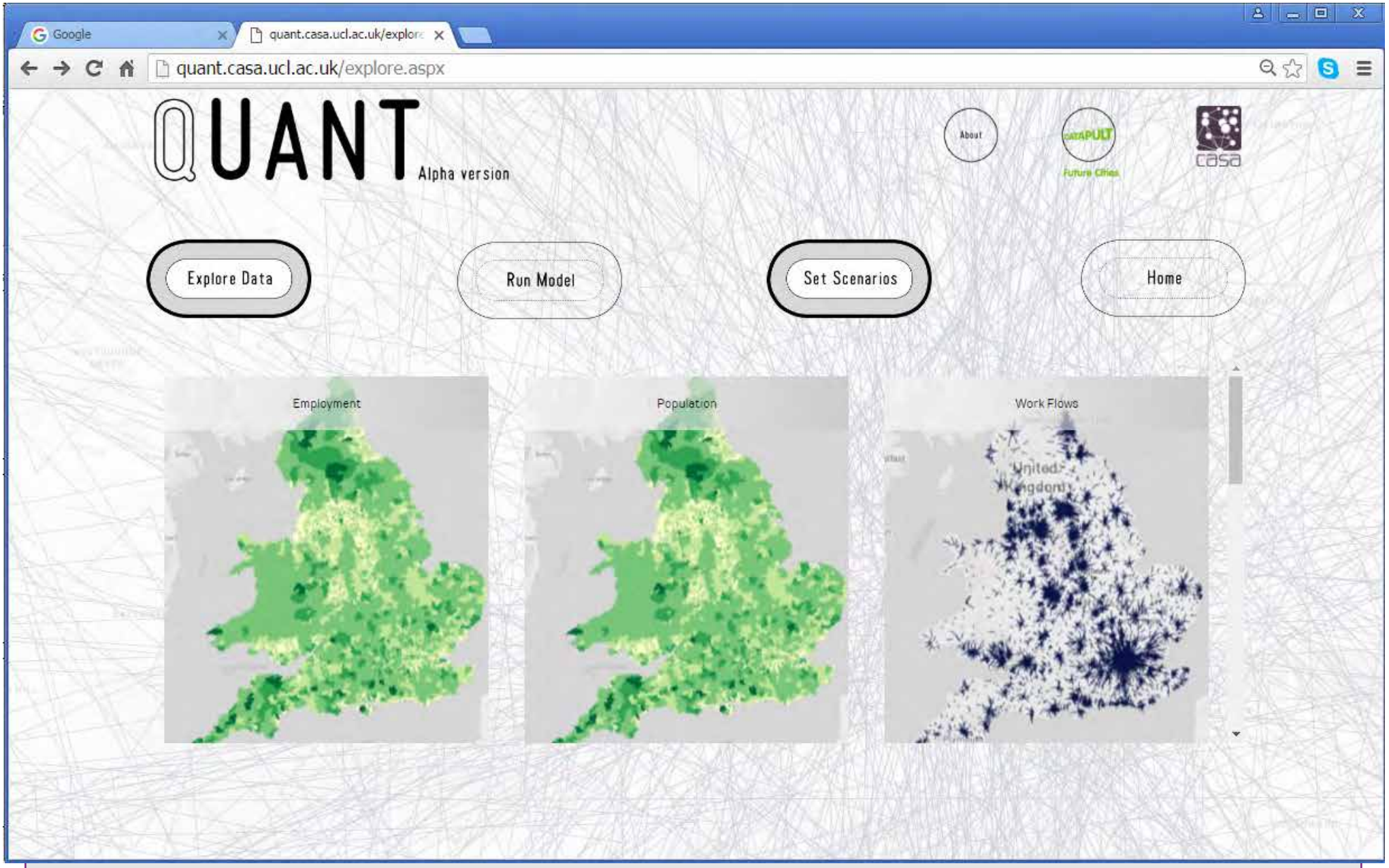

quant.casa.ucl.ac.uk/exploremodel.aspx
QUANT
Alpha version
About
CASA
Explore Data
Run Model
Set Scenarios
Home
Maps
Graphs
Charts
Population
This is total working population in each zone of the model. This is the employment which lives in residential zones.
Observed Population
Predicted Population
Population Differences
Population Density
This is working population divided by the land area of the residential zone.
Observed Population Density
Predicted Population Density
Population Density Differences

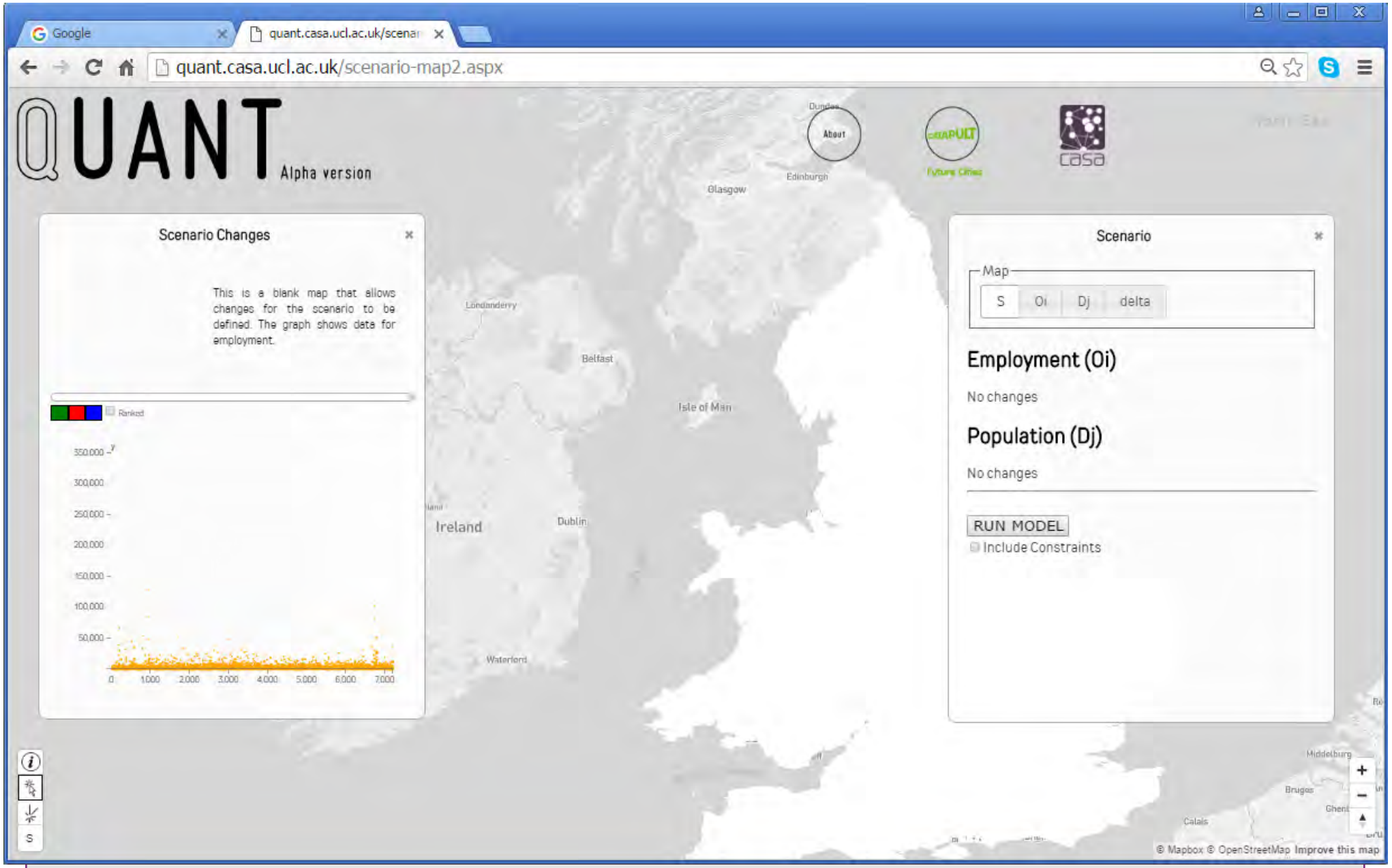

QUANT
Alpha version
Explore Data
Run Model
Set Scenarios
Home
Maps
Graphs
Charts
zone to the employment zone.
Accessibilities Origins
This is the accumulated weighted distance from each individual workplace zone to all residential zones.
Observed Job Accessibility
Predicted Job Accessibility
Job Accessibility Differences
Accessibilities Destinations
This is the accumulated weighted distance from each individual residential zone to all workplace zones.
Observed Housing Accessibility
Predicted Housing Accessibility
Housing Accessibility Differences



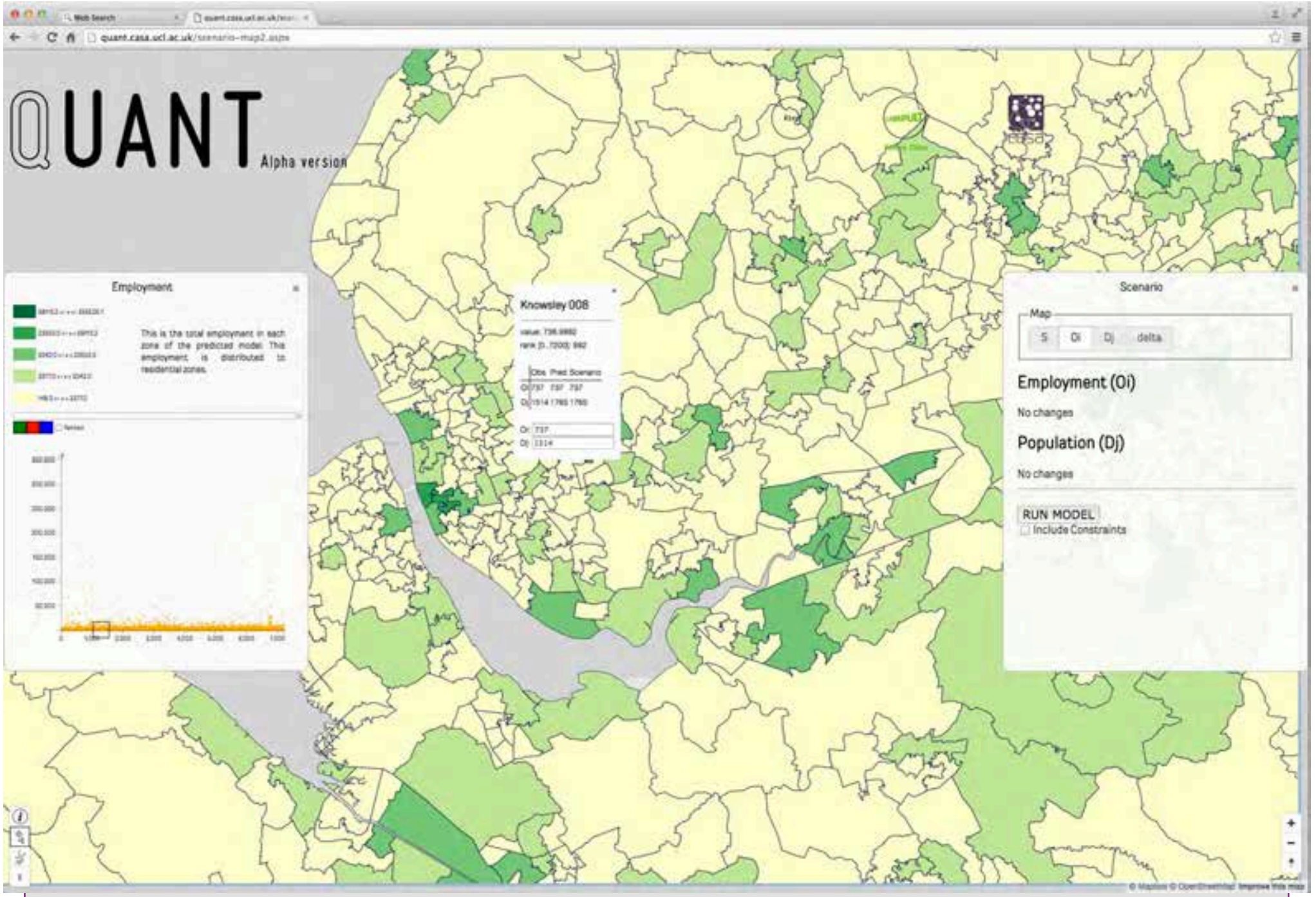
Simulating Very Large Scale Spatial Systems


