



Research Article

## Scaling simulation models for spatially heterogeneous ecosystems with diffusive transportation

Qiong Gao<sup>1,2\*</sup>, Mei Yu<sup>2</sup>, Xiusheng Yang<sup>3</sup> & Jianguo Wu<sup>4</sup>

<sup>1</sup>MOE Key Lab of Environmental Change and Natural Disaster, Institute of Resources Science, Beijing Normal University, Beijing, China; <sup>2</sup>Laboratory of Quantitative Vegetation Ecology, The Chinese Academy of Sciences, <sup>3</sup>Department of Natural Resources Management and Engineering The University of Connecticut, Storrs, USA; <sup>4</sup>Department of Life Sciences, Arizona State University West, Phoenix, USA; \*(Author for correspondence at current address: Department of Botany, Duke University, Durham, NC 27708, USA; e-mail: qgao@duke.edu)

Received 6 June 2000; Revised 28 October 2000; Accepted 6 November 2000

*Key words:* ecosystem simulation, grassland, landscape, resolution, songnen plain

### Abstract

The behavioral dependence of vegetation simulation models for spatially heterogeneous grasslands on simulation resolution was investigated. The dependence can be largely attributed to the non-linearity of the models. We showed that increasing scale or decreasing spatial resolution tended to overestimate the changing rate of an ecosystem using our landscape simulation model for alkaline grasslands in northeast China. A technique for scaling up simulation models with diffusive transportation was developed in this study by means of expanding the nonlinear driving functions in the model. The analysis showed that a simulation model for spatially heterogeneous landscapes might necessitate modification of both its mathematical structure and parameterization when applied to different scales. The scaling coefficients derived in this study were shown to be proportional to the variances or covariance of the spatially referenced variables, and can be estimated by running the model at a fine resolution for selected samples of the coarser grid cells. The technique was applied to a grassland landscape in northeast China and the results were compared with five-year observations on community succession. The comparison indicated that the proposed technique could effectively reduce overall scaling error of the model by as much as 80%, depending on the scaling difference between the fine and the coarse resolutions as well as the sampling scheme used.

### Introduction

Scaling is one of the key issues in simulation studies on spatially heterogeneous landscape ecosystems (Allen and Hoekstra 1991; Allen et al. 1994; Fuhlendorf and Smeins 1996; Lawton 1987; Levin 1992; Maurer 1987; Wiens and Milne 1989). Dynamic modeling of ecosystems at landscape scales very often, if not always, involves integration of nonlinear functions of spatially referenced variables. Since numerical integration is a summation of the products of the mean value of an analytical function over each grid cell and the corresponding cell area for a finite time step, the model is valid only for the resolution at which the parameterization is done. General approaches for cross-

scale ecosystem modeling are in great need and have stimulated interests of landscape and system ecologists (Auger 1986; Costanza and Maxwell 1993; Fitz et al. 1996; King 1991; Maurer 1987). While the spatial heterogeneity of landscape ecosystems and the dependence of the spatial simulation models on its calibration scales have been long recognized (Allen and Wileyto 1983; Collins 1995; Luce and Narens 1987; Wu et al. 1997), a hierarchical theory of landscape ecology has been developed to incorporate the scaling effects into consideration (Bonner 1973; Collins and Glenn 1990, 1991; O'Neill et al. 1989, 1992; Wu and Loucks 1995).

Scaling up of nonlinear functions from local scale measurements to large scales remains a challenge for

landscape modelers. In some studies, the scaling effects are totally ignored by directly using locally measured instantaneous quantities, such as carbon assimilation characteristics of individual plants, in regional or global models. In these applications, functions that describe the dependence of the measured variables on environmental conditions for individual plants are also directly employed with hourly, daily or monthly time steps of integration, allowing only a calibration constant to take care of the errors. Such treatments may cause not only cumulated calculation errors in simulation, but also misinterpretation of the results.

The objective of this research was to develop a rational technique to incorporate the scaling effects into a grassland simulation model at landscape scales. The capability of the method is demonstrated by comparing grassland observations against model outputs with and without the proposed scaling treatment.

### Theoretical considerations

We started with the assumption that dynamics of a landscape can be described by the following model:

$$\frac{\partial u_i}{\partial t} = \frac{\partial}{\partial x} \left( \alpha_i \frac{\partial u_i}{\partial x} \right) + \frac{\partial}{\partial y} \left( \alpha_i \frac{\partial u_i}{\partial y} \right) + S_i(p_j, u_r, v_k), \quad (1)$$

where  $u_i$  or  $u_r$  ( $i, r = 1, 2, 3, \dots, n$ ) is the vector of the state variables of the landscape ecosystem;  $\alpha_i$  is the diffusive coefficient describing spatial transportation of energy and mass between adjacent subsystems within the landscape;  $S_i(p_j, u_r, v_k)$  is a source/sink term used to describe local processes, such as plant growth and change in soil water content, at a point within the landscape;  $p_j$  ( $j = 1, 2, \dots, n_p$ ) is a parameter vector with  $n_p$  components; and  $v_k$  ( $k = 1, 2, \dots, m$ ) is a vector of a set of spatially referenced auxiliary variables including the environmental driving functions.

The model, with  $S_i(p_j, u_r, v_k)$  deterministic, is regarded as 'exact' at infinitesimal spatial-temporal resolution. The discretized form of the model, however, is an approximation at the resolutions at which the model parameters are determined in light of observations and experimental results. The numerical solution of Equation (1) inevitably involves integration of  $S_i(p_j, u_r, v_k)$ , or a summation of the products of the function evaluated at finite grid cells and the area of the respective cells. If  $S_i(p_j, u_r, v_k)$  is non-linear, as it is in most cases, changing resolution

clearly produces different integration results. In other words, if the model is reasonably accurate at a certain resolution, simulations at other resolutions using the same governing equations and the same set of parameters are anticipated to introduce certain magnitude of computation errors in the results.

To gain more insight into the scaling characteristics in the model above, let us further assume that the model and parameters are valid at a fine resolution, which we will refer to as micro resolution, with grid size  $\Delta x \Delta y$ , and temporal increment  $\Delta t$ . Equation (1), with constant  $\alpha_i$ , can be expressed in the following finite form

$$\frac{\partial u_i}{\partial t} = \alpha_i \left( \frac{u_i(x + \Delta x, y, t) - 2u_i(x, y, t) + u_i(x - \Delta x, y, t)}{(\Delta x)^2} + \frac{u_i(x, y + \Delta y, t) - 2u_i(x, y, t) + u_i(x, y - \Delta y, t)}{(\Delta y)^2} \right) S_i(p_j, u_r, (x, y, t), v_k(x, y, t)), \quad (2)$$

where  $u_i(x, y, t)$  and  $v_k(x, y, t)$  represent the values of  $u_i$  and  $v_k$ , respectively, at a point  $(x, y)$  within a grid cell at a specific time  $t$ . However, in an area-based model, a variable is in fact often evaluated as the mean value of the variable in the grid cell. That is,  $u_i(x, y, t)$  and  $v_k(x, y, t)$  represent the mean values of  $u_i$  and  $v_k$  respectively in the grid cell with a geometric center  $(x, y)$  and an area  $\Delta x \Delta y$ , within the time interval  $\Delta t$  that include time  $t$ . Equation (2) has an implicit assumption that values of the parameter  $p_j$  were also obtained in the same spatio-temporal scale, so that the integration of  $S_i$  is done at the best accuracy.

