RECONSTRUCTING THE HISTORICAL FREQUENCY OF FIRE:
A MODELING APPROACH TO DEVELOPING AND TESTING METHODS

by

Joseph Gordon Fall
B.Sc. University of Victoria 1991

RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF NATURAL RESOURCES MANAGEMENT

in the
School of Resource and Environmental Management

Report No. 225

© Joseph Gordon Fall 1998

SIMON FRASER UNIVERSITY

August 1998

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
APPROVAL

Name: Joseph Gordon Fall

Degree: Master of Natural Resources Management
in the School of Resource and Environmental Management

Report No.: 225

Title of Research Project:
Reconstructing The Historical Frequency Of Fire:
A Modeling Approach To Developing And Testing Methods

Examining Committee:

Senior Supervisor: Ken Lertzman
Associate Professor
School of Resource and Environmental Management
Simon Fraser University

Committee Member: Alton Harestad
Associate Professor
Department of Biological Sciences
Simon Fraser University

Committee Member: Emily Heyerdahl
Research Associate
School of Resource and Environmental Management
Simon Fraser University

Date Approved:

ii
Abstract

Fire is a prevalent natural disturbance in most of British Columbia’s forest ecosystems. Recently, scientists and forest managers have recognized the important role fire plays in regulating forest ecosystems and maintaining biodiversity. In response, B.C. Government initiatives propose to use an ecosystem’s historical disturbance dynamics for guiding forest management. Gaining an understanding of the methods used to estimate historical fire frequency, along with the limitations of these methods, and the sources of uncertainty and magnitude of bias in such estimates, will be critical for developing such ecosystem-based management objectives.

In Chapter 2, I review the published fire history literature, focusing particularly on the methods, underlying models, and calculations used to estimate historical fire frequency. This review is presented as an interactive tutorial, to aid a novice reader gain an understanding of some of the more difficult aspects of fire frequency reconstruction and interpretation. Some sample pages and a description of the tutorial are provided along with instructions on how to obtain the complete package.

All fire history studies rely on a series of inferences based on a set of physical evidence left by fire. This physical evidence contains inherent errors, most often of unknown magnitude. In addition, other errors are introduced when a researcher samples this evidence to create a data set, and estimates the history of fire occurrence from this data set. In Chapter 3, I present a methodology for quantifying the level of confidence that should be placed in an estimate of historical fire frequency made from tree-ring based fire interval data. In this approach, I use a spatial simulation model of the fire regime to generate synthetic fire histories. I propose and use new techniques to model the formation and survivorship of fire evidence in the tree-ring record. These models introduce errors into the synthetic fire histories based on the types of errors thought to be present in the physical data. Finally, a spatial model of fire history sampling is used to simulate errors introduced by the researcher.
I use Monte Carlo simulation to derive a confidence interval for the empirical estimate of fire frequency made in a recent fire history study. The results indicate that it is possible to reliably estimate historical fire frequency from fire interval data. However, the greatest source of uncertainty in this estimate is the probability with which fire evidence is formed in the tree ring record. This source of error has received little attention in the literature, and so I conclude by recommending that this problem be given serious study. In the mean time, I recommend that researchers minimize this source of uncertainty by collecting samples from several trees at each sampling point in the landscape.
Acknowledgments

I would like to thank Brigitte Dorner for her helpful insights and infinite patience; Ken Lertzman for his constant inspiration and encouragement; and my brother, Andrew Fall, for all of his help with SELES and the trickier bits of math — they each deserve much credit for this work, except for the errors, which are all mine. In addition, this project would not have been possible without the case study data set on which it is based. I thank Emily Heyerdahl for bravely submitting her Dugout Creek data to the tortures of my analyses, and give her all the credit for designing and carrying out such an excellent empirical fire history study. This research was supported by a Natural Sciences and Engineering Research Council Post-Graduate Scholarship (NSERC PGS A), and grants from Forest Renewal British Columbia.
# Table of Contents

<table>
<thead>
<tr>
<th>APPROVAL</th>
<th>ii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
</tbody>
</table>

## Chapter 1: Background and Problem Definition

1. The relevance of fire regimes to resource management 1
2. Key factors driving fires and fire regimes 3
3. Characterizing fire regimes 5
   - Sources of uncertainty in estimates of historical fire frequency and extent 7
4. Overview of Chapter 2 8
5. Overview of Chapter 3 10

## Chapter 2: An Introductory Tutorial on Common Methods for Determining Fire Frequency

1. Introduction 12
2. Motivation and Scope 12
3. Example Tutorial Pages
   - Sample Tutorial Page 1: *Tutorial Map* 14
   - Sample Tutorial Page 2: *Introduction* 15
   - Sample Tutorial Page 3: *What is Fire Frequency?* 16
   - Sample Tutorial Page 4: *Poisson Tutorial* 17
   - Sample Tutorial Page 5: *Poisson Overview* 18
   - Sample Tutorial Page 6: *Poisson Simulation* 19
   - Sample Tutorial Page 7: *MFRI Tutorial* 20
   - Sample Tutorial Page 8a: *Fire Frequency Model* 21
   - Sample Tutorial Page 8b: *Fire Frequency Model (cont.)* 22
   - Sample Tutorial Page 9: *Interpretation* 23
   - Sample Tutorial Page 10: *Assumptions* 24
# Chapter 3: Testing Methods for Estimating Fire Frequency from Fire Interval Data

## Introduction
- Sources of uncertainty in estimates of historical fire frequency derived from fire interval data

## Models and Methods
- REFR -- A Spatially Explicit, Stochastic Fire Regime Simulation Model
- EVA -- a Stochastic Model of Error Sources in Fire History Data

## Case Study -- A model for Dugout Creek.
- Simulating the Dugout Creek Fire Regime
- Simulating the Dugout Creek Fire History Sampling
- Summary of REFR and EVA model parameter values for Dugout Creek:

## Results and Discussion
- Consistency of fire regime parameters
- Variability in fire regime
- Confidence interval for point fire frequency
- Primary sources of uncertainty in fire history studies
- Evaluating the Sampling Design
- The Effect of Fire Frequency on Fire Frequency Estimates

## Conclusions

## Figures

## Appendix A: Transforming an exponential distribution into a histogram

## Appendix B: Computing the probability of a fire scar forming
- Application problem
- Equivalent counting problems to derive an estimate of $N$
- Estimating a value for $p_r$, the probability of recording a fire
- Assumptions
- Proof: $f(t) = 1 - (1-p)^t$

## Literature Cited
List of Tables

Table 1. Parameters for the EVA sub-sampling models discussed in the text. 43
Table 2. Results from the different sampling scenarios described in Table 1. 45
Table 3. Analysis of variance on the point fire interval distributions for several replicates of the Base Case scenario. 49
List of Figures

Figure 1. Conceptual diagram for the project. 64
Figure 2. Conceptual diagram of REFR simulation. 65
Figure 3. Conceptual diagram of EVA model. 66
Figure 4. The distribution of fire years per decade at Dugout Creek. 67
Figure 5. Distribution of intervals between fire years at Dugout Creek. 68
Figure 6. The distribution of fire extents at Dugout Creek. 69
Figure 7. Sampling design used in the EVA sampling model (Figure 3). 70
Figure 8. The cumulative distribution of survival times for fire evidence at Dugout Creek. 71
Figure 9. (a) The proportion of fires recorded by trees sampled at Dugout Creek; and (b) the function G(t) fit to the empirical estimate G, for all Dugout Creek sample sites (see Appendix B). 72
Figure 10. Point fire interval distributions for (a) the empirical data set for Dugout Creek and (b) the synthetic Base Case scenario replicate Run73. 73
Figure 11. Distribution of mean fire return intervals resulting from 100 replicates of the CompleteLong (a), Complete (b), and Base Case (c) scenarios. 74
Figure 12. Synthetic point fire interval histograms from three of the 255 year replicates. 75
Figure 13. Distribution of biases in the estimates of MFRI for the 100 replicates of the Base Case scenario, showing the 95% interval for the histogram. 76
Figure 14. Distribution of biases in MFRI estimates for 100 replicates of the EVA sampling scenarios that vary the probability of recording fire evidence. 77
Figure 15. Distribution of biases in MFRI estimates for 100 replicates of the EVA sampling scenarios that vary the number of trees collected at each plot. 78
Figure 16. Correlation between the true MFRI and the magnitude of bias in the MFRI estimate. 79
Chapter 1: Background and Problem Definition

The relevance of fire regimes to resource management

“The more that managed forests resemble the forests that were established from natural disturbances, the greater the probability that all native species and ecological processes will be maintained.” (B.C. MoF 1995)


Recently, scientists and forest managers have recognized that these changes in fire regime often result in undesirable effects on the species and communities that have adapted to the forests and fire regimes of the previous centuries (Hansen et al. 1991, Agee 1993, Forest Ecosystem Management Assessment Team (FEMAT) 1993, Bunnell 1995). This developing awareness of the important roles fire plays in regulating forest ecosystems and maintaining biodiversity is creating new challenges for managers who wish to use an ecosystem’s natural dynamics for guiding management (Agee and Johnson 1988, Hansen et al. 1991, Forest Ecosystem Management Assessment Team (FEMAT) 1993, Hunter 1993, Swanson et al. 1993, Bunnell 1995, Johnson et al. 1995, DeLong and Tanner 1996, Duinker and Euler 1997, Fule et al. 1997, Lertzman et al. 1997). B.C. Government initiatives, such as the Forest Practices Code (FPC) Biodiversity Guidebook, and the

Gaining an understanding of the historical fire regime and range of natural variability with which forest ecosystems have evolved will be critical for developing ecosystem-based management objectives (Swanson et al. 1993, Johnson et al. 1995, DeLong and Tanner 1996, Duinker and Euler 1997, Fule et al. 1997, Lertzman et al. 1997, Parminter 1998). For example, in order to develop effective objectives and strategies to restore or maintain “natural” (historic or desired) conditions in parks and wilderness areas, managers will often need to understand how the historical fire regime shaped the landscape. In addition, managers of smaller, conservation-oriented reserves (usually called ecological reserves in B.C.) will need to know whether these systems will remain stable over time, and thus serve their purpose, in the presence, or absence, of a particular fire regime. Forest health managers will also require knowledge about the relationship between fire and outbreaks of insects or pathogens, to determine appropriate fire and pest management strategies. Even recent initiatives in timber management use historical fire regimes as models for anthropogenic disturbances (Forest Ecosystem Management Assessment Team (FEMAT) 1993, British Columbia Ministry of Forests 1995, Scientific Panel for Sustainable Forest Practices in Clayoquot Sound 1995, Duinker and Euler 1997).

In B.C., the FPC Biodiversity Guidebook (British Columbia Ministry of Forests 1995) establishes target seral stage distributions for all landscapes managed for timber in the province. These targets are designed to reflect the seral stage distributions of “natural,” or pre-industrial, forests, and are thus based on an estimate of historical disturbance frequencies. To determine the target distribution from the estimate of historical disturbance frequency, the Biodiversity Guidebook makes use of a common model of fire frequency and age-class structure (the negative exponential model, discussed in Chapter 2). While the usefulness of this model is currently being questioned by some fire researchers (Lertzman et al. 1998, Andison pers. comm.), others are applying it to
guide forest management. For example, recent research within the B.C. Ministry of
Forests (MoF) Nelson Region makes use of this and other fire frequency models to argue
for increased harvest levels in the region (Pollack et al. in prep.). Clearly, forest managers
require at least a basic understanding of the models, assumptions, and implications of the
various methods used to reconstruct historical fire frequency. In addition to
understanding how an estimate of historical fire frequency was derived, they will also
require information about the historical range of variability, along with some measure of
the uncertainty in these estimates. These requirements form the primary motivation for my
project.

**Key factors driving fires and fire regimes**

A fire regime describes the general, long-term pattern of fire occurrence and
effects in an ecosystem. Individual fires are commonly described by both their physical
characteristics, called fire behaviour (e.g., fire line intensity, rate of spread, etc.), and by
their ecological effects (e.g., proportion of trees killed, depth of burn, etc.; Rothermel
1972, Johnson 1992, Agee 1993, Whelan 1995). Similarly, fire regimes may also be
described by their physical characteristics (e.g., fire frequency, mean fire extent, etc.), and
by their ecological effects (e.g., typical fire severity, influence on species composition and
competitive advantage, etc.; Heinselman 1973, Barrett et al. 1991, Bergeron 1991,
Johnson 1992, Agee 1993, Heyerdahl 1997). However, while individual fires are of
limited duration and occur over a discrete, identifiable spatial area, a fire regime
summarizes the cumulative, typical, or statistical characteristics and effects of many fires
occurring over some larger region of space and time. This definition raises an important
point — whereas the spatio-temporal domain of an individual fire has a physical definition
(e.g., the fire burned for 10 days over an area of 1000 ha.), the spatio-temporal domain
used to define a fire regime is largely determined by the observer. This point has
significant implications for the interpretation of historical fire regimes because the
conclusions that a researcher draws from a fire history will be greatly influenced by the
spatio-temporal domain chosen for the study (Lertzman et al. 1998).
The key factors driving the behaviour of individual fires are well understood. These variables include fire weather (wind speed and direction, temperature, and relative humidity); topography (elevation, slope, and aspect), and fuels (fuel load, size, moisture content, and continuity) (Johnson 1992, Agee 1993, Whelan 1995, Pyne et al. 1996). In fact, given a set of values for these variables, the behaviour of many fires can be predicted with some accuracy (Rothermel 1972, Keane et al. 1990, Green et al. 1995, Finney 1996). The effects of any particular fire are largely a function of the fire’s behaviour, coupled with the pre-fire state of the ecosystem (i.e., depth of duff layer, status of seed banks, etc.). In general, the effects of fire on soils, hydrology, wildlife and vegetation, both at the individual and community level, are also fairly well understood (Keane et al. 1990, Johnson 1992, Agee 1993, Whelan 1995, Turner et al. 1997). However, considerable heterogeneity in both the physical environment during the fire and in vegetative responses after the fire generally result in effects that are quite variable and thus more difficult to predict accurately (Agee 1993, Huff 1995, Turner et al. 1997).

The characteristics of a fire regime, on the other hand, are primarily determined by climate (Johnson 1992, Agee 1993, Whelan 1995), although other larger scale physical attributes of the ecosystem, such as the soil mosaic, physiography, and the scale and types of heterogeneity in the system may also play important roles. Climate is such an important factor because it encompasses and drives so many physical and ecosystem processes. For example, climate has a direct influence on the fire regime by determining the weather and duration of the fire season (Johnson and Wowchuk 1993, Bessie and Johnson 1995). These climatic characteristics, in turn, drive a host of critical fire related variables, such as fuel moisture content, and the frequency of lightning ignitions. Climate also has important indirect influences on the fire regime because it largely determines the potential dominant vegetation types for the system. The type of vegetation, in turn, determines the quantity, quality, and spatial distribution of fuels, and thus plays a key role in driving the fire regime.

The fire regime plays many important roles in structuring a forest. In many ecosystems the species composition and distribution (both overstory and understory) can be at least partially explained by the fire regime (Agee 1993, Whelan 1995). This
mechanism has interesting implications since it indicates that there may be a feedback interaction between the vegetation type and the fire regime (Parminter 1998). The fire regime is also often primarily responsible for determining the age structure of the vegetation (Johnson 1979, Huff 1995). Other attributes of the forest affected by the fire regime include the landscape patch-size distribution, the availability, adjacency, and connectivity of different habitat types through time, nutrient cycling, productivity, and the physical properties of the soil, including the depth of the organic layer, and erosion potential (Heinselman 1973, Stark 1977, Eberhard and Woodard 1987, Dyrness et al. 1989, Hansen et al. 1991, Ruggerio et al. 1991, Agee 1993, Turner et al. 1994, Bunnell 1995, DeLong 1997, Parminter 1998). The degree of variability in the fire regime and heterogeneity in the system will play a large part in determining the stability of each of the above factors through time (Lertzman and Fall 1998).

**Characterizing fire regimes**

The physical characteristics of a fire regime are generally described by three primary parameters. The *frequency* is a measure of how often fires occur (e.g., interval in years between fires); the *extent* is a measure of the typical size of fires (e.g., mean area burned in hectares); and the *magnitude* is a measure of the intensity, severity, or effects of the fires (e.g., proportion of trees killed). Other attributes of the fire regime, such as the seasonality of fires and typical fire behaviour (e.g., surface fires vs. crown fires), help to complete this characterization (Agee 1993).

