Cellular Automata in Urban Simulation: A Progress Report

Michael Batty, David O' Sullivan
& Paul Torrens
University College London, University of Auckland,
University of Maryland

m.batty@ucl.ac.uk d.osullivan@auckland.ac.nz
torrens@geosimulation.com

You can find this presentation as a pdf at

http://www.complexcity.info/

As the latest post at the top level

Themes

- Defining Urban Models, Defining Cellular Automata
- Urban Processes the Rules for Transition, and the Definition of Neighbourhoods
- The Physical Bias
- Simple and Complex
- Scale and Multilevel Modelling
- Verification, Calibration, Validation
- Merging CA into Other Models: ABM and LUTI
- Where to Next

The editorial 11 years ago that set us thinking that a progress report on CA was needed and this was the conference. We will write it up properly after the meeting

Environment and Planning B: Planning and Design 2001, volume 28, pages 163 – 168

DOI:10.1068/b2802ed

Editorial

Cellular automata and urban simulation: where do we go from here?

Introduction

> "He figured out that mathematicians, unlike carpenters, only needed to have one tool in their toolbox, if it were the right sort of tool. Turing realized that it should be possible to build a meta-machine that could be reconfigured in such a way that it would do any task you could conceivably do with information. It would be a protean device that could turn into any tool you could ever need. Like a pipe organ changing into a different instrument every time you hit a preset button."
>
> Stephenson (1999, page 197)

For some time now, cellular automata (CA) have been in popular use for urban simulation. Many of the most significant contributions to this new, emerging field have featured in *Environment and Planning B*, with a special issue devoted to the topic

in truly realizing the potential of CA and related modeling approaches. Hopefully, researchers in the field will feel encouraged to report future developments in these areas in the pages of *Environment and Planning B: Planning and Design*.

Paul M Torrens, David O'Sullivan

References

Batty M, Couclelis H, Eichen M, 1997, "Editorial: urban systems as cellular automata" *Environment and Planning B: Planning and Design* **24** 159 – 164

Benenson I, 1999, "Modelling population dynamics in the city: from a regional to a multi-agent approach" *Discrete Dynamics in Nature and Society* 3 149 – 170

Defining Urban Models, Defining Cellular Automata

These are computer models that explain and predict urban outcome such as land uses and activities therein

They are built on mathematical structures in the broadest sense of the word that range from formal and analytical equation systems to rule-based logics

They are applicable at urban scales from the metro-region to the local district level but rarely outside these scales

They either predict spatial outcomes at a cross section in time or outcomes that change in time: these types are called static or dynamic

Now cellular automata models are much more generic in that they are defined prior to their application to urban and regional systems, to land use cover problems etc

They were stated first by Stanislav Ulam at Los Alamos in 1944 where he worked with von Neumann on the Manhattan project.

These were picked up by both von Neumann and his RA Arthur Burkes who sort of popularised them in the 1960s at Michigan and they were then widely popularised by Martin Gardner in his 1971 Scientific American column where he talked about Conway's Game of Life. Note Burkes and Tobler were at Michigan and I think they knew each other. Small world

Now we can define two types of CA cellular automata models, first ***strict CA*** and second ***non strict CA*** where the former are based on completely local neighbourhoods and the latter on non-local – I will define these in a minute but the latter are often called ***cell space models***

They depend on five features and for strict CA these are ***universal properties –*** they apply to all cells

1. They operate on a set of cells which is usually some form of regular tessellation

2. Around each cell there is a neighbourhood of cells which is only those cells which are local – ie adjacent

3. Each cell is defined by state and the simplest states are off or on, undeveloped or developed

4. Around each cell there is a neighbourhood of cells which is only those cells which are local – i.e. adjacent

5. Cells change state at each time period by examining those cells in the neighbourhood of each cell and defining these in terms of some transition rule which is also uniform and universal across the space

This leads to emergence – in this sense, the local pattern is replicated at larger scales

The game of life is one of the simplest cellular automata -

It is composed of these simple rules – births, deaths and continued existence, a cell is spawned if there are 3 cells alive in its neighbourhood, anything less or more the cell dies, and if the cell is alive, it stays alive of there are 3 cells alive in its neighbourhood, I think.

