Environ Ecol Stat (2009) 16:271–289
DOI 10.1007/s10651-007-0088-y

The fire frequency analysis branch of the pyrostatistics tree: sampling decisions and censoring in fire interval data

Max A. Moritz · Tadashi J. Moody · Lori J. Miles · Matthew M. Smith · Perry de Valpine

Received: 1 March 2006 / Revised: 1 October 2006 / Published online: 28 February 2008

Abstract Statistical characterization of past fire regimes is important for both the ecology and management of fire-prone ecosystems. Survival analysis—or fire frequency analysis as it is often called in the fire literature—has increasingly been used over the last few decades to examine fire interval distributions. These distributions can be generated from a variety of sources (e.g., tree rings and stand age patterns), and analysis typically involves fitting the Weibull model. Given the widespread use of fire frequency analysis and the increasing availability of mapped fire history data, our goal has been to review and to examine some of the issues faced in applying these methods in a spatially explicit context. In particular, through a case study on the massive Cedar Fire in 2003 in southern California, we examine sensitivities of parameter estimates to the spatial resolution of sampling, point- and area-based methods for assigning sample values, current age surfaces versus historical intervals in generating distributions, and the inclusion of censored (i.e., incomplete) observations. Weibull parameter estimates were found to be roughly consistent with previous fire frequency analyses for shrublands (i.e., median age at burning of ∼30–50 years and relatively low age dependency). Results indicate, however, that the inclusion or omission of censored observations can have a substantial effect on parameter estimates, far more than other decisions about specifics of sampling.

Keywords Age dependency · Cedar Fire · Fire ecology · Hazard function · Spatial analysis · Weibull model

M. A. Moritz (✉) · T. J. Moody · L. J. Miles · M. M. Smith · P. de Valpine
Environmental Science, Policy, and Management Department, University of California, Berkeley, CA, USA
e-mail: mmoritz@nature.berkeley.edu

1 Introduction

For millennia, fire has been one of the most important natural processes in terrestrial ecosystems across the planet. Many plant species have developed adaptations to recurring fires, such as thick bark that provides protection from heat and combustion, and seed banks that require heat or chemical cues from fire for germination (Bond and van Wilgen 1996). Since Euro-American settlement, North American ecosystems have been subject to a variety of management regimes, including logging, grazing, fire suppression, elimination of Native American ignitions, and expansion of the urban-wildland interface, all of which have directly and indirectly affected the process of fire on the landscape. Today, it is accepted that an understanding of past fire patterns can provide important guidance for contemporary management and restoration of ecosystems. As such, quantitative understanding of past fire occurrence as a cyclical natural process is often the focus of fire-related studies (Landres et al. 1999; Allen et al. 2002), and many statistical methods may be applicable. In this article we give a general introduction to fire ecology for statisticians, summarize application of survival analysis to fire frequency data, and use an unusually complete data set to explore how sensitive parameter estimates are to choices about sampling and data generation from mapped fire histories.

Natural processes produce, and are in turn affected by, landscape patterns (Turner 1989), and fire is a prime example. Fire dynamics are complex, exhibiting varying patterns at different scales of space and time (e.g., Turner et al. 1993; Moritz et al. 2005), leading to substantial challenges for statistical analysis. Fire regime (cf., Gill 1975) characteristics that are often quantitatively estimated include fire frequency, seasonality, spatial extent, and intensity (i.e., heat release per unit time) (Gill 1975; Romme 1980; Heinselman 1981; Bond and van Wilgen 1996). A distinction is usually made between regimes of frequent low-intensity surface fires, which consume leaf litter and other biomass on or near the ground, and infrequent high-intensity crown fires that burn or kill the majority of dominant plants. This difference may be captured through the notion of low and high "severity," referring to the ecological effects of a fire, as opposed to physical measurements related to the fire itself. Severity can be a misleading term, however, as high-intensity crown fires are natural and ecologically beneficial events in many ecosystems, and a low-intensity event could have negative ecological consequences in these plant communities. Variation in fire regimes in space and time has been termed *pyrodiversity* (Martin and Sapsis 1992), and this natural heterogeneity is important for maintenance of biodiversity in many fire-prone ecosystems (e.g., Allen et al. 2002). One would therefore like to estimate probability distributions of all fire regime parameters over a long period of record, allowing characterization of each and how it may covary with others (e.g., season and spatial extent).

Of the statistical methods used in fire-related analyses, collectively termed *pyrostatistics* here, survival analysis (Smith 2002; Lawless 2003) of fire interval data has become one of the most common. Survival analysis is typically called *fire frequency analysis* in the fire science literature, where fire events are considered "deaths" and fire intervals represent survival times (Johnson and Gutsell 1994). Much fire research has focused on a central tendency of fire intervals, most commonly the average fire interval or period between successive fires. This emphasis may reflect an assumption that the chance of burning (the hazard function) increases with time since the last fire, which may correspond to the age of the forest stand, so that less frequent fires will result in greater fuel accumulation, higher fire intensities, and possibly larger spatial extents burned. However this assumption is not necessarily true, particularly for ecosystems that naturally experience high-intensity fires during extreme fire weather episodes or other climatic anomalies, so that chance of burning is not primarily driven by time since the last fire (e.g., Johnson 1992; Moritz 2003; Schoennagel et al. 2004). In these crown fire ecosystems, there is a basic tradeoff in the relative importance of fire regime controls (Bessie and Johnson 1995; Moritz 2003), meaning that the age and spatial patterns of vegetation are less of a constraint to fire ignition and spread under certain weather and climatic conditions. Many fire history analyses have thus been aimed at understanding the strength (or lack thereof) of fuel age in driving fire hazard, to move beyond single central tendency measures to more complex understanding of variation in fire intervals and heterogeneity in fire interval distributions.

Due to the growing use of survival analysis in fire-related research, coupled with increasing availability of mapped fire atlas data in geographic information systems (GIS), our goal is to review and to examine some of the issues encountered in applying these methods to mapped fire interval data. We start by summarizing the types of data and some historical background of fire frequency analysis. Then we summarize use of the Weibull model for fire frequency analysis and give a case study on fire interval distributions related to the massive Cedar Fire of 2003 in southern California. In particular, we assess possible sensitivities to sampling densities, censoring, and different spatially explicit methods of data generation for estimating fire interval distributions. By raising awareness of these issues, we hope to stimulate discussion and to make a small contribution to fire frequency analysis as a major branch of the pyrostatistics tree.