In areas where accurate historical records are not available, there are two dominant approaches to estimating the historical *frequency* of fire from tree-ring based evidence:

1. In areas with high-severity, stand-replacing fires, even-aged cohorts of post-fire regeneration are dated and mapped for the entire study area to produce a “time-since-fire” map. By assuming that the forest landscape age-class distribution is relatively stable over time (e.g., a steady-state shifting mosaic; Baker 1989), and that the frequency of disturbance can be modeled as a spatially homogeneous, stationary Poisson process, an estimate can be made of the average time between fires for all points in the study area (Johnson and Van Wagner 1985, also see Chapter 2).
1. In areas with low-severity, stand-maintaining fires, fires may kill a portion of the cambium on some trees, while not killing the tree itself. This creates a scar on the tree, which can be dated dendrochronologically (Madany et al. 1982, McBride 1983). A single tree may record many such fire events, and the distribution of intervals between fire events can be used to derive an estimate of the fire frequency (Agee 1993, also see Chapter 2).

Similar methods have also been used to characterize the typical spatial extent of fires. Using “time-since-fire” methods, the annual percent burned (the average proportion of the study area burned each year) can be estimated from the frequency measure (Johnson and Gutsell 1994). In conjunction with an estimate of the average number of fires per year, this estimate can be used to derive a mean fire extent, but will not yield a distribution of fire extents (see Chapter 2). On the other hand, various methods have been used to reconstruct the extent of individual fires directly from fire scar data (e.g., Morrison and Swanson 1990, Heyerdahl 1997). These reconstructions can then be used to estimate the distribution of fire extents over time for the study area.

The magnitude of fire is likely the most heterogeneous of the three primary fire regime parameters. It is variable both between different fires and within a single fire event (Romme and Despain 1989, Morrison and Swanson 1990, DeLong and Tanner 1996). In addition, while evidence of fire occurrence may remain for centuries, evidence of the fire’s intensity is most often obscured in a relatively short period. Thus, the historical magnitude of fire is usually described qualitatively (e.g., surface fire vs. crown fire) from anecdotal evidence or written historical record. To my knowledge, no quantitative methods for deriving a measure of the magnitude of historical fire have been developed to date. For the purposes of my study, a “fire” must be severe enough to potentially scar or kill a tree. Historical fires of lower severity cannot be detected and are thus excluded from analysis. Other than this observation, I will not deal with fire magnitude further.
Sources of uncertainty in estimates of historical fire frequency and extent

All fire history studies rely on a series of inferences based on a set of physical evidence left by fire. This evidence includes even-aged, post-fire regeneration cohorts (e.g., Johnson 1979, Masters 1990); anomalies in the tree-ring structure of individuals, such as suppression-release radial growth signals and fire scars (e.g., Kilgore and Taylor 1979, Heyerdahl 1997); and charcoal, found in both soil (e.g., Gavin et al. 1996, Gavin et al. 1997) and lake sediments (e.g., Cwynar 1987, Long et al. 1997). All of these sources of physical evidence contain inherent errors, often of unknown magnitude. In addition, other errors are introduced when a researcher samples this evidence to create a data set, estimates the history of fire occurrence from this data set, and makes inferences about the historical fire regime from this history.

Both dominant methods for reconstructing fire history, time-since-fire and fire-interval, yield estimates of fire frequency and extent with some level of uncertainty. In applying time-since-fire methods, this uncertainty is primarily a result of substantial heterogeneity in forest systems and variability in fire regimes, which violate the assumptions of the method and models (Lertzman et al. 1998). In particular, many studies indicate that landscape age structure cannot be assumed to be temporally stable (Romme 1982, Baker 1989, Sprugel 1991, Turner et al. 1993, Andison 1996, Cumming et al. 1996). In addition, spatial and temporal autocorrelation and variability in the extent and timing between fires can introduce a substantial error to fire frequency estimates derived from age-class distributions (Boychuk et al. 1997, Lertzman et al. 1998). Compounding these problems, the difficulty in detecting and accurately aging small, old stands may cause the disturbance frequency to be overestimated (Finney 1995).

In the case of reconstructions from fire scars, uncertainty is primarily due to our inability to detect all past fires. A historical fire may go undetected for several reasons:

- a fire may fail to leave a record in every location it burned (i.e., did not produce a fire scar);
- a severe fire may erase the evidence of previous fires; or
- the sampling scheme may be insufficient to detect all fires that did leave a record.
While an intensive survey may be done to estimate the error introduced by the sampling scheme (e.g., Morrison and Swanson 1990), the level of uncertainty introduced by deficiencies in the physical data are difficult to assess from the fire history data itself. In other areas of ecology, replication and time-series analysis may be used to estimate errors or noise in the physical evidence of a process. However, a fire historian cannot use replication (i.e., each forest landscape is unique), and, of course, there is no way of ever knowing, and thus making a comparison with, the true, long-term fire history for an area.

While there has been a great deal of fire history research, both theoretical (e.g., Johnson and Van Wagner 1985) and empirical (e.g., Agee 1991), it is often difficult to compare values for the primary fire regime parameters between study areas because methods for data collection and analysis are not consistent among investigators. This problem largely arises because the appropriate methods and measures vary for different forests types and disturbance regimes. However, there is also little consensus amongst investigators as to the particular methods and set of measures that should be derived to adequately depict the three primary parameters of a fire regime. In addition, few studies have been undertaken to quantify the uncertainty in estimates of fire regime parameters. Thus, it is difficult to determine if two differing estimates of some parameter may in fact have been produced from similar fire regimes. A more consistent framework for characterizing disturbance regimes and estimating the uncertainty in the measures used is required to guide management activities. Thus, the purposes of this study are two-fold:

1. to present a clear and comprehensive description of common methods for reconstructing fire frequency (Chapter 2); and

1. to develop a method for quantifying the uncertainty in estimates of fire frequency derived from fire interval data; and to apply this method to a case study (Chapter 3).

**Overview of Chapter 2**

Chapter 2 serves as a review of the published tree-ring based fire history literature, focusing particularly on the methods, underlying models, and calculations used to estimate historical fire frequency. This review is presented as an interactive tutorial, programmed in Excel 5 for Windows. The material is presented in this format to help a novice reader
gain access to some of the more difficult aspects of this literature. Some sample pages from this tutorial are provided in this document, with the complete, interactive tutorial is available on diskette, or on the World Wide Web at www.rem.sfu.ca/frstgrp/.

The tutorial starts with some basic definitions and an introduction to the concept of fire frequency. This portion of the tutorial is intended to give the reader some understanding of how the occurrence of fire over time translates into fire frequency. It covers the types of evidence typically employed in analyses of fire frequency, along with a description of how these data are used to build models of fire frequency. The introduction also includes a map of the tutorial, to aid the reader in finding particular sections, a glossary of commonly used terminology, and a bibliography, which serves as an introduction to the literature on fire frequency reconstruction.

The primary subject material is structured as a set of individual, but related, sub-tutorials, each on a particular model or method used in fire frequency analyses. For example, the first tutorial covers the Natural Fire Rotation (NFR) method, as presented by Heinselman (1973). This tutorial presents Heinselman’s data for the Boundary Waters Canoe Area and demonstrates how the NFR is computed. The reader can compute the NFR over different time periods.

The next tutorial introduces a Poisson model of fire occurrence. It describes the underlying assumptions of this model and explains why the model is so important for fire frequency analysis. Two interactive pages allow the reader to control a random Poisson process. The outcomes of this random Poisson process are then analyzed as fire occurrences over space on one page and as fire occurrences over time on the other. These data are then transformed into a measure of frequency to demonstrate the relationship between the Poisson process as a model of fire occurrence, and the computation of fire frequency.

The third sub-tutorial examines methods for computing fire frequency from fire interval data. The data set from Dugout Creek, Oregon (Heyerdahl 1997) is presented and analyzed. A simple computation of the Mean Fire Return Interval is followed by a more complex analysis of the fire interval distribution. This tutorial demonstrates how
two estimates of fire frequency, using different methods, can be derived from the same data set. Included is an interactive section that demonstrates, and alerts the reader to the presence of, some of the scale dependent properties of fire interval data.

The fourth sub-tutorial presents the negative exponential and Weibull models of fire frequency. This tutorial includes an interactive section that allows the reader to examine the equations and graphs for these two models, and to see the effect of changing the model parameters on the shape of the distributions. The final sub-tutorial presents a fire cycle analysis using the negative exponential model. In this tutorial, a time-since-fire data set is presented, and the negative exponential model is fit to the cumulative distribution. Although all the methods tutorials contain a section on assumptions and limitations, these topics are dealt with most thoroughly for this method.

While the scope of this tutorial is somewhat limited, and covers only one of the many skills required to conduct a fire frequency analysis, it does cover the particular subject matter in some depth. The material is also presented in such a way as to make it accessible to non-expert readers. This is especially important in B.C., where current forest policy uses the historical frequency of fire to make determinations about harvest levels, and requirements for seral stage distributions. Thus, people from outside the field require the ability to assess and interpret the models and methods used for historical fire frequency reconstructions.

**Overview of Chapter 3**

The primary purpose of Chapter 3 is to provide answers to the following two general questions:

1. What level of confidence should be placed in estimates of fire frequency derived from fire interval data?
2. Which sources of error in the fire history data and which aspects of the sampling scheme have the most significant influence on the uncertainty in these estimates?

To answer these questions, I:

- developed a spatial statistical simulation model of fire occurrence;
- developed models of the formation and survivorship of fire evidence (i.e., fire scars);
developed a spatial model to simulate field sampling of fire evidence;
estimated the parameters for the fire occurrence and fire evidence survivorship models from the Dugout Creek fire history data;
ran a Monte Carlo simulation of the Dugout Creek fire regime, to produce multiple synthetic realizations of fire histories that could have resulted from this fire regime.
produced realistic, synthetic fire history data sets, by applying the fire evidence formation, survivorship, and sampling models to these synthetic fire histories.
analyzed these synthetic fire history data sets to quantify a confidence interval for the original estimate of fire frequency; and
performed sensitivity analyses that vary the fire evidence formation and sampling models to determine the magnitude of effect on the confidence interval for each factor.

The results indicate that the sampling protocol applied in Dugout Creek was sufficient to provide a reliable estimate of the historical fire frequency. The results also clearly demonstrate that the greatest source of uncertainty in this estimate is the probability with which fire evidence is formed in the tree ring record. This is an aspect of fire interval analyses that has received little attention, and so I conclude by recommending that this problem be given serious study. Until a better understanding of these mechanisms exists, I recommend that researchers minimize this source of uncertainty by collecting records from several fire scarred trees at each sample point to create a more complete record of fire occurrence.
Chapter 2: An Introductory Tutorial on Common Methods for Determining Fire Frequency

Introduction

The purpose of this chapter is to present a clear and comprehensive description of common methods for computing historical fire frequency. This chapter is structured as an interactive tutorial, programmed in Excel 5.0 for Windows. Some sample pages of the tutorial are provided below and the complete, interactive tutorial is available on diskette, or on the World Wide Web at www.rem.sfu.ca/frstgrp/. The following topics are covered in the tutorial:

- An introduction to fire regimes, fire frequency, the evidence and data used to estimate fire frequency, measures of fire frequency, and fire frequency models, including a glossary and bibliography.
- A tutorial on the Natural Fire Rotation method for estimating fire frequency.
- A tutorial on the Poisson model of fire frequency.
- A tutorial on the Fire Return Interval method for estimating fire frequency.
- A tutorial on potential pitfalls of working with Fire Interval Data.
- A tutorial on the Negative Exponential and Weibull Fire Frequency Models.
- A tutorial on the Fire Cycle method for estimating fire frequency.

Motivation and Scope

I was originally motivated to produce this tutorial because I found it difficult to compare different methods for computing fire frequency presented in the literature. I also found that working with these published data sets in a consistent framework helped me to grasp some of the concepts and computations involved in estimating historical fire frequency. While the tutorial covers only a single aspect of fire history reconstruction (i.e., the mechanics of computing fire frequency), I hope that it will serve to make this portion of the science accessible to others who do not make it their career. In this section I will further outline the objectives, relevance, and scope of the tutorial.
I have two primary objectives for this tutorial:
1. to review the published fire history literature and determine the methods, underlying models, and calculations being used, along with the statistical and ecological assumptions and interpretations being made in fire frequency reconstructions; and
1. to synthesize and convey this material in a format that will help a novice reader understand what has been done, how it was done, and why it was done that way.

It is important to note that this tutorial is not meant to advocate any particular method, nor is it meant to propose new or emerging approaches to fire frequency reconstruction. It is simply a review of published material that is meant to help aid comprehension of that material.

This tutorial is particularly relevant at this time in B.C. because of new initiatives and projects underway at the Ministry of Forests. In particular, the Forest Practices Code Biodiversity Guidebook (British Columbia Ministry of Forests 1995) uses a rough fire cycle analysis employing the negative exponential fire frequency model to estimate the seral stage distributions required to meet biodiversity objectives. Any forest manager who wishes to understand how these distributions were derived, and to evaluate the applicability of this model, needs to have a clear understanding of the fire cycle method, the negative exponential model, and their assumptions and interpretation. A project currently underway in the Nelson Forest Region provides a good example (Pollack et al. in prep.). The authors of that study use a fire cycle analysis based on the Provincial forest cover maps (the FIP/SEG database) and, based on this analysis, suggest that the seral stage requirements for the region should be altered and the Annual Allowable Cut (AAC) for the region should be increased. It is very important that the people who will review this material and make decisions about these proposed changes should have access to a clear presentation of the underlying models that were used in the original analysis.

Finally, the scope of this tutorial is limited to the computation of fire frequency. An understanding of these computations is only one of the many tools and skills required to reconstruct a fire history. In this respect, it would be useful to include this fire frequency tutorial within a larger package of fire ecology and fire behaviour tutorials.
However, for the time being it is important to keep in mind that the scope of this tutorial is limited, and that it is designed only to provide access to understanding fire frequency analyses. The tutorial is not sufficient to equip people with all the tools they would require to undertake a fire frequency analysis.

**Example Tutorial Pages**

The following ten sample pages are representative of the material contained in the interactive tutorial. These ten pages were selected to complement references to this material in Chapter 3. Each sample page has been scaled to fit on one paper page, and thus do not reflect the actual text size in the tutorial (all actual text is in 12 point font).
Sample Tutorial Page 1: Tutorial Map

This page shows the logical layout for the tutorial and allows the user to “jump” directly to any page.
Sample Tutorial Page 2: Introduction

This page introduces the purpose of the tutorial and serves as a focal point for selecting which part of the tutorial to view next.
What is Fire Frequency?

Fire frequency simply refers to the recurrence of fire in a given area over time, or in other words: how often fires burn. You can think of fire frequency as the number of forest fires that occur over some fixed time interval, or as the average number of years between successive fires.

Obviously, the frequency of forest fire occurrence over the last few hundred or thousand years is not directly observable. So fire ecologists must employ a variety of evidence and lines of reasoning in order to reconstruct the fire dates and estimate the historical fire frequency. See the next pages for a description of the types of evidence and lines of reasoning typically employed.

The diagram below provides a conceptual model for fire frequency. This figure shows a hypothetical landscape as it evolves over time. The different colored patches represent areas that burned in different fire years (the year of the fire for each patch is given). The tables following this diagram illustrate the two types of fire history data that may be obtained by a fire ecologist working in 1990. Which data is actually available will depend on the fire regime. (See the next pages for details).

**A Conceptual Diagram of Fire Recurrence over Time**

![Diagram showing the evolution of a landscape over time, with different colored patches indicating areas burned in different fire years.](image)

**Time-Since-Fire Data**
(For this data we look only at the 1990 map)

<table>
<thead>
<tr>
<th>Current Year</th>
<th>Fire Year</th>
<th>Area in 1990 (hectares)</th>
<th>Time-Since-Fire</th>
<th>Percent of total area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>1540</td>
<td>1000</td>
<td>450</td>
<td>7.69%</td>
</tr>
<tr>
<td>1745</td>
<td>1800</td>
<td>2200</td>
<td>190</td>
<td>16.92%</td>
</tr>
<tr>
<td>1835</td>
<td>2000</td>
<td>3000</td>
<td>155</td>
<td>23.08%</td>
</tr>
<tr>
<td>1880</td>
<td>1500</td>
<td>110</td>
<td>15.38%</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>1500</td>
<td>25</td>
<td>11.54%</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>13000</td>
<td>100.00%</td>
<td></td>
</tr>
</tbody>
</table>

**Fire Interval Data**
(For this data we look at the sampling points A, B, C, & D through time)

<table>
<thead>
<tr>
<th>Data from Pnt. A</th>
<th>Data from Pnt. B</th>
<th>Data from Pnt. C</th>
<th>Data from Pnt. D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fire Year Interval</td>
<td>Fire Year Interval</td>
<td>Fire Year Interval</td>
<td>Fire Year Interval</td>
</tr>
<tr>
<td>1540</td>
<td>1560</td>
<td>1465</td>
<td>1695</td>
</tr>
<tr>
<td>340</td>
<td>240</td>
<td>280</td>
<td>140</td>
</tr>
<tr>
<td>1880</td>
<td>1800</td>
<td>1745</td>
<td>1835</td>
</tr>
<tr>
<td>165</td>
<td>165</td>
<td>165</td>
<td>1835</td>
</tr>
</tbody>
</table>

**Sample Tutorial Page 3: What is Fire Frequency?**

An example of the basic, introductory material that is provided as background for novice readers.
Fire over Time as a Poisson Process

Abstract Concept:
If the occurrence of fire over time is a Poisson process, then the number of fires occurring in time periods of equal length, \( t \), should have a Poisson distribution with parameter \( \lambda t \). Furthermore, if the numbers of fires in each period are independent, then the distribution of elapsed time between two successive fires is a negative exponential with parameter \( \lambda \). This is the basis for the negative exponential fire frequency model!

