



WORKING PAPER SERIES

http://www.casa.ucl.ac.uk/working_papers/paper79.pdf

© Copyright CASA, UCL

ISSN: 1467-1298



Visualization in Spatial Modeling

Michael Batty
Philip Steadman
Yichun Xie



www.casa.ucl.ac.uk

T +44 (0) 20 7679 1782 • T +44 (0) 20 7679 1813 • F +44 (0) 20 7813 2843 • E casa@ucl.ac.uk

Centre for Advanced Spatial Analysis • University College London • 1 - 19 Torrington Place • Gower St • London • WC1E 7HB

Visualization in Spatial Modeling

Michael Batty¹, Philip Steadman², and Yichun Xie³

March 27, 2004

Abstract

This paper deals with issues arising from a central theme in contemporary computer modeling – visualization. We first tie visualization to varieties of modeling along the continuum from iconic to symbolic and then focus on the notion that our models are so intrinsically complex that there are many different types of visualization that might be developed in their understanding and implementation. This focuses the debate on the very way of ‘doing science’ in that patterns and processes of any complexity can be better understood through visualizing the data, the simulations, and the outcomes that such models generate. As we have grown more sensitive to the problem of complexity in all systems, we are more aware that the twin goals of parsimony and verifiability which have dominated scientific theory since the ‘Enlightenment’ are up for grabs: good theories and models must ‘look right’ despite what our statistics and causal logics tell us.

Visualization is the cutting edge of this new way of thinking about science but its styles vary enormously with context. Here we define three varieties: visualization of complicated systems to make things simple or at least explicable which is the role of pedagogy; visualization to explore unanticipated outcomes and to refine processes that interact in unanticipated ways; and visualization to enable end users with no prior understanding of the science but a deep understating of the problem to engage in using models for prediction, prescription, and control. We illustrate these themes with a model of an agricultural market which is the basis of modern urban economics – the von Thunen model of land rent and density; a model of urban development based on interacting spatial and temporal processes of land development – the *DUEM* model; and a pedestrian model of human movement at the fine scale where control of such movements to meet standards of public safety is intrinsically part of the model about which the controllers know intimately.

Acknowledgments

This paper was first presented at the ESLab International Workshop on *The Study of Artificial Complex Environments*, 2-4 April 2004, San Servolo, Venice, Italy.

¹ CASA, University College London, 1-19 Torrington Place, WC1E 6BT, UK
m.batty@ucl.ac.uk <http://www.casa.ucl.ac.uk/>

² Bartlett Graduate School, University College London, 1-19 Torrington Place, WC1E 6BT, UK
j.p.steadman@ucl.ac.uk <http://www.vermeer.co.uk/>

³ Institute for Geospatial Research & Education, Eastern Michigan University, 125 King Hall,
Ypsilanti, Michigan 48197, USA
yxie@emich.edu <http://ceita.emich.edu/>

1 Defining Visualization

Visualization is a term that gained widespread currency in the mid-1980s when for the first time, computer graphics was linked to supercomputer processing, particularly in scientific contexts such as astrophysics where it was essential for the results of such processes to be absorbed and understood visually (Kaufmann and Smarr, 1993). This was explicit recognition that many large data sets, whether produced for input to computer models or as outputs or simply as raw data captured by digital instruments, needed to be understood holistically using the synthetic properties of the mind and eye in unaided form. In parallel, our interaction with computers was becoming more visual and although this began prior to the micro-revolution, it was spurred on by the immediacy of interaction which the PC enabled. The ultimate outcome is that we now interact with computers almost exclusively using graphical user interfaces (GUIs) and this, in itself, has broadened the concept of visualization to most aspects of human-computer interaction.

Visualization is so broad a term that to define its role in spatial modeling, we first need to stand back a little and examine the kinds of models that the visualizations that we present here pertain to. Classifications of models go back to the 1950s and 1960s when the term became popular and it is instructive to note, for example in papers such as Ira Lowry's (1965) "A Short Course in Model Design", that the starting point was defined as continuum of models from iconic to symbolic embracing analog along the way. Iconic models are physical versions of the real thing, usually scaled down to toy-like proportions such as architects' block models, while analog represent the system of interest using another but different system which may be either physical or digital. Good examples represent the movement of pedestrians in streets using analogies with hydrodynamic flow theory. Symbolic models are those which replace the physical or material system by some logical-mathematical structure, usually algebraic, with computer, hence digital, representation and manipulation a central feature of such simulation. Models in all these senses are of course simplifications where key features of their system relevant to their users are emphasized, often to the exclusion of many other features.

These issues are being rapidly influenced by a sea change which is occurring in how we view models in science and social science. Fifty years ago, modeling was parallel to classical science in that it was based on implementing good theory in models in the most parsimonious way possible. Good theories and models were those that could explain the data in the simplest, most efficient way, notwithstanding the fact that there were often critical issues which most would agree form part of the system being modeled, left out. This is now changing in every discipline and domain. The cutting edge of theory and modeling in the spatial-geographical domain, particularly where this involves human systems, is embracing ever more diverse and richer model structures. These structures are never likely to be validated in their entirety against data, they are too rich, and the data required for their testing too poor for any complete assessment. Many of the models that we introduce here follow this tradition in that there may only be a few points at which their data and processes touch the real world in terms of the data available. Science is now much more comfortable with this notion of a theory or model which is only partly testable than it has been hitherto. There is increasing recognition that our systems of interest are intrinsically complex and must be handled rather differently from those on which classical science has been founded.

A further twist to the visualization paradigm involves the way the model is represented. As digital computation has become all pervasive, symbolic models no longer represent the sole focus. Iconic and analog models are also increasingly digital with the key issue being the mix of symbolism, analogy and iconic representation that can be employed for a single system where the simulation involves passing between any of these styles of modeling. The best examples involve relatively real renditions of spatial systems based on digital modeling of the physical appearance of the objects of interest. In this case, the appearance may only be used as the visual container in which analysis takes place. Such is the case in 3-d GIS (geographic information systems) where the 3-d container is a digital version of the physical fabric. Increasingly 3-d is being used to display patterns which can then be projected back onto the digital iconic model, or even onto a physical model of the system itself as in the *Tangible Media* projects at MIT (<http://tangible.media.mit.edu/>). Moreover once a digital iconic model is developed, it can be aggregated into various forms, put into other digital contexts in semi-recursive fashion, and even used to manufacture actual

physical models as hard copy print versions in traditional wood or plastic. An early version of such a mix of media is embodied in the hypothesis that Vermeer used a primitive camera to generate ‘some’ of his paintings (see Steadman’s (2001) book *Vermeer’s Camera* at <http://www.vermeerscamera.co.uk/>). Animating such a mix and using this digitally as part of the story line is what Tufte (1997) calls a ‘visual confection’.

In this paper, we will concentrate on what has come to be called ‘scientific visualization’ whose main focus is on the inputs, processes, and outputs associated with symbolic or mathematical models, in this case urban and spatial systems focused on the human-built environment. In the next section, we will introduce a generic modeling process and show how this can be linked to planning, management and action. It is this nexus of explanation, simulation, forecasting, design and control which provides the wider context for visualization and we will thus identify the key types that map onto this spectrum of possibility. We will then develop a generic form for visualization in the spatial modeling field, which tie these various possibilities together in what we call the ‘space of visualization’. This sets the scene for three distinct demonstrations: the first is pedagogic and focuses on an explanation of a well-known theory of land use due to von Thunen, the second enables a model of dynamic (temporal) urban development to be explored, and the third shows how important it is to develop models with strong visual content which enable designers and decision-makers to use models to generate effective designs and policies.

2 Defining Spatial Modeling

Computer models are structured in many different ways but the standard sequence follows digital processing which involves manipulating a series of inputs which drive the model to a series of outputs, thus reflecting the various functions that tie the elements of the model together. This sequence reflects the logic chain that any model is built around, with inputs defining the exogenous or independent variables that dictate how endogenous or dependent variables are conditioned. Many models involve elaborate causal chains which may be activated many times, recursively

through different kinds of time which in turn enable the model's outputs to become stable. This *model processing* is usually nested within a wider process of *model fitting*, estimation, or calibration which enables parameters of the model – macro variables that usually have global significance – to be tuned to values that connect the inputs to the outputs in the most satisfactory way. This kind of prediction is enabled so that confidence in the way the model reproduces a known situation as reflected in the input data and the independently corroborated data associated with the models outputs or predictions, is assured. The third process is *model use*, in conditional prediction which often mirrors calibration, but is part of a wider phase in which the model can be used in design, management, and control.

