Urban growth simulation from "first principles"

Claes Andersson*
EES-6 MS T003, Los Alamos National Laboratory, Los Alamos, New Mexico 87545
and Physical Resource Theory, Chalmers University of Technology, 412 96 Göteborg, Sweden

Kristian Lindgren†
Physical Resource Theory, Chalmers University of Technology, 412 96 Göteborg, Sweden

Steen Rasmussen‡
EES-6 MS T003 and T-CNLS MS B258, Los Alamos National Laboratory, Los Alamos, New Mexico
and Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501

Roger White§
Department of Geography, Memorial University, St. John's A1B3X9 Newfoundland, Canada
(Received 27 September 2001; published 12 August 2002)

General and mathematically transparent models of urban growth have so far suffered from a lack in microscopic realism. Physical models that have been used for this purpose, i.e., diffusion-limited aggregation, dielectric breakdown models, and correlated percolation all have microscopic dynamics for which analogies with urban growth appear stretched. Based on a Markov random field formulation we have developed a model that is capable of reproducing a variety of important characteristic urban morphologies and that has realistic microscopic dynamics. The results presented in this paper are particularly important in relation to "urban sprawl," an important aspect of which is aggressively spreading low-density land uses. This type of growth is increasingly causing environmental, social, and economical problems around the world. The microdynamics of our model, or its "first principles," can be mapped to human decisions and motivations and thus potentially also to policies and regulations. We measure statistical properties of macrostates generated by the urban growth mechanism that we propose, and we compare these to empirical measurements as well as to results from other models. To showcase the open-endedness of the model and to thereby relate our work to applied urban planning we have also included a simulated city consisting of a large number of land use classes in which also topographical data have been used.

DOI: 10.1103/PhysRevE.66.026204 PACS number(s): 89.75.−k, 61.43.Hv, 89.65.Lm, 89.75.Da

I. INTRODUCTION

Since the dynamics of city growth is governed by mechanisms that to a large extent take place on a scale in time and space that is in the range of casual human perception, it is a naturally occurring complex dynamical system that many people have a strong connection to in their day-to-day life. The urban system is an integral and important part of our lives, and the problems that follow from rapid urban growth affect not only those living in the cities. For researchers it is a system that is challenging and relevant in a variety of aspects: Physicists study abstract models with theoretical understanding in almost exclusive focus [1–8] while, on the other end of the spectrum, other urban researchers (geographers, etc.) are designing tuned predictive models with deployment for actual use in urban planning as the primary goal [9–18]. In the latter case, little understanding of the underlying dynamics can generally be derived. (The goal is generally to embed knowledge rather than to derive it.) We choose to use the term "first principles" since the interactions in our model are intended to depict human actions and decisions, which is what ultimately drives the process of urban growth. It is also on that level of description that an understanding of urban dynamics is the most portable to its areas of application. The microscopic formulation of our model is highly macroscopic compared to the level of description that is traditionally used in statistical physics, from which we have borrowed much of the formulation. Because of this, the use of concepts from physical models has to be viewed in light of the lack of rigorous methods for selecting the complex fundamental objects of an urban system.

The human intellect conceptualizes the world hierarchically with different concepts on different levels. To see this we do not need to look further than to the concepts of meronyms and holonyms in human language: block, neighborhood, city part, city, country, nation, and continent. This mode of simplification is borrowed into our model to reduce the computational complexity of maintaining latticewide site interactions through mean-field approximations. The concept of using a hierarchy of scales for describing urbanization is well-established in geography through central place theory [19–21], and renormalization is a standard method in statistical physics. Spatial interaction models have been used ex-

*Email address: andersson@lanl.gov; claesand@fy.chalmers.se
†Email address: frtkl@fy.chalmers.se; www.frt.fy.chalmers.se/folks/kristian/kristianeng.html
‡Email address: steen@lanl.gov; www.lanl.gov/home/steen
§Email address: roger@plato.ucs.mun.ca

tensively in the past for social modeling and introduced the concept of using interactions between activities on a lattice in models of urban growth and transportation [22–30]. These are in a sense similar to the approach developed here—for example, both assume that spatial structure and spatial interaction are mutually determined. However, the emphasis of these models is primarily on predicting interzonal flows rather than achieving an understanding of spatial structure.

The problems caused by the modern phenomenon that is commonly referred to as ''sprawl'' provide ample reasons for understanding the fundamental factors behind urban growth. Urban sprawl is the cause of many urban miseries such as biotope fragmentation, long transportation times, smog, traffic congestion, destruction of fertile farmland, and other environmental issues. Because of this, it also attracts much attention from researchers and policy makers [31–33]. The question of how to shift development towards ''smart growth'' instead of sprawl is hard to address if the underlying dynamics remain a mystery. The ''unwilling neighbor'' (UN) rule that we present within our framework in this paper is arrived at by extracting the simplified essence out of economic factors governing a common mode of development. It is based on the assumption that there (i) is a benefit in being a part of the infrastructure network and (ii) land price is generally related to development density. We motivate our rule in more detail in Sec. III A 1 and discuss the validation process in Sec. III A 3.

II. MODEL DEFINITION

Each model component corresponds to elements of the real world urban system of which we are representing the dynamics. In each of the following sections we argue for the basic structure of our framework.

A. Basic dynamics

Formally, our simulation framework consists of a modified two-dimensional (2D) Markov random field (MRF) representation of the site-to-site interactions using a recursive mean-field approach to take into account interactions from not only neighboring sites, but from all lattice sites. The state transitions at the individual sites are determined by a global (probabilistic) selection criterion, as in evolutionary selection, and not by a local selection criterion as in the classical MRF.

