

Multiresolution Assessment of Forest Inhomogeneity

Katja Ickstadt
Robert L. Wolpert

ABSTRACT The spatial distribution of dominant tree species in an undisturbed mature stand tends to be regular and even, often exhibiting less variation than a simple Poisson model would suggest; in contrast the spatial distribution of species in a recovering or transitional stand would be expected to display considerable spatial variation. This paper studies the spatial distribution of hickory trees within the Bormann research plot of Duke Forest in an attempt to assess the degree of variation, as an indicator for forest maturation, using models recently introduced in [WI95]. A data augmentation scheme and Markov chain Monte Carlo methods are employed to evaluate Bayesian posterior distributions.

Key words: Bayesian hierarchical models, data augmentation, Markov chain Monte Carlo, spatial extensibility.

1 Introduction

In 1951 Frank Bormann selected a 140m square portion of the Durham subdivision of Duke Forest in the Piedmont region of North Carolina as a research plot for examining the effect of plot size and shape on empirical ecological studies, describing it as a “relatively undisturbed immature climax forest” (by “undisturbed” he means unaffected by recent cutting, burning, or grazing); he and members of the Duke University Department of Botany took a census of the species, position, and diameter of every tree within the plot in 1951–52 [Bor53]. In 1974 Norm Christensen of the Duke University Department of Botany remapped the area in a study of forest maturation, seeking to characterize the changes that occur during this process [Chr77]; his analysis suggests that the plot had been less “undisturbed” than Bormann believed.

The present paper is an attempt to shed light on this issue, using data at a single time point (1974), by assessing the regularity of the distribution of a subdominant species (hickory) in this predominantly oak stand. Irregularities would reveal that the subdominant species is receding from or encroaching into a stand recovering from a disturbance, while regularity would suggest a more stable and undisturbed recent history. It is intended as an introductory step toward a multi-species spatio-temporal analysis

of forest maturation intended to predict features of the plot important to forest management, such as susceptibility to damage from fire or insect infestation.

The most common methods in use for studying possibly-correlated spatial count data are Gaussian random field models in which the count data, following a variance-stabilizing transformation, are treated as having normal distributions with a partially-unknown covariance structure (this is closely related to the spatial interpolation method of Kriging introduced in [Mat63]), and log-linear Poisson/log-normal models, in which the data are treated as Poisson with intensities whose logarithms are modeled with a Gaussian random field. Unfortunately Gaussian models [Cre93, CC89] fail to respect the discrete nature of the count data, while log-linear models [BPBG96, BYM91, CK87, RMG⁺95] do not scale properly under aggregation and refinement of partitions: the logarithmic structure leads to products rather than sums for the Poisson means for unions of neighboring partition elements, forcing a full reanalysis of the data and reelicitation of any needed prior distributions for each level of refinement of interest.

Other methods often applied to spatial point count data include auto-Poisson models [Bes74] and Strauss models [Str75], which cannot reflect the positive associations that might be expected for nearby reporting units (all count correlations must be negative), and conjugate Poisson/gamma hierarchical models [CK87], under which all counts are independent under both prior and posterior. More elaborate Markov point processes can, of course, be constructed to exhibit positive local correlation but will not be considered further here. Recently we introduced a flexible class of hierarchical random field models for possibly spatially-dependent count data (generalizing the conjugate Poisson/gamma models), intended for applications in which the precise location of every point is observed and available for analysis [WI95].

In the present paper we study hickory tree count data for discrete reporting units of prescribed spatial location and extent. We partition the Bormann plot into a rectangular quadrat grid and study the degree of dependence among hickory tree counts for nearby quadrats. The data may be studied at many different levels of aggregation (i.e., the rectangular grid of quadrats may be of any size). The counts are modeled in a three-stage hierarchical model based on the random field models of [WI95], with a much simpler computational scheme designed for discretized data sets (drawing a few gamma variates each iteration instead of the thousands of points needed to simulate a gamma random field) and a novel model-independent prior elicitation scheme, leading to a coherent analysis at a range of levels of aggregation.

The data set is introduced in Section 2. In Section 3 the models are described and prior distributions are elicited; in Section 4 posterior distributions, expectations, and correlations for the unobserved Poisson means are presented. Conclusions and directions for further research are discussed

in Section 5, followed by an Appendix describing the computational method (a hybrid Gibbs/Metropolis scheme, featuring data augmentation).

2 Hickory Trees

The underlying data set consists of a complete census count of every tree in the $140m \times 140m$ Bormann research plot in Duke Forest, Durham, North Carolina, with a diameter of at least $2.5cm$ at “breast height” ($1.5m$). It includes the diameter (within $\pm 0.1cm$), spatial coordinates (within $\pm 0.1m$), and an indicator of the tree’s condition, at the initial Bormann 1951–52 census [Bor53] and at subsequent updates performed in 1974, 1982, 1989, and 1991 variously by Norm Christensen of Duke University and Robert Peet of the University of North Carolina at Chapel Hill in a continuing study of forest maturation [Chr77]. The complete data set includes records for 7,859 trees of more than twenty species, predominantly oak and hickory. As suggested by [Chr77] we treat the four hickory species present (pignut, sweet pignut, shagbark, and mockernut hickory) as a single group, and omit trees located close to the boundary to minimize edge effects; the present analysis concerns only the locations of the 57 hickory trees located at least $10m$ from the boundary with diameter at least $2.5cm$ in the 1974 census (see Table 1.1), and makes no use of the trees’ locations within quadrats nor of their conditions or diameters.

