Chapter 2
A Generic Framework for Computational Spatial Modelling

Michael Batty

Abstract We develop a generic framework for comparing spatial models whose dynamics range from comparative static equilibrium structures to fully dynamic models. In the last 40 years, a variety of spatial models have been suggested. Until the mid 1980s, most models were static in structure and tended to embrace detailed mechanisms involving spatial economics and social physics. Typical examples were Land Use Transportation Interaction (LUTI) models that embraced theories of spatial interaction and discrete choice modelling. During this earlier period, the problems of making these models dynamic and more disaggregate was broached but progress was slow largely because of problems in collecting requisite data and problems of increasing the complexity of such models to the point where they could be properly validated in traditional ways. 20 years or more ago, new modelling approaches from very different sources came onto the horizon: in particular, dynamic models based in Cellular Automata (CA) which were largely physical in nature and Agent-Based Models (ABM) providing explicit behavioural processes that often rested alongside these automata. Systems Dynamics Models (SDM), Spatial Econometric Models (SEM) and Microsimulation Models (MM) all informed the debate. It is tempting to see these models as all being of different genera but here we attempt to see them as part of an integrated whole, introducing a framework for their elaboration and comparison. After the framework is introduced, we review these six model types and choose three – CA, ABM and LUTI models – that we then work up in more detail to illustrate these comparisons. We conclude with the conundrums and paradoxes that beset this field.

M. Batty (✉)
Centre for Advanced Spatial Analysis (CASA), University College London (UCL), London, UK
e-mail: m.batty@ucl.ac.uk

A.J. Heppenstall et al. (eds.), *Agent-Based Models of Geographical Systems*,
DOI 10.1007/978-90-481-8927-4_2,

2.1 Antecedents: The Origins of Spatial Models

Digital computers appeared in the late 1940s largely as a result of developments in the logic of computing and the notion that large-scale numerical processing could be massively speeded up by reducing routine tasks to binary equivalents operating on equivalent electrical devices. Right from the beginning, scientific applications involved spatial as well as temporal problems and by the mid-1950s, rapid advances in digital computation led to computable problems in the human applications domain involving spatial systems such as cities and transportation. Mathematical theories of such spatial systems were slowly developing prior to the invention of the digital computer but there had been little focus on how such theories might be operationalised, tested through validation, and then used in forecasting. Digital computers were to provide the spark for such applications and in 1955, the first models of traffic flow were implemented in a digital environment in the Chicago Area Transportation Study (Plummer 2007).

These first models, unlike many if not most that have followed them, were specifically tailored to the problems in question and the way those problems were perceived. Transport flows were critical as the problems in question involved providing for new transport capacity, while land use location too was essential in a period of relatively rapid economic growth which involved the search for new locations for urban development. These early models were equilibrium-seeking rather than dynamic, aggregate at the level of populations involving spatial interactions, and built on conceptions of the city articulated using ideas from urban economics and social physics. They are usually now referred to as Land Use Transportation Interaction (LUTI) models. From these early attempts, as computers and their software developed, new generations of computable spatial models have become more generic in that the software developed for general classes of model has become ever more significant, thus elevating generic ideas about modelling through their software to a point where specific model types now tend to defer to generic modelling styles. In this chapter, indeed in this book, this notion of generic models and generic software is very much to the fore because agent-based models (ABM) and their close relatives cellular automata (CA) models represent classes and styles that are much wider in scope and applicability than the sorts of spatial systems to which they are applied.

Here we will outline as wide an array of spatial models as is possible in an integrated fashion, setting the scene for many of the more specific applications and developments in the chapters that follow. As it is rare in this field to see highly standardised applications which barely differ from case to case, each model application tends to be tailored in some specific way to the problem and its context such that model styles and structures become mixed. However what we will do is identify six distinct styles of spatial model that cover most of this array beginning with the original social physics and urban economic models that kick-started the field half a century ago. But before we introduce specific model types and show how these relate and evolve from one another, we will begin this review by examining model structures, identifying the key characteristics and themes that dominate model development. We will first focus on questions of abstraction and representation, noting the difference between the substantive components of any spatial model which we define as its population in contrast to the environment with which it interacts. In one sense, all models can be so defined and this serves as a basis on which to characterise the way populations which provide the objects or components of the spatial system under question, interact with one another and with their environment through a series of key processes. We will examine issues of representing spatial and temporal scale, aggregation, and constraints, and then we will look at processes of change, feedback, and dynamics. Many of these features and themes merge into one another and to an extent, any such categorisation of the key characteristics of spatial models is arbitrary. But these categories do enable us to sketch out the array of ideas that dominate the field which appear time and again in this book. Once we have introduced these ideas to set the context, we will examine six model types beginning with the simplest cellular automata, defining agent-based models, noting econometric, systems dynamics and microsimulation all of which involve generic approaches, concluding with notions about specific models that contain their own styles and features such as those that were the first to be developed in the land use transportation domain. To give focus to this review, we will then outline examples of CA, ABM and LUTI models in more detail, providing the reader with ideas about how such models are designed and used in practice.

2.2 Modelling as Computation: Abstraction and Representation

Half a century ago, the idea of a model was in its infancy. Scientific theory essentially was based on formal and systematic theories, often represented mathematically, whose testing was confined either to controlled experiments in the laboratory or to various categories of thought experiment. Computation changed all that. The idea that a scientific theory could then be translated into an intermediate form – called a 'model' – represented a way of enabling controlled experiments to be carried out not on the actual system of interest but on a computable abstraction of that system. The term model quickly entered the lexicon and it is now widely used to describe any kind of experimental context in which the computer is used as the medium for its exploration and testing. In fact, the term is now used even more generally to refer to any kind of abstraction that represents an obvious 'simplification of the real thing' and in this sense its meaning is no longer exclusively associated with computation (Lowry 1965; Batty 2007).

When computer models were first developed, the general assumption was that these were simply representations of the system on which testing would take place so that the theory on which the model was based could be tested against data. In general, it was assumed that the traditional canons of the scientific method in which theory was successively refined to withstand its falsification and to engender greater parsimony of explanation, would apply. Most spatial models from the 1950s onwards were predicated on the basis that their predictions would be tested against data taken from the system of interest and that the model would be tuned in such a way as to reproduce the system of interest within the computational environment in a way that was closest to the real thing. Goodness of fit was the main means of validation while calibration of the parameter values ensured that the model might be tuned most effectively to the system in question. The quest originally was thus to find some minimalist explanation for the system of interest in the belief that models should be as simple as possible while also generating predictions closest to our observations of the 'true' system. In fact, as in all science, this involves a trade-off.

Yet the complexity of human systems has meant that right from the first applications, there was continued pressure to develop greater and greater detail – to disaggregate the model's variables to the point where sufficient heterogeneity of the system might be represented in a manner useful to those who sought to use the model to make predictions. There were limits on what computation could offer and data concerning social systems has always been a problem but as computers got more powerful and as the world moved to a point where computation became all pervasive, our ability to model in detail changed by an order of magnitude. As the world moved online, new and richer data sources are becoming ever more available and this computational power combined with access to new and different data, meant that what we could model and represent began to change. Moreover, the key challenge in social systems is to know how much detail to represent and it would appear that the sorts of average behaviour that are characteristic of physical systems are rather different in the social world. Heterogeneity and hence greater detail is what seems to be required so that ever more plausible models can be constructed.

At the same time, as bigger and richer models have been built, their software has become more generic with general purpose simulation processes being articulated in software that can be adapted to many different types of problem. All this is fast leading to significant doubt that the scientific method taken from the classical traditions of physics has the same relevance to the social world as it does in the physical. Indeed even in science itself there is substantial questioning of the traditional canons of scientific inquiry as the quest for parsimony, simplicity, and homogeneity is increasingly being confronted by the need for plausibility, richness, and heterogeneity. The question turns on whether or not a simple, parsimonious model that can completely explain a limited set of system characteristics is as useful as one which contains many characteristics which are plausible in terms of the functioning of the system but cannot be proven as being of definitive explanatory value. In fact the problem is complicated by the predictability of many parsimonious models that are able to explain spatial behaviour as it can be observed but are unable to predict future behaviours which do not admit the same stability as those that are observed in the past. This is a deep problem that suggests that what we observe is considerably more ordered and structured at any point in time than that same set of observations at a future time. This is not just a problem in dynamics or equilibrium but one which is intrinsic to our ability to disentangle true explanation from the way we observe the world. Currently the received wisdom is that different models apply to different kinds of problem and problem context and that in the last analysis, models are useful to inform the debate through crystallising ideas.

In designing any model, the builder must decide what constitutes the structure of the system as distinct from the environment in which the system functions. In fact, this boundary problem is highly significant for it defines how the system relates to other systems and to the rest of the world in general. Very often the same model can be applied to different conceptions of the same system which is defined differently with respect to its environment. Here we will define the term environment rather more narrowly than its general use in systems theory where it refers to the rest of the world or the problem context. We will make a distinction between the wider environment within which the system sits relative to the rest of the world, and the local environment of the system which is the space-time nexus that pertains to the functions in question. In short, the system's environment here is the spatial tessellation of its cells or its locational referents which change through time. In contrast, we define the system in terms of its population, meaning its components and their functions that operate within this local environment. In essence, it is the population that constitutes the structure of the system and its functioning which operates in its space-time environment. The functioning takes place between the population and its environment and there are feedbacks in both directions, that is the population can influence the environment just as the environment can influence the system but these two aspects of the model are qualitatively quite different as we will see. In terms of how this *population-environment system* relates to the outside world often called the environment too, then the usual assumption is that although the environment of the outside world can influence the system, the system does not influence the outside world in terms of the operation of its model. This is the usual convention in systems theory.

