UCL CENTRE FOR ADVANCED SPATIAL ANALYSIS

UCL

WORKING PAPERS SERIES

Paper 164 - Mar 11

Visually-Driven Urban Simulation: Exploring Fast and Slow Change in Residential Location

ISSN 1467-1298

casa

Centre for Advanced Spatial Analysis University College London 1 - 19 Torrington Place Gower St London WC1E 7HB
Tel: +44 (0)20 7679 1782 casa@ucl.ac.uk www.casa.ucl.ac.uk

Visually-Driven Urban Simulation: Exploring Fast and Slow Change in Residential Location

Michael Batty

Centre for Advanced Spatial Analysis (CASA)
University College London (UCL)
90 Tottenham Court Road, London W1T 4TJ, UK
m.batty@ucl.ac.uk www.casa.ucl.ac.uk

Monday, 21 March 2011

Abstract

We are developing a large scale residential location model of the Greater London region in which all stages of the model-building process from data input, analysis through calibration to prediction are rapid to execute while presenting both the structure of the model and the region to which it has been applied in the most visually accessible and immediate fashion. The model is structured to distribute trips across competing modes of transport from employment to population locations. It is cast in an entropy-maximising framework which has been extended to measure actual components of energy – travel costs, free energy and unusable energy (entropy itself) and these provide indicators for examining future scenarios based on changing the costs of travel in the metro region. Although the model is comparative static thus simulating an equilibrium at a cross-section in time, we interpret the changes that come from using the model predictively in terms of fast and slow processes – fast relating to changes in transport mode and slow relating to changes in location. After developing the model and showing how this level of spatial complexity can be handled using appropriate visual analytics, we test a scenario in which road travel costs double, showing that mode switching is considerably more significant than shifts in location which are minimal. We then discuss how these changes can be interpreted through changes in our energy and related cost indicators.

Acknowledgements

Steve Evans was instrumental in assembling the locational data and Stuart Barr and Ali Ford developed the transportation network data and analysis. Richard Milton programmed the interface to *Google Earth*. This project was part funded by the Tyndall Centre for Climate Change Phase 2: Cities Research Theme, 2006-2008.

Introduction: The Logic of Simulation

Urban simulation models appeared in the mid 1950s as computers were first used for large scale transactions processing in business and government. These models were developed against a background of belief that good and accurate predictions could be made for systems as complex as cities in terms of the impacts of urban growth and new transportation infrastructure on their form and functioning. A decade or so later, the sobering consequences of this experience provided many salutary lessons in developing large scale simulation. In short, the early models were judged to be either 'too simple', or 'not simple enough' to grapple with the daunting complexity of cities (Brewer, 1973; Lee, 1973); and so began a long period of reflection, extension and reworking of model structures in the quest to make such models more applicable and relevant to policy making. Although progress has been made, many problems remain (Timmermans, 2006).

Two key themes have dominated model development since that time. Urban models by their nature tend to treat the city system 'comprehensively' and there has been a long line in developing models that treat ever more detailed representations and functions, largely through disaggregating activities and adding new sectors. This has involved the representation of markets which balance the simulated demand and supply of urban activities at ever finer scales of disaggregation to the level of firms and households. Such models now fashion explicit links to quasi-independent transportation models or embed these models directly within their structure. They still remain largely cross-sectional and equilibrium-seeking, but tend to simulate urban change between two or more points in time, often using combinations of model types ranging from spatial interaction to micro-simulation. These models still appeal to the tradition of being 'large-scale' in that they are complicated to set up, take time to run, and are the product of teams of analysts rather than the work of individuals. The reasons for their continued development revolve around the notion that the complexity or variety of any particular application requires equal complexity or variety in the model system while policy-makers usually demand the kind of detail that such models are able to supply. Good reviews of the state of the art are provided by Timmermans (2006), Hunt, Miller and Kriger (2005), and Iacono, Levinson, and

El-Geneidy (2008). Typical of current developments is the series of models developed by Echenique (2004).

In contrast to this tradition of making simpler models more complex, there has been a less organised quest to develop models that are simpler than their predecessors. This has proceeded by decoupling submodels and developing these in more detail or by fashioning different elements of comprehensive models into individual models that are used as elements in a tool box of techniques. Many planning support systems are constructed in this fashion (see Brail, 2008) although the quest to develop simpler models is dwarfed by the wider trend of extending models to embrace new developments in information technologies and better and richer sources of data. Perhaps a clearer way of impressing this difference between simplicity and complexity, between the small and the large scale, is to adopt Bankes' (1993) distinction between 'consolidative' and 'exploratory' models (or modelling styles). The consolidative style tends to focus on the construction of models that might ultimately provide accurate, or focussed predictions in contrast to exploratory that will never do so but are used to define salient characteristics and to 'inform' the debate over particular problems. The large scale tradition in urban modelling very definitely relies upon the former while the notion of using simpler models over and over again, often modifying their structure in countless ways, accords to the latter, more exploratory viewpoint.

The model we will present and apply here is clearly in this newer exploratory tradition but it originates from the earlier large scale modelling movement. It is based on simulating interactions from work to home through a residential location model with four transport modes where the modes in question are based on road, heavy rail, tube and light rail, and bus networks. It sits squarely in the tradition of aggregate spatial interaction modelling as a singly (origin), semi-destination constrained model where flows of workers to residential zones across the four competing modal networks determine the population which locates in each of the destination zones. In this sense, it is both an interaction and a location model. This version was first developed as one stage in an integrated assessment of climate change for the London region based on a series of coupled models beginning with a national-regional input-output model (Hall et al., 2009). These employment forecasts were then scaled to small areas of the urban

region, feeding to the residential location model, the subject of this paper. These residential populations were then reduced to an even finer spatial scale using a model reflecting physical constraints on land development as incorporated in GIS, reminiscent of more physically-based urban development models in the cellular automata tradition (Batty, 2009). This enables these predictions of population to be tested with respect to the flood risk derived from hydrological models geared to account for sea level rises in the Thames and its Estuary, consistent with forecasts from the UK Climate Impacts Programme (UKCIP) for the next 50 and 100 years (Dawson, et al., 2009).

In this paper, we develop the model in a way that makes its use in assessing the impact of abrupt changes in energy/travel costs immediate, before one's very eyes, so-to-speak. This immediacy is a major requirement for communicating the modelling outcomes to a range of stakeholders who are non-expert in the particular model design. To this end, our model system is visually-driven so that the greatest amount of information about the model and its predictions can be communicated as effectively as possible to diverse audiences. We will first define a series of strict criteria that the model must meet before we begin outlining its structure. We will then examine the implicit dynamics of the model which is a cross-sectional equilibrium structure before presenting its derivation using entropy/utility maximising. This sets the method for consistently calibrating, validating, and evaluating the model where we focus on how the model handles energy use in the urban system. All these methods are then embodied in the visually-driven interface which we present as a series of shapshots of how the model is implemented. We then reach the position where we can examine the impact of abrupt and rapid change in energy costs which are encapsulated in fast changes in interaction patterns through mode shift and much slower changes in location patterns reflecting redistribution of the population. We finally evaluate these changes using changes in the energy-entropy balance consistent with the model's structure and this sets the context for generalising this method for the rapid assessment of future scenarios.

Requirements for Urban Simulation

Rapid Execution and Visually-Driven Prediction in the Dialogue with Stakeholders

The kind of predictions that this model must address are posed as the outcomes to "What If" types of question. These assume an immediate reaction or impact on the system of interest or a reaction that takes place over an unspecified time; that is for very short or very long time horizons where the outcomes in either case assume the city system will adjust to some equilibrium state which is the focus of interest. Thus accurate predictions are not the goal of this kind of model for the predictions may take many years to realise in terms of a generating a long term equilibrium. This, in fact, can never occur due to unforeseen changes and adaptations that will take place on the trajectory towards this state. In this sense, the model predictions are designed to inform the debate and engender learning amongst the stakeholders. As Epstein (2008) so cogently argues, this style of model can "… discipline the dialogue about options and make unavoidable judgements more considered". In terms of the style of model developed, this is quite consistent with evaluating predictions for 50 or 100 years which embody significant impacts due to climate change but as we shall see, the model is also capable of examining much more rapid change and its consequences.

The model must be capable of being used over and over again so that a rapid dialogue can be maintained between model builders and users. This puts an upper limit on the time required to run the model which must be in an environment that generates predictions in a matter of seconds or at most minutes, suggesting a desktop or web-based environment in which outcomes can be communicated through visual analytics such as maps, bar charts, tree diagrams, flow networks and so on, all of which might be represented in different dimensions and through various animations. These requirements are essential to a diverse community of stakeholders which we need to be clear about. In using the model for integrated assessment which consists of chaining different models together across different spatial and temporal scales, different kinds of expertise are required and visual media makes it easier to communicate model structures and outcomes to other scientists involved in models that require different disciplinary and professional expertise. Our focus then extends to stakeholders, professionals involved in policy-making of various sorts from those

trained in cognate professional and scientific disciplines but not involved in developing models *per se* all the way through to policy-analysts and advisors. Finally we consider the more visual products used to present model outcomes should also be intelligible to those involved in the policy process who are non-expert but at least informed, affected by, and instrumental in resolving the problems in question.

Models whose outcomes can be generated quickly and disseminated rapidly must also be capable of being reconfigured to embrace different features of the problem solving context that become important during analysis. This suggests that these kinds of models should be modular in some sense. Although the model structure developed here is relatively simple without any extensive modularity *per se* (although its structure can be easily replicated for other subsystems of the city system), the manner in which its outcomes can be communicated involve extensive modularity with respect to the tool kit of visual analytics available for its communication. Modularity is also essential in integrating different model types into sequences of predictions that are coupled over different spatial scales with different kinds of science being used at each stage, and this again reinforces the need for a common medium of communication between different models and model-builders. Visualisation is by far the most effective medium in which to communicate different kinds of model outcomes, thus posing additional requirements about the need for rapid and quickly repeatable model runs that can be generated *in situ* in the presence of relevant scientists, decision-makers, and stakeholders.

