Uncovering Spatial Feedbacks At Alpine Treeline Using Spatial Metrics In Evolutionary Simulations

George P. Malanson and Yu Zeng

Department of Geography
University of Iowa
Iowa City, IA 52242 USA
Telephone: +1 319-335-0540
Fax: +1 319-335-2725
Email: george-malanson@uiowa.edu

Abstract

The recognition of feedbacks between plants and their physical environment is important in biogeography. At alpine treeline, positive feedbacks for trees versus tundra include warming the canopy via lower albedo, increasing nutrients and water by increasing local atmospheric deposition, and reducing transpiration and abrasion by slowing wind; negative feedbacks include shading and cooler soil. To study the relations between the spatial pattern of trees, the feedbacks that ensue, and the potential change in the spatial pattern that results, we apply genetic algorithms to a cellular automaton model to simulate the advance of trees into tundra. In one model, spatial metrics on trees in the neighborhood of a tundra cell determine its probability of becoming a tree cell and are used in the fitness function. In a second model, the genetic algorithm addresses a polynomial series for the feedbacks and the frequency distribution of a spatial metric is the fitness function. The simulations presented here support the interpretation of self-organization in the alpine treeline ecotone. Moreover, in self-organizing systems, evolutionary computation has the potential to help understand the broad form of the functions describing system behavior.

1. Introduction

The spatial patterns of ecosystems in landscapes are important to their functioning (Turner 1989). Exchanges of energy, matter, and individual organisms and species across landscapes are in part determined by pattern, and it is this exchange, along with vertical fluxes, that creates the pattern. Central to our understanding of pattern, spatial fluxes across landscapes, and changing pattern are ecotones, the transition zones between adjacent ecosystems (Risser 1995). Ecotones are primarily characterized by the change is species composition, but are also seen as areas of interaction among the ecosystems. The change in species may be important because it indicates the limits for a species range, at least locally, and can be indicative of the species relation to the environment and other species. Fluxes may be higher here because of steeper gradients in the environment that coincide with the difference in species.

1.1 Treeline ecotones

Treelines are the ecotones most studied. Most treeline studies are of those where forest borders grassland or tundra (rather than tree-shrub boundaries). The contrast in physiognomy has lead to a number of hypotheses of treeline formation, location and pattern. All recognize multiple causes, but have different emphases. Stevens and Fox (1991) pose stature and growth-related hypotheses. The stature-related hypotheses involve the problems that a large plant encounters when resources are spatially and temporally limited, such as inability to accumulate enough resources in its space. Other stature related problems are those of exposure of the terminal buds and/or photosynthetic tissue to damage. Growth-related hypotheses are related to the respiration costs of a large woody plant when resources are few in time (i.e., seasonally) and/or space. Analyses of carbon balance indicate that calculations of photosynthesis minus respiration are a fair indicator of treeline location (Cairns and Malanson 1998), but physiological studies indicate that treeline carbon storage may be more flexible than previously thought (Hoch & Korner 2003). These studies did not, however, consider alpine treeline advance and its spatial configuration from the perspective of complexity. Our recent work indicates that endogenous fractal dynamics may be the underlying mechanism (Zeng & Malanson, in review).

Important to this understanding is recognition of feedbacks between trees and their physical environment. Wilson and Agnew (1992) theorized that positive feedback occurs where trees can modify their local environment in their own favor or cause environmental deterioration nearby, and that forbs or grasses can do the same. Negative feedback has not been as well theorized for treelines in general, but shading the soil is one (Korner 1998). With feedbacks comes the recognition of the role of spatial pattern in affecting the process. Feedbacks are local. They occur in the immediate vicinity of the trees, so that the pattern of trees determines the pattern by which the underlying and surrounding environment is modified; this process affects the pattern in a feedback loop (Malanson 1997).

Alpine treeline, also called the forest-alpine tundra ecotone, typifies such ecotones (Figure 1). Positive feedbacks for trees versus tundra include warming the canopy via lower albedo, increasing nutrients and water by increasing local atmospheric deposition, and reducing transpiration and abrasion by slowing wind. The boundary is between adjacent ecosystems of very different physiognomy. The macro-scale controls are clearly related to climate, given the elevation gradient of treeline from the equator to high latitudes. The local elevations and pattern are not so simple, and so may help us understand the interaction of pattern and process.

