Landscape Ecol (2011) 26:805–822
DOI 10.1007/s10980-011-9610-6

RESEARCH ARTICLE

A Bayesian MCMC approach to reconstruct spatial vegetation dynamics from sparse vegetation maps

Imelda Somodi · Klára Virágh · István Miklós

Received: 22 March 2010 / Accepted: 20 April 2011 / Published online: 4 May 2011

Abstract In studies of vegetation dynamics, data points describing the changes are often sparse, because changes were not recognized in early stages or investigations were part of different projects. The snapshots at hand often leave the nature of the dynamics unrevealed and only give a rough estimation of the directions of changes. Extrapolation of the dynamics with traditional cellular automaton modeling is also complicated in such cases, because rules often cannot be deduced from field data for each interaction. We developed a Bayesian MCMC method, using a discrete time stochastic cellular automaton model to reconstruct vegetation dynamics between vegetation maps available and provide estimation of vegetation pattern in years not surveyed. Spread capability of each vegetation type was characterized by a lateral spread parameter and another for establishment from species pool. The method was applied to a series of three vegetation maps depicting vegetation change at a grassland site following abandonment of grazing in north-eastern Hungary. The Markov chain explored the missing data space (missing maps) as well as the parameter space. Transitions by lateral expansion had a greater importance than the appearance of new vegetation types without spatial constraints at our site. We estimated the trajectory of change for each vegetation type, which bore a considerable non-linear element in most cases. To our best knowledge, this is the first work that tries to estimate vegetation transition parameters in a stochastic cellular automaton based on field measurements and provides a tool to reconstruct past dynamics from observed pattern.

Keywords Cellular automata · Monte Carlo Markov chain · Neighborhoods · Non-linear dynamics · Vegetation maps · Vegetation dynamics · Transition probabilities

Electronic supplementary material The online version of this article (doi:10.1007/s10980-011-9610-6) contains supplementary material, which is available to authorized users.

I. Somodi (✉) · K. Virágh
Institute for Ecology and Botany of the Hungarian Academy of Sciences, 2-4 Alkotmány utca, 2163 Vácrátót, Hungary
e-mail: somodi@botanika.hu; imelda.somodi@gmail.com

I. Somodi
Department of Plant Taxonomy and Ecology, Institute of Biology, Eötvös Loránd University, 1/C Pázmány Péter sétány, 1117 Budapest, Hungary

I. Miklós
Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, 13-15 Reáltanoda utca, 1053 Budapest, Hungary

Introduction

Interest in deducing landscape scale dynamics including natural vegetation dynamics from observed spatial pattern has grown greatly recently (Brook and Bowman 2006; Whited et al. 2007; Gibon et al. 2010;

Jenerette and Potere 2010), especially because of their relevance in predicting climate change effects (Schumacher et al. 2004; Banfai and Bowman 2006; Rickebusch et al. 2007). Nevertheless, due to the large spatial and corresponding temporal extents most of the studies of landscape-scale vegetation dynamics are either based on theoretical assumptions (Schumacher et al. 2004; Karau and Keane 2007) or evaluate only a few snapshots of change. Typical sources of information for the latter are consecutive remote-sensed images (e.g., Flamenco-Sandoval et al. 2007; Whited et al. 2007; Jenerette and Potere 2010) or repeated mapping especially if vegetation is in the focus (e.g., Horváth and Csontos 1992; Blackstock et al. 1995; Somodi et al. 2004). Nevertheless, the sparseness of such snapshots—however informative—leaves the course of the dynamics between these states hidden and makes extrapolation difficult. This is a problem not only for the reconstruction of vegetation dynamics, but is shared by other landscape dynamics studies as well (Jenerette and Wu 2001; Gibon et al. 2010; Jenerette and Potere 2010).

Although remote sensing tends to provide detailed data from recent years, there are rather sparse remotely sensed images from the past especially in the spatial and spectral resolution the recognition of vegetation mapping would require (Cochrane 2000; Schmidt and Skidmore 2003; Pu 2009). Ground-based vegetation maps are even less possible to be obtained on an annual basis. The reason of this is partly the considerable field work requirement, partly the possibility of human error in case of transitions between natural vegetation entities. Thus the process has to be reconstructed, in which we need methods to connect the observed states.

Markov models have been applied to connect known states of vegetation (Baker 1989; Scanlan and Archer 1991; Callaway and Davis 1993), however the traditional form of these disregard spatial information. To handle spatial dependence, a special form of cellular automata was introduced: spatio-temporal Markov chains (STMC, Silverton et al. 1992; Baltzer et al. 1998; Baltzer 2000). In these SMTCs the space is handled as a grid. Each cell of the grid contains one vegetation state (a vegetation type for example), the content of the cell is refreshed in yearly steps. Whether the cell changes from one year to the other is determined by transition rules, which depend at least on the status of the neighboring cell (Baltzer et al. 1998). One of the major parameterization challenges of cellular automata therefore is, to provide probabilities conditional on neighbors as inputs. In most cases, only a limited number of neighborhood configurations arise in reality, therefore it is not possible to base our estimations on observations, but expert or experimental knowledge (Silverton et al. 1992; Dieckmann et al. 1997) or theoretical estimations have to be used (as in Turner 1988; Gassman et al. 2000; Colasanti et al. 2007). The centrality of this problem is also demonstrated by efforts of objectively estimating such rules (Dieckmann et al. 1997; Jenerette and Wu 2001). All the theoretical considerations, however, introduce a large amount of our preconceptions into the model. If too many initial assumptions are included, results of the simulation might reflect our own preconceptions only.

Monte Carlo Markov chains are relatively new additions to the landscape ecologists' methodological repertoire. Despite this, there have already been applications of the method for parameter estimation in related fields (land cover issues: Jenerette and Potere 2010, population distributions: Clark 2005), although not in the form of a mechanistic vegetation dynamics model. Our approach complements this, by applying the technique to a problem of vegetation dynamics.

Our main aim in this study is to provide a tool for reconstructing vegetation dynamics between existing vegetation maps, which

- is spatially explicit
- makes the fewest possible assumptions about the nature of the dynamics
- allows uneven intervals between observations
- gives conditional transition probabilities as outputs and does not require them as inputs.

As a test we applied our method to a series of three vegetation maps depicting vegetation dynamics of an abandoned pasture in Hungary. It is possible to form some hypotheses about the expected behavior of the vegetation entities (vegetation types) in this case study. We expected clearly different patterns of dynamics per entities with different species composition. Some of the vegetation entities are species rich, where the dynamics emerges as a result of many species' individual dynamics. This emergent property can be viewed as typical to that entity (Austin and Smith 1989), which by definition cannot even be

predicted from individual species' behavior (O'Neill 1986; Gaudet and Keddy 1995). Due to their complex nature, we also expected a lower importance of lateral expansion and rather a joint establishment without spatial constraints at favorable sites. Some other vegetation entities are overwhelmingly dominated by one species, therefore, we expected their dynamics to be driven by the population dynamics of that species. Such is the vegetation type dominated by *Calamagrostis epigejos*. This species is known to spread by clonal growth by the phalanx strategy in dry grasslands, because it cannot germinate in dry places (Rebele and Lehmann 2001). Therefore, we expected a strong lateral element in the spread dynamics of the vegetation dominated by this species. Another example is *Elymus repens*, which is also clonal, but is more of a guerilla strategist, i.e., spreads by long runners vegetatively, besides it can sexually reproduce as well (Palmer and Sagar 1963; Fitter and Peat 1994), also under the local conditions. Therefore, we hypothesized that the lateral spread component, though still pronounced, will play a smaller role in its case than for *C. epigejos*-dominated vegetation. Finally, the predominant species of another vegetation type, *Chrysopogon gryllus* is a tussock grass, which is the least likely to be limited to lateral spread.

Methods

Model

Biological and conceptual basis

The theoretical assumptions we make are in common with cellular automata simulation of plant populations or landscape dynamics (e.g., Barkham and Hance 1982): we allow establishment from the neighborhood and without spatial constraints, which correspond to vegetative spread and establishment by seeds respectively. A difference between the simulations of population and vegetation dynamics, however is, that besides establishment by seeds a directed change in the abundance relationships of species can also lead to vegetation transition without spatial constraints (e.g., wooded steppe → lower abundance of woody species, higher abundance of light-demanding undergrowth → steppe). In the current model only the direct neighbors can be colonized by the first method, because based on previous knowledge (see hypotheses, the description of the case study and Somodi et al. 2004, 2008, 2010) none of the dominant or subordinated species in the system modeled are likely to spread laterally by vegetative growth more than 2 m a year, which is the cell size used in the simulations. This is, however, not an inherent limitation of the approach. Long-distance establishment of new patches is accounted for by a spatially unconstrained element of vegetation transitions.

