Network + cellular automaton → land values and much more

Claes Andersson
Chalmers University of Technology

The background of the work

Dissatisfaction with urban growth models based on CA (or well...), including those developed by myself and co-workers (Also Batty, Xie, White, Engelen, Clarke etc).

Batty originally acted on the insight that CAs could capture the distributed and local-biased dynamics of urban growth.

However, no matter how hard I tried, the CA:s could never reproduce any of the scaling relationships that are typical of urban clusters and systems.

So…

We did some thinking, and the result was a hybrid between a classical cellular model and a scale-free network.

We then obtained a land value map of Sweden (100X100 meter cells), which at the time was remarkably complete and detailed.

The result...

...was a model that actually reproduced a good number of statistiscal features of urban systems that were known, along with a number that weren't known.

Most importantly, it did so with very few parameters and assumptions (as urban growth modeling goes at least) and on many levels of observation at the same time.

The model

Land values are typically seen as proportional to the ability of the tenant to pay rent, which it typically does due to income streams.

- Income streams, fundamentally, involve trade (economic exchange).
- Trade happens over distance.
- The best guess, with no extra knowledge, of the growth rate of an activity, is preferential growth.

So...

Network

...therefore, the idea that a preferential attachment network would be a suitable model was not far fetched: preferential growth with the potential to bias attachement with distance.

The degree of the nodes, then are taken to represent a measure of economic activity, which should be observable (on average) as land value.

The scheme...

All nodes exist from start and are arranged in a 2D grid. The nodes are the cells from the CA perspective.

In order to handle the introduction of new development, a CA-style scheme had to be used.

As a CA, cells can take 3 states:

1. Undeveloped
2. Perimeter
3. Developed

Network growth

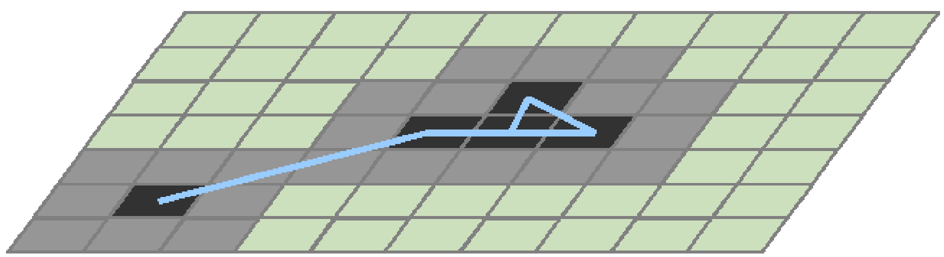
Growth takes place in two steps: primary and secondary growth.