Figure 1. Patterns at treeline include advancing fingers, isolated patches, and openings remaining within contiguous forest.

We wish to study the relationship between the spatial pattern of trees, the feedbacks that ensue, and the potential change in the spatial pattern that results. This study is part of a project that includes other modeling approaches. Having started with a topography-based empirical model (Brown et al. 1994), we now use a physiologically mechanistic model that calculates photosynthesis, respiration, and the allocation of carbon (Cairns & Malanson 1998), a forest dynamics model to simulate the population dynamics of trees (Bekker et al. 2001), and a cellular automaton (CA), wherein the weights of neighboring trees on a cells tree/tundra status are based on calculation of the physiological effects (Alftine & Malanson, in review). However, there are two obstacles. First, landscape processes and their interactions with spatial patterns are complex and thus it is difficult to identify specific mechanisms. Second, even if an expression can be used, it is difficult to optimally quantify any set of parameters. Here, we attempt to determine if the weights for the neighboring effects of trees on the status of a cell in a CA can be determined using evolutionary computation.

1.2. Evolutionary Computation
Evolutionary computation is a means to produce aspects of computer programs that achieve specified ends. Commonly, EC is applied to optimization problems in geography, wherein the locations of facilities or land-use sites are optimized relative to multiple criteria. It is essentially an engineering approach that models the evolutionary process to search a large parameter space and is used to avoiding being trapped in local minimum while approximating the global maximum because of its genetic operations and stochastic nature under selection pressure. The inherent parallelism in EC can explore a large solution space in an efficient way (Mitchell 1996). These techniques have been pioneered in human geography for location-allocation (e.g., Hosage & Goodchild 1986, Xiao et al. 2002), and for spatial modeling models – including pattern recognition – by Openshaw (1992, 1995, 1998, Diplock & Openshaw 1996, Turton & Openshaw 1998). They have been applied in land use (Matthews et al. 1999, Bennett et al. 1999) and transportation (e.g., Pereira 1996). Manson (200?) used genetic programming to search for local farmers' land use strategies. Applications related to biogeography are few (e.g., Noever et al. 1996) and those with spatial detail are rare (e.g., Hraber & Milne 1997, Giske et al. 1998). The issue of how to make use of this optimization method for scientific simulation needs to be further explored. Here we evaluate genetic algorithms (GA), which is one of the methods under the umbrella of EC, to help us in this task.

2. Analysis with three spatial metrics
2.1 Methods A

We apply genetic algorithms (GA) to a CA model of alpine treeline to simulate the advance of trees into tundra. We begin with a CA in which the cells are either tree or tundra (cf. Noble 1993, Malanson et al. 2001). We assume that the presence of trees in a neighborhood will increase the probability that a tundra cell will become a tree cell based on the positive feedback mechanism (Wilson & Agnew, 1992; and others). The number or trees in the neighborhood can be considered in several ways. Landscape ecologists have developed numerous spatial metrics for quantifying landscape pattern (McGarigal & Marks 1995). Here, we chose three metrics that we though should be important: number of trees, size of patches, and total number of cells. Because we do not know the correct weights or parameters in an expression of feedback, we adopted evolutionary computation to search for the most likely ones. GA essentially is a goal-directed random search program and we can take advantage of its implicit parallelism in maintaining a large set of search points at any time. Here, the task of the GA is to find the weights that determine the effect of these metrics on a cell. Another consideration to adopt GA is also related to the uncertainty of the exact form of feedbacks. It is possible that the feedback can be expressed in different but equivalent forms. Even though the exact forms of feedbacks are uncertain or unknown, we can allow a GA to approach the best results that can be obtained under different forms of transition rules.