In the proposed model, establishment without spatial constraints is allowed in all the cases, nevertheless, invasion from the neighborhood is either directly the function of weights in the neighborhood transition matrix (if there are at least four different vegetation types in the neighborhood) or is further weighted by frequency in the neighborhood in a simple way (see mathematical details below). In the former case, weights are applied to all the vegetation types present in the eight-cell neighborhood. In the latter case we assume that only those two vegetation types can invade which are most common in the neighborhood, therefore their weights are doubled and other vegetation types are excluded. There are two reasons for using this approach. (1) A practical one: there were nine vegetation types in our experiments (see Table 1). If the likelihood had been recalculated after a parameter change by recalculating all 290,814 transition probabilities on the maps (15,306 positions on each map and 19 transitions), the running time would have been more than 700 times longer than with the proposed approach. The latter however only uses 414 different transition types: a particular position can be occupied by any of the 9 vegetation types, there are 36 possible pairs of the two most frequent neighbor vegetation types, all 9 types might be the only neighbor of a particular position or there might be more than 4 vegetation types around the position. In the software we implemented, the number of transitions for all 414 transition types is counted, and the likelihood of a series of maps is calculated using this weighting approach. This gives a significant acceleration when a parameter is changed and the likelihood has to be recalculated. However, when we did attempt to weight neighbors directly according to their frequency ("full neighborhood enumeration approach") starting from our empirical maps, we could not reach

Table 1 Vegetation types and some of their characteristics

Symbol	Name	Dominant species	Habitat character	Species diversity, physiognomy
F	*Festuca* type	*Festuca rupicola*	Dryer variant of wooded steppe meadows	Diverse, rich in broad-leaved forbs; grasses are of medium height.
Si	*Sieglingia* type	*Sieglingia decumbens* together with *Festuca rupicola*	Mesophilous wooded steppe meadow	The most diverse, rich in broad-leaved forbs; grasses are of medium height.
D	*Danthonia* type	*Danthonia alpina* together with *Festuca rupicola*	Wooded steppe meadow	Diverse, rich in broad-leaved forbs; grasses form two layers of different heights.
Sh	Shrubs	*Crataegus monogyna, Quercus pubescens* (shrubs)	Groups of shrubs and trees, the species composition corresponds to the wooded steppe	A mixture of generalist shrub species and individuals of tree species that were present in the original open forest
Ch	*Chrysopogon* type	*Chrysopogon gryllus*	Wooded steppe meadow overtaken by this species; slightly degraded	Less diverse; grasses form two layers of different heights, taller than the first three types.
E	*Elymus* type	*Elymus repens* with a full transition range towards *Elymus hispidus,* possibly including the hybrid as well	Moderately to completely degraded	Species poor stands of medium height, weeds are present.
B	*Bothriochloa* type	*Bothriochloa ischaemum*	Degraded	Species poor, with a short grass layer
Ca	*Calamagrostis* type	*Calamagrostis epigejos*	Degraded	Almost monodominant stands of the tall grass
L	*Leontodon* type	*Leontodon hispidus*	Degraded	Species-poor, dwarf stands dominated by forbs, numerous weed species present.

Degradation levels correspond to the following scale: close-to-natural means that the stands of the type resemble stands known to have been undisturbed by humans. Slightly degraded → moderately degraded → degraded refers to stages of a degradation gradient, where first the abundance pattern, then the species composition as well departs from the undisturbed form of the closest plant species association

convergence even after 1.5 months pure computer time. This clearly makes that approach impossible to use in practice.

Besides practical reasons, the proposed simulation approach of vegetative spread is well-founded by ecological considerations as well (2). Vegetation dynamics and pattern are known to be spatially autocorrelated. A vegetation type does not spread by individual spatial units (e.g., 2×2 m^2 cells in our case), but in an aggregated way. To preserve this feature in vegetation map simulations Weaver and Perera (2004) has already advised to introduce spatial autocorrelation into spatio-temporal Markov chains and also provided examples which supported their point. The increased weights we applied to the two most common neighbors serve the goal of preserving the spatial autocorrelation structure.

To exclude any possible doubts, however, we carried out a pilot project comparing neighborhood weighting approaches, too, using the same vegetation maps as in the main study. Maps used in the pilot project, though, were coarsened to 10 m cell size and 5-year intervals were used instead of one-year updates, to assure reasonable convergence time even with the full neighborhood enumeration approach. We performed modeling with both the proposed simulation approach, and the full neighborhood enumeration approach and compared the outcomes the same way as in the case of the main simulations.

The mathematical construction

We consider a discrete time Markov model, in which the states are changed in each year. The model is

parameterized by two matrices. The first matrix is $V = \{v_{i,j}\}$, where $v_{i,j}$ describes the ability of vegetation type g_j to occupy the neighbor position in the next year that is currently occupied by vegetation type g_i. The larger the value, the greater the probability that vegetation type g_i will replace g_j. The second matrix is $A = \{a_{i,j}\}$, where $a_{i,j}$ tells the ability of vegetation type g_j to invade vegetation type g_i from a common "pool" without spatial constraints. The state space consists of the possible configurations of vegetation types on a given grid. A particular configuration is called a map, M_t denotes a map in year t. The first-order Markov property means that a map in year t, M_t, depends only on the map in the immediately previous year, M_{t-1}, and does not depend on years farther before. The probability that a particular entry $m_{i,j,t+1}$, the entry in position (i,j) in map M_{t+1}, has vegetation type g_k depends on $m_{i,j,t}$ only and its one-unit neighborhood, namely, its eight closest neighbors. This probability of $P(m_{i,j,t+1} = g_k)$, where k refers to one of the nine vegetation types, follows a distribution proportional to weights w_k that are calculated in the following way in the proposed model. If $m_{i,j,t}$ has neighbors belonging to more than four different vegetation types, then we set all w_k to 0. Otherwise, let the vegetation type of $m_{i,j,t}$ be v_k, and for the two most frequent vegetation types g_{k_1} and g_{k_2}, we set $w_{k_1} = v_{k,k_1}$ and $w_{k_2} = v_{k,k_2}$.We increase w_k by $v_{k,k}$, i.e., by the self-replacement probability (w_k might be either 0 or already $v_{k,k}$), and each w_l is also increased by $a_{k,l,}$ where l refers to the vegetation type that is considered as an invader of the cell. The probability of $P(m_{i,j,t+1} = g_k)$ is then

$$P\left(m_{i,j,t+1} = g_k\right) = \frac{w_k}{\sum_l w_l} \tag{1}$$

If it is clear from the text that we are talking about the current realization of a map entry, we will omit g_k, and we will denote the above probability with $P(m_{i,j,t})$. Each $m_{i,j}$ entry in a map develops independently, therefore the likelihood of the entire map is the product of the probabilities of individual entries.

As a contrast, in the pilot simulation we also applied another method for calculating the transition weight, which we refer to as the "full neighborhood enumeration approach". Here weights describing establishment from neighborhood were set to be directly proportional to the number of pixels occupied by each vegetation type in the eight-cell neighborhood.

The probability of a series of maps is the product of transition probabilities of all entries in all maps, which depends on the set of parameters $\theta = \{\mathbf{V}, \mathbf{A}\}$.

Bayesian MCMC

Theoretical background

We denote by

$$P(\mathbf{M}|M_0, \theta) \tag{2}$$

the probability of observing maps $\mathbf{M} = \{M_1, M_2, \ldots M_n\}$ in years $t = 1, 2, \ldots n$ given model parameters θ and that we observed map M_0 in year $t = 0$. We recall that this is the product of probabilities for each $m_{i,j,t}$ defined in Eq. 1, namely

$$P(\mathbf{M}|M_0, \theta) = \prod_{(i,j)\in I} \prod_{t=1}^{n} P\left(m_{i,j,t}\right) \tag{3}$$

where I denotes the index set of the map on which data is available. We do not have a priori information about the parameter set θ, therefore we set a flat prior distribution $P(\theta)$: each entry in both matrices $\mathbf{A}$ and $\mathbf{V}$ is uniformly distributed between 0 and 1.

Only some of the maps amongst $\mathbf{M} = \{M_1, M_2, \ldots M_n\}$ are observed, the remaining maps must be estimated. Let $\mathbf{M}^*$ denote the set of observed maps, and let $\mathbf{M}^+ = \mathbf{M} \backslash \mathbf{M}^*$ denote the missing (hidden) data. Here '\' denotes set subtraction.

Our aim is to sample from the posterior distribution of the parameters given by the Bayes theorem:

$$P(\theta|\mathbf{M}^*, M_{t_0}) = \frac{\sum_{\mathbf{M}^+} P(\mathbf{M}|M_{t_0}, \theta)P(\theta)}{\int_{\theta'} \sum_{\mathbf{M}^+} P(\mathbf{M}|M_{t_0}, \theta')P(\theta')d\theta'}, \tag{4}$$

where θ' refers to θ running over all possible parameter value. Since direct sampling from the posterior distribution is intractable, we need a classical data augmentation. The joint posterior distribution of hidden data and parameters are defined by the Bayes theorem:

$$P(\theta, \mathbf{M}^+|M_{t_0}, \mathbf{M}^*) = \frac{P(\mathbf{M}^* \cup \mathbf{M}^+|M_{t_0}, \theta)P(\theta)}{\int_{\theta'} P(\mathbf{M}^* \cup \mathbf{M}^+|M_{t_0}, \theta')P(\theta')d\theta'} \tag{5}$$

Note that $\mathbf{M}^* \cup \mathbf{M}^+ = \mathbf{M}$, and thus the map probabilities on the right-hand side of Eq. 5 is the same as given in Eq. 3. The likelihood calculation for the augmented data, $\mathbf{M}$ is easy, on the other hand, the partition function (the denominator in the Bayes theorem) is hard to calculate. Hence, the joint posterior of parameters and hidden data can be calculated only up to an unknown normalizing constant. However, this is enough in Bayesian MCMC, as described below.