In terms of interpretations, outcomes hence model structures should be intelligible enough to associate causes with effects. In complex models, predictions may be the result of emergent processes that, some argue, may not be explicable in terms of our knowledge of how the model works. But here we impose the requirement that whatever the outcomes, these must be traceable to changes in input values, notwithstanding counter-intuitive impacts. In short, such counter-intuition must be ultimately explicable in terms of the model's functioning for the essence of informed debate is to recognise when and why such effects can occur. In terms of the role of the model here in integrated assessment, the organisational environment in which these kinds of models are developed is one that is highly fractured with different scientific expertise located in different places. Thus the need to stitch models together and even

to assemble elements of the same models that are built in different places requires common media of communication. Resource constraints and different expectations dictate the need for simplicity in design and communication, and all this implies, fast, simple, visual, and accessible models.

Dynamics and Comparative Statics: Equilibrium in Terms of Fast and Slow Change

The argument that cities should be treated as equilibrium structures is based on wide agreement that most cities display a similar generic spatial structure and morphology, notwithstanding distinct differences at detailed levels. Such structures also appear to persist over decades and longer and this gives power to Harris's (1970) point that such clear evidence of an equilibrium should provide the prime focus for simulation. This view does not reject the notion that urban change must also be a focus but that models that attempt to replicate urban form and structure must simulate this equilibrium prior to any additional features that such models may take on. This is given added weight in that idealised plans for cities often redefine this equilibrium in ways that are quite different from that which is observed, and thus the first stage in understanding must be to simulate what already exists.

Models that do not reflect an explicit dynamics simulate what is observed at a cross section in time and make the assumption that whenever a prediction is made, the outcomes from the model reflect the fact that the system will have moved to a new equilibrium within the given time period. Lowry (1964) in his model for Pittsburgh referred to this as an "instant metropolis" (p. 39) and suggested that forecasts made with such models must be seen '…as "quasi-predictions" of the emerging spatial structure …'. (p. iv) How appropriate this is depends on the time period in question as well as the processes of change involved. In contrast, the alternate view is that cities are forever in disequilibrium and thus simulation must focus not on replicating a static urban structure but changes to this structure, thus reflecting dynamic processes that unfold in time, and thus destroying any equilibrium that might be assumed to exist at a more aggregate level. In this respect, dynamic models tend to be more complex than cross-sectional, in that processes of change are integral to the model design. To reconcile these competing views however, the context and the question being addressed become the ultimate arbiters.

Equilibrium models can deal with both short and long term change if the intricate dynamics of the way this change works itself out in the city system does not need to be explicit. Very long term change over periods of 50 years or more as, for example, those that relate to climate change, lead to a new equilibrium that is clearly only one from a multitude of future states. The long term outcome that is predicted is purely notional in that the sheer scale of adaptation that would take place between the current and future prediction dates would be such as to destroy any idea that this outcome would ever take place. In these instances, forecasting with such models simply provides a perspective that informs the debate about the long term future which is also the case with any radical change. Very short term change however shows what might happen immediately if the prediction could be borne out assuming no other constraints on the outcome. But this too is unlikely for there are many constraints that only become explicit when an outcome is emerging. Adaptation usually happens even in the very short term and it is usually unclear how this works itself out.

There is a distinction between slow, medium and fast processes of change in urban systems first formally noted by Wegener, Gnad, and Vannahme (1986) with the slowest relating to changes in infrastructure, particularly transportation networks and the built environment, medium relating to demographic, economic and related processes, and the fastest relating to mobility ranging from local migration to flows on many different scales of network. This continuum can be further elaborated from changes in physical structures including land use which are slow, to changes in population and labour markets through redistribution and migration which are faster. In terms of the spatial interaction-location models developed here, Wilson (2008) identifies fast change in interactions, in this case, the journey to work, which is a diurnal cycle in contrast to population change in terms of the supply of housing which is more likely to take place over years. Spatial interaction-location models are usually formulated in cross-sectional terms and when used in a predictive context, it is assumed that the flows generated begin to change immediately while the ultimate locational redistribution takes longer to work itself out. In fact this process of working out is implicit and the ultimate equilibrium that occurs is a product of both fast and slow processes which have no explicit time scale. The assumption is that the outcome that is predicted would take place if all other conditions were kept the same, thus representing an ultimate steady state which would only occur under the idealised

conditions of no other change. As Wegener, Gnad, and Vannahme (1986) so persuasively note, buried within this mix of slow and fast processes there are strong contradictions as to the best ways in which these intricate processes might be modelled.

If the model is constructed for simulating changes in the demand for interaction and location, then the new equilibrium that results is one that assumes that demand is met with entirely elastic supply. We know that this will never be the case in real systems and this is thus another way in which predictions made with such equilibrium models represent an idealised future state. In most instances, changes in demand will be moderated by supply and the ultimate equilibrium will be composed of a complex process of demand and supply adapting to one another and to other exogenous constraints. It is in this sense then that predictions with this model are to be used in wider processes of planning support to inform the debate and to pose immediate answers to 'What If' types of question. To this end, we require fast and accessible models of the kind that we will now describe.

The Residential Location Model

Specification and Derivation Using Entropy Maximising

We cast the model to be developed here in the most parsimonious form possible where we explain flows between workplaces (called origins) and residential areas (called destinations) as a function of strictly physical quantities. That is, the variables that we wish to model – flows (or trips) – are measured in terms of persons but we explain them entirely with respect to physical quantities that are determined by the size and scale of the system itself. This is phrased in terms of the technological limits on how people are able to interact which relate ultimately to the geometry of the system, albeit expressed in units of cost of travel, and in terms of the land area associated with these flows. The model is in the tradition of spatial interaction (Wilson, 1970) but will be expressed in an explicit energetic framework, consistent with its application here.

Flows are defined as T_{ij}^k which are movements from origin zones $i, 1, 2, ..., I$ to destinations zones $j, 1, 2, ..., J$ with respect to the mode of travel $k, 1, 2, ..., K$. The numbers of zones is in the 100's, here 633 for both origins and destinations in contrast to a handful of modes, four in all comprising road, heavy rail, tube and light rail, and bus. We will in fact derive the model in terms of the density of trips T_{ij}^k / A_j destined for a particular zone j with residential land area A_j but it will be expressed in terms of trip volumes. The model will be subject to two physical constraints, the first based on the total cost of travel by each mode C^k which is defined as

$$\sum_i \sum_j T_{ij}^k c_{ij}^k = C^k \quad , \tag{1}$$

where c_{ij}^k is the energy expended, measured in terms of travel costs using the modal technology k to move from i to j. The second constraint is on the origin activity measured as the number of jobs E_i which provides the overall dimensioning of person activities in the system. Then

$$\sum_j \sum_k T_{ij}^k = E_i \quad . \tag{2}$$

The total number of trips in the system T is fixed implicitly by equation (2) which can be written explicitly as

$$\sum_i \sum_j \sum_k T_{ij}^k = \sum_i E_i = T \quad . \tag{3}$$

It is worth noting the particular structure of this model. The modal costs in equation (1) are constrained so that each mode is distinct in terms of the energy it uses whereas this is not the case when the trips are summed across modes with respect to their origins. This implies that the model simulates competition between modes, an essential criterion for handling switches between modes. As the basic model is a singly- or origin-constrained spatial interaction model, besides the flow matrix, the

main predictor from the model is activity destined for each residential location which is working population P_j derived as

$$\sum_i \sum_k T_{ij}^k = P_j \qquad , \tag{4}$$

but other volumes might be predicted such as employment and population by mode at origins and destinations.

To derive the model, we follow the well-established method of defining and maximising the entropy S of the distribution associated with $\{T_{ij}^k\}$ as first popularised by Wilson (1970) amongst others for transportation models. In fact, we use a more consistent definition for entropy which is the discrete approximation to the continuous form proposed by Batty (1974, 2010) which is given as

$$\begin{aligned} S &= -\sum_i \sum_j \sum_k T_{ij}^k \log \frac{T_{ij}^k}{A_j} \\ &= -\sum_i \sum_j \sum_k T_{ij}^k \log T_{ij}^k + \sum_i \sum_j \sum_k T_{ij}^k \log A_j \end{aligned} \qquad . \tag{5}$$

We next perform a maximisation of this entropy by forming a Lagrangian L using equations (1), (2) and (3)

$$\begin{aligned} L = - = &-\sum_i \sum_j \sum_k T_{ij}^k \log T_{ij}^k + \sum_i \sum_j \sum_k T_{ij}^k \log A_j \\ &+ \sum_i \lambda_i \sum_j \sum_k \left\{ T_{ij}^k - E_i \right\} + \lambda \sum_i \sum_j \sum_k \left\{ T_{ij}^k c_{ij}^k - C^k \right\} \end{aligned} \tag{6}$$

which we set equal to zero for the maximisation condition. This leads to

$$\frac{\partial L}{\partial T_{ij}^k} = -\log T_{ij}^k - 1 + \log A_j - \lambda_i + \lambda^k c_{ij}^k = 0 \qquad , \tag{7}$$

from which the model can be easily derived. Note that we will incorporate the -1 term in the multipliers λ_i without loss of generality, and we will not define new variables. The model can be stated first in log form, and then in normal form as

$$\left.\begin{aligned} \log T_{ij}^k &= -\lambda_i + \log A_j - \lambda^k c_{ij}^k \\ T_{ij}^k &= \exp(-\lambda_i) A_j \exp(-\lambda^k c_{ij}^k) \end{aligned}\right\} \quad . \tag{8}$$

There are two properties that are worth noting for they arise in the subsequent discussion. First we can produce an interesting form for the normalising equation (2) if we substitute the model in (8) into this constraint. From this is derived the value for λ_i as

$$\lambda_i = \log\left\{\frac{\sum_j \sum_k A_j \exp(-\lambda^k c_{ij}^k)}{E_i}\right\} \quad . \tag{9}$$

Equation (9) is a log sum accessibility which appears extensively as a measure of benefit in consumer analysis particularly related to transportation and as we shall see, it is related directly to the free energy in the system. An equivalent expression cannot easily be produced for the modal cost parameters λ^k. Second if we compare any two modes as modelled from equation (8), then these produce a particularly simple form of competition. Then taking the ratio of the relevant model equations for say $k = 1$ and $k = 2$, then

$$\frac{T_{ij}^{k=1}}{T_{ij}^{k=2}} = \frac{\exp(-\lambda^{k=1} c_{ij}^{k=1})}{\exp(-\lambda^{k=2} c_{ij}^{k=2})} \tag{10}$$

and this implies that modal split is, in logarithmic form, a direct function of the ratio of the relevant costs of travel , that is $\lambda^{k=1} c_{ij}^{k=1} / \lambda^{k=2} c_{ij}^{k=2}$. These issues are important

here for the development of this model is largely for purposes of comparing changes in energy costs, that is changes in c_{ij}^k.