2.1.1 General design of the simulation system

There are three modules in this system: GA, CA, and spatial metrics calculation (Figure 2). Operationally, they are organized into three layers: the first layer is the GA which creates a population of values in parameter space expressed by sets of weights according to the genetic operations of mutation, crossover, and selection. The second layer is the CA model that accepts a set of weights into its transition rule, which drives the treeline advance. The third layer is the spatial metrics module that is called by CA model to calculate the spatial metrics in a designated neighborhood. The results of the treeline spatial patterns are in turn fed back to the GA module for the purpose of computing fitness a value in the genetic operations and producing the next generation of parameters or sets of weights for future CA steps.
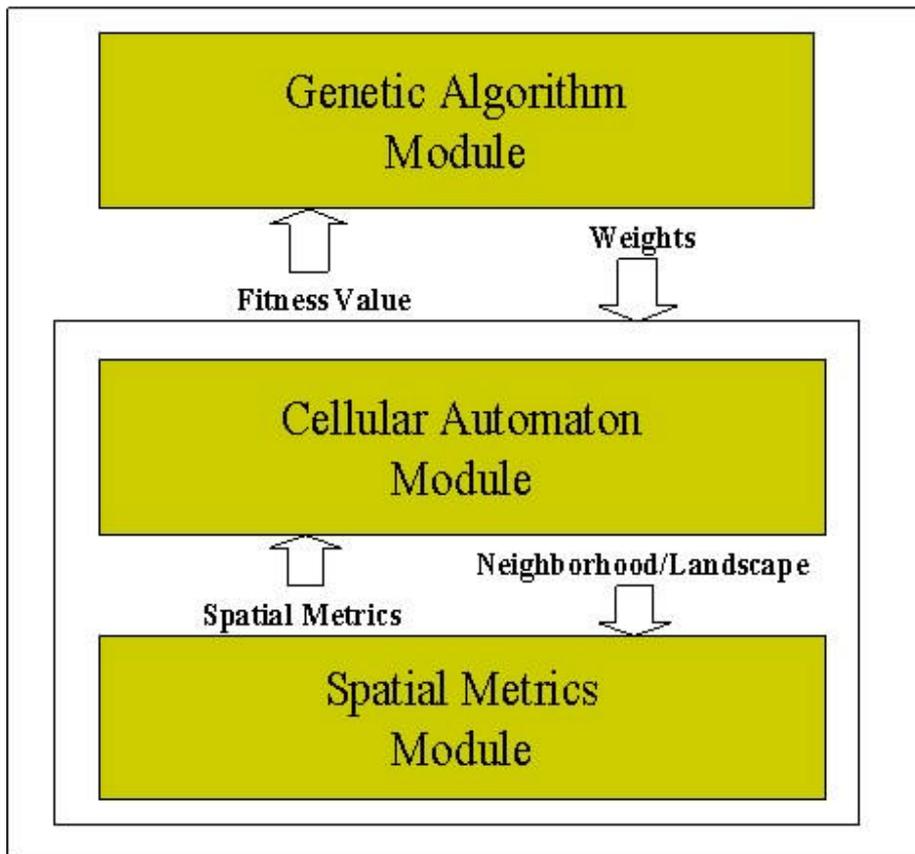
Module Diagram

Figure 2. Outline of the interaction between components of our EC model for alpine treeline.

2.1.1.1 CA model

Our CA is a 2-D lattice with each cell having 2 possible states, tree or tundra. The lattice represents a long gradual slope of tundra but initially with trees near the bottom. Tundra cells can change to trees, but not the reverse. The lattice is initialized with a gradient of probability of a tundra cell changing to a tree cell. The initial probability of a cell changing to tundra is modified by trees in its neighborhood. The amount of change in the probability is calculated as a function of selected spatial metrics.

2.1.1.2 Spatial metrics

Our fitness function is defined as the combination of spatial metrics that best match the observed pattern at an alpine treeline. Landscape ecology is largely founded on the notion that environmental patterns strongly influence ecological processes (Turner 1989). Landscape patterns can be quantified in a variety of ways depending on the type of data collected, the manner in which it is collected, and the objectives of the investigation (McGarigal & Marks 1995). At each iteration the pattern of trees and tundra is analyzed. We are trying to determine if the clumping of trees has an effect different from the

number of trees. We construct patches of contiguous trees and quantify the number and size of patches and the density of trees. We ran the simulation using different neighborhood window sizes ranging from 3×3 to 9×9; the results do not show any significant difference.

