Spatial Patterns of Modern Period Human-Caused Fire Occurrence in the Missouri Ozark Highlands

Jian Yang, Hong S. He, Stephen R. Shifley, and Eric J. Gustafson

Abstract: The spatial pattern of forest fire locations is important in the study of the dynamics of fire disturbance. In this article we used a spatial point process modeling approach to quantitatively study the effects of land cover, topography, roads, municipalities, ownership, and population density on fire occurrence reported between 1970 and 2002 in the Missouri Ozark Highland forests, where more than 90% of fires are human-caused. We used the AIC (Akaike information criterion) method to select an appropriate inhomogeneous Poisson process model to best fit to the data. The fitted model was diagnosed using residual analysis as well. Our results showed that fire locations were spatially clustered, and high fire occurrence probability was found in areas that (1) were public land, (2) within 6 km to 17 km of municipalities, and (3) <500 m from roads where forests are accessible to humans. In addition, fire occurrence probability was higher in pine-oak forests on moderate (<25 degree) slopes and xeric aspects and at higher (>270 m) elevations, reflecting the effects of natural factors on fire occurrence. The results serve as a provisional hypothesis for expanding fire risk estimation to surrounding areas. The spatial scale of analysis (approximately 1 ha) provides new information to guide planning and risk reduction efforts. FOR. SCI. 53(1):1–15.

Keywords: fire occurrence, human-caused fires, spatial point pattern, inhomogeneous Poisson process, wildfire

THE SPATIAL PATTERN of fire occurrence is a major focus in the study of the dynamics of forest fire and its role in determining landscape structure and vegetation community composition (Turner and Romme 1994, Guyette et al. 2002, Ryan 2002, Larjavaara et al. 2005, Mermoz et al. 2005). Although there is an inherent stochasticity in fire occurrence, there are many biotic and abiotic factors that affect where forest fires occur. Many studies have found that fuel characteristics (type, loading, moisture, and inflammability) and topography (elevation, slope, and aspect) are prominent factors in shaping spatial patterns of natural-caused fires (e.g., Kushla and Ripple 1997, Diaz-Avalos et al. 2001, Latham and Williams 2001, Wotton and Martell 2005). However, the effects of these factors on fire occurrences can vary among ecosystems and across spatial and temporal scales. For example, Tanskanen et al. (2005) showed that differences in the moisture regime of surface fuels between stands of different age classes dominated by *Picea abies* (L.) Karst. or *Pinus sylvestris* L. can result in significantly different ignition conditions in boreal forests of southern Finland. Contrarily, Moritz et al. (2004) analyzed several hundred wildfires over a broad expanse of California shrub lands, and their results revealed that there was generally not a strong relationship between fuel age and fire probabilities in their study area. These apparently contradictory results suggest that fire occurrence is a complex process in which fuel characteristics and abiotic factors cannot provide the full explanation.

Fire ecologists and historians have found that fire regimes of many forest ecosystems are anthropogenic and shaped largely by human settlement and management (e.g., Veblen et al. 1999, Guyette et al. 2002, Bergeron et al. 2004, Hessburg et al. 2005). They have identified, mainly using dendrochronology, temporal stages (e.g., pre-European-settlement and modern eras) of fire regimes related to human population density levels and shifting cultural behaviors. There are usually more human-caused fires than natural-caused fires in modern fire regimes. For example, lightning-caused fires contribute only 35% of the fires reported in the boreal forest of Canada (Weber and Stocks 1998), and less than 1% of the fires reported in Missouri between 1970 and 1989 (Westin 1992).

Despite of the prevalence of human-caused fires in modern fire regimes, only a few researchers have sought to assess human factors influencing wildfires (e.g., Donoghue and Main 1985, Cardille et al. 2001, Prestemon and Butry 2005). Humans directly influence fire regimes both by igniting fires and by suppressing the spread of those fires (Sturtevant et al. 2004). Studies have shown that socioeconomic variables reflecting land use, fuel management, law enforcement, and human access are significant predictors of human-caused wildfires, and wildfires caused by different reasons respond differently to these

Jian Yang, School of Natural Resources, University of Missouri–Columbia, 203 ABNR Bldg., Columbia, MO 65211—Phone: (573) 489-8707; Fax: (573) 882-1977. Hong S. He, School of Natural Resources, University of Missouri–Columbia, 203 ABNR Bldg., Columbia, MO 65211—Phone: (573) 489-8707; Fax: (573) 882-1977. Stephen R. Shifley, USDA Forest Service, North Central Research Station, 202 ABNR Bldg., University of Missouri–Columbia, Columbia, MO 65211—Phone: (573) 875-5341 ext. 232. Eric J. Gustafson, USDA Forest Service, North Central Research Station, 5985 Highway K, Rhinelander, WI 54501-9128—Phone: (715) 362-1152; Fax: (715) 362-1166.

Acknowledgments: The authors gratefully acknowledge Dr David Larsen, Dr Richard Guyette, and Dr Tim Wright for providing comments on an earlier version of this article. We thank two anonymous reviewers and an associate editor for helpful comments throughout the review process. We are grateful to Lyn Carpenter and Margaret Olson of the Mark Twain National Forest for providing us their compiled fire occurrence data.

Manuscript received July 17, 2006, accepted August 21, 2006

factors (Prestemon et al. 2002, Mercer and Prestemon 2005). Most of these studies focused on the temporal pattern of daily fire ignition counts at a broad spatial scale. For example, Butry and Prestemon (2005) evaluated the spatio-temporal pattern of wildland arson ignition counts at a daily temporal scale and a spatial scale related to US census tracts. These broad spatial scale analyses hold promise for aggregate fire risk assessment, but modeling at finer scales is also needed to improve our understanding of spatial variation in human-caused fires (Prestemon et al. 2002).

The growing body of contemporary data about where, when, and how individual fires occur provides new opportunities to study the spatial pattern of fire occurrence in landscapes at spatial scales ranging from square kilometers to hectares in resolution. Spatial point pattern (SPP) data analysis can be useful in modeling the spatial pattern of fire ignition locations. SPP data are a set of event locations, often referred to as a point pattern. Point pattern and marked point pattern analysis methods have been previously used in forestry, particularly for describing the variability of forest stands in which the "points" are tree locations and the "marks" are tree characteristics such as tree species and dbh (e.g., Larsen and Bliss 1998, Stoyan and Penttinen 2000, Kokkila et al. 2002). Few researchers (e.g., Podur et al. 2003, Genton et al. 2006) have applied SPP in characterizing the spatial clustering pattern of forest fire occurrences by using the *K* function, *L* function, and kernel intensity estimation. The nonparametric statistics they used were exploratory and offered limited power for spatially specific inferences. However, recent theoretical development within the realm of SPP, such as the provision of formal likelihood-based methods of inference for a wide range of models, provides tools for statistically rigorous modeling of spatial patterns of fire occurrence.

In this study we used both nonparametric and parametric statistical analysis to describe and model spatial point patterns of fire occurrences over a long modern period for a landscape in the Missouri Ozark Highlands. We considered both human-caused and lighting-caused wildfires reported between 1970 and 2002. The primary objectives of our study were to (1) characterize spatial point patterns exhibited in the fire occurrence data, (2) model the human-caused fire occurrence process with the consideration of both environmental heterogeneity (including human effects) and interactions with neighboring fire occurrences (if there are any), (3) estimate the spatial distribution of fire occurrence probability, and (4) determine how specific spatial covariates contribute to the probability of fire occurrence.

Study Area and Fire Occurrence Data

The study area, located in the Missouri Ozark Highlands, is about 1,287 km^2 in size (Figure 1). The geographical bounding coordinates are: 91.5°W, 91.1°W, 37.0°N, and 36.7°N. A large proportion (86%) of the study area is forested, and 71% of the area is in the Mark Twain National Forest. White oak (*Quercus alba* L.), post oak (*Quercus stellata* Wangenh.), black oak (*Quercus velutina* Lam.), and shortleaf pine (*Pinus echinata* Mill.) are the dominant tree species. Topographic variation is high, with elevations averaging 260 m and ranging from 140 m to 350 m. The Eleven Point River crosses the study area, and the landscape is dissected by a dense stream network creating ridges and valleys, some with steep slopes. There were 1,299 fires reported to either the Mark Twain National Forest or the Missouri Department of Conservation (MDC) in the study area between 1970 and 2002. Only 16 (1.2%) of the fires were caused by lightning, 970 (74.7%) fires were caused by arson (including grudge fire and pyromania), and the others were accidental fires caused by escaped campfires, railroads, smoking, debris-burning, or for unspecified (unknown) reasons. The annual temporal variations of fire counts across different ignition sources are shown in Figure 2. The years 1976 and 1986–88 were the worst fire years and strongly correlated with dry climate (Patrick Guinan, Missouri State Climatologist, personal communication, Apr. 24, 2006). For our research objectives, we studied only aggregated spatial patterns over the three-decade period.

Methods

The overall modeling approach is shown in Figure 3. Reported fire ignition (origin) locations, which were geographically referenced using the UTM (Universal Transverse Mercator) coordinate system, and spatial covariates such as roadways, land cover, and topography were mapped and processed using a GIS. Nonparametric spatial statistics such as kernel intensity estimation and Ripley's *K* function were calculated to characterize spatial patterns (e.g., clustering or regularity) within the fire occurrence data. The fire occurrence process was then modeled as an inhomogeneous Poisson process. That is, we assumed there were no interactions with neighboring fire occurrences (i.e., having a fire in one location did not affect the chance of fire occurrence in the neighboring area), and spatial patterns of fire occurrence resulted solely from the effects of environmental heterogeneity in the landscape. We considered a wide range of inhomogeneous Poisson models that include different sets of spatial covariates (with transformation) and used residual analysis and AIC methods to select a best model to fit the data. A maximum pseudolikelihood method was used to estimate coefficients of each candidate model. Finally, the assumption of an inhomogeneous Poisson process was tested by calculating the generalized *K* function for the best-fit model.