QUANT
Alpha version
Job Accessibility Differences
This is the difference between the predicted and observed distribution of population around a job location.
Ireland
Amsterdam
The Netherlands
Belgium
Paris

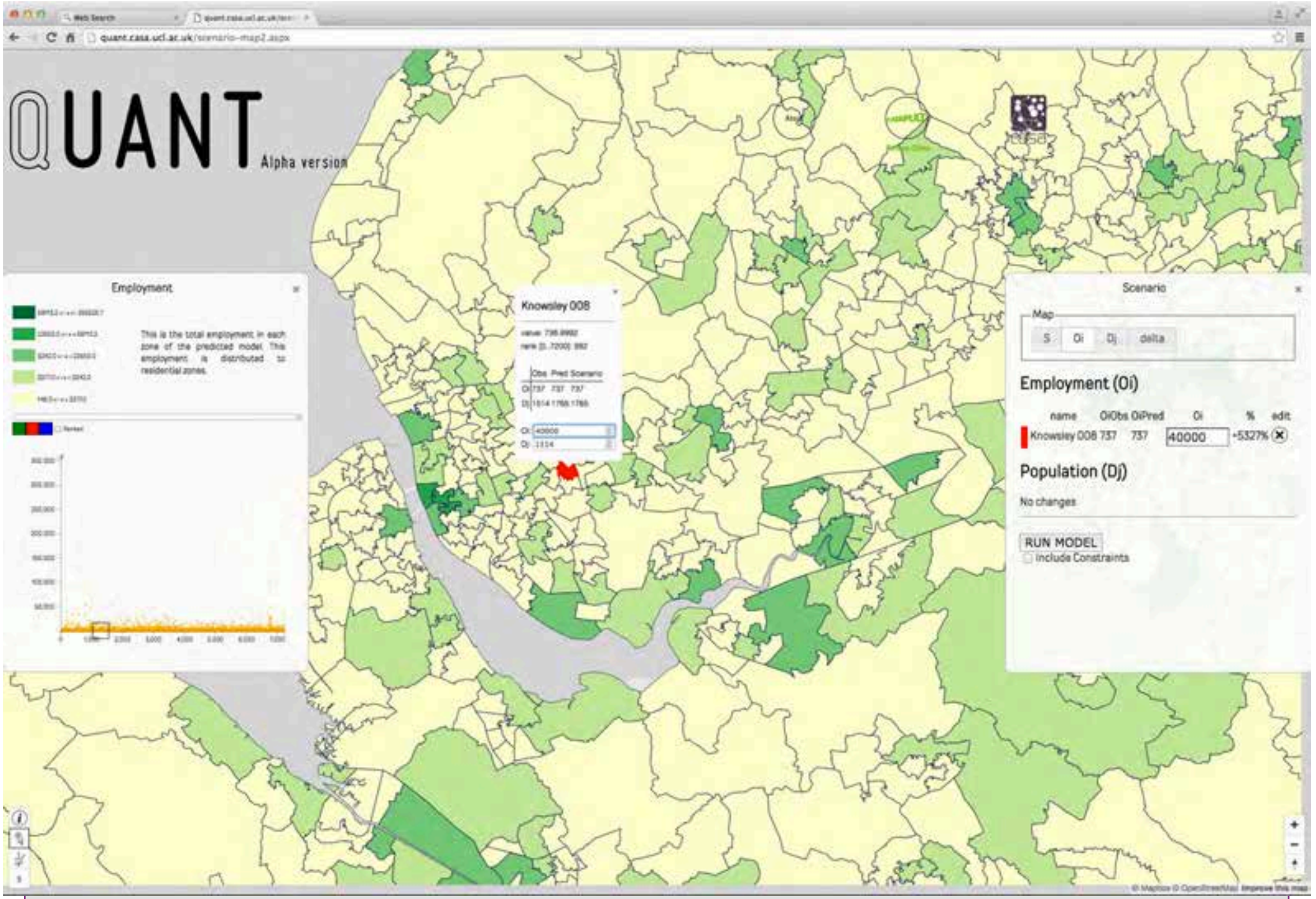

Google
quant.casa.ucl.ac.uk/explore.aspx
quant.casa.ucl.ac.uk/explore.aspx
QUANT
Alpha version
About
Explore Data
Run Model
Set Scenarios
Home
Employment
Population
Work Flows



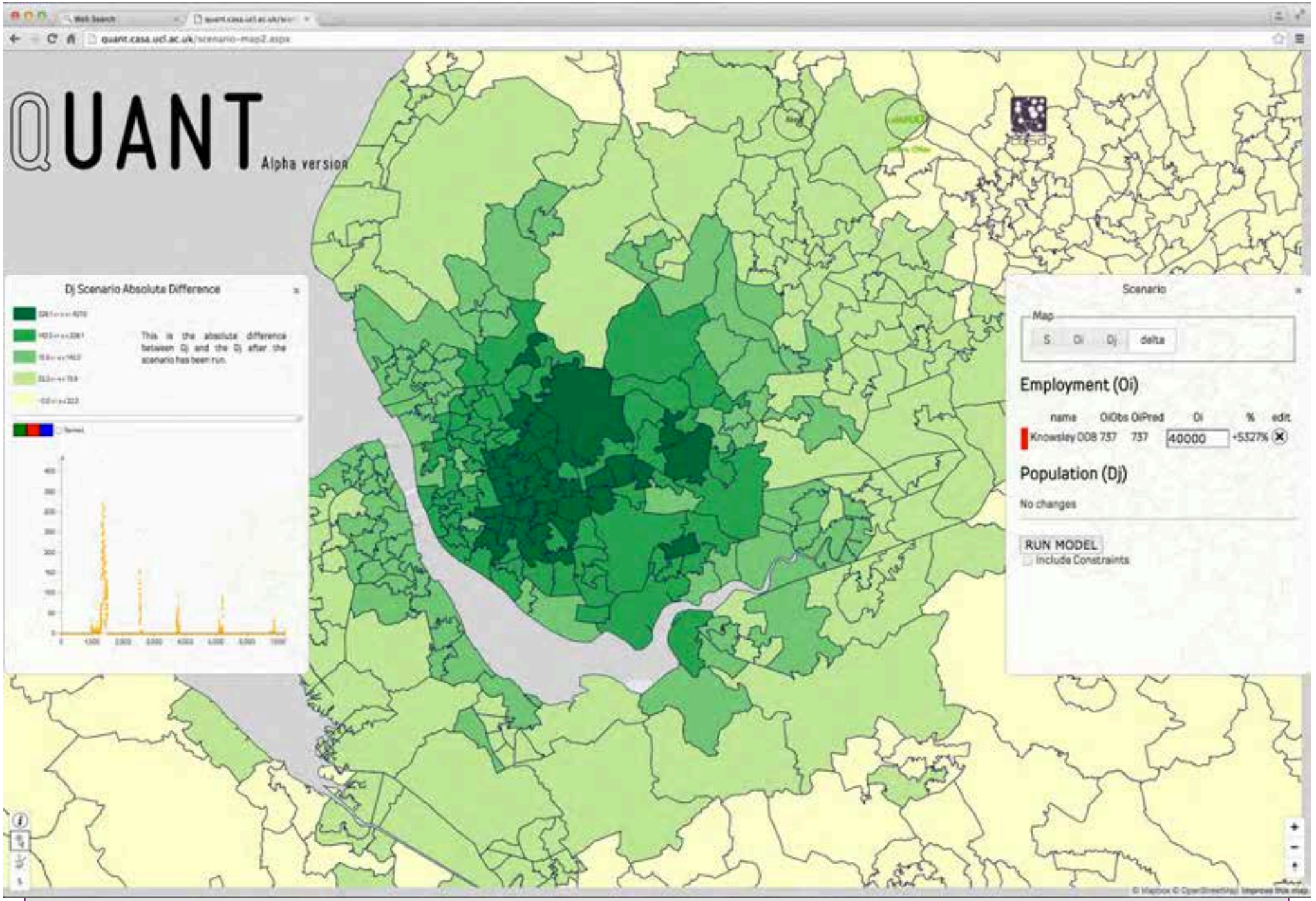
Simulating Very Large Scale Spatial Systems



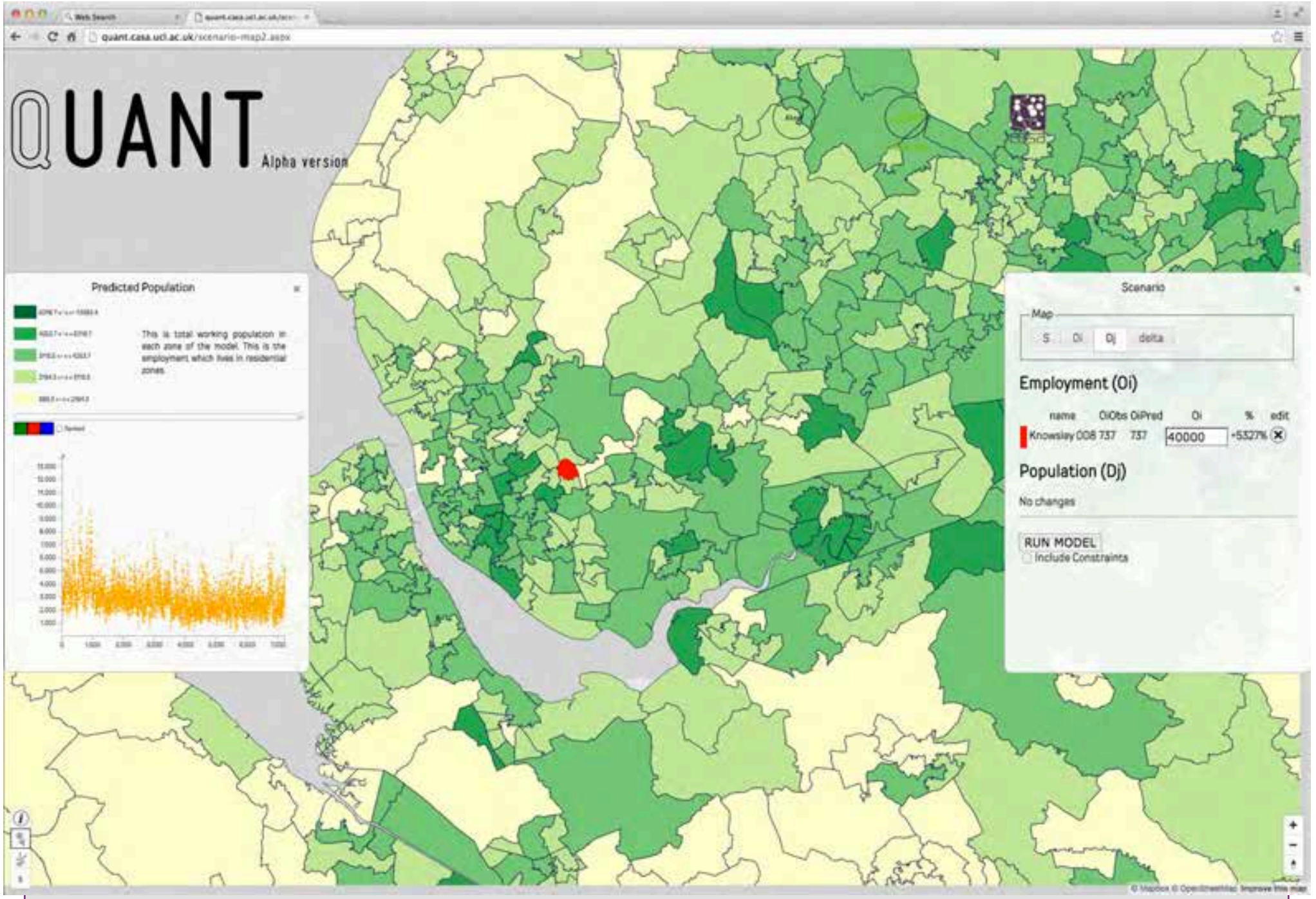
Simulating Very Large Scale Spatial Systems



