ESTIMATION OF TEMPORAL VARIATIONS IN FIRE FREQUENCY FROM TIME-SINCE-FIRE DATA

W.J. REED, C.P.S. LARSEN, E.A. JOHNSON
AND G.M. MACDONALD

DMS-738-IR

June 1996

ESTIMATION OF TEMPORAL VARIATIONS IN FIRE FREQUENCY FROM TIME-SINCE-FIRE DATA

W.J. Reed[1], C.P.S. Larsen[2], E.A. Johnson[3] and G.M. MacDonald[4]

[1]Department of Mathematics, University of Victoria, Victoria BC, V8W 2Y2, Canada

[2]Department of Geography, SUNY Buffalo, Buffalo NY, 14261-0023, USA

[3]Division of Ecology, Department of Biological Sciences and Kananaskis Field Stations, University of Calgary, Calgary AB, T2N 1N4, Canada

[4]Department of Geography, University of California, Los Angeles, Los Angeles CA, 90024-1524, USA

ABSTRACT

This paper introduces a method to estimate the fire frequency in a number of different temporal periods (epochs) and to assess whether these fire frequencies differ significantly. The method is applicable to time-since-fire maps, and unlike previous methods, includes the different hazards of burning to which a unit has been subject since its establishment. Maximum likelihood estimates of the fire frequencies are obtained, and likelihood ratio methods are used to obtain confidence intervals and to test the significance of postulated change points. The test is not strictly valid for change points suggested from exploratory data analysis and assumes that a change point to be tested is selected without prior reference to the data. The model used assumes that the hazard of burning does not depend on age or location, although it includes a contagion effect modelled by the inclusion of an overdispersion parameter, and the construction of a quasi log-likelihood function.

The method is applied to published fire history data from the Kananaskis Valley. The results indicate that, as was to be expected, previous analyses have overestimated the fire frequency for the most recent epoch.

Keywords: change point, disturbance, fire cycle, fire frequency, fire history, quasi-likelihood, maximum likelihood, overdispersion, contagion effect.

INTRODUCTION

This paper develops a statistical method to estimate the fire frequency in different time periods (epochs) and to assess whether the fire frequencies are significantly different. The method assumes that older forests have been subject to the changing hazard of burning in each subsequent epoch. This is in contrast to previous methods (*e.g.* Johnson and Larsen 1991) which assumed that older forests were subject only to the hazard of burning during the epoch in which they originated and had been immune to the hazards in subsequent epochs. Note that we use the term 'hazard of burning' (Johnson & Gutsell, 1994) for what statisticians would call the 'hazard rate' which, loosely speaking, is the per annum probability of burning conditional on previous survival (see Reed, 1994 for precise definition). The hazard of burning should not be confused with the forestry term "fire hazard" which is the potential fire behaviour based only on the fuel structure but not the fuel moisture. It is equivalent to the fire frequency, which is a forestry term. The inverse of the hazard of burning is the fire cycle, which is the number of years required to burn an area equal in size to the study area.

The data used in this method are time-since-fire maps of a complete study area. A time-since-fire map consists of units of vegetation which are demarcated by the last year in which they burned. Each unit is labelled with the time-since-fire. The creation and use of time-since-fire maps was introduced by Heinselman (1973).

Two general conclusions from Heinselman's (1973) study have been supported by subsequent fire history studies. The first conclusion is so obvious today that we tend to forget its importance, namely that no unit in most northern coniferous forests has survived for long without burning. In Heinselman's (1973) study, every unit had burned at least once during the last 378 years. This finding has had serious implications for models of vegetation succession which require long periods with no disturbance during which competitive exclusion could occur (*e.g.* Connell and Slatyer 1977).

The second general conclusion of Heinselman's (1973) research was that the empirical distribution of the forest area in each time-since-fire age-class (Fig. 1 upper panel) could be used to estimate the survivorship from forest fires and hence the historical hazard of burning. This conclusion has since been formally situated within the statistical theory of survival analysis

(Johnson and Van Wagner 1985, Johnson and Gutsell 1994, Reed 1994) and is developed further in this paper.

Heinselman (1973) also speculated that the population of units which made up his study area could be subdivided into subpopulations, each of which could have different hazards of burning. For example, Heinselman (1973) suggested variously that topographic location, vegetation age, vegetation composition, and past periods of differing land-use might affect the hazard of burning. Most studies have implicitly considered the overall survivorship distribution to be the superposition of the survivorship distributions of these subpopulations with different fire hazards. This overall survivorship can be termed a mixed or mixture survival distribution. A number of studies (McCune 1983, Clark 1990, Johnson *et al.* 1990, Masters 1990, Bergeron 1991, Johnson and Larsen 1991) have found mixed survival distributions to be due to temporal changes in hazard rates. In this paper we will only discuss issues regarding temporal variations in survivorship rates.

The separation of these temporal-component survival distributions to estimate their respective hazards of burning has been problematic. The separation technique should provide methods to determine the number and location of change points in the hazard of burning and methods to estimate parameters of each component distribution. In addition, the technique should be based on an underlying ecological model. Note that the time period between significant change points will be referred to as an epoch.

Change points have typically been identified using graphical analysis of the survival distribution (*e.g.* Heinselman 1973, Van Wagner 1978, Masters 1990, Bergeron 1991, Johnson and Larsen 1991). A change point is suggested at the point of intersection of two graphically identified line segments fitted to the semilog plot of the survivorship (Fig. 2). Two different methods have been used to estimate the hazard of burning in different epochs: a graphical separation of the component distributions followed by a maximum likelihood estimation (MLE) of the hazard of burning (*e.g.* Johnson *et al.* 1990, Johnson and Larsen 1991), or estimation of the hazard of burning directly from the mixed survival distribution using linear regression (*e.g.* Masters 1990, Bergeron 1991).

