Environ Ecol Stat (2007) 14:399–410
DOI 10.1007/s10651-007-0027-y

Analyzing spatial count data, with an application to weed counts

Willem Kruijer · Alfred Stein · Willem Schaafsma · Sanne Heijting

Received: 1 April 2005 / Revised: 1 January 2006 / Published online: 6 September 2007

Abstract Count data on a lattice may arise in observational studies of ecological phenomena. In this paper a hierarchical spatial model is used to analyze weed counts. Anisotropy is introduced, and a bivariate extension of the model is presented.

Keywords Markov random fields · Anisotropy · Multivariate count data · Precision agriculture

1 Introduction

In precision agriculture, an efficient and environmentally friendly treatment of weeds is a major challenge (Rew and Cousens 2001; Kropff et al. 1997; Stein 2001). Weeds in a field are a threat for valuable crops, as their quality may be affected, either by the presence of damaging substances or by the reduction in yield. To optimize yield, spraying of herbicides is often required, which in most cases is still done by applying the same dosage to the whole field.

W. Kruijer (✉)
Department of Mathematics, Vrije Universiteit Amsterdam,
Boelelaan 1081a, 1081 HV, Amsterdam, The Netherlands
e-mail: kruijer@few.vu.nl

A. Stein
International Institute for Geo-Information Science and Earth Observation,
P.O. Box 6, 7500 AA, Enschede, The Netherlands
e-mail: stein@itc.nl

W. Schaafsma
Department of Mathematics, Rijksuniversiteit Groningen,
P.O. Box 800, 9700 AV, Groningen, The Netherlands

S. Heijting
Department of Crop and Weed Ecology, Wageningen University,
P.O. Box 430, 6700 AK, Wageningen, The Netherlands
e-mail: sanne.heijting@wur.nl

Environmental and economical considerations, however, require a minimization of herbicide use and hence dose amounts that vary with location.

Precision agriculture may contribute to this aim by using high-technology equipment to optimize herbicide application. A requirement is that weeds are properly characterized, both in species and in location. Spatial statistics allows the characterization of weed patterns based on a quantitative analysis. Attention has also been given to an integrated approach toward site specific weed management (SSWM) (Kropff et al. 1997). As many weeds occur in patches, it is natural to separate the abundance into trend and noise components as in Hess et al.'s (2001) investigation of the weed *Striga hermonthica* (Del.) Benth. in a Sahelian farmer's field. Such an approach is of interest in precision agriculture, as weeds with a uniform pattern require a uniform treatment, whereas obvious site-dependencies will require a site-specific application of herbicides. By making this distinction, a reduction in herbicide use may be achieved.

We extend this model by introducing anisotropy and by linking results for different weed species, in an attempt to model weed occurrence in a concise, economically more viable way. The aim of this study is to develop a spatial statistical model that is suitable for describing weed occurrence in a farmer's field. Our study will be illustrated with data of three different weed species observed in 2001 in a single field in Wageningen, The Netherlands.

2 Methodology

Our starting point is the model used by Hess et al. (2001), which is a modified version of the model introduced by Besag et al. (1991). As concerns notation, stochastic variables are denoted by capitals and their realizations by lower case characters. We analyze data collected on contiguous quadrats (Diggle 2003, p. 23). The set of quadrats S is labeled $S = \{1, 2, \ldots, n\}$. To specify spatial dependence, a pre-defined neighborhood structure is used. For every pair (i, j), the quadrats associated with i and j are neighbors ($i \sim j$) if they are adjacent in two orthogonal directions, say the y- and the x-directions. For the set of i's neighbors we write $S_i = \{i_1, i_2, \ldots, i_{m_i}\}$. The summation $\sum_{i=1}^{n} \sum_{j \in S_i}$ denotes the summation over all pairs of neighboring quadrats (i, j). Note that in this summation every pair is included twice. Let the elements of an $n \times n$ matrix A be denoted by a_{ij}, with A^t as its transpose. Finally, m_i equals the number of neighbors of quadrat i, i.e., $m_i = 2$ if the i^{th} quadrat is a corner quadrat, $m_i = 3$ if it is an edge quadrat and $m_i = 4$ otherwise.

2.1 Model specification

Our model includes a spatial autoregressive process $U = (U_1, U_2, \ldots, U_n)$ representing spatial variation, and a noise vector $V = (V_1, V_2, \ldots, V_n)$ to allow for extra variation. Given $U = u$ and $V = v$, the observed counts y_i are realizations of independent Poisson random variables Y_i, with intensities

$$\lambda_i = \beta \cdot e^{u_i + v_i}. \tag{1}$$

The parameter β expresses the overall mean, whereas the factors $e^{u_i+v_i}$ account for local deviations from this mean. In disease mapping, β is often referred to as the relative risk. In this paper it will be fixed, and assumed to be equal to the mean of the observed counts.