Now let us consider the same problem at a larger scale with grid cells of size  $\Delta X \Delta Y$  and time step  $\Delta T$ . We will refer to this as the macro resolution, in contrast to the micro resolution. Discretization of Equation (1) into this resolution requires us to evaluate the average value of  $S_i(p_j, u_r, v_k)$ , or  $\bar{S}_i$ , based on our knowledge of its behavior at the micro scale. A simple mathematical analysis can provide us with

$$\bar{S}_i = \frac{1}{\Delta X \Delta Y \Delta T} \sum_{c=1}^{N_t} \sum_{e=1}^N S_i [p_j, u_r(x_e, y_e, t_c), v_k(x_e, y_e, t_c)] \Delta x \Delta y \Delta t \neq S_i(p_j, \bar{u}_r, \bar{v}_k),$$

where  $\bar{u}_r$  and  $\bar{v}_k$  are average values of  $u_r$  and  $v_k$  within the grid cells at the macro-resolution,  $N$  is the number of micro grid cells in a macro grid cell,  $N_t$  is the number of micro time intervals within  $\Delta T$ ,  $x_e$  and  $y_e$  are

the spatial coordinates of the center of the  $e$ 'th macro grid cell, and  $t_c$  is the time at  $c$ 'th micro time interval. A truncated Taylor expansion of  $S_i$  with respect to  $u_r$  and  $v_k$  gives

$$\begin{aligned} \frac{\partial \bar{u}_i}{\partial T} &= \alpha_i \\ &\left( \frac{\bar{u}_i(X+\Delta X, Y, T) - 2\bar{u}_i(X, Y, T) + \bar{u}_i(X-\Delta X, Y, T)}{(\Delta X)^2} \right. \\ &\quad \left. + \frac{\bar{u}_i(X, Y+\Delta Y, T) - 2\bar{u}_i(X, Y, T) + \bar{u}_i(X, Y-\Delta Y, T)}{(\Delta Y)^2} \right) \\ &S_i(p_j, \bar{u}_r, (X, Y, T), \bar{v}_k(X, Y, T)) + \\ &+ \sum_{\xi=1}^M \sum_{\zeta=1}^M \frac{\partial^2 S_i}{\partial V_\xi \partial V_\zeta} \theta_{\xi\zeta}, \end{aligned} \quad (3)$$

where  $V_\xi$  is a vector resulting from concatenation of  $u_r$  and  $v_k$ , with vector length  $M = m + n$  (the first  $n$  members are of  $u_r$  and the rest  $m$  members are of  $v_k$ ), and  $\theta_{\xi\zeta}$  is a coefficient defined as:

$$\begin{aligned} \theta_{\xi\zeta} &= \frac{1}{2\Delta T \Delta X \Delta Y} \int_{T-\Delta T/2}^{T-\Delta T/2} \int_{X-\Delta X/2}^{X-\Delta X/2} \int_{Y-\Delta Y/2}^{Y-\Delta Y/2} \\ &(V_\xi - \bar{V}_\xi)(V_\zeta - \bar{V}_\zeta) d\vartheta d\rho d\tau \\ &= \frac{1}{2N_i N} \sum_{c=1}^{N_i} \sum_{e=1}^N (V_\xi(x_e, y_e, t_c) - \bar{V}_\xi) \\ &(V_\zeta(x_e, y_e, t_c) - \bar{V}_\zeta). \end{aligned} \quad (4)$$

Thus  $\theta_{\xi\zeta}$  is one half of the variance ( $\xi = \zeta$ ) or covariance ( $\xi \neq \zeta$ ) of the state and auxiliary variables within macro grid cells. Note that the first order term of the Taylor expansion does not appear in Equation (3) because the  $S_i$  was expanded at mean values of  $u_r$  and  $v_k$  within  $(\Delta X, \Delta Y, \Delta T)$ , and the integration of the deviations from the mean value is zero.

Note that in Equation (3), the parameter vector  $p_j$  is hinged to the previous micro resolution as in Equation (2). However, Equation (3) now has additional terms on the right hand side resulting from a scaling up from a relatively fine to a coarse resolution, referred as scaling terms. This result implies that cross-scale modeling may necessitate modifications of both the mathematical structure and the model parameters. Equation (3) also implies that additional parameters  $\theta_{\xi\zeta}$ , which we will refer as *scaling coefficients* hereafter, have to be determined to scale a model from a relatively fine scale up to a large scale with a coarser resolution.

The scaling coefficients  $\theta_{\xi\zeta}$  are in general a function of spatial coordinates and time in macro-resolution, i.e.,  $\theta_{\xi\zeta} = \theta_{\xi\zeta}(X, Y, T)$ . When the macro-resolution  $(\Delta X, \Delta Y, \Delta T)$  is quite much different from the micro-resolution  $(\Delta x, \Delta y, \Delta t)$ , we propose to sample the macro grids to select a subset of cells in macro-resolution. By regarding each selected macro grid cell as a sub-domain of simulation that contains a number of micro grid cells, we can run the model in micro-resolution for each of the sub-domains to obtain a sampled subset of  $\theta_{\xi\zeta} = \theta_{\xi\zeta}(X, Y, T)$  using Equation (4). Interpolation of the sample into all macro grid cells can then be used to scale the model up from the micro- to the macro-resolution.

### Application of the scaling algorithm to a grassland landscape in northeast China

#### The simulation model

With 5-year observations from 1989 to 1993 on spatial patterns of plant communities in a one-hectare alkaline grassland landscape in northeast China, Gao et al. (1996) constructed a model for the landscape to simulate the process of community succession in response to soil alkalization and de-alkalization. The original model was slightly modified in this study to improve the continuity of the functions with respect to state variables. While the detailed mathematical equations are given in Appendix, a brief description is provided here. The model included coverages of five types of plant communities within the 1-ha grassland landscape, dominated respectively by *Calamagrostis epigeios* (CE), *Aneurolepidium chinense* (AC), *Puccinellia tenuiflora* (PT), *Aluopus litorolis* (AL) and *Suaeda corniculata* (SC), and soil alkali, as state variables. We will refer to each type of these communities by its dominant species name hereafter, as the behavior of a community type was assumed to be describable by its dominant species in this study. Competition among plant communities, migration of plant species, and interactions between soil alkali and communities were considered in the model. The mathematical formulation of the model gave 6 coupled partial differential equations for  $C_i$ , the coverage of 5 types of plant communities, for  $i = 1, 2, \dots, 5$ , and  $N_a$  for soil alkali. These state variables corresponds to the variables of  $u_i$  in Equation (1), with  $u_i = C_i$  for  $i = 1, 2, \dots, 5$ , and  $u_6 = N_a$ .