Each of these processes can be represented using different forms of visualization. Indeed every aspect of model operation and use can be visualized as the ultimate structure of the model is digital whose location in computer memory can be mapped in some way to the 2-dimensional screen. As all our models are spatial, hence some of their inputs and outputs are mappable, then associations between inputs and outputs with respect to map pattern provide an obvious form. There are many ways in which such inputs and outputs can be linked - offline or online in terms of showing how inputs are converted into outputs and a classic strategy of visual comparison is to array these maps as separate and comparable layers, as 'small multiples' in the manner suggested by Tufte (1990). 3-dimensional forms can be widely exploited too as in the standard manner where such representations are portrayed using the three Euclidean dimensions. But at the level of abstraction used in this style of modeling, the third dimension is more likely to be employed for scientific visualization of the phase space of model solution rather than for more literal, or iconic visualizations associated with built form or rural landscapes.

These visualizations almost assume that what is being modeled is static in structure where outputs occur at a single point in time but many spatial models are dynamic and thus sequences of inputs and outputs need to be visualized. This is the space-time process and although small multiples are useful, animation in 2-d or 3-d is often employed. We will return to this when we deal with calibration below but animation also constitutes a way of linking inputs to outputs, thus revealing model functions or processes. There are also different dynamics from the routine where the focus is on

showing how objects move in space to longer-term migrations where the focus is in comparing map patterns at different points in time. Movement is most simply conceived of in terms of animation but there are a variety of ways in which such animations can be linked to more abstract properties of the map patterns and the space-time movement of related objects.

A time-honored strategy for spatial visualization is to use multiple windows with different phenomena in each, some spatial in 2-d or 3-d but some aspatial or non-spatial and to hot-link these windows so that change in each can be related. The process of linking inputs to outputs involves the functioning of the causal chains which form the core of the model. If the model is dynamic, visualization may be built around space-time in the literal sense but if the process is recursive for the model to converge on stable outputs, this too might be visualized in much the same way. In cases where the model is both temporally dynamic and recursive in terms of its path to solution, then a combination of both is possible.

Spatial models are often built around aspatial or non-spatial processes which although touching the spatial system at some point, can be represented using visual traditions very different from the 2-d map or 3-d surface. For example, the spatial economic model that we first introduce below is conceived in terms of demand and supply curves and only then mapped onto a simplified spatial landscape. In fact, one of the great powers of scientific visualization is to make such links between non-spatial, aspatial and spatial representations, as much for pedagogic purposes as for use in more practical contexts. However such mapping from one visual media to another in terms of representation always needs to be determined before visualization takes place. All this means, is that visualization is a creative process. It is only as good as our imagining of how different elements of a model relate to one another and to the wider context in which they sit.

The second environment in which spatial models are formed involves the process of *model fitting*. This connects up directly to searching for pattern in data but in particular for pattern in the input data which is exploited by tuning the model's processes and functions to explain as much of this pattern as possible. This is usually accomplished using a process of trial and error fitting with successive improvement to

the best fit. To state the long standing analogy, the process might be visualized as climbing a hill where the surface terrain represents the different performance of the model with respect to different input parameters. The process of model fitting is then the process of climbing this hill and reaching some global optima, ensuring that the process does not get stuck in some local optima, some hillock in the landscape of hills. In many traditional models which are parsimonious in the extreme such as those embodying spatial interaction (Batty, 1976), the structure is such that a unimodal performance surface can be ensured if the model is formulated with mathematical correctness. Standard procedures can then be used to reach the global optima. In fact, visualizations of this process showing the climb across the terrain have been used quite widely since the 1960s.

Most of our models are much more complex in that the phase space which embodies this terrain can no longer be mapped due to the very large number of parameters that contemporary spatial models contain. This is especially true of simulation models incorporating new notions of cells and agents. There the process is often much more partial in that the calibration might be visualized using exploratory procedures which do not aim to find any global optima. In such cases, there still needs to be structure to the process. In fact, the calibration phase of spatial modeling is pushed one stage back in these more exploratory models due to the fact that the very formulation of the model itself comes under scrutiny as soon as spatial data begins to be explored. Exploratory spatial data analysis (ESDA) which became popular after the first wave of scientific visualization had been established in the early 1990s, threw up the notion that the model should emerge naturally from an exploration of pattern in data. Although there have been many demonstrations, the focus has been more on inductive generalization than on the development of explicitly deductive models which arise from such analysis. Pattern and error in the data is a key consequence of such ESDA but there have been very few examples where such analysis has then proceeded to build models round these patterns, other than those which are based on inductive statistics. Although visualization may help generate new models, this is likely to be over a much longer time span than the modeling process itself and although the notion that the model emerges naturally from such exploration is an attractive one, almost Eureka-like in its impact, this is unlikely to be the case in most modeling efforts. Usually the model is proposed in advance and all the focus is on tuning it to a real

situation, understanding that situation, and using the model predictively and prescriptively to solve some problem or implement some new design.

The third environment in which spatial models sit involve their practical use in prediction and policy. In the simplest sense, similar visualizations might be built around prediction and prescription as around data exploration and model calibration. Early software embodying GIS functionality in standard cross-sectional urban models illustrates this principle where standard sets of visualization functions apply to any of the four model stages – data assembly and checking, analysis, prediction, prescription (Batty, 1992). However this process raises the question of purpose and engagement in terms of what the model for, who is it for, and how is it to be used? In short, models are rarely for the indulgence of the model-builder or scientist, more likely for the persons who commissioned it in the first place for practical use. In our examples below, we present three models; the first is for educating ourselves, while although the second is for exploration of urban development processes, it is conceived as being applicable to practical problems and policies involving urban sprawl. The third is quite definitely for the stakeholders involved in solving problems in the local environment.

In these applied and policy contexts, visualization is likely to be a little different from the kinds of scientific visualization we have been implying so far. In fact it is likely to reflect a much looser interpretation of model inputs, processes and outputs and may be linked to media that do not form part of the model in the first place. This is no less rigorous but it does change the kind of engagement that modelers and stakeholders have in the process. Visualization thus becomes an essential part of communicating complex ideas to a non-expert clientele, and in this sense, it probably involves developing procedures for involving this clientele at different stages of the modeling process. The notion that the model is delivered, scientists explore and tune it using visualization, and its outputs are then pictured in conventional scientific form, is not necessarily the most appropriate procedure in situations where stakeholders are involved in using models directly. We will present such applications as part of our third example below.

3 The Space of Visualization

Visualization as a style and set of activities is almost impossible to classify for every aspect of computer modeling and its application can be subject to representation using digital pictures. Nevertheless it is useful to begin to organize the field with respect to the models we will illustrate here, if only to show how the kinds of visualization employed depend intimately on purpose, the system of interest, and the environment in which the application exists. We will define a generic space which is organized on two levels: first with respect to the purposes for which the model, thence its visualization is designed, and second, in terms of the key techniques used to implement its visualization. We could have developed a third level based on different media but in all the cases we conceive of here, the media are conventional pencil and paper and their digital equivalents. The panoply of VR and tangible media have not yet been invoked in any of the models that we present here although this is an important direction in which these more abstract models should be developed.

We define four distinct purposes: education, exploration, explanation, and engagement. These purposes are not mutually exclusive of each other, nor are they arrayed orthogonally; more likely a model and its visualization tend to stress these four purposes in different ways, often with one purpose dominating. For example, visualization for education can be both narrow and wide although in the sense used here, we will be taking a narrower view. Of course all model building and applications involve education of ourselves and of our clients but in this context, we are specifically thinking of visualization for the prime purpose of getting the message over of how a model actually works. In this sense then, we see visualization as enabling an understanding which would not be possible without pictorial help. This kind of visualization makes the operation and meaning of the model much clearer than any other form of communication.