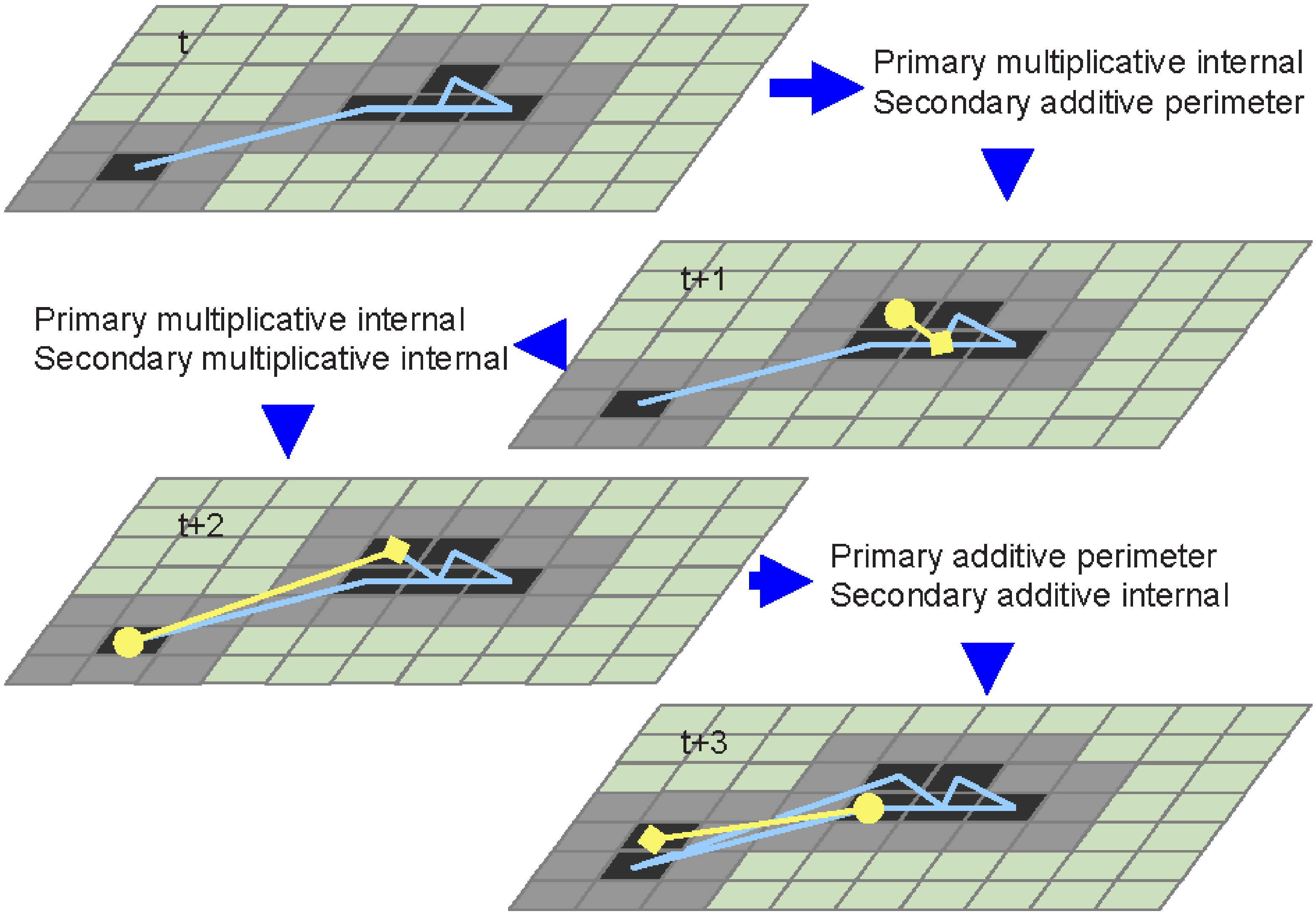
These can be either multiplicative or additive.

Additive growth can mean many things, but it is really all reasons why anyone should build something from scratch.

Simple local infrastructure availability model

External	Perimeter	Internal
Minimum availability of Infrastructure: low “a_i”	Medium availability of Infrastructure: medium “a_i”	Maximum availability of Infrastructure: high “a_i”
• Low additive growth • No multiplicative growth	• Medium additive growth • No multiplicative growth	• High additive growth • Multiplicative growth




t
Primary multiplicative internal
Secondary additive perimeter
t+1
Primary multiplicative internal
Secondary multiplicative internal
t+2
Primary additive perimeter
Secondary additive internal
t+3

Primary growth

$$\Pi_i^{1,mul} = q_2 \frac{x_i}{\sum_i x_j},$$

$$\Pi_i^{1,add} = q_1 \frac{a_i}{\sum_j a_j}.$$

Index *i* denotes a cell/node. *q1+q2=1 controls the relative amount of additive and multiplicative growth.*

The *a parameters denotes availability of infrastructure in cells. For Internal, a_i=1, for Perimeter a_i=b (a parameter) and for External* $a_i = b\epsilon \frac{n_t^{(P)}}{n_t^{(E)}}$

$$\Pi_i^{1,add} = q_1 \frac{\delta_i^{(D)} + b\delta_i^{(P)} + b\epsilon \frac{n_t^{(P)}}{n_t^{(E)}} \delta_i^{(E)}}{\sum_j \left(\delta_j^{(D)} + b\delta_j^{(P)} + b\epsilon \frac{n_t^{(P)}}{n_t^{(E)}} \delta_j^{(E)} \right)} = q_1 \frac{\delta_i^{(D)} + b\delta_i^{(P)} + b\epsilon \frac{n_t^{(P)}}{n_t^{(E)}} \delta_i^{(E)}}{n_t^{(D)} + b(1+\epsilon) n_t^{(P)}}$$

Secondary growth

$$\Pi_{ij}^{2,mul} = q_2 \frac{D_{ij} x_i}{\sum_k D_{kj} x_k}$$

$$\Pi_{ij}^{2,add} = q_1 \frac{D_{ij} a_i}{\sum_k D_{kj} a_k}$$

$$\Pi_{ij}^{2,add} = q_1 \frac{D_{ij} \left(\delta_i^{(D)} + b\delta_i^{(P)} + b\epsilon \frac{n_t^{(P)}}{n_t^{(E)}} \delta_i^{(E)} \right)}{\sum_k D_{kj} \left(\delta_k^{(D)} + b\delta_k^{(P)} + b\epsilon \frac{n_t^{(P)}}{n_t^{(E)}} \delta_k^{(E)} \right)}$$

$$D_{ij} = (1 + cd(i,j))^{-\alpha}$$

$d(i,j)$ is the Euclidean distance between sites i and j.