The graphical technique used by Johnson and Larsen (1991) is based on a method originally proposed by Kao (1959). Once a change point has been identified on the survival distribution, the mixture is partitioned by assuming that each epoch represents the total population (see Johnson and Gutsell 1994 for details). Statistically, this partitioning is equivalent to the assumption that older forests are not subject to the hazard of burning that created the younger forests. An ecological reason for this argument would be that some factor (*e.g.* fuel, forest structure, microclimate) changes with forest age and leads to older forest being more likely to burn than younger forest (*e.g.* Heinselman 1973, Romme 1982, Minnich 1983, Renkin and Despain 1992). Recent work has shown that in some situations, forest age is far less important than climate in affecting flammability (Bessie and Johnson 1995).

The estimation of hazards of burning directly from the different portions of the survival distribution using least-squares linear regression (*e.g.* Masters 1990, Bergeron 1991) is a purely *ad hoc* technique and of questionable statistical validity. For example neither of the assumptions of independence and homoscedasticity, required for least-squares regression, are satisfied for cumulative survival distribution data.

In this paper we develop a statistical model which recognizes that older forest units have successively been subject to the hazard of burning in each subsequent epoch. Initially it is assumed that each unit burns or survives independently of other units, but the model is then extended to take account of the fact that fire is a contagious process *i.e.* the probability of a unit burning is increased by adjacent units burning. Using the model we first obtain maximum likelihood estimates of the hazards of burning in different epochs. Secondly we present likelihood ratio methods for obtaining confidence intervals. Thirdly we show how to test for differences in the hazard of burning between epochs, assuming that the postulated change point was not determined from exploratory data analysis. We do not present a method for determining the change points. This is a complicated problem in the statistical field of "model selection" and we defer discussion of this problem to a subsequent paper.

Data

The data required for the method is a complete inventory of the study area (time-since-fire map). More precisely, on a time-since-fire map the time-since-last-fire is identified everywhere within the region studied and the data are *areas* belonging to disjoint classes defined by their time-since-fire, *i.e.* years since the last fire. The ecological meaning of "time-since-fire" depends on the objective of the study (*e.g.* canopy replacing fire, surface fire, etc.). If the hazard of burning has been a constant, λ, over all times and locations, the distribution of areas in age-classes will be related to the negative exponential survival distribution (see Johnson & Van Wagner, 1985) with survivor function

$$S(t) = \exp(-\lambda t) \tag{1}$$

In fact, under the assumption of units burning independently of one another, the distribution of areas in age-classes will follow a multinomial distribution (see Reed, 1994); if contagion effects are present the distribution of areas in age-classes will be overdispersed relative to a multinomial distribution. This is further discussed later in the paper. First, however, we consider the modelling of temporally distinct hazards of burning.

Temporally Distinct Hazards

Here we will consider the situation in which there are several (N) change points in the fire frequency distribution and consequently $N + 1$ distinct hazards of burning. These periods of constant hazard will be called epochs. We change the name of the periods from components (*cf.* Johnson & Van Wagner 1985) to epochs because we assume the study has already shown that the mixed distribution is due to a temporal change in hazard, not a spatial change in hazard.

The important ecological assumptions that the model makes are as follows. The first assumption is that the change in the hazard of burning is rapid between two distinct epochs. While the literal validity of this may be questioned, it nonetheless provides a convenient approximation to the situation where the change in hazard is short relative to the length of epochs.

The second assumption is that the areas which originated from fire some years ago have been subject to every subsequent hazard. Suppose changes in the hazard of burning occurred P_1 and P_2 years ago where $P_1 < P_2$ and that the hazards were λ_1 between the present and P_1 years ago,

λ_2 between P_1 and P_2 years ago and λ_3 more than P_2 years ago. These periods of constant hazard of burning will be called respectively the youngest epoch, middle epoch, and oldest epoch. A unit whose age puts it in the oldest epoch will be subject to the hazard of burning during the oldest epoch and in turn to the hazards of the middle and most recent epochs as it passes through those periods. Further, a unit in the middle epoch would be subject sequentially to the hazards of burning of the middle and most recent epochs. Finally, a unit in the youngest epoch would be subject only to the hazard of burning of the youngest epoch.

Thus we assume that the overall hazard of burning $\lambda(t)$ is a piecewise constant function, constant over each epoch. The survivor function is then:

$$S(t) = \begin{cases} \exp[-\lambda_1 t], & 0 \le t < P_1 \\ \exp[-\lambda_1 P_1 - \lambda_2(t-P_1)], & P_1 \le t < P_2 \\ \exp[-\lambda_1 P_1 - \lambda_2(P_2-P_1) - \lambda_3(t-P_2)], & P_2 \le t \end{cases} \tag{2}$$

Equation (2) explicitly reflects the two assumptions: instantaneous change points and the pattern of hazards. How the change points work is obvious in equation (2); but let us explain how the pattern of hazards works. The survivorship of units of age $0 \le t < P_1$ will only have been subject to hazard of burning λ_1. The survivorship of units of age $P_1 \le t < P_2$ will have been subject to hazard of burning λ_2 during the interval t to P_1 years ago and also the hazard of burning λ_1 during the interval from P_1 years ago to the present. Similarly, the survivorship for units of age $P_2 \le t$ will have been subject to hazard of burning λ_3 (during the interval t to P_2 years ago), λ_2 (during the interval P_2 to P_1 years ago) and λ_1 (from P_1 years ago to the present).

The Likelihood Function for Parameter Estimation

We consider now how the various hazards of burning in distinct epochs can be estimated from data on areas of forest in different age-classes. Age classes are used because time-since-fire data are not always accurate to the nearest year. Each age class will have length T year. The forest area (A_j) in age class j ($j = 1,2,\cdots,m-1$) will have originated between $(j\text{-}1)T$ and $jT\text{-}1$ years ago. Age-class m (with area A_m) will be a "collector" class comprising areas of age $(m\text{-}1)T$ years or greater.

We develop a likelihood function for estimating the epochal hazards of burning, firstly under the assumption that stands burn or survive independently of one another. This is mainly for expositional purposes. The more realistic situation, when there is a contagion effect in burning, can be dealt with using a fairly simple extension of this basic model.

