
Generating settlement structures: a method for urban planning and analysis supported by cellular automata

Reinhard Koenig, Christian Bauriedel

Faculty of Architecture, Bauhaus-University Weimar, Belvederer Allee 1, D-99423, Germany;
e-mail: koenig@entwurforschung.de, christian.bauriedel@archit.uni-weimar.de

Received 5 March 2007; in revised form 29 October 2007; published online 26 January 2009

Abstract. Previous models for the explanation of settlement processes pay little attention to the interactions between settlement spreading and road networks. On the basis of a dielectric breakdown model in combination with cellular automata, we present a method to steer precisely the generation of settlement structures with regard to their global and local density as well as the size and number of forming clusters. The resulting structures depend on the logic of how the dependence of the settlements and the road network is implemented to the simulation model. After analysing the state of the art we begin with a discussion of the mutual dependence of roads and land development. Next, we elaborate a model that permits the precise control of permeability in the developing structure as well as the settlement density, using the fewest necessary control parameters. On the basis of different characteristic values, possible settlement structures are analysed and compared with each other. Finally, we reflect on the theoretical contribution of the model with regard to the context of urban dynamics.

1 Introduction

When examining urban growth we focus on the question of how the underlying principles of the development of human settlements can be transferred into mathematical models and algorithms. This is not only about the explanation of a certain condition but also about the temporal development as well as the transitions of one condition to another, and finally about the influence of political, social, economic, and ecological frameworks, which are reflected in the control parameters. Generating or growing a watched phenomenon in silicio is considered by Epstein and Axtell (1996) as a new method of scientific discussion, as a generative explanation:

“Perhaps one day people will interpret the question ‘Can you explain it?’ as asking ‘Can you grow it?’” (Epstein and Axtell, 1996, page 20).

The agent-based model built by Dean et al (2000) to throw light on the development of the prehistoric Anasazi culture in a restricted area of northwest Arizona is a good example of the combination of ecological, economic, social, and religious conditions with changes in the corresponding settlement pattern. We will come back to the possible consequences of these different influences at the end of this paper.

This paper contributes to the search for a general generative explanation of the evolution of certain regional and urban settlement structures. In the development of the model particular attention was given to the control parameters that allow one to manipulate the artificial growth processes. The principle of parsimony suggests that the number of control parameters be reduced as far as possible. However, this is often associated with the difficulty of linking the control parameters of an explanation model of urban development with the social, economic, ecological, and political conditions of reality. Take, for example, the aggregation models by Benguigui (1995) or Andersson et al (2002), in which urban growth depends largely on a noise or random value. Indeed the principle of parsimony is demonstrated in an exemplary fashion here, though exactly how this highly aggregate parameter can be interpreted remains unclear. Andersson et al (2002) interpret this parameter in a pragmatic but unsatisfactory

way as a part of a system for which no model currently exists. Batty (2005) regards such random variables as equivalent to suddenly appearing innovations.

Avoiding a rash reduction in complexity of the problem in this context, it seems reasonable to examine more closely the principle of plenitude. Consequently, in the following description, we will repeatedly address the question of a sensible connection between the control parameters and the conditions of reality.

2 State of the art

Dendritic growth without meshing is a general characteristic of models based on the principle of a diffusion limited aggregation (DLA) process. In Batty et al (1999) we find an example of a DLA model which describes the growth of several clusters whose respective roads do not connect within a cluster or between clusters, but which form dendritic structures. To achieve a meshing of the dendrites in aggregation models, further rules can be added. For example, a road network can be enlarged at each node in one of four directions. By executing this rule repeatedly, an automatically meshed network of orthographic, equally long street segments results (Batty and Xie, 1997). If the lengths of the segments are varied and the possible directions are not restricted to an orthogonal grid, a disproportionate amount of dead ends arise and the networking of the roads arises by coincidental overlapping. With their 'path node generator' Erickson and Lloyd Jones (1997) show another example, in which a test mechanism is introduced, which after adding a new street element checks if another street element is nearby and links the two where necessary. The meshing of the road system is essentially controlled by the rate at which house elements are placed along the roads. In this model the development of the settlement is determined by the successive aggregation of street elements, which can be attached either in front, or to the left or right, following a defined weighting. The ability to vary the size of the area enclosed by roads as well as the density of the settlement is limited in this model, which is why the resulting settlement patterns are always alike and reminiscent of informal structures of unplanned settlements.

The models specified so far were all based on an orthogonal cell grid and the cells can be in different states, essentially corresponding to the three land uses: free area, building, and street. Sembolini (2000) introduces a model with an irregular cell structure in which the cell edges and areas as well as the neighbourhood relations are defined by means of Delaunay triangulation or a Voronoi diagram. The centres of the cells form the basic framework of the road system, which is generated as an independent structure along the Delaunay triangulation. The expansion of the road network is subject to two conditions. Firstly, if at least one built-upon cell is located in the neighbourhood at the end of a street, the existing route is continued as straightly as possible. Secondly, if a node is already part of a street, and the number of street nodes in the wider surroundings falls below a certain threshold value, and the neighbourhood exhibits at least three service cells, two new street sections are attached to the existing road on the left and on the right in as orthogonal an orientation as possible. These rules yield a road system with exclusively four-way crossroads and with a geometry that is predefined by the arrangement of the cells. The land-use development is the generator for the expansion of the road network, since under both growth conditions at least one built-upon cell exists in the surrounding area. Another unusual feature of this model is the dynamically changing cell structure. Under certain conditions further cells are inserted into the system and the geometry of the surrounding cells is adapted. This can be understood as an approach towards the division of a plot that is not quite correct, as an original cell is not divided but rather a new cell is inserted between two existing cells. Although, by means of the

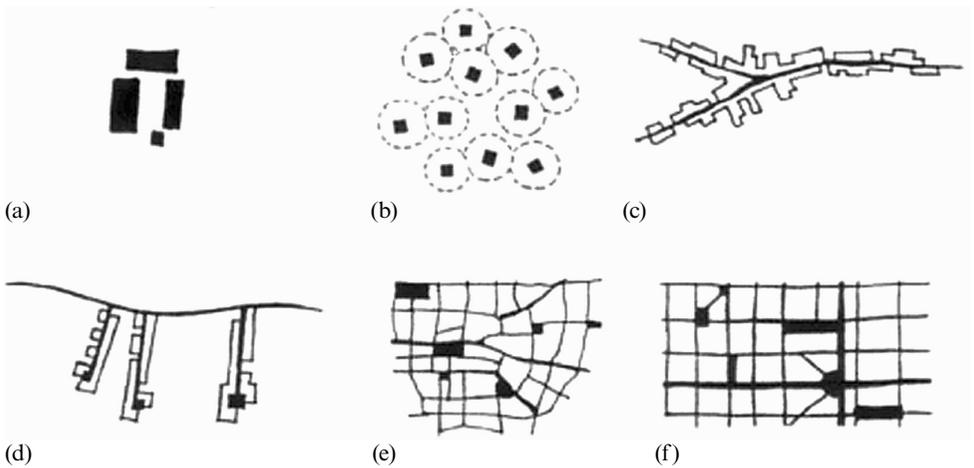


Figure 1. The six field types: (a) nucleus, (b) cluster, (c) highwayman, (d) boom type, (e) interlink, (f) plan (Humpert, 1992).

irregular Voronoi diagram, the topological neighbourhood relationships of a settlement can be expressed better, a sensible combination with the roadside environment is missing. Despite the relatively complex method of Delaunay triangulation, the geometric representation of the spatial arrangement is just as insufficient as with conventional regular grid structures. Both methods have their own advantages and disadvantages.

A further relevant approach for a generative theory is based on the ‘shape grammar’ principle. Compared with the additive logic of the aggregation models, the ‘shape grammar’ method offers the ability to divide a defined urban area into patches and to declare the resulting edges as streets. Duarte et al (2007) have extended the ‘shape grammar’ technique to an ‘urban grammar’ and have demonstrated the generation of an urban fabric using the example of a quarter in Marrakesh. This method produces very realistic results, whose geometry is not bound to any basic grid, but requires the implementation of a variety of micro rules for all combination possibilities of streets and plots. In Duarte et al’s (2007) paper, we can count sixty-five such rules, and accordingly these require the definition of many control parameters. Finally, it should also be noted that, with the ‘urban grammar’ method, the development of a settlement structure is primarily determined by the spread of the road system.

The theory of field types (Humpert, 1992; 1997) (figure 1) integrates some essential principles of urban development and offers a suitable systematisation for the model we will develop below. The nucleus type describes the beginning of a settlement and can be understood as a kind of seed. The cluster type corresponds to the informal agglomeration processes of unplanned settlements and also contains the division of a plot by the inclusion of further buildings or occupancies. The highwayman type describes passive settlement along existing roads and the boom type extends the highwayman type for active development of the hinterland through newly laid out roads. The transition process from the original pattern of fields and farm roads to the urban structure is described by the interlink type. Finally, the plan type describes the man-made design of the environment.

3 The road system

In the first step we deal with the generation of a road system. As a starting point we use a variation of the Eden (1960) model. In the middle of a regular cell grid one cell is occupied with an initialisation node (nucleus). With every step thereafter further nodes

are placed on randomly selected neighbouring cells and linked by an edge with existing nodes within a certain radius. For the formal notation we assign to every cell H an index $i \{i = 1, 2, \dots, n_H\}$. The cell in the middle of the grid is given the index c . Consequently, a cell H occupied by a node $N \{l = 1, 2, \dots, n_N\}$ is labelled with $N_l^{H_i} = 1$ and the state of a cell is written as S^{H_i} . Now the initial configuration can be expressed as:

$$N^{H_c}(0) = 1, \quad N^{H_i}(0) = 0, \quad \forall i \neq c; \quad S^H(0) = 0. \tag{1}$$

The cells in which further nodes can be placed are marked by their state. The states are defined as $S^H = 0$ for empty cells that do not neighbour an occupied or closed cell (empty), $S^H = 1$ for cells occupied with a node (occupied), $S^H = 2$ for cells which neighbour an occupied cell (closed), and $S^H = 3$ for empty cells that neighbour a closed cell (candidates). Further nodes can be placed only on candidate cells ($S^H = 3$).