If we think of undeveloped lattice points as being in the empty state, only addition is modeled in this paper. Since modern cities grow rapidly, both from an increasing population and as a result of a larger fraction of the population living in cities, removal is also relatively uncommon in cities of appreciable sizes. Basically, the model employs an evolutionary dynamics with allocation of new urbanization at the most ''fit'' location [34]. A generalization of this model including removal (transitions to the empty state) is discussed in [35].

B. Lattice and land use

We represent land by a two-dimensional grid consisting of N square cells of equal size. Each cell corresponds to an area and its (discrete) state represents what is on it—its land use. This means that we divide all possible uses of land into classes that are assumed to be homogeneous within each cell. In the simplest case, only two classes are considered: Built and rural. However, this setup can accommodate any resolution of land use since there is no limit to how many states a cell can take.

C. Markov random field

A classical MRF [34] representation of a 2D land use dynamics may be defined as follows: From a set of c land uses, consider two different land uses $a,b \in \{1,2,...,c\} = C$. The maximum radius of land use to land use influence is R. The potential (''energy'') of land use class a at a given location x is

$$E_a(x) = \left(\sum_{d \leq R} \sum_{b \in C} w_{ab}(d) \right), \tag{1}$$

where $w_{ab}(d)$ is the positive or negative influence (''energy contribution'') from land use class b on land use class a at distance d.

We want to translate potential energies into probabilities and be able to continuously tune the model's sensitivity to the state. A reasonable way of realizing this is to use the Gibbs weight function that gives us the probability of finding a site at position x in state a as

$$p_a(x) = \frac{F(E_a(x))}{\sum_{b \in C} F(E_b(x))}, \tag{2}$$

where F is a Boltzmann transformation

$$F(E_a(x)) = e^{-\beta E_a(x)}, \tag{3}$$

where β is a free parameter that corresponds to $1/T$ where T can be viewed as the temperature of the system. Please see Secs. IV and III A 3 for discussion about interpretations of temperature for this system. Equations (1) through (3) define a Gibbs random field (or spin glass).

Although the fundamental idea behind our approach is as described above, our model differs in two significant ways: (i) We take the entire geographic region into account in the interaction rather than just a small radius neighborhood and (ii) the state transition probabilities are defined globally and not locally. This is described in the following section.

D. Extended MRF for simulation of urban growth

1. Modifications

In an effort to make the model as simple as possible (but not simpler than that) we have extended the MRF model to use transition intensities (demand model) that are defined externally rather than internally. The reason for this is that the supply and demand cycle of urban economics is too complicated to be captured by a simple model using only pairwise site interactions. If no regional resource market is introduced in the energy function (which could be done by, e.g., using multicell interaction), the simulations are highly unstable. The problem of simulating growth by using internally generated growth intensities is currently being investigated by the authors. We will in the following paragraphs give a background to modifications and choices that have been made.

In a MRF model, the action of making a transition is decided from the viewpoint of the lattice point in which it takes place. For a number of reasons, generation of urban morphology is better viewed as a process of regional allocation:

(i) Information can be assumed to travel over the lattice at time scales far shorter than that of a lattice update, so in each choice of development site every lattice point is a candidate.

(ii) We simulate only a small part of a larger system, so there will be a background that drives the model.

The incorporation of long-range interactions (with some metric) is intuitively a sound addition for simulating urban growth. Correlations in cities are clearly longer than nearest-neighbor and the system is not changing on a time scale that would allow longer-range interactions to emerge from the dynamics. [Note that information in a nearest-neighbor cellular automaton (CA) travels via state transitions.] The use of mean fields is motivated by a combination of reasons: (i) Intuitively, hierarchical scales are used by humans for conceptualizing successively larger areas. (ii) Its applicability in the urban growth context is indicated by central-place theory [19–21] in which this conceptualization is formalized. (iii) It is a well-established mathematical modeling technique.

2. Definition of modifications

With a reformulation of Eq. (2), which is an often-used trick for grid updates in MRF models, we get the probability $q(x)$ for a given lattice site x to be the next site that is updated,

$$q(x)=\frac{\sum_{a\in C} e^{-\beta E_a(x)}}{\sum_x \sum_{b\in C} e^{-\beta E_b(x)}}. \tag{4}$$

The probability for this lattice point x to both be the next site to be updated, and that the transition will be into state a is then given by

$$q_a(x)=\frac{e^{-\beta E_a(x)}}{\sum_x \sum_{b\in C} e^{-\beta E_b(x)}}. \tag{5}$$

This can be written as

$$q_a(x)=\tilde{p}_a(x)I(a), \tag{6}$$

where

$$\tilde{p}_a(x)=\frac{e^{-\beta E_a(x)}}{\sum_x e^{-\beta E_a(x)}}, \tag{7}$$

$$I(a)=\frac{\sum_x e^{-\beta E_a(x)}}{\sum_x \sum_{b\in C} e^{-\beta E_b(x)}}. \tag{8}$$

Here, $\tilde{p}_a(x)$ is the allocation submodel and defines the relative (to other sites) intensity of change of lattice site x into land use class (state) a in the next update. The intensity function $I(a)$ is the demand submodel, and a global intensity for the development events that result in land use a. In this paper we have used an externally defined constant, $I(a)=k_a, a\in C$, in the simulations since the assumptions we would have to make to embed a model of endogenously defined intensities (a market model) would only add unnecessary detail. This demonstrates the modifications that we have made from the original MRF formulation.

