

Space–time modelling of lightning-caused ignitions in the Blue Mountains, Oregon

Carlos Díaz-Avalos, David L. Peterson, Ernesto Alvarado, Sue A. Ferguson, and Julian E. Besag

Abstract: Generalized linear mixed models (GLMM) were used to study the effect of vegetation cover, elevation, slope, and precipitation on the probability of ignition in the Blue Mountains, Oregon, and to estimate the probability of ignition occurrence at different locations in space and in time. Data on starting location of lightning-caused ignitions in the Blue Mountains between April 1986 and September 1993 constituted the base for the analysis. The study area was divided into a pixel–time array. For each pixel–time location we associated a value of 1 if at least one ignition occurred and 0 otherwise. Covariate information for each pixel was obtained using a geographic information system. The GLMMs were fitted in a Bayesian framework. Higher ignition probabilities were associated with the following cover types: subalpine herbaceous, alpine tundra, lodgepole pine (*Pinus contorta* Dougl. ex Loud.), whitebark pine (*Pinus albicaulis* Engelm.), Engelmann spruce (*Picea engelmannii* Parry ex Engelm.), subalpine fir (*Abies lasiocarpa* (Hook.) Nutt.), and grand fir (*Abies grandis* (Dougl.) Lindl.). Within each vegetation type, higher ignition probabilities occurred at lower elevations. Additionally, ignition probabilities are lower in the northern and southern extremes of the Blue Mountains. The GLMM procedure used here is suitable for analysing ignition occurrence in other forested regions where probabilities of ignition are highly variable because of a spatially complex biophysical environment.

Résumé : Des modèles linéaires généraux mixtes (GLMM) ont été utilisés pour étudier l'effet du couvert végétal, de l'altitude, de la pente et des précipitations sur la probabilité d'allumage dans les Blue Mountains, en Oregon, et pour évaluer la probabilité qu'un allumage se produise à différents endroits dans l'espace et le temps. Des données sur la localisation des allumages causés par la foudre dans les Blue Mountains entre les mois d'avril 1986 et septembre 1993 ont servi de base à l'analyse. L'aire d'étude a été divisée en un réseau de pixel–temps. Une valeur de 1 ou 0 a été attribuée à chaque endroit correspondant à un pixel–temps selon qu'au moins un allumage est survenu ou non. Les informations associées à chaque pixel ont été obtenues à l'aide d'un système d'information géographique. Les GLMM ont été ajustés selon une structure bayésienne. De plus fortes probabilités d'allumage étaient associées aux types de couverts suivants : plantes herbacées subalpines, toundra alpine, pin lodgepole (*Pinus contorta* Dougl. ex Loud.), pin à blanche écorce (*Pinus albicaulis* Engelm.), épinette d'Engelmann (*Picea engelmannii* Parry ex Engelm.), sapin subalpin (*Abies lasiocarpa* (Hook.) Nutt.) et sapin grandissime (*Abies grandis* (Dougl.) Lindl.). Pour chaque type de couvert végétal, les probabilités d'allumage étaient plus élevées à plus faible altitude. De plus, les probabilités d'allumage étaient plus faibles aux extrémités nord et sud des Blue Mountains. La procédure GLMM utilisée ici convient pour analyser l'occurrence des allumages dans d'autres régions couvertes de forêt où les probabilités d'allumage sont très variables à cause d'un environnement biophysique dont la configuration spatiale est complexe.

[Traduit par la Rédaction]

Received September 15, 1999. Accepted May 2, 2001.
Published on the NRC Research Press Web site at
<http://cjfr.nrc.ca> on September 6, 2001.

C. Díaz-Avalos.¹ Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Apartado Postal 20-726 Del. Alvaro Obregón, México D.F., C.P. 0100.

D.L. Peterson. U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Cascadia Field Station, P.O. Box 352100, Seattle, WA 98195, U.S.A.

E. Alvarado. College of Forest Resources, University of Washington, P.O. Box 352100, Seattle, WA 98195, U.S.A.

S.A. Ferguson. USDA Forest Service, Pacific Northwest Research Station, 4043 Roosevelt Way NE, Seattle, WA 98105, U.S.A.

J.E. Besag. Department of Statistics, University of Washington, P.O. Box 354322, Seattle, WA 98195, U.S.A.

¹Corresponding author (e-mail: carlos@sigma.iimas.unam.mx).

Introduction

Fire is the most important natural disturbance in forests of western North America (Peterson 1998; Schmoldt et al. 1999) and a critical component of forest dynamics and biogeochemical cycling across a wide range of spatial scales (Rogers 1996). In the Pacific Northwest region of North America, fire was historically the dominant disturbance, with much shorter fire-return intervals in low-precipitation eastern British Columbia, Washington, and Oregon than in the high-precipitation west (Agee 1990).

During the 19th century, fire was such a common occurrence in northeastern Oregon and southeastern Washington that the Blue Mountains (Fig. 1) derived their name from the presence of smoke, which partially obscured the mountains and lingered in adjacent valleys (Agee 1993). By the early 20th century, fire-return intervals increased and forest area burned decreased, because fire was largely excluded from these forests because of a variety of human influences, in-

Fig. 1. Geographic location of the Blue Mountains area.



cluding fire suppression and fuel modification caused by timber harvest, roads, and agricultural land (Mutch et al. 1993; Quigley et al. 1996). As a result, stand densities in forests of the Blue Mountains have gradually increased. This has led to management concerns regarding the decreased resistance of current stands to insects and pathogens (Hessburg et al. 1994; Johnson 1994). In addition, longer fire-return intervals have led to higher fuel loads in many areas, thereby increasing the potential for stand-replacement fires, as opposed to understory fires, which were more common prior to the 20th century (Maruoka 1994; Heyerdahl 1997).

The United States Forest Service is now using prescribed fire on a limited basis to reduce fuel loads and stem densities in the Blue Mountains (Mutch et al. 1993; Johnson 1994) and other areas of the Pacific Northwest (Agee 1998), but the best and safest way to reintroduce fires in those forests is still not clear.

The spatiotemporal dynamics of disturbance (Boychuk et al. 1997; Lertzman and Fall 1998; Lertzman et al. 1998; McKenzie 1998) is an important consideration in large-scale, long-term resource management in the Blue Mountains, particularly on public lands where forest landscapes with "natural" disturbance characteristics are often a management objective (DeLong 1998). Resource managers need to know where and when fires are most likely to occur, and which biological and physical factors are related to the presence of ignitions. Fire-history studies (e.g., Agee et al. 1990;

Swetnam 1993; Johnson and Wowchuk 1993; Johnson and Gutsell 1994; Bessie and Johnson 1995) and other reconstruction methods (Reed 1994; McKelvey and Busse 1996) are typically used to quantify the spatial and temporal domains of fire in forest ecosystems. Fire-return intervals based on fire-history studies have been calculated for many of the common forest associations in the Blue Mountains (Maruoka 1994; Heyerdahl 1997). However, fire-history data do not necessarily reflect actual fire occurrence, particularly in situations in which fires are small and highly dispersed (Johnson and Gutsell 1994; McKenzie et al. 2000), because the proportion ignitions/strikes does not necessarily correlate with the magnitude of fire effects, as registered in the fire history. Therefore, data on fire occurrence (e.g., Cunningham and Martell 1973) and on the relationship of the biophysical environment to ignitions (Alvarado et al. 1998) are needed to provide a more complete understanding of fire disturbance as an input to fire-management planning (Mills and Bratten 1982; Bratten 1984). As a result of the stochastic nature of the spatiotemporal occurrence of lightning-caused fires, a statistical modeling approach is the best option to analyse fire incidence data, because inferences based on raw data are potentially misleading (Besag et al. 1991). One point of interest is to test if the observed spatial pattern of fire occurrences can be explained by means of other variables showing spatial variation. For example, it is known that some vegetation types are more prone to fire than others

and that, for a given vegetation type, fire probabilities vary because of factors such as elevation. If this is true, then we can use this information to predict the probability of a lightning-caused fire at a given spatial location. This prediction process is potentially useful for areas where information about lightning density, temperature, and other variables more closely related to lightning ignitions is not available.

In this study, we analysed the initiation of lightning-caused fires in the Blue Mountains as a space–time stochastic process. The data we analyzed refer to lightning-caused fires that were ignited, detected, and reported in the study area. Although these fires are a subset of all fires that actually ignited in the Blue Mountains, we will refer to them as “ignitions”. The primary objectives of our study were to (i) calculate the probability of ignition at different locations in space and time at intermediate scales and (ii) determine how specific features of the biophysical environment are linked to the probability of ignition. The analysis was used to develop statistically derived maps that can be used to quantify fire disturbance and landscape patterns and as a potential input to fire-management decisions (e.g., deployment of fire suppression resources; Mees 1978). First, we present the basic assumptions of a stochastic mechanism describing the process of forest ignition in space and in time. Second, we describe a model-fitting process under the Bayesian paradigm that uses Markov chain Monte Carlo simulations. Finally, we discuss the value of our modeling approach for quantifying the distribution of lightning-caused ignitions in the Blue Mountains and in other forested regions where ignition occurrence data are available.

Statistical approach and assumptions

Thunderstorms are the generators of lightning, but not every lightning strike ignites a fire because of variability in the moisture, bulk density, and depth of fuels (Kourtz and Todd 1992). Thus, the number of lightning strikes that result in an ignition is correlated to these fuel quality factors. Despite the presence of variability in fuels at large spatial scales, it is possible to assume homogeneity at smaller scales (Rothermel 1972; Andrews 1986). For example, forest stands with comparable vegetation, elevation, and other biophysical characteristics can be expected to have similar fuel characteristics and, therefore, reasonably similar response to lightning strikes, resulting in a similar probability of ignition within a stand (Bratten 1984; Quinby 1987).

Based on this general assumption, we divided the time and the space defined by the study area into pixels. Thus, the system under study is represented by a three-dimensional space–time array of N pixels observed at times $t = 1, 2, \dots, T$. For each pixel–time combination we assume the existence of a random variable Y_{it} taking the value 1 if at least one ignition occurred at pixel i at time t and 0 otherwise. We use the general relationships $P(Y_{it} = 1) = \mu_{it}$ and $P(Y_{it} = 0) = 1 - \mu_{it}$ to denote the probabilities of ignition occurrence and their complement, respectively. It is assumed that the probability of ignition is constant within each pixel. Our goals are to estimate the probabilities of ignition μ_{it} and to investigate their relationships with a set of biological and topographical variables.

The dichotomous nature of the observations in this approach allows the use of generalized linear models of the binomial family. Generalized linear models relate the expected value of the response variable Y to a set of p covariates through a linear predictor $\xi = \mathbf{Z}\boldsymbol{\beta}$, where \mathbf{Z} is a matrix of covariates and $\boldsymbol{\beta}$ is a vector $p \times 1$ of coefficients, and a link function g , with $g(\mu) = \mathbf{z}^T\boldsymbol{\beta}$.