Figure 1 shows the data (dots) and empirical density plots at six levels of discretization, with darker shading indicating higher density. Evidently different spatial patterns emerge at different resolutions, with differing apparent levels of inhomogeneity; fixing a partition in advance might lead to an inference specific to the particular (and arbitrary) partition chosen. The spatially extensible hierarchical Poisson/gamma models described here permit the data set to be analyzed at a range of differing levels of aggregation using the same models and a common prior elicitation.

3 Methodology

The simplest generalization of conjugate Poisson/gamma models that can reflect possible positive association for neighboring regions is to allow the Poisson means to be, not independent gamma variates, but linear combinations of a collection of independent gamma variates, with overlapping linear combinations leading to dependence. We use the following three-stage hierarchical generalized Poisson/gamma model.

At the lowest stage of the hierarchy the counts $\{N_i\}_{i \in I}$ of hickory trees in elements $\{\mathcal{R}_i\}_{i \in I}$ of a rectangular quadrat partition $\mathcal{R} = \cup_{i \in I} \mathcal{R}_i$ of the research plot (cf. Figure 1) are represented by independent Poisson ran-

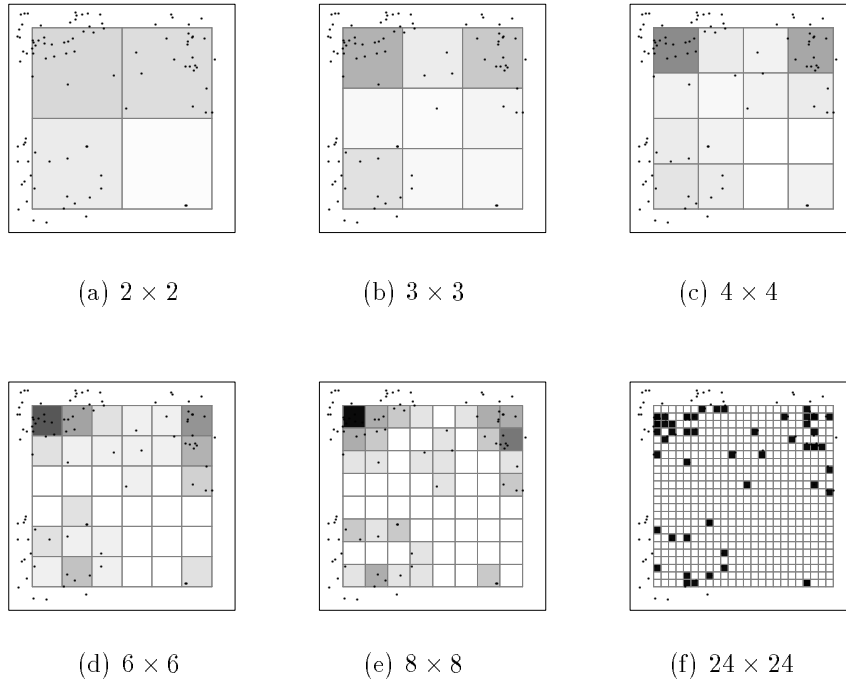


FIGURE 1. Empirical hickory tree density plots for various discretizations.

x	y	x	y	x	y	x	y
31.4	10.3	33.5	18.4	38.7	14.1	48.7	17.3
112.4	12.3	113.4	12.4	11.6	23.6	56.3	23.4
56.2	32.2	12.2	47.7	21.5	41.4	33.6	43.2
46.6	51.1	46.7	51.1	72.9	76.4	117.4	77.5
126.2	73.6	126.7	89.8	10.3	97.6	34.0	92.3
64.7	98.6	83.2	99.9	19.2	109.8	104.9	109.2
112.4	104.4	115.5	104.1	118.0	101.6	119.2	104.8
120.0	103.5	10.2	115.8	10.2	115.9	10.8	112.6
11.1	120.2	12.5	118.0	18.5	119.6	23.3	118.7
24.8	110.6	32.8	112.5	38.3	114.2	79.4	113.5
113.9	119.1	117.2	110.2	10.6	121.5	14.4	121.3
18.2	123.7	31.6	120.2	34.3	120.7	37.4	122.7
44.3	127.5	51.2	126.8	58.0	129.6	99.9	122.6
114.0	120.8	114.6	126.0	117.0	123.8	117.6	124.0
125.4	122.9						

TABLE 1.1. Hickory tree locations, relative to SW corner of 140m square plot.

dom variables. At the second stage of the hierarchy the Poisson means are taken to be products of the quadrats' areas A_i and unit-area intensities, linear combinations $\Lambda_i \equiv \sum_{j \in J} k_{ij} \Gamma_j$ of unobserved independent gamma-distributed impulses $\{\Gamma_j\}_{j \in J}$ with uncertain nonnegative coefficients $\{k_{ij}\}_{i,j \in I,J}$. In the present examples $I = J$ and the areas $\{A_i\}_{i \in I}$ are all equal to the total area divided by the number of quadrats, but other applications (see Discussion) will require the greater flexibility of an arbitrary index set J and unequal areas $\{A_i\}_{i \in I}$.

