Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 345 (2005) 227–244
PHYSICA A
www.elsevier.com/locate/physa

A spatial network explanation for a hierarchy of urban power laws

Claes Andersson*, Alexander Hellervik, Kristian Lindgren

Department of Physical Resource Theory, Chalmers University of Technology, 412 96 Göteborg, Sweden

Available online 8 August 2004

Abstract

Power laws in socioeconomic systems are generally explained as being generated by multiplicative growth of aggregate objects. In this paper we formulate a model of geographic activity distribution with spatial correlations on the level of land lots where multiplicative growth is assumed to be dominant but not exclusive. The purpose is to retain the explanatory power of earlier models due to Simon, Gibrat and others while attaining some additional properties that are attractive for both empirical and modelling purposes. In this sense, the model presented here is a combination of the two factors that have been identified as central to urban evolution but rarely appear unified in the same model: transportation costs and multiplicative growth. The model is an elaboration of a previously reported complex network model of geographical land value evolution. We reproduce statistical properties of an empirical geographical distribution of land values on multiple hierarchical levels: land value per unit area, cluster areas, aggregated land value per cluster and cluster area/perimeter ratios. It is found that transportation effects are not strong enough to disturb the power law distribution of land values per unit area but strong enough to sort nodes to generate a new set of power laws on a higher level of aggregation. The main hypothesis is that all these relations can be understood as consequences of an underlying growing scale-free network of geographic economic interdependencies.
© 2004 Elsevier B.V. All rights reserved.

PACS: 61.43.Hv; 89.75.Da; 89.75.Hc; 89.65.Gh

Keywords: Complex networks; Urban growth; Land values; Spatial networks; Power laws; Random multiplicative growth

*Corresponding author. Tel.: +46-31-7721513; fax: +46-31-7723827.
E-mail addresses: claes@santafe.edu (C. Andersson), f98alhe@dd.chalmers.se (A. Hellervik), frtkl@fy.chalmers.se (K. Lindgren).

0378-4371/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2004.07.027

1. Introduction

Macroscopic regularities are common in complex dynamical systems with distributed interacting agents. Maybe, the most well-known type of such regularities is the presence of fat-tailed distributions. It is widely agreed that fat-tailed distributions arise as a consequence of fundamental properties of the microscopic mechanisms including stochastic multiplicative growth and optimizing behavior. Stochastic multiplicative growth is an attractive candidate for modelling various mechanisms in economic systems and in economic geography, and multiplicative growth also serves as the foundation of the prevailing models; see e.g. Ref. [1]. Although generally considered to be derivations of Herbert Simon's model of urban dynamics [2], it can be argued that the genealogy of such models date back to Yule's model of the distribution of species in plant genera [3].

In economic geography, the rank-size rule, as in the generalized form of Zipf's Law, says that the size of a city has a power law relation to its rank. The importance of such observations probably reside primarily in the constraint that they put on the generating micro mechanisms. Understanding the mapping between micro mechanisms and phenomena on the macro scale in socioeconomic systems is of great importance; it is, for example, at the micro level that policies aimed at controlling such systems must be implemented. We argue for a view analogous to the situation with a series expansion where lower-order terms must be included before higher-order terms. According to this logic, broad and general properties of systems such as the presence of fat-tailed distributions are interesting because we can ask in which way, and why, the behavior of a system deviates from that predicted by a very simple generative model.

2. Empirical urban power laws

Empirical investigations of scaling laws in urban systems have generally been carried out observing population rather than, as we do here, observing land value. In Fig. 1 it can be seen that the relation between cluster population and aggregated cluster land value is essentially linear which implies that results apply also to the previous work where population distributions have been investigated. The most widely known urban scaling law is likely Zipf's Law for city sizes that in its generalized form, known as the rank-size rule, states that the size of a city is proportional to its rank to the power of a constant. The exponent differs between different parts of the world, as has been reported by e.g. [4]. Furthermore, fractal properties of internal cluster structures have been studied, among these the relationship between the perimeter and area of clusters [5]. In this paper we use this measure to compare some aspects of the internal structure of emergent urban clusters in the model with empirical observations. A thorough review of empirical work and models concerning urban scaling laws is provided by Pumain [6].

The empirical data with which we compare model results consists of geographical maps of land values and population in Sweden from year 2000 aggregated into 400×400 m cells. The underlying database delivered by Sweden Statistics consists of

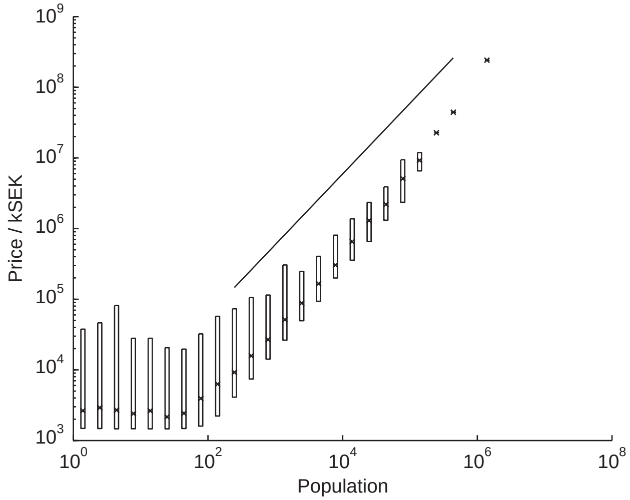

Fig. 1. Shown in this figure is a double-logarithmic plot of aggregated cluster land value as a function of cluster population. The vertical interval of the boxes contains 90% of the observations in the bin, and the crosses indicate the medians of the prices in the bins. For bins with only one observation, the boxes have no vertical extension. 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.

estimates of market value for 2.9 million taxing units. We identify clusters using a simple clustering algorithm that identifies contiguous areas of land value above a threshold value, chosen to be 1425 kSEK per cell (10 SEK $\approx$ 1 *USD*). This method was preferred over other methods because it can be universally and equally applied to the entire system and to simulated and empirical data. Simulation results will be discussed in a later section, so we will now only consider the empirical observations. Fig. 2 shows that empirical probability density for land value per unit area for Sweden is well described by a power law. This has been shown to be the case also for land value data from Japan [7].