Guiding Hypotheses

Because most fires in this area are human-caused and spatial distribution of roadways determines the human access to forests, we hypothesized that locations close to roads should have a higher probability of human-caused fires. We obtained the digital roadway coverage published by the Missouri Department of Transportation (MoDOT) in 2004.

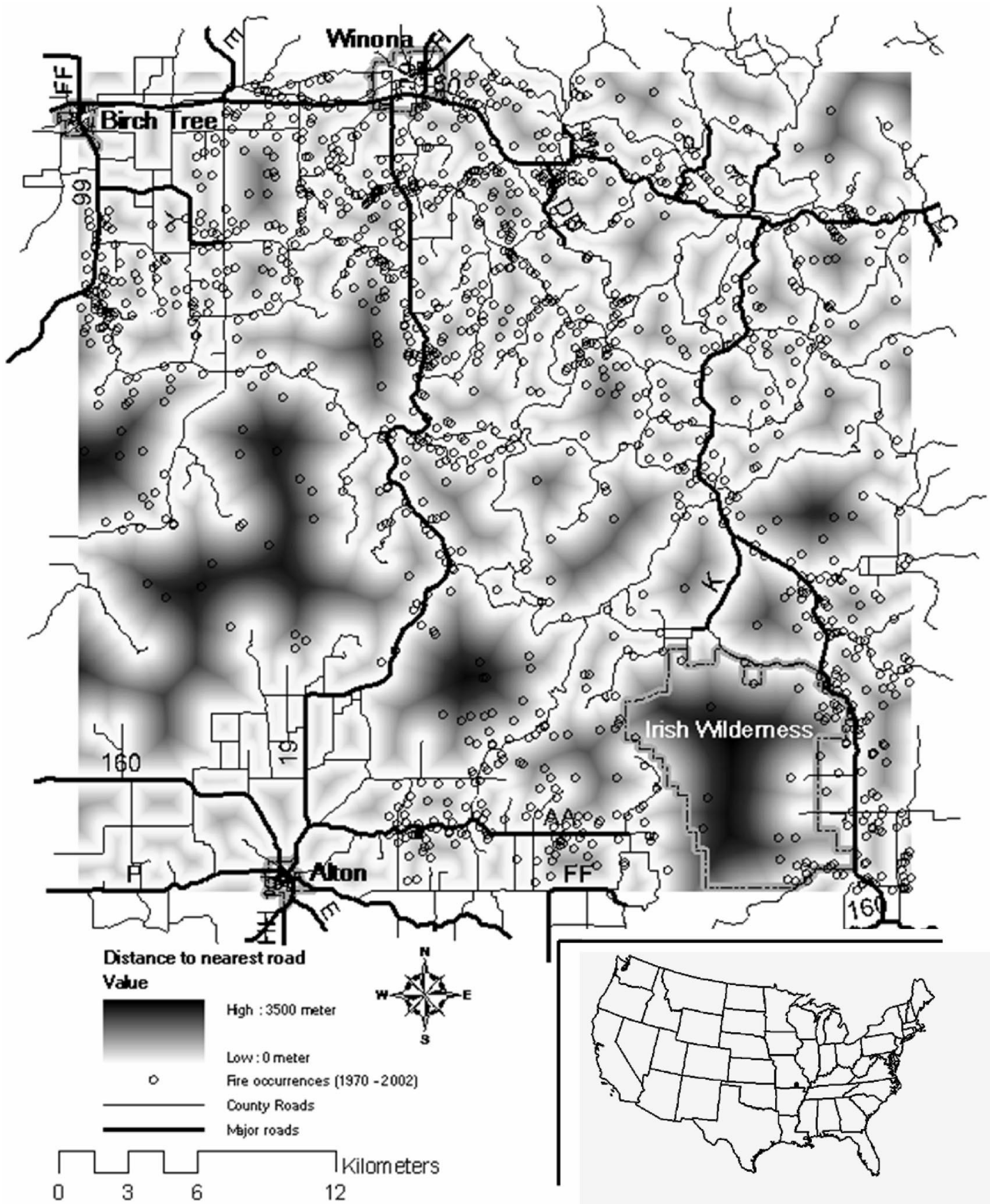

Figure 1. Study area with reported fire locations, roadway coverage, and proximity to roads. Major roads include US highways 60 and 160, numbered state highways 19 and 99, and lettered county highways (e.g., route K). The three towns (Alton, Birch Tree, and Winona) are located in the boundary of the study area. The Irish wilderness is located in the southeast of the study area.

The data layer contains all US highways, state highways (Missouri numbered routes and Missouri lettered routes), county roads, and city streets. The spatial accuracy is ± 3 m. We calculated the proximity to roads as a measure of human accessibility. This was done by calculating the Euclidean distance from each cell (30-m resolution) to the nearest road (Figure 1), a function provided in the ArcGIS Spatial Analyst tool.

The area consists of three different ownership groups: (1) the land outside the MTNF (Mark Twain National Forest) purchase unit boundary (14%), (2) private land within the MTNF purchase unit (25%), and (3) public land within the MTNF purchase unit (61%). We believed that ownership would affect firefighting and arson behaviors and influence the spatial pattern of fire locations. The polygon coverage of ownership data for this region published by MSDIS (Missouri Spatial Data Information Service) was converted into a grid (60-m resolution) and used as a categorical spatial covariate.

Our study area is one of the most rural places in Missouri. The mean human population density calculated from 2000 census block-level data is about 4.0/km^2, the median is only about 1.5/km^2, and the 95th percentile is still less than 15/km^2. Although the population density in

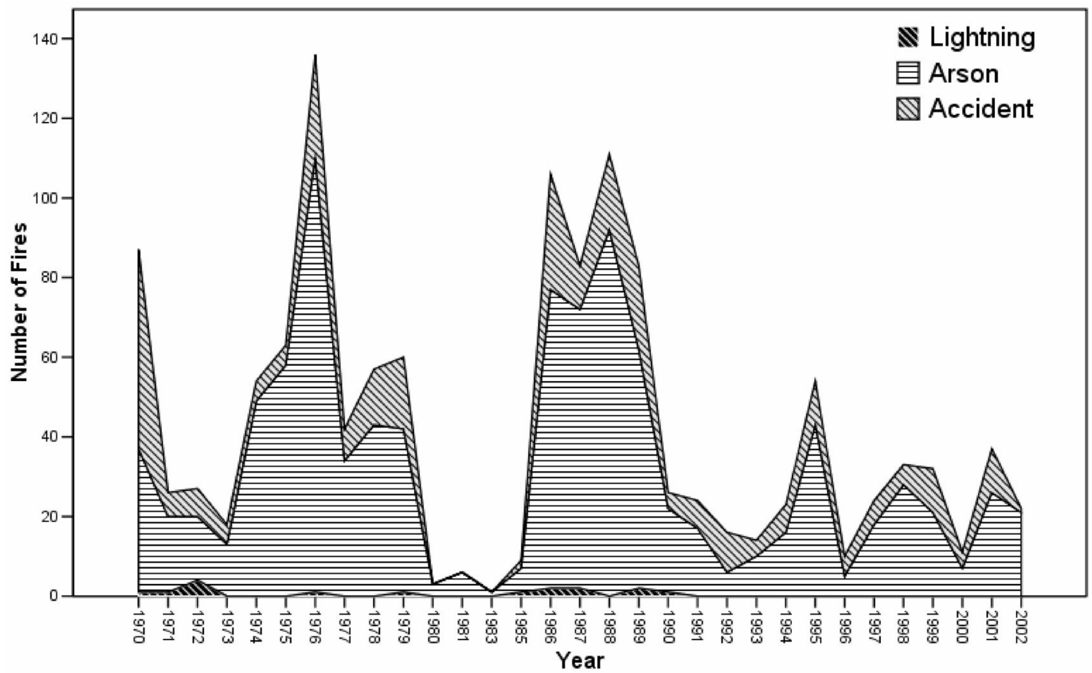

Figure 2. The annual temporal variation of fire counts.

most of our study area is relatively low and exhibits little spatial variation, we still included it as a potentially important factor expressing the ubiquitous effects of human population.

We also believed that the three towns located in our study area affected the spatial pattern of human-caused fires by providing potential arsonists to the nearby forest. We used proximity to the nearest towns (population is about 700, 600, and 1,300 for the three nearby towns Alton, Birch Tree, and Winona, respectively) as one human factor in the analysis. The proximity of a town could also influence fire detection and fire suppression efficiency.

Land cover was also hypothesized to be an important determinant of fire occurrence because different forest types can be related to different fuel types. We used a 30-m resolution grid of the land cover of the state of Missouri published by the Missouri Resources Assessment Partnership (MoRAP) in 1999. This data set is an amalgam of classified Landsat Thematic Mapper (Landsat TM) satellite image data ranging from 1991 to 1993. There are 12 classes of land cover in our study area, and we aggregated them into four classes (Table 1) because some land cover classes (e.g., open water, urban impervious) were rare in the study area. The land cover within the study area was assumed to be constant between 1970 and 2002.

We hypothesized that topography helps to determine the likelihood of fire occurrence because it can influence fuel loading and moisture and limit human accessibility. We used a 60-m resolution grid of digital elevation model (DEM) data published by MSDIS in 1999. The slope and aspect were calculated from the DEM data using the surface analysis provided in the ArcGIS Spatial Analyst tool. Later in our model selection process, the calculated aspect azimuth was further transformed into three aspect categories: mesic for low potential solar insolation (NE, NW, N, and E), xeric for high potential solar insolation (SW, SE, S, and W), and flat.