The parameters of the model were tuned for 2 m  $\times$  2 m resolution using a nonlinear least square algo-

rithm (Gao et al. 1996) to produce a minimum sum of squared differences between the simulated and observed total coverage of the 5 types of communities, i.e.,

$$ME = \sqrt{\frac{1}{4n} \sum_{i=1}^5 \sum_{t=2}^5 [U_i(t) - \hat{U}_i(t)]^2}, \quad (5)$$

where  $ME$  stands for grand mean error;  $\hat{U}_i(t) = \iint_A C_i(x, y, t) dx dy$ , with  $A$  denoting the domain of the landscape, is the simulated total coverage of the  $i$ th type of plant community; and  $U_i(t)$  is the same quantity obtained from the observed data. The mean error for an individual community type is

$$ME_i = \sqrt{\frac{1}{4n} \sum_{t=2}^5 [U_i(t) - \hat{U}_i(t)]^2}. \quad (6)$$

The patterns at the first year (year 1989) were used as the initial conditions. The model was solved with periodic boundary conditions, i.e., all fluxes were assumed to be zero at all the boundaries. This of course was a simplification of real boundary conditions.

#### Scaling runs of the model

We used the  $2 \text{ m} \times 2 \text{ m}$  grids as the micro-resolution ( $R_0$ ). Three macro-resolutions,  $4 \text{ m} \times 4 \text{ m}$ ,  $10 \text{ m} \times 10 \text{ m}$ , and  $20 \text{ m} \times 20 \text{ m}$ , named  $R_1$ ,  $R_2$  and  $R_3$ , respectively were used to test the algorithm. The one hectare ( $100 \text{ m} \times 100 \text{ m}$ ) landscape was divided into 2500 micro grid cells in  $R_0$ , 625 grid cells in  $R_1$ , 100 grid cells in  $R_2$  and 25 grid cells in  $R_3$ , respectively. Each grid cell in  $R_1$ ,  $R_2$  and  $R_3$  contains 4, 25, and 100 grid cells in  $R_0$ , respectively. The scaling terms were derived according to Equations (3) and (4) and are given in the Appendix. We first ran the model for all resolutions with the same set of parameter values to obtain the outputs of the model at the 4 resolutions without incorporating the scaling terms into simulation. The outputs of these runs were termed as *unscaled model outputs*.

To compute the scaling terms, we further assumed that the scaling coefficients  $\theta_{\xi\zeta}$  were functions of spatial locations in the macro-resolutions but largely independent of time in this 5-year period, i.e.,  $\theta_{\xi\zeta} = \theta_{\xi\zeta}(X, Y)$ . We then select grid cells in  $R_1$ ,  $R_2$  and  $R_3$ , to run the model at  $R_0$ . The selection was done with the following four systematic sampling schemes as defined by sampling interval (SI):

- (1) SI = 1, or every row and every column, all cells in macro resolutions were samples;

- (2) SI = 2, or every the other row and every the other column;
- (3) SI = 3, or every one in three rows and every one in three columns;
- (4) SI = 5, or every one in five rows and every one in five columns; and finally
- (5) SI = 1', or every row and every column of the initial patterns of the first year observations.

Macro grid cells at 4 corners of the landscape were forced to be sampled for all sampling schemes. Scaling coefficients were computed for these selected macro grid cells, and averaged over years for SI = 1, 2, 3, and 5, and interpolated linearly over all the grid cells in respective macro-resolutions for SI = 2, 3, and 5.

The model was run with scaling terms (Equation 3) for macro-resolutions  $R_1$ ,  $R_2$ , and  $R_3$  to obtain scaled outputs for different sampling schemes. The scaled outputs were then compared with the unscaled outputs and the observations.

## Results and discussion

### *Unscaled output: effects of resolution variation on model behavior*

Figure 1 compares the simulated coverage patterns of the two most important community types dominated by *Aneurolepidium chinense* and *Suaeda corniculata* respectively, against the observed patterns for the years of 1990 through 1993 (indicated as year = 1, 2, 3, and 4, respectively), at the micro-resolution. *A. chinense* is the major grazed plant species at normal local soil conditions and the communities dominated by *A. chinense* are known as major communities. *S. corniculata*, on the other hand, is an indicator of serious soil alkalization, and thus the communities dominated by *S. corniculata* are results of serious degradation of grassland landscapes in Songnen Plain. The two types of communities together occupied more than 80% of the total area in this study, leaving only less than 20% coverage for other three types of communities (*C. epigeous*, *A. littoralis*, and *P. tenuiflora*). The figure indicated that the model is in good agreement with the observations.

Unscaled model outputs for all resolutions were illustrated in Figures 2 and 3. While Figure 2 plots simulated total coverage of the two important types of communities (a, b) and soil alkali (c) cross all resolution in comparison with observations, Figure 3 depicts