For clusters we use two ways of measuring size: area and aggregated land value. The cluster area distribution is shown in Fig. 3, and the cluster aggregated land value distribution can be seen in Fig. 4. Both distributions are in good agreement with the rank-size rule for city sizes, but they show somewhat different exponents. Thus, power law relations between land value and frequency of occurrence can be found both on the level of lots (cells) and on the level of urban clusters. Furthermore, to characterize the geometries of the clusters we have measured the relationship between cluster areas and perimeters; see Fig. 5.

3. Theoretical and operational models

Models following Simon (henceforth "Simonoid models") have provided a framework in which to understand power laws in economic geography and their

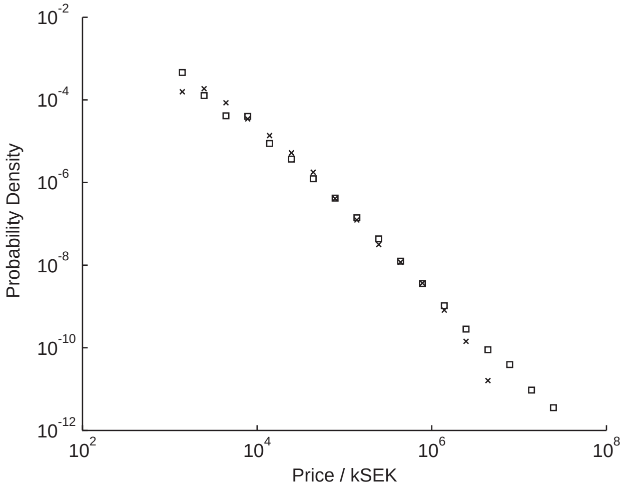

Fig. 2. Shown in this figure are double-logarithmic probability density plots of empirical (×) and simulated (□) land values per unit area. The empirical land values are aggregated into 400 m × 400 m sized cells. Only values above the threshold of 1425 kSEK/cell are shown.

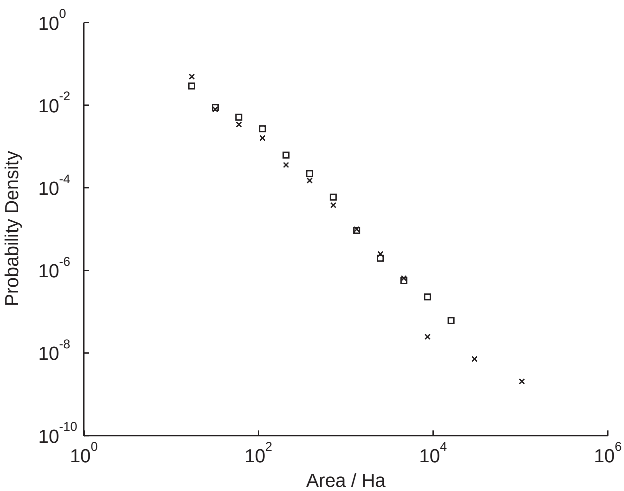

Fig. 3. Shown in this figure are double-logarithmic probability density plots of empirical (×) and simulated (□) cluster areas. Empirical land values were aggregated to 400 m × 400 m cells, and a threshold of 1425 kSEK/cell was applied. All contiguous (8-cell neighborhood) areas above this threshold were identified as clusters.

most important benefit is their conceptual simplicity (see e.g. Refs. [1,2,8]). Simonoid models typically describe the time dynamics of urban systems on the level of cities and they are not geographical models in the sense that no interactions between cities

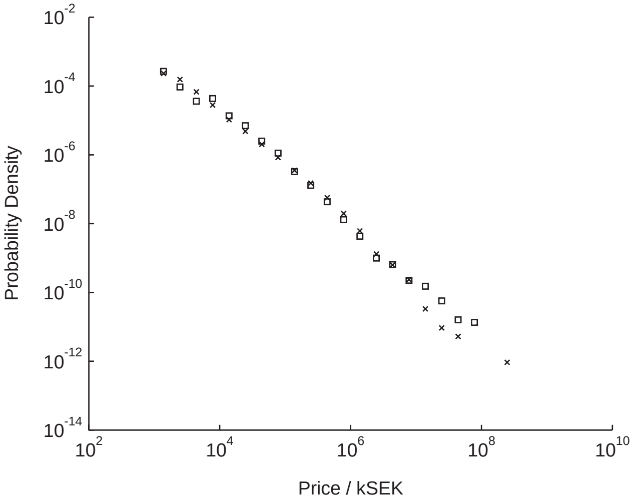

Fig. 4. Shown in this figure are double-logarithmic probability density plots of empirical (×) and simulated (□) land values per cluster. Empirical land values were aggregated to 400 m × 400 m cells, and a threshold of 1425 kSEK/cell was applied. All contiguous (8-cell neighborhood) areas above this threshold were identified as clusters.