We assume that U follows a multivariate normal distribution with covariance matrix A^{-1}, where A is symmetric with $a_{ii} = \tau_1 m_i$ on the diagonal, $a_{ij} = a_{ji} = -\tau_1 \gamma$ if $i \sim j$ and $a_{ij} = 0$ otherwise. The parameter γ specifies the strength of spatial dependence and τ_1 is the precision of the process U. We only consider the case that $\gamma \in (0, 1)$, leading to an invertible

precision matrix A. The vector U is a *conditional autoregressive* (CAR) process (see, for example, Cliff and Ord 1980, Sect. 6.2) with density

$$\frac{\sqrt{|A|}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}u^t Au\right) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\left(\tau_1 \sum_{i=1}^{n} m_i {u_i}^2 - \tau_1 \gamma \sum_{i=1}^{n} \sum_{j \in S_i} u_i u_j\right)\right\} \quad (2)$$

The conditional distribution of U_i given its occurrence at the other locations, therefore, is the following univariate normal distribution :

$$U_i \mid \{U_j = u_j, j \neq i\} \sim N\left(\frac{\gamma}{m_i} \sum_{j \in S_i} u_j, \frac{1}{m_i \tau_1}\right). \quad (3)$$

In contrast to U, V has no spatial structure. The V_i are *i.i.d* normal with precision τ_2, and independent of U.

One could also directly consider $U + V$, which is normally distributed with covariance matrix $A^{-1} + \frac{1}{\tau_2} I_n$. The present decomposition into U and V however, facilitates the estimation of τ_1 and τ_2. A Bayesian approach is taken to obtain estimates of τ_1,τ_2,γ and β. In Sect. 2.4 we specify priors $p(\tau_1)$, $p(\tau_2)$ and $p(\gamma)$, and describe the Gibbs sampler we use to sample from

$$\begin{aligned} &p(u, v, \beta, \gamma_1, \gamma_2, \tau_1, \tau_2 \mid y) \propto \\ &\quad p(u \mid \gamma_1, \gamma_2, \tau_1) p(v \mid \tau_2) p(\gamma_1, \gamma_2) p(\tau_1) p(\tau_2) \prod_{i=1}^{n} p(y_i \mid u_i, v_i), \end{aligned} \quad (4)$$

Also estimates of u and v can be obtained from the Gibbs sampler.

2.2 Introducing anisotropy

Given values $U_j = u_j$, $j \neq i$, the distribution of U_i is normal with mean $\frac{\gamma}{m_i} \sum_{j \in S_i} u_j$, as is stated by (3). In this sum, all neighbors u_j receive weights $\frac{\gamma}{m_i}$. In this section, we vary these weights. The motivation to do so is that the spatial dependence in one direction may be stronger than that in the other direction. It would then be reasonable to assign larger weights to the direction with the strongest dependence. In this case we replace γ by γ_1 and γ_2, and write $i \overset{x}{\sim} j$ if i and j are adjacent in the x-direction and $i \overset{y}{\sim} j$ if they are adjacent in the y-direction. Our aim is to define A such that for every quadrat i not located on an edge or corner, the y-neighbors of i receive weights $\frac{\gamma_1}{4}$ and the x-neighbors of i receive weights $\frac{\gamma_2}{4}$. We could define A as the sum of A_1 for the y-direction and A_2 for the x-direction, with A_1 and A_2 defined as in Sect. 2.1. For A_1, for example, $a_{ij} = a_{ji} = -\tau_1 \gamma_1$ if $i \overset{y}{\sim} j$. Up to a factor τ_1, the diagonal contains the number of neighbors in the y–direction. With a similar definition for A_2 however, a complication occurs on the edges. In fact, a quadrat with two y-neighbors and only one x-neighbor has a sum of weights that could be larger than one. Take, for example, $\gamma_1 = 1.7$ and $\gamma_2 = 0.2$. Then the sum equals $\frac{1.7+1.7+0.2}{3} = 1.2$, suggesting that the distribution might not exist. To obtain a positive definite matrix, we modify the γ_1 and γ_2 by a varying γ_{ij}. For all inner quadrats i and j, we define $\gamma_{ij} = \gamma_1$ if $i \overset{y}{\sim} j$ and $\gamma_{ij} = \gamma_2$ if $i \overset{x}{\sim} j$. It is assumed that $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$ are constants such that $\gamma_1 + \gamma_2 = 2 - 2\delta$ for a small constant $\delta > 0$. In our application we set $\delta = 0.005$. First, consider the case $\gamma_1 \geq \gamma_2$. For quadrats i and j both located on either the minimal or maximal x-edge, we define $\gamma_{ij} = {\gamma_1}' = \frac{3-3\delta-\gamma_2}{2}$. For quadrats i and j both located on either the minimal or

maximal y-edge, we define $\gamma_{ij} = \gamma_2{}' = \frac{3-3\delta-\gamma_1}{2}$. The latter modification is not necessary to obtain positive definitess as discussed in Sect. 4. The case $\gamma_2 > \gamma_1$ is handled similarly: $\gamma_{ij} = \gamma_2{}'$ is defined if i and j are on the minimal or maximal y-edge, and $\gamma_{ij} = \gamma_1{}'$ is defined if i and j are on the minimal or maximal x-edge. These modifications are motivated by Geršgorin's disc theorem (Sain and Cressie 2003).