For the definition of cells for a possible aggregation we deviate from the Eden model. If there is already a node in the Moore neighbourhood of a considered cell, it is closed so that the road system does not get too dense and a wider variation of the angles becomes possible [figure 2(a)]. As potential cells for aggregation (candidates), empty cells are indicated that have a closed cell in the Moore neighbourhood $U(H)$ with $k = 1$ (figure 3). The rule for the transition of a cell state is denoted as a state transition rule and is as follows:

$$S^{H(t+1)} = \begin{cases} 1, & \text{if } N^H(t) = 1, \\ 2, & \text{if } \sum_G \{1 | G \in U(H), H \notin G, N^G > 1\}, \\ 3, & \text{if } \sum_G \{1 | G \in U(H), H \notin G, S^G = 2\} > 1 \text{ and } S^H(t) = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

For the correct transformation of the states, rule (2) must be executed twice in succession. The various cell states are illustrated in figure 2(a). With each step t a new node is added on a randomly selected candidate cell, as defined by the development rule:

$$\begin{aligned} &\text{if } S^H(t) = 3 \text{ and } S^H(t) + e^H(t) = \max, \\ &\text{then } N^H(t+1) = 1, \end{aligned} \tag{3}$$

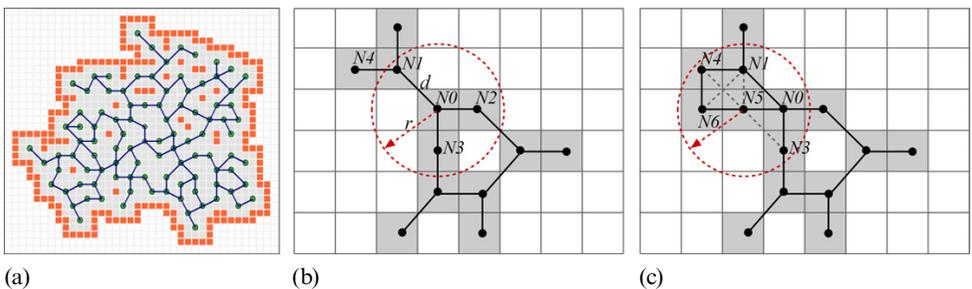


Figure 2. [In colour online, see <http://dx.doi.org/10.1068/b34025>] (a) States of the cells: empty (white), closed (grey), and candidate (red). The nodes are represented as points and the edges as lines. (b) and (c) Networking rule: to link two nodes they must be within a radius r and the shortest path between them via the existing graph must be greater than X steps.

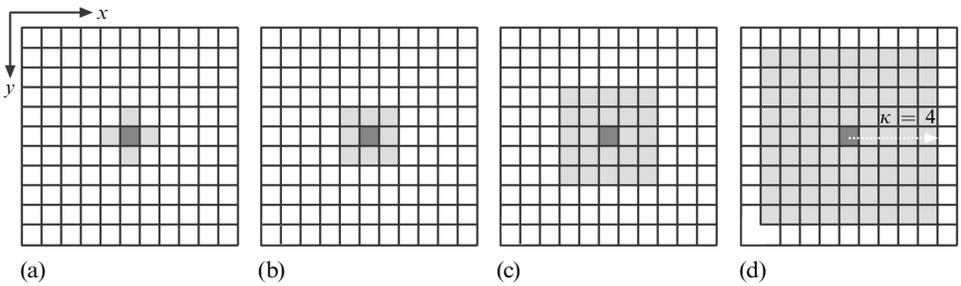


Figure 3. Different neighbourhood definitions $U(H)$. (a) Von Neumann neighbourhood; (b)–(d) Moore neighbourhoods with different neighbourhood sizes: (b) $k = 1$, (c) $k = 2$, and (d) $k = 4$.

where ε^H is a noise value in the unit interval $[0, 1]$. A networking rule [figures 2(b) and 2(c)] defines how a newly added node is linked to existing nodes and can be expressed formally as:

$$\text{if } (d < r \text{ and } K > X) \text{ then } E_{(Ni, Nj)}, \quad (4)$$

where d is the distance between the nodes. If there is another node located within a radius R , the rule checks whether the number K of nodes one has to pass through on the shortest path in a graph between start and target node is larger than X . The value X specifies the minimum required distance between two nodes Ni and Nj before they are linked with one another and an edge E is drawn between them. To determine the shortest path we use the A* algorithm. The first edge is simply connected to the nearest node. After the networking rule is executed, the procedure starts again at rule (2). To simplify the process the radius r , within which existing nodes are searched for, can be adapted to an extended Moore neighbourhood (figure 3).

With this simple model we can generate networks with different mesh sizes by varying the parameter X (figure 4). The problem of poor networking of dendritic structures discussed earlier is overcome by using the networking rule (4). A more detailed analysis of the control parameter X and the resulting networks using the measurement of the mean length (m) follows in the analysis section (figure 8).

The development of a road network never proceeds without necessity, but is coupled to the land-use distribution. The mutual dependency of land-use development and road system is discussed in the next section.

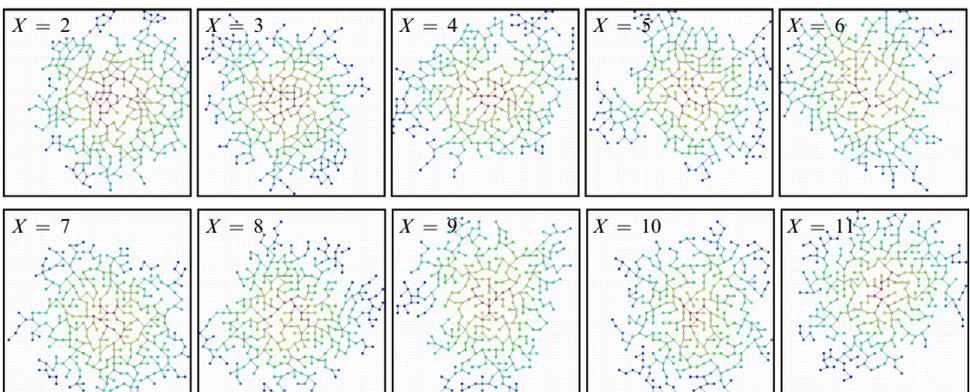


Figure 4. [In colour online.] Road networks with different mesh sizes ($X = 2$ –11).

4 Land-use development

Because of the costs involved, infrastructures such as streets are only built for a reason. As such, road infrastructure develops in conjunction with uses and their interrelationships. In the next step we therefore examine the settlement process in the framework of our model.

The traditional theory of urban economy assumes in its description of land-use development that there is an exogenously predefined road infrastructure along which the settlement samples develop, thereby saving transport costs and the costs of road infrastructure construction. Most classical land-use models adhere to this assumption. Clarke et al (1997) use this principle in their cellular automata model by defining the areas along existing streets, just like areas in the surroundings of agglomerations, as gravitation areas for further settlements.

Even today, despite an institutionalised planning process, it is not always clear what the driving force is behind urban development: an attractive place to settle in or an existing road infrastructure? It is equally unclear whether, at the outset of the urbanisation process, the settlement of a region followed the pattern of available paths or whether the paths resulted from movement between populated areas. The question is therefore whether streets form the driving force for the settlement of the adjacent areas or whether the distribution of settlements in the region determines the formation of the connecting paths. Certainly one can cite examples for both positions. A plausible suggestion is that both scenarios apply with different weights which may shift over the course of time. These shifts can be a consequence of technological developments as well as cultural and social changes.

The following discussion attempts to do justice to both possible development logic alternatives by using two model configurations that are based on the sample principle. According to Humpert's field types the first configuration is denoted by a cluster model in which the land-use development is the primary driving power (figure 6). The 'polynucleated urban landscapes' model (Batty, 2005) is based on a similar concept. The interlink model describes the second configuration, in which streets dominate the pattern of settlement development (figure 7). This second configuration can be compared with the concept of the attractive effect of a street presented by Clarke et al (1997), though in the latter case streets were provided externally. In Batty and Xie's (1997) model the settlement process and the development of the roads are coupled with each other but the development of the roads is subordinate.

To enable an interaction between the system of routes and the settlement process, we introduce another cell grid with a higher resolution, which at the same time determines the scale of the system. In the existing model, 3×3 smaller cells are placed per cell H so that in a cell H there are nine smaller cells $M \{j = 1, 2, \dots, n_M\}$, denoted by $V(H)$. $H(M)$ specifies the primary cell, in which the cell M under construction is located. This partition follows the idea of a multilevel grid (Andersson et al, 2002; Liu and Andersson, 2004). Furthermore, a diffusion process based on the coarse grid defines the potential field and is explained in more detail below. We use a DLA model transferred to the urban context as a basis (Batty, 1991; 2005; Batty and Longley, 1994), which can also be described as a dielectrical breakdown model (DBM) (Schweitzer, 2003). The possible states of the cells M are defined as $S^M = 0$ for empty cells, $S^M = 1$ for roads, and $S^M = 2$ for settled cells. The diffusion field determines the potential P^H of a cell to be built up.

4.1 Cluster model

We start with the development that is driven by the settlement. Subject to a potential field the cells are settled and any populated area is connected to the existing road infrastructure. The model is initialised with:

$$\begin{aligned} N^{Hc}(0) &= 1, \quad N^{Hi}(0) = 0, \quad \forall i \neq c; \\ S^H(0) &= 0, \quad P^H(0) = 0, \quad S^M(0) = 0. \end{aligned} \quad (5)$$

The growth rate is predefined and corresponds to the number of cells M that are settled per step at the place with the highest respective potential:

$$\begin{aligned} \text{if } S^M(t) &= 0 \text{ and } P^M(t) + \varepsilon^M(t) = \max, \\ \text{then } S^M(t+1) &= 2, \end{aligned} \quad (6)$$

where the factor ε^H provides a random selection of several cells with the same potential. If the primary cell $H(M)$ is not yet developed, a new node is put in its centre:

$$\text{if } N^{H(M)}(t) = 0, \text{ then } N^{H(M)}(t+1) = 1. \quad (7)$$

The new node is then linked to the nearest existing node:

$$\text{if } d_{(Hi, Hj)} = \min, \text{ then } E_{(Ni, Nj)}. \quad (8)$$

If several existing nodes are possible for the first link, one is chosen randomly. Afterwards the networking rule (4) is executed.