2.1.1.3 Genetic algorithm

To find the optimal weights for the spatial metrics used to set the probabilities in the CA, we used a GA. Population level operations included 1) Roulette Wheel Selection; and subsequent chromosome level operations: 2) Reproduction by N-point Crossover; 3) Mutation; and 4) Statistics of the average, highest and lowest fitness value. We use a distance measure in the space defined by the three spatial metrics as the fitness function. Similarity is measured to the pattern observed in a classified ADAR (aerial multispectral 1-m spatial resolution) image of a section of treeline on Lee Ridge, Glacier National Park, MT (Figure 3).
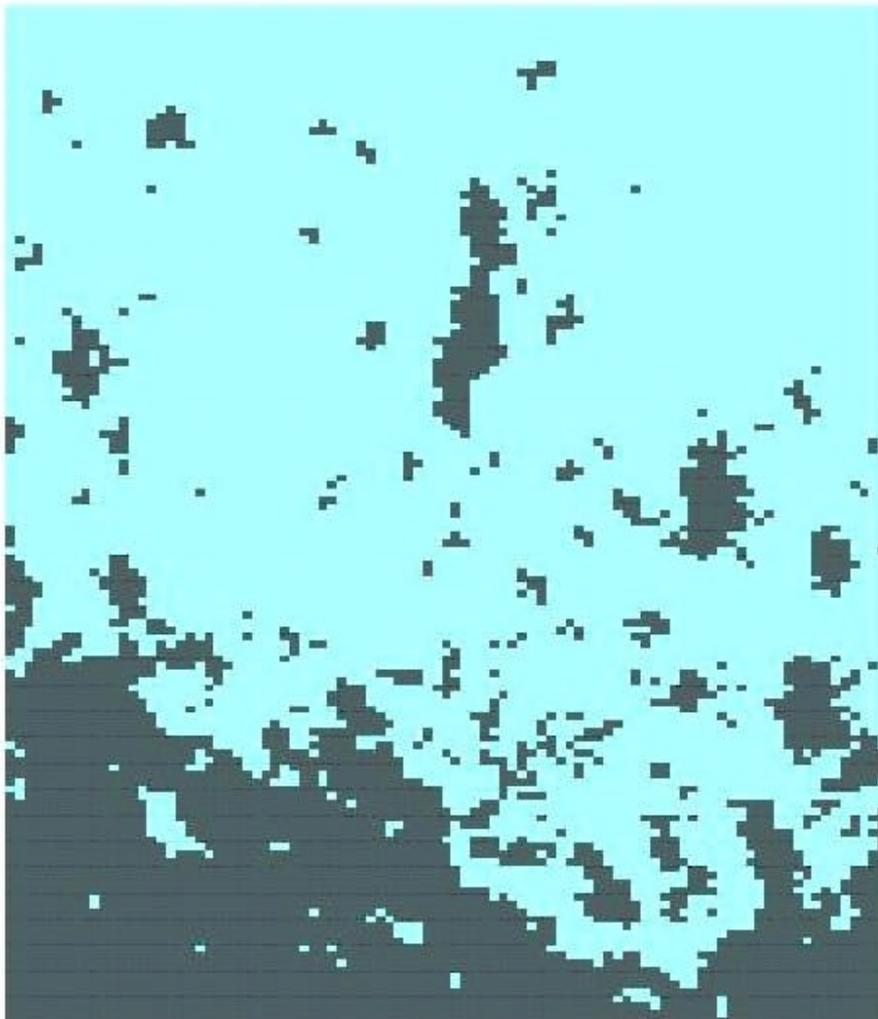

Figure 3. Classified ADAR multispectral aerial photograph (1 m resolution) of a portion of Lee Ridge, Glacier National Park, MT, USA.

2.1.1.4 Monte Carlo runs

Because of high variances of the resultant spatial patterns from the same set of weights in the transition rule, we adopted a Monte Carlo approach to try to increase the reliability of the fitness comparison and the next generation of weights used in the model. We tried the variance reduction technique of antithetic variables but without success. We believe that the variance of the resultant spatial patterns is equivalent to different stages of treeline advance and can be considered as different realizations of the same system and that it arises from self-organization within this system, as discussed below.

2.1.2 Tests

We used EC to determine the weights in the CA for two conditions, a linear model:

p = weight1 * patch number + weight2 * average patch size + weight3*total cell number

and a nonlinear model:

p = weight1 * patch number + weight2 * (average patch size)^2 + weight3*total cell number

Mean fitness is calculated by averaging the fitness values of one generation of chromosomes, which represents how fit this generation is and how close it to the global maximum. The mean fitness series also tells the progress of the GA.

Elite fitness measures the best chromosome of the current generation, producing the highest fitness value. In this program, we use elitism to keep the best chromosome in next generation to accelerate the evolutionary process of optimization.

