

University of Denver

Digital Commons @ DU

Electronic Theses and Dissertations

Graduate Studies

1-1-2017

Quantifying and Simulating Fine-Scale Spatial Patterns in Plant Populations

Darin Schulte
University of Denver

Follow this and additional works at: <https://digitalcommons.du.edu/etd>



Part of the [Ecology and Evolutionary Biology Commons](#)

Recommended Citation

Schulte, Darin, "Quantifying and Simulating Fine-Scale Spatial Patterns in Plant Populations" (2017).
Electronic Theses and Dissertations. 1311.
<https://digitalcommons.du.edu/etd/1311>

This Dissertation is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu, dig-commons@du.edu.

Quantifying and Simulating Fine-Scale Spatial Patterns in Plant Populations

Abstract

Plant ecology as a discipline has increasingly acknowledged the importance of fine-scale spatial patterns in developing our understanding of community/population dynamics. These spatial patterns are largely determined by direct and indirect interactions between plants and their immediate neighbors. Such interactions thus play an important role in the structure and function of plant communities. Study of these types of local interactions has greatly benefitted from simulation based approaches. One such simulation method, agent-based modeling, has increasingly been identified as a useful tool for simulating these fine-scale interactions, and for investigating theoretical descriptions of underlying processes. Similarly, statistical techniques aimed at quantifying and comparing spatial patterns across a range of spatial scales are an active area of research, and have served to greatly increase our understanding of plant communities.

Typically underlying these statistical and simulation methods, is a simplified representation of individuals as grid cells, points or circles. Recent work has illustrated that fine-scale spatial patterns may be misrepresented when such assumptions are made, and researchers are increasingly developing methods that do not rely on such geometric simplifications. The work presented in this dissertation shows that important inter-annual changes in spatial pattern occur in *Bouteloua gracilis* populations at multiple, sub-meter scales, reinforcing the belief that local interactions are important factors in community structure (Chapter 1). It further illustrates that the very notion of 'local' is markedly influenced by the particular data type chosen to represent individuals, and that geometric simplifications change how neighborhood composition is described (Chapter 2). The dissertation presents an extension to traditional point-pattern analysis techniques that allows for more complex geometries in describing randomness, clustering and/or regularity in spatial patterns down to the scale of individual plants (Chapter 3). The proposed method extends a recent advance in the literature for quantifying spatial patterns in polygon data consisting of irregularly shaped objects by considering the physical space occupied by competing individuals, rather than simply the density of neighbors. This provides a useful metric of competition intensity experienced by individuals within a population. Finally, this dissertation presents a proof-of-concept agent-based model that extends previous models by allowing individual plants to respond to local conditions by dynamically changing size and shape (Chapter 4).

Document Type

Dissertation

Degree Name

Ph.D.

Department

Biological Sciences

First Advisor

Anna Sher, Ph.D.

Second Advisor

Robin Tinghitella

Third Advisor

Nathan Sturtevant

Keywords

Agent-based modeling, Point-pattern analysis, Polygons, Simulation, Spatial dynamics, Spatial patterns

Subject Categories

Ecology and Evolutionary Biology

Publication Statement

Copyright is held by the author. User is responsible for all copyright compliance.

Quantifying and Simulating Fine-Scale Spatial Patterns in Plant Populations

A Dissertation

Presented to

the Faculty of Natural Sciences and Mathematics

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

Darin Schulte

June 2017

Advisor: Dr. Anna Sher

Author: Darin Schulte

Title: Quantifying and Simulating Fine-Scale Spatial Patterns in Plant Populations

Advisor: Dr. Anna Sher

Degree Date: June 2017

ABSTRACT

Plant ecology as a discipline has increasingly acknowledged the importance of fine-scale spatial patterns in developing our understanding of community/population dynamics. These spatial patterns are largely determined by direct and indirect interactions between plants and their immediate neighbors. Such interactions thus play an important role in the structure and function of plant communities. Study of these types of local interactions has greatly benefitted from simulation based approaches. One such simulation method, agent-based modeling, has increasingly been identified as a useful tool for simulating these fine-scale interactions, and for investigating theoretical descriptions of underlying processes. Similarly, statistical techniques aimed at quantifying and comparing spatial patterns across a range of spatial scales are an active area of research, and have served to greatly increase our understanding of plant communities.

Typically underlying these statistical and simulation methods, is a simplified representation of individuals as grid cells, points or circles. Recent work has illustrated that fine-scale spatial patterns may be misrepresented when such assumptions are made, and researchers are increasingly developing methods that do not rely on such geometric simplifications. The work presented in this dissertation shows that important inter-annual changes in spatial pattern occur in *Bouteloua gracilis* populations at multiple, sub-meter scales, reinforcing the belief that local interactions are important factors in community structure (Chapter 1). It further illustrates that the very notion of 'local' is markedly influenced by the

particular data type chosen to represent individuals, and that geometric simplifications change how neighborhood composition is described (Chapter 2). The dissertation presents an extension to traditional point-pattern analysis techniques that allows for more complex geometries in describing randomness, clustering and/or regularity in spatial patterns down to the scale of individual plants (Chapter 3). The proposed method extends a recent advance in the literature for quantifying spatial patterns in polygon data consisting of irregularly shaped objects by considering the physical space occupied by competing individuals, rather than simply the density of neighbors. This provides a useful metric of competition intensity experienced by individuals within a population. Finally, this dissertation presents a proof-of-concept agent-based model that extends previous models by allowing individual plants to respond to local conditions by dynamically changing size and shape (Chapter 4).

SUMMARY

The work presented here provides a valuable set of techniques for addressing the challenges of quantifying and simulating fine-scale spatial patterns in systems where individuals vary in size and shape. Results suggest that important shifts in spatial pattern occur across a range of spatial scales from multi-plant groupings down to the scale of the smallest individuals, and that a multi-scale approach to quantifying pattern shifts is able to detect large scale changes in cover while remaining sensitive to changes at very fine-scales (Chapter 1). Neighborhood composition was found to vary markedly depending on the data type used to represent individuals (Chapter 2). The final two chapters of this dissertation present novel extensions to current methods for quantifying local spatial patterns when individuals are mapped as irregular

polygons (Chapter 3), and for agent-based modeling of interacting individuals that vary in size and shape while responding to local competitive interactions (Chapter 4).

TABLE OF CONTENTS

ABSTRACT.....	ii
SUMMARY.....	iii
CHAPTER 1: HIERARCHICAL DECOMPOSITION METHODS FOR FINE-SCALE SPATIAL PATTERN ANALYSIS OF A <i>Bouteloua gracilis</i> POPULATION	1
CHAPTER 2: POINTS, CIRCLES AND POLYGONS: THE INFLUENCE OF DATA TYPE ON DEFINING INDIVIDUAL PLANTS AND NEIGHBORHOODS	24
CHAPTER 3: AN AREA-BASED EXTENSION TO THE PAIR-CORRELATION FUNCTION FOR POLYGON DATA	37
CHAPTER 4: A DYNAMIC VECTOR AGENT MODEL OF PLANT-PLANT AND PLANT-ENVIRONMENT INTERACTIONS: A PROOF OF CONCEPT MODEL	57
LITERATURE CITED	80
APPENDIX A.....	87

LIST OF FIGURES

Figure 1: Venn diagram of mutual information interpretation.	8
Figure 2: Example of mapped <i>Bouteloua gracilis</i> plants. Plot dimensions 1m x 1m.	10
Figure 3: Mutual information and uncertainty coefficient spectra for a two-map comparison of simulated random patterns. Black dots indicate no significant difference between maps at the given partition level, while grey dots indicate significant differences ($P \leq 0.05$).13	13
Figure 4: Mutual information box plot for 14 between-map comparisons of random patterns. .	14
Figure 5: Mutual information and uncertainty coefficient spectra for two examples of between-map comparisons of spatially autocorrelated maps. Black dots indicate non-significant difference between maps at the given partition level. Grey dots indicate significant differences ($P \leq 0.05$)	15
Figure 6: Mutual information and uncertainty coefficient box plots for 14 between-map comparisons of spatial autocorrelation simulations.	15
Figure 7: Mutual information and uncertainty coefficient spectra for between-map comparisons of <i>B. gracilis</i> basal area. Plots represent a single plot comparisons between 1997- 1998 (left) and 1998-1999 (right). Point colors are the same as those for previous figures. ...	16
Figure 8: Mutual information boxplots for <i>B. gracilis</i> data by quadrat.	17
Figure 9: Uncertainty coefficient boxplots for <i>B. gracilis</i> data by quadrat.	17
Figure 10: Examples of large composition changes between years for two quadrats (unun_5b 2001-2002 (left) and unun_5b 2004-2005 (right)), and local maxima in uncertainty coefficient spectra at fine partition levels (right). Note, y-axis scales for uncertainty coefficient plots differ.....	18
Figure 11: Example of continued disproportionate composition in maps for a single plot between 2003 and 2004. Note the reduced range in the y-axis for the uncertainty coefficient plot.	22
Figure 12: Illustration of relationship between polygon area, shape irregularity (Shape Complexity Index) and perimeter from mapped basal areas of <i>Bouteloua gracilis</i> plants in the shortgrass steppe of Colorado, USA (Data source: Chu et al. 2013). For a given polygon size (area), as the shape becomes increasingly irregular (increasing Shape Complexity Index values), perimeters increase.	27
Figure 13: Illustration of polygon centroids (blue crosshatches) and a centroid appearing outside of the intended polygon due to its shape (red crosshatch). Data represent mapped basal areas of <i>Bouteloua gracilis</i> plants. (Data source: Chu et al. 2013).	28
Figure 14: Polygon data for a quadrat exhibiting A) a high degree of size/shape variability, and B) a low degree of size/shape variability. Polygons represent mapped basal areas of individual plants. Quadrat dimensions: 1m x 1m.	31
Figure 15: Examples of the three versions of spatial data created for the high size/shape variability quadrat. Quadrat dimensions: 1m x 1m.	32

Figure 16: Histograms of Shape Complexity Index (SCI) for the high size/shape variance quadrat (left), and the low size/shape variance quadrat (right). Vertical and horizontal scales are equivalent between figures.	33
Figure 17: Estimate of the mean distance to the k^{th} nearest neighbor for (left) the high size/shape variance quadrat, and (right) the low size/shape variance quadrat. Solid lines represent the Local Polynomial Regression Fit (loess). Grey bands indicate 95% CI for estimates of the means.	34
Figure 18: Two potential scenarios for the level of competition experienced by an individual. Both scenarios illustrate individuals at equivalent edge-to-edge distance from each other but with varying degrees of space occupancy and therefore intensity of competition.....	40
Figure 19: Mapped basal areas (solid outlines) and centroids (blue crosses) of <i>Bouteloua gracilis</i> individuals from an ungrazed plot in 1998. Plot dimensions: 1m x 1m.....	43
Figure 20: Point approximation for polygons by centroids (blue crosses). The red box identifies a polygon for which the center of mass was outside the polygon boundary. Green box identifies a multi-part polygon with an outer and inner ring.	49
Figure 21: 3D density plot for point densities of <i>B. gracilis</i> individuals.	50
Figure 22: Estimated Pair Correlation Function $g(r)$ for point pattern data. Confidence envelope (dark grey) generated from 500 randomizations of the point pattern. Point locations were derived from polygon centroids. Red line indicates $g(r)$ under the null hypothesis of CSR. Black line indicates $g(r)$ for the observed field data. Note the y axis is truncated for better visualization.....	51
Figure 23: Estimated area-based polygon pair correlation function. Dark grey bands indicate 95% CI generated from 500 randomizations of the original polygon data set for the null hypothesis of CSR. Black horizontal line derived from the mean of 500 simulations and represents $gArea(r) = 1$	52
Figure 24: Illustration of shape dynamics methods.....	60
Figure 25: Example of mapped basal areas from LTER data set.....	62
Figure 26: Density plot of LTER data set Shape Complexity Index values (excluding the smallest square polygons as described in the text).	75
Figure 27: Density plot of the median SCI values from the final round of parameter estimation simulation runs. Median SCI was calculated for five runs with the same parameter values.	76
Figure 28: Examples of simulation runs with final parameter values.....	77

CHAPTER 1: HIERARCHICAL DECOMPOSITION METHODS FOR FINE-SCALE SPATIAL PATTERN ANALYSIS OF A *Bouteloua gracilis* POPULATION

ABSTRACT

External environmental effects and local biotic interactions play important roles in determining the growth, survival, and reproduction of individual plants. The varying influence of these factors result in observable differences in spatial patterns within plant communities through time and at different spatial resolutions. Observed patterns serve as tools to increase our understanding of the relative contributions of these effects. Quantifying spatial pattern at a single resolution may not adequately capture the varying importance of external and internal drivers as their effects manifest at different scales. Recently developed hierarchical decomposition methods for quantifying spatial pattern across multiple scales have been developed to address this issue, and allow for the comparison of nested thematic and spatial patterns between multiple data sets. To date, however, these methods have only been applied at the region or landscape scales. This chapter explores the utility of hierarchical decomposition methods at fine spatial resolutions in a system where plant-plant interactions are known to be important drivers of community structure. This is accomplished by applying a hierarchical decomposition model developed for the analysis of gridded data, to quantify inter-annual variation in spatial pattern for a dominant grass in the short-grass steppe of Colorado. Results

suggest that important shifts in spatial pattern between years occur across a range of spatial scales.

INTRODUCTION

Interactions between plants and their immediate neighbors play an important role in the growth, survival, and reproduction of individuals. These interactions along with larger scale environmental factors directly and indirectly influence the spatial distribution of individuals throughout a population and/or community. While external drivers tend to influence spatial pattern at a community-wide scale or beyond, those at finer spatial resolutions are additionally a result of interactions between individuals with their immediate neighbors (Silvertown et al. 1992). Changes in spatial pattern at one scale do not necessarily translate linearly to those at another, and as a result, it is valuable to consider spatial pattern across a hierarchy of scales, and not at a single, fixed resolution (Levin 1992; Wu and David 2002; Dale and Fortin 2014).

Spatial data sets from sources such as satellite imagery and aerial photography are increasingly available for ecological research, and serve as valuable tools to address questions regarding patterns in vegetation across multiple scales. Generally, these questions pertain to the proportions of various vegetation types or classes (composition) and their spatial distribution (configuration) (Boots 2003). Multiple techniques for comparing composition/configuration patterns between two data sets (e.g., maps or images) at multiple scales have been developed in landscape ecology and the analysis of remotely sensed data. Differences in proportional cover of mapped categories (such as vegetation versus bare ground) can be compared among maps to determine compositional shifts at the map scale, while changes in category values at specific locations (pixels or cells), through techniques such as map differencing, identify variability at

finer resolutions (Csillag and Boots 2005). Similarly, summaries of measures such as patch size, fragmentation and shape complexity describe differences in spatial configuration at the map scale (Boots and Csillag 2006), while quantifying change in spatial qualities of individual patches (e.g., expansion, shrinking and division) provides information about spatial configuration at a finer-scale (Dale and Fortin 2014).

In semi-arid grasslands, measures of spatial pattern have informed hypotheses regarding disturbance intensities, competition for limiting resources, and facilitation via microclimate amelioration (Herben et al. 2000; Adler, Raff, and Lauenroth 2001; Berger et al. 2008). Plants interact with their immediate neighbors, and studying patterns at relatively fine spatial scales provides information regarding the type and strength of those interactions. However, external drivers can override these local interactions and lead to pattern shifts at coarser spatial resolutions (Aguilera and Lauenroth 1993). Rarely do patterns observed at one spatial resolution directly relate to those at more coarse or fine scales even under the assumption of stationarity (Legendre and Fortin 1989; Purves and Law 2002; Wagner and Fortin 2005).

Work described here, focuses on categorical maps of *Bouteloua gracilis* basal areas from a fourteen-year study in the short-grass steppe of Colorado, USA. To investigate shifts in spatial pattern within a range of relatively fine spatial scales, a hierarchical decomposition approach was used to quantify pattern changes between map pairs. Specifically, sequential pair-wise map comparisons were performed for five permanent plots in which the basal areas of individual plants were mapped annually. The degree of similarity between spatial patterns for each plot in sequential years was quantified, as well as the scales at which observable shifts in spatial

pattern took place. The chapter concludes with a discussion of interpreting pattern change across a range of fine-scale spatial partitions.

INFORMATION-BASED APPROACHES

Claude Shannon introduced the concept of entropy as a measure of the amount of information in a message (Shannon 1948). Considering a string of English letters, Shannon entropy refers to the uncertainty in being able to predict what the next letter in the string will be. Messages with a larger number of letters, that are equally represented, make such predictions more uncertain, and thus have higher entropy. Entropy, as a measure of species diversity (i.e., Shannon-Wiener Diversity; H'), has long been an important tool for ecologists in describing community composition at various scales and among different locations (Legendre and Fortin 1989; Legendre, Borcard, and Peres-Neto 2005; Roe et al. 2012). Given an individual sampled from a community, species diversity indices such as H' describe the relative uncertainty in being able to identify the species of the sampled individual (Jost 2006). The larger the number of species (species richness), and the more evenly represented they are (evenness), the higher the entropy. For this reason, measures of information such as Shannon entropy are viewed as indices of species diversity (Jost 2006).

The application of information-based, or entropy measures has also played a major role in the field of image analysis and spatially explicit simulations. Images or maps containing categorical values directly lend themselves to measures such as Shannon entropy to classify the degree of spatial (or spatio-temporal) complexity or heterogeneity (Parrott 2005). These concepts have been extended to allow for comparisons between maps (or images) and are often employed for tasks such as medical image registration (Pluim, Maintz, and Viergever 2003).

Central to these applications is the concept of mutual information between images, which serves as a general measure of how related two data sets are to each other (Song, Langfelder, and Horvath 2012). Mutual information has been applied in image/map analysis to understand the similarity or difference between areas and/or the same area across time intervals (Rommel and Csillag 2006). In effect, mutual information describes the degree to which knowledge of one pattern can predict another (Pluim, Maintz, and Viergever 2003; Rommel and Csillag 2006). As such, mutual information between a map and itself is maximized.

HIERARCHICAL DECOMPOSITION

While useful for general comparisons between maps, mutual information only describes compositional similarity, and not configuration. The concept has been extended to allow for comparisons of compositional similarity across multiple, nested, scales as a proxy for configuration (Rommel and Csillag 2006). By comparing mutual information between hierarchically nested regions of maps, one can assess not only compositional differences between maps, but also the spatial scales at which significant shifts in composition occur. Such methods are common in image analysis for tasks such as image registration (Pluim, Maintz, and Viergever 2003) and feature detection (Pennekamp and Schtickzelle 2013) and provide a logical and mathematical framework for partitioning space into hierarchically nested areas.

This dissertation applies the hierarchical decomposition model proposed by Rommel & Csillag (2006) for assessing pattern shifts at landscape scales, to test its ability to discern shifts at fine-spatial scales. Rommel & Csillag (2006) developed a flexible model able to incorporate multiple levels of hierarchical structure for each variable. Information across all levels of spatial partitioning provide vectors of values (i.e., spectra) describing changes in categorical

heterogeneity between subsequent partitions or between data sets (Csillag and Boots 2005; Remmel and Csillag 2006). Following their structure, I employed a complete quad-tree method of spatial partitioning in which a map was partitioned into four equal area quadrants, and each quadrant was recursively partitioned down to the level of individual pixels or cells. This method resulted a completely nested hierarchy, represented as a multidimensional array.

A simplified form of the Remmel & Csillag (2006) model was developed in which the only variable with a hierarchical nature referred to the nested spatial partitions. Derivation of mutual information makes use of its close relationship with joint entropy (Russakoff et al. 2004). Joint entropy of two maps A and B across all spatial partitions (Y) is defined as:

$$H(A, B, Y) = - \sum_{i=1}^{4^k} \sum_{a,b} \frac{1}{4^k} p_{AB}(a, b, y_i) \log_2 \left(\frac{1}{4^k} p_{AB}(a, b, y_i) \right) \quad (1)$$

where $y_i \in \{1, 2, 3, \dots, 4^k\}$ and p_{AB} is the joint probability distribution of pixel values for the two maps (a representing pixel values for map A , and b for pixel values in map B). For the case of exactly equivalent maps, joint entropy is minimized, and increases as the pixel values begin to differ between maps (Russakoff et al. 2004).

Mutual information considers joint entropy $H(A, B, Y)$ and the individual map (marginal) entropies $H(A, Y)$ and $H(B, Y)$. Mutual information is thus defined as:

$$I(A, B, Y) = H(A, Y) + H(B, Y) - H(A, B, Y) \quad (2)$$

Equation (2) demonstrates the intuition that mutual information increases as the joint entropy is minimized.

The G^2 test statistic (likelihood-ratio chi-square test) was used to test for significant differences in mutual information between maps (Csillag and Boots 2005; Remmel and Csillag 2006). Such a model construction and statistical test, allow for comparisons between maps across the entire spectra of measured values (i.e., across all levels of spatial partitioning).

MUTUAL INFORMATION AND UNCERTAINTY COEFFICIENT

Mutual information between two maps can be thought of as the extent to which information about one pattern describes the other. The concept is analogous to the intersection between two events in set theory (Pluim, Maintz, and Viergever 2003), and a simple case can be visualized with a Venn diagram (Figure 1). For two events (i.e., maps) A and B, mutual information $I(A, B)$ is represented by the region of overlap between marginal entropies $H(A)$ and $H(B)$. This illustrates that I is symmetric ($I(A, B) = I(B, A)$), and that the more similar a pair of images, the greater the overlap between $H(A)$ and $H(B)$, and thus, the greater the mutual information. Mutual information thereby relates to the amount of information A contains about B , and vice-versa. Quantifying mutual information at nested spatial scales allows for the investigation of resolutions at which the similarity between images shifts.

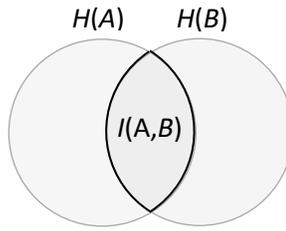


Figure 1: Venn diagram of mutual information interpretation.