Initially let us adopt the position that every time period (of length T years) has its own distinct hazard of burning which we will denote by $\lambda^{(j)}$ for time period $j = 1,2,\cdots,m$. Thus for period $j = 1$ (between time 0 and T-1 years ago), the hazard of burning is $\lambda^{(1)}$; for period $j = 2$ (between T years ago and $2T$-1 years ago), the hazard is $\lambda^{(2)}$; while during period $j = m$ (more than $(m$-1$)T$ years ago), the hazard is $\lambda^{(m)}$.

Now, following from equation (2), let

$$S^{(j)} = \exp(-\lambda^{(j)}T) \qquad j = 1,2,\cdots,m-1 \tag{3}$$

represent the conditional probability that an area which having survived unburned to the time jT years ago, survives the next T years *i.e.* to the time $(j$-1$)T$ years ago.

Under the assumption that units burn or survive unburnt independently of one another the probability of finding the observed areas in the various age classes would be a multinomial probability, with corresponding likelihood function:

$$L \propto [S^{(m-1)}S^{(m-2)}\ldots S^{(1)}]^{Am} \prod_{j=1}^{m-1} [(1-S^{(j)})S^{(j-1)}\ldots S^{(1)}]^{Aj} \tag{4}$$

or

$$L \propto \left[(S^{(1)})^{\Sigma_{i=2}^{m}A_i}(1-S^{(1)})^{A_1}\right]\left[(S^{(2)})^{\Sigma_{i=3}^{m}A_i}(1-S^{(2)})^{A_2}\right]\ldots\ldots\left[(S^{(m-1)})^{A_m}(1-S^{(m-1)})^{A_{m-1}}\right] \tag{5}$$

Now assume that $S^{(1)},S^{(2)},\cdots,S^{(m)}$ are to be expressed in terms of other parameters, in our case the hazards of burning for distinct epochs. The likelihood in equation (5) would then be a function of these other parameters. For example, if we assume that a single hazard of burning prevailed during the lifetime of all surviving units ($\lambda^{(1)} = \lambda^{(2)} = ... = \lambda^{(m)}$) we would have conditional per period survival probabilities of, say

$$S^{(1)} = S^{(2)} = ... = S^{(m)} = \exp(-\lambda_0 T) = S_0, \tag{6}$$

and likelihood function

$$L(S_0) = S_0^{\Sigma_{j=1}^{m-1} B_j} (1 - S_0)^{\Sigma_{j=1}^{m-1} A_j} \tag{7}$$

where

$$B_j = \sum_{i=j+1}^{m} A_i$$

is the *cumulative* area of forest older than jT years. If there is a change point between pT - 1 and pT years ago dividing the past into two epochs and if the hazards of burning are constant within each epoch, then

$$\begin{aligned} S^{(1)} = S^{(2)} = ... = S^{(p)} = e^{-\lambda_1 T} = S_1 \text{ say} \\ \text{and } S^{(p+1)} = S^{(p+2)} = ... = S^{(m)} = e^{-\lambda_2 T} = S_2 \text{ say} \end{aligned} \tag{8}$$

and the likelihood is:

$$L(S_1, S_2) = S_1^{\Sigma_{j=1}^{p} B_j} (1 - S_1)^{\Sigma_{j=1}^{p} A_j} S_2^{\Sigma_{j=p+1}^{m-1} B_j} (1 - S_2)^{\Sigma_{j=p+1}^{m-1} A_j} \tag{9}$$

In either case maximum likelihood estimates of S_0 (and hence λ_0) or of S_1 and S_2 (and hence λ_1 and λ_2) can be found as the values that maximize the corresponding likelihood ((7) or (9)). However we do not consider details of the maximization here, but rather defer it to the next section in which a more general overdispersed model (and quasi-likelihood function) corresponding to a contagion effect in fire probability is developed.

An Overdispersed Model for the Contagion Effect

Equation (9) is based on the assumption that units burn or survive independently of one another. Forest fires of course burn continuously over an area, resulting in forest units that are adjacent to ones that are burning having a higher hazard of burning than other units which are far from those which are burning. Thus there is a contagion effect. Reed (1994) has modeled this for fire frequency using the Dirichlet distribution. Here we adopt a simpler approach using a quasi-likelihood function for an overdispersed distribution (*e.g.* McCullah and Nelder 1989).

This method is based on the observation that, while the contagion effect will change the distribution of areas from a multinomial form, it should not change the expected values. Contagion will inflate the variances and covariances (*overdispersion*) but since one can reasonably assume that the inflation factor is the same for all variances and covariances, the overdispersion can be incorporated by introducing a scalar *overdispersion parameter* to multiply the covariance matrix. This enables the specification of the first and second order terms of the log-likelihood function in terms of the model parameters and the overdispersion parameter. Since most of the asymptotic theory connected with the likelihood is based only on these terms, it is sufficient to treat them as if they constituted a full log-likelihood. Such an 'approximate' log-likelihood is referred to as a *quasi-likelihood* (see McCullagh and Nelder 1989, Ch. 9). This approximation should be good for "large samples", *i.e.* when a large study area is being considered.

Thus by using a quasi-likelihood approach one can easily include a contagion effect without being too explicit about the detailed mechanism through which it operates. This is convenient in the current context since it finesses the need to develop a complex stochastic spatial model for the spread of fires.