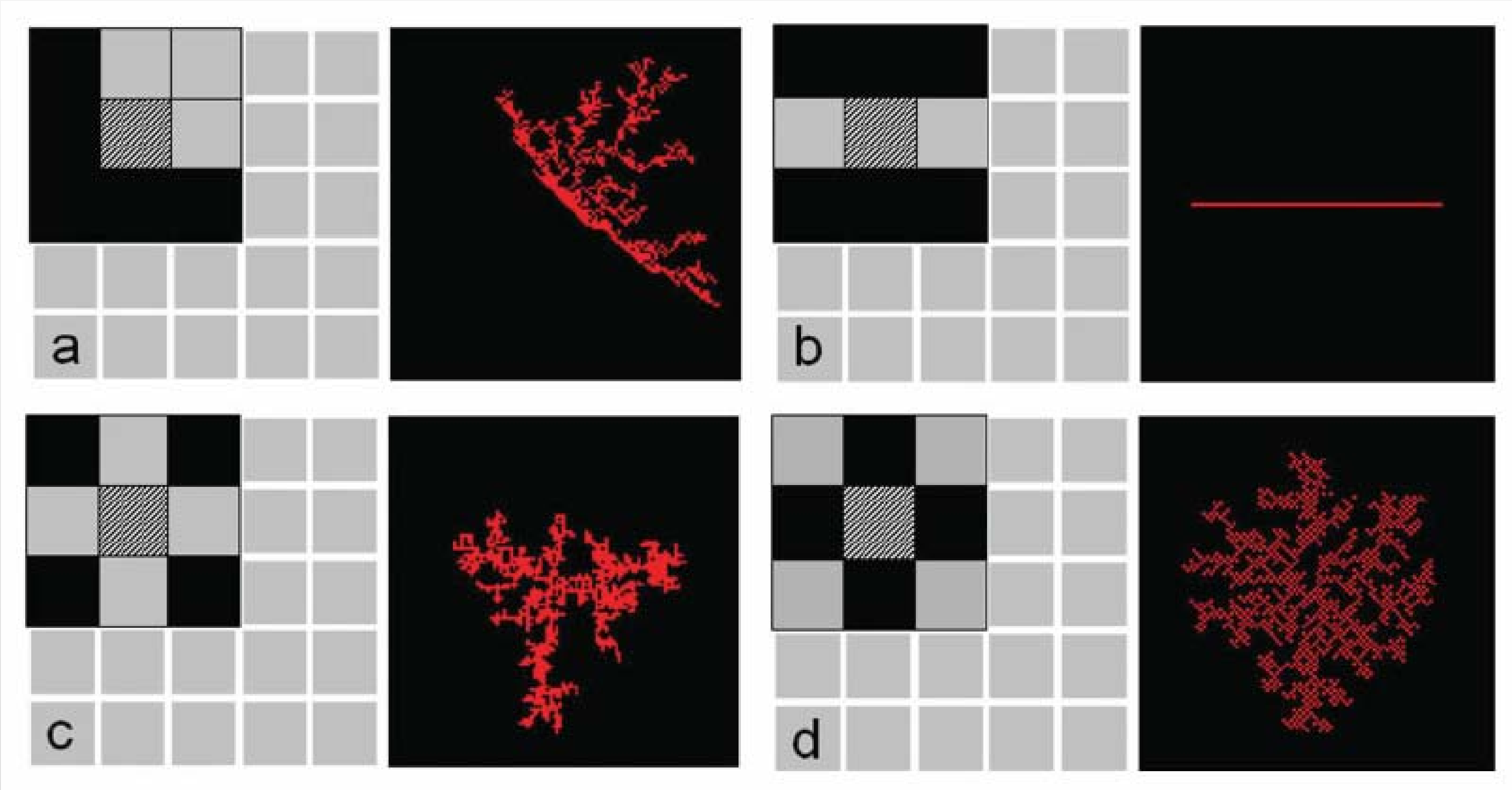
The formulation of a CA is as follows

$$A_{it+1} = \beta \sum_{j \in Z_I} f(A_{jt})$$

some rule based logic effects this transition function

The transition rules encoded in a linear equation, whose parameters are chosen using some form of optimisation along the lines we have been discussing at the conference – regression, pattern swarm, ANN and so on

These kinds of CA models lead to emergence and I show some simple examples below from a paper I wrote in AD four years ago.