Characterizing Spatial Point Patterns

We began by computing nonparametric statistics K function and the kernel smoothed intensity function to characterize the spatial structure of the fire occurrence point data. A spatial point pattern is a data set $\mathbf{x} = \{x_1, \ldots, x_n\}$, with n points observed in a bounded region D of the plane R^2. The pattern is often the result of a mixture of both *first-order* and *second-order* effects. First-order effects are related to variation in the mean value of the spatial process, which is a global spatial trend. Second-order (local) effects result from the spatial correlation structure or the spatial dependence in the process (Bailey and Gatrell 1995). The first-order properties in spatial point pattern data are described by the intensity function $\lambda(u)$, which is the mean number of events per unit area at the point u. For a general location s and the given data $\mathbf{x}$ in the region, the kernel-smoothed intensity at s is estimated by

$$\hat{\lambda}_b(s) = \frac{1}{C_b(s)} \sum_{i=1}^{n} K_{ib}\,(s - x_i).$$

where $K_b(\,)$ is a kernel with band $b > 0$, and $C_b(\,)$ is an edge correction factor (Diggle 2003).

The second-order properties of a spatial point pattern can be characterized by Ripley's K function. It is defined as

$$k(h) = \left. \begin{array}{c}\text{[mean number of events}\\ \text{within the lag distance } h\\ \text{of an arbitrary event]}\end{array} \right/ \lambda.$$

The K function is widely used to describe the small-scale

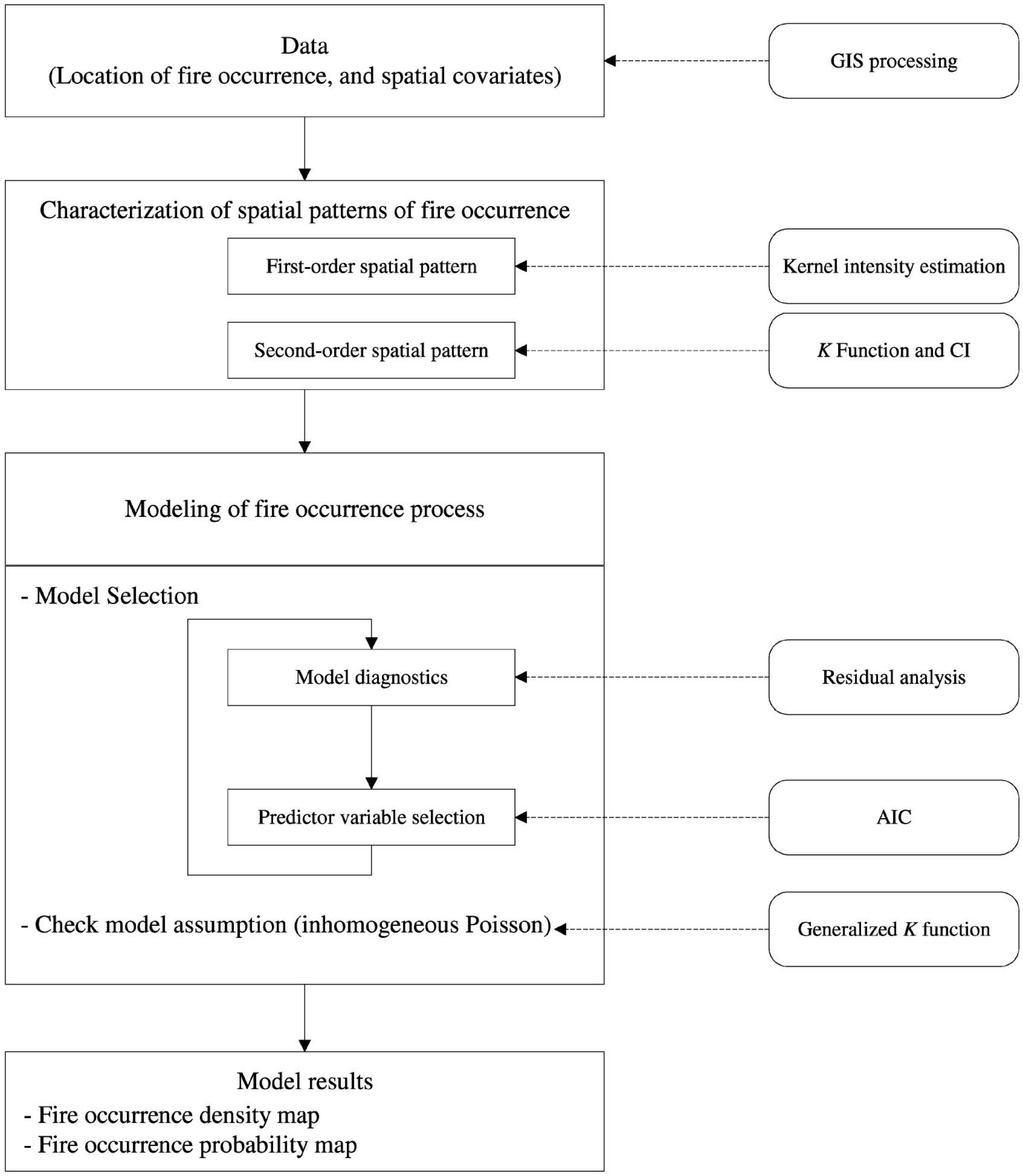

Figure 3. The flow chart of the overall modeling approach. Each rectangle stands for a specific procedure, and each rounded rectangle stands for the method used in the corresponding procedure.

Table 1. Land cover classes from the Missouri Resources Assessment Partnership data and the aggregated classes used in the analysis

Land cover class	Percentage	Aggregated class
Urban impervious	0.09	Urban
Urban vegetated	0.10	
Barren or sparsely vegetated	0.03	
Grown crops	0.01	
Glade complex	0.10	
Open water	0.04	
Cool-season grassland	15.42	Grassland
Eastern redcedar-deciduous forest/woodland	2.02	Deciduous
Deciduous woodland	3.32	
Upland deciduous forest	51.91	
Shortleaf pine-oak forest/woodland	16.77	Pine-oak
Shortleaf pine forest/woodland	10.19	
Sum	100	

spatial correlation structure of a point pattern. The estimate of K is usually compared to the true value of K for the null model of complete spatial randomness (CSR), which is $K(h) = \pi h^2$. Deviations between the empirical and theoretical K curves may suggest spatial clustering or spatial regularity: for a regular pattern $K(h) < \pi h^2$, whereas under clustering $K(h) > \pi h^2$ (Bailey and Gatrell 1995). Edge effects (i.e., a circle centered at the point too close to the boundary to be evaluated without a bias) can be mitigated by edge-correction methods (Diggle 2003). We also used the clustering index (CI) proposed by Genton et al. (2006) to measure the overall amount of clustering across various lag distances. The CI is defined as

$$\mathrm{CI}(h) = \int_0^h \left(\sqrt{K(h)/\pi} - h \right) \mathrm{d}h.$$

A larger CI indicates a higher degree of clustering.

Spatial Point Process Modeling

A spatial point process (e.g., Poisson, Cox, and Strauss process) is any stochastic mechanism that generates the spatial point pattern $\mathbf{x}$. The point process models fitted to the data are often formulated in terms of their Papangelou conditional intensity $\lambda_\theta(u; \mathbf{x})$, which may be loosely interpreted as the conditional probability of having an event at point u given that the rest of the point process coincides with $\mathbf{x}$ (Baddeley and Turner 2000). For the Poisson process, the conditional intensity function is the same as the intensity function $\lambda(u; \mathbf{x}) = \lambda(u)$ because of the independence properties of the Poisson process. In practice, the conditional intensity is often specified through a log-linear regression model with two components:

$$\log \lambda_\theta(u; \mathbf{x}) = \theta_1 B(u) + \theta_2 C(u, \mathbf{x}), \tag{1}$$

where $\theta = (\theta_1, \theta_2)$ are parameters that can be estimated using the maximum pseudolikelihood method (Baddeley and Turner 2000).

The trend term $B(u)$ depends only on the spatial location u, so it represents spatial trend or spatial covariate effects. The interaction term $C(u, \mathbf{x})$ depends on not only the point u, but also the configuration of $\mathbf{x}$, hence it represents stochastic interactions between the points. The term $C(u, \mathbf{x})$ is reduced to zero for the Poisson process.

Residual Analysis

Residual analysis for spatial point processes, recently formulated by Baddeley et al. (2005), is considered a milestone in the development of point pattern statistics in that it provides an excellent technique for determining the need for model refinement or for diagnosing the quality of the model-fitting for spatial point processes in a statistically rigorous way. The method provides various innovation measures and residual measures, which are analogous to the errors and residuals in nonspatial linear models. For example, the Pearson residual measure for any subregion B, defined as

$$\mathrm{R}(B, \hat{\theta}) = \sum_{x_i \in \mathbf{x} \cap B} \frac{1}{\sqrt{\hat{\lambda}(x_i; \mathbf{x})}} - \int_B \sqrt{\hat{\lambda}(u; \mathbf{x})}\, \mathrm{d}u,$$

is analogous to the usual Pearson residuals for linear models in that the variance is standardized, ignoring effects of parameter estimation. We used a technique called the lurking variable plot to investigate the presence of spatial trends in point processes. The technique is to plot the residual against a spatial covariate (or one of the Cartesian coordinates). For a spatial covariate $Z(u)$, which must be spatially continuous in the study region D, we may evaluate the residual measure on each subregion defined by

$$\mathrm{B}(z) = \{u \in D : Z(u) \leq z\},$$

yielding a cumulative residual function

$$\mathrm{A}(z) = \mathrm{R}(\mathrm{B}(z), \hat{\theta}).$$

The function $A(z)$ should be approximately zero if the fitted model is correct, and any systematic pattern in the lurking variable plot suggests an appropriate modification of the model to better account for that spatial covariate.

Model Selection

The compiled fire occurrence data and GIS-processed spatial covariates were analyzed and modeled using a spatial point pattern analysis package (*spatstat*) of R (Baddeley and Turner 2005). The major task at this step was to identify a best choice of model that would balance fit to the data, model size (parameters within the model), and model complexity. The goal was to (1) determine the interaction term $C(u, \mathbf{x})$ specified in Equation 1, and (2) select predictor variables (Cartesian coordinates, spatial covariates, and their transformations) for the trend term $B(u)$. Three major tools were used in this stage: residual analysis, the generalized K function K_i, and the Akaike information criterion (AIC). The AIC (Akaike 1974) is widely used as a measure for selecting the best among competing models for a fixed data set. It is defined as

$$\mathrm{AIC} = -2 \max(\text{log-likelihood}) + 2\,(\text{number of parameters}).$$

The model with the smaller AIC is considered to be a better fit to the data. It has been emphasized that AIC can be used for comparisons among non-nested models (Ogata 1988).