To test the hypothesis that fires occur as a Poisson process, we would need to collect a long record of fire occurrence at a single point, divide the record into equal sized periods, and count the number of fires in each period. Unfortunately, the record of fires at a point location is very limited, and thus it is impossible to actually gather the data we would need to model point fire frequency in this way. However, by pooling the fire dates over a wide area, we can obtain a suitable area-based data set for this example.

Please note: This example is to illustrate a concept only! Because the fire dates are pooled over a wide area, the results must not be interpreted as a point fire frequency. See the special tutorial on Working with Fire Interval Data for details.

Even though this method is not practical in reality, it is a very useful exercise because it clearly demonstrates the relationship between fire as a Poisson process and the negative exponential model of fire frequency!

Basic Method:
I. Estimate the empirical distribution of "fires per decade":
   1) Collect a complete inventory of fire dates for a location (the location's fire chronology).
   2) Divide the time period spanned by this fire chronology into decades (or some other equal sized intervals), and count the number of fires that occurred in each decade.
   3) Count the number of decades in which zero, one, two, etc. fires burned.

II. Fit a Poisson model to this "fires per decade" histogram:
   The Maximum Likelihood Estimate (M.L.E.) for the Poisson parameter, \( \lambda t \), is simply the mean number of fires per decade from the empirical data set. Use the M.L.E. \( \lambda t \) to produce the Poisson model.

III. Graph the "fires per decade" histogram as: \( n \times p(n) \)
   (where \( n \) = fires per decade; \( p(n) \) = proportion of decades with \( n \) fires).
   along with the Poisson model distribution.

IV. Do a \( \chi^2 \) goodness-of-fit test to determine if we should accept or reject the hypothesis that the fires per decade were drawn from this Poisson distribution.

V. Make an interpretation of the model based on:
   1) the model parameter (\( \lambda \)); and
   2) the model's coincidence with the original data (\( \chi^2 \) fit).

Method Name: Poisson Process Model
References: Agee (1993, ch.4) mentions it - No other known references.

Introduction:
The foundation of the fire frequency models we examine in other sections of the tutorial is the Poisson process. Although modeling fire occurrence directly as a Poisson process can yield an estimate of fire frequency, other models derived from the Poisson process allow for a more comprehensive analysis. Nonetheless, it is worth understanding the Poisson model as background for these other models.

In this section, we will examine two methods for directly fitting a Poisson model to a fire interval data set and interpreting the fire frequency from this model. In addition, we provide an overview of the Poisson process and distribution, in case you need a refresher!!

Sample Tutorial Page 4: Poisson Tutorial
This page outlines a basic method for computing fire frequency that relies directly on modelling fire as a Poisson process (see also sample tutorial page 5 and 6).
An Overview of the Poisson Process

The Poisson Process over Time

In this example, we will model the occurrence of events over time as a Poisson process. In this case, our “bins” (from above) are time periods of equal size \( t \) (e.g., one year). Each “ball” is some specific event that will occur zero or more times in the time period, (e.g., a lightning storm).

- We take a sample of equal sized time periods (e.g., \( t_0 \) to \( t_{20} \)) and count the number of events that occurred in each period.
- Then, we count the proportion of periods in which zero, one, two, etc. events occurred.
- When we graph the histogram of “Number of Events” X “Proportion of Periods”, we should get an approximation of the shape of the distribution from which the samples were drawn.
- The Maximum Likelihood Estimate (M.L.E.) of the Poisson parameter, \( \lambda \), is the average number of events in a time period.

When we graph the histogram of “Number of Events” X “Proportion of Periods”, we should get an approximation of the shape of the distribution from which the samples were drawn.

- We can construct the M.L.E. Poisson distribution with this parameter \( \lambda \) and use a \( \chi^2 \) (chi-squared) goodness-of-fit test to determine if we should accept or reject the hypothesis that the samples were in fact drawn from this Poisson distribution.

We have constructed an interactive simulation of this process:

- In this simulation you will:
  1) select a random sample of time periods from a “model” Poisson distribution;
  2) graph the histogram of “empirical” samples along with the “model” Poisson distribution from which they were drawn; and
  3) do a \( \chi^2 \) goodness-of-fit test between the “model” and “empirical” distributions.

Below, we look at two examples of data modeling using the Poisson distribution...

- First, the classic Poisson process “experiment”:
  - You have a large number of bins and a large number of balls.
  - You randomly drop each ball into a bin such that every bin has an equal probability of receiving each ball.
  - After all the balls are dropped, you count the number of bins with zero balls, the number of bins with one ball, two balls, etc.

Then:

- The number of balls in a bin, \( X \), is a Poisson random variable.
- The number of bins with \( X \) balls follows a Poisson distribution, \( p(X) = \) the probability of finding exactly \( X \) balls in a bin.
- The expected number of balls per bin is the simple mean \( E(X) = \Sigma X/N \) (where \( X \) is the number of balls in the \( i^{th} \) bin, and \( N \) is the number of bins).

The Poisson distribution (probability density function) is:

\[
p(x) = \frac{e^{-\lambda} \lambda^x}{x!}
\]

where \( \lambda \), the distribution’s parameter, is the rate per unit time or area with \( E(x) = \lambda! \)

Sample Tutorial Page 5: Poisson Overview

This page serves as a brief refresher on Poisson processes.
Sample Tutorial Page 6: Poisson Simulation

This page runs an interactive simulation of a Poisson process using a random number generator. The user can select the parameter, $\lambda$, for the Poisson process, and generate a small sample of Poisson random numbers. The graph shows the theoretical Poisson distribution along with the “empirical” sample. Users can get an idea of the range of variability exhibited by a sample of this size.
**Method Name:** Mean Fire Interval Method


**Abstract Concept:**
We have an area with relatively frequent fires and an extensive record of past fires. The dates of multiple fires are recorded in individual tree ring records (e.g. as fire scars) throughout the study area. We would like to know how many years, on average, between fires at any given point in the study area. The statistical population we are interested in is the complete inventory of fire intervals at every point in the area. If we knew this, we could fit a distribution to the intervals and compute the average interval between fires for the area, or the mean fire interval.

To obtain the data for this method we take a random sample of fire intervals across the study area and use this sample to estimate the fire frequency for the whole study area. Fire interval data are usually derived from evidence of fire found in tree-ring records. The tree-ring sections must be collected in the field and then prepared, and cross-dated in the lab.

**Basic Method:**
I. Select a set of randomly located, equal sized “sample plots”. Collect a complete inventory of fire dates for each plot (the plot's “fire chronology”).

II. For each plot, compute the fire intervals from this fire chronology. Pool the fire intervals from all plots to construct a composite fire interval histogram (distribution of fire intervals).

III. Graph the fire interval distribution as: $t \times f(t)$

IV. [where $t$=interval in years; $f(t)$=Proportion of occurrences of that interval in the composite fire interval histogram.]

V. Fit a theoretical model to the fire interval distribution:
   - Common models used are: the empirical (approx. Normal) and Weibull. Each of these models has a probability density function, $f(t)$, which is fit to the fire interval distribution.
   - The data may require spatial or temporal subdivision of data to get significant fit. The model will only give a good fit in spatio-temporal regions that are relatively homogeneous with respect to their fire frequency.

VI. Make an interpretation of the model based on:
   1) the model parameters (scale & shape); and
   2) the model's coincidence with the original data (fit).

**Sample Tutorial Page 7 : MFRI Tutorial**
This page introduces and outlines the method for computing fire frequency from fire interval data (as in Chapter 3). Sample tutorial pages 8 through 10 are taken from the fire-interval tutorial.
The Mean Fire Return Interval is very easy to compute: \[ \text{MFRI} = \frac{\sum x_i}{N} = 14.48 \text{ years} \]

However, if the distribution of fire intervals is not Normal, this may be a poor estimate of the central tendency. Fitting a model of Fire Frequency to the data also allows for a more complete interpretation.

**Fire Frequency Model:**

Fitting a theoretical fire interval distribution to the data gives a frequency model of mortality, \( f(t) \). The fire interval distribution function, \( f(t) \), is called the mortality function because for any interval, \( t \), \( f(t) \) gives the probability of experiencing a fire in an interval of length \( t \).

To the right, we use a Maximum Likelihood Estimation (M.L.E.) technique to fit the data to a Weibull model of mortality:

\[ f(t) = \left( \frac{ct^{c-1}}{b^c} \right) \exp \left( -\frac{t}{b} \right)^c \]

This best-fit Weibull function is superimposed on the original interval histogram in the graph below.

For more details about the Weibull mortality function, run the special tutorial on Fire Frequency Models.

Sample Tutorial Page 8a: *Fire Frequency Model*

This page, along with sample tutorial page 8b, shows the Weibull fire frequency model fit to an empirical fire interval distribution. It gives the simple computation of mean fire return interval (MFRI) and a description of the process used to fit the fire frequency model.
Sample Tutorial Page 8b: Fire Frequency Model (cont.)

This page is a continuation from the right-hand side of sample tutorial page 8a. It shows some of the data used in the computation of the Weibull maximum likelihood estimate (MLE), along with the formulas used for doing so.
Sample Tutorial Page 9 : Interpretation

This page shows some interpretations of the fire frequency model shown on sample tutorial page 8, including the goodness of fit test. (Note that the data for the Chi-squared test is not shown for brevity’s sake!)
Assumptions, Strengths and Limitations of the Mean Fire Interval Method:
The following items should be considered before selecting the mean fire interval
method, and fully accounted for when interpreting the results of a fire interval analysis.

Assumptions:
The major assumptions of this method and the models it uses are:

- The record of all past fires is reasonably well preserved at each sample site.
- For the exponential model: All stands have equal flammability, regardless of age -- ignition locations are drawn at random from a uniform distribution.
- For the Weibull model: Stands have differential flammability based on age only -- ignition locations are drawn at random from a polynomial distribution based on age (where age is defined as time since last fire).

Strengths:
The main advantage of this method is that it works directly with a series of fire dates at each location. This has several important implications:

- The researcher is able to fit the fire interval distribution (from the fire frequency model) directly to the data, making modeling and model interpretation fairly straight forward.
- Fire frequency can be computed for small areas (in fact it is computed separately for each sample location!).
- The data gives the researcher both spatial and temporal resolution on the fire frequency. Thus changes in fire frequency over time, or differences in fire frequency between two locations are easily tested in a formal hypothesis testing framework.

The models employed by this method are based on modeling fire as a Poisson process, and thus have some nice properties (see Poisson tutorial).

Limitations:
This method suffers from the following limitations:

I. The models require a large fire interval data set, with multiple events recorded at each sample location.
   Many forests experience high-intensity fires that effectively “erase” evidence of previous events. This approach will not be useful in forests that do not maintain a long record of multiple fire events at each sample location.

II. A significant bias and/or misinterpretation may result if not all fire dates at a site are discovered.
   The major assumption of this method is that, for each location, every past fire has been detected. Since the evidence of past fires is at least partially erased by subsequent fires, and trees are not perfect recording instruments (i.e. not every fire leaves a scar on every tree), the data required for this method is often incomplete. There will be a tendency to underestimate fire frequency, and this tendency will increase as for older time periods.

III. Pooling fire dates from several trees to overcome the above problem must be done with great care, and the results can sometimes be difficult to interpret.
   There are a number of potential difficulties in analyzing composite fire interval data. For a detailed description of these problems, run the special tutorial on working with fire interval data.

Sample Tutorial Page 10: Assumptions
This page covers the assumptions that were made in the analysis of the fire interval dataset, along with some of the strengths and weaknesses of this method.
Chapter 3: Testing Methods for Estimating Fire Frequency from Fire Interval Data

Introduction

All studies of fire history rely on a series of inferences based on a set of physical evidence left by fire. In forests where the fire regime is dominated by frequent, low-severity fires, the physical evidence of fire occurrence is usually limited to anomalies in the ring structure of individual trees, such as fire scars, where patches of cambium have been killed, and/or patterns of suppression and release in radial growth (Arno and Sneck 1977, Kilgore and Taylor 1979, Laven et al. 1980, Madany et al. 1982, McBride 1983, Dietrich and Swetnam 1984). This tree-ring record contains inherent errors, often of unknown magnitude. These deficiencies in the physical evidence are due to both the inconsistency with which fires are recorded in the tree rings and the limited lifespan of this record. Other errors are introduced when the tree-ring record is sampled to create a data set, the history of fire occurrence is estimated from this data set, and inferences are made about the historical fire regime from this history.

Estimates of the historical frequency of fire in forests with a low-severity fire regime are usually computed using some variant of the Fire Interval method discussed in Chapter 2 (e.g., Arno and Sneck 1977, Kilgore and Taylor 1979, Dietrich 1980, Grissino-Mayer 1995, Heyerdahl 1997, Riccius 1998). Most researchers acknowledge uncertainty in these estimates and caution readers to the potential for bias in their results, but few have attempted to estimate the magnitude of this uncertainty. By contrast, in forests with a high-severity fire regime, time-since-fire samples or maps are typically employed to compute fire frequency using the Fire Cycle method discussed in Chapter 2 (e.g., Johnson 1979, Masters 1990, Johnson and Larsen 1991). Several authors have used simulation models to demonstrate or estimate the uncertainty in measures of fire frequency derived from time-since-fire data (e.g., Van Wagner 1978, Baker 1992, Finney 1995, Boychuk et al. 1997, Li et al. 1997, Lertzman et al. 1998).
In this chapter, I develop an approach to assessing the uncertainty in estimates of fire frequency derived from fire interval data. The foundation of this approach is a model of fire frequency, based on the statistical characteristics of fire regimes often described in fire history studies, coupled with a simple, spatial model of typical fire history sampling schemes. In addition, I propose statistical models of fire scar formation and survivorship. Using these models, I construct a series of computer simulations, based on a fire history study in Oregon’s Blue Mountains (Heyerdahl and Agee 1996, Heyerdahl 1997). These simulations are used to create synthetic fire interval data sets similar to those collected in the fire history study. Analyses of these synthetic data sets are, in turn, used to determine the level of uncertainty that we should expect in the original estimates of fire frequency for the study area. A sensitivity analysis of the model parameters provides an indication of the most important factors contributing to this uncertainty.

**Sources of uncertainty in estimates of historical fire frequency derived from fire interval data**

Trees are not perfect recorders of fire. While the mechanisms of fire scarring are fairly well understood in the laboratory (e.g., Gutsell and Johnson 1996), little is known about the probability of a fire scar forming in the field. A low-severity fire sometimes kills a portion of the cambium at the base of a tree, without killing the tree itself. As this wound heals, a “fire scar” is formed in the tree-rings that can be accurately dated using dendrochronology (Stokes and Smiley 1968, Madany et al. 1982, McBride 1983). There is some evidence that the probability of a fire initially scarring a previously unscarred tree is very low (personal communication Peter Impara). Once a tree has been scarred, however, the thinner bark covering this scar makes the tree more susceptible to being scarred again by a subsequent fire (see references in Gutsell and Johnson 1996). Thus, fire history researchers generally sample multiply scarred trees to reconstruct the occurrence of fire over time (e.g., Arno and Sneck 1977, Kilgore and Taylor 1979, Grissino-Mayer 1995, Heyerdahl 1997, Riccius 1998). These multiply scarred trees are often referred to as “recorder trees”, but little is known about the susceptibility of these trees to scarring,
and thus recording the occurrence of successive fires. In this study, I assume that trees with more than one scar will record some, but not necessarily all, subsequent fires.

Two key issues arise from the previous discussion that have important implications for this study. First, the probability of a tree recording a fire (i.e., forming a fire scar) for the first time is different from the probability of the tree recording subsequent fires. For simplicity, I assume in this study that there are an adequate number of recorder trees distributed across the landscape at all times. With this assumption, each fire that occurs has a similar opportunity to leave a record of its occurrence. Since a fire history sampling scheme will concentrate on these recorder trees, we need only be concerned with the probability that a previously scarred tree records a fire. However, this raises the second issue — the probability that a previously scarred tree will actually record the occurrence of a particular fire is most often unknown. While many researchers assume this probability is less than one, few have attempted to estimate its actual value (e.g., Dietrich 1980, Arno and Petersen 1983, Agee 1993). Note that fire evidence can never be recorded unless there was a fire (barring the misinterpretation of non-fire related tissue damage). Thus, the physical evidence is biased towards under-representing the occurrence of fire, and the magnitude of that bias is often unknown.