Exploration is more geared to investigating how model structures translate inputs into outputs. This is an essential quest in learning about how the model works. The more complicated the simulation, the more likely that exploration is required to test the limits of the model, and to enable researchers to be sensitized to the impacts of their

scientific decisions. All modeling involves some exploration but in the development of models separate from immediate practice, then exploration is of the essence, especially where causal structures cannot be analyzed using mathematical formalisms. Exploration might simply be trial and error based on trying to find out how the model behaves, or it may be more systematic as, for example, in the process of calibration.

In contrast, explanation involves using visualization to confirm or falsify some theory which is embodied within the model, and the usual processes of comparing pattern in the input and output data is central to this. A tricky issue however is to visualize how the various processes linking inputs to outputs match what we know about the operation of such processes in real life. Often visualization as explanation is rather partial, being based solely on a comparison of outputs from the model with those that observed in the real world. Every model which is built afresh, requires some sense of how well it explains the reality to which it is being applied, although this is more likely to be to the fore in applications which are removed from practice.

The last purposive activity we define is engagement. Rather than define purposes which involve forecasting for policy-making, forecasting to test design impacts, management or control, we prefer to simply note that models which are developed for purposes other than science *per se*, involve the engagement of non-modeling experts. In fact, as models are often built by large teams whose expertise differs markedly between team members, then it might be supposed that visualization might be used to engage the team in assembling the best model. This is indeed the case but here the contrast is greater between scientists and non-scientist stakeholders who need to be involved in the process in rather different ways. Essentially engaging stakeholders and non-scientific experts involves different kinds of visualization and dissemination which probably requires more non-scientific information to be assembled and related to the model and its application. Furthermore, the process of using visualization tools becomes significant when diverse groups are involved in this kind of communication.

The second level of visualization we define is based on a limited number of techniques. Much ingenious visualization is one-off and cannot be classified generically. This is because visualization tends to involve some insight which is produced idiosyncratically and then pictured in some meaningful way. There are no

formulas for creating such graphics although there are some simple and obvious methods for taking spatial media and representing it visually. We define the 2-d map and the 3-d icon – surface, iconic physical shape, and so on – as being key elements of the way we visualize map pattern. Spatial process is harder to fashion but the notion of a process occurring in space and time can be illustrated using small multiples (of pictures) which provide a sense of change in space and time; animation is often simply arranging these multiples as frames in sequence. We also invoke the parameter space as being a vehicle which controls the operation of the model specifically through calibration. These five features can be arrayed against one another and combined in diverse ways to give real substance to the idea of modeling as visual confection. The use of hot-linking through multiple windows is simply one of the ways in which such visualizations can be animated and linked.

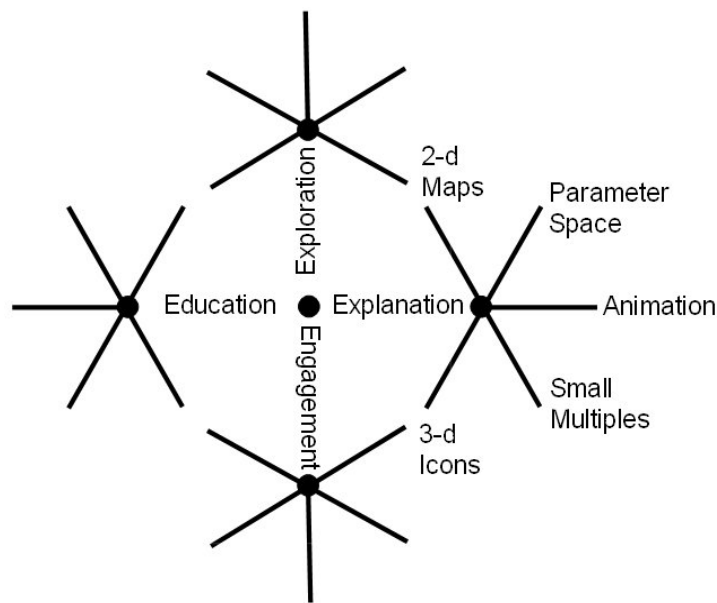


Figure 1: Elements of Visualization: Purpose and Technique

These two levels of visualization by purpose and technique lead to a tree of possibilities that we have arrayed in Figure 1. This provides a simple means of classifying the types of visualization in any spatial model application and we will use it to illustrate the relative differences between the three applications which follow. In any full scale application, many branches of this tree will be invoked: models are defined to educate, explore and explain with these activities often taking place, even though the main purpose is to engage those stakeholders who control the policy and

design process. Our first example below clearly illustrates this point although the three applications we have chosen are very different from one another.

4 Visualizing Theoretical Explanations: The von Thunen Model of Market Land Rent and Transport

Count Johann Heinrich von Thunen is accredited with devising the first land use model in 1826 which explained the spatial distribution of crop cultivation (and/or animal husbandry), using evidence from his farm in East Prussia (Hall, 1996). His model essentially examines the trade-off between the productivity of agricultural land with respect to the yield and price associated with selling the crop at the market and the transportation cost necessary to move it there. In essence, this tradeoff fixes the level of rent that the farmer is able to pay and the use of the land is thus determined by the crop that generates the highest rent. The model involves comparing several crops and determining, with respect to one market center, the actual crop cultivation on each land parcel at different distances from the market. If there is more than one market center, the allocation depends on a comparison of possibilities but the solution is still stable for the selection of the land use simply depends on the maximum rent payable from whatever market.

For a single market with uniform transport costs which imply location on a homogeneous plain, land uses arrange themselves concentrically around the center. This can be easily extended to several centers. If physical distortions due to the transport network are introduced, then the pattern of location is affected, with fast transport routes having higher accessibility than the areas between them. This model is simple in form but the spatial outcomes from its generalization to many centers and to many different transport routes are often difficult to anticipate, hence the need for some simple demonstration. Moreover, when the price of the good or its composite yield, and the relative transport costs are changed, the pattern of land uses shifts. This is the essence of the model for which we have produced a very simple but effective visualization. Of course, the importance of the model does not lie in its application to an agricultural or rural system but in its generalization to the urban realm where it

forms the basic equilibrium model of the urban land market, underpinning contemporary urban economics (Alonso, 1964).

To illustrate how tricky the model is to non-mathematicians, we first define a location with respect to a single market center as j , the distance from the center to j as d_j , and the cost of transport for the commodity in question, k , as β^k . The quantity produced per unit of land or yield is assumed to be uniform – on a homogeneous plain, and is defined as Q^k with its price at the market as ρ^k . We will also assume a fixed cost of production for each commodity called c^k . The rent which a farmer cultivating a crop k at j is called R_j^k and is calculated as

$$R_j^k = Q^k [\rho^k - c^k] - \beta^k d_j \quad . \quad (1)$$

This is a linear equation, sometimes called a bid rent curve, in that it shows what a farmer cultivating crop k can bid for renting the land at different distances j from the market center. In the model, we have built, we assume that one element of the yield, price and fixed cost in equation (1) can be varied and thus this equation can be written as

$$R_j^k = \alpha^k - \beta^k d_j \quad . \quad (2)$$

For a single market center, the use of each parcel of land j is computed as the maximum rent payable over all crops, that is $R_j^l = \max_k R_j^k$. For any pair of land uses/crops, it is also easy to compute points in the landscape where the rent payable is the same, that is the breakpoints between crops with respect to their distance from the market. This occurs where $R_j^k = R_j^l$ and this equation can be solved for any k and l to yield the break point $d_j(k:l)$ as

$$d_j(k:l) = \frac{\alpha^k - \alpha^l}{\beta^k - \beta^l} \quad . \quad (3)$$

Finally where there are two or more market centers called i , the land rents need to be computed using a modified form of the bid rent equations which are now indexed with respect to each center as

$$R_{ij}^k = Q^k [\rho^k - c^k] - \beta^k d_{ij} \quad . \quad (4)$$

We choose the land use for each location j which maximizes the rent as

$$R_j^l = \max_{ki} R_{ij}^k .$$

The visualization necessary in this model is largely so that the user can understand how the land use equilibrium occurs. The software essentially enables the user to show how the bid rent curves for three land uses – milk, grain and livestock production – can vary in price and transport cost, thus changing the intercepts and slopes of the linear bid rent function, α^k and β^k . The guts of the visualization are a blank canvas – the map – onto which one can draw transport routes and locate market centers. It is also possible to define constraints as unproductive land. The canvas is initially an homogeneous plain but background maps can be attached to it so that the user can draw on features that pertain to some real situation. The second type of canvas but within the same GUI, reflects the bid rent curves, one canvas for changing the slope of these, the other for the intercept. When these are changed, the distribution of land uses appears immediately and thus there is a direct association from the parameter space to the real space, with the parameter space in fact being constituted as a cross section through the hypothetical real space. The final feature of the visualization is the same canvas but with the distribution of land uses portrayed in 3-d as a wire frame or rendered image. There are two sliders to change the orientation of the x - y and z dimensions associated with this visualization.