Uncertainty about the parameters for the gamma distributions of the impulses $\Gamma_j \sim \text{Ga}(\alpha_j, (\beta_j)^{-1})$ and about the kernel k_{ij} is modeled at the third stage through dependence of α_j^θ , β_j^θ , and k_{ij}^θ on an uncertain parameter $\theta \in \Theta$, regarded (in the Bayesian paradigm) as random and accorded a prior density $\pi(\theta)$ with respect to Lebesgue measure $d\theta$ on $\Theta \subset \mathbb{R}^p$ ($p = 3$ in our application); dependence on θ is indicated by superscripts. Thus, the model class can be summarized:

$$\begin{array}{llll}
 \text{Parameter:} & \theta & \sim & \pi(\theta) \\
 \text{Impulses:} & \{\Gamma_j\}_{j \in J} | \theta & \stackrel{\text{ind}}{\sim} & \text{Ga}(\alpha_j^\theta, (\beta_j^\theta)^{-1}) \\
 \text{Intensities:} & \Lambda_i & \equiv & \sum_{j \in J} k_{ij}^\theta \Gamma_j \\
 \text{Point counts:} & \{N_i\}_{i \in I} | \theta, \{\Gamma_j\} & \stackrel{\text{ind}}{\sim} & \text{Poi}(A_i \Lambda_i).
 \end{array} \tag{1.1}$$

Conjugate Poisson/gamma hierarchical models are the special case of (1.1) in which $I = J$ and k_{ij} is diagonal (typically the identity matrix δ_j^i); nondiagonal kernels will lead to spatial correlation under prior and posterior distributions. The mean vectors and covariance matrices of the point counts, conditional on $\theta \in \Theta$, follow by routine computation:

$$\begin{aligned}
 \text{E}^\theta[N_i] &= \sum_{j \in J} A_i k_{ij}^\theta (\beta_j^\theta)^{-1} \alpha_j^\theta \\
 \text{Cov}^\theta[N_i, N_{i'}] &= \sum_{j \in J} (\delta_{i'}^i + A_{i'} k_{i'j}^\theta (\beta_j^\theta)^{-1}) A_i k_{ij}^\theta (\beta_j^\theta)^{-1} \alpha_j^\theta.
 \end{aligned} \tag{1.2}$$

Unconditionally the $\{N_i\}_{i \in I}$ are distributed as sums of independent negative-binomial random variables.

3.1 Prior Elicitation

We consider two essentially different model kernel functions k^θ (a sparse nearest-neighbor kernel and a dense distance-based kernel), at a variety of levels of spatial aggregation. To highlight the role of spatial extensibility in these models we elicit prior information before introducing and studying specific models and, in particular, do not frame the elicitation exercise in the context of a particular parametric model. Instead we consider three questions about distances (consensus answers given by our elicitees are shown in parentheses):

Q₁: How far would a randomly-selected point $x \in \mathcal{R}$ be from the nearest hickory tree of diameter at least 2.5cm , on average Γ

$$(D_1 = 7.5\text{m} \pm 50\%)$$

Q₂: How far would a randomly-selected hickory tree be from the nearest hickory tree of diameter at least 2.5cm , on average Γ

$$(D_2 = 5.0\text{m} \pm 50\%)$$

Q₃: How far would y have to be from a point $x \in \mathcal{R}$ for knowledge of the tree density at y to be only half as helpful in guessing the density at x as would knowledge of the density at a point only 1m from x Γ

$$(D_3 = 3.0\text{m} \pm 50\%)$$

If hickory tree locations are independent (for example, if trees are well modeled by a Poisson point process with uniform intensity in \mathcal{R}) then $D_1 = D_2$, while a tendency of hickory trees to cluster would lead to $D_2 < D_1$; see [Rip81, Ch. 7,8]. From the three answers above we can infer features of the prior distributions about tree densities, over-dispersion, and interaction distance.

3.2 Models

For each model below let d be an integer, I a set of d^2 indices, and for $i \in I$, N_i the count of hickory trees in the i^{th} partition element \mathcal{R}_i in a $d \times d$ rectangular quadrature $\mathcal{R} = \cup_{i \in I} \mathcal{R}_i$. In each of the models we also take $J = I$ and, for $j \in J$, take $\alpha_j^\theta = e^{\theta_0}$ and $\beta_j^\theta = e^{-\theta_1}$.

In the first model M_S (“S” for *sparse*) we regard \mathcal{R}_i and \mathcal{R}_j as neighbors (and write $i \sim j$) if the quadrats are adjacent; let nn_i be the number of neighbors of the \mathcal{R}_i , excluding itself (four in the interior, three on the edges, two in the corners) and set $k_{ij}^\theta \equiv 1$ for $i = j$, $k_{ij}^\theta \equiv e^{\theta_2}/\text{nn}_i$ for $i \sim j$. Since each $\text{nn}_i \leq 4$, the matrix k^θ is relatively sparse with fewer than $5d^2$ nonzero elements.