Theorem (Geršgorin) *Let T be an $n \times n$ matrix with complex elements t_{ij}. For $1 \le i \le n$ define $R_i(T) = \sum_{1\le j\le n, j\neq i} |t_{ij}|$. Then the eigenvalues of T are located in the union of discs given by $\bigcup_{i=1}^{n}\{|z - t_{ii}| \le R_i(T)\}$.*

A proof can be found in many linear algebra or matrix analysis textbooks, see for example Lancaster and Tismenetsky (1985, pp. 371–372).
For our matrix A, the condition

$$m_i > \sum_{j\in S_i} \gamma_{ij} \qquad \forall i$$

guarantees that all eigenvalues are positive, implying that A and A^{-1} are positive-definite and non-singular. For all quadrats i and j, the condition $m_i + m_j \le 6$ is equivalent to i and j being *both* located on an edge. Using this fact we summarize the above definitions. For $\gamma_1 \ge \gamma_2$,

$$\begin{aligned}\gamma_{ij} = \frac{-a_{ij}}{\tau_1} = \frac{-a_{ji}}{\tau_1} \;&= \gamma_1{}' 1_{\{(i \overset{y}{\sim} j)\wedge(m_i+m_j\le 6)\}} + \gamma_1 1_{\{(i \overset{y}{\sim} j)\wedge(m_i+m_j\ge 7)\}} \\ &\quad + \gamma_2{}' 1_{\{(i \overset{x}{\sim} j)\wedge(m_i+m_j\le 6)\}} + \gamma_2 1_{\{(i \overset{x}{\sim} j)\wedge(m_i+m_j\ge 7)\}} \end{aligned} \tag{5}$$

and for $\gamma_2 > \gamma_1$,

$$\begin{aligned}\gamma_{ij} = \frac{-a_{ij}}{\tau_1} = \frac{-a_{ji}}{\tau_1} \;&= \gamma_2{}' 1_{\{(i \overset{x}{\sim} j)\wedge(m_i+m_j\le 6)\}} + \gamma_2 1_{\{(i \overset{x}{\sim} j)\wedge(m_i+m_j\ge 7)\}} \\ &\quad + \gamma_1{}' 1_{\{(i \overset{y}{\sim} j)\wedge(m_i+m_j\le 6)\}} + \gamma_1 1_{\{(i \overset{y}{\sim} j)\wedge(m_i+m_j\ge 7)\}} \end{aligned} \tag{6}$$

For $\gamma_2 = \gamma_1$, the isotropic case, (5) and (6) both imply that $\gamma_{ij} = 1 - \delta$ for all $i \sim j$. In the anisotropic case, we have $\frac{1}{m_i}\sum_{j\in S_i}\gamma_{ij} = 1 - \delta$ for any quadrat i, except for the four corners. For the covariance structure described above, the distributions given by (2) and (3) can be refined to

$$p(u) = \frac{1}{(2\pi)^{n/2}\sqrt{|A^{-1}|}} \exp\left\{-\frac{1}{2}\left(\tau_1 \sum_{i=1}^{n} m_i {u_i}^2 - \tau_1 \sum_{i=1}^{n}\sum_{j\in S_i} \gamma_{ij} u_i u_j\right)\right\} \tag{7}$$

$$U_i \mid \{U_j = u_j, j \neq i\} \sim N\left(\frac{1}{m_i}\sum_{j\in S_i}\gamma_{ij}u_j, \frac{1}{m_i\tau_1}\right). \tag{8}$$

2.3 The bivariate model

In this section a bivariate model is described. We are given counts Y_i^A and Y_i^B, and given $U^A = u^A$, $U^B = u^B$, $V^A = v^A$ and $V^B = v^B$, these are assumed to be independent Possion random variables with intensities

$$\lambda_i^A = \beta_A \cdot e^{u_i^A + v_i^A}, \quad \lambda_i^B = \beta_B \cdot e^{u_i^B + v_i^B}. \tag{9}$$

We refer to model (9) as the bivariate model, in contrast to the univariate model defined by (1). The processes $V^A = \{V_i{}^A\}_{i=1,2,\ldots,n}$ and $V^B = \{V_i{}^B\}_{i=1,2,\ldots,n}$ are independent and normally distributed with precisions $\tau_2{}^A$ and $\tau_2{}^B$. Also U^A and U^B have normal distributions, with parameters specified later in this section. This model is motivated by the idea that the conditional specification in expression (8) should be replaced by a bivariate normal distribution. It has variance $\frac{1}{\tau_1 m_i}$, and we can write the exponent in the right hand side of (7) as

$$-\frac{1}{2}\left(\tau_1 \sum_{i=1}^{n} m_i {u_i}^2 - \tau_1 \sum_{i=1}^{n} \sum_{j \in S_i} \gamma_{ij} u_i u_j\right) = -\frac{1}{2}\left(\sum_{i=1}^{n} m_i u_i \left(\frac{1}{\tau_1}\right)^{-1} u_i - \sum_{i=1}^{n} \sum_{j \in S_i} \gamma_{ij} u_i \left(\frac{1}{\tau_1}\right)^{-1} u_j\right).$$

Let the correlation between A and B be controlled by the parameter $c \in (-1, 1)$, and define

$$\Gamma = \begin{pmatrix} \frac{1}{\tau_1{}^A} & \frac{c}{\sqrt{\tau_1{}^A \tau_1{}^B}} \\ \frac{c}{\sqrt{\tau_1{}^A \tau_1{}^B}} & \frac{1}{\tau_1{}^B} \end{pmatrix}.$$