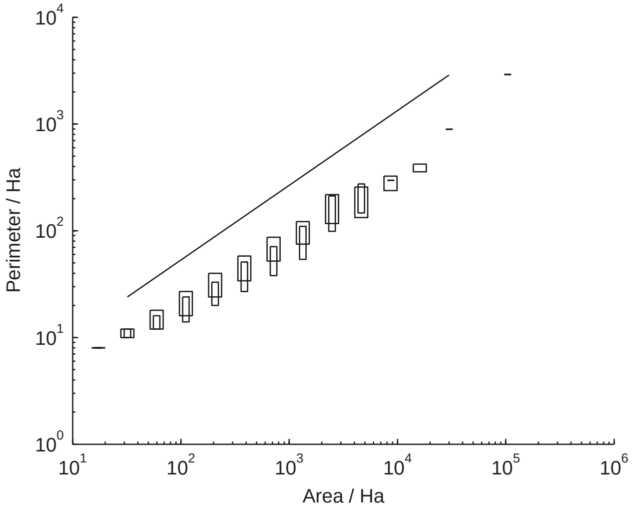

Fig. 5. This figure shows cluster area plotted against exponentially binned cluster perimeters for empirical (thin boxes) and simulated (broad boxes) results. The vertical interval of the boxes contains 90% of the perimeters in the bin. The reference line has a slope of 0.7.

and no geographical extent of the cities are modelled. Cities are problematic to use as objects in geographical models for a number of reasons. Firstly, cities grow together and new sub-centers may emerge over time. Furthermore, cities are not atomic objects even during their lifetime since parts of two different cities may interact more with each other than these parts do with some parts of the cities to which they belong. Another shortcoming is that interactions are not incorporated in these models although it is generally agreed that transportation diseconomies are defined in urban evolution; see e.g. Refs. [9,10]. This means that Simonoid models can provide little guidance for policies aimed at controlling the behavior of these systems and that they become very hard to garnish with additional detail for application to specific scenarios.

Since neoclassical microeconomic models, where transportation plays a central role, have great difficulties in modeling evolutionary growing systems and Simonoid models lack the geographic aspect, urban modeling has turned to cellular automata as a modeling framework [11–13]. While cellular automata models of urban systems capture the geographical form of the system they do not provide much in terms of avenues for systematic analysis.

There are also a number of physics-inspired models aimed at explaining urban power laws. The correlated percolation model [14] is capable of reproducing power law distributed cluster areas, but it suffers from a need of specifying both an a priori center and a concentration gradient. Marsili and Zhang suggested that the pair-wise interaction between individuals might explain the power law distribution on the level of cities [8]. No explicit spatiality or interaction between cities is introduced, which means that the explanatory power is similar to that of a Simonoid model. Zanette and Manrubia used a multiplicative stochastic growth model with spatial diffusion to produce a power law population distribution both on the cell and cluster level [15].

4. A complex network model of economic geography

4.1. Introduction to the model

It has sometimes been said that power laws are over identified [6], and, indeed, power laws are generated by a wide range of models. Thus, the presence of power law distributions does not uniquely identify underlying processes. Because of this it is reasonable that a model, in addition to producing the desired macroscopic behavior, should be argued for in terms of the microscopic properties of the target system.

The aim here is to create a model that combines the qualities of previous models while solving some of their problems along the lines outlined in the previous section. A complex network model can be embedded in a cellular space; it can incorporate a model of transportation in the mechanism by which connections are formed, and it may employ multiplicative growth along with other types of growth. For reviews of complex network models, see Refs. [16–18].

To employ network models to explain urban power laws, we argue that there is, on average, a linear relation between node degrees of the underlying economic interaction (trade) network and urban land values. This relation will be motivated later on in the text. Furthermore, as identified in Fig. 1, there is a linear relation between land price and population, so results also apply to earlier empirical work where population rather than land price has been observed. In the spirit of Simonoid models we then start out by specifying a model of network growth without a geographical component. Quite in line with such models, we show that as far as multiplicative growth is a possible explanation for the observed power law of land values, transportation costs need not be a part of this explanation. On the contrary, when we add geography to the model (a bias against long connections) to explain statistical properties of urban clusters, we have to verify that the bias is weak enough not to sufficiently disturb the scale-free properties of the growing network.

Thus, we explain the power law distribution of land values per unit area as a consequence of multiplicative network growth. On top of this, we introduce a spatial bias making long connections more rare than short connections to the end of arranging the nodes into geographical clusters of high node degrees. This spatial bias is sufficiently weak not to cause major deviations from a true power law distribution of node degrees but strong enough to result in another set of power laws on the level of geographical clusters. Clusters are measured entities and are not used as objects in their own right in the model—they are emergent structures. Transportation costs arguably being the strongest geographical bias, it appears likely that further biases can be introduced to fit empirical data even better without destroying the fundamental properties of the model.

4.2. Ontology

One of the aims with the presented model is to move the formulation of a Simonoid multiplicative growth model "one step down" to a less aggregated level. By doing this, as previously stated, we intend to create a geographical model that is open-ended with regard to additional detail needed for operational modeling and that is well founded in behavioral propositions on the micro level. A precondition for this is that the objects and relations in the formulation sufficiently meet criteria for individuality. As stated previously in the section "Theoretical and operational models", models formulated in terms of city objects face some operational problems as geographical models. To avoid this we take nodes in the network to be cells in a cellular space representing the geographical area to be simulated. We will use the terms "cells" and "nodes" interchangeably throughout the following text depending on context.