2.2 Results A

The landscapes resulting from nonlinear transition rule are much closer to the observed one than those from the linear transition rule after optimization (Figure 4). Even the elite fitness of the linear model did not perform as well as the mean fitness of the nonlinear model. The overall results do not show improvement through time after the early generations. These results suggest the nonlinear positive feedback is more likely to be the driving force in creating alpine treeline patterns. This result does not inform us about the usefulness of EC, but instead points to another line of research, wherein EC may be useful, i.e., the determination of the value for the exponent in the nonlinear function.
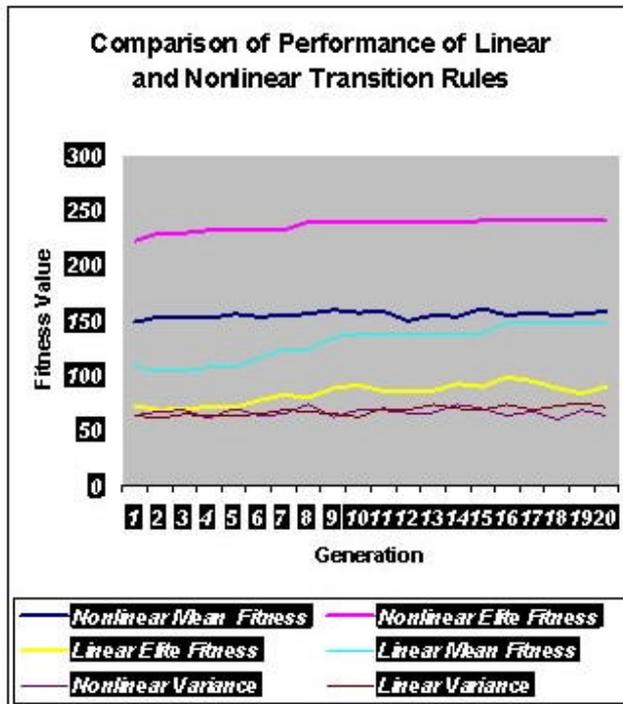
Figure 4. Fitness performance of the models based on spatial metrics with and without a nonlinear function and with and without using elite fitness.

However, the variances from both linear and nonlinear simulations are indistinguishable. That suggests that the inherent stochasticity is independent of the choice of transition rules and is difficult to be tackled by evolutionary optimization techniques alone. How best to apply optimization methods for such a stochastic system needs further study. The variance may also be due to the spatial metrics chosen to describe the landscape. Perhaps these cannot exactly quantify the internal complexity within this system, which causes the "moving target" problem for optimization.

The results show that it is difficult to distinguish different transition rules in terms of different sets of weights. The variances in results are very large and the Monte Carlo approach also cannot reduce them. The underlying cause has been explored and it seems self-organized complexity may be responsible for such phenomenon; the nearly steady state after a few generations supports this idea. The endogenous self-organization process makes the system insensitive to quantitative changes of internal mechanisms, which is consistent with universality in self-organized complex systems. We suggest that the quantitative modification of transition rules may not be a good way of searching for appropriate mechanisms.

3. Analysis with polynomial feedback
3.1 Methods B
Close examination of the spatial patterns generated using the above method reveals that there are still significantly fewer big patches and more small patches than the observed

landscape. To further explore the usefulness of GA in a self-organizing system, we turn to a description of establishment and mortality probabilities that take into account both positive and negative feedbacks that are linear or nonlinear. Taking the Weierstrass Approximation Theorem, i.e., that a continuous function on a bounded and close interval can be approximated on that interval by polynomials, we chose polynomials of the form to represent the establishment process:

$$\Pr obability(X) = \sum_{i=1}^{N} w_i X^i$$

X is the average size of patches in neighborhood for establishment. Mortality is a function of tree age and positive feedback of neighboring trees.

$$Mortality(Z) = \left(1 - \frac{1}{1 + a\,e^{-z}}\right) \Big/ bM$$

Z is the age of the target tree; a and b are coefficients.

We then rewrote the GA program so that the GA searches the function space of the polynomials through the searching for the parameter space. According to our previous work on exploring Self-Organized Complexity (Turcotte & Rundle 2002) in alpine treeline (Zeng & Malanson, in review), we used the difference between the power law exponents of frequency-size distribution of observed and simulated landscapes as the fitness function. In this way we can see if more similar spatial patterns can be generated and identify the relative importance of the two processes. The rationale is that, even if we do not know what the specific functions for establishments and mortality are, the GA will search for the best results of the two processes, which can then be compared. This is different from and a conceptual improvement over the Monte Carlo approach in that we have used our theoretical understanding of the system and optimization power of genetic algorithms to search for the best performance to see if better spatial patterns can be obtained and identify the relative importance.