The column vector $(U_i{}^A, U_i{}^B)^t$ is denoted $U_i{}^{A,B}$. Similarly we write $u_i{}^{A,B}$ and $y_i{}^{A,B}$; $U^{A,B}$ denotes the column vector $(U_1{}^A, U_1{}^B, \ldots, U_n{}^A, U_n{}^B)^t$. For the joint density of U^A and U^B the extension of (2) is defined as

$$\begin{aligned} p(u^{A,B}) &\propto \exp\left\{-\frac{1}{2}((u^{A,B})^t \Sigma (u^{A,B}))\right\} \\ &= \exp\left\{-\frac{1}{2}\left(\sum_{i=1}^{n} m_i (u_i{}^{A,B})^t \Gamma^{-1} (u_i{}^{A,B}) - \sum_{i=1}^{n} \sum_{j \in S_i} \gamma_{ij} (u_i{}^{A,B})^t \Gamma^{-1} (u_j{}^{A,B})\right)\right\}, \end{aligned} \tag{10}$$

where Σ is the $2n \times 2n$ block-matrix consisting of 2×2 blocks $\Sigma_{ii} = m_i \Gamma^{-1}$ and $\Sigma_{ij} = -\gamma_{ij} \Gamma^{-1}$ for $i \neq j$. For every $i \sim j$, the terms $\frac{\gamma_{ij}}{2} (u_i{}^{A,B})^t \Gamma^{-1} (u_j{}^{A,B})$ and $\frac{\gamma_{ij}}{2} (u_j{}^{A,B})^t \Gamma^{-1} (u_i{}^{A,B})$ occur in (10). Hence the conditional density can be written as

$$\begin{aligned} p(u_i{}^{A,B} \mid \{u_j{}^{A,B}\}_{j \neq i}) &\propto \exp\left\{\frac{1}{2} (u_i{}^{A,B})^t \left(\frac{1}{m_i}\Gamma\right)^{-1} (\bar{u}_i^{A,B}) \right. \\ &\qquad \left. + \frac{1}{2} (\bar{u}_i^{A,B})^t \left(\frac{1}{m_i}\Gamma\right)^{-1} (u_i{}^{A,B}) - \frac{1}{2} (u_i{}^{A,B})^t \left(\frac{1}{m_i}\Gamma\right)^{-1} u_i{}^{A,B}\right\} \\ &\propto \exp\left\{-\frac{1}{2} (u_i{}^{A,B} - \bar{u}_i^{A,B})^t \left(\frac{1}{m_i}\Gamma\right)^{-1} (u_i{}^{A,B} - \bar{u}_i^{A,B})\right\}, \end{aligned} \tag{11}$$

in which $\bar{u}_i^{A,B} = (\frac{1}{m_i} \sum_{j \in S_i} \gamma_{ij} u_j{}^A, \frac{1}{m_i} \sum_{j \in S_i} \gamma_{ij} u_j{}^B)^t$. This is proportional to the bivariate normal density with mean $\bar{u}_i^{A,B}$ and covariance matrix $\frac{1}{m_i}\Gamma$. The process $U^{A,B}$ is a two-dimensional Gaussian Markov Random Field (MRF). Such processes were studied by

Mardia (1988) and used in various Bayesian models by Pettitt et al. (2002), Gelfand and Vounatsou (2003), Sain and Cressie (2003), and by Jin et al. (2005). In case the spatial correlation between more than two variables is studied, extension of the bivariate model to a multivariate model is straightforward. The bivariate normal density defined by (11) then becomes a multivariate normal density.

2.4 Gibbs sampling

We now turn to Gibbs sampling. In Sect. 2.4.1 we address the univariate case. The extension to the bivariate case is discussed in Sect. 2.4.2.

2.4.1 Univariate sampling

We assume the following prior distributions the parameters β,τ_1,τ_2 and γ_1. The anisotropy parameter γ_1 is discretized , and uniformly distributed on the set

$G = \{0, 0.01, \ldots, 1.99\}$. As we assumed that $\gamma_2 = 1.99 - \gamma_1$, γ_2 also has a uniform distribution on G. Note that in this construction, either $\gamma_1 \geq 1 > \gamma_2$ or $\gamma_2 \geq 1 > \gamma_1$. Alternatively we could choose a discretization with, for example, 201 points, containing the case $\gamma_1 = \gamma_2 = 0.995$. The parameters τ_1 and τ_2 are given identical exponential priors with mean $\frac{1}{\mu}$. In the present application we set $\mu = 1$. Given τ_1, τ_2, and γ_1, U has the density given by (7) and the V_i are uncorrelated normal random variables with mean 0 and variance $\frac{1}{\tau_2}$. For every quadrat $i \in \{1, \ldots, n\}$, the vector $u_{S_i} = (u_{i_1}, \ldots, u_{m_i})$ contains the values of u at adjacent quadrats. To sample from (4) the following conditional distributions need to be sampled.

$$p(u_i \mid u_{S_i}, y, \gamma_1, \tau_1) \propto p(y_i \mid u_i, v_i)p(u_i \mid u_{S_i}, \gamma_1, \tau_1) \tag{12}$$

$$p(v_i \mid y) = p(v_i \mid y_i) \propto p(y_i \mid u_i, v_i)p(v_i \mid \tau_2) \tag{13}$$

$$p(\gamma_1 \mid u, \tau_1) \propto p(u \mid \gamma_1, \tau_1)p(\gamma_1) \tag{14}$$

$$p(\tau_1 \mid u, \gamma_1) \propto p(u \mid \gamma_1, \tau_1)p(\tau_1) \tag{15}$$

$$p(\tau_2 \mid v) \propto \prod_{i=1}^{n} p(v_i \mid \tau_2)p(\tau_2)$$

The conditional density in (12) is proportional to

$$\begin{aligned}(\beta e^{v_i} e^{u_i})^{y_i} \exp\left\{-\beta e^{v_i} e^{u_i} - \frac{m_i \tau_1}{2}(u_i - \bar{u}_i)^2\right\} \\ \propto \exp\left\{y_i u_i - \beta e^{v_i} e^{u_i} + (m_i \tau_1 \bar{u}_i)u_i - \frac{m_i \tau_1}{2} {u_i}^2)\right\}\end{aligned} \tag{16}$$