The data were first fitted to the null model (a homogeneous Poisson model), and we used residual analysis (especially lurking variable plots) to find out whether the point pattern depended on tested spatial covariates. We then assumed the fire occurrence process was an inhomogeneous process (hence $C(u, \mathbf{x})$ was zero) and included the spatial covariates that were revealed to affect the fire occurrence patterns into the trend term $B(u)$ of the log-linear regression model. The included spatial covariates could be transformed, and we used a polynomial function (up to a power of three provided that convergence in maximum likelihood estimation is attained) to transform spatial continuous covariates or Cartesian coordinates. Unlike stepwise methods that use a restricted search and use a dubious hypothesis testing method for choosing among models, AIC methods typically involve a wider search and compare models in a preferable manner (Faraway 2002). We fitted all alternative models and chose the best one according to AIC and model diagnostics. Finally, we computed the generalized K function K_i, which can be used to evaluate the second-order pattern after taking into account first-order effects (Baddeley et al. 2000) for the fit model to check the Poisson process assumption. If there were still any second-order patterns after we weighted the trend fitted from the model, then the Poisson process assumption was considered to be unsound, and other point processes (e.g., Strauss process) might be assumed.

Results

Characteristics of Fire Occurrence Patterns

The kernel-smoothed intensity (Figure 4) estimated from the fires reported between 1970 and 2002 showed a "hot bed" (where fire occurrence density is high) in the north center of the study area (near the town of Winona) and a "cold bed" (where fire occurrence density is low) in the southwest (near the town of Alton), indicating that the fire occurrence process is not completely random. The estimated K function (Figure 5) is larger than that of theoretical CSR, indicating a spatial clustering pattern. However, it doesn't necessarily suggest that there is a strong dependence among fire ignition points, since the clustering of environmental factors (e.g., camping sites, vegetation distribution) may also explain this clustered pattern. Our model-fitting exercises in which fire occurrence is modeled as an inhomogeneous Poisson process further showed that the spatial clustering pattern could indeed be explained by the environmental heterogeneity. The amount of clustering varied by ignition source: CI (h = 5 km) = 4.57 km^2 for arson; 3.40 for accident; and 2.94 for lightning.

Model-Fitting and Diagnostics

The null model (homogeneous Poisson) fitted to the data has a constant intensity, valued 1.01e−06 m^{-2} over the entire region. This constant intensity can be interpreted as an average of 1.01 fire occurrences every 1 km^2 in our study area over 33 years (1970–2002). We plotted the cumulative Pearson residuals against the five continuous spatial covariates (distance to nearest road, slope,

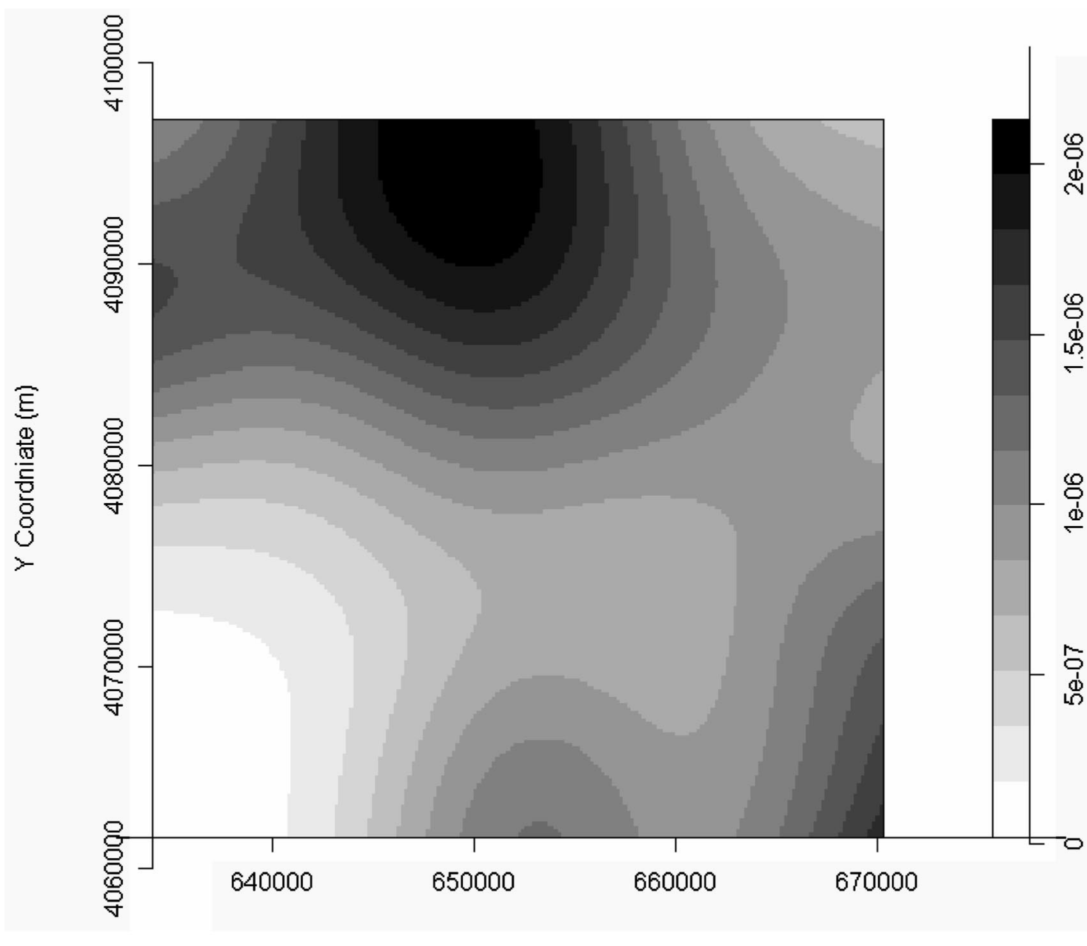

Figure 4. Kernel smoothed intensity (expected number of fires per meter) estimated from the fire occurrence data reported in 1970–2002.

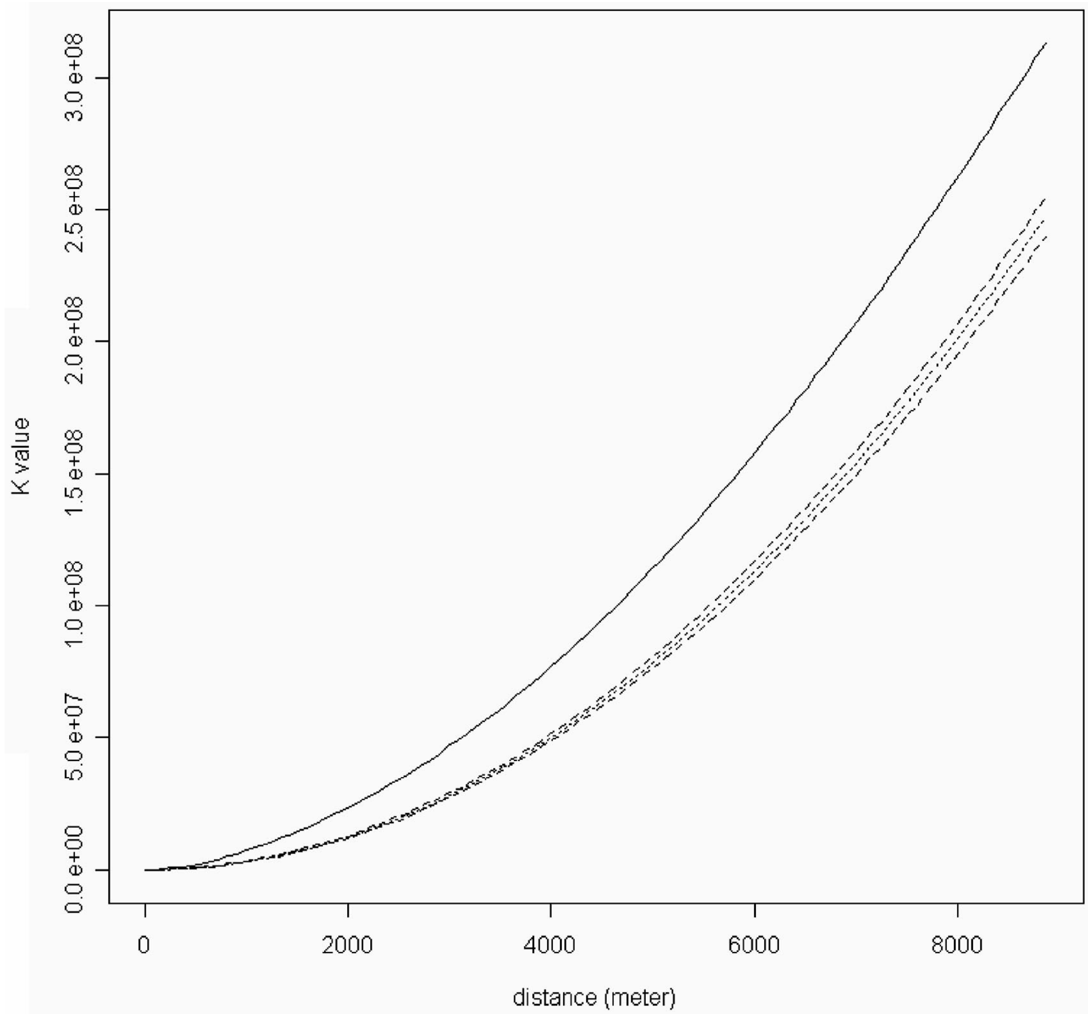

Figure 5. Estimated *K* function for the observed fires in 1970–2002.

elevation, population density, and distance to nearest town) and the two Cartesian coordinates (x and y) for the null model (some are shown in Figure 6). The cumulative residual function defined by Equation 2 should approximate zero if the null model adequately addresses effects of the spatial covariate. In Figure 6a, the cumulative Pearson residuals are larger than the $+2\sigma$ limit when the distance to the nearest road $<$1,500 m. This suggests the null model underestimates fire occurrence at this scale. In other words, there are more fires occurring at spatial locations that are close to roads than the null model predicts. The steepest increases in the curve are observed at $<$150 meters from roads, implying that many human-caused fires occurred very close to roads (refer to Figure 1). The peak in Figure 6b occurs for slopes of about 25 degrees and the steepest increase occurs within the range of slopes of 10–20 degrees, suggesting that more fires occur on gentle slopes than steeper slopes. There are obvious systematic patterns related to elevation (Figure 6c). The nadir of the curve is much less than the -2σ limit of error bounds, and there is a steep increase of cumulative residuals after the nadir point (270–330 m). These patterns suggest that the null model overestimates intensity of fire occurrence for spatial locations at $<$270 m of elevation, and more fires are ignited at higher elevations. Figure 6d shows that there is more than the average number of fires at the scale of 6 km to 17 km distance from the nearest towns. The decrease of cumulative residuals when the distance is $<$4 km or $>$14 km suggests that fewer fires occur either close to or far away from the towns.