The road infrastructure is represented on the one hand as a graph and on the other hand is transferred to the cell grid to enable an interaction with the settlement dynamics. The vector lines of the graph are transferred to the cells M by means of Bresnham's line algorithm (Hearn and Baker, 1996), whereby the cells between the respective start and destination nodes are allocated the status of a street:

$$E_{(Ni, Nj)} \rightarrow S^M = 1. \quad (9)$$

In a further step the potential is determined on the basis of the developed cells:

$$\begin{aligned} \text{if } C^{V(H)}(t) &> 0, \\ \text{then } P^H(t+1) &= \left(\frac{C^{V(H)}(t)}{9} \right)^{1-P^H(t)}, \end{aligned} \quad (10)$$

where

$$C^{V(H)}(t) = \sum_G \{1 | G \in V(H), S^G = 2\}. \quad (11)$$

$C^{V(H)}/9$ indicates the fraction of occupied cells M per cell H . To ensure that a cell with $C^{V(H)} = 1$ is not automatically given a low potential value, the density of the neighbour cells is taken into account, by raising the fraction of occupied cells ($C^{V(H)}/9$) by the power of the potential value of cell H . The potential field can now be calculated by averaging the potential values of the von Neumann neighbourhood $U(H)$, including the cell under consideration:

$$P^H(t+1) = \frac{\sum_{G \in U(H)} P^G(t)}{5}. \quad (12)$$

So that the field approaches an equilibrium, equation (12) is executed repeatedly (five times in the following examples). Finally, the potential values P^H of the coarse grid are

assigned to the subordinate cells $V(H)$:

$$P^M(t+1) = P^{H(M)}(t+1). \quad (13)$$

At this point all procedures for a given step t have been executed and the process starts again from instruction (6) with the next step $t+1$.

4.2 Interlink model

The cluster algorithm essentially corresponds to an adapted DBM and can now be modified so that areas can be settled only when a road is already nearby. The extension of the road network is defined by a potential field, which is produced by the populated areas. For the initialisation of the model two nodes are placed at H_{c1} and H_{c2} and are linked with an edge:

$$\begin{aligned} N^{Hc1}(0) &= 1, \quad N^{Hc2}(0) = 1, \quad N^{Hi}(0) = 0, \quad \forall i \neq \{c1, c2\}; \\ S^H &= 0, \quad P^H(0) = 0, \quad S^M(0) = 0. \end{aligned} \quad (14)$$

Subsequently conditions (8) and (9) are executed.

The settlement of empty cells M happens in the same manner as in the cluster model—that is, at the place with the highest potential. However, in this case we consider only cells that have a street cell as a neighbour or have a built-up cell in the neighbourhood. Furthermore, a final condition is introduced whereby the potential value of the considered cell has to exceed a certain threshold value Ψ_{dev} , with which the extent of the settlement in the hinterland without road infrastructure can be controlled. Therefore equation (6) is enlarged as follows:

$$\begin{aligned} \text{if } S^M(t) = 0 \text{ and } [C^{U(M)} \geq 1 \text{ or } P^M(t) > \Psi_{dev} \text{ and } B^{U(M)} \geq 1] \\ \text{and } P^M(t) + \varepsilon^M(t) = \max, \end{aligned} \quad (15)$$

$$\text{then } S^M(t+1) = 2,$$

where

$$C^{U(M)}(t) = \sum_G \{1 | G \in U(M), S^G = 1\}, \quad (16)$$

$$B^{U(M)}(t) = \sum_G \{1 | G \in U(M), S^G = 2\}.$$

The potential values for the developed (or settled) cells are then subsequently updated by means of instructions (10) and (11), and the potential field is calculated and transferred using equations (12) and (13).

The road network is extended when a new node is placed on a cell with the highest potential, but only if the potential exceeds the threshold Ψ_{node} :

$$\text{if } S^H(t) = 3 \text{ and } P^H(t) > \Psi_{node} \text{ and } P^H(t) + \varepsilon^H(t) = \max, \quad (17)$$

$$\text{then } N^H(t+1) = 1.$$

The newly added node is linked with the existing nodes using the networking rule (4). At this point all procedures for a given step t have been executed and the process starts again from instruction (15) with the next step $t+1$.

For both methods the density of the settlements is determined primarily by the size of the noise value $\varepsilon^M(t)$. The larger the noise value, the lower the probability that one or several agglomerations will form. In an extreme case the settlement areas spread out purely by accident. Correspondingly, the density of the settlement can be controlled using a scaling factor $\tau\varepsilon^M(t)$. Figure 5 shows five examples of structures of the cluster

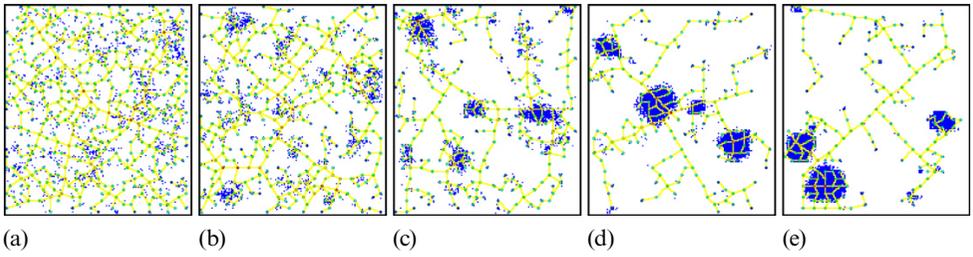


Figure 5. [In colour online.] Cluster structures with different scalings of the noise value τ , where the grids each have $n_H = 50 \times 60$ cells and $n_M = 150 \times 180$ cells. (a) $\tau = 400$, (b) $\tau = 200$, (c) $\tau = 100$, (d) $\tau = 50$, (e) $\tau = 25$.

type for the spectrum of τ . It can clearly be seen that, as the scaling factor τ decreases, the agglomerations get larger and their borders become clearer. In the agglomeration clusters the density is normally highest in the centres and decreases only towards the edges.

4.3 Centripetal and centrifugal forces

For more precise control of the spreading behaviour of the settlement than is possible using the scaling of a noise factor, we draw on a concept by Krugman. Krugman (1996) introduces two contrary forces, where the first, F_{petal} , has a centripetal or centralising effect which represents the needs of the population and the economic advantages (economies of scale) of settling close to one other, and the second, F_{fugal} , has a centrifugal, decentralising effect which represents people's desire to occupy as much vacant space as possible and to avoid competition with one other. This principle can be transferred to the model represented above. The two forces are described as follows:

$$\begin{aligned} F_{\text{petal}}^M(t) &= A\{1 - \exp[\omega_p P^M(t)]\} + \varepsilon^M(t), \\ F_{\text{fugal}}^M(t) &= B\{1 - \exp[\omega_f P^M(t)]\} + \varepsilon^M(t). \end{aligned} \quad (18)$$

The factors A and B are scaling parameters which define the strength of the centripetal or the centrifugal force and ω_p and ω_f affect the impact of the potential of a cell. The resulting force F_{res} can be regarded as a trade-off between the two contrary forces:

$$F_{\text{res}}^M(t) = F_{\text{petal}}^M(t) - F_{\text{fugal}}^M(t). \quad (19)$$

Equation (6) for the decision in which a settlement takes place can now be modified as follows:

$$\begin{aligned} \text{if } S^M(t) = 0 \text{ and } F_{\text{res}}^M(t) + \varepsilon^M(t) = \max, \\ \text{then } S^M(t+1) = 2. \end{aligned} \quad (20)$$

For the sake of simplicity, the scaling factor τ does not appear in this equation, since in the following we set $\tau = 1$ and control the impact of the noise factor $\varepsilon^M(t)$ by means of the scaling factors A and B .

It seems reasonable to suppose that agglomerations will take place only if A is sufficiently large relative to B , but that multiple subcentres are possible only if B is sufficiently large with respect to A . In accordance with Krugman (1996), the interesting range for the parameters A , B , ω_p , and ω_f is where:

$$\frac{\omega_p}{\omega_f} > \frac{A}{B} > \frac{\omega_f}{\omega_p}. \quad (21)$$

Transferring this to the present model, the relationship of the four parameters can be described verbally as follows: as soon as the last two quotients A/B and ω_f/ω_p

converge towards one another, the generated structure becomes more and more homogeneous and fewer clusters form. If, on the other hand, both of the first quotients ω_p/ω_f and A/B converge, larger and larger agglomerations result and the number of clusters decreases until all areas finally combine to form one large cluster. This connection is based on the fact that the resulting force F_{res} causes a preferential treatment of certain areas of the potential gradient for settlement (figure 6 and figure 7). If the maximum value of F_{res} lies in the lower interval of the possible potential values (which lie between 0 and 1), the density of the generated agglomerations decreases, though the clusters are preserved providing there is a corresponding strength of centripetal force. It is this property that represents the major advantage of using centripetal and centrifugal forces over the more coarse control of using a scaled noise factor.

5 Examples

“A picture may be worth a thousand words, but in this case a pattern is worth a million numbers.”

Couclelis (1997)

In this chapter some selective generated structures are presented with the help of qualitative illustrations. The next paragraph describes the associated quantitative statistical evaluations. For the examples described here some basic settings were fixed. The growth rate $\lambda_{\text{dev}} = 5$ means that, per time step, five cells are developed or settled by repeating, five times, instruction (6) for the cluster type and instruction (15) for the interlink type. The scaling factor for the noise value is $\tau = 1$ and the minimum distance of the nodes is $K = 3$. In particular, for the interlink type we define the parameters for the threshold value at the development of a cell $\Psi_{\text{dev}} = 0.1$ and for the threshold value at the occupation of a cell with a node $\Psi_{\text{node}} = 0.1$. The parameters for the centripetal force and the centrifugal force are specified separately in each of the individual examples. For all examples the initial configuration is as indicated in equation (5)—that is, a simulation always starts with an empty cell grid with merely one node and one street cell as the nucleus in the middle of the grid. The simulations of the cluster type and interlink type are based on a resolution of $n_H = 50 \times 60$ cells as well as $n_M = 150 \times 180$ cells and were executed until step $t = 500$.