This visualization is to educate the user through explanation. It was devised as part of the Open University First Level Course on Technology and Cities and the software was distributed free to students registered on this course (Roberts, 1999). It also

contains elements of exploration but it is strictly designed to explain and educate. In terms of our characterization in Figure 1, we show this form of visualization in Figure 2(a) (where we also compare this with the two other examples reproduced here) and this shows immediately that we are using minimal but effective visualization techniques. The essence of this is a small portable piece of software which embodies a kind of sketch explanation or sketch modeling. Only one window is used and there is no hot-linking but the control over the model is so quick and direct that this is an example of extremely parsimonious visualization which is pretty effective.

a) *Example 1: Explanation* b) *Example 2: Exploration* c) *Example 3: Engagement*

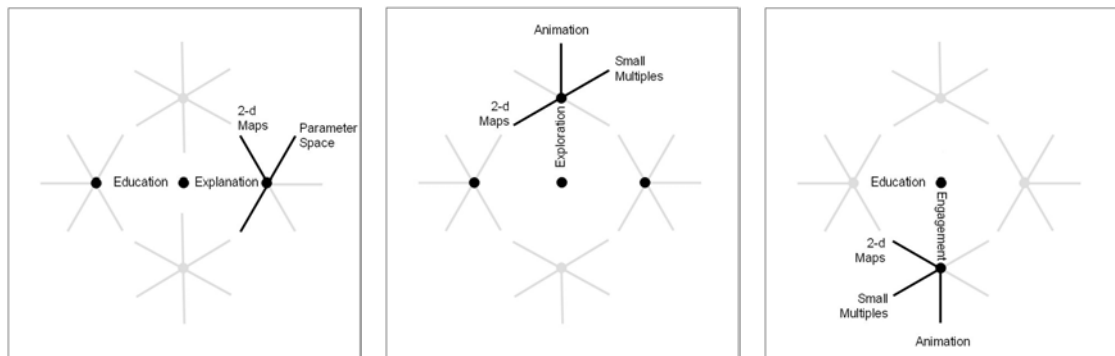


Figure 2: Structure of Visualization in the Three Example Models

We show two examples of the use of this software in Figures 3 and 4. In Figure 3, we show the single market center with no distortions associated transport routes. This is the homogenous plain example which appears everywhere in the location theory literature. The concentric symmetric ring structure around the market is clearly shown. In Figure 4, we have taken a map of Chicago and its railroads in 1861 from Cronon's (1991) magnificent book and use this to impress the fact that land use around Chicago is influenced by these routes. Note that we define Lake Michigan as 'unproductive land'. The resulting land use pattern shows the classic distortion posed by differential transport routes. The 3-d surface also shows the limits to the degree of distortion in that the picture is a little too confused. Nevertheless this does show how a model can be moved from theory to practice, from hypothesis to reality, albeit that the realities we choose are more caricatures of the real thing than the sorts of model

reality we present below. There are other bells and whistles that are invoked in this software – fuzzy boundaries and precise distances, scaling of bid rent curves to reflect actual distances, and so on – but the essence of the visualization is shown in Figures 3 and 4 which provide a complete picture of this approach to explaining location theory and its relationship to the micro-economy in a spatial setting.

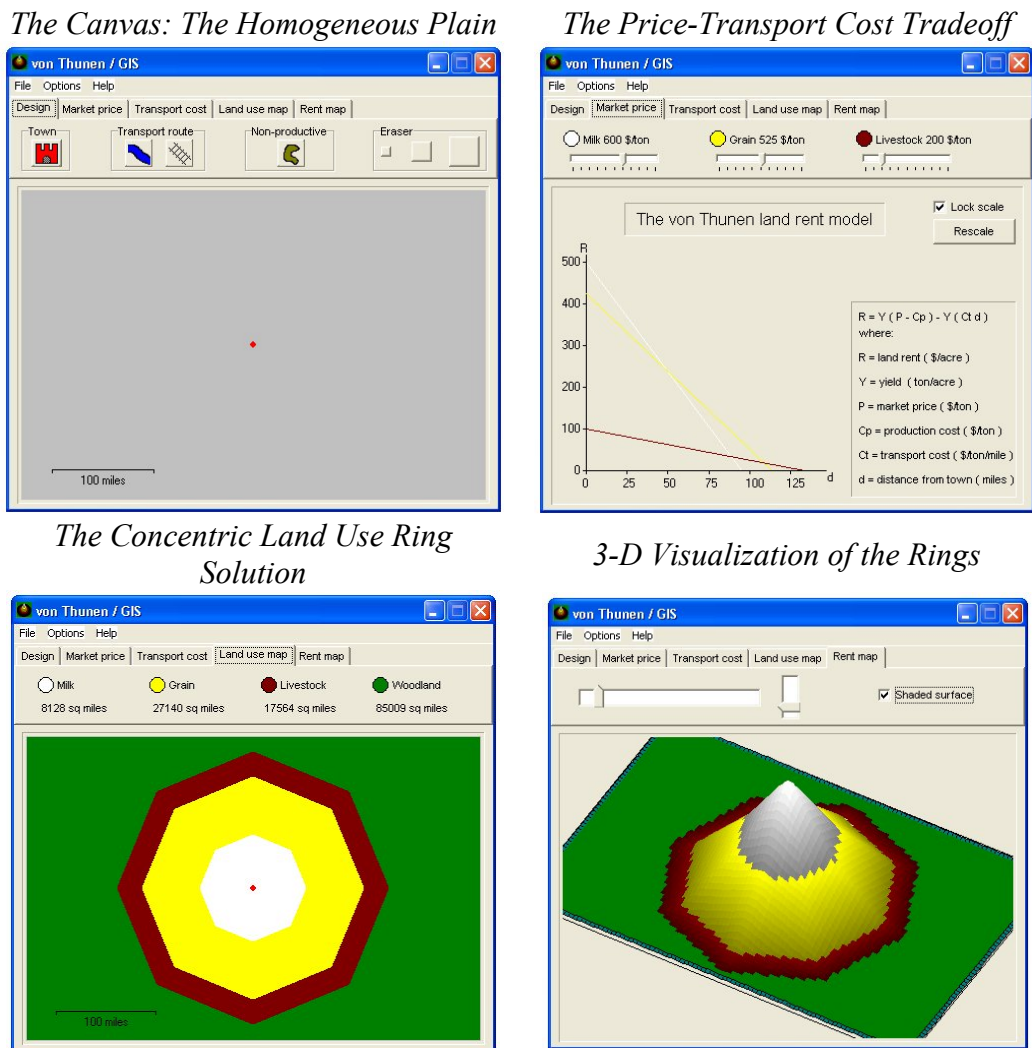


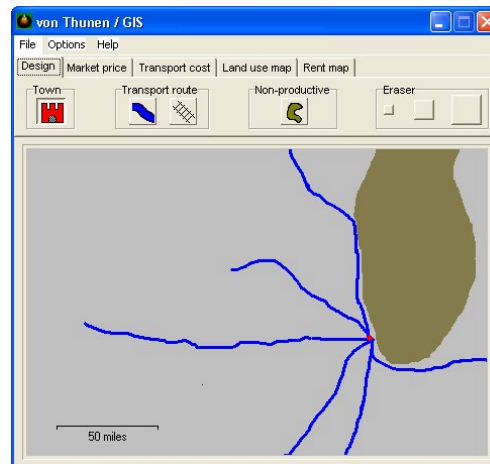
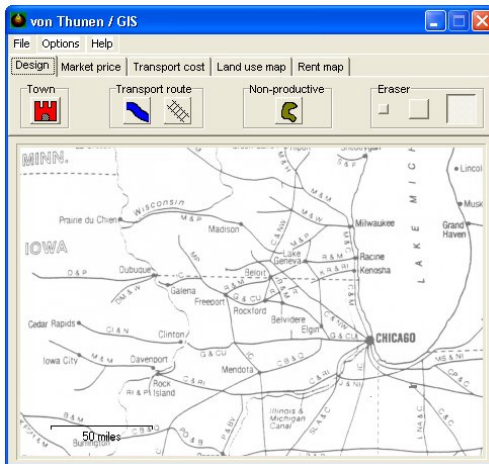
Figure 3: von Thunen's Model