The Markov chain

We applied the Metropolis–Hastings algorithm (Metropolis et al. 1953; Hastings 1970) with random window proposals (Green 1995; Lunter et al. 2005). This algorithm makes a random walk on the Descartes-product of hidden data, $\mathbf{M}^+$, and parameters, θ, defined by the following two steps:

(i) (proposal) In case of choosing parameter modification, a new parameter set, θ_{new} is chosen from a proposal distribution $T(\theta_{new}|\theta)$, as defined below. In case of modifying the hidden data, a random window, w, is cut out from the hidden data that contains the same rectangle from maps from consecutive years. Data is modified in this window as described below.

(ii) (acceptance) The new parameters or hidden data are accepted with probability

$$\min\left\{1, \frac{P(\theta_{new}, \mathbf{M}^+|M_{t_0}, \mathbf{M}^*)T(\theta|\theta_{new})}{P(\theta, \mathbf{M}^+|M_{t_0}, \mathbf{M}^*)T(\theta_{new}|\theta)}\right\} \quad (6)$$

in case of parameter changing step and with probability

$$\min\left\{1, \frac{P(\theta, \mathbf{M}^+_{new}|M_{t_0}, \mathbf{M}^*)P(\mathbf{M}^+|\mathbf{M}^+_{new}, w_{new})P(w_{new})}{P(\theta, \mathbf{M}^+|M_{t_0}, \mathbf{M}^*)P(\mathbf{M}^+_{new}|\mathbf{M}^+, w)P(w)}\right\} \quad (7)$$

in case of hidden data changing step. If the proposal is not accepted, the Markov chain remains in the current state.

Here w_{new} is a window in the modified hidden data, $\mathbf{M}^+_{new}$, whose content has to be modified to transform $\mathbf{M}^+_{new}$ back to $\mathbf{M}^+$. The globally stable stationary distribution of this random walk is $P(\theta, \mathbf{M}^+|M_{t_0}, \mathbf{M}^*)$ (Metropolis et al. 1953; Lunter et al. 2005).

The parameter changing step systematically tries to change all parameters of the model. For each parameter, the proposal distribution is the uniform distribution around the current parameter value with a fixed span, with a possible cut-off of forbidden values, namely,

$$U[\max\{x_{\min}, x - \Delta x\}, \min\{x_{\max}, x + \Delta x\}], \quad (8)$$

where $x_{\min}$ and $x_{\max}$ are the minimal and maximal values that the current parameter might take, x is the actual value, Δx is the span, and $U[a,b]$ denotes the uniform distribution on the interval $[a,b]$. In our program, Δx was set to 0.01. Thus, the ratio of proposal and backproposal probabilities in Eq. 6 when a parameter $x \in \theta$ is replaced to $x_{new} \in \theta_{new}$ is

$$\frac{T(\theta|\theta_{new})}{T(\theta_{new}|\theta)} = \frac{\min\{x_{\max}, x + \Delta x\} - \max\{x_{\min}, x - \Delta x\}}{\min\{x_{\max}, x_{new} + \Delta x\} - \max\{x_{\min}, x_{new} - \Delta x\}} \quad (9)$$

The hidden data changing step is parallel to the parameter changing. First a random window, w is selected. This window contains a rectangular part of the maps, the lengths of the two parallel edges of the rectangle are uniformly distributed between 1 and 40 (giving the number of 4 m^2 cells involved from the map). The window spans over a random number of years, the span is uniformly distributed between the possible $[t_k,t_{k+1}]$ intervals and the length of the span is also uniformly distributed between 1 and $t_{k+1} - t_k$. Since the probability of choosing a window does not depend on the current maps, and w_{new} is located in the same position as w, the probability of choosing w and w_{new} is the same. Therefore, $P(w_{new})/P(w)$ always cancels out in Eq. 7.

Hidden data is changed in the selected window in the following way. For each vegetation type and each matrix entry of the window, we calculate the a priori probabilities of the particular vegetation type occupying the actual matrix entry in the model. These probabilities are treated as weights, and we add further values to these weights, increasing weights of vegetation types that are expected to appear in the actual matrix entry based on the map from the previous year and the map from the next year, when observation is available. The proposed vegetation

type is drawn from a probability distribution that is proportional to the final weights. These probabilities appear in the proposal and backproposal probabilities, $P(\mathbf{M}^{+}_{new}|\mathbf{M}^{+}, w)$ and $P(\mathbf{M}^{+}|\mathbf{M}^{+}_{new}, w_{new})$ in Eq. 7.

We would like to mention that the proposal distribution changes only the speed of the convergence of the Markov chain towards the prescribed stationary distribution, but not the stationary distribution itself: any change in the proposal distribution causes a change in the acceptance probabilities in Eqs. 6 and 7 and the two changes keep the stationary distribution the same.

Simulation and output

We applied the above described method on field data at two resolutions (pilot and main simulations) and at two levels of incorporating observations (full and artificially incomplete). In each run, we performed 10,000 steps as burn-in, and then 300 samples were taken with 1000 random steps between two samplings. The convergence was checked based on the log-likelihood trace of the Markov chain. The model output is a distribution of map configurations for each year within the investigation period and two distributions of matrices containing likelihood of establishment from neighborhoods and without spatial constraints respectively.

Uncertainty of the model in estimating parameters and missing vegetation maps was assessed the following ways. In the case of the parameters, the standard deviations of the values emerging in the full run were calculated for each of the cells of the transition matrices. Regarding maps, we assessed the uncertainty on a year-by-year basis by calculating the Shannon Entropy (equivalent to Shannon diversity) of the estimated vegetation types per map cells.

Case study

Site description and vegetation maps

The three vegetation maps used in this study originate from a formerly grazed loess steppe at the foot of the Northern Medium Mountains in Hungary (47°54′ N, 20°35′ E, Fig. 1). The original vegetation of the area must have been a wooded steppe, but forests have been removed at least 400 years ago. Afterwards the

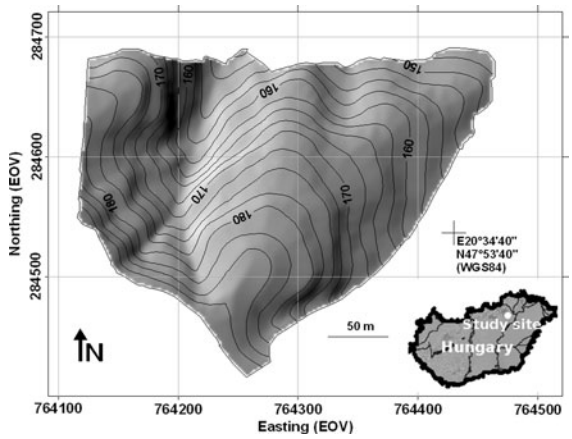

Fig. 1 Location ad topography of our study site in Hungary

area was subjected to different levels of grazing, which was abandoned in the late 1970s. Diverse geomorphology and grazing preferences resulted in a mosaic pattern of plant assemblages dominated or co-dominated by different grasses. The mosaic was mapped in 1983 (Virágh and Fekete 1984), 1988 and 2002 (Somodi et al. 2004) and these maps provide the data for our simulation. Base maps were rasterized with 2 m cell size. Vegetation entities were defined at a finer level than the association, therefore we refer to our categories as vegetation types. These were named after the dominant or one of the co-dominant species and were demonstrated to possess distinct species composition and abundance pattern (Somodi et al. 2004, Table 1 updated from Somodi et al. 2004). Vegetation types can be grouped according to naturalness. Some can be regarded as remnants of the grassland component of the wooded steppe vegetation: *Danthonia* type, *Sieglingia* type, partially *Festuca* type also. The latter has a more xerophyl character than the typical steppe. The vegetation type referred to as shrubs is composed of the original woody species of the wooded steppe. At the beginning of the study it was dominated by the shrub species and was shorter than a forest. Today its physiognomy is variable: in some places shrubs still dominate, other patches are more like groups of trees and start to acquire the structure of the wooded steppe forests. The rest of the vegetation types have in some way departed from the natural vegetation. *Leontodon* and *Bothriochloa* types were open and short grasslands adapted to grazing. In *Calamagrostis*, *Elymus* and *Chrysopogon* types one species has become

overwhelmingly dominant. *Calamagrostis epigejos* is almost monodominant in stands of the type named after the species, at the other end of the dominance gradient, *Chrysopogon* type is less overtaken by its dominant species and thus is the closest to natural conditions from the three types mentioned. Somodi et al. 2004 explored the changes and the regularities in those based on the three vegetation maps. Here we only give a very brief summary of the changes: there were very little differences between the map in 1983 and in 1988. The most apparent was the start of the expansion of *Calamagrostis* type. In contrast, there were considerable changes until 2002. *Leontodon* and *Bothriochloa* types almost completely disappeared, the area occupied by *Sieglingia* and *Festuca* types greatly decreased. At the same time, *Danthonia, Chrysopogon, Calamagrostis* and *Elymus* types and shrubs spread considerably.