5.1 Cluster-type examples

For the cluster type the settlement areas are generated on the basis of a potential field and the newly developed areas are subsequently attached to the existing road network. Previously we stated that the density of the agglomerations depends on the area of the potential gradient that is preferred for settlement. This preference is indicated by the maximum value of F_{res} , which is given by the ratio of the control parameters of the centrifugal and centripetal forces. In this section we examine some examples from the spectrum of structures that can be generated by varying the value of F_{res} . The examples primarily exemplify the control of the number as well as the density of the agglomeration clusters forming. Figure 6(a) shows the structures resulting when a centrifugal force dominates, as shown on the graphs below the structures, where the high point of the potential gradient is at the lower potential values on the x -axis. This means that lower densities are preferred for settlement. The weight of the potential field is $\omega_p = 10.0$ for the centripetal force and $\omega_f = 8.0$ for the centrifugal force. For all five variants the maximum is in the range of $0 < F_{\text{max}} < 0.2$. Despite the rise of the strength of the centripetal force achieved by increasing parameter A from 1.8 to 2.2, the density of the agglomerations remains relatively low. An increase in the maximum value of F_{res} from approximately 0.05 to approximately 0.31 results in the formation of

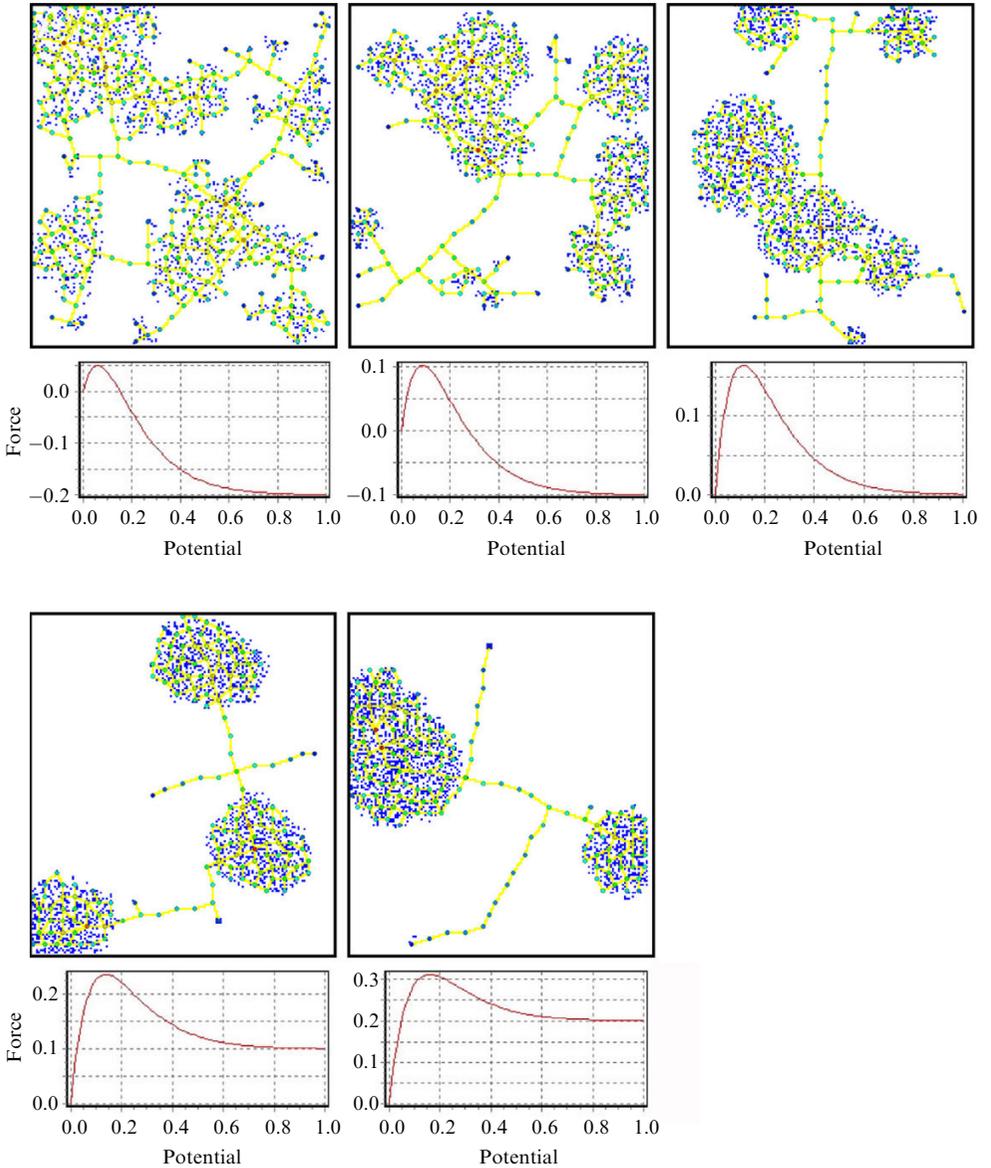
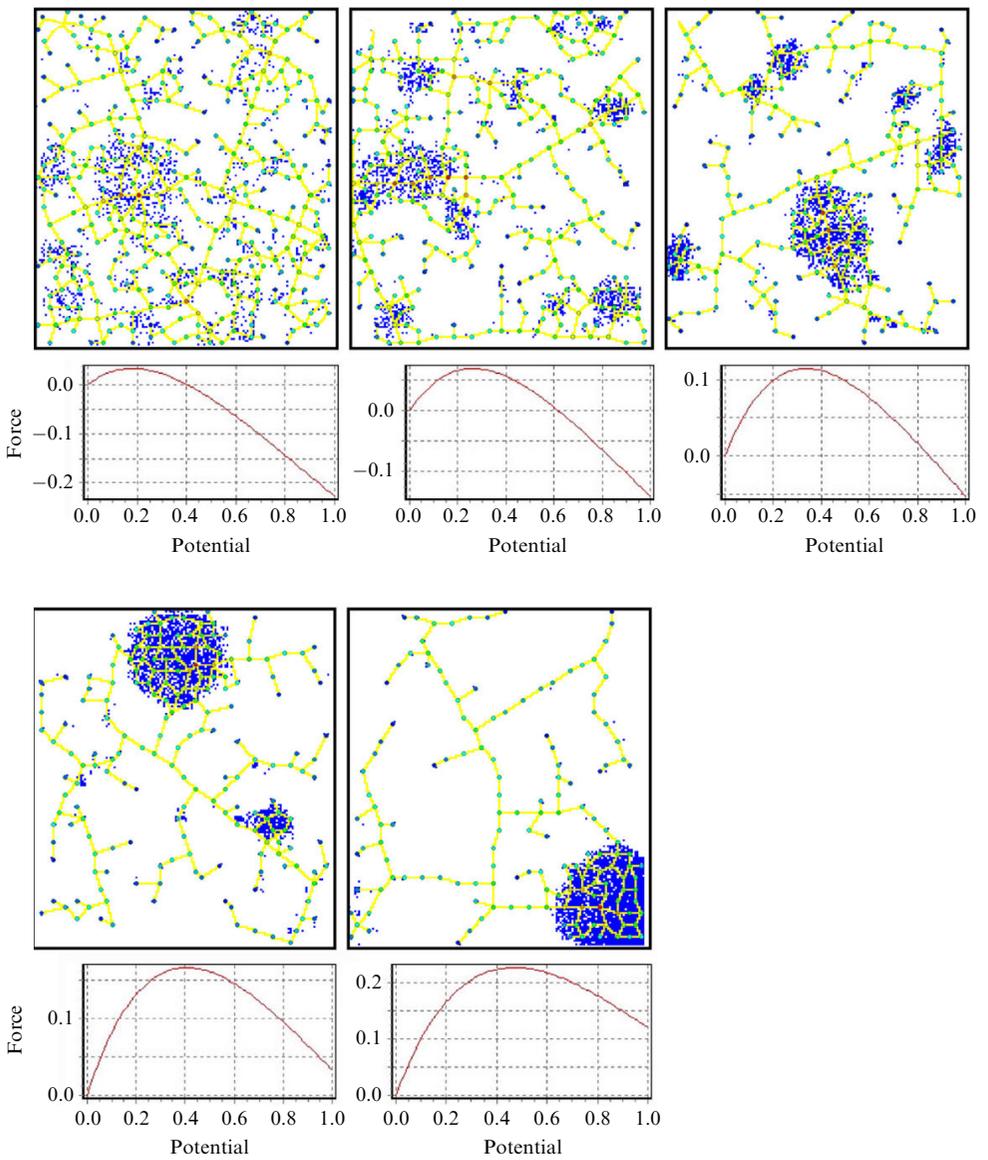


Figure 6. [In colour online.] Cluster types with settings for agglomerations of different densities at $t = 500$. (a) Top left to bottom right: $A = 1.8, 1.9, 2.0, 2.1, 2.2$; $B = 20$; $\omega_p = 10.0$; and $\omega_l = 8.0$. (b) Top left to bottom right: $A = 1.2, 1.3, 1.4, 1.5, 1.6$; $B = 2.0$; $\omega_p = 2.0$; $\omega_f = 1.0$. (c) $A = 2.0$; $B = 2.0$; $\omega_f = 1.0$; top left to bottom right: $\omega_p = 1.2, 1.3, 1.4, 1.5, 1.6$. The graphs below the pictures show F_{res} against potential, and thus represent the preferred potential areas.

fewer and fewer clusters. The proportions of the quotients ω_p/ω_f , A/B , and ω_f/ω_p are summarised for all rows of figure 6 in table 1.

In row B of figure 6 the weight of the potential field was reduced for the centripetal force to $\omega_p = 2.0$ and for the centrifugal force to $\omega_f = 1.0$. In this case the maximum value and the strength of F_{res} are in the range $0.2 < F_{\text{max}} < 0.5$, and A was increased incrementally from 1.2 to 1.6, with $B = 2.0$. In figure 6 one can see clearly from left to right how the number of clusters decreases and the density of the agglomerations increases.