2 Fire interval data

The most common types of data for fire frequency analyses can be generated from point samples or from area-based map data. Fire interval distributions are often derived from stand age data on static time-since-fire maps, which provide a snapshot of the landscape at a specific time (e.g., Johnson and Larsen 1991; Reed et al. 1998). This approach is most applicable to fire regimes that are dominated by stand-replacing events (e.g., crown fire ecosystems), so that the age of trees is equivalent to the time since the last high intensity fire. In addition, fire interval distributions have been generated directly from fire scars in tree ring data, which can record the time between successive fire events (e.g., Clark 1990; Grissino-Mayer 1999). This approach is most useful in low intensity surface fire ecosystems where enough trees faithfully record fire events to characterize a study area. In some studies investigators have used small sample sizes or sought trees that seemed to provide rich data sets (i.e., old trees with many recorded fire scars), but these approaches have been criticized as possibly biased and non-random samples of the fire record (Johnson and Gutsell 1994; Baker and Ehle 2001). Overlapping mapped fire events in a GIS also allow one to examine the entire population of areas burned at different fire intervals (e.g., Baker 1989; Polakow and Dunne 1999; Moritz 2003), an opportunity that we take advantage of later in this article.

Fire interval data can be complete (i.e., bounded on both ends with known dates), but this is not always the case (Polakow and Dunne 1999). For example, successive fire scars in tree ring data can provide complete and uncensored fire intervals, as can the times between mapped fire events in a GIS. However, data on time since last fire on a static stand age map represent right-censored intervals, as the next fire has not yet occurred. Right-censored intervals are also obtained by the time from the beginning of a known fire history to the first known burn (e.g., first fire scar on a tree or first mapped fire event in a fire atlas), which again gives a (right-censored) lower bound on a fire interval.

3 Fire frequency analysis

Earlier research on fire frequency analysis (e.g., Van Wagner 1978; Johnson 1979; Yarie 1981; Johnson and Van Wagner 1985; Clark 1989, 1990) has established the role of the Weibull distribution. One useful aspect of the Weibull is that it has the exponential distribution as a special case, a non-unique feature that is shared by the Gamma. The exponential distribution has a constant hazard function and thus represents a null hypothesis that hazard of burning is independent of time since last fire, a hypothesis that has led to substantial debate among fire ecologists. The Weibull has also played a role in models of fire as a disturbance agent, based on renewal theory and survival analysis (e.g., Cox 1962; Pielou 1977; Cox and Oakes 1984). Despite the popularity of the Weibull model, and recent software allowing a limited form of fire frequency analysis on tree-ring data (Grissino-Mayer 2001), its use is primarily phenomenological. McCarthy et al. (2001) proposed other models based on theoretical considerations, such as the fact that the Weibull can not represent an asymptotic maximum hazard of burning with age, which may be biologically realistic for some cases; however, they point out that in practice distinguishing between types of distributions will be challenging. Schoenberg et al. (2003) have also provided a non-parametric estimation of the probability of burning versus the age of fuels, in addition to estimates for other variables at a particular location (i.e., temperature, fuel moisture, and precipitation).

3.1 The Weibull model

Several current papers cover the basic application of the Weibull model to fire interval data (Johnson and Gutsell 1994; Grissino-Mayer 1999; Polakow and Dunne 1999; Moritz 2003) so we will give only a brief description of methods here. The Weibull cumulative density function, or probability that a fire interval is less than time span t, is

$$F(t) = 1 - \exp(-(t/b)^c), \tag{1}$$

with $b > 0$, $c > 0$, which has density function

$$f(t) = \frac{ct^{c-1}}{b^c} \exp(-(t/b)^c) \tag{2}$$

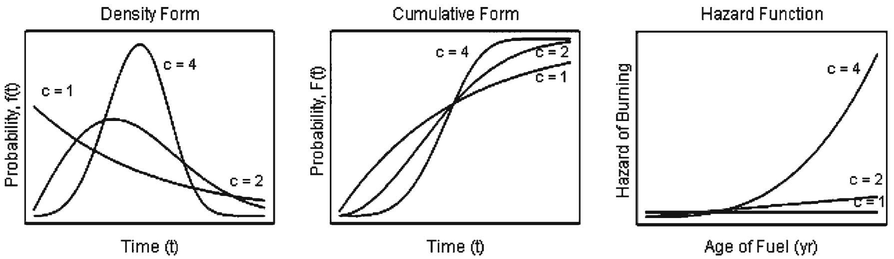

Fig. 1 Fire interval distributions and the Weibull model (parameter b fixed arbitrarily). Higher values of c reflect lower probabilities of burning in younger fuels and the majority of burning in middle-aged fuels; less of the landscape survives beyond the middle-aged "peak" of burning, so probabilities are again reduced for older fuels (left panel). Parameter c also controls the shape of the cumulative form and the hazard function (middle and right panel)

The survival function, $1 - F(t)$, or probability that a fire interval is at least t, is conventionally denoted $A(t)$ in fire frequency analysis (e.g., Johnson and Gutsell 1994). The hazard function, $\lambda(t)$, which gives the instantaneous rate of burning at t given burn interval of at least t, is called the "hazard of burning," partly to distinguish it from the forestry term "burn hazard" (Reed et al. 1998). For the Weibull,

$$\lambda(t) = \frac{f(t)}{A(t)} = \frac{ct^{c-1}}{b^c} \tag{3}$$

Several parameter interpretations are of interest for fire frequency analysis (Fig. 1). The scale parameter, b, gives the 63.2 percentile of fire intervals (i.e., $F(b) = 63.2$) and thus represents a measure of the characteristic time scale of a fire regime. A question of great ecological interest has been whether hazard of burning depends on stand age. The case $c = 1$, corresponding to an exponential distribution (often called "negative exponential" in the fire frequency literature, cf. Johnson and Gutsell 1994) of fire intervals, has constant hazard of burning (i.e., not dependent on stand age). The range $c > 1$ describes increasing hazard of burning over time, with $c = 2$ reflecting linear growth in hazard with time since the last fire. The range $0 < c < 1$ is generally not considered realistic for fire interval distributions, although some vegetation types could conceivably display a decreasing hazard beyond some point in their development (e.g., McCarthy et al. 2001; Odion et al. 2004). An additional metric often used for comparison is the median Weibull fire interval (MEI; Grissino-Mayer 1999; Murthy et al. 2004), given by,

$$\mathrm{MEI} = b(\ln 2)^{(1/c)}$$

It is important to note that when all fire intervals from all (randomly sampled) locations are included in the analysis, then $f(t)$ represents the probability density *over space and time* of fire intervals for a given landscape. This distribution reflects a complex stochastic process of ignition, spread, and regrowth dynamics that can create a mosaic of overlapping burned areas from different fires through time. To address some of the complexities inherent in fire frequency analysis, application of survival time models to fire interval data has seen steady advancement of methods. Early applications did not include censored data in the likelihood function. Polakow and Dunne (1999) incorporated censored data, and Polakow and Dunne (2001) considered the case where all data are censored using recurrence-time ideas. Reed et al. (1998) and Reed (2001) considered change points that are known or estimated, respectively, in fire frequency parameters for time-since-fire data. Their models, and that of Reed and Johnson (2004) for fire-scar intervals, combined a survival time model (exponential) for time-to-ignition with overdispersed binomial models for the number of trees that burn given an ignition; these studies represent steps toward incorporating mechanistic considerations into statistical fire models.