There is one last point that will change the detail of these equations but we will not show this formally here. In some versions of the model, constraints on destination activities – in short on population densities – have been imposed and this turns the model from an origin to an origin-semi-destination constrained model in which the following constraint is imposed:

$$\sum_i \sum_k T_{ij}^k \leq P_j^{\max} \quad , \tag{11}$$

where $P_j^{\max}$ is the maximum residential population allowed in zone j. Only a subset of zones are so constrained in that in many, this constraint is purely notional in that there is so much space in some zones that it is unlikely the constraint would be breached. However in inner and more dense areas, equation (11) can be critical. If this is breached, a new parameter λ_j must to be introduced to ensure equation (11) is met and then the model needs to be solved and iterated in a slightly more elaborate fashion. Many of the above and subsequent equations would need to be modified to account for this new constraint but in the subsequent example that follows, we report the case of the pure singly constrained model.

Calibration, Validation and Evaluation

There are several ways of determining the parameter values, all of which revolve around the fact that the two key constraints in equations (1) and (2) must be met. We first write the model in its full form making explicit the origin constraint as

$$T_{ij}^k = E_i \frac{A_j \exp(-\lambda^k c_{ij}^k)}{\sum_j A_j \sum_k \exp(-\lambda^k c_{ij}^k)} \tag{12}$$

from which it is easy to see we could find the modal parameters λ^k starting with some reasonable estimates such as $\lambda^k = 1.5/C^k$, and then checking how close the predicted

value of these costs is to the observed. The iteration is then performed by changing the values of these parameters with respect to the differences between the predicted and observed costs until convergence This is akin to solution of the model using maximum likelihood or by actually maximising the entropy equation directly. In some respects this is simply dimensioning the model to the data in that no attempt is made to actually optimise the goodness of fit other than ensuring that the total travels costs for each mode generated by the model are the same as those that are observed (for an extended discussion of these issues, see Batty, 1976).

There are a number of indicators that can be generated from the model that pertain to the energy used in spatial interaction. First we will write the entropy from equation (5) by substituting equations (8) into the standard form which results in the following simplifications:

$$\begin{aligned} S &= -\sum_i \sum_j \sum_k T_{ij}^k (-\lambda_i + \log A_j - \lambda^k c_{ij}^k) + \sum_i \sum_j \sum_k T_{ij}^k \log A_j \\ &= \sum_i \lambda_i E_i + \sum_k \lambda^k C^k \qquad (13) \\ &= \log \sum_j \sum_k A_j \exp(-\lambda^k c_{ij}^k) + \sum_k \lambda^k C^k \end{aligned}$$

where it is clear that the land area term cancels from the maximisation, indicating that the way this is introduced into the equation is purely for purposes of dimensioning the distribution. Entropy S in generic terms has a structure that associates it with unusable energy in the system and in this context it is usually assumed to be equal to the actual energy C less the free energy F. In short we might write the equation as $S = C - F$. In fact equation (13) is not quite in this form as the values of the parameters from the maximisation are assumed to be negative and of course in terms of the normalising constraint on origins, this is likely to be positive. Hence equation (13) might be interpreted as $S = -F + C$ from which it is clear that total energy $C = S + F$, unusable energy plus free energy (Atkins, 1994), and this links to more formal analogies between urban structure and statistical thermodynamics (Wilson, 2009; Morphet 2010).

The real value in thinking in terms of different measures of energy becomes significant when changes in the input variables – specifically costs and employment – are made. Note that changes in land supply have no effect in this formulation because these are purely introduced to dimension the system. If land area is instrumental in bringing energy benefits or costs to bear on the system, then they need to be formulated as constraints in the manner of equations (1) and (2). If we assume only changes in travel costs, let us say each travel cost for each mode can change by an increment or decrement Δ_{ij}^k as $c_{ij}^k(2) = c_{ij}^k(1) + \Delta_{ij}^k$, then we can compute the change in entropy as a change between free and actual energy, that is $\Delta S = \Delta C - \Delta F$. Using the above definitions, the actual computation can be written as

$$\begin{aligned}\Delta S &= S(2) - S(1) \\ &= \log \sum_j \sum_k A_j \exp(-\lambda^k c_{ij}^k(2)) - \log \sum_j \sum_k A_j \exp(-\lambda^k c_{ij}^k(1)) \\ &+ \sum_k \lambda^k C^k(2) - \sum_k \lambda^k C^k(1)\end{aligned} \quad (14)$$

Equation (14) can be further simplified to

$$\Delta S = \log \frac{\sum_j \sum_k A_j \exp(-\lambda^k [c_{ij}^k + \Delta_{ij}^k])}{\sum_j \sum_k A_j \exp(-\lambda^k c_{ij}^k)} + \sum_k \lambda^k \left\{ \frac{\sum_i \sum_j T_{ij}^k(2)(c_{ij}^k + \Delta_{ij}^k)}{\sum_i \sum_j T_{ij}^k(1) c_{ij}^k} \right\}. \quad (15)$$

The free energy term which is the first on the right hand side of equation (15) is reminiscent of consumer surplus and there is a degree of intuitive sense in this derivation. It might even be possible to simplify these measures further but the form in which they are is probably the most useful for our analysis. The crucial issue is to examine actual changes in the three overall measures – ΔS entropy, ΔF free energy and ΔC actual energy – and all of these measures will be used in the subsequent analysis to show how changes in travel costs have repercussions on accessibility, benefits as well as the total amount of energy used or distance travelled.

The Visually Driven Interface

Principles for Visualisation

The general principle that we ascribe to here is to put as much information used and generated by the model as possible into the display device used to communicate the model's data and predictions as well as its implementation for the user. In this context, our display device is essentially the desktop, possibly the desktop linked to internet through a web browser, as well as more conventional media in terms of paper products. There are three key stages in the model-building process that these models are constructed around: first exploration of the model's input data which reflects the way the system of interest is articulated in terms of the model, second the calibration or fine-tuning of the model to this data (as well as its interpretation through validation and verification), and third the generation of model outcomes as predictions. In each of these three stages which drive the process forward in sequential fashion, the same graphics tools are used to display information visually in the form of maps, flows, networks, tree diagrams and so on as we illustrate below. This is a process we have developed before using more primitive graphics (Batty, 1983; Batty, 1992).

The interface is organised as one main window controlling key operations and displaying key outputs and two kinds of tool bar: first the main tool bar which strings each stage of the modelling process together in an implied sequence of progression from data to prediction, and a second tool bar that is launched at each stage when graphical outputs of various sorts are required. The main tool bar begins with data input, normalisation of the data, and then data exploration which launches the second tool bar for display which is central to the process of exploratory spatial data analysis. Then the particular model variant can be chosen and this leads immediately to a window that is launched in which the model is fine tuned or calibrated to observed statistics, followed immediately by the second tool bar from which the model outcomes at calibration can be explored visually in terms of their goodness of fit. Finally predictions with the model can then be activated from the main toolbar: scenarios can be either imported from file or constructed on-the-fly which involves altering locational data concerning employment, floorspace, network characteristics, and travel costs for the various modes. Once the scenario is built, predictions are

generated and then these can be once again explored from the second tool bar which provides similar graphical display capability as at the data input and calibration stages.

The graphics tools which are accessed through the second tool bar mainly display media in the form of 2D thematic maps, maps that produce 'desire lines' recording interaction from origins and destinations in proportion to their flow volumes, histograms or bar maps showing activity volumes, tree maps that display the hierarchy of activities in proportion to their volume in small areas and the next level of hierarchy which in this case is the London boroughs, and scatter graphs of trips and travel costs. All these data can be displayed as either counts or densities, and there are several derived maps that are built from comparing one activity to another in ratio form. To enrich the analysis by comparing model data and predictions with spatial data that cannot be imported into the model, we have enabled all the data produced by the model in map, flow or histogram form to be exported on-the-fly to **Google Earth** where it can be compared against network data such as roads and rail lines and a variety of raster based data such as topographic and climate layers. Data is exported to **Google Earth** in XML format which enables KML files to be constructed from the vector data that the model generates. This is accomplished in real time when the model is running so the user can display data in 3D, flying through it to gain a rich and detailed impression of how the data used and generated by the model compares to other features of the region in question input to **Google Earth**.