Connections (network edges) are taken to be streams of trade exchanges between activities residing in the cells. In this context, we define trade as any profitable interaction between two economic agents. For instance, the economic interaction between a household and a workplace is considered to be a trade of labor for currency and benefits that is perceived as profitable to both end-points. Network loops (a connection between a node and itself) are allowed as they represent trade between activities in such close physical proximity as being located in the same geographical cell. A trade stream, assuming both end-points have agreed to its establishment, will on average result in an economic benefit (profit) to both parties. Without identifying activities as sellers or buyers we may also make the simplifying assumption that both end-points to a connection will benefit equally. Furthermore, we define connections to be of equal and unit value to both end-points such that large trade streams would be represented by a bundle of connections. Following this logic, long-range connections are identical in size with short-range connections but they form at a slower rate to reflect that fewer potential trades remain profitable over long distances. To conclude, the system state is described completely by an undirected network where multiple edges and loops are allowed.

Using fixed-size non-overlapping cells as nodes carries with it some important implications. First of all, such nodes will by definition remain the same throughout the simulation. The only difference between a node at different points in time will be its connection to other nodes. If we would have used cities instead of cells as nodes, this would have no longer been true, since nodes would divide and fuse over time and the problem of even defining and delimiting a node would be much greater. Ultimately, this problem becomes especially pressing when attempting to compare model states with empirical states because we would need to somehow devise a method for node identification that applies both to the real system and the simulated system. Using cells, this problem becomes much less restricting.

Secondly, because the area of a cell is by definition constant, any addition of connections will translate to an increase in profits per unit area in the cell in question. If we assume market pricing of land to be a process that is fast compared to the process of urban growth, we can therefore invoke the left-over principle from urban economics to interpret the degree of a node to be proportional to the rent that the land owner can charge tenants residing therein [19]. This principle states that, in a competitive land market, rent equals the amount of money left after all expenses (except rent) are paid. This amount of money available equals the sum of all trade benefits at the site.

From the interpretation of node degrees as land rents, a first approximation of land value in turn being proportional to land rent is reasonable. Capitalizing indefinite periodic rent income from cell i gives land value $v_i = r_i/\rho$, where r_i is the land rent and ρ is the interest rate [10].

The connection between node degree and land value thus consists of two proportionalities. For node i we get

$$v_i \propto r_i \propto x_i \,, \tag{1}$$

where v_i is the value of the corresponding land area, r_i is the land rent, and x_i is the node degree representing the total number of trade connections as outlined above. This line of reasoning ties model quantities to empirically measurable quantities and thereby allows for a comparison to be made between model outcomes and the state of the real world system.

In the current model we do not specify the type of economic activity that takes place at a site, only the activity level, i.e., node degree. Thus, activities are considered to be average activities and they all have equal probabilities for establishing interactions with each other. Although units of activity (connection end points) are of no particular type they are not assumed to be all of the same type. Rather, it is assumed that a sufficient mix of activities are present in all areas. In this respect, it can be said that activities are "unobserved".

The network grows by successive addition of edges and nodes to the system. The addition of a new end-point to a node corresponds to an increase in the profit generated in this land area. Addition of new nodes to the network represent development of previously undeveloped sites, i.e., transition from non-urban to urban land use. Empirically, we define this to correspond to the land value rising above a prescribed threshold. Note that, in principle, transactions should be possible to represent a resolution where interactions will be pair-wise. The connection of new edges to nodes is modelled through additive and multiplicative growth. The situation where growth events are either multiplicative or additive should be viewed as a model simplification. More likely, individual growth events in the real system will have a more or less strong multiplicative component, as defined in the next paragraph.

Multiplicative (or preferential) growth can be said to involve processes that act per unit of development. For example, if an additional unit of economic activity arises in one location and we seek the geographical location of its trading partner. Assuming all units of activity to be "unobserved", and after taking transportation costs into account, the likelihood of a given node to contain this trading partner is proportional to the amount of development (degree) in that node. Furthermore, technological improvements will often be applied to production processes that are scaled to a suitable level. For example, an activity in a location will often consist of a number of copies of some production process and improvements to this production process will thus be multiplicative with respect to the present scale of production. This means that, in absolute terms, had the activity level been half of the present amount, the benefit of some efficiency improvement would also be half. This line of reasoning assumes that, on average, the relative size of the improvement is independent of the current profitability of the process. Additive growth is growth that does not take place at a rate that is proportional to the present amount of development. For example, policies intended to stimulate growth in areas of low development will be non-preferential. Furthermore, roads connect areas of high development but are likely to also run through areas of low development. Improvements of the quality of infrastructure because of this will inadvertently boost areas that happen to lie along the extent of roads regardless of the amount of development that is currently present. Another important additive effect might be when the location of new activity is multiplicative on a large scale but not on a small scale. For example, a new unit of development might be allocated to a small area with low development inside a larger area with relatively high development. This effect seems to apply particularly well to areas near the urban perimeter where there is a good supply of nearby undeveloped land.

4.3. The non-spatial model

Several variations and extensions of the initial Barabási–Albert model [20] have been proposed; see Refs. [18,21,22] for reviews. A simple, but fairly general version can be formulated formally as follows: consider a growing, enumerated set of nodes, $\{1, 2, \ldots, N\}$ connected by undirected edges. There are no restrictions on multiple edges between two nodes and a node is allowed to have any number of connections to itself. The degree of node i is denoted by x_i.