Let us note...

...that power law distributed land values are predicted without spatiality.

Hence, the question as far as they are concerned is rather: do they <u>survive</u> spatiality.

It turns out that they do...

Let us also note...

That most parameters could be approximated directly from the empirical data and did not have to be just tuned by hand. Please see paper for details.

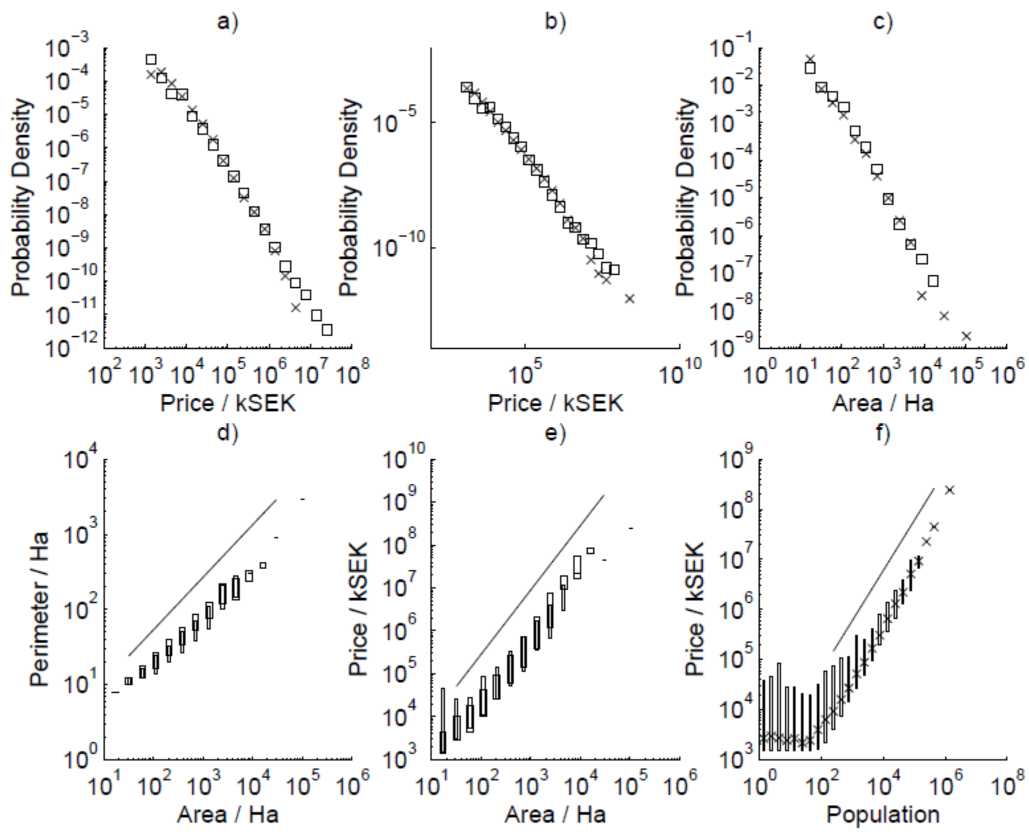
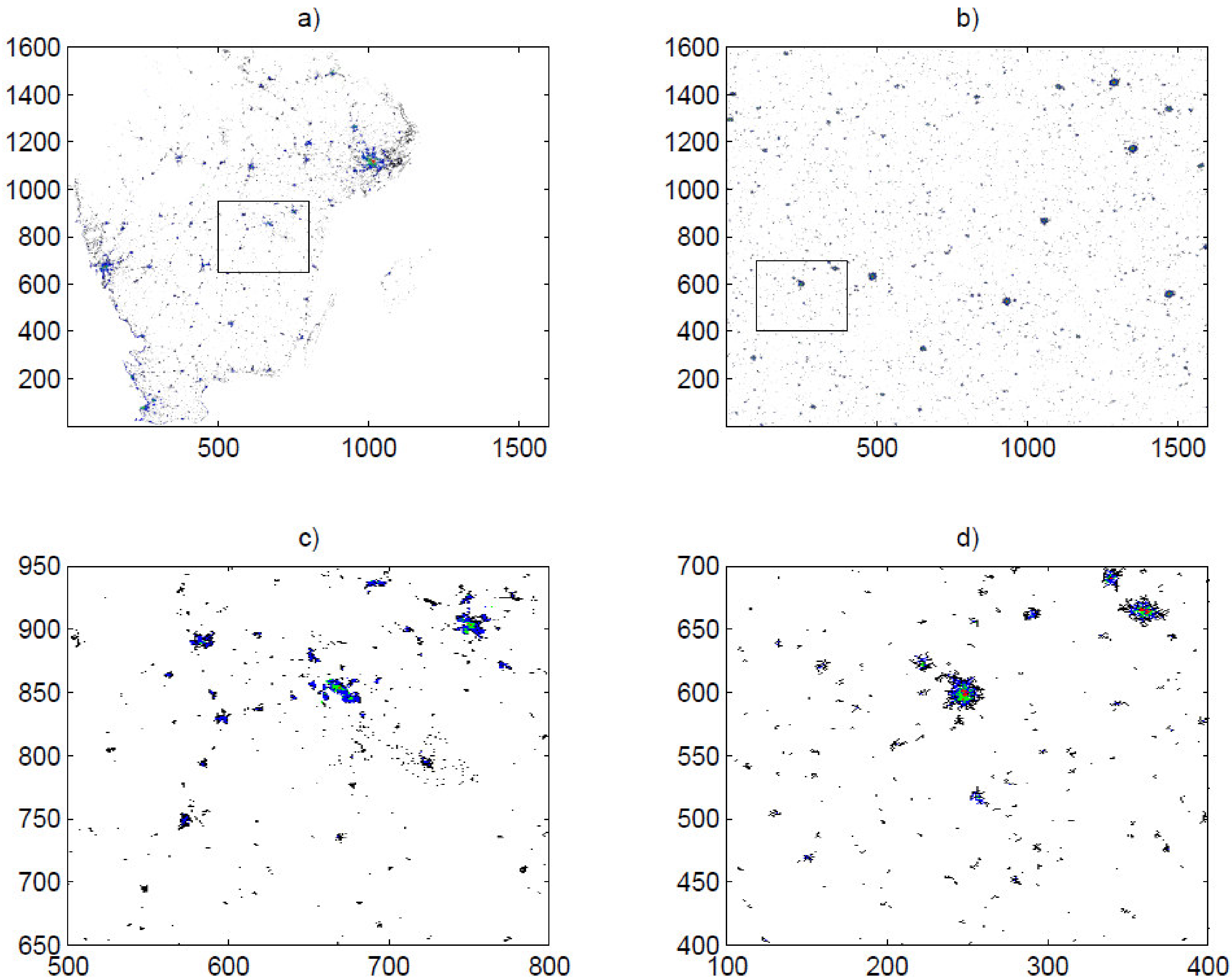


