SPATIAL ANALYSIS OF GRASSHOPPER DENSITY AS INFLUENCED BY ANTHROPOGENIC HABITAT CHANGES

Bahman Shafii
William J. Price
Dennis J. Fielding
Merlyn A. Brusven

Follow this and additional works at: http://newprairiepress.org/agstatconference

Part of the Agriculture Commons, and the Applied Statistics Commons

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.

Recommended Citation

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.
Spatial Analysis of Grasshopper Density as Influenced by Anthropogenic Habitat Changes

Bahman Shafii and William J. Price
Statistical Programs
Dennis J. Fielding and Merlyn A. Brusven
Division of Entomological Sciences
College of Agriculture
University of Idaho
Moscow, ID 83844

ABSTRACT

The rangeland environment in southern Idaho has been heavily impacted by human activities. Invasion by exotic plant species, frequent fires, grazing pressure, and other ecological disturbances have greatly affected the structure and dynamics of grasshopper populations. Quantification of spatial patterns of grasshopper density and species composition is important in order to determine their influence on grassland ecosystems, as well as evaluating managerial decisions concerning vegetation manipulations, grazing practices, and spraying programs. A spatial statistical approach to modeling the heterogeneity of grasshopper populations is presented, and the impact of vegetation and grazing treatments on grasshopper density is investigated. Empirical applications are demonstrated with reference to repeated field surveys conducted over several years in south central Idaho.

Keywords: Spatial dependence, modeling, variogram, grasshoppers

I. INTRODUCTION

Grasshoppers are considered one of the most abundant arthropods in the intermountain sagebrush ecoregion of the western United States. Due to their influence on grassland ecosystems (Mitchell and Pfadt, 1974), as well as their impact on ranching-based economic systems (Davis, et.al., 1993), the potential effects of range management practices on grasshoppers are of great concern to those interested in the health of the rangeland environments.

Ecologically diverse plant communities representing a mosaic of macro and micro environments exist on the rangeland of south central Idaho. In addition to this natural variation, the region has a history of anthropogenic disturbances, including grazing, fire, and introduction of exotic plant species (Young and Evans, 1978), particularly two species of grasses, crested wheatgrass (Agropyron desertorum Fisch ex. Link) and cheatgrass (Bromus tectorum L.). Crested wheatgrass is a perennial bunchgrass that is frequently seeded to help rehabilitate degraded rangelands. Cheatgrass is an invasive annual grass that has rapidly spread throughout the intermountain region.
The above ecological disturbances have greatly altered the spatial heterogeneity of grasshopper environments. Several studies have investigated the grasshopper community composition as influenced by changes in habitat [Evans (1988), Quinn and Walgenbach (1990), Kemp et. al. (1990), Fielding and Brunsven (1993)]. However, no attempt has been made to quantify the underlying spatial structure of ecological processes. Such a strategy will involve modeling a given response characteristic after the spatial dependence has been modeled. This will in turn exploit the correlation structure to obtain more precise estimators of the specified model parameters.

The purpose of this study is to present a statistical approach to modeling the spatial structure of grasshopper populations and to determine the impact of ecological disturbances such as vegetation and grazing on grasshopper density in south central Idaho.

II. METHODS

Variogram Estimation and Fitting

The first step in modeling the spatial dependence is the specification and estimation of the sample variogram. Let \( \{x_1, \ldots, x_n\} \) represent the set of coordinate positions (spatial locations) in the field (index set) \( D \) with corresponding observed data \( \{z(x_1), \ldots, z(x_n)\} \). Then the regionalized variable (Matheron, 1971) \( z(x_i) \) can be considered a realization of the set of random variables \( Z(X) \) for all \( x_i \) in \( D \),

i.e.: \( \{Z(X) : X \in D\} \), \( D \subset \mathbb{R}^d \).  