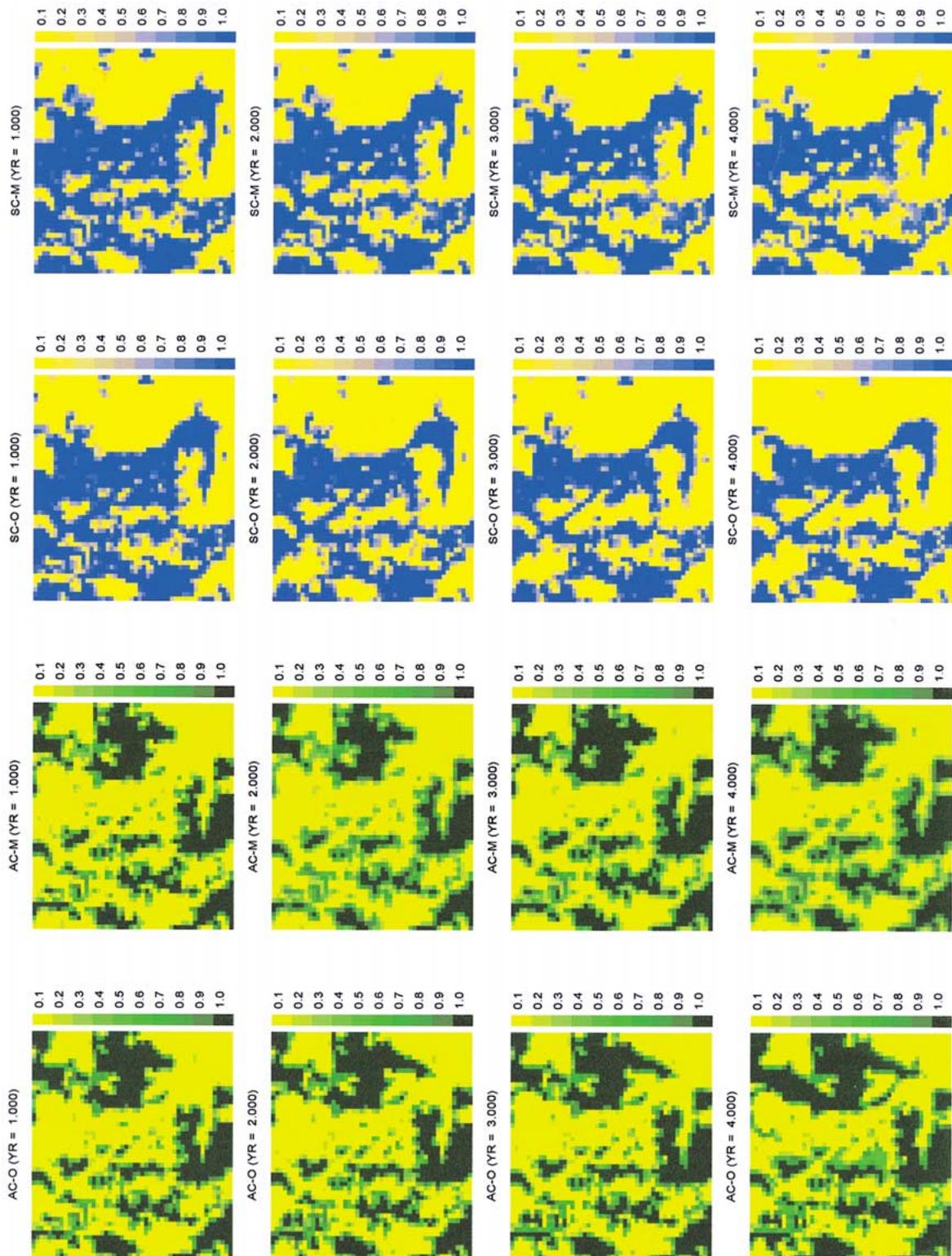


Figure 1. Comparison between observed and simulated community coverage in a one-hectare alkaline grassland landscape at 2 m x 2 m resolution. AC, *A. chinense* communities; SC, *S. corniculata* communities. Suffix 'O' denotes observations, and 'M' denotes simulation.

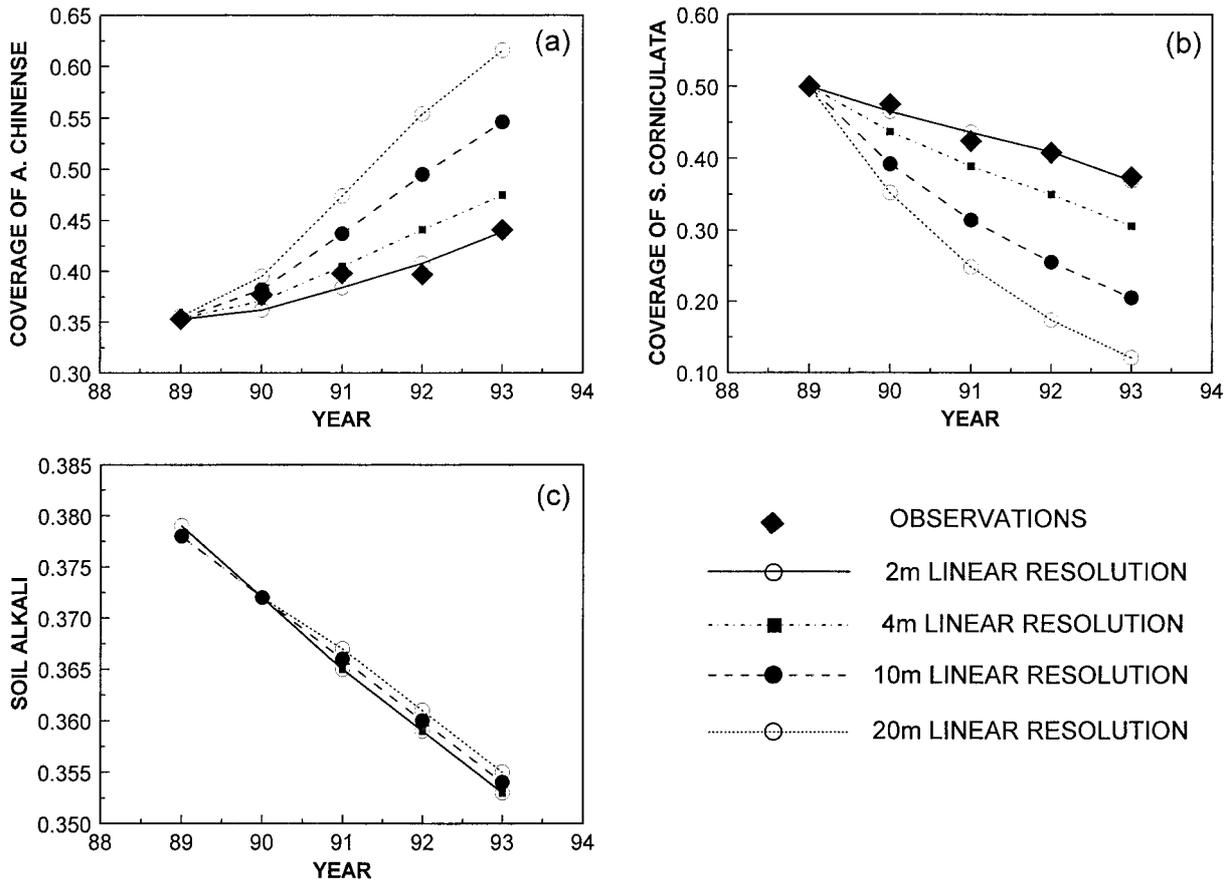


Figure 2. Unscaled simulation outputs of the model, averaged over the domain of the alkaline grassland landscape, in comparison with observation. (a) *A. chinense* communities, (b) *S. corniculata* communities, and (c) soil alkali.

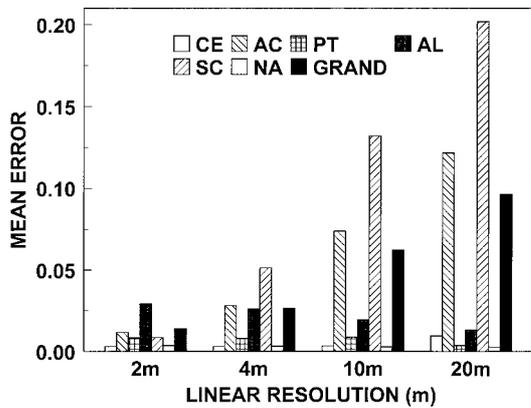


Figure 3. Error associated with the unscaled simulation outputs for all the state variables of the grassland model. CE, AC, PT, AL and SC stand for *C. epigeios*, *A. chinense*, *P. tenuiflora*, *A. litorolis* and *S. corniculata* communities, NA is soil alkali, and GRAND is the total grand average error.

the corresponding errors of the model without scaling treatment.