3. Mean fields

The fundamental lattice with N lattice points is referred to as the level-0 grid and it is the grid with the highest resolution. Grids at higher levels l of aggregation have cells that are mean fields (aggregates) of progressively larger concentric portions of the level-0 grid. Thus, an l-level cell has contributions from 3^2 times as many level-0 cells as a $(l-1)$-level cell. Starting from the most coarse grained, or aggregated, level L where the whole lattice is aggregated, 3^2 new subgrids are generated for each recursion and thus $N=(3^2)^L$ and $L=1/2\log_3 N$. This indicates that $1/2\log_3 N$ recursive lattice averaging operations are needed for the update of each site. L then defines the depth of the lattice.

Starting at level 0 we define $c_a^{(0)}(i,j)$ as the cell count of activity a at location (i,j). At level 0, each lattice site has one state s, which indicates the land use class to which the site belongs. Hence, for level 0 lattice sites, one activity, $a=s$, will be unity and all other activities, $a\neq s$, zero. Then

$$c_a^{(1)}(i,j)=\sum_{k=-1}^{1}\sum_{m=-1}^{1} c_a^{(0)}(i+k,j+m) \tag{9}$$

defines the cell count of activity a at level 1 and by induction it is seen that

$$c_a^{(l)}(i,j)=\sum_{k=-1}^{1}\sum_{m=-1}^{1} c_a^{(l-1)}(i+k3^{l-1},j+m3^{l-1}) \tag{10}$$

expresses the number of cells that carry activity a at level l defined from the cells at level $l-1$. The algorithm we use for updating the grid has a time complexity in $O(N\log N)$.

E. Activity interactions

The value of the energy function depends on the number of built cells in the neighborhood and the interaction function. By changing the energy function we can have detailed control over the microdynamics. The growth rule that we use in this paper can be described in simple terms as edge growth in addition to mutual inhibition. In reality, this corresponds to the advantage of hooking into existing infrastructure, see Sec. III A 1. The amount of penalty received for developing away from an edge is tuned with a parameter and thus in the special case where no penalty is given, edge growth is not preferred. For the analysis carried out in this paper we have used two land use classes (states): undeveloped and developed. In the equations these are referred to as states 0 and 1, respectively. Under the unwilling neighbor (UN) rule, two distinctly different types of interactions are modeled: One for nearest-neighbor interactions and one for all distances beyond. More generally, the form of Eq. (13) is used and the parameter ξ can be specified for combinations of land use classes and distances. For nearest-neighbor influence in the UN rule, the energy function is defined as follows:

$$h_1^{(1)}(c_1^{(1)}=0)=-\ln\epsilon,$$

$$h_1^{(1)}(c_1^{(1)}>0)=-\ln(1-\epsilon)-\xi c_1^{(1)},$$

$$h_0^{(1)}(c_1^{(1)}=0)=H-\ln\epsilon,$$

$$h_0^{(1)}(c_1^{(1)}>0)=H-\ln(1-\epsilon). \quad (11)$$

Here, ϵ is the parameter that controls the extent to which developing near other development is beneficial. As discussed in Ref. [35], ϵ is a number that is typically very small. Note that $\epsilon=0.5$ corresponds to no preference for edges and as ϵ tends to 0 we get only edge growth. Depending on whether we want the microscopic dynamics to correspond to mutual inhibition or stimulation of further development we choose $\xi>0$ or $\xi<0$, respectively. We do not model transitions from developed land to undeveloped land and assume the energy penalty for such a transition to be sufficiently large, $H\rightarrow\infty$.

For long-range interactions on levels $1<l\leqslant L$ [Eq. (10)], we let the interaction strength decay exponentially with distance. This exponential decay $d^{(l)}$ [3,36] is common for all land use classes and is defined as

$$d^{(l)}=3^{-2l}, \quad (12)$$

which exactly accounts for the exponential increase of cells at each recursive level of interaction. We thus define the energy contribution to the long-range interactions for $l>1$ as

$$h_1^{(l)}(c_1^{(l)})=d^{(l)}\xi c_1^{(l)} \quad (13)$$

with $1<l\leqslant L$, which corresponds to w_{ba} in Eq. (1). Here, $c_1^{(l)}$ is the count of cells in the built state at mean-field level l.

F. Transitions

Conceptually, the activity addition dynamics is defined through a global selection of cells based on their "fitness." We calculate $E_a(x)$ as a sum over the energy contributions $h^{(l)}$ from all mean-field levels l, as described in Eqs. (12) through (13). The transition probabilities are then obtained using the modified MRF formulation in Eqs. (4) through (8).

III. GROWTH PATTERNS

Because urban growth has a preference for taking place around the edges of already urbanized areas, it has a natural connection to statistical physics of clustering. For example, diffusion-limited aggregation, dielectric breakdown and correlated percolation [3] have all contributed to the understanding of urban growth by providing minimal abstract models that capture important aspects of the target system. As noted by Makse *et al.* in Ref. [3], the DLA model has many shortcomings when adopted to urban systems. Most notably, the components of the model lack intuitive counterparts in the real system and, as pointed out by Makse *et al.* it predicts a single cluster. Instead, correlated percolation was proposed as a model that more realistically depicts the dynamics resulting from how growth attracts further growth. The results presented by Makse *et al.* are in better agreement with empirical measurements than what is the case for the simplistic DLA and DBM models. However, the microscopic dynamics are still not consistent with common knowledge about urban growth; an urban core still has to be defined and the configuration is primed with a density that decays from the core.