We ran two simulations. One has establishment as a random process and uses the mortality with feedback described above, while the second reverses this comparison and has mortality as a random process and establishment with feedback.

3.2 Results B
This approach produces useful insights into the system. Figure 5 shows the dynamics of the mean and elite fitness values from the two simulations. It shows that both mean and elite fitness values with random establishment and spatially explicit mortality are significantly higher than those from the reverse case. This result indicates that mortality process is the most essential process and more important than the establishment process in generating the observed spatial patterns. Establishment without spatial positive feedback in mortality does not generate any significant patches. The highest fitness would be when both processes include spatially explicit feedback. The differences among the runs are, however, informative. These results indicate that ongoing mortality,

tempered by the ameliorated environment of nearby trees, is the part of the process that most influences the spatial pattern of an advancing treeline.
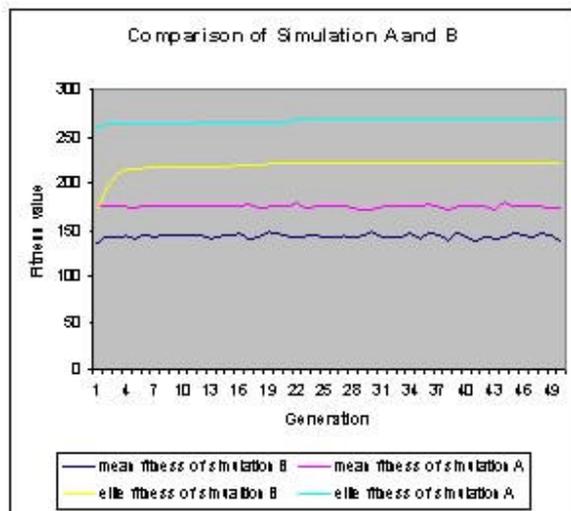

Figure 5. Fitness performance of the models based on polynomials with either establishment or mortality random and the other with feedback and with and without using elite fitness.

The GA is not able to sharpen our understanding of stem behavior by focusing in on particular values for the coefficient in the polynomials. It does, however, set some boundaries. At lower levels of establishment or higher levels of mortality, trees do not advance at all, while when the process is reversed the advance does not allow the development of spatial pattern. These results in general are trivial, but this is where quantitative specificity may be most useful.

4. Discussion and conclusions

This research supports the theory that feedback is important in structuring the tree establishment pattern found at alpine treeline. Further research is necessary to determine the relative importance of various components of feedback, including directional forces. The overall balance between negative and positive feedbacks and the roles of establishment and mortality need additional investigation. Trees or krummholz create a negative feedback by cooling soil temperatures below and immediately adjacent to themselves (Korner 1998), but nearby create positive feedbacks by increasing canopy temperature and mitigating wind. The balance of these two forces in creating pattern varies with the wind and the canopy structure, which depend on each other. The two will affect establishment and seedling mortality in different directions.

As a result of these simulations, we tested a theoretical model of treeline advance (Zeng & Malanson, in review). Here, feedback modifies establishment probability as a logistic function that captures the positive neighborhood feedback or facilitation that is reduced by negative neighborhood feedback when too many trees are nearby (and shading