It is worth noting that fires caused by different sources may respond to the spatial covariates differently. This can be verified from lurking variable plots. Take elevation as an example. The cumulative residuals are below the error bounds at elevations $<$300 m for arson (Figure 7a), but are within the error bounds for accident fires (Figure 7b) and above error bounds at elevations $<$220 m for lightning fires (Figure 7c). This finding motivated us to model the fire occurrence process not only for the aggregation of all kinds of fires, but also for arson, accident, and lightning fires separately. We considered a wide range of alternative models that included all possible combinations of the potential spatial covariates. Table 2 gives AIC values and formulas of the best models identified in the model selection stage for different fire types. The identified set of predictor variables and their transformations are different for different causes. When we did not include Cartesian coordinates as predictor variables in the log-linear regression models, the best model fit to the aggregated fire data, denoted as H_1,

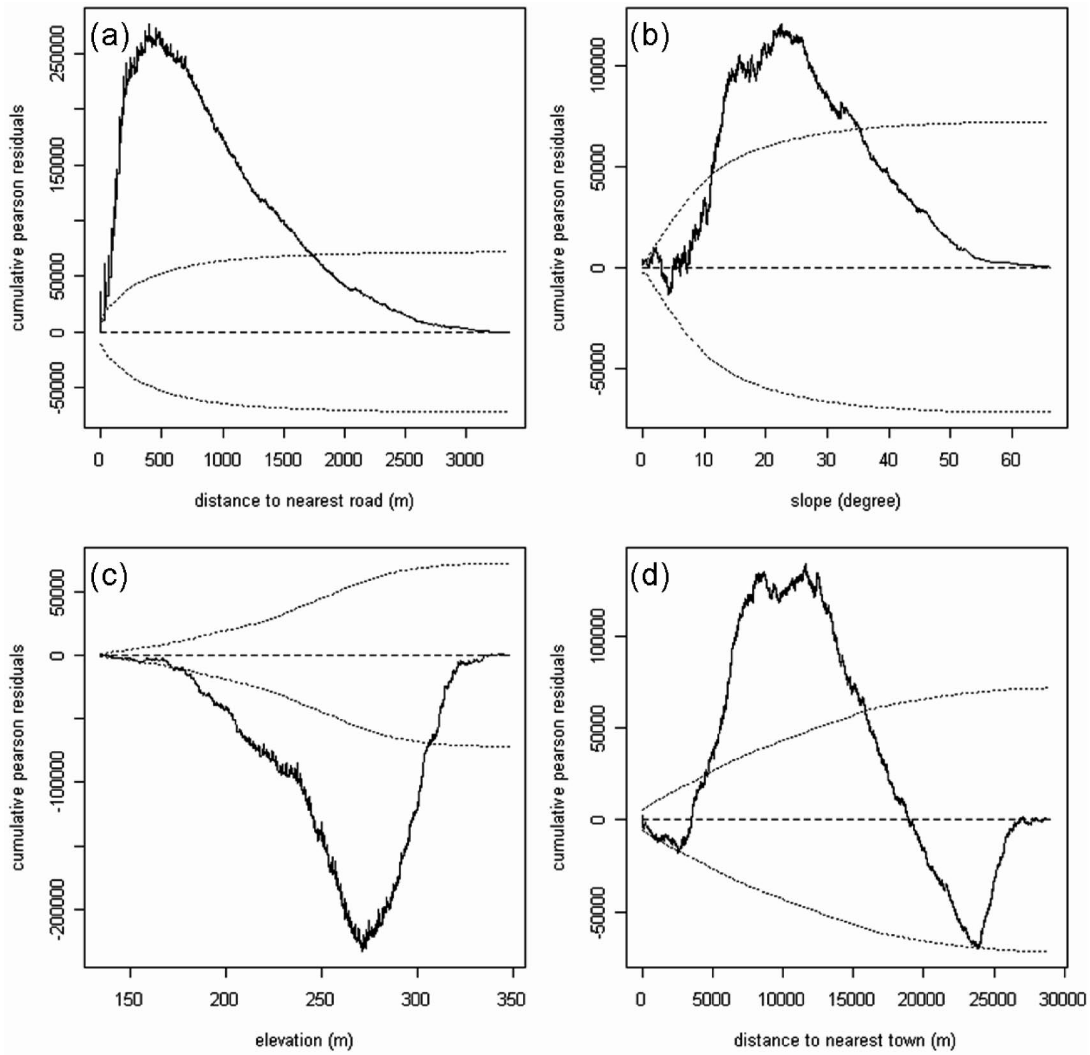

Figure 6. Lurking variable plots against (a) distance to nearest road, (b) slope, (c) elevation, and (d) distance to nearest town for the null model of cause-aggregated fire data. The solid lines are empirical curves of cumulative Pearson residuals. The dotted lines denote two-standard-deviation error bounds.

identified the predictor variables as the distance to the road, slope, aspect, distance to nearest town, ownership, and land cover. Elevation and population density were dropped. Nevertheless, there was still glaring violation of error bounds in the lurking variable plots against Cartesian coordinates for the model H_1, which implied the presence of a residual spatial trend not captured in the model. We then added Cartesian coordinates into the model. The selected model H_2, having 24 parameters, (Table 2) was more complicated than H_1, but also showed greater improvement in terms of model diagnostics. Including Cartesian coordinates also decreased AIC for the arson and accident fires, but not for the lightning-caused fires.

We calculated the generalized K function for an inhomogeneous Poisson point process for the final full model H_2. The calculated K_i function was within the envelopes of 99 simulations of an inhomogeneous Poisson point process with the same spatial distribution of intensity as the one observed for model H_2 (Figure 8), indicating that our assumption of an inhomogeneous Poisson process was valid.

Distribution of Fire Occurrence Probability in Space

The spatial intensity calculated from our point process model can be loosely interpreted as fire occurrence density, which can be used in calculating fire occurrence probability. We used the model H_2 as our final full model to derive fire occurrence probability in space. The study area was represented as a grid of 1 ha (100 $\times$ 100 m) cells. The normalized fire occurrence density (λ, number of fires per m^2 per decade) was then estimated using the model H_2. The fire occurrence probability is defined as the probability of having at least one fire occurrence within the cell over a decade. It is calculated from the Poisson probability density function as

$$\mathrm{P}(x \geq 1) = 1 - \mathrm{P}(x = 0) = 1 - e^{-\lambda * 1000},$$

where x denotes number of fire occurrences within the cell. The calculated fire occurrence probability for each cell is shown in Figure 9.

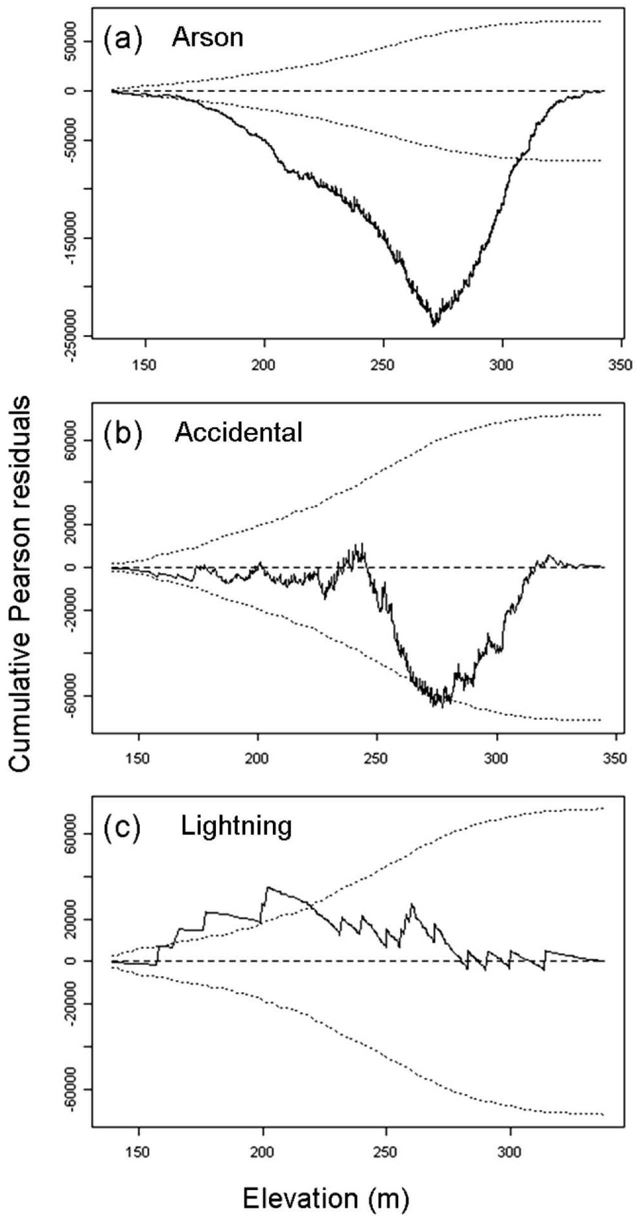

Figure 7. Lurking variable plots against elevation for the null models of (a) arson, (b) accident, and (c) lightning fire data.