Of course there are many variants of these kinds of model and these are catalogued by the modern popularise Stephen Wolfram in his collected papers and also in his book ***A New Kind of Science***

Frankly I have never understood his notion that the world is one large CA and everything can be seen in this way, notwithstanding that CA is a universal machine or can be likened to one which Turing supposed and was demonstrated best by Conway

I suppose then that in generic terms we can fashion the world as a dynamic system using CA as an exemplar but I am not sure this is that helpful. No matter. Let me move on.

Urban Processes the Rules for Transition, and the Definition of Neighbourhoods

This is where the rubber hits the road. CA depend on an action at a distance principle that is entirely local – it is an excellent model for diffusion where one cell influences an adjacent cell but not one that is non adjacent – the influence is only on its nearest neighbours

This means that there is no proper action at a distance and we know in urban systems that there must be.

This is where the rule in CA about neighbourhood size is relaxed. This essentially destroys emergence.

Or rather, it destroys the sense that we can get emergence easily. The modularity is still there but neighbourhoods overlap in complex ways from which the emergent pattern can be hard to find

However it is other things relating to the way the pattern is transformed through transition rules that really destroys. In strict CA it is the morphology of the cells in the neighbourhood that determines emergence. Once we add attributes to the cells then these are no longer universal in form and this definitely destroys emergence.

It may still happen but it is impossible to see. This may be no bad thing but it does project most urban CA models away from their origins in complexity theory

Of course the real issue here and this is what O'Sullivan and Torrens said in their editorial 11 years ago was that – is that the transition rules are where these models connect up to how the urban system functions.

And so far these have been pretty simple minded processes – hardly processes per se but changes of state occasioned largely by accessibility and land suitability. There are very few urban CA models that introduce positive feedbacks in terms of the way the state of the system at time t influences in a positive way the state at time t+1 – of course t+1 is built on t but to really produce this kind of feedback one would need to change the transition rules as the system evolved – and this would really throw the model away from CA.

The Physical Bias

We noted yesterday that CA are rooted in physical space – usually 2 dimensional in terms of cities, perhaps 3 but cities mainly vary in 2-d not 3-d – and distance is all important in terms of their functioning.

This makes it hard to build in spatial things – not impossible but because of the physical adjacency criterion, the functions tend to be articulated physically. It might be possible to break out of this mould in urban applications but these have been rare. Most of the models at this symposium are based on 2-d neighbourhoods which are physically linked by distance.

This means that the network rather than the flow system is intrinsic to urban CA. In turn this means that things such as travel time, even travel cost which lie at the heart of LUTI models, for example, are rarely considered.

In fact transport is one of the major features that is lacking in CA models. A corollary to this is that transport depends on activities and most CA models do not operate at the level of activities. In the spatial interaction paradigm, activities generate flows and one needs this to be able to recast CA to embrace this.

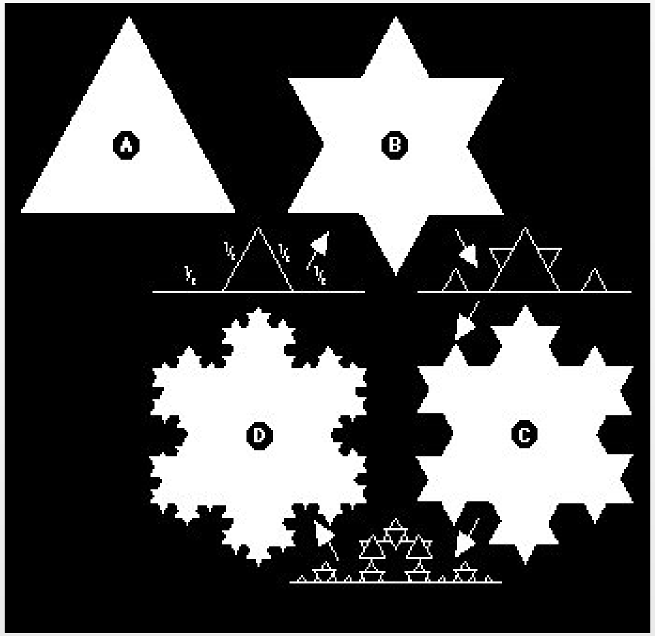
It is possible to do this but no one really has. Mostly CA in so far as they consider transport, loosely couple the model to other transport and related models