The classical variogram estimator is defined as

\[
2\gamma(h) = \frac{1}{N(h)} \sum \{z(x_i) - z(x_i+h)\}^2
\]

which is estimated from the sample by

\[
2\hat{\gamma}(h) = \frac{1}{N(h)} \sum \{z(x_i) - z(x_i+h)\}^2, \ h \in \mathbb{R}^d.
\]

The quantity \( 2\hat{\gamma}(\cdot) \) [conceptually, mean-squared difference or variance] is called an estimated variogram (Matheron, 1962), and \( \hat{\gamma}(\cdot) \) is called an estimate of the semivariogram; \( N(h) \) is the number of distinct pairs of observations \( [z(x_i) - z(x_i+h)] \) separated by the distance/lag vector \( h \).

Fitting a theoretical model to the estimated variogram is the next step in quantifying the spatial dependence. While a number of parametric variogram models have been suggested [Journal and Huijbregts (1978)], only the following isotropic [i.e., when \( 2\gamma(h) \) depends only on the magnitude and not the direction of vector \( h \)] models were considered in this study:
Linear

\[ 2\gamma(h) = \begin{cases} 
  c_0 + \beta h, & 0 < h \leq r_L \\
  c_0 + c_L, & h \geq r_L 
\end{cases} \]  
(4)

Where \( c_0 \geq 0 \) is the nugget effect, \( c_0 + c_L \) is the sill, and \( r_L \) is the range of the variogram;

Power

\[ 2\gamma(h) = \begin{cases} 
  0, & h = 0 \\
  c_0 + \beta_p(h)^\lambda, & h \neq 0 
\end{cases} \]  
(5)

where \( \lambda < 2 \);

Gaussian

\[ 2\gamma(h) = \begin{cases} 
  0, & h = 0 \\
  C_0 + C_0[1 - \exp\{-h/\theta_0\}^2], & h \neq 0 
\end{cases} \]  
(6)

Spherical

\[ 2\gamma(h) = \begin{cases} 
  0, & h=0 \\
  c_0 + c_s[(3/2)h/\theta_s - (1/2)(h/\theta_s)^3], & 0 < h \leq \theta_s \\
  c_0 + c_s, & h \geq \theta_s 
\end{cases} \]  
(7)

where \( c_0 \geq 0, c_s \geq 0, \theta_s \geq 0 \); and

Wave (hole-effect) (Cressie, 1991)

\[ 2\gamma(h) = \begin{cases} 
  0, & h=0 \\
  c_0 + c_w[1 - r_w\sin(h/r_w)/h], & h \neq 0 
\end{cases} \]  
(8)

where \( c_0 \geq 0, c_w \geq 0, r_w \geq 0 \).

Variogram models were fitted to the specified estimators using linear and nonlinear least square methods.
Response Model Estimation and Inference

The basic linear model for the spatial data may be written as

\[ Z = X\beta + \delta \tag{9} \]

Where

\[ Z = [z(x_1), \ldots, z(x_n)]', \]
\[ X = n \times p \text{ matrix of explanatory variables}, \]
\[ \beta = (\beta_1, \ldots, \beta_p)', \]
\[ \delta = [\delta(x_1), \ldots, \delta(x_n)]'. \]

Under classical (non spatial) assumption of independent and identically distributed errors, the ordinary least squares estimator of \( \beta \) is given by

\[ \hat{\beta}_{OLS} = (X'X)^{-1}X'Z \]

which is equivalent to the maximum likelihood estimator if \( \delta(\cdot) \) is a Gaussian process. However, when the error process exhibits spatial correlation, the generalized least square (GLS) estimator given by

\[ \hat{\beta}_{GLS} = (X'\Sigma X)^{-1}X'\Sigma XZ \tag{10} \]

provides the appropriate estimators which is often more efficient (Searle, 1971). In the spatial context, \( \Sigma = \text{Var}(\delta) \) is an \( n \times n \) symmetric, nonnegative definite matrix whose elements are determined according to the underlying spatial structure of the scientific problem (see the application section).

The general spatial model (9) implies different models (possessing different variance-covariance structures) depending on the choice of \( \Sigma \). This will in turn affect the resulting variance estimate for the GLS estimator (10) given by

\[ \text{Var}(\hat{\beta}_{GLS}) = (X'\Sigma X)^{-1} \tag{11} \]

Let \( W \) be an \( n \times n \) positive definite weight matrix. Then it follows that \( W = \Sigma, W = \text{diag}(\Sigma), \) and \( W = I \) represent respectively, the full spatial model, the heteroskedastic model, and the classical model. Note that the classical model is a special case of the spatial model for which \( \Sigma = \sigma^2 I \).