When the response is dichotomous, the most commonly used link functions are the logit, probit, and complementary log–log (McCullagh and Nelder 1989). In our analysis we used the logit link function, $\xi = g(\mu) = \log(\mu/(1 - \mu)) = \mathbf{z}^T\boldsymbol{\beta}$, because it permits interpretation of the risk of ignition in terms of odds. The relationship between the probability of ignition and the linear predictor is given by $\mu = g^{-1}(\xi) = e^{\mathbf{z}^T\boldsymbol{\beta}}/(1 + e^{\mathbf{z}^T\boldsymbol{\beta}})$. In this paper, we assume that the logit of ignition is linearly related to factors such as elevation, vegetation, slope, aspect, and precipitation. Although factors such as stand age, fuel bulk density, fuel depth, lightning strike density, and air temperature are closely related to ignition occurrence, lack of data about them for the study area precluded their use in the model. Nevertheless, some information about them is obtained through the covariates included in the analysis because of the presence of correlation between them. For example, a moist fuel bed may be hit by a lightning strike without starting an ignition, so fuel moisture and lightning strike density are negatively associated. Similarly, the number of lightning strikes that ignite a fire is different for different vegetation covers (Kourtz and Todd 1992). Because of the time scale of the observed data, we consider that vegetation cover for any pixel is constant over the time window of the analysis. The same is obviously true for elevation, slope, and aspect, whereas precipitation has considerable temporal variation.

The logistic model in its original form assumes that the responses at the different pixels are independent (Nelder and Weddeburn 1972). Although this assumption may be adequate in some instances, neighbouring pixels may have similar vegetation cover, and they may also be influenced by the same macroclimatic factors, making it reasonable to suspect the existence of spatial association between neighbouring pixels. To accommodate this in the model, we include a random spatiotemporal effect ψ_{it}^* . Thus, for example, a model that includes only main effects of vegetation (V), elevation (E), slope (S), aspect (A), precipitation (R), and the spatiotemporal effect plus the grand mean (α) has the form

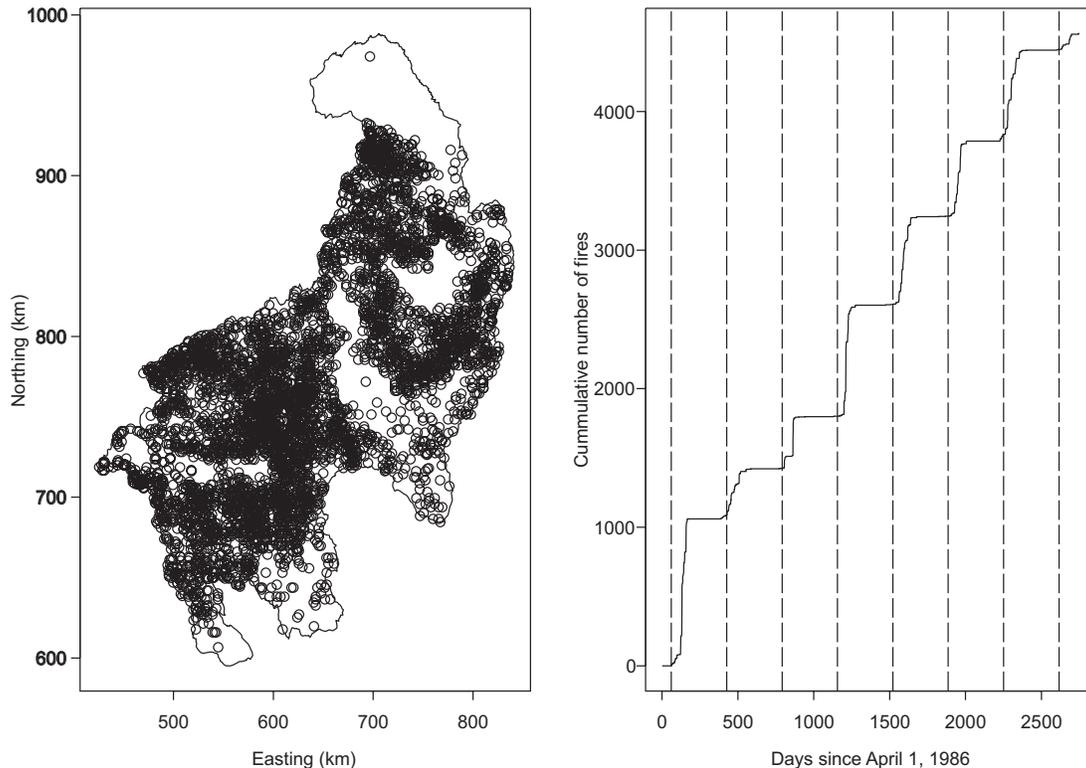
$$\begin{aligned}\xi_{it} &= \alpha + V_i + E_i + S_i + A_i + R_{it} + \psi_{it}^* \\ &= V_i + E_i + S_i + A_i + R_{it} + \psi_{it}\end{aligned}$$

or in the more common notation:

$$[1] \quad \xi_{it} = \mathbf{z}_i^T\boldsymbol{\beta} + \gamma R_{it} + \psi_{it}$$

where $\boldsymbol{\beta}$ is a vector whose components are the coefficients related to vegetation, elevation, slope, and aspect and γ is the coefficient for the effect of precipitation on the linear predictor. Note that $\psi_{it} = \alpha + \psi_{it}^*$. The spatiotemporal term ψ_{it} can be considered as a surrogate for unobserved variables that are correlated in space and time (Besag et al. 1995). Equation 1 corresponds to the generalized linear mixed models (GLMM), a model class useful in problems that involve the mapping of risks (Clayton and Kaldor 1987). It states that

Fig. 2. Spatial pattern of ignition occurrence from April 1, 1986, to September 9, 1993 (a), and cumulative distribution in time of ignition occurrence (b) in the Blue Mountains area. Vertical lines correspond to June 1 of every year. Latitude and longitude scales are in kilometers.



each pixel has its own baseline risk of ignition $e^{\psi_{it}} / (1 + e^{\psi_{it}})$ and that the combined effect of the covariates and time is to multiply the odds of ignition by $e^{z_i^T \beta + \gamma R_{it}}$. Equation 1 may be compared statistically against models with either a more complicated or simpler structure by comparing the changes in deviance. Likelihood-based inferences in GLMM require numerical integration techniques to obtain the maximum likelihood estimates, the score equations, and the information matrix (Clayton 1997). Standard errors for the parameter estimates are obtained from asymptotic approximations (Breslow and Clayton 1993). For data with a high proportion of individuals with null response, the standard error of the estimates as well as the usual goodness-of-fit tests are not adequately computed, making the results of the analysis unreliable (McCullagh and Nelder 1989; Hosmer and Lemeshow 1989). The hierarchical nature of GLMM makes the use of Bayesian procedures attractive (Clayton and Kaldor 1987). An appealing feature of the Bayesian approach in cases such as our application is that the standard error of the estimates does not depend on asymptotic assumptions. This diminishes the influence of zeroes in the vector of responses on the reliability of the results. The Bayesian approach also allows to assess the uncertainty in estimated random effects and functions of model parameters through posterior distributions.

Methods

We analysed data on the starting locations of lightning-caused fires occurring in the Blue Mountains area (Fig. 1) between April 1986 and June 1993. These data were compiled from the interior

Columbia River basin integrated assessment (Quigley et al. 1996). The total number of lightning-caused fires between those dates was 4482, representing >80% of the total number of fires (the others being human caused) and >94% of the area burned in the Blue Mountains. Starting location and estimated date of ignition are available for each fire. Clustering in both space and time is apparent; most of the ignitions in the Blue Mountains occur during the summer months (Fig. 2), with low incidence during the rest of the year (Agee 1994).

The exact locations at which the lightning ignitions occurred were known, and a point process model (Karr 1986; Diggle 1983) with an intensity function $\lambda(x,y)$ depending on covariates was also considered as a possible approach. Point processes have been suggested as an alternative to avoid the problem of obtaining scale-dependent patterns when analyzing occurrence of events in space and time (Wolpert and Ikstadt 1998). However, when the intensity function includes covariates the resulting maps depend on the resolution of the covariate information. Another drawback is that hypothesis tests for the parameter estimates of the intensity functions are based on asymptotic assumptions valid only as the study area size becomes infinite, and restrictions have to be imposed to $\lambda(x,y)$ to ensure that the likelihood function has only one maximum (Cressie 1993).

The study area was partitioned in pixels of 2.5×2.5 km, resulting in a total of $N = 8089$ pixels for the Blue Mountains. The time axis was divided into quarterly intervals as follows: 1 March – 31 May (spring), 1 June – 31 August (summer), 1 September – 30 November 30 (fall), 1 December – 28 February (winter), resulting in a total of $T = 30$ quarterly intervals. For each pixel–time combination we assigned the value 1 if at least one ignition occurred there and 0 otherwise. Note that this only implies that fire was present in that space–time location and not that the whole pixel burned. Statistical methods to predict the burned area are beyond the scope of this paper.

Using digital elevation and vegetation maps available for the area, we used a geographic information system (GIS) to attach data about vegetation type, elevation, and slope at each pixel center. Precipitation was incorporated from a data base of mean monthly precipitation for the area, estimated with the PRISM model (Daly et al. 1994), which uses regression and digital elevation maps to interpolate data from existing weather stations over complex terrain. These interpolated precipitation data have been used in a wide variety of coarse-scale models (Lenihan et al. 1998; Ohmann and Spies 1998; McKenzie and Halpern 1999; McKenzie et al. 2000).

Choosing an adequate pixel size is a problem similar to that of choosing class length in histograms. Large pixel size produces loss of information about multiple ignition occurrences, whereas too small pixel size makes hard to detect any spatial pattern in the data. In our case, the selection of the 2.5×2.5 km scale was done to match the resolution of the digital vegetation map, which was the coarsest covariate map available. The idea was to make a trade-off between resolution and spatial homogeneity for the covariates used in the model. Digital maps available for the other covariates used in the analysis were converted to the 2.5×2.5 km scale using GIS software. On the time axis, quarterly time intervals were chosen to differentiate periods with high ignition incidence (summer), low ignition incidence (winter), and two transition intervals (fall and spring). Multiple ignition occurrences in a single pixel–time combination were never greater than two. In the context of our model this produces negligible differences in the estimated ignition probabilities.

There are 21 vegetation types represented in the Blue Mountains, based on the classification used by the USDA Forest Service for the Columbia River basin (Clarke and Bryce 1997). Because the categorical nature of this variable would result in an unacceptably large number of coefficients for reasonable model interpretation, we grouped vegetation types into classes to reduce the number of parameters (Table 1). Aggregation of vegetation classes is based on similarities in physiognomy, fuel properties, and fire effects.

Because the possible loss of representativity of the covariates at the scale we used, we performed correlation tests between the covariate values associated to the actual location of each ignition in the data base and the covariate values of the associated pixels for each quarter. The results gave correlation values between 0.62 and 0.87 ($p < 0.005$), from which we conclude that the pixelized data base is providing a fair amount of information about the actual values of the covariates.