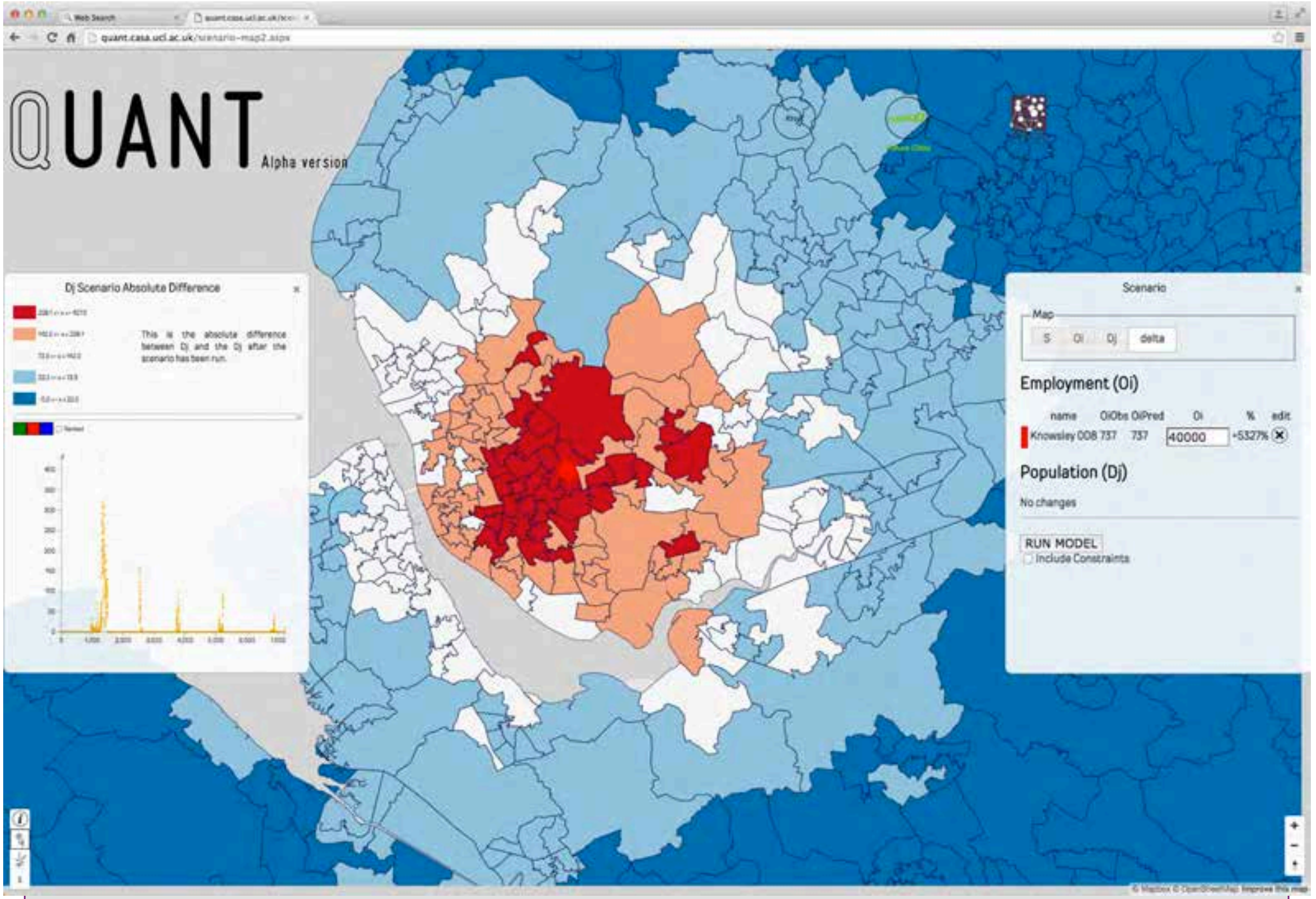
Simulating Very Large Scale Spatial Systems



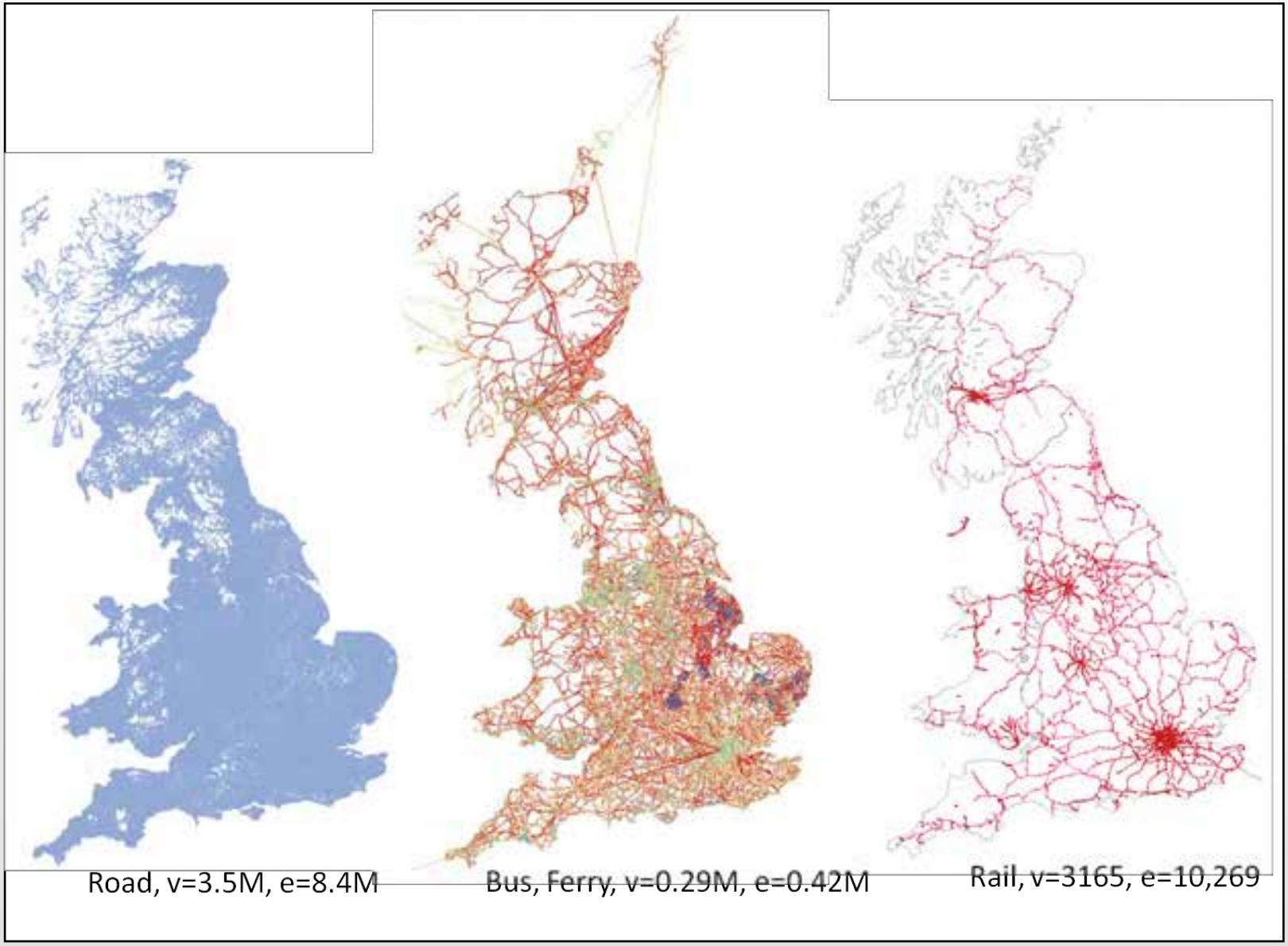
Simulating Very Large Scale Spatial Systems