The simulated coverage of the two community types were close to observations at 2 m resolution ( $R_0$ ), but deviated from observations systematically as the resolution goes from fine to coarse. The largest difference between model and observation was seen for  $R_3$  (20 m resolution). Soil alkali was shown to be less sensitive to variation of resolution. The different behavior between soil alkali and community coverage can be explained in terms of the model equations. The source/sink terms of the model for community coverages are nonlinear with respect to both soil alkali and community coverage. On the other hand, the source/sink term for soil alkali is linear (see Appendix) with respect to the state variables. Even though the nonlinearity of community coverage equations is coupled with soil alkali dynamics, the coupled effect on soil alkali seemed indirect, and affected little the simulated soil alkali across all the resolutions.

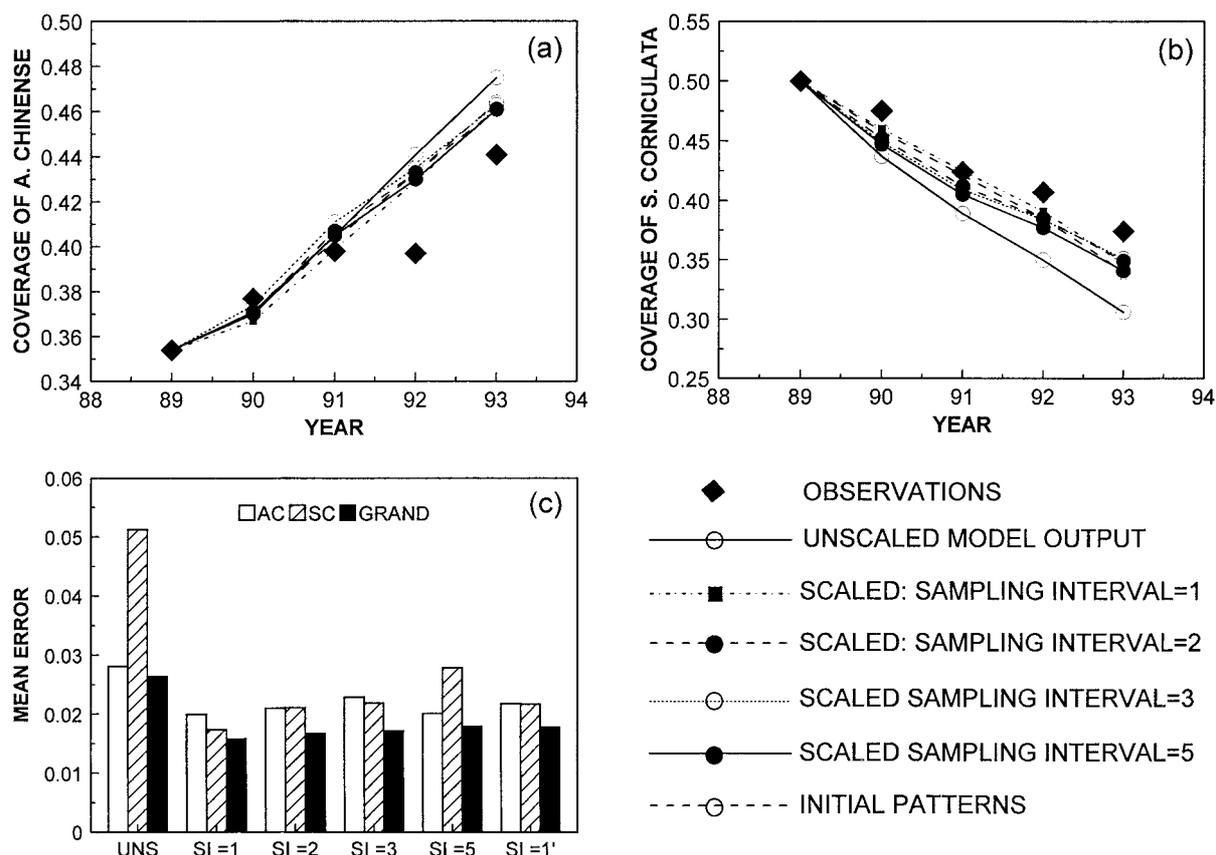


Figure 4. Scaled model outputs at  $4\text{ m} \times 4\text{ m}$  resolution. (a) *A. chinense* coverage; (b) *S. corniculata* coverage; (c) Mean error between model outputs and observations. Codes in (c): UNS = unscaled, SI = sampling interval. Other codes are the same as in the previous figures.

Another important result shown in Figures 2 and 3 is that scaling up from a finer to a coarser resolution tended to overestimate the rates of major ecosystem processes in this model. While the observed coverage of *A. chinense* communities increased from 35% to 44% during the period, the corresponding coverage of *S. corniculata* communities decreased from 50% to 37%. The dynamic coverage of these two types of communities was closely simulated at 2 m resolution. But the simulated increases in the coverage of *A. chinense* communities were from 35% to 46%, 55% and 62%, and the estimated decreases in the coverage of *S. corniculata* communities were from 50% to 31%, 21%, and 12%, for  $R_1$ ,  $R_2$  and  $R_3$ , respectively. Figure 3 indicates that the mean errors for those two types of communities and the grand mean error were more than tripled as resolution went from 2 m to 20 m, whereas the mean errors for other types of communities and soil alkali remained relatively constant.

The reason for the resolution-dependent deviation was the overestimation of the absolute rates of ecosystem processes in the macro-resolutions. The overestimation became much more serious in this model as the resolution became coarser.