Mutual information spectra monotonically increase, and represent absolute gain in information (Remmel and Csillag 2006). The rate of increase (slope) of the spectra describes the relative similarity / difference between categorical distributions (composition) in the pair of maps between partition steps.

The hierarchical decomposition model output also contains the uncertainty coefficient. Uncertainty coefficient values are also a measure of information, but quantify relative information gain rather than absolute information gain (Pluim, Maintz, and Viergever 2003; Remmel and Csillag 2006). It is expressed as a percentage, and defined to be (Csillag and Boots 2005):

$$UC = \frac{I(A,B)}{H(B)} * 100$$

The uncertainty coefficient describes the relative reduction in uncertainty (as a percentage) at each given partition level. As an example, mutual information will always increase between a coarse partition level and the subsequent finer one. Uncertainty coefficient spectra may closely resemble that of mutual information, however, if variation in class composition is high (heterogeneous) at a coarse level and decreases substantially at the subsequent finer level, the

uncertainty coefficient will be relatively large at the finer partition. If they occur, local maxima in the uncertainty coefficient spectra indicate important spatial scales when comparing maps. Depending on the mapped data and the range of partitions used, multiple local maxima are possible, indicating important pattern changes at multiple scales.

METHODS

STUDY AREA AND DATA DESCRIPTION

Data used in this analysis consisted of shapefiles in which polygons represent the mapped basal areas of individual plants. Data were collected at the Shortgrass Steppe Long-Term Ecological Research (SGS-LTER) site in Northern Colorado (40°49'N latitude, 107°47'W longitude) from 1997 through 2010 (Chu et al. 2013) as a part of a long-term grazing experiment. The study consisted of four blocks each with a different grazing treatment, however, for the work described here, only the block which was ungrazed prior to, and throughout the entire study period was considered. Within this treatment, five permanent 1²m plots were established, and the basal areas of all plants within each plot were mapped annually in July (Figure 2). Field data were collected with a pantograph (Chu et al. 2013), and digitized maps are publicly available as shapefiles through the LTER data portal (LTER, 2009).

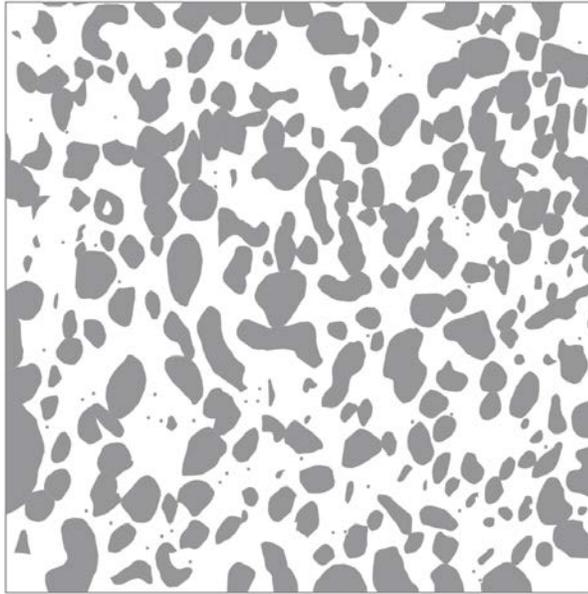


Figure 2: Example of mapped *Bouteloua gracilis* plants. Plot dimensions 1m x 1m.

The dominant vegetation in the shortgrass steppe are long-lived C₄ grasses (Lauenroth and Burke 2008). *Bouteloua gracilis* and *Buchloe dactyloides*, both perennial caespitose (i.e., growing in clumps or tufts) grasses, constitute the majority of basal coverage in the data set. Due to its marked dominance within the data set, only *B. gracilis* was considered for this study.

DATA PROCESSING AND ANALYSIS

Hierarchical decomposition was performed on the five plots within the ungrazed treatment. For some of these plots, field data for 2000 were not available, and between-map comparisons skip from a 1998-1999 comparison to a 2001-2002 comparison. All shapefiles were filtered to retain only *Bouteloua gracilis* polygons and converted to raster format in QGIS (QGIS 2016) with a grid resolution of 256 X 256 cells. Quadrat boundaries varied slightly between

individual shapefiles, thus standardized extents of (0,1) for both x and y dimensions (in meters) were imposed.

Comparisons between maps were performed in the *hdeco* package (Remmel et al. 2015) in R (R Development Core Team 2016). The current release of *hdeco* requires that images be of the same resolution, and the complete quadtree method further requires that the total cell count for a map be 2^L , with L being a positive integer. To ensure that the smallest individual plants were retained in the rasterization process, and polygon sizes changed as little as possible, we set L equal to 8, creating raster maps of dimension 256 x 256. Raster maps were converted to numeric matrices in R. Mutual information and uncertainty coefficient spectra were derived for all sequential between-map comparisons for individual quadrats throughout the study period (e.g., Map 1 = plot A_{t-1} , Map 2 = plot A_t , where $t \in \{2, 3, 4, \dots, T\}$ and $T =$ *number of surveyed years*).

PATTERN SIMULATION

Before analyzing the mutual information and uncertainty coefficient spectra in the field data, it is useful to consider hierarchical decomposition results from simulated data to demonstrate several descriptive patterns. Mutual information and uncertainty coefficient values for simulated random patterns and patterns exhibiting spatial autocorrelation illustrate several characteristic behaviors of the two spectra (i.e., mutual information and uncertainty coefficient spectra), and serve as a reference for understanding the spectra obtained from the mapped basal areas data set. A first-order conditional autoregressive model was used to simulate a series of 500 binary maps with random spatial distribution of classes, and a separate series of 500 binary maps exhibiting clustering within classes. All simulated maps had a resolution of

256x256 pixels, and were defined to have approximately equal representation for both classes. Fifteen maps were randomly sampled from the simulations for each pattern type to illustrate spectra behaviors for between-map comparisons.

RESULTS

RANDOM PATTERNS

Between map comparisons for simulated random patterns exhibited no spatial autocorrelation between the two classes. At coarse partition levels, mutual information and uncertainty coefficient values were low, and remained relatively low until fine partition levels were reached (Figure 3). At the finer resolutions, both spectra began to increase exponentially. Between-step differences in either metric were rarely significant at coarser partition levels, typically becoming so at the third or fourth step (Figure 3).

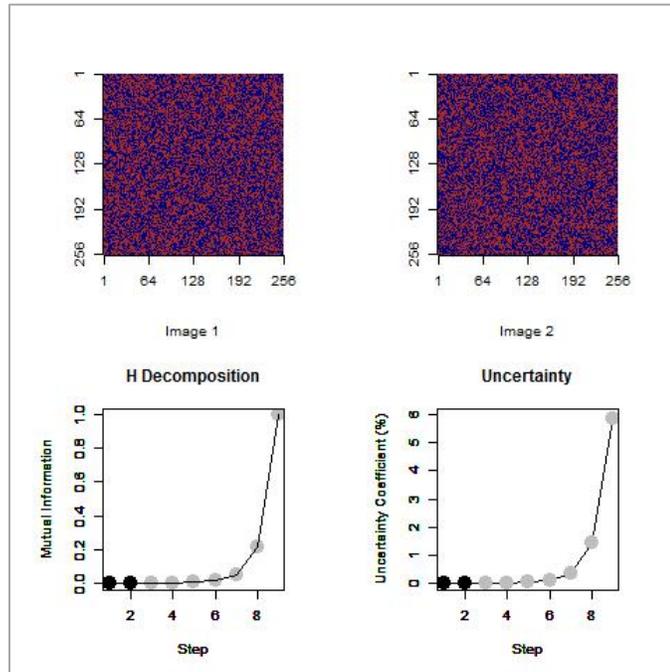


Figure 3: Mutual information and uncertainty coefficient spectra for a two-map comparison of simulated random patterns. Black dots indicate no significant difference between maps at the given partition level, while grey dots indicate significant differences ($P \leq 0.05$).

Variation in mutual information (Figure 4) and uncertainty coefficient values (not shown) remained low throughout all partition levels.

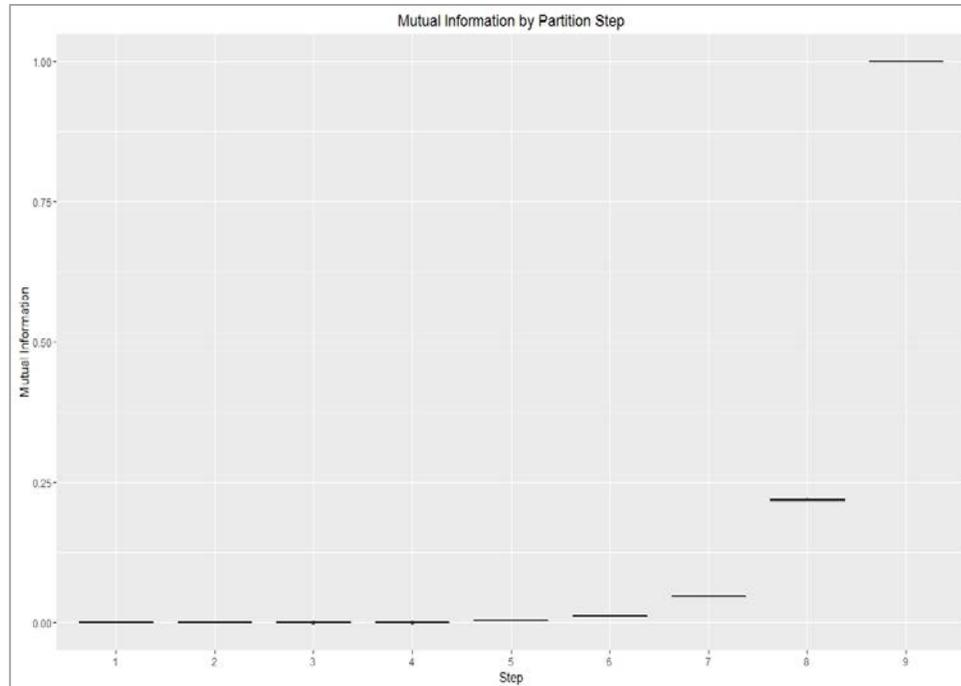


Figure 4: Mutual information box plot for 14 between-map comparisons of random patterns.

AUTOCORRELATION IN PATTERNS

Simulated patterns exhibiting autocorrelation produced maps with increased clustering among classes relative to simulated random patterns. Within-class autocorrelation led to the development of relatively homogenous patches at varying spatial scales in all maps (Figure 5). Autocorrelation was limited to North-South and East-West directions (no diagonal autocorrelation between cell values) and equally weighted between directions. As a result, patches that developed were randomly distributed throughout the maps, and exhibited no directional trend. Mutual information spectra increased at earlier (coarser) partition steps than those for random patterns, and tended towards more linear forms (Figure 6). Uncertainty coefficient spectra also displayed greater increases at coarser partition steps relative to random pattern comparisons. Greater increases between steps were observed at coarse and fine levels,

while more moderate shifts occurred at intermediate steps (Figure 5). Within partition variation in both metrics at all but the coarsest and finest partition levels increased relative to that observed in the random pattern comparisons (Figure 6).

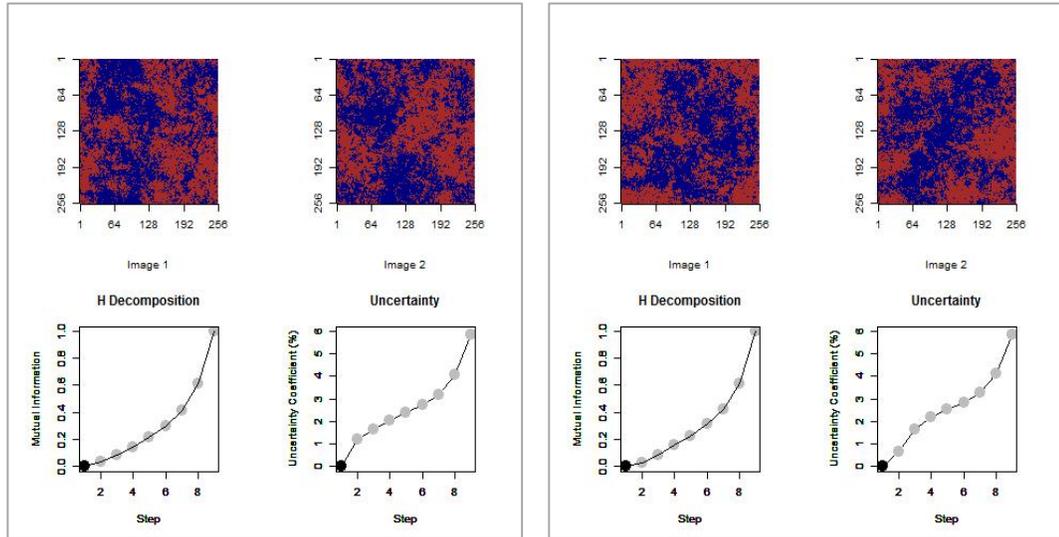


Figure 5: Mutual information and uncertainty coefficient spectra for two examples of between-map comparisons of spatially autocorrelated maps. Black dots indicate non-significant difference between maps at the given partition level. Grey dots indicate significant differences ($P \leq 0.05$).

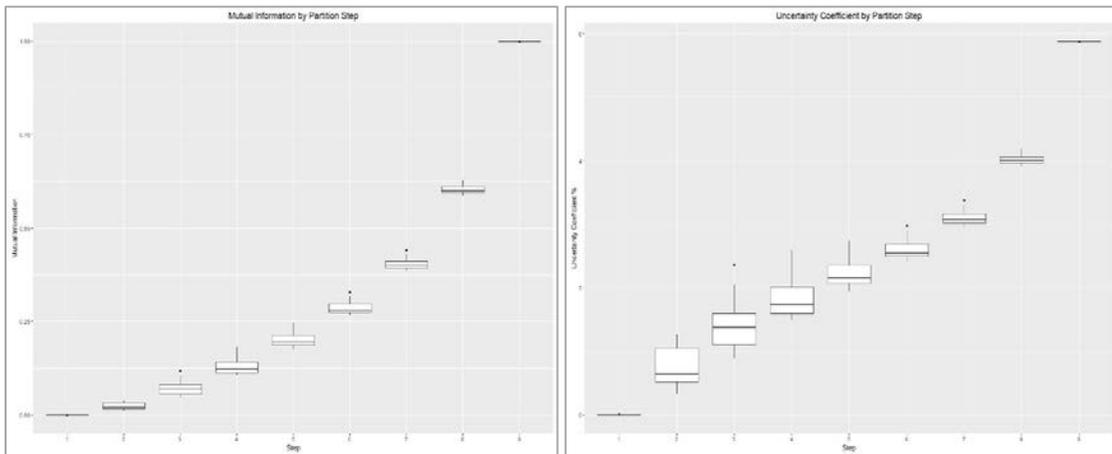


Figure 6: Mutual information and uncertainty coefficient box plots for 14 between-map comparisons of spatial autocorrelation simulations.

PATTERNS IN *Bouteloua gracilis* BASAL AREA

Mutual information for sequential between-map comparisons in the LTER data set suggested a general trend of significant differences between partition levels at most (typically all) steps with large shifts in mutual information between subsequent steps generally occurring at intermediate partition levels (Figure 7).

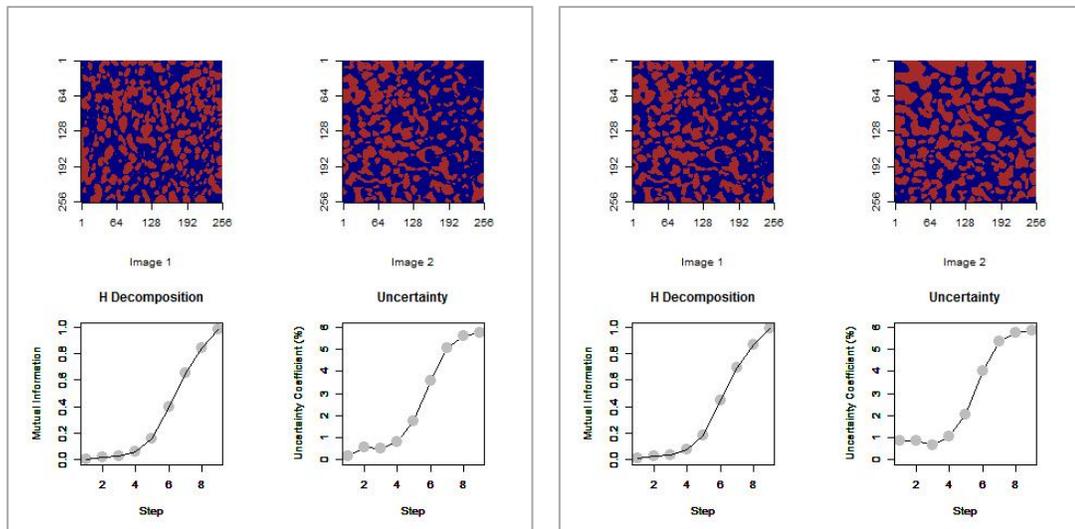


Figure 7: Mutual information and uncertainty coefficient spectra for between-map comparisons of *B. gracilis* basal area. Plots represent a single plot comparison between 1997- 1998 (left) and 1998-1999 (right). Point colors are the same as those for previous figures.

Mutual information spectra tended toward a sigmoidal form for map comparisons in which the two maps did not differ largely in either composition (proportional representation of bare ground / *B. gracilis* cells) or configuration. The maps contain information about discrete objects (plants) and thus exhibit a high degree of autocorrelation in a gridded representation. Relatively large variation in mutual information and uncertainty coefficient values were observed at most partition levels for all quadrats in the data set (Figure 8, Figure 9).

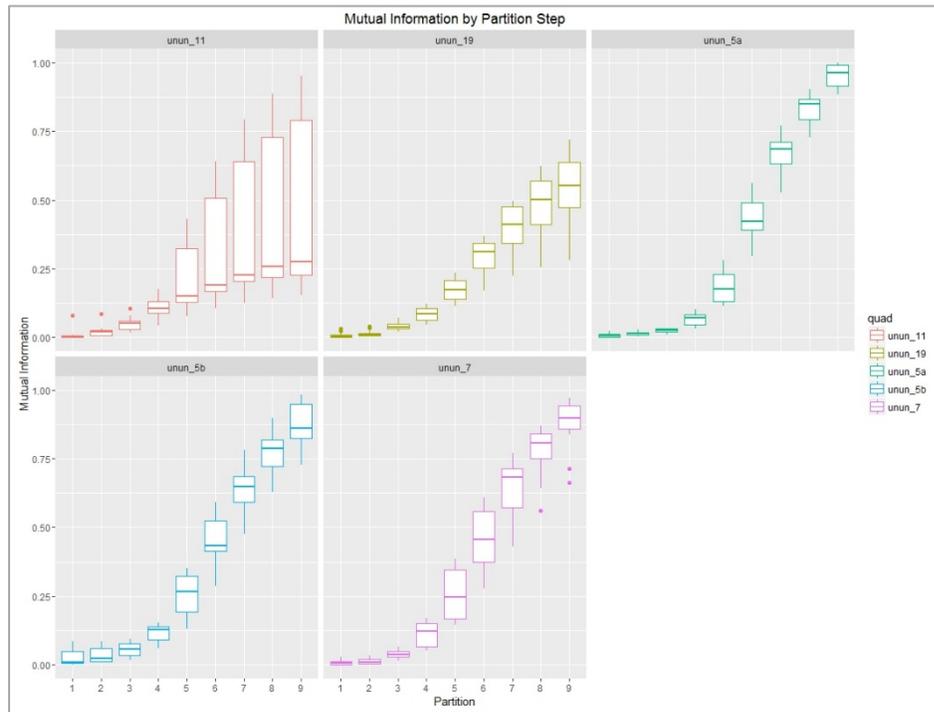


Figure 8: Mutual information boxplots for *B. gracilis* data by quadrat.

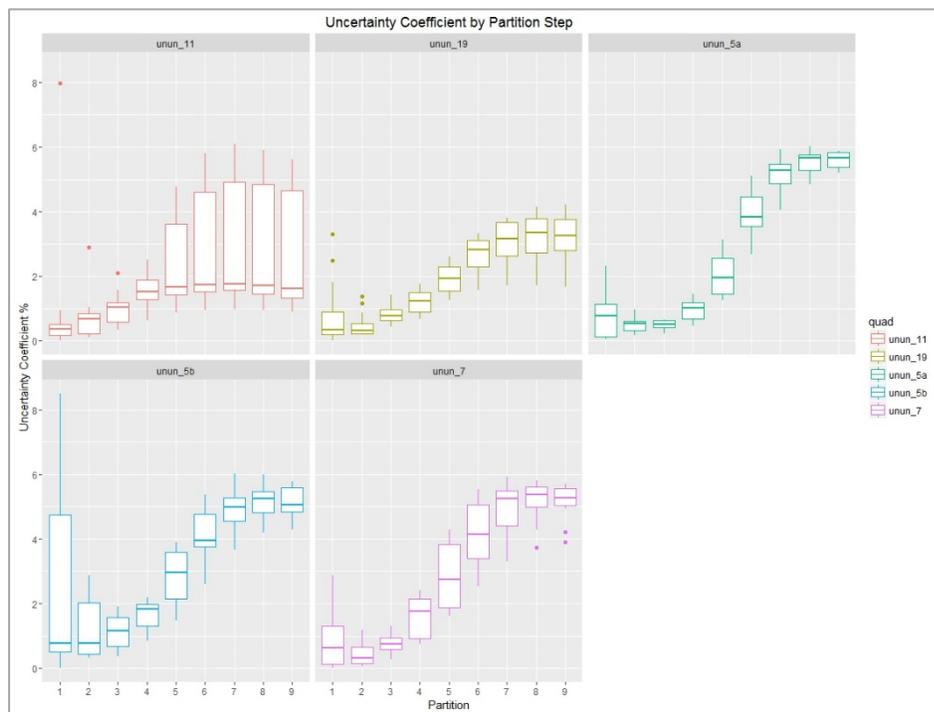


Figure 9: Uncertainty coefficient boxplots for *B. gracilis* data by quadrat.

