

## Cellular automata models for vegetation dynamics

Heiko Balzter \*, Paul W. Braun, Wolfgang Köhler

*Biometry and Population Genetics, Justus-Liebig-University, Giessen, Germany*

Accepted 19 November 1997

---

### Abstract

A short review of cellular automata models in ecology is given. Introduction of a spatial dimension into a discrete-time Markov chain leads to a class of cellular automata called spatio-temporal Markov chains (STMC). The behaviour of the model is determined by its temporal and spatial orders. It has been proven that STMC models have ergodic distributions in certain cases for spatial order 0. Simulations suggest that STMC models of higher spatial order also have ergodic distributions. The model of Dytham (1995) is demonstrated to be an STMC. Modelling population dynamics of three plant species on a lawn, two STMC models of different order are compared. The model validation shows a good agreement for *Glechoma hederacea*, but large deviations for *Lolium perenne* and *Trifolium repens*. The species-dependent performance of the models can be explained by selective grazing. Modifications of the transition matrices are used to examine possible causes of the deviations. © 1998 Elsevier Science B.V. All rights reserved.

**Keywords:** Markov chains; Cellular automata; Spatio-temporal modelling; Spatially explicit models; Vegetation dynamics

---

### 1. Introduction

In ecology there is an increasing interest in spatially explicit models. On a large scale they are used to model landscape dynamics and changes in land use or vegetation cover, while on a small

scale single or several populations of animal or plant species are of interest. Spatially explicit models are expected to increase our ability to accurately model populations subject to complex processes, to create basic ecological knowledge and to improve the application of landscape ecology to conservation and land management problems (Dunning et al., 1995; Kareiva and Wennergren, 1995; Turner et al., 1995).

Vegetation cover plays an important role in the human environment, being the nutritional basis

---

\* Corresponding author. Present address: Biometrie und Populationsgenetik, Universität Giessen, Ludwigstrasse 27, D-35390 Giessen, Germany. Tel.: +49 641 9937540/9937548; fax: +49 641 9937549; e-mail: Heiko.Balzter@agrar.uni-giessen.de

for higher organisms such as animals and man. Because species in plant communities generally show aggregated distributions at one scale or more and competition between sessile organisms occurs mainly between neighbours, attention has to be paid to the spatial arrangements of plants in the field (Silvertown et al., 1992). Dynamic models of ecological communities should not neglect the spatial aspect, though most of them do so (Green, 1989). The simplification of using a non-spatial model may lead to invalid forecasts, because spatial relationships can radically change conditions for persistence and coexistence (Czárán and Bartha, 1992). Bascompte and Solé (1995) point out that quite complex patterns may emerge from simple spatio-temporal models. Explaining complexity from simple local rules is particularly attractive for modellers. Applying only non-spatial analyses, these complex patterns may easily be misinterpreted.

## 2. Cellular automata

Cellular automata are dynamic models that are discrete in time, space and state. A simple cellular automaton  $A$  is defined by a lattice  $L$ , a state space  $Q$ , a neighbourhood template  $\delta$  and a local transition function  $f$  (Adamatzky, 1994):

$$A = \langle L, Q, \delta, f \rangle.$$

Each cell of  $L$  can be in a discrete state out of  $Q$ . The cells can be linked in different ways. In the simplest case they are connected geometrically according to a spatial order, such as in a one- or two-dimensional square grid or in hexagonal plots. The cells are indexed by numbers  $s \in I$ , where  $I$  is the index set.

Cells can change their states in discrete time-steps. Usually cellular automata are synchronous, i.e. all cells change their states simultaneously. The fate of a cell is dependent on its neighbourhood and the corresponding transition function  $f$ . Frequently used neighbourhood templates for two-dimensional square grids are the Moore-neighbourhood consisting of the central cell and eight adjacent cells, and the von Neumann-neighbourhood, containing the central cell and four

adjacent cells. In a slight modification of Li (1995) we call the Moore-neighbourhood ‘second order’ and the von Neumann-neighbourhood ‘first order’. The corresponding transition rules can be deterministic or stochastic (Czárán and Bartha, 1992) and are denoted in the form

$$a_{t+1}^s = f(a_t^{s-r}, \dots, a_t^s, \dots, a_t^{s+r})$$

where  $a_t^s$  represents the state of cell  $s$  at time  $t$ ,  $r$  is the range of the neighbourhood of cell  $s$ , and  $f$  is the local transition function representing the transition rules. The set of values  $\{a_t^s | \forall s \in I\}$  is also called a configuration of the cellular automaton at time  $t$ . Probably the most popular cellular automaton is Conway’s LIFE (Gardner, 1971).

If several rules are used, the order of processing these rules may be particularly important for the model results (Gilpin, 1990; Ruxton, 1996).

Discussing the significance of cellular automata to ecological theory, Phipps (1992) draws the conclusion that, as far as applications to natural systems are concerned, stochastic rules usually have a better analogy to the system than deterministic ones, though sometimes their heuristic value is less. For theoretical contributions to the understanding of cellular automata see Wolfram (1983, 1984) and Adamatzky (1994).

Cellular automata have a broad range of applications, for example modelling chemical reactions with spatial diffusion, the development of spiral galaxies, for phase transitions and crystal growth, but also quite often to model biological and ecological systems (Wolfram, 1983; Molofsky, 1994). Extensive literature exists today about applied modelling with cellular automata. Table 1 shows some selected references in alphabetical order. Further references concerning applications to ecology are given by Phipps (1992) and Balzter et al. (1996).

Cellular automata sometimes appear under different names, such as individual-based dynamical automaton model (IBDA (Wiegand et al., 1994)), cellular space, modular computer, iterative computer and tessellation automaton (Smith III, 1971). These belong to the class of methods of emergent computation (Olson and Sequeira, 1995) which process changes of several objects simultaneously.