#### Simulated plant community coverages with scaling algorithm

Figures 4, 5, and 6 show the simulated coverage of the two community types by employing the scaling algorithm developed in this study, using the same sampling schemes at  $R_1$ ,  $R_2$  and  $R_3$  resolutions respectively. Adding the scaling terms to the model resulted in significant reductions in the mean simulation errors for each community type as well as the grand mean simulation error of the model. The reduction in the mean errors of the simulations varied from 11% to 81%, depending on the resolution and the sampling interval (SI). In general, the relative error reduction is more evident for a coarser resolution than for a finer resolu-

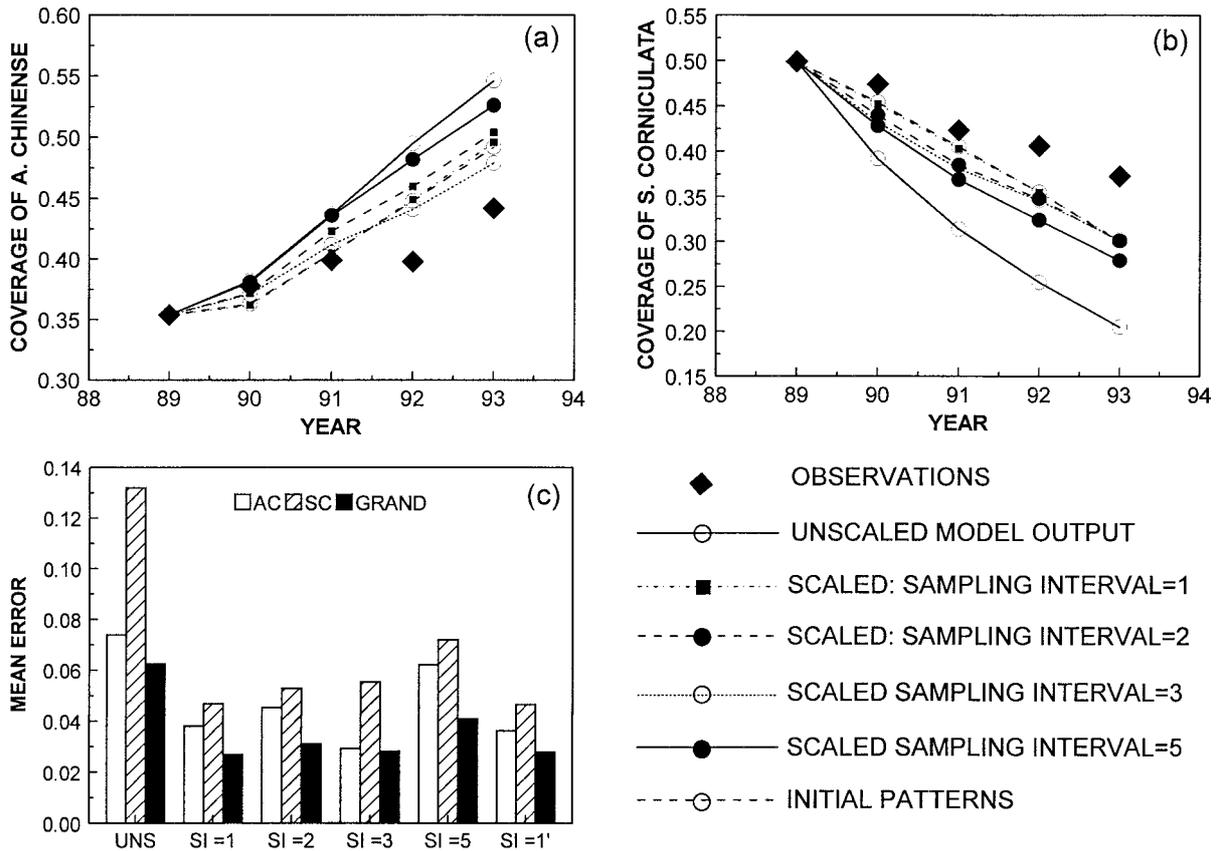


Figure 5. Scaled model outputs at 10 m x 10 m resolution. Codes and legends are the same as in the previous figures.

tion, and the reduction in mean error was less as the sampling interval became larger, for a larger sampling interval usually means less samples taken. Fewer samples are in general less effective to represent the true distribution of the scaling coefficients. One exception was the results for the 20 m resolution, where mean errors of the *A. chinense* community type for SI = 3 and 5 were smaller than for SI = 1 and 2. The reason for this exception might be due to the effects of sampling. As we had only a total of 25 grid cells in the sampling pool at this resolution, SI = 3 and 5 generated only 9 and 4 grid cells in the sample, respectively. The sampling schemes might incidentally give more accurate estimation of the scaling coefficients than the full sampling scheme (SI = 1).

There are still some systematic errors that were not corrected by the proposed scaling algorithm. The remaining errors may be attributed to the truncation in the Taylor expansion, and may also have something to do with the employed numerical method. Even if the differential equations of a model are linear with re-

spect to state variables, solutions of the equations can be nonlinear (King 1991). The errors associated with this kind of nonlinearity cannot be reduced by means of our method. Hence the results of this research only represented an incremental advancement for a better scaling. Further studies are needed to improve the methodology.

**Summary and conclusions**

We demonstrated the dependence of ecosystem simulation on the resolution of numerical computation. We also developed a numerical algorithm to reduce the errors associated with the scaling up to improve the results of cross scale simulation. The technique was based on Taylor expansion of nonlinear functions in the model. The scaling coefficients were derived and shown to be proportional to the variance or covariance of the spatially referenced variables and can be estimated by running a model at micro resolution for