Effects of Biotic, Abiotic, and Human Factors

The lurking variable plots against spatial continuous covariates for the null model have already provided some rough ideas on how these factors affect fire occurrence. The coefficients of the chosen models estimated from the fire occurrence data provided further quantitative information on how each spatial covariate (both continuous and categorical) affects fire occurrence density, and consequently fire occurrence probability. Table 3 gives estimated parameters of the spatial covariates for the model H_2. The model is a log-linear regression of intensity as shown in Equation 1, where the conditional intensity is equivalent to the intensity for Poisson point processes. The intensity is a product of several multiplicative components, and each component represents the contribution of a predictor variable. Variables with positive coefficients have positive contributions to fire occurrence density and hence fire occurrence probability, while the ones with negative coefficients have negative contributions. The effects of most continuous variables were curvilinear (quadratic or cubic). For example, log-intensity is modeled as a quadratic function with respect to slope for the aggregated fires. Using the estimated corresponding parameters and basic calculus, one can derive that the log-intensity contributed from the slope factor increases from zero to the maximum 0.16 as slope increases from zero to 15.59 degrees; then it decreases as slope increases further. It becomes a negative value when slope is $>$31.17 degrees. This approximates the curve shown in the corresponding lurking plot (Figure 6b). The ranking of levels of a categorical variable can be easily sorted by the order of their corresponding coefficients. The ranking of land cover class in terms of fire occurrence probability is urban $<$ grassland $<$ deciduous forest $<$ pine-oak mixed forest. The ranking is non-MTNF $<$ private land $<$ public land for ownership, and mesic $<$ xeric $<$ flat for aspect (Table 3).

Discussion

Studies of dendrochronological fire histories from the Missouri Ozarks have suggested that cultural effects such as roads and agricultural development are replacing fuel and topography as determinants of the spread, frequency, and size of fires during the modern era (Guyette et al. 2002). Our results further showed that fire occurrences are spatially clustered, and human factors such as ownership, municipalities, human accessibility, and population density are important determinants of spatial locations of human-caused fires. Proximity to roads is one aspect of human accessibility, and the places close to roads ($<$500 m) are generally associated with higher fire occurrence probabilities. The roads are located either on the ridge tops or in the valley bottoms in this region, but most fires are set along the roads on ridge tops (elevation $>$270 m asl), as shown in Figure 6c. This pattern can be partly explained by the fact that moisture is greater in lower elevations in the landscape. This decreases flammability of fuels, increases the rate of fuel decomposition, and shifts the species composition from pine-oak mixed forests, which tend to accumulate surface fuels, to more mesic white oak forests (Batek et al. 1999). Slope angle is also a limiting factor. Fires go upslope faster with higher intensity, but steep slopes imply greater topographic roughness (Guyette et al. 2002) and are more difficult for humans to gain access. The differential effects of various land cover classes (substitute for fuel types) on human-caused fire occurrence probability can also be explained by the characteristics of human behaviors. If an arsonist intends to set a fire, he or she is more likely to experience success in flammable fuel types such as shortleaf pine and oak forests in this region. Therefore, fuel type and topography, which are often modeled in lightning-caused fire occurrence processes (e.g., Wotton and Martell, 2005), may still be prominent ingredients in human-caused fire occurrence processes, although the explanations of such effects are different. This finding

Table 2. **Comparison of Poisson point process models fitted to the fire occurrence data**

Model	AIC	Number of parameters	Formula
Null model (homogeneous Poisson with a constant intensity over entire area)			
H_0	38,469	1	1
H_0.a	29,293	1	1
H_0.b	10,161	1	1
H_0.c	617	1	1
Models with minimum AIC that only include spatial covariates			
H_1	37,486	15	$P(D_1, 2) + P(D_3, 2) + F(D_4) + P(D_5, 3) + F(D_7) + F(D_8)$
H_1.a	28,465	14	$P(D_1, 1) + P(D_2, 1) + P(D_3, 1) + F(D_4) + P(D_5, 3) + F(D_7) + F(D_8)$
H_1.b	9,972	13	$P(D_1, 2) + P(D_3, 2) + P(D_5, 3) + P(D_6, 3) + F(D_7)$
H_1.c	615	3	$P(D_2, 1) + P(D_3, 2)$
Models with minimum AIC that also include Cartesian coordinates			
H_2	37,401	24	$P(D_1, 2) + P(D_3, 2) + F(D_4) + P(D_5, 3) + F(D_7) + F(D_8) + P(x, y, 3)$
H_2.a	28,324	23	$P(D_1, 1) + P(D_2, 1) + P(D_3, 1) + F(D_4) + P(D_5, 3) + F(D_7) + F(D_8) + P(x, y, 3)$
H_2.b	9,959	22	$P(D_1, 2) + P(D_3, 2) + P(D_5, 3) + P(D_6, 3) + F(D_7) + P(x, y, 3)$
H_2.c	615	3	$P(D_2, 1) + P(D_3, 2)$

Model H_0 denotes the null model for aggregation of all kinds of fire data; H_0.a denotes the null model for arson, H_0.b for accidental fires, and H_0.c for lightning fires. Similarly, H_1 denotes the best model fit to aggregated data when Cartesian coordinates were not included, . . . , H_2.c denotes the best model fit to the lightning fires when Cartesian coordinates were included.
$D_1, D_2, \ldots, D_8$ denote the spatial covariates: distance to nearest road, elevation, slope, aspect, distance to nearest town, population density, ownership, and land cover, respectively; P(cov, *n*) denotes a polynomial function up to the order *n* of a continuous covariate cov; and F(cov) denotes the covariate cov is categorical (i.e., a factor).

Table 3. **Coefficients of the predictor variables of the final models**

Variables	Fire causes			
	Aggregated	Arson	Accident	Lightning
Intercept	4.13e + 06 (4.40e + 06)	5.44e + 05 (5.37e + 06)	1.27e + 07 (9.65e + 06)	−1.7e + 01*** (3.01e + 00)
D_1	−1.61e − 03*** (2.12e − 04)	−9.98e − 04*** (1.31e − 04)	−1.96e − 03*** (3.73e − 04)	
D_1^2	3.17e − 07*** (1.09e − 07)		5.43e − 07*** (1.69e − 07)	
D_2		1.04e−02*** (2.51e−03)		−1.0e − 02 (9.84e − 03)
D_3	2.07e − 02 (1.33e − 02)	−5.00e − 03 (5.84e − 03)	5.17e − 02** 2.50e − 02	
D_3^2	−6.64e − 04** (3.08e − 04)		−1.03e − 03 (5.57e − 04)	−3.71e − 03 3.18e − 03
D_4: mesic	0 (N/A)	0 (N/A)		
D_4: xeric	1.43e − 01* (7.47e − 02)	2.39e − 01 (8.96e − 02)		
D_4: flat	4.51e − 01 (5.11e − 01)	5.11e − 01 (6.22e − 01)		
D_5	−0.63e − 05 (6.97e − 05)	6.93e − 05 (8.63e − 05)		
D_5^2	−9.83e − 09** (4.88e − 09)	−2.21e − 08*** (6.08e − 09)		
D_5^3	3.19e − 13*** (1.14e − 13)	5.74e − 13*** (1.42e − 13)		
D_6			−1.74e − 02 (2.75e − 02)	
D_6^2			7.50e − 04 (6.50e − 04)	
D_6^3			−2.75e − 06 2.75e − 06	
D_7: non-MTNF	0 (N/A)	0 (N/A)	0 (N/A)	
D_7: private	2.57e + 00*** (4.25e − 01)	2.22e + 00*** (4.78e − 01)	5.52e + 00* (3.18e + 00)	
D_7: public	3.23e + 00*** (4.33e − 01)	2.98e + 00*** (4.88e − 01)	5.76e + 00* (3.17e + 00)	
D_8: urban	0 (N/A)	0 (N/A)		
D_8: grassland	6.14e − 01 (1.34e + 00)	1.05e + 01 (2.51e + 02)		
D_8: deciduous	8.85e − 01 (1.34e + 00)	1.07e + 01 (2.51e + 02)		
D_8: mixed	8.92e − 01 (1.34e + 00)	1.06e + 01 (2.51e + 02)		
Observations	7,703	5,874	1,917	645

$D_1, D_2, \ldots, D_8$ denote the spatial covariates: distance to nearest road, elevation, slope, aspect, distance to nearby town, population density, ownership, and land cover, respectively.
*** Indicates the *t*-test is rejected at 1% significance, ** at 5%, and * at 10%.

suggests that in an anthropogenic fire regime, vegetation dynamics can still be an important determinant of fire occurrence.