When composition, configuration, or both markedly varied between maps, uncertainty coefficient spectra diverged from sigmoidal form. In cases where substantial changes in composition occurred between images, uncertainty coefficient values were large at the coarsest partition level, dropped immediately at the following step, and increased throughout the remaining steps (Figure 10). In several comparisons, additional local maxima occurred at fine partition levels.

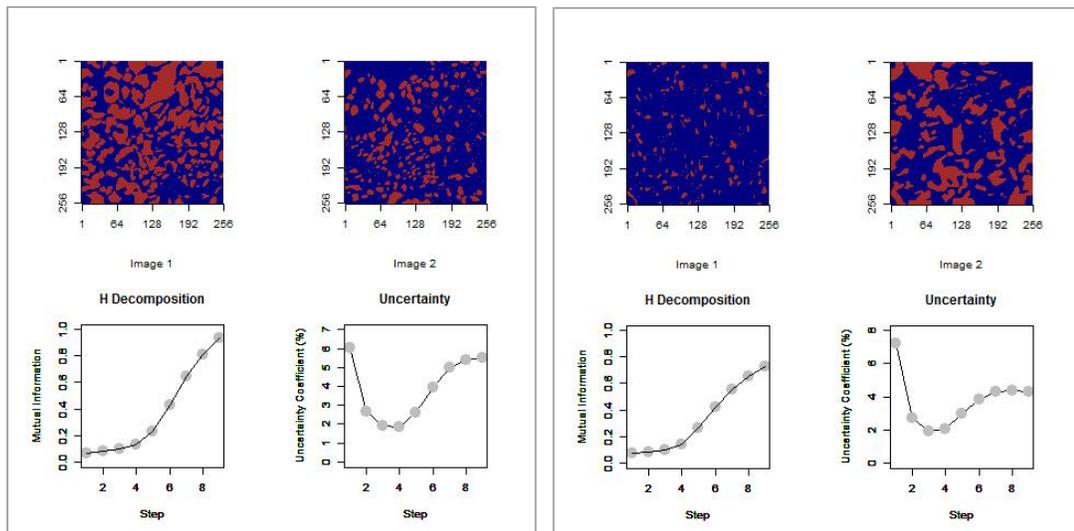


Figure 10: Examples of large composition changes between years for two quadrats (unun_5b 2001-2002 (left) and unun_5b 2004-2005 (right)), and local maxima in uncertainty coefficient spectra at fine partition levels (right). Note, y-axis scales for uncertainty coefficient plots differ.

DISCUSSION

Fine-scale spatial patterns in semi-arid grasslands inform us about the underlying processes of mortality, competition, facilitation, and recruitment. The spatial distribution of individuals can vary dramatically through time with mortality and recruitment, and the space which individuals themselves occupy changes as they grow and respond to their immediate environment. This leads to quantifiable changes in spatial pattern from the scale of the smallest

individuals to populations (Silvertown et al. 1992). Results from a hierarchical analysis of spatial pattern on a long-term field study illustrate significant interannual shifts in spatial pattern across a range of fine-scale partitions, and provide a valuable method for quantifying multi-scale phenomena in plant ecology.

Hierarchical decomposition methods allow for measurement of pattern change across a range of spatial scales by quantifying composition within fully nested subareas as a proxy for spatial configuration. Information-based metrics allow for comparison between patterns in multiple maps or images by defining a common unit of measurement (bits) (Rempel and Csillag 2006). This can provide useful information regarding changes in spatial pattern among different areas, or for the same area at different times by comparing sequential maps for a given area. Simulated patterns that conditionally control the degree of spatial autocorrelation among classes provide intuition into characteristic patterns in mutual information and uncertainty coefficient spectra and serve as a useful reference when interpreting patterns observed in the environment.

RANDOM PATTERNS

Map comparisons of random patterns display several common behaviors regarding mutual information and uncertainty coefficient spectra (Figure 3, Figure 4). As the patterns represent random configurations of pixel values (classes) with approximately equal representation (i.e., 50% class 1, 50% class 2), mutual information remains relatively low until very fine spatial partitions are considered. At increasingly finer partition levels, varying distributions of composition emerge. This result is illustrated by the exponential form of both mutual information and uncertainty coefficient spectra (Figure 3). Such forms are common

when comparing random patterns and suggest that the compared maps are not significantly different until fine spatial scales are considered (Remmel and Csillag 2006). As partitions consider increasingly finer areas, local clustering of classes emerges, and significant differences between the maps are observed. Similarly, variation in either metric at individual partition levels is low (Figure 4).

AUTOCORRELATION IN PATTERNS

As class values exhibit autocorrelation, and depart from randomness, mutual information and uncertainty coefficient values increase at coarser partition levels (Figure 5). Mutual information spectra become less exponential, and generally tend toward more linear forms. Uncertainty coefficient spectra begin to illustrate that relative information gain increases at coarser partition levels, and varies in magnitude throughout the full range of partition levels. This is a result of homogenous patches developing due to within class autocorrelation. Variation in both metrics at each partition level increases relative to that of random patterns as heterogeneity among nested quadrats varies due to local aggregation of pixel values (Figure 6).

PATTERNS IN *Bouteloua gracilis* BASAL AREA

Patterns observed in mapped basal areas of *Bouteloua gracilis* represent a higher degree of autocorrelation within categories (bare ground and *B. gracilis*) than seen in the simulated patterns primarily due to two reasons. First, groups of pixels refer to individual objects (plants) and form distinct clusters thus exhibiting a high degree of spatial autocorrelation in class values. Second, the simulated patterns with autocorrelation were constructed with approximately equal representation of classes. In the *B. gracilis* data, the two classes are rarely represented equally at the global (plot) level, and the composition changes (at

times substantially) from one year to the next. For map pairs with relatively small differences in composition, mutual information and uncertainty coefficient spectra tended toward a sigmoidal form (Figure 7). Such a form for uncertainty coefficient spectra describe the relative similarity between maps at coarse partition levels, and increasing differences at finer scales. This behavior depicts the lack of coarser scale changes between images, while capturing the shifts in composition at finer partition levels. This is a common behavior as recruitment of new (small) individuals occurs between years.

Within each of the five plots, a relatively high degree of variation in mutual information and uncertainty coefficient values was observed at most partition levels (Figure 8, Figure 9). A portion of this variation was due to changes in class proportions and the autocorrelated nature of the mapped data as described earlier. For several mapped quadrats, many individuals disappeared from one year to the next. Disturbance events at or beyond the quadrat extent such as drought or fire, or the establishment of a competitive species, could result in such shifts. Their effects are twofold. First, the drastic change in composition from the first map to the second map result in large uncertainty coefficient values at the coarsest partition level (Figure 10), demonstrating the coarse scale differences between maps. Second, for several years following the shift, classes (bare ground and *B. gracilis*) were disproportionately represented (i.e., there are far more bare ground pixels). Disproportionate representation of classes restricts the range of the uncertainty coefficient to lower values (Figure 11). If *B. gracilis* recruitment was low in the following years, composition remained disproportionate. This effect can be seen in lower values for uncertainty coefficient spectra for plots where this is the case relative to plots where no large-scale shifts in composition occurred or *B. gracilis* reestablished (Figure 9).

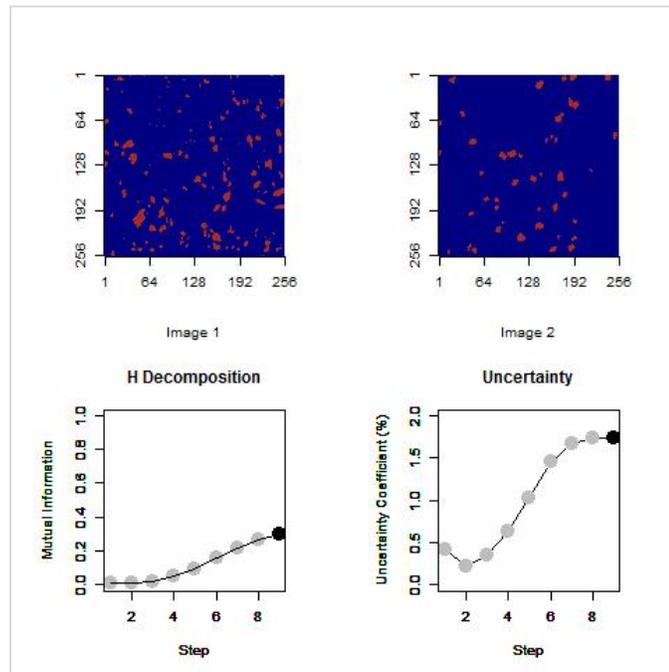


Figure 11: Example of continued disproportionate composition in maps for a single plot between 2003 and 2004. Note the reduced range in the y-axis for the uncertainty coefficient plot.

CONCLUSIONS

Hierarchical decomposition methods provide a useful tool for detecting important shifts in composition and configuration at relatively fine spatial scales. Such a technique may prove useful for testing hypotheses concerning global and local processes driving observed spatial patterns in plant populations. The methods described here can readily be extended to multi-species assemblages, comparisons of more than two maps, and a hierarchy of external factors defining groups of maps.

ACKNOWLEDGEMENTS

Data sets were provided by the Shortgrass Steppe Long Term Ecological Research group, a partnership between Colorado State University, United States Department of Agriculture, Agricultural Research Service, and the U.S. Forest Service Pawnee National Grassland. Significant funding for these data was provided by the National Science Foundation Long Term Ecological Research program (NSF Grant Number DEB-1027319 and 0823405).

CHAPTER 2: POINTS, CIRCLES AND POLYGONS: THE INFLUENCE OF DATA TYPE ON DEFINING INDIVIDUAL PLANTS AND NEIGHBORHOODS

ABSTRACT

Fine-scale spatial patterns in plant communities are often largely a result of interactions between individuals and their immediate neighbors. When studying these patterns, the type of data used to represent individuals can have considerable impact on our definition of local neighborhoods and resulting interpretations of plant-plant interactions. In areas where the size and shape of individuals varies to a large degree, data types that reduce the spatial extent of individuals to point locations or simplified geometries (e.g., circles) influence interplant distances, and may misrepresent the number of individuals with which a focal plant interacts. This chapter discusses several possible vector-based data types common in spatial data sets and illustrates the effect of defining individuals as points, circles, and complex polygons on neighborhood representation. Data from a long-term study mapping basal areas of individual plants in the shortgrass steppe of Colorado, USA are presented as a case study to demonstrate scenarios in which the choice of data type can have varying influence when quantifying neighborhood structure.

INTRODUCTION

Interactions between herbaceous plants and their immediate neighbors play an important role in the growth, survival, and reproductive success of individuals. These interactions, at the spatial scale of several centimeters, have been shown to be important drivers of overall population / community patterns (Purves and Law 2002; Benot et al. 2013). A '*plant's-eye view*' approach (Turkington and Harper 1979) posits that plants respond more to biotic and abiotic conditions within relatively short distances than community-wide general conditions as described by a *mean-field assumption* (Purves and Law 2002; Bolker, Pacala, and Neuhauser 2003). These interactions with, and responses to, local conditions may lead to markedly different spatial patterns at fine versus coarse spatial resolutions. When studying spatial patterns down to the scale of individuals, the type of data used to represent individuals can be a particularly important decision.

Neighborhood composition (i.e., the frequency and density of individuals a plant may interact with) and the spatial arrangement of neighboring individuals have been shown to be equally important in determining outcomes of interactions such as competition (Silvertown et al. 1992). The spatial extent of this neighborhood strongly depends on the size and shape of individuals. Proxies such as point locations, circles, or explicitly mapped basal area may be used to define the region from which an individual may extract resources and would thus compete with other individuals. However, the choice of proxy can have important effects on measures of spatial pattern.

As the size of an individual plant increases, so too does the number of expected neighbors within a given search distance r (Aguilera and Lauenroth 1993), and the same

expectation holds for individuals of a fixed area but increasingly irregular shapes. Consider the case where individuals are mapped as polygons. Numerous measures exist for quantifying the regularity / irregularity of shapes (Brinkhoff et al. 1995) with the simplest shape for an individual being a circle. Many such measures used in landscape ecology and forestry are based on some form of a perimeter-to-area ratio (McGarigal and Marks 1995; Perry et al. 2002) with lower values associated with simple geometries (i.e., circles). One such measure, the *Gap Shape Complexity Index* (referred to here as SCI) defined as

$$SCI = Perimeter / \sqrt{4\pi * Area}$$

assigns a complexity value of 1 for circular shapes and increases as the shape becomes increasingly irregular (Getzin, Nuske, and Wiegand 2014).

The SCI illustrates the impact that shape irregularity can have on defining the neighborhood of a polygon. For a polygon of fixed area, as the shape becomes more irregular, the perimeter increases (Figure 12), thus increasing the likelihood of encountering a neighbor at distance r from the focal plant (though this effect is reduced as size continues to increase) (Nuske, Sprauer, and Saborowski 2009).

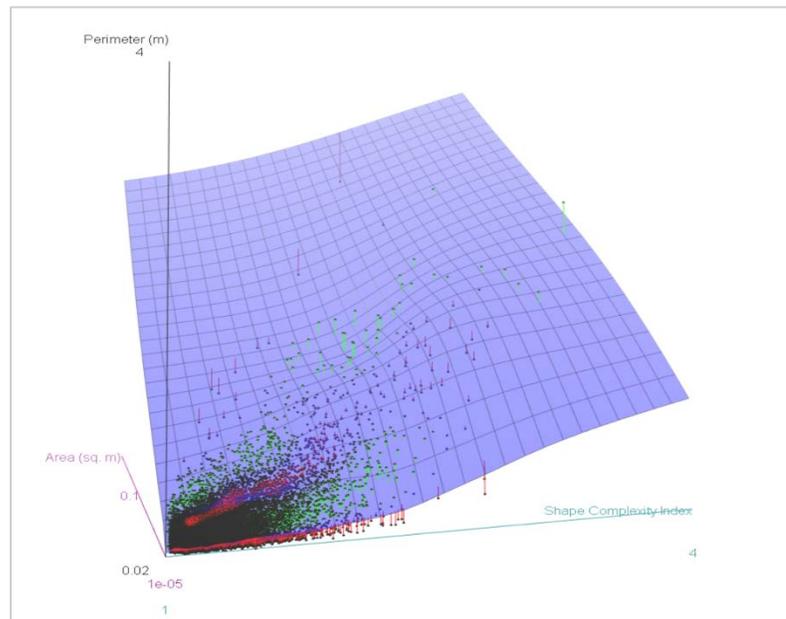


Figure 12: Illustration of relationship between polygon area, shape irregularity (Shape Complexity Index) and perimeter from mapped basal areas of *Bouteloua gracilis* plants in the shortgrass steppe of Colorado, USA (Data source: Chu et al. 2013). For a given polygon size (area), as the shape becomes increasingly irregular (increasing Shape Complexity Index values), perimeters increase.

If individuals are known to interact with each other over short distances, it is important to choose an appropriate data type (e.g., points, circles, polygons) for representing both the individuals themselves and distances between them prior to quantifying a neighborhood (Wiegand et al. 2006; Nuske, Sprauer, and Saborowski 2009; Getzin, Nuske, and Wiegand 2014). When individuals are represented as polygons, distances are commonly measured between polygon centroids (Nuske, Sprauer, and Saborowski 2009). When shapes exhibit a range of size

and shape variation, this may have undesirable effects. Several methods exist for defining the 'center' of a polygon. Depending on the chosen method and the shape of a polygon, the result may not be as intended. An example can be seen for polygons with a 'C' shape (Figure 13) where the centroid is placed outside the intended polygon, or when polygons have 'holes' and the centroid is placed within a hole (i.e., at a location not occupied by the plant).

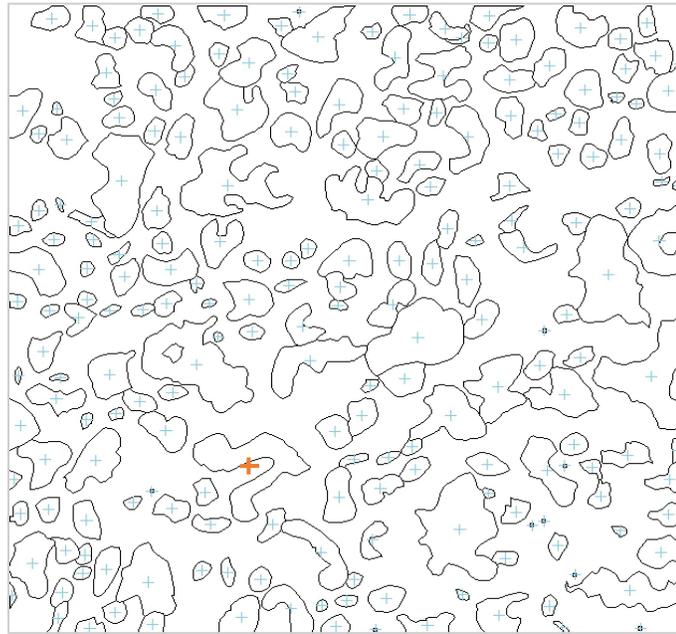


Figure 13: Illustration of polygon centroids (blue crosshatches) and a centroid appearing outside of the intended polygon due to its shape (red crosshatch). Data represent mapped basal areas of *Bouteloua gracilis* plants. (Data source: Chu et al. 2013).

A more desirable distance measure would be the shortest distance between two polygons (edge-to-edge distance). Due to the computational difficulty in calculating such a measure, centroid-to-centroid distance has served as the predominant method for calculating distances between polygons. Recently however, several software programs allow for the calculation of edge-to-edge distances. This presents an opportunity to more accurately define local neighborhoods for data consisting of irregular polygons. To illustrate the disparity between

spatial data choice (points, circles and polygons) and distance measures on neighborhood calculations, we present a brief case study using field data from a long-term field study in the shortgrass steppe of Colorado, USA. The case study shows that the choice of data type used to represent individuals can have profound implications for the perceived extent and intensity of interaction between individuals in a particular study.

The following specific questions are addressed in this chapter; 1) Will the measured intensity of competition differ depending on whether individuals are represented as points, circles or polygons? and 2) does the degree of shape irregularity impact these differences? Data type choice can represent the physical space occupied by an individual differently, and it is expected that these differences will have an influence on distances between individuals, and thus intensity of competition. It is further expected that these differences will vary depending on the degree of distortion introduced with different representations.

METHODS

STUDY AREA AND DATA DESCRIPTION

Case study data come from a 14-year study in Northern Colorado (40°49'N latitude, 107°47'W longitude) and consist of mapped basal areas of individual plants. The data set was collected between 1997 and 2010 as a part of a long-term grazing exclusion study (Chu et al. 2013) at the Shortgrass Steppe Long-Term Ecological Research site (SGS-LTER). Basal areas of individual plants were mapped annually (typically in early July) with a pantograph in permanent 1m² quadrats, and the entire data set was digitized and made publicly available in shapefile format through the LTER data portal (LTER, 2009).

In this region, C₄ grasses represent the dominant plant species, with the perennial grasses *Bouteloua gracilis* and *Buchloe dactyloides* constituting the majority of individuals within the data set. For each sampled quadrat at each time period, two shapefiles were created with individuals represented as either points or polygons depending on their growth form (Chu et al. 2013). The original polygons as drawn in the field were highly variable in size and shape both within and between quadrats, thus providing an opportunity to explore the effects of different data types on defining neighborhood structure.

DATA PROCESSING

Intensity of competition was measured in terms of the mean distance to the k^{th} nearest neighbor, that is, the number of *cm* to the first, second, third nearest neighbors and so forth. The shorter the distance, the greater the competition intensity. Greater rates of increase in the distances with neighbor index shows larger distances between individuals and their neighbors. To determine whether the degree of shape irregularity affects this relationship, data from two quadrats were selected; one that demonstrated a wide range of size and shape variability, and another (in the same grazing treatment) that provided an example of relatively minor variation in shapes among mapped individuals (Figure 14).

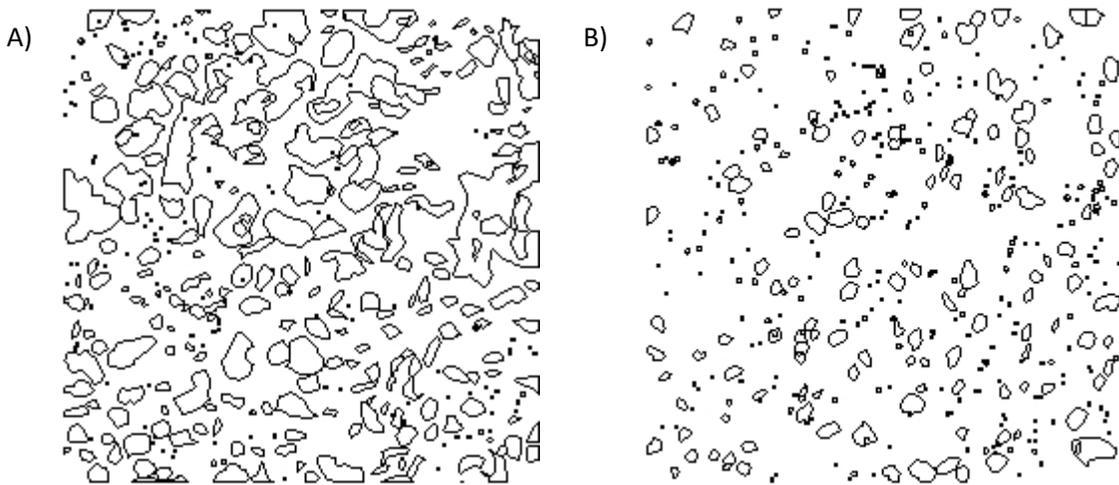


Figure 14: Polygon data for a quadrat exhibiting A) a high degree of size/shape variability, and B) a low degree of size/shape variability. Polygons represent mapped basal areas of individual plants. Quadrat dimensions: 1m x 1m.

As data for each plot consisted of two shapefiles (a point shapefile and a polygon shapefile) we combined them into a single data file. To do so, we created 0.5cm buffers around each individual point in the point shapefile and combined the results with the polygon shapefile into a single data object (a *spatialPolygonsDataFrame* in R). Three versions of the data for each quadrat were created, with the first being the original polygon data (with points converted to circle polygons). The second version converted all polygons to circles centered on their respective polygon centroid locations, with areas equivalent to those in the original (polygon) data, while the third version contained only the centroid locations for each polygon (Figure 15).