Simple software demonstrates how the tradeoff between product yield and transport cost gives rise to land use competition and stable spatial organization.

You can download the software for this application from <http://www.casa.ucl.ac.uk/vonthunen/>

*The Canvas: Chicago 1861
(from Cronon, 1991)*

Railroads and Lake



Land Use Solution: Distorted Rings

3-D Visualization

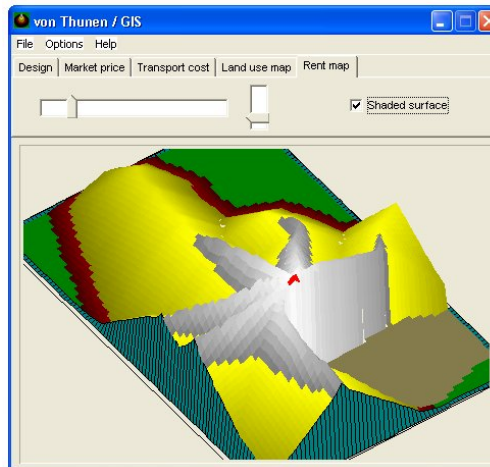


Figure 4: Nature's Metropolis: How Railroad Structure Distorts the Concentric Pattern in 19th C Chicago

5 Visualizing Model Exploration and Calibration: Urban Development Using Cellular Automata

Our second and third visualizations involve models whose outputs can be examined as they are executed. In a sense, this is true of the von Thunen model but as this is a comparative static structure, its operation is immediate. Models which can be examined as they run are usually temporally dynamic with the time simulation synchronized with simulation in computer time, notwithstanding any additional processing involving trial and error calibration during the simulation time itself. In the case of our first model of this genre – a dynamic model of urban land development using cellular automata principles – as the land area becomes more developed, more

and more functions are invoked to examine distance relationships and comparative links relating different spatially specific land uses to one another. Thus the computer time taken slows in linear proportion to this additional processing. This is not a problem in visualizing the structure and dynamics of the model but it gives the wrong impression in terms of real time. Thus movies must be made of the structure based on different frames at specific times if an accurate impression through simulation time is to be presented.

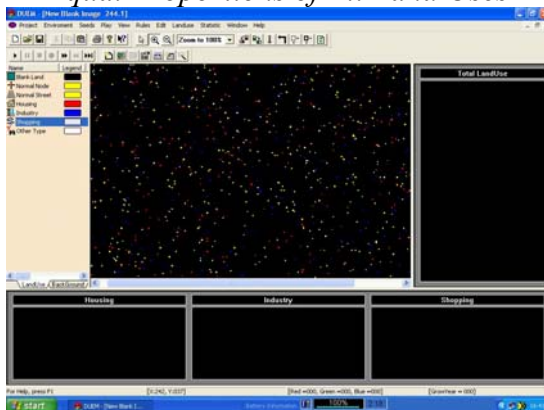
The model was developed by Xie (1996). It is called *DUEM* (Dynamic Urban Evolutionary Model) and is based on classic CA principles in which land is developed as a function of what other land uses exist within the neighborhood of each site being considered. All we can do here is sketch its rudiments for it is complicated, being part of a wider project which is aimed at putting the model into a web-based context and disaggregating the cellular spaces to enable agents to be specifically represented (Xie and Batty, 2004). The model contains five land uses – housing, industry, and services/commercial, as well as two kinds of street – junctions and segments. Junctions are needed to connect segments, and housing, industry and commerce cannot develop without there being streets nearby, within some neighborhood. Streets are a function of what gets developed in terms of these first three land uses.

Each land use is considered as being in three states, reflecting its aging: new, mature, and declining with new land uses being the seeds that motivate further growth. When land uses pass through their cycle to declining, they disappear and the land vacated becomes available for new development. Three scales of neighborhood are defined: first the small strict cellular neighborhood which is mainly used to ensure streets and land uses are connected, second a wider district neighborhood in which the distance from the cell in question at its center is considered with respect to other uses in the district, determining what use the central cell changes to, and third a regional neighborhood in which constraints on development are imposed. The model is thus richly constituted with respect to its life cycling and the representation of spatial scale.

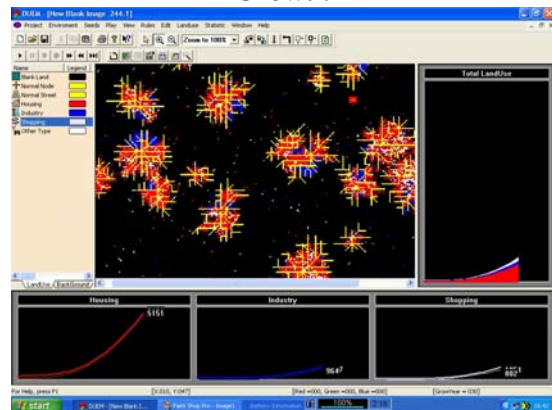
We have designed the interface in an entirely visual way based on two kinds of windows: the map canvas and four related windows which show the trajectories of growth and decline for housing, industry and commerce, and all three of these. The

user can switch elements on or off in each of these windows. The key reason for arranging the visualization of urban dynamics in this way is to present two features: the capacitated growth of the spatial system which is reached when the map canvas fills up, thus illustrating logistic growth and oscillations around the equilibrium, with even the possibility, yet to be seen, of simulating chaotic growth; and the wave effects seen when land uses age through their cycle. Waves of change cycle across the map, eventually dissipating as the system gets older. Because these kinds of model are highly diverse and contain many parameters, visualizations must be highly tuned to particular purposes. In this case, we have not produced any visualization of the parameter space for our focus is more on showing how different morphologies of development can result from very different initial and boundary conditions. Again one of our quests is to show how scale makes a difference.