On taking logarithms, the quasi-likelihood corresponding to the general form (5) is

$$Q = \frac{1}{\sigma^2}[B_1 \ln S^{(1)} + A_1 \ln(1 - S^{(1)}) \\ + B_2 \ln S^{(2)} + A_2 \ln(1 - S^{(2)}) \\ + B_{m-1} \ln S^{(m-1)} + A_{m-1} \ln(1 - S^{(m-1)})] \tag{10}$$

where σ^2 is the over-dispersion parameter. If there is a constant hazard rate, the quasi-likelihood corresponding to (7) is:

$$Q = \frac{1}{\sigma^2}\left[\left(\sum_{j=1}^{m-1} B_j\right) \ln S_0 + \left(\sum_{j=1}^{m-1} A_j\right) \ln(1 - S_0)\right] \tag{11}$$

If there is one change point and two hazards, the quasi-likelihood corresponding to (9) is:

$$Q = \frac{1}{\sigma^2}\left[\left(\sum_{j=1}^{p} B_j\right) \ln S_1 + \left(\sum_{j=1}^{p} A_j\right) \ln(1 - S_1) \\ + \left(\sum_{j=p+1}^{m-1} B_j\right) \ln S_2 + \left(\sum_{j=p+1}^{m-1} A_j\right) \ln(1 - S_2)\right] \tag{12}$$

Maximum Likelihood Estimates

Maximum likelihood estimates (MLE) of parameters can be found in the usual way, *i.e.* treating the quasi-likelihood as an ordinary log-likelihood. The point estimates will be exactly the same as they would be for the multinomial likelihood with no contagion. Thus if a single homogeneous hazard of burning is assumed, the MLE of the corresponding survival probability is (from differentiating (11)):

$$\hat{S}_0 = \frac{\sum_{j=1}^{m-1} B_j}{\sum_{j=1}^{m-1} (B_j + A_j)} \\ = \frac{\sum_{i=2}^{m} A_i + \sum_{i=3}^{m} A_i + \dots + \sum_{i=m}^{m} A_i}{\sum_{i=1}^{m} A_i + \sum_{i=2}^{m} A_i + \dots + \sum_{i=m-1}^{m} A_i} \tag{13}$$

The MLE of λ_0 is simply

$$\hat{\lambda}_0 = -\frac{1}{T}\ln \hat{S}_0. \tag{14}$$

Similarly if a single change point is assumed, the MLE's of the two survival probabilities are (from differentiating (12)):

$$\hat{S}_1 = \frac{\sum_{j=1}^{p} B_j}{\sum_{j=1}^{p} (B_j + A_j)} = \frac{\sum_{i=2}^{m} A_i + \sum_{i=3}^{m} A_i + \dots + \sum_{i=p+1}^{m} A_i}{\sum_{i=1}^{m} A_i + \sum_{i=2}^{m} A_i + \dots + \sum_{i=p}^{m} A_i} \tag{15}$$

and

$$\hat{S}_2 = \frac{\sum_{j=p+1}^{m-1} B_j}{\sum_{j=p+1}^{m-1} (B_j + A_j)} = \frac{\sum_{i=p+2}^{m} A_i + \sum_{i=p+3}^{m} A_i + \dots + \sum_{i=m}^{m} A_i}{\sum_{i=p+1} A_i + \sum_{i=p+2}^{m} A_i + \dots + \sum_{i=m-1}^{m} A_i} \tag{16}$$

The estimates (13), (15) and (16) have a simple interpretation. They all represent the proportion of units facing "fire trials" in the appropriate epoch which survive the trial. A "fire trial" corresponds to a unit surviving or burning over a T-year period.

Likelihood ratio methods for confidence intervals and for testing the significance of a change point.

While the point estimates are unaffected by the presence of contagion, the standard errors of the estimates are affected and depend on the value of the overdispersion parameter σ^2. This parameter cannot be estimated by maximizing the quasi-likelihood. Customarily it is estimated separately either from the Pearson statistic or the residual deviance, suitably normalized (McCullagh & Nelder 1989, Ch. 9).

The covariance matrix of the MLE's can be obtained from the Hessian matrix of second derivatives of the quasi-likelihood in the same way as for ordinary likelihood. The only difference is that the covariance matrix includes, as a scaling factor, the overdispersion parameter σ^2, which must be estimated. Using one or other of the estimates of σ^2 mentioned above provides an estimated covariance matrix and hence estimated standard errors and correlations between parameter estimates.

Likelihood ratio methods can be used for obtaining confidence intervals and for testing hypotheses concerning hazard rates. Essentially the methods are the same as for ordinary likelihoods, and can be expressed in terms of the *deviance*. For a particular model M, with parameters Θ_M (*i.e.* in which the probabilities $S^{(j)}$, $j = 1,\cdots,m-1$ are expressed in terms of the parameters Θ_M) the *quasi-deviance* (see McCullagh & Nelder 1989, Sec. 9.2) is defined as:

$$D_M = 2\sigma^2\left[Q(\hat{\Theta}_M) - \hat{Q}_S\right], \tag{17}$$

where $Q(\hat{\Theta}_M)$ is the quasi-likelihood maximized over the parameters Θ of the model M, and Q_S is the maximized quasi-likelihood corresponding to the *saturated model*, *i.e.* Q (in (10)) maximized over the m-1 parameters $S^{(j)}$, $j = 1,\cdots,m-1$. In other words, Q is evaluated at the MLE's

$$\hat{S}^{(j)} = \frac{\sum_{i=j+1}^{m} A_i}{\sum_{i=j}^{m} A_i},\ j = 1,\cdots,m-1. \tag{18}$$

Note that the quasi-deviance does not depend on the overdispersion parameter σ^2. A form of the deviance which does depend on σ^2 is the *scaled deviance*, which for model M is defined as:

$$D_M^* = \frac{D_M}{\sigma^2} = 2\left[Q(\hat{\Theta}_m) - \hat{Q}_S\right] \tag{19}$$

Deviances are non-negative and additive. They play a similar role to sums of squares in the statistical theory of Gaussian linear models. In consequence, for a sequence of nested models, one can construct an *analysis of deviance* table analogous to an analysis of variance table in standard linear model theory.

To illustrate, let us consider a test of the hypotheses

$$H_0: S^{(1)} = S^{(2)} = ... = S^{(m)} \quad (= S_0, \text{ say}) \; vs$$
$$H_1: S^{(1)} = S^{(2)} = ... = S^{(p)} \quad (= S_1, \text{ say}) \neq S^{(p+1)} = S^{(p+2)} = ... = S^{(m)} \quad (= S_2, \text{ say}),$$

i.e. that the hazard of burning was constant at all times in the past, against the alternative that it changed at some *pre-specified* change point, pT years ago.