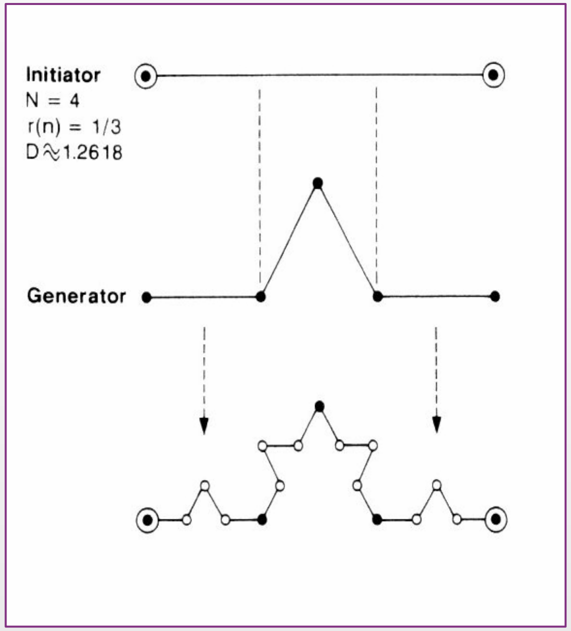
Simple and Complex

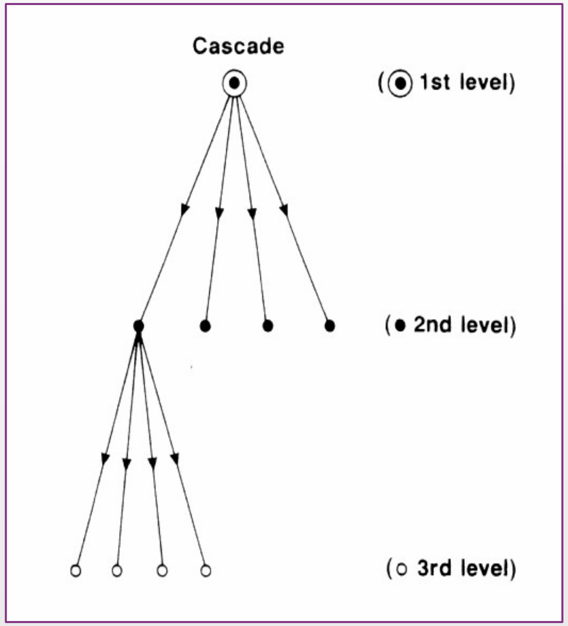
The great feature of CA is its comparative simplicity – CA models are easy to state and understand – perhaps too easy?

CA models build on the simple bottom up principle of modularity that generates emergent pattern from the operation of simple rules – or at least strict CA does this.

In fact, CA is essentially the algorithm that generates a two dimensional fractal and it can be specified in very simple terms as the recursive application of a set transition rules to an initiator – a basic morphology and the reiteration/recursion of these rules to the evolving structure. For example:







A Way of Generating Such Complexity

This is how we generate self-similar objects across many scales this developing fractal objects. This is the famous Koch curve. Let me explain how it works. All the paradoxes of fractals are contained within this curve.

This modularity is still the cornerstone of CA dynamics notwithstanding that the emergent structure is no longer clear in most applications.

This however does still provide the prospect for surprising emergent behaviour. Again one of the hallmarks of complexity theory but there has not been much on this so far in urban CA.

In short although in principle CA models – even non strict CA or cell space models can generate surprise given their structure, there has not been much on this in terms of what is simulated. Jenni Partanen's paper was the closest we saw to this in this meeting

Scale and Multilevel Modelling

I am not going to say much about scale except that CA models are intrinsically scaling – we do not know at what scale they work as their processes scale up.

If an object is scaling it has the ***same form*** across many ***orders of magnitude***. And this means that it is very hard to say what is the best scale.

There has been some discussion in the meeting about the notion of the grid and its match to land parcels but no serious testing of this issue for if we move away from the grid to irregular tessellations, then we sort of abandon one of the major simplifying principles of CA.