Estimation of Mean Effects

The following analysis of variance (ANOVA) model was used to estimate large-scale treatment differences:

\[ Z_{ijk} = \mu + \tau_k + \beta_{jk} + \epsilon_{ijk} \tag{12} \]
Where \( Z_{ijk} \) = response on the ith unit of the jth replicate in the kth treatment,
\[
\mu = \text{constant mean},
\]
\[
\tau_k = \text{effect due to the kth treatment},
\]
\[
\beta_{jk} = \text{effect due to the jth replicate of the kth treatment},
\]
\[
\epsilon_{ijk} = \text{random error with mean 0 and covariance matrix } \sigma^2 \Sigma \text{ given in (10).}
\]

Based on the above development, a general ANOVA table can be constructed as follows (Wong, 1989):

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>( kj - 1 )</td>
<td>( \beta_{GLS}X'W^{-1}Z - [(1'W^{-1}1)^{-1}(Z'W^{-1}1')W^{-1}Z] )</td>
</tr>
<tr>
<td>Treatment</td>
<td>( k - 1 )</td>
<td>( SS_{GLS}(Z_{ijk} = \mu + \tau_k + \epsilon_{ijk}) )</td>
</tr>
<tr>
<td>Block</td>
<td>( k(j - 1) )</td>
<td>By Subtraction</td>
</tr>
<tr>
<td>Residual</td>
<td>( n - kj )</td>
<td>( Z'W^{-1}Z - \beta_{GLS}X'W^{-1}Z )</td>
</tr>
</tbody>
</table>

\[ C. \ Total \ n - 1 \quad Z'W^{-1}Z - [(1'W^{-1}1)^{-1}(Z'W^{-1}1')W^{-1}Z] \]

The GLS residuals defined as
\[
\epsilon_{GLS} = Z - \hat{X}\beta_{GLS} = [W - (X'W^{-1}X)^{-1}X']W^{-1}Z \quad (13)
\]
may be then used for the purpose of model validation.

### III. Application

**Description of Data**

The study area included portions of nine pastures within four allotments managed by the USDI Bureau of Land Management under a rest-rotation grazing system. The sampling plan was implemented according to a fixed hexagonal grid system (\( n = 72 \)) with about 0.8 km between rows and sample points within rows covering a total of ca. 4,800 ha. The grid was designed to reflect the scale of grazing patterns on the landscape, with 3-4 sampling points across each pasture.

Density of all grasshoppers were estimated by counting the number of grasshoppers flushed from 20, 0.5 m² rings set out the previous day in an approximate 50 m diameter circular transect. Proportions of individual grasshopper species were estimated by sweep net samples and visual counts. Vegetation was sampled by estimating the percentage ground
cover by plant species within 20, 0.25 m² quadrats during June and August of each year. Forage utilization by livestock was also estimated at each sampling point. Visual estimates were made of the percentage of aboveground biomass removed from each of 40 plants in each of the two or three dominant plant species at a sampling point.

Observations were made in June and August of 1991, 1992, and 1993, however, only the August data are considered here. Vegetative cover and grazing intensity were used as a means of classifying the study area into vegetation (cheatgrass, crested wheatgrass) or grazing (high, low) treatments based on the medians of these variables in each year (Figure 1). Due to limited space, detailed spatial analyses are given for 1991 only, while results are summarized for all three years. *Melanoplus sanguinipes* F. was the dominant species of grasshopper in the study area, comprising 40 to 50% of the total grasshopper population. Accordingly, *M. sanguinipes* F. was the only species treated in this report.

All statistical computations were carried out using SAS/STAT (1991) or SAS/IML (1990).