The tree-ring record of fire occurrence also has a limited lifespan. Trees die and, barring fossilization, decay. The record of fire in a badly decayed log is often irrecoverable. Thus, the maximum lifespan of a tree, plus the time required for such a tree to decay sufficiently, limits the temporal depth of a tree-ring record. In actuality, the record will vary in length from tree to tree, with the depth of record in each sample being no greater than this theoretical upper limit. In addition, a particularly severe fire may burn and obliterate the evidence of previous fires. This “record erasure” (Weisburg 1997) occurs when a recorder tree or log is consumed by a fire, or when the scar lobe of a previous fire is burned off. Thus, the physical evidence experiences “mortality”, and we should be able to construct mortality and survivorship models for the evidence itself.

These errors are inherent in the physical evidence left by fire: further errors may be introduced if the fire history sampling scheme is insufficient to capture the full range of
variability in fire frequency, or to detect all fires that did leave a record. While a spatially
intensive survey may be employed to estimate the error introduced by the sampling
scheme (e.g., Morrison and Swanson 1990), the degree of uncertainty introduced by
deficiencies in the physical data is difficult to assess from the fire history data itself.

**Models and Methods**

I built two computer simulation models to manufacture synthetic fire interval data
sets similar to the empirical data sets typically collected and used to reconstruct fire
history for a low-severity fire regime. The first model simulates a low-severity fire regime
and records the dates and locations of all fires that burned during the simulation. The
second model censors and samples this complete synthetic fire record. It simulates the
formation and survival of fire scars in the tree-ring record and then sub-samples this
record based on a fire history sampling strategy. Together, these two models produce a
“synthetic fire history” data set that can be analyzed in an identical manner to an empirical
fire interval data set (Figure 1).

These two models use probability functions to capture the variability apparent in
the system. Using Monte Carlo simulation to generate a number of “replicate” fire
histories, I derive an estimate of the expected variability and bias in fire frequency
measures computed from the simulated system and sampling strategy. The following
subsections describe in detail the development of these models.

**REFR -- A Spatially Explicit, Stochastic Fire Regime Simulation Model**

*Fire regime models* should be distinguished from *fire models*. The latter are
generally mechanistic models concerned with predicting the behaviour and/or effect of fire
over a discrete period of time, for a fixed set of fuel and weather conditions (e.g.,
Rothermel 1972, Keane et al. 1990, Finney 1996). By contrast, fire regime models are
generally stochastic, and more concerned with the potential range of long term spatial and
temporal dynamics, given a general range of vegetative and climatic conditions (e.g.,
Baker 1992, Boychuk et al. 1997). Efforts to incorporate the mechanistic details of a fire
model into a larger scale fire regime model are underway, but the data requirements and computational overhead involved make this challenging (McKenzie et al. 1996).

Here, I propose a simple fire regime model based on statistical distributions that are often described in empirical fire history studies. The key aspect of this model is that it reproduces the long-term, statistical characteristics of the fire regime under study. The foundation of this model lies in the relationship between fire extent and fire frequency, described by equation 1.

\[
MFRI = \left( \frac{R \text{ years}}{E \text{ ha}} \right) * A \text{ ha}
\]

where \( MFRI \) is the Mean Fire Return Interval (1 / point fire frequency) in years; \( R \) is the average number of fire-free years between years in which a fire burned anywhere in the study area (Return Interval for fire on the study area); \( E \) is the average area burned per fire year in hectares (Fire Extent); and \( A \) is the total size of the study area in hectares.

The Natural Fire Rotation (NFR), Fire Cycle (FC), and Mean Fire Return Interval (MFRI) (described in Chapter 2) are commonly reported measures of point fire frequency -- that is the inverse of average time between fires at a point in the study area. Simple point fire frequency, however, is not a good model of fire occurrence over time and space because it does not describe the spatial contagion and temporal discreteness of fires. Equation 1 shows that point fire frequency can be decomposed into two more basic parameters -- the average fire extent, \( E \), and the average time between fire years in the study area as a whole, \( R \). The size of the study area, \( A \), enters equation 1 because both \( E \) and \( R \) are scaled by this value (for details, see Chapter 2 “Working with Fire Interval Data”, or Dietrich 1980, Arno and Petersen 1983, Agee 1993).

If we imagine \( E \) is actually the complete distribution of fire extents and that \( R \) is the distribution of intervals, then the left hand side of equation 1 would give us a distribution of point fire intervals over time and space. The two distributions, \( R \) and \( E \), along with some concept of how individual fires occupy space (e.g., their shape and contiguity), provide a more complete model of fire occurrence over time. To actually construct the distribution of point fire intervals requires stochastic simulation, and some assumptions about the inter-dependence of the distributions \( R \) and \( E \).
I constructed a computer simulation of the Return-Extent Fire Regime (REFR) model, above, using SELES (Fall and Fall 1998). REFR is a raster-based, spatially explicit, stochastic simulation model. The parameters for the model can be estimated from empirical fire-interval data sets. Each of these parameters generally represents a set of ecological and physical processes acting at a particular scale. While these processes are not modeled directly, their composite influence on the fire regime is captured by a simple function or probability distribution. The REFR requires 4 such composite parameters:

1. The **Return Time (RT)** parameter specifies the distribution of intervals between fire years in the study area, or the “area frequency” (*sensu* Agee 1993). This parameter is highly scale dependent (see Chapter 2 or Agee 1993) and the data used to estimate its value must be matched with the scale of the model. The RT is driven by large-scale processes responsible for the occurrence of fire through time. This parameter may be thought of as the frequency of fire-conducive weather, or ignition sources, on an annual scale.

1. The **Event Extent (EE)** specifies the distribution of total area burned per fire year. Each fire has a set of ignition points, called openings, and fires spread from these ignition points until, cumulatively, the required extent is burned. To mitigate edge effects, opposite edges of the landscape are contiguous (i.e., when a fire burns off one edge of the landscape, it wraps around to the opposite edge, as in Boychuk 1997). The EE is driven by the range of meso-scale conditions, such as physiography, weather and fuel conditions, responsible for determining the duration and extent of fire events.

1. Together two parameters determine the spatial pattern of the area burned in a single fire year: the **Event Openings (EO)** specifies the number openings per fire event; and the **Fire-Shape-Complexity (SC)** specifies the complexity of edge shape for each opening. Rather than simulating specific processes of fire behaviour, these parameters simply control the number of start points and the number of neighbours to which a burning fire spreads. These two parameters represent the fine-scale processes responsible for the configuration and pattern of a burned area. This pattern may be
related to spatial heterogeneity in the fuel, topography, or canopy structure, in addition to the specific behaviour of the fire event, such as spotting or crowning.

Figure 2 shows a conceptual diagram of the REFR simulation model. The time-since-fire is stored for each cell in the model. When a fire burns in a cell, the time and location of the fire is recorded, and the time-since-fire for that cell is reset to zero. During the simulation, the parameters described above are treated as independent, random variables, where the value of the variable is drawn from the specified distribution whenever it is required. Ignition locations and fire spread directions are chosen at random. This gives us a simple starting point, however it is important to keep in mind that in the real system these parameters may be inter-dependent and the behaviour of fire may be dependent on the time-since-last-fire. While it is possible to construct a SELES model that incorporates these interactions, I use only the simple case in this project.

REFR produces a complete, spatially referenced record of fire occurrence for the duration of the simulation. This synthetic record of fire occurrence is called the “Complete Event Record” in Figure 1. I used this complete record to verify the REFR model by reconstructing the realized Return Time and Event Extent distributions for a number of different simulation scenarios. I then compared these realized distributions to the original input model parameters to verify that the model was behaving correctly.

EVA -- a Stochastic Model of Error Sources in Fire History Data

The EVA model employs two sub-models to construct a realistic, synthetic data set from the complete event record produced by REFR (Figures 1 and 3). The complete event record is considered an ideal fire record because from it we can obtain a complete list of fire years for any point on the landscape, or a complete list of locations burned for any given fire year. However, this is not the record available to a fire ecologist in the field. In reality, the length of the fire record is limited by the period covered in the tree ring record. Furthermore, some fires fail to leave evidence (e.g., a fire scar) at every location they burned. I constructed a stochastic model of these processes to degrade the ideal fire record. A second model samples this degraded record based on a specified
sampling strategy, to yield a synthetic data set similar to the empirical one. Figure 3 shows a conceptual diagram of these two EVA sub-models, each of which is described in more detail below.

**A Model of the Formation and Survival of Fire Evidence in the Tree-Ring Record**

The “EVA Fire Record Degradation” sub-model removes a sub-set of the fire dates from the ideal record produced by REFR (Figure 3). Two composite parameters are required for this sub-model. The “survivorship function”, $S$, specifies the distribution of survival times, into the past, for individual tree ring records of fire. Note that this function does not describe the expected time a newly created record will survive into the future. Rather, it describes the expected number of years that a fire record will extend into the past. This function can be estimated from the distribution of ages, in years before present, of the oldest scar on each fire history sample. The “fire scar recording rate”, $p$, specifies the probability that evidence of a fire will be recorded in the tree ring record (e.g., probability of a fire scar forming on a previously scarred tree). I have been unable to locate any studies or methods that might be used to estimate this parameter. Thus, I developed a method based on the same principles used to estimate population sizes in animal ecology using mark-recapture methods (e.g., Krebs 1989). This method uses the relationship between the number of trees sampled at a point and the total number of unique fire years identified on all such samples to estimate the total number of fires that burned at that point over the period common to all samples. The proportion of fires recorded by each tree yields an estimate of the fire scar recording rate for that tree. A second method, based on the same relationship, directly yields a single estimate of the fire scar recording rate averaged over all samples. See appendix B for a derivation of these two methods.
A Model of the Fire Interval Sampling Strategy

The “EVA Fire History Sampling” sub-model simulates the data collection and analysis stages of a fire history study (Figure 3). The three parameters for this sub-model describe the sampling strategy. The “sampling density” specifies the number of sample sites at which fire records will be collected, while the “sample layout” specifies how these sample sites are distributed in space (i.e., randomly vs. systematically). Together, these two parameters are used to build a list of cells in the REFR raster to serve as sampling sites. Note that the size of each cell in the REFR raster must be the same as the size of the sample sites! A third parameter, $N_r$, specifies the number of trees to be sampled at each sample site.

The EVA model treats each cell in the REFR raster as a potential fire history sampling site, with $N_r$ recorder trees on each site. Initially, each of these recorder trees contains the ideal or complete fire record for its site. These complete records are then degraded independently, based on the probability functions defined for the evidence survivorship, $S$, and the fire scar recording rate, $p_r$. In other words, for each recorder tree, a length of record is randomly selected from $S$, and all fire dates older than this are removed from that tree’s record; each of the remaining fire dates is removed from the tree’s record with a probability of $1 - p_r$. The set of fire dates that remain form the record of fire for each recorder tree. The fire dates from all $N_r$ trees on a sample site are then pooled to form the synthetic record of fire for that site. This record represents the best possible history of fire available to the field ecologist. A spatial sample of this synthetic fire record yields the synthetic data set for the model run (Figures 1 and 3). This synthetic data set can then be analyzed in a manner equivalent the analysis of the original empirical fire history data set (e.g., the point fire frequency for the simulation might be estimated from the synthetic data set). Note that this model assumes that all fire evidence is dated correctly. No provisions are made for errors introduced by mis-dating fires, or mis-interpreting non-fire related scars. (See Chapter 2, or Gara et al. 1986, Agee 1993, for details on these avoidable sources of error.)
Case Study -- A model for Dugout Creek.

A recent fire history study at Dugout Creek, in Oregon’s Blue Mountains, (Heyerdahl and Agee 1996, Heyerdahl 1997) provided me with an ideal data set with which to develop a case study for these methods. The rigorous methods used to collect and analyze these data made it very appropriate as a basis for a trial of this methodology. In this section, I describe the Dugout Creek study area, the fire history data set for the area, and the data modelling necessary to derive the REFR and EVA model parameters from this data set. I then use these models of the Dugout Creek fire regime and sampling strategy to answer the following questions:

1. Are the fire regime parameters reconstructed from the original data internally consistent (i.e., can we reasonably expect to replicate the reconstructed fire regime with a model parameterized from the empirical fire history data)?

1. What is the expected range of variability in the fire regime over time?

1. What degree of confidence should be placed in the point frequency estimate computed from the original data?

1. What factors have the largest influence on our confidence in the parameter estimates?

1. Was the sampling design used to collect the fire history data optimal?

The Dugout Creek study area is approximately 21,000 acres (51,900 ha) and is located in an area with gentle topography and predominantly dry ponderosa pine (Pinus ponderosa) and Douglas-fir (Pseudotsuga menziesii) forests. These forests historically experienced a low-severity, stand-maintaining fire regime (Heyerdahl and Agee 1996, Heyerdahl 1997). The fire history sampling was conducted in 72 one acre plots distributed in a regular pattern over the study area. In each plot, an average of three fire scarred trees were sampled. The fire scars were crossdated and fire years from all trees on the same one acre plot were pooled together to give the history of fire for that plot. The 255-year period from 1645 to 1900 was deemed be fairly homogeneous, with respect to fire frequency, and to have a sufficiently rich record of fire to allow further analysis. This “period of reliability” is used for all the data modelling described below. These data
provide an estimate of the point fire frequency (mean fire return interval or MFRI) during this period for the study area, by way of the Fire Interval Method outlined in Chapter 2.

**Simulating the Dugout Creek Fire Regime**

This section describes the data modelling that was undertaken to parameterize the REFR model for the Dugout Creek fire regime. Because the REFR model is a realization of the temporal Poisson process described in Chapter 2 (Poisson Models), I first needed to determine if the recurrence of fire over time in Dugout Creek could be modeled as such. I constructed a histogram of the number of fires per decade (Figure 4). A Chi-squared goodness of fit test did not reject the null hypothesis that the number of fires per decade had a Poisson distribution ($p > 0.9$). Thus, I assume that the arrival of fires over time in the study area can be adequately modeled as a Poisson process.

I reconstructed the **Return Time (RT)** distribution by pooling the record of all fire years detected anywhere in the study area. The intervals between successive fire years were computed, and a histogram of interval sizes constructed. If a Poisson process is an adequate model for the occurrence of fire over time (as above), then this interval distribution should be approximately exponential (Devore 1982). Figure 5 shows the empirical RT interval histogram for Dugout creek along with the negative exponential fit to these data, derived from the Maximum Likelihood Estimate (MLE). (See appendix A for a derivation of the MLE exponential histogram used in the simulation.) A Chi-squared goodness of fit test did not reject the null hypothesis that the intervals were drawn from this distribution ($p > 0.9$).

Similarly, I reconstructed the **Event Extent** distribution by using Maximum Likelihood Estimation to derive parameters for a Weibull distribution from the raw fire extent data. Since I had no expectation for the shape of this distribution, I selected a Weibull distribution because of its flexibility. Figure 6 shows the empirical fire extent histogram for Dugout creek (in 1000 acre size classes), along with the MLE Weibull distribution for these data. A Kolmogorov-Smirnov goodness of fit test did not reject the null hypothesis that the original fire extents were drawn from this distribution ($p ≅ 0.5$).
Finally, I used maps of the historic fires (Heyerdahl and Agee 1996) to estimate the emergent spatial heterogeneity in fire behaviour. The distribution of Event Openings shows that in almost 70% of the fire years the burned area represented a single, large, contiguous patch. To simplify the model, only one opening was created per fire event. I do not expect this to significantly affect the results. In addition, a qualitative assessment of the shape of the individual fires was used to estimate an appropriate value for the Fire-Shape-Complexity parameter. The value selected results in fires that are similar in shape to those reconstructed for Dugout Creek — not simple squares, but not highly convoluted either.