Simulating Very Large Scale Spatial Systems



Simulating Very Large Scale Spatial Systems

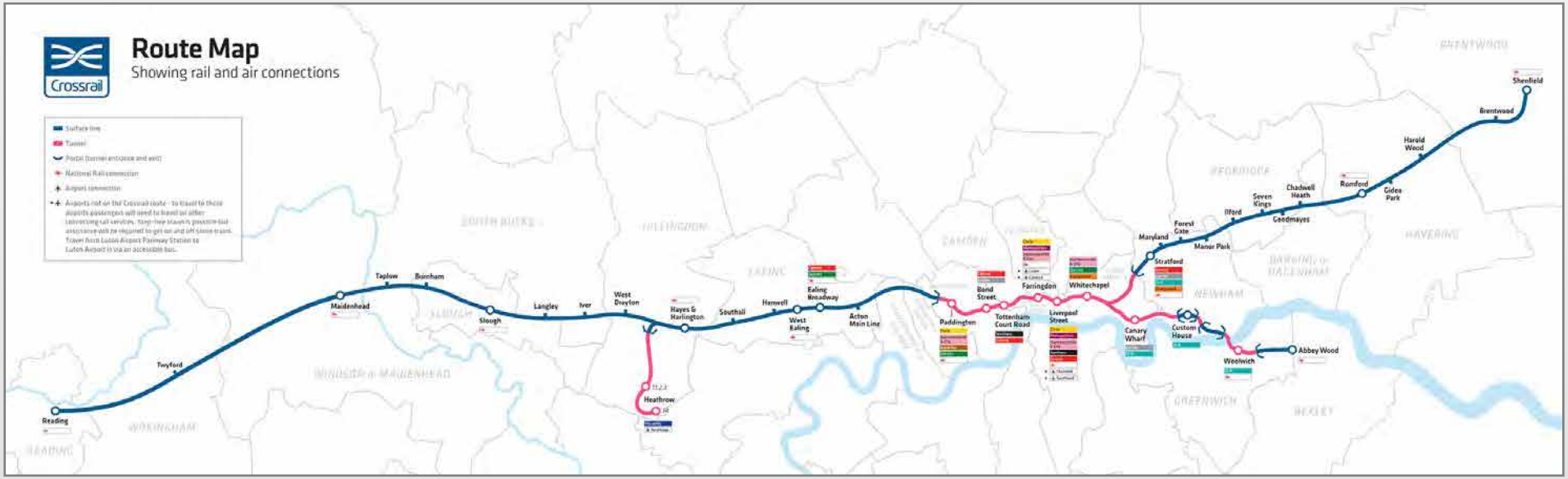

Road, v=3.5M, e=8.4M
Bus, Ferry, v=0.29M, e=0.42M
Rail, v=3165, e=10,269

Crossrail1: New Tube Line in Central London

Reading, Heathrow, Shenfield, Abbey Wood

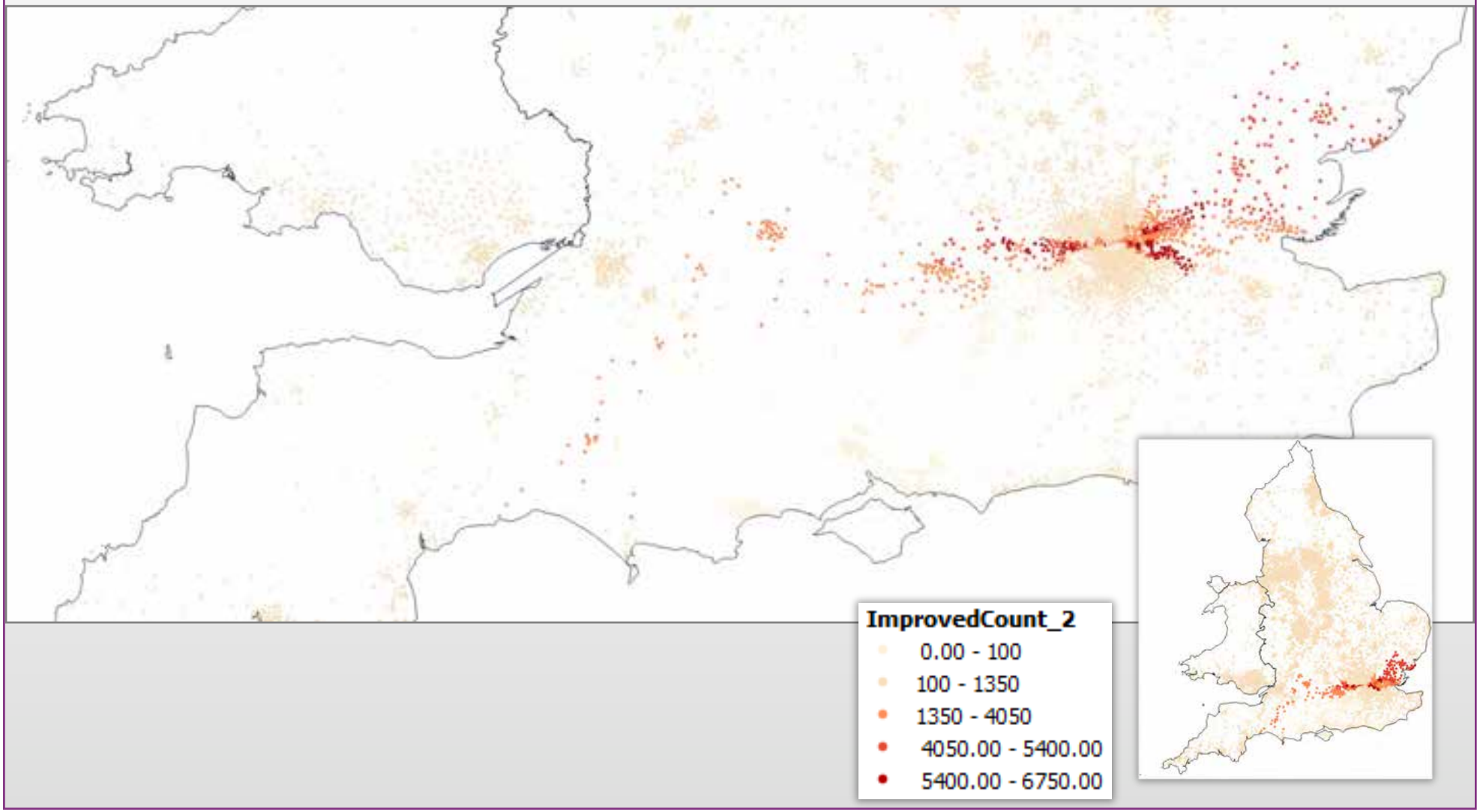






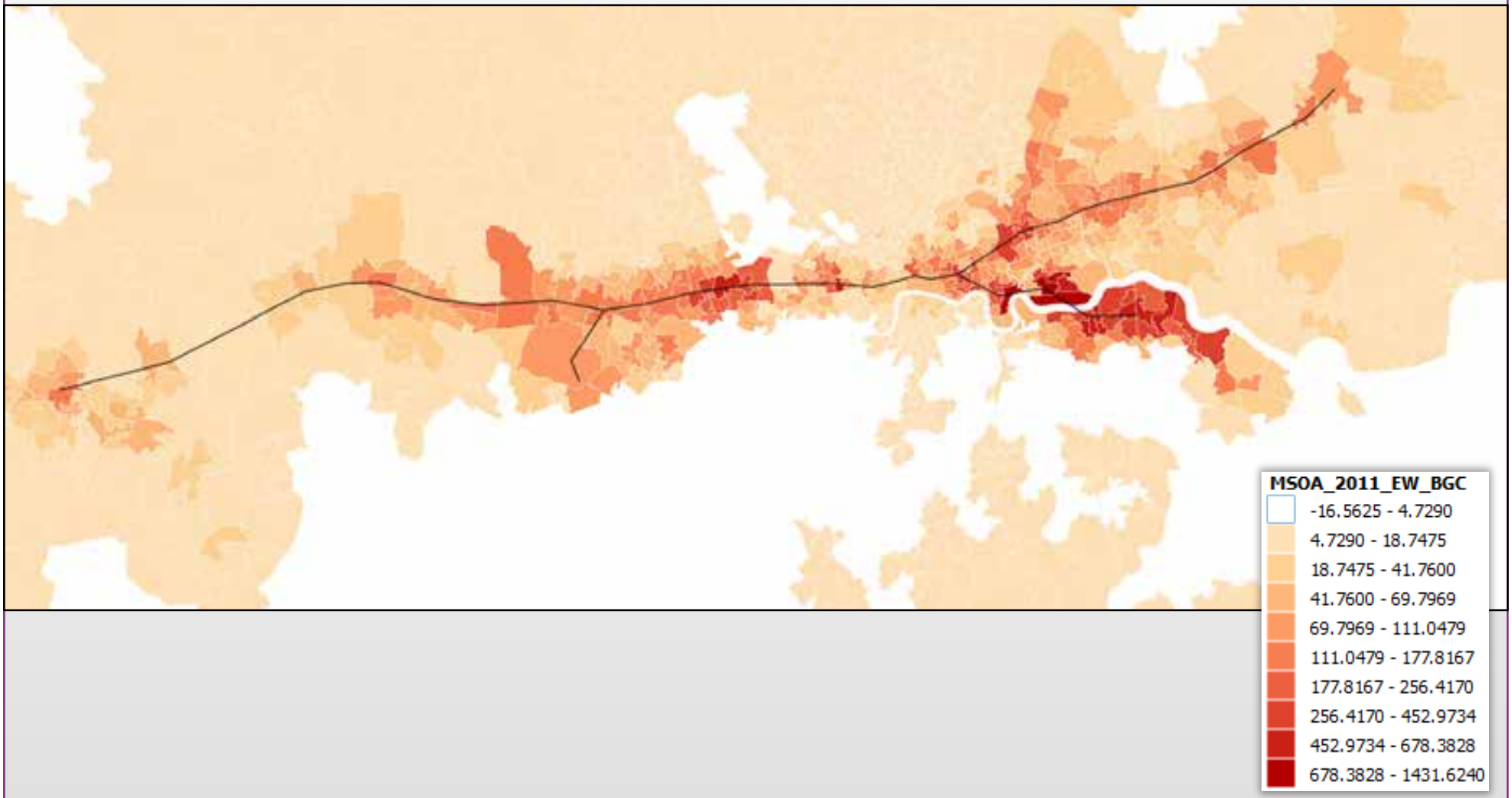
Crossrail

Number of Improved Journeys (n_i)



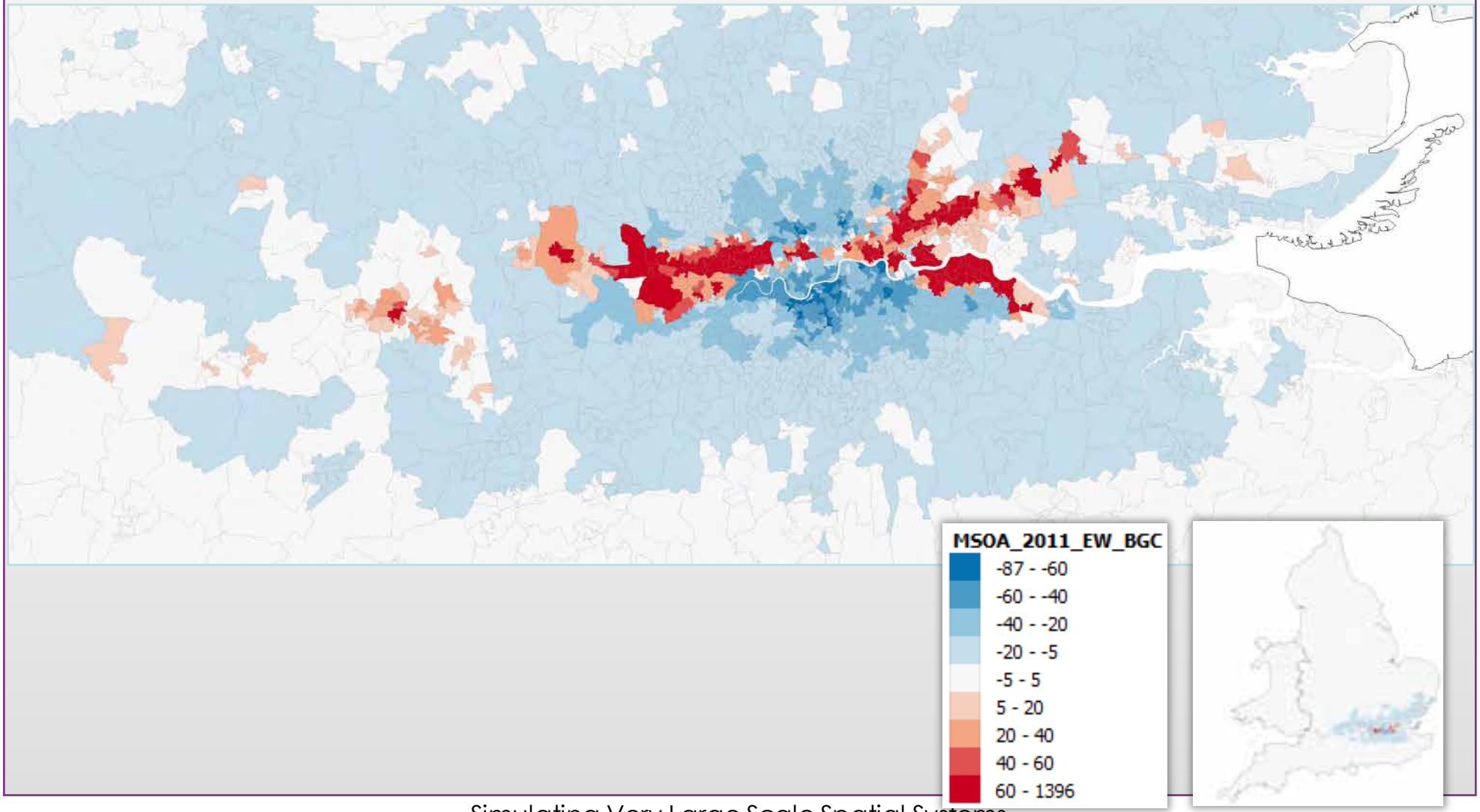
Crossrail

Population change (rail mode only)