The network is initialized by connecting n_0 nodes so that each has a degree of x_0. At each iteration t the network is updated by the addition of one edge between two nodes that are chosen independently according to the following rules:

(1) With probability q_1 the node is chosen uniformly between existing nodes. The probability of node i to be selected is

$$\Pi_i^u = \frac{1}{N} \; . \tag{2}$$

(2) With probability q_2 the node is chosen preferentially, which corresponds to the uniform selection of an edge end-point in the system and the subsequent location of its node. The probability of node i to be selected is

$$\Pi_i^p = \frac{x_i}{\sum_j x_j} \; . \tag{3}$$

(3) With probability q_3 a new node is added to the network. This node will get a degree of 1.

The parameters q_1, q_2 and q_3, fulfill $q_1 + q_2 + q_3 = 1$ and are assumed to be constant during the evolution of the network.

The time evolution of the expected degree for node i follows:

$$x_i(t+1) = x_i(t) + 2q_1 \frac{1}{N(t)} + 2q_2 \frac{x_i(t)}{\sum_j x_j(t)} \; , \tag{4}$$

where $N(t) = 2q_3 t$ is the expected number of nodes developed after t iterations. By using the continuous-time method introduced in Ref. [23], we find that after a sufficiently long time, the degree distribution approaches the form

$$P[x_i = x] \sim (x + B)^{-\gamma} \; , \tag{5}$$

with

$$B = \frac{q_1}{q_2 q_3} \tag{6}$$

and

$$\gamma = 1 + \frac{1}{q_2} \; . \tag{7}$$

According to Eq. (5), and if land value is linearly related to node degree, as motivated in the previous section, this simple stochastic model can reproduce the power law distribution of land values shown in Fig. 2. In the context of urban power laws, the next step is to introduce some notion of distance that inversely affects the propensity of the nodes to connect. In this way, the probability for a new node to connect to an existing node becomes positively dependent on the degree of the existing node and negatively dependent on the distance between the new node and the existing node.

4.4. The spatial model

In the non-spatial model, activities will trade at random with other activities according to some scheme with the only property of a cell visible to the process being its degree. Of course, this specification has a number of serious drawbacks and needs to be elaborated on. Most notably, without any geographic incentive to remain in a cell with high development, intelligently administered activities would of course spread out to avoid high rents. Thus, the non-spatial case is useful only to demonstrate that, quite in line with Simonoid models, transportation is not needed as such for explaining the scale-free distribution of activities over nodes. All that is needed is some incentive to remain in highly developed nodes.

When introducing geography into the model we need to distinguish between a first and a second end-point when forming a connection. This requirement is of little practical importance and also seems innocuous since it appears reasonable that most exchanges involves one searching and one passive party: one that offers and one that browses and accepts or rejects. A complementary interpretation is that this represent the approximation that the random economic perturbations do not happen at both the supply and demand side at the same time. We call the creation of the first and second end-points in a new connection, "primary growth" and "secondary growth", respectively. After the connection has been established, it is considered undirected.

If we assume constant fractions q_1 and q_2 for uniform and preferential growth respectively, with $q_1 + q_2 = 1$, the probability of selecting a node i preferentially in primary growth is

$$\Pi_i^{1,pref} = q_2 \frac{x_i}{\sum_j x_j} , \tag{8}$$

which is similar to the non-spatial case.

In the non-spatial model we treated the case of node addition as a third, separate case, which turns out to be unnecessary in the spatial model. It can now be handled as uniform growth in a previously undeveloped cell, if the probability of selecting this cell is dependent on the availability of basic infrastructure. To keep it simple, we divide the cells into three categories—developed cells, perimeter cells and external cells. If we ascribe a weight a_i to each cell (dependent on to which category it belongs) and pose the condition that the sum of all probabilities for uniform primary growth should be q_1, we get

$$\Pi_i^{1,uni} = q_1 \frac{a_i}{\sum_j a_j} . \tag{9}$$

The developed cells are defined to be the base case, and $a_i = 1$ for these. Perimeter cells are all undeveloped cells adjacent to a developed cell (8-cell neighborhood), and all these get $a_i = b$, where b is a parameter. In most realistic cases $b < 1$, because there is, compared to a developed cell, a less probability for an average perimeter cell to have adequate infrastructure. Government growth control will also have the effect of decreasing b.

External cells are cells that are neither developed nor on the urban perimeter. In this case, usually only cells with direct access to roads can be considered for development. To avoid the (for the purposes of this model unnecessary) details in specifying a realistic model for the development of a road network, we simply assume that all external cells, with some probability have access to relevant roads and other infrastructure (thus effectively making it a perimeter cell). This probability must surely grow with the system, and we represent this by, at each iteration t, taking it proportional to the ratio between the number of perimeter cells, $n_t^{(P)}$, and the number of external cells, $n_t^{(E)}$. We then get, for external cells, $a_i = b\varepsilon n_t^{(P)}/n_t^{(E)}$, where ε is a constant parameter describing the relative density of infrastructure. For this to be reasonable, we must assume that the lattice is not too crowded, i.e., $\varepsilon n_t^{(P)} < n_t^{(E)}$.