The map of fire occurrence probability estimated from the model can be directly used in fire risk analysis by forest managers. For example, it can feed directly into forest planning and spatially explicit assessment of fire risk (Mark Twain National Forest 2005). The map helps us not only to pinpoint "hot spots" (high fire occurrence probability), but also to identify areas that should be considered for greater attention with respect to fire risk. For example, The Irish Wilderness, a 66 km^2 wilderness area established in 1984 and a part of the Mark Twain National Forest, is located in the southeast of our study area. The area is primarily covered by shortleaf pine and oak forests/woodlands, and deciduous forests. There are no drivable roads in the wilderness area and motorized or mechanical equipment, including bicycles, are prohibited. Hence the fire occurrence probability in this area is very low. However, the fire occurrence

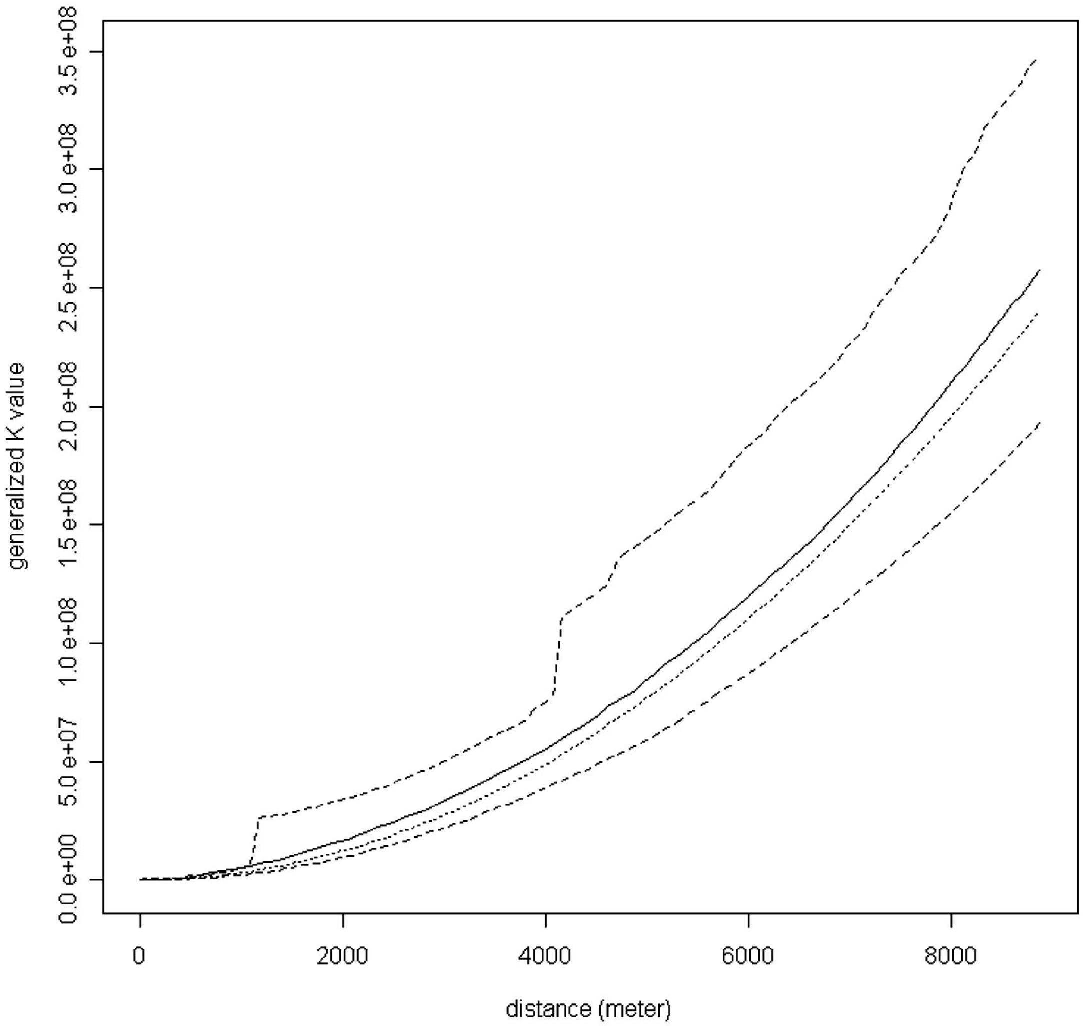

Figure 8. Generalized K function, K_i for the final full inhomogeneous Poisson point process model H_2. The solid line represents the observed K_i, the dotted line is the theoretical K_i function, and the two dashed lines are the upper and lower bounds calculated from 99 simulations.

probability surrounding this area (especially to the east and west) is high (Figure 9). This makes the chance for the wilderness area to be burned by the propagation of fire ignited from surroundings relatively high, and access for fire suppression is problematic. Our historical fire database showed that there were only 14 fires ignited in this area during 1970–2002, but three of them were as big as 100 ha in fire size.

Another important application of our results is in forest landscape modeling of succession and fire disturbance. Simulation models of forest landscape change are increasingly used in long-term forest planning (He and Mladenoff 1999, Keane et al. 2004, Shifley et al. 2006). Fire occurrence simulation is an important component in fire simulation for long-term, large-scale planning. Fire occurrence is usually modeled as two consecutive stages: fire ignition and fire initiation (e.g., Antonovski et al. 1992, Davis and Burrows 1994, Li 2000, Yang et al. 2004). For these forest landscape models, the number of ignitions is often simulated from a Poisson distribution whose parameter intensity (i.e., fire ignition density) is provided by users as a model input. The most common parameterization of the fire ignition density is based on fire frequency estimates that are uniform over large areas. It therefore fails to incorporate the effects of roads and topography on the fire occurrence process. We can derive a fire ignition density map from our modeling results and use it to refine the input data for these models of landscape change in response to fire and other disturbances. That will greatly improve simulation realism in terms of spatial patterns of fire occurrence.

We used a spatial point process modeling approach to analyze spatial patterns of human-caused fires in this landscape. We were able to quantify the effects of biotic (land cover), abiotic (topography), and human factors (ownership, proximity to roads and towns, and population density) on fire occurrence by using an inhomogeneous Poisson model. The inclusion of Cartesian coordinates in the final models suggests that there are still some important explanatory factors omitted from our model. Because the extent of our study area is relatively small, the spatial variations of many other potential factors such as law enforcement, weather, and climate patterns are

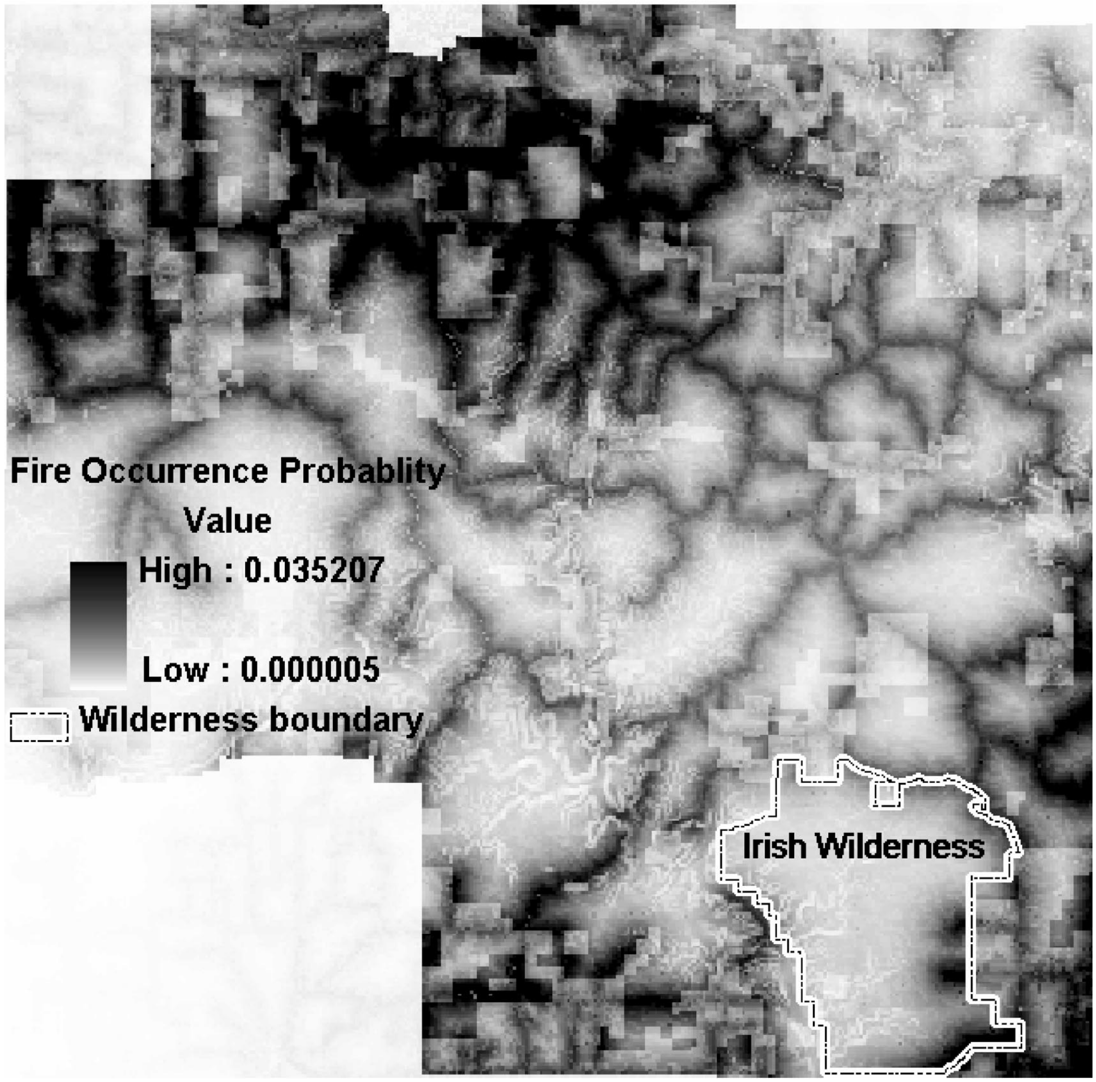

Figure 9. The grid map of estimated fire occurrence probability, defined as the probability of having at least one fire occurrence for the given 1 ha-size cell over a decade.

difficult to measure and incorporate. We intend to include more variables at larger spatial scales in future studies. Cartesian coordinates are often used to substitute for unobservable factors. This practice is useful and common in applications of point process modeling (e.g., Ogata 1988). The residual analysis proved to be a very powerful tool in selecting predictor variables and characterizing the effects of spatial covariates. The approach used here is statistically rigorous and can be applied elsewhere in the modeling of human-caused or natural-caused fire occurrence process and other point processes in forestry. The approach allows use of spatial grid data with different resolutions (e.g., the resolution of DEM data is 60 m, while the resolution of land cover data is 30 m), but it requires all the spatial covariates to be either spatially continuous (e.g., distance to nearest road) or lattice data (e.g., land cover). It is possible in future studies that a newly added explanatory variable is observed only in the locations of fire occurrence. In this case, one could use geostatistical methods to interpolate from the observed values to the entire region (Rathbun 1996).