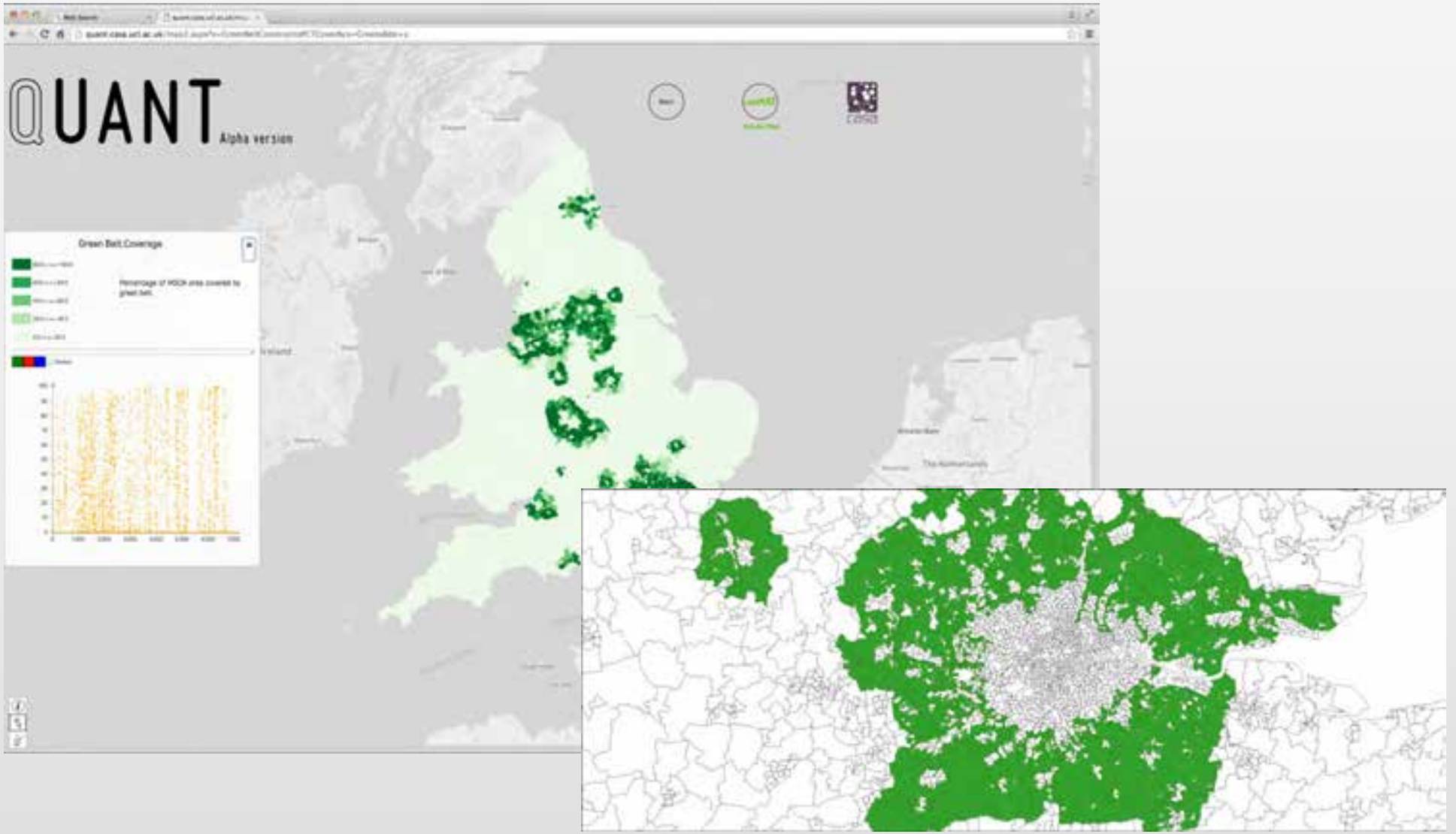
Crossrail

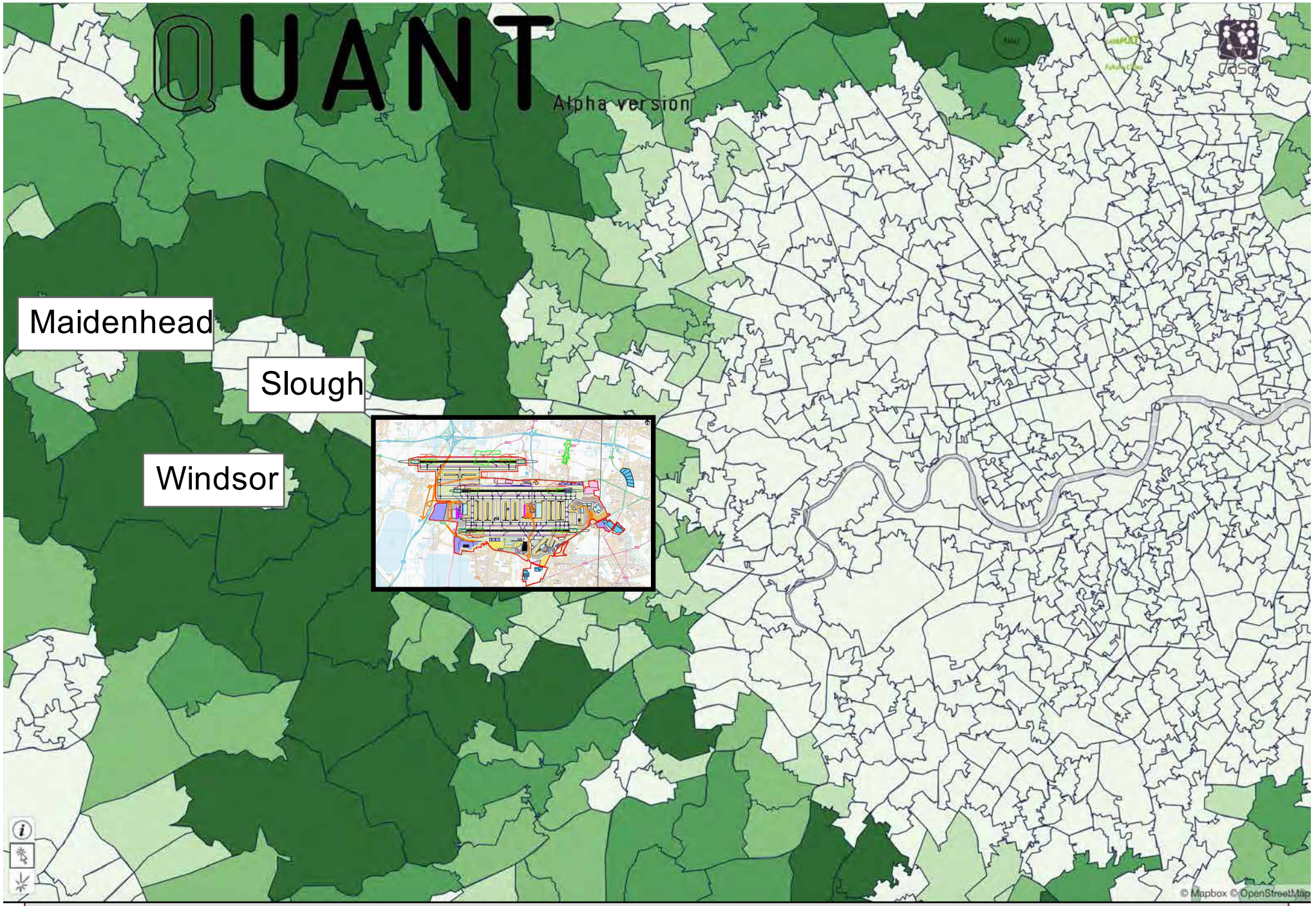
Population Change (all modes)