Our approach to model fitting is in the Bayesian context, that is, if θ is a vector of k components containing all the parameters in the model, statistical inferences about θ are based on the posterior distribution:

$$[2] \quad f(\theta | y) = \frac{L(\theta, y) \pi(\theta)}{\int_{\Theta} L(\theta, y) \pi(\theta) d\theta} \propto L(\theta, y) \pi(\theta)$$

where $\pi(\theta)$ is the prior distribution of the model parameters, reflecting the researcher’s “prior” belief about plausible values for θ without knowledge of the data, and $L(\theta, y)$ is the likelihood of the data given θ (Bernardo and Smith 1994). The statement in eq. 2 is known as Bayes theorem. Bayesian inferences about the i th component of θ are done using the marginal posterior distribution, obtained by integrating eq. 2 over the remaining components of θ . In a univariate case for example, we may use the mode of the posterior distribution as a point estimate of θ , and an interval based on percentiles of such a distribution (known as credible interval) can be used to summarize its variability.

Obtaining the posterior distribution of θ analytically is not always possible, and most of the time one must resort to numerical integration to compute expectations and other functions of the model parameters. One method is by simulating samples from the

Table 1. Vegetation classes used in the analysis are derived from vegetation types used in the Columbia River basin assessment (Clarke and Bryce 1997).

Vegetation class	Vegetation type
1	Mixed grass, agricultural land, and shrubs
2	Seral shrub, aspen
3	Subalpine herbaceous, alpine tundra
4	Engelmann spruce–subalpine fir, lodgepole pine, whitebark pine
5	Interior Douglas-fir
6	Grand fir
7	Interior ponderosa pine
8	Bluebunch wheatgrass, western juniper – big sagebrush – bluebunch wheatgrass, Idaho fescue – bluebunch wheatgrass, crested wheatgrass, Utah juniper – big sagebrush – bluebunch wheatgrass
9	Mountain big sagebrush, Wyoming big sagebrush, low sagebrush, salt desert shrub

posterior distribution of the parameters and approximating the necessary integrals numerically using Markov chain Monte Carlo (MCMC) techniques (Gilks et al. 1996). With those simulated samples, point estimates of the model parameters as well as their standard errors are obtained by computing their sampling moments. For a good introduction to the use and theoretical fundamentals of MCMC methods, the reader is referred to Gelman et al. (1995).

In the first part of the analysis, we fitted a model of the form of eq. 1 for each one of the 30 quarters, that is, we fitted the following model:

$$[3] \quad \xi_i = \mathbf{z}_i^T \boldsymbol{\beta} + \gamma R_i + \mathbf{x}'_i \mathbf{u} + \psi_i$$

where $\boldsymbol{\beta}$ is a vector with the coefficients related to the main effects of vegetation, elevation, slope and aspect, γ is the coefficient for the effect of precipitation, and \mathbf{u} is a vector of coefficients related to the first-order interactions between vegetation, elevation, slope, aspect, and precipitation for each quarter. This transversal exploratory analysis was done to detect possible trends and seasonality in the parameters and to gain insight about the form of a general model for the space–time domain. We assumed a flat, non-informative prior distribution (Box and Tiao 1973) for the non-spatial parameters in our model, which allows them to be assigned any arbitrary initial values. The spatiotemporal term is considered as a surrogate of unobserved covariates showing smooth spatial variation. Equation 3 implies that the odds of ignition are affected multiplicatively by factors acting at large scale (e.g., covariates) as well as factors acting at shorter scales (spatial term). Following Besag et al. (1991), we assumed a Gaussian pairwise difference prior distribution for the spatiotemporal term, with precision λ :

$$[4] \quad \pi(\psi) \propto \lambda^{0.5N} |\mathbf{W}|^{0.5} e^{-0.5\lambda\psi^T \mathbf{W}\psi}$$

where $\boldsymbol{\psi} = (\psi_1, \psi_2, \dots, \psi_N)$ is the vector of spatiotemporal components, \mathbf{W} is a matrix with $W_{ii} = \nu$, $W_{ij} = -1$ if pixels i and j are neighbours and $W_{ij} = 0$ otherwise, and $|\mathbf{W}|$ denotes the product of the nonzero eigenvalues of \mathbf{W} . We are weighting equally each direction, because the geographic scale we are working it is not detailed enough to detect possible anisotropies. The prior density in eq. 4 belongs to the class of nonstationary Gaussian intrinsic auto-regressions and may be considered as the stochastic equivalent of linear interpolation (Besag et al. 1991). Because both the columns and rows of \mathbf{W} add to zero, eq. 4 is improper. However, the full conditional densities necessary to make statistical inferences about

the ψ_i , $i = 1, 2, \dots, N$, are well defined (Besag et al. 1995). For the precision we assumed a $\Gamma(1, 1)$ prior density, which allows initially low values for λ and, therefore, high variability in ψ . The model is completely specified by further assuming independence between the components of $\theta = (\beta, \gamma, \mathbf{u}, \psi, \lambda)$ and by assuming that the observations Y_i are conditionally independent. The posterior density of the parameters is proportional to

$$[5] \quad \left(\prod_{i=1}^N \frac{e^{y_i(\mathbf{z}_i^T \beta + \gamma R_i + \mathbf{x}_i^T \mathbf{u} + \psi_i)}}{1 + e^{\mathbf{z}_i^T \beta + \gamma R_i + \mathbf{x}_i^T \mathbf{u} + \psi_i}} \right) \times (\lambda^{0.5N} |\mathbf{W}|^{0.5} e^{-0.5 \lambda \psi^T \mathbf{W} \psi}) \times (\lambda^{a-1} e^{-b\lambda})$$

where the first term is the likelihood for the parameters given the sample. MCMC methods are usually more efficient if one works with the full conditional distribution of one set of parameters given the others. For any component θ_k of the vector parameter θ , these full conditionals are proportional only to those terms of eq. 5 that include θ_k . From eq. 5, the full conditional distributions are

$$[6] \quad \pi(\beta, \mathbf{u} | \beta, \mathbf{u}^-) \propto \left(\prod_{i=1}^N \frac{e^{y_i(\mathbf{z}_i^T \beta + \mathbf{x}_i^T \mathbf{u})}}{1 + e^{\mathbf{z}_i^T \beta + \gamma R_i + \mathbf{x}_i^T \mathbf{u} + \psi_i}} \right)$$

$$[7] \quad \pi(\gamma | \gamma^-) \propto \left(\prod_{i=1}^N \frac{e^{y_i(\gamma R_i)}}{1 + e^{\mathbf{z}_i^T \beta + \gamma R_i + \mathbf{x}_i^T \mathbf{u} + \psi_i}} \right)$$

$$[8] \quad \pi(\psi_i | \psi_i^-) \propto \left(\frac{e^{y_i(\psi_i)}}{1 + e^{\mathbf{z}_i^T \beta + \gamma R_i + \mathbf{x}_i^T \mathbf{u} + \psi_i}} \right) e^{-0.5 v_i \lambda (\psi_i - \bar{\psi}_i)^2}$$

$$[9] \quad (\lambda | \lambda^-) \sim \Gamma \left(a + 0.5N, b + 0.5 \sum_{i=1}^N v_i (\psi_i - \bar{\psi}_i)^2 \right)$$

where the notation $(\cdot | \cdot^-)$ means the conditional distribution of the one set of parameters given the rest of the components in the model. In general, if the full conditionals are proportional to the likelihood of the observations and if not all the responses are zero, they are proper densities (Zellner and Rossi 1984; Ibrahim and Laud 1991). When using a flat noninformative prior for the coefficients of the linear predictor, the resulting full conditionals are proportional to the likelihood of the data. Thus, eqs. 6 and 7 are proper densities. Equation 8 corresponds to a density that is proportional to the likelihood of the responses times a Gaussian kernel and is, therefore, proper. A FORTRAN program was written to simulate samples from the full conditionals. To avoid numerical problems during the MCMC computations, elevation values were normalized to

$$[10] \quad d = \frac{\text{elevation} - 1450}{396.36}$$

where the value 1450 is the mean elevation of the study area in metres and 396.36 is the standard deviation for the elevation in metres. We used the Hastings (1970) algorithm to update the nonconjugate distributions (eqs. 6–8), generating the candidate values with a Gaussian density centered at the current value for each parameter. The dispersions of these candidate generators were tuned to get acceptance rates in the 25–60% range, but because of the different number of ignitions at each quarter, the tuning values were different for each time period. We ran five independent chains for each time period, each one with a length of 10 000 iterations. The burning in time was taken as 5000, and we kept every fifth observation of the remainder of each chain for the posterior analysis. Thus, our posterior analysis is based on five samples of size 1000 for each parameter. Inferences based on MCMC simulation assume

that the Markov chain converges to the target distribution, and it is critical to assess if we have such convergence. In our application, we checked convergence using the method described by Gelman and Rubin (1992) and Besag et al. (1995).

In the Bayesian framework, model selection and goodness of fit assessment are done based on the posterior and the predictive distribution of a discrepancy D , respectively (Gelman et al. 1996). For binary variables, the discrepancy measure more commonly used is the deviance. Denoting by $\eta^{(l)}$ the linear predictor value at the l th iteration of the MCMC, sample values from the posterior deviance were computed as

$$D^{(l)} = -2y^T \eta^{(l)} - 2 \sum_i \log(1 - \mu_i^{(l)})$$

The posterior deviance values were used to compute credible intervals to compare between models with different covariate structures. The use of posterior deviance samples allows comparison between non-nested models. Model selection started with a model with main effects plus first-order interactions, followed by deletion of nonsignificant terms one at a time. A term in the model was removed only if its removal did not produce a significant increase in the posterior deviance.

We assessed model adequacy by computing the Bayesian p values (p_b) for the predictive deviance using the algorithm described by Gelman et al. (1996). According to these authors, values of p_b higher than 0.1 indicate an acceptable fit. Significance tests for the model parameters were based on credible intervals obtained from the MCMC simulation. These credible intervals also allow us to make paired comparisons between the model parameters whenever relevant.