Kareiva and Wennergren (1995) point out that a strength of these models is their capability of

Table 1  
Use of cellular automata

Authors	Topics
Balman (1993)	Economic rural development
Balzter et al. (1996)	Three plant populations in a lawn
Bockstael et al. (1995)	Land use, ecological and economic development
Colasanti and Grime (1993)	Succession on resource gradients
Crawley and May (1987)	Competition between annual and perennial plant species
Darwen and Green (1996)	Extinction of a population
Dytham (1995)	Coexistence between two plant species
Ellison and Bedford (1995)	Impacts of a power plant to a wetland community
Epperson (1995)	Distribution of genotypes of individuals
Flamm and Turner (1994)	Land cover
Gerhardt et al. (1990)	Excitable media in physics
Green (1989)	Effects of fire, dispersal and spatial pattern on a forest
Hassell et al. (1991)	Host-parasite interactions in insect populations
Karlson and Jackson (1981)	Community structure of sessile invertebrates
Kummer et al. (1994)	Spread of rabies in fox populations
Liu et al. (1995)	Effects of a forest management plan on a bird species
Loh and Hsieh (1995)	Succession in a savanna landscape
Marsula and Ratz (1994)	Effect of fire on a serotine plant species
Molofsky (1994)	Pattern formation in theoretical populations
Ratz (1994)	Effect of fire in boreal forests
Silvertown et al. (1992)	Five competing grass species
Van Tongeren and Prentice (1986)	Succession of heathland after fire
Wiegand et al. (1994)	Semi-arid shrub ecosystem
Wilkie and Finn (1988)	Land use change
Winkler et al. (1994)	Dry grassland community

modelling ecosystems with respect to spatial segregation, which is often a condition for coexistence of predator and prey populations.

Silvertown et al. (1992) model interspecific competition between five grass species. The cellular automaton consists of a  $40 \times 40$  square grid and uses a first order neighbourhood. Its rules are time-invariant and based upon the observed extent of invasion across boundaries between the species grown in hexagonal plots in an experiment taken from the literature. Silvertown et al. take the transition matrix of a Markov chain and restore the spatial dimension important in real plant communities. The transition probabilities are weighted by the number of adjacent cells occupied by the invading species, but not for the persistent species. Thus, cells tend to remain in a given state longer than would be expected if the self-transition probabilities were also weighted. For an alternative way of weighting the probabilities see Acevedo et al. (1995).

The model of Silvertown et al. (1992) combines two important properties, the stochasticity of Markov chains and the spatial dimension of cellular automata.

### 3. Stochastic spatio-temporal models

#### 3.1. Introducing a spatial dimension to a Markov chain

Consider a discrete-time Markov chain. The elements of the state space  $Z$  are possible outcomes of a group of discrete random variables  $X_{t_i}$  ( $t_i \in T$ ) that are in certain states at several equidistant times  $t \in T$ , where  $T = \{t_0, t_1, \dots\}$  is the parameter space of the Markov chain.

In terms of the Markov model the transition probabilities from one state to another are denoted as

$$P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n) = p_{i_n j_{n+1}} \quad (1)$$

the probability that random variable  $X_{t_{n+1}}$  will be in state  $i_{n+1}$  at time  $t_{n+1}$  under the condition that  $X_{t_n}$  was in state  $i_n$  at the previous time-step. The Markov property is then expressed by Eq. (2).

$$\begin{aligned}
 P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n, X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} \\
 = i_1, X_{t_0} = i_0) = P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n).
 \end{aligned}
 \tag{2}$$

In other words the probability distribution of the outcomes of a single trial depends solely on the outcome of the immediately preceding trial and this temporal first order dependence is the same for all  $t_i \in T$ . A Markov chain is completely defined by (i) the initial state probability distribution  $\underline{p}(t_0)$  and (ii) the matrix of transition probabilities  $P = [p_{ij}]$ . Here we consider the transition probabilities as time independent and the time-steps as approximately equidistant, so that they can be enumerated by  $\{0, 1, \dots, T\}$  instead of  $\{t_0, t_1, \dots, t_n\}$ .

A state probability distribution  $\underline{p} = \{p_j\}$  is called stationary if

$$p_j = \sum_i p_i p_{ij}.$$

Let  $p_{ij}^{(t)}$  denote the probability that an object in state  $i$  will be in state  $j$  after  $t$  time-steps. Then every irreducible aperiodic Markov chain belongs to one of the following two classes (Feller, 1957; Stewart, 1994): (i) either the states are all transient or all null states; in this case

$$\lim_{t \rightarrow \infty} p_{ij}^{(t)} = 0 \quad \forall(i, j)$$

and no stationary distribution exists, or (ii) all states are ergodic, that is

$$\lim_{t \rightarrow \infty} p_{ij}^{(t)} = \pi_j > 0 \quad \forall(i, j),$$

where  $\{\pi_j\}$  is the unique stationary distribution. For an irreducible aperiodic Markov chain, however, a limiting distribution always exists. The limiting distribution is ergodic if it is independent of the initial state distribution vector  $\underline{p}(0)$ . It is calculated using Eq. (3) (Stewart, 1994).

$$\Pi = \lim_{t \rightarrow \infty} (\underline{p}(0) \cdot P^t) \tag{3}$$

where  $\Pi$  is a matrix consisting of equal rows,  $\underline{p}(0)$  is the initial state probability vector and  $P^t$  is the transition matrix raised to the  $t$ -th power. The vector of the limiting distribution  $\underline{\pi}$  is any row of the matrix  $\Pi$ .

Including a spatial dimension of the data into the model, we let every site pass separately

through the Markov chain, noting its location. This results in spatial maps of states for each time-step. This Markov chain is spatial, because the location of each site (plot, object) that moves through the chain is considered important.