FIG. 4 Diagrams (a), (b), and (c) show double-logarithmic histograms with exponentially binned empirical (×) and simulated (□) observables: (a) - land value per 400m×400m sized cell, (b) - aggregated cluster land value, (c) - cluster area. For empirical cluster measurements, land values were aggregated to 400m×400m cells, and a threshold of 1425 kSEK/cell was applied. All contiguous (8-cell neighborhood) areas above this threshold were identified as clusters. In diagrams (d) and (e), empirical (broad boxes) and simulated (thin boxes) results for cluster area are plotted against exponentially binned cluster perimeters (d) and aggregated cluster land values (e). The vertical interval of the boxes contains 90% of the observations in the corresponding bins. The reference lines have slopes 0.7 in (d) and 1.5 in (e). Diagram (f) shows empirical cluster population plotted against exponentially binned empirical aggregated cluster land values. The vertical interval of the boxes contains 90% of the observations in the corresponding bins, and the crosses indicate the median cluster land values in the bins. The reference line has a slope of 1.0, which indicates that there is a near linear relationship between cluster price and population, for clusters with a population larger than 100.


a)
b)
c)
d)
1600
1400
1200
1000
800
600
400
200
500
1000
1500
950
900
850
800
750
700
650
500
600
700
800
650
550
450
400
100
200
300
400