In the second model M_D (“D” for *dense*) set $k_{ij}^\theta \equiv \exp(-(|x_i - x_j|/e^{\theta_2})^2)$ for all $i, j \in I$, where $x_i \in \mathbb{R}^2$ denotes the center of quadrat \mathcal{R}_i (relative to the center of \mathcal{R}); now all d^4 elements of k^θ are nonzero.

Under both kernels the data become approximately independent as the quantity e^{θ_2} becomes close to zero; thus large values of θ_2 suggest spatial correlation and irregular distribution of hickory trees. The nearest-neighbor kernel structure in model M_S suggests links with Markovian (Gibbs) models, but we do not pursue that connection here.

3.3 Prior Distributions

Now we derive probability distributions for θ in each model we consider (M_S , M_D) at each level of aggregation ($d = 2, 4, 8$) consonant with the

elicited features.

The mean density (in hickory trees per square meter; see (1.2)) for each model is $E[\Lambda_i] = \exp(\theta_0 + \theta_1) \sum_{j \in J} k_{ij}^\theta$, while from the prior elicitation it should be about $1/4D_1^2$; thus $e^{\theta_0 + \theta_1} (1 + e^{\theta_2}) \approx 1/225$ for M_S .

A second nonlinear function of θ is given by the nearest-neighbor ratio D_1/D_2 , related to the K -function of Bartlett and Ripley ([Bar64]); a variety of related measures of over-dispersion are described in [Rip81, Ch. 6]. For the present analysis we use the Index of Clumping [DM54], the sample variance/sample mean ratio minus one, estimated for hickory tree counts in the Bormann plot in random quadrats $5m$ on a side by [Chr77] to be about $ICS = 0.12$. In model M_S this ratio for quadrats with sides of length s should be approximately $ICS \approx s^2 e^{\theta_1} (1 + e^{2\theta_2}/4)/(1 + e^{\theta_2})$ (note this depends on quadrat size).

If correlation falls off approximately exponentially with rate β then averages over neighboring quadrats would have correlation $\rho = (1 - e^{-x})^2 / 2(e^{-x} - 1 + x)$, for $x = 120\beta/d$, while in model M_S the correlation would be approximately $\rho = 2e^{\theta_2} / (4 + e^{2\theta_2})$; from this relation we can infer an implicit value of θ_2 from the elicited value of $D_3 = \log 2/\beta$, a third equation in the three components of θ .

From this third relation for model M_S with $d = 4$ we see $\theta_2 \approx -1.775$; from the second we have $0.12 = s^2 e^{\theta_1} (1 + e^{2\theta_2}/4)/(1 + e^{\theta_2})$, or $\theta_1 \approx -5.190$; from the first we have $\theta_0 + \theta_1 \approx -5.573$, or $\theta_0 \approx -0.383$. Thus for a 4×4 quadrature we take gamma variates to have shape parameters $\alpha_j^\theta = \exp(\theta_0) \approx 0.682$ and inverse scale parameters $\beta_j^\theta = \exp(-\theta_1) \approx 179.4$, and take a kernel with off-diagonal elements $k_{ij}^\theta = \exp(\theta_2)/4 \approx .042$ for modal values. We used independent normal prior distributions for the components of θ , centered for the 4×4 resolution at the value $\hat{\theta} = (-0.383, -5.190, -1.775)$ suggested above with standard deviations of $(1.0, 1.0, 1.0)$ to allow misspecification of approximately $\pm 50\%$.

Models M_S and M_D differ only in their kernel matrices k_{ij}^θ , which depend on θ only through θ_2 ; accordingly for M_D we use the same prior distributions as in M_S for θ_0, θ_1 but center the prior for θ_2 at the point where $\exp(-(D_3/e^{\theta_2})^2) \approx 0.5$, i.e., $\theta_2 \approx \log(D_3/\sqrt{\log 2}) \approx 1.28$.

4 Posterior Analysis

The interest in this study lies in the degree of spatial variation or association. One measure of this in both models is the value of $\exp(\theta_2)$, the sum of the nearest-neighbor kernel terms in model M_S and the interaction distance in M_D ; values close to zero in either model suggest that there is little spatial effect and that the quadrat counts are nearly independent, while strongly positive values suggest association among neighboring cells and a lack of independence. A second measure is the variance/mean ratio

	M_S	M_D	ICS
2×2	5.190	5.283	5.193
4×4	4.015	3.949	5.137
8×8	1.034	1.109	1.535

TABLE 1.2. Extra-Poisson Variation: $\text{Var}[A_i\Lambda_i]/\text{E}[A_i\Lambda_i]$, $\text{ICS}=(S_{N_i}^2/\overline{N_i}) - 1$.

$\text{Var}[A_i\Lambda_i]/\text{E}[A_i\Lambda_i]$, approximately equal to the expected clumping index ICS in each model; values near zero suggest near-independence (i.e., nearly homogeneous Poisson behavior), while large values suggest evidence in favor of inhomogeneity.