Extension of the model to the space-time domain

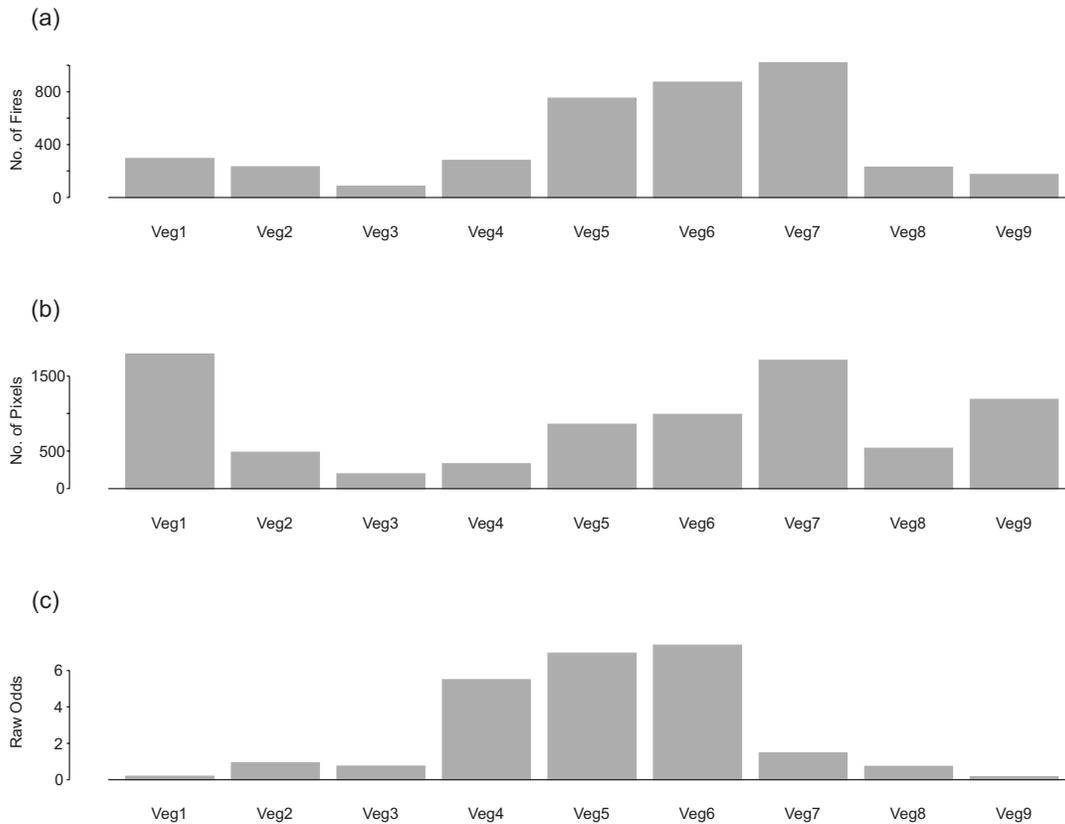
A model that is useful for short-term predictions of ignition probability requires extending eq. 3 to the space-time domain. The transversal analyses discussed in the previous section provide an overview of the dynamic nature of ignition probability in the Blue Mountains. To construct this space-time model, it is important to consider the possibility of time dependence of the estimates, because the effects of location and time on ignition probability do not seem to be independent. Another option is to consider these parameters as fixed in time and to accommodate the space-time interaction through the inclusion of a spatiotemporal term, ψ_{it} , in the model. This leads to a model of the form of eq. 1. To fit this model, we assume again a flat noninformative prior density for the covariate effects. We assume a Gaussian pairwise prior for the spatiotemporal effect, now in three dimensions. The conditional density of ψ_{it} given the rest of the ψ is proportional to

$$[11] \quad \lambda^{0.5} e^{-0.5 v_{it} \lambda (\psi_{it} - \bar{\psi}_{it})^2}$$

In this case $\bar{\psi}_{it}$ is a weighted average of the values of the spatiotemporal effect at neighbouring pixels in space and in time, and v_{it} is the number of those pixels. For seasonal data, a second option perhaps more adequate is to consider neighbours of order S over the time axis. We used this option with $S = 4$ to test the predictive performance of the model. For the descriptive part of the analysis, however, we used $S = 1$ to avoid the loss of the first four and the last four quarters. We again assume a $\Gamma(1, 1)$ prior density for the precision λ . The rest of the model structure is the same as the one described previously. The full conditionals in this case are

$$[12] \quad \pi(\beta, \mathbf{u} | \beta^-) \propto \left(\prod_{i=1}^N \prod_{t=1}^T \frac{e^{y_{it}(\mathbf{z}_i^T \beta + \mathbf{x}_i^T \mathbf{u})}}{1 + e^{\mathbf{z}_i^T \beta + \gamma R_{it} + \mathbf{x}_i^T \mathbf{u} + \psi_{it}}} \right)$$

Fig. 3. Presence of lightning-caused ignitions in the Blue Mountains: (a) proportion of ignitions by vegetation class; (b) proportion of area covered by vegetation class; and (c) empirical odds of ignition by vegetation class.



$$\begin{aligned}
 [13] \quad \pi(\gamma | \Upsilon) &\propto \left(\prod_{i=1}^N \prod_{t=1}^T \frac{e^{y_{it}(\gamma R_{it})}}{1 + e^{\mathbf{z}_i^T \boldsymbol{\beta} + \gamma R_{it} + \mathbf{x}_i^T \mathbf{u} + \psi_{it}}} \right) \\
 [14] \quad \pi(\Psi_{it} | \Psi_{it}^-) &\propto \left(\frac{e^{y_{it}(\Psi_{it})}}{1 + e^{\mathbf{z}_i^T \boldsymbol{\beta} + \gamma R_{it} + \mathbf{x}_i^T \mathbf{u} + \psi_{it}}} \right) e^{-0.5 v_i \lambda (\Psi_{it} - \bar{\Psi}_{it})^2} \\
 [15] \quad (\lambda | \lambda^-) &\sim \Gamma \left(a + 0.5NT, b + 0.5 \sum_{i=1}^N \sum_{t=1}^T v_{it} (\Psi_{it} - \bar{\Psi}_{it})^2 \right)
 \end{aligned}$$

The samples from these full conditional distributions were simulated in the same manner as for the nonspatial model.

An attractive feature of the Bayesian approach is the potential for predicting ignition probability directly from the MCMC algorithm. We now consider predictions over the next three quarters $T + 1$, $T + 2$, and $T + 3$ by updating all the parameters in the model for the system of $8089 \times T$ pixel-quarter combinations at each cycle of the MCMC procedure. At the end of the l th basic $N \times T$ MCMC cycle, we have $\boldsymbol{\beta}^{(l)}$, $\gamma^{(l)}$, and $\psi_{it}^{(l)}$ and compute the corresponding $\mu_{it}^{(l)}$ from $\xi_{it}^{(l)}$. Because the predictive distribution of the response in pixel i at times $T + j$ is Bernoulli (μ_{iT+j}) and because the effect of the covariates is constant over time, for $j = 1, 2, 3$ it suffices to generate a sample from the distribution $\psi_{iT+j}^{(l)} \sim N(\bar{\Psi}_{iT+1}^{(l)}, 1/\lambda^{(l)})$, where

$$[16] \quad \bar{\Psi}_{iT+j}^{(l)} = \frac{1}{v_i + 1} \left(2\psi_{iT+j-s}^{(l)} - \psi_{iT+j-2s}^{(l)} + \sum_{\{k \in \delta_i\}} \psi_{iT+j}^{(l)} \right)$$

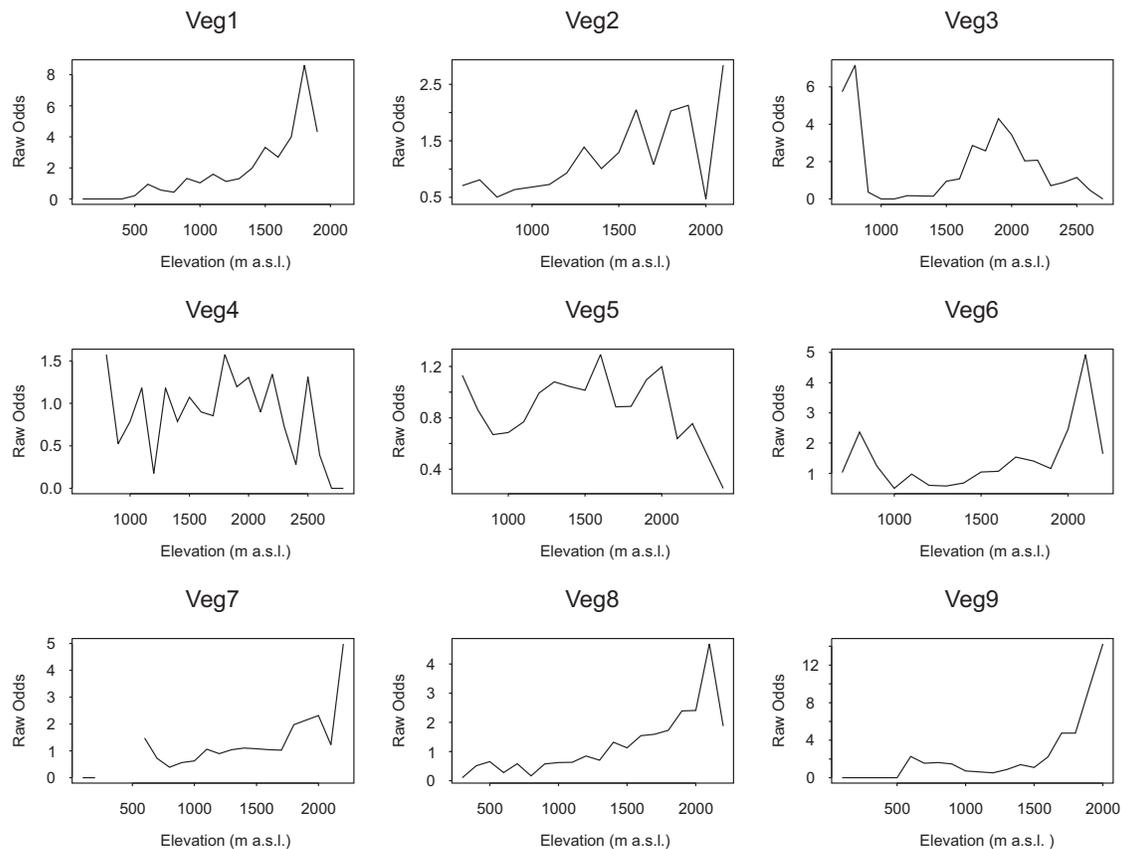
and from here compute the corresponding realization of the predicted probability of ignition using the inverse logit transformation.

Here, $\{k \in \delta_i\}$ denotes the set of pixels in the neighbourhood of pixel i , v_i is the number of those pixels and S is a seasonal lag with $S = 1$ if there is not seasonal effect. The predictive performance of our model was tested using the data of the first 27 quarterly periods as a training set and predicting the logarithm of odds (ξ_{it}) for $T = 28, 29$, and 30 , which correspond to the spring, summer, and fall of 1993. These predictions were compared with the values of ξ_{it} estimated using the complete data set.