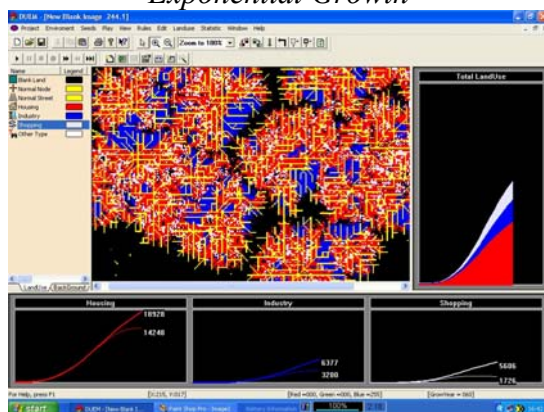
Initial Conditions: Random Seeds with Equal Proportions of All Land Uses



Growth to Time = 30: Early Exponential Growth



Growth the Time = 60: Exponential Growth



Growth to Time = 120: Capacitated Growth and Logistic Oscillations

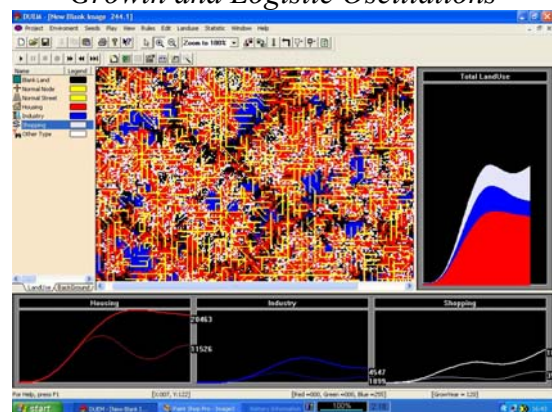
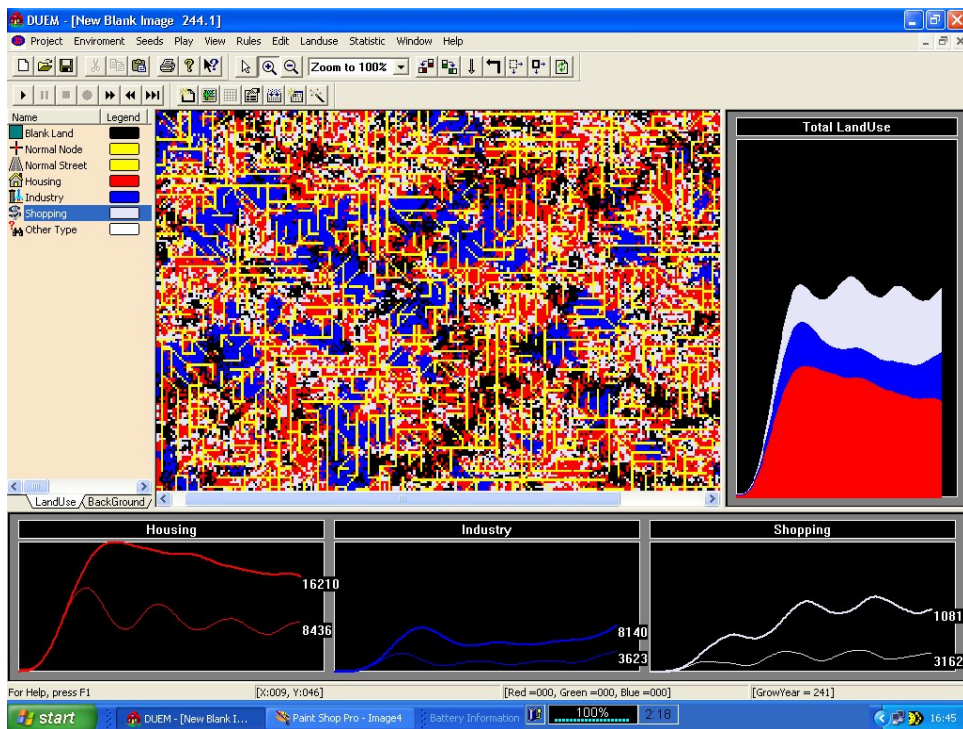


Figure 5: Small Multiples of Urban Growth Processes

Oscillating Growth at Time = 240



Exploring the Steady State to Time = 1500: Industry Takes Over Indicating that the Default Rules are Badly Specified

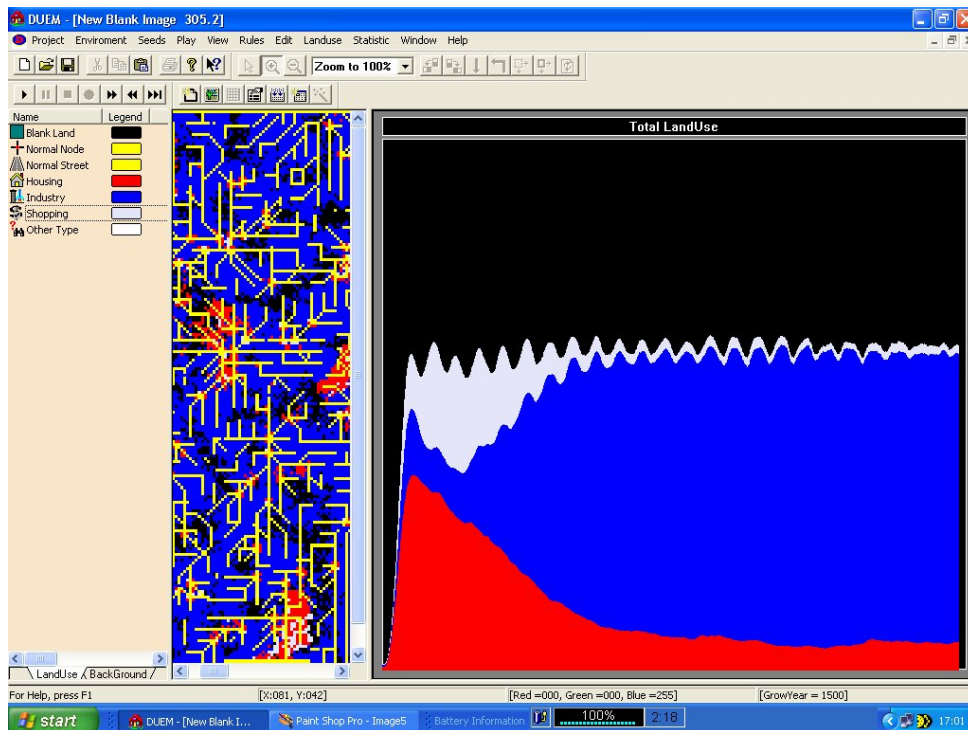


Figure 6: Exploring Trajectories to the Steady State

Behind these windows, access to the phase space within which the parameters are set is accessible using various dialogues, also called up as windows. Essentially by doubling clicking on each land use, we are able to bring up a series of dialogues with respect to how neighborhoods are configured, how distance influences the operation of state changes in each neighborhood, and how long different land uses remain in their different life cycle stages. We have not provided any means for visualizing how these changes in parameters might impact on the different morphologies produced. Offline, we can use any of the usual strategies to explore the parameter space using small multiples of maps and linking the parameter values in its phase space to these spatial differences. In fact, it is easy enough to make animations from the simulations which can be linked to different values in the parameter space but our focus here is more exploratory. The framework is so diverse in terms of scale that an obvious approach is to see how different kinds of initial conditions can be simulated. Comparisons from these thus become important, and again techniques such as small multiples and layering are clear ways to visualize these differences. Finally these kinds of models can be explored on-the-fly. As the model is running, we are able to explore changes to the morphology either directly or by stopping the model and changing parameters.

To show the power of this visualization which is charted in Figure 2(b), we present a typical run of the model in Figure 5. This shows the various windows and the development of a system of cities to the capacity of the map canvas as illustrated through the model's trajectories. The effect of different waves of growth in the residential sector can also be seen as distinct gaps in the development by the time the simulation reaches $t=120$ but these traveling waves are much clearer in the animation. In Figure 6, we show what happens in the steady state, how the land use totals oscillate but also how industry gradually encroaches on the other land uses, thus indicating that the default rules, no matter however plausible, are badly specified. This is the classic finding of our exploration: that models although plausible can be quite unrealistic when pushed to their limit. This, of course, is an essential diagnostic in developing a more realistic model with visualization being essential to this process.

There are many elaborations on this system that we might make, especially as the visualization occurs in computer time and the location of activity can be changed on-

the-fly. Within the package is a drawing capability, a little like that contained in the von Thunen model, and this enables the user to interact with the model in direct fashion. In fact, in all three packages we have used, there is drawing capability that lets the user interact with the model physically, notwithstanding that the data that drives these models is by and large numerical.

6 Visualizing Model Operation with Stakeholder Involvement: Pedestrian Movement and Public Safety

Our last example builds on the cellular dynamics approach of the last section but develops a model at a much finer spatial scale with a real problem driving the application. We are working with several varieties of active walker model which combine agent-based with cellular modeling using the cells to represent the landscape on which the agents – the pedestrians – move. The generic model essentially combines movement in the default direction of forward with some random perturbation and with obstacle avoidance, the direction being fixed according to a walker attraction surface which is formed from a synthesis of the multiple forces determining why people wish to walk and for what purpose (Batty, Desyllas, and Duxbury, 2003). In a shopping mall, for example, this surface would reflect the kinds of goods and their locations in shops that walkers wish to purchase.