Here the parameter space under H_0 is of dimension one (*i.e.* there is one free parameter, the common S_0), while under H_1 it is of dimension two (two free parameters, S_1 and S_2). Furthermore, H_0 is nested in H_1. The difference in quasi-deviances

$$\nabla D = D_0 - D_1 \tag{20}$$

for the two models gives a measure of the discrepancy of the data from H_0 in the direction of H_1. The distribution of ∇D depends on the overdispersion parameter σ^2. However, asymptotically the difference in *scaled* deviances

$$\nabla D^* = D_0^* - D^* = \frac{\nabla D}{\sigma^2} \tag{21}$$

has a χ_1^2 distribution under H_0 (the one degree of freedom corresponding to the difference between the dimensionality of the parameter space under H_1 and H_0). Of course ∇D^* depends on the overdispersion parameter σ^2. However one can use an estimate of σ^2 in the expression ∇D^*, and arrive at the test statistic

$$\nabla \hat{D}^* = \frac{\nabla D}{\hat{\sigma}^2} \tag{22}$$

To estimate σ^2 one can use either the residual deviance divided by its degrees of freedom[1]

---

[1]The degrees of freedom here equal the number of age classes less one (m-1) minus the number of free parameters ($\nu_1 = 2$) under H_1.

$$\hat{\sigma}_D^2 = \frac{D_1}{m-3}, \tag{23}$$

or the Pearson statistic similarly scaled

$$\hat{\sigma}_P^2 = \frac{1}{m-3} \sum_{i=1}^{m} \frac{(A_i - \hat{A}_i)^2}{\hat{A}_i (1 - \hat{S}_i)} \tag{24}$$

where $\hat{A}_i$ and $\hat{S}_i$ are respectively the estimated expected area and proportion in age-class i using the parameter estimates under H_1. Usually the two estimates of $\hat{\sigma}_D^2$ and $\hat{\sigma}_P^2$ are very similar. Furthermore, they both have χ^2_{m-3} distributions asymptotically. It follows that the test statistic

$$\begin{aligned} \nabla\hat{D}^* = -\frac{2}{\hat{\sigma}^2} \Bigg[& (B_1 + B_2 + \cdots + B_p) \ln \frac{\hat{S}_0}{\hat{S}_1} + (A_1 + \cdots + A_p) \ln \left(\frac{1-\hat{S}_0}{1-\hat{S}_1} \right) \\ & + (B_{p+1} + \cdots + B_{m-1}) \ln \left(\frac{\hat{S}_0}{\hat{S}_2} \right) + (A_{p+1} + \cdots + A_{m-1}) \ln \left(\frac{1-\hat{S}_0}{1-\hat{S}_2} \right) \Bigg] \end{aligned} \tag{25}$$

(where $\hat{S}_0$, $\hat{S}_1$ and $\hat{S}_2$ are the MLE's in (13)-(16)) follows approximately an $F_{1,(m-3)}$ distribution under H_0 and thus that an approximate P-value for testing H_0 *vs.* H_1 can be calculated by comparing the observed value of $\nabla\hat{D}^*$ with an $F_{1,(m-3)}$ distribution.

The same procedure can be followed in more complicated situations provided H_0 is nested within H_1. The test statistic would again be the difference D_0 - D_1 in quasi-deviances for the two models divided by its degrees of freedom, over the estimate of σ^2 under H_1. This estimate would now have divisor (degrees of freedom) $m - 1 - \nu_1$ and the degrees of freedom of the reference F-distribution would be $((\nu_1 - \nu_0), (m - 1 - \nu_1))$ where ν_0 is the dimension of the parameter space under H_0 and ν_1 is the corresponding dimension under H_1. The P-value would be

$$P = Pr\left(F_{\nu_1 - \nu_0, m-1-\nu_1} \geq \nabla\hat{D}^*_{\text{obs}}\right) \tag{26}$$

This procedure is analogous to the F-test in ANOVA for a Gaussian linear model with the estimate of σ^2 and the difference in quasi-deviances being analogous respectively to the residual mean square and the extra sum of squares. The main difference is that in the linear Gaussian case the test would be exact (subject of course to correct model specification) whereas here the F-test is only valid asymptotically (for large n).

The issue of sample size raises an important question not so far addressed, *viz.* what is the "correct" size of a "unit area." It would seem at first sight that one could make n as large as one wished, simply by defining a unit to be of sufficiently small area. The question then is whether this would have any effect on the MLE's and test statistics discussed above. In fact it does not because a change in the size of a unit by a factor α (*e.g.* from 10 ha to 100 ha) would simultaneously change the area variables A_i and $\hat{A}_i$ etc. by a factor of α^{-1} and hence change both the quasi-deviances and the estimates of σ^2 by a similar factor. In consequence any LR test statistic $\nabla\hat{D}^*$ (22) would be unaffected. Similarly since the MLE's (equations (13), (14), (15), (16)) depend only on ratios of areas, they would also be unaffected. Provided that the study area is large the use of asymptotic inference procedures should provide good approximations.

The likelihood ratio procedure can be used in the customary way to obtain approximate confidence intervals, or confidence regions, for model parameters, exploiting the fact that a $100(1-\alpha)\%$ confidence interval for a parameter θ can be thought of as a set of values θ_0 for which the null hypothesis $H_0{:}\theta = \theta_0$ would not be rejected at the level α in favor of the alternative $H_1{:}\theta \neq \theta_0$. For example, for a model with a single homogeneous hazard of burning a $100(1-\alpha)\%$ confidence interval for the common survival probability S_0 could be obtained as

$$\left\{S_0 : \nabla\hat{D}^* \leq F_{1,m-2;\alpha}\right\} \tag{27}$$

where the estimated difference in scaled deviances is

$$\nabla\hat{D}^* = \left(\frac{1}{\hat{\sigma}^2}\right)\left[\sum_{j=1}^{m-1}\left\{B_j \ln\left(\frac{S_0}{\hat{S}_0}\right) + A_j \ln\left(\frac{1-S_0}{1-\hat{S}_0}\right)\right\}\right] \tag{28}$$

with $\hat{S}_0$ given by (13) being the MLE of the common per period survival probability, and $\hat{\sigma}^2$ being the estimate of the overdispersion parameter for this estimate. Of course the inequality in (27) must be solved numerically to obtain the explicit confidence interval. The corresponding confidence interval for the hazard of burning λ_0 or the fire cycle can be found by simply applying the transformation from S_0 to λ_0 (14) to the confidence limits of S_0.