Figure 1 illustrates the basic template – the main window and the two tool bars where the window in this case is the panel that displays the data input. From this panel the user can explore the data numerically, can query the map for the location of zones and higher level units and can get some sense of the correctness and the dimensionality of the data in context. On launching the model, the splash screen first occupies this window and once the data input has been displayed a window controlling the normalisation of the model data and thence the choice of the model and its calibration are launched within the template, always leaving the location map to the right on screen to keep the user orientated. Once calibration is finished, the main window is refreshed to enter the stage of defining and thence generating predictions which we illustrate in the next section.

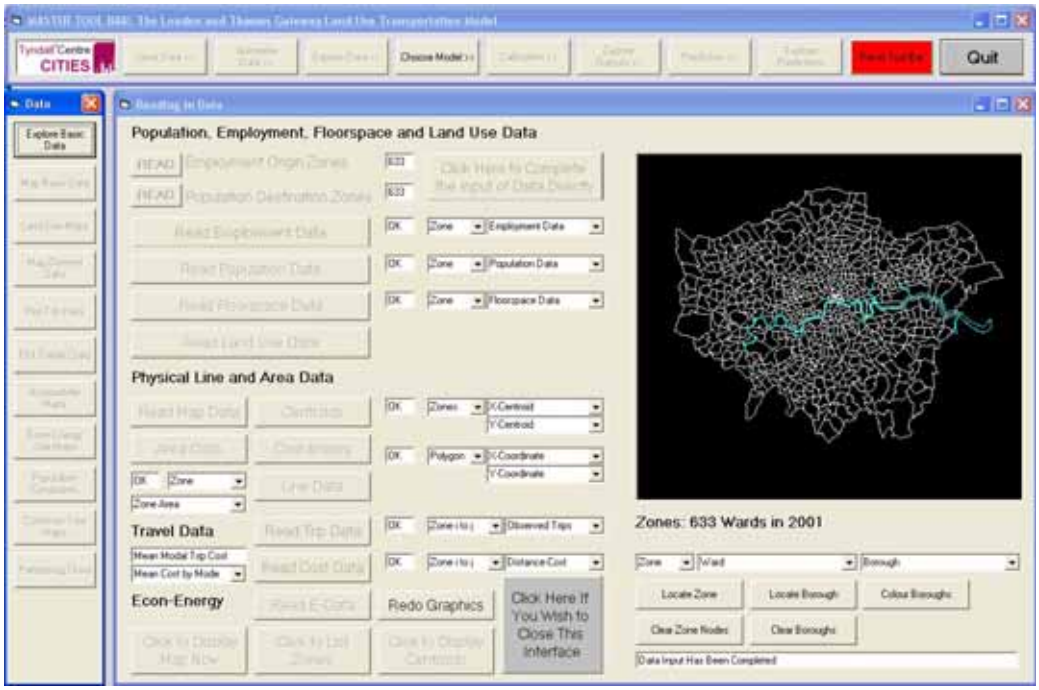

Figure 1: Windows Comprising the Basic Interactive Model Template

Exploratory Data Analysis, Calibration and Generating Scenarios in the Desktop Environment

The main tool bar is straightforward in that it drives and directs the user to the sequence of stages defining the model building process from data through to prediction but the second tool bar controls the graphics and is launched at each of the three key stages. We will examine the tools that are available at the data analysis stage but these in fact are similar to those used for the model's calibrated predictions and scenario predictions at the second and third stages. The tool bar contains 10 key display types, seven of which are maps of various kinds: histograms showing activity volumes by location, thematic maps showing the same volumes by area, and flow maps showing different flow volumes from any origin to all destinations or vice versa. These displays can also be queried with respect to individual locations and individual flows in interactive form although for the most part, the data is completely mapped each time a map is drawn. One key distinction is between count and density data, so, for example, population $\{P_j\}$ is plotted as an absolute count which can then be compared against its density $\{P_j / A_j\}$ where A_j is land area devoted to residential

use in zone j. All other variables can so be defined as densities from the land use areal data that is available for different land use types.

The first button enables to user to query individual location and flow data for population, employment (densities and counts) and for trip data $\{T_{ij}^k\}$ from each origin or destination over any mode of travel k. The second button enables the user to plot these as complete maps which can then be exported to display within **Google Earth**. Buttons displaying each land use as a thematic map, and then derived data such as activity rates such as P_j / E_j, follow and then the user can plot trip data such as $\{T_{ij}^k / E_i P_j\}$ against travel costs $\{c_{ij}^k\}$ as scatter graphs. Buttons activating maps as cost surfaces from any origin to all destinations for any mode and vice versa followed by detailed accessibility surface maps based on potential, consumer surplus and related indices, again for specific origins or destinations by mode, can then be plotted. The final maps reflect wages, and house prices data not utilised in any of the applications here, and then population constraints are displayed in terms of land availability. Last but not least, we let the user plot any of this location data as tree maps which effectively represent the volumes of activity in each zone tagged to each higher level unit – the borough in this case– in terms of proportional rectangles.

In Figure 2, we present a small collage of these kinds of display, showing population density, employment counts, road trips from zone 6 (Heathrow Airport), travel costs, accessibilities, and the tree map for residential land area. We use these maps to learn about the region in terms of its structure which, from Figures 2a and 2b, is strongly monocentric with respect to employment densities and counts. Note how the congestion charge zone is picked out in terms of the road accessibility from Heathrow but is not featured in any way in the travel cost map from the centre of town (Charing Cross) in terms of the tube and light rail network. To supplement this visualisation, for a number of these layers we can export the input data to **Google Earth** and in Figure 3, we show how we can plot employment counts as histograms, population density in thematic map layer and the flow data over the road from zone 6 in 3D, successively updating this visualisation as we continue to generate the data from various display produced from the model. There is much else we can explore using this capability but Figure 3 gives some sense of this potential.

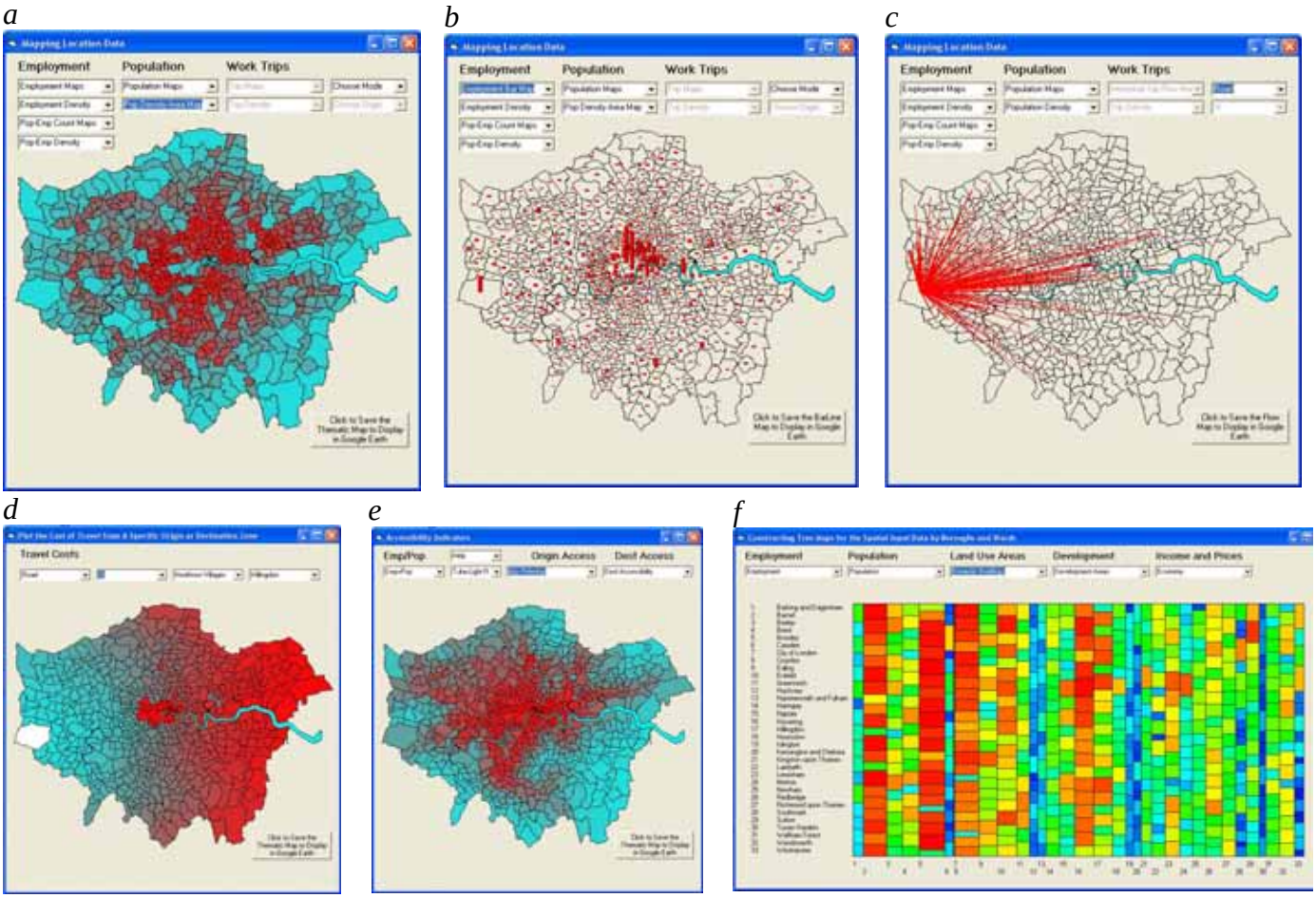

Figure 2: Small Multiples of Graphic Output from Exploration of the Model Data

a) population density b) employment counts c) road trips from Zone 6 Heathrow d) travel costs from Heathrow e) accessibility potential on the Tube system f) Tree map showing residential areas at Borough and Ward levels