Results and discussion

Figure 3 shows, for the study area and time period, the frequency of ignitions (Fig. 3a), the number of pixels at risk (Fig. 3b) and the empirical odds of ignition for each of the 9 vegetation classes (Fig. 3c). Because these empirical odds are computed from the proportion, number of ignitions/number of pixels, they represent an analogous quantity to the per-capita risk in the epidemiology framework. Classes 1, 2, 8, and 9 showed a relatively high frequency of ignitions, but the number of pixels at risk in those classes is also high, resulting in low values for the odds of fire. On the other hand, class 4 showed a relatively high number of ignitions spread in a comparative low number of pixels at risk, raising the value for the odds of ignition in this class. Classes 5 and 6 showed a similar relationship between number of ignitions and the number of pixels at risk. Despite showing the highest frequency of ignition occurrences, the number of pixels in class 7 is also high, decreasing the value of the empirical odds of ignition. Figure 4 shows the relationship between the empirical odds of ignition and the elevation for each veg-

Fig. 4. Empirical odds of ignition by elevation for the nine vegetation classes. Plots are based on a 30-bin partition of the elevation range obtained from a digital elevation model for the study area.



etation class. Except for classes 3, 4, and 5, the rest show an overall positive trend with elevation. This is, in part, the result of a low number of pixels at risk and perhaps the increased incidence of lightning strikes at high elevations.

The temporal variation of the empirical odds of ignition for the nine vegetation classes aggregated by quarterly periods is shown in Fig. 5. For all the classes it is evident the presence of a seasonal component, with high values during the summer and minima in winter. The first peak corresponds to the summer of 1986. The height of the peaks is somewhat stable along the time period except for 1986 and 1989. These were years of high incidence of lightning-caused ignitions in the Blue Mountains. Because inferences based on the empirical odds of ignition are potentially misleading, we will focus on the results of the statistical model fitted to the data.

The model selection procedure led in all the cases to models that included only the main effects for vegetation, elevation, slope, and precipitation plus the spatial term ψ . The p_b values were greater than 0.16 in all cases. It is surprising that terms related to aspect were absent in the final models, because this covariate is related to radiation received and, therefore, to fuel moisture. Exclusion of aspect from the final models may be a consequence of the low resolution for this variable imposed by the pixel size used in the analysis.

Analysis of risk factors

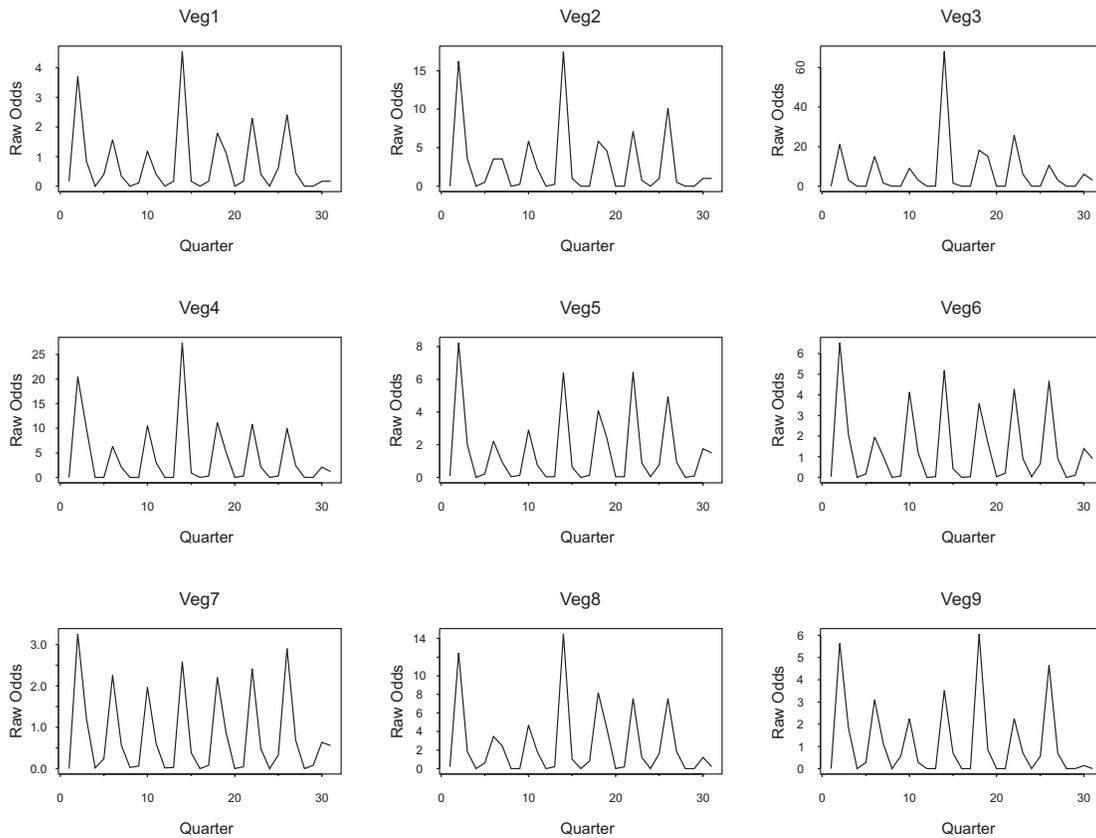
The time trend for the 95% posterior credible intervals for the different vegetation classes is shown in Fig. 6. None of

the plots shows a seasonal pattern, and the coefficient estimates mostly fluctuate around a particular mean level of each class. When significant, they imply that for a given pixel the effect of being covered by vegetation class V_k is to either increase or decrease that pixel's odds of ignition by a factor of e^{β_k} , provided we keep the other variables at the same level.

Table 2 shows the average over time of the factor by which each coefficient affects the odds of ignition for the different vegetation classes. For pixels with equivalent elevation, slope, and precipitation, the overall effect of classes 1, 7, and 9 is to reduce the odds of ignition to 33.0, 61.4, and 10.8%, respectively. For the remaining vegetation classes, their effect generally is to increase the odds of ignition as much as 955%. Class 4 has, by far, the highest risk of lightning ignitions.

Figure 3c shows that, before correcting for the other covariates, the ranking of the vegetation class in terms of odds of ignition is $\{9 < 1 < 3 < 8 < 2 < 7 < 4 < 5 < 6\}$. From the data in Table 2 it is possible to rank the vegetation classes in terms of their risk of ignition. In this case, the overall ranking of vegetation classes for ignition probability ($9 < 1 < 7 < 5 < 8 < 6 < 2 < 3 < 4$) was maintained through nearly all quarters. The different ranking obtained from Table 2 is due to the correction that the model fitted makes for the effect of the rest of the covariates as well as the spatial term in the model. Filtering this effect is important for cases where the data will be used in decision making processes and is a benefit obtained from fitting a model.

Fig. 5. Seasonal variation of the empirical odds, based on a 31-quarter partition of the time window. The first quarter corresponds to spring 1986.



Values in Table 2 are directly related to how likely a pixel is to be disturbed by fire and initiate a successional process. The resulting plant composition after disturbance will depend on the original plant composition as well as on the seed bank. The odds for a lightning-caused ignition are highest in pixels classified as lodgepole pine (*Pinus contorta* Dougl. ex Loud.), whitebark pine (*Pinus albicaulis* Engelm.), and Engelmann spruce (*Picea engelmannii* Parry ex Engelm.) – subalpine fir (*Abies lasiocarpa* (Hook.) Nutt.), which together comprise class 4. In these forest types, successful postfire germination produces forests with similar species composition (Keane et al. 1996b), so vegetation class of the pixels tends not to change much over time. For other vegetation types, fire disturbance may alter the dominant vegetation considerably; for example, grand fir (*Abies grandis* (Dougl.) Lindl.) (class 6) typically is dominant only if fire is excluded (Hall 1977). Our results show that pixels covered by this vegetation class have high odds of ignition, suggesting that the relative presence of these species in the Blue Mountains would decrease in the future if fire occurrence increases.

The 95% credible intervals for elevation, slope, and precipitation are shown in Fig. 7. The estimates are always negative for elevation and precipitation, indicating that a unit increase in those variables decreases the odds of ignition. These coefficients for elevation oscillate around a mean level of -2.0 , with the highest deviations from this value occurring during spring and fall 1988 and 1989. For the model-fitting process, elevations were standardized by creating the covariate d in eq. 10. Because the coefficients for elevation

are always negative, increases in elevation result in decreased odds of ignition. The highest values of this coefficient always occurred in the summer months, indicating that the probabilities of a lightning ignition during warm weather are minimally affected by moderate changes in elevation. The lowest values, with median values as low as -9.026 in 1989, generally are associated with spring. The highly negative values during this season are related to the presence of snow cover and higher precipitation at high elevations. There is no significant difference between the estimates for spring and fall.

The estimates for slope were mostly nonsignificant, and this term should have been removed most of the time by our model selection procedure. However, the slope effect was retained in the models to allow a comparison of the estimates over time, because they were significant on some occasions. The lowest estimates for this covariate also occurred in the spring quarter. The effect of precipitation oscillates around -0.04 , with the highest estimates during the summer and lowest estimates in fall and spring. The spatial term in the model was significant most of the time and showed a clear seasonal pattern over time and for fixed locations. The trajectories for a random sample of 10 pixels showed a seasonal pattern, but their peaks and valleys did not always coincide in time, suggesting the presence of an interaction between location and time.

Space–time model results

Table 3 shows the results of the model selection procedure for the space–time extended model. These results are consis-

Fig. 6. The 95% posterior credible intervals for the coefficients related to the different vegetation classes, as estimated from the longitudinal analysis. The white square inside each interval corresponds to the posterior median. The sequences SFS... denote summer, fall, spring, starting with summer 1986. Winter is not included.

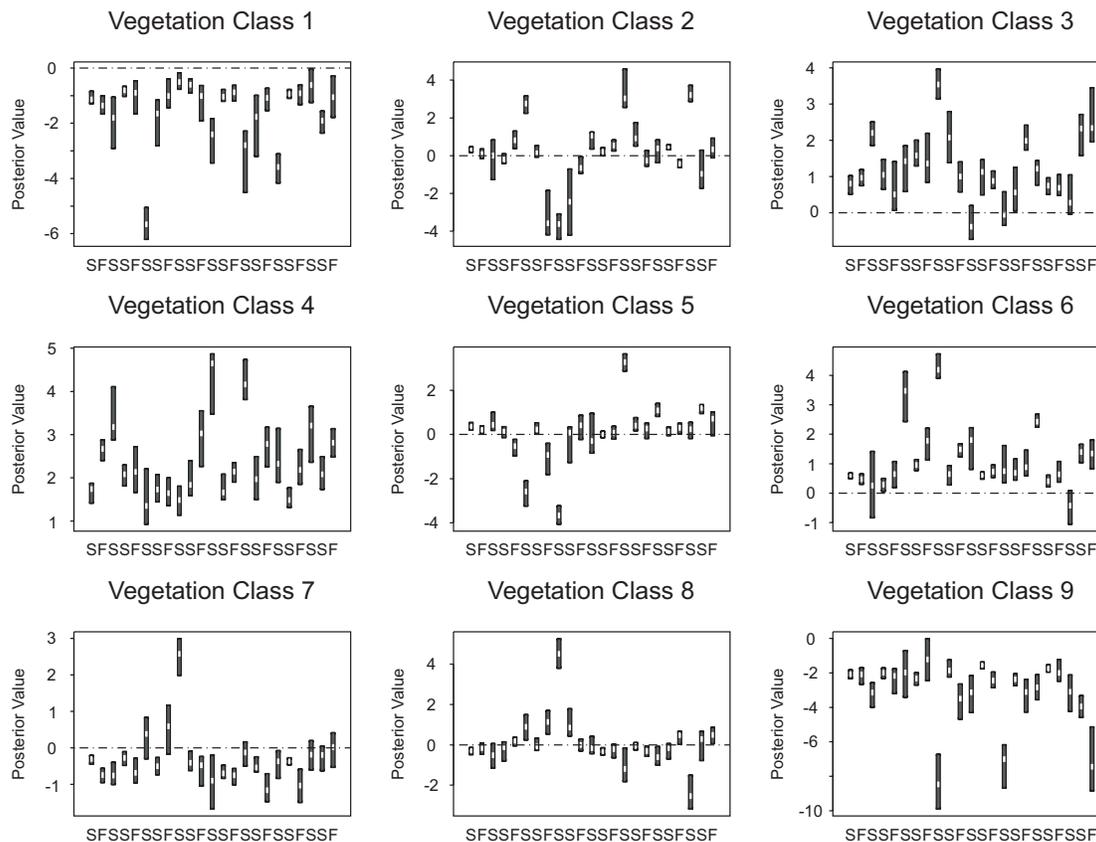


Table 2. Summary statistics for the median values of the exponential of the coefficients related to the vegetation classes over time.