Relatively little work has been done with spatial covariates that may affect hazard of burning in fire frequency analysis, with an early exception by Clark (1990), who used a Cox proportional hazards model for effects of slope and aspect. Some other types of fire studies use shorter, more recent, highly detailed records of observed fires to model probability of ignition as a function of local (e.g., fuel load) and regional (e.g., weather) variables, which requires consideration of spatial covariates and non-independence (Wagner and Fortin 2005). This has been handled, for example, by Chou et al. (1993) by including a neighborhood measure of burned units in the model. Others have included additional variables by including a smooth non-parametric spatial effect on probability of burning (Preisler et al. 2004). Most fire history analyses attempt to control for potential spatial covariation by stratifying sampling of sites by environmental variables (e.g., for tree-ring data) or by constraining the study area to relatively homogeneous regions (e.g., for mapped fire perimeter data). Further incorporation of spatial covariates into fire frequency analysis could add important insights into understanding of fire patterns.

The examples in this article are not intended to extend modeling methods but rather to examine how the most commonly used method—the basic Weibull model—may be sensitive in a real data set to sampling density, inclusion of censored data, and use of historically rich versus only the most recent fire data. Next we summarize the Weibull likelihood function, followed by our sensitivity comparisons.

3.2 Data and Weibull likelihood

Data are typically a combination of uncensored and censored fire intervals. Let T_i be the time for measurement i and δ_i be a censoring indicator equal to 0 if T_i is an uncensored fire interval and 1 if T_i is a (right) censored lower bound on a fire interval, such that fire interval i is at least T_i. Data may be censored for two reasons: (1) if the last time a tree or landscape patch burned is known, and it has not burned again at the time of data collection, then the time-since-burn (T_i) is a right-censored fire-interval; and (2) the fire interval (T_i) from the beginning of the period of record until the first-recorded burn is also right-censored, because it burned previously at an unknown point in time. The case where a measured observation did not burn during the entire study period gives double-censored data (Polakow and Dunne 1999; Lawless 2003), and is not considered here or in our case study.

The factor contributed to the likelihood from datum i is,

$$L_i = \begin{cases} ct^{c-1}b^{-c}\exp\left[-(t/b)^c\right] & \delta_i = 0 \\ \exp\left[-(t/b)^c\right] & \delta_i = 1 \end{cases} \tag{4}$$

giving the total likelihood,

$$L = \prod_{i=1}^{N} L_i \tag{5}$$

The factors in (4) from uncensored data ($\delta_i = 0$) are the Weibull pdf for known values of T_i, while the factors from censored data ($\delta_i = 1$) are probabilities that an interval is at least T_i. This is a standard likelihood for Type 1 right-censored data (Polakow and Dunne 1999; Lawless 2003). Lawless (2003) shows that although the randomness of the censoring process must also be considered, it drops out of the likelihood for estimating the survival time distribution in many cases. It is interesting that for fire intervals, the censoring is related to a fire interval prior to the one being measured, which may lead to consideration of recurrent event models, but these have not been use much to date (but see Polakow and Dunne 2001). The examples here use maximum likelihood estimation of b and c from maximizing (5).

4 The Cedar Fire of 2003

4.1 Study area and fire history

Our study area consisted of ~110,000 ha of land that burned within the perimeter of the Cedar Fire of 2003, east of the city of San Diego (Fig. 2), which is dominated by chap-

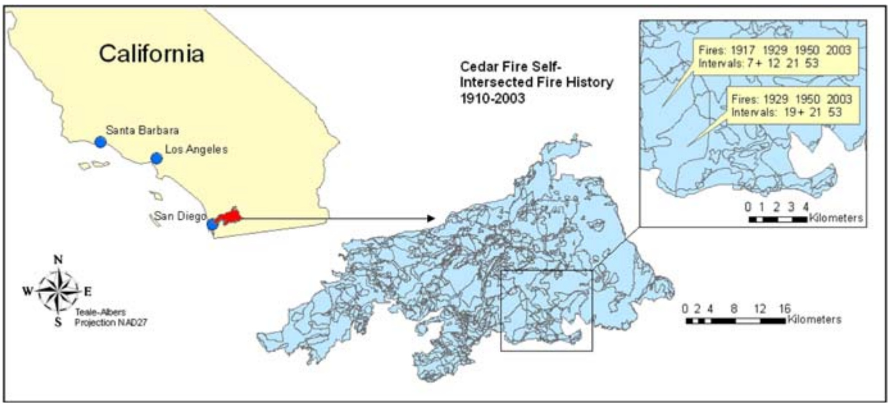

Fig. 2 Study area map and the Cedar Fire data. All past overlapping fire polygons were overlaid and intersected to create a single layer of smaller polygons, each with its own fire history (see inset example). Fire interval distributions could then be generated by using all historical events, or only just the most recent areas burned in 2003

arral shrubland species. The climate for this area is generally of Mediterranean type and characterized by warm, dry summers and mild, wet winters. Continental air masses influence climate more in the interior uplands, while maritime climates dominate near the coast, with a gradient of influence between. Elevation within the fire perimeter ranges from 40 m to 1,979 m. Between the months of September and November this area typically experiences several Santa Ana wind events, which are characterized by low relative humidities, high temperatures, and sustained high velocity winds from the east (Schroeder et al. 1964).

We acquired digital data in GIS format for the Cedar Fire, including fire perimeter and progression maps, from the California Department of Forestry and Fire Protection (CDF), Fire Resource and Assessment Program (FRAP). Accuracy of the Cedar Fire perimeter data had reportedly been verified by fire agency personnel who worked on the fire. We also acquired the statewide mapped fire history database through 2004 from FRAP, though only fires prior to and including the Cedar Fire were included in our analysis. The period of record for this database varies across the state, but the area of the Cedar Fire has been reliably mapped since ~1910. Fires in the statewide database have been compiled from federal agencies with minimum size of 10 acres (~4 ha), and from CDF records with a 300 acre (~120 ha) minimum size (CDF, 2004), meaning that the minimum mapping unit for a particular study area may vary. Although this is a source of uncertainty in the data, the smallest fires do not contribute much to the overall area burned, and their inconsistencies and omission are assumed not to have a substantial effect.

4.2 Scenarios examined

4.2.1 How different are the fire interval distributions from all past overlapping fires versus those from only the most recent fire intervals?

In theory, for a large landscape in a stochastic equilibrium such that the distribution of fire intervals across the landscape changes little, one might expect that similar fire interval distributions would be generated from overlapping fires through time, versus using only the latest fire intervals measured via the present age surface. In reality, study areas are limited, and intervals due to a single large event such as the Cedar Fire might be expected to have a different distribution than that generated from a record of all past intervals through time in the same location (Fig. 2). Indeed, Johnson and Gutsell (1994) suggested caution in interpreting time-since-fire data with more than 1/3 of the area burned in a single event. In practice, fire ecologists often have only one type of data or the other. The unusually complete Cedar Fire data allows us to compare how different fire interval distributions would be from historically more complete overlapping event data versus only the most recent intervals burned in 2003.

To make this comparison, all past historical fire perimeter polygons within the boundary of the Cedar Fire extent were self-intersected to create a layer of new, non-overlapping polygons, each "splinter" of the landscape with a unique fire history (Fig. 2). This complex layer formed the basis for two types of data: (1) An age surface dataset (AGESURF), with each polygon reflecting the associated age of vegetation at the time of the Cedar Fire in 2003; and (2) An overlapping fire interval dataset (OVERLAP), with each polygon recording all past fires (and the corresponding fire-free intervals) from the beginning of the period of record up to and including the Cedar Fire. The AGESURF fire interval distribution (Fig. 3a) is analogous to the time-since-fire map often used in studies of boreal forest fire regimes, which are stand-replacing (Johnson and Gutsell 1994). (Note that in this case, the time-since-fire distribution actually equals the age-at-burn distribution, since all areas burned in 2003). The OVERLAP

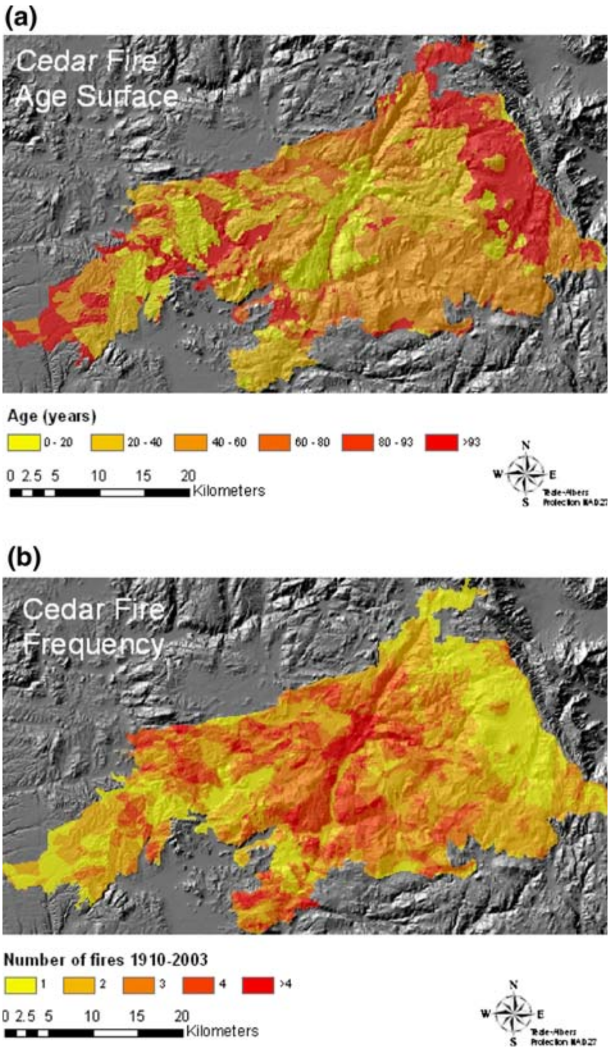

Fig. 3 Mapped fire history patterns within the Cedar Fire of 2003. Panel A represents the age surface of patches burned in 2003, the only data used in generating the AGESURF fire interval distribution. Panel B indicates the number of times burned since 1910 in the same study area, with each successive overlapping fire generating a new set of intervals to add to the OVERLAP distribution

fire interval distribution is representative of datasets used in many other fire regime analyses (e.g., Baker 1989; Grissino-Mayer 1999; Polakow and Dunne 2001; Moritz 2003), where intervals are based on repeated fire events in the past (Fig. 3b).

4.2.2 How different are fire interval distributions when censored data are omitted versus included?

In theory, if a record of complete fire intervals has good coverage of most of the true distribution, then including or omitting censored intervals from the beginning and ends of the data could amount to only a sample size issue; more data should be a bit more informative, but perhaps not to much practical effect if the complete data are a representative sample. In contrast, if the complete intervals undersample the long end of fire intervals, then the only information about long fire intervals may be contained in the censored samples. In that case, omitting the censored samples would lead to marked underestimation of the fire interval distribution.

Many fire frequency analyses have omitted certain intervals. For example, in tree ring-based studies using the common FHX2 tree ring analysis software (Grissino-Mayer 2001), the first interval from pith to fire scar is usually ignored, but this is because it is assumed the tree was not susceptible to scarring during that period. It is also possible to omit the last incomplete (censored) interval in FHX2 (e.g., commonly bounded by the year 1880; Grissino-Mayer 1999). Fire frequency analyses of mapped fire data have also omitted censored observations (e.g., Baker 1989; Moritz 2003), as have some that are based on stand age maps (e.g., Finney 1995). The assumption in omitting censored observations is presumably that they add little information or that the interval data from uncensored observations is representative of the fire regime as a whole. Therefore it is of practical interest to examine how big a difference including or omitting censored data makes, so we analyzed the Cedar Fire data with and without censored data, for both the OVERLAP and AGESURF scenarios.

A substantial portion of the area that burned in the Cedar Fire had not burned in the period of record, resulting in censored observations that were at least 93 years of age (1910–2003; Fig. 3a), so it is reasonable to suspect that censored samples might add important information to the data. These censored observations, which are all relatively old, are relevant to both the AGESURF and OVERLAP data scenarios. The OVERLAP scenario can also include censored observations from the beginning of the known fire record (i.e., 1910) to the first mapped fire in a given location, as well as all subsequent fire (complete) intervals up to the present. Since all data end with the Cedar Fire, there were no doubly censored intervals from the end of the data (i.e., no locations with unknown time of next burn).

4.2.3 How does spatial sampling density affect accuracy of fire interval distributions?

One of the basic practical problems facing fire researchers is the amount of effort required for data collection. In theory, a sparse sampling scheme should lead to more variance in parameter estimates relative to a dense sampling scheme; however, this

Table 1 Scenarios examined in case study

Interval generation methods	Grid spacings (m)	Grid sample locations	Handling of incomplete observations
Age surface, overlap	100, 400, 1,200	Center point, cell majority	Include censored, omit censored

would add little or no bias, assuming the scheme is well-designed. In practice, it can be hard to know a priori how dense a sampling scheme must be to obtain usefully accurate estimates. Given the essentially complete data for the Cedar Fire, we addressed this question by simulating grid sampling schemes with grid spacings of 100, 400, or 1,200 m. From a random point of reference, a grid at each resolution was overlain on the self-intersected fire history, and each cell was assigned either a single interval value (most recent for AGESURF scenarios), or multiple values (all past intervals for OVERLAP scenarios). Additionally, two different grid cell value assignment methods were tested, as there are different possible choices for this process in a GIS. The center location method applies the value at the center of the cell, making each a point observation. The majority method applies the value of the fire interval that covers the most area in the cell. The AGESURF and OVERLAP data sets were created with both methods to investigate possible sensitivities of these models to sample assignment methods.

All combinations of the above interval sampling, censoring, and spatial sampling density scenarios were examined in the case study and are summarized in Table 1. For each combination, Weibull parameters c and b were estimated and MEI was calculated. For the very fine grid resolution of 100 m, the sample was treated as essentially the complete population of fire intervals. For the larger resolutions, confidence intervals (CI) were estimated by bootstrapping for the sampled datasets. For CI generation, the Weibull model was fit repeatedly ($n = 1,000$) to random samples of the original data (with replacement), and 2.5% and 97.5% quantiles were then identified from the resulting fitted parameter values. Although possibly useful for ecological and management interpretation, differences between distributions were not examined through goodness-of-fit tests (e.g., Kolmogorov–Smirnov), because our goal was to assess sensitivities and shifts in individual Weibull parameters under different scenarios.

5 Results

Parameter estimates for the scenarios we examined are shown in Table 2, along with the 95% bootstrapped CIs where appropriate.

5.1 Sensitivity to sampling methods and spacing

Scale parameter b, shape parameter c, and median Weibull fire interval MEI were quite insensitive to the cell assignment method (i.e., center point or area majority) used to

Table 2 Parameter values for all scenarios from the Cedar Fire 2003 dataset. The 100 m resolution constitutes the population of fire events, so parameter estimations do not have CI around them

*	Grid spacing (m)	Grid sampling	Censored observa-tions	c	95% CI for c	b (yr)	95% CI for b	MEI (yr)
Interval generation method: age surface								
A	100	Center	Include	1.49	–	66.12	–	51.75
B	100	Center	Omit	2.14	–	46.72	–	39.37
C	100	Majority	Include	1.50	–	66.11	–	51.75
D	100	Majority	Omit	2.14	–	46.71	–	39.37
E	400	Center	Include	1.50	1.47–1.54	66.02	64.82–67.28	51.71
F	400	Center	Omit	2.16	2.09–2.22	46.68	46.09–47.24	39.38
G	400	Majority	Include	1.50	1.47–1.53	66.05	64.74–67.29	51.73
H	400	Majority	Omit	2.15	2.09–2.21	46.69	46.13–47.33	39.39
I	1,200	Center	Include	1.49	1.39–1.60	67.36	63.93–71.46	52.66
J	1,200	Center	Omit	2.12	1.96–2.30	47.18	45.18–48.90	39.69
K	1,200	Majority	Include	1.56	1.46–1.67	67.14	63.71–71.19	53.06
L	1,200	Majority	Omit	2.30	2.13–2.50	47.32	45.59–48.95	40.35
Interval generation method: overlap								
M	100	Center	Include	1.43	–	60.68	–	46.99
N	100	Center	Omit	1.77	–	38.58	–	31.38
O	100	Majority	Include	1.43	–	60.69	–	46.00
P	100	Majority	Omit	1.78	–	38.58	–	31.39
Q	400	Center	Include	1.46	1.44–1.48	59.00	59.10–60.87	46.66
R	400	Center	Omit	1.78	1.75–1.81	38.55	38.13–39.00	31.38
S	400	Majority	Include	1.45	1.42–1.46	61.05	60.11–61.92	47.34
T	400	Majority	Omit	1.79	1.76–1.82	38.81	38.33–39.28	31.61
U	1,200	Center	Include	1.41	1.35–1.48	61.05	58.56–63.86	47.10
V	1,200	Center	Omit	1.75	1.67–1.83	38.45	37.06–39.84	31.17
W	1,200	Majority	Include	1.47	1.41–1.54	62.86	60.41–65.95	48.98
X	1,200	Majority	Omit	1.86	1.77–1.95	39.73	38.41–41.07	32.60

* *Note:* Scenarios have arbitrarily been assigned letters in this column for ease in referencing them in the text

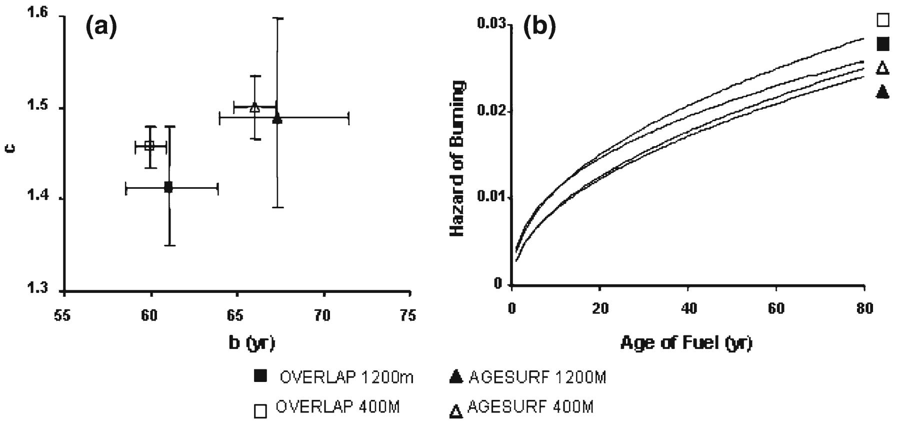

Fig. 4 Scenario comparison for grid spacing (1,200 m and 400 m) and fire interval generation method (age surface and overlapping events). These scenarios are for center sampling with censored observations (i.e., Table 2 rows: E, I, Q, U), primarily demonstrating the relatively small effect of changing sample spacing and the larger effect of interval generation method. A—Weibull parameters. B—Hazard of burning

designate each sample value. For a given treatment of censoring and sample spacing in a scenario, fitted parameter values were often nearly identical between assignment methods, varying by only a few hundredths (e.g., Table 2 rows: A vs. C, R vs. T). This insensitivity is not very surprising at finer scales of resolution (from 100 m to 400 m), as relatively small portions of the landscape that had burned at different intervals are more likely to be captured through finer spatial sampling, regardless of the cell assignment method. The consistency in parameter estimates was also maintained after scaling up to the 1,200 m sampling distance, but with larger confidence intervals reflecting smaller sample sizes. Parameter *c* showed somewhat more sensitivity to cell assignment method than either *b* or MEI (e.g., Table 2 rows: I vs. K), but even these changes in *c* are less than ∼5%.

Similar to our findings on the choice of cell assignment methods, we observed a general insensitivity to the spatial scale of sampling. For example, Fig. 4 demonstrates this insensitivity to sample spacing for two scenario comparisons using 400 m vs. 1,200 m. A consistent result across scenarios, however, was that the coarsest spatial sampling of 1,200 m produced wider CI for both *b* and *c* parameters (e.g., Table 2 rows: E vs. I, S vs. W).

There were more substantial sensitivities and consistent trends observed between scenarios that used fire intervals generated from age surface data as of 2003, in comparison to overlapping fire events up through 2003. Estimates for parameters *b* and *c* and MEI were all consistently lower for the OVERLAP fire interval dataset than for the AGESURF one (i.e., Table 2 rows: E–L vs. Q–X), and this difference was almost always significant according to the estimated 95% CI. Fig. 4 also demonstrates this somewhat greater sensitivity to fire interval generation method than to sample spacing.

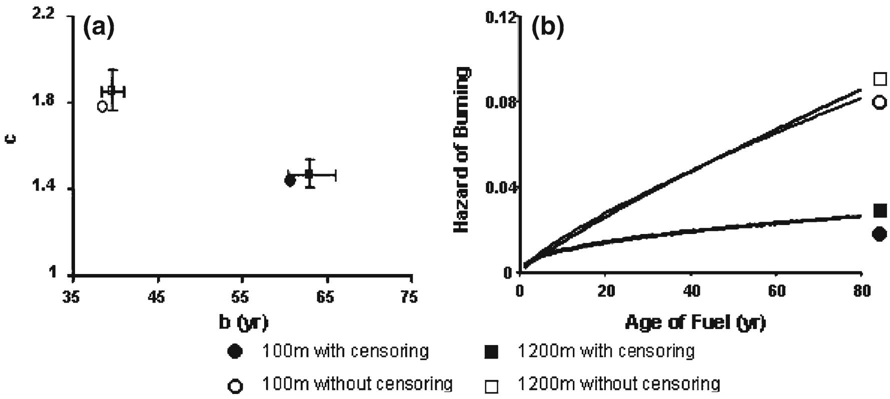

Fig. 5 Comparison of parameter values, *b* and *c*, at the 100 m and 1,200 m grid spacings with and without censored observations. Inclusion of censored observations significantly affected both parameters in all scenarios by lowering *c* and increasing *b*. Notice the insignificant effect of grid spacing here on *b* and *c* (i.e., OVERLAP data with majority sampling method; Table 2: rows O, P, W, X). A—Weibull parameters. B—Hazard of burning

5.2 Sensitivity to inclusion of censored observations

The most striking result of our study was that parameter estimates were highly sensitive to censoring in fire interval distributions. The inclusion of censored observations significantly influenced both Weibull parameters in all scenarios by increasing the scale parameter *b* and lowering the shape parameter *c*; large corresponding increases in MEI were also observed when censored observations were accounted for in parameter estimation (e.g., Table 2 rows: E vs. F). It is interesting to note that inclusion of censored observations consistently made CIs for *b* wider for a given scenario, while making those for *c* consistently narrower. The effect of including censored observations was also more pronounced when fire intervals were generated from the AGESURF method, in comparison to the OVERLAP method (e.g., Table 2 rows: A vs. B compared to M vs. N).

To demonstrate how sensitivity to censoring compares to other factors, consider Figs. 5 and 6, which include the largest increment possible in sample spacing (100–1,200 m) and changes due to fire interval generation method (AGESURF and OVERLAP), respectively. In both cases, parameter shifts after accounting for censoring far outweigh those due to the other factors.

6 Discussion

6.1 Spatial independence and spatial patterns

In our case study, it appears that fine-scale mapped patterns in fire history have minimal influence, regardless of how they are accounted for in sample generation. At increasingly coarse scales of sampling, both the center point and majority area cell assignment methods are likely to filter out smaller patches within each grid cell, smoothing sub-

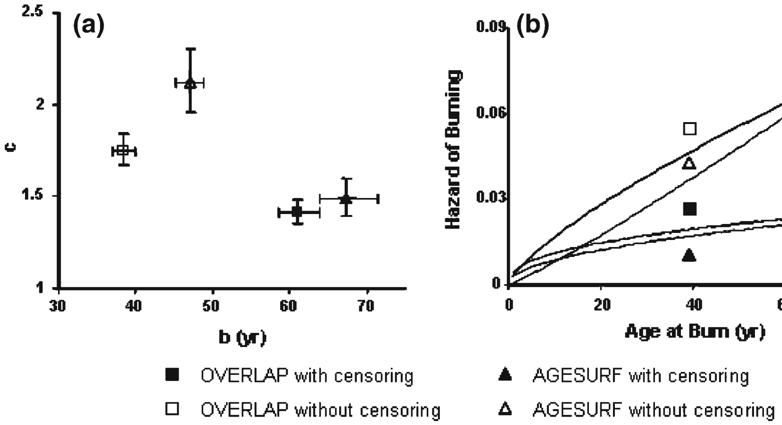

Fig. 6 Comparison of parameter values, *b* and *c*, of age surface and overlapping datasets with and without censored observations. The differences between fire interval distributions generated from age surface data versus overlapping fire data are more pronounced when censoring is not included in the analysis, as shown here in the 1,200 m spacing scenarios (i.e., center point sampling method; Table 2: rows I, J, U, V). A—Weibull parameters. B—Hazard of burning

grid cell heterogeneity in fire interval data; however, the area majority method will always result in the dominant age class, while the center point method may still select the age class of a small patch by chance. Regardless, it would seem that the key relationship of interest here is the characteristic size of mapped fire polygons in relation to the size of sampling units used in an analysis. For example, in the CDF mapped fire history database, some areas are subject to a ∼120 ha minimum mapping unit for fire events to be included. Because this size limit is roughly equivalent to the 1,200 m grid cell spacing, it is possible that the influence of cell assignment method would be more pronounced at spatial sampling distances larger than this already quite coarse resolution. Even so, a sensitivity to the possible inclusion of small patches should have been most obvious for the comparison of OVERLAP versus AGESURF methods for fire interval generation at the coarsest scale. This is because the overlapping event dataset creates and uses many more small “slivers” of the landscape that could contribute to fire interval distributions (Figs. 2 and 3b). Differences between cell assignment methods were relatively small, though, even if close to being outside 95% CIs for *b* and *c* (e.g., Table 2 rows: U vs. W).

The importance of statistically valid sampling designs is clear and well-supported for fire frequency analyses (Johnson and Gutsell 1994; Reed et al. 1998; Reed 2001; Reed and Johnson 2004), but how to deal with the scale of spatial dependence is not. Although sampling design decisions are not as critical when one has the population of fire history mapped, as we do here, spatially independent sampling is a key concern in landscape ecology research. From our case study, one could apparently sample at any spatial scale ranging over an order of magnitude (i.e., 100–1,200 m) and ignore many of the smaller patches on the landscape (i.e., using area majority assignment), with relatively minor impacts on final parameter estimates. We do not know how our results would change with coarser sample spacing or on landscapes with much smaller patch sizes. Reed (2001) also examined a similar issue in case studies of different fire interval data subsets (from a complete mapped population), finding that sensitivities varied by location and typical patch size on the landscape.

Another scale-related question is whether one needs to deliberately account for patches of different sizes that make up the fire history population. Polakow and Dunne (1999) discuss the rationale for reducing landscape patches to point observations and for eliminating patches with identical fire histories as pseudoreplicates. While this seems logical for studies with many observations and relatively similar patch sizes on the landscape, such a sampling approach will result in losses of very large amounts of data in some situations, and it ignores the importance of larger patches in shaping fire interval distributions (e.g., large patches in Fig. 3a). In general, one should not be faced with decisions about eliminating patches of various sizes when one has the mapped population of fire events. This issue does raise questions about what constitutes an observation, if one is resampling data for some reason (e.g., to change spatial scale or to perform bootstrapping), and it highlights an area for future work.

6.2 Ecological interpretation

As Polakow and Dunne (1999) have pointed out, censored observations can contain important information and should be accounted for in parameter estimation. In our case study using data from the 2003 Cedar Fire, we also found that inclusion of censored data had a relatively large influence on results. Unlike Polakow and Dunne (1999), however, we found the shape parameter c to decrease after inclusion of censored observations, regardless of the scenario in question. Sensitivity to censored data was especially pronounced when many very old observations were added (e.g., the long unburned patches on the eastern edge in Fig. 3a) in the age surface method of fire interval generation.

The inclusion of older censored observations increases the estimated typical age at burn, reflected by higher b values. The observed lower c, which drives the rate at which hazard of burning increases with age, implies that the chance of burning in a given year increases more slowly with time since the last fire (e.g., Figs. 5 and 6). In terms of the shape of a fire interval distribution, it demonstrates that adding censored observations can tend to flatten out the density function $f(t)$ (Fig. 1), pushing the distribution closer to $c = 1$ (i.e., age-independent). As has been shown in previous analyses (e.g., Moritz 2003; Moritz et al. 2004), burning through younger age classes can result in lower c values and thus a lower degree of age dependency in a fire regime; we see here that adding older, censored observations can also result in the same interpretation.

6.3 Concluding thoughts

The increasing availability of digital fire atlases with overlapping fire events has the potential to add significantly to our understanding of how fire has operated on the landscape (Morgan et al. 2001). Our case study has shown, however, that there can be substantial sensitivities to details of how fire interval distributions are generated and analyzed. Admittedly, the generality of our findings is not known. Parameter estimates for the area burned by the Cedar Fire of 2003 were in the general range of previous fire frequency analyses on chaparral-dominated shrublands in southern California (Polakow et al. 1999; Moritz et al. 2004), reflecting a less age-dependent dynamic than many have traditionally assumed. Regardless, several questions remain in regard to these statistical methods. If we decide to use incomplete observations in fire frequency analyses, how do we interpret the results? Will current age surfaces adequately characterize the fire frequency for an area, or do overlapping past intervals improve these estimates? There is widespread use of survival analysis in examining fire interval distributions, and our hope here is to have added some insight into the effects of sampling decisions related to spatially explicit fire history data. Because results can be interpreted for ecosystem management and have real implications, a better understanding of inherent sensitivities in these techniques is essential.

Acknowledgements We thank the Joint Fire Sciences Program for funding this work (Project 04-1-2-01). We also appreciate the use of code for MLE fitting from DA Polakow and TT Dunne, as well as comments from FP Schoenberg, ML Gumpertz, and two anonymous reviewers on the manuscript.

References

Allen CD, Savage M, Falk DA, Suckling KF, Swetnam TW, Schulke T, Stacey PB, Morgan P, Hoffman M, Klingel JT (2002) Ecological restoration of southwestern ponderosa pine ecosystems: a broad perspective. Ecol Appl 12:1418–1433

Baker WL (1989) Effect of scale and spatial heterogeneity on fire interval distributions. Can J Forest Res 19:700–706

Baker WL, Ehle D (2001) Uncertainty in surface-fire history: the case of ponderosa pine forests in the western United States. Can J Forest Res 31:1205–1226

Bessie WC, Johnson EA (1995) The relative importance of fuels and weather on fire behavior in subalpine forests. Ecology 76:747–762

Bond WJ, van Wilgen BW (1996) Fire and plants. Chapman & Hall, London

California Department of Forestry and Fire Protection (CDF) (2004) Mapped fire history database for California; public version and detailed methodology. http://frap.cdf.ca.gov/projects/fire_data/fire_perimeters/. Accessed Dec 2007

Chou YH, Minnich RA, Chase RA (1993) Mapping probability of fire occurrence in San Jacinto Mountains, California, USA. Environ Manage 17:129–140

Clark JS (1989) Ecological disturbance as a renewal process: theory and application to fire history. Oikos 56:17–30

Clark JS (1990) Fire and climate change during the last 750 years in northwestern Minnesota. Ecol Monogr 72:1102–1118

Cox DR (1962) Renewal theory. Wiley, London

Cox DR, Oakes D (1984) Analysis of survival data. Chapman & Hall, London

Finney MA (1995) The missing tail and other considerations for the use of fire history models. Int J Wildland Fire 5:197–202

Gill AM (1975) Fire and the Australian flora: a review. Aust For 38:4–25

Grissino-Mayer HD (2001) FHX2: Software for the analysis of fire history from tree rings. Tree Ring Res 57:113–122

Grissino-Mayer HD (1999) Modelling fire interval data from the American Southwest with the Weibull distribution. Int J Wildland Fire 9:37–50

Heinselman ML (1981) Fire intensity and frequency as factors in the distribution and structure of northern ecosystems. In: Mooney HA et al. (eds) Proceedings of the conference, fire regimes and ecosystem dynamics, US Forest Service General Technical Report WO-26, pp 7–57

Johnson EA (1979) Fire recurrence in the subarctic and its implications for vegetation composition. Can J Botany 57:1374–1379

Johnson EA (1992) Fire and vegetation dynamics: studies from the North American boreal forest. Cambridge University Press, Cambridge
Johnson EA, Gutsell SL (1994) Fire Frequency Models, Methods and Interpretations. Adv Ecol Res 25:239–287
Johnson EA, Larsen CPS (1991) Climactically induced change in fire frequency in the southern Canadian Rockies. Ecology 72:194–201
Johnson EA, Wagner CEvan (1985) The theory and use of two fire history models. Can J Forest Res 15:214–219
Landres PB, Morgan P, Swanson FJ (1999) Overview of the use of natural variability concepts in managing ecological systems. Ecol Appl 9:1179–1188
Lawless JF (2003) Statistical Models and Methods for Lifetime Data. Wiley, Hoboken
Martin RE, Sapsis DB (1992) Fires as agents of biodiversity: Pyrodiversity promotes biodiversity. In: Proceedings of the Symposium on Biodiversity of Northern California, Univ. California, Berkeley, pp 150–157
McCarthy MA, Gill AM, Bradstock RA (2001) Theoretical fire-interval distributions. Int J Wildland Fire 10:73–77
Morgan P, Hardy CC, Swetnam TW, Rollins MG, Long DG (2001) Mapping fire regimes across space and time: understanding coarse and fine-scale patterns. Int J Wildland Fire 10:329–342
Moritz MA (2003) Spatiotemporal analysis of controls on shrubland fire regimes: age dependency and fire hazard. 2003. Ecology 84:351–361
Moritz MA, Keeley JE, Johnson EA, Schaffner AA (2004) Testing a basic assumption of shrubland fire management: how important is fuel age?. Front Ecol Environ 2:67–72
Moritz MA, Morais ME, Summerell LA, Carlson JM, Doyle J (2005) Wildfires, complexity and highly optimized tolerance. In: Proceedings of the national academy of sciences USA, vol 102, pp 17912–17917
Murthy DNP, Xie M, Jiang R (2004) Weibull Models. Wiley, Hoboken
Odion DC, Strittholt JR, Jiang H, Frost EJ, DellaSala DA, Moritz MA (2004) Patterns of fire severity and forest conditions in the Klamath Mountains, Northwestern California, USA. Conserv Biol 18:927—936
Pielou EC (1977) Mathematical Ecology. Wiley, New York
Polakow DA, Dunne TT (1999) Modelling fire-return interval T: stochasticity and censoring in the two-parameter Weibull model. Ecol Model 121:79–102
Polakow DA, Dunne TT (2001) Numerical recipes for disaster: changing hazard and the stand-origin-map. Forest Ecol Manage 147:183–196
Polakow D, Bond W, Lindenberg N, Dunne T (1999) Ecosystem engineering as a consequence of natural selection: methods for testing Mutch's hypothesis from a comparative study of fire hazard rates. In: Proceedings of Australian bushfire conference. http://www.csu.edu.au/special/bushfire99/papers/polakow/. Accessed Dec 2007
Preisler HK, Brillinger DR, Burgan RE, Benoit JW (2004) Probability-based models for estimation of wildire risk. Int J Wildland Fire 13:133–142
Reed WJ (2001) Statistical inference for historical fire frequency using the spatial mosaic. In: Johnson EA, Miyanishi K (eds) Forest fires: behavior and ecological effects. Academic Press, San Diego
Reed WJ, Johnson EA (2004) Statistical methods for estimating historical fire frequency from multiple fire-scar data. Can J Forest Res 34:2306–2313
Reed WJ, Larsen CPS, Johnson EA, MacDonald GM (1998) Estimation of temporal variations in historical fire frequency from time-since-fire map data. Forest Sci 44:465–475
Romme W (1980) Fire history terminology: report of the ad hoc committee. In Stokes MA, Dieterich JH (eds) Proceedings of the fire history workshop. US Forest Service General Technical Report RM-81, pp 135–137
Smith PJ (2002) Analysis of failure and survival data. Chapman & Hall/CRC Press, Boca Raton
Schoenberg FP, Peng R, Huang Z, Rundel P (2003) Detection of non-linearities in the dependence of burn area on fuel age and climatic variables. Int J Wildland Fire 12:1–6
Schoennagel T, Veblen TT, Romme WH (2004) The interaction of fire, fuels, and climate across rocky mountain forests. Bioscience 54:661–676
Schroeder MJ et al. (1964) Synoptic weather types associated with critical fire weather. A.D. 449–630. US Department of Commerce, National Bureau of Standards, Institute for Applied Technology, Washington D.C., USA

Turner MG (1989) Landscape Ecology: The Effect of Pattern on Process. Annu Rev Ecol Syst 20:171–197
Turner MG, Romme WH, Gardner RH, O'Neill RV, Kratz TK (1993) A revised concept of landscape equilibrium: disturbance and stability on scaled landscapes. Landsc Ecol 8:213–227
Wagner CEvan (1978) Age-class distribution and the forest fire cycle. Can J Forest Res 8:220–227
Yarie J (1981) Forest fire cycles and life tables: a case study from interior Alaska. Can J Forest Res 11:554–562
Wagner HH, Fortin MJ (2005) Spatial analysis of landscapes: concepts and statistics. Ecology 86:1975–1987

Author Biographies

Max A. Moritz received his Ph.D. in Biogeography and Spatial Ecology from the University of California, Santa Barbara in 1999. He is currently the University of California Cooperative Extension Specialist in Wildfire, as well as Adjunct Assistant Professor at the University of California, Berkeley. His interests include the statistical analysis of ecological patterns and processes, especially those relating to natural disturbances.

Tadashi J. Moody received his M.S. in 2005 in the Environmental Science, Policy, and Management Department at the University of California, Berkeley. His interests include the fire regime of mixed conifer forests in the Sierra Nevada mountains of California and methods for fire history analysis in general.

Lori J. Miles received her B.S. in Botany and Biology from California State University, Humboldt in 2004. She is currently pursuing a M.S. in the School of Life Sciences at the University of Nevada, Las Vegas; her work is on invisibility patterns in extreme environments and rare plant communities.

Matthew M. Smith received his B.S. in horticulture from the University of Connecticut in 1997. His interests include the use of web-based technologies in conservation, and he is currently leading a long-term study on the role of plants and animals in everyday life.

Perry de Valpine received his Ph.D. in Ecology from the University of California, Davis, in 2000. He is currently an Assistant Professor at the University of California, Berkeley. His research interests in Ecological Modeling and Statistics include population dynamics, time-series analysis, state-space and mixture models, and computational statistical methods.