Here, another approach of introducing a spatial dimension into a Markov model will be mentioned. Acevedo et al. (1995) model forest dynamics by a semi-Markov process. They parameterise the semi-Markov model with previously carried out simulations of the deterministic forest model ZELIG. A spatial dimension is then introduced, whereby environmental factors (temperature, soil moisture and soil fertility) with values ranging from 0 to 1 are stored in a geographic information system (GIS), used to adjust the parameters of the semi-Markov model for each cell. Results are transferred back to the GIS for display and analysis. Each of the cells is modelled as a mosaic of smaller, gap-scale plots and its state is given as the proportion of area covered by several cover types. The three environmental factors are multiplied to calculate a compound limiting factor for each role and every cell. Transition probabilities are weighted with these limiting factors. To include neighbourhood effects new variables are obtained by averaging the layers of cover maps for each type  $i$  of all cells surrounding the cell. The transition probabilities are then multiplied by a correction factor to ensure that the probabilities of each row sum up to one (Acevedo et al., 1995).

### 3.2. Cellular automata as spatio-temporal Markov chains

Let us now consider a cellular automaton with a first order neighbourhood. The transition of a cell during a time-step is thus dependent on the configuration of states of nine cells one time-step ago. First we modify the time index of a simple Markov chain to a purely spatial index, then we consider a spatio-temporal cellular automaton model. If not stated otherwise we assume the time and space indices to be equidistant.

Regarding the index set  $I = \{0, 1, \dots, T\}$  of a Markov chain as a spatial index leads to the theory of Markov random fields (Guttorp, 1995; Li, 1995). The Markov property for one spatial

dimension is then formulated by Eq. (4) and is independent of the direction of the indices  $t \in I$ .

$$\begin{aligned}
 P(X_t = i_t | X_{t-1} = i_{t-1}, \dots, X_{t-k} = i_{t-k}, X_{t+1} \\
 = i_{t+1}, \dots, X_{t+m} = i_{t+m}) \\
 = P(X_t = i_t | X_{t-1} = i_{t-1}, X_{t+1} = i_{t+1}). \quad (4)
 \end{aligned}$$

Writing  $\underline{X}_Z$  to denote,  $\{X_s, s \in Z\}$ ,  $\underline{X}$  is a Markov random field, if  $\underline{X}$  is a  $Z^G$ -valued random vector satisfying Eq. (5).  $\underline{X}_{-s}$  indicates the vector of random variables at all sites except site  $s$  and  $\underline{i}_{-s}$  is the vector of states of all sites except site  $s$ .

$$\begin{aligned}
 P(X_s = i_s | \underline{X}_{-s} = \underline{i}_{-s}) = P(X_s = i_s | \underline{X}_{\delta_s} = \underline{i}_{\delta_s}) \\
 = p_s(x_s | x_{\delta_s}) \quad (5)
 \end{aligned}$$

where  $p_s$  are local characteristics at site  $s$ , that can be site specific, but here are considered as stationary (identical for each  $s$ ), and  $\delta_s$  is the neighbourhood of  $s$ . Note that the neighbourhood of  $s$  in a Markov random field does not include site  $s$ .

Markov random fields are used e.g. for image restoration, smoothing, interpolation and edge detection. Considering rectangular lattices a Markov random field is called first order if the neighbourhood consists of four adjacent cells, second order if it consists of eight adjacent cells, and third order if the neighbourhood contains 24 cells (Li, 1995). Markov random fields are equivalent to Gibbs random fields (Li, 1995).

The concept of a Markov random field does not refer to time, which shall now be introduced. Furthermore, the Moore-neighbourhood of cellular automata includes both its eight neighbours and the central cell one time-step previously, while the Markov random field forecasts the state of the central cell depending on its eight neighbours at one time. Extending the time index  $t$  to both a spatial and a temporal dimension, we get  $\{X_{s,t}\}$  as a set of random variables that are both spatially ( $s = \{0, 1, 2, \dots\}$ ) and temporally ( $t = \{0, 1, 2, \dots\}$ ) ordered.  $s$  can be a one-, two or more-dimensional space. Let us redefine  $\delta_s$  as the set of cells of the Moore-neighbourhood of cell  $s$ ,  $i_{s,t}$  as the state of cell  $s$  at time  $t$  and  $\underline{i}_{\bullet,t}$  the vector of the states of all cells of the grid at time  $t$ . Then the spatio-temporal Markov property is expressed by Eq. (6).

$$\begin{aligned}
 P(X_{s,t+1} = i_{s,t+1} | \underline{X}_{\bullet,t} = \underline{i}_{\bullet,t}, \underline{X}_{\bullet,t-1} \\
 = \underline{i}_{\bullet,t-1}, \dots, \underline{X}_{\bullet,0} = \underline{i}_{\bullet,0}) \\
 = P(X_{s,t+1} = i_{s,t+1} | \underline{X}_{\delta_s,t} = \underline{i}_{\delta_s,t}). \quad (6)
 \end{aligned}$$

The transition probabilities of a cell  $s$  thus only depend upon the values in its neighbourhood one time-step previously. This type of model is called a spatio-temporal Markov chain (STMC). It is temporally first order and spatially second order, denoted STMC(1,2). Clearly, it may be extended to higher orders. The introduction of space in a Markov chain assuming spatial independence as it was described in Section 3.1 is a special case of such an STMC model, namely STMC(1,0).

In conclusion, any cellular automaton satisfying the following conditions is an STMC:

1. spatial dependence: the future of cell  $s$  depends on a neighbourhood  $\delta_s$ .  $\delta_s$  can be restricted to the central cell  $s$ , resulting in an STMC( $t,0$ ) which is a simple, non-spatial Markov chain;
2. temporal dependence: the future states of cells must depend on the past states. The structure of dependence must be Markovian, i.e. the transition rules must be of the form stated in Eq. (6).

## 4. Applications

### 4.1. Competition of two plant species

In a theoretical approach, Dytham (1995) uses a cellular automaton to model the persistence of two plant species in environments with different habitat destruction patterns. Dytham (1995) assumes that two competing species can only coexist if one is the competitive dominant (species  $c$ ) and the other possesses superior dispersal abilities (species  $d$ ). There are four possible patch states in his model: patch permanently destroyed; patch empty but available for colonisation; patch occupied by species  $c$ ; patch occupied by species  $d$ . Colonisation has a range of only one patch. Therefore, the probability of a patch being colonised is higher if more of its direct neighbours contain the species. During each discrete time-step an occupied cell can become extinct with proba-

bility 0.1. The rate of colonisation is 0.2 for the superior and 0.5 for the inferior competitor, and each of the eight adjacent patches can be colonised. Species *c* always wins against species *d*, while species *d* can only colonise empty cells. Extinctions and colonisations occur during the same time-step, but for a newly colonised patch it is impossible to immediately produce propagules. The impact of four habitat destruction patterns on population dynamics is simulated on a grid of  $50 \times 50$  cells. Dytham's results indicate that species' coexistence heavily depends not only on the amount of available habitat, but also on the destruction pattern.