Vegetation class	Minimum	First quantile	Median	Mean	Third quantile	Maximum
1	0.003	0.153	0.330	0.501	0.405	6.609
2	0.007	0.591	1.298	3.269	2.284	24.810
3	0.263	1.810	2.788	5.352	4.700	34.410
4	3.327	5.853	9.554	22.790	22.000	118.400
5	0.025	0.617	1.189	2.441	1.498	26.430
6	0.242	1.652	2.094	5.930	4.328	66.050
7	0.099	0.468	0.614	1.203	0.810	13.160
8	0.078	0.739	0.905	5.314	1.642	91.400
9	<0.001	0.034	0.108	0.203	0.184	6.389

tent with those of the transversal analysis in the sense that the selected model includes only the main effects plus the space–time term. The Bayesian *p* value for the full model was 0.184, which suggests that the model provides a reasonable description of the data (Gelman et al. 1996). Residual analysis (not shown here) does not indicate a systematic pattern, and although 1582 space–time locations have residuals that could be considered as outliers, they represent a small proportion (0.65%) of the total. The outliers correspond to pixels covered by vegetation classes 6 and 7 during summer but with small posterior probability of ignition, because their neighbours belong to low-probability classes. A high resid-

ual would be expected when a positive response occurs in a cell with a low estimate of μ_{it} (Agresti 1978), as is the case here.

The 95% posterior confidence intervals for the covariate coefficients are shown in Table 4. Using this spatiotemporal model, the ranking of vegetation classes in terms of ignition probability is $9 < 1 < 8 < 2 < 7 < 5 < 3 < 6 < 4$ (compared with $9 < 1 < 7 < 5 < 8 < 6 < 2 < 3 < 4$ in the transversal analysis and to $9 < 1 < 3 < 8 < 2 < 7 < 4 < 5 < 6$ with the raw data.). Because this space–time model assumes that the covariate effects are constant over time, the ignition-risk order holds for any time. Thus, for a given pixel, the term

Fig. 7. The 95% posterior credible intervals across time for the coefficients related to elevation, slope and precipitation, computed from the longitudinal analysis. The open square inside each interval corresponds to the posterior median. The sequences SFS... denote summer, fall, spring, starting with summer 1986. Winter is not included.

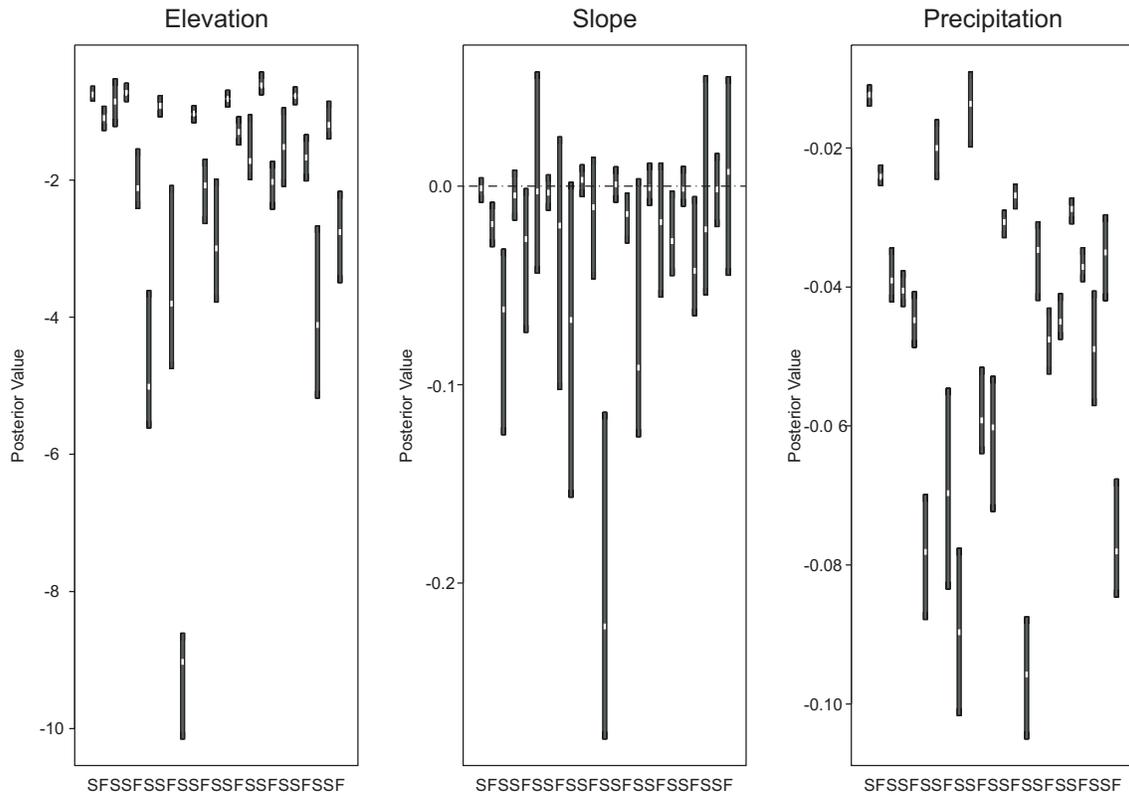


Table 3. Posterior quantiles for the deviance of some space–time models tested.

Model	5% quantile	Median	95% quantile	Bayesian p value
ψ	22 487	25 016	25 877	0.681
$V + E + S + \psi$	20 539	21 262	21 812	0.137
$V + E + R + \psi$	20 806	21 189	21 629	0.441
$V + S + R + \psi$	21 820	22 173	22 917	0.167
$E + S + R + \psi$	23 179	23 793	24 570	0.119
$V + E + S + R + \psi$	19 205	19 606	19 924	0.184

$e^{V_i+E_i+S_i}$ is the contribution to the odds of ignition by the ecological and environmental characteristics of the pixel. The term $e^{R_i+\psi_i}$ captures the contributions of local characteristics (e.g., precipitation) as well as other unobserved variables with spatiotemporal variation.

Distribution of ignition probability in space and time

As expected, the risk of ignition is highest during the summer months because of low fuel moisture and high capture of solar energy by ground fuels and live vegetation. The shading pattern in the posterior probability maps shows substantial interannual variation, except in summer quarters, in which the pattern is constant (Fig. 8). The posterior probability maps show the presence of areas with low risk throughout the year. These low risk areas are associated with vegetation classes 1, 7, and 9 and elevations below 1300 m,

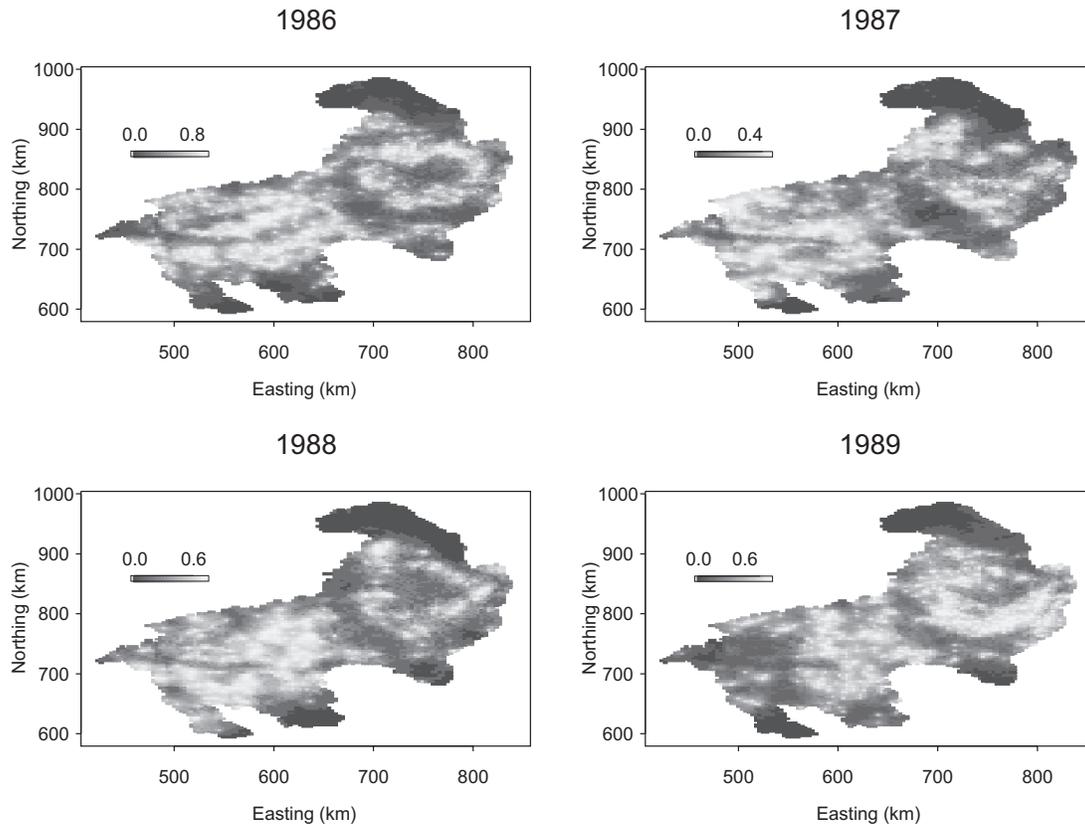
Table 4. Posterior quantiles for the fixed effects and the precision in the spatiotemporal model.