The parameters described above (estimated parameters, $A_0$, in Figure 1) provide the complete specification for the REFR model. One-hundred replicates of this fire regime model were run on a landscape of 140 x 150 (= 21,000) one acre cells, to approximate the size and shape of the empirical study area. This landscape was homogeneous with respect to fire occurrence and spread, because the gentle topography at Dugout Creek was thought to have little influence on these processes. Each replicate was run for 500 years, with year 500 representing 1994; and the 255 year period from year 151 to year 406 representing the period of reliability for Dugout Creek, 1645 to 1900. Each of these 100 replicates is one instance of the stochastic process defined by the REFR model. Similarly, the empirical history of fire occurrence could be considered a single instance of the stochastic process defined by the fire regime acting at Dugout Creek. If we assume that the REFR model yields a reasonable representation of variability in this fire regime, then the 100 simulation replicates provide a means for studying its statistical properties (i.e., we can interpret the variability exhibited by the 100 simulation replicates as an estimate of the expected variability in the natural system). For most of the analyses described in the Results section, the complete fire records from the 100 replicate histories are sub-sampled using the EVA models described below.
Simulating the Dugout Creek Fire History Sampling

This section describes how the EVA model was parameterized for the Dugout Creek fire history study. The parameters for the fire history sampling sub-model were easily obtained directly from the sampling strategy used in Dugout creek. In the real study, an average of three fire scarred trees were sampled in each plot. Although the actual number of trees sampled varied from two to five per plot, in the model exactly three trees are sampled in each plot. This provides the same number of samples, but they are distributed slightly differently across space. In the model, fire scarred trees are distributed evenly across space, all have equal probability of recording fires, and all have the same survivorship function. Since these assumptions may or may not hold true in reality, it is possible that the different sampling intensity at each site in the empirical study either introduced or compensated for a bias. This potential source of error was not investigated in this study.

The sample plot density of 72 one-acre plots and the systematic layout of the plots (as opposed to random) were duplicated in the model. However, in the empirical study these plots were allocated across space at two scales (fine and coarse). To simplify the model, the plots were evenly distributed across space (Figure 7). I do not expect this difference to have any significant impact on the results.

The two remaining parameters for the fire record degradation sub-model were estimated from the empirical data. I used a Weibull function to represent the fire evidence survivorship parameter, S, because the Weibull is commonly used as a survivorship model (Cox and Oakes 1984) and gives a good fit to the data (Figure 8). For each sample tree, the date of the first fire scar indicates the establishment date of the record for that sample. I used these establishment dates to construct an empirical cumulative survivorship function (\textit{sensu} Johnson and Gutsell 1994). This function, \( Y = P(X) \), gives the proportion of samples, \( Y \), that survived at least \( X \) years \textit{into the past} (thus in terms of typical failure time analysis, the establishment date actually represents the failure time for each record). Figure 8 shows the empirical survivorship data along with the MLE Weibull distribution fit to these data. Neither a Chi-squared nor a Kolmogorov-Smirnov goodness of fit test
rejected the null hypothesis that the survivorship data were drawn from this distribution (p>0.9 for both tests). This theoretical function can now be interpreted as giving the probability that a randomly drawn sample tree will have recorded fires for at least X years.

Fortunately, the structure of the sampling strategy used at Dugout Creek also allowed me to estimate $p_r$, the probability that a recording tree would record a subsequent fire. Specifically, both methods for estimating $p_r$ require that multiple trees be collected at a single site. In the first method, the fire dates on each recorder tree are treated as a random sample from the “population” of years in which fire burned the site. A Schnabel mark-and-recapture calculation (Krebs 1989) is used to estimate the size of this population (i.e., the number of fires that burned on the site over the period of study). An estimate of $p_r$ can then be derived for each tree on the site by dividing the number of fires recorded on the tree by the number of fires that burned on the site. This calculation was performed for each tree on each of the 72 sites sampled at Dugout Creek. Figure 9a shows the distribution of proportion of fires recorded by each sample tree along with the MLE normal distribution fit to this data. Only sites where at least 10 fires burned are included in this graph, to avoid spurious values caused by extremely small sample sizes.

In the second method, I rely more directly on the relationship between the number of trees sampled on a site, $t$, and the total number of unique fire dates observed from those trees, $U_t$. Note that the number of new unique fire dates detected (and add to $U$) with each new tree sampled will decrease until all the fires that burned the site are detected, after which no new fire dates are added with any additional sample trees. This relationship may be described in theory by the function:

$$G(t; p) = \frac{(1 - (1-p)^t)}{(p^t)}$$

and in practice by the equivalent empirical quantity:

$$G_t = \frac{U_t}{F_t}, \quad \text{[for } t = 2, 3, 4...]\); \text{ where }$$

$F_t$ is the total number of fire scars found on $t$ trees from the same site; and $U_t$ is the number of unique fire dates from all $F_t$ scars.

The parameter, $p$, of $G(t)$ fit to the empirical values $G_t$ yields a single estimate of $p_r$ for all trees and time periods included in the computation. Figure 9b shows the function $G(t)$ fit
(via least-squares) to the empirical data, $G_i$, for Dugout Creek. See Appendix B for a complete development and derivation of these two methods.

To my knowledge, neither of these two methods has been tried before, and the results need to be empirically tested. Thus, the parameter value $\hat{p}_r$ cannot be assumed robust, yet it is a critical parameter for the EVA model. While it is encouraging (if not somewhat expected) that both methods yield the same estimate, $\hat{p}_r=0.56$, I use the estimate $\hat{p}_r$ as a “best guess” only.

Summary of REFR and EVA model parameter values for Dugout Creek:

The Return-Extent Fire Regime model was parameterized as follows:

- Model size, $A = 140 \times 150 = 21,000$ one acre cells (homogeneous landscape)
- Return-Time is drawn from an Exponential histogram: $p(\text{RT}) = \text{Exp}(\text{RT}; \lambda)
  \begin{align*}
p(\text{time to next fire is } x \text{ years, where } x=1,2,3,...) = \text{Exp}(x; 3.68) &= 
  e^{-(x-1)/\lambda} - e^{-x/\lambda} \\
\text{where } &\lambda = \frac{-1}{\ln(1-1/3.68)} \quad \text{(see Appendix A)}
\end{align*}
- Event-Extent is drawn from a Weibull distribution: $p(\text{EE}) = w(\text{EE}; \alpha, \beta)$
  \begin{align*}
p(\text{fire burning } x \text{ cells}) = w(x; 0.9, 4915) &= 
  0.9 \times \frac{0.9-1}{4915} \times e^{-(x/4915)^{0.9}}
\end{align*}
- Event-Openings is always one: $\text{EO} = 1$.
- Fire-Shape-Complexity is set to give moderately complex fire shapes:
  \begin{align*}
\text{SC} = 0.5 \quad \text{(spread fire to half of neighboring cells, on average)}
\end{align*}

The parameters for the EVA sub-sampling model that were derived from the empirical data and sampling design used at Dugout Creek form the “Base Case” sampling scenario:

- Evidence survivorship is drawn from a cumulative Weibull: $p(S) = W(S; \alpha, \beta)$
  \begin{align*}
p(\text{sample tree recording for at least } x \text{ years}) = W(x; 4.37, 331) = e^{-(x/331)^{4.37}}
\end{align*}
Probability of a tree recording a fire is assumed to be constant for all trees:
\[ p(\text{individual tree recording a particular fire}) = 0.56 \]

Number of trees per sampling plot, \( N_t = 3 \)

Sampling density = 72 one acre plots

Sample layout = uniform

I use the 100 synthetic fire history data sets that result from this “Base Case” sampling scenario to answer questions about the quality of the empirical MFRI estimated for Dugout Creek (Figure 1). A few notes of caution are warranted: I have assumed that the parameters reconstructed from the empirical data set, and subsequently used to parameterize REFR, are a fairly good approximation of the true fire regime parameters. An inadequate empirical sample or a mis-match between the scale of the sampling design and the scale of the true fire regime, for example, would generate misleading results. I have also assumed that the EVA model incorporates into the synthetic samples all of the important sources of error and bias present in the empirical sample. If the errors in the synthetic samples are not distributed similarly to those in the empirical sample, then the measures of uncertainty that I propose tell us little about the quality of the empirical observations.

To answer other questions about the expected variability in the fire regime, the relative importance of various sources of uncertainty, and the adequacy of the sampling design, I use a number of other sampling scenarios. The EVA parameters for each of these other scenarios are listed in Table 1. The “Complete” scenario yields the complete fire record (Figure 1) over the 255 period of study, for all 21,000 cells with no sub-sampling. This record can be used to compute the true fire frequency realized for each of the replicate simulations. The three “Spatial” sampling scenarios yield similar complete, un-degraded records for a subset of 36, 72, and 144 of the model cells. These records can be used to determine the impact of different sampling densities in isolation from other sub-sampling mechanisms. The remainder of the scenarios vary the value of a single parameter while holding the value of all other parameters equal to those of the “Base Case” scenario. The “BaseCase36” and “BaseCase144” scenarios vary the number of sample sites (from
72 sites collected in the Base Case, to 36 and 144 sites respectively). The “RecRate.25”, “RecRate.75”, and “RecRate1” scenarios vary the fire scar recording rate (probability of a fire scarring a previously scarred tree). These scenarios use a recording rate of 0.25, 0.75, and 1 respectively, as compared to the Base Case value of 0.56. In the “Trees1”, “Trees2”, and “Trees4” scenarios one, two, and four trees, respectively, are collected at each of the 72 sample sites, as opposed to the three trees collected at each site in the Base Case scenario. These eight scenarios form a sensitivity analysis for the results of the analysis on the Base Case scenario.

Each of these scenarios takes its samples from the same set of 100 replicate simulations, at the same sampling locations, over the same 255 year period. Thus, any differences between the samples is purely an artifact of the sub-sampling mechanisms employed. It is the magnitude of these differences that allows me to determine the relative importance of each individual source of uncertainty.

Three other sampling scenarios were run over an extended temporal period. The “CompleteLong”, “SpatialLong”, and “BaseCaseLong” scenarios are identical to the Complete, Spatial72, and Base Case scenarios described above, except that the temporal extent of the record is 500 years, as opposed to 255 years in all the other scenarios. These scenarios allow us to examine the effect of the limited temporal extent of the period of study in isolation from other factors.
Table 1. Parameters for the EVA sub-sampling models discussed in the text. Each sub-sampling model is applied to the same 100 replicate Complete fire records produced by the Monte Carlo runs of the REFR model (see Figure 1).

<table>
<thead>
<tr>
<th>EVA Sampling Scenario</th>
<th>Years in Period</th>
<th># Sample Sites</th>
<th>Trees per Site</th>
<th>Probability of Recording</th>
<th>Evidence Survivorship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>255</td>
<td>21,000</td>
<td>-</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>Spatial72</td>
<td>255</td>
<td>72</td>
<td>-</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>Spatial36</td>
<td></td>
<td>36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial144</td>
<td></td>
<td>144</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base Case</td>
<td>255</td>
<td>72</td>
<td>3</td>
<td>0.56</td>
<td>W(4.37, 331)</td>
</tr>
<tr>
<td>BaseCase36</td>
<td></td>
<td>36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BaseCase144</td>
<td></td>
<td>144</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RecRate1</td>
<td>255</td>
<td>72</td>
<td>3</td>
<td>1.0</td>
<td>W(4.37, 331)</td>
</tr>
<tr>
<td>RecRate.25</td>
<td></td>
<td>36</td>
<td></td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>RecRate.75</td>
<td></td>
<td>144</td>
<td></td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Trees1</td>
<td>255</td>
<td>72</td>
<td>1</td>
<td>0.56</td>
<td>W(4.37, 331)</td>
</tr>
<tr>
<td>Trees2</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trees4</td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CompleteLong</td>
<td>500</td>
<td>21,000</td>
<td>-</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>SpatialLong</td>
<td></td>
<td>72</td>
<td>-</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>BaseCaseLong</td>
<td></td>
<td>72</td>
<td>3</td>
<td>0.56</td>
<td>W(4.37,331)</td>
</tr>
</tbody>
</table>

**Results and Discussion**

Figure 11 shows the point fire interval distributions, reconstructed from three of the “Base Case” scenario replicates, along with the true point fire interval distributions for the corresponding “Complete” replicate. The mean of each fire interval distribution gives an estimate of the MFRI for the replicate (see Chapter 2). The bias in the MFRI estimate is computed by comparing it to the true MFRI for the replicate (MFRI Bias = Estimated MFRI - True MFRI; thus, a positive bias indicates an overestimated MFRI). For each of the sampling scenarios listed in Table 1, the MFRI and MFRI Bias were computed for, and summarized over, the 100 replicates.
Table 2 gives the summary results for each scenario. For example, the first row in Table 2 presents the results for the “Complete” scenario, which express the “true” MFRI’s realized over the 255 year period of study for the 100 replicate simulations. The left-hand portion of the table shows that the replicate with the highest point fire frequency (Min. MFRI) realized a MFRI of about 11 years, while the replicate with the lowest point fire frequency (Max. MFRI) has a MFRI of just over 20 years. The average “true” realized MFRI over all 100 replicates is 14.7 years with a standard deviation of 2.22 years. Geary’s test for normality fails to reject that the MFRI’s over the 100 replicates may be drawn from a normal distribution -- in other words, the MFRI’s for the 100 replicates appear to be distributed normally about the mean. The right-hand side of the table yields a similar set of descriptive statistics for the bias in the MFRI estimates for each scenario. Because the “Complete” scenario represents the true MFRI for each replicate, there is no bias in this scenario. However, the fifth row of Table 2 gives the results for the “Base Case” scenario, and shows that the bias in the MFRI estimates from the Base Case samples range from one to just over six years. The average bias over all replicates in this scenario is about two and a half years with a standard deviation of 0.81 years. Geary’s test for normality does reject that the biases are distributed normally (p<0.05), and an examination of the biases for this scenario confirms that the distribution is actually skewed to the right of the mean (see Figure 13).

The “Mean MFRI” and “Mean bias” columns of Table 2 are discussed extensively in the following sections. All of the sampling scenarios started with the same 100 replicate fire histories represented by the “Complete” scenario. Thus, differences in average MFRI between scenarios are attributable solely to the sub-sampling mechanisms employed by the scenario. The average difference between the MFRI estimate from a sampling scenario and the true MFRI from the “Complete” scenario yields the “Mean bias” for the scenario. The magnitude of this bias indicates the relative impact of the scenario’s sub-sampling mechanisms on our ability to accurately reconstruct the fire history.
Table 2. Results from the different sampling scenarios described in Table 1.

MFRI = point mean fire return interval. Bias = estimated(MFRI) - true(MFRI).

Mean and StdDev columns give the average and standard deviation over 100 replicates, while Max and Min give the maximum and minimum value that occurred for a single replicate. Geary’s Z is a test statistic, where |Z| ≥ 1.96 rejects the hypothesis that the data are normally distributed, at the α=0.05 level. The full distribution of MFRI’s for the 100 replicates (CompleteLong, Complete, and Base Case scenarios) is shown in Figure 11. The full distribution of biases for the 100 replicates (Base Case, RecRate, and Trees scenarios) are shown in Figures 14 and 15.