A sensitivity analysis of the model to the starting proportions of the two species is also carried out by Dytham (1995). For the gradient habitat destruction pattern the results are not significantly different for proportions of 0.1, 0.25 and 0.5 occupation of the remaining habitat for each species (ANOVA  $P > 0.05$ ). The number of empty cells also does not differ significantly, if the model runs for 1000, 1500 or 2000 time-steps (ANOVA  $P > 0.05$ ).

Dytham's model is an STMC(1,2). The model's state space is  $Q = \{x, e, c, d\}$ , where *x* means 'destroyed', *e* 'empty', *c* 'occupied by species *c*' and *d* 'occupied by species *d*'. The model uses the second order neighbourhood of nine cells. It is not possible to estimate all transition probabilities accurately, because obviously there are 262144 possible state vectors  $\underline{X}_{\delta,t}$  of the neighbourhood from which transitions can occur, and the lattice *L* is comprised of only 2500 cells. Nevertheless, some probabilities can be inferred from the concept of the model, for example  $P(X_{s,t+1} = c | \underline{X}_{\delta,t} = (x, x, x, x, e, x, x, c, x)) = 0.2$ , which is the probability of an empty cell being colonised by species *c*, given that at time *t* all neighbours are permanently destroyed except the one colonised by species *c*. Although a detailed mathematical analysis of the model was not carried out, the acceptance of the  $H_0$ -hypotheses in the sensitivity analysis is presumably caused by the Markov property. Simulations by the authors show that STMC models often tend towards a limiting distribution, which in some cases is ergodic. Without mathematical proof, we propose that in Dytham's

case an ergodic distribution exists, independent of the initial configuration. If this is true, convergence should be achieved in far less than 1000 time-steps, which would explain the model behaviour.

#### 4.2. Vegetation dynamics

Vegetation samples were taken by the authors from a 616 m<sup>2</sup> lawn each June from 1993 to 1996. The study site is located at the Agricultural Nursery at Justus-Liebig-University in Giessen, Germany. Vegetation sampling followed the point-quadrat method, yielding spatial data for  $\approx 40$  plant species. Study site and methodology are described more thoroughly in Balzter et al. (1995). Population dynamics of three plant species are modelled with a cellular automaton. The species are *Lolium perenne*, the perennial ryegrass, a widely used forage grass which is also common in lawns in sports fields, *Trifolium repens*, white clover, belonging to *leguminosae* and also used as a forage plant, and *Glechoma hederacea*, the ground ivy, a common ruderal species usually found on the edges of tracks. All three species spread generatively and by above-ground stolons. Percentage cover is used to characterise population dynamics. It is defined as the proportion of bare ground that would be shaded by a plant species if the sun were to shine in directly perpendicular to the ground surface. Because different plant species usually overlap, percentage cover is usually  $> 100\%$  in dense communities and  $< 100\%$  in sparse communities.

The data are spatially ordered in 12 columns and ten rows resulting in a lattice *L* of 120 cells. Two modelling approaches are conducted: an STMC(1,0) and an STMC(1,2) model. The state space of both models is  $Q = \{0, 1\}$ , state 0 meaning 'species absent' and 1 meaning 'species present'. For STMC(1,0) the transition probability from state *i* into state *j* is estimated as the number of cells, that are in state *i* at time  $t-1$  and in state *j* at time *t*, divided by the number of cells in state *i* at time  $t-1$ . The transitions are pooled over time, i.e. transitions from 1993 to 1994 and 1994 to 1995 are counted together. Table 2 gives the transition matrices of this

Table 2  
Transition matrices for the STMC(1,0) model and the three plant species

From	<i>Lolium perenne</i>		<i>Trifolium repens</i>		<i>Glechoma hederacea</i>	
	To					
	0	1	0	1	0	1
0 (Absent)	0.55	0.45	0.53	0.47	0.92	0.08
1 (Present)	0.45	0.55	0.23	0.77	0.62	0.38

0, Species absent; 1, species present.

model. Preliminary results are given in Balzter et al. (1996)<sup>1</sup>.

In the STMC(1,2) model there are 512 different possible neighbourhood configurations. To ensure sufficient accuracy for the probability estimation, given the low number of observations, the set of possible neighbourhoods is grouped into five intervals according to the sum of cells in the neighbourhood occupied by the species (Table 3). Afterwards, pooled transitions from group *i* to state *j* are counted in the way described for the STMC(1,0) model above. The resulting 5 × 2 transition matrices are shown in Table 4. At the edges a problem arises with incomplete neighbourhoods, whereby cells are ‘reflected’ over the border to give the same number of neighbours. Reflection is presumably more appropriate to vegetation data than other methods (see Haefner et al. (1991) for a comparison of methods), because the surroundings of the area are heterogeneous (a track, various hedges and cherry trees).

<sup>1</sup> The STMC(1,2) model was called ‘Moore neighbourhood model’ and the STMC(1,0) was called ‘spatial Markov chain’.

Table 3  
Grouping of neighbourhood states for the STMC(1,2) model

Neighbourhood group	<i>Lolium perenne</i>	<i>Trifolium repens</i>	<i>Glechoma hederacea</i>
0	0	0	0
1	1–2	1–2	1
2	3–4	3–4	2
3	5–7	5–7	3
4	8–9	8–9	4–9

Numbers are the sums of neighbouring cells occupied by the species. Because percentage cover is smallest for *Glechoma hederacea*, followed by *Lolium perenne* and finally *Trifolium repens*, the classification is species-dependent.