We can initiate the same learning cycle with respect to comparing the predictions from the calibration to the data using similar map layers as well as direct comparisons of deviations between observed and predicted activities while similar displays are available in exploring the impact of various activity and interaction-network changes. However in driving the process forward, we first need to normalise the travel cost data to relate it to other costs and to represent very large costs in an appropriate way, and then we have the option of choosing the attractor for the model which is a function of the land area A_j. In the model demonstrated here, we simply use the raw variable as implied in equations (8) and (12) above. We initiate the calibration using an iterative method to ensure that the trip lengths in equations (1) are reproduced and we use a damped Newton-Raphson (hill-climbing) method to ensure that this

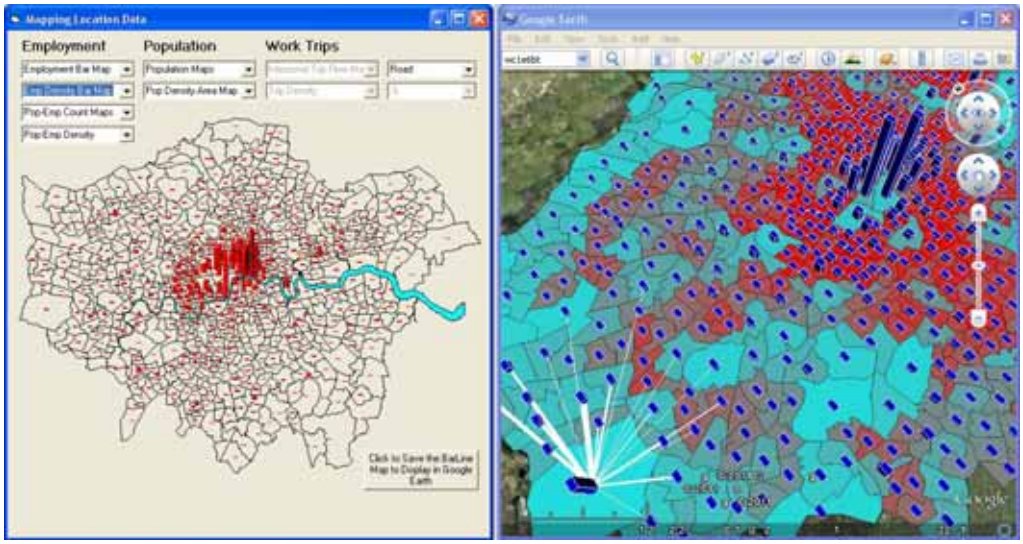

Figure 3: Visualising Thematic Map Layers, Flows, and Histograms Using Google Earth as an External Viewer Linked to the Desktop Interface

convergence takes place as quickly as possible (see Batty, 1976). In terms of the fit of the model that we report here, this is quite modest in that some 62 percent of the variation in the population and 43 percent in terms of the overall trip matrix explained but as the purpose of this paper is to introduce the framework and not comment on the substantive results, we show the simplest possible unconstrained version of the model, not the one that generates the best goodness of fit. In Figure 4, we show the typical fit of the model in terms of population counts and densities.

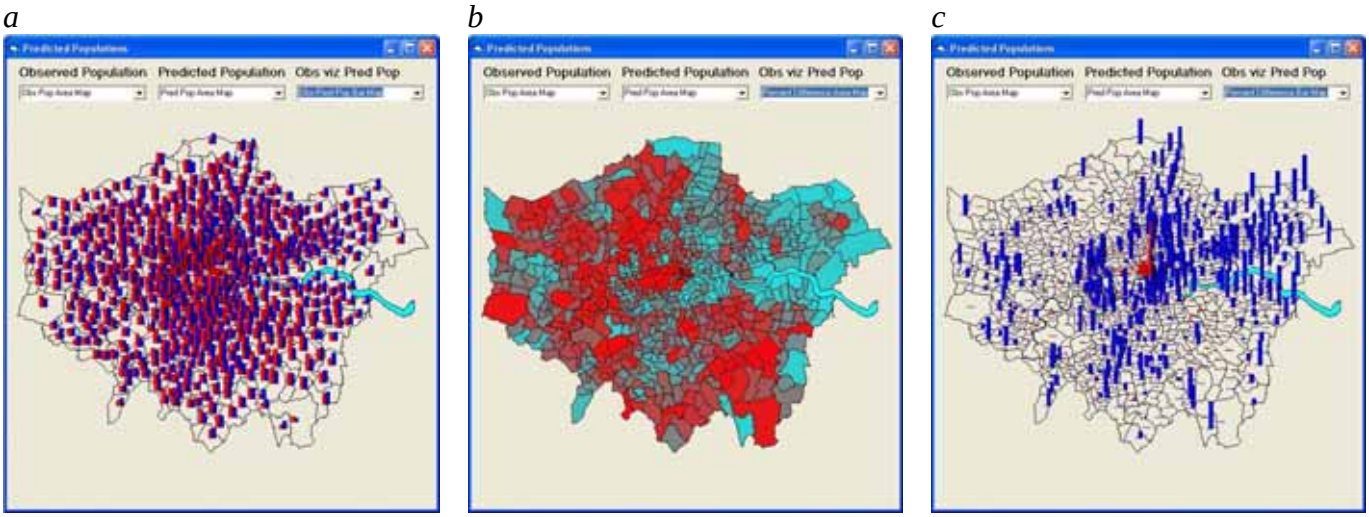

Figure 4: Predictions of Residential Population from the Model

a) Predicted and Observed Populations b) Percentage Differences (red overprediction, blue underprediction) c) Bar Graph of Percentage Differences (red – over and blue – under)

The last stage in this process involves testing the impact of scenarios. These are framed in terms of changes to the input variables – employment, travel costs, land area constraints and so on. Users can import new data files which contain these scenarios or can develop them directly in the visual interface screen activated in the main window. We show one such screen in Figure 5 which illustrates how the user can add to the network by drawing a new transport route – in this case a line from Heathrow to Kings Cross – a Cross-Rail – which when input to the model enables the shortest routes to be recalculated for that mode. Many such changes defining a scenario can be input for any set of locations and modes and in this way, the future that is to be tested, is assembled. We are then able to generate the relevant predictions and explore these using the second tool bar which activates the graphics. We now have three possible sets of comparisons to make: between observed and calibrated data, between observed and scenario predictions, and between calibrated and scenario predictions. In fact, the appropriate comparison is the latter one in that any errors introduced by the model need to be factored in so it is wiser to compare calibrated rather than observed with predicted when generating some measure of change or impact.

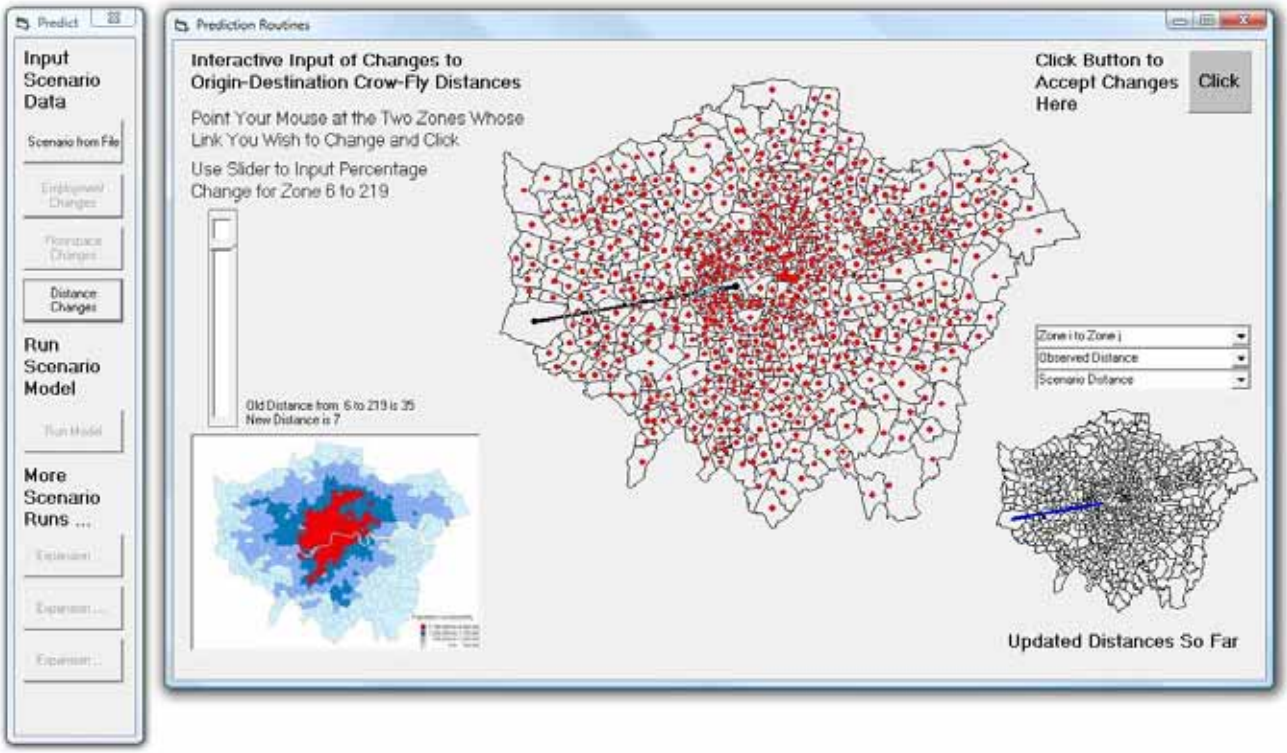

Figure 5: Building Scenarios on-the-Fly: Inputting a New Heavy Rail Line from the Airport at Heathrow to the West End

Slow and Fast Change: The Impact of Urban Energy Costs

Although this genus of model essentially simulates a world in equilibrium, using the model to predict future change implies a variety of dynamics that are assumed to work themselves out completely by the time the new equilibrium is established. Lowry's (1964) original idea of the 'instant metropolis' was predicated on the basis that contained within the existing equilibrium, were emergent structures that would be revealed when the model was calibrated but would only truly show themselves when predictions to a future state were made with the model. Thus the notion of comparing the calibrated against a future state would be one of comparing the implied equilibrium of the present (not the same as the actual present) with a future equilibrium once new changes embodied in the scenario had worked themselves out. The meaning of these predictions in terms of simulating relevant changes thus turn on the processes that are implicit in the internal dynamics of such models.