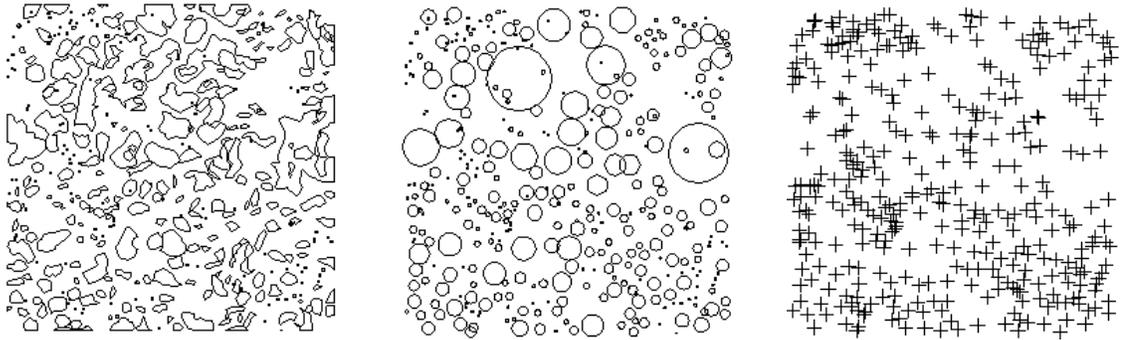


Figure 15: Examples of the three versions of spatial data created for the high size/shape variability quadrat. Quadrat dimensions: 1m x 1m.

When considering distances between individuals, regardless of model, neighborhood estimates will be biased for objects close to quadrat boundaries as the spatial extent of individuals extending beyond the quadrat boundaries is truncated, and the presence of plants beyond the quadrat boundaries is unknown. To address this, an inner buffer of 20cm was constructed from the quadrat bounding box. Objects (polygons, circles, or points) that were entirely outside the inner buffer did not have their neighborhoods calculated, but remained as potential neighbors for polygons that were within the inner buffer (including those that intersected the inner buffer). For each object (polygon, circle or point), distances were calculated to their first 20 neighbors. Distances for the polygon and circle data sets were calculated as shortest edge-to-edge distances, with point-to-point distances being calculated for the points data set. The process of calculating distances for the point data set is equivalent to what would result from calculating centroid-to-centroid distances for the polygon data set. Distributions of distances to the k^{th} nearest neighbors were compared between data types for each quadrat. All data processing and analyses were performed in the R statistical package version 3.3.1 (R Core Team, 2016).

RESULTS

The two quadrats exhibited similarly skewed distributions for the Shape Complexity Index (SCI) (Figure 16) with median values of approximately 1.13 for both, meaning that many individuals were slightly irregular in shape, but not extremely so. Quantile regression with bootstrapped estimation of standard errors, revealed no significant difference in median values ($p = 0.887$), however variance in SCI was significantly different between quadrats (Levene test for homogeneity of variance (median centered), $Df = (1, 308)$, $F = 9.802$, $p = 0.002$).

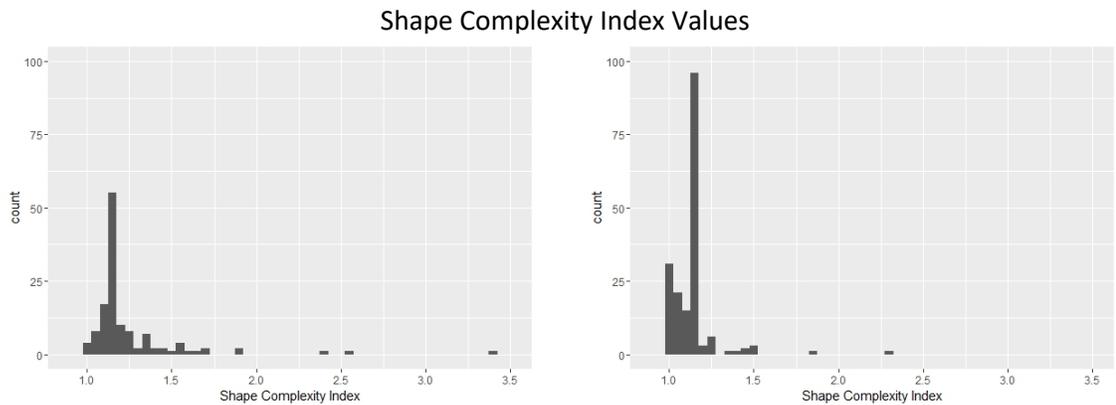


Figure 16: Histograms of Shape Complexity Index (SCI) for the high size/shape variance quadrat (left), and the low size/shape variance quadrat (right). Vertical and horizontal scales are equivalent between figures.

Defining objects as either points, circles or irregular polygons had a marked effect on estimates of the mean distance to the k^{th} nearest neighbor. For each quadrat, the mean distance to the k^{th} nearest neighbor was estimated with a Local Polynomial Regression Fit (loess). Differences between data types were evident in both quadrats, with larger differences occurring in the quadrat with greater size/shape variance than in the quadrat with less variance (Figure 17). In both quadrats, the distance to the k^{th} nearest neighbor was greater (i.e.,

perceived competition intensity was less) for the point representation than it was for either the polygon or circle data types. Similarly, for the quadrat with high size/shape variance, the circle representation had larger mean distances to the neighbors for $k > 3$. The low size/shape variance quadrat had no significant differences between polygon and circle data (Figure 17).

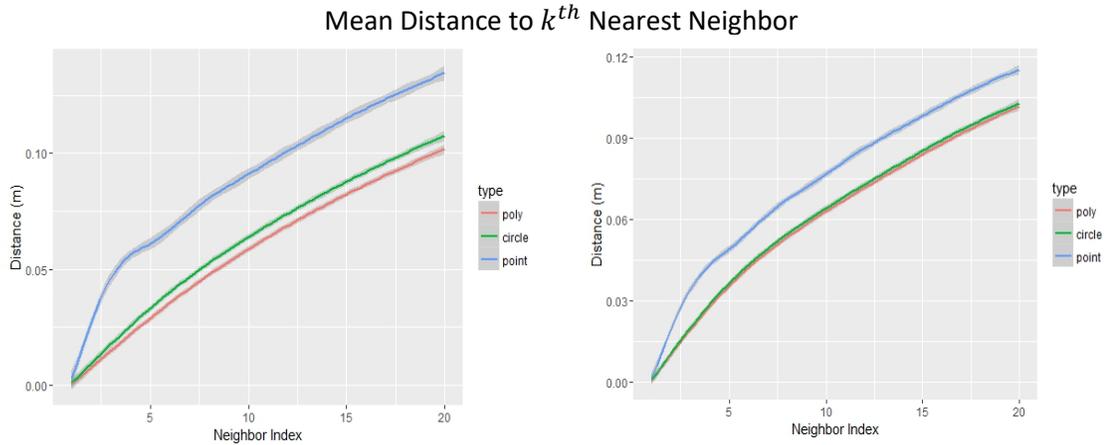


Figure 17: Estimate of the mean distance to the k^{th} nearest neighbor for (left) the high size/shape variance quadrat, and (right) the low size/shape variance quadrat. Solid lines represent the Local Polynomial Regression Fit (loess). Grey bands indicate 95% CI for estimates of the means.

DISCUSSION

A comparison between multiple strategies for constructing a spatial data set from mapped basal areas provides insight into when data type choice may influence how neighborhood composition is defined. For objects of varying size and shape irregularity, point approximations influence measures of distance between individuals by artificially increasing interplant distances relative to the actual shortest distance between pairs of individuals. Results showed that a greater competition intensity was detected when individuals were represented as polygons or circles, versus when they were represented as points. It is intuitive that polygon (or circle) edge-to-edge distances are shorter than centroid-to centroid differences, as was

observed in both quadrats. However, it is useful to note that for polygons and circles (of equivalent area), these interplant distances were influenced by the degree of shape irregularity in mapped individuals.

Due to the relatively limited spatial extent of quadrats within the data set, neighborhood analysis beyond approximately 20cm should not be considered meaningful. A further potential confounding influence of point approximations can be seen when methods of assigning point locations to irregularly shaped objects such as centroid approximations, leads to locations being set outside the actual extent of an individual (Figure 13). In extreme situations, a point location may be assigned to an individual when that location lies outside of the region in which the plant would be acquiring resources or being influenced by neighboring individuals.

These observations suggest that in systems where individuals vary in size and shape, and the distances at which they interact with neighboring individuals are small relative to their size, geometric simplifications should be cautiously applied, and their influence on defining spatial patterns carefully considered. Similarly, when attempting to model plant-plant interactions through spatially explicit simulations (e.g., agent-based modeling) models that fail to account for shape irregularity may fail to adequately capture the processes they seek to represent. Multiple studies have pointed towards the importance of considering several spatial statistics when quantifying spatial pattern (Perry et al. 2002; Wiegand, He, and Hubbell 2013; Dale and Fortin 2014). For particular study systems, it may also be important to consider multiple representations of individuals and the influence those choices have on measures of spatial pattern.

ACKNOWLEDGEMENTS

Data sets were provided by the Shortgrass Steppe Long Term Ecological Research group, a partnership between Colorado State University, United States Department of Agriculture, Agricultural Research Service, and the U.S. Forest Service Pawnee National Grassland. Significant funding for these data was provided by the National Science Foundation Long Term Ecological Research program (NSF Grant Number DEB-1027319 and 0823405).

CHAPTER 3: AN AREA-BASED EXTENSION TO THE PAIR-CORRELATION FUNCTION FOR POLYGON DATA

ABSTRACT

Fine-scale spatial patterns in semi-arid plant communities provide important information regarding underlying biotic and abiotic processes and may also act as drivers of such processes at local scales. Individual plant growth, survival and reproduction are governed more so by local interactions with immediate neighbors than overall population densities. Further, the net effect of these local interactions is largely influenced by the area (or volume) of space occupied by other individuals, for which neighborhood density may not always be an adequate proxy.

Point pattern analysis techniques have been successfully used to study local interactions, however, when large variation in size and shape exists among individuals, alternative methods may be more appropriate. An extension to traditional point pattern analysis techniques that accounts for size and shape variation has recently been proposed in which inter-plant distances are measured as polygon edge-to-edge distances rather than centroid-to-centroid distances. This method has advantages over traditional point pattern analysis techniques when size/shape variation exists, but still relies on quantifying local densities, and not the proportion of space occupied by neighboring plants.

This chapter proposes an extension to the polygon-based pair correlation function that provides estimates of departure from randomness in terms of occupied space rather than number of individuals. The approach builds on previous methods by calculating proportional overlap between sequential buffers and neighboring objects. The process is demonstrated on data from a long-term field study in which the basal areas of individual grass plants were mapped and the results are compared with traditional point pattern analysis measures of spatial pattern.

INTRODUCTION

Quantifying fine-scale spatial patterns in plant communities can provide valuable information regarding underlying ecological processes, however, modeling approaches that represent individuals as simplified geometries (e.g., points or circles) may misrepresent those patterns (Nuske, Sprauer, and Saborowski 2009). Second-order point pattern analysis techniques, wherein individuals are modeled as point locations, serve to investigate variation in local neighborhood densities and test hypotheses of plant-plant interactions at fine-scale spatial resolutions (Law et al. 2009). While representing individuals as points is appropriate in certain cases, recent research suggests that doing so may misrepresent spatial patterns when substantial variation in individual size and shape exists (Wiegand et al. 2006; Nuske, Sprauer, and Saborowski 2009). Reducing irregularly shaped objects to circles or point locations can result in inaccurate descriptions of locations and extents of individuals and distort the distances between them. When fine-scale inter-plant interactions are of interest, the bias introduced by geometric simplifications may have undesirable effects on how neighborhoods are defined.

A recent extension of point pattern analysis techniques, considers object-to-object distances between irregular polygons, and does not rely on representing individuals with centroid (point) approximations (Nuske, Sprauer, and Saborowski 2009; Getzin, Nuske, and Wiegand 2014). By explicitly considering size and shape variation, this approach more accurately captures the shortest distance between objects, and the density of individuals around irregularly shaped objects.

Distortions of inter-plant distances can be particularly important when interactions are limited to local neighborhoods. Grasses interact with their neighbors over relatively short distances, often only several centimeters (Benot et al. 2013). The competitive pressure experienced by a plant is a function of not only the number of neighboring individuals, but also the proportion of the target plants' neighborhood they occupy. When the area (or volume) from which an individual may extract resources extends beyond their mapped representation (e.g., horizontal root distributions extending beyond the mapped basal area) and these areas are known to overlap with those of neighboring individuals, local density measures may not adequately describe the neighborhood structure with respect to competitive pressures experienced by an individual. As a simple example, one may envision two spatial arrangements in which individuals competing for a common resource are at a given (edge-to-edge) distance from each other (Figure 18). The competitive pressure experienced by a target individual results not only from the local neighborhood density (the number of neighbors within a given distance), but also the proportion of space occupied by neighbors.

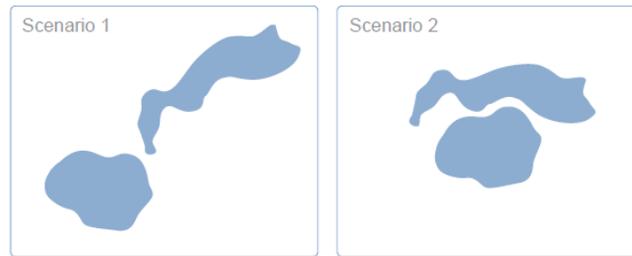


Figure 18: Two potential scenarios for the level of competition experienced by an individual. Both scenarios illustrate individuals at equivalent edge-to-edge distance from each other but with varying degrees of space occupancy and therefore intensity of competition.

Here, a novel process for quantifying the proportion of neighborhood occupied across a range of distances is described. The method combines the concepts of edge-to-edge distances and sequential buffering of polygon data, to provide an areal analogue to the traditional pair-correlation function in point pattern analysis.

REPRESENTING INDIVIDUALS AND DISTANCES

Representing individuals with point locations allows for the use of a number of well-developed statistical software programs / packages for point pattern analysis (e.g., (Baddeley and Turner 2005)). When simple geometries (e.g., circles) serve as adequate representations of the actual space occupied by an individual, centroid approximations may serve as an appropriate abstraction. However, centroid approximations exaggerate inter-plant distances when size and shape irregularity varies substantially (Nuske, Sprauer, and Saborowski 2009). Similarly, representing irregular shapes with circle geometries may result in overlapping objects, which if not present in the original data, can produce artificial pattern characteristics.

Recently proposed polygon-based approaches consider shortest edge-to-edge distances between polygons, rather than centroid-to-centroid distances, and thus provide a more accurate measure of the distances between neighboring objects (Nuske, Sprauer, and Saborowski 2009; Whigham 2013; Getzin, Nuske, and Wiegand 2014). Doing so has several advantages. First, the true size and shapes of individuals (as accurately as can be mapped) are explicitly retained in the data, rather than being relegated to attributes of point data as in the study of *marked point patterns* (Penttinen, Stoyan, and Henttonen 1992; Illian et al. 2008; Law et al. 2009; Chang et al. 2013). Second, the degree to which interplant distances are distorted by centroid approximations is not always clear, and varies depending on the shapes of individuals. Indeed, for extreme cases, the centroid approximation may fall outside of the actual polygon extent.

QUANTIFYING FINE-SCALE SPATIAL PATTERN

In traditional point pattern analysis, first-order statistics relate to the intensity λ (the number of objects within a given region, i.e., density) of a given point pattern and describe the variation in λ within the study area. Second-order statistics, such as *Ripley's K function* and more recently the pair-correlation function, describe local variation in point densities and have become increasingly popular measures of spatial pattern in plant ecology (Wiegand, He, and Hubbell 2013).

Nuske et al. (2009) proposed a polygon-based extension to the pair-correlation function that quantifies the local densities of neighboring objects within increasing search radii of a target objects' perimeter. The approach is applied in quantifying spatial patterns in forest

canopy gaps to address questions regarding competition for space and light within the study system.

This dissertation proposes a novel extension to the polygon-based pair-correlation function that accounts for the proportional overlap between sequential buffer 'rings' around individuals and neighboring polygons, rather than the density of neighbors. Results describe the relative departure from Complete Spatial Randomness (CSR) by comparing field data with simulated random polygon patterns. Further, quantified patterns for polygon data are compared with measures from the traditional (point-pattern) pair-correlation function. In study systems where interactions are at least partially governed by occupied space, the proposed method retains the benefits of point pattern analysis techniques, while providing a description of the observed patterns in terms of areal overlap.

METHODS

STUDY AREA AND DATA DESCRIPTION

The field data were collected by researchers at the Shortgrass Steppe Long-Term Ecological Research (SGS-LTER) site in Northern Colorado (40°49'N latitude, 107°47'W longitude) from 1997 through 2010 (Chu et al. 2013) as a part of a long-term grazing experiment. Basal areas of individual plants were mapped annually in permanent, 1m x 1m plots using a pentagraph. Field data were digitized and provided as shapefiles through the LTER Data Portal (LTER 2009). Long-lived C₄ grasses compose the dominant vegetation in the shortgrass-steppe, of which the perennial caespitose grasses, *Bouteloua gracilis* and *Buchloe dactyloides*, constitute the majority (Lauenroth and Burke 2008).

The intent of the work described here is to illustrate the application of an area-based approach for quantifying spatial pattern. The method is demonstrated on data from a single plot collected in 1998 (Figure 19). *Bouteloua gracilis* individuals accounted for approximately 82% of mapped individuals within the shapefile. Due to its dominance within the shortgrass steppe, intraspecific competition (typically via inhibition of seedling recruitment (Manuel O. Aguilera and Lauenroth 1993; M. O. Aguilera and Lauenroth 1993)) is the main form of competition experienced by *B. gracilis* individuals. Thus, to the demonstration presented here is limited to *B. gracilis* individuals. All data processing and analysis steps were performed in R (R Development Core Team 2016).

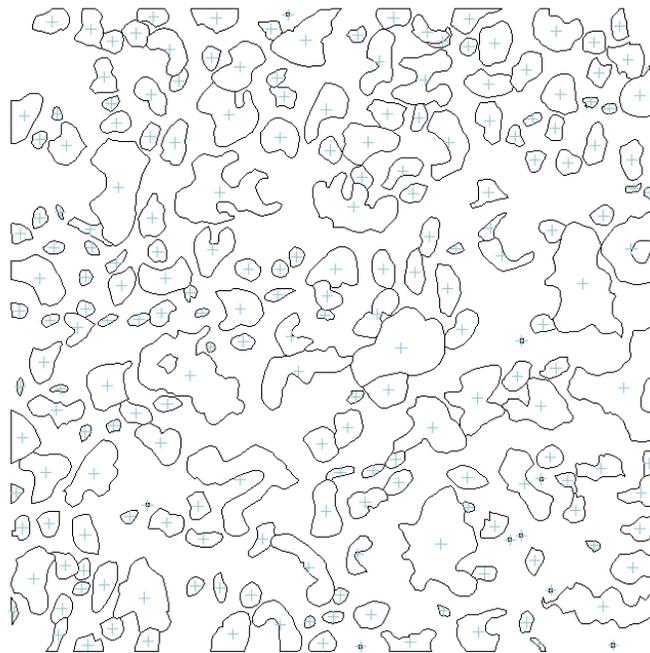


Figure 19: Mapped basal areas (solid outlines) and centroids (blue crosses) of *Bouteloua gracilis* individuals from an ungrazed plot in 1998. Plot dimensions: 1m x 1m.

PAIR-CORRELATION FUNCTION EXTENSION

For traditional point pattern analysis, the estimated pair correlation function $\hat{g}(r)$ is defined as

$$\hat{g}(r) = \sum_{i=1}^n \sum_{j=1, i \neq j}^n \frac{\omega(r_{ij} - r)}{\hat{\lambda}^2 2\pi r s(r)}, \quad r > 0 \quad (3)$$

(Penttinen, Stoyan, and Henttonen 1992) where $\hat{\lambda}$ is the estimated intensity (mean number of individuals per unit area), r the search radius, and r_{ij} is the distance between points i and j in the pattern. A kernel function, $\omega(\cdot)$ weights points within a given distance of r to account for points approximately at distance r from the focal point, and $s(r)$ serves as an edge correction factor (Nuske, Sprauer, and Saborowski 2009). For a given distance r from a target point, $\hat{g}(r)$ provides an estimate of the number of points within a small region around r (defined by the kernel function $\omega(\cdot)$) relative to the expected number of points within that region for a homogeneous Poisson process of intensity λ . Values of $\hat{g}(r) = 1$ suggest complete spatial randomness (CSR), with $\hat{g}(r) > 1$ and $\hat{g}(r) < 1$ suggesting clustering and regularity respectively (Penttinen, Stoyan, and Henttonen 1992).

Application of the pair correlation function to polygon edge-to-edge distances produces a biased estimator of $\hat{g}(r)$ due to the challenges posed in estimating the expected number of polygons at a given distance r . Nuske et al. (2009) proposed a bias correction factor $c(r)$ for polygon data as the mean biased estimator of a Monte Carlo simulation of the null model. The bias corrected estimate for polygon data is thus

$$\hat{g}(r) = c^{-1} \sum_{i=1}^n \sum_{j=1, i \neq j}^n \frac{\omega_E(r_{ij} - r)}{\hat{\lambda}^2 2\pi r p_{ij}}, \quad r > 0 \quad (4)$$

where $\omega_E(x)$ represents the commonly applied Epanechnikov kernel (Silverman 1986; Stoyan and Stoyan 1994), and p_{ij} the proportion of the perimeter of a buffer centered at point i within the study region. Polygon j is thus weighted by the inverse of p_{ij} assigning more weight to points at distance r from polygons close to an edge (i.e., the number of points at distance r cannot be determined as the area searched extends beyond the study region) (Ripley 2004).