In this review, we will attempt to represent all our models no matter how different using the same notational structure and to this end, we define an index of space as i or j and any interaction or relation between them as ij while we use k to define some attribute or feature of the population which pertains to different sectors. Time is indexed as t. Where we need to refer to more than two locations or two attributes or two time periods, we will define appropriate additional symbols as we proceed. We first define a spatial unit i at time t within the environment as A_{it}, and then an attribute or segment of the population at the same coordinates as N_{it}. The two matrices $\mathbf{A}$ and $\mathbf{N}$ contain the key elements of the system which interact with one another in ways that we make specific when we detail models of how populations function, interact and change and how these relate to the spatial system. We can write these feedback loops as $\mathbf{A} \Leftrightarrow \mathbf{N}$ to give some sense of the symmetry of these relations but at the same noting that $\mathbf{A}$ and $\mathbf{N}$ are generically different.

We can easily aggregate these discrete quantities into larger spatial units that we call Z_I where I is a spatial index to the number of cells i that are within Z, or into larger temporal units Q_T where T is the aggregate temporal index. Note that there are continuity and contiguity constraints that we need to be aware of when we aggregate over space and/or time. We thus define the appropriate units at larger scales as

$$\left.\begin{aligned} A_{IT} &= \sum_{i\in Z_I}\sum_{t\in Q_T} A_{it} \qquad and \\ N_{IT} &= \sum_{i\in Z_I}\sum_{t\in Q_T} N_{it} \end{aligned}\right\} \tag{2.1}$$

where there are likely to be conservation constraints in terms of size such as $A = \sum_i A_{it}, \forall t$ and $N = \sum_i \sum_t N_{it}$, the particular form of which are usually specified when the model is implemented. Functions defined on the population and the environment and the relations between them constitute the structure of the system and usually specify the dynamics of change through time. However to provide some sense of closure to this rather abstract form of representation, at any cross section in time, it is possible to define interactions between these components over space. For example, the populations might interact which we can specify in the following way without detailing the mechanisms. Then the interaction between spaces i and j can be written as $A_{ij} = A_{it} \otimes A_{jt}$ where the concatenation is specified according to some behavioural or physical principle embodied in the model.

It is worth noting that functions like this tend to be specified in systems theory independently of time so that the structure of the system is laid bare. There may be many such functions and before anything further can be said about a model structure, the mechanisms must be specified. What is important is that this framework is seen as being generic in that it can apply to a variety of different problems and problem contexts, to different systems be they physical or human, material or conceptual but with a slight bias towards the subject matter of this book which is agent-based models in the social sciences, particularly the geographical social sciences. Whether or not this is the best representation is not particularly relevant. Each model is developed in its own formal style and the purpose of this framework is to provide a template for assessing how different the array of models that we define here are from one another, not in terms of their substantive or behavioural similarities or differences. In this sense, the population and the environment can be very different. The only common point of reference is the fact that we make this distinction between these two sides of the model and specify space and time in the formal notation of cells and time instants, rather than in the continuous fashion that is often used to couch more theoretical statements of spatial models.

2.3 Feedback, Dynamics and Processes of Change

During the sweep of history over which spatial models have evolved, there has been a shift from simple, parsimonious models that simulate systems at a cross section in time and represent populations in aggregate form to more complex, richer models that deal directly with the time dimension and specify model functionality in terms of processes of change at a much more disaggregate level than their earlier counterparts. The switch has been occasioned by many forces. Already we have noted the growth in computation and the emergence of online data sources which have made much richer models possible but there has also been a sea change in the way we think about human systems. Complexity theory has raised the notion that systems are never in equilibrium, in fact their predominant condition is far-from-equilibrium and disequilibrium is their normal state. Moreover human systems have become ever more complex due to technological change, the demographic transition and increasing wealth, at least in the west and many newly developed countries. This has made spatial behaviours more complex, certainly in terms of movement and communication as well as locational preferences. All in all, dynamics has come firmly onto the agenda while the notion of explanation has shifted from aggregates to much more heterogeneous populations composed of individuals and groups that need to be understood at a much finer level of detail.

In generating behaviours, feedback is an important mechanism where we might specify this in functional terms as a dependence of population on itself or on the environment, that is $A_{it+1} = f(N_{it})$, $N_{it+1} = f(N_{jt})$, or $N_{it+1} = f(A_{it})$ and so on. Negative feedback tends to damp activity so that departures from some norm are restored, the classic example being a thermostat which controls the heat from a boiler to some environment. Positive feedback on the other hand accelerates the degree of change, sometimes with catastrophic consequences, but usually with beneficial impacts if some quantity such as income or even population is increasing. The best way to illustrate the effect of feedback is in terms of population growth and the basic equation which can be used to simulate positive feedback is

$$N_{it+1} = \alpha N_{it}, \tag{2.2}$$

where α is the rate of change defined as N_{it+1} / N_{it}. If the growth rate α is greater than 1, then this leads to exponential growth as we show in Fig. 2.1. If less than 1 then this leads to a decline to zero population but in both cases, the change is due to the compounding effect which can be easily seen if we generate a recursion on Eq. 2.2 up to time $t+T$ as

$$N_{it+T} = \alpha^T N_{it}. \tag{2.3}$$

Negative feedback can be shown when change is damped according to some threshold but it is more appropriate to show this as a moderation of exponential growth as encapsulated in the logistic equation. Then if we define a limit to population as say $\bar{N}_i$, then we write the logistic as

$$N_{it+1} = \beta N_{it}(\bar{N}_i - N_{it}), \tag{2.4}$$

where the rate β is moderated with respect to the scale of the growth. We also show this form in Fig. 2.1 where it is clear that the population grows exponentially at first and is then damped by the effect of the constraint $\bar{N}_i$. In fact if the damping effect is lagged leading to an oscillation around the limit value of $\bar{N}_i$, then the growth of population mirrors the sort of behaviour characteristic of systems dynamics models that were developed by Forrester (1969) in cases where resource limits dominate.

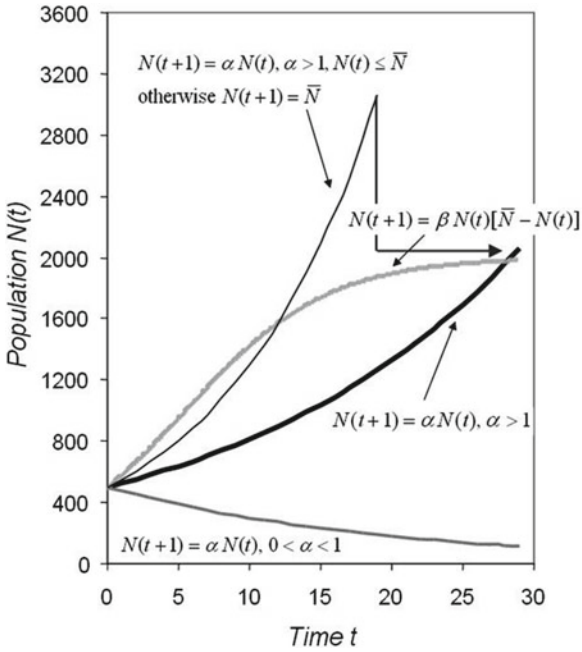

Fig. 2.1 Simple dynamics based on positive feedback

Another form of dynamics relates to variations across space in the manner we illustrated for spatial interaction in the previous section. If we add time to this kind of dynamics which involves spatial relations, associations, correlations or movements, then we can represent these as flows from i to j between times t and $t+T$. In fact the interaction N_{ij} which we associate with cross-sectional static models in the previous section does take place through time although the time is much shorter than the usual periods that are associated with spatial modelling. Only quite recently has our concern in understanding cities shifted to thinking of cities in real time for such a real time focus has previously been captured as a static snapshot of movements in the city as, for example, in transport and traffic modelling. However longer time periods are associated with flows such as migration where the variable $N_{i,j,t,t+T}$ is now associated directly with time. Mechanisms for such models are only specified when the precise form of model is defined and these are often based on activity patterns, distance, travel time and related cost structures that determine spatial associations. In fact, flows of this kind are also associated with networks which scale from topological relations down to physical infrastructures. Currently there is substantial activity in embedding such flow structures in their networks and this is beginning to be reflected in spatial models as is implicit in some of the contributions in this book. In the three examples we use to illustrate the computational model types below, flows and networks are significant. It is worth noting too at this point, that in spatial modelling, most focus has been on measurable physical and hence observable quantities that change through time but increasingly there are hidden flows and relations associated with the electronic world that are influencing how spatial systems change and develop. This too is a major challenge for spatial modelling.

Before we move the discussion on to classify different kinds of dynamics, it is worth noting that all the variables that we have introduced so far can be disaggregated down from their aggregate populations to more disaggregate components. Ultimately the disaggregation is down to individuals where we denote such atomic elements by the subscript k on populations, that is N_{it}^k is the $k'th$ individual or group in the appropriate space and at the relevant time in the system. In fact this notation can also be extended to any group which is a subset of the aggregate population such that the sum of the groups and/or individuals adds to the relevant aggregate variables, that is $\sum_k N_{it}^k = N_{it}$. In the models, particularly the agent-based models that follow, individuals will form the focus of the simulation where processes are specified for a class of individuals but are operated at the level of each individual in the simulation, usually specific to the space and time within which the individual is located as well as individuals 'proximal in some way' to the object in question.

As the concept of equilibrium has fallen into disrepute and as the spatial models have become more explicitly dynamic, different kinds of time scale and change have been identified which characterise spatial systems. In particular the notion of smooth change has given way to systems that clearly have discontinuities in their behaviour through time (as well as space) where such discontinuities represent thresholds that are crossed, for example, the step function also shown in Fig. 2.1. Here once the growing population reaches the limit, it precipitously declines to its initial value. This classification of dynamics extends all the way to behaviours that generate endogenous discontinuities as is characteristic of catastrophe and bifurcation theories. This portfolio of dynamic behaviours has also been enriched by smooth changes that lead to chaos, systems that behave in entirely unpredictable ways in terms of their initial conditions but are nonetheless deterministic and portray smooth and continuous change. Into this nexus has come the notion that change can generate surprising and novel behaviours. For example, edge cities that suddenly appear around well established metropolitan areas, segregation patterns that do not appear to be embedded in the logic of change but suddenly manifest themselves, and repercussions from changes in one element of the system that cascade and grow as they diffuse to other sectors are all examples of the sort of changes that many models of spatial systems now take as routine.