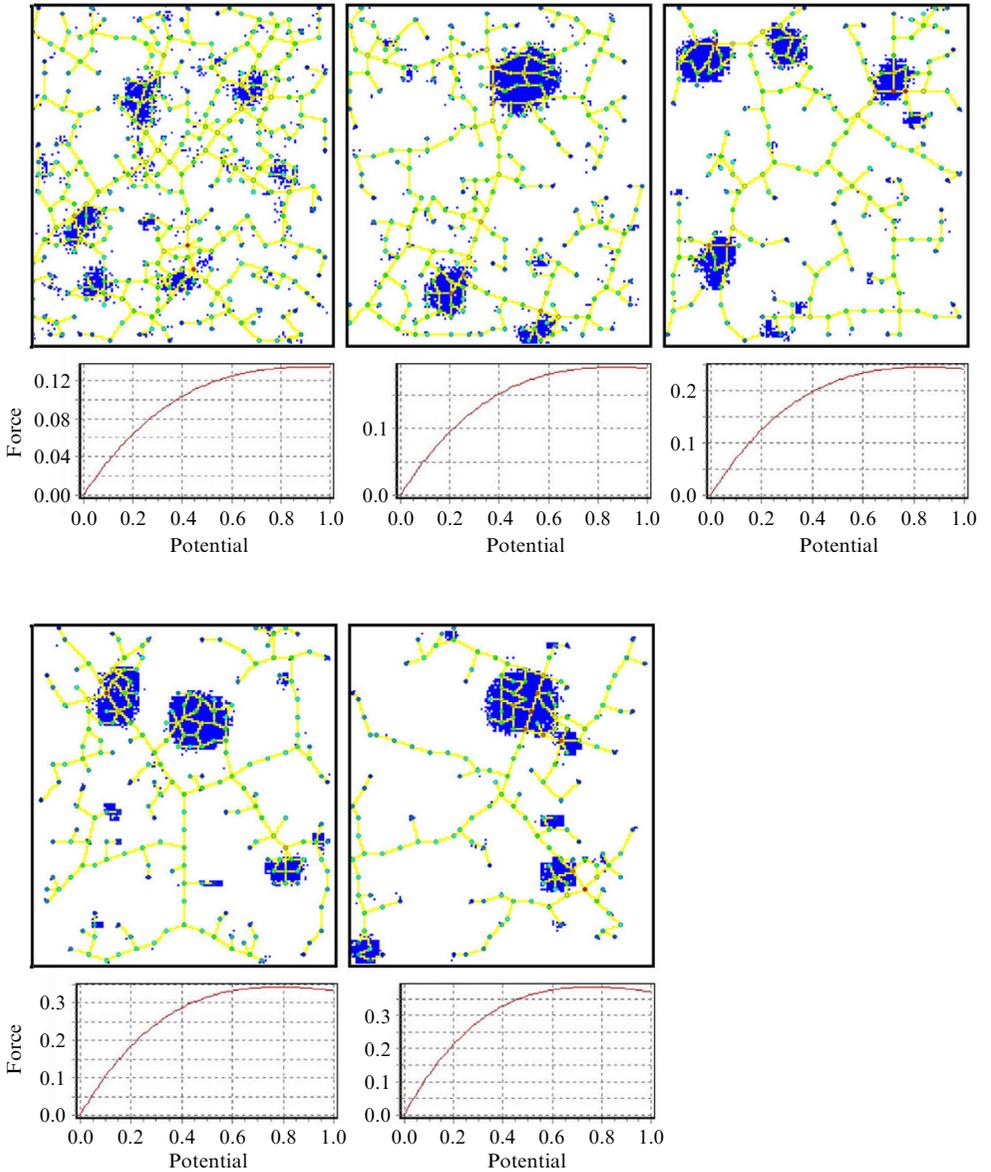


(b)

Figure 6 (continued).

Table 1. Proportions of the quotients for figure 6 and figure 7.

	ω_p / ω_r	A/B	ω_r / ω_p
Row A	1.25	0.90–1.10	0.80
Row B	2.00	0.60–0.80	0.50
Row C	1.20–1.60	1.00	0.83–0.62



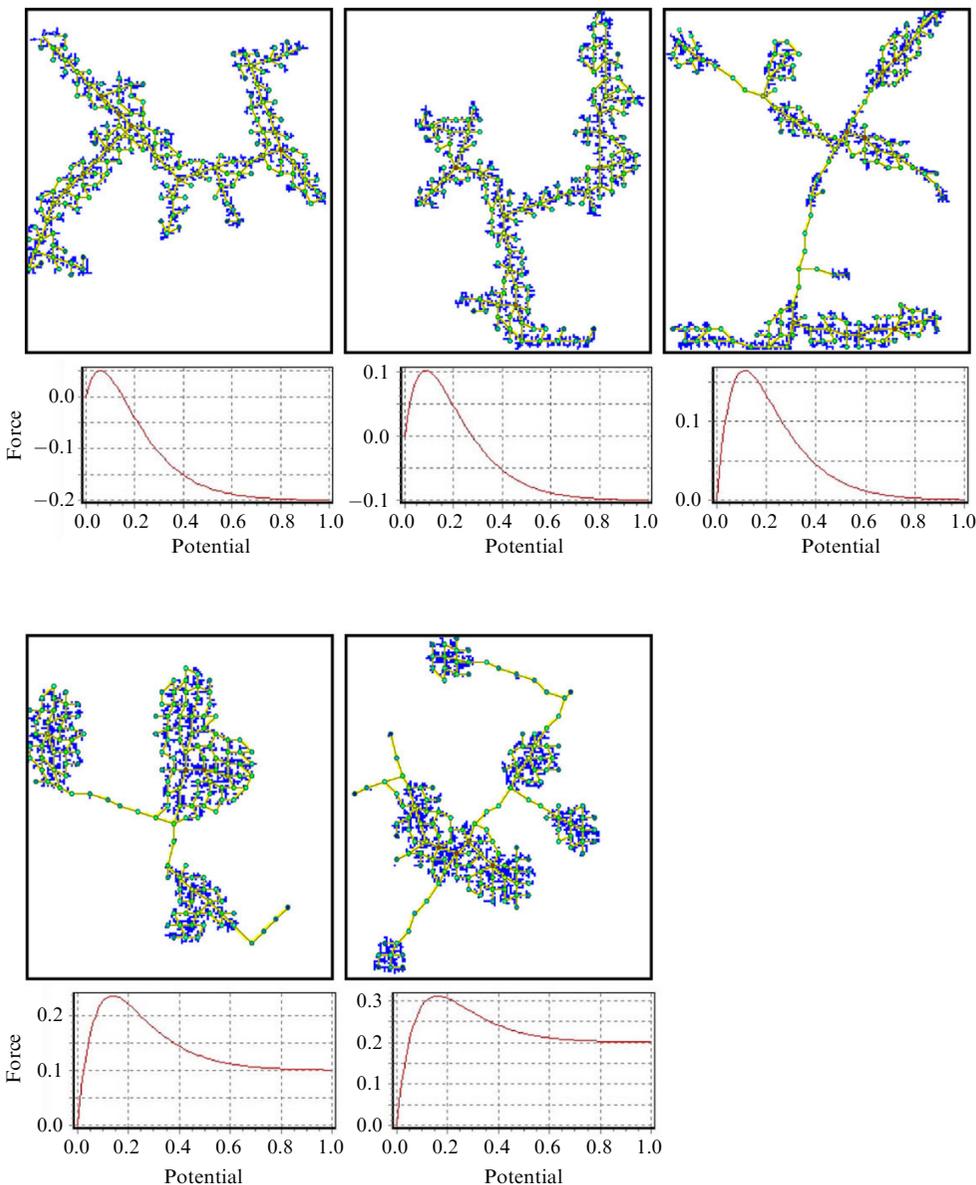
(c)

Figure 6 (continued).

Finally, in figure 6, row C the structures are dominated by the centripetal force. The strengths of the two forces were set to be equal using the parameters $A = B = 2.0$. The weight of the potential field was gradually increased from 1.2 to 1.6 for the centripetal force ω_p and was set to $\omega_f = 1.0$ for the centrifugal force. The graphs below the structures show that the maximum value of F_{res} increases from approximately 0.14 to approximately 0.40, resulting for the most part in a decrease in the number of clusters forming and an increase in their size.

For all variants the number of resulting clusters depends primarily on the level of the maximum value of F_{res} , and therefore on the strength of the resulting force.

On the other hand, the density of the agglomerations depends on the position of the maximum value of F_{res} , and consequently on the preferred region of the potential gradient for settlement. With the help of the four parameters A , B , ω_p , and ω_f for the regulation of the two forces, the position and the level of the maximum value of F_{res} can be determined. Using this apparatus we can control all major characteristics for the generation of settlement structures using the cluster type.



(a)
Figure 7. [In colour online.] Interlink types with settings for agglomerations of different densities at $t = 500$. (a) Top left to bottom: $A = 1.8, 1.9, 2.0, 2.1, 2.2$; $B = 2.0$; $\omega_p = 10.0$; and $\omega_f = 8.0$. (b) Top left to bottom right: $A = 1.2, 1.3, 1.4, 1.5, 1.6$; $B = 2.0$; $\omega_p = 2.0$; $\omega_f = 1.0$. (c) $A = 2.0$; $B = 2.0$; $\omega_f = 1.0$; top left to bottom right: $\omega_p = 1.2, 1.3, 1.4, 1.5, 1.6$. The graphs below the pictures show F_{res} against potential, and thus represent the preferred potential areas.

5.2 Interlink-type examples

For the interlink type we used the same parameter settings as for the cluster type. This enables us to compare the two principles with each other. Since we used the same settings as above, we will not repeat them here but instead discuss the essential differences of the results.