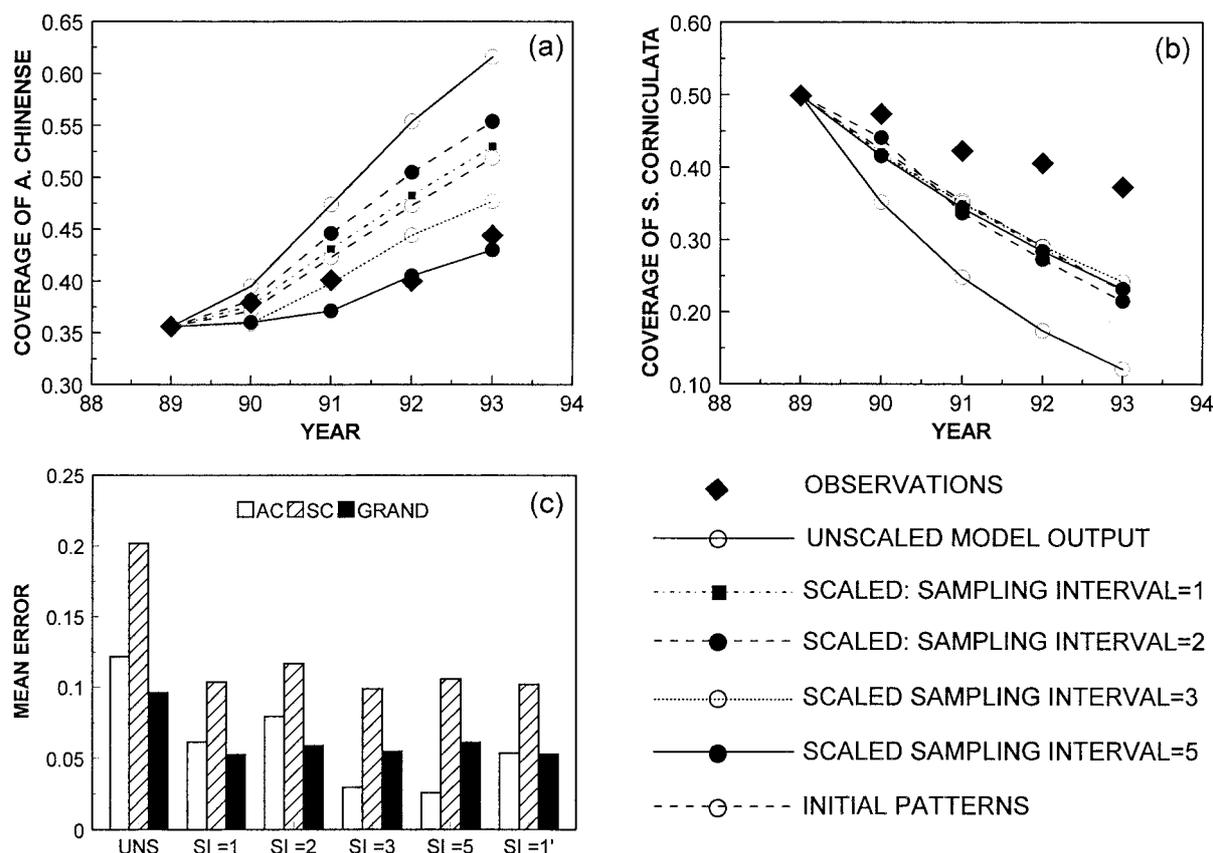


Figure 6. Scaled model outputs at  $20\text{ m} \times 20\text{ m}$  resolution. Codes and legends are the same as in the previous figures.

selected macro grid cells. The following conclusions can be drawn from the results:

- (1) The mathematical formulation and parameters of an area-based model for spatially heterogeneous ecosystems are always hinged to a particular spatio-temporal scale. The dependence of model behavior on scale is attributed to the nonlinearity of the model. Changing the scale of a model may necessitate modification of both mathematical equations and parameterization. The common practice of calibration used in many global and regional models may not be adequate for cross-scale modeling.
- (2) Decreasing resolution of computation or increasing size of grid cells tended overestimate the rates of major ecosystem process in our alkaline grassland landscape model.
- (3) The scaling algorithm developed in this study can effectively reduce the modeling error by as much as 80%, depending on the scaling distance (the difference between two resolutions) and the sampling

frequency at which the sites in a coarse resolution are selected to run the fine scale model.

Application of the proposed scaling technique to our grassland model used explicit analytical differentiation, because the model was mathematically simple enough. For applications to complicated simulation models for which analytical derivatives with respect to state/auxiliary variables are not feasible, numerical differentiation may be used.

#### Acknowledgements

This research was jointly supported by the Natural Science Foundation of China under Grant numbers 39725006, 39770133 and 39899370, the University of Connecticut, and Arizona State University.

## References

- Allen, T.F.H., King, A.W. and Milne, B.T. 1994. The problem of scaling in ecology. *Evol. Trends Plants* 7: 3–8.
- Allen, T.F.H. and Hoekstra, T.W. 1991. Role of heterogeneity in scaling of ecological systems under analysis. *In Ecological Heterogeneity*. pp. 47–68. Edited by J. Kolasa and S.T.A. Pickett. Springer-Verlag, New York, NY, USA.
- Allen, T.F.H. and Wileyto, E.P. 1983. A hierarchical model for the complexity of plant communities. *J. Theor. Biol.* 101: 529–540.
- Auger, P. 1986. Dynamics in hierarchically organized systems: a general model applied to ecology, biology and economics. *System Res.* 3: 41–50.
- Bonner, J. 1973. Hierarchical control programs in biological development. *In Hierarchical Theory: the Challenge of Complex Systems*. pp. 49–70. Edited by H.H. Pattee. George Braziller, Inc., New York, NY, USA.
- Collins, S.L. 1995. The measurement of stability in grasslands. *Trends Ecol. Evol.* 10: 95–96.
- Collins, S.L. and Glenn, S.M. 1990. A hierarchical analysis of species abundance patterns in grassland vegetation. *Am. Nat.* 135: 633–648.
- Collins, S.L. and Glenn, S.M. 1991. Importance of spatial and temporal dynamics in species regional abundance and distribution. *Ecology* 72: 654–664.
- Costanza, R. and Maxwell, T. 1993. Resolution and predictability: an approach to the scaling problem. *Landscape Ecol.* 9: 47–57.
- Fitz, H.C., Debelleve, E.B., Costanza, R., Boumans, R., Maxwell, T., Wainger, L. and Sklar, F.H. 1996. Development of a general ecosystem model (GEM) for a range of scales and ecosystems. *Ecol. Modeling* 88: 263–295.
- Fuhlendorf, S.D. and Smeins, F.E. 1996. Spatial scale influence on long term temporal patterns of a semi-arid grassland. *Landscape Ecol.* 11: 107–113.
- Gao, Q. 1996. Dynamic modeling of ecosystems with spatial heterogeneity, a structured approach implemented in Windows environment. *Ecol. Modelling* 85: 241–252.
- Gao, Q., Li, J. and Zheng, H. 1996. A dynamic landscape simulation model for the alkaline grasslands on Songnen Plain in northeast China. *Landscape Ecol.* 11: 339–349.
- King, A.W. 1991. Translating models across scales in the landscape. *In Quantitative Methods in Landscape Ecology*. pp. 479–518. Edited by M.G. Turner and R.H. Gardner. Springer-Verlag, New York, NY, USA.
- Lawton, J.H. 1987. Problems of scale in ecology. *Nature* 325: 206.
- Levin, S.A. 1992. The problem of pattern and scale in ecology. *Ecology* 73: 1943–1967.
- Luce, R.D. and Narens, L. 1987. Measurement scales on the continuum. *Science* 236: 1527–1523.
- Maurer, B.A. 1987. Scaling of biological community structure: a systems approach to community complexity. *J. Theor. Biol.* 127: 97–110.
- O'Neill, R.V., Johnson, A.R. and King, A.W. 1989. A hierarchical framework for the analysis of scale. *Landscape Ecol.* 3: 193–205.
- O'Neill, R.V., Gardner, R.H. and Turner, M.G. 1992. A hierarchical neutral model for landscape analysis. *Landscape Ecol.* 7: 55–61.
- Wiens, J.A., and Milne, B.T. 1989. Scaling of landscapes in landscape ecology, or landscape from a beetle's perspective. *Landscape Ecol.* 3: 87–96.
- Wu, J. and Loucks, O.L. 1995. From balance of nature to hierarchical patch dynamics: a paradigm shift in ecology. *Quart. Rev. Biol.* 70: 439–466.
- Wu, J., Gao, W. and Tueller, P.T. 1997. Effects of changing spatial scale on the results of statistical analysis with landscape data: a case study. *Geogr. Inf. Sci.* 3: 30–41.