Dynamics in all these senses has added to the burden of modelling. Like disaggregation, dynamics enriches the model in that data demands become severe and often much of the change that needs to be simulated is hard to observe and match to data. In fact, the notion that dynamics leads to surprising changes is part and parcel of the insights that are coming from complexity theory where the routine operation of space-time processes from the bottom up leads to emergent patterns that only in hindsight can be explained. Such unanticipated behaviour is quite counter to the traditions of well-behaved dynamic systems that tend to converge to an equilibrium or steady state, that is where $N_{it} \to N_i$ in the limit of t.

The last feature of dynamics that we need to note before we begin to classify spatial models in these terms involves relationships to the external environment, either to the rest of the world or indeed exogenous changes to the population and/or environment that comprise the system in question, for it is by these means through which unusual dynamics can be stimulated. For example the population equation might be subject to external shocks, that is from Eq. 2.1 we might add a shock such as X_{it} leading to

$$N_{it+1} = \alpha N_{it} + X_{it} \tag{2.5}$$

which basically removes a degree of predictability from this particular model, dependent on the size and frequency of the external input. If the external input is once-for-all, its effects may die away but sometimes these kinds of shocks feed on one another and are enough to push the system into uncharted waters with quite unpredictable consequences. Moreover changes in the environment of the system, such as the addition of new capacity a_{it} in terms of land available, say, which we might mirror as

$$\left.\begin{aligned} A_{it+1} &= A_{it} + a_{it} \\ N_{it+1} &= f(A_{it+1}) \end{aligned}\right\} \tag{2.6}$$

can lead to equivalent unpredictability. Even in these simple cases, we can easily complicate the dynamics through additional functions that immediately show that any movement to a steady state is likely to be the exception rather than the rule.

2.4 Six Styles of Spatial Model

It is exceptionally hard to provide a completely comprehensive overview of spatial models in the human domain even with as narrow a focus as we adopt here which is mainly on cities. This is largely because model types shade into one another and many of the features that we have identified in the previous sections appear in more than one model. Different modelling styles merge into one another. Nevertheless various researchers have attempted to classify such models and it is worth noting some of these attempts before we outline our own focus on this field. In general as noted earlier, there has been a sea change from aggregate cross-sectional comparative static models of spatial systems to models that are disaggregate and dynamic. This has marked the transition from *land use transportation interaction* models (LUTI) to *cellular automata* (CA) and *agent-based models* (ABM). This has also represented a change in scale and focus and in the case of CA models, these shift the focus from social and economic processes to physical land development. ABM models are more generic still but in terms of urban modelling, most applications are at the fine spatial scale at the level of pedestrians, for example, and local movement, with only a handful of such models being developed for metropolitan areas. In fact as LUTI models have been disaggregated, then some of these such as ILUTE and UrbanSim have features that can be described as agent-based (Hunt et al. 2005).

The other four types of model that we will classify and define here are those based on less well entrenched applications and methodologies. *Spatial econometric models* (SEM) have been widely applied but often at a larger scale involving regions while *systems dynamics models* (SDM) have been proposed and implemented in some contexts but these have not found widespread application largely because they have not been generalised to spatial systems in any consistent manner. Last but not least there are *microsimulation models* (MM) of which there are several spatial variants and these also tend to merge into ABM at one level of specification. There are no general reviews of all six modelling styles but the author (Batty 2008) provides a discursive discussion of how LUTI models made the transition to CA and ABM during the last 30 years. The short review of LUTI, ABM and CA models also by Batty (2009) focuses on their structure, dynamics and aggregation properties. There are comprehensive reviews of ABM, CA, SDM, MM and some LUTI models by Haase and Schwartz (2009) and there are a series of reviews of operational land use models mainly in the US agencies such as the EPA (see Southworth 1995 for example). However apart from the review of CA models by Liu (2008), most of the reviews tend to be of LUTI models. In particular the chapters by Wegener (2005), Iacono et al. (2008) and Hunt et al. (2005) are good summaries of the state of the art to which the reader is referred. The essence of the models which are the subject of this book – mainly ABM, CA and MM – are contained in the relevant chapters in this section by Birkin and Wu (2011) (MM), Dearden and Wilson (2011) (LUTI-spatial interaction), Iltanen (2011) (CA) and Crooks and Heppenstall (2011) (ABM). In fact the focus is much more strongly on ABM than any other model type in this book although CA models, as we will see, provide an implicit form of ABM. This chapter and more generally this section, do however provide a useful overview of the field with the focus very much on situating ABM in the wider context of spatial modelling.

We will begin with generic models and only when we have reviewed most of these will we look at specific models with methodologies that are precisely configured to the systems and problems at hand. We will treat each model in terms of the eight characteristics which we identified in the previous two sections, namely, environment and population, scale and aggregation, conservation and constraint, disaggregation, feedback in space-time, dynamic type, emergence and convergence, and external inputs, and we will begin with CA models which are by far the simplest. In fact CA models are explicit and simple spatial dynamic models with little or no presumption about the form of the dynamics and rather simple notions about the effect of space. In their strictest form they simulate the spatial diffusion around a point where the diffusion is to immediate neighbours and time and space are treated as one. In this sense, the environment is treated as being synonymous with the population with each state of the system – i.e. the population – being directly associated with a spatial location at a point in time, in short $A_{it} = N_{it}$. Scale and level of temporal and spatial aggregation tend to be quite flexible in these models although for urban and land cover systems, both scales are large – often land parcels and census tracts and above, while temporal intervals are at least for one yearly periods. This however is not a major constraint. Such models do not strictly conserve quantities of population in that there is nothing intrinsic to such models that limits their growth or decline although often such models are subject to more macro-constraints provided by other models in their wider environment. The models can be fairly disaggregate but most applications divide the cell states into land use types limited to no more than a dozen. Feedback in space is extremely simplistic and often unrealistic in that the CA nearest neighbour influence principle which is essential for physical diffusion processes is often not a good analogue for spatial effects where there is action-at-a-distance. Such models do not tend to fall into any particular dynamic class, for if they produce unusual and discontinuous dynamic behaviours, this is likely to be due to external inputs rather than anything built into the model dynamics. Emergence is possible with such models, indeed essential to their original formulation although in urban applications this is generally not a specific focus. All in all, such models tend to simulate land development processes from the supply side or at best models of the balance between demand for and supply of land. They are not strongly socio-economic in that they do not embrace detailed demographics, and in this sense are essentially physicalist in tone.

ABM models have many of the characteristics of CA models except that the environment and population sides of the system are kept apart. The population sector is essentially that which contains these agents whose behaviour is specified in considerable detail. Agents tend to be mobile in a spatial sense and even if they do not physically move in space, they can be associated with different spaces and their change over time can reflect an implicit process of movement. In this sense, the environment is treated more passively than the population with the population driving any change in the environment, although in principle there is no priority for one or the other. A detailed specification of ABM in these terms is contained in Batty's (2005) book where the idea of an agent having a specific behavioural profile and acting on this purposively is central to their definition. In terms of aggregation and scale, ABMs tend to be at smaller scales than the region or the metropolis although some land cover models based on ABM are predicated at these larger scales. They tend not to be constrained in terms of conserving any key quantity although they may be structured to generate or conserve a certain level of population, especially if the focus is on movement in a fixed space as in pedestrian models. Their dynamics and relationships to the wider environment are similar to CA and they tend to be highly disaggregate down to the point where individuals constitute their basic units. Problems emerge when individuals are aggregated to groups or when the agents become agencies for then such models tend to be of more conceptual interest than of predictive practical use.

Like CA and ABM models, microsimulation models (MM) tend to be loosely structured in terms of their dynamics. Such models may even be cross-sectional rather than dynamic but the fact that the populations tend to be represented in terms of their basic units means that such models are usually temporally dynamic, i.e. individuals are represented in terms of their behaviour which is intrinsically dynamic. These kinds of models work on the premise that a population is described in terms of a distribution of characteristics – for example, an income distribution, and individuals are then selected from that distribution so such models are essentially random samples from a much larger universe or population. In this sense, the models can be at any scale but the distributions are usually composed of individuals in that any point sample from a distribution is associated with an individual. Point samples can of course be aggregated into large groups in space and time. There is not much more that can be said about such models for all their other characteristics will depend on the specific model characterisation once it has been worked up to the system in question. Quantities do tend to be conserved and sampling can be subject to some constraints while feedbacks depend on how different sectors in the model are configured in relation to one another. The model dynamics again tends to be straightforward and most models to date (see Birkin and Wu 2012) do not tend to reflect discontinuities of the kind associated with emergence of new structures. External inputs into such models are usually extensive as many of the drivers of such behaviour are reflected in the wider environment. Microsimulation models are essential tools for sampling large-scale populations where it is impossible to represent all the individuals explicitly and where some sense of the heterogeneity of the population needs to be represented in the model. The MoSeS model designed by Birkin and Wu (see this volume) is a good example of how MM is applied to human spatial systems where the focus is on demographics and its relationships to the provision of health and related social facilities at a fine spatial scale.

Spatial econometric models (SEMs) have been widely developed in the tradition of aggregate modelling (Anselin 1988). To an extent such models do not really distinguish between population and environment although the focus in such models is more on subsuming the environment into the population than the other way around in contrast to CA models. Such models are usually developed at a scale where statistical averages are stable and this means that the spatial and temporal units must be such that the data are appropriate for standard statistical inference. Quantities in such models tend to be conserved but within statistical limits although increasingly constraints are put on statistical models where it is essential to keep predictions within bounds. SEMs tend to be structured along rather formal lines where the standard model is linear, often simultaneous in that feedbacks between different model sectors are associated with different model equations, and the dynamics is often well-defined with the equilibrium properties of such models being well-known in terms of their stationarity. Emergent behaviours are not usually a feature of such models but the distinction between exogenous and endogenous variables as in much economic modelling is strong. In this book, these kinds of models are not reported although occasionally, econometric techniques are used in ABM, SDM, and MM.