Example

To demonstrate the use of the methods described above we use the data from the time since fire map of Johnson and Larsen (1991). These data are for a 495 km^2 area of the Kananaskis watershed in the southern Canadian Rockies. Attempts by Johnson and Larsen to divide the map into spatial subunits with distinct hazards of burning were unsuccessful. Using the method of Kao (1959), they identified graphically a change point around 1730 CE and estimated the hazard of burning (and corresponding fire cycle) for the resulting two epochs. Remember that the partitioning method assumed forest originating during the older epoch was not subjected to the subsequent hazards of burning. Johnson and Larsen (1991) estimated the fire cycles (time required to burn an area equivalent to the study area) as 50 years for greater than 250 years ago and 90 years for 250 years ago to the present.

We now illustrate the methods in the preceding section using the Kananaskis data. For these data $T = 10$, $m = 40$.

The data were plotted on semilog graph paper (see Figure 2). The graph suggests that the data are from a mixed distribution, with a change point near to that identified by Johnson and Larsen, and that the survivorship of each of the two epochs appear to follow negative exponential distribution (*i.e.* straight lines on the semilog plots). Assuming a change point at $p = 25$ ($\sim$1730) the MLE's for S_1 and S_2 are:

$$\hat{S}_1 = \frac{\sum_{i=2}^{m} A_i + \sum_{i=3}^{m} A_i + \dots + \sum_{i=p+1}^{m} A_i}{\sum_{i=1}^{m} A_i + \sum_{i=p+2}^{m} A_i + \dots + \sum_{i=p}^{m} A_i} = 0.9332$$

$$\hat{S}_2 = \frac{\sum_{i=p+2}^{m} A_i + \sum_{i=p+3}^{m} A_i + \dots + \sum_{i=m}^{m} A_i}{\sum_{i=p+1}^{m} A_i + \sum_{i=p+2}^{m} A_i + \dots + \sum_{i=m-1}^{m} A_i} = 0.8067.$$

The corresponding estimated hazards of burning (from (14)) and fire cycles (the inverse of the hazard) are:

Epoch 1 (1730-1980) $\lambda_1 = 0.00692$ $FC_1 = 144.6$ years

Epoch 2 (before 1730) $\lambda_2 = 0.02148$ $FC_2 = 46.6$ years

The estimate of 46.6 years for the earlier epoch agrees with that of Johnson and Larsen, while that of 144.6 years for the more recent epoch is considerably longer than the Johnson and Larsen (1991) estimate of 90 years. Since the Johnson and Larsen (1991) estimate was based on the assumption that units originating in an earlier epoch were not vulnerable to fire in the more recent epochs, it is not surprising that compared with the current method it would underestimate the survival probability (and hence underestimate the fire cycle) in Epoch 1.

Likelihood ration (LR) 95% confidence intervals[2] were obtained by solving the equivalent of (27) numerically, with the following results:

Epoch 1 (1730-1980) FC_1: 99.2 - 222.3 years

Epoch 2 (1580-1729) FC_2: 19.9 - 151.2 years

These confidence intervals are somewhat asymmetrical about the MLEs. This is due to taking reciprocals of the confidence limits for the hazards of burning, λ_i, which in contrast are quite symmetrical. This fact, coupled with the invariance with respect to re-parameterization of the

[2]These calculations used the Pearson estimate of σ^2. The corresponding intervals using the residual deviance estimate of σ^2 are 99.3 - 222.0 and 20.0 - 150.0. As expected there is little difference between the estimates.

LR statistic indicates that the confidence intervals for *FC* should be fairly accurate.

The confidence intervals are quite wide, especially that for the earlier epoch. The latter imprecision reflects the fact that there is a relatively small amount of data for the earlier epoch, with only 15% of the total forest area (about 75 km^2) surviving from before 1730 CE. Even for the more recent epoch (for which the data are more abundant) there is considerable imprecision. In the hazard-rate scale the confidence limits for the later epoch are respectively about 30% below and 40% above the MLE. This degree of imprecision simply reflects the fact that the estimation of probabilities with high precision requires very large amounts of data. If the change point, pT, is *pre-specified* and not selected by examining the data, the correct LR statistic for testing the statistical significance of the change point, *i.e.* for testing

$$H_0 : \lambda_1 = \lambda_2 \ vs.\ H_1 : \lambda_1 \neq \lambda_2$$

where λ_1 and λ_2 are the hazards of burning before and after the change point, is (from (25))

$$\nabla \hat{D}^* = \frac{D_0^* - D_1^*}{\hat{\sigma}^2} = 3.55$$

where D_0^* and D_1^* are the scaled deviances corresponding to H_0 and H_1 respectively, and $\hat{\sigma}^2$ is an estimate of the overdispersion parameter.

An approximate *P*-value for testing H_0 *vs.* H_1 can be obtained by comparing the observed value of $\nabla \hat{D}^*$ (= 3.55) with an $F_{1,37}$ distribution. This yields an approximate *P*-value of 0.067.[3] Thus the evidence of a change point 250 years ago is weak at best.

In fact there is a slightly smaller *P*-value for the significance of change point at $p = 24$ (*i.e.*

---

[3]This *P*-value is calculated using the Pearson estimate of σ^2. If instead the residual deviance estimate is used the test statistic has a value $\nabla \hat{D}^*_{\text{obs}} = 3.57$ and the corresponding *P*-value is again 0.067.

around 1740). The F-statistic in this case has a value 4.17 (or 4.26 using residual deviance), leading to a P-value of 0.048 (or 0.046 with residual deviance). Thus there appears to be some evidence of a change point around 1740. With a change point at this time the estimated fire cycles, with 95% confidence intervals are:

Epoch 1 (1740-1980) $FC_1 = 147.8$ years (100.9-229.1)

Epoch 2 (1580-1739) $FC_2 = 48.1$ years (22.1-137.2)

Although the change points tested were suggested from exploratory data analysis, the use of the likelihood ratio F-test above could be justified from the fact that a change around the middle of the 18th century has been identified for other regions in Western Canada (*e.g.* Glacier and Kootenay National Parks - see Johnson Fryer and Heathcott (1990) and Masters 1989).