For further details on site conditions, vegetation types and observed changes consult Virágh (1982) and Somodi et al. (2004). Species names follow Flora Europaea (Tutin et al. 1964–1993).

The pilot study

First, we carried out a pilot project to explore the effects of the two neighborhood weighting approaches on results. Since the full neighborhood enumeration approach applied to the maps in the original temporal and spatial resolution would have required an unrealistic amount of computing time, all three field vegetation maps were coarsened from 2 to 10 m cell size. Time steps investigated were set to be 5 years instead of one as in the main simulation. As the time interval between the second and the third vegetation map is 14 years, the last "5-year" interval was set to 4 years in practice, but we believe that this small change did not influence the outcome of the pilot study due to its already coarse temporal and spatial grain.

We proceeded in the pilot project exactly the same way as in the main simulation, except for using not only the neighborhood weighting approach proposed by us, but also the full neighborhood enumeration approach. Results were analyzed further for major trends in the same way as in the case of the main simulation exercise (see below) except for hot-spot detection which needs the resolution of the full run to be meaningful.

The reconstruction of vegetation dynamics

In the main simulation all three original vegetation maps available were used to reconstruct vegetation dynamics at our study site and to demonstrate the potential of the proposed method ("three-map simulation"). As raw outputs contain a great amount of complex information, we extracted synthetic descriptors to make the outputs tractable.

Firstly, we averaged entries in the transition matrices over the distribution supplied by the simulation. Secondly, average area occupied by each vegetation type was calculated for each year. Again averaging was needed to give a synthetic picture of the distribution of maps provided by the method. Thirdly, synthetic maps were drawn to visualize the distributions available for years. Pixels in the synthetic maps are painted proportionally to the four most likely vegetation types for that pixel in the distribution. This coloring provides a rough overview of proportions of vegetation types predicted for different parts of the maps.

Establishment areas were visualized for *Chrysopogon* type as an example. In this visualization the proportion of cells predicted to contain the type is shown for an excerpt of the investigation period. Hot spots of changes in general were identified based on the dissimilarities of the proportion vectors of vegetation types in one cell of the estimated maps. The Manhattan metric was used as dissimilarity measure, which is the sum of the absolute differences in the vectors. Dissimilarities of proportion vectors in consecutive years only were taken into account (instead of a full dissimilarity matrix), because these are the values that refer to subsequent changes in the maps. For illustration, we created maps of the intensity of changes (i.e., hotspots) by cellwise averaging the dissimilarity values over the first five and the last 14 years separately.

Validation

Another simulation using an artificially incomplete dataset, namely only the first and last observed map, was also conducted in order to assess the reconstructive capacity of the proposed technique. We further refer to this simulation as the "two-map simulation". A comparison of this simulation with the one based on all three maps provides a test of how well the model

predicts lacking data. In order to assist comparison, average areas occupied by different vegetation types were calculated and synthetic maps were produced for this run in the same manner as for the run based on the complete dataset. Uncertainties found for the full and this second run were compared by paired t-tests. There were separate tests carried out for estimated maps in each year (as many comparisons as years) and for transition matrices (1 comparison). In case of the estimated maps, the two uncertainty values in the same cell in the two different runs were treated as a pair. Calculations were carried out on a cell-by-cell basis in each investigation year. In case of transition matrices the matrix entries were paired given the two simulations with different input data. Uncertainties were visualized in a map for the year 1988, when the comparison with the observed map ensured deeper insight.

Finally, direct evaluation of model estimations from this run was possible for 1988 using the observed pattern as a basis of validation. This was carried out by taking the frequency of individual types in one pixel as an estimate of probability of that type occurring there and used the area under the Receiver Operating Characteristic (ROC) curve (AUC; Swets 1988) to assess how well these probabilities matched the observed data.

Results

A powerful method

The two simulations with different neighborhood weighting approaches yielded largely the same results (Figs. 2, 3) in the pilot study. As the coarsened version of all three maps were used in these simulations with 5 year intervals, two intermediate years arose, for which maps were estimated in the pilot project: 1993 and 1998. There were no tendencies of systematically over- or underestimating the average area occupied by any vegetation types in these maps (Fig. 3). This provides grounds to use the computationally simpler approach in the more complex simulations. Even though the simulated dynamics are coarse due to the large pixel size and temporal gaps, results can be well related to that of the final simulations. For example, the striking increase of the average area occupied by *Danthonia* type, followed by a decrease mirrors the bell-shaped dynamics found in the full simulation (Figs. 2, 4).

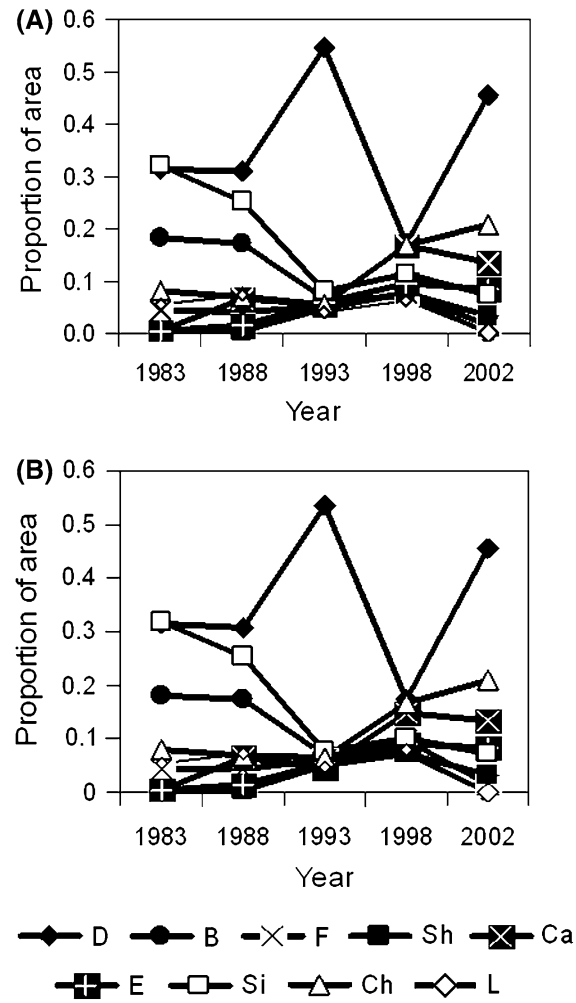

Fig. 2 Simulated area dynamics of the nine vegetation types in the study area in the pilot project. The underlying simulation was based on data from 1983, 1988 and 2002. **a** Neighborhood weighting by the full neighborhood enumeration approach, **b** neighborhood weighting by the method proposed in the article. In the pilot project there is an estimation available for each fifth year only, therefore lines between symbols have only a function to lead the eye from one symbol to the other and are not supposed to imply any linear change. *Letters* refer to vegetation types, for more information see Table 1

The complete simulation produced diverse and realistic expansion or shrink dynamics for the different vegetation types while keeping the observed vegetation maps constant (Fig. 4). It allowed a considerable diversity of area gain or loss patterns: saturation, continuous area loss or gain, bell-shaped and U-shaped curves. The full ("three-map") and the

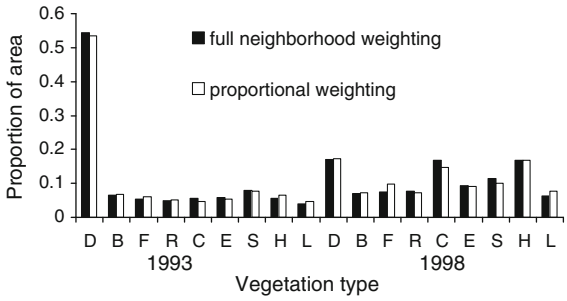

Fig. 3 Differences between the relative areas occupied by the various vegetation types in the 2 years for which the pilot simulation provided results according to the two neighborhood weighting approach