Now, when all three categories of cells have been considered, Eq. (9) can be written as

$$\begin{aligned} \Pi_i^{1,uni} &= q_1 \frac{\delta_i^{(D)} + b\delta_i^{(P)} + b\varepsilon(n_t^{(P)}/n_t^{(E)})\delta_i^{(E)}}{\sum_j(\delta_j^{(D)} + b\delta_j^{(P)} + b\varepsilon(n_t^{(P)}/n_t^{(E)})\delta_j^{(E)})} \\ &= q_1 \frac{\delta_i^{(D)} + b\delta_i^{(P)} + b\varepsilon(n_t^{(P)}/n_t^{(E)})\delta_i^{(E)}}{n_t^{(D)} + b(1+\varepsilon)n_t^{(P)}}\,, \end{aligned} \tag{10}$$

where $n_t^{(D)}$ is the number of developed cells and $\delta_j^{(D)} = 1$ if cell j is developed and $\delta_j^{(D)} = 0$ otherwise. The meanings of $\delta_j^{(P)}$ and $\delta_j^{(E)}$ are analogous to $\delta_j^{(D)}$, with P referring to perimeter cells and E referring to external cells. A cell must belong to one and only one of the categories at a time, which means that for each j one and only one of $\delta_j^{(D)}$, $\delta_j^{(P)}$ and $\delta^{(E)}$ is equal to 1 while the other two equal 0.

After selecting a primary cell, using one of the two mechanisms described above, the secondary cell should be selected with a probability decreasing with distance from the primary one. To accomplish this, we define D_{ij} to be the spatial interaction strength between cells i and j. We have the restrictions $D_{ij} \leqslant 1$ and $D_{ij} = D_{ji}$.

The probability of secondary preferential growth at cell i as a consequence of primary growth at cell j is

$$\Pi_{ij}^{2,pref} = q_2 \frac{D_{ij}x_i}{\sum_k D_{kj}x_k} \tag{11}$$

and analogical for secondary uniform growth, it is

$$\Pi_{ij}^{2,uni} = q_1 \frac{D_{ij}a_i}{\sum_k D_{kj}a_k} \tag{12}$$

and with the same site categories as for primary growth, this probability can be written as

$$\Pi_{ij}^{2,uni} = q_1 \frac{D_{ij}(\delta_i^{(D)} + b\delta_i^{(P)} + b\varepsilon(n_t^{(P)}/n_t^{(E)})\delta_i^{(E)})}{\sum_k D_{kj}(\delta_k^{(D)} + b\delta_k^{(P)} + b\varepsilon(n_t^{(P)}/n_t^{(E)})\delta_k^{(E)})} \,. \tag{13}$$

Several choices for D_{ij} are possible, but it is clear that the interaction strength should decay with increasing transportation costs. We have for the results reported here, used

$$D_{ij} = (1 + cd(i,j))^{-\alpha} \,, \tag{14}$$

where $d(i,j)$ is the Euclidean distance between cells i and j. The non-negative parameters c and α control the impact of spatiality. The choice of functional form and parameters for D_{ij} is not of critical importance.

This simple model of trade network evolution aims to capture some mechanisms driving the system: (i) random perturbations of the economy giving rise to new opportunities for profit, (ii) the seizing of such opportunities, and, (iii) spatial correlation due to profit resulting from trade and trade incurring transportation costs.

The model of trade relation end-point formation is implemented, as specified earlier, as a combination of uniform and preferential growth; see Eqs. (8), (10), (11) and (13). Costs of transportation and other distance-related diseconomies between a pair of nodes $\{i,j\}$ are captured by a connection length bias, D_{ij}, that makes long connections form at a slower rate than short connections. This distance bias modifies secondary uniform and preferential growth rates; see Eqs. (11) and (13).

With this formulation we can keep the useful definition of an edge representing unit profit and node degree proportional to land rent; see Eq. (1). This can primarily be viewed in two ways. Firstly, once a connection is formed it is always of unit benefit to its end-points but the further apart two lots are located the lower will the rate of connections forming between them be. This is because fewer and fewer potential trades will remain profitable when transportation costs are deducted. Secondly, larger than unit profit trade streams will have to be represented as bundles of unit profit connections. Stretched out over a longer distance, such connection bundles will grow thinner as profits are reduced by transportation costs.

The model for availability of infrastructure is also simple and in addition also local. We simply distinguish between three cases: internal, perimeter and external. Internal availability to infrastructure is maximal and is used as a unit. Perimeter availability applies to undeveloped lots adjacent to developed lots. External lots are lots that are undeveloped and whose geographic neighbor lots are also without development. External infrastructure availability is determined by the parameter ε and reflects the existence of an ambient infrastructure network serving long-range interactions. Such infrastructure spans undeveloped areas sparsely and may be "tapped into", thus making what will appear as new clusters of developed cells possible. See Eqs. (10) and (13).

4.5. Analysis

To simplify the analysis of the model it is useful to assume that development of a new land (addition of a connection to a previously undeveloped cell) takes place at a constant rate q_A (compared to other types of growth, not as a function of physical time). To motivate the assumption, let us consider the growth of developed clusters and their perimeters. It is true in simulations of the model, and it can be empirically verified (see Figs. 3 and 5), that the cluster area distribution is close to a simple power law with density function $f(A) \sim A^{-\beta}$, and that the relation between cluster perimeter P and cluster area A has the form $P \sim A^{\lambda}$, with $\lambda < 1$. From this we observe that for the entire system of clusters we have

$$\frac{n_t^{(P)}}{n_t^{(D)}} \sim \frac{\int_1^\infty A^{-\beta} A^{\lambda} dA}{\int_1^\infty A^{-\beta} A dA} = \frac{\beta - 2}{\beta - \lambda - 1} , \tag{15}$$

assuming that $\beta > 2$ and $\lambda < 1$. This means that it can be expected that the ratio between the total number of perimeter nodes and developed nodes is approximately constant. We now define q_1' as the fraction of primary activity increments that occur on developed nodes under uniform growth, and Eq. (10) gives

$$q_1' = \sum_i \delta_i^{(D)} \Pi_i^{1,uni} = q_1 \left(1 + b(1 + \varepsilon) \frac{n_t^{(P)}}{n_t^{(D)}} \right)^{-1} , \tag{16}$$

which, because of Eq. (15), is approximately constant. This means that the rate of primary node activation, $q_A = q_1 - q_1'$, can also be considered constant.