Measurements on simulated and real configurations show (see Figs. 4 and 5) [1,37–39] that scaling is present over some orders of magnitude. This relation would indicate some distributed growth process capable of producing such characteristics. It has been suggested that the growth mechanisms result in self-organized criticality [1,6,7,40], but it need also be noted that urban development is planned and executed at all levels, from the building of a new garage to the restructuring of entire city parts [41], something that is a reflection of a hierarchical structure in the decision-making system. This is also a credible mechanism by which correlations over many length scales can be introduced.

A. Counteracting inhibition and stimulation: Unwilling neighbors

1. Rationale and introduction

The UN rule is based on counteracting inhibition and stimulation on different length scales. The stimulating influence from proximity to edges comes from the coordination benefits gained by being attached to the rest of the structure and the inhibiting influence comes from local competition. In the real urban system, these forces correspond to the benefit of being attached to infrastructure such as roads and utilities, and the disadvantage of high land prices in highly developed areas. The reason, the authors believe, that this has had an increasing impact on urban growth dynamics is that the means of personal communication have become considerably cheaper and more efficient over the last century while building infrastructure still is far beyond the budget of any but the largest companies.

The aggregative nature of the urban cluster has been noted earlier, and physical models of aggregation that are

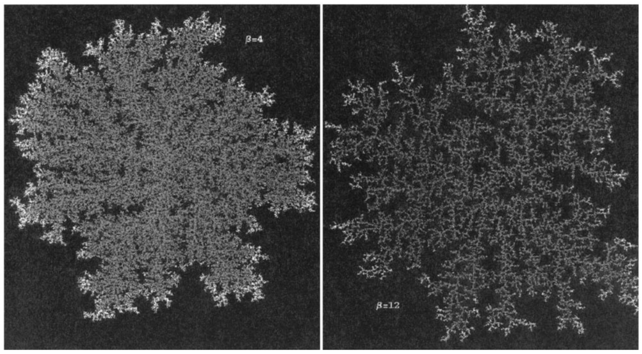

FIG. 1. The images shown above have been used for calculating the scaling relationship between radius and area of the agglomerate. This measure is often referred to as the radial fractal dimension. The effects of increasing the degree of randomness in the mapping, by varying the inverse temperature β, can be seen in the density of the structures. However, it is evident (see Fig. 4) that the fractal dimension does not change. To visualize the growth dynamics the color of added cells in the figure is gray up to a certain point in time after which they are white. This serves to visualize how new sites are selected given knowledge about the present state and how the β parameter affects this dynamics. The parameters used are $\xi=1$, $\epsilon=10^{-9}$.

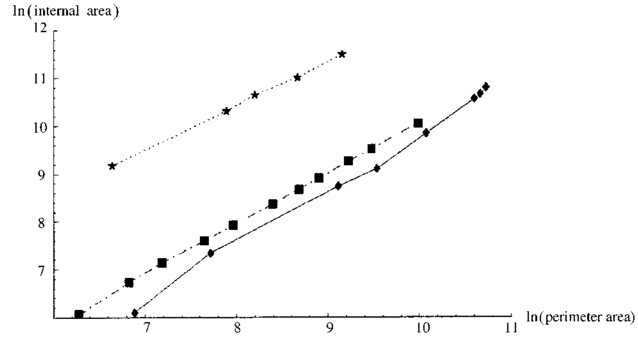

FIG. 2. A comparison between how developed area increases as a function of the perimeter. The curves belong, from the topmost downwards, to measurements from: Sioux Falls, USA, simulation with UN rule and Washington/Baltimore, USA. Area is defined here as the number of cells that are in the developed state and the perimeter is all nondeveloped cells that are adjacent to developed cells. Along the lines of our model, perimeter constitutes the cells that constitute the primary growth zone. This is because they do not receive a penalty from the ϵ rule according to Eq. (12) and Ref. [2]. A scaling exponent that is not trivial (such as an expanding disk) requires a distributed mechanism that makes the structure sparse. We have used data from the growth of Sioux Falls and Washington/Baltimore to study how cities grow in this fashion. The parameters used for the model are $\beta=4$, $\xi=1$, $\epsilon=10^{-9}$, $N=315\times315$. On the X axis is the logarithm size of the perimeter and on the Y axis is the logarithm size of the developed area.

capable of generating fractal noncompact clusters such as diffusion-limited aggregation (DLA) and correlated percolation models have been employed to generate urban morphology. Using the UN rule in isolation (as the only developed land use class) the results are very similar to DLA (see Fig. 1) and a combination between several classes yield the more compact and realistic clusters of the correlated percolation used by Makse *et al.* (see Fig. 2). The important difference is that our model is defined from a microscopic formulation while the models mentioned earlier are motivated by preknown macroscopic semblance and a rather loose microscopic similarity in that they are aggregation models. The detailed microdynamics are very far from that of real city growth, i.e., there is no correspondence between the random walkers in a DLA and the mechanism by which development demand is allocated to new lots in the real urban system. Another important effect of the formulation we use is that seeding is unnecessary for growth to start. In the simplest case, the first settlement will take place anywhere with equal likelihood if no development is present to break the symmetry. In more complex setups (see Fig. 2) other factors such as topography and roads are present and affect the growth distribution even when no development is present. The impact that this has no the model validation is discussed further in Sec. III A 3.

2. Dynamics

As was noted in the preceding section, configurations produced using our model with the UN rule have a striking resemblance to configurations generated with DLA models (see Fig. 1) until the lattice is crowded, at which point the characteristics of the growth dynamics changes (see Fig. 3).