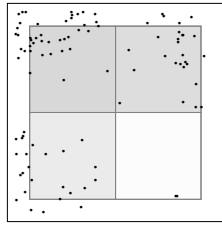
Estimated posterior mean densities $\text{E}[\Lambda_i]$ based on 100,000 Markov chain Monte Carlo (MCMC) replications (after a burn-in period of 1,000 steps) are presented in Figure 2 for both models and compared to the data at each of the three discretizations (darker shading indicates higher tree densities). The computational method used is described in detail in the Appendix. Conventional diagnostics suggest that MCMC convergence was attained after the first 10,000 iterations.

Both models reproduce gross features of the data well, fitting the overall density (57/14,400 trees/ m^2) within $\pm 0.5\%$ at all three levels of discretization and picking up the apparent density peaks near the northwest and northeast corners and the smaller one in the southwest corner; in common with all hierarchical models they act to smooth the data somewhat, exhibiting somewhat less variation and more regular variation than the raw data. Figure 2 seems to suggest that the dense model M_D smoothes somewhat less, fitting the data somewhat better than does M_S .

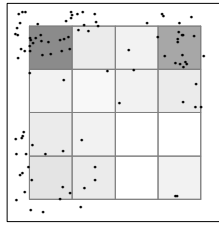
In a doubly-stochastic Poisson model $N_i \sim \text{Poi}(A_i\Lambda_i)$ with $A_i\Lambda_i$ random, with some mean μ and variance σ^2 , the mean and variance of N_i are μ and $\mu + \sigma^2$, respectively, so the Index of Clumping (recall Section 3.3) will be approximately $\text{ICS} \approx \sigma^2/\mu$; Table 1.2 shows the empirical ICS along with the ratio of sample variance to sample mean for the Poisson means $A_i\Lambda_i$ for each of the models and aggregation levels considered.

Evidently both models show slightly less variation, as quantified by ICS, than do the raw data (another suggestion of smoothing), but in each case there remains the clear suggestion that hickory trees are not distributed uniformly in this stand. This also justifies the choice of models capable of reflecting positive spatial association, rather than those (such as Strauss and auto-Poisson models) suitable for modeling inhibitory processes with negative correlation between neighboring sites.

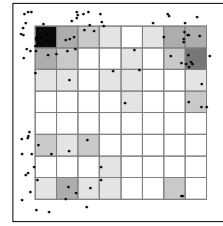
Figure 3 shows the prior (curve) and the estimated posterior (histogram) density function for parameters θ_0 , θ_1 , and θ_2 of model M_D with a 4×4 discretization, with posterior means $(-0.02, -5.28, 1.54)$, and priors and posteriors for the related quantities $\alpha_j^\theta = \exp(\theta_0)$ and $\beta_j^\theta = \exp(-\theta_1)$, the parameters of the gamma distributions for the impulses Γ_j , and the associ-



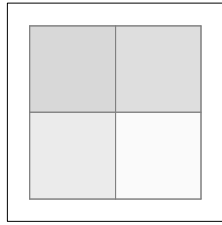
(a) Data, 2×2



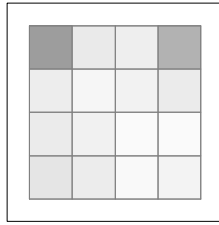
(b) Data, 4×4



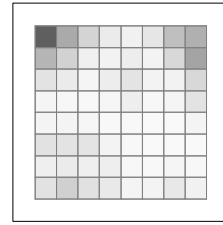
(c) Data, 8×8



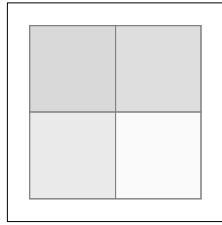
(d) M_S , 2×2



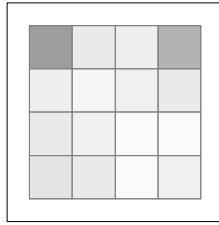
(e) M_S , 4×4



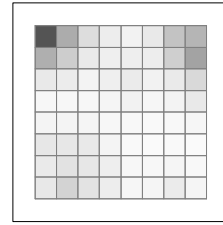
(f) M_S , 8×8



(g) M_D , 2×2



(h) M_D , 4×4



(i) M_D , 8×8

FIGURE 2. Empirical and posterior hickory tree density plots for two models at three discretizations.