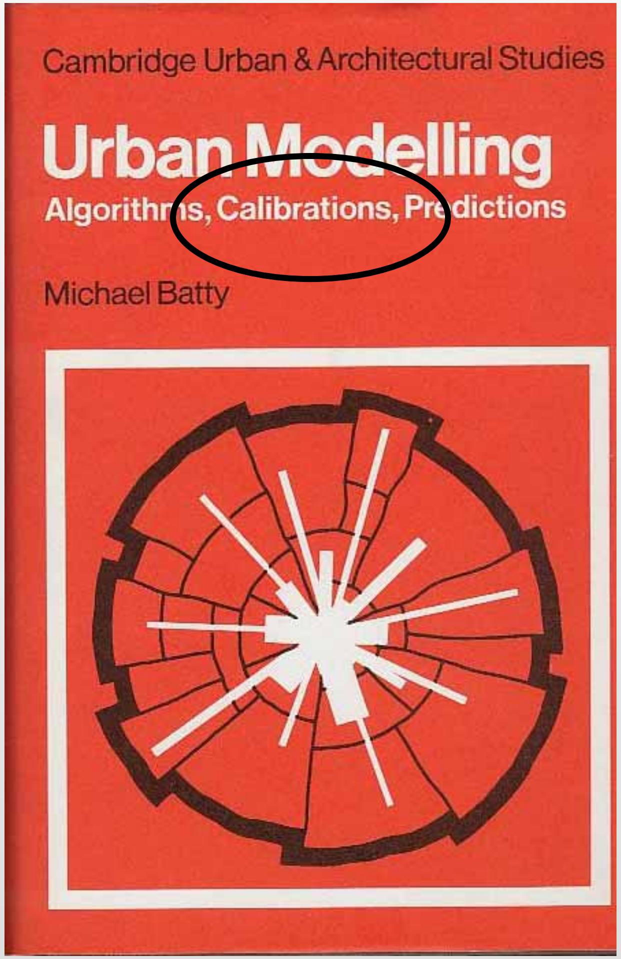
In terms of dealing with different scales – multilevel modelling if you like – I tend to think that we have made very little progress in this area in urban and spatial systems.

Nice idea that processes work at different levels and that we need to somehow dovetail all these together in some workable fashion but generally this is only made operational by adding actions at one level and displaying them at another – for example, in enabling the CA to work at say a grid level and then examining its outputs at supergrid level and so on up the hierarchy. I don't think we have seen much that operates up and down the hierarchy in some ingenious way for this would imply different transitions at different scales.

Verification, Calibration, Validation

Ok let me come to one of the major issues in CA. But before I do let me define what I think is now meant by this triumvirate of terms pertaining to making the model reproduce in some way what we observe.

In my youth, we simply called all of this calibration after the engineering term, or estimation after the statistical term. In fact I used the title calibration to cover all these things in my book ***Urban Modelling***



I think ***Verification*** means making sure the model works – ie runs – we used to say in the old days that the model had run but it didn't work – it the program ran but it produced incorrect results. Verification is making sure this doesn't happen

Calibration means fine tuning the parameters to reproduce some known statistical characteristics in the observed or synthetic data – this is in effect estimation but it focuses very often on the mechanisms used to do this in highly non linear and compute intensive systems. ***Validation*** is measuring how good the fit is – it may be closely related to calibration for calibration may be judged on the basis of validation bit often is based on additional measure of the goodness of fit to those used to calibrate the model.

The big question in CA modelling is should we expect them to be validated – the argument runs as follows:

CA are pedagogic devices to show emergence – their advantages lie in their simplicity – therefore we should not expect them to be simulation devices for systems that we know are more complex than the CA which is being used.

I have tended to ascribe to this view. DUEM is more in this tradition than SLEUTH or GeoDynamica and Explorer and Metronamica. I now tend to be more relaxed about this but still am quite critical particularly of SLEUTH for pretty blind applications that really hardly relate to policy at all which they claim.

I think the main issue is – what is being validated and if the processes are rather simple minded, then …

Also the discussion turns on whether or not it is the processes of change that are being validated or just some proxy for these at a crude level. For example if we think of CA as being a mirror of the land development process, then this is a well documented process of supply by developers and demand by consumers of land use etc. Most CA models do not model this and certainly don't test it – they proxy it by some sort of overall index that sets off the transition, an index that is a sort of amalgam of demand and supply and which does not change through the model process – which of course land development behaviour does.