## Appendix 1. Governing equations of the grassland model

The governing equations in our alkaline grassland landscape model are given in the following:

$$\begin{aligned}\frac{\partial C_i}{\partial t} &= \alpha_i \left( \frac{\partial^2 C_i}{\partial x^2} + \frac{\partial^2 C_i}{\partial y^2} \right) + g_i G_i W_i, \quad i = 1, 2, \dots, 5, \\ \frac{\partial N_a}{\partial t} &= \alpha_n \left( \frac{\partial^2 N_a}{\partial x^2} + \frac{\partial^2 N_a}{\partial y^2} \right) + g_n W_n,\end{aligned}\quad (7)$$

where  $C_i$  is the coverage of the  $i$ th community type ( $\text{m}^2 \text{m}^{-2}$ );  $N_a$  is the soil alkali, defined as the fraction of the exchangeable in the total cations in 100 grams of dry soil;  $\alpha_i$ ,  $g_i$  and  $\alpha_n$ ,  $g_n$  are model parameters. The term  $g_i G_i W_i$  describes the change in the coverage of community type  $i$  due to local ecological processes. Similarly,  $g_n W_n$  is the local net source term for  $N_a$ , representing the processes that bring up alkaline solutes from deep to surface soil (source) or flush solutes downward into the deep soil (sink). The variable  $G_i$  is the coverage response function of community type  $i$  to soil alkali, and was assumed to have the following form

$$G_i(N_a) = 4 \exp \left[ - \left( \frac{N_a - N_{a0i}}{1.865 R_i} \right)^2 \right] - 3, \quad i = 1, 2, \dots, 5, \quad (8)$$

where  $N_{a0i}$  represents the optimal soil alkali for species  $i$ . This response function obtains its maximum value of 1 at  $N_a = N_{a0i}$ . The quantity  $R_i$  in Equation (8) stands for the tolerance range of species  $i$  to soil alkali. When soil alkali is greater than  $N_{a0i} + R_i$  or smaller than  $N_{a0i} - R_i$ , the local coverage increasing rate will be zero or negative, bounded by  $-3$ . We assumed that the ecological characteristics of a plant community could be represented by the behavior of its dominant species.

Function  $W_i$  depends on community coverage  $C_i$  and was assumed to be in the following form:

$$W_i = \begin{cases} C_i \left( 1 - \sum_j C_j \right), & \text{if } G_i > 0, \\ C_i, & \text{otherwise.} \end{cases} \quad (9)$$

Thus, the local rate of coverage variation took the form of the classical Logistic model. The competition among plant community types was reflected in the term within the parentheses in Equation (9). The decreasing rate of  $C_i$ , however, was assumed to be proportional to  $C_i$ . Equation (9) also indicates that the sum of coverage of all communities at any point in the simulation domain cannot be greater than 1.

The function  $W_n$  describes the dependence of the local rate of change in  $N_a$  on surface vegetation conditions, as:

$$W_n = \begin{cases} N_{a \max} - N_a, & \text{if } \sum_i C_i < C_{\min}, \\ -N_a, & \text{otherwise,} \end{cases} \quad (10)$$

where  $C_{\min}$  is the pivot value of the total vegetation coverage and  $N_{a \max}$  is the upper bound of  $N_a$ .

## Appendix 2. Scaling terms

To apply the proposed scaling method to this particular model, we first match up variables and parameters in Equation (1) to the variables in Equations (7)–(10). Our grassland model had no auxiliary variables. Hence  $n = 6$  and  $m = 0$ . The state variable  $u_i$  in Equation (1) is  $C_i$  in Equation (7) for  $i = 1-5$ , and  $u_6$  in Equation (1) is  $N_a$  in Equation (7). The source/sink term in Equation (1) are matched to functions in (7) as follows:  $S_i = g_i G_i W_i$  for  $i = 1-5$ , and  $S_6 = g_n W_n$ . The following differentiation was carried out for the scaling terms in Equation (3):

$$\frac{\partial^2 S_i}{\partial C_i \partial C_j} = \begin{cases} \begin{cases} -2g_i G_i, & G_i > 0 \\ 0, & \text{otherwise} \end{cases}, & i = j \\ \begin{cases} -g_i G_i, & G_i > 0 \\ 0, & \text{otherwise} \end{cases}, & i \neq j \end{cases}, \quad i = 1, 2, \dots, 5 \quad (11)$$

$$\frac{\partial^2 S_i}{\partial N_a \partial C_j} = \left\{ \begin{array}{l} \left\{ \begin{array}{l} -2.3g_i \exp \left[ - \left( \frac{N_a - N_{a0i}}{1.865R_i} \right)^2 \right] \times \left( 1 - \sum_{j=1}^5 C_j - C_i \right) \left( \frac{N_a - N_{a0i}}{R_i^2} \right), \quad G_i > 0 \\ -2.3g_i \exp \left[ - \left( \frac{N_a - N_{a0i}}{1.865R_i} \right)^2 \right] \times \left( \frac{N_a - N_{a0i}}{R_i^2} \right), \quad \text{otherwise} \end{array} \right\}, \quad i = j \\ \left\{ \begin{array}{l} -2.3g_i \exp \left[ - \left( \frac{N_a - N_{a0i}}{1.865R_i} \right)^2 \right] \times \left( \frac{N_a - N_{a0i}}{R_i^2} \right), \quad G_i > 0 \\ 0, \quad \text{otherwise} \end{array} \right\}, \quad i \neq j \end{array} \right. \quad (12)$$

$$i = 1, 2, \dots, 5,$$

where all the second order derivatives of  $S_6$  with respect to the state variables were zero.

### Appendix 3. Field observations and model parameterization

From 1989 to 1993, one hectare of seriously alkalized grassland was fenced up in southern Songnen Plain, northeast China to study the recovery dynamics of the plant communities. A map of spatial distribution of plant communities within the fenced area was drawn in August in each year with 10-cm sampling resolution (Gao et al. 1996). The maps were digitized, aggregated into 2 m resolution for this study, and were used to obtain  $a_i$ ,  $g_i$ ,  $a_n$ , and  $g_n$  by means of a nonlinear least square algorithm (Gao 1996). The parameterization was done by the following steps:

- (1) A set of  $a_i$ ,  $g_i$ ,  $a_n$ , and  $g_n$  was arbitrarily selected to run the model at 2 m resolution for 5 years at 1 year time increment. The field observed community distribution pattern in 1989 was used as initial conditions;
- (2) The simulated coverage patterns in the successive 4 years were then compared with observations to compute the sum of error squares;
- (3) The well-known Gauss–Newton algorithm was used to modify and update parameter  $a_i$ ,  $g_i$ ,  $a_n$ , and  $g_n$  for the successive runs of the model to produce a smaller sum of error squares;
- (4) The simulation rerun repeatedly with modified  $a_i$ ,  $g_i$ ,  $a_n$ , and  $g_n$ . The model was parameterized with the acceptable value of the sum of the error squares.