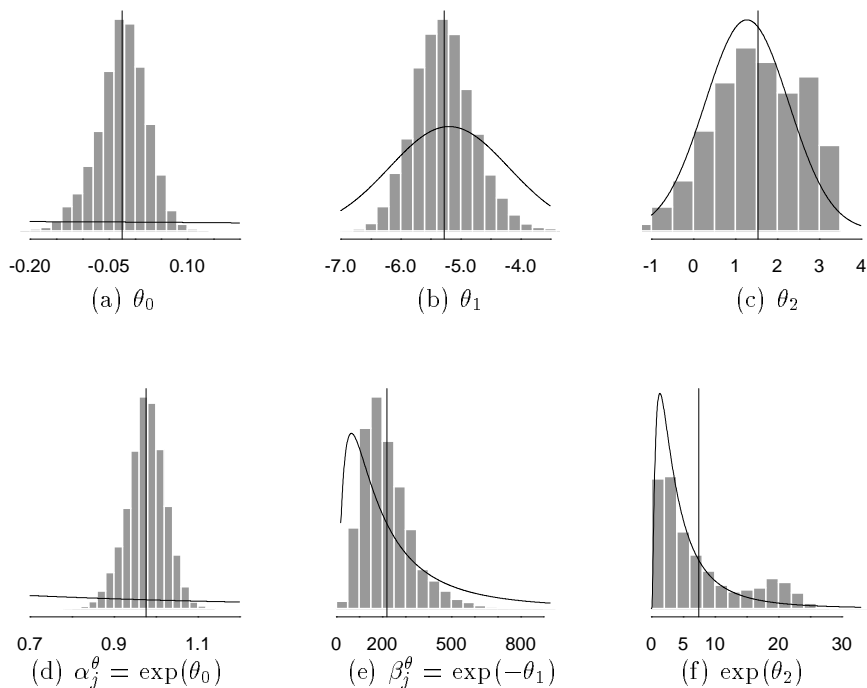


FIGURE 3. Prior (solid line) and posterior (histogram) parameter plots and posterior means (vertical lines) for M_D , 4×4 .

ation measure $\exp(\theta_2)$, with posterior means (0.977, 220.8, 7.449) indicated by vertical lines.

The posterior distribution of $\exp(\theta_2)$ (Figure 3(f)) appears to be bimodal, with most of the mass in both modes well above zero (the sample median is $q_{.50} = 4.708$) but with enough mass near zero to make the case for inhomogeneity ambiguous (the sample 5% quantile is only $q_{.05} = 0.78$); although intended to represent substantial uncertainty, the prior standard deviation of 1.0 for θ_2 appears to make the prior distribution comparatively sharp compared to the likelihood. To ensure that it is the data and not the prior that lead the posterior to favor high values we also tried reparametrizing the model and placing an improper flat prior on the quantity $\exp(\theta_2)$, maintaining the same prior distribution for θ_0 and θ_1 ; the result was a posterior mean of 17.35 and standard deviation of 6.70 for $\exp(\theta_2)$, with quantiles $q_{.50} = 19.25$ and $q_{.05} = 3.21$, revealing much stronger support for inhomogeneity than the normal prior had allowed. Table 1.3 gives the prior and posterior means and standard deviations for $\exp(\theta_2)$ for the different discretizations, with the normal prior distribution on each component of θ .

	Prior	Posterior
2×2	0.128 ± 0.168	0.114 ± 0.108
M_S 4×4	0.279 ± 0.366	0.281 ± 0.156
8×8	0.643 ± 0.843	1.149 ± 0.924
2×2	5.941 ± 7.788	6.309 ± 6.714
M_D 4×4	5.941 ± 7.788	7.449 ± 6.764
8×8	5.941 ± 7.788	14.977 ± 4.227

TABLE 1.3. Posterior and prior mean \pm SD for spatial effect $\exp(\theta_2)$.

5 Discussion

Each model and discretization (except perhaps for the sparse model at 2×2 discretization) shows a strong spatial effect, with $\exp(\theta_2)$ well above zero. Both models show increasing spatial correlation as the partition is refined, suggesting that even more refined partitions may be of interest. An interesting alternative, pursued in [WI95], is to dispense with discretization altogether and model the precise locations of each event; another would be to use formal methods (e.g., Bayes factors) to compare models at different levels of refinement. In doing so one may wish to penalize the additional complexity of more refined models, since our MCMC scheme with data-augmentation (see Appendix) for a $d \times d$ dense model would include d^4 latent variables in addition to θ . Of course the present methods are applicable to other lattice structures (hexagonal, triangular, etc.), and to discrete settings without a lattice structure at all—disease mapping or population modeling, for example, where in place of quadrats we might have counties or regions (whose areas A_i would vary) and where the concern about arbitrary level of discretization may not arise.

We considered two essentially different models, at three levels of aggregation, and found a clear suggestion of spatial irregularity in hickory tree density. The present methods did not permit us either to compare the models and quantify the evidence in favor of one or the other, nor to quantify the evidence against uniformity. Bayes factor calculations and hypothesis tests would be a natural extension.

The present analysis did not include covariates—soil chemistry, elevation and aspect, distance from water source, etc.; an interesting direction for further development would be to extend the models to include covariates known to influence forest composition in Duke Forest ([RPPW93]) in the kernel structure (either through a simple proportional-hazard approach or the more sophisticated penalized likelihood approach of [BC93]). The present data set does not include such covariates, but the quadrat locations could serve as a surrogate for any smoothly varying covariate with a linear spatial trend, or elevation and soil measurements could be taken. The tree density Λ_i is represented in these models as a linear combina-

tion (with weights k_{ij}^θ) of latent variables Γ_j . One way of extending the models to reflect the uncertain dependence on (possibly spatially-varying) covariates is to take for J a larger index set than I , and include among the $\{\Gamma_j\}_{j \in J}$ variables representing possible influences on the tree density—soil features, point sources of water or nutrients, etc.