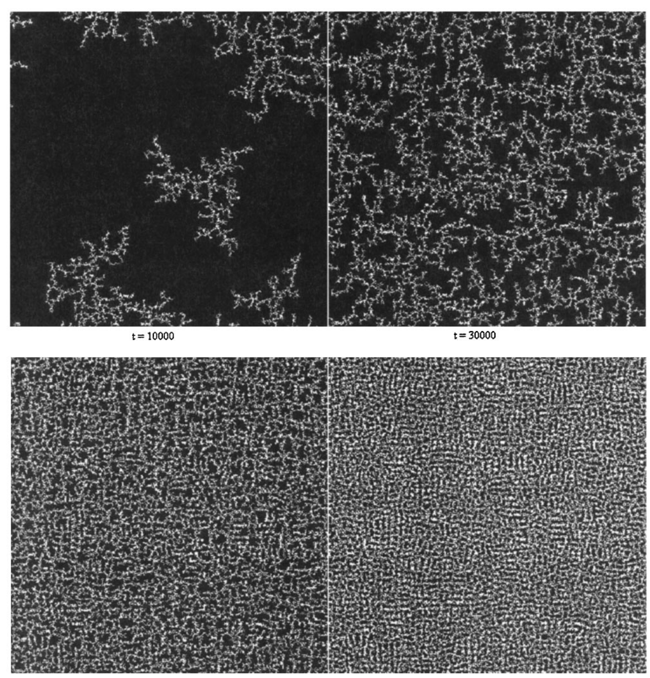

FIG. 3. When the lattice gets crowded, the growth dynamics changes from one where growth takes place from a central core and outwards to one where the lattice is successively filled uniformly over the whole lattice. The first image that is the least filled is similar to a DLA while the following gradually deviate from that to approach a two-dimensional pattern. The parameters used are $\beta=16$, $\xi=1$, $\epsilon=10^{-9}$, $N=415\times415$.

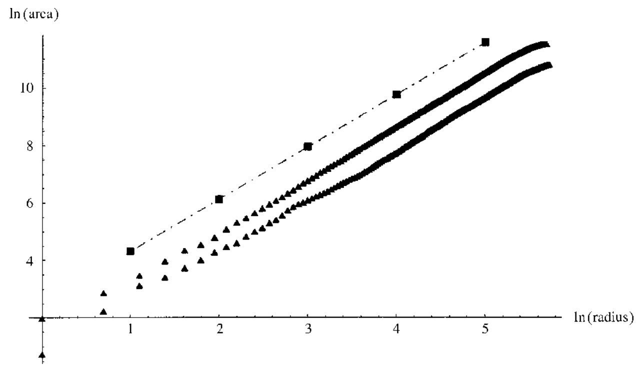

FIG. 4. A double logarithmic (ln) plot of area as a function of distance for the images in Fig. 1 shows that the scaling relation is independent of β. From top to bottom (triangles), β values of 4 and 12 are plotted, respectively. The scaling relation between area and radius is often referred to as the radial fractal dimension and provides us with a measure on the density of the cluster. To be able to conveniently measure this property we have here used configuration grown from a central seed, even though our model is capable of tuning the probability for addition away from the cluster. Plotted as a reference is a line with a slope of 1.815 indicating that the fractal dimension of the sets is close to this value. It can further be noted that the same measure applied to an unbiased DLA yields very similar results.

However, the radial fractal dimension (area as a function of radius) is not sensitive to the value of β (see Figs. 1 and 4).

Comparisons between measurements of the radial dimension of DLA and UN configurations confirms the visual similarity. DLA captures well the short-range attraction by aggregation of development on infrastructure (transition of liquid to solid in DLA) but the dendritic structure that it produces as a result of diffusion-limitation is coincidental since no conceivable units in the urban system behaves similarly to the random walkers in the DLA model. Apart from having a seemingly more realistic microdynamics, the UN rule is formulated in a framework that also allows the seamless integration of other rules that can be ascribed to other growth dynamics (see Fig. 2).

3. Validation

The assumptions on which we base the inhibition/stimulation growth rule that we call the UN rule is discussed earlier in the paper in Sec. III A 1. The aim of validating the model's macroscopic behavior is to provide credibility to the correctness of the microscopic rule in question. Although the model is capable of combining an arbitrary number of individually behaving land use classes (see Fig. 2), we study the case where the UN rule determines the growth of the sole growing land use class on a uniform background. An automatic method for comparing macrostates of an urban model will be blunt compared to similar measures for physical systems since relevant state variables are hard to formulate and measure; the analogies we use are far from being as powerful as they are for physical systems. However, despite this fact there are observables that have proven useful for doing this.

So far, one type of observable has proven useful for such notorious aggregate systems as rivers, cities, and crystals: specifically the scaling relationships in their geometry. Just as in the case of Euclidean objects, growing structures in nature often exhibit a scaling relationship between their dimensions [42]. However, rather than simple relationships with integer exponents, they exhibit scaling with noninteger exponents—this is called a fractal geometry and urban clusters are among the systems that have been shown to exhibit this property. Specific discussions about fractal urban geometry and what causes it can be found in, for example, Refs. [37,38,43,44]. We have selected two observables that seem particularly meaningful because there is a clear connection between them and important aspects of the growth dynamics. They have also been used by other groups, so there are values in the literature with which we can compare our results.