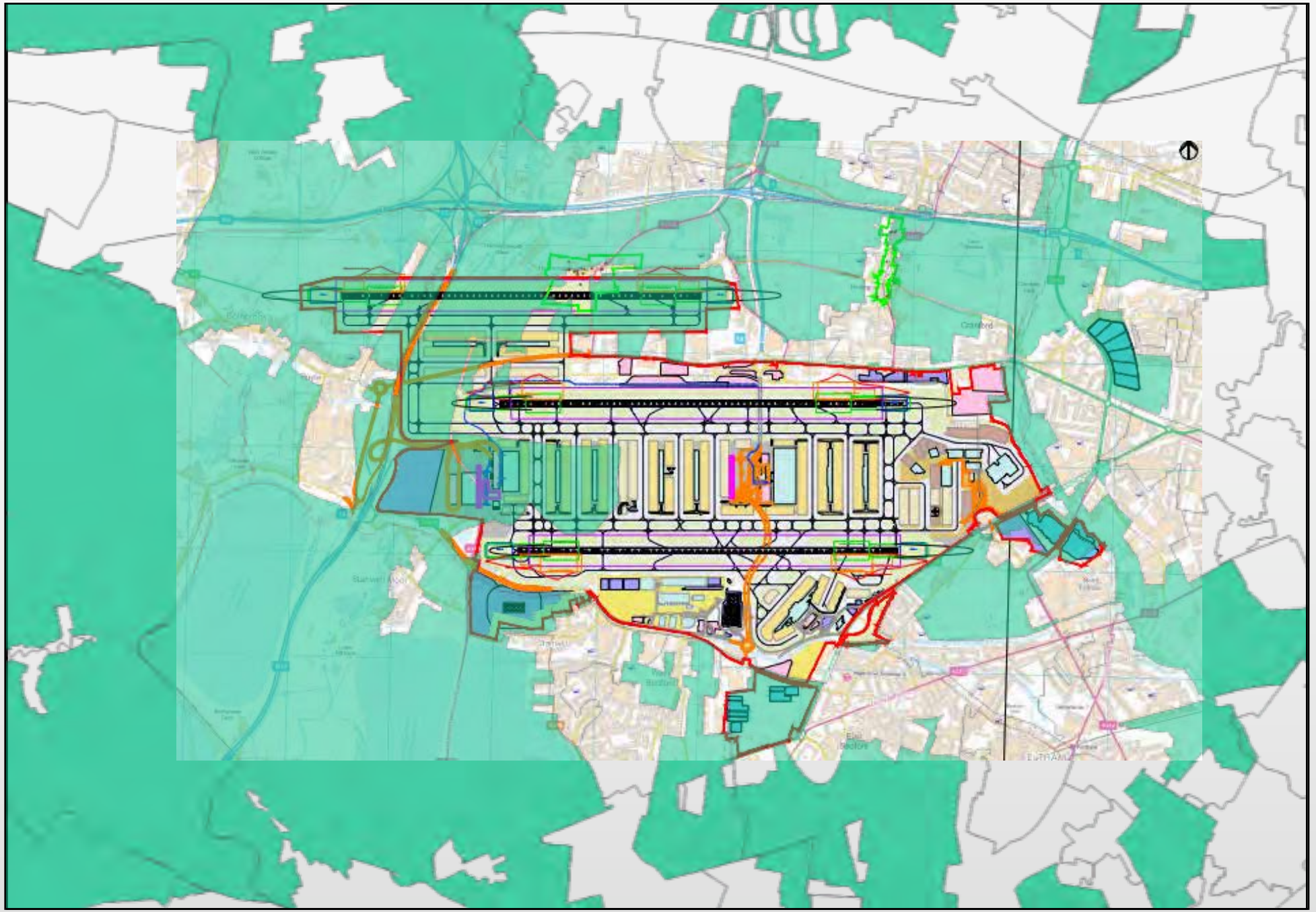
Third Runway – Heathrow, 50,000 Jobs

one the Green Belt





Simulating Very Large Scale Spatial Systems

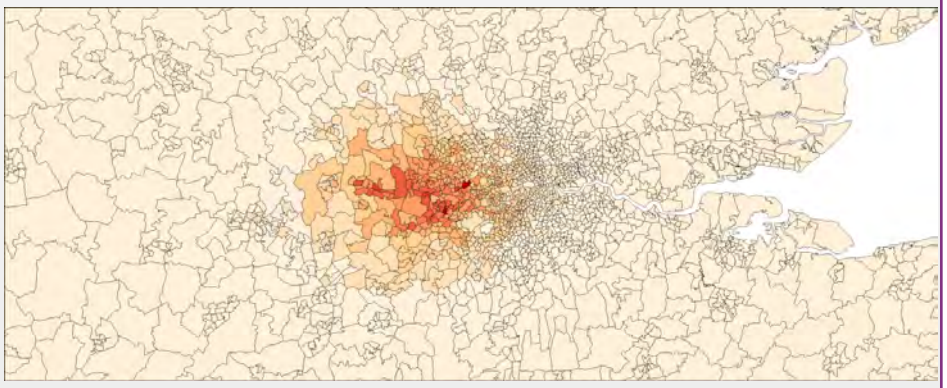


Simulating Very Large Scale Spatial Systems

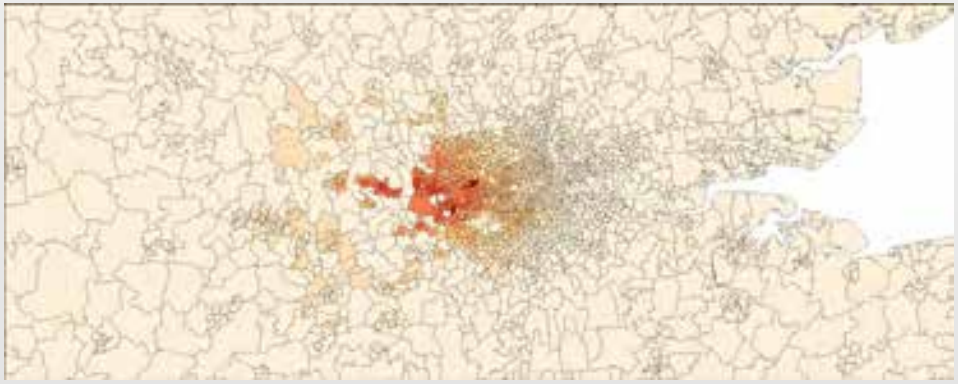
We have Tested the Impact of the Third Runway on E&W in terms of Where the Population from an Extra 50000 Jobs will Locate



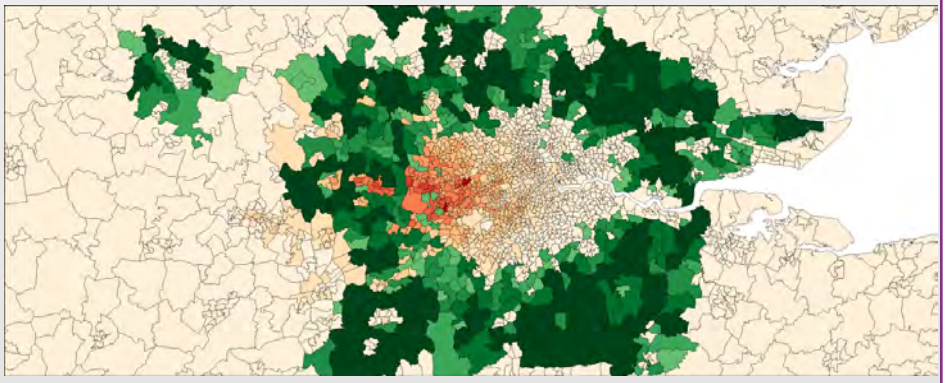
Green Belt (GB)