The data suggest that hickory density is not uniform—rather, that it is higher near the edges of the research plot, evidence that hickories may be urging into this predominantly oak stand or receding from it. An interesting object for further study will be the evolution and prediction of spatial patterns within the plot for each species.

Appendix

Here we describe the data augmentation and Markov chain Monte Carlo (MCMC) computational scheme needed for the posterior analysis presented in Section 4. While the likelihood function for θ from (1.1) is unavailable analytically, we can study the intractable posterior distribution of the uncertain θ and the $\{\Gamma_j\}_{j \in J}$ by simulating steps from an ergodic Markov chain with the posterior for its stationary distribution [GS90, GRS96, Tie94]. This simulation approach requires sampling from the complete conditional distribution of the impulses $\{\Gamma_j\}_{j \in J}$ (given θ and the observed counts $\{N_i = n_i\}_{i \in I}$), a mixture distribution from which direct sampling is cumbersome; following [TW87], we introduce latent “augmentation” variables $\{N_{ij}\}_{i,j \in I,J}$ which can be sampled easily and which lead to conjugate (gamma) complete conditional distributions for the $\{\Gamma_j\}_{j \in J}$.

Conditional on the values $\{\gamma_j\}_{j \in J}$ of $\{\Gamma_j\}_{j \in J}$ and $\{n_i\}_{i \in I}$ of $\{N_i\}_{i \in I}$ set $\lambda_{ij} \equiv k_{ij}^\theta \gamma_j$, $\lambda_{i+} \equiv \sum_{j \in J} \lambda_{ij}$, $p_{ij} \equiv \lambda_{ij} / \lambda_{i+}$, and let $\{N_{ij}\}_{i \in I}$ have independent multinomial $\text{MN}(n_i, p_{i\cdot})$ distributions; note that the data can be recovered as the row-sums $n_i = N_{i+} \equiv \sum_{j \in J} N_{ij}$. The random variables N_{ij} might be regarded as the portions of count n_i associated with the j^{th} impulse, $j \in J$; sometimes these impulses may be interpreted as causes for the events (nutrients, water, or light resources, for example). Conditional on the augmented data the impulses have independent gamma distributions $\Gamma_j \stackrel{\text{ind}}{\sim} \text{Ga}(\alpha_j^\theta + n_{+j}, (\beta_j^\theta + k_{+j}^\theta)^{-1})$, where $n_{+j} \equiv \sum_{i \in I} n_{ij}$ but $k_{+j}^\theta \equiv \sum_{i \in I} A_i k_{ij}^\theta$, leading to the hybrid Gibbs/Metropolis MCMC scheme:

0. Select a parametric family $\{\alpha_j^\theta, (\beta_j^\theta)^{-1}\}$ of shape and scale parameters for impulses’ gamma distributions $\Gamma_j \sim \text{Ga}(\alpha_j^\theta, (\beta_j^\theta)^{-1})$ and a parametric kernel $\{k_{ij}^\theta\}_{i,j \in I,J}$, a prior density $\pi(\theta)$ on Θ , a transition probability density $q(\theta, \theta^*)$, and initial values $\theta^0 \in \Theta$ and $\{n_{ij}^0\}_{i,j \in I,J} \subset \mathcal{N}$ satisfying $n_{i+}^0 = n_i$. Generate successive points starting at $t = 1$ as follows:
 1. Gibbs step to update the impulses variables:

Given $\{n_{ij}^{t-1}\}_{i,j \in I,J}$ and θ^{t-1} ,

- set $\alpha_j^t \equiv \alpha_j^{\theta^{t-1}} + n_{+j}^{t-1}$ and $\beta_j^t \equiv \beta_j^{\theta^{t-1}} + k_{+j}^{\theta^{t-1}}$;
- generate $\gamma_j^t \stackrel{\text{ind}}{\sim} \text{Ga}(\alpha_j^t, (\beta_j^t)^{-1})$;
- set $\lambda_{ij}^t \equiv k_{ij}^{\theta^{t-1}} \gamma_j^t$, $\lambda_{i+}^t \equiv \sum_{j \in J} \lambda_{ij}^t$, and $p_{ij}^t \equiv \lambda_{ij}^t / \lambda_{i+}^t$.

2. Gibbs step to update the augmentation points:

Given $\{n_i\}_{i \in I}$, $\{\gamma_j^t\}_{j \in J}$, and θ^{t-1} ,

- generate $n_{ij}^t \sim \text{MN}(n_i, p_{ij}^t)$, independent for $i \in I$.