The first observable is the scaling relationship between area and radius, which for a structure on a plane will be between 1 and 2, or, in other words, between a line and a disc. While cities, especially large urban agglomerates, are definitely multicentered, much growth takes place from a central core and outwards. The actual exponent values turn out to vary too much between cities to be of intelligible significance, it is rather the fact that all published measurements do exhibit good scaling properties that is important. The scaling properties of our simulations are robust and independent of β values, see Fig. 4 [38] which is also the case for DLA. Note, however, that this does not mean that configurations grown with different temperatures are equivalent, in fact the structures easily become compact because the side of a cell is fairly large compared to the entire area that is being simulated.

What this means is, among other possible interpretations, that for a structure in which the builders' behaviors are close to the behavior predicted by the UN rule (low temperature), we will see more low-density growth whereas if factors external to the rule are important (high temperature), we would expect more deviation from it, i.e., dense structures. Such an interpretation would also make sense intuitively: A city that is centrally planned (typically older parts of cities) is much more dense than one where individual builders independently can attempt to maximize their investments (see Fig. 5).

The other observable that we study is the relation between the perimeter and the inner area of a configuration. This is related to the former measure but is aimed at the state of the growth rather than at the state of the cluster. We compare the time evolution of our simulations with that of real urban regions: Sioux Falls, USA and Washington/Baltimore, USA. This is a scaling relation that has been verified from empirical data [37] and is repeated in our measurements. The presence of scaling is also found in our simulations and the exponent is similar to that of the real regions (see Fig. 6) for which we have made measurements. It should be noted that this is true despite the fact that an actual urban footprint incorporates all growth mechanisms that are active while the compared output from our model only employs the rather idealized UN rule (see Fig. 7).

IV. DISCUSSION

The correspondence between model components and entities of the real world needs to be carefully considered, es-

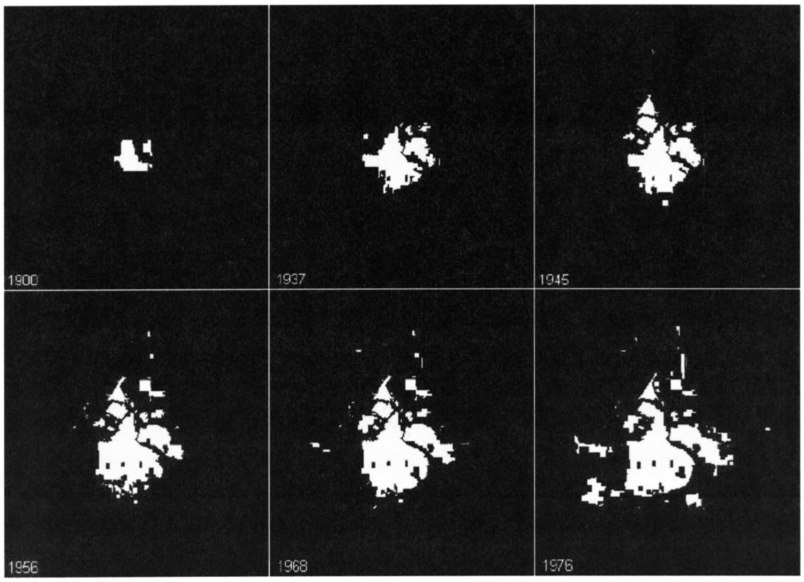

FIG. 5. As a real-life comparison to the model for measuring the scaling relationship between area and perimeter in the growth of a city, we show here the growth of Sioux Falls between 1900 and 1976. See Fig. 6. (The data used are courtesy of N. Goldstein, Dept. of Geography, UCSB.)

pecially when models are used over the boundaries of research disciplines. In physical models, model components are always carefully founded in knowledge about the system that is being modeled; this does not, however, translate as easily as one would wish to complex aggregate systems. It is, therefore, not surprising that a social model would need to be similar, but not identical, to a physical model that more exactly describes a much simpler system. It also should be of little surprise that the formulation could be on a level that is intuitively very easy to understand since we are indeed tuned as biological beings for grasping concepts on the social scale; we can more easily relate to a sibling situation than to a covalent bond between two atoms. Along the same lines, we can easily relate to the process of selecting a place to build a house whereas the process of an electrical discharge requires education to understand.

In the model presented here, we have used analogies from physics for performing basic actions on an appropriate scale in time and space: Demand for land adaptation is met by allocation according to a measure of goodness by which can-