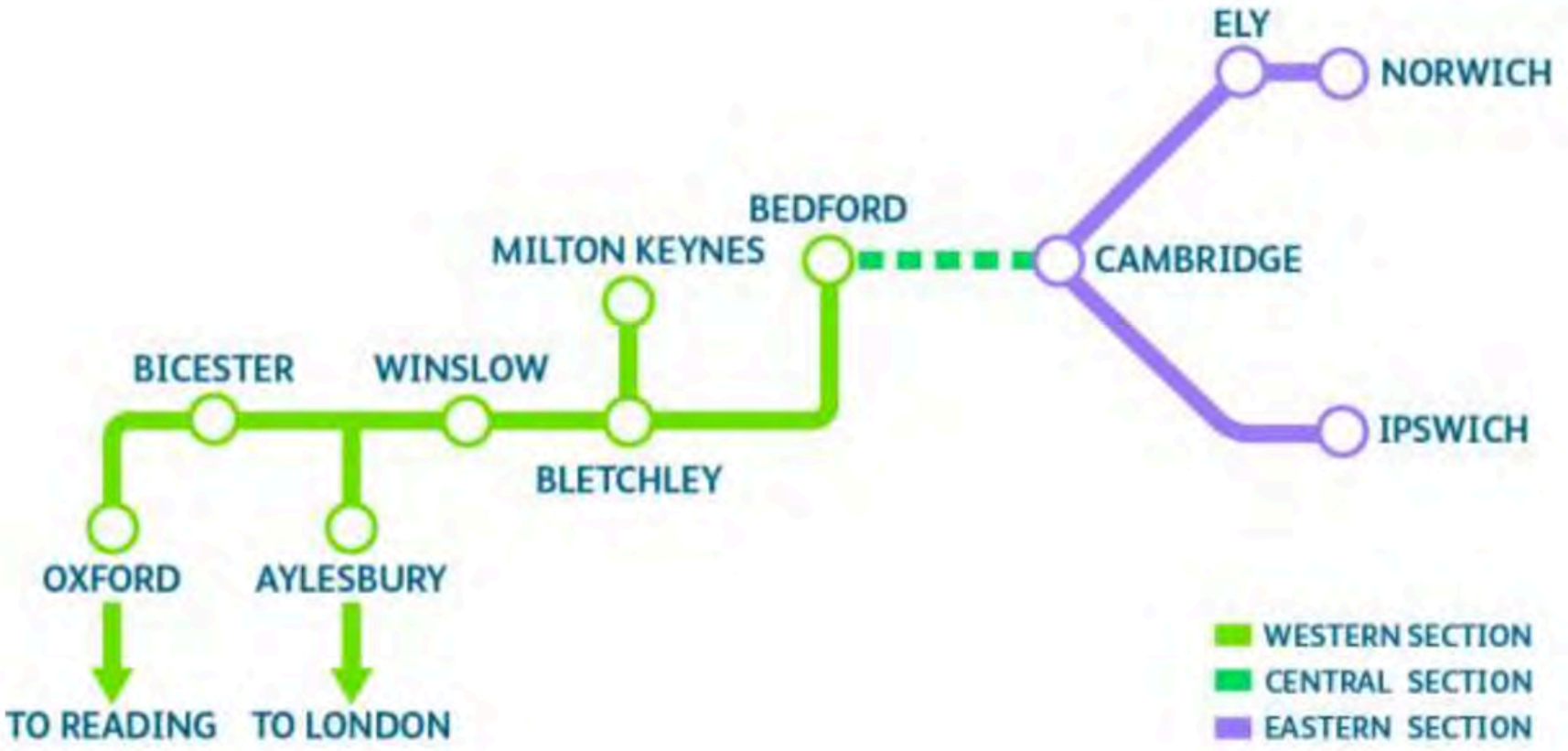
Population Change No GB

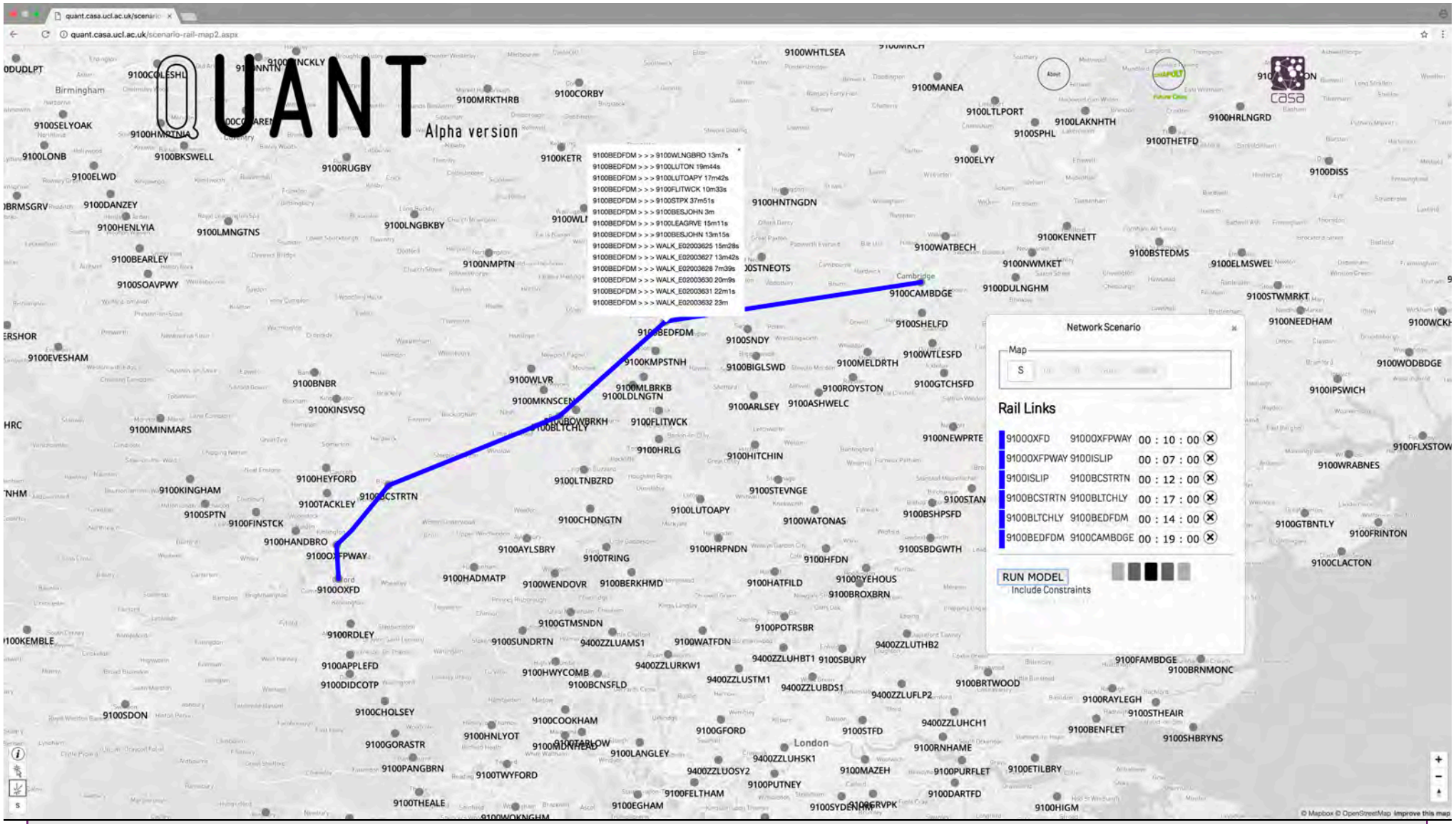


Population Change With GB

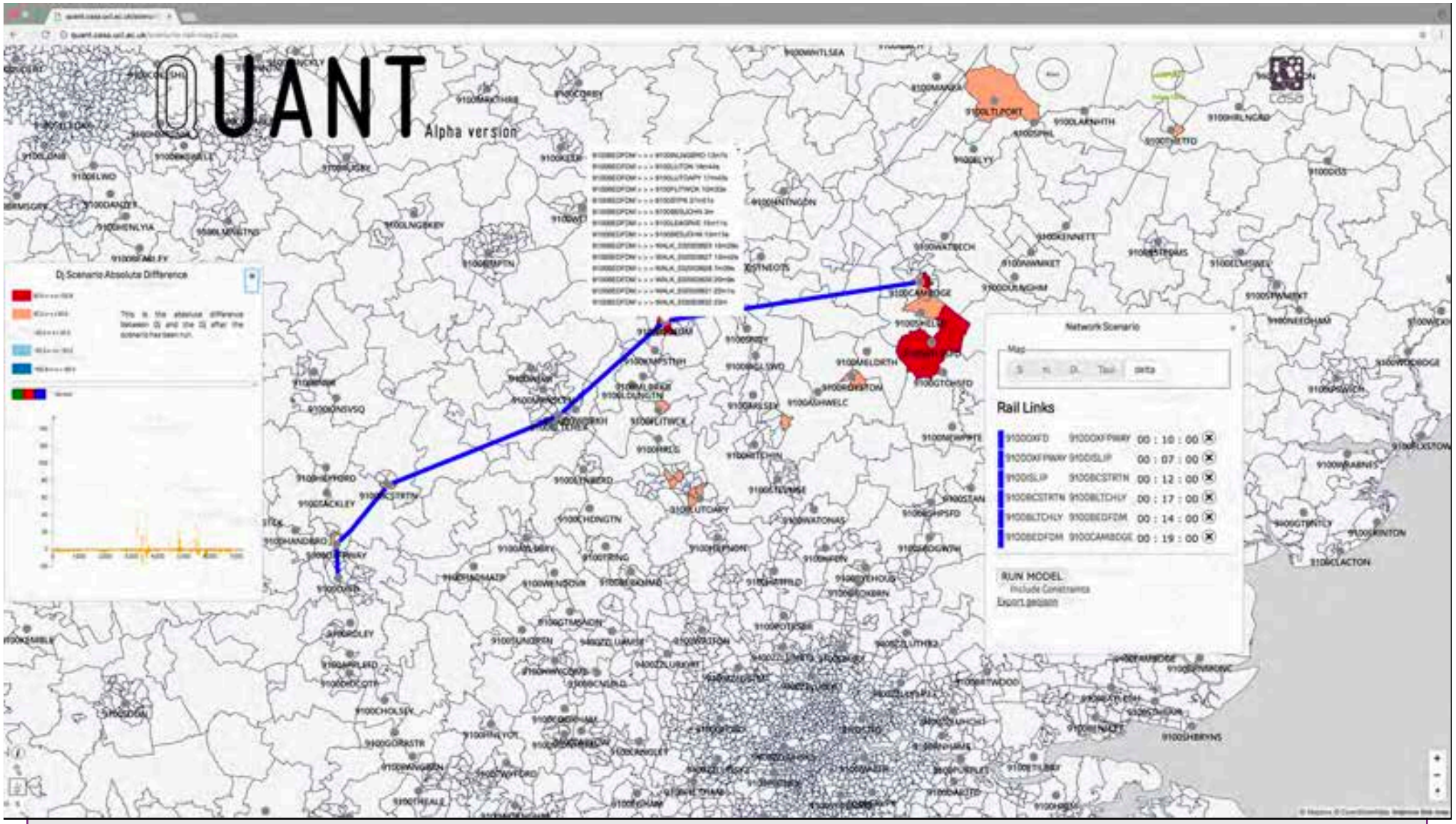


The New East West Rail: The CAMKOX Corridor

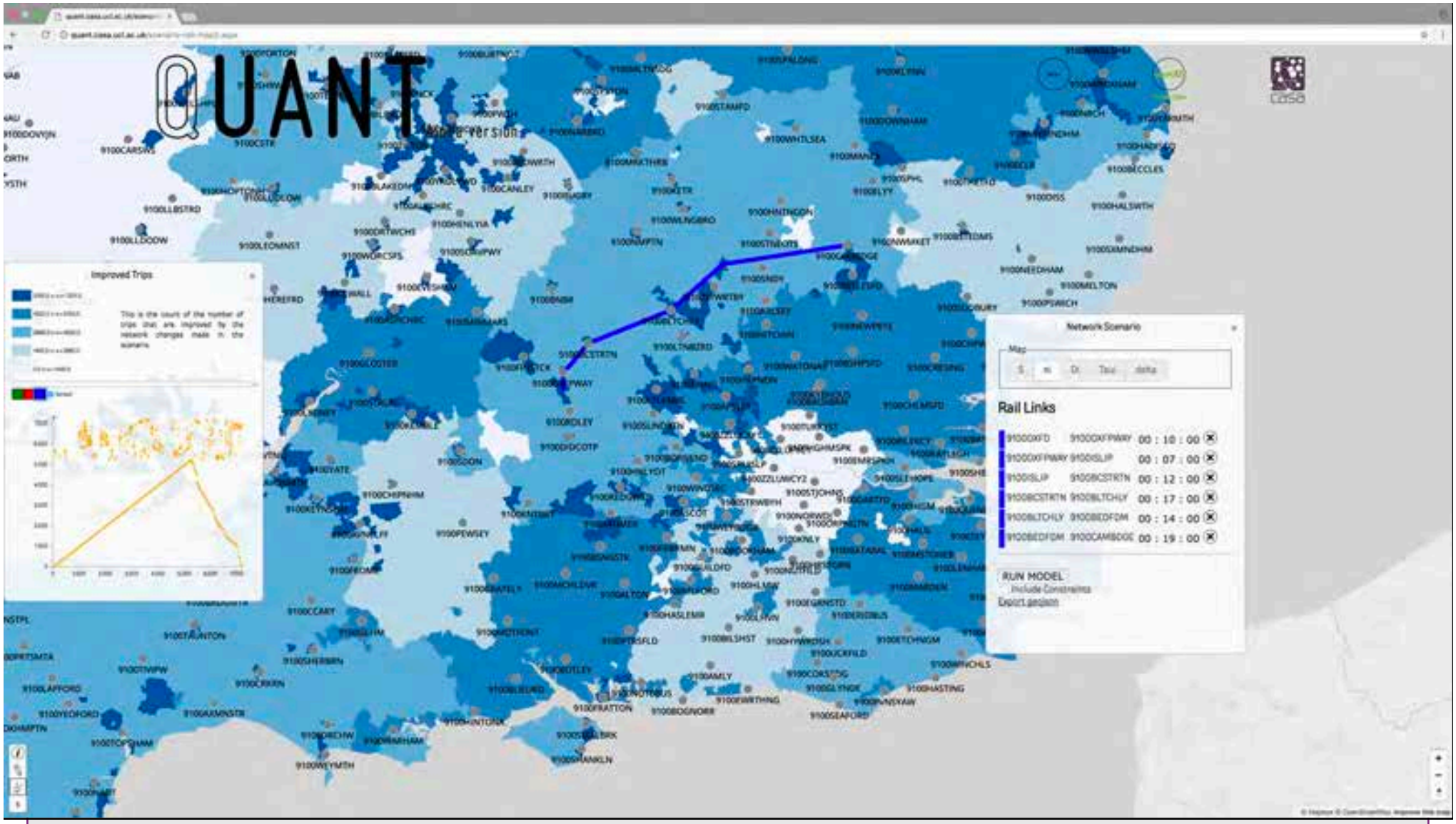




Simulating Very Large Scale Spatial Systems



Simulating Very Large Scale Spatial Systems



Simulating Very Large Scale Spatial Systems

Challenges to Be Addressed viz the Model

There are many challenges: we need to expand the modular structure and always improve the running times. In my room in the Crown Plaza I ran the model and it took about 20 seconds (while I watched the World Cup Final). It took 5 runs to calibrate but with housekeeping, this took 1 minute 20 seconds.

Now if we were to add 3 sectors then we can assume the run time of the model will increase to1 minute. But there are many variants to get an equilibrium allocation and loads of iteration can be used to solve the model.

However the key problem is the question of the spatial resolution that pertains to the network. The mismatch between the model and network resolution is key

We need to solve shortest route problems very time the network changes. Now the rail network has 3165 vertices and we can solve this problem in seconds. But the road network has 3.5 million and this takes a long time.

We need to work on much faster methods for the really big networks also noting that any scenario only affects a proportion of the network. Various decomposition techniques need to be invoked and this is a problem we need to get onto the agenda quickly.

We have many new visualisation tools to enhance the interface and we need to test the interface with users – extensively – this is a massive project as it stands but we also need to integrate this model type with other models. We will deal with these now ….

The National Infrastructure Models

There are various national infrastructure models developed in association with UKCRIC – the UK Collaborative for Research into Infrastructure and Cities – and a new facility to integrate them in a one-stop shop is being developed by the Research Councils at the STFC at the Rutherford-Appleton Lab (Harwell).

These models are highly spatial in that they deal with water, waste, pollution, transport, utilities, telecoms etc. across the UK (England, Wales and Scotland)

They are being developed by some 12 University institutions in the UK – led by Jim Hall who is a hydrologist Oxford Uni but also Professor of Climate Change there in Geography

There are various national infrastructure models developed in association with UKCRIC – the UK Collaborative for Research into Infrastructure and Cities – and a new facility to integrate them in a one-stop shop is being developed by the Research Councils at the STFC at the Rutherford-Appleton Lab (Harwell).

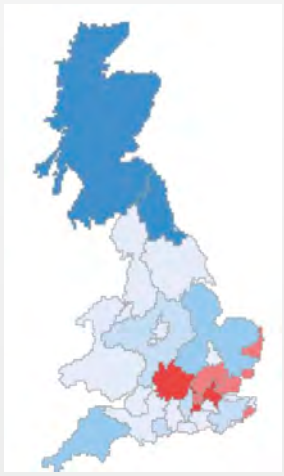
These models are highly spatial in that they deal with water, waste, pollution, transport, utilities, telecoms etc. across the UK (England, Wales and Scotland)

They are being developed by some 11 University institutions in the UK – led by Jim Hall (Oxford Uni)

The facility being developed to link these models is called **DAFNI** – ***Data and Analytics Facility for National Infrastructure***

Let me give a quick impression of the models in question – first locations models, water, socio-economics, fine scale buildings and energy











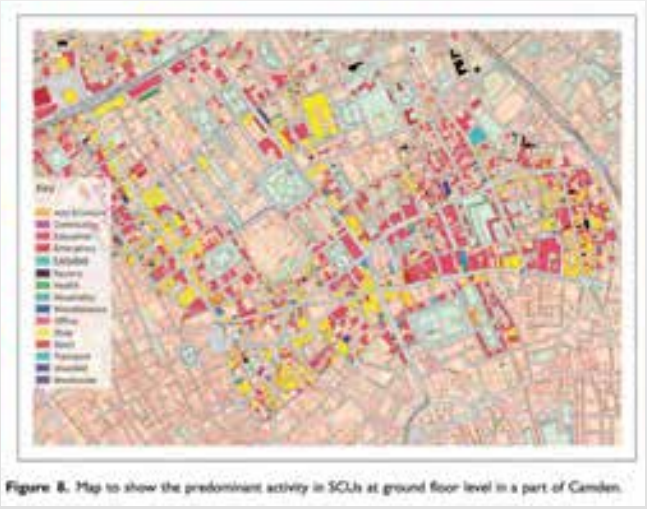


Figure 8. Map to show the predominant activity in SCUs at ground floor level in a part of Camden.