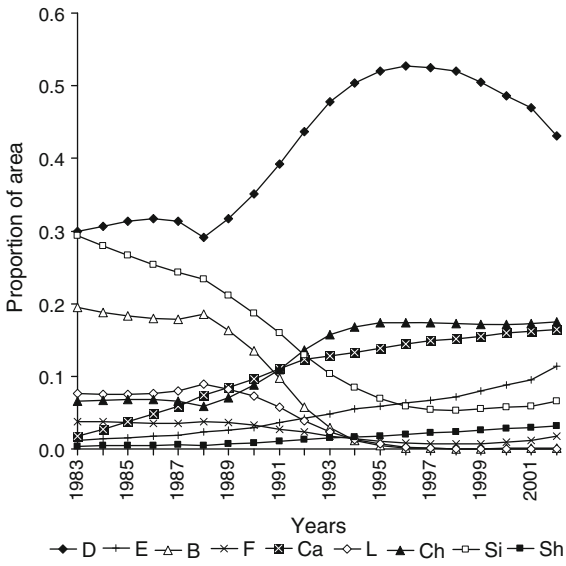

Fig. 4 Simulated area dynamics of the nine vegetation types in the study area. The underlying simulation was based on data from 1983, 1988 and 2002. *Letters* refer to vegetation types, for more information see Table 1

artificially incomplete ("two-map") simulations interpolated basically the same dynamics for the various vegetation types (Fig. 5) and the predicted maps in the two modeling scenarios were surprisingly similar (Appendix 1 in Electronic Supplementary Material). There were cases where even the shape of area change curve was precisely captured by the model based on two observations only (Fig. 5a). Nevertheless, even in other cases it was only the timing of the change which differed and not the shape of the curve (Fig. 5b, c). There were two types, however, which did not loose area or even increased their representation between 1983 and 1988 and only declined afterwards in reality, in contrast to estimations of the two-map simulation, which estimated a sharp decline right from the beginning (Fig. 5b).

The comparison of the observed map of 1988 and the one estimated in the artificially incomplete ("two-map") run also supports the reliability of the method. All the AUC values show good (0.7–0.9) to excellent (above 0.9) agreement using the scale of Swets (1988; Table 2). AUC were relatively lower for those two types, *Leontodon* and *Bothriochloa* types, which have almost disappeared by 2002.

Uncertainty levels in maps were low compared to the theoretical maximum (2.19 for 9 categories). The highest value for any year in the three-map simulation was 1.25 and 1.28 in the two-map simulation. Uncertainty in map estimates of the simulation based on two maps only was larger than that of the other run (Fig. 6). The difference between the average uncertainty in the estimated maps was significant for each pairs of years ($p < 0.001$ in each comparison).

As a contrast, the uncertainty was significantly lower ($p < 0.001$) in the two-map simulation when neighborhood dependent transitions were examined, while there was no significant difference ($p = 0.$ 0.699) in the uncertainty of the neighborhood-independent transitions (Appendix 2 in Electronic Supplementary Material).

The reconstructed maps add to the dynamics expressed by area changes by locating the areas where these occurred. Blurred areas between observations show simultaneous establishment followed by lateral extension, while it is also apparent, where types only expanded into their direct neighborhoods (Appendix 1 in Electronic Supplementary Material). Maps showing proportions of estimated presence in the distribution of maps for individual years help to identify these areas (Fig. 7). Hot spots of changes in general are also readily identifiable for any chosen periods (Fig. 8). The maximum average dissimilarity of subsequent years was higher in the first five years studied than in the next 14.

The case study

Establishment probabilities from neighborhoods were considerably larger than those without spatial constraints for each vegetation types (Appendix 2 in Electronic Supplementary Material, Table 3). Nevertheless, there were considerable differences between

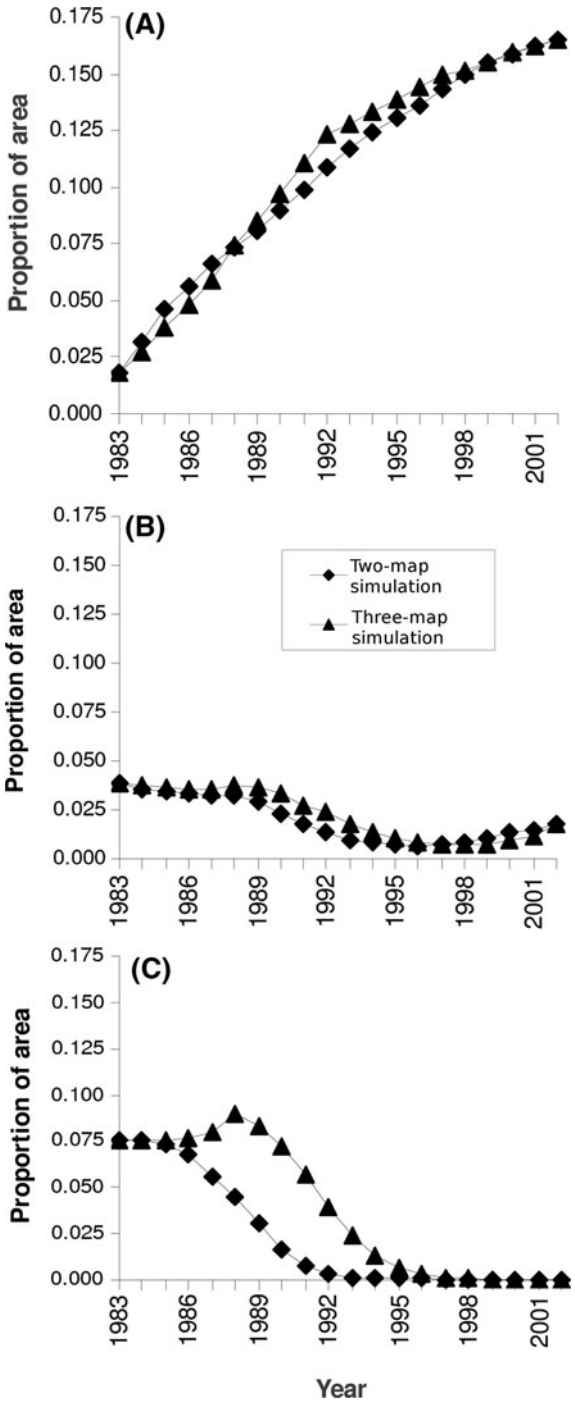

Fig. 5 Three examples of area dynamics interpolated for vegetation types with the two modeling scenarios. **a** *Calamagrostis* type, **b** *Festuca* type, **c** *Leontodon* type

vegetation types regarding the relative importance of establishment without spatial constraints in their dynamics. Types that lost area during the

Table 2 Agreement between the observed and predicted map in 1988

Veg. type	AUC
D	0.9610
E	0.9460
B	0.8814
F	0.9872
Ca	0.9699
L	0.8720
Ch	0.9724
Si	0.9459
Sh	0.9998

Agreement is measured by the Area Under the ROC Curve (AUC). For details of vegetation types see Table 1

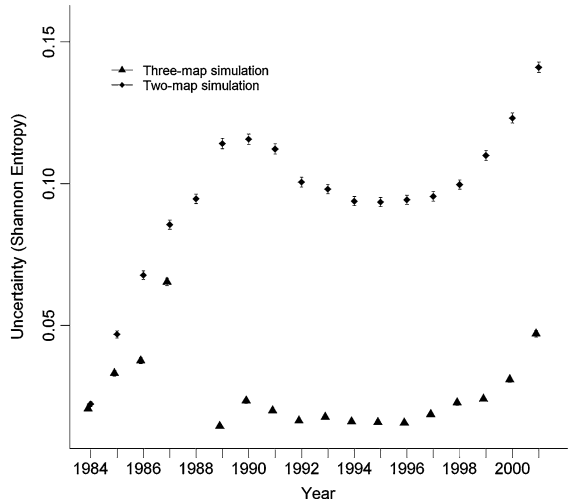

Fig. 6 Uncertainty of the estimated maps. Uncertainty is measured by Shannon Entropy, *error bars* represent standard error of the mean. *Three map simulation*-simulation using all the three observed maps available, *Two-map simulation*-simulation using only the first and last map as fixed data. There is no estimation for 1988 in the three-map simulation, because this map was fixed there

investigation period (*Bothriochloa*, *Leontodon* and *Sieglingia* types) cannot be characterized by this ratio, since the means of their loss of area was rather determined by the type that took them over. Nevertheless, for types that have become more widespread during the investigation period increased weights away from the diagonal carry valuable information. Establishment without spatial constraints (including that from common species pool) was most important for *Danthonia* and *Elymus* types, while played only a

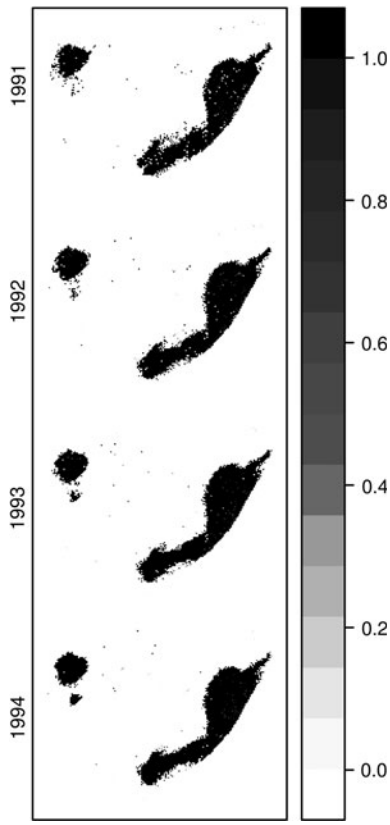

Fig. 7 An excerpt of the changes in the proportion of presence per cells estimated by the full simulation for *Chrysopogon* type. The proportion of presence is calculated from the 300 maps that resulted from the simulation for each year