In fact, I think much much more could be done by first embedding explicit development processes in CA models, and then testing these but it needs considerable model development and this has not happened.

Two additional things here – First the focus on CA has been very much guided by traditions in GIS and remote sensing, but not by urban theory as Roger pointed out.

Second, policy. There is little question that what policy makers want are numbers so they can fix a plan and CA is more qualitative than this. Nesting CA as part of a suite of models is the obvious way forward in the short term. Developing them to embrace activities is one of the longer term goals in my view. And it can be done, easily.

Let me dwell a little longer on policy before I begin to round off this progress report.

I don’t think that CA models are very good at dealing with what is the predominant change in urban systems which is largely redevelopment, regeneration, invasion and succession of existing land and buildings. In UK cities most change –probably 95% -- is this kind of change, not new growth.

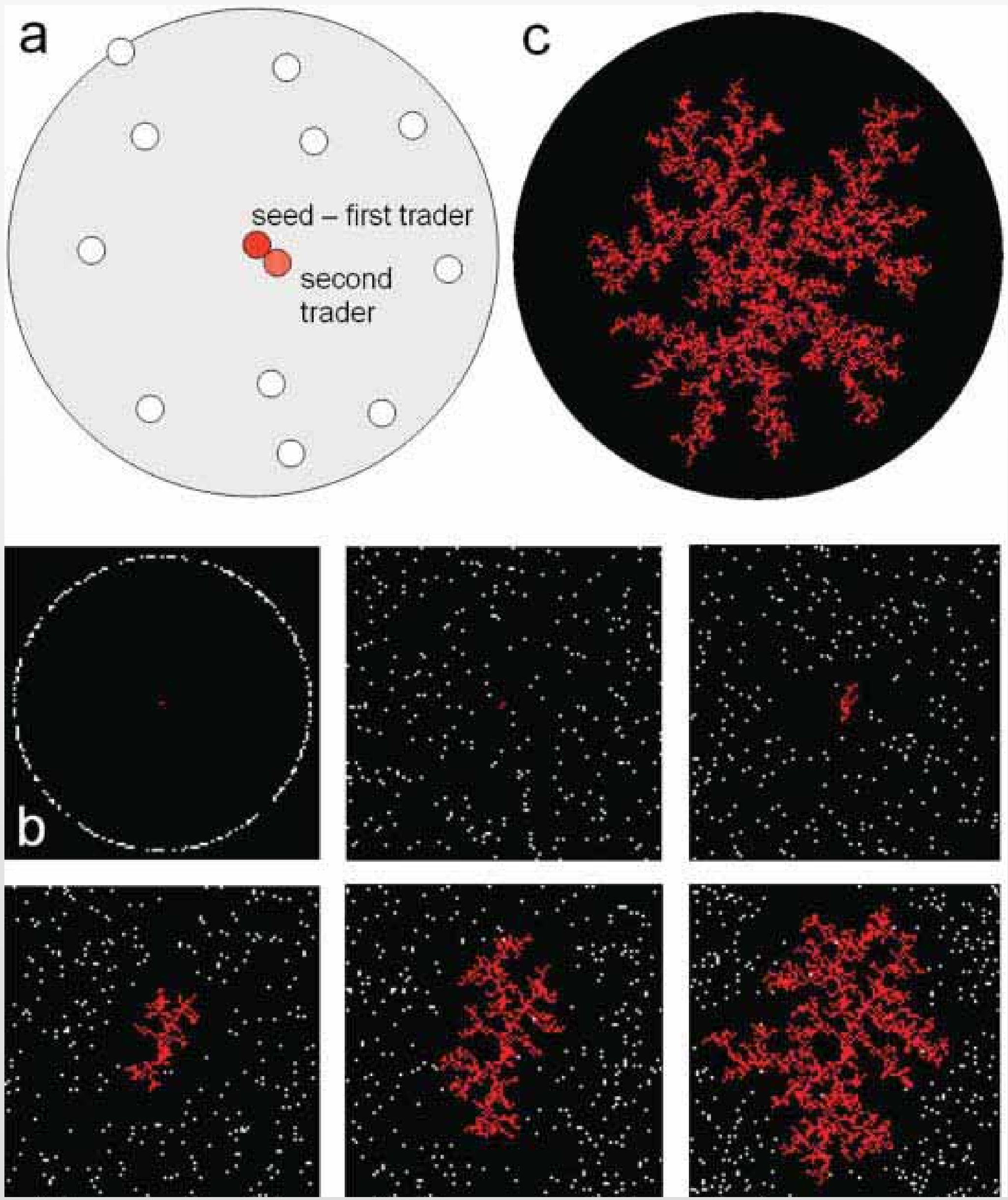
CA models barely touch this – they operate best in situation of rapid growth – leading to sprawl like patterns. They could be fashioned in terms of existing development but with exceptions – David O’Sullivan’s thesis. They should be and this could give CA models a new life