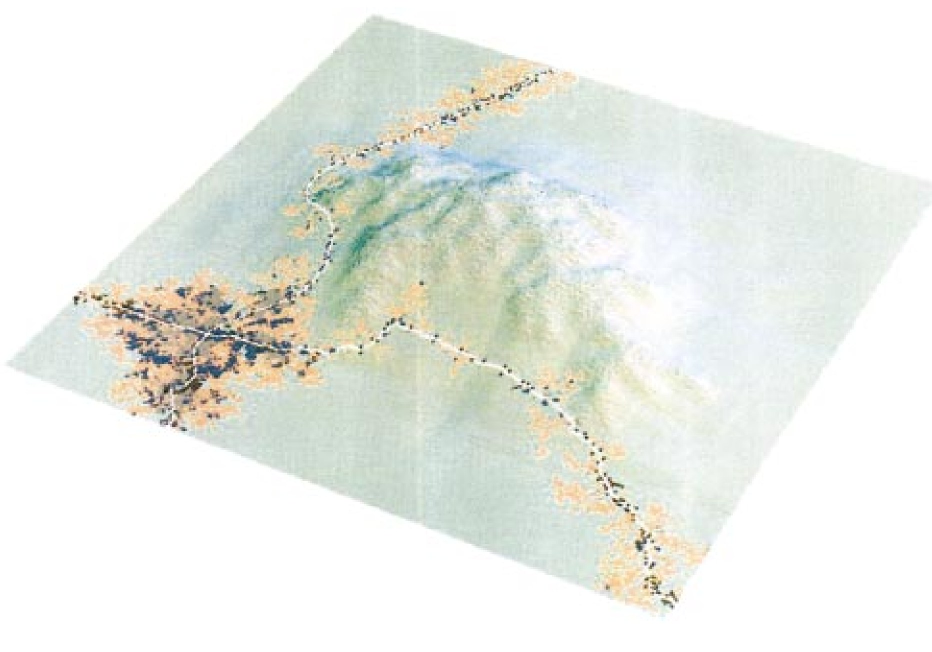

FIG. 6. (Color) The framework is capable of incorporating a wide range of additional information such as topographical maps and statical land use classes that are static throughout the simulation such as the limited-access highways in this example. The dynamics of the above configuration can be explained by investigating each land use and the interaction functions that are used. In particular, Eq. (13) is generalized such that ξ can be separately defined for each mean-field level and land use pair. Some of the land uses used in this image are stimulating to other land use classes on some distances and inhibiting on other, a detailed account of this is beyond the scope of this article and this image is only used as a visualization of the open-endedness of the model. Legend: orange is residential, gray is central business district, blue is commercial, black is industry, white is highways, and green/shaded relief is undeveloped land.

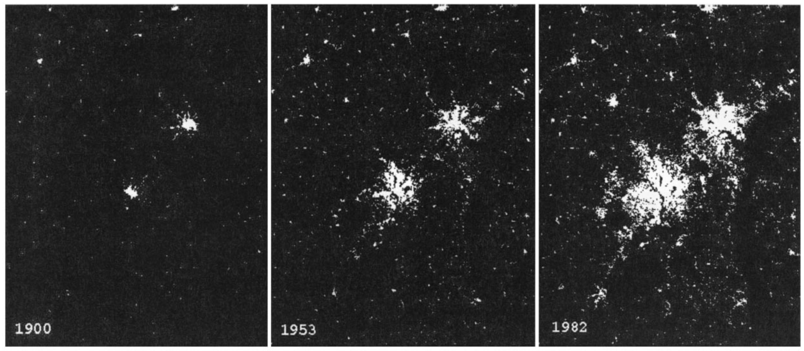

FIG. 7. Shown here is the urban extent of the Washington/Baltimore area over time [18]. These images along with others in the same series were used to obtain the area-perimeter scaling relationship of the growth of the urban agglomeration. See Fig. 6. Note that an actual urban footprint, such as those shown here (and in Fig. 5), results from very complicated growth dynamics. This is in contrast to the rather simplistic UN rule and more complex rules are needed to generate urban patterns that also visually look like a complete city. Compare with Fig. 2 for a realistic (and complex) example of simulated urban growth.

didate sites can be compared. This is done in a fashion that is commensurate with existing theory and because of this we can also realize the meaning of details of the model in its new context. For example, the parameter β has a meaning in the Boltzmann transformation when applied to physical systems and should have a meaningful interpretation also in an urban system if we use it there. Due to obvious differences between models, the connection to thermodynamic definitions of temperature is hard to make and it lies closer to the interpretation of temperature as a characterization of information deficiency. This would translate into saying that temperature (noise) accounts for the parts of the system for which we do not have a model.

A concrete example in the context of urbanization would be the following: Consider two sites whose energy turns out to be identical under the energy function of the model and that are located some distance apart. Now, in a real scenario, picture a family evaluating candidate sites for building or buying a house and that the two mentioned sites are at the top of their list. If we further assume that they happen to work at a third site that is closer to either of the two under consideration, they would probably select that site. What this means is that the model we use to measure the suitability of sites is an approximation of their internal model; each agent has detailed criteria that are unique to them. The application of temperature corresponds to the concept of a maximum-entropy model formulation [23].

It should also be noted that the fact that the temperature is nonzero is absolutely essential for bringing about interesting dynamics. A zero temperature would mean that the site that is the most suitable according to the model would be selected with probability 1, or $1/n$ if there are n equally suitable places. In the case of a the "unwilling neighbor" rule it is easy to see that this would result in growing straight lines since the edge pixels are those that are furthest from the rest of the structure. Actually, the transition from $D>1$ to $D=1$ growth for DLA does not take place in the limit of no randomness but rather much earlier [45].

A further benefit from the compartmental approach we have used to define the model is that a generalization to other types of colonies is conceivable. The dual inhibition/stimulation of a common framework is universal to many situations in nature where limited resources have to be coutilized by many individuals. Sprawling forms of colonies are abundant in nature. Colonies of sessile marine animals such as corals and barnacles are obvious examples, but so also bird colonies and grazing herds may be examples where a counteracting inhibition/stimulation on different length scales exists.

V. CONCLUSION

The model presented here reproduces realistic macroscopic characteristics of real cities, similar to earlier published results based on aggregation models. However, rather than being based on potentially coincidental macroscopic similarity, the presented model is built from the bottom-up situation based on a microscopic formulation. It thereby serves to validate the hypothesis that frustrations caused by a combination of stimulation and inhibition, resulting from an interplay of development intensity and distance, might be in part responsible for the growth of urban sprawl. The model can provide an urban growth simulation framework that is configurable, scalable, and capable of rich dynamics while still being mathematically transparent in its formulation.