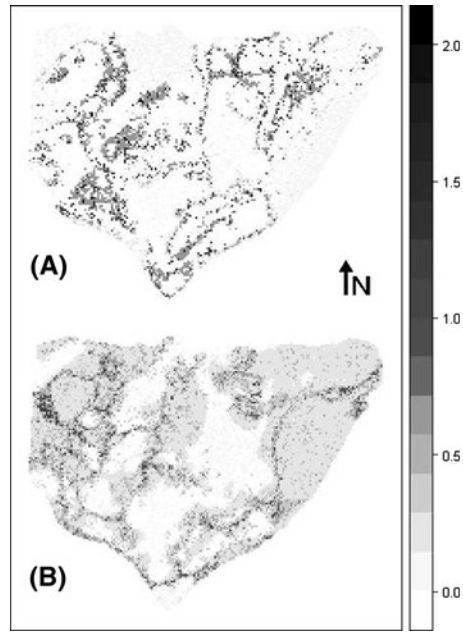

Fig. 8 Hotspots of changes in the two investigation periods. **a** 1983–1988, **b** 1988–2002. Hotspots typically signify sites of expanding vegetation types. A hotspot in the middle of the left-hand side of "map A" is due to the expansion of *Calamagrostis* type, for example. Intensity of changes was measured by the dissimilarity of the proportion vectors of estimated vegetation types in each pixel between subsequent years. Dissimilarity was measured by the Manhattan metrics. Values were averaged over the period in question

minor role in the case of *Calamagrostis* type. In the case of shrubs, establishment from neighborhood clearly outweighed establishment without spatial constraints. On the other hand, the typical transition towards *Chrysopogon* type (from *Bothriochloa* type) was characterized by the lowest neighborhood-dependent/independent ratio is in the study (Table 3).

The model interpolated continuous increase in the proportion of area occupied by *Calamagrostis* and *Elymus* types and shrubs (Figs. 4, 5). In the case of *Chrysopogon* type a rapid expansion was estimated after 1988 followed by a period of stable area occupation. For *Leontodon* and *Bothriochloa* types, which disappeared meanwhile, a continuous decline was interpolated, but a considerable lag was also apparent before the loss. Both models (using 3 vs. 2 vegetation maps) showed that there possibly was a maximum of *Danthonia* type occupation already back in time and the present distribution must be a snapshot of slow decline. At the same time the most restricted distribution is possibly behind for *Festuca* type and what we observe as a restricted distribution compared to 1988 is part of an already growing pattern. A similar pattern was discernable, however less obvious, in the case of *Sieglingia* type.

According to the series of maps, establishment of vegetation types can often not be exactly located regarding cells in the maps, but rather assigned a certain area (Appendix 1 in Electronic Supplementary Material). Maps assist in visualizing the scale and pattern in the modeled dynamics. An example is shown for *Chrysopogon* type in the period of intensive establishment (1991–1994; Fig. 7). General hotspots of change can be well coupled with the major establishment sites of expanding vegetation types. For example, in the first 5 years a major hotspot appeared where patches of *Leontodon* type merged into one large patch (Northeast on the map, Fig. 8a, Appendix 1 in Electronic Supplementary

Table 3 Ratio of average weights of individual transitions between vegetation types by the two establishment mechanisms

	D	E	B	F	Ca	L	Ch	Si	Sh
D	30.58	76.61	*369.59*	**319.73**	94.70	*193.91*	**175.19**	*272.49*	**873.52**
E	24.70	15.60	*80.08*	**101.83**	76.96	*1921.20*	39.65	*281.93*	**406.14**
B	12.64	23.53	*12.70*	23.98	**113.43**	*651.08*	13.32	*8.90*	90.21
F	23.42	17.83	*81.43*	9.41	48.36	*437.99*	**166.01**	*51.20*	**107.26**
Ca	20.02	67.28	*168.79*	72.51	23.82	*747.44*	**280.39**	*93.97*	**2371.68**
L	5.55	**248.61**	*492.79*	**610.79**	32.19	*7.29*	46.17	*18.04*	**214.95**
Ch	45.09	56.76	*45.92*	**372.13**	**213.32**	*448.57*	20.87	*139.76*	**247.08**
Si	27.81	61.82	*67.01*	**193.23**	37.20	*85.54*	93.37	*16.79*	**597.26**
Sh	23.47	48.13	*65.59*	45.26	**488.72**	*195.69*	67.83	*145.43*	14.05
Mean	23.70	68.46	*153.77*	**194.32**	**125.41**	*520.97*	**100.31**	*114.28*	**546.91**

Weights averaged over the distribution resulting from the full simulation run can be found in Appendix 2. Values in this table express how many times the neighborhood-dependent weights are larger than the weights of establishment without spatial constraints. The ratio is less informative for types that have lost area, these are indicated by italics. High values for other vegetation types are marked above the arbitrary limit of 100 by bold. Columnwise means are also shown, which give a rough assessment of the relative importance of establishment from neighborhood for the vegetation type of the column. Letters refer to vegetation types, for further information consult Table 1

Material). In the same period, another hotspot in the middle of the left-hand side of the map is due to the expansion of *Calamagrostis* type. Reconstructed maps make it obvious that *Calamagrostis* type conquers the direct neighborhood of its patches. This is also reflected in the hotspot representation, by the patches, that have been already occupied in 1988 undergoing virtually no change in the second period (Fig. 8b). On the other hand, shrubs, *Chrysopogon* and *Elymus* types are rather characterized by random establishments in appropriate hotspots followed by rapid lateral expansion from these cores. This also means that their most probable establishment sites cannot be given exactly, only the broader area, where this phenomenon occurred (Appendix 1 in Electronic Supplementary Material). These broader areas are the hotspots of the second period, the northeastern part of the area, as well as the northwestern side (Fig. 8).

As for the lessons of the comparison of the estimated and observed map regarding the study site, uncertainties in the two-map simulation accumulated mostly in the eastern part in 1988, where the change from *Bothriochloa* to *Chrysopogon* type as well as the intermittent expansions of *Leontodon* type posed special challenge to the simulation (Fig. 9). The higher level uncertainty can be regarded as a warning of the simulated maps being potentially less accurate. In this case the timing of these changes was not perfectly captured. Therefore some of the changes

Fig. 9 Uncertainty of map estimates for the year 1988 as measured by the Shannon Entropy of the proportion vector of the estimated vegetation types per map cells. The estimates come from the "two-map" simulation, in which only the first and last observed vegetation map was used. High uncertainty values appear at the edge of some of the major vegetation patches and in the Northeastern corner. In the latter area changes were intense in vegetation types, where the lack of the intermittent vegetation map caused a delay in the transition in the simulation compared to observations

manifest themselves already in some of the maps of the simulated distribution of 1988, while in some others there is (correctly) no change yet, thus the Shannon Entropy of the estimates increase. The same effect is shown in the accuracy assessment: the area occupied of *Bothriochloa* and *Leontodon* types were estimated much less precisely than that of others.

Discussion

Evaluation of the method

The MCMC simulation as a whole showed good results in general, making it an important candidate for use among methods aiming at the reconstruction of vegetation dynamics. MCMC approaches have already proven to be useful in estimating parameters in other aspects of landscape dynamics (Clark 2005; Jenerette and Potere 2010), our method complements on these. Results of the pilot study confirmed that the proposed neighborhood weighting method can be applied instead of full neighborhood weighting for speeding up the simulation without the danger of changing the outcome, since it provided essentially similar results to an attempt using direct proportionality of vegetation types in the neighboring cells. This way, computing time becomes reasonable and qualitative results do not change. Furthermore, the difference is likely to preserve the ecologically relevant autocorrelation structure (Weaver and Perera 2004).

As for the full simulation method introduced in this paper, there appeared numerous advantages of its use. Firstly, its data requirements are modest: any number of vegetation maps with arbitrary intervals between them can be used. This is a distinct advantage over traditional Markov Chains, where either only two maps could have been used as bases (Horváth and Csontos 1992; Callaway and Davis 1993; Kadmon and Harari-Kremer 1999) or maps had to have equal spacing, or spacing of a lowest common multiple integer among them (Gibson et al. 1983; Lough et al. 1987; Scanlan and Archer 1991). If the nature of the transition allows, this opportunity can be exploited for other landscape dynamics studies as well (in similar situations to Jenerette and Wu 2001, for example).