<table>
<thead>
<tr>
<th>Sampling Scenario</th>
<th>Min. MFRI</th>
<th>Max. MFRI</th>
<th>Mean MFRI</th>
<th>StdDev MFRI</th>
<th>Geary's Z</th>
<th>Min. bias</th>
<th>Max. bias</th>
<th>Mean bias</th>
<th>StdDev (bias)</th>
<th>Geary's Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>10.9</td>
<td>20.7</td>
<td>14.7</td>
<td>2.2</td>
<td>0.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Spatial72</td>
<td>10.9</td>
<td>20.7</td>
<td>14.7</td>
<td>2.3</td>
<td>0.9</td>
<td>-0.4</td>
<td>0.7</td>
<td>0.0</td>
<td>0.1</td>
<td>-4.1</td>
</tr>
<tr>
<td>Spatial36</td>
<td>11.1</td>
<td>20.8</td>
<td>14.7</td>
<td>2.3</td>
<td>1.1</td>
<td>-0.7</td>
<td>0.6</td>
<td>0.0</td>
<td>0.2</td>
<td>-2.3</td>
</tr>
<tr>
<td>Spatial144</td>
<td>11.0</td>
<td>20.6</td>
<td>14.7</td>
<td>2.2</td>
<td>0.9</td>
<td>-0.2</td>
<td>0.3</td>
<td>0.0</td>
<td>0.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Base Case</td>
<td>12.4</td>
<td>26.9</td>
<td>17.1</td>
<td>2.7</td>
<td>-0.2</td>
<td>1.1</td>
<td>6.2</td>
<td>2.4</td>
<td>0.8</td>
<td>-2.5</td>
</tr>
<tr>
<td>BC36</td>
<td>12.3</td>
<td>27.0</td>
<td>17.2</td>
<td>2.8</td>
<td>-0.5</td>
<td>0.5</td>
<td>6.3</td>
<td>2.5</td>
<td>0.9</td>
<td>-2.4</td>
</tr>
<tr>
<td>BC144</td>
<td>12.5</td>
<td>26.3</td>
<td>17.1</td>
<td>2.7</td>
<td>-0.2</td>
<td>0.5</td>
<td>5.6</td>
<td>2.4</td>
<td>0.7</td>
<td>-1.4</td>
</tr>
<tr>
<td>RecRate1</td>
<td>10.8</td>
<td>21.7</td>
<td>14.7</td>
<td>2.3</td>
<td>0.2</td>
<td>-0.8</td>
<td>1.0</td>
<td>-0.1</td>
<td>0.3</td>
<td>-3.0</td>
</tr>
<tr>
<td>RR.25</td>
<td>19.6</td>
<td>41.6</td>
<td>26.4</td>
<td>4.1</td>
<td>-1.7</td>
<td>7.7</td>
<td>20.9</td>
<td>11.7</td>
<td>2.3</td>
<td>-2.0</td>
</tr>
<tr>
<td>RR.75</td>
<td>11.4</td>
<td>23.2</td>
<td>15.5</td>
<td>2.4</td>
<td>0.2</td>
<td>-0.2</td>
<td>2.5</td>
<td>0.8</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Trees1</td>
<td>17.5</td>
<td>36.9</td>
<td>23.8</td>
<td>3.7</td>
<td>0.4</td>
<td>5.9</td>
<td>16.2</td>
<td>9.1</td>
<td>1.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Trees2</td>
<td>14.0</td>
<td>27.3</td>
<td>19.1</td>
<td>3.0</td>
<td>0.5</td>
<td>2.4</td>
<td>6.9</td>
<td>4.4</td>
<td>1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>Trees4</td>
<td>12.1</td>
<td>24.1</td>
<td>16.2</td>
<td>2.5</td>
<td>0.0</td>
<td>0.4</td>
<td>3.4</td>
<td>1.5</td>
<td>0.5</td>
<td>-1.3</td>
</tr>
<tr>
<td>CmptLong</td>
<td>12.0</td>
<td>20.5</td>
<td>15.1</td>
<td>1.8</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>SplitLong</td>
<td>12.0</td>
<td>20.7</td>
<td>15.1</td>
<td>1.8</td>
<td>0.0</td>
<td>-0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>BCLong</td>
<td>13.2</td>
<td>29.5</td>
<td>17.4</td>
<td>2.5</td>
<td>-2.0</td>
<td>-0.4</td>
<td>9.3</td>
<td>2.3</td>
<td>1.3</td>
<td>-5.0</td>
</tr>
</tbody>
</table>
Consistency of fire regime parameters

From equation 1, we would expect that the empirical MFRI at Dugout Creek should be roughly equal to 15.10 years \[\text{mean}(RT)/\text{mean}(EE) \times A = \frac{3.68 \text{ years}}{5,170 \text{ acres}} \times 21,213 \text{ acres}\]. The actual empirical mean point fire return interval for the 255 year period 1645 to 1900 is 15.02 years (computed from 991 fire intervals collected at 72 sites). Furthermore, a Weibull model fit to the fire interval distribution yields a MLE MFRI of 15.09 years (Figure 10a). If these three estimates of MFRI were not similar, it would indicate that the \text{Return Time}, \text{Event Extent}, and \text{MFRI} parameters were not internally consistent, and thus at least one of them was in error. For Dugout creek, these three estimates of the MFRI are virtually identical (given the precision of the data), and thus suggest that \text{RT} and \text{EE} form a consistent set of parameters with which to model the fire frequency for the area. In addition, note that the \text{Return Time} for fires in the study area, 3.68 years, is consistent with the mean of the Poisson distribution of the arrival of fire over time (10 years per decade / 2.69 fires per decade \(\approx 3.7 \text{ years per fire}\); Figure 4). This result supports my assumption that Poisson is an adequate model for the arrival of fire over time in Dugout Creek.

One of the questionable, yet critical, assumptions of the REFR model is that fire occurrence is not dependent on time-since-last-fire. The correlation coefficient between the number of fire-free years preceding a fire (empirical RT) and the subsequent fire size (empirical EE) is 0.058 \(r^2=0.003, p=0.6, N=79;\) for the period 1645 to 1900), indicating no evidence that fire extent is dependent on the time since the previous fire year. Thus, it was reasonable to treat these two distributions as independent. By contrast, the empirical point fire interval distribution shows few intervals shorter than 10 years (Figure 10a), indicating that the probability of fire occurrence at a particular location may increase with the time-since-fire at that point. This hypothesis of increasing hazard of burning with time-since-fire is supported by the MLE Weibull distribution shape parameter, \(c\), being substantially greater than 1 (see Chapter 2, fire frequency models, or Johnson and Gutsell 1994, for details). We can use the synthetic point fire interval distributions to determine if...
this value should be regarded as significant evidence for an increasing hazard of burning function operating at Dugout Creek.

Many of the synthetic fire histories also exhibited a skewed and/or leptokurtic point fire interval distribution, similar to the empirical distribution for Dugout Creek. For example, Figure 10 shows both the empirical and a synthetic point fire interval histogram. Both histograms exhibit a similar shape and degree of variability, and both have a modal value close to 10 years. However, the expected shape of the fire interval distribution under a constant hazard of burning is negative exponential (Van Wagner 1978, Johnson and Van Wagner 1985, Johnson and Gutsell 1994). Because the hazard of burning is constant in the REFR model, some other mechanism is responsible for the skewed shape of the synthetic point interval distributions. One hypothesis is that this phenomenon results from the contagious fires in the model, which violates one of the assumptions of the Poisson model. Additional simulation experiments would be required to test this hypothesis.

While it seems evident that some mechanism other than an increasing hazard of burning may be responsible for at least some of the skewed shape of the empirical fire interval distribution, the Weibull shape parameter clearly distinguishes the empirical distribution from all of the synthetic distributions. The shape parameters for the 100 synthetic fire interval distributions range between about 0.98 and 1.31, whereas the shape parameter for the empirical distribution is 1.72, well outside the range exhibited by the simulation. These pieces of evidence suggest that while there is likely no time-since-fire dependency at the scale of the entire study area, there may be a local effect whereby a recently burned area has a lower probability of reburning. Thus, a more realistic fire regime model for Dugout Creek should incorporate an increasing hazard of burning with time-since-fire in each cell.

**Variability in fire regime**

A substantial degree of variability in measurements of MFRI is introduced as an artifact of the limited temporal and spatial extent of a study. Assuming constant
conditions over a very long period (or a very large area), the MFRI for a system should converge to that of the generating processes, which in the case of our simulation model is a Poisson process with a MFRI of 15 years. However, even when averaged over 500 years (more than 30 fire rotations!) and 21,000 sample locations, the MFRI still shows substantial variability among the 100 replicate simulated landscapes (Table 2 and Figures 11 and 12). As the period examined (or size of the study area) is reduced, this apparent variability in the MFRI among replicates increases.

Figure 11 depicts this scale effect clearly, with the distribution for the CompleteLong scenario (MFRI is averaged over 500 years, 21,000 sites) exhibiting less variance than that for the Complete scenario (where the MFRI is averaged over about half the number of years). This increasing variance would continue as we reduced the temporal extent of our observations. On the other hand, over very long periods, we’d expect that every replicate would converge to a MFRI of 15 years, with little variance among replicates.

The variance introduced by the temporal scale of the observations has important implications because it implies that substantially different historical MFRI’s can arise from the same set of generating processes (where by historical I simply mean those reconstructed over some finite region of space and time). Figures 11 and 12 show this effect clearly — substantial differences are apparent between replicate synthetic histories, although each was created by an identical generating process. Thus, investigators should avoid the temptation to infer a difference in the processes driving the fire regime based solely on a finding of significant difference between two fire interval distributions (see also Lertzman et al. 1998). Such an inference would be inappropriate, for example, when comparing two periods, because the observed difference indicates only that the historical occurrence of fires differed between periods, not necessarily that the fire regime changed. An analysis of variance between replicates from the same scenario illustrates this cautionary point clearly (Table 3). While each of the 100 replicate fire interval distributions was driven by an identical process, it is not difficult to find statistically significant differences between them. In fact, even a very small difference between the
means of two distributions may result in a significant difference (e.g., Run39 vs. Run85 in Table 3) because the number of intervals typically collected in a fire history study makes this a high power test.

**Table 3. Analysis of variance on the point fire interval distributions for several replicates of the Base Case scenario.**

<table>
<thead>
<tr>
<th>Replicate Fire Interval Distributions Compared</th>
<th>Number of Intervals</th>
<th>Mean Point Fire Intervals (years)</th>
<th>Kruskal-Wallis test significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run9 vs. Run100</td>
<td>932 &amp; 647</td>
<td>16.5 &amp; 21.7</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Run39 vs. Run85</td>
<td>1037 &amp; 932</td>
<td>14.1 &amp; 15.1</td>
<td>p = 0.001</td>
</tr>
<tr>
<td>Run9 vs. Run85</td>
<td>932 &amp; 932</td>
<td>16.5 &amp; 15.1</td>
<td>p = 0.133</td>
</tr>
</tbody>
</table>

In general, even the “true” or complete point fire interval distributions (FID) showed a surprising degree of variability, both within and between replicates. Based on the constant hazard of burning model used in the simulation (see Chapter 2 or Johnson and Gutsell 1994, for details on hazard model), I expected these distributions to exhibit a negative exponential shape, and while many did, the majority did not. In about 20% of the replicates, the mode of the FID was one year (as in a negative exponential), but in over 50%, the mode was five or more years. A few of the distributions had a bi-modal shape, whereas others were very flat, with one or two single interval “spikes” accounting for up to 15% of the total intervals each. Most had an irregular skewed-normal shape. This result indicates, as discussed in the previous section, that some mechanism other than an increasing hazard of burning can skew the shape of the FID. These observations contradict the findings of Boychuk et al. (1997).

Figure 12 shows some typical FID’s for a few of the replicates in the Complete and Base Case scenarios. This figure illustrates both the range of apparent variability due solely to the temporal scaling inherent in the observations (Complete scenario) and the way in which bias is introduced by sub-sampling these observations (Base Case scenario). Notice that the Base Case scenario tends to underestimate the proportion of small intervals (~<10 years) and overestimate the proportion of larger intervals.
Although the synthetic FID’s exhibited a fairly wide range of shapes (Figure 12), they almost all differed from the empirical FID for Dugout Creek in one respect. The maximum interval in the synthetic FID’s tended to be much longer than the maximum point fire interval found at Dugout Creek (64 years). Only one of the 100 Base Case replicate samples contained no intervals longer than 64 years (Max. interval in Run41 = 58 years). The other 99 replicates had maximum intervals of between 69 and 190 years, with an average maximum interval of 106 years for the Base Case scenario. Because the EVA model incorporated all mechanisms thought to censor the FID in the empirical sample, long fire intervals should have been detected in the empirical sample with a likelihood equal to that of the Base Case sampling scenario. If this is true, then the empirical observation of a maximum point fire interval of 64 years was very unlikely (1 in 100, or $p \cong 0.01$).

This line of reasoning suggests that either: i) long point fire intervals are more rare at Dugout Creek than in the REFR model; or ii) the evidence of longer fire intervals that did occur at Dugout Creek was censored by some mechanism not included in the EVA model. The first hypothesis is supported by the Weibull model fit to the empirical FID (Figure 10), which suggests an increasing hazard of burning with increasing age at Dugout (see Chapter 2 or Johnson and Gutsell 1994). Alternatively, the second hypothesis might suggest that fire severity increases with time-since-fire, such that the evidence of long intervals is destroyed by subsequent severe fires. This hypothesis would be supported by evidence of stands dominated by young, even-aged forest that contained few fire scarred recorder trees. While a number of such stands were found at Dugout Creek, they also lacked any evidence of a previous forest (logs, snags, rootballs, etc.). Thus, these stands were actually thought to have been un-forested prior to fire suppression (Emily Heyerdahl, personal communication). The second hypothesis would also be supported if long inter-fire intervals were less likely to be recorded for some other reason. For example, thicker bark developed in the fire-free period might reduce the probability of subsequent fires being recorded on a recorder tree. Distinguishing between these two hypotheses will require further investigation.
While the Poisson process that drives the simulated fire regime in the REFR model is admittedly simplistic, the variability it exhibits may serve as a first estimate of the expected range of variability for the processes driving the fire regime at Dugout Creek. Based on the variability exhibited by my model of Dugout Creek, the particular historical sequence of fires observed at Dugout Creek is one of a fairly wide range that could have been generated by the physical fire regime, and, conversely, could have been generated by a number of distinct fire regimes. Thus, the observed sequence of fire may be typical for the processes acting at Dugout Creek, as we have assumed, or it may represent the tail of the distribution and not the average behaviour of these processes. Due to the limited temporal extent of any tree-ring record, and/or the limited period over which the driving processes can be assumed to be relatively stable (Masters 1990, Johnson and Larsen 1991), it will be difficult to distinguish between these two cases.

This result indicates that while it may be possible to reconstruct the historical sequence of fires for an area with some accuracy, making inferences about the fire regime from this historical sequence is more difficult. In fact, it is difficult to find a meaningful, statistically satisfying definition for the term “fire regime” because the processes driving fire regimes vary over a broad range of scales (Sprugel 1991, Lertzman et al. 1998). For example, the fire regime at Dugout Creek exhibits annual variability (Figures 5 and 10a), and decadal variability (Figure 4) over the relatively short period of study. Other fire histories over longer periods of study exhibit variability between centuries (Masters 1990), and millennia (Swetnam 1993, Long et al. 1997). Interestingly, while the range of variability in fire frequency is critical for understanding the processes driving the fire regime and the longer-term (centuries to millennia) ecological consequences of the fire regime, the simple, historical sequence of fires may be of greater importance for understanding the shorter-term (decades to centuries) ecological effects of the fire regime.

**Confidence interval for point fire frequency**

Ideally, I would like to know what set of fire regimes might have resulted in the observed historical sequence of fires at Dugout Creek. Because a direct answer to this
question is intractable, I have instead asked “what set of observed MFRI estimates could result from a given fire regime?” The REF model described above is the “given fire regime”, and the 100 replicates of the Base Case sampling scenario yields a “set of observed MFRI estimates”. The errors in these synthetic observations are given by the MFRI bias (Table 2). If we assume that the errors in the synthetic observations are distributed similarly to those in the empirical observation, then the average synthetic error approximates the expected error in the empirical observation (see notes of caution in the parameter summary section above!). Furthermore, the range into which 95% of the synthetic errors fall yields a measure of how likely it is that the error in the empirical estimate also falls within that range. Since this range gives us a measure of confidence in the empirical observation, I will refer to it as the 95% confidence interval.¹

To compute a 95% confidence interval for the empirical MFRI estimate for Dugout Creek, I constructed the distribution of biases in the Base Case MFRI estimates. The bias for each synthetic sample is computed by subtracting the true MFRI for the simulation (calculated from the Complete record) from the estimated MFRI (calculated from the sub-sampled “Base Case” record). The distribution of biases for the Base Case sample appears to be a skewed or truncated normal distribution (Figure 13; Geary’s test

¹ More formally:
Define $A_{\text{nature}}$ as the true value of parameter $A$ in nature;
$A_{\text{empirical}}$ as the empirically estimated parameter value;
$A_{\text{synthetic}}$ as the true value of parameter $A$ for the synthetic fire histories;
$A_i$ as the parameter value estimated from synthetic fire history $i$; and
$B_{\text{empirical}} = A_{\text{empirical}} - A_{\text{nature}}$ as the true bias in the empirical estimate of $A$.

Given the set \{ $A_i | i=1..N_r$\}, then the bias in each synthetic estimate is
$B_i = A_i - A_{\text{synthetic}}$, and the expected bias in any synthetic estimate is
$\bar{B} = \sum B_i / N_r$. If we assume that the errors in $A_i$ are distributed similarly to those in $A_{\text{empirical}}$, then the expected bias in the empirical estimate of $A$ is: $E(B_{\text{empirical}}) = \bar{B}$.

Furthermore, define $C_{a,b}$ as the proportion of synthetic errors that fall into the range $[a..b]$: $C_{a,b} = \left| \left\{ B_i | a \leq B_i \leq b \right\} \right| / N_r$. Since $C_{a,b}$ yields an estimate of the expected proportion of trials in which the error falls in the interval $[a..b]$, I will refer to this interval as the $C_{a,b} \times 100\%$ confidence interval.
for normality rejects that the biases were drawn from a normal distribution, p<0.05). The mean bias was 2.4 years with a standard deviation of 0.8 (Base Case in Table 2). The range into which 95% of the MFRI biases fall is [1.1, 4.0] (Figure 13). I use this range, coupled with the empirical estimate of the MFRI at Dugout Creek, 15 years, to compute a 95% confidence interval for the true MFRI at Dugout Creek of 11 to 13.9 years, with an expected value of 12.6 years.