Systems dynamics models (SDM) are very much in the tradition of the discrete population models that we illustrated earlier in Eqs. 2.1–2.4. In fact these models are based largely on coupled difference equations whose structure is such that they lead to exponential growth followed by damped oscillations around fixed resource limits. In this sense such models are heavily constrained. They can be quite disaggregate dealing with different sectors but the environment is entirely absorbed in the population as there is usually no spatial variation although some models have simply applied what happens in one space to many others. In terms of feedbacks, the entire behaviour of these models is structured around damped logistic growth reflecting repercussions through the model structure which leads to oscillations around the resource limits. In this sense, the dynamic behaviour of these models is well-defined. Links to the wider environment are structured in terms of control over resource limits. Progress with these models has been quite slow with only a limited number of applications largely due to the difficulty of articulating space within their structure. In fact as soon as space is introduced, these models begin to look rather different from traditional SDM and in this sense, they change in focus. Many of these model structures are more like model methodologies that can be merged together in the construction of more elaborate models, as for example, in models such as UrbanSim.

Our last class of models – land use transportation interaction or LUTI models – are quite different in structure. These models are essentially fashioned around ideas in spatial interaction and discrete choice theory, merged with notions about economic input-output analysis, multipliers and demographic modelling that all come together in what are largely aggregate cross-sectional model structures simulating the location of activities and their interactions at a single point in time. These models, like SEM, tend to merge environment into population and since their inception, they have become more disaggregate. Spatial constraints and the concatenation of activities are central to such structures. Various feedbacks between the sectors are incorporated but these usually reflect spatial not temporal effects. In terms of dynamics, such models struggle to embrace the wider portfolio of possibilities being, at best, incremental which essentially involve static models being applied to increments of time. That is, static model structures are used to model incremental change and such models do not attempt to explore longer term dynamics. In fact there are extensions of such models into dynamic frameworks such as those developed by Wilson (2008) but in general, the practicalities of limited temporal data have constrained such models in terms of dynamic simulation. This is an important issue as most of the other models we have described in this section simply assume that the lack of temporal data is not a constraint on their specification and application. In short, LUTI models build on social physics and urban economics which are essentially atemporal.

These model types and styles provide a wide range of possible structures from which to select appropriate forms for specific problems. Our summary shows at a glance the array of model types that we might draw upon in simulating spatial systems in the human domain. In the rest of this review, we will not detail all of these but we will focus on CA, ABM and LUTI models to give some flavour of how they might be developed and the way they are calibrated, validated, and verified in practice. This will set the scene for the rest of the review chapters in this section which take these models types further and develop specific issues with respect to their design and construction.

2.5 Cellular Automata: Physical Simulation Models of Urban Morphologies

CA models are by far the simplest of any urban model in that they merge entirely their populations with their environment. In essence, the components of the environment are identical to the objects comprising the population in the sense that the locational spaces that define the environment at any point in time, are equivalent to each elements of the population. In the simplest case, one cell in the environment is equivalent to one object in the population which in formal terms means that $A_{it} = N_{it}$. Now each cell in a CA model can take on more than one state which means that the population object can vary in its attributes. Again, the simplest form is that a cell can take on one of two states – it can be switched on or off which in urban terms might be compared to the cell being developed or not developed. This is often represented as

$$A_{it} = \begin{cases} 1 \; \textit{if} \; i \text{ is developed} \\ 0, \text{otherwise} \end{cases} \tag{2.7}$$

In slightly more complicated CA models, there may be more than one population object in one cell but this probably is the interface between CA and ABM. If a cell has one population object only but that object can take on different attributes or changes in state, then this is still a CA model. In short, when a cell can take on more than two states, then this is usually used to reflect different changes in land cover such as land use types but it could also be associated with different changes in the population object such as its level of income, its age and so on. The formulation is entirely generic.

CA models in their strict sense have no action-at-a-distance except in the most restrictive sense. A cell is deemed to influence or be influenced by its nearest neighbours where near is defined as physically adjacent if the application is to some spatial system. This is the only way in which emergence can be charted in such models in that if the field of influence is wider than nearest neighbours in a regular sense, then it is impossible to trace any emergent effects on the ultimate spatial structure. Essentially CA in this manner is used to implement procedures that lead to fractal structures where patterns repeat themselves at different scales which only emerge when the system in question grows and evolves. We can illustrate strict CA in the following way. Assume that the set Z_I is the set of immediate neighbours on a regular square lattice. The usual neighbourhood is defined as the Moore neighbourhood – all cells at the eight compass points around the cell in question or the von Neumann neighbourhood which are the cells N, S, E and W of the central cell. Then we define a function F_{it} as the concatenation of effects in the Z_I neighbourhood, and if this function takes a certain value, this generates a change in state of the cell in question, cell i. Imagine that the rule – and there can be many, many different rules – is that if this function is greater than a certain threshold Ψ which is a count of the developed cells in the neighbourhood, then the cell changes state.

In the simplest case, it is developed if it is not already developed or its stays developed if already developed. Using the definition in Eq. 2.7, then

$$F_{it} = \sum_{j \in Z_I} A_{jt} \qquad \text{and} \tag{2.8}$$

$$if \quad F_{it} > \Psi \quad then \quad A_{it+1} = 1. \tag{2.9}$$

It is very easy to show that this process leads to a regular diffusion starting from a single cell. If we assume that the threshold $\Psi = 1$, all the cells in original Moore neighbourhood around the seed cell get developed first, then all cells around those that have just been developed, and so on with the recursion simply leading to the growth of a square cellular region around the starting cell. In fact in this instance, space and time are collapsed into one which is the key criteria of regular physical diffusion. These ideas are developed in more detail in Batty (2005) to which the reader is referred for many illustrations of such basic strict CA models.

If the CA models are slightly more complicated in terms of their neighbourhood rules then various geometric fractals result while there can be key spatial orientations and biases introduced into the structures that are generated. However it is usual in CA modelling for the neighbourhoods, the rules and the process of generation to be entirely uniform. As soon as the notion of varying neighbourhoods over space and varying rules over time is introduced, the models are no longer CA. In fact many urban applications are not strictly CA models at all but cell-space models, motivated by physical land development problems and raster based GIS map algebras in that they do not generate emergent patterns in any recognisable form and they usually relax the constraints placed on both size of neighbourhood and uniformity of cell transition rules. In Fig. 2.2, we show three typical CA models generated using the Moore neighbourhood. The first is the simple diffusion from a source where any development in any adjacent cell spurs development of the cell in question, the second is simple diffusion from a source using a fractal generating rule where the pattern of cells developed determines the rule, and the third is based on a more complicated pattern of cells in the neighbourhood that steers the growth which in this instance is stochastic in a given direction. These are the kinds of structures that form the basis of such automata and all applications to real systems contain mechanisms of recursion built along the same lines as those used to generate the patterns in Fig. 2.2.

There are several ways in which the strict CA model has been relaxed in developing spatial applications. First it is easy to control the growth of developed cells by imposing some sort of growth rates with respect to different cells. If growth is one unit cell, then various external constraints can be used to control the growth but as in all cases where the homogeneity rules are relaxed, then the CA no longer can generate emergent patterns in quite the simple way in which those in Fig. 2.2 are generated. Moreover to introduce variety and heterogeneity into the simplest

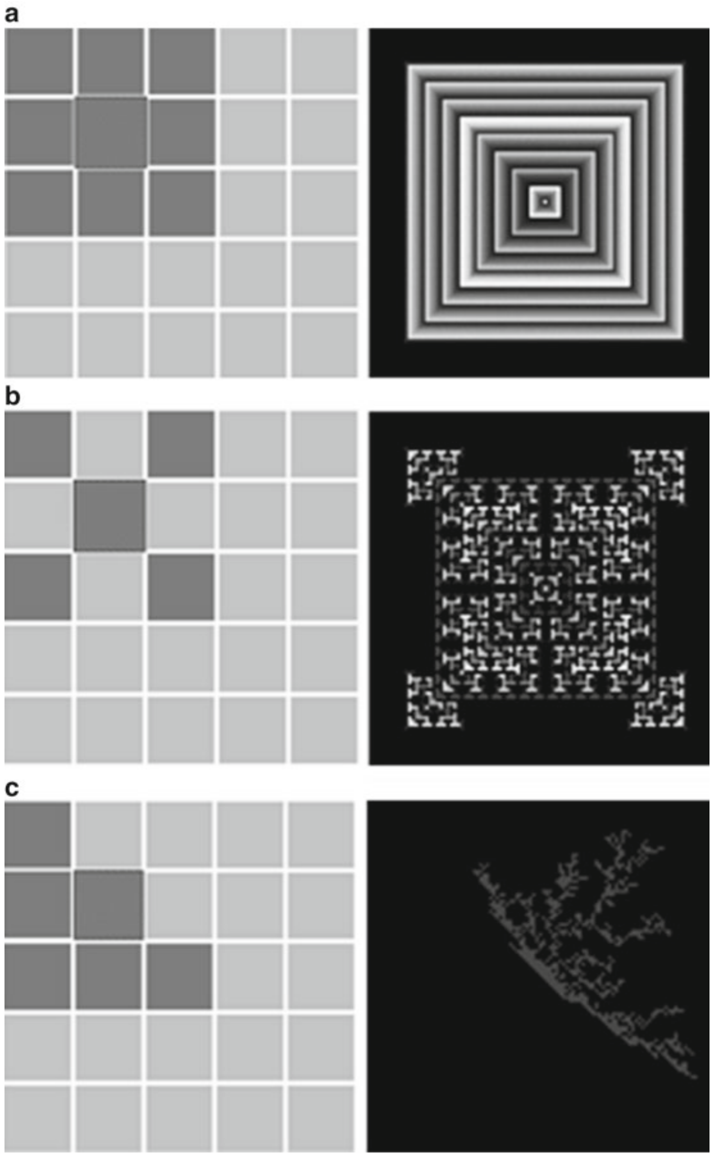

Fig. 2.2 Classic CA models (**a**) Nearest neighbour physical diffusion on a grid (**b**) Koch-like fractal diffusion (**c**) Oriented diffusion limited aggregation

models directly, sometimes the cellular count or concatenation of cells performed in the neighbourhoods is converted to a probability function which is then used to condition the development using a random number generator. For example the structure in Eqs. 2.8 and 2.9 now becomes

$$P_{it} = \sum_{j \in Z_I} A_{jt} / 8, \quad 0 \le P_{it} \le 1 \qquad \text{and} \tag{2.10}$$

$$if \quad rand\,(\Psi) < P_{it} \quad then \quad A_{it+1} = 1 \tag{2.11}$$

where P_{it} is a probability of development and $rand(\Psi)$ is a random number between 0 and 100, say which if less than the probability, implies the land should be developed. There are many adaptations that can be made in this manner but the most significant is related to relaxing the strict neighbourhood rule replacing this with some sort of action-at-a-distance. For example replacing F_{ij} in Eq. 2.8 with the gravitational expression for accessibility leads to

$$F_{it} = \sum_{j} A_{jt} \big/ d_{ij}^2 \tag{2.12}$$

and this provides a model which can predict development in proportion to accessibility, that is

$$A_{it+1} \propto F_{it} \tag{2.13}$$

This almost converts this cellular automata model to an accessibility potential models which lies at the core of spatial interaction theory and was first developed for these purposes at the very inception of land use transportation modelling (Hansen 1959). The question of course is how such a model might related to the extensive tradition of LUTI models that are in general far superior in their explanatory and predictive power than these kinds of CA model.