To account for area overlap rather than polygon counts, we extended the work of Nuske et al. (2009) by considering a series of sequential buffers $\gamma_i(r)$ constructed around polygon i at distance r , and $\alpha(r)_{ij}$ as the area of overlap between $\gamma_i(r)$ and polygon j . Eq. (4) is initially modified to be

$$\hat{g}_{Area}(r) = c^{-1} \sum_{i=1}^n \sum_{j=1, i \neq j}^n \frac{\alpha(r)_{ij}}{\hat{\lambda}_\alpha^2 \nu_\alpha(\gamma_i(r)) p_{ij}}, \quad r > 0. \quad (5)$$

Implicit in this definition is an analogue of the simple rectangular kernel function in point pattern analysis

$$\omega(x) = \begin{cases} \frac{1}{2\Delta}, & \text{if } -\Delta \leq x \leq \Delta \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

which weights all points equally if they fall within the region defined by $r \pm \Delta$. The term is removed from Eq. (5) as overlap between buffer rings and neighboring polygons is directly calculated when it exists, and by definition, 0 otherwise. In the denominator of Eq. (5) the area based intensity, $\hat{\lambda}_\alpha$ is defined as the proportion of the study area occupied by all

polygons. $v_\alpha(\gamma_i(r))$ represents the area of the buffer ring $\gamma_i(r)$ and p_{ij} the proportion of $\gamma_i(r)$ that overlaps the study area. The bias correction factor $c(r)$ remains as defined previously.

Given this definition, $\hat{g}_{Area}(r)$ can be interpreted as estimating the ratio of observed areal overlap between buffer rings at distance r around polygons i and neighboring polygons j to that expected under the null hypothesis of a Poisson process with edge correction. A Monte Carlo method for simulating a Poisson process was used to determine the bias correction factor and estimate confidence intervals for the null hypothesis.

MONTE CARLO METHOD

A Monte Carlo simulation was developed to evaluate the significance of departure from the null model of CSR. Previous studies involving gridded approximation of plants (Wiegand et al. 2006) and polygon representations (Nuske, Sprauer, and Saborowski 2009; Whigham 2013) simulated patterns by randomly rotating and relocating all polygons within the study area boundary. Within the LTER data set, polygons representing plants that extended beyond the plot boundary were clipped to the bounding box edge. As a result, the true spatial extent of those individual plants is unknown, and their rotation and relocation required a customized process.

For each simulated pattern, a set of random points were generated within the plot boundaries to serve as potential locations for a polygon. Overlap between polygons was not present in the original data set, and thus not allowed in the simulated patterns. As a result, not every random point was guaranteed to successfully serve as a location for a reassigned polygon.

To address this, $2n$ random points were generated (where n is the number of polygons in the original data). For polygons completely within the plot boundary (i.e., not intersecting an edge), a random rotation angle between 0 and 359 was selected. The rotated polygon was shifted to the location of one of the randomly generated points (i.e., the centroid coordinates of the polygon were reassigned as the coordinates of the random point). Conditions were checked for overlap (allowing for intersection) with the plot boundaries, and any previously placed polygons. If no overlap was found, the rotation and reassigned location were accepted. If overlap was present, a new random rotation angle and location were selected and overlap checks performed. If a new location was not accepted within 1,000 attempts, the simulation was abandoned. Polygons were selected by decreasing size to limit failed randomizations.

Polygons that intersected the plot boundaries (and were thus clipped to the plot edges in the original data) were treated in a similar manner, though the rotation and location reassignment was performed such that rotated and reassigned polygons would remain along an edge of the bounding box. Specifically, the index of the intersecting bounding box edge was determined (the bounding box was a square with edge indices from 1 to 4 defined clockwise from the left most vertical edge), and the polygon was rotated by a randomly selected increment of 90° . Depending on the chosen rotation angle and the initial bounding box edge index, the rotated polygon was reassigned to a random location on the appropriate bounding box edge. As an example, if the polygon was initially intersecting the first bounding box edge (i.e., the left most vertical edge) and the random rotation angle was 180° , the polygon was moved to a random location along the third bounding box edge. Random locations along a bounding box edge were restricted to prevent the polygon from being placed too close to a corner of the bounding box and thus extending beyond the study area, or intersecting more

than one bounding box edge. The new location was accepted conditional on non-overlap with any previously placed polygons.

Approximate confidence envelopes at a given significance level α were estimated as the $(k + 1)\alpha/2$ and $k - ((k + 1)\alpha)/2 + 1$ lowest values of $\hat{g}_{Area}(r)$ from the k simulations (Besag and Diggle 1977; Stoyan and Stoyan 1994; Nuske, Sprauer, and Saborowski 2009). The upper and lower limits for the estimated envelope ($\alpha = 0.05$) were derived from the 12th largest and smallest values of $\hat{g}_{Area}(r)$ for each value of r from 500 randomized patterns. Five hundred randomizations have been suggested as adequate to estimate significance envelopes for $\alpha = 0.01$ (Diggle 2003; Perry, Miller, and Enright 2006).

The traditional point pattern pair correlation analysis was performed using the spatstat package in R (Baddeley and Turner 2005), also with 500 randomizations. To perform the point pattern analysis, point locations for individual plants were estimated with their respective centroids (i.e., center of mass) (Figure 20).

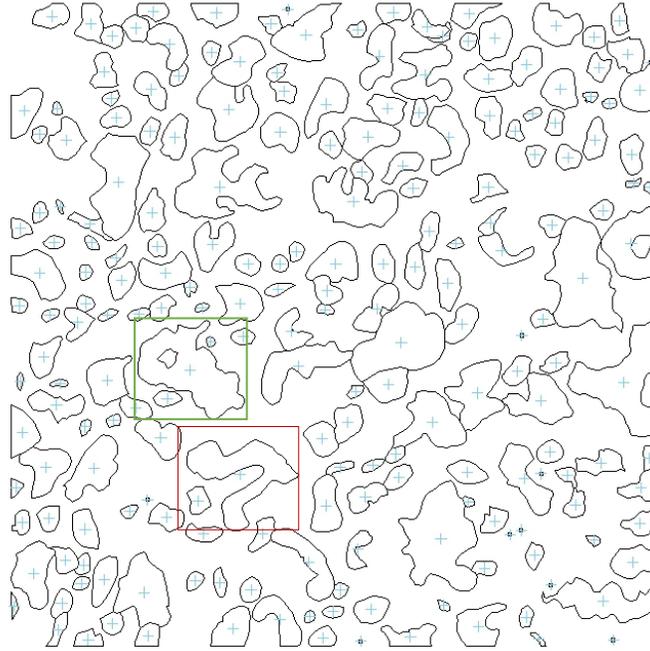


Figure 20: Point approximation for polygons by centroids (blue crosses). The red box identifies a polygon for which the center of mass was outside the polygon boundary. Green box identifies a multi-part polygon with an outer and inner ring.

This is a common approach when converting polygon data to point locations, however, when polygons are irregularly shaped, their center of mass is not guaranteed to be within the given polygon. This occurred for a single polygon in the data set (Figure 20 red box). Although methods exist for approximating polygons with point locations with the explicit requirement that the point be within the polygon boundaries (e.g., geodesic center) (Pollack, Sharir, and Rote 1989), the single occurrence was not addressed and no correction was performed in this analysis.

RESULTS

Field data contained polygons for 215 *Bouteloua gracilis* individuals ($n = 215$).

Exploratory analysis of polygon centroid locations suggested variation in local densities within the study area (Figure 21).

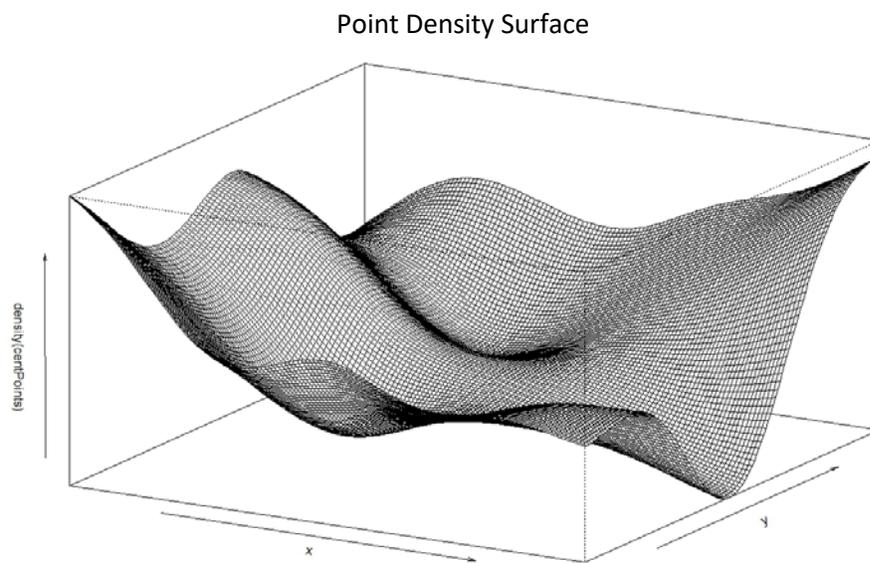


Figure 21: 3D density plot for point densities of *B. gracilis* individuals.

Quantitative descriptions of local spatial patterns were obtained for both the point approximation and polygon data with their respective pair correlation function estimates. Estimates of the pair correlation function for both the point and the polygon data exhibited clustering, randomness and regularity at multiple distances. Both approaches suggested

regularity at shorter distances, with clustering at short to intermediate distances, though the distances at which these trends were observed differed.

The estimated pair correlation function for the point pattern $\hat{g}(r)$ derived from 500 random simulations suggested significant regularity at approximately $r \leq 0.04m$ (Figure 22). Significant clustering was observed at values of $r \approx 0.06m$. No significant departures from CSR were detected for values of $r \geq 0.07m$. The range of r was automatically set to approximately $\frac{1}{4}$ the vertical and/or horizontal dimension of the observation window by default in spatstat (Baddeley and Turner 2005).

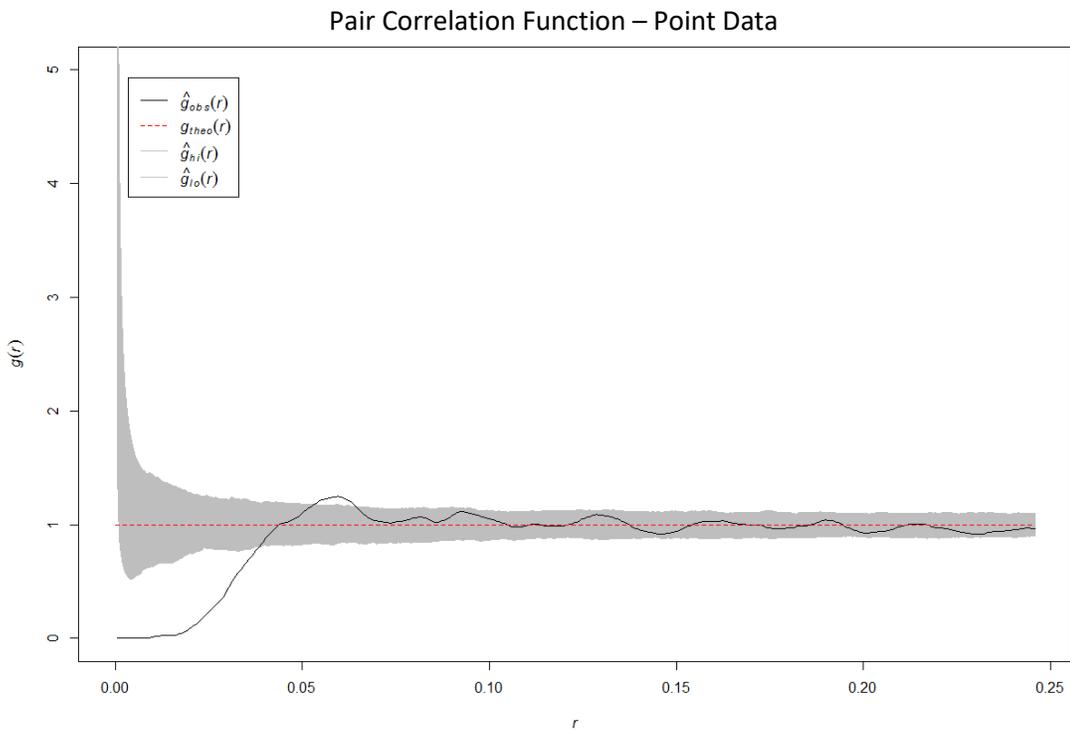


Figure 22: Estimated Pair Correlation Function $\hat{g}(r)$ for point pattern data. Confidence envelope (dark grey) generated from 500 randomizations of the point pattern. Point locations were derived from polygon centroids. Red line indicates $\hat{g}(r)$ under the null hypothesis of CSR. Black line indicates $\hat{g}(r)$ for the observed field data. Note the y axis is truncated for better visualization.

Estimates for the area-based polygon pair correlation function $\hat{g}_{Area}(r)$ suggest significant clustering at the shortest distance considered ($r = 0.005m$), and again at larger distances ($0.08m \leq r \leq 0.095m$) (Figure 23). Significant regularity was detected at distances of approximately $0.015m \leq r \leq 0.025m$. An upper limit to r for the polygon approach was set to $0.20m$ to cover a similar range as in the point pattern analysis approach, while taking into consideration the fact that buffer rings would encounter an edge earlier for the polygon data than would be the case when only point locations were considered. Additional clustering was detected towards the upper buffer distance limit ($r > 0.18m$), however this was likely influenced by the size of polygons relative to the study area, and the choice of edge correction (see Discussion).

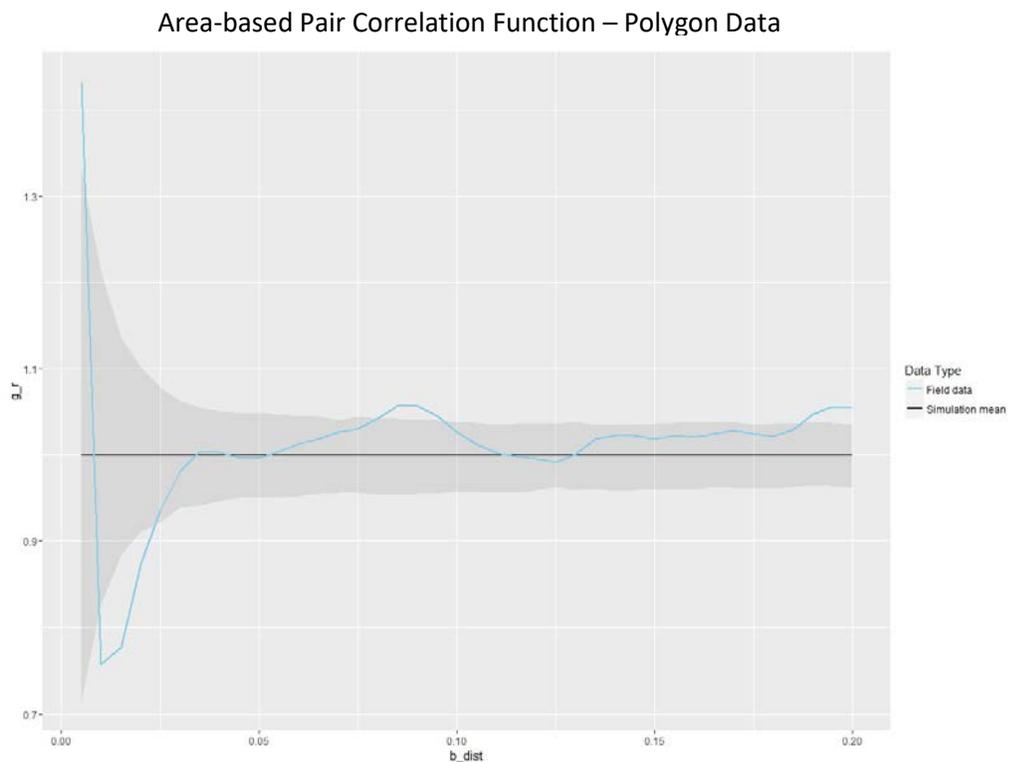


Figure 23: Estimated area-based polygon pair correlation function. Dark grey bands indicate 95% CI generated from 500 randomizations of the original polygon data set for the null hypothesis of CSR. Black horizontal line derived from the mean of 500 simulations and represents $\hat{g}_{Area}(r) = 1$.

DISCUSSION

The model presented here demonstrates a method for extending the traditional pair-correlation function to polygon data and how measures of spatial pattern (clustered, regular, random) vary depending on the type of data used to represent individuals. Measures of fine-scale spatial patterns can serve as useful tools for inference regarding underlying ecological processes. The choice of measure should be informed by shape characteristics of the organism(s) under study and the scale at which interactions with neighboring individuals takes place. The strength of those interactions varies depending on interplant distances and local neighborhood structure (Law and Dieckmann 2000), and incorporating areal overlap into a metric of fine-scale spatial pattern may serve as an ecologically meaningful extension to commonly used point pattern metrics.

The common functional form between the method described here and the traditional pair correlation function allows for a similar description of spatial pattern (clustering, regularity, randomness) with the added benefit of accounting for areal overlap rather than simply local density. As applied in this study, both measures illustrated clustering, regularity, and randomness in spatial pattern at varying distances, however the area-based extension suggested clustering at the shortest inter-plant distance measured (Figure 23).

This can be understood by considering the alternative representations of objects in the two estimates (points versus polygons). The effect of approximating areal data with point locations (e.g., centroid coordinates) on estimates of neighborhood structure is well documented, and can be seen in a simple example in which objects are represented as non-overlapping circles. It has been noted that in such a case, estimates of local densities are

inherently biased and may falsely suggest regularity at short distances (Nuske, Sprauer, and Saborowski 2009; Cressie 2015). Strict assumptions of non-overlapping individuals constrain the minimum distance at which a neighbor can exist to twice the radius of the smallest individual circle and suggest a soft or hard-core process when in fact it is not present (Cressie 2015). In many cases, similar behavior can be seen when considering point estimates of non-overlapping, irregular polygons in that the shortest distance between individuals is constrained to twice the shortest centroid to (self) polygon edge distance observed in the data. (It is interesting to note however that, as described earlier, depending on the degree of irregularity of polygons, centroid locations (i.e., center of mass) are not guaranteed to be within the target polygon boundaries and thus complicate the relationship.) In studies where such behavior is not preferred, polygon edge-to-edge distances, either calculated directly, or approximated with sequential buffering, may provide more accurate measures of spatial pattern at the smallest distances considered.

Computationally efficient ways for calculating edge-to-edge distances for irregular polygons have recently become available in modern statistical packages (e.g., rgeos package in R), though in cases where polygons contain holes (Figure 20 green box), such distances are typically only considered from the bounding polygon. Sequential buffering, though more computationally intensive and by definition a discrete approximation of inter-plant distances, may be of particular use in situations where objects may develop holes and it is possible for other individuals to occupy that space. The ability to apply both outer and inner buffer rings has the benefit of describing spatial pattern within polygon holes. Sequential buffering also provides an intuitive way to calculate overlapping areas directly or as a proportion of a given buffer ring area.

The research described here presents an initial attempt to extend traditional measures of fine-scale spatial pattern by incorporating areal information and explicitly considering size and shape variation. Several assumptions of this approach should receive attention when applying it in practice. As with all measures of local spatial pattern relying on discretized distance measurements, the buffer ring width should be appropriately defined for the system under study. Buffer distances that are too large or small may misrepresent the actual spatial pattern present in the data.

Further, due to the increased perimeter to area ratio for irregular polygons (relative to a circle of the same area), a polygon buffer is accounting for a greater proportion of the observation area (Nuske, Sprauer, and Saborowski 2009), and as a result, some form of bias correction is necessary to interpret the results in a similar fashion as those derived from point patterns. Similarly, and potentially of greater importance, the method of edge correction for sequential buffers of irregular polygons should be carefully considered. As in traditional point pattern analysis, edge correction that weights the area of overlap between a buffer ring and neighboring polygons (or the count of neighbors in the density approach) by the proportion of the ring within the study area, may have undesirable consequences when a large number of objects are distributed close to the boundaries of the study area (Perry, Miller, and Enright 2006; Dale and Fortin 2014). At larger buffer distance values, such an edge correction approach will likely overestimate clustering. This is thought to at least partially explain the significant clustering observed in the analysis presented here (Figure 23).

It is not immediately clear how size and shape irregularity of objects will affect the measure of spatial pattern, and the sensitivity of the given measure should be determined for multiple edge correction approaches. The use of a guard area in which an extended area is

sampled and objects within a fixed distance (inner-buffer) of the survey area boundary are excluded from the set of objects considered as target individuals (though retained to count as neighbors) is one possible approach to reduce the bias of edge correction choice (Perry, Miller, and Enright 2006). However, this results in a loss of data (potentially even more so when objects are large and irregular polygons) and may limit the potential for analysis on existing data sets, or the collection of additional data may be cost prohibitive.

ACKNOWLEDGEMENTS

Data sets were provided by the Shortgrass Steppe Long Term Ecological Research group, a partnership between Colorado State University, United States Department of Agriculture, Agricultural Research Service, and the U.S. Forest Service Pawnee National Grassland. Significant funding for these data was provided by the National Science Foundation Long Term Ecological Research program (NSF Grant Number DEB-1027319 and 0823405).

CHAPTER 4: A DYNAMIC VECTOR AGENT MODEL OF PLANT-PLANT AND PLANT-ENVIRONMENT INTERACTIONS: A PROOF OF CONCEPT MODEL

ABSTRACT

This chapter describes an agent-based modeling approach that attempts to capture size and shape dynamics of individuals responding to local conditions. We outline a generic agent construction that builds on the recently proposed Dynamic Vector Agents (DVA) approach by incorporating interaction with a grid-based environment. The proposed model allows for the simulation of polygon agents, that can change their geometry (shape) by two general methods; 1) by selecting from a set of rules governing how growth takes place (node displacement, edge displacement, point displacement), and 2) by changing the responsiveness to local environmental conditions (directional growth). Simulation results suggest that such an approach is capable of generating polygon agents varying in shape complexity, and serves as a proof of concept model for incorporating size/shape dynamics into models of plant competition/facilitation.