It can be argued that expected secondary growth behaves very similar to primary growth, even when the impact of spatiality is strong [24]. This means that the time evolution of expected activity on a developed site i can be approximated by

$$x_i(t+1) = x_i(t) + 2q_1' \frac{1}{n_t^{(D)}} + 2q_2 \frac{x_i(t)}{\sum_j x_j(t)} , \tag{17}$$

which is similar to the non-spatial case and the degree distribution can be found in the same way, and thus, after a long time, can be expected to approximately approach a generalized power law, $P[x_i = x] \sim (x + B)^{-\gamma}$, with

$$B = \frac{q_1'}{q_2 q_A} = \frac{q_1'}{q_2(1 - q_2 - q_1')} \tag{18}$$

and

$$\gamma = 1 + \frac{1}{q_2} . \tag{19}$$

4.6. Model parameters

If we want the model to reproduce the statistical properties of a specific economic geographic system, then most of the parameters in the model can be roughly estimated from empirical data, saving us from extensive parameter tuning. For instance, because ε controls the ratio of external development to perimeter development, and external growth for the most part gives a remaining new cluster, we have the approximate relationship

$$\frac{\varepsilon}{1+\varepsilon} \simeq n_t^{(C)}/n_t^{(D)}\,, \tag{20}$$

where $n_t^{(C)}$ is the number of clusters. The exponent γ in the distribution of land prices gives information about the size of q_2 (and q_1) via Eq. (19). The perimeter parameter b controls the ratio of perimeter development and internal uniform growth, and it can then be determined by

$$q_A \simeq \frac{n_t^{(D)}}{\sum_i x_i(t)} \tag{21}$$

and with the use of Eq. (16) we get

$$b \simeq q_A \left((q_1 - q_A)(1+\varepsilon)\, \frac{n_t^{(P)}}{n_t^{(D)}} \right)^{-1} . \tag{22}$$

The fact that the observed current state of the system constrains the value of these parameters does not mean that they are purely empirical in nature. They are aggregated measures of fundamental properties of the current economic system (and region) of interest, and in the general case they may change with time. If the model should be used for prediction, instead of explanation as is the case here, then more elaborate ways of estimating current (and future) parameters are needed.

The spatial parameters c and α in Eq. (14) reflect the statistics of transport characteristics of the economic configuration. Their effect on the statistical properties of the system are somewhat subtle and thus they are not as easily estimated from data as the parameters mentioned above. Anyhow, results seem robust to their exact values and the functional form of D_{ij}.

4.7. Simulation results

Figs. 2–5 reflect data from the same Monte Carlo simulation run of the model. Together, the figures show that the model reproduces a number of statistical properties of the target system concurrently on different levels of emergence. The power law distribution of land prices is predicted already by the non-spatial model and when spatiality is introduced Eq. (14) the node degrees are spatially sorted into clusters with the desired properties while retaining the per-cell distribution from the non-spatial model. To reproduce the power law exponent in the empirical distribution of land values, $q_1 = 0.2$ was used. The parameters $b = 0.34$ and $\varepsilon =$

0.18 were then estimated from empirical cluster data as described above. The spatial parameters were $c = 0.1$ and $\alpha = 2$. Investigation of sensitivity shows that exponents and proportions change slowly and smoothly with all parameters. A square grid of 1600×1600 cells (each cell representing a 16 ha square) was used to match the land area of Sweden (41,093,400 ha). The number of iterations was $270,000$, to give the final simulated configuration the same total price as the aggregated empirical price of developed sites (760,000,000 kSEK). In comparison with empirical land values, the model land values were taken as $1425x_i$ kSEK (degree 1 then corresponds to the threshold land value for developed cells).

5. Summary

Operational models of urban growth today are not founded on more general conceptual models [25–27]. Conceptual models, such as Simon's model are not geographical and their insights can thus not easily be extended to models of geographical distributions. The model presented here aims to provide a synthesis between operational and conceptual models. In terms of previous models, the present model is a random multiplicative growth model in the tradition of Simon but is formulated on a lower level: a level where a better case for individuality of the included objects can be made, where the microeconomic argument can be made more transparent, and where geography can be included.

We have chosen to represent the urban system as a collection of land lots with trade relations between them. Maintaining the importance in principle of preferential growth as a base-line model of economic change, such a network model essentially becomes a scale-free network model following Barabási et al. In essence we say that when a new trade relation appears, sufficiently often the probability of it having an end-point in a given node is proportional to the amount of activity already present in that node. In addition, we say that new (previously isolated) nodes are added to the network with a certain rate. In terms of the target system this corresponds to the development of land areas that previously had no urban activity.

This model, provided our argument for the mapping between real-world objects and nodes and connections in the model is convincing, bears a scale-free distribution of node degrees as a consequence. Thus, the scale-free distribution of land values per unit area in empirical data can be explained as stemming from a preferential growth of the underlying trade network. This forms the basic model which is basically a Barabási–Albert model with interpretations similar to Simonoid models put on its components and mechanisms.