Another advantage of the proposed method is, that it is spatially explicit, without requiring specification of neighborhood interaction rules in advance. It is an advantage both over Markov chains, which are non-spatial models and spatio-temporal Markov chains, where neighborhood rules have to be estimated before the modeling starts (Baltzer et al. 1998). Therefore, our method allows the use of datasets, where expert knowledge is not capable to give rules. Furthermore, mechanistic establishment rules are provided as an output, which allows both to explore mechanisms of vegetation dynamics and thus neighborhood interactions as well as to test hypotheses about these, rather than requiring a priori estimations. Although, efforts to estimate neighborhood rules were made earlier as well, our model has a unique advantage not shared by those (Dieckmann et al. 1997 for example), namely that transition rules are estimated globally, not independently from each other.

Thirdly, the model incorporates no restrictions on the emergent dynamics proved by the diverse area change curves found in our case study. It is indeed a major advantage of the modeling framework, that does not require preliminary knowledge of the dynamic behavior of the modeled entities. There are several indications that vegetation dynamics have non-linear characteristics (Leps 1987; Fuhlendorf and Smeins 1998; Gassman et al. 2000; Somodi et al. 2004). Nevertheless, if only a few maps are at hand we have no baseline of what non-linearities to expect between the observations. Our model reconstructs these by providing a whole distribution of likely courses of dynamics, which is possible to be synthesized into mean expected area curves, hotspots of changes, dispersion measures or any other statistics describing the dynamics.

Finally, there were positive indications that the interpolated dynamics are reliable. The artificial simulation yielded similar results to the full simulation, even if with higher uncertainty in the map estimates, but with good accuracy in the year, where validation was possible. Thus, the distribution of vegetation maps, or rather possible vegetation for each pixel in each year, provides useful estimates of intermittent stages. These can be used as past templates if analyses would require data from specific years, where observations are not available.

Major limitations of the model are due to two of its assumptions: homogeneity regarding time and space. Our model is time-homogeneous because it searches for two establishment matrices that fit best the observations and are constant during the modeled period. Markov chain based studies often argue that homogeneity assumptions are unrealistic (Usher 1981; Scanlan and Archer 1991; Baltzer 2000) and efforts have been made to overcome these limitations (Li 1995; Childress et al. 1998). Nevertheless, our study differs from Markov chain approaches by having two

transition matrices addressing different aspects of establishment mechanisms. These are most likely to be inherent to assemblages, although weather conditions can still influence these. Most models developed on the conceptual basis of Markov chains are spatially homogeneous. Transition rules can be influenced, however, by environmental heterogeneity in space as well (Kadmon and Harari-Kremer 1999; Augustin et al. 2001). Although the present version of our model does not treat these dependencies, the conceptual design allows the incorporation of establishment matrices conditional on both temporally and spatially varying environmental factors.

Discussion of the results of the case study

Our hypothesis, that vegetation types with different composition will show characteristic emergent dynamic properties has been confirmed by our results. There were diverse non-linear dynamics observed, similar to those found in other simulation studies on species' population dynamics (Silverton et al. 1992; Tilman 1994). We have found saturation curves just like those for superior competitor species with mortality rates equal to the inferiors. Such curves might indicate that the dynamics of the vegetation type is determined by the dominant species, which coincides with field experience and literature in post-abandonment situations especially for *Calamagrostis* type (Rebele and Lehmann 2001). The pattern of decline of types previously favored by grazing (*Leontodon* and *Bothriochloa* types) is similar to that of inferior competitors in two-species—simulations. Nevertheless, there are non-linear patterns different from these basic ones, which give valuable information on potential expectations as well. Bell-shaped and U shaped parts alternate for *Festuca* and *Danthonia* type.

Furthermore, most of these simulated dynamics carry very different messages than the traditional Markov chain approach (Somodi et al. 2004). Although, there are a few types for which the two approaches coincide, differences are more typical. The most striking distinction, also with a conservation implication, is that for the natural and low-competitive *Festuca* and *Sieglingia* type, Markov chains predict a continuous decline, while the simulations foreshadow an increase.

As we expected, the importance of lateral expansion was overwhelming above expansion without spatial constraints. This partly follows from our study system. Grasslands are usually rich in clonal species, therefore local establishment by lateral spread is a predominant way of the expansion of species (Fenner and Thompson 2005) and therefore that of their assemblages. This has been clearly mirrored in the weights in the transition matrices showed that such mechanisms prevail at our site as well. The reconstructed series of maps further revealed that even accidental establishment is typically followed by lateral expansion from the new cores, which also manifests in the broad establishment areas on the hotspot maps. This draws our attention to the importance of allowing both mechanisms to operate for each unit in the model (similarly to Barkham and Hance 1982), rather than assigning clear strategies for types as were done in cellular automata models of population dynamics (e.g., Crawley and May 1987) even if one or another strategy dominates the behavior of one type (Somodi et al. 2010).

Nevertheless, we did expect differences in the lateral vs. non-lateral expansion potential of different vegetation types as well, which has been confirmed by the results. Shrubs have been identified as expanding mostly into their own neighborhoods, which is in good accordance with previous findings (Wild and Winkler 2008). As we also hypothesized, *Calamagrostis* type was mostly expanding laterally, rather than establishing without being present in the neighborhood earlier. This is in good accordance with its dominant species' ecology in Central Europe (Rebele and Lehmann 2001) and with previous findings at our study site (Somodi et al. 2010). *Elymus* type on the other hand, was among the least dependent on establishment from the neighborhood, though we expected considerable influence of neighboring cells due to the dominant species' characteristics (Palmer and Sagar 1963; Fitter and Peat 1994). Appearance of a type by abundance restructuring can play a role here, as opposed to the process in the case of *Calamagrostis* type, and can thus increase the spatially independent appearance of patches of *Elymus* type. Namely, its dominant and subordinate species are present in lower abundance all over the study site, therefore the vegetation can react to a

change in the environmental conditions quickly by the restructuring of the abundance pattern so that *Elymus* type manifests itself. *Chrysopogon* type was also expected to expand relatively much without spatial constraints. Here the most frequent transition (*Bothriochloa* → *Chrysopogon* type) was indeed pronouncedly less influenced by neighborhoods than transitions in general in the simulation. In accordance with expectations again, establishment was the least dependent on neighborhood relationships for the species rich *Danthonia* type among all the types studied.

Examining the interpolated maps helps us to further refine our view of appearance patterns of vegetation types. We already know that different drivers influence the expansion of different types (Somodi et al. 2010). Nevertheless, besides ratios of transition weights, interpolated maps also imply that even types for which proximity of previous patches plays a minor role in determining future patches tend to expand laterally from cores. It is true for *Chrysopogon* type, for example, which has been shown earlier to be primarily determined by geomorphology and not by proximity relationships (Somodi et al. 2010). Therefore, inclusion of spatial dependence seems to be vital in vegetation dynamics modeling, just as Weaver and Perera (2004) pointed out.

As a result of the establishment pattern detailed above, interpolated maps help to identify the establishment area of individual types, rather than giving exact locations. If discrete maps were needed, field rules of mapping can transform these areas into patches and finally create vector vegetation maps. The establishment areas, if multiple vegetation types are regarded together, are also the hotspots of changes. Identification of the location of establishment hotspots of species and thus hotspots of vegetation change has received increased interest very recently (Franklin 2010; Van Uytvanck et al. 2010). Our hotspot maps also showed that although most of the changes occurred after a lag period at our site (Somodi et al. 2004), some of the quickest transformations between vegetation types indeed occurred in the first 5 years.

Finally, our evaluation of the simulations for this case study is far from exhaustive. Many other features of the dynamics can be extracted from the full distributions of transition weights or type distributions within pixels, which can assist vegetation science and nature conservation in various ways. The modes of extractions will certainly be determined by the scientific or practical question.

Conclusion

To our best knowledge, the proposed method is the first work that tries to estimate vegetation transition parameters in a stochastic cellular automaton based on field measurements and so provides a tool to reconstruct past dynamics. Thus it may become a useful tool for vegetation ecology. Furthermore, we believe that the approach can be adapted to landscape dynamics studies, focused on other land cover classes than vegetation as well.

The method can give an estimate of the timing of changes, and also provides a mechanism-based extrapolation possibility. Its moderate data requirements, various outputs and the low-level and clear built-in constraints make it an ideal candidate for the reconstruction of vegetation and landscape dynamics. Its major advantage is that while being spatially explicit it requires no a priori considerations of neighborhood rules and allows the use of observations from different time intervals. Outputs include whole distributions of interpolated maps for each year and those of the two mechanistic transition matrices. Limitations of the model are due to its spatio-temporal homogeneity.

Specific results for the study site pointed out the overwhelming dominance of expansion dynamics into the neighborhoods of existing patches regarding all vegetation types. Estimated maps revealed previous maximum and minimum points of expansion dynamics of vegetation types, which could not have been documented. As a consequence, our expectations about the future dynamics of these types changed thoroughly, which also has an outreach towards nature conservation.

Acknowledgments Klára Virágh was supported by the Hungarian Research Fund (OTKA) grant no.SUP062338, István Miklós by that of no. F61730. Imelda Somodi was supported by both listed grants, as well as by the Hungarian Research Fund (OTKA) grant no. NI68218. The authors are grateful to Beáta Oborny for her support during model development.