INTRODUCTION

Agent-based models (ABMs) have seen increasing use in plant ecology since the mid 1970's (DeAngelis and Grimm 2014). Such models attempt to simulate system dynamics from a 'bottom-up' perspective by defining sets of relatively simple rules that govern the behaviors of

individuals (i.e., agents). From these simple rules, complex phenomena at the population/community level may emerge without a deterministic model framework being imposed. Spatial representation of agents within simulations is often simplified to either a grid or point data type. Such simplifications are advantageous in terms of computing resources when running simulations involving large numbers of interacting agents. However, for systems in which interactions between individuals take place at fine-spatial scales (relative to the sizes of individuals), incorporating the ability to simulate increasingly complex shapes may be advantageous. In this work, we describe a generic agent based model that defines agents as polygons that change size and shape in response to local availability of a simulated resource.

PLANT COMPETITION MODELS

Cellular Automata (CA) models in which individuals (or collections of individuals) were modeled as individual grid cells in a matrix environment (Silvertown et al. 1992; Balzter, Braun, and Köhler 1998; Matsinos and Troumbis 2002; Wang et al. 2003) were among the earlier spatially explicit simulations to gain popularity in ecology. Individual responses to states of adjacent grid cells were governed by sets of rules that allowed for efficient simulation of the complex interactions between numerous individuals, and the quantification of community/population wide, emergent patterns. While useful in generating and/or exploring hypotheses, CA models are limited to discretized representations of space and a fixed spatial resolution (Berger et al. 2008). Agent-based models (or Individual-based models) allow for the location and extent of individuals to be modeled in continuous space, and further allow for interactions between agents defined in continuous space and discrete space.

More recently, ABMs of plant populations/communities typically represent individuals as two-dimensional circular objects and interactions between individuals (competition/facilitation) take place when the areas of individuals overlap (Berger et al. 2008). In such models, the circular area attributed to an individual plant is often referred to as a *Zone Of Influence (ZOI)* (Berger et al. 2008; C.-J. Chu et al. 2008; Railsback and Grimm 2011) and represents the area from which the plant may acquire resources (e.g., light, water, nutrients). This concept has been further extended with the *Field of Neighborhood (FON)* approach which allows for the zone around an individual to be represented with a decay function such that the competitive effect of an individual decreases with greater distance from its central location (Bauer et al. 2004; Berger et al. 2008).

Circular geometries most often serves as an abstraction of either basal area in grassland systems or canopy cover in forest systems. Such a representation is certainly appropriate in many situations, however, when a high degree of variability in size and shape exists among individuals, simplified geometries may misrepresent the actual distances at which individuals interact (Wiegand et al. 2006; Whigham 2013). Such a misrepresentation is particularly concerning when fine-spatial scale interactions between neighboring plants are of interest and modeling plants with simplified geometries does not accurately represent distances between plants, or the area from which they may attain resources. As a result, it is reasonable to expect that simulations constructed with such simplifications may fail to capture a potentially important spatial quality of plant communities/populations.

DYNAMIC VECTOR AGENTS

Hammam et al. proposed a method for incorporating shape change into agent behavior with their concept of Dynamic Vector Agents (DVA) (Hammam, Moore, and Whigham 2007). This construct allows for agents to increase in size not simply by incremental increases in area of circular objects (e.g., by a defined allometric relationship between stem diameter and canopy area for a tree species), but by probabilistically selecting from multiple growth strategies. The three strategies discussed here (adapted from the DVA approach) include (Figure 24); 1) node displacement; in which one of the nodes in a polygon representing an agent is moved, 2) point displacement; whereby an individual edge is split at its midpoint and the point resulting from the split is placed some distance away from the edge. The new point is then connected to the two nodes of the original edge, and the original edge is removed, and 3) edge displacement; in which an edge of the polygon is moved.

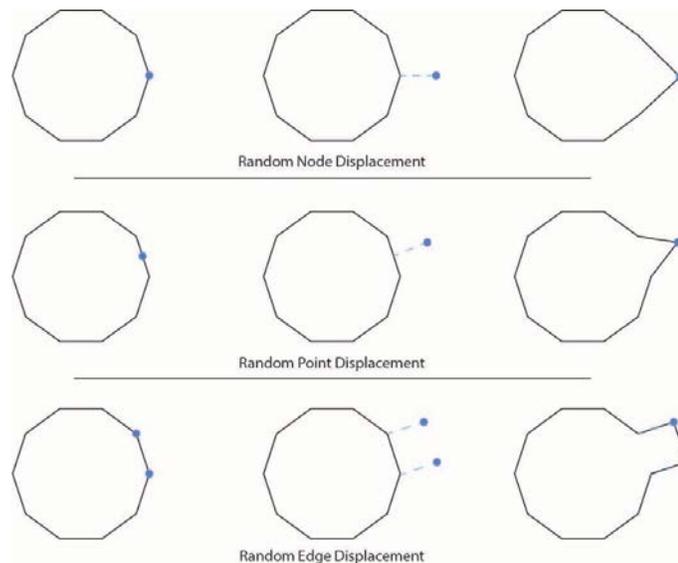


Figure 24: Illustration of shape dynamics methods.

The DVA approach demonstrates the ability to generate polygons of a wide range of shape complexity ranging from simple expansion of an initial shape (e.g., expanding square) to highly complex fractal-like shapes (Hammam, Moore, and Whigham 2007). Here, we extend the DVA approach by incorporating interaction between agents directly through checks for spatial overlap, and indirectly through their respective influences on local availability of a simulated resource. This approach allows for the simulation of fine-scale interactions between organisms which may vary in size and shape complexity and extends the ZOI and FON concepts to objects of irregular shape.

ODD DESCRIPTION

In what follows, we adhere to the ODD protocol for reporting agent-based models (Grimm et al. 2006; Grimm et al. 2010), and organize the model description into the following sections: 1) Purpose; 2) Entities & State Variables; 3) Process Overview & Scheduling; 4) Design Concepts; 5) Initialization; 6) Input data; and 7) Sub-models.

PURPOSE

The model described in this work has two specific aims; 1) to describe the Dynamic Vector Agents approach to an ecology audience, and illustrate its application in modeling irregularly shaped plant agents, and 2) to extend the concept by incorporating plant-environment interaction in a fashion analogous to the Zone of Influence and Field of Neighborhood approaches. As such, the model allows for direct interactions between polygon agents (plants) and a grid environment simulating an abstract level of resource availability, as well as indirect interactions between polygon agents via their local influence on the grid environment.

The model is intended to serve as a proof-of-concept model, and not necessarily to explicitly model a particular plant community or population. However, the motivation for developing such a model came from a unique, long-term dataset from the Long Term Ecological Research (LTER) station in the shortgrass steppe of Northern Colorado, USA. Between 1997 and 2010, researchers mapped the basal areas of individual plants using a pantograph in a series of 1x1m permanent plots, as a part of a larger grazing exclusion study (C. Chu et al. 2013). Hand-drawn data were digitized and made publicly available via the LTER Data Portal (LTER 2009) in shapefile format. Data from these plots demonstrate a wide range of basal area shape complexity for a number of species, in particular the dominant grass *Bouteloua gracilis* (Figure 25).

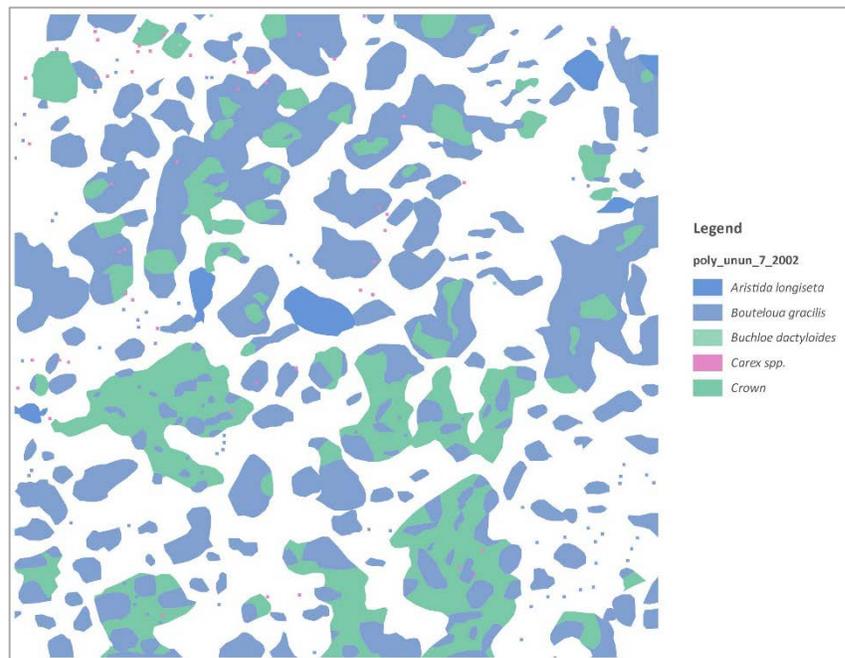


Figure 25: Example of mapped basal areas from LTER data set.

Plants in this system are interacting with other individuals within their immediate neighborhoods (M. O. Aguilera and Lauenroth 1993), and as a result, these local interactions are

important in determining growth, reproduction, seedling establishment, and mortality (Manuel O. Aguilera and Lauenroth 1993). The majority (approximately 75%) of the roots of a typical *B. gracilis* individual occur within 5cm horizontally of the plant, and approximately 10cm vertically (Coffin and Lauenroth 1991). As such, basal area should serve as a reasonable proxy for horizontal distribution of roots for *B. gracilis* individuals. However, it also makes it clear that the area within which an individual can impact local resource levels, extends beyond the mapped basal area. Given that *B. gracilis* roots tend to remain fairly close to the basal 'footprint' of an individual, it is reasonable to suggest that plants with markedly different basal area shapes would differ in root distribution, and potentially have different levels of interaction with neighboring plants.

We present the proof-of-concept model here, inspired by this long-term study, to illustrate a method for simulating shape irregularity among individuals, and the ability of individuals to impact resource availability locally beyond their mapped representation. The aim is not to explicitly reproduce observed patterns as resource levels were not monitored within the plots, and thus the incorporation of plant-environment interaction is generic in nature. We demonstrate the range of shape dynamics the model is able to produce, and discuss different parameterizations, however we do not undertake an extensive Pattern Oriented Modeling (POM) approach (Thiele, Kurth, and Grimm 2014) to learn parameter values, as no data exist for environmental conditions which largely drive shape dynamics within the model. However, the described model should serve as a robust base on which to build more complex models for systems where appropriate field data exist. The current version of the model was constructed in the GAMA modeling and simulation platform version 1.7 (Taillandier et al. 2010).

ENTITIES & STATE VARIABLES

Three primary types of entities exist in the model; 1) the global agent in which global variables and actions are defined, 2) grid agents representing the background environment (habitat cells), and 3) polygon agents representing individual plants. GAMA requires that dimensions be provided for the global environment. For the model described here, the global environment was defined as a toroidal world with dimensions 100x100 units, and represents the extent of the simulation and thus the bounds on all agents (whether point, polygon, or grid agents). The toroidal construction allows for plant agents that grow across the edge of the simulation environment to continue on the opposite side, effectively creating a world that wraps around on itself.

The grid agent was defined relative to the global extent by specifying the number of cells horizontally and vertically to be created. For the model demonstrated here, we set the grid agent to be 50x50 cells. Polygon agents were initialized as circular polygons with a radius of 0.5 units and 10 nodes along their perimeter. Polygon agents are initialized with random locations throughout the global environment. State variables are quantities which can vary for individual entities throughout a model run though in some cases they are treated as constants. Each agent type (global, grid, polygon) contained a set of state variables that defined important quantities for actions they perform during the simulations. Here we describe some of the most important variables that determine agent behaviors or initialization, and provide a complete list of all state variables (Table 1) as well as the model code (APPENDIX A).

Table 1: State variables and parameters for the model.

Variable	Initial value	Description
Global Agent		
worldDimension	100	Variable for global dimensions
shape	square(worldDimension)	Defines the shape of the world
simStep	20	Number of steps until a polygon shape is simplified
rNodeProb	Varied	Probability of shape change by displacing a polygon node
rEdgeProb	Varied	Probability of shape change by displacing a polygon edge
randInt	Varied	Controls the influence of resource dependent growth direction for polygon agents
Grid Agent		
maxResource	100	Maximum resource level for grid cells
inputResource	Random integer < 10	Level of resource input at each time step
resourceLoss	Random integer < 10	Level of resource loss at each time step
availableResource	Random integer < 100	Level of resource available in a grid cell (derived value after first step)
Polygon Agent		
pointsOnShape		List of polygon nodes
randomNode		Index of polygon node to reference for shape changes
previousRandNode		Additional node index in the event that the last point in the polygon point list is chosen
nextRandNode		For edge displacement. This is the index of the node after <i>randomNode</i>
segmentAngle		Used for defining angle of displacement
outSegmentAngle		Similar use to <i>segmentAngle</i>
angleToDisplace		Angle by which to displace a node or edge
randAngle	90	Defined to displace node or edge orthogonally
polyNeighbors		List of neighboring polygon agents (used to reduce the number of polygons an agent needs to check for overlap)
offsetScale		Scale for the distance of node or edge offset at each step
polyNodes		List of polygon nodes (updated after shape simplification and/or displacement)
nodeToDisplace		Node that is to be displaced
translateAngle		Angle by which to translate a node or edge
angleCos		For angle displacement calculations
angleSin		For angle displacement calculations
tempShape		Used when checking for intersections
randomPointTest		Stores new point location (after translation)
randomLineTest		Used when translating polygon edge
crossedTest		Tests for crosses with self and other polygons
sci		Shape Complexity Index
bufferNeighbors		List of grid cells that are within a given buffer distance of a polygon agent
growthNodes		List of nodes that could be translated

Grid agents: state variables

The grid agent consists of cells each of which has a level of resource availability. The *maxResource* variable defines the maximum amount of a resource a given cell may hold (set to 100) while the *inputResource* and *resourceLoss* variables control the resource fluctuations independent of interaction with polygon (plant) agents. Each time step, these two variables are random integer values between 0 and 10. Resource dynamics are encapsulated in the *availableResource* variable which updates each time step.

Polygon agents: state variables

The version of the model described here is focused primarily on implementing shape dynamics for polygon agents, and interactions between polygon and grid agents to determine the strategy for shape change. Thus, the current model does not attempt to model growth with allometric or similar equations as is common in plant ABMs (C.-J. Chu et al. 2009; C.-J. Chu et al. 2010; Lin et al. 2013; Lin et al. 2014). Growth equations could be implemented within this framework, however the particular method of shape change chosen and degree of shape irregularity of a polygon would require special consideration when constructing allometric rules for growth. Currently, the distance at which a point, node or edge (depending on shape change choice) is offset is one unit, scaled by a, randomly selected factor of 1, $1/2$, $1/3$, or $1/4$.

Polygon agents interact directly with overlapping grid cells and those within a specified distance. Each node of a polygon agent identifies the grid cell it overlaps and stores its *availableResource* value. At each time step, a given polygon will create a sorted list of these node values. The degree to which underlying grid values influence directional growth is determined by selecting from among the first n elements of the sorted list. This behavior is

controlled with the *randInt* parameter which in effect, governs directional growth. As an example, for *randInt* = 1, the node with the highest associated grid resource level is chosen as the node to displace (i.e., the location along the polygon perimeter that will move outward expanding its shape). For a *randInt* value of 10, one of the first ten nodes in the sorted list is randomly chosen for displacement.

At each time step, upon selecting a node, point, or edge for displacement, a polygon agent undergoes a series of checks to ensure that the new shape (regardless of displacement method) remains topologically correct (e.g., that it does not cross itself). Polygon agents have multiple state variables that prevent self-crossing and crossing neighboring polygon agents.

PROCESS OVERVIEW & SCHEDULING

Upon beginning a simulation run, the global agent is initialized first, and defines general properties of the simulation and global variables for agents. The grid agent is initialized next, with any explicitly defined initial values for its variables. Lastly, the polygon agents are initialized. Agent actions are performed first for grid agents, and then for polygon agents. Though customizable scheduling is possible in GAMA, the current implementation of this model uses the default scheduling in which agents are called in the order in which they were created. Time in the simulation is modeled in discrete steps, but no explicit relationship to time units (e.g., daily or monthly time steps) is implied. If for a given time step, a topological exception is encountered (e.g., if a polygon crosses itself when performing a displacement) the agent does not undergo any shape change during that time step. This has important ramifications on the simulation. A variety of sizes and shape irregularity results during simulation runs, however,

agents that remain small or constant for a number of steps are doing so because of topological exceptions, and not explicitly due to simulated energy reserves. Incorporating reserves that would inhibit growth is a logical candidate for an extension to this model.

Each time step, grid agents update their *availableResource* attributes by drawing random values for the *inputResource* and *resourceLoss* variables, as well as accounting for their resource reduction due to polygon agent actions. Polygon agents reduce the amount of the *availableResource* variable each time step for those grid cells they overlap, and to a lesser degree, those within a defined buffer distance (stored in the *bufferNeighbors* variable of polygon agents). This process may take the form of a decay function reducing the resource levels by decreasing amounts relative to their distance from the polygons, though doing so requires the storage of a large number of cells for each polygon agent, and may lead to dramatically increased run times. In the case described here, the grid cell resolution is such that plants were only modeled as impacting grid cells within 5 units of their location (recall that the global environment is 100x100 units).

Polygon agents undergo a *grow* action each time step which consists of the following processes. The polygon creates the sorted list of resource values for each of its nodes, and selects one of the nodes depending on the definition of the *randInt* variable which controls how many nodes will be considered as candidates to displace. With a node selected, the polygon selects one of the three growth strategies (random point displacement, random node displacement, or random edge displacement). Each strategy is assigned a value between 0 and 1 and the choice is made based on the normalized distribution of those values. These parameters can be manually adjusted during simulation runs, or set within GAMA as parameters to be explored in batch runs.

DESIGN CONCEPTS

Basic Principles

The model described here aims to incorporate the principles of Dynamic Vector Agents with interaction between agents and a grid-based environment. This combination allows for Zone of Influence type modeling scenarios to be applied in simulations in which agents dynamically change geometries based on those interactions, and (indirectly) through interactions with other polygon agents. While the current representation is aimed at applying these principles in a proof-of-concept context, and not at reproducing observed patterns per-se, it serves as a base on which to build more biologically / ecologically realistic simulations in the future.

Emergence

The primary form of emergent phenomena in the current model is in the form of shape complexity dynamics of polygon agents in response to interactions with other agents and their environment. Depending on the selection of simple growth rules a range of shape irregularity is possible. The growth actions are probabilistic in nature, and the observed size/shape distributions for polygon agents are a result of both 'built-in' rule assignments (topological checks) as well as interactions with neighbors and local conditions.

Adaptation, Objectives, Learning & Prediction

In the current basic implementation of the model, no processes of adaptation, objective based actions, learning or prediction occurs within polygon or grid agents. While several recent ABMs dealing with changes in aboveground/belowground biomass allocation have been

developed (Lin et al. 2014), it is rare in plant based ABMs to implement learning or strategy based behaviors within agents during a simulation.

Sensing

Polygon agents are able to access the grid values for cells they directly overlap, as well as those within a user-defined buffer. While functionally, there is no reason that an individual plant cannot access all grid values, the decision regarding the spatial extent of their access should be based on knowledge of the species being modeled (e.g., based on known spatial distribution of root systems). The querying of grid cell values is explicitly coded in the polygon agents behaviors, and can be refined to provide a higher degree of biological realism.

Interaction

Polygon agents directly interact with grid agents they overlap, and those are within the defined buffer distance. Polygon agents interact both directly (by checking for collision with other polygons) and indirectly (by effects on underlying grid resource values) with other polygon agents.

Stochasticity

Currently, initial locations of plant individuals are randomly assigned. Likewise, grid values for the resource level are randomly assigned to create heterogeneity within the grid environment. Growth strategies are probabilistic in nature, but the probabilities are fixed within a simulation run, and are not considered truly random in this context.

Collectives

No form of collective behavior or information sharing exists in the current version of this model.

Observation

Data can be collected or followed about any variable or attribute for individual polygon agents. In this model version, reporting on individual Shape Complexity Index values (SCI) was coded in the model, and can be written to files for further analysis. The SCI was defined as $SCI = Perimeter / \sqrt{4\pi * Area}$, taking a value of 1 for circles, and larger values with increasing shape irregularity (Nuske, Sprauer, and Saborowski 2009). As the purpose of this current version of the model was to incorporate dynamically changing shapes of polygon agents, no measures of mortality, reproduction or other variables were observed or reported on, though if defined within the model, GAMA allows for reporting on all variables of interest.

INITIALIZATION

All global agent properties were constant through multiple simulation runs. Grid agent values were randomly assigned at initiation and take values between 0 and 1 for resource levels. Polygon agents were randomly located upon initialization, and subsequent simulation runs have different random positions selected for polygon agents.

INPUT DATA

No input data were utilized in the current version of the model, though GAMA does allow for GIS data to be provided to models. A useful extension to this model could be seen in providing shapefiles with actual measured objects as the initial configuration of polygon agents, or raster files containing data about soil properties.

SUBMODELS

Polygon Growth

GAMA refers to an action performed every time step as a 'reflex'. The *grow* reflex in this model consists of several submodels/routines depending on the method of displacement chosen. The reflex begins with a node being selected for a given polygon agent. This choice is influenced by the *randInt* variable and the underlying grid values as described in the PROCESS OVERVIEW & SCHEDULING section. Once a node is selected, a polygon may grow by either displacing that node, displacing the edge between the selected node and the subsequent node in the polygon, or finding the midpoint of that edge and displacing that point while simultaneously adding the new point as a node in the polygon. For the current version of the model, displacement occurs orthogonally with respect to adjacent polygon edges.

Depending on the probabilities associated with the various growth strategies, artificially complex or 'jagged' shapes may result. This may not be an issue in some systems, but a method was defined to simplify polygon geometries to moderate this phenomenon. The *simShape* parameter allows the user to define the number of time steps to allow before shape simplification is performed. This process involves reducing the number of nodes in a given polygon to a user defined number (which can also be variable depending on the number of nodes in the current polygon). Reducing the number of nodes in general tends to limit extremes in jagged edges. This is left as a parameter, as it further influences shape dynamics, and its value should be informed by the target organisms being simulated.