**Variogram Estimation and Modeling**

Directional trends in grasshopper densities were examined within each of the vegetation and grazing classes. The north-south trend indicated the most prominent pattern for each class and was, therefore, chosen as the primary direction. Using the east-west direction as replication, averages (mean and median) for the north-south trends were computed for each classification variable and plotted over distance (Fig 2). Median-based removal was then applied to obtain stationary residual densities (Fig 3).

The sample variogram estimator given in (2) was calculated separately for each vegetation and grazing classification using the median removed residuals as the response. All classifications showed the presence of spatial variability. Specifically, both vegetation classes and the high grazing treatment had strong increasing trends, while this was not evident in the low grazing treatment.

The isotropic models given in (4), (5), (6), (7), and (8) were fitted individually to each classification sample variogram using linear and nonlinear least squares. Figure 4 gives the fitted models for each treatment in 1991. The number of lags varied among treatments and in most cases, the last two or three lags were dropped due to the limited spatial information contained in these points. Each classification resulted in a unique variogram model. Cheatgrass was determined to be best estimated by the Gaussian model (6), while crested wheatgrass was best estimated with the Power model (5). The high and low grazing treatments were estimated using the Gaussian and Spherical (7) models, respectively.

**Construction of the Weight Matrix**

The following covariance structure was assumed for both the vegetation and grazing classifications:

$$\text{var}(z) = \sigma^2 W = \sigma^2 \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}.$$  

The weight matrix, $W$, is a $72 \times 72$ block diagonal matrix where the size of each block, $\Sigma_i$, $i=1, 2$, is determined by the size of the $i$th classification level. The respective elements of
the specified submatrices were determined based on the estimated variogram models as follows:

Vegetation

Crested Wheatgrass:
\[
\begin{align*}
\sigma^2 \Sigma_1 &= \text{cov}(z(x_i), z(x_{i+1})) \\
&= \sigma^2(0.354 - \hat{\gamma}_{cw}(\|x_i - x_j\|)),
\end{align*}
\]

Cheatgrass:
\[
\begin{align*}
\sigma^2 \Sigma_2 &= \sigma^2(1.75 - \hat{\gamma}_{CO}(\|x_i - x_j\|)),
\end{align*}
\]

High Grazing:
\[
\begin{align*}
\sigma^2 \Sigma_1 &= \sigma^2(3.22 - \hat{\gamma}_{H}(\|x_i - x_j\|)),
\end{align*}
\]

Low Grazing:
\[
\begin{align*}
\sigma^2 \Sigma_2 &= \sigma^2(3.65 - \hat{\gamma}_{L}(\|x_i - x_j\|)).
\end{align*}
\]

Modeling Grasshopper Density

Following the analysis of variance model given in (12), the generalized least squares technique was used to estimate the effects of vegetation complexes or grazing intensity on grasshopper densities. For the purpose of comparison, a standard ANOVA model was also fitted. The results are summarized in Table 1. Both vegetation and grazing showed significant reductions in the magnitude of the residual mean square (RMS) under the spatial model. For example, the standard analysis for vegetation had an RMS of .77, while the spatial analysis reduced this to .28. The vegetation effect in both cases was highly significant (p < .0062). A similar reduction in the RMS for grazing resulted in a change in the significance of the grazing effect from the standard analysis (p = .2254) to the spatial analysis (p = .0385), indicating an effect of grazing on grasshopper densities. Spatial analyses also provided least squares means with magnitudes similar to those of the standard analysis, but with reduced standard errors in most cases. Cheatgrass, for example, had a least square mean estimate of .91 and a standard error of .15 in the standard analysis, while the corresponding spatial analysis resulted in a similar least square mean estimate with a reduced standard error of .08. Plots of standardized residuals versus predicted values also confirmed the increase in overall precision as well as a reduction in heterogeneity with the spatial model (Fig 5). Under the standard ANOVA model, vegetation residuals showed a spreading or fanning pattern, while the same plot under the spatial model indicated a negligible pattern of residuals and a significant reduction in the magnitudes. Residual plots for grazing also showed a similar reduction in magnitudes under the spatial model with a less evident pattern of heterogeneity.