Given the number of assumptions made in these analyses, and the range of natural variability exhibited by the system, this magnitude of bias is quite reasonable and, I suspect, most researchers would be pleased with such a result. However, for critical applications of the Dugout Creek study (e.g., the development of a conservation plan for an endangered species that is fire dependent) it may be important to consider that there is a 1 in 100 chance that the empirical MFRI estimate may be biased by 40% or more (e.g., the MFRI estimate for Run82 was biased by 6 years). While providing an absolute measure of confidence in the empirical MFRI estimate for Dugout Creek, this analysis has broader consequences because it can be used to quantify the relative magnitude of a number of sources of uncertainty that have likely had a more substantial impact on other fire history studies.

**Primary sources of uncertainty in fire history studies**

To determine the primary sources of uncertainty in the estimates of fire regime parameters, I performed a sensitivity analysis on the EVA sampling model. Table 1 shows the parameter settings for the different sub-samples used in the sensitivity analysis. In each case, the sub-sample proceeds as for the Base Case, except that one parameter of interest is varied about its Base Case value. The following results simply examine the effect of varying each EVA parameter on the bias and uncertainty in the reconstructed MFRI, as compared to the Base Case scenario.

**Sampling Density**

I ran two scenarios to determine the effect of varying sampling density. While the Base Case scenario had 72 plots (1 plot per 295 acres), the BaseCase36 and BaseCase144
scenarios had half and twice that density of plots, respectively. Although I expected the sample size to influence the amount of variance in the MFRI estimates, in fact varying the sampling density had little effect (Table 2). This insensitivity likely arises because the fires tended to be much larger than the spacing between sample points (mean fire size = 5,170 acres). Thus, even at half the sampling density (1 plot per 590 acres), most fires are still sampled adequately. Note that sampling density is usually defined, as it is here, with respect to the size of the landscape. When designing a sampling strategy, it would be more useful to think about the density with respect to typical fire size (which, of course, is usually unknown before sampling begins). I expect that reducing sampling density to a degree that is low with respect to the typical fire size (e.g., one plot per \( X \) acres, where \( X \) is the average fire size) would have a significant influence on the estimate of fire frequency.

**Evidence Survivorship**

While I did not run any scenarios that varied the parameters of the evidence survivorship function, two of the scenarios demonstrate that it did not have a direct impact on the fire frequency estimate for Dugout Creek. The Spatial72 sampling scenario has perfect evidence survivorship (it is simply a spatial subsample of the complete fire record at 72 one acre sample cells), while the RecRate1 scenario applies only the survival function as a censoring mechanism to this spatial sample. There was very little difference between the MFRI estimates from these two samples (Table 2), indicating that the evidence survivorship played a minor role in affecting this estimate. This insensitivity makes intuitive sense because the tree-ring records for Dugout tend to extend further back in time than the period of analysis. Approximately 30% of the records extended to 350 or more years before present, which is the maximum extent of the period of analysis. Because three trees were sampled at each site, there was a high probability that at least one of these trees would contain a record extending to at least the beginning of the period. In fact, in a carefully analyzed fire history study, as in Dugout, this is exactly how the period of reliability is generated in the first place. So the evidence survivorship typically
plays an indirect role in censoring the fire history by primarily determining the period of reliability (see section on “Temporal Censoring” below).

It is useful to note that the RecRate1 scenario, where “evidence survivorship” is the sole censoring mechanism, is the only scenario that introduced a negative mean bias into the MFRI estimate (i.e., the MFRI tended to underestimate; Table 2). This underestimation occurs because long intervals are more “susceptible to mortality”, and are thus underrepresented in the sample. In other words, a record of fire is more likely to begin following (or end just prior to) a long interval, simply because there are more years in a long interval than a short one. This effect is present in all the sampling scenarios, it is just overwhelmed by the positive bias introduced from other sources. If the size of the longer intervals approaches or exceeds the temporal extent of the period of study, these longer intervals will not be detected. While evidence survivorship plays an insignificant role in the Dugout case study, it will be a critical factor in study areas where fire intervals are long relative to the period of analysis, or in studies that attempt to make inferences about the fire frequency outside the period of reliability.

Temporal Censoring

While the “evidence survivorship” parameter did not play a direct role in influencing the MFRI estimate at Dugout Creek, it plays a substantial indirect role by limiting the temporal extent of the analysis. As discussed above, temporal censoring of the fire history introduces both variability and a bias into measurements of fire frequency. For example, differences between the Complete and CompleteLong scenarios indicate that the MFRI’s realized over 255 years exhibit a higher variance than those computed over 500

\[ \text{Consider a short interval of } S \text{ years and a long interval of } L \text{ years. If the probability that a fire record will begin is equal in each of the } N=S+L \text{ years, then the probability that it will begin in the long interval is } \frac{L}{N}, \text{ and in the short interval, } \frac{S}{N}. \text{ The probability } \frac{L}{N} \text{ is greater than } \frac{S}{N} \text{ because } L > S, \text{ and thus it is more likely that the record will begin during the long interval. This is an oversimplification of the process of fire record establishment, but it seems likely that these statistical properties should hold in any case. Because the interval preceding the start of a fire record is not known and not included in} \]

\[ \]
years (Figure 11; \( s^2 = 4.9 \) vs. \( s^2 = 3.2 \) respectively; Table 2). Also, while on average the realized point fire frequencies in the CompleteLong scenario matched the expected input point frequency of 1/15 years, the realized frequencies for the 255 year period of interest were slightly higher, 1/14.7 years. Obviously, for the Dugout study, this 0.3 year temporal censoring effect is insignificant. It plays a minor role primarily because the 255 year period of analysis was much longer than even the longest intervals between successive fires at a point, which was about 106 years, on average. However, analyses over periods that are short with respect to the MFRI will introduce a more substantial bias because they will miss or under represent longer fire intervals.

Note that the period of reliability is primarily a property of the physical data that cannot generally be controlled by the researcher. Thus, while researchers may have good reason to examine only a short period (e.g., they suspect that the processes driving the fire regime are not temporally stable), they should be aware of this source of bias and uncertainty in their findings, although it will typically be overwhelmed by errors from other sources.

*Probability of Recording Fires in the Tree-Ring Record*

It is apparent that the only significant source of bias and uncertainty in the empirical estimate of MFRI for Dugout Creek arises from the censoring of individual fire dates within the period of reliability. While the “probability of recording” \((p_r)\) is primarily responsible for the censoring itself, the “number of trees sampled at each plot” \((N_t)\) is equally important in determining the magnitude of its effect. Increasing the value for one of these two parameters can compensate for uncertainty introduced by a lower value in the other.

Figure 14 shows the distribution of biases for the 4 scenarios in which the \(p_r\) parameter was varied. The RecRate.25 and RecRate.75 scenarios applied a recording rate of about half and twice that used in the base case, respectively, whereas in the RecRate1 analyses, there is a higher likelihood that longer intervals will be excluded from an analysis than short ones.
scenario, all fires at each site are recorded \((p_r = 0.25, 0.75, \text{and } 1.0\) respectively; Table 1). The RecRate.25 samples exhibited, by far, both the largest bias and variance between replicates of any of the scenarios. By contrast, the RecRate.75 samples exhibited the lowest bias and variance of all scenarios that applied some record degradation (Table 2 and Figure 14). The RecRate1 scenario shows that the bias introduced from all sources other than the recording rate is very small (Table 2 and Figure 14). These results suggest that the magnitude of bias and uncertainty in a fire history study is highly dependent on the true value of \(p_r\). For example, if the recording rate at Dugout is 0.25 rather than 0.56, then it is very likely that the MFRI estimate of 15 years may be out by a factor of two or more (Max. MFRI estimate = 42 years in RecRate.25 scenario; Table 2).

Further research is required to determine if the estimate, \(\hat{p}_r = 0.56\), is sound, and what range of values constitutes a reasonable confidence interval for the estimate. In addition, the “probability of recording” is likely different for each recorder tree and variable among and within different fire events. A simple simulation experiment could be used to determine if it makes a difference that this parameter was treated as a constant in this study, rather than as a “random” variable as it more likely is in nature. In any case, the probability with which fires are recorded in the tree-ring record appears to be critically important for estimating the bias in a fire history reconstruction. It is surprising, therefore, that this topic has received virtually no attention in the literature to date.

**Number of Sample Trees per Plot**

The Trees1, Trees2, and Trees4 scenarios sample one, two, and four trees in each sample cell, respectively, compared with three trees sampled in the Base Case (Table 1 and Figure 15). Relatively little is gained or lost, in terms of bias or variance in the MFRI estimate, by sampling two or four trees, rather than three trees, at each site. However, when only a single tree is sampled at each site, the bias and uncertainty in the MFRI estimate becomes very high (Figure 15). This observation likely holds true for any recording rate, \(p_r < 1\), because a single tree sample will miss \((1-p_r)\times100\%\) of the fires, on average, at each site. Each missed fire represents a compound error in the estimate of
MFRI because two true short intervals are removed from the FID in addition to one false long interval being added. For any recording rate, $p_r > 0.5$, a sample of at least two trees has a reasonable chance of detecting most fires. Because the parameter $N_t$ is under the control of the researcher, multiple trees should always be sampled at each site if one wishes to reconstruct the history of fire with greatest certainty.

**Evaluating the Sampling Design**

In terms of simply estimating fire frequency, the sampling design employed at Dugout Creek was close to optimal, although a small improvement might have been possible, in theory and with hindsight. Reducing the sampling density made very little difference to the bias in the MFRI estimate (BC36 scenario, Table 2), and so I assume that an empirical sample of 36 plots with 4 trees sampled at each plot would yield a result similar to that of the Trees4 scenario (mean bias in MFRI estimate = 1.47 years; Table 2). Thus, assuming a fixed budget (in terms of number of trees sampled), a less biased MFRI estimate likely would have resulted from a strategy that sampled more trees per plot at fewer plots. In reality, this may not have been feasible (e.g., there may not have been a suitable number of recorder trees available at each site), nor desirable from the perspective of achieving the other objectives of the study (e.g., more accurate estimates of fire extent).

**The Effect of Fire Frequency on Fire Frequency Estimates**

Holding everything else constant, the frequency of fire itself has a direct impact on the magnitude of both the bias and uncertainty in the estimates of MFRI. This effect occurs because, over the same period, an area with lower fire frequency (higher MFRI) will simply have fewer fire intervals, and thus a lower $N_t$ for the MFRI calculation. Thus, a decrease in fire frequency has an effect very similar to that of reducing the period of analysis, discussed above. Because each of the 100 replicate fire histories in my study realized a slightly different fire frequency over the 255 year period of analysis, I was able to examine this effect over the limited range of fire frequencies covered by the replicates. Figure 16 shows the relationship between MFRI and the bias in the Base Case estimate of
MFRI. The regression between MFRI and bias is significant (correlation coefficient $R^2 = 0.27$, $N = 100$, $p < 0.001$), indicating that such a relationship can be identified even over the limited range examined. Further analysis of this relationship over a wider range of conditions may yield some general principles about the uncertainty in fire frequency estimates that could help guide fire history researchers. It is also worth noting that the variance in bias, and thus uncertainty in the MFRI estimate, also appears to increase with MFRI (examine the magnitude of deviation of the data points from the trend line Figure 16). Although this relationship was not formally tested, it makes sense both intuitively (variance usually increases with increasing mean), and for some of the reasons discussed above under temporal censoring.

**Conclusions**

The results of this study indicate that there are a number of techniques that fire researchers can employ to minimize the bias in their estimates of historical fire frequency. A number of important considerations also arose that have direct implications for forest managers who wish to use the results of a fire history study to guide their management actions. In the following sub-sections, I make a set of specific recommendations that highlight the important points of my results for each of these two groups. In addition, I review the plethora of outstanding questions related to testing fire history methods raised by my analyses.

**Recommendations for Fire History Researchers**

The case study for Dugout Creek shows that it is possible to reconstruct the historical frequency of fire from fire interval data with some precision. However, the sensitivity analysis of the EVA model indicates that a sampling design that is insufficient or not well matched with the scale of the fire regime can produce very inaccurate results. Two features of the Dugout Creek sampling design substantially reduced the bias in the empirical MFRI estimate — the collection of multiple samples of the fire record at each point and the use of an iterative approach in determining the sampling density. These two
approaches have not been consistently applied in studies of fire history from fire interval data.

While it may be difficult to determine the exact probability with which fire is recorded in the tree-ring record, there is little doubt that this probability is, in general, substantially less than one. The error introduced by the fire recording rate is inherent in the physical data, and thus not under the control of the researcher. However, the collection of samples from multiple recorder trees at each sample site in Dugout Creek not only allowed me to estimate the fire recording rate, but also compensated for the errors introduced by it, and thus played a key role in minimizing bias in the empirical MFRI estimate. I strongly recommend that in all studies of fire history from fire interval data, researchers should sample multiple trees at each sampling point. The size of a “point” on the landscape (e.g., one acre at Dugout Creek) must be large enough to encompass several fire scarred trees, yet small enough that it can be assumed to be acting as a single unit with respect to fire occurrence.

Another strong feature of the sampling strategy applied at Dugout Creek was its use of an iterative approach—data analysis from the first field season was used to redesign the sampling strategy for subsequent seasons. In general, the budget for a study will, to some extent, pre-determine the number of samples that can be taken. A pilot study should be used to help determine how to best allocate these samples across space. The goal of such a pilot study should be to gain an estimate of the spatial scale and heterogeneity of the fire regime, along with an estimate of the probability with which trees record the passage of fire (see Appendix B for details). This information can then be used to determine a plot density compatible with the typical fire extent, and the number of trees that need to be sampled in each plot to achieve an acceptable level of certainty in the results. At Dugout Creek, the results from the first season indicated, quite correctly, that the investigator should reduce plot density and sampled a wider area in subsequent field seasons (Emily Heyerdahl, personal communication).

Finally, temporal censoring, due to the limited lifespan of fire evidence, plays a key role in increasing the variance apparent in the fire interval distribution. Fire history
researchers will find themselves in a dilemma, caught between wanting to expand the temporal period of analysis (to reduce the variance in their MFRI estimate), yet needing to restrict the temporal period of analysis such that they have an adequate sample at each spatial location over the entire period. The choice of period of analysis is further complicated if the processes driving fire regimes are hypothesized to have changed over the period of study (Johnson and Gutsell 1994, Lertzman and Fall 1998). This is a difficult problem for which I have no direct recommendations, other than to select a period of analysis by objective criteria that are sensible for the type of question being asked.

Implications for Management

It will be difficult to formulate timber harvesting objectives based on low-severity, stand-maintaining disturbance regimes because the fires that characterize these regimes tend to remove the unmerchantable trees, while leaving most of the mature trees intact. The Biodiversity Guidebook (British Columbia Ministry of Forests 1995) designates the Interior Douglas Fir and Ponderosa Pine Biogeoclimatic zones as Natural Disturbance Type 4 (NDT4). According to this designation, “frequent, stand-maintaining fires” dominate the disturbance regime in these forests (see p. 39 of Biodiversity Guidebook). The return interval for low-intensity surface fires in NDT4 is estimated between 4 and 50 years for these forests. However, the seral stage distribution for these forests are modeled on a 250 year average return interval high-intensity, stand destroying fires. While my research indicates that uncertainty exists in the fire frequency estimates for these forest types, the designation of a 250 year disturbance return interval for NDT4 forests blatantly ignores the fire history research for these forests. I recommend that the goals for timber management in NDT4 systems be revised to better reflect the reality of the disturbance regime in these forests. These goals should include maintenance of forest structure, species, and communities, which in turn will require that harvesting rates and techniques are adapted to replace the role of fire. Such management strategies will need to be based on fire history research and will need to account for the uncertainty in the disturbance frequency estimates.
Furthermore, the range of variability in NDT4 systems makes the traditional approach of working with mean disturbance intervals questionable. My research demonstrates that these forests experience a wide range of inter-fire intervals, across both space and time. This variance is likely as important to the ecology as the mean, because it promotes a diversity of species and vegetative responses to fire. This diversity may, in turn, allow these ecosystems to adapt to a change in fire regime (e.g., brought on by a change in climate). The Biodiversity Guidebook recognizes this point by emphasizing seral stage distributions rather than a fixed rotation period. However, there have been many difficulties interpreting the meaning of the seral stage distributions in the Biodiversity Guidebook, and, in practice, variability is not being implemented across the landscape — each forest block is being operationalized with the mean (John Nelson, personal communication). More research is required to determine how to best base a management strategy on a low-severity fire regime, and how to incorporate variation into the forest operations.

My finding that there may be an increased hazard of burning with time-since-fire, or that severe fires may ensue after long periods without fire has important fire management implications for Dugout Creek, in which only two significant fires have burned since 1900. The role of fire suppression in increasing fuel loads and thus the potential for large, catastrophic fires is well documented in the literature and deserves serious consideration by forest managers.