Before the procedures of the interlink type begin, three repetitions of the cluster type are executed to initiate a basic road structure. This basic structure provides the framework for the aggregation of further settlements areas (figure 7). In reference to the urban development theory mentioned above, this basis framework could also derive from a predefined structure of paths. As the same parameter settings have been used as

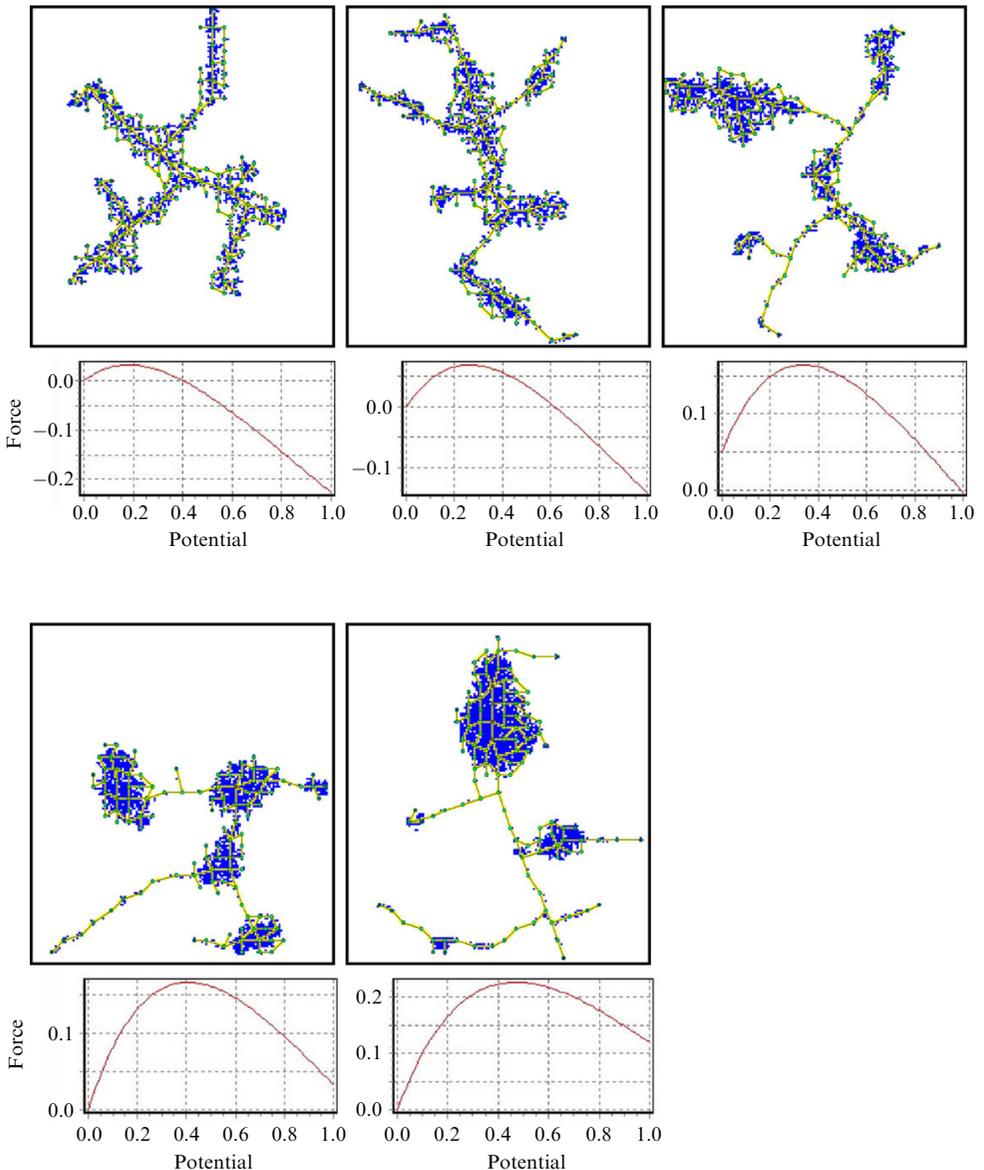
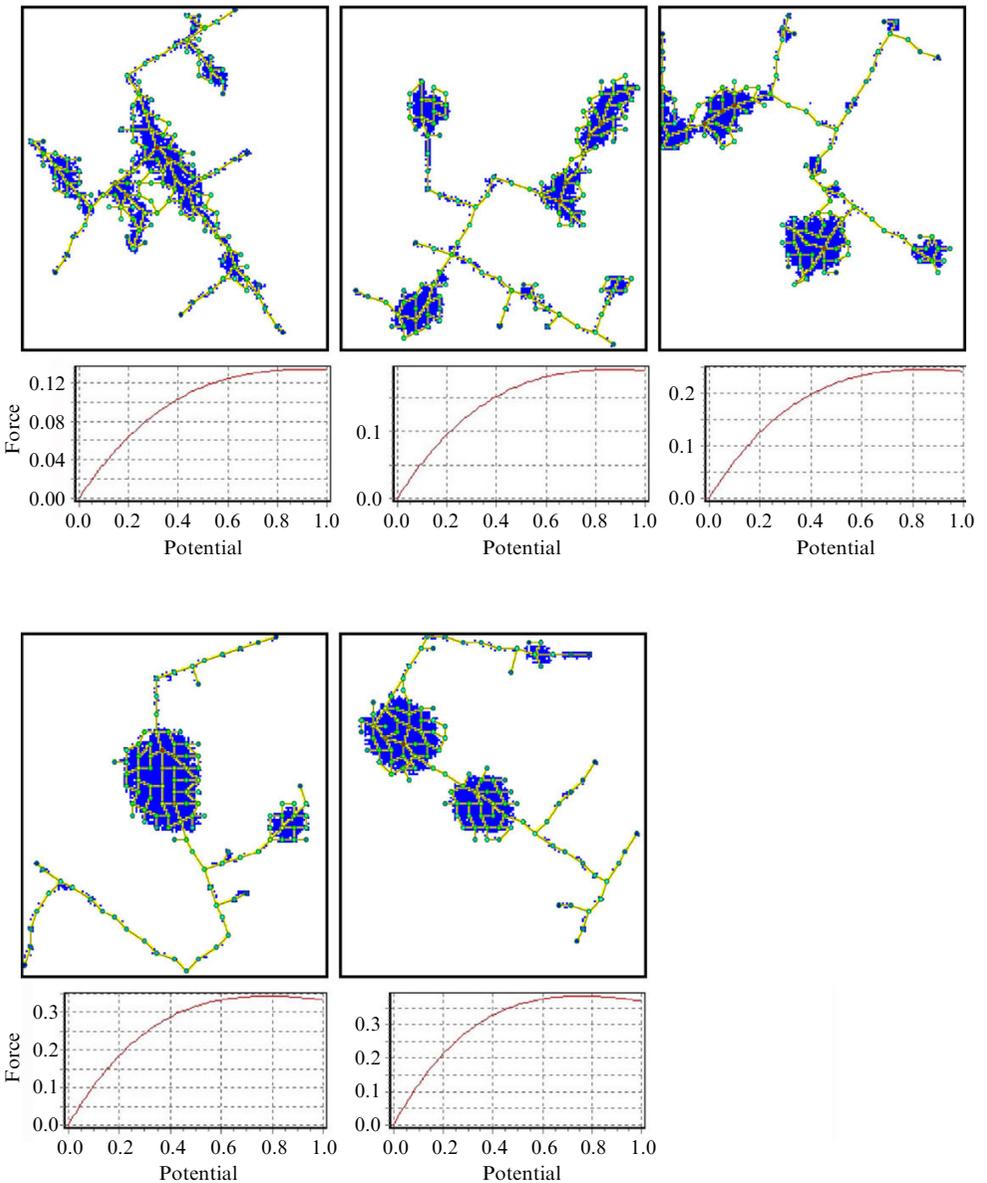


Figure 7 (continued).

in the cluster-type examples, the graphs also show the same series for the resulting force F_{res} of the interlink type in figure 7, and the proportions of the quotients ω_p/ω_f , A/B , and ω_f/ω_p correspond to those given in table 1.

The obvious differences between the agglomerations in figure 7 and those in figure 6 result from the fact that, in figure 7, settlements can arise only along streets or in places next to areas already populated. Nevertheless, we can also see here that the density of the clusters depends on the position of the maximum value of F_{res} and that the number of clusters is dependent on the strength of F_{res} . In comparison with the cluster type, the essential difference is that, with the interlink type, the formation of



(c)

Figure 7 (continued).

agglomerations is only possible along predefined roads. As a result of this restriction of settlement possibilities, the agglomerations with a low density in the left column of figure 7 are more densely populated than the comparable structures in the left column of figure 6. In comparison, no notable differences can be recognised at higher densities. We shall look at these distinctions in more detail in the analysis section with the help of quantitative measuring. To summarise, one can see that, by means of the four parameters A , B , ω_p , and ω_f , all major characteristics for the generation of settlement structures can be controlled for the interlink type as well as for the cluster type.

The field types highwayman and boom, not yet discussed, are inherent to the interlink type and can be generated by means of the three parameters K , Ψ_{dev} , and Ψ_{node} . By increasing Ψ_{node} to greater than or equal to 1, no more streets are added to the structure and the hinterland is not developed any further. If we also increase Ψ_{dev} to greater than or equal to 1, no cells are settled that do not directly border on a street—that is, this corresponds to the highwayman type. The boom type results from using relatively low values for Ψ_{node} and a high value for K , which hinders networking of the road branches that lead into the hinterland.

6 Analysis

A visual evaluation is essential in order to understand the morphological characteristics and to facilitate a quick overview of the different structures generated and their main characteristics. It is, however, informative to examine the consequences of the parameter variations in more detail using a variety of qualitative measuring techniques: the fractal dimension; the rank–size analysis; the segregation index, which differentiates between developed and free cells; and the R^2 value for determining correlation between potential and density and between potential and the connectivity of the road pattern. Using these characteristic values it is possible to compare the structures and to identify special features of particular structures or of variations of the parameters. In the following discussion the different methods of analysis are explained briefly and then applied to the examples shown earlier.

6.1 Methods

For the determination of the fractal dimension D_b , the so-called box-counting or grid-counting method was used (Batty and Longley, 1994; Bovill, 1996; Frankhauser, 2002). For this a grid is placed over a picture or similar and the mesh size s of the grid is scaled gradually. With each iteration the number of grid cells N_s is ascertained in which, in our case, at least one populated cell can be found. Afterwards the total number of occupied fields is compared with the number of empty fields. For each scaling step the relationship

$$D_b = \frac{\log_{10} N_s}{\log_{10} \frac{1}{s}}, \quad (22)$$

is drawn in a log–log diagram. The slope of the regression line gives an estimate of the box-counting dimension X . For $\frac{1}{s}$ the number of meshes in the lower row of the counting grid is used.

For the rank–size analysis the sizes of the different clusters are measured and sorted with regard to their size S . After a rank r has been assigned to every cluster, where the largest cluster has a value of $r = 1$, the rank–size rule means that the cluster sizes have an inverse relationship to their rank:

$$S_r = S_1/r, \quad \text{or} \quad S_r = S_1 r^\alpha. \quad (23)$$

The distribution of the clusters follows the power law, which is known also as a Pareto distribution.