dominates). Tree mortality is a function of tree age and the number of trees in its immediate neighborhood. We analyzed the relation between the time series of landscape potential and that of the exponent of frequency distribution of patch size (Figure 6). When lagged by 5 years, the exponent of frequency distribution of patch size is negatively correlated with landscape potential (-0.4186; p<0.001), suggesting that spatial structure may change ahead of landscape potential and exert a positive impact on the latter.
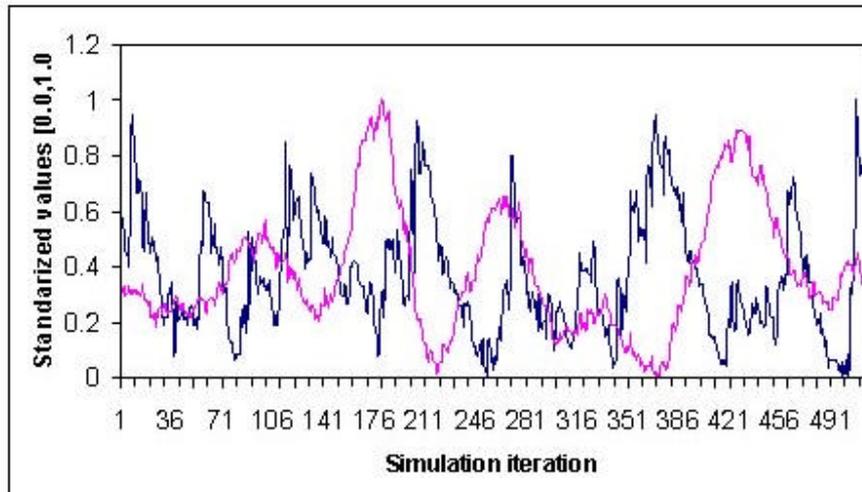

Figure 6. The frequency distribution of patch size is a useful indicator of the spatial structure of alpine treeline and it is correlated with the potential of the landscape to support new advance based on positive feedback.

The rules specify the nonlinear positive and negative feedbacks between pattern and process at a local scale, but these local interactions diffuse stochastically across the landscape. Global patterns and a linear correlation between global pattern and process with temporal and spatial fractal scaling properties emerge from these dispersed, localized interactions through a cross-scale self-organizing process.

Fractal dynamics driven by dispersed, localized pattern-process interactions collectively are capable of self-organizing to create long-term, landscape scale correlations. When new tree patches form or existing patches expand, localized pattern-process interactions are established across the landscape. The spatial pattern of trees increases environmental heterogeneity and establishes landscape connectivity at various scales diffusively. These interactions cause fluctuations at small and medium scales at most times, but occasionally the interactions create connectivity at the landscape scale through long-term and large-scale feedbacks in a coalescence of patches. A second-order phase transition between high and low fractal states is triggered that collapses the environmental variability and connectivity extends across the landscape. We propose that self-organized complexity (sensu Turcotte and Rundle (2002), not self-organized criticality sensu Bak et al. (1998)) is an organizing theory that covers this pattern-process interaction.

The GA-CA simulations presented here support the interpretation of self-organization in the alpine treeline ecotone. Moreover, in self-organizing systems, EC has the potential to help understand the broad form of the functions describing system behavior. Within the form of the functions, however, the exact quantities of coefficients may not be honed in on by EC, but neither may they be so important. The nature of self-organization and the basic optimization purpose of EC are not harmonious. The particular coefficients that we use in our model are not very important, and the overall system behavior is robust across some range. Future research will be aimed at assessing the range of the coefficients, the extent to which the range can be explained in terms of biophysical processes, and, finally, whether EC can be used to determine the general form of the function if not specific coefficient values.

5. Acknowledgements

6. References

BAK, P., TANG, C. & WIESENFELD, K. 1988. Self-organized criticality. *Physical Review A,* 38, 364-374.

BEKKER M.F., MALANSON, G.P., ALFTINE, K.J. & CAIRNS, D.M. 2001. Feedback and pattern in computer simulations of the alpine treeline ecotone. In Millington, A.C., Walsh, S.J. & Osborne, P.E., eds. *GIS and Remote Sensing Applications in Biogeography and Ecology*. Kluwer, Dordrecht, pp. 123-138.

BENNETT, D.A., WADE, G.A. & ARMSTRONG M.P. 1999. Exploring the solution space of semi-structured geographical problems using genetic algorithms. *Transactions in GIS,* 3, 89-109.

BROWN, D.G., CAIRNS, D.M., MALANSON, G.P., WALSH, S.J. & BUTLER, D.R. 1994. Remote sensing and GIS techniques for spatial and biophysical analyses of alpine treeline through process and empirical models. In W.K. Michener, S. Stafford, and J. Brunt, eds. *Environmental Information Management and Analysis: Ecosystem to Global Scales*. Taylor and Francis, Philadelphia, pp. 453-481.

CAIRNS, D.M. & MALANSON, G.P. 1998. Environmental variables influencing carbon balance at the alpine treeline ecotone: a modeling approach. *Journal of Vegetation Science,* 9, 679-692.