References

Augustin NH, Cummins RP, French DD (2001) Exploring spatial vegetation dynamics using logistic regression and a multinomial logit model. J Appl Ecol 38:991–1006

Austin MP, Smith TM (1989) A new model for the continuum concept. Plant Ecol 83:35–47

Baker WL (1989) Landscape ecology and nature reserve design in the boundary waters canoe area, Minnesota. Ecology 70:23–35

Baltzer H (2000) Markov chain models for vegetation dynamics. Ecol Model 126:139–154

Baltzer H, Braun PW, Köhler W (1998) Cellular automata models for vegetation dynamics. Ecol Model 107:113–125

Banfai DS, Bowman DMJS (2006) Forty years of lowland monsoon rainforest expansion in Kakadu National Park, Northern Australia. Biol Conserv 131:553–565

Barkham JP, Hance CE (1982) Population dynamics of the wild daffodil (*Narcissus pseudonarcissus*) III. Implications of a computer model of 1000 years of population change. J Ecol 70:323–344

Blackstock TH, Stevens JP, Howe EA, Stevens DP (1995) Changes in the extent and fragmentation of heathland and other semi-natural habitats between 1920–1922 and 1987–1988 in the Llyn Penninsula, Wales, UK. Biol Conserv 72:33–44

Brook BW, Bowman DMJS (2006) Postcards from the past: charting the landscape-scale conversion of tropical Australian savanna to closed forest during the 20th century. Landscape Ecol 21:1253–1266

Callaway RM, Davis FW (1993) Vegetation dynamics, fire, and the physical environment in coastal central California. Ecology 74:1567–1578

Childress MW, Crisafulli CM, Rykiel EJ Jr (1998) Comparison of Markovian matrix models of a primary successional plant community. Ecol Model 107:93–102

Clark JS (2005) Why environmental scientists are becoming Bayesians. Ecol Lett 8:2–14

Cochrane MA (2000) Using vegetation reflectance variability for species level classification of hyperspectral data. Int J Remote Sens 21:2075–2087

Colasanti RL, Hunt R, Watrud L (2007) A simple cellular automaton model for high-level vegetation dynamics. Ecol Model 203:363–374

Crawley MJ, May RM (1987) Population dynamics and plant community structure: competition between annuals and perennials. J Theor Biol 125:475–489

Dieckmann U, Herben T, Law R (1997) Spatio-temporal processes in plant communities. In: Lepenies W (ed) Yearbook 1995/1996. Institute for Advanced Study Berlin, Nicolaische Verlagsbuchhandlung, Berlin, pp 296–326

Fenner M, Thompson K (2005) The ecology of seeds. Cambridge University Press, Cambridge

Fitter AH, Peat HJ (1994) The ecological flora database. J Ecol 82:415–425

Flamenco-Sandoval A, Ramos MM, Masera OR (2007) Assessing implications of land-use and land-cover change dynamics for conservation of a highly diverse tropical rain forest. Biol Conserv 138:131–145

Franklin J (2010) Vegetation dynamics and exotic plant invasion following high severity crown fire in a southern California conifer forest. Plant Ecol 207:281–295

Fuhlendorf SD, Smeins FE (1998) The influence of soil depth on plant species response to grazing within a semi-arid savanna. Plant Ecol 138:89–96

Gassman F, Klötzli F, Gian-Reto W (2000) Simulation of observed types of dynamics of plants and plant communities. J Veg Sci 11:397–408

Gaudet CL, Keddy PA (1995) Competitive performance and species distribution in shortline plant communities: a comparative approach. Ecology 76:280–291

Gibon A, Sheeren D, Monteil C, Ladet S, Balent G (2010) Modelling and simulating change in reforesting mountain landscapes using a social-ecological framework. Landscape Ecol 25:267–285

Gibson CW, Guilford TC, Hambler C, Sterling PH (1983) Transition matrix models and succession after release from grazing on Aldabra atoll. Vegetatio 52:151–159

Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82:711–732

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109

Horváth F, Csontos P (1992) Thirty-year-changes in some forest communities of Visegrádi Mts., Hungary. In: Teller A, Mathy P, Jeffers JNR (eds) Responses of forest ecosystems to environmental changes. Elsevier, Barking, pp 481–488

Jenerette GD, Potere D (2010) Global analysis and simulation of land-use change associated with urbanization. Landscape Ecol 25:657–670

Jenerette GD, Wu J (2001) Analysis and simulation of land-use change in the central Arizona–Phoenix region, USA. Landscape Ecol 16:611–626

Kadmon R, Harari-Kremer R (1999) Landscape-scale regeneration dynamics of disturbed Mediterranean maquis. J Veg Sci 10:393–402

Karau EC, Keane RE (2007) Determining landscape extent for succession and disturbance simulation modeling. Landscape Ecol 22:993–1006

Leps J (1987) Vegetation dynamics in early old-flied succession: a quantitative approach. Vegetatio 72:95–102

Li B-L (1995) Stability analysis of a non-homogeneous Markovian landscape model. Ecol Model 82:247–256

Lough TJ, Wilson JB, Mark AF, Evans AC (1987) Succession in a New Zealand alpine cushion community: a Markovian model. Vegetatio 71:129–138

Lunter GA, Miklós I, Drummond A, Jensen JL, Hein J (2005) Bayesian coestimation of phylogeny and sequence alignment. BMC Bioinform 6:83

Metropolis N, Rosenbluth AW, Teller MN, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21:1087–1092

O'Neill RV (1986) A hierarchical concept of ecosystems. Princeton University Press, Princeton

Palmer JH, Sagar GR (1963) Biological flora of the British Isles: *Agropyron repens* (L.) *Beauv*. (*Triticum repens* L.; *Elytrigia repens* (L.) *Nevski*). J Ecol 51:783–794

Pu RL (2009) Broadleaf species recognition with in situ hyperspectral data. Int J Remote Sens 30:2759–2779

Rebele F, Lehmann C (2001) Biological flora of Central Europe: *Calamagrostis epigejos* (L.) *Roth*. Flora 196:325–344

Rickebusch S, Gellrich M, Lischke H, Guisan A, Zimmermann NE (2007) Combining probabilistic land-use change and forest dynamics modelling to simulate global change responses at the Alpine tree-line. Ecol Model 209:157–168

Scanlan JC, Archer S (1991) Simulated dynamics of succession in a north American subtropical Prosopis savanna. J Veg Sci 2:625–634

Schmidt KS, Skidmore AK (2003) Spectral discrimination of vegetation types in a coastal wetland. Remote Sens Env 85:92–108

Schumacher S, Bugman H, Mladenoff DJ (2004) Improving the formulation of tree growth and succession in a spatially explicit landscape model. Ecol Model 180:175–194

Silverton 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

Somodi I, Virágh K, Aszalós R (2004) The effect of the abandonment of grazing on the mosaic of vegetation patches in a temperate grassland area in Hungary. Ecol Complex 1:177–189

Somodi I, Virágh K, Podani J (2008) The effect of the expansion of a clonal grass, *Calamagrostis epigejos* on the species turnover of a semiarid grassland. Appl Veg Sci 11:187–194

Somodi I, Virágh K, Székely B, Zimmermann NE (2010) Changes in predictor influence with time and with vegetation type identity in a post-abandonment situation. Basic Appl Ecol 11:225–233

Swets JA (1988) Measuring the accuracy of diagnostic systems. Science 240:1285–1293

Tilman D (1994) Competition and biodiversity in spatially structured habitats. Ecology 75:2–16

Turner MG (1988) A spatial simulation model of land use changes in a piedmont county in Georgia. Appl Math Comp 27:39–51

Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA (eds) (1964–1993) Flora Europaea, vols 1–5. Cambridge University Press, Cambridge

Usher MB (1981) Modelling ecological succession, with particular reference to Markovian models. Vegetatio 46:11–18

Van Uytvanck J, Van Noyen AV, Milotic T, Decleer K, Hoffman M (2010) Woodland regeneration on grazed former arable land: a question of tolerance, defence or protection? J Nat Cons 18:206–214

Virágh K (1982) Vegetation dynamics induced by some herbicides in a perennial grassland community I. Acta Bot Hung 28:427–447

Virágh K, Fekete G (1984) Degradation stages in a xeroseries: composition, similarity, grouping, coordination. Acta Bot Hung 30:427–459

Weaver K, Perera AH (2004) Modelling land cover transitions: a solution to the problem of spatial dependence in data. Landscape Ecol 19:273–289

Whited DC, Lorang MS, Harner MJ, Hauer R, Kimball JS, Stanford JA (2007) Climate, hydrologic disturbance and succession: drivers of floodplain pattern. Ecology 88: 940–953

Wild J, Winkler E (2008) Krummholz and grassland coexistence above the forest-line in the Krkonoše Mountains: grid-based model of shrub dynamics. Ecol Model 213:293–307