First u^*, the mode of $p(u_i \mid u_{S_i}, y, \gamma_1, \tau_1)$ is determined numerically. We use a second order Taylor approximation around u^*, $e^{u_i} \approx e^{u^*}(1+(u_i-u^*)+\frac{1}{2}(u_i-u^*)^2)$. Instead of sampling (12) exactly, we take the approximating normal density with mean $(b - ce^{u^*}(1-u^*))/(2a + ce^{u^*})$ and variance $1/(2a + ce^{u^*})$, for $a = \frac{\tau m_i}{2}$, $b = y_i + \tau m_i \bar{u}_i$ and $c = \beta e^{v_i}$. A similar normal approximation is used for (13).

The main difficulty in sampling (14) is the determinant of A. It is more convenient to work with $Q = \frac{1}{\tau_1}A$, not depending on τ_1. For all $\gamma_1 \in G$, $|Q|$ is calculated before running the Gibbs sampler, and stored in a table. For the current values of u and τ_1, $p(u \mid \gamma_1, \tau_1) \propto \sqrt{{\tau_1}^n |Q|} \exp\{-\frac{\tau_1}{2} u^t Q u\}$ is then evaluated for all $\gamma_1 \in G$. After normalization, a new γ_1 is

drawn according to this vector. The density given by (15) is proportional to

$$e^{-u^*\tau_1}\sqrt{{\tau_1}^n|Q|}\exp\left\{-\frac{1}{2}u^t\tau_1 Qu\right\} \propto {\tau_1}^{\frac{n}{2}}\exp\left\{-(u^*+\frac{1}{2}u^tQu)\tau_1\right\},$$

also a gamma density. The same holds for the conditional density of τ_2.

2.4.2 Bivariate sampling

Using the same priors for the parameters, updating the ${v_i}^A$'s, ${v_i}^B$'s and the parameters τ_2^A,τ_2^B and γ_1 can be done exactly as in the univariate sampler. The parameter c is discretized, and sampled from the set $\{-0.99, -0.98, \ldots, 0.99\}$, on which a uniform prior is assumed. Additionally, we need to sample from

$$p({u_i}^A|u_{S_i}^A, u_{S_i}^B, {y_i}^A, \gamma_1, {\tau_1}^A, {\tau_1}^B, c) \propto p({y_i}^A \mid {u_i}^A)p({u_i}^A \mid u_{S_i}^A, u_{S_i}^B, \gamma_1, {\tau_1}^A, {\tau_1}^B, c)$$
$$p({\tau_1}^A \mid u^{AB}, {\tau_1}^B, c, \gamma_1) \propto p(u^{AB} \mid \gamma_1, {\tau_1}^A, {\tau_1}^B, c)p({\tau_1}^A). \quad (17)$$

Similar factorizations hold for $p({u_i}^B \mid u_{S_i}^A, u_{S_i}^B, {y_i}^B, \gamma_1, {\tau_1}^A, {\tau_1}^B, c)$ and $p({\tau_1}^B \mid u^{AB}, {\tau_1}^A, c, \gamma_1)$. Since $p({u_i}^A \mid {u^A}_{S_i}, u_{S_i}^B, \gamma_1, {\tau_1}^A, {\tau_1}^B, c)$ is normally distributed, (17) is of the same form as (16). The density $p({\tau_1}^A \mid {\tau_1}^B, c, u^{AB}, \gamma_1)$ is of the form $(\tau_1^A)^{\frac{n}{2}}e^{-c_1\tau_1^A - c_2\sqrt{\tau_1^A}}$, with constants c_1, c_2 depending on u^{AB},c,τ_1^B and μ, and can be sampled using a rejection scheme.

2.5 Software

All calculations were performed in R. The code is available on request to the first author.

3 Application

3.1 Data

Weeds were observed within an arable field of 64 m wide and 281 m long, on a clay soil located near Wageningen, The Netherlands, in 2001 (Heijting et al. in preparation). The field was planted with maize. Weed plants were counted within a rectangular plot of 50.25 m long (y-direction) and 12 m wide (x-direction). This plot was partitioned into a contiguous grid of $0.75\times$ 0.75 m quadrats, corresponding in size to the 0.75 m spacing between the rows of maize plants. The plot contained $n = 1,072$ quadrats with 67 and 16 quadrats in the y- and x-direction, respectively. Quadrats are the units for sampling. Because of the strong trend in the data, we restricted our analysis to the plot of 16×16 quadrats, located along the Northern boundary of the observation area.

Weed plants were counted from 18 to 21 June. More than 20 different weed species were found in the plot. For practical reasons, such as computation time, we restrict ourselves in this study to three of these: *Chenopodium album* L. (Fat hen), *Solanum nigrum* L. (Black nightshade) and *Taraxacum officinale* Weber (Dandelion). Figure 1 contains a graphical representation of the data. The orientation is such that the left sides of the images are facing the West. Descriptive statistics of the counts can be found in Table 1.