We have applied this model to a situation of crowding which is associated with a major street festival, the Notting Hill Carnival, which is held once a year for 2 days in west central London. There are major problems of public safety associated with this event and the model has been built to show how people enter the area, flock to the attractions, namely the street bands and parade, visiting a series of events which are located in a small area of around 3 square kilometers. The model essentially walks visitors through the street system to the carnival attractions according to the existing controls on the event managed by the police and other services. The problem of crowding is severe in that there is a serious conflict between the parade and the street bands, the bands being inside the parade route which is circular and continuous. Visitors crossing into the area where the bands are located are in conflict with the parade and there are more general problems of crowding at different points within the

area where streets are narrowest. Problems of crime have grown as the carnival has gained in popularity and reducing crowd densities is seen as a way of making the event safer.

The critical focus of this model is that it is designed to help in alleviating crowding by showing how streets might be closed, barriers erected, sound (band) systems moved and the route of the parade changed. All these elements can be controlled within the model and posed as 'what if' questions. Most models can be put into a 'what if' context but in the case of this particular model, the situation is so highly controlled already that it is impossible to think of the model as simulating some relatively uncontrolled situation and then adding controls to meet some objectives. This kind of problem is quite unlike the problem of optimal city design where it is assumed that most cities develop organically from the bottom up and that planning control is imposed to direct growth rather specifically in situations where such direction is lacking or ineffective. This is not the case in something like the Notting Hill Carnival as throughout the history of the event, there has been strong control and management.

This suggests that those involved in managing the event and who know it best be intimately involved not only in the use of the model but in its design. Moreover, in a situation where there is high control, it is useful to think of model calibration as reflecting various degrees of control, for example, by beginning with a relatively uncontrolled situation and then adding controls one by one. To do this effectively, stakeholders who know what controls are effective should be involved so that the process of model calibration and use in problem-solving and plan-making is simply a natural extension of the model fitting process. In the case of the Notting Hill model, this process can be extended even further back with the data needed to operate the model being in itself simulated in cases where it is difficult to observe.

To illustrate these various stages, visual interfaces are necessary with the software being user friendly and interactive as in all the programs so far in this paper. However it is debatable as to whether stakeholders and non-experts should be involved in the software *per se* as the simulations can be captured as animations and pictured using small multiples. This may be enough to engage in debate although this is uncharted territory in so far as communicating the model to a wider set of non-scientific experts

is concerned. The actual model developed begins with data which is observed where people enter the carnival and are finally destined for. This is site specific and the first thing that is done is to model the tracks that pedestrians make from their entry points

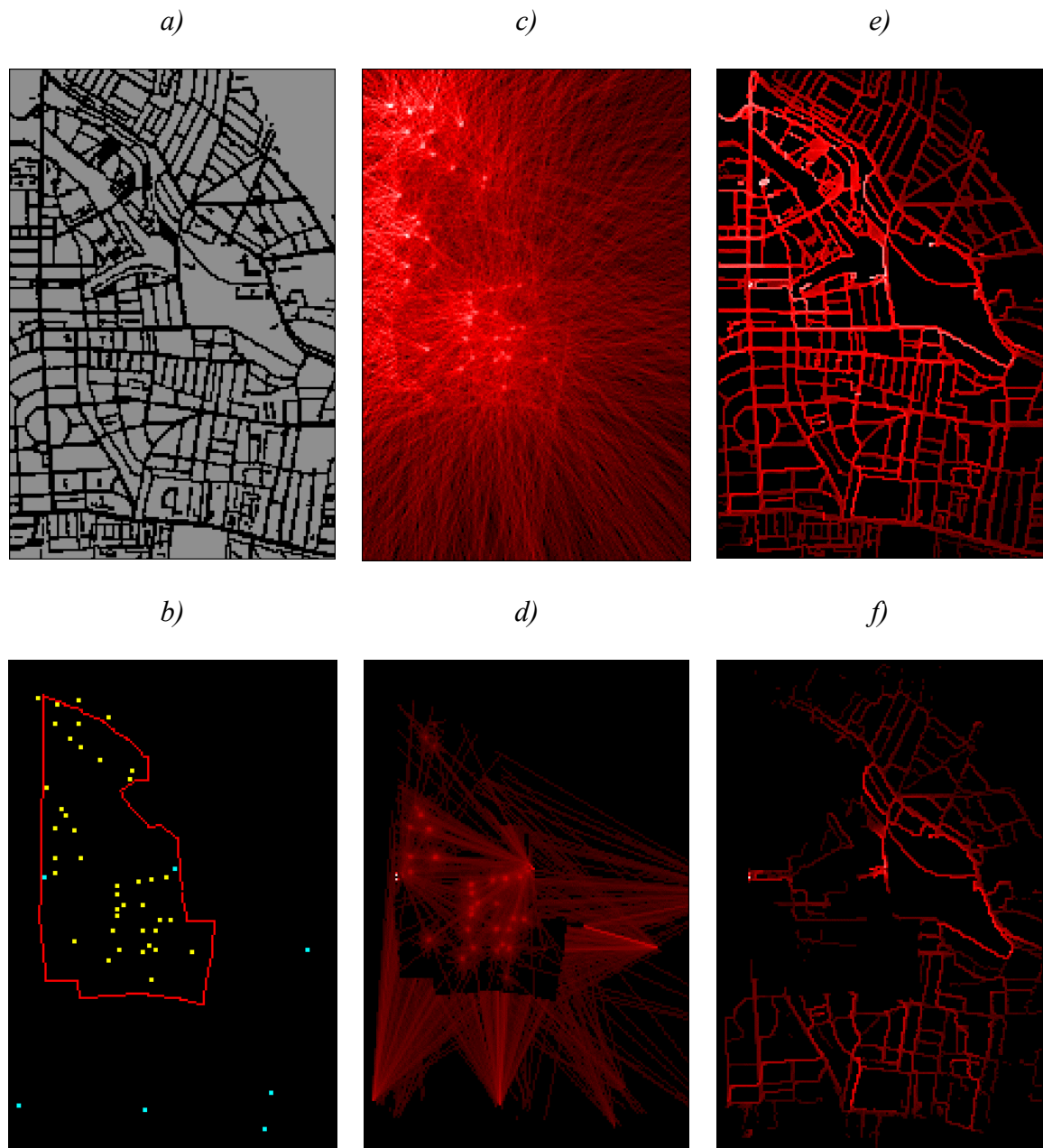


Figure 7: Exploration of the Street System in Notting Hill

- a, The street geometry b, The parade route (red), the static sound systems (yellow) and the tube stations (blue) c, Accessibility from parade and sound systems without streets d, Shortest routes to tubes without streets e, Accessibility in streets f, Shortest routes in streets.