STMC(1,0) and STMC(1,2) yield similar results. Some graphical presentations are shown in Fig. 1, suggesting the existence of ergodic distributions. This is proven for the STMC(1,0) model, because all states are aperiodic classes of positive recurrent states and the state space is finite (Heller et al., 1978). Table 5 shows the ergodic distributions for the STMC(1,0) and the limiting distributions for the STMC(1,2) model. Whether the limiting distributions of the STMC(1,2) model are ergodic, i.e. independent of the initial configuration, is not clear, although there is some evidence from simulation runs with different initial configurations, all of which result in the same limiting distribution.

The STMC(1,2) model is now approximated by a simple Markov chain with a new state space, further abbreviated as the APPROX model. Considering the neighbourhood configurations of the STMC(1,2) model as stated in the APPROX model, we define new random variables  $Y_t$  as the sum over the neighbourhood of cell *s* at time *t*:

$$Y_t = \sum_{k \in \delta_s} X_{k,t} \tag{7}$$

Table 4

Transition matrices for the STMC(1,2) model using five neighbourhood groups for the three plant species.

From neighbourhood group	<i>Lolium perenne</i>		<i>Trifolium repens</i>		<i>Glechoma hederacea</i>	
	To					
	0	1	0	1	0	1
0	0.91	0.09	1.00	0.00	0.99	0.01
1	0.52	0.48	0.54	0.46	0.73	0.27
2	0.54	0.46	0.50	0.50	0.86	0.14
3	0.46	0.54	0.27	0.73	0.29	0.71
4	0.41	0.59	0.18	0.82	0.33	0.67

The state space is then  $Q_{\text{APPROX}} = \{0, 1, \dots, 9\}$ . An example for a transition matrix of the APPROX model is given by Table 6. The APPROX model is used to simplify the calculation of limiting distributions for the STMC(1,2) model. The APPROX model has ergodic distributions for *Lolium perenne* and *Trifolium repens*. For *Glechoma hederacea* the transition matrix is modified, because there are no transitions observed, which leave states 5, 7, 8 and 9, but there are transitions leading into states 5 and 7. States 8 and 9 could thus be eliminated and the transitions from  $i = 5$  and  $i = 7$  are assumed to be random, i.e.  $P_{ij} = 0.125$  for all  $j$ .

Although the assumption of spatial independence is violated, the ergodic distributions of the APPROX model come fairly close to the predictions of the STMC(1,2) model for *Lolium perenne* and *Trifolium repens* (Table 6). The deviating result for *Glechoma hederacea* is probably caused by the modification of the transition matrix.

Model validation is carried out using the data from June 1996, which is equivalent to the prediction of time-step 1. The APPROX model is not validated, because it is simply an approximation of the STMC(1,2) model. The results are shown in Table 7.  $\chi^2$ -tests are carried out in two steps, first testing STMC(1,0) versus STMC(1,2) versus observations in a  $3 \times 2$  table and, if in this test  $H_0$  was rejected, testing all three pairwise comparisons in  $2 \times 2$  tables. The first  $\chi^2$ -tests indicate significant differences for *Lolium perenne* ( $P < 10^{-15}$ ) and *Trifolium repens* ( $P < 10^{-15}$ ), but homogeneity for *Glechoma hederacea* ( $P = 0.38$ ). For *Lolium perenne* and *Trifolium repens* STMC(1,0) and STMC(1,2)

do not differ significantly ( $P = 1$  and  $0.49$ ), but both models are different from the observations in 1996 (all  $P < 10^{-14}$ ). Balzter et al. (1996) argue that the STMC(1,2) model predictions for *Glechoma hederacea* are unlikely to match reality in the long-run, because the limiting distribution is higher than is expected from ecological intuition. However, this model seems to produce reasonable output for a short time. The limiting distribution and thus ecological equilibrium is probably never reached in reality anyway, because in most ecosystems the environment is likely to change over longer time intervals, influencing the dynamics of the organisms and the underlying probabilities of the stochastic processes (Lippe et al., 1985). Successions in ecological communities seldom reach equilibrium or climax states. Instead communities tend towards a relative climax until environmental conditions change. For grazed communities the climax depending on the herbivore population is called the zootic climax (Daubenmire, 1968). Possibly the model gives a good fit for the near future of *Glechoma*, but fails to predict long-term changes. This is due to a violation of the model assumption that the transition probability matrices be stationary over time, including constancy of environmental conditions.

Assuming a preference of rabbits for forage plants supports our hypothesis. It could explain the species-dependent validity of the model results. To improve the model fit the transition matrices for *Lolium* and *Trifolium* are changed arbitrarily. In order to obtain predicted values for 1996 as small as the observed 5% for *Lolium* and 16% for *Trifolium*, the transition matrices must be presented



as in Table 8. The probabilities of changing the state from a colonised plot (or a plot with many colonised neighbours) to an empty plot are higher than the estimates from Tables 2 and 3. Catastrophic reductions in population size must have

occurred, for the survival probabilities (present to present) to decrease and the extinction probabilities (present to absent) to increase so significantly between 1995 and 1996.

## 5. Discussion

Finding appropriate rules for the cellular automaton model is a crucial task. Either the rules are improved iteratively (bottom-up approach (Kummer et al., 1994)), or rules are identified by a specific algorithm (top-down approach (Adamatzky, 1994)). Great care must be taken when the structure of spatio-temporal dependence is defined. Neighbourhood definition affects model results in a species-specific way. Maybe the spatio-temporal correlation structure of the examined variables can give hints for selection of the correct temporal and spatial orders of the model. Correlograms for different time and space lags could be applied to determine the correct neighbourhood and the appropriate temporal dependence.

A major problem with stochastic population models is the influence of unobserved (and often unobservable) variables that affect the probability structure over time. Although proper planning of the sampling scheme can sometimes avoid this difficulty, exhaustive sampling of all relevant factors is either not always possible or too expensive. In our study of the lawn dynamics in Giessen, the size of the rabbit population is suspected to be of particular importance. Although this was taken into account by the authors as a possible cause for changing transition probabilities over time (Balzter et al., 1996), the rabbit population was initially expected to take a longer period than 3 years to influence vegetation dynamics so drastically. However, the interpretation of transition probabilities provides knowledge about the examined system. Considering the strong decrease of *Lolium* and *Trifolium*, the true probabilities of colonising an empty plot must have been near 0.