The distribution of the weed counts is highly skewed. Having a maximum equal to 44 and a minimum equal to 0, the median of the *C. album* counts is only 2, whereas for the two other species the median equals 0. Large peak densities occur for *C. album* and *S. nigrum*. In an

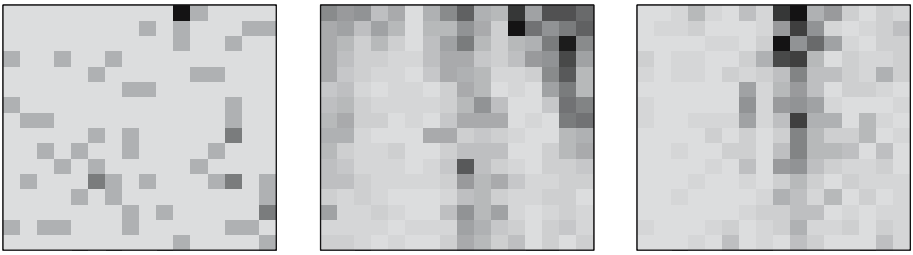

Fig. 1 Numbers of weed plants of *T. officinale* (*left*), *C. album* (*middle*) and *S. nigrum* (*right*), in 2001. The *dark grey* tones represent high counts

Table 1 Quantiles, minimum, maximum and median and mean of the numbers of weed plants observed on 16×16 quadrats in 2001

Species	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
T. officinale	0	0	0	0.2305	0	4
C. album	0	2	5	7.887	11	44
S. nigrum	0	0	1	2.004	2	19

experimental plot in Canada, Goudy et al. (2001) observed, among others, *C. album* and *T. officinale*. The former had a density of 0.09 plants m^{-2} in 1998 and 1.30 plants m^{-2} in 1999. For *T. officinale*, a mean density of 0.62 plants m^{-2} was observed in 1998. Note that in the present study, weeds are counted within areas sized $0.75m \times 0.75m = 0.5625m^2$.

3.2 Results

3.2.1 Estimation of the processes

The 2001 counts of *T. officinale* were analyzed using the univariate model. In addition, the bivariate model was applied to the *C. album* and *S. nigrum* counts. In both cases the Gibbs sampler was run for 70,000 iterations, after a burn-in period of 1,000 iterations. Every 70th iteration was taken as a sample from the posterior distribution, hence 1,000 samples were stored. The posterior density estimates are given in Fig. 2. The posterior means of the parameters can be found in Table 2. We notice that the distributions are relatively wide, although mostly showing a sharp mode, which may be caused by the fact that the underlying process does not fully obey the conditions imposed by the statistical model. The estimates of τ_1 and τ_2 measure the relative magnitudes of spatial and non—spatial variation, respectively. For all species the spatial variation is considerably larger than the non—spatial variation. The anisotropy apparent in the counts of *C. album* and *S. nigrum* is reflected in the estimate of γ_1. The posterior mean of c is 0.5747, substantially larger than the correlation between the actual counts, which is 0.2088.

Figure 3 displays images of the posterior means of U, V and $\beta \exp(U + U)$, which are denoted as $\tilde{u}$, $\tilde{v}$ and $\tilde{\lambda}$, respectively. The image of $\tilde{v}$ is merely included as a check on randomness. Indeed sudden variations appear in $\tilde{v}$, although for *S. nigrum* it still possesses some spatial structure. Generally $\tilde{u}$ and $\beta \exp(\tilde{u} + \tilde{v})$ have the same spatial structure as is apparent in the data, but in a more smoothed fashion.