3. Metropolis step to update the parameter θ :

Given $\{n_{ij}^t\}_{i,j \in I,J}$, $\{\gamma_j^t\}_{j \in J}$, and θ^{t-1} ,

- generate a new candidate $\theta^* \sim q(\theta^{t-1}, \theta^*) d\theta^*$;
- calculate the Hastings acceptance probability

$$\begin{aligned}
 P^* &\equiv \frac{\pi(\theta^*) q(\theta^*, \theta^{t-1})}{\pi(\theta^{t-1}) q(\theta^{t-1}, \theta^*)} \prod_{i,j \in I,J} \left[\frac{k_{ij}^{\theta^*}}{k_{ij}^{\theta^{t-1}}} \right]^{n_{ij}^t} \\
 &\times \prod_{j \in J} \left[\frac{(\gamma_j^t)^{\alpha_j^{\theta^*} + n_{+j}^t - 1} (\beta_j^{\theta^*})^{\alpha_j^{\theta^*}} \Gamma(\alpha_j^{\theta^{t-1}})}{(\gamma_j^t)^{\alpha_j^{\theta^{t-1}} + n_{+j}^t - 1} (\beta_j^{\theta^{t-1}})^{\alpha_j^{\theta^{t-1}}} \Gamma(\alpha_j^{\theta^*})} \right] \\
 &\times \exp \left(- \sum_{j \in J} (\beta_j^{\theta^*} - \beta_j^{\theta^{t-1}} + k_{+j}^{\theta^*} - k_{+j}^{\theta^{t-1}}) \gamma_j^t \right)
 \end{aligned}$$

- generate $\theta^t \sim \begin{cases} \theta^* & \text{with probability } \min(1, P^*) \\ \theta^{t-1} & \text{otherwise.} \end{cases}$

4. Increment $t \leftarrow t + 1$ and return to step 1.

Acknowledgments

The authors would like to thank Norm Christensen and Robert Peet for sharing data and insight, Merlise Clyde for helpful conversations, and the editors and an anonymous referee for helpful suggestions. This analysis was supported by U.S. E.P.A. grant CR822047-01-0, U.S. N.S.F. grant DMS-9626829, and the Deutsche Forschungsgemeinschaft; the data management and refinement by Robert Peet was supported by U.S. N.S.F. grants BSR-8502430, BSR-8905926, and BSR-9107357.

6 REFERENCES

- [Bar64] M. S. Bartlett. The spectral analysis of two-dimensional point processes. *Biometrika*, 51:299–311, 1964.
- [BC93] Norman E. Breslow and David G. Clayton. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421):9–25, March 1993.
- [Bes74] Julian Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B (Methodological)*, 35(2):192–236, 1974. (With discussion).
- [Bor53] Frank H. Bormann. The statistical efficiency of sample plot size and shape in forest ecology. *Ecology*, 34:474–487, 1953.
- [BPBG96] Luisa Bernardinelli, Cristian Pascutto, Nicola G. Best, and Wally R. Gilks. Disease mapping with errors in covariates. To appear in *Statistics in Medicine*, 1996.
- [BYM91] Julian Besag, Jeremy York, and Annie Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43:1–59, 1991. (With discussion).
- [CC89] Noel A. C. Cressie and Ngai H. Chan. Spatial modeling of regional variables. *Journal of the American Statistical Association*, 84(406):393–401, June 1989.
- [Chr77] Norman L. Christensen. Changes in structure, pattern and diversity associated with climax forest maturation in Piedmont, North Carolina. *American Midland Naturalist*, 97:176–188, 1977.
- [CK87] David G. Clayton and John Kaldor. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43(3):671–681, September 1987.
- [Cre93] Noel A. C. Cressie. *Statistics for Spatial Data*. John Wiley & Sons, New York, NY, USA, 1993.
- [DM54] F. N. David and P. G. Moore. Notes on contagious distributions in plant populations. *Annals of Botany*, 18:47–53, 1954.
- [GRS96] Wally R. Gilks, Sylvia Richardson, and David J. Spiegelhalter, editors. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, New York, NY, USA, 1996.

- [GS90] Alan E. Gelfand and Adrian F. M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409, June 1990.
- [Mat63] Georges Matheron. Principles of geostatistics. *Economic Geology*, 58:1246–1266, 1963.
- [Rip81] Brian D. Ripley. *Spatial Statistics*. John Wiley & Sons, New York, NY, USA, 1981.
- [RMG⁺95] Sylvia Richardson, Christine Monfort, Martyn Green, Gerald Draper, and Colin Muirhead. Spatial variation of natural radiation and childhood leukaemia incidence in Great Britain. *Statistics in Medicine*, 14:2487–2501, 1995.
- [RPPW93] Rebecca A. Reed, Robert K. Peet, Mike W. Palmer, and Peter S. White. Scale dependence of vegetation-environment correlations: A case study of a North Carolina piedmont woodland. *Journal of Vegetation Science*, 4:329–340, 1993.
- [Str75] David J. Strauss. A model for clustering. *Biometrika*, 62(2):467–476, 1975.
- [Tie94] Luke Tierney. Markov chains for exploring posterior distributions. *Annals of Statistics*, 22(4):1701–1762, December 1994. (With discussion).
- [TW87] Martin A. Tanner and Wing H. Wong. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–550, June 1987.
- [WI95] Robert L. Wolpert and Katja Ickstadt. Gamma/Poisson random field models for spatial statistics. Discussion Paper 95-43, Duke University ISDS, USA, December 1995.