One of the major developments of these cellular models is to specify different cell states in terms of different land uses which we will disaggregate and notate as k, A_{it}^k being the appropriate land use k in cell i at time t. In several models, these land uses relate to one another as linkages which determine, to an extent, the locational potential for a site to be developed. Then we might write the change in state of the cell in question as a function of several land uses in adjacent cells where we use a functional notation to simply indicate that the change in question has to be specified in more detail once the model application is implemented. Then the new state of cell i at time t would be

$$A_{it+1}^k = f\left(A_{jt}^{\ell}, d_{ij}\right) \quad \forall \ell \tag{2.14}$$

where $j \in Z_I^k$ is a neighbourhood defined entirely generically and the field over which distance is defined is again specific to the zone in question. In fact this relaxes the strict CA quite dramatically and is characteristic of many applications (for reviews see Batty 2005, and Liu 2008). It is worth noting that the rules to define land use transitions generally vary the definition of the neighbourhood from the strict no action-at-a-distance principle to the gravitational one. This links different land use states and their densities and types to each land use in question, and also relates these links to different action-at-a-distances effects. These rules also pertain to constraints which are hard and fast on whether a cell can be developed or not. Above a given level, they define how land uses cannot relate to one another. Rules extend to the development of transport links in cells that ensure land use is connected, and structure the regeneration of cells according to various life cycle effects. All of these rule sets are featured in CA models and they are central for example to the SLEUTH, DUEM, METRONAMICA and related model packages that have been developed (Batty, and Xie 2005). They will feature in our brief reference to the DUEM model below.

A more generic CA like structure which is a lot closer to the differential model that dominates the dynamics of physical phenomena at much finer scales is based on a reaction-diffusion structure which might be written in the following way:

$$A_{it+1} = \alpha A_{it} + \beta \sum_{j \in Z_I} A_{it} + (1 - \alpha - \beta) X_{it} \tag{2.15}$$

where α and β are normalising parameters between 0 and 1 and X_{it} is an exogenous variable that reflects changes from the wider rest of the world environment that might be treated as error or noise in the system but more usually is treated as an exogenous shock or as an input that is not predictable by the model. To operationalise this structure, it may be necessary to impose various other constraints to ensure that variables remain within bounds but the essence of the structure is one where the first term on the right hand side is the reaction, the second the diffusion and the third the external input or noise. If we assume that $X_{it} = 0$, the evolution or growth is purely a function of the trade-off between how the system reacts and how activity within it diffuses. In fact, this is rather an artificial structure as change in absolute terms always needs to be controlled and in this sense, external inputs are always likely to be the case. Many CA models do not explicitly adopt this more general structure and a lot of applications have tended to simply scale the outputs of the developed cells to meet exogenous forecasts rather than introducing such exogeneity in more consistent and subtle ways as in the reaction diffusion model in Eq. 2.15.

There are many variants of CA models, examples of which are contained in the last section of this book but as we will see these do tend to merge into ABM. To conclude this section it is worth outlining a model that the author has worked with (see Batty et al. 1999, and Batty 2005). This is the Dynamic Urban Evolution Model (DUEM) which is a fine scale cellular model with several cells states reflecting land use as well as transport and a series of decision rules for changing states that relate one land use to another through its density and accessibility as well as their position in the life cycle of development. The model is largely a pedagogic tool rather than one which can be finely tuned to real situations although a number of applications have been made to the Ann Arbor region and the wider region of South East Michigan which is largely metro Detroit. The model is based on several land uses – residential, commercial, industrial, open space, vacant land and transport/road space – which are functions of the different density and accessibility rules as well as plot sizes which determine how land is developed. We have developed the model for

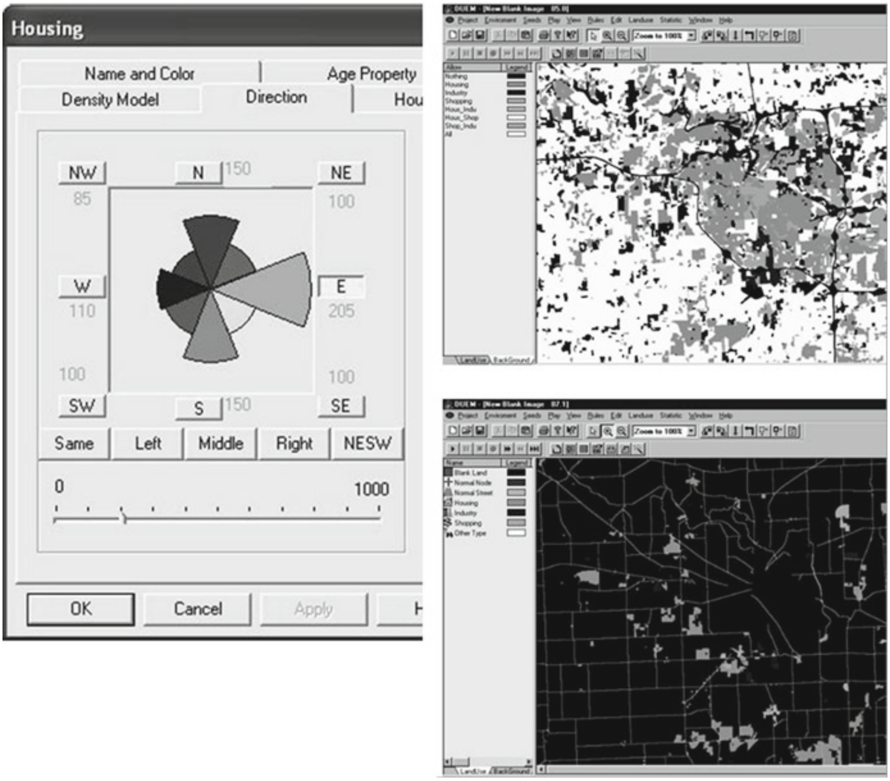

Fig. 2.3 Application of a typical CA model to simulating land use change 1985–1990 and 1990–1995 in Ann Arbor MI

the changes between 1985 and 1990, calibrating the model in a rather crude way. The rule set is large and thus we have not engaged in any kind of exhaustive calibration to find the best fit, although the fit to 1990–1995 from the calibration between 1985 and 1990 is reasonable. We show a segment of the typical interface to the models, showing developed land use in Ann Arbor in 1990, and changes predicted by the model from 1990 to 1995 in Fig. 2.3.

The real critique of CA models relates to their highly physicalist approach to urban structure and dynamics. Essentially these are supply side models, simulating the supply of land based on physical constraints. The notion of demand is foreign to these models as is the notion of interaction as reflected in transport. By abandoning the principles of uniformity, restricted neighbourhoods and homogeneity of states which it is often necessary to do once one applies these ideas, then the models often become poor equivalents to LUTI and other models. However in their favour is the fact that they are explicitly dynamic although dynamic processes other than physical land development do not feature very much in their formulations. Their dynamics is also rather straightforward and if surprising and novel forecasts do emerge, this is more by accident than design in the sense that these models tend to simulate a relatively smooth dynamics. There are not many points at which the kinds of disequilibrium crises and discontinuities which plague the modern world can enter into such models. They also owe a lot to GIS and remote sensing and it is no accident that they have been almost entirely developed by a very different set of researchers from those still working with traditional urban models in the LUTI tradition.

2.6 Agent-Based Models: Purposive Behaviour, Physical Movement and Temporal Change

As we have argued, at one level CA models can be seen as simplified varieties of ABM where the cells form the agents and the states their attributes. Unlike ABM, however, cells do not move and if they change their state, this change might be attributed to some movement but this movement remains implicit and is not formally simulated. ABM implies some form of movement or at least change between agents. Agents as objects in the population are defined individually as k but are made specific in terms of the locations where they exist i at time t. In fact agents may not physically move or indeed in non-spatial models, they may not even be defined in terms of location. If the model is simply one of examining relations between agents at a cross section in time, then such relations might solely be defined in terms of say N^k and N^l, the relation between them defining a link in a social network $N^{k\ell} = f(N^k, N^\ell)$. In fact throughout this book, the agents that are defined by various authors, exist in terms of location and time but very different kinds of relations exist across space. These imply movement or interaction from i to j, from time t to $t+1$ or a later time period $t+T$, from individual object k to ℓ as we have just defined in terms of social network links and any higher order combinations such as: links across space and time, space and different individuals, time and different individuals and across all three – space, time and individuals.