A third characteristic value for the generated structures is the index of dissimilarity or ID (Duncan and Duncan, 1955). The value of the ID can be regarded as the fraction of cells that have to change their position or their state to produce a homogeneous uniform distribution.

The two final characteristic values specify the coefficient of determination r^2 , taken from the correlation r , firstly of the potential values P^H of the local density and secondly of the potential values P^H of the relative connectivity or mean length (m) of the nodes N^H of the road map.

The connectivity (Haggett, 1991) of a node results from the sum of the minimal distances to all other nodes. These distances are measured in terms of the nodes one has to cross on the shortest path, which is calculated by means of the Dijkstra algorithm. If we divide the value of the connectivity by the number of the remaining nodes, we obtain the mean length m of the node under consideration. Using this characteristic value it is possible to compare the different networks from figure 3 with each other [figure 8(a)]. The behaviour of m when comparing networks of different sizes is shown on the graph in figure 8(b). As expected m increases with X for large networks ($n_N = 300$ nodes), since the degree of networking decreases and larger detours become necessary. There is an increase in m in cases in which networks are expanding, as longer paths are necessary to reach all nodes. The relations between X , n_N , and m are roughly linear (figure 8).

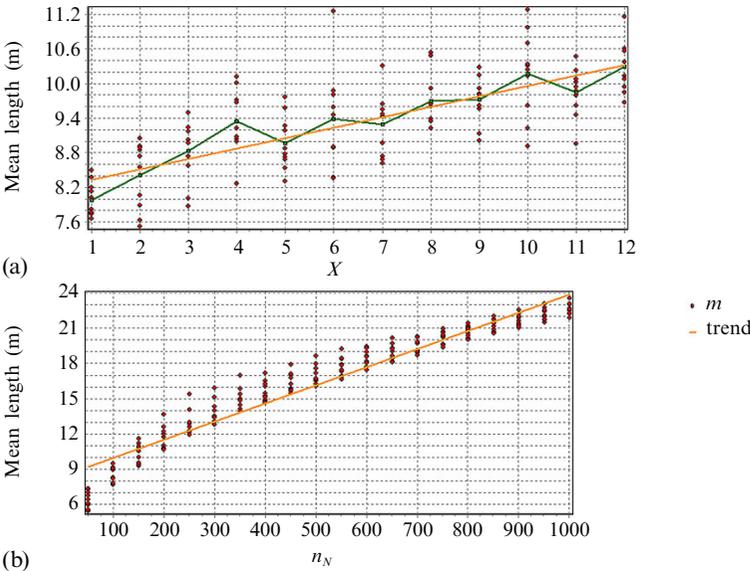


Figure 8. [In colour online.] (a) Mean length m against X ($r^2 = 0.81$) for networks with $n_N = 300$; (b) mean length m against n_N ($r^2 = 0.93$).

6.2 Analysis of land-use development

Using the measuring methods described above, the corresponding characteristic values can now be determined for each constellation of parameters. To compensate for coincidental outlying values, five measurements are taken using the same parameters. The graphs in figure 9 connect the mean average values of the five measurements, respectively. The measurement of a generated structure is carried out every 300 time steps.

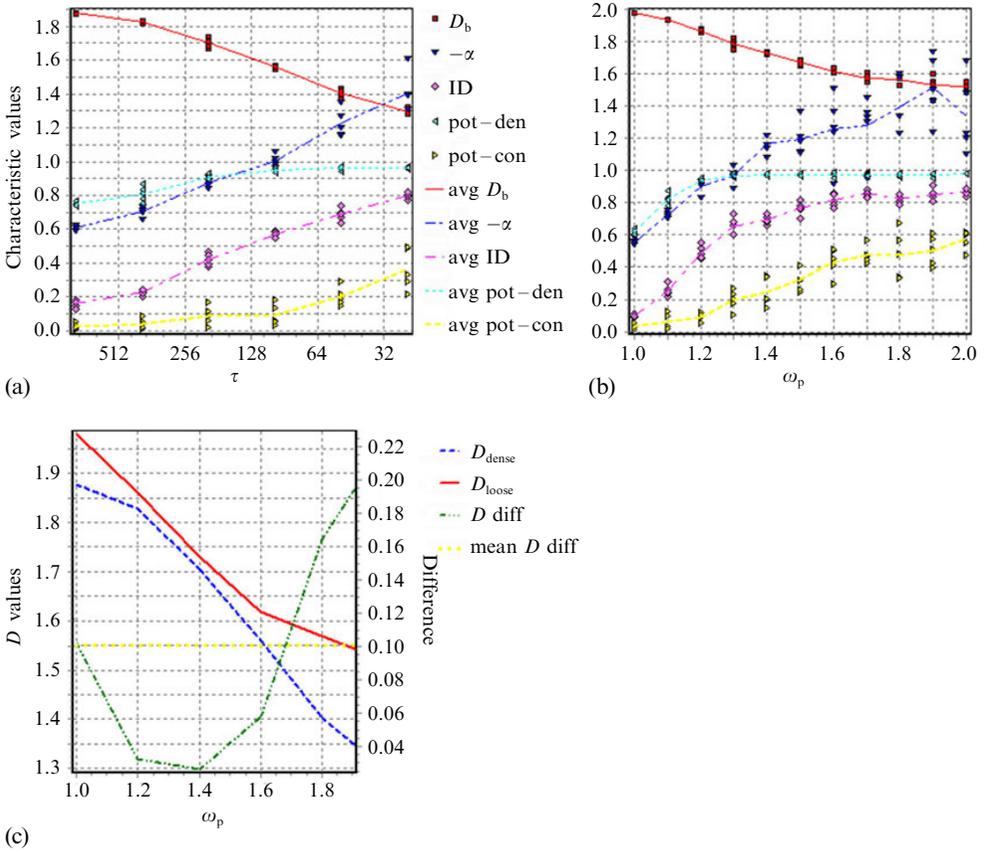


Figure 9. [In colour online.] Diagrams showing the characteristics of the different structures. (a) Characteristics of the structures in figure 5, as a function of noise factor. (b) Characteristics of the structures in figure 6(b), as a function of ω_p . (c) D_b values from (a) and (b), the difference between the two, and the mean of that difference. (d) Characteristics of the loosely structured cluster-type structures in figure 6(a), as a function of ω_p . (e) Characteristics of the densely structured cluster-type structures in figure 6(a), as a function of ω_p . (f) D_b values from (d) and (e), the difference between the two, and the mean of that difference. (g) Characteristics of the loosely structured interlink-type structures in figure 7(a), as a function of A . (h) Characteristics of the densely structured interlink-type structures in figure 7(c), as a function of ω_p . (i) D_b values from (g) and (h), the difference between the two, and the mean of that difference. D_b = fractal dimension calculated using the box-counting method, α = Pareto coefficient, ID = index of dissimilarity, pot/den = r^2 of correlation between potential and density, pot-con = r^2 of correlation between potential and connectivity.

In figure 9 the upper row shows how the quantitative characteristic values of the version controlled by the scaling of a noise value (figure 5) differ from those of the version controlled by the weighted influence of the centripetal and centrifugal forces (figure 6). The progression of the respective series of measurements looks similar for the two diagrams in figures 9(a) and 9(b). The differences are in the details of the structures—for example, in the density of the forming agglomerations, which can be measured primarily by the characteristic values of the fractal dimension D_b and the ID. Since the measurement of the fractal dimension has been established as a common value for the quantification and comparison of settlement structures (Batty and Longley, 1994; Bovill, 1996; Frankhauser, 2002), we concentrate on this value. In order that the different series of measurements are comparable, the x values of figure 9(a) have

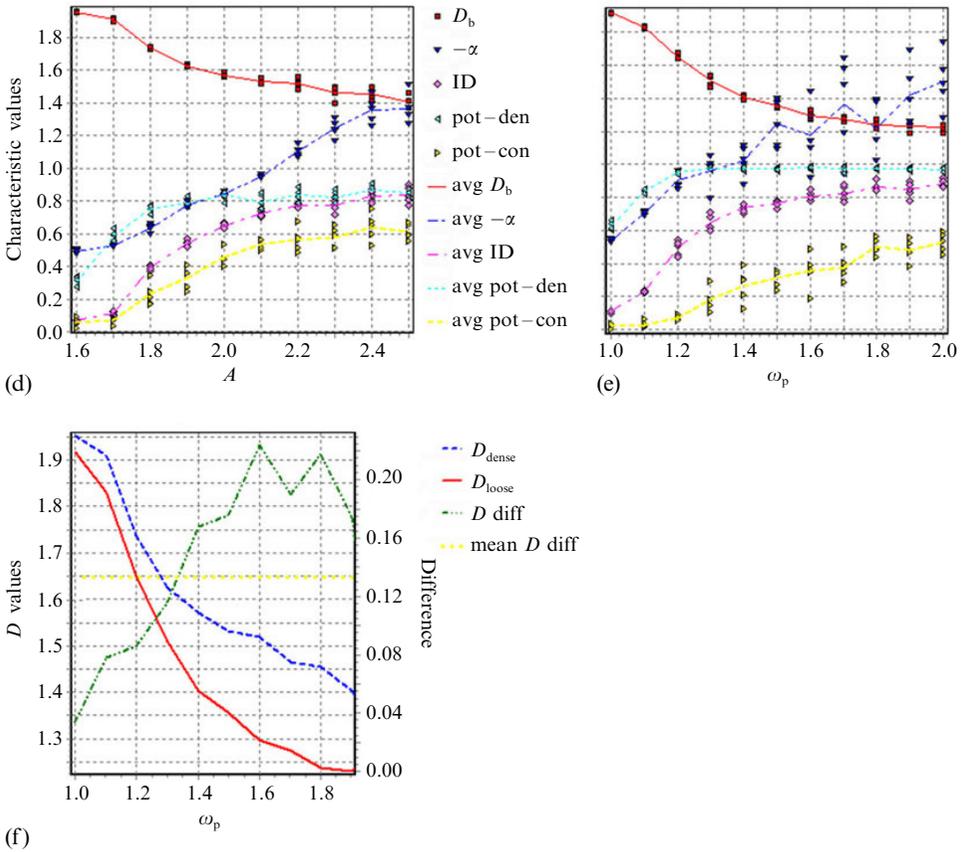


Figure 9 (continued).

been scaled accordingly. In figure 9(c) the values of D_b from figures 9(a) and 9(b) are compared and the differences of each of the measured values is plotted, yielding an average of $\Delta D_b = 0.10$. In figures 9(d)–(f) the corresponding characteristic values are given for cluster-type structures with low [figure 9(d)] and high [figure 9(e)] densities. The average difference of the fractal dimension is $\Delta D_b = 0.13$ [figure 9(f)]. Finally, in figures 9(g)–(i) the measured values are represented for the interlink structures at low and high densities. Particularly noticeable are the clear deviations of the Pareto exponent α . The average difference of the fractal dimension is again indicated in the right column [figure 9(i): $\Delta D_b = 0.10$].