The low observed values in 1996 for *Trifolium repens* and *Lolium perenne* can be explained by an unexpectedly strong increase in the size of the rabbit population from 1993 to 1996. We assume

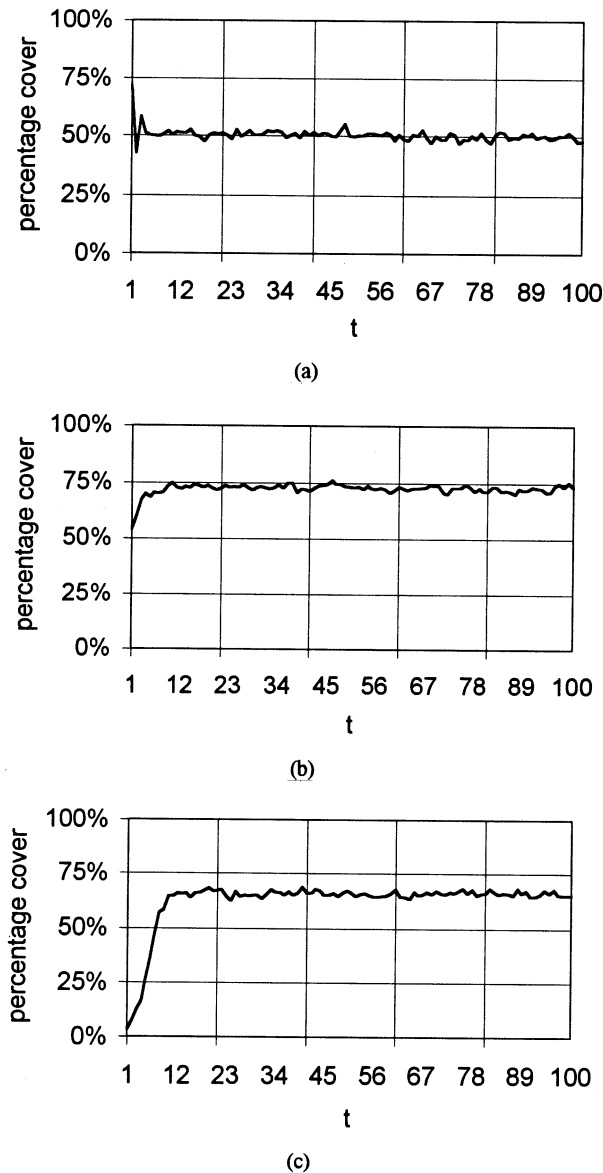


Fig. 1. Predictions of percentage cover over 100 time-steps predicted by the STMC(1,2) model. Means of 10 simulation runs. (a) *Lolium perenne*, (b) *Trifolium repens*, (c) *Glechoma hederacea*.

Table 5

Ergodic distributions of the STMC(1,0) model and the approximating model to the STMC(1,2) model (APPROX) and limiting distributions of the STMC(1,2) model

State:	<i>Lolium perenne</i>		<i>Trifolium repens</i>		<i>Glechoma hederacea</i>	
	0	1	0	1	0	1
STMC(1,0)	0.50	0.50	0.33	0.67	0.89	0.11
APPROX	0.49	0.51	0.26	0.74	0.79*	0.21*
STMC(1,2)	0.50	0.50	0.28	0.72	0.34	0.66

0, Absent; 1, present.

Values for STMC(1,0) and APPROX are calculated by Eq. (3), while values for STMC(1,2) are obtained by simulations from the means of the results of ten simulation runs and time-steps 81...100.

\* Indicates a modified transition matrix, see text.

Table 6

Example for a transition probability matrix of the APPROX model for *Trifolium repens*

Occupied cells:	0	1	2	3	4	5	6	7	8	9
0	0.29	0.29	0.14	0.00	0.14	0.00	0.14	0.00	0.00	0.00
1	0.50	0.00	0.00	0.25	0.00	0.00	0.25	0.00	0.00	0.00
2	0.00	0.00	0.05	0.40	0.00	0.25	0.15	0.10	0.05	0.00
3	0.00	0.00	0.11	0.15	0.19	0.04	0.15	0.15	0.11	0.11
4	0.00	0.00	0.10	0.26	0.16	0.07	0.23	0.07	0.00	0.13
5	0.00	0.00	0.03	0.09	0.09	0.06	0.17	0.20	0.23	0.14
6	0.00	0.00	0.00	0.00	0.03	0.21	0.21	0.26	0.10	0.21
7	0.00	0.00	0.00	0.03	0.08	0.18	0.21	0.23	0.10	0.18
8	0.00	0.00	0.00	0.00	0.00	0.13	0.30	0.13	0.22	0.22
9	0.00	0.00	0.00	0.00	0.07	0.00	0.27	0.20	0.13	0.33

Relative frequencies of transitions from the sums of one to another second order neighbourhood.

that certain plant species (forage plants as *Lolium perenne* and *Trifolium repens*) are preferred food plants for these herbivores compared with others and that selective grazing can facilitate growth of other populations like *Glechoma hederacea*. This mechanism could cause changes in the transition matrices, which can no longer be assumed to be

Table 7

Comparison of percentage cover values forecasted for 1996 by the STMC(1,0) and the STMC(1,2) model versus observed values

	<i>Lolium perenne</i>	<i>Trifolium repens</i>	<i>Glechoma hederacea</i>
STMC(1,0)	51%	66%	11%
STMC(1,2)	51%	70%	17%
Observed	5%	16%	16%

stationary in time. Examples for strong effects of grazing by small mammals on vegetation are well known. Batzli and Pitelka (1970) exclude small rodents (meadow mice) from grasslands in California and find that, in the two years following exclosure, food plants like *Lolium multiflorum* and *Avena fatua* contribute > 60% to the vegetational volume on areas without grazing but < 30% on grazed areas. The decrease in forage species under rodent grazing is highly significant (ANOVA  $P < 0.01$ ). They also observe a strong influence of rainfall (ANOVA  $P < 0.0005$ ) resulting in increased growth of *Lolium multiflorum* and *Avena fatua*, which nevertheless is still greater on ungrazed areas. Ricklefs (1979) mentions the rapid regeneration of native pine *Callitris* in New South Wales following the introduction of the myxomatosis virus that dramatically reduced the rabbit population.