Acknowledgments

This research was supported by NSERC grant OGP 7252 to W.J. Reed; by a University of Waterloo Institute for Risk Research grant and a McMaster University Graduate Scholarship to C.P.S. Larsen; and by NSERC grants to E.A. Johnson and to G.M. MacDonald.

References

Bergeron, Y. 1991.The influence of island and mainland lakeshore landscapes on boreal forest fire regimes. Ecology 72: 1980-1992.

Bessie, W.C. and Johnson, E.A. 1995. The relative importance of fuels and weather on fire behavior in subalpine forests. Ecology 76: 747-762.

Clark, J.S. 1990. Fire and climate change during the last 750 years in northwestern Minnesota. Ecological Monographs 60: 135-159.

Connell, J.H. and Slatyer, R.O. 1977. Mechanisms of succession in natural communities and their role in community stability and organization. American Naturalist 111: 1119-1144.

Cox and Hinkley. 1974.

Engelmark, O., Kullman, L. And Bergeron, Y. 1994. Fire and age structure of Scots pine and Norway spruce in northern Sweden during the past 700 years. New Phytologist 126: 163-168.

Heinselman, M.L. 1973. Fire in the virgin forests of the Boundary Waters Canoe Area. Quaternary Research 3: 329-382.

Johnson, E.A., Fryer, G.I. and Heathcott, M.J. 1990. The influence of man and climate on frequency of fire in the interior wet belt forest, British Columbia. Journal of Ecology 78: 403-412.

Johnson, E.A. and Gutsell, S.L. 1994. Fire frequency models, methods and interpretations. Advances in Ecological Research 35: 239-287.

Johnson, E.A. and Larsen, C.P.S. 1991. Climatically induced change in fire frequency in the southern Canadian Rockies. Ecology 72: 194-201.

Johnson, E.A. and Van Wagner, C.E. 1985.Theory and use of two fire history models. Canadian Journal of Forest Research 15: 214-220.

Kao, J.H. 1959. A graphical estimation of mixed Weibull parameters in life-testing of electron tubes. Technometrics 1: 389-407.

Masters, A.M. 1990. Changes in forest fire frequency in Kootenay National Park, Canadian Rockies. Canadian Journal of Botany 68: 1763-1767.

McCullagh and Nelder. 1989.

McCune, B. 1983. Fire frequency reduced two orders of magnitude in the Bitterroot Canyons, Montana. Canadian Journal of Forest Research 13: 212-218.

Minnich, R.A. 1983. Fire mosaics in southern California and northern Baja California. Science 219: 1287-1294.

Reed, W.J. 1994.Estimating the historic probability of stand-replacement fire using the age-class distribution of undisturbed forest. Forest Science 40: 104-119.

Renkin, R.A. and Despain, D.G. 1992. Fuel moisture, forest type, and lightning-caused fire in Yellowstone National Park. Canadian Journal of Forest Research 22: 37-45.

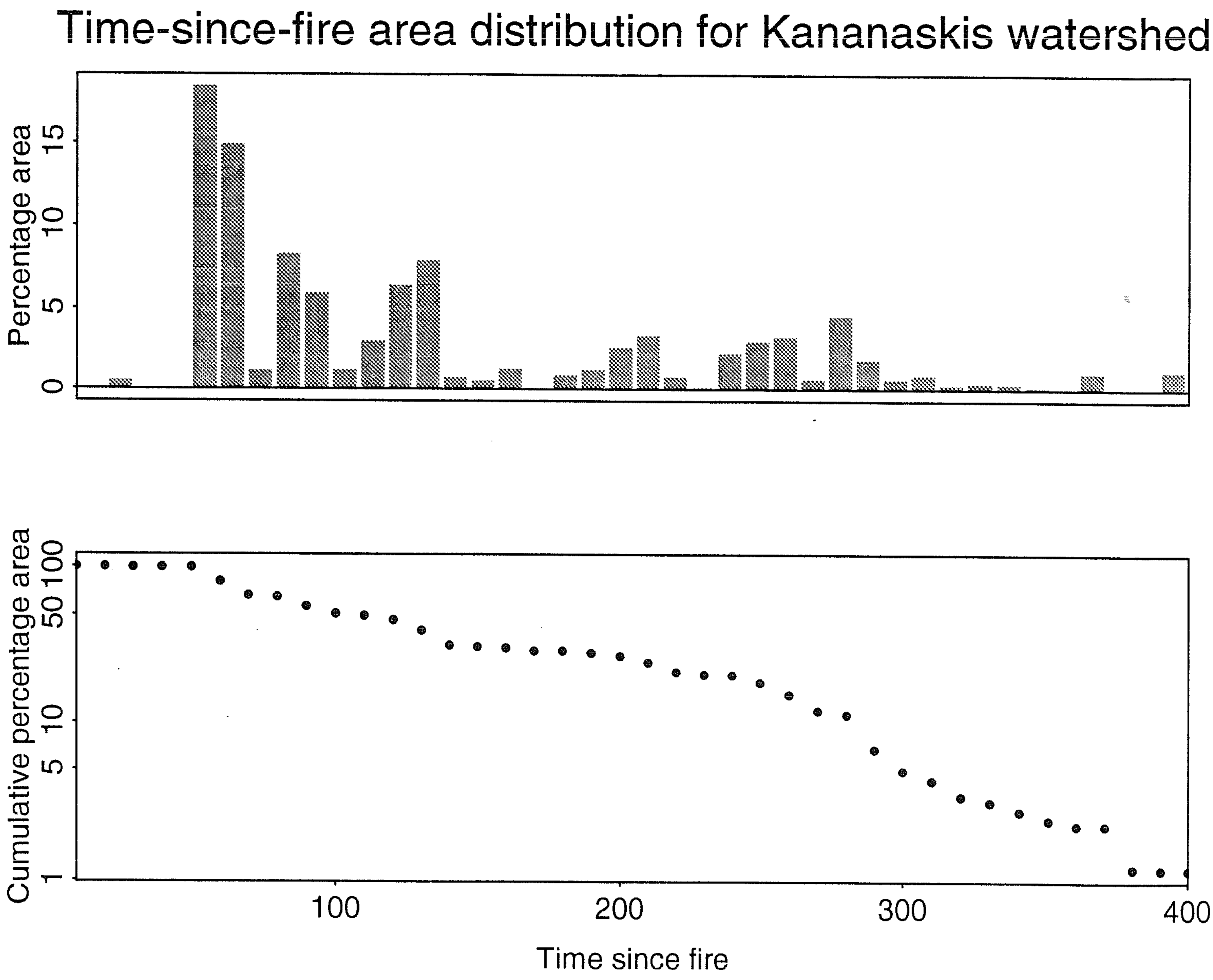
Romme, W.H. 1982. Fire and landscape diversity in subalpine forests of Yellowstone National Park. Ecological Monographs 52: 199-221.