There are clearly an array of different dynamics that are involved in any changes in location and trip-making decisions. Our focus here is on changing costs and in terms of trip-making, if cost increase the response is likely to be rapid in that trip-makers will switch to lower cost modes. If the capacity of the mode which appears more cost effective is limited, switching to such a mode might increase congestion, thereby increasing costs to the point where the original mode switch is discounted. This process might take time to work itself out but it is likely to be a lot faster than other types of location change. Changing locations with response to such costs clearly takes longer as there can be no immediacy in making a residential move and it is likely that changes in jobs (which are not part of this model) have a longer dynamics. The real issue is those effects that are second, third order and so on which are in fact longer term adaptations that we know little or nothing about because they are often obfuscated by other changes.

We can strictly differentiate here between changes in modal split and changes in location. Changes in both however ultimately translate themselves into changing infrastructures which tend to be slow – that is changes in the built environment – in contrast to fast changes that involve people using the same infrastructure but in

different ways generating different volumes of activity and indeed involving different individuals (Wegener, Gnad, and Vannahme, 1986). In fact our model was originally designed to examine the impact of very long term changes in climate specifically rises in sea level on the locational pattern of the population in Greater London over the next 100 years. There will be substantial adaptation to and indeed mitigation of these effects over this period which would clearly lead to a future state very different from the equilibrium that this model would predict.

In fact the equilibrium model is important so that these other changes that will inevitably occur can be filtered out. Such models are classic 'What-If' types of instrument in that they are used to pose and answer questions of the kind "assuming every else remains the same and *x* changes, what is the effect on the system of interest?". In short, the model can be used to generate the causal chains that are exposed by this usage in terms of definite and explicit impacts, in this case on interaction and location. Although the model can define how much change is due to changes in interaction versus location, the balance in fact cannot be attributable to anything other than the structure of the entire model that contains many such causal chains. Changes in infrastructure however are harder to gauge because individuals switching transport mode, route and location in response to such changes clearly take place over much longer intervals. Our key example here in fact will be changes in cost not infrastructure and we will examine one from many such possibilities. In July 2008, gasoline reached $148 a barrel and at this point, a phase transition was triggered which became immediately evident in places where the predominant mode of travel was by car. In Los Angeles, flocks of cyclists appeared on the freeways as car users (and riders) swapped one mode of transport for another only to disappear again very shortly thereafter when the price of gasoline fell. In this context, what we will do here is examine the impact of a doubling of the cost of road transport relative to all other modes: that is we will increase the cost of road travel uniformly over the system doubling the unit cost and keeping the costs of the other three modes – heavy rail, tube and light rail, and bus constant. The residual mode which caters for all other users remains as a residual to the analysis for it is not formally modelled. Formally the change from state (1) to state (2) in transport cost for the car mode $k = 1$ is written as $c_{ij}^{k=1}(2) = 2c_{ij}^{k=1}(1) = c_{ij}^{k=1}(1) + c_{ij}^{k=1}(1)$. If we substitute this into the modal split

comparator equation (10), we can see that the relative shift in trips between any mode is a simple function of the previous time step, that is

$$\frac{T_{ij}^{k=1}(2)}{T_{ij}^{k\neq 1}(2)} = \frac{T_{ij}^{k=1}(1)}{T_{ij}^{k\neq 1}(1)} \exp[-\lambda^{k=1} c_{ij}^{k=1}(1)] \tag{16}$$

where it is clear that as the cost gets greater for the mode with the increased cost, the percent shift gets greater. This is logical given that trips decline exponentially with respect to travel costs.

When we double costs in this manner, the model predicts shifts in all modes as travellers seek to travel on more cost effective routes and as the model is singly constrained, there will be shifts with respect to their residential locations. There are two key indicators – first the total average travel costs which we would expect to rise for road travel (although modal switches are likely to mean this cost will not rise by the exogenous change of 100 percent), and second modal split. These statistics are presented in Table 1 where it is clear that the overall average trip costs rise by 17 percent of which by far the largest component of this is the increase in road trip costs by 27 percent. Rail and tube only rise between 2 and 3 percent while the average cost of bus travel drops by slightly less than 2 percent. These changes are almost the inverse of shifts in modal split where car ridership decreases by 46 percent in contrast to bus transport which increases by some 42 percent. Heavy rail and tube ridership also increase substantially by 35 and 21 percent respectively. Such big shifts might be expected to be associated with big shifts in residential location activity which we will now examine. In fact we must note that such large shifts would not actually occur for they would require big increases in rail infrastructure and in a massive extension to the bus fleet. However they are indicative of the pressures in the system and in this sense, quite consistent with the idea of using equilibrium models to explore such future possibilities.

Changes in trip volumes between the existing and new states (1) and (2) lead directly to changes in activity at residential destinations through equation (4). In difference terms, these changes can be portrayed as

$$\sum_i \sum_k T_{ij}^k(2) - \sum_i \sum_k T_{ij}^k(1) = P_j(2) - P_j(1) \tag{17}$$

where it is clear that as the total number of trips is conserved by definition as $\Sigma_{ijk} T_{ij}^k(2) = \Sigma_{ijk} T_{ij}^k(1)$, then the sum of the differences between residential activities across all locations is zero, that is

$$\sum_j P_j(2) - \sum_j P_j(1) = 0 \quad . \tag{18}$$

In short, changes in costs simply lead to a redistribution of existing activities and the new equilibrium predicted by the model – which is clearly a first order equilibrium in that there is nothing in the model to predict further order effects – is composed of these locational changes and the shifts in mode split shown in Table 1.

We can compute two rather graphic illustrations of these locational shifts. First we compute the absolute proportion of all population moving as

$$\Phi = 100 \frac{\sum_j \left| P_j(2) - P_j(1) \right|}{\sum_j P_j(1)} \quad , \tag{19}$$

and we can also partition the system at any point into two sets of zones that call Z_1 and Z_2 where the entire set of zones is $Z = Z_1 \bigcup Z_2$. Thus

$$\sum_{j \in Z_1} [P_j(2) - P_j(1)] = -\sum_{j \in Z_2} [P_j(2) - P_j(1)] \quad , \tag{20}$$

Equation (20) means that we can partition the system into any two sets of zones and examine the flow from one subset of zones to another. In this way, we can examine if the effects of locational constraints are spatially-biased towards any locations in the system, specifically in this case either towards the inner city or the outer suburbs which reflect major differences in terms of car usage (Batty, Hall, and Starkie, 1974).

Table 1: Changes in Average Trip Costs and Modal Split

Mode	Observed Mean Trip Cost	Predicted Mean Trip Cost	Percent Difference	Observed Modal Share	Predicted Model Share	Percent Difference
Road	38.668	49.157	27.124	0.389	0.210	–45.899
Heavy Rail	77.780	79.591	2.328	0.122	0.165	34.997
Tube	59.662	61.196	2.570	0.331	0.400	20.988
Bus	14.659	14.428	–1.576	0.158	0.224	41.926
All Modes	47.600	55.621	16.851			

The most surprising prediction from the model is that the percentage of the working population shifting residential locations Φ is only 2.4 percent which involves some 110736 persons. This is extremely low and it is a measure of the resilience of the system to changing transport costs. In fact overall costs might rise substantially as we will show below but actual second order costs due to potential shifts in residential location are likely to be much less than might be expected. To an extent, this result is simply indicative of the fact that there are many more degrees of freedom with respect to potential changes in interactions than in locations. Figure 6 shows the pattern of these shifts in location from which it is clear that the city tends to compact slightly with loss of population from many of the suburban areas all around the city with the exception of the relatively vibrant western corridor which attracts some population. This is probably due to the configuration of employment in the west, in and around Heathrow, and the relatively prosperous belt of commuters with good access to transport infrastructure in west and south west London. In Figure 7, by trial and error we show a non-contiguous partition of the system that leads to a high flow of population across these boundaries. The user can choose any partition by clicking on

the zones at the fine or broad scale – wards or boroughs and in terms of the 633 wards there are 633! possible combinations to consider. Clearly some intuition about the workings of the model and the structure of the spatial system is required to use this tool.