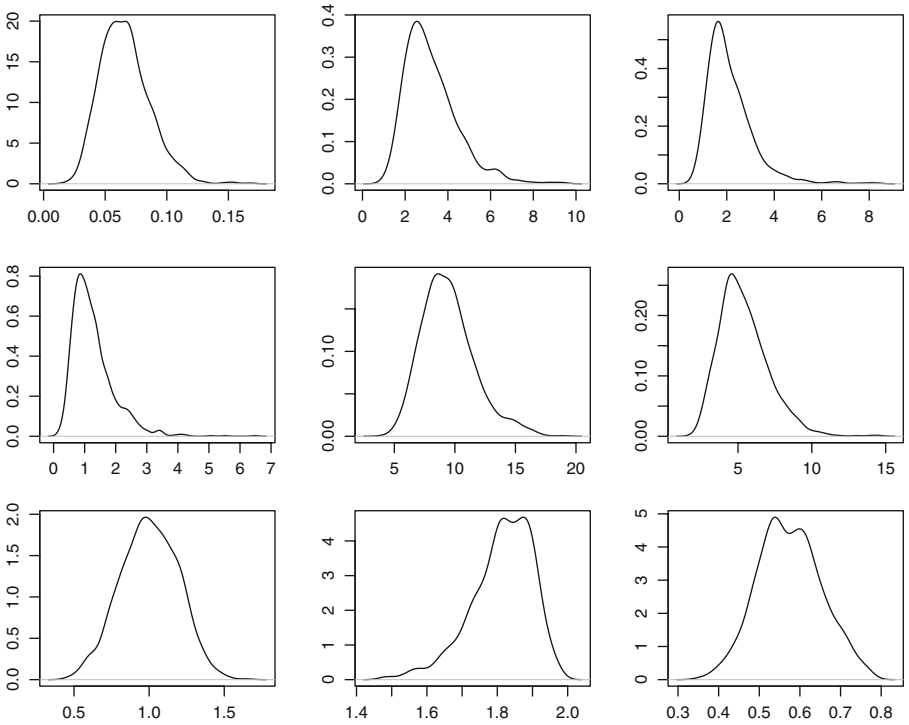

Fig. 2 Posterior densities of the parameters τ_1 (*top line*), and τ_2 (*middle line*) for *T. officinale* (*left*), *C. album* (middle) and *S. nigrum* (*right*). On the bottom line: the posterior densities of γ_1 for *T. officinale* (*left*), γ_1 for *C. album* and *S. nigrum* (*middle*) and the posterior of the correlation c between *C. album* and *S. nigrum* (*right*)

Table 2 The posterior means of γ_1, τ_1, τ_2 and c

Species	τ_1	τ_2	γ_1	c
T. officinale	0.0661	1.2641	1.0017	
C. album	3.2201	9.4576	1.8130	0.5747
S. nigrum	2.1353	5.3999	1.8130	0.5747

In the bivariate model, *C. album* and *S. nigrum* have the same γ_1 and c

4 Discussion

The models described in this paper can be useful in the analysis of count data on a lattice. For further application, the approach taken in this paper could be extended by incorporating covariables.

For quadrats on the edges the γ_{ij}'s can be defined in various ways. For instance if $\gamma_1 \geq \gamma_2$, and i and j are located on the minimal or maximal y-edge, $\gamma_{ij} = \gamma_2$ could be defined instead of $\gamma_{ij} = \gamma_2' = \frac{3-3\delta-\gamma_1}{2} > \gamma_2$. Numerical experiments with isotropic test data indicate, however, that this model only performs well on square lattices, but is biased for rectangular lattices. More precisely, this model favors a stronger spatial dependence in the longest

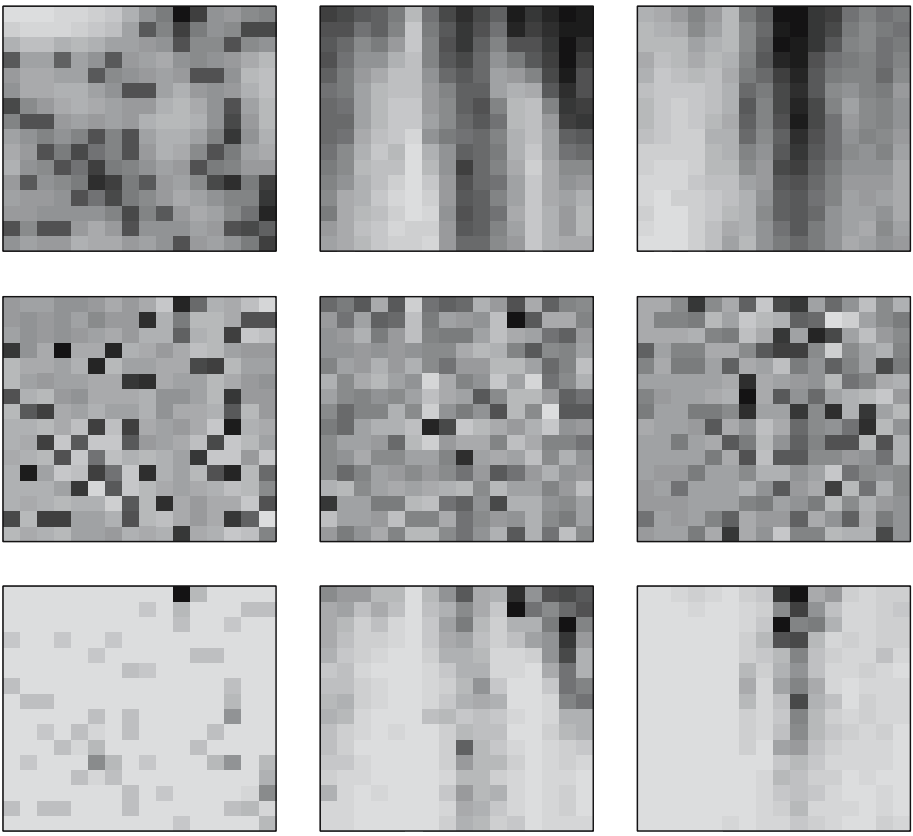

Fig. 3 The posterior mean of U (*top line*), the posterior mean of V (*second line*) and the estimated Poisson intensities $\beta \exp(u + v)$ (*bottom line*), for the 2001 counts of *T. officinale* (*left*), *C. album* (*middle*) and *S. nigrum* (*right*). The *dark grey* tones represent large values

direction of the field. With the current definition, this problem does not occur. Although our dataset is not rectangular, this improves the applicability of the model.

The models presented in Sect. 2 can be used to quantify the spatial- and non-spatial variation in the data. For the application of SSWM, images of $\tilde{u}$ are easier to interpret than images of the original data, which exhibit much small scale variation. This may lead to a more efficient use of resources.