As described earlier, displacement actions that result in a polygon edge crossing itself, are not allowed in the model. Likewise, crosses with neighboring polygons are not allowed. This

has important implications on the development of size variance among polygon agents, as no growth occurs at time steps where crosses occur for a polygon. While this is not biologically motivated, it is necessary to maintain the correct definition of polygons as a set of closed line segments. Further model development should address this by allowing polygons to continue selecting nodes for displacement until a successful change occurs or an allowable amount of size increase occurs for a given step. This is not a trivial matter, as agents may need to be constructed such that they are allowed multiple actions in a given time step, and area calculations will need to be repeatedly performed by each agent to ensure the proper growth amount at a given step. Despite these challenges, these are important additions to be made to better simulate observed systems.

METHODS

To demonstrate the ability of the model to simulate observed shape dynamics, we estimated parameter values in relation to their ability to reproduce patterns in the Shortgrass Steppe LTER data set. This is not meant to serve as a rigorous example of parameter estimation, rather to illustrate that the model is capable of producing shape irregularities similar to those observed in the field data.

In the LTER data set, polygons representing the smallest *B. gracilis* individuals were arbitrarily assigned square shapes with side-length of 0.25cm during digitization (C. Chu et al. 2013). There were multiple years in which large numbers of small individuals dominated the data. As the purpose of this demonstration is to illustrate the model's ability to reproduce irregular shapes, and the simulated individuals start from a circular rather than square shape, we excluded those individuals when defining metrics used for parameter estimation. The

shapefiles were further subset to include only those polygons that were completely within the plot boundaries (i.e., those that were not truncated by the bounding box), as the full extent and shape of polygons intersecting the plot boundaries were unknown, and thus SCI calculations for those polygons would not be useful. Due to the toroidal structure of the simulations, this is not an issue in the ABM output.

To estimate parameter values that result in similar SCI values in ABM runs, we used the squared difference between ABM median SCI values and the LTER median SCI value as a value to minimize. For each set of parameter combinations, simulations of 10 polygons were run for 100 time steps, after which SCI values were recorded. Each parameter set was run five times, and the median SCI value was calculated for that given set of parameter values. Options for parameter values were initially selected from a coarse-level subset of possible values that spanned the full range of values a parameter could take. If they existed, subranges for each parameter that performed better in approximating the field data were identified and complete enumeration of parameter values within their given subranges was performed.

RESULTS

Shape Complexity Index values for polygons in the LTER data set exhibited a right-skewed distribution with a median SCI value of approximately 1.165 and standard deviation of 0.194 (Figure 26). In general, this suggests that many polygons were slightly irregular in shape, with some having greater shape complexity. It should be noted that very large (and typically among the more complex) polygons often intersected plot boundaries, and were thus not represented in this distribution.

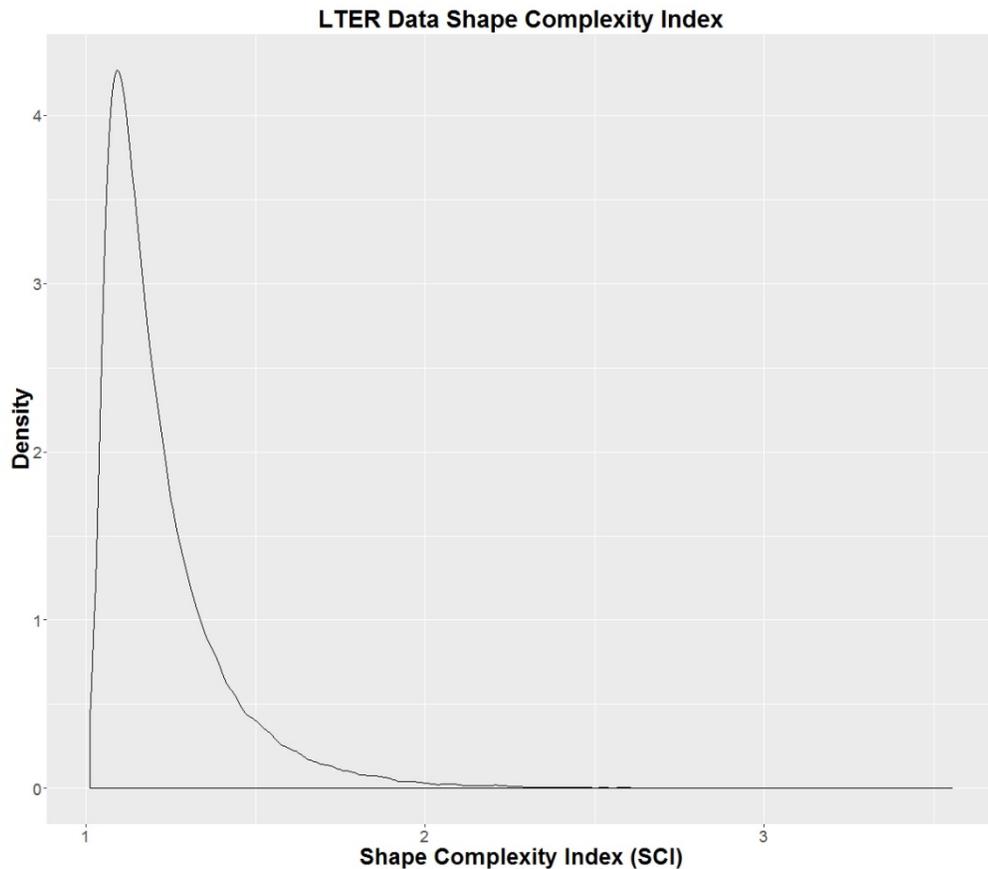


Figure 26: Density plot of LTER data set Shape Complexity Index values (excluding the smallest square polygons as described in the text).

It is likely that the median would shift to a larger value if the shape of those individuals were known, however, the aim of this stage of model development is more concerned with developing the mechanism for irregular shape development, rather than reproducing the observed patterns with a high degree of accuracy. Sampling parameter values at coarse-scales quickly identified that only the lowest values of random edge and random point displacement parameters (0.0-0.2) should be subsampled, while the full random node displacement parameter range was considered in the final round of parameter estimation. The shape simplification parameter was sampled from 1 to 50 in increments of 10 for the final round.

Median SCI values for the second round of parameter estimation exhibited a bimodal distribution (Figure 27).

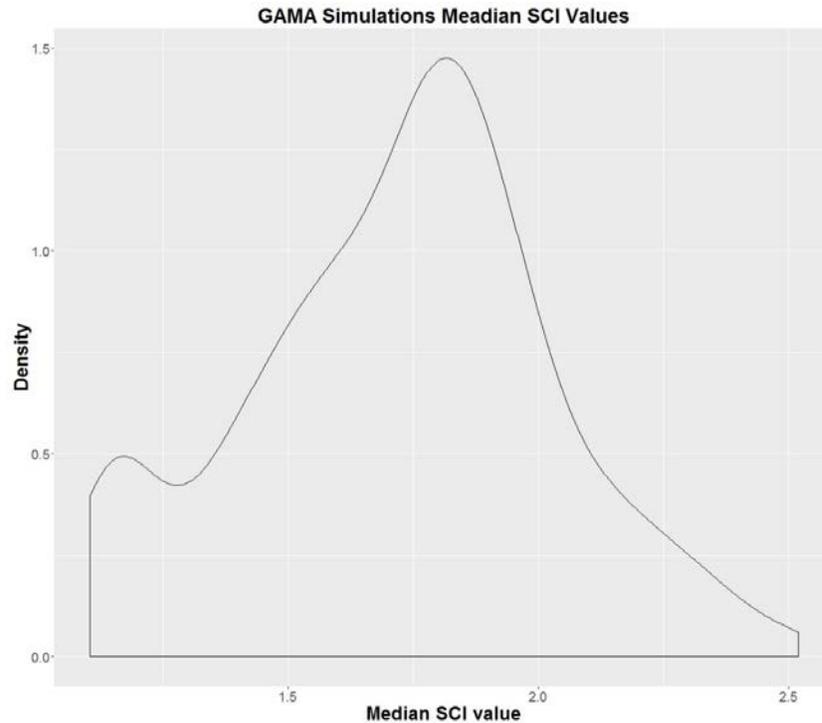


Figure 27: Density plot of the median SCI values from the final round of parameter estimation simulation runs. Median SCI was calculated for five runs with the same parameter values.

In its current form, minimizing the squared error between the modeled SCI values and the LTER values suggests that both the random edge and random point displacement strategies were not important in generating similar shapes to those observed in the field data. Random edge displacement was not used at all in any of the top ten performing parameter sets. Random point displacement was similar, only being assigned a value (0.1) in any of the top ten performing parameter sets. Squared error was minimized (sq. error = $1.6304e-07$) with a

parameter set of; Shape simplification = 20 steps, random node displacement = 0.3, and the remaining parameters = 0.0 (Figure 28).

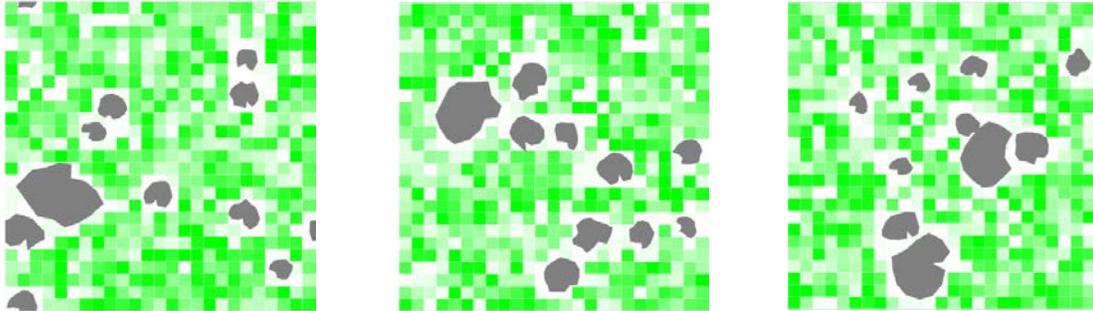


Figure 28: Examples of simulation runs with final parameter values.

DISCUSSION

The model presented here illustrates a method for constructing agent-based models in which agents change size and shape dynamically in response to local conditions. This approach potentially benefits modeling efforts where these dynamics are thought to be of ecological importance, such as plant-plant interactions in grasslands. We described a modeling framework that combined aspects of the Dynamic Vector Agents and Zone of Influence concepts to allow for indirect interactions between agents via their effects on local conditions.

While the model was capable of producing similar irregular shapes as those observed in a long-term field study of mapped basal areas for *Bouteloua gracilis*, this should be interpreted with caution. The metric used for parameter estimation (median SCI) showed that while exhibiting some irregularity, the mapped basal areas in the LTER data were not extremely complex (e.g. fractal-like). This likely contributed to the limited to no importance of two of the shape dynamics parameters (random edge and random point displacement). Both strategies can

quickly result in increasingly fractal-like shapes and within the LTER data, the prominent form of irregularity was from elongation of polygons, and not protruding extensions.

While the focus for this work was primarily on polygon agents, the method of simulating resource dynamics in the grid agent played an important role in determining shape dynamics of polygon agents. The grid agent was initialized with random levels of resource availability, and this immediately led to the lack of importance for the way in which we structured directional growth for polygon agents. The importance of this parameter would likely change with simulated gradients in resource availability and/or with increasing density of individuals. Gradual elongation of polygon agents rarely occurred unless they were in close proximity to other polygons, at which point they would stop growth towards each other while continuing growth in the remaining directions. While similar behavior appeared to be happening in the LTER data, it is likely that the randomization method for grid resources and the low polygon density in the simulations obscured the relationship between some of the parameters.

Agent-based modeling has become a useful tool for ecologists studying spatially explicit processes and has seen increasing use in recent decades (DeAngelis and Grimm 2014). Our hope with the modeling approach presented here is to call attention to a method developed in other disciplines for modeling Dynamic Vector Agents and illustrate a proof-of-concept model that combines the DVA approach with concepts from recent plant ABMs. The current model construction suggests several interesting lines of inquiry for refinement and increasing its performance in simulating plant-plant interactions at fine-spatial scales.

ACKNOWLEDGEMENTS

Data sets were provided by the Shortgrass Steppe Long Term Ecological Research group, a partnership between Colorado State University, United States Department of Agriculture, Agricultural Research Service, and the U.S. Forest Service Pawnee National Grassland. Significant funding for these data was provided by the National Science Foundation Long Term Ecological Research program (NSF Grant Number DEB-1027319 and 0823405).

LITERATURE CITED

- Adler, P., D. Raff, and W. Lauenroth. 2001. "The Effect of Grazing on the Spatial Heterogeneity of Vegetation." *Oecologia* 128 (4): 465–79. doi:10.1007/s004420100737.
- Aguilera, Manuel O., and William K. Lauenroth. 1993. "Seedling Establishment in Adult Neighbourhoods--Intraspecific Constraints in the Regeneration of the Bunchgrass *Bouteloua Gracilis*." *Journal of Ecology* 81 (2): 253–61. doi:10.2307/2261495.
- Aguilera, M. O., and W. K. Lauenroth. 1993. "Neighborhood Interactions in a Natural Population of the Perennial bunchgrass *Bouteloua Gracilis*." *Oecologia* 94 (4): 595–602. doi:10.1007/BF00566977.
- Baddeley, Adrian, and Rolf Turner. 2005. "Spatstat: An R Package for Analyzing Spatial Point Patterns." *Journal of Statistical Software* 12 (6): 1–42.
- Balster, Heiko, Paul W. Braun, and Wolfgang Köhler. 1998. "Cellular Automata Models for Vegetation Dynamics." *Ecological Modelling* 107 (2–3): 113–25. doi:10.1016/S0304-3800(97)00202-0.
- Bauer, Silke, Tomasz Wyszomirski, Uta Berger, Hanno Hildenbrandt, and Volker Grimm. 2004. "Asymmetric Competition as a Natural Outcome of Neighbour Interactions among Plants: Results from the Field-of-Neighbourhood Modelling Approach." *Plant Ecology* 170 (1): 135–45. doi:10.1023/B:VEGE.0000019041.42440.ea.
- Benot, M. -L., A. -K. Bittebiere, A. Ernoult, Bernard Clement, and Cendrine Mony. 2013. "Fine-Scale Spatial Patterns in Grassland Communities Depend on Species Clonal Dispersal Ability and Interactions with Neighbours." *JOURNAL OF ECOLOGY* 101 (3): 626–36. doi:10.1111/1365-2745.12066.
- Berger, Uta, Cyril Piou, Katja Schiffers, and Volker Grimm. 2008. "Competition among Plants: Concepts, Individual-Based Modelling Approaches, and a Proposal for a Future Research Strategy." *Perspectives in Plant Ecology, Evolution and Systematics* 9 (3–4): 121–35. doi:10.1016/j.ppees.2007.11.002.
- Besag, Julian, and Peter J. Diggle. 1977. "Simple Monte Carlo Tests for Spatial Pattern." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 26 (3): 327–33. doi:10.2307/2346974.

- Bolker, Benjamin M., Stephen W. Pacala, and Claudia Neuhauser. 2003. "Spatial Dynamics in Model Plant Communities: What Do We Really Know?" *The American Naturalist* 162 (2): 135–48. doi:10.1086/376575.
- Boots, Barry. 2003. "Developing Local Measures of Spatial Association for Categorical Data." *Journal of Geographical Systems* 5 (2): 139–60. doi:10.1007/s10109-003-0110-3.
- Boots, Barry, and Ferko Csillag. 2006. "Categorical Maps, Comparisons, and Confidence." *Journal of Geographical Systems* 8 (2): 109–18. doi:10.1007/s10109-006-0018-9.
- Brinkhoff, Thomas, Hans-Peter Kriegel, Ralf Schneider, and Alexander Braun. 1995. "Measuring the Complexity of Polygonal Objects." In *ACM-GIS*, 109. Citeseer. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.73.1045&rep=rep1&type=pdf>.
- Chang, Ya-Mei, Adrian Baddeley, Jeremy Wallace, and Michael Canci. 2013. "Spatial Statistical Analysis of Tree Deaths Using Airborne Digital Imagery." *International Journal of Applied Earth Observation and Geoinformation* 21 (April): 418–26. doi:10.1016/j.jag.2012.04.006.
- Chu, Cheng-Jin, Fernando T. Maestre, Sa Xiao, Jacob Weiner, You-Shi Wang, Zheng-Hu Duan, and Gang Wang. 2008. "Balance between Facilitation and Resource Competition Determines Biomass–density Relationships in Plant Populations." *Ecology Letters* 11 (11): 1189–1197. doi:10.1111/j.1461-0248.2008.01228.x.
- Chu, Chengjin, John Norman, Robert Flynn, Nicole Kaplan, William K. Lauenroth, and Peter B. Adler. 2013. "Cover, Density, and Demographics of Shortgrass Steppe Plants Mapped 1997–2010 in Permanent Grazed and Ungrazed Quadrats." *Ecology* 94 (6): 1435–1435. doi:10.1890/13-0121.1.
- Chu, Cheng-Jin, Jacob Weiner, Fernando T. Maestre, You-Shi Wang, Charles Morris, Sa Xiao, Jian-Li Yuan, Guo-Zhen Du, and Gang Wang. 2010. "Effects of Positive Interactions, Size Symmetry of Competition and Abiotic Stress on Self-Thinning in Simulated Plant Populations." *Annals of Botany* 106 (4): 647–52. doi:10.1093/aob/mcq145.
- Chu, Cheng-Jin, Jacob Weiner, Fernando T. Maestre, Sa Xiao, You-Shi Wang, Qi Li, Jian-Li Yuan, Lu-Qiang Zhao, Zheng-Wei Ren, and Gang Wang. 2009. "Positive Interactions Can Increase Size Inequality in Plant Populations." *Journal of Ecology* 97 (6): 1401–1407. doi:10.1111/j.1365-2745.2009.01562.x.
- Coffin, Debra P., and William K. Lauenroth. 1991. "Effects of Competition on Spatial Distribution of Roots of Blue Grama." *Journal of Range Management* 44 (1): 68–71. doi:10.2307/4002642.
- Cressie, Noel. 2015. *Statistics for Spatial Data*. 2 edition. Hoboken, NJ: Wiley-Interscience.

- Csillag, Ferko, and Barry Boots. 2005. "Toward Comparing Maps as Spatial Processes." In *Developments in Spatial Data Handling*, 641–52. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/3-540-26772-7_48.
- . 2016. "Toward Comparing Maps as Spatial Processes." In *SpringerLink*, 641–52. Springer Berlin Heidelberg. Accessed November 8. http://0-link.springer.com.bianca.penlib.du.edu/chapter/10.1007/3-540-26772-7_48.
- Dale, Mark R. T., and Marie-Josée Fortin. 2014. *Spatial Analysis: A Guide For Ecologists*. Cambridge University Press.
- DeAngelis, Donald L., and Volker Grimm. 2014. "Individual-Based Models in Ecology after Four Decades." *F1000Prime Reports* 6 (June). doi:10.12703/P6-39.
- Diggle, Peter J. 2003. *Statistical Analysis of Spatial Point Patterns*. 2 edition. London; New York: Hodder Education Publishers.
- Getzin, Stephan, Robert S. Nuske, and Kerstin Wiegand. 2014. "Using Unmanned Aerial Vehicles (UAV) to Quantify Spatial Gap Patterns in Forests." *Remote Sensing* 6 (8): 6988–7004. doi:10.3390/rs6086988.
- Grimm, Volker, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss-Custard, et al. 2006. "A Standard Protocol for Describing Individual-Based and Agent-Based Models." *Ecological Modelling* 198 (1–2): 115–26. doi:10.1016/j.ecolmodel.2006.04.023.
- Grimm, Volker, Uta Berger, Donald L. DeAngelis, J. Gary Polhill, Jarl Giske, and Steven F. Railsback. 2010. "The ODD Protocol: A Review and First Update." *Ecological Modelling* 221 (23): 2760–68. doi:10.1016/j.ecolmodel.2010.08.019.
- Hammam, Yasser, Antoni Moore, and Peter Whigham. 2007. "The Dynamic Geometry of Geographical Vector Agents." *Computers, Environment and Urban Systems, Geospatial Analysis and Modeling*, 31 (5): 502–19. doi:10.1016/j.compenvurbsys.2007.08.003.
- Herben, Tomáš, Heinjo J. During, Richard Law, and others. 2000. "Spatio-Temporal Patterns in Grassland Communities." *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, 48–64.
- Illian, Dr Janine, Prof Antti Penttinen, Dr Helga Stoyan, and Dietrich Stoyan. 2008. *Statistical Analysis and Modelling of Spatial Point Patterns*. 1 edition. Chichester, England; Hoboken, NJ: Wiley-Interscience.
- Jost, Lou. 2006. "Entropy and Diversity." *Oikos* 113 (2): 363–75. doi:10.1111/j.2006.0030-1299.14714.x.
- Lauenroth, William K., and Ingrid C. Burke, eds. 2008. *Ecology of the Shortgrass Steppe: A Long-Term Perspective*. 1 edition. Oxford; New York: Oxford University Press.