[1] S. C. Manrubia, D. H. Zanette, and R. V. Sole, Fractals **1**, 1 (1999).
[2] R. White and G. Engelen, in *Self-Organization of Complex Structures: From Individual to Collective Dynamics*, edited by Frank Schweitzer (Gordon and Breach, London, 1997), pp. 519–535.
[3] H. A. Makse, J. S. de Andrade, M. Batty, S. Havlin, and H. E. Stanley, Phys. Rev. E **58**, 7054 (1998).
[4] H. A. Makse, S. Havlin, and H. E. Stanley, Nature (London) **377**, 608 (1995).
[5] M. Batty, Envir. Plan. A **23**, 511 (1991).

[6] D. H. Zanette and S. C. Manrubia, Phys. Rev. Lett. **79**, 523 (1997).
[7] D. H. Zanette and S. C. Manrubia, Intermittency model for urban development. Phys. Rev. E **58**, 295 (1998).
[8] L. Benguigui, Physica A **219**, 13 (1995).
[9] F. Wu, Environ. Plan. B: Plan. Des. **25**, 731 (1998).
[10] F. Wu, Int. J. Geograph. Inf. Sci. **12**, 63 (1998).
[11] G. Engelen, R. White, and I. Uljee, in *Decision Support Systems in Urban Planning*, edited by H. Timmerman (Spon, London, 1997), pp. 125–155.
[12] T. W. Foresman, http://biology.usgs.gov/luhna/chap5.html.
[13] K. C. Clarke and L. Gaydos, Int. J. Geograph. Inf. Sci. **12**, 699 (1998).
[14] K. C. Clarke, S. Hoppen, and L. Gaydos, Environ. Plan. B: Plan. Des. **24**, 247 (1997).
[15] J. Candau, S. Rasmussen, and K. C. Clarke, in Proceedings of 4th International Conference on Integrating Gis and Environmental Modeling (GIS/EM4) Banff, Alberta, Canada, 2000 (unpublished).
[16] H. Couclelis, Envir. Plan. A **V17**, 585 (1985).
[17] M. Takeyama, *Artificial Worlds Urban Studies* 347 (1996).
[18] I. Benenson and J. Portugali, Agent-based Simulations of a City Dynamics in a Gis Environment, COSIT 1997, pp. 501, 502 (unpublished).
[19] W. Christaller, *Central Places in Southern Germany* (Prentice-Hall, Englewood Cliffs, NJ, 1933).
[20] R. White, Geogr. Anal. **10**, 201 (1978).
[21] R. White, Geogr. Anal. **9**, 227 (1977).
[22] D. Helbing, *Quantitative Sociodynamics: Stochastic Methods and Models of Social Interaction Processes* (Kluwer, Dordrecht, 1995).
[23] A. G. Wilson, *Entropy in Urban and Regional Modeling* (Pion Ltd., London, 1970).
[24] M. Batty, *Urban Modelling: Algortihms, Calibrations, Predictions* (Cambridge University Press, Cambridge, U.K., 1976).
[25] R. G. Golledge and R. J. Stimson, *Spatial Behavior: A Geographic Perspective* (The Guilford Press, New York, 1997).
[26] W. Weidlich, *Sociodynamics: A Systematic Approach to Mathematical Modelling in the Social Sciences* (Harwood Academic Publishers, Chur, Switzerland, 2000).
[27] W. Weidlich, Discrete Dyn. Nat. Soc. **3**, 137 (1997).
[28] K. Nagel, S. Rasmussen, and C. L. Barrett, in *Self-Organization of Complex Structures*, edited by F. Schweitzer (Gordon and Breach, London, 1997).
[29] D. Yamins, S. Rasmussen, and D. Fogel, *Netw. Spatial Econ.* (to be published).
[30] Transims: http://transims.tsasa.lanl.gov/.
[31] W. Fulton, R. Pendall, M. Nguyen, and A. Harrison, Who Sprawls Most? How Growth Patterns Differ Across the U.S., The Brookings Institute Survey Series, July (2001) (available from Brookings Center on Urban and Metropolitan Policy).
[32] B. McCann, STPP—Surface Transportation Policy Project, 2001 (unpublished).
[33] J. G. Mitchell, Natl. Geogr. **200**, 48 (2001).
[34] Contemp. Math. **1**, (1980), edited by R. Kindermann and J. L. Snell.
[35] C. T. Andersson, S. Rasmussen, and R. White, Envir. Plan. B (to be published).
[36] C. Clark, J. R. Stat. Soc. Ser. A. Gen. **114**, 490 (1951).
[37] M. Batty and P. Longley, *Fractal Cities: A Geometry of Form and Function* (Academic Press, London, 1994).
[38] R. White and G. Engelen, Envir. Plan. A **25**, 1175 (1993).
[39] H. E. Stanley, L. A. N. Amaral, S. V. Buldyrev, A. L. Goldberger, S. Havlin, H. Leschhorn, P. Maass, H. A. Makse, C.-K. Peng, M. A. Salinger, M. H. R. Stanley, and G. M. Viswanathan, Physica A **231**, 20 (1996).
[40] P. Bak, *How Nature Works* (Springer-Verlag, New York, 1996).
[41] A. Daffertshofer, H. Haken, and J. Portugali, Environ. Plan. B: Plan. Des. **28**, 89 (2001).
[42] B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman, New York, 1988).
[43] G. Haag, Chaos, Solitons Fractals **4**, 519 (1994).
[44] L. Benguigui and M. Daoud, Geogr. Anal. **23**, (1991).
[45] M. B. Hastings, XXX LANL Condensed Matter Archive, 2001 (unpublished).