Merging CA into Other Models: ABM and LUTI

But perhaps the most promising development lie in linking CA models to ABM and even to LUTI models. Let me finish with two examples from my own work.

First as I have implied here, CA algorithms lie at the heart of generating 2-d fractal patterns, and an obvious extension is to use the cellular landscape as the canvas on which to simulate how objects can move between cells of this landscape. The best example is the diffusion limited aggregation model and let me show how this works.

Essentially we have a grid of cells between which actors or agents move to be near the growing city.



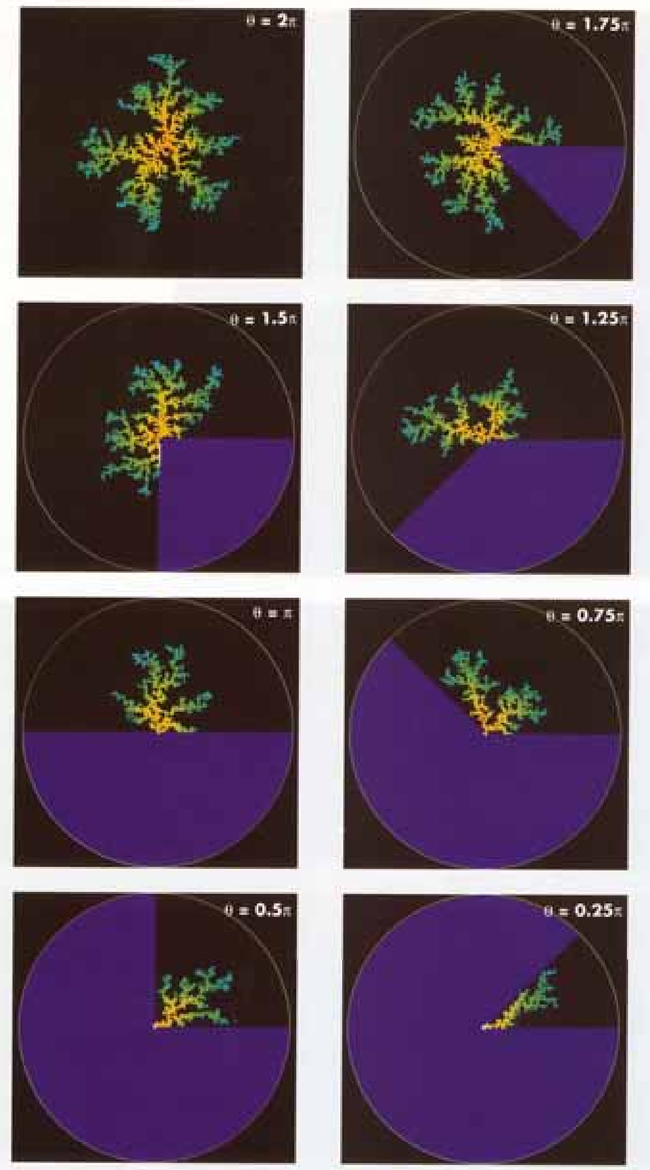
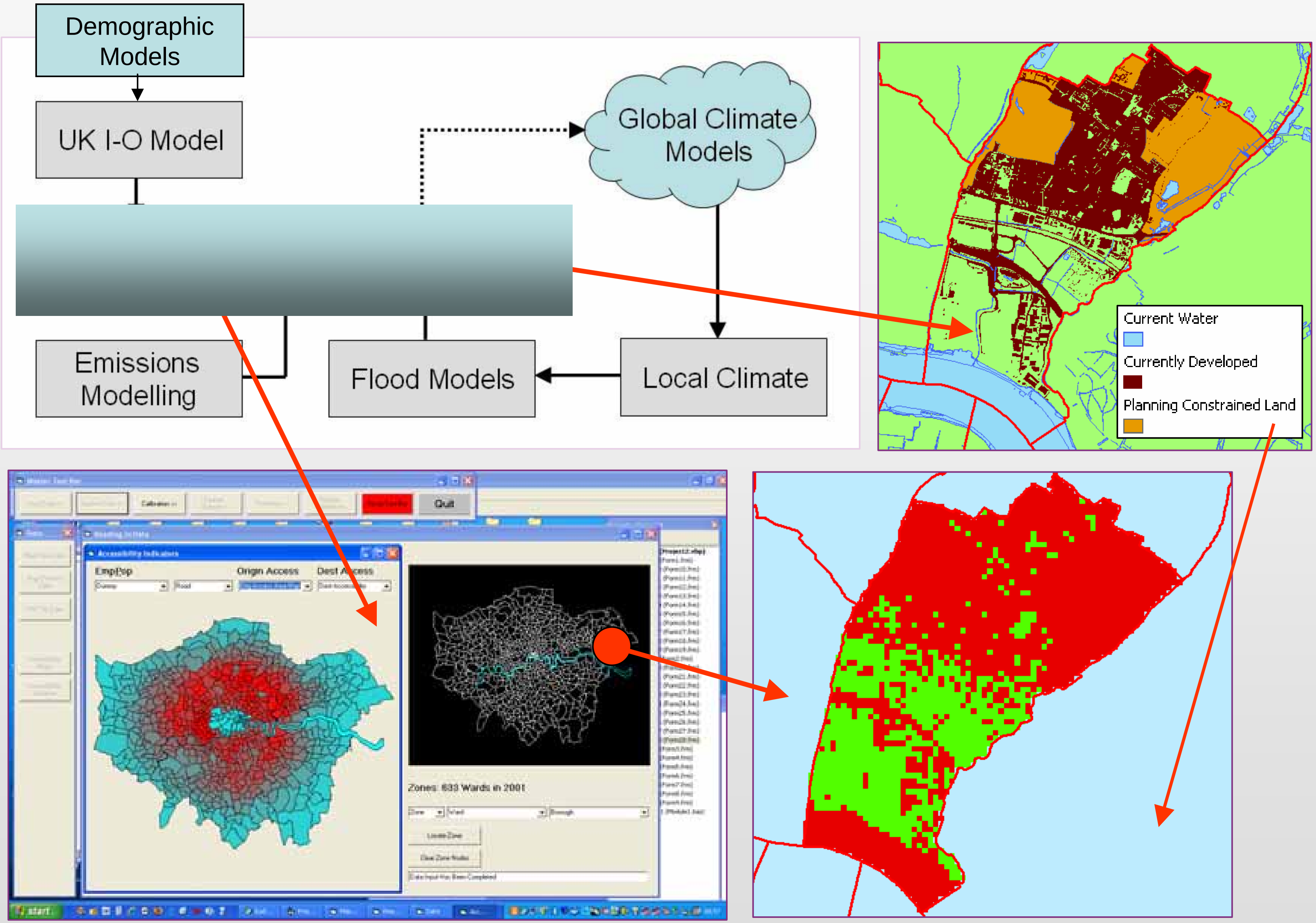


Plate 8.2 Physically Constrained DBM Simulations

Relations to suites of models – we have developed an integrated assessment of climate change in London which embeds a LUTI models followed by a CA like model as we downscale from wards which the LUTI models works at to 50 metre grid squares where the CA models works at

In fact the entire suite scales from national to regional to metropolitan to site specific to even finer levels where the hydrological flooding models work.

Here is the flow chart and some pictures of what goes on


Demographic Models
UK I-O Model
Global Climate Models
Emissions Modelling
Flood Models
Local Climate
Current Water
Currently Developed
Planning Constrained Land
Quit
EmpPop
Origin Access
Zones: 633 Wards in 2001

Where to Next

- Richer theory is needed and it can be done
- Nesting CA models in suites of related models
- Building real urban processes inside them
- Handling the neighbourhood-field issue much more satisfactorily
- Handling the irregular parcel issue and the grid cell much more cleverly
- Simpler schemes for parameterisation and calibration

You can find this presentation as a pdf at

http://www.complexcity.info/