The key difference between CA and ABM is that the system is driven by the ABM where each individual object is endowed with purposive behaviour which conditions their specific and individual behaviour in contrast to aggregate models where this behaviour is part of an aggregate or collective. In this sense, the environment of the system is the space-time frame A_{it} which is relatively passive in comparison to the behaviour of the agents N_{it}^k. Nowhere in such models does $A_{it} = N_{it}$ or vice versa but as we have already implied earlier there are certainly feedback loops $A_{it} \Leftrightarrow N_{it}$ as well as the core loops between agents themselves which we define generically as $N_{it}^k \Leftrightarrow N_{jt+1}^\ell$. We assume in ABM models for spatial systems that the environment is not purposive, that is, no loops such as $A_{it} \Leftrightarrow A_{jt+1}$ exist. If such loops are required then the model would need to be reformulated and part of the environment may then enter the population. The movement of an agent is particularly important in spatial models because whereas in CA, these models tend to be bereft of spatial interaction, ABM models have found extensive application as models of fine scale movement at the pedestrian level for example (Batty 2003). We can formulate such a model in functional terms as

$$N^k_{jt+1} = f(N^k_{it}, N^\ell_{jt}, A^a_{it}, A^a_{jt+1}, z \in Z_I) \tag{2.16}$$

where the superscript a relates to some characteristic attribute of the cell. The functional Eq. 2.16 suggests that agents move through space across time but are influenced by other agents and other locations during such a move. The object N^ℓ_{jt} is in a different location from the moving object k and when the move takes place, a whole series of relations might exist between these two objects such as the visibility of one from another, avoidance of physical contact between one and the other, the clustering of the two or more objects through some social network, or the attributes of the other object being of importance to the locational move, and so on. In terms of the cells themselves, then an object moving from one cell to another would also take account of related cells in the system, usually in the neighbourhood of the move itself.

A good example might be shopping behaviour. An agent enters a shopping centre with a specific purpose to buy goods, encounters other agents along the way, avoids them, or follows them in terms of the crowd. The agent would be influenced by the provision of goods in different cells of the system and in this sense would move in relation to the existence of materials and products that were located in different cells of the system. This kind of characterisation can provide a baseline for movement with visibility, obstacle avoidance, the search for a location which matches the purpose for which the object or agent is moving, and so on. The agent may have a budget and when visiting different cells would exhaust this budget and end the trip once the movement had achieved its purpose. In terms of other moves, then if the agent were migrating over a longer time span in search of a job or house, then the characteristics of the job or house location would be encoded into the environment, in A^a_{it} but the job itself and maybe the actual house would also be part of the set of agents. In this sense, an agent need not be a human individual but an object in the built environment that in and of itself might be subject to change in type and location.

It is worth sketching a simple model of the development process to show how generic this kind of thinking can be. First we make a distinction between consumers k and producers ℓ with N^k_{it} the individual demanding to be housed and N^ℓ_{jt} the developer producing or supplying the housing. The characteristics of the site or cell under consideration for the production of housing is defined as A^a_{zt} where z is a different location but all the locations i, j, z define the cells in the system where consumers and producers carry out their activities. The sequence of actions in any one time period can be orchestrated as follows: first a producer examines all the sites in question which in terms of each site can be represented by $N^\ell_{jt} \leftrightarrow A^a_{jt}$. The decision to produce a house in cell j is then made with respect to the attributes of j but also the potential demand for site j which might be based on previous demand at that site N^k_{jt-1}. The decision is made and the house produced which alters the characteristics of the site A^a_{jt+1}. The production of the house at this site can be defined as a unit of development or level of development D_{jt+1} which a potential house buyer – consumer – will now react to. When the house has been developed, potential residents will examine its location and then decide to occupy it or not, that is $N^k_{jt+1} \rightarrow D_{jt+1}$ and if an evaluation threshold is crossed then the individual will occupy the house, that is the house will be occupied O_{jt+1}. Formally the consumer might evaluate a function which works out a new level of the attribute of the site A^a_{jt+1} which can be formalised as

$$A^a_{it+1} = \phi \sum_{j \in Z_I} O_{jt} + \theta \sum_{j \in Z_I} A^a_{jt} + \vartheta \sum_{j \in Z_I} D_{jt+1} + \varepsilon_{jt+1} \tag{2.17}$$

where the parameters ϕ, θ, ϑ determine the relative weighting and normalisation while the error term ε_{jt+1} is a way of introducing some noise or uncertainty into the locational choice. If the cell attribute value is now above a certain threshold Γ, then the house is occupied; if not it remains unoccupied and the systems move into the next time phase where the process begins once again. Then

$$O_{it+1} = \begin{cases} 1, & if \quad A_{it} = 0 \quad and \quad \sum_{j \in Z_I} A_{jt+1} \geq \Gamma \\ 0 & otherwise \end{cases} \tag{2.18}$$

In this way demand adjusts to supply and vice versa if the system is well specified. Of course this simple model could not be programmed from this formulation for there are other decisions that need to be made to make the process computable but this sketch suffices to show how demand and supply agents interact with their cell space environment to produce and then consume housing. Immediately it is clear that in such a model, although the rules are quite plausible, it is extremely difficult to collect data on such a decision-making process. Moreover at this level of disaggregation, there are many features of the development process that cry out for specification; for example, issues about housing finance and finance for land development, issues about distance from home to work and to other facilities, provision of budgets, life style issues, all crowd into such a model. In a sense, this is why ABMs are so hard to build and test because once this level of detail is broached, it is hard to control the aggregation in such a way as to produce testable propositions. It is worth noting that spatial interaction effects fall out of this model quite easily, thus connecting ABM directly to the LUTI models that we will deal with in the next and final section of this review. The gravitational model of trips can be specified in agent form as

$$T^{k\ell}_{ijt} = \frac{N^k_{it} N^\ell_{jt}}{d^2_{ij}} \tag{2.19}$$

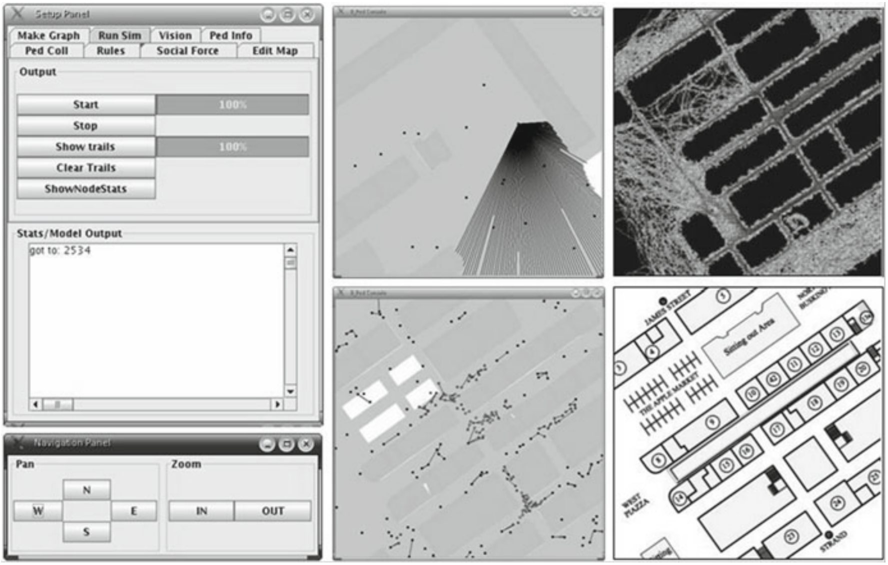

Fig. 2.4 Simulations from Ward's ABM pedestrian model of Covent Garden. *From top left to bottom right*: model control panel, visual lines of sight from a single pedestrian, flow intensity of all pedestrians, navigation panel, interactions of walkers, part of the map of the stalls in the shopping area

where we define k in terms of residence and o in terms of workplace. T_{ijt}^{ko} is the flow from i to j at the cross section and this can be lagged across time if so specified. We can also sum trips over i and j in terms of spatial interaction accounting and this serves to link these models to their aggregate equivalents. In fact, a sequence of locational decisions involving work and residence location in terms of an ABM might actually generate trips of these kinds through individual decisions rather than through this aggregate distance model. This does show that it is possible to begin to introduce social physics ideas into ABM with such connections to discrete choice modelling and microsimulation appearing extremely promising. Similar ideas of movement and spatial interaction are briefly introduced in Batty (2005) to which interested readers are referred.

The last thing we will do in this section is illustrate a typical example of ABM at the pedestrian movement level. In Fig. 2.4, we show a model built for the Covent Garden Market complex in central London by Ward (2007). This model is based on a simple social forces model in which agents have certain tasks to perform such as shopping and entertainment. They have two specific functions: to navigate in search of their goals which involves either purchasing entertainment or goods as efficiently as possible; and to move around the complex in more casual fashion. Most behaviour in this market is a combination of the casual and the formal but a key feature is the learning behaviour that must be built into navigating the space. This set of agents – walkers – is divided into different types dependent on purpose and how much exposure they have already had to the area. A substantial proportion of walkers are tourists. We do not have time to detail the model but it is clear that the nature of the problem imposes quite substantial differences between this application and others in terms of the composition of the agent population and the nature of the facilities in the complex. Again ABM is appropriate because of the rule-based behaviours in this kind of context and because navigation, obstacle avoidance and visibility calculations are important in simulating this type of mobility.

2.7 Land Use Transportation (LUTI) Models: Aggregate Behaviour in Spatial Equilibrium

Our last examples which are not part of the mainstream applications in this book except in the contribution below from Dearden and Wilson (2012), revert back to the origins of computational modelling in spatial systems which are in a rather different tradition from the new paradigms explored in the various contributions that follow. In fact, LUTI models have continued to be developed and strengthened and as we noted earlier, there has been a long quest to retain the advantages of simple aggregate models that can be calibrated against available data in contrast to the need for ever greater detail through disaggregation with the specification of temporal dynamics which move these models outside the equilibrium paradigm. In essence, when the time dimension is suppressed, the representation of environment and system is greatly simplified. The environment is simply indexed by space as A_i while the population is indexed as P_i^k where different activities k now refer to aggregates of populations covering employment, residential population, retail activity and so on. Just as CA models collapse population into the environment, LUTI models tend to collapse the environment into population: all the action in such models is, like ABM, focused on the aggregate with the environment in terms of cells, or zones as they are commonly called, being only relevant when various constraints on land availability and physical features of the space influence the simulation. In short, we can represent such models purely in terms of populations although distance and the attributes of space do occasionally enter the model framework from the environment.