DIPLOCK, G. 1998. Building new spatial interaction models by using genetic programming and a supercomputer. *Environment and Planning A,* 30, 1893-1904.

DIPLOCK, G. AND OPENSHAW, S. 1996. Using simple genetic algorithms to calibrate spatial interaction models. *Geographical Analysis,* 28, 262-279.

GISKE J, HUSE G, FIKSEN O. 1998. Modelling spatial dynamics of fish. *Reviews in Fish Biology and Fisheries,* 8, 57-91.

HOCH, G. & KORNER C. 2003. The carbon charging of pines at the climatic treeline: a global comparison. *Oecologia,* 135, 10-21.

HOSAGE, C.M. & GOODCHILD, M.F. 1986. Discrete space location-allocation solutions from genetic algorithms. *Annals of Operations Research,* 6, 35-46.

HRABER PT, MILNE BT. 1997. Community assembly in a model ecosystem. *Ecological Modelling,* 103, 267-285.

KORNER, C. 1998. A reassessment of high elevation treeline positions and their explanation. *Oecologia,* 115, 445-459.

MALANSON, G.P. 1997. Effects of feedbacks and seed rain on ecotone patterns. *Landscape Ecology,* 12, 27-38.

MALANSON, G.P., XIAO, N. & ALFTINE, K.J. 2001. A simulation test of the resource averaging hypothesis of ecotone formation. *Journal of Vegetation Science,* 12, 743-748.

MANSON, S.M. 2003. The SYPR integrative assessment model: complexity in development. In: *Final Frontiers: Understanding Land Change in the Southern Yucatan Peninsular Region.* Eds. B. L. Turner II, D. Foster and J. Geoghegan. Claredon Press, Oxford, in press.

MATTHEWS, K.B., SIBAALD, A.R. AND CRAW, S. 1999. Implementation of a spatial decision support system for rural land use planning: integrating geographic information system and environmental models with search and optimisation algorithms. *Computers and Electronics in Agriculture,* 23, 9-26.

MCGARIGAL, K. AND MARKS, B.J. 1995. *FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure.* Oregon State University, Forest Science Department, Corvallis, OR.

MITCHELL, M. 1996. *An Introduction to Genetic Algorithms.* MIT Press, Cambridge, MA.

NOBLE, I.R. 1993. A model of the responses of ecotones to climate change. *Ecological Applications,* 3, 396-403.

NOEVER DA, BRITTAIN A, MATSOS HC, BASKARAN, S. AND OBENHUBER, D. 1996. The effects of variable biome distribution on global climate. *Biosystems,* 39:135-141.

OPENSHAW, S. 1992. Some suggestions concerning the development of artificial intelligence tools for spatial modeling and analysis in GIS. *Annals of Regional Science,* 26, 35-51.

OPENSHAW, S. 1995. Developing automated and smart pattern exploration tools for geographical information systems applications. *Statistician,* 44, 3-16.

OPENSHAW, S. 1998. Neural network, genetic, and fuzzy logic models of spatial interaction. *Environment and Planning A,* 30, 1857-1872.

PEREIRA, A.G. 1996. Generating alternative routes by multicriteria evaluation and a genetic algorithm. *Environment and Planning B,* 23, 711-720.

RISSER, P.G. 1995. The status of the science examining ecotones. *BioScience,* 45:318-325.

STEVENS, G.C. & FOX, J.F. 1991. The causes of treeline. *Annual Review Ecology & Systematics,* 22,177-191.

TURCOTTE, D.L. & RUNDLE, J.B. 2002. Self-organized complexity in the physical, biological, and social sciences. *Proceedings of the National Academy of Sciences, USA,* 99, 2463-2465.

TURNER, M.G. 1989. Landscape ecology – the effect of pattern on process. *Annual Review of Ecology & Systematics,* 20, 171-197.

TURTON AND OPENSHAW, S. 1998. High-performance computing and geography; development, issues, and case studies. *Environment and Planning A,* 30, 1839-1856.

WILSON, J.B. & AGNEW, A.D.Q. 1992. Positive-feedback switches in plant communities. *Advances in Ecological Research,* 23, 263-336.

Xiao, N.C., Bennett, D.A. & Armstrong, M.P. 2002. Using evolutionary algorithms to generate alternatives for multiobjective site-search problems. *Environment and Planning A,* 34, 639-656.