Literature Cited

AKAIKE, H. 1974. A new look at the statistical model identification. *IEEE Trans. on Automatic Control* 19:716–723.

ANTONOVSKI, M.Y., M.T. TER-MIKAELIAN, AND V.V. FURYAEV. 1992. A spatial model of long-tem forest fire dynamics and its applications to forests in western Siberia. P. 373–403 in *A systems analysis of the global boreal forest*, Shugart, H.H., R. Leemans, and G.B. Bonan (eds.). Cambridge University Press.

BADDELEY, A., J. MØLLER, AND R. WAAGEPETERSEN. 2000. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica* 54: 329–350.

BADDELEY, A., AND R. TURNER. 2000. Practical maximum pseudolikelihood for spatial point patterns. *Austral. NZ J. Statist.* 42(3):283–315.

BADDELEY, A., AND R. TURNER. 2005. Spatstat: An R package for analyzing spatial point patterns. *J. Statist. Software* 12(6):1–42.

BADDELEY, A., R. TURNER, J. MØLLER, AND M. HAZELTON. 2005. Residual analysis for spatial point processes. *J. R. Statist. Soc. B.* 67(5):617–666.

Bailey, T.C., and A.C. Gatrell. 1995. *Interactive spatial data analysis*. Longman Scientific & Technical, Harlow. 436 p.

Batek, M.J., A.J. Rebertus, W.A. Schroeder, T.L. Haithcoat, E. Compas, and R.P. Guyette. 1999. Reconstruction of early nineteenth-century vegetation and fire regimes in the Missouri Ozarks. *J. Biogeography* 26(2):397–412.

Bergeron, Y., S. Gauthier, M. Flannigan, and V. Kafka. 2004. Fire regimes at the transition between mixed wood and coniferous boreal forest in Northwestern Quebec. *Ecology* 85(7):1916–1932.

Butry, D.T., and J.P. Prestemon. 2005. *Spatio-temporal wildland arson crime function*. Paper presented at the Annual Meeting of the American Agricultural Economics Association, July 26–29, 2005. Providence, RI. 18 p.

Cardille, J.A., S.J. Ventura, and M.G. Turner. 2001. Environmental and social factors influencing wildfires in the Upper Midwest, United States. *Ecol. Applic.* 11(1):111–127.

Davis, F.W., and D.A. Burrows. 1994. Spatial simulation of fire regime in Mediterranean-climate landscapes. P. 117–139 in *The role of fire in Mediterranean-type ecosystems*, Talens, M.C., W.C. Oechel, and J.M. Moreno (eds.). Springer-Verlag, New York.

Diaz-Avalos, C., D.L. Peterson, E. Alvarado, S.A. Ferguson, and J.E. Besag. 2001. Space-time modelling of lightning-caused ignitions in the Blue Mountains, Oregon. *Can. J. For. Res.* 31:1579–1593.

Diggle, P. 2003. *Statistical analysis of spatial point patterns*, 2nd ed. Arnold, London. 168 p.

Donoghue, L.R., and W.A. Main. 1985. Some factors influencing wildfire occurrence and measurement of fire prevention effectiveness. *J. Environ. Manage.* 20:87–96.

Faraway, J. 2002. Practical regression and ANOVA using R. Available online at www.maths.bath.ac.uk/~jjf23/book; last accessed Dec. 12, 2006. 212 p.

Genton, M.G., D.T. Butry, M. Gumpertz, and J.P. Prestemon. 2006. Spatio-temporal analysis of wildfire ignitions in the St. Johns River Water Management District. *Int. J. Wildland Fire* 15:87–97.

Guyette, R.P., R.M. Muzika, and D.C. Dey. 2002. Dynamics of an anthropogenic fire regime. *Ecosystems* 5:472–486.

He, H.S., and D.J. Mladenoff. 1999. Spatially explicit and stochastic simulation of forest landscape fire disturbance and succession. *Ecology* 80(1):81–99.

Hessburg, P.F., J.K. Agee, and J.F. Franklin. 2005. Dry forests and wildland fires of the inland Northwest USA: Contrasting the landscape ecology of the pre-settlement and modem eras. *For. Ecol. Manage.* 211(1–2):117–139.

Keane, R.E., G.J. Cary, I.D. Davies, M.D. Flannigan, R.H. Gardner, S. Lavorel, J.M. Lenihan, C. Li, and T.S. Rupp. 2004. A classification of landscape fire succession models: Spatial simulations of fire and vegetation dynamics. *Ecol. Model.* 179:3–27.

Kokkila, T., A. Makela, and E. Nikinmaa. 2002. A method for generating stand structures using Gibbs marked point process. *Silva Fennica* 36(1):265–277.

Kushla, J.D., and W.J. Ripple. 1997. The role of terrain in a fire mosaic of a temperate coniferous forest. *For. Ecol. Manage.* 95(2):97–107.

Larjavaara, M., T. Kuuluvainen, and H. Rita. 2005. Spatial distribution of lightning-ignited forest fires in Finland. *For. Ecol. Manage.* 208(1–3):177–188.

Larsen, D.R., and L.C. Bliss. 1998. An analysis of structure of tree seedling populations on a Lahar. *Landscape Ecol.* 13:307–332.

Latham, D., and E. Williams. 2001. Lightning and forest fires. P. 376–418 in *Forest fires: Behavior and ecological effects*, Johnson, E.A., and K. Miyanishi (eds.). Academic Press, San Diego.

Li, C. 2000. Reconstruction of natural fire regimes through ecological modelling. *Ecol. Model.* 134:129–144.

Mark Twain National Forest. 2005. Fire risk assessment, Appendix G of final environmental impact statement to accompany the 2005 land and resource management plan (2005 forest plan). Mark Twain National Forest, U.S. Dept. Agric. Forest Service, Rolla, MO. Available at www.fs.fed.us/r9/forests/marktwain/projects/forest_plan_revision/; last accessed Dec. 12, 2006.

Mercer, D.E., and J.P. Prestemon. 2005. Comparing production function models for wildfire risk analysis in the wildland-urban interface. *For. Policy Economics* 7(5):782–795.

Mermoz, M., T. Kitzberger, and T.T. Veblen. 2005. Landscape influences on occurrence and spread of wildfires in Patagonian forests and shrub lands. *Ecology* 86(10):2705–2715.

Moritz, M.A., J.E. Keeley, E.A. Johnson, and A.A. Schaffner. 2004. Testing a basic assumption of shrubland fire management: How important is fuel age? *Frontiers Ecol. Environm.* 2:67–72.

Ogata, Y. 1988. Statistical models for earthquake occurrence and residual analysis for point processes, *J. Am. Statist. Assoc.* 83:9–27.

Podur, J., D.L. Martell, and F. Csillag. 2003. Spatial patterns of lightning-caused forest fires in Ontario, 1976–1998. *Ecol. Model.* 164(1):1–20.

Prestemon, J.P., and D.T. Butry. 2005. Time to burn: Modeling wildland arson as an autoregressive crime function. *Am. J. Agric. Econ.* 87(3):756–770.

Prestemon, J.P., J.M. Pye, D.T. Butry, T.P. Holmes, and D.E. Mercer. 2002. Understanding broad scale wildfire risks in a human-dominated landscape. *For. Sci.* 48:685–693.

Rathbun, S.L. 1996. Estimation of Poisson intensity using partially observed concomitant variables. *Biometrics* 52:226–242.

Ryan, K.C. 2002. Dynamic interactions between forest structure and fire behavior in boreal ecosystems. *Silva Fennica* 36(1):13–39.

Shifley, S.R., F.R. Thompson, W.D. Dijak, M.A. Larson, and J.J. Millspaugh. 2006. Simulated effects of forest management alternatives on landscape structure and habitat suitability in the Midwestern United States. *For. Ecol. Manage.* 229(1–3):361–377.

Stoyan, D., and A. Penttinen. 2000. Recent applications of point process methods in forestry statistics. *Statist. Sci.* 15(1):61–78.

Sturtevant, B.R., P.A. Zollner, E.J. Gustafson, and D.T. Cleland. 2004. Human influence on the abundance and connectivity of high-risk fuels in mixed forests of northern Wisconsin, USA. *Landscape Ecol.* 19:235–253.

Tanskanen, H., A. Venalainen, P. Puttonen, and A. Granstrom. 2005. Impact of stand structure on surface fire ignition potential in *Picea abies* and *Pinus sylvestris* forests in southern Finland. *Can. J. For. Res.* 35(2):410–420.

Turner, M.G., and W.H. Romme. 1994. Landscape dynamics in crown fire ecosystems. *Landscape. Ecol.* 9(1):59–77.

Veblen, T.T., T. Kitzberger, R. Villalba, and J. Donnegan. 1999. Fire history in northern Patagonia: The roles of humans and climatic variation. *Ecol. Monogr.* 69(1):47–67.

Weber, M.G., and B.J. Stocks. 1998. Forest fires and sustainability in the boreal forests of Canada. *Ambio* 27(7):545–550.

Westin, S. 1992. *Wildfire in Missouri*. Missouri Department of Conservation, Jefferson City. 161 p.

Wotton, B.M., and D.L. Martell. 2005. A lightning fire occurrence model for Ontario. *Can. J. For. Res.* 35(6):1389–1401.

Yang, J., H.S. He, and E.J. Gustafson. 2004. A hierarchical fire frequency model to simulate temporal patterns of fire regimes in LANDIS. *Ecol. Model.* 180:119–133.