The most obvious deviation from pure preferential growth in an urban system stems from costs of transportation. An activity in some location is more likely to trade with nearby lots because of cost related to transportation. Hence, the probability of interaction, due to increasing costs of transportation, can be assumed on an average to taper off gradually as a function of distance.

As argued in Ref. [24], a distance bias would have effects similar to the effects of node fitness [28,29]. This means that we should expect transportation costs to make the distribution of land values per unit area multi fractal. That is, it should become a sum of power laws with different exponents. However, although the spatial bias is strong enough to arrange the nodes into clusters spatially, it is not strong enough to cause a major deviation from a pure power law. Thus, the model ''survives'' the introduction of a transportation cost to perturb the dynamics.

Furthermore, we add some additional constraints governing the recruitment of new nodes to also capture the arrangement of nodes into urban clusters. We retain the previous behavior of the model, i.e., that local land values are distributed similarly to observed empirical land values, and, in addition, some properties on the emergent level of geographical clusters are also captured. These properties are: cluster area, cluster-aggregated land value and the ratio between cluster areas and perimeters. We also verify empirically that there is essentially a linear relation between the aggregated land value of a cluster and its population which indicates that results should apply also to observations of population rather than land value.

Acknowledgements

We thank Anders Hagson and Jonas Tornberg, City and Mobility, Chalmers University of Technology, for providing preprocessed empirical data. We also thank Koen Frenken, Martin Nilsson, Steen Rasmussen and Anders Eriksson for valuable discussions and input.

References

[1] X. Gabaix, Zipf's law for cities: an explanation, Quart. J. Econ. 114 (1999) 739–767.
[2] H.A. Simon, On a class of skew distribution functions, Biometrika 42 (1955) 425–440.
[3] G. Yule, A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, F.R.S. Philos. Trans. Roy. Soc. London Ser. B 213 (1925) 21–87.
[4] Y.M. Ioannides, H.G. Overman, Zipf's law for cities: an empirical examination, Regional Sci. Urban Econ. 33 (2) (2002) 127–137.
[5] M. Batty, P. Longley, Fractal Cities: A Geometry of Form and Function, Academic Press, New York, 1994.
[6] D. Pumain, Scaling laws and urban systems, Santa Fe Institute pre-prints: 04-02-002, www.santafe.edu/sfi/publications/wpabstract/200402002.
[7] T. Kaizoji, Scaling behavior in land markets, Physica A 326 (2003) 256–264.
[8] M. Marsili, Y.-C. Zhang, Interacting individuals leading to zipf's law, Phys. Rev. Lett. 80 (12) (1998) 2741–2744.
[9] M. Fujita, P. Krugman, T. Venables, The Spatial Economy, MIT Press, Cambridge MA, 1999.
[10] A. O'Sullivan, Urban Economics, McGraw Hill Higher Education, 2002.
[11] W. Tobler, Cellular geography, in: S. Gale, G. Olsson (Eds.), Philosophy in Geography, D. Reidel Publishing Company, Dordrecht, Holland, 1979, pp. 379–386.
[12] H. Couclelis, Cellular worlds: a framework for modeling micro-macro dynamics, Environ. Plan. A 17 (1985) 585–596.
[13] P.M. Torrens, D. O'Sullivan, Cellular automata and urban simulation: where do we go from here?, Environ. Plan. B 28 (2001) 163–168.
[14] H.A. Makse, J.S. de Andrade, M. Batty, S. Havlin, H.E. Stanley, Modeling urban growth patterns with correlated percolation, Phys. Rev. E 58 (1998) 7054–7062.

[15] D.H. Zanette, S. Manrubia, Role of intermittency in urban development: a model of large-scale city formation, Phys. Rev. Lett. 79 (1997) 523–526.
[16] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47–97.
[17] S. Dorogovtsev, J. Mendes, Evolution of networks, Adv. Phys. 51.
[18] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2) (2003) 167–256.
[19] W. Alonso, Location and Land Use, Harvard University Press, 1964.
[20] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[21] A.-L. Barabási, H. Jeong, R. Ravasz, Z. Néda, T. Vicsek, A. Schubert, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47–97.
[22] S. Dorogovtsev, J. Mendes, Evolution of networks, Adv. Phys. 51 (2002) 1079–1187.
[23] R. Albert, A.-L. Barabási, H. Jeong, Mean-field theory for scale-free random networks, Physica A 272 (1999) 173–187.
[24] C. Andersson, A. Hellervik, K. Lindgren, A. Hagson, J. Tornberg, The urban economy as a scale-free network, Phys. Rev. E 68 (036124).
[25] R. White, G. Straatman, G. Engelen, Best Practices in Spatially Integrated Social Science, Oxford University Press, Oxford, pp. 420–442.
[26] R. White, G. Engelen, I. Uljee, The use of constrained cellular automata for high-resolution modeling of urban land use dynamics, Environ. Plan. B, vol. 24 (1997) 323–343.
[27] K.C. Clarke, S. Hoppen, L. Gaydos, A self-modifying cellular automaton model of historical urbanization in the San Francisco bay area, Environ. Plan. B: Plan. Design 24 (1997) 247–261.
[28] G. Bianconi, A.-L. Barabási, Competition and multiscaling in evolving networks, Europhys. Lett. 54 (4) (2001) 436–442.
[29] G. Ergün, G.J. Rodgers, Growing random networks with fitness, Physica A 303 (2002) 261.