Table 8  
 Changed transition matrices to improve model fit for *Lolium perenne* and *Trifolium repens*

From neighbourhood group	STMC(1,2)		From state	STMC(1,0)	
	To			To	
	0	1		0	1
0	0.99	0.01	0	0.99	0.01
1	0.90	0.10	1	0.80	0.20
2	0.80	0.20			
3	0.80	0.20			
4	0.70	0.30			

Unobserved variables can be incorporated into the model in different ways: (i) the variable is roughly estimated and taken as a covariate on an ordinal scale. The transition probabilities are corrected according to a specified rule and the improvement of the new model is examined; (ii) the stochastic processes can be made more complex by introducing unobserved states that themselves follow a Markov chain and that affect the states of vegetation. This methodology of hidden Markov models is applied to precipitation data by Zucchini and Guttorp (1991); (iii) the influence of the unobserved variable can be determined in separate field trials under controlled conditions. Regression analyses or related statistical methods quantify this influence deterministically; (iv) the deviation of the predictions from observations can stimulate the generation of new hypotheses, which can then be tested in future studies. Scanlan and Archer (1991) provide another way of including external factors into Markovian models. They model vegetation dynamics with simple Markov chains, but take into account the weather dynamics by assuming two weather states (dry and wet) and estimating two transition matrices separately. Which matrix governs the next transition is determined by randomly choosing the next weather state.

The extension of Markov chain theory to spatio-temporal models is a promising field. As modelling becomes more complex, the need for a theoretical framework of these models increases. We argue that model behaviours from the Silver-town model and the Dytham model (and hypo-

thetically many others) could be explained mathematically if theorems on the behaviour of spatio-temporal Markov chains (STMC) were found, where the behaviour of the distribution in the limit is of particular interest. This would allow important conclusions to be drawn regarding observed ecological phenomena, like competition, coexistence, succession or spatial patchiness.

#### Acknowledgements

Heiko Balzter is supported by a grant from the Federal State of Hesse. Our colleague Rod Snowdon corrected the English version of the manuscript. We would also like to thank Marcus Hoffmann for some field work. Last but not least we thank the Royal Statistical Society, London, for organising the RSS 96 conference in Guildford, where this paper was first presented.

#### References

- Acevedo, M.F., Urban, D.L., Ablan, M., 1995. Transition and gap models in forest dynamics. *Ecol. Appl.* 5, 1040–1055.
- Adamatzky, A., 1994. Identification of Cellular Automata. Taylor and Francis, London.
- Balmann, A., 1993. Modellierung regionaler Agrarstrukturentwicklungen mittels des Konzepts zellulärer Automaten. *Zeitschrift für Agrar-informatik* 2(93), 34–41.
- Balzter, H., Braun, P., Köhler, W., 1995. Detection of Spatial Discontinuities in Vegetation Data by a Moving Window Algorithm. In: Gaul, W., Pfeifer, D. (Eds.), *From Data to Knowledge: Theoretical and Practical Aspects of Classification, Data Analysis and Knowledge Organization*. Springer, Berlin, pp. 243–52.