Future Research

My analyses raised a number of interesting questions and novel applications for the modelling framework that I developed for this project. This framework could easily be applied to other study areas, and could also be used to:

- determine the cause of the skew and kurtosis common in the synthetic point fire interval distributions and relate this effect to hazard of burning function in models of fire frequency (e.g., Weibull);
- determine the optimal sampling density with respect to typical fire size;
• determine the minimum period of reliability with respect to the mean and/or longest fire intervals, required to adequately reconstruct fire frequency;

• generate the expected range of variability for two reconstructed fire regimes to yield a measure of variance that can be used to test for differences between two fire regimes.

• simulate a fire regime on a heterogeneous landscape and/or changes in fire regime over time to determine effectiveness of methods in detecting these differences;

• develop methods for recognizing empirical samples that may be extremely biased by analyzing a set of synthetic samples that are outliers on the MFRI bias distribution for common characteristics;

As a final note, it may be possible to derive an analytic relationship between the probability of recording, the number of trees in each sample, and the expected bias in the MFRI estimate, holding all else equal (e.g., MFRIbias = f(MFRIestimate, RecRate, TreesPerSample)). As a first approximation, the models described herein could be used to generate an “empirical” approximation of such a function. Having such a function would greatly assist researchers in designing optimal sampling strategies, and estimating the bias in their reconstructed estimates of fire regime parameters, without having to resort to the Monte Carlo analyses conducted in this project.
Figure 1. Conceptual diagram for the project. Each box in this diagram represents some object of study. The arrows should be read as "produces via" the mechanism specified for the arrow. The empirical fire history study is depicted above the heavy dashed line. Through a set of ecological and physical processes, the fire regime leaves evidence of fire occurrence in the tree rings. A researcher samples this evidence to create an empirical fire history data set. This data set is, in turn, used to estimate the parameters of the fire regime, $A_0$. The models used in this project are shown below the heavy dashed line. The empirical parameters, $A_0$, are used in the REFR model to simulate the fire regime and produce a synthetic realization of fire occurrence. The EVA model then censors and samples this complete record to yield a synthetic data set similar to the empirical one. Synthetic parameters $A_{s1}$ ($i=1..N$) may be derived by the same methods used to derive $A_0$. These synthetic parameter estimates may then be used to form a confidence interval for the empirical estimate, $A_0$. 

![Diagram of empirical fire history study and corresponding models](Image)
Figure 2. Conceptual diagram of REFR simulation. When the simulation starts, the time of the first fire year is selected from the Return Time distribution (Figure 5). When a fire year occurs, a fire size is selected from the Event Extent distribution (Figure 6). A random start location is chosen for the fire, and it spreads out from this cell until it burns the number of cells required by its Extent. The time until the next fire year is then selected from the Return Time distribution. The complete record of all cells burned by each fire is the “Complete Event Record” in Figure 1. (A map of time-of-last-fire is also kept but not used in this project.)
Figure 3. Conceptual diagram of EVA model. Both the uncertain states of nature and the sampling options are parameters for this model. The “Complete Event Record” is produced by the REFR fire regime model (Figure 2). The EVA record degradation sub-model introduces errors into this record based on the types of error thought to be present in the physical data. The EVA fire history sampling sub-model then selects a sub-sample of the spatial locations to simulate data collection by the fire history researcher.
Figure 4. The distribution of fire years per decade at Dugout Creek between 1645 and 1900. The Maximum Likelihood Poisson distribution is also shown. Note that 2.69 fires per decade translates into approximately 3.7 years between fires (10/2.69 ≈ 3.7).
Figure 5. Distribution of intervals between fire years at Dugout Creek between 1645 and 1900. The Maximum Likelihood negative exponential distribution is also shown. Random numbers are selected from this “Return Time” distribution in the REFR simulation to determine the time to next fire year on the landscape (Figure 2).
Figure 6. The distribution of fire extents at Dugout Creek between 1645 and 1900. The Maximum Likelihood Weibull distribution is also shown. Random numbers are selected from this “Event Extent” distribution in the REFR simulation to determine the spatial extent of each fire event (Figure 2).
Figure 7. Sampling design used in the EVA sampling model (Figure 3). The inset on the lower right shows the approximate layout of plots in empirical study at Dugout Creek for comparison.
Figure 8. The cumulative distribution of survival times for fire evidence at Dugout Creek. The Maximum Likelihood Weibull survivorship model is also shown. Random numbers are selected from this distribution in the EVA record degradation sub-model to truncate the synthetic complete fire records (Figure 3).
Figure 9. (a) The proportion of fires recorded by trees sampled at Dugout Creek; and (b) the function $G(t)$ fit to the empirical estimate $G_t$ for all Dugout Creek sample sites (see Appendix B). In (a), only those sets of trees with at least 10 total fires in their common period of record are included, yielding $N=190$ sets. In (b), all 72 sites are included, but there are fewer sites where four and five trees shared a common period (i.e., there were few sites where $G_4$ and $G_5$ could be estimated). Thus, the least-squares fit function, $G(t)$, is most heavily influenced by $G_2$ and $G_3$. 

$p_r = \text{Proportion of Fires Recorded by a Tree (mean = 0.558, std. dev. = 0.203)}$
Figure 10. Point fire interval distributions for (a) the empirical data set for Dugout Creek and (b) the synthetic Base Case scenario replicate Run73.
Figure 11. Distribution of mean fire return intervals (MFRI) resulting from 100 replicates of the CompleteLong (a), Complete (b), and Base Case (c) scenarios. None of these distributions were rejected as being normal by Geary's Z (p>0.4 in all cases).
Figure 12. Synthetic point fire interval histograms from three of the 255 year replicates. Each graph shows the Complete interval histogram (grey bars) and the reconstructed interval distribution from the Base Case sampling scenario for the same replicate (black bars), along with the MLE Weibull distribution fit to the Base Case distribution (rejected with p<=0.01 in...
each case). Also shown are the number of point intervals in the sample (N), the mean and maximum point interval in the sample, and the proportion of intervals not shown on the graph (> 64 years, the maximum interval found at Dugout). The MLE Weibull parameters and the Kolmogorov-Smirnov goodness of fit test values for each replicate are also given.

Figure 13. Distribution of biases in the estimates of MFRI for the 100 replicates of the Base Case scenario, showing the 95% interval for the histogram. Note that this is slightly different (skewed right) from the 95% interval for MLE normal distribution.
Figure 14. Distribution of biases in MFRI estimates for 100 replicates of the EVA sampling scenarios that vary the probability of recording fire evidence. A perfect estimate has a bias of zero. The wider the distribution, the higher the uncertainty resulting from samples of this type. The further the distribution's mean is shifted from zero, the greater the expected bias from samples of this type. (An approximation to the MLE normal is also shown for each scenario. See Table 2 for the exact mean, standard deviation, and significance of fit.)
Figure 15. Distribution of biases in MFRI estimates for 100 replicates of the EVA sampling scenarios that vary the number of trees collected at each plot. A perfect estimate has a bias of zero. The wider the distribution, the higher the uncertainty resulting from samples of this type. The further the distribution's mean is shifted from zero, the greater the expected bias from samples of this type. (An approximation to the MLE normal is also shown for each scenario. See Table 2 for the exact mean, standard deviation, and significance of fit.)
Figure 16. Correlation between the true MFRI and the magnitude of bias in the MFRI estimate. The true MFRI is computed from the Complete sample, over 255 years, for each of the 100 replicates. The MFRI estimate is computed from the Base Case samples. The bias is the difference between the true and estimated MFRI. While there is a substantial amount of variation, the trend itself is significant (N=100, $R^2=0.27$, F=35.8, p<0.001).
Appendix A

Transforming an exponential distribution into a histogram

There is a complication with using an exponential distribution to represent the Return Time parameter, RT. The empirical return times are discrete integers and never have a value less than one, due to the resolution of tree ring data. Thus, the complete histogram of return times forms a probability mass function (pmf) with mean $\mu$:

$$h(x; \mu) = P(X = x) = \text{proportion of RT intervals equal to } x \text{ years } \{ x = 1, 2, 3, \ldots \}$$

The expected value of $h(x)$ is:

$$E(h(x)) = \sum_{x=1}^{\infty} x * h(x) = \mu \quad (A1)$$

The probability distribution used to represent RT in the simulation, $p(x)$, should have the same properties as $h(x)$ — that is it returns integer values $\geq 1$ with the same shape and expected value, $E(p(x))=\mu$. Although the exponential model is suggested by the Poisson process and gives a good fit to $h(x)$ (i.e. it is the correct shape), it is a continuous probability density function (pdf) with mean $\lambda$, defined on $[0 \leq x < \infty]$:

$$p(x; \lambda) = \frac{1}{\lambda} e^{-x/\lambda} \quad \text{for } x \geq 0 \quad (A2)$$

Two questions arise: first, what is the best method to fit the continuous exponential model $p(x; \lambda)$ to the discrete pmf $h(x; \mu)$?; and second, how can we draw random integer values on $[1, \infty]$ from $p(x; \lambda)$ with an expected value of $\mu$? To solve these problems, we require a discrete probability histogram, $h'(x) = f(p(x; \lambda))$, such that $E(h'(x)) = \mu$. In words, we need a theoretical pmf (histogram) with an exponential distribution and the same mean as the empirical pmf. To find such an $h'(x)$, we note from equation A1 that:

$$E(h'(x)) = \sum_{x=1}^{\infty} x * f(p(x)) = \mu \quad (A3)$$

Because $p(x)$ is a continuous distribution, an obvious choice for $f(p(x))$ is:

$$f(p(x)) = \int_{x-1}^{x} p(x'; \lambda) \, dx \quad \text{[for } x=1,2,3\ldots \] \quad (A4)$$

By substituting in the r.h.s. of A4, and solving the equality in Equation A3, we get:
\[
\lambda = \frac{-1}{\ln(1 - 1/\mu)} \quad \text{(see proof at end)} \quad (A5)
\]

With \( \lambda \) defined as in A5, we get the MLE pmf fit to the empirical pmf:

\[
h'(x; \mu) = \int_{x-1}^{x} \frac{1}{\lambda} e^{-x'/\lambda} dx = e^{-(x-1)/\lambda} - e^{-x/\lambda} \quad \text{[for x=1,2,3,...]} \quad (A6)
\]

We can select random values from \( h'(x) \) during the simulation, and the resulting RT distribution will have an expected shape and mean of that of the empirical distribution.

**Proof:**

\[
\lambda = \frac{-1}{\ln(1 - 1/\mu)}
\]

We defined the pdf, \( h'(x) \), such that:

\[
E(h'(x)) = \mu = \sum_{x=1}^{\infty} x \cdot P(X = x),
\]

where \( P(X=x) \) is the proportion of intervals of size \( x \), which is the area under the histogram bar \( x \) (see figure).

We note that for each integer \( i \), an exponential pdf, \( p(x; \lambda) \) defines:

\[
A_i = \int_{i-1}^{i} \frac{1}{\lambda} e^{-x/\lambda} \quad [i=1,2,3,...]
\]

where \( A_i \) is the area under \( p(x) \) on the interval \((i-1, i)\) (see figure).

We require \( p(x; \lambda) \) such that \( A_\chi = P(X=x) \). To obtain the desired \( \lambda \), we simply substitute \( P(X=x) = A_\chi \) in the first equation above, and solve the equality:
\[ \mu = \sum_{i=1}^{\infty} i \cdot A_i \]
\[ = \sum_{i=1}^{\infty} i \cdot \int_{i-1}^{i} \frac{1}{\lambda} e^{-x/\lambda} \]
\[ = \sum_{i=1}^{\infty} i \cdot \left( e^{-(i-1)/\lambda} - e^{-i/\lambda} \right) \]
\[ = \left( e^{-0/\lambda} - e^{-1/\lambda} \right) + \left( 2e^{-1/\lambda} - 2e^{-2/\lambda} \right) + \left( 3e^{-2/\lambda} - 3e^{-3/\lambda} \right) + \ldots \]
\[ = \sum_{x=0}^{\infty} e^{-x/\lambda} \]
\[ = \sum_{x=0}^{\infty} \left( e^{-1/\lambda} \right)^x \]
\[ \quad \text{Sum of a geometric series } \sum_{0}^{\infty} r^x = \frac{1}{1 - r} \]
\[ \Rightarrow \]
\[ \mu = \frac{1}{1 - e^{-1/\lambda}} \]
\[ e^{-1/\lambda} = 1 - \frac{1}{\mu} \]
\[ -\frac{1}{\lambda} = \ln(1 - 1/\mu) \]
\[ \lambda = \frac{-1}{\ln(1 - 1/\mu)} \]
Appendix B

Computing the probability of a fire scar forming

Application problem:
• Given a set of \( T \) trees on a site, that were all fire recorders for a period of \( Y \) years, yielding a list of fire years recorded by each tree over the period,
• assume that for each fire year in the period, the whole site burned;
• assume that in each of \( N \) fire years during the period, each tree acted as an independent recording device to record the fire with some unknown probability, \( p_r \).
• We would like to know the values of \( N \) and \( p_r \).

Equivalent counting problems to derive an estimate of \( N \):

In the case where \( T = 2 \), this problem is very similar to a classic counting problem (adapted from Constantine 1987):
• Persons \( A \) and \( B \) independently proofread a book (all errors are assumed to be independent and equally likely to be found).
• Person \( A \) finds \( a \) errors and \( B \) finds \( b \) errors, with \( c \) errors spotted by both \( A \) and \( B \).
• What is the number of errors, \( N \), in the book?

The solution is fairly simple and intuitive:
⇒ Note that the probability \( A \) finds a randomly selected error is \( a/N \), and for \( B \) it’s \( b/N \).
⇒ The number of errors found by both will be, on average, \( (a/N) \times b \) [because \( A \) should find approximately \( a/N \) of those errors found by \( B \)].
⇒ Thus, we solve \( c \equiv a/N \times b \) for \( N \) and get: \( N \equiv a \times b / c \)

Ecologists may recognize this solution as the Petersen method for estimating population abundance using the mark-and-recapture technique. In this case, \( a \) is the number of individuals caught and marked in the first capture, \( b \) is the number of individuals caught in the second capture, and \( c \) is the number of marked individuals, re-
captured in the second capture (Krebs 1989). \( N = a*b / c \) gives us an estimate of the total population size. We may also calculate a confidence interval for \( N \) (see Krebs 1989).

For the application problem, \( a \) is the number of fire years recorded on tree \( A \) over the period \( Y \), \( b \) is the number of fire years recorded on tree \( B \) over the same period, and \( c \) is the number of fire years recorded by both \( A \) and \( B \). \( N = a*b / c \) gives us an estimate of the total number of fire years for the site, over the period \( Y \).

In fact, there may be more than two sample trees on a site (i.e., \( T > 2 \)). In this case, the Schnabel mark-and-recapture method (Krebs 1989) serves as an appropriate model. This method simply treats the \( T \) samples as a series of Petersen samples, and estimates the population size with a weighted average of Petersen estimates:

\[
N \cong \frac{\sum_{t=1}^{T} \left( C_t * M_{t-1} \right)}{\sum_{t=1}^{T} R_t}
\]

where \( C_t \) is the total number of fire years recorded by tree \( t \) (e.g., \( C_A = a \)), \( M_{t-1} \) is the number of unique fire years in the pooled record of \( t-1 \) trees, and \( R_t \) is the number of “re-captured” fire years on the \( t \)th tree (already in \( M_{t-1} \)).

**Estimating a value for \( p_t \), the probability of recording a fire:**

It is possible to use the estimate of \( N \), above, to derive an estimate of \( p_t \) for each tree, \( \hat{p}_t \), by simply dividing by \( N \) the number of fires recorded by the tree during the period \( Y \) (e.g., \( \hat{p}_t = C_t / N \)). Because different trees cover different periods, each tree, \( t \), may have several different estimates of \( \hat{p}_t \), so the final value of \( \hat{p}_t \) would actually be a weighted average of these estimates.

In contrast, the following method derives \( p_t \) directly, without relying on \( N \). Although the derivation below uses only trees from a single site and period, as above, the final curve fitting may be done using the values from any number of sites and/or periods. This method yields a single, “average” value of \( p_t \) for all trees and periods included in the computation, rather than a separate probability, \( \hat{p}_t \), for each tree, \( t \).
Problem Development:

- Define $C_t$, $R_t$, and $M_t$ as above.
- Let $S_t$ = the set of all fire dates “captured” by $t$ trees (including duplicates), and let $U_t$ = the set of unique fire years in $S_t$. Then, $\sum C_t$ = the size of set $S_t$, and $M_t$ = the size of set $U_t$.
- Then, define $f_t = M_t / N$, to give the proportion of fire years found after the $t^{th}$ tree is examined.
- When $t = 1$, $M_t \equiv p \times N$ and $f_t \equiv p$.
- As $t \to \infty$, $M_t \to N$, and $f_t \to 1$.
- Furthermore, the theoretical function $f(t)$ yields an estimate $f_