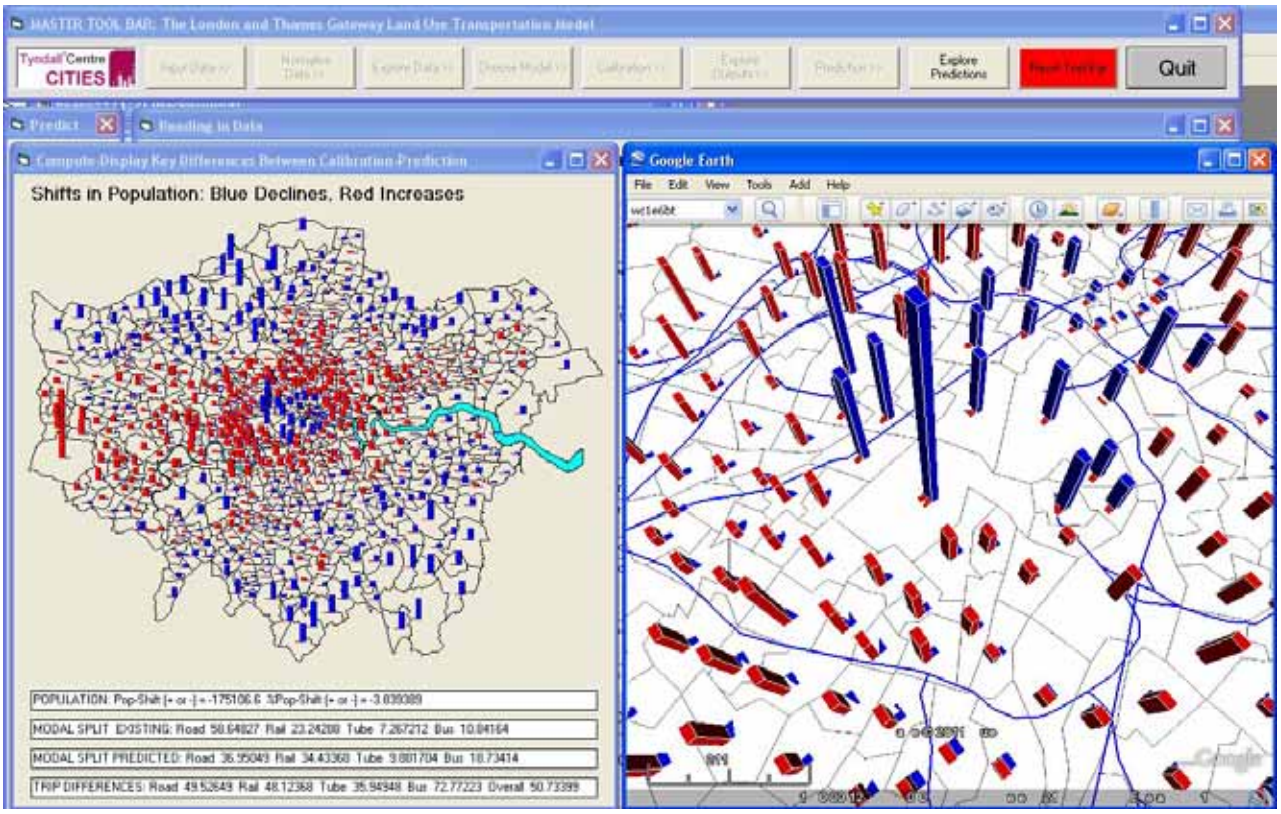

Figure 6: The Impact of a Doubling of Road Travel Costs for Private car on Location
(red is increase in population, blue decrease)

The energy equations that we introduced earlier also provide a useful if somewhat oblique perspective on these results. The units in which energy and entropy are measured bear no resemblance to the units in which the data for the model is input which is minutes of travel time. This is because entropy measures in equations (13) to (15) are computed in terms of total trips, not probabilities, as formulated in traditional versions of entropy-maximising derivations of these models (Wilson, 1970). Moreover as equation (13) makes clear, travel costs by mode are normalised by the appropriate travel parameter and then summed to produce a composite cost. However the relative weighting of these measures gives some sense in which the system changes from the first state (1) to the scenario state (2). We show all these values in Table 2 where the percent changes are only indicative. The critical issue is that the

total energy in cost terms massively increases due to the external imposition of the 100 percent change in road costs and this leads to an equivalent increase in free energy while the entropy increases only slightly. Clearly the travel costs by mode sum to the total costs and thus can be seen in the disaggregations. The actual travel costs are the normalised minutes travelled where the friction of distance parameter is applied to the total minutes travelled in the system; insofar as there is a standard unit used in this analysis, it is a kind of normalised trip. In fact we make no apologies for these tortuous interpretations which are compounded by the fact that our model has a complex disaggregation in which modes compete with one another. We have not explored the disaggregation of the entropy equation by different modes because this is distorted by the fact that the free energy equation cannot be so broken up as it contains the coupling mechanism that is needed to ensure the model acts as one. However this is an emerging and active area of research, all the more important because of our current concern for energy costs in terms of problems of resource depletion and climate change. Somewhat mysteriously the substantive interpretations of the energy in these entropy maximising models has remained dormant since their inception some 40 years ago, and only now is there any effort to ground these concepts in real measurements (Batty, 2010). Various interpretations of the values in Table 2 are possible and the reader is referred to the recent papers by Wilson (2009) and Morphet (2010).

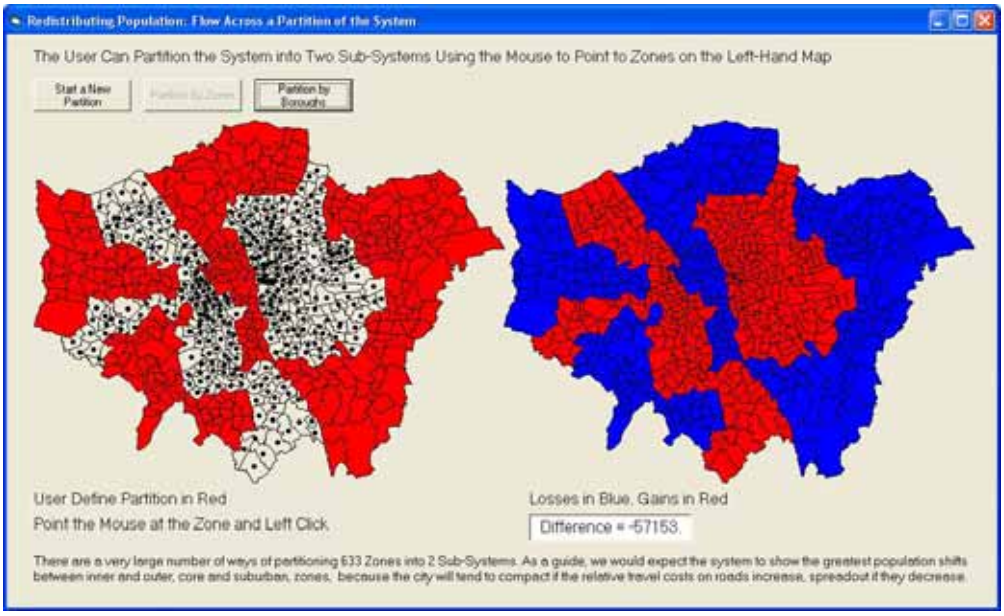

Figure 7: A Non-Contiguous Partition of the System Leading to Population Relocation

To complete the picture, we will examine the changes in accessibility that are occasioned by this 100 percent rise in cost of travel by road. In Figure 8, we show changes in accessibility based on computing the standard log sum term which is the first component of the entropy in equation (13) and we apply this to the accessibility of origins, meaning that this is accessibility to the location of employment. We can restate the change equation for each destination as follows:

$$\Delta V_i^k = \log \frac{\sum_j A_j \exp[-\lambda^k c_{ij}^k (2)]}{\sum_j A_j \exp[-\lambda^k c_{ij}^k (1)]} \quad , \qquad (21)$$

but it is clear that no changes are communicated from mode to mode through this form of accessibility: for three of the modes, the travel costs are the same for both the *before-(1)* and *after-(2)* states. However we can examine changes in the road accessibility and in Figure 8, we show the before, after and ratio of these two sets of accessibilities as computed from equation (21) with $k = 1$. These two accessibility surfaces are mapped in rank form, that is the highest accessibilities are the darkest colour (red) and the lowest the lightest blue. The surface tends to contract a little between the two states, that is, the surface tends to draw itself close to the centre but a better illustration is the ratio of the two surfaces as in Figure 8c. This shows that there is a loss of accessibility relatively in the south west and the congestion charge area becomes much more accessible largely because if costs are increased uniformly across the board, then the higher cost areas (central London) become more advantageous even through they still have the highest unit costs.

Next Steps in Model Development

We are currently extending the model in terms of the number of sectors modelled and the number of zones defining the size of the urban region. Our criterion for visually accessible, rapid operation of the model is still a major objective in its development but to achieve these changes in scale, we are developing a slimmed down desktop version of the model and a much faster web-based version built in state of the art software, in this case using the ECLIPSE Integrated Development Environment

(http://www.eclipse.org/). The expansion of the residential model to link with retail and local services location models mirrors developments of more integrated models elsewhere (Batty, 2009). These will include disaggregation by activity type as well as mode and will be interfaced with various capacity constraints on location and on the transport network. In this sense, the new model will include assignment of trips to the various networks and the assessment of related capacity constraints reflecting cost of transport.

Table 2: Changes in Entropy, Energy and Costs

Energy Value	Calibrated-Observed State (1)	Scenario State (2)	Percent Change
Entropy S	4550379	4551015	0.014
Free Energy F	9243178	50642320	448
Total Trip Costs $\Sigma_{ijk}\lambda^k C^k$	4692799	46091305	882
Road Costs $\lambda^1 C^1$	3657600	1654163	-55
Rail Costs $\lambda^2 C^2$	3142503	6351686	102
Tube Costs $\lambda^3 C^3$	1343076	10268510	665
Bus Costs $\lambda^4 C^4$	1100000	32367970	2843

a *b* *c*

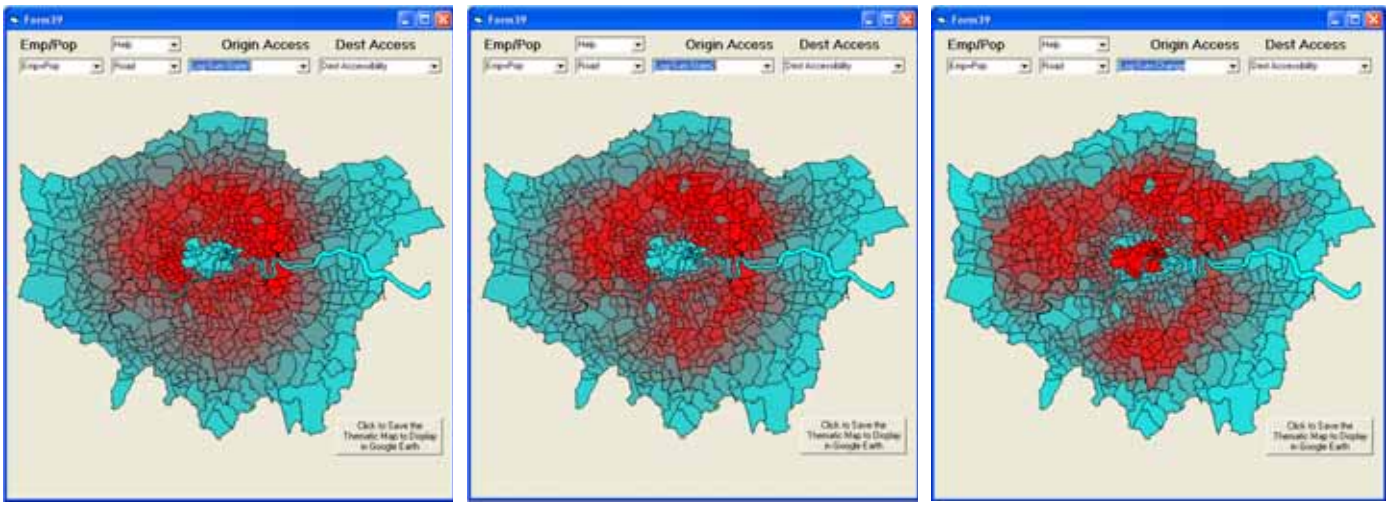

Figure 8: Before (a), After (b) Accessibility by Road and Their Ratio(c)