We already have a simple form of LUTI model where spatial interactions are implicit in our development of CA in an earlier section. Equation (2.12) determines the function that converts a cell from one state into another, from undeveloped to developed for example, in terms of gravitational potential and we can write this more generally for any sector k as

$$N_i^k = \xi \sum_j A_j d_{ij}^{-\lambda^k} \tag{2.20}$$

where ξ is the relevant scaling constant, and λ^k is the friction of distance parameter for the gravitational potential. Equation (2.20) might apply to any sector although it is strongly physicalist in form being a function of only land (cell or zone) area A_j and geometric distance (or travel time/cost) d_{ij}. Without any obvious coupling, any LUTI model composed of several different population sectors such as types of residential housing, employment and so on would simply be a series of disconnected models. The most obvious way to connect sectors is to make each sector a function of all others in terms of composite accessibilities that might be written as

$$N_i^k = \xi \sum_{\ell} \sum_j N_j^{\ell} d_{ij}^{-\lambda^{\ell}} \tag{2.21}$$

where we note that the scaling constant is suitably adjusted and that the summation over sectors ℓ may or may not include the self-sector k, a decision that would depend on the precise model specification. In this sense then, the sectors are coupled through their relative spatial distributions.

In fact most LUTI models developed in the last 40 years have specified population as a function of explicit spatial interactions although the first models such as Lowry's (1964) were based on accessibility potentials as in Eqs. 2.20 and 2.21. Using an explicit spatial interaction model, then one of the simplest forms can be written as

$$N_i^k = \sum_j T_{ij}^k = \xi N_i^k \sum_{\ell} \sum_j N_j^{\ell} d_{ij}^{-\lambda^{\ell}} = \xi \sum_{\ell} \sum_j N_i^k N_j^{\ell} d_{ij}^{-\lambda^{\ell}} \tag{2.22}$$

We should note again that the summation is over sectors, that the scaling constant must be suitably adjusted and that there is immediate circularity in the model as the predicted variable appears on both sides of the equation. We do not have time here to dwell on this circularity but it can be resolved in many ways through model specification, balancing and iteration but in essence it reflects the reality of breaking into the spatial system at a cross section in time. In fact, in real applications, the use of appropriate balancing constraints resolves the issue (Batty 1976, 2008).

However the usual way of coupling such models is by assuming that the self-sector is not a function of the model or using another variable such as land area of the zone or cell A_i^k. Then substituting this for N_i^k in Eq. 2.22 and noting now that we will specify a two sector model where $k=1$ is the first sector and $\ell=2$, the second sector, then we can write equations for these two sectors as

$$\left.\begin{aligned} N_i^1 &= \sum_j T_{ij}^1 = \xi^1 A_i^1 \sum_j N_j^2 d_{ij}^{-\lambda^2} + X_i^1 \\ N_i^2 &= \sum_j T_{ij}^2 = \xi^2 A_i^2 \sum_j N_j^1 d_{ij}^{-\lambda^1} + X_i^2 \end{aligned}\right\} \tag{2.23}$$

Here we have extended the coupled model even further adding an exogenous input to each sector in the same manner that we did for the reaction-diffusion model earlier in Eq. 2.15 for the CA model. This structure is generic. It can be extended to many other sectors and it is at the basis of a whole class of LUTI models. For example the extended MEPLAN models developed by Echenique (2004) are based on this structure where there are explicit links to input-output models. The original extensions to the Lowry (1964) model were couched in these terms. The first equation in (2.23) was defined for total employment N_i^1 where X_i^1 was basic employment and the second equation was defined for total population N_i^2 where there was no exogenous population, that is $X_i^2 = 0, \forall i$. In short, this is the model structure suggested by Garin (1966) and Batty (1976).

This structure has been exploited in many ways. First it has been disaggregated to embrace many different classes of population with respect to residential population, housing and house types, industrial employment, retailing, commercial and related sectors such as education and health care. Second, relationships between the environment and population have been made in terms of land and density constraints, while third, the spatial interaction models have been extended in terms of utility maximising and route choice building on much more disaggregate individual-based models. In this sense, versions of LUTI models such as UrbanSim (Waddell 2002), ILUTE (Miller 2004) and DELTA (Simmonds 1999) begin to approach ABM illustrating that the line between modelling types and styles can become very blurred. Fourth, the models have been disaggregated to treat ever more zones and spatial units but of course, once these approach ABM, then locations are collapsed directly into individuals within the population and the notion of agents defined by zones has less relevance. Fifthly in many of these models, rule-based algorithms to sort out allocation as in CA models appear alongside more formal equation systems that determine locational distributions. Particularly where demand and supply are explicitly represented, then market clearing and the determination of prices that indicate how the model is balancing are often structured through rule-based mechanisms. As these models have extended their scope, then their formal parsimonious structures have been compromised. Their operation has become more ad hoc and pragmatic which appears to be a consequence of adding more and more detail and more and more sectors.

Dynamics has also been added to such models. At first, such static models were applied to forecast increments of change; that is the static model structure is used to assume that increments or decrements of change observed between two points in time such as $\Delta N_i = N_{it+1} - N_{it}$ become the focus of the prediction. In fact this is often simply a matter of scaling the equations to deal with net change. Many variants of this structure have been developed but there has not been much attention to breaking up the static structure into activities with different propensities to move. There are no models (to the authors knowledge, that is) where populations are divided into movers and stayers and these components dealt with in comparative static terms as different specifications of the equilibrium. Most extensions to dynamics have thus been ad hoc and in fact, there have been few developments of nonlinear dynamics of the kind described earlier involving catastrophes and bifurcations embedded directly into the structure of these models. There are examples where static models are embedded into dynamic frameworks but these are largely for pedagogic use and have never been fitted to real systems (see Dearden and Wilson 2012, this volume). The same might be said of Allen's (1997) work where embedding spatial interaction models into dynamics that lead to bifurcating behaviours in terms of locations are largely illustrative.

In terms of applications, the dominant model in urban and transport planning is still the LUTI model variant, largely because it deals explicitly with transport and housing in terms of their markets and the way they clear. Urban sprawl, for example, which CA models have attempted to simulate is highly dependent on transport and thus LUTI models are preferable as they deal directly with the drivers of sprawl. In North America, the dominant model was DRAM-EMPAL until quite recently when UrbanSim appears to have been more widely applied. Elsewhere MEPLAN and TRANUS have been developed, particularly in South America (Echenique 2004) while in Europe, there has been a mix of models. The focus is less on growth there and thus engagement with these kinds of formal model has been less intense although recently new waves of such models are being applied particularly in the London region. We will conclude our review with a brief summary of some of these models.

The MEPLAN structure developed as the LASER model has been used for 20 years for examining major transport proposals in the South East of England and this is now being supplemented with the LonLUTI model built on the back of the Delta model by Simmonds (1999). We have been developing residential location models as part of the integrated assessment of climate change, specifically flooding and pollution issues, in the Greater London region. This model is a standard structure of the kind presented here with a focus on heavy visualisation. A screen shot of typical output is shown at the top left of Fig. 2.5 where the focus on trip movements and their modal split is clear. It has now been extended using the structure in Eq. 2.23 where there are now three sectors being handled: population, retail and internal population-orientated employment with exogenous employment handled as a separate sector. This model is applied to the outer metropolitan area based on nearly 2,000 zones making the model quite large in spatial scale. The focus is still on fast and immediate visualisation and the current plan is for the model to be disaggregated and different modes to be added. The model is subject to capacity constraints in all sectors including trips and in this sense is quite comprehensive. We show a screen shot of the region in Fig. 2.5 at the top right and below, where it is clear that we are dealing with a complex polynucleated urban system based on a world city with some 14 million population. In contrast to the sort of pictures that we showed earlier for CA models in Ann Arbor (Fig. 2.3), it is clear that these models operate at a higher spatial scale although in the climate change applications, a CA-like model at 50 m grid square scale has been added to the integrated assessment to deal with populations at a much finer spatial scale than the LUTI configuration which is based on zones with an average of 10,000 persons. There is much more we could say about these models but interested readers are referred to this detail in Batty (2011), and Batty et al. (2011).

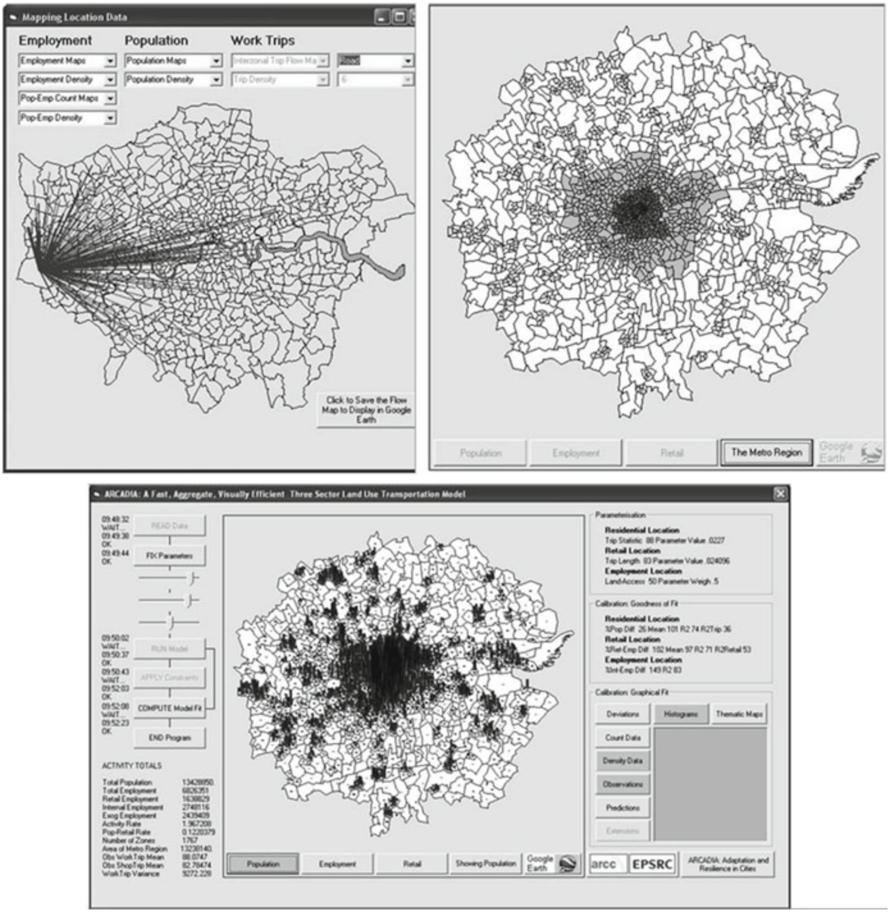

Fig. 2.5 LUTI models for the London region. *Top left*: Work trips from Heathrow in the Greater London residential location model: *Top right*: The nested model applications: *Bottom*: The interface to the 1767 zone London region model showing population histograms