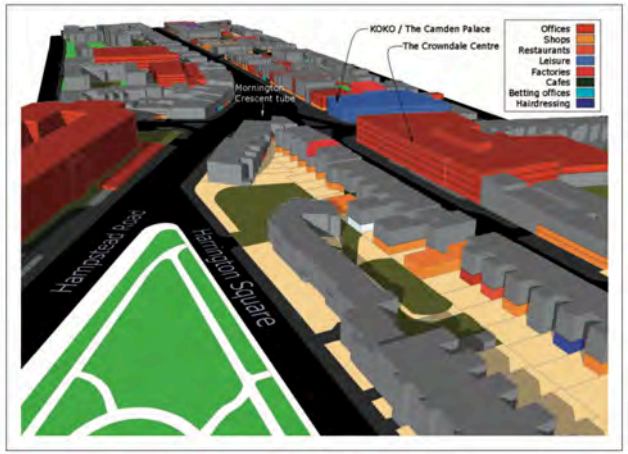
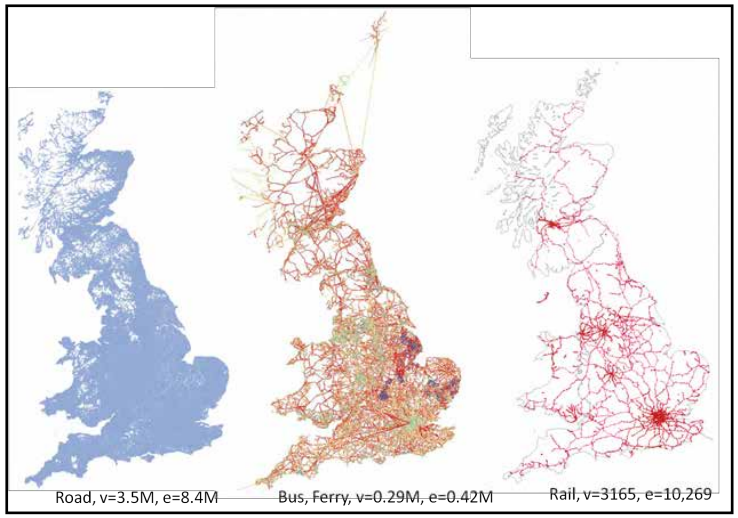
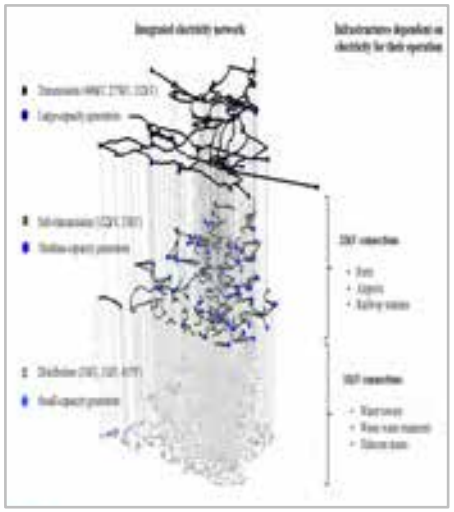
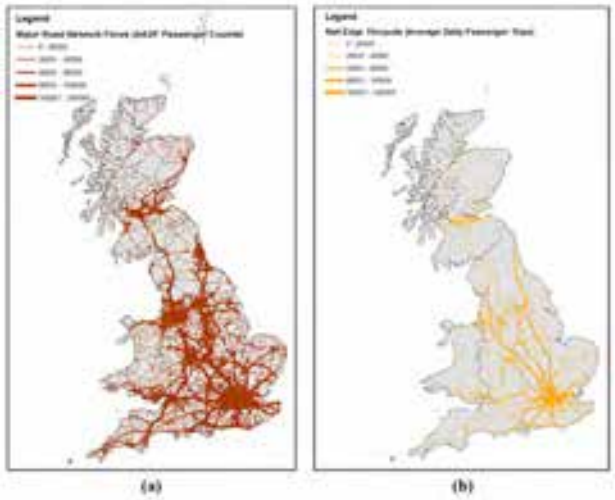


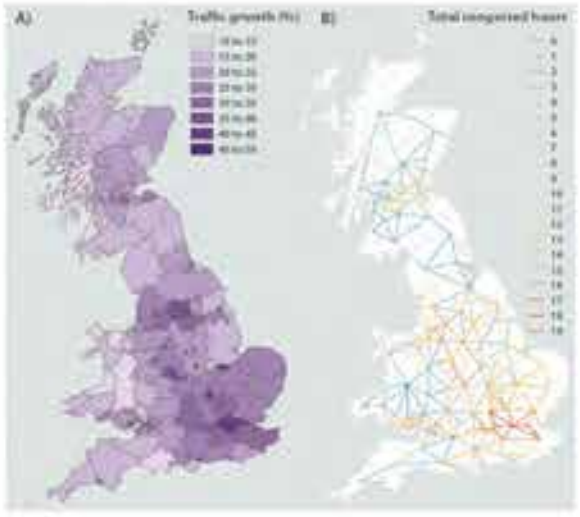
Figure 10. Three-dimensional visualisation of part of the Camden model for an area along Camden High Street. The dominant activity in each floor of each SCU is colour coded. Hereditaments that are not non-domestic (hence probably domestic) are shown in grey.

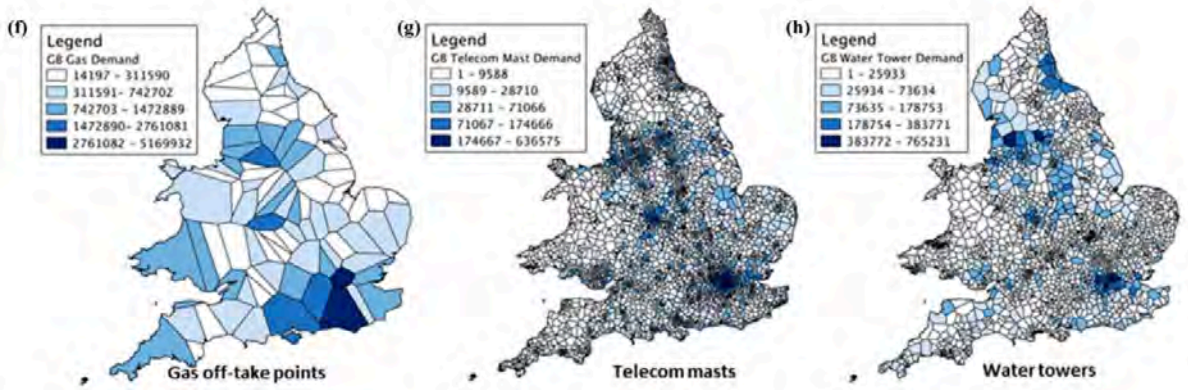
Second network models – roads, rail, electricity, gas etc











The Hardware-Software Facility at Rutherford-Appleton: DAFNI at STFC

Ok this is being built so that all these kinds of models can be shared – first and foremost, but second so they can be linked, integrated.

Not clear as yet whether a user whether that be an academic or practitioner user will be able to fashion their own models from the various components of all these models we have noted which will be available at DAFNI

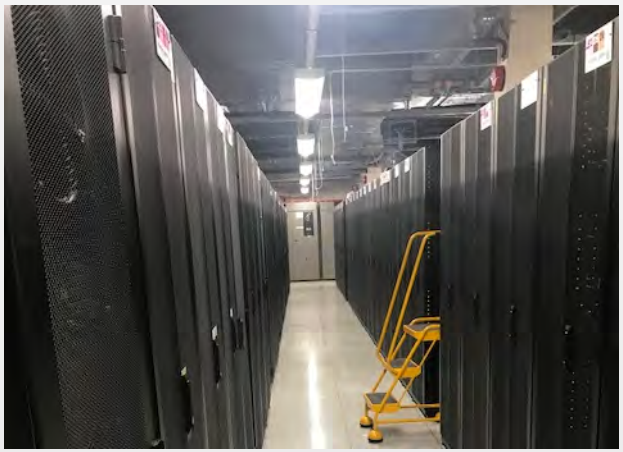
What is DAFNI – well it is a big project of dedicated hardware and software and user advice in the place where the UK Research Councils have a lot of resource – at the Science and Technologies Facilities Centre (STFC).

https://www.dafni.ac.uk/



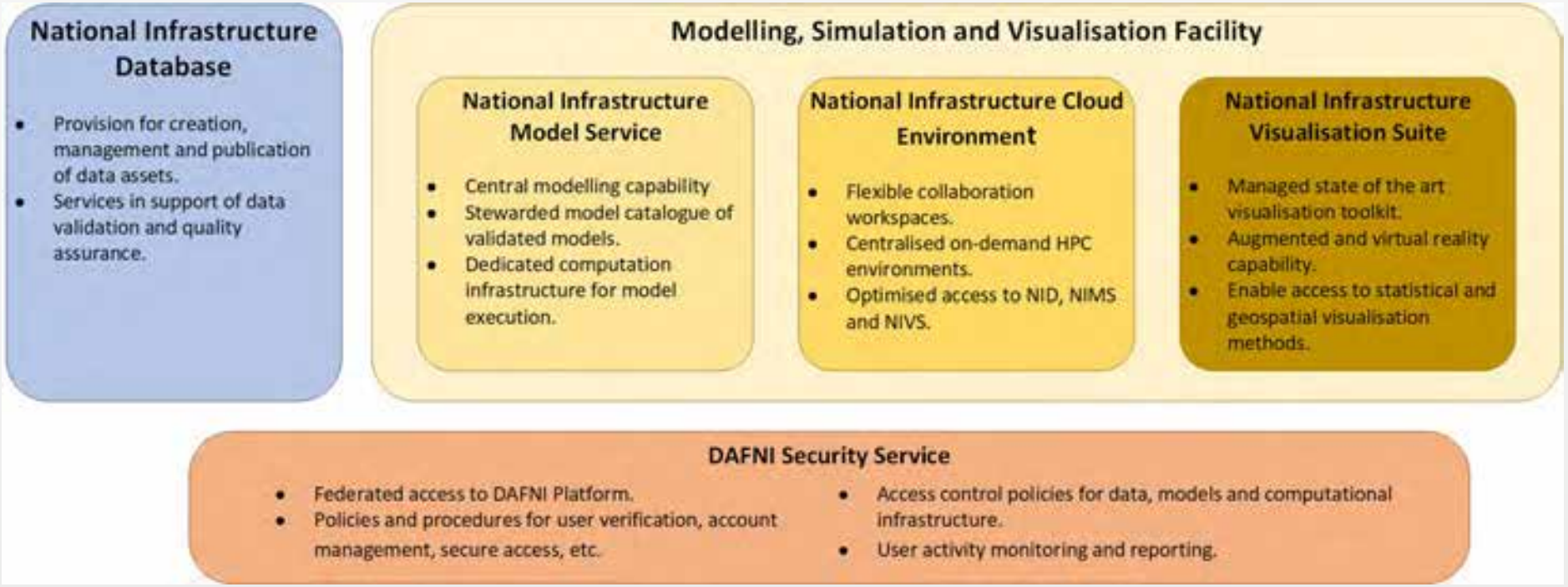






This is a professional software centre supporting and developing infrastructure for academics, primarily such physicists working at places like CERN. They have experience with JASMIN, a facility supporting climate change & environmental sciences

DAFNI Project Architecture



Currently the team are evaluating 8 pilots, building up their portfolio of models from the various ones indicated earlier. It is not clear to me precisely what the DAFNI team will actually do but they will first stitch user friendly interfaces onto these models but also try and develop a generic database from existing bases.

The Software Conundrum

The key issue is to me is – who develops the software? And how software developers interface with domain based experts such as ourselves.

There are some big questions about these relationships. They pertain to highly qualitative issues as to how our science interfaces with our users and the extent to which geospatial science in terms of software is for software to develop ever better geospatial science or for software to enable users other than us to develop applications using state of the art science.

The user interface issue requires new forms of cooperation which can be helped by portals such as those developed in DAFNI but they also require resources to do this. As well as an organisation.

Thanks

Acknowledgements

CASA: Richard Milton, Lyz Cortes,

FCC: Ewen Gibb, Scott Cain, Borkur Sigurbjornsson

DAFNI: Jim Hall, Stuart Barr

m.batty@ucl.ac.uk
@jmichaelbatty

And almost,
You do it on your phone .....



DAFNI
Data & Analytics Facility for National Infrastructure

UKCRIC
UK Collaboratorium for Research in Infrastructure & Cities

CATAPULT
Future Cities

Infrastructure and Projects Authority

National Infrastructure Commission