Parameter	5% quantile	Median	95% quantile
Class 1	-0.724	-0.586	-0.471
Class 2	-0.157	0.018	0.137
Class 3	0.093	0.310	0.590
Class 4	0.489	0.571	0.735
Class 5	0.211	0.290	0.377
Class 6	0.200	0.340	0.507
Class 7	-0.031	0.037	0.121
Class 8	-0.287	-0.157	0.009
Class 9	-7.106	-2.228	-2.280
Elevation	-3.663	-3.234	-2.650
Slope	-0.013	-0.006	0.000
Precipitation	-0.014	-0.011	-0.008
λ	0.109	0.119	0.121

encompassing ponderosa pine, shrublands, grasslands, and agricultural land.

The highest posterior probabilities of ignition occur in pixels with vegetation class 4 (lodgepole pine, whitebark pine, Engelmann spruce, subalpine fir) and at elevations in the range 1400–1800 m. Although the tree species found in this vegetation class range up to 2300 m, ignition probability was found to be higher at the lower elevational range of their distribution. During fall and spring, the probability of igni-

Fig. 8. Posterior mean estimates for the probability of ignition in the Blue Mountains for the summer months of 1986–1989 using the space–time model.



tion is in general higher in the southern Blue Mountains than in the northern part, most likely because of generally lower precipitation, lower snow-cover duration, and higher occurrence of lightning strikes in the south (Kridler et al. 1980; Maruoka 1994; Heyerdahl 1997).

The 95% posterior predictive median for ξ_{i28} , ξ_{i29} , and ξ_{i30} (plus signs) in a sample of four pixels, computed from eq. 16, are shown in Fig. 9 when seasonality is ignored ($S = 1$, plus signs) and when seasonality is considered ($S = 4$, solid circles). The asterisks in the plots are the posterior means for ξ_{it} ; $t < 28$, obtained using the complete data set, and the solid lines are the 95% credible bands for the complete data set. Predictions computed without seasonality induce a linear trend for ξ_{it} . The potential problem with this is that the predicted risk of ignition may be over- or underestimated for two- and three-step-ahead predictions. For the one-step-ahead predictions, the credible bands contain the predicted estimates in the four cases shown here, but the two- and three-step-ahead predictions are over- and underestimated in three of the four cases. This is a well-known problem when using second-order differences to make forecasts (Wei 1990). Including seasonality in the predictive distribution of ξ_{it} improves the quality of the predictions, in the sense that predictions calculated assuming seasonality retain the “up–down” pattern obtained with the full data set, and the credible bands now contain a higher number of the posterior predictive means for ξ_{it} .

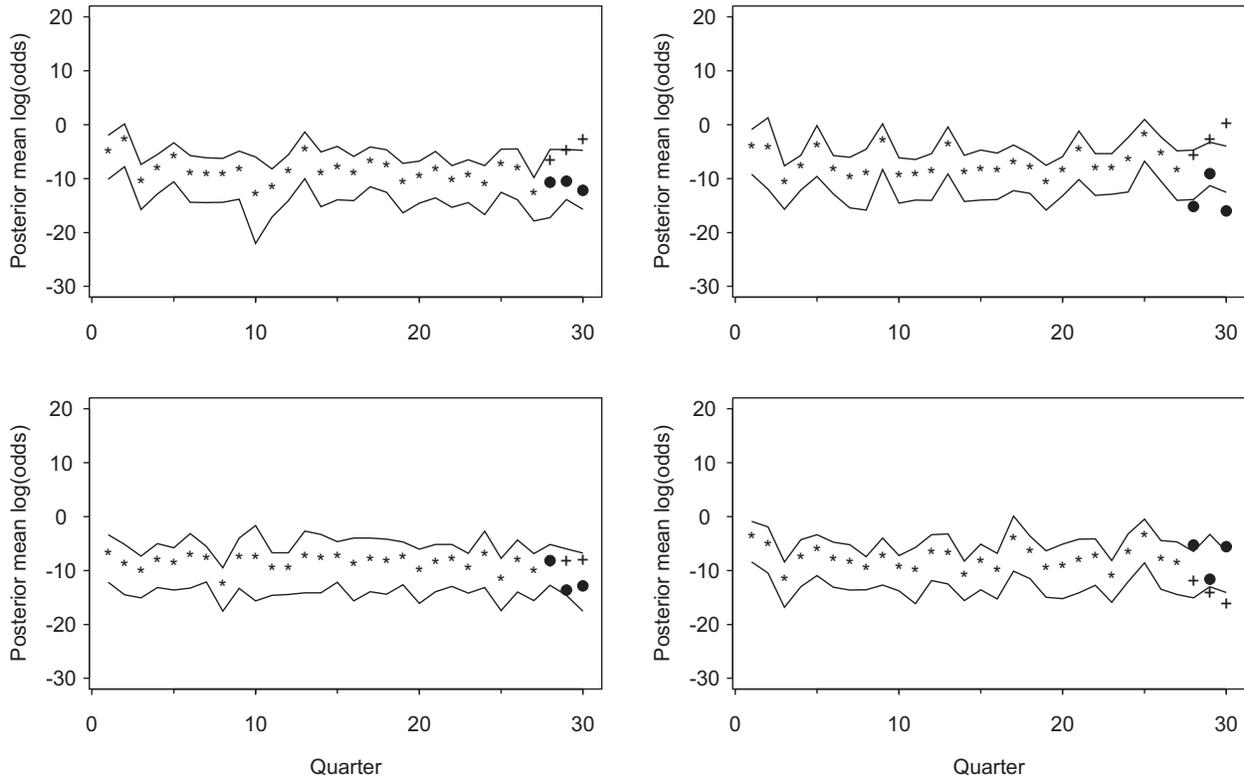
Although not shown here, the predictive posterior maps for 1993 resemble the map estimated for that year with a full

data set. Because we are predicting an unobserved variable, there is no direct way to test the quality of the predictions obtained with the model, and therefore, these predictions, like those obtained with other models, must be used with some caution.

Conclusions

The model we have proposed in this paper is of general applicability to spatiotemporal data for which one has $i = 1, 2, \dots, N$ locations in space and $t = 1, 2, \dots, T$ time intervals. The advantage of our modeling approach is that the shape of the pixels is irrelevant for the statistical estimation procedure. The only relevant information is ignition occurrence at that location in a fixed time interval and the neighbourhood structure of the data. Thus, models of the form of eq. 1 are still applicable for ignition data associated with irregularly shaped areas (e.g., counties, administrative units). As with any gridded data set, one should be aware that the risk maps obtained with models of the type proposed here may be sensitive to changes in interval size of pixels and time. This problem is similar to that of changing the number of classes in a histogram. However, because our interest is in mapping the risk of ignition and not the exact locations of ignitions, it is important to determine the spatiotemporal pattern, a primary objective of our model. Another point to consider is that, because of the inclusion of spatial covariates, alternative methods such as point processes produce maps of equivalent resolution because of the dependence of the intensity function on the covariate scale.

Fig. 9. Posterior predictive 95% credible bands for quarters 28, 29, and 30 using quarters 1–27 as a training set for a sample of four pixels. Predictions made with $S = 1$ (no seasonality) are shown as plus signs. Predictions made with seasonality ($S = 4$) are shown as dots. Asterisks are the posterior mean estimate obtained with the training data set.



Our analysis of ignition-occurrence data for the Blue Mountains provides important quantitative information on how several covariates, or risk factors, affect the probability of ignition for a given location and time. Despite the high spatiotemporal variability in ignition probability, the changes in the posterior deviance show that the inclusion of covariates accounts for a significant amount of this variability, and the inclusion of a spatiotemporal correlated term improves the model fit. The assumed presence of unobserved variables is not uncommon, but previous applications of logistic regression in forest-ignition modelling (e.g., McKelvey and Busse 1996) have ignored the possibility of a spatially correlated term in the model. Their approach was to take a random sample of pixels to meet the independence assumptions for the response, but this approach loses all the information contained in the neighbouring structure of the data. Whenever the presence of spatial and temporal correlation is suspected, it is generally advisable to include a spatial term in the model. For the vegetation classes analysed, the different odds of ignition reflect their differential susceptibility to lightning strikes due to fuel characteristics and other factors that affect a fire start. The absence of interactions between the covariates in the models implies that the ranking obtained for the different vegetation classes is consistent across a range of different elevations, slopes, and precipitation and is independent of location in the Blue Mountains.

The probability of ignition in the Blue Mountains has considerable spatial and temporal variability. Spatial variability is related to the distribution of different vegetation types and to topographic patterns as well as other unobserved variables

not included in the model. The highest probabilities of ignition after correcting for elevation, slope, and precipitation are associated with pixels covered by vegetation classes 4 (lodgepole pine, whitebark pine, Engelmann spruce, subalpine fir), 3 (subalpine herbaceous, alpine tundra), and 6 (grand fir). Although vegetation class 4 is the most flammable, it covers high elevation and precipitation zones, resulting in moderately high raw odds of ignition. Because raw odds do not consider the effect of the covariates, vegetation 4 shows the highest fire probability after the corrections induced by the model. Probability of ignition decreases with increasing elevation and slope, despite the fact that the number of lightning strikes increases with elevation. This is probably a consequence of higher precipitation and lower temperatures at high elevations. In the time domain, variation in ignition probability is related partly to variation in precipitation. Another portion of the spatiotemporal variation is explained by the spatial term included in the model. This effect may be related to unobserved variables (e.g., temperature, fuel quantity, etc.), so it is difficult to infer a specific biophysical cause.

The second-difference prior assumed for the predictive distribution of the space–time component in the model has some potential disadvantages, but as we showed, it can be replaced by other constructions such as seasonal second differences. The resulting predictions captured the seasonal pattern observed with the estimates from the full data set and, in general, are useful for short-term forecasts. A potential improvement would be to replace the second-difference structure by a structure with random coefficients; these coef-

ficients could be considered random variables with a prior distribution of their own.

The predictive potential of the space–time model presented here has applications for forest managers who need information on the distribution of ignition occurrence and associated fire effects (Boyчук et al. 1997; Lertzman et al. 1998). The resulting ignition probabilities can be linked with other models that require such probabilities as an input (Mills and Bratten 1982; Keane et al. 1996a; Lenihan et al. 1998). A potential drawback of the model is the amount of computational time using the MCMC method to estimate parameters. However, a one-time analysis of a large data set can provide the basis for incorporating fire disturbance in long-term management for large geographic areas, thereby making it a good investment of time and effort.

Acknowledgments

We thank Donald McKenzie, Eduardo Gutierrez Peña, and the anonymous reviewers for their thorough review of the statistical analysis and other portions of the manuscript. The first author was supported in part by the USDA Forest Service, Pacific Northwest Research Station, and by the Dirección General de Asuntos del Personal Académico of the National University of México.