- Balzter, H., Braun, P., Köhler, W., 1996. Modelling Population Dynamics with Cellular Automata. In: Mowrer, T., Czaplewski, R.L., Hamre, R.H. (Eds.), *Spatial Accuracy Assessment in Natural Resources and Environmental Sciences: Second International Symposium*. May 21–23 1996. General Technical Report RM-GTR-277, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, pp. 703–12.
- Bascompte, J., Solé, R.V., 1995. Rethinking complexity: modelling spatio-temporal dynamics in ecology. *TREE* 10, 361–366.
- Batzli, G.O., Pitelka, F.A., 1970. Influence of meadow mouse populations on California grassland. *Ecology* 51, 1027–1039.
- Bockstael, N., Costanza, R., Strand, I., Boynton, W., Bell, K., Wainger, L., 1995. Ecological economic modeling and valuation of ecosystems. *Ecol. Econ.* 14, 143–59.
- Colasanti, R.L., Grime, J.P., 1993. Resource dynamics and vegetation processes: a deterministic model using two-dimensional cellular automata. *Funct. Ecol.* 7, 169–176.
- Crawley, M.J., May, R.M., 1987. Population dynamics and plant community structure: competition between annuals and perennials. *J. Theor. Biol.* 125, 475–489.
- Czárán, T., Bartha, S., 1992. Spatio-temporal dynamic models of plant populations and communities. *TREE* 7, 38–42.
- Darwen, P.J., Green, D.G., 1996. Viability of populations in a landscape. *Ecol. Model.* 85, 165–171.
- Daubenmire, R., 1968. *Plant Communities*. Harper and Row, New York.
- Dunning, J.B., Stewart, D.J., Danielson, B.J., Noon, B.R., Root, T.L., Lamberson, R.H., Stevens, E.E., 1995. Spatially explicit population models: current forms and future uses. *Ecol. Appl.* 5, 3–11.
- Dytham, C., 1995. The effect of habitat destruction pattern on species persistence: a cellular model. *OIKOS* 74, 340–344.
- Ellison, A.M., Bedford, B.L., 1995. Response of a vascular plant community to disturbance: a simulation study. *Ecol. Appl.* 5, 109–123.
- Epperson, B.K., 1995. Spatial structure of two-locus genotypes under isolation by distance. *Genetics* 140, 365–375.
- Feller, W., 1957. *An Introduction to Probability Theory and Its Applications*, vol. I, 2nd ed. Wiley, New York.
- Flamm, R.O., Turner, M.G., 1994. Alternative model formulations for a stochastic simulation of landscape change. *Landsc. Ecol.* 9, 37–46.
- Gardner, M., 1971. On cellular automata, self-reproduction, the garden of Eden and the game 'life'. *Sci. Am.* 224, 112–117.
- Gerhardt, M., Schuster, H., Tyson, J.J., 1990. A cellular automaton model of excitable media including curvature and dispersion. *Science* 247, 1563–1566.
- Gilpin, M.E., 1990. Extinction of Finite Metapopulations in Correlated Environments. In: Shorrocks, B., Swingland, I.R. (Eds.), *Living in a Patchy Environment*. Oxford University Press, New York, pp. 177–86.
- Green, D.G., 1989. Simulated effects of fire, dispersal and spatial pattern on competition within forest mosaics. *Vegetatio* 82, 139–153.
- Guttorp, P., 1995. *Stochastic Modelling of Scientific Data*. Chapman and Hall, London.
- Haefner, J.W., Poole, G.C., Dunn, P.V., Decker, R.T., 1991. Edge effects in computer models of spatial competition. *Ecol. Model.* 56, 221–244.
- Hassell, M.P., Comins, H.N., May, R.M., 1991. Spatial structure and chaos in insect population dynamics. *Nature* 353, 255–258.
- Heller, W.-D., Lindenberg, H., Nuske, M., Schriever, K.-H., 1978. *Stochastische Systeme*. De Gruyter, Berlin.
- Kareiva, P., Wennergren, U., 1995. Connecting landscape patterns to ecosystem and population Processes. *Nature* 373, 299–302.
- Karlson, R.H., Jackson, J.B.C., 1981. Competitive networks and community structure: a simulation study. *Ecology* 62, 670–678.
- Kummer, G., Jeltsch, F., Brandl, R., Grimm, V., 1994. Kopplung von prozessen auf lokaler und regionaler skala bei der tollwutausbreitung: ergebnisse eines neuen modellansatzes. *Verh d Ges f Ökologie* 23, 355–364.
- Li, S.Z., 1995. *Markov Random Field Modeling in Computer Vision*. Springer, Tokyo.
- Lippe, E., De Smidt, J.T., Glenn-Lewin, D.C., 1985. Markov models and succession: a test from a heathland in the Netherlands. *J. Ecol.* 73, 775–791.
- Liu, J., Dunning, J.B., Pulliam, H.R., 1995. Potential effects of a forest management plan on Bachman's sparrow (*Aimophila aestivalis*): linking a spatially explicit model with GIS. *Conserv. Biol.* 9, 62–75.
- Loh, D.K., Hsieh, Y.-T.C., 1995. Incorporating rule-based reasoning in the spatial modeling of succession in a Savanna landscape. *AI Appl.* 9, 29–40.
- Marsula, R., Ratz, A., 1994. Einfluß von Feuer auf die populationsdynamik von serotinen pflanzen-ein modell. *Verh d Ges f Ökol* 23, 365–372.
- Molofsky, J., 1994. Population dynamics and pattern formation in theoretical populations. *Ecology* 75, 30–39.
- Olson, R.L., Sequeira, R.A., 1995. Emergent computation and the modeling and management of ecological systems. *Comput. Electron. Agric.* 12, 183–209.
- Phipps, M.J., 1992. From Local to Global: The Lesson of Cellular Automata. In: DeAngelis, D.L., Gross, L.J. (Eds.), *Individual-Based Models and Approaches in Ecology: Populations, Communities and Ecosystems*. Chapman and Hall, New York, pp. 165–87.
- Ratz, A., 1994. Modellierung feuererzeugter räumlicher Strukturen in borealen Wäldern. *Verh d Ges f Ökol* 23, 373–381.
- Ricklefs, R.E., 1979. *Ecology*, 2nd ed. Chiron Press, New York.
- Ruxton, G.D., 1996. Effects of the spatial and temporal ordering of events on the behaviour of a simple cellular automaton. *Ecol. Model.* 84, 311–314.
- Scanlan, J.C., Archer, S., 1991. Simulated dynamics of succession in a North American subtropical *Prosopis Savanna*. *J. Veg. Sci.* 2, 625–634.

- Silvertown, J., Holtier, S., Johnson, J., Dale, P., 1992. Cellular automaton models of interspecific competition for space—the effect of pattern on process. *J. Ecol.* 80, 527–534.
- Smith III, A.R., 1971. Cellular automata complexity trade-offs. *Inf. Control* 18, 466–482.
- Stewart, W.J., 1994. *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, Princeton, NJ.
- Turner, M.G., Arthaud, G.J., Engstrom, R.T., Hejl, S.J., Liu, J., Loeb, S., McKelvey, K., 1995. Usefulness of spatially explicit population models in land management. *Ecol. Appl.* 5, 12–16.
- Van Tongeren, O., Prentice, I.C., 1986. A spatial simulation model for vegetation dynamics. *Vegetatio* 65, 163–173.
- Wiegand, T., Milton, S.J., Wissel, C., 1994. Ein räumliches Simulationsmodell für eine Pflanzengemeinschaft in der südlichen Karoo Südafrika. *Verh der Ges f Ökol* 23, 407–416.
- Wilkie, D.S., Finn, J.T., 1988. A spatial model of land use and forest regeneration in the Ituri forest of northeastern Zaire. *Ecol. Model.* 41, 307–323.
- Winkler, E., Klotz, S., Wissel, C., 1994. Mathematische Modellierung der Phytozönosedynamik eines Trockenrasens. *Verh d Ges f Ökol* 23, 451–458.
- Wolfram, S., 1983. Statistical mechanics of cellular automata. *Rev. Mod. Phys.* 55, 601–644.
- Wolfram, S., 1984. Universality and complexity in cellular automata. *Physica* 10D, 1–35.
- Zucchini, W., Guttorp, P., 1991. A hidden Markov model for space-time precipitation. *Water Resour. Res.* 27, 1917–1923.