Similar analyses were carried out for August data in 1992 and 1993. Plots of fitted variogram models for these years are given in Figure 6 and Figure 7, respectively. Although the modeling process was essentially the same in all years, the crested wheatgrass treatment for 1992 differed in that it showed no spatial variability. Construction of the weight matrix for this classification was therefore carried out assuming constant variability based on the
ACKNOWLEDGEMENTS

Contribution from the College of Agriculture, University of Idaho, Idaho Agricultural Experiment Station paper number 9501.
REFERENCES


Table 1. Standard and Spatial Analyses for Vegetation and Grazing Treatments in 1991.

### 1991 Vegetation Analyses

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Veg</td>
<td>1</td>
<td>6.46</td>
<td>8.4</td>
<td>.0062</td>
</tr>
<tr>
<td>rep(Veg)</td>
<td>33</td>
<td>0.49</td>
<td>0.6</td>
<td>.8956</td>
</tr>
<tr>
<td>error</td>
<td>37</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. Total 71 0.72

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Veg</td>
<td>1</td>
<td>8.53</td>
<td>30.1</td>
<td>.0001</td>
</tr>
<tr>
<td>rep(Veg)</td>
<td>33</td>
<td>0.73</td>
<td>2.6</td>
<td>.0030</td>
</tr>
<tr>
<td>error</td>
<td>37</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. Total 71 0.61

<table>
<thead>
<tr>
<th>LSMEAN</th>
<th>Std Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheatgrass</td>
<td>0.92 0.15</td>
</tr>
<tr>
<td>Crested Wheat</td>
<td>0.22 0.16</td>
</tr>
</tbody>
</table>

### 1991 Grazing Analyses

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graze</td>
<td>1</td>
<td>1.07</td>
<td>1.5</td>
<td>.2254</td>
</tr>
<tr>
<td>rep(Grz)</td>
<td>26</td>
<td>0.73</td>
<td>1.0</td>
<td>.4621</td>
</tr>
<tr>
<td>error</td>
<td>44</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. Total 71 0.72

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graze</td>
<td>1</td>
<td>1.55</td>
<td>4.5</td>
<td>.0385</td>
</tr>
<tr>
<td>rep(Grz)</td>
<td>26</td>
<td>2.13</td>
<td>6.3</td>
<td>.0001</td>
</tr>
<tr>
<td>error</td>
<td>44</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. Total 71 1.01

<table>
<thead>
<tr>
<th>LSMEAN</th>
<th>Std Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.46 0.14</td>
</tr>
<tr>
<td>Low</td>
<td>0.82 0.17</td>
</tr>
</tbody>
</table>
Table 2. LSMEANS and Standard Errors for Vegetation and Grazing Treatments in 1991, 1992, and 1993.

<table>
<thead>
<tr>
<th>Vegetation Types Cheatgrass (CG) and Crested Wheatgrass (CW) (#/m²)</th>
<th>1991</th>
<th>1992</th>
<th>1993</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.E.</td>
<td>Pr&gt;F</td>
</tr>
<tr>
<td>CG</td>
<td>0.91</td>
<td>0.086</td>
<td>.0001</td>
</tr>
<tr>
<td>CW</td>
<td>0.21</td>
<td>0.083</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grazing Levels High (H) and Low (L) (#/m²)</th>
<th>1991</th>
<th>1992</th>
<th>1993</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.E.</td>
<td>Pr&gt;F</td>
</tr>
<tr>
<td>H</td>
<td>0.45</td>
<td>0.153</td>
<td>.0385</td>
</tr>
<tr>
<td>L</td>
<td>0.93</td>
<td>0.073</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Spatial arrangement of vegetation and grazing treatments in 1991.
Figure 2. Plots of mean and median grasshopper density for the vegetation and grazing treatments in 1991.
Residual Density Trends

Figure 3. Median removed residual trends for vegetation and grazing treatments in 1991.
Figure 4. Fitted variogram models for vegetation and grazing treatments in 1991.
Figure 5. Standardized residuals plotted vs predicted value of grasshopper density for vegetation and grazing treatments in 1991.
Figure 6. Fitted variogram models for vegetation and grazing treatments in 1992.
Figure 7. Fitted variogram models for vegetation and grazing treatments in 1993.