Our method is not restricted to grids of contiguous quadrats, and also applicable to data obtained with discrete area sampling, or to observations at grid points. The grid does not need to be rectangular. In order to define first- and higher-order neighbors, however, distances between the observation areas, or points, need to be regular or at least, it must be possible to distribute them over distance classes.

Christensen and Waagepetersen (2002) takes a geostatistical approach, considering weed counts as point observations. Instead of an autoregression process U, defined on a lattice, he assumes a geostatistical process, defined for every point in the area. Prior distributions for the regression parameters are specified, and posterior distributions are estimated by MCMC simulation. An advantage of this approach is that predicting at unsampled locations is

straightforward. Hrafnkelsson and Cressie (2003) compare approaches based on lattice and geostatistical processes and describe how $MCMC$-simulations can be made.

Another open question is the possibility of having different degrees of anisotropy for the two species. Parameters ${\gamma_{ij}}^A$'s and ${\gamma_{ij}}^B$'s have to introduced such that Σ remains positive-definite.

Acknowledgements We want to thank M.N.M. van Lieshout (CWI, Amsterdam) and A.W. van der Vaart (Vrije Universiteit Amsterdam) for giving useful comments and suggestions.

References

Agarwal DK, Gelfand AE, Citron-Pousty S (2002) Zero-inflated models with application to spatial count data. Environ Ecol Stat 9(4):341–355

Besag JE, York J, Mollié A (1991) Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math 43(1):1–59

Christensen OF, Waagepetersen R (2002) Bayesian prediction of spatial count data using generalized linear mixed models. Biometrics 58:280–286

Cliff AD, Ord JK (1980) Spatial processes models & applications. Pion, London

Diggle PJ (2003) Statistical analysis of spatial point patterns, 2nd edn. Academic Press, New York

Gelfand AE, Vounatsou P (2003) Proper multivariate conditional autoregressive models for spatial data analysis. Biostatistics 4:11–15

Goudy HJ, Bennett KE, Brown RB, Tardif FJ (2001) Evaluation of site-specific weed management using a direct-injection sprayer. Weed Sci 49:359–366

Hess D, van Lieshout MNM, Payne B, Stein A (2001) A review of spatio-temporal modelling of quadrat count data with application to striga occurrence in a pearl millet field. Int J Appl Earth observ Geoinform 3(2):133–138

Heijting S, Werf W, van der, Stein A, Kropff MJ. Are weed patches stable in location? Application of an explicitly two-dimensional methodology. Accepted by Weed Research (in press)

Hrafnkelsson B, Cressie NA (2003) Hierarchical modeling of count data with application to nuclear fall-out. Environ Ecol Stat 10(2):179–200

Jin X, Carlin BP, Banerjee S. (2005) Generalized hierarchical multivariate CAR models for areal data. Biometrics 61:950–961

Kropff MJ, Wallinga J, Lotz LAP (1997) Precision agriculture: spatial and temporal variability of environmental quality. Wiley, Chichester (Ciba Foundation Symposium 210) pp 182–204

Lancaster P, Tismenetsky M (1985) The theory of matrices, 2nd edn. Academic Press

Mardia KV (1988) Multi-dimensional multivariate Gaussian Markov random fields. J Mult Anal 24:265–284

Pettitt AN, Weir IS, Hart AG (2002) A conditional autoregressive Gaussian process for irregularly spaced multivariate data with application to modelling large sets of binary data. Stat Comput 12(4):353–367

Rew LJ, Cousens RD (2001) Spatial distribution of weeds in arable crops: are current sampling and analytical methods appropriate?. Weed Res 41:1–18

Sain SR, Cressie NA (2003) A spatial model for multivariate lattice data, preprint. Available online: http://math.cudenver.edu/~ssain/mrf.pdf

Stein A (2001) Editorial: spatial statistics for production ecology and resource conservation. Environ Ecol Stat 8:293–295

Biographical sketches

Willem Kruijer received the MSc degree in statistics from Groningen University, The Netherlands, after which he spent 6 months at the ITC International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands. He is currently a PhD-student at the Vrije Universiteit Amsterdam, where his research focusses on nonparametric mixture models, Bayesian asymptotics and information theory.

Alfred Stein received the MSc degree in mathematics and information science, with a specialization in applied statistics from Eindhoven University of Technology, Eindhoven, The Netherlands, and a PhD in agricultural and environmental sciences at Wageningen University, The Netherlands.

He is currently a Professor of Mathematical and Statistical methods for geodata at the ITC International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands, where he heads the department of Earth Observation Science. He is also chief editor of the *International Journal of Applied Earth Observation and Geoinformation*. His main interest is in spatial statistics and data mining, with a focus on spatial data quality. Applications emerge from remote sensing and a range of agricultural and environmental fields.

Willem Schaafsma is Professor Emeritus at Groningen University, where he received his PhD in 1966. His main areas of research are discriminant analysis, testing hypotheses with restricted alternatives, distributional inference and the foundations of statistics. Besides his work in mathematical statistics, he is interested in a wide range of applications, with ongoing cooperations with researchers from various fields such as ornithology, pathology and physical anthropology.

Sanne Heijting is currently finishing her PhD thesis on "Spatial analysis of weed patterns" at Crop and Weed Ecology group at Wageningen University. The thesis will include papers on techniques to characterize weed spatial patterns and their temporal behaviour on arable fields. It furthermore focusses on relations between weed patterns and heterogeneity of underlying soil parameters and dispersal of plant material by machinery.