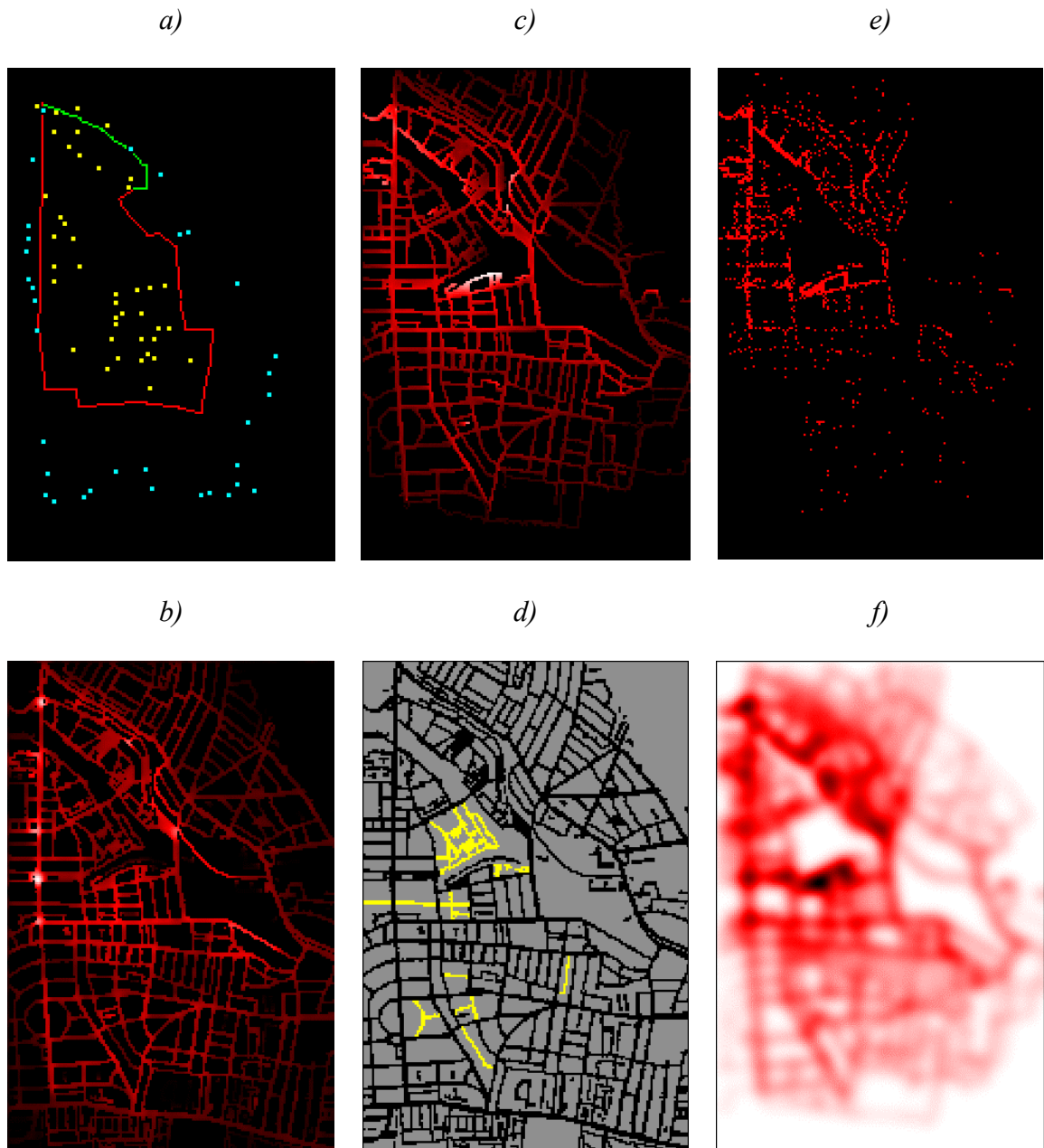


Figure 8: The Full Modeling Sequence and Identification of Vulnerable Locations

a, The 2001 parade route (red and blue) with the proposed 2002 route in red, the static sound systems (yellow) and the entry points (blue) b, The composite accessibility surface from stage 1 c, Traffic density from stage 2 d, Areas closed by the police used in stage 3 e, Location of walkers in the stage 3 steady state f, Vulnerable locations predicted from stage 3.

to the destinations at the carnival itself. These tracks can be found as shortest routes from entry points to attractions through the street system and a swarm algorithm is used to find these (Bonabeau, Dorigo, and Theraulaz, 1999). This is a rather technical stage of the model design but once completed, an attraction surface is formed from these shortest routes and the second stage invoked. This surface is used to direct

walkers from entry points to the carnival attractions and it is at this point that crowding is assessed. We start with a situation of no control and then gradually introduce controls until safety levels are reached. This involves running the model through many stages. Ideally it is during this process that those who best know the carnival should be involved. In this stage, it is entirely possible that the current situation is replicated but in fact it is likely that the current situation will be found wanting in some way, as we know it is, hence the rationale for this style of modeling.

In short, this kind of modeling involves using a model in such a way that the expertise of those who know the problem best is gradually added into the simulation. This is why we do not define this as a planning model or even a forecasting model but a model which engages those who know the situation best. The model can thus be seen as the product of many decisions from those who know the event and this naturally leads to a rather different style of ‘what if’ analysis and a rather different kind of policy making process. We illustrate two of the stages of the model using small multiples – in Figures 7 and 8. In Figure 7, we show how the walkers swarm out of the carnival area in search of destination points which is part of the early stage of generating an appropriate and realistic data set for the model. In Figure 8, we show the second stage where the walkers climb the surface of attraction, generating crowds and leading to an analysis of key problems of safety which need to be resolved. Animations of these processes are essential in visualizing how crowds move and thus how they might be controlled and in this sense, the model has a usage in almost real time.

7 Next Steps: A Paradigm for Visual Modeling

We are very conscious that we have not mapped out here a comprehensive framework for visualization in spatial modeling. This is largely because so much visualization is characterized by ingenious solutions which involve putting unlike pictures together, by large scale simulations that depend on very sophisticated software in the search for pattern, and by the very nature of the models themselves and how they are formulated. All this is influenced by the imagination we bring to bear on the pictorial world.

Equally well there are many insights to be made using other intellectual media – verbal discourse and numerical reasoning. In getting a complete picture of how our models can be understood and best applied, all these strategies are required. In fact we have shown here that rather than visualization *per se*, what is evolving are visual models: models that cannot be designed in any other way than using the visual medium. In our last example, this was even more pointed in that to involve non-scientists in developing such models, the visual medium is essential.

What we have not addressed here but is something that needs to be pursued is the physical media for spatial modeling. As our models are digital, they can be manipulated in countless ways. For example, GIS is often used not for spatial analysis but for paper map-making with the physical product and its perfection being the rational for digital representation of the reality in the first place. In the same way, models of the built environment can be printed in 3-d using the new generation of hard copy printers, thus simply aiding the manufacture of iconic models in their traditional physical form. The notion of building ‘models of models’ – simulacra as Baudrillard (1994) refers to them – is also a useful way forward and in the examples illustrated here, some elements of this recursion do permeate the model building process. In short, scientific visualization is increasingly being informed by physical visualization but with the digital representation being central and stable to this entire media. These are exciting developments and there is an urgent need to engage a debate about such possibilities in spatial modeling and analysis.

8 References

Alonso, W. (1964) *Location and Land Use*, Harvard University Press, Cambridge, MA.

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

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

Batty, M., Desyllas, J., and Duxbury, E. (2003) The Discrete Dynamics of Small-Scale Spatial Events: Agent-Based Models of Mobility in Carnivals and Street Parades, *International Journal of Geographic Information Science*, 17, 673-687.

Baudrillard, J. (1994) *Simulacra and Simulation*, University of Michigan Press, Ann Arbor, MI.

Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999) *Swarm Intelligence: From Natural to Artificial Systems*, Oxford University Press, New York.

Cronon, W. (1991) *Nature's Metropolis: Chicago and the Great West*, W. W. Norton and Company, New York.

Hall, P. (Editor) (1966) *Von Thunen's Isolated State*, Pergamon, Oxford, UK.

Kaufmann, W. J., and Smarr, L. L. (1993) *Supercomputing and the Transformation of Science*, Scientific American Library, No 43, W. H. Freeman & Co, San Francisco, CA.

Lowry, I. S. (1965) A Short Course in Model Design, *Journal of the American Institute of Planners*, 31, 158-166.

Roberts, G. K. (Editor) (1999) *The American Cities and Technology Reader: Wilderness to Wired City*, Routledge, London.

Stedman, P. (2001) *Vermeer's Camera*, Oxford University Press. Oxford, UK

Tufte, E. R. (1990) *Envisioning Information*, Graphics Press, Cheshire, CN.

Tufte, E. R. (1997) *Visual Explanation*, Graphics Press, Cheshire, CN.

Xie, Y. (1996) A Generalized Model for Cellular Urban Dynamics, *Geographical Analysis*, 28, 350-373.

Xie, Y. and Batty, M. (2004) Integrated Urban Evolutionary Modeling, in Fu, W., Atkinson, P., et al. (Editors) *GeoDynamics*, CRC Press, Boca Raton, FL, forthcoming.