- Law, Richard, and Ulf Dieckmann. 2000. "A Dynamical System for Neighborhoods in Plant Communities." *Ecology* 81 (8): 2137–48. doi:10.2307/177102.
- Law, Richard, Janine Illian, David F. R. P. Burslem, Georg Gratzer, C. V. S. Gunatilleke, and I. a. U. N. Gunatilleke. 2009. "Ecological Information from Spatial Patterns of Plants: Insights from Point Process Theory." *Journal of Ecology* 97 (4): 616–28. doi:10.1111/j.1365-2745.2009.01510.x.
- Legendre, Pierre, Daniel Borcard, and Pedro R. Peres-Neto. 2005. "ANALYZING BETA DIVERSITY: PARTITIONING THE SPATIAL VARIATION OF COMMUNITY COMPOSITION DATA." *Ecological Monographs* 75 (4): 435–50. doi:10.1890/05-0549.
- Legendre, Pierre, and Marie Josée Fortin. 1989. "Spatial Pattern and Ecological Analysis." *Vegetatio* 80 (2): 107–138.
- Levin, Simon A. 1992. "The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture." *Ecology* 73 (6): 1943–67. doi:10.2307/1941447.
- Lin, Yue, Uta Berger, Volker Grimm, Franka Huth, and Jacob Weiner. 2013. "Plant Interactions Alter the Predictions of Metabolic Scaling Theory." *Plos One* 8 (2): e57612. doi:10.1371/journal.pone.0057612.
- Lin, Yue, Franka Huth, Uta Berger, and Volker Grimm. 2014. "The Role of Belowground Competition and Plastic Biomass Allocation in Altering Plant Mass-Density Relationships." *Oikos* 123 (2): 248–56. doi:10.1111/j.1600-0706.2013.00921.x.
- Matsinos, Yiannis G., and Andreas Y. Troumbis. 2002. "Modeling Competition, Dispersal and Effects of Disturbance in the Dynamics of a Grassland Community Using a Cellular Automaton Model." *Ecological Modelling* 149 (1): 71–83.
- McGarigal, Kevin, and Barbara J. Marks. 1995. "FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure." <http://www.treearch.fs.fed.us/pubs/download/3064.pdf>.
- Nuske, Robert S., Susanne Sprauer, and Joachim Saborowski. 2009. "Adapting the Pair-Correlation Function for Analysing the Spatial Distribution of Canopy Gaps." *Forest Ecology and Management* 259 (1): 107–16. doi:10.1016/j.foreco.2009.09.050.
- Parrott, Lael. 2005. "Quantifying the Complexity of Simulated Spatiotemporal Population Dynamics." *Ecological Complexity* 2 (2): 175–84. doi:10.1016/j.ecocom.2004.11.004.
- Pennekamp, Frank, and Nicolas Schtickzelle. 2013. "Implementing Image Analysis in Laboratory-Based Experimental Systems for Ecology and Evolution: A Hands-on Guide." *Methods in Ecology and Evolution* 4 (5): 483–92. doi:10.1111/2041-210X.12036.
- Penttinen, Antti, Dietrich Stoyan, and Helena M. Henttonen. 1992. "Marked Point Processes in Forest Statistics." *Forest Science* 38 (4): 806–24.

- Perry, George L. W., Ben P. Miller, and Neal J. Enright. 2006. "A Comparison of Methods for the Statistical Analysis of Spatial Point Patterns in." *Plant Ecology* 187 (1): 59–82. doi:10.1007/s11258-006-9133-4.
- Perry, J. N., A. M. Liebhold, M. S. Rosenberg, J. Dungan, M. Miriti, A. Jakomulska, and S. Citron-Pousty. 2002. "Illustrations and Guidelines for Selecting Statistical Methods for Quantifying Spatial Pattern in Ecological Data." *Ecography* 25 (5): 578–600. doi:10.1034/j.1600-0587.2002.250507.x.
- Pluim, J.P.W., J.B.A. Maintz, and M.A. Viergever. 2003. "Mutual-Information-Based Registration of Medical Images: A Survey." *IEEE Transactions on Medical Imaging* 22 (8): 986–1004. doi:10.1109/TMI.2003.815867.
- Pollack, R., M. Sharir, and G. Rote. 1989. "Computing the Geodesic Center of a Simple Polygon." *Discrete & Computational Geometry* 4 (6): 611–26. doi:10.1007/BF02187751.
- Purves, D. W., and R. Law. 2002. "Fine-Scale Spatial Structure in a Grassland Community: Quantifying the Plant's-Eye View." *Journal of Ecology* 90 (1): 121–129. doi:10.1046/j.0022-0477.2001.00652.x.
- Railsback, and V. Grimm. 2011. *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton University Press. <http://books.google.com/books?id=tSI2DkMtWQC>.
- Rommel, Tarmo K., and Ferenc Csillag. 2006. "Mutual Information Spectra for Comparing Categorical Maps." *International Journal of Remote Sensing* 27 (7): 1425–52. doi:10.1080/01431160500419303.
- Ripley, Brian D. 2004. *Spatial Statistics*. Hoboken, N.J: Wiley-Interscience.
- Roe, Cailin M., Graham C. Parker, Annika C. Korsten, Christina J. Lister, Sam B. Weatherall, Rachael H.E. Lawrence Lodge, and J. Bastow Wilson. 2012. "Small-Scale Spatial Autocorrelation in Plant Communities: The Effects of Spatial Grain and Measure of Abundance, with an Improved Sampling Scheme." *Journal of Vegetation Science* 23 (3): 471–482. doi:10.1111/j.1654-1103.2011.01375.x.
- Russakoff, Daniel B., Carlo Tomasi, Torsten Rohlfing, and Calvin R. Maurer. 2004. "Image Similarity Using Mutual Information of Regions." In *Computer Vision - ECCV 2004*, 596–607. Springer, Berlin, Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-24672-5_47.
- Shannon, C. E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27 (3): 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
- Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. Boca Raton: Chapman and Hall.

- Silvertown, Jonathan, Senino Holtier, Jeff Johnson, and Pam Dale. 1992. "Cellular Automaton Models of Interspecific Competition for Space--The Effect of Pattern on Process." *Journal of Ecology* 80 (3): 527–33. doi:10.2307/2260696.
- Song, Lin, Peter Langfelder, and Steve Horvath. 2012. "Comparison of Co-Expression Measures: Mutual Information, Correlation, and Model Based Indices." *BMC Bioinformatics* 13: 328. doi:10.1186/1471-2105-13-328.
- Stoyan, Dietrich, and Dr Helga Stoyan. 1994. *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. 1 edition. Chichester ; New York: Wiley.
- Taillandier, Patrick, Duc-An Vo, Edouard Amouroux, and Alexis Drogoul. 2010. "GAMA: A Simulation Platform That Integrates Geographical Information Data, Agent-Based Modeling and Multi-Scale Control." In *Principles and Practice of Multi-Agent Systems*, edited by Nirmal Desai, Alan Liu, and Michael Winikoff, 242–58. Lecture Notes in Computer Science 7057. Springer Berlin Heidelberg. http://0-link.springer.com.bianca.penlib.du.edu/chapter/10.1007/978-3-642-25920-3_17.
- Thiele, Jan C., Winfried Kurth, and Volker Grimm. 2014. "Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and R." *Journal of Artificial Societies and Social Simulation* 17 (3): 11.
- Turkington, R., and John L. Harper. 1979. "The Growth, Distribution and Neighbour Relationships of *Trifolium Repens* in a Permanent Pasture: I. Ordination, Pattern and Contact." *Journal of Ecology* 67 (1): 201–18. doi:10.2307/2259345.
- Wagner, Helene H., and Marie-Josée Fortin. 2005. "SPATIAL ANALYSIS OF LANDSCAPES: CONCEPTS AND STATISTICS." *Ecology* 86 (8): 1975–87. doi:10.1890/04-0914.
- Wang, Jihuai, M.J. Kropff, B. Lammert, S. Christensen, and P.K. Hansen. 2003. "Using CA Model to Obtain Insight into Mechanism of Plant Population Spread in a Controllable System: Annual Weeds as an Example." *Ecological Modelling* 166 (3): 277–86. doi:10.1016/S0304-3800(03)00164-9.
- Whigham, Peter A. 2013. "Extending Point-Pattern Analysis to Polygons Using Vector Representations." <http://otago.ourarchive.ac.nz/handle/10523/4218>.
- Wiegand, Thorsten, Fangliang He, and Stephen P. Hubbell. 2013. "A Systematic Comparison of Summary Characteristics for Quantifying Point Patterns in Ecology." *Ecography* 36 (1): 92–103. doi:10.1111/j.1600-0587.2012.07361.x.
- Wiegand, Thorsten, W. Daniel Kissling, Pablo A. Cipriotti, and Martín R. Aguiar. 2006. "Extending Point Pattern Analysis for Objects of Finite Size and Irregular Shape." *Journal of Ecology* 94 (4): 825–37. doi:10.1111/j.1365-2745.2006.01113.x.

Wu, Jianguo, and John L David. 2002. "A Spatially Explicit Hierarchical Approach to Modeling Complex Ecological Systems: Theory and Applications." *Ecological Modelling* 153 (1–2): 7–26. doi:10.1016/S0304-3800(01)00499-9.

APPENDIX A

Complete GAMA model code

```
/**
 * Name: shape_dyn_test_SQUARES
 * Author:
 * Description:
 * Tags: Tag1, Tag2, TagN
 */

model shape_dyn_test_SQUARES

global torus: true {

    // Define world shape and size
    float worldDimension <- 100#m;
    geometry shape <- square(worldDimension);
    int simStep;
    float rPointProb;
    float rNodeProb;
    float rEdgeProb;

    // Initialize an agent of species: poly
    init {
        create poly number: 10;
    }

    // Save sci to csv file at 100 steps and end model run
    reflex save_sci_vals when: cycle = 100 {
        ask poly {
            save [name, simStep, rNodeProb, rEdgeProb, rPointProb, sci]
            to: "../sciTEST.csv" type: "csv";
        }
    }
}

// Grid species definition
grid habitat width: 25 height: 25 use_regular_agents: false use_individual_shapes: false
{
    int maxResource <- 100;
    int inputResource <- rnd(10);
    int resourceLoss <- rnd(10);
    int availableResource <- rnd(100) update: availableResource + inputResource -
resourceLoss
        max: maxResource min: 0;
    rgb color <- rgb((255 * (1-(availableResource*0.01))), 255, (255 * (1-
(availableResource*0.01))))
    update: rgb((255 * (1-(availableResource*0.01))), 255, (255 * (1-
(availableResource*0.01))));

    reflex resourceChange {
        inputResource <- rnd(10);
        resourceLoss <- rnd(10);
    }
}

// Dynamic polygon species definition
species poly {
    geometry shape;
    list<point> pointsOnShape;
    geometry polyEdge;
    int num_sides;
    int randomNode;
```

```

int previousRandNode;
int segmentAngle;
int outSegmentAngle;
int angleToDisplace;
list<point> pointsOnLine;
point randomPointOnLine;
int randAngle <- 90;
list<geometry> polyNeighbors;

//Scale the amount of offset at each step.
float offsetScale <- (1/ rnd(1, 4, 1));
list<point> polyNodes;
point nodeToDisplace;
int translateAngle;
float angleCos;
float angleSin;
geometry tempShape;
geometry intersectPolys;
geometry intersectingLines;

point randomPointTest;
geometry randomLineTest;
bool crossedTest;
bool crossedPoly;
list listEdgeLengths;
float min_edge_length <- 0.0005; // minimum edge length to apply randomPoint or edge
displacements
float sci;
list<float> sciList;
float medianSCI;

// Grid interaction variables
list<habitat> bufferNeighbors;
float maxResource;
int maxResourcePointIndex;
list<int> growthNodes;

init {
    // Initial polygon shapes are circles
    shape <- circle(0.5);

    // Initially generate points along the circle
    pointsOnShape <- shape points_on (shape.perimeter/10);
    shape <- polygon(pointsOnShape);
    polyNeighbors <- poly at_distance 20;
}

// _____
// Actions for changing shape
// _____

// Action to get a random edge of a polygon
geometry getPolyEdge (point pt1, point pt2) {
    geometry pEdge <- line([pt1, pt2]);
    return pEdge; // Returns an edge of the polygon
}

// Action for calculating random displacement coordinates (point translation)
// Input variables:
// angle_1: the base angle relative to the x-axis. It takes different meanings in the randomPointDisplace
// and nodeDisplace actions, and is calculated there.
// splitRange: for randomPointDisplace range does not need to be split since it is relative to the target
// polygon edge. For nodeDisplace, it does need to be split since the base angle is 1/2 the external
// angle between the adjacent edges (relative to the x-axis).

```

```

point translatePoint (int angle_1, bool splitRange, point pointToTranslate)
{
  if splitRange = true {
    translateAngle <- 0;
  } else {
    translateAngle <- randAngle;
  }

  angleCos <- cos(angle_1 + translateAngle)*offsetScale;
  angleSin <- -sin(angle_1 + translateAngle)*offsetScale; // negative since (0, 0) is
upper left corner of 'world'

  // Scale offset distance by random value
point translatedPoint <- point(pointToTranslate translated_by {angleCos,
angleSin});
  return translatedPoint;
}

// Action to calculate coordinates for line translation.
// Input variables:
// angle_1: the angle of the selected edge relative to the x-axis
// lineToTranslate: the polygon edge that will be translated.
geometry translateLine (int angle_1, geometry lineToTranslate) {
  translateAngle <- randAngle;
  angleCos <- cos(angle_1 + translateAngle)*offsetScale;
  angleSin <- -sin(angle_1 + translateAngle)*offsetScale; // negative since (0, 0) is
upper left corner of 'world'
  geometry trLineTemp <- lineToTranslate translated_by {angleCos, angleSin};
  return trLineTemp;
}

// _____ Choose Point Action _____
// Action to choose node with highest associated grid resource values
// get list of habitat cells that the shape point overlaps and total their values for each point.
list<int> choosePoint (list<point> pointOptions) {
  list<list<habitat>> pointHabitatCells;
  list<float> pointResources;
  bufferNeighbors <- (habitat overlapping(self));

  loop i from: 0 to: length(pointOptions) - 2 {
    add (bufferNeighbors where ((each distance_to pointOptions[i]) <
5)) to: pointHabitatCells;
    add (sum(pointHabitatCells[i] collect each.availableResource)) to:
pointResources;
  }

  // Create a map to sort by resource value, but preserve the original index
map<int, float> ptResMap;

  loop i from: 0 to: length(pointResources)-2 {
    ptResMap <+ i::pointResources[i];
  }

  // Get index values (keys) for map sorted by value
list<int> sortedMap <- ptResMap.keys sort_by ptResMap[each];
sortedMap <- reverse(sortedMap);
return sortedMap;
}

// _____ Random Point Displacement _____
// Action for growing by displacing a random point along a polygon edge.
action randomPointDisplace {

  // Create a line object that represents the (randomly) selected edge of the polygon
  // Using the getPolyEdge action.

```

```

+ 1));

polyEdge <- getPolyEdge(shape.points[randomNode], shape.points[randomNode
only
if (polyEdge.perimeter > min_edge_length) {

    // Generate x number of points along the given line segment.
    pointsOnLine <- polyEdge points_on (polyEdge.perimeter/2); // midpoint

    // Choose one of the points along the line segment
    randomPointOnLine <- pointsOnLine[1]; // the midpoint in this case

    // Calculate the angle of the polygon edge relative to the x-axis
    // This is used later to set the angle of displacement for the point
    segmentAngle <- angle_between({polyEdge.points[0].location.x,
polyEdge.points[0].location.y},
{polyEdge.points[1].location.x,
polyEdge.points[1].location.y},
{polyEdge.points[0].location.x + 10,
polyEdge.points[0].location.y});

    // Translate the point using translatePoint Action
    randomPointTest <- translatePoint(segmentAngle, false,
randomPointOnLine);

    // Create a polyline from the adjacent polygon nodes to the new point to test for
intersection
tempShape <- polyline({polyEdge.points[0].x,
polyEdge.points[0].y}, {randomPointTest.x, randomPointTest.y},
{polyEdge.points[1].x, polyEdge.points[1].y});

    //Get the intersection of the new triangle, and the original shape.
    crossedTest <- tempShape crosses shape;
    list intersectTest <- polyNeighbors overlapping(tempShape);
    if !(crossedTest) and (length(intersectTest) = 0) {
        // Update the polygon shape with the new node location
        // Note: The first / last point in the polygon will not be changed in this action (see
pointsOnLine section above), so
not an issue here (see nodeDisplace action
        // dealing with the first/last point duplication in the shape.points list is
        // for an example of when it is an issue).
        polyNodes <- list(shape.points);
        polyNodes[randomNode + 1] <-- ({randomPointTest.x,
randomPointTest.y});
        shape <- polygon(polyNodes);
    }
}

// _____ Node Displacement _____
// Action for displacing a node in the polygon
action nodeDisplace {

    // Get the point associated with that index
    nodeToDisplace <- shape.points[randomNode];

    // Get the external angle between lines adjacent to the node and divide by 2.
    // This is used later to set the angle of displacement for the node.
    if (randomNode = 0) {
        previousRandNode <- length(shape.points)-2;
    } else {
        previousRandNode <- randomNode -1;
    }
}

```

```

        segmentAngle <- angle_between({shape.points[randomNode].location.x,
shape.points[randomNode].location.y},
        {shape.points[previousRandNode].location.x,
shape.points[previousRandNode].location.y},
        {shape.points[randomNode + 1].location.x, shape.points[randomNode +
1].location.y});
        segmentAngle <- int (segmentAngle/2);

        // Get angle of outgoing line relative to the x-axis
        outSegmentAngle <- angle_between({shape.points[randomNode].location.x,
shape.points[randomNode].location.y},
        {shape.points[randomNode + 1].location.x,
shape.points[randomNode + 1].location.y},
        {shape.points[randomNode].location.x + 10,
shape.points[randomNode].location.y});

        // Calculate base offset angle for node. This splits the angle between the adjacent lines and accounts
for their
        // angle relative to the x-axis. NOTE: if baseOffsetAngle is > 360 it is corrected automatically (i.e.,
baseOffsetAngle - 360)
        int baseOffsetAngle <- segmentAngle + outSegmentAngle;

        // Translate the node using translatePoint Action
        // Note: in this case, splitAngle parameter needs to be set to 'true'
        randomPointTest <- translatePoint(baseOffsetAngle, true, nodeToDisplace);

        // Create a polyline from the adjacent polygon edges to the new point to test for intersection
        tempShape <- polyline({{shape.points[previousRandNode].location.x,
shape.points[previousRandNode].location.y},
        {randomPointTest.location.x, randomPointTest.location.y},
        {shape.points[randomNode + 1].location.x, shape.points[randomNode +
1].location.y}});

        //Get the intersection of the new triangle, and the original shape.
        crossedTest <- tempShape crosses shape;
        list intersectTest <- polyNeighbors overlapping(tempShape);
        if !(crossedTest) and (length(intersectTest) = 0) {

                // Update the polygon shape with the new node location
                polyNodes <- list(shape.points);
                if (randomNode = 0) {
                        polyNodes[randomNode] <- ({randomPointTest.x,
randomPointTest.y});
                        polyNodes[length(polyNodes) - 1] <- ({randomPointTest.x,
randomPointTest.y});
                } else {
                        polyNodes[randomNode] <- ({randomPointTest.x,
randomPointTest.y});
                }
                shape <- polygon(polyNodes);
        }
}

//_____ Edge Displacement_____
// Action for displacing an edge of the polygon
action edgeDisplace {

        // Create a line object that represents the (randomly) selected edge of the polygon |
        // Using the getPolyEdge action.
        int nextRandNode;
        if (randomNode = length(shape.points) - 2) {
                nextRandNode <- 0;
        } else {
                nextRandNode <- randomNode + 1;

```

```

    }
    polyEdge <- getPolyEdge(shape.points[randomNode],
shape.points[nextRandNode]);

    if (polyEdge.perimeter > min_edge_length) {

        // Calculate the angle of the polygon edge relative to the x-axis
        // This is used later to set the angle of displacement for the edge
        segmentAngle <- angle_between({polyEdge.points[0].location.x,
polyEdge.points[0].location.y},
{polyEdge.points[1].location.x,
polyEdge.points[1].location.y},
{polyEdge.points[0].location.x + 10,
polyEdge.points[0].location.y});

        // Translate the line using translateLine action
        randomLineTest <- translateLine(segmentAngle, polyEdge);

        // Create a polyline from the adjacent polygon edges to the new point to test for
intersection
        tempShape <- polyline([shape.points[randomNode].location.x,
shape.points[randomNode].location.y},
{randomLineTest.points[0].location.x,
randomLineTest.points[0].location.y},
{randomLineTest.points[1].location.x,
randomLineTest.points[1].location.y},
{shape.points[randomNode + 1].location.x,
shape.points[randomNode + 1].location.y}]);

        bool crosses <- tempShape crosses shape;
        bool inters <- randomLineTest overlaps shape;
        list intersectTest <- polyNeighbors overlapping(tempShape);
        if !(crosses) and !(inters) and (length(intersectTest) = 0) {
            // Update the polygon shape with the new node location
            polyNodes <- list(shape.points);
            polyNodes[randomNode + 1] +<- ({randomLineTest.points[0].x,
randomLineTest.points[0].y});
            polyNodes[randomNode + 2] +<- ({randomLineTest.points[1].x,
randomLineTest.points[1].y});
            shape <- polygon(polyNodes);
        }
    }
}

//-----
reflex grow {

    growthNodes <- choosePoint(shape.points);
    int randInt <- rnd(0, (length(growthNodes)/4), 1); // random int between 0 and
(value)

    randomNode <- growthNodes[randInt];
    int actionChoice <- rnd_choice([rPointProb, rNodeProb, rEdgeProb]);

    // Compute Shape Complexity Index
    sci <- shape.perimeter/(2*#pi*(sqrt(shape.area/#pi)));

    if (actionChoice = 0) {

        do randomPointDisplace;

    } else if (actionChoice = 1) {
        do nodeDisplace;

    } else {
        do edgeDisplace;
    }
}

```

```

// _____
//           Reflex for polygon simplification
if every(simStep) {
  pointsOnShape <- shape points_on
(shape.perimeter/(length(shape.points) * 1));
  shape <- polygon(pointsOnShape);
}

ask (habitat overlapping shape) {
  availableResource <- availableResource - 30;
}

ask bufferNeighbors {
  availableResource <- availableResource - 10;
}

}

aspect standard_aspect {
  draw shape empty: false color: #grey;
}
}

// _____

// Batch Experiment code

experiment edgeDisplace_experiment type:batch keep_seed: false repeat:5 until: (time =
101){
  parameter "Steps until shape simplification: " var: simStep among: [1, 10, 20, 30,
40, 50] init: 20;
  parameter "Random Node Probability: " var: rNodeProb min:0.1 max: 1.0 step:0.2
init: 0.6;
  parameter "Random Edge Probability: " var: rEdgeProb min:0.0 max: 0.2 step:0.1
init: 0.1;
  parameter "Random Point Probability: " var:rPointProb min:0.0 max:0.2 step:0.1
init: 0.1;
  //parameter "Directional Growth Parameter" var:randInt min: 1 max: 18 step: 3 init: 6;
}

```