7 Discussion and conclusion

Using the model presented in this paper, road networks can be generated with different qualities, controlled primarily by a networking parameter. By introducing a cell grid with a higher resolution it becomes possible to define the system scale with a relationship between road meshes and cell sizes. The disadvantage of the method shown for generating a road network based on a grid is that it constitutes a geometric restriction when compared with Duarte et al’s (2007) ‘urban grammar’ method. However, using our model we can generate topologically extensive structures with a comparatively simple principle.

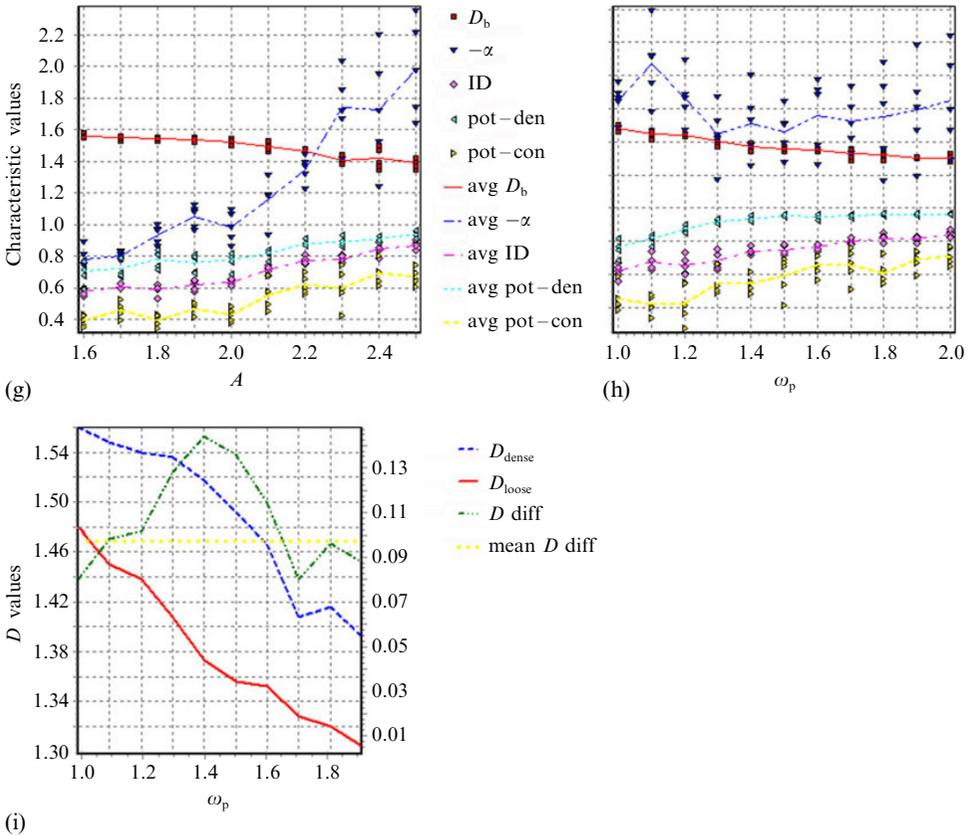


Figure 9 (continued).

Using four parameters for the two opposite effective forces F_{petal} and F_{fugal} , it is possible to precisely steer the generation of settlement structures with regard to their global and local density as well as the size and number of forming clusters. If we compare the results from figure 6 and figure 7 with Batty's (2005, pages 156–182) structurally similar model, one of the striking qualities of the model is the precise controllability of the settlement density using the resulting force F_{res} .

As mentioned earlier, the introduction of centripetal and centrifugal forces as control parameters allows one to achieve a compromise that on the one hand takes advantage of central places, and on the other follows the need for private free space as spaciouly as possible, which can only be achieved outside the centres. Generally speaking, it is evident that the value of a location depends on its closeness or distance to other locations. On the basis of adaptation from Franck (2002), the closeness or distance can be rated in two different ways: firstly with regard to the travel expenses, which are connected to the choice of the location, and secondly with regard to the quality of the view, the air, the landscape, and the absence of sources of disturbances and impairments. If we assume that travelling expenses, the scarcity of areas available for building in the centres, and the rates of spatial discounting⁽¹⁾ equalise, the question arises as to what governs the location choice priorities for particular sections of society, commercial facilities, or production facilities, and how and why these preferences change.

⁽¹⁾ Discounting means the assessment of spatial and temporal distance according to the distance to the here and now. This distance is independent of the assessment of space as utilisable volume and from the assessment of time as utilisable hour (Franck, 1992, page 3).

The average preference for a location between the centre and open landscape corresponds in our model to the resulting force F_{res} , whose graphs are shown below the illustrations of the generated structures in figures 6 and 7.

In more recent decades the development of prospering urban regions was indicated primarily by a strong tendency towards suburbanisation and secondly by an increasing demand for central locations and a corresponding increase in rents and land prices in town centres. These contrary trends can be explained on the one hand by increasing income and with it an increasing demand for residential accommodations in areas of low building density. A further prerequisite for suburban development is the availability of a corresponding transport technology and traffic infrastructure which facilitates the bridging of distances between places of residence and work in a relatively short time.

A primary objective of the methodology introduced here is, after Epstein (2006), the development of a generative urbanism which enables one to examine different urban development theories as well as possible consequences of planning schemes in silicio in an urban laboratory.

References

- Andersson C, Lindgren K, Rasmussen S, White R, 2002, "Urban growth simulation from 'first principles'" *Physical Review E* **66** 1–9
- Batty M, 1991, "Generating urban forms from diffusive growth" *Environment and Planning A* **23** 511–544
- Batty M, 2005 *Cities and Complexity. Understanding Cities with Cellular Automata, Agent-based models, and Fractals* (MIT Press, London)
- Batty M, Longley P, 1994 *Fractal Cities: A Geometry of Form and Function* (Academic Press, London)
- Batty M, Xie Y, 1997, "Possible urban automata" *Environment and Planning B: Planning and Design* **24** 175–192
- Batty M, Xie Y, Sun Z, 1999, "Modeling urban dynamics through GIS-based cellular automata" *Computers, Environment and Urban Systems* **23** 205–233
- Benguigui L, 1995, "A new aggregation model: application to town growth" *Physica A* **219** 13–26
- Bovill C, 1996 *Fractal Geometry in Architecture and Design* (Birkhäuser, Boston, MA)
- Clarke K C, Hoppen S, Gaydos L, 1997, "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area" *Environment and Planning B: Planning and Design* **24** 247–261
- Couclelis H, 1997, "From cellular automata to urban models: new principles for model development and implementation" *Environment and Planning B: Planning and Design* **24** 165–174
- Dean J S, Gumerman G J, Epstein J M, Axtell R L, Swedlund A C, Parker M T, McCarroll S, 2000, "Understanding Anasazi culture change through agent-based modeling", in *Dynamics in Human and Primate Societies: Agent-based Modeling of Social and Spatial Processes* Eds T Kohler, G Gumerman (Oxford University Press, Oxford) pp 179–205
- Duarte J P, Rocha J D M, Soares G D, 2007, "Unveiling the structure of the Marrakech Medina: a shape grammar and an interpreter for generating urban form" *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **21** 317–349
- Duncan O D, Duncan B, 1955, "A methodological analysis of segregation indexes" *American Sociological Review* **20** 210–217
- Eden M, 1960, "A two-dimensional growth process", in *Proceedings of Fourth Berkeley Symposium on Mathematics, Statistics, and Probability* (University of California Press, Berkeley, CA) pp 223–239
- Epstein J M, 2006 *Generative Social Science: Studies in Agent-based Computational Modeling* (Princeton University Press, Princeton, NJ)
- Epstein J M, Axtell R, 1996 *Growing Artificial Societies: Social Science from the Bottom Up* (MIT Press, Cambridge, MA)
- Erickson B, Lloyd-Jones T, 1997, "Experiments with settlement aggregation models" *Environment and Planning B: Planning and Design* **24** 903–928
- Franck G, 1992 *Raumökonomie, Stadtentwicklung und Umweltpolitik* (Kohlhammer, Stuttgart)
- Franck G, 2002, "Soziale Raumzeit", in *Raumzeitpolitik* Eds D Henckel, M Eberling (Leske and Budrich, Opladen) pp 61–80

-
- Frankhauser P, 2002, "Fractal behaviour of urban patterns on different scales", in *Fundamental Principles of Urban Growth* Eds K Humpert, K Brenner, S Becker (Müller and Busmann, Wuppertal) pp 158–175
- Haggett P, 1991 *Geographie: Eine Moderne Synthese* (Ulmer, Stuttgart)
- Hearn D, Baker P, 1996 *Computer Graphics, C Version* (Prentice-Hall, Upper Saddle River, NJ)
- Humpert K (Ed.), 1992 *Das Phänomen der Stadt: Berichte aus Lehre und Forschung Städtebauliches Institut*, University of Stuttgart, Stuttgart
- Humpert K, 1997 *Einführung in den Städtebau* (Kohlhammer, Stuttgart)
- Krugmann P, 1996 *The Self-organizing Economy* (Blackwell, Oxford)
- Liu X H, Andersson C, 2004, "Assessing the impact of temporal dynamics on land-use change modeling" *Computers, Environment and Urban Systems* **28** 107–124
- Schweitzer F, 2003 *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences* (Springer, Berlin)
- Semboloni F, 2000, "The growth of an urban cluster into a dynamic self-modifying spatial pattern" *Environment and Planning B: Planning and Design* **27** 549–564

Conditions of use. This article may be downloaded from the E&P website for personal research by members of subscribing organisations. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.