2.8 Conclusions: Modelling Using Generic or Purpose-Built

The model framework developed in this chapter is designed so that readers might see the connections between a variety of model types at different levels of sectoral and temporal disaggregation. It is almost a non-sequitur that static cross-sectional models tend to be simpler to notate than dynamic models but what dynamic models add in terms of temporal richness, static models tend to compensate for in terms of sectoral feedback and strongly coupled activities. The framework we have introduced is certainly generic for the distinction between environment which is the space-time nexus and population which tends to be the driving force of all these models, is common to all spatial models of the kind developed in this book.

The level of aggregation although exceptionally important in terms of applications, is less important in terms of model structure. What we have not done here is dwell on methods of fitting different models within this framework to data and it is worth concluding with some remarks for this serves to polarise differences between the various models.

As the level of detail in terms of sectors, spatial-locational resolution, and temporal resolution increases, data demands generally increase and models become increasingly difficulty to validate in terms of being able to match all the model hypotheses – functions – to observed data. As temporal processes are added, this can become exceptionally difficult but even with cross-sectional static models, when we add mechanisms for coupling and for market clearing as is the case in many LUTI models, we face a severe problem of validation. Many processes in these models cannot be observed and in principle some of these may simply be unobservable. Thus the model-builder faces problems of convincing client and stakeholder groups, which may comprise other scientists, of the veracity of their simulations. This tends to force modelling back to the traditional canons of scientific inquiry where parsimonious and simple models are the main goal of scientific explanation. Occam's razor may still be the ultimate quest but in many social systems, evident complexity is so great that plausibility rather than validity may be the real quest. This tension is felt very heavily throughout this book although it is broached only gently by many of the authors who are clearly conscious of the weight of scientific credibility that these new approaches to social systems impose.

In fact cutting across this dilemma is the notion that as we improve our understanding of spatial systems, we might be able to generalise models to the point where generic software becomes dominant. In fact, quite the opposite is happening. As we learn more we consider each problem context to be more individualistic where the model has to be specifically tailored to the task in hand. Software engineers have in fact sought to develop ever more generic packages but these are often frameworks which guide the modeller rather than establish complete frameworks for the development of a specific model. Most general frameworks for ABM for example such as RePast and Netlogo, even MATLAB and Mathematica, do not extend to the point where very detailed spatial models can be built within their structures. LUTI models are a case in point. 30 years ago when spreadsheets were first developed it was perfectly possible to develop pedagogic versions of such models using that software but no real application would ever fit into such structures. To date, there is no standard software for such models. In fact herein is the dilemma. Most serious applications rather than proofs of concept or pedagogic demonstrations require specific software applications. Insofar as generic software can be used, this provides many of the basic routines but these still have to be assembled in situ by skilled programmers, notwithstanding the fact that downstream applications may emerge which are generic. But then such applications tend to be pedagogic, showing what has been done and any new application requires purpose-built software development. It is hard to see this situation changing in that the problems that we need to engage with always seem to outstrip previous applications and software already developed for these. The various contributions on this book clearly demonstrate this point.

References

Allen, P. M. (1997). *Cities and regions as self-organizing systems: Models of complexity*. London: Taylor and Francis.

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Dordrecht: Kluwer.

Batty, M. (1976). *Urban modelling: Algorithms, calibrations, predictions*. Cambridge: Cambridge University Press.

Batty, M. (2003). Agent-based pedestrian models. In P. A. Longley & M. Batty (Eds.), *Advanced spatial analysis: The CASA book of GIS* (pp. 81–106). Redlands: ESRI Press.

Batty, M. (2005). *Cities and complexity: Understanding cities with cellular automata, agent-based models, and fractals*. Cambridge, MA: The MIT Press.

Batty, M. (2007). Model cities. *Town Planning Review, 78*(2), 125–151.

Batty, M. (2008). Fifty years of urban modelling: Macro statics to micro dynamics. In S. Albevrio, D. Andrey, P. Giordano, & A. Vancheri (Eds.), *The dynamics of complex urban systems: An interdisciplinary approach* (pp. 1–20). Heidelberg: Physica.

Batty, M. (2009). Urban modeling. In R. Kitchin & N. Thrift (Eds.), *International encyclopaedia of human geography* (Vol. 12, pp. 51–58). Oxford: Elsevier.

Batty, M. (2011). Visually-driven urban simulation: Exploring fast and slow change in residential location. CASA working paper 164, University College London. Available at http://www.casa.ucl.ac.uk/working_papers/paper164.pdf

Batty, M., & Xie, Y. (2005). Urban growth using cellular automata models. In D. J. Maguire, M. Batty, & M. F. Goodchild (Eds.), *GIS, spatial analysis, and modeling* (pp. 151–172). Redlands: ESRI Press.

Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environments and Urban Systems, 23*, 205–233.

Batty, M., Vargas-Ruiz, C., Smith, D., Serras, J., Reades, J., & Johansson, A. (2011). Visually-intelligible land use transportation models for the rapid assessment of urban futures. CASA working paper 163, University College London. Available at http://www.casa.ucl.ac.uk/working_papers/paper163.pdf

Birkin, M., & Wu, B. (2012), A review of microsimulation and hybrid agent-based approaches. In A. J. Heppenstall, A. T. Crooks, L. M. See & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 51–68). Dordrecht: Springer.

Crooks, A. T., & Heppenstall, A. J. (2011). Introduction to agent-based modelling. In A. J. Heppenstall, A.T. Crooks, L. M. See & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 85–105). Dordrecht: Springer.

Dearden, J. & Wilson, A. (2012). The relationship of dynamic entropy maximising and agent-based approaches in urban modelling. In A. J. Heppenstall, A. T. Crooks, L. M. See & M. Batty (Eds.), *Agent-based models of Geographical Systems* (pp. 705–720). Dordrecht: Springer.

Echenique, M. H. (2004). Econometric models of land use and transportation. In D. A. Hensher, K. J. Button, K.E. Haynes, & P. R. Stopher (Eds.), *Handbook of transport geography and spatial systems* (pp. 185–202). Amsterdam: Pergamon.

Forrester, J. W. (1969). *Urban dynamics*. Cambridge: The MIT Press.

Garin, R. A. (1966). A matrix formulation of the Lowry model for intrametropolitan activity location. *Journal of the American Institute of Planners, 32*, 361–364.

Haase, D., & Schwarz, N. (2009). Simulation models on human–nature interactions in urban landscapes: A review including spatial economics, system dynamics, cellular automata and agent-based approaches. *Living Reviews in Landscape Research*, *3*, 5–35. Available at http://www.livingreviews.org/lrlr-2009-2

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners, 25*, 73–76.

Hunt, J. D., Kriger, D. S., & Miller, E. J. (2005). Current operational urban land-use-transport modelling frameworks: A review. *Transport Reviews, 25*(3), 329–376.

Iacono, M., Levinson, D., & El-Geneidy, A. (2008). Models of transportation and land use change: A guide to the territory. *Journal of Planning Literature, 22*, 323–340.

Iltanen, S. (2011). Cellular automata in urban spatial modelling. In A. J. Heppenstall, A. T Crooks, L. M. See & M. Batty (Eds.), *Agent-based models of Geographical Systems* (pp. 69–84). Dordrecht: Springer.

Liu, Y. (2008). *Modelling urban development with geographical information systems and cellular automata*. Boca Raton: CRC Press.

Lowry, I.S. (1964). *A model of metropolis*. Memorandum RM-4035-RC. Santa Monica: The Rand Corporation.

Lowry, I.S. (1965). A short course in model design. *Journal of the American Institute of Planners, 31*, 158–165.

Miller, E. J. (2004). Integrated land use transport model requirements. In K. J. Button, K. E. Haynes, P. Stopher, & D. A. Hensher (Eds.), *Handbook of transport geography and spatial systems* (Handbooks in Transport, Vol. 5, pp. 147–162). New York: Elsevier Science.

Plummer, A.V. (2007). The Chicago area transportation study: Creating the first plan (1955–1962): A narrative. Available at http://www.surveyarchive.org/Chicago/cats_1954-62.pdf

Simmonds, D.C. (1999). The design of the DELTA land-use modelling package. *Environment and Planning B: Planning and Design, 26*, 665–684.

Southworth, F. (1995). *A technical review of urban land use-transportation models as tools for evaluating vehicle travel reduction strategies, ORNL-6881*. Washington DC: Center for Transportation Analysis Energy Division.

Waddell, P. (2002). UrbanSim: Modelling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association, 68*, 297–314.

Ward, J. (2007). *Urban movement: Models of pedestrian activity*. Unpublished PhD thesis, CASA, University College London.

Wegener, M. (2005). Urban land use transportation models. In D. J. Maguire, M. F. Goodchild, & M. Batty (Eds.), *GIS, spatial analysis, and modeling* (pp. 203–220). Redlands: ESRI Press.

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.