References

- Agee, J.K. 1990. The historical role of fire in Pacific Northwest Forests. *In* Natural and prescribed fire in Pacific Northwest forests. *Edited by* J.D. Walstad, S.R. Radosevich, and D.V. Sandberg. Oregon State University Press, Corvallis. pp. 25–38.
- Agee, J.K. 1993. Fire ecology of Pacific Northwest forests. Island Press, Covelo, Calif.
- Agee, J.K. 1994. Fire and weather disturbances in terrestrial ecosystems of the eastern Cascades. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-320.
- Agee, J.K. 1998. The landscape ecology of Western forest fire regimes. *Northwest Sci.* **72**(Spec. Issue): 4–23.
- Agee, J.K., Finney, M., and deGouvenain, R. 1990. Forest fire history of Desolation Peak, Washington. *Can. J. For. Res.* **20**: 350–356.
- Agresti, A. 1978. *Categorical data analysis*. Wiley, New York.
- Alvarado, E., Sandberg, D.V., and Pickford, S.G. 1998. Modeling large forest fires as extreme events. *Northwest Sci.* **72**(Spec. Issue): 66–75.
- Andrews, P.L. 1986. BEHAVE: fire behavior prediction and fuel modeling system—BURN subsystem, part 1. USDA For. Serv. Gen. Tech. Rep. INT-194.
- Bernardo, J.M., and Smith, A.F.M. 1994. *Bayesian theory*. Wiley, New York.
- Besag, J., York, J., and Mollie, A. 1991. Bayesian image restoration, with two applications in spatial statistics (with discussion). *Ann. Inst. Stat. Math.* **43**: 1–59.
- Besag, J., Green, P., Higdon, D., and Mengersen, K. 1995. Bayesian computation and stochastic systems (with discussion). *Stat. Sci.* **10**: 3–66.
- Bessie, W.C., and Johnson, E.A. 1995. The relative importance of fuels and weather on fire behavior on subalpine forests. *Ecology*, **76**: 747–762.
- Box, G.E.P., and Tiao, G.C. 1973. *Bayesian inference in statistical analysis*. Addison-Wesley, London.
- Boyчук, D., Perera, A.H., Ter-Mikaelian, M.T., Martell, D.L., and Li, C. 1997. Modelling the effect of spatial scale and correlated fire disturbances on forest age distribution. *Ecol. Modell.* **95**: 145–164.
- Bratten, F.W. 1984. Fire occurrence probabilities in the northern Rocky Mountain – Intermountain zone: an estimation technique. USDA For. Serv. Res. Note PSW-366.
- Breslow, N.E., and Clayton, D.G. 1993. Approximate inference in generalised linear mixed models. *J. Am. Stat. Assoc.* **88**: 9–25.
- Clarke, S.E., and Bryce, S.A. (*Editors*). 1997. Hierarchical subdivisions of the Columbia Plateau and Blue Mountains ecoregions, Oregon and Washington. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-395.
- Clayton, D. 1997. Generalised linear mixed models in statistics. *Statistician*, **41**: 327–328.
- Clayton, D., and Kaldor, J. 1987. Empirical Bayes estimates of age-standardised relative risks for use in disease mapping. *Biometrics*, **43**: 671–682.
- Cressie, N.A. 1993. *Statistics for spatial data*. Wiley & Sons, New York.
- Cunningham, A.A., and Martell, D.L. 1973. A stochastic model for the occurrence of man-caused forest fires. *Can. J. For. Res.* **3**: 282–287.
- Daly, C., Neilson, R.P., and Phillips, D. 1994. A statistical-topographical model for mapping climatological precipitation over mountainous terrain. *J. Appl. Clim.* **33**: 140–158.
- DeLong, S.C. 1998. Natural disturbance rate and patch size distribution of forests in northern British Columbia: implications for forest management. *Northwest Sci.* **72**(Spec. Issue): 35–48.
- Diggle, P.J. 1983. *Statistical analysis of spatial point patterns (mathematics in biology)*. Academic Press, New York.
- Gelman, A., and Rubin, D.B. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**: 457–472.
- Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 1995. *Bayesian data analysis*. Chapman & Hall, New York.
- Gelman, A., Meng, X., and Stern, H. 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* **6**: 733–807.
- Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (*Editors*). 1996. *Markov chain Monte Carlo in practice*. Chapman & Hall, New York.
- Hall, F. 1977. Ecology of natural underburning in the Blue Mountains of Oregon. USDA For. Serv. Pac. Northwest Reg. Tech. Rep. R6-ECOL-79-001.
- Hastings, W.K. 1970. Monte Carlo sampling using Markov chains and their applications. *Biometrika*, **57**: 97–109.
- Hessburg, P.F., Mitchell, R.G., and Filip, G.M. 1994. Historical and current role of insects and pathogens in eastern Oregon and Washington forested landscapes. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-327.
- Heyerdahl, E.K. 1997. Spatial and temporal variation in historical fire regimes of the Blue Mountains, Oregon and Washington. Ph.D. dissertation, University of Washington, Seattle.
- Hosmer, D.W., and Lemeshow, S. 1989. *Applied logistic regression*. Wiley & Sons, New York.
- Ibrahim, J.G., and Laud, P.W. 1991. On Bayesian analysis of generalized linear models using Jeffrey's prior. *J. Am. Stat. Assoc.* **86**: 981–985.
- Johnson, C.G. 1994. Forest health in the Blue Mountains: an ecologist's perspective on ecosystem process and biological diversity. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-339.
- Johnson, E.A., and Gutsell, S.A. 1994. Fire frequency models, methods, and interpretations. *Adv. Ecol. Res.* **25**. pp. 239–287.
- Johnson, E.A., and Wowchuk, D.R. 1993. Wildfires in the southern Canadian Rocky Mountains and their relationships to mid-tropospheric anomalies. *Can. J. For. Res.* **23**: 1213–1222.

- Karr, A.F. 1986. Point processes and their statistical inference. Marcel Decker Inc., New York.
- Keane, R.E., Ryan, K.C., and Running, S.W. 1996a. Simulating effects of fire on northern Rocky Mountain landscapes with the ecological process model FIRE-BGC. *Tree Physiol.* **16**: 319–331.
- Keane, R.E., Long, D.G., Menakis, J.P., Hann, W.J., and Bevins, C.D. 1996b. Simulating coarse-scale vegetation dynamics using the Columbia River Basin Succession Model—CRBSUM. USDA For. Serv. Gen. Tech. Rep. INT-GTR-340.
- Kourtz, P., and Todd, B. 1992. Predicting the daily occurrence of lightning-caused forest fires. *For. Can. Petawawa Nat. For. Inst. Info. Rep.* PI-X-112.
- Krider, E.P., Noggle, R.C., Pifer, A.E., and Vance, D.L. 1980. Lightning direction finding systems for forest fire detection. *Bull. Am. Meteorol. Soc.* **61**: 980–986.
- Lenihan, J.M., Daly, C., Bachelet, D., and Neilson, R.P. 1998. Simulating broad-scale fire severity in a dynamic global vegetation model. *Northwest Sci.* **72**(Spec. issue): 91–103.
- Lertzman, K., and Fall, J. 1998. From forest stands to landscapes: spatial scales and the roles of disturbances. *In Ecological scale: theory and applications. Edited by D.L. Peterson and V.T. Parker.* Columbia University Press, New York. pp. 339–367.
- Lertzman, K., Fall, J., and Dorner, B. 1998. Three kinds of heterogeneity in fire regimes: at the crossroads of fire history and landscape ecology. *Northwest Sci.* **72**(Spec. Issue): 4–23.
- Maruoka, K.R. 1994. Fire history of *Pseudotsuga menziesii* and *Abies grandis* stands in the Blue Mountains of Oregon and Washington. Master's thesis, University of Washington, Seattle.
- McCullagh, P., and Nelder, J.A. 1989. Generalised linear models. Chapman & Hall, New York.
- McKelvey, K.S., and Busse, K.K. 1996. Twentieth century fire patterns on Forest Service lands. *In Sierra Nevada ecosystem project: final report to Congress.* Centers for Water and Wildland Resources, University of California, Davis. pp. 1119–1153.
- McKenzie, D. 1998. Fire, vegetation, and scale: toward optimal models for the Pacific Northwest. *Northwest Sci.* **72**(Spec. Issue): 49–65.
- McKenzie, D., and Halpern, C.B. 1999. Modeling the distributions of shrub species in Pacific Northwest forests. *For. Ecol. Manag.* **114**: 293–307.
- McKenzie, D., Peterson, D.L. and Agee, J.K. 2000. Fire frequency in the Columbia River basin: building regional models from fire history data. *Ecol. Appl.* **10**: 1497–1516.
- Mees, R.M. 1978. Computing arrival times of firefighting resources for initial attack. USDA For. Serv. Gen. Tech. Rep. PSW-27.
- Mills, T.J., and Bratten, F.W. 1982. FEES: design of a fire economics evaluation system. USDA For. Serv. Gen. Tech. Rep. PSW-65.
- Mutch, R.W., Arno, S.F., Brown, J.K., Carlson, C.E., Ottmar, R.D., and Peterson, J.L. 1993. Forest health in the Blue Mountains: a fire management strategy for fire adapted ecosystems. USDA Forest Service Gen. Tech. Rep. PNW-310.
- Nelder, J.A., and Weddeburn, R.W.M. 1972. Generalised linear models. *J. R. Stat. Soc. Br.* **135**: 370–384.
- Ohmann, J.L., and Spies, T.A. 1998. Regional gradient analysis and spatial pattern of woody plant communities of Oregon forests. *Ecol. Monogr.* **68**: 151–182.
- Peterson, D.L. 1998. Large-scale fire disturbance: from concepts to models. *Northwest Sci.* **72**(Spec. Issue): 1–3.
- Quigley, T.M., Haynes, R.W., and Graham, R.T. (Editors). 1996. An integrated scientific assessment for ecosystem management in the interior Columbia Basin and portions of the Klamath and Great basins. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-382.
- Quinby, P.A. 1987. An index to fire incidence. *Can J. For. Res.* **17**: 731–734.
- Reed, W.J. 1994. Estimating the historic probability of stand-replacement fire using the age-class distribution of undisturbed forest. *For. Sci.* **40**: 104–119.
- Rogers, P. 1996. Disturbance ecology and forest management: a review of the literature. USDA For. Serv. Gen. Tech. Rep. INT-GTR-336.
- Rothermel, R.C. 1972. A mathematical model for predicting fire spread in wildland fuels. USDA For. Serv. Res. Pap. INT-115.
- Schmoldt, D.L., Peterson, D.L., Keane, R.E., Lenihan, J.M., McKenzie, D.M., Weise, D.R., and Sandberg, D.V. 1999. Assessing the effects of fire disturbance on ecosystems: a scientific agenda for research and management. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-455.
- Swetnam, T.W. 1993. Fire history and climate change in giant sequoia groves. *Science (Washington, D.C.)*, **262**: 885–889.
- Wei, W.W.S. 1990. Time series analysis: univariate and multivariate methods. Addison-Wesley, London.
- Wolpert, R.L., and Iktadt, K. 1998. Poisson/gamma random field models for spatial statistics. *Biometrika*, **85**: 251–267.
- Zellner, A., and Rossi, P.E. 1984. Bayesian analysis of dichotomous quantal response models. *J. Econometrics*, **25**: 365–393.