The biggest potential change to the structure presented in this paper is in terms of the form of the residential location model. Wegener (2008) argues that each sector

modelled in the urban system is likely to be subject to very different explanations in terms of economic dynamics. Retailing, he argues, is a process of rapid response to changes in demand and supply which can be seen largely in terms of travel costs and accessibility while residential location is based much more on the trade-off between house prices and travel costs which depend on wages. We have a version of the current model which replaces the constraints on travel costs with a budget equation that links incomes to costs through a new constraint. Equation (1) thus becomes

$$\sum_i \sum_j T_{ij}^k (w_i - c_{ij}^k - r_j)^2 = \sigma^k \qquad , \tag{22}$$

where w_i is the average wage earned at location i and r_j is the average house price (or rent suitably discounted to the appropriate time period) at location j. Equation (22) is in fact a variance that ensures that the majority of trips will cluster around the mean value based on the difference between wages earned and the costs of travel and housing. This assumes that the probability of where people will locate increases as they move closer towards exhausting their budget. The model that is generated using this constraint and the usual constraint on origins in equation (2) thus becomes

$$T_{ij}^k = \exp(-\lambda_i) A_j \exp[-\lambda^k (w_i - c_{ij}^k - r_j)] \qquad . \tag{23}$$

There are many variants of this model in that we can formulate the budget equation in different ways, making this mode specific if data is available and of course disaggregating the equation to deal with different employment groups and housing types. This is the form we are taking forward in our three sector enlarged model where local service employment and retailing are treated using specific and different models from that used for residential.

When we examine changes to travel costs in the current model, the shift in population locations is of an order of magnitude less than the shifts in modal split. Our casual

knowledge of urban systems and the way people react to such changing costs suggests that the order of the locational shift – some 2.5 percent of the working population – is too low. This shift is almost the first order change that might take place although the model makes no such distinctions. The new model in equation (23) which we have already experimented with a little gives much larger shifts as travel costs are directly compared to house prices. If travel costs increase by 100 percent on road journeys which was the scenario tested here, then using equation (23) the shift in population is of the order of 12 percent. This is directly due to the fact that as travel costs increase for road users, they have less to spend on housing and consequently seek cheaper houses which is a direct locational effect. In our extended model, this kind of structure will be central to such location and thus the impact of changing costs will be much more realistic. In fact in future applications of this and the extended model, we propose to explore many different but related scenarios so that we can examine sensitivity to changing travel costs as well as actual impacts. As the model can be run rapidly, hundreds of such scenarios can be generated and this raises the prospect of some means of disciplining such actions so that the user can explore future solution spaces easily and effectively.

Our last foray into future developments concerns the underpinning of this and similar models using the entropy-energy framework. This needs substantially more effort in making the proper connections and interpretations between city systems and the way energy flows through their spatial fabric. We need a clearer explanation of free entropy and a detailed analysis of the way energy flows in systems that are coupled such as through the modal choice model we are working with here. One of the problems in measuring such energies is that the entropy maximising framework is, for historical reasons, difficult to dimension and the roles of the travel cost parameters and partition (normalisation) functions need to be worked out consistently for such coupled systems. These are all issues that are under active development in the context of the extended model in which the impacts of changes in energy will continue to be a central motivation for their continued development.

References

Atkins P W 1994 **The 2nd Law: Energy Chaos and Form** (W H Freeman and Company New York)

Bankes S 1993 "Exploratory Modeling for Policy Analysis" **Operations Research 41** 435-449.

Batty M, 1974 "Spatial Entropy" **Geographical Analysis 6** 1-31.

Batty M, 1976 **Urban Modelling: Algorithms Calibrations Predictions** (Cambridge University Press, Cambridge UK).

Batty M, 1983 "Linear Urban Models" **Papers of the Regional Science Association 53** 5-25.

Batty M, 1992 "Urban Modeling in Computer-Graphic and Geographic Information System Environments" **Environment and Planning B 19** 663–685.

Batty M, 2009 "Urban Modeling" in R Kitchin, N Thrift, Editors, **International Encyclopaedia of Human Geography** (Volume 12 Elsevier Oxford) 51–58.

Batty M, 2010 "Space Scale and Scaling in Entropy-Maximising" **Geographical Analysis 42** 395–421.

Batty M, Hall P G, Starkie D N M, 1974 "The Impact of Fares-Free Public Transport upon Urban Land Use and Activity Patterns" **Transportation Research Forum 15** 347–353.

Brail R K, Editor, 2008 **Planning Support Systems for Cities and Regions** (Lincoln Institute of Land Policy Cambridge MA).

Brewer G D, 1973 **Politicians Bureaucrats and the Consultant: A Critique of Urban Problem Solving** (Basic Books New York).

Dawson R, Hall J., Barr S, Batty M, Bristow A, Carney S, Dagoumas A, Evans S, Ford A, Harwatt H, Kohler J, Tight M, Walsh C, Zanni A, 2009 "A Blueprint for the Integrated Assessment of Climate Change in Cities" in K Tang, Editor, **Green Citynomics: The Urban War Against Climate Change** (Greenleaf Publishing Chippenham UK) 32-51.

Echenique M H, 2004 "Econometric Models of Land Use and Transportation" in D A Hensher, K J Button, K E Haynes, P R Stopher, Editors, **Handbook of Transport Geography and Spatial Systems** (Pergamon Amsterdam) 185-202.

Epstein J M, 2008 "Why Model?" **Journal of Artificial Societies and Social Simulation 11** no. 4 12 available at http://jasss.soc.surrey.ac.uk/11/4/12.html

Hall J W, Dawson R J, Walsh C L, Barker T, Barr S L, Batty M, Bristow A L, Burton A, Carney S, Dagoumas A, Evans S, Ford A C, Glenis V, Goodess C G, Harpham C,

Harwatt H, Kilsby C, Köhler J, Jones P, Manning L, McCarthy M, Sanderson M, Tight M R, Timms P M, Zanni A M 2009 **Engineering Cities: How Can Cities Grow Whilst Reducing Emissions and Vulnerability?** (Department of Civil and Geomatic Engineering Newcastle University Newcastle-upon-Tyne UK).

Harris B, 1970 "Change and Equilibrium in the Urban System" **Highway Research Record No 309** (Highway Research Board Washington DC) 24-33.

Hunt J D, Miller E J, Kriger D S, 2005 "Current Operational Urban Land-Use-Transport Modelling Frameworks: A Review" **Transport Reviews 25** 329-76.

Iacono I. Levinson D. and El-Geneidy A. 2008 Models of Transportation and Land Use Change: A Guide to the Territory **Journal of Planning Literature 22** 323-340.

Lee D B, 1973 "Requiem for Large-scale Models" **Journal of the American Institute of Planners 39** 163-178.

Lowry I S, 1964 **A Model of Metropolis** (Memorandum RM-4035-RC The Rand Corporation Santa Monica CA).

Morphet R 2010 "Thermodynamic Potentials and Phase Change for Transport Systems" Working Paper 156 Centre for Advanced Spatial Analysis UCL London, available at http://www.casa.ucl.ac.uk/working_papers/paper156.pdf

Timmermans H, 2006 "The Saga of Integrated Land Use and Transport Modelling: How Many More Dreams Before We Wake Up?" In K W Axhausen, Editor, **Moving Through Nets : The Physical and Social Dimensions of Travel** (Elsevier Oxford UK) 219-248.

Wegener M, 2008 "Multi-Scale Spatial Models: Linking Macro to Micro" paper presented to the Centre for Advanced Spatial Analysis University College London, available at http://www.spiekermann-wegener.de/pro/pdf/MW_CASA_090108.pdf

Wegener M, Gnad F, Vannahme M, 1986 "The Time Scale of Urban Change" in B Hutchinson, M Batty, Editors, **Advances in Urban Systems Modelling** (North Holland Amsterdam) 175-197.

Wilson A G, 1970 **Entropy in Urban and Regional Modelling** (Pion Press London).

Wilson A G, 2008 "Boltzmann Lotka and Volterra and Spatial Structural Evolution: An Integrated Methodology for Some Dynamical Systems" **Journal of the Royal Society Interface 5** 865-871.

Wilson A G, 2009 "The 'Thermodynamics' of the City: Evolution and Complexity Science in Urban Modelling" in A Reggiani, P Nijkamp, Editors, **Complexity and Spatial Networks** (Springer Berlin) 11-31.