Van Wagner, C.E. 1978. Age-class distribution and the forest fire cycle. Canadian Journal of Forest Research 8: 220-227.

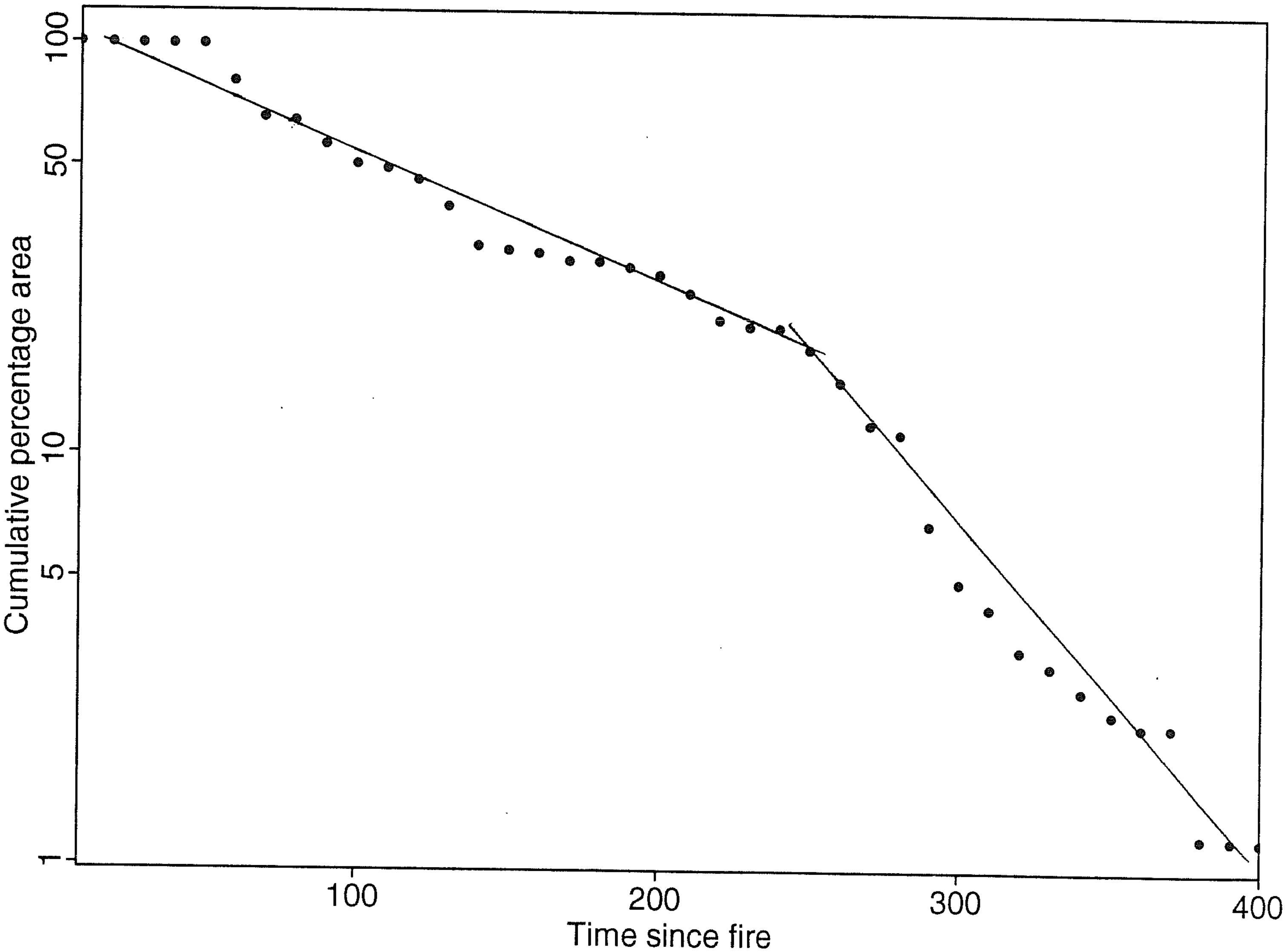
Figures

Figure 1. A time-since-fire distribution. The upper panel is a histogram of areas in time-since-fire classes; the lower panel is a semi-log cumulative frequency plot *i.e.* a plot of cumulative percentage area (on a logarithmic scale) against time-since-fire.

Figure 2. A semi-log cumulative frequency plot showing a graphical partition into two epochs.



Time-since-fire area distribution for Kananaskis watershed
Percentage area
0
5
10
15
Cumulative percentage area
1
5
10
50
100
100
200
300
400
Time since fire



Cumulative percentage area
100
50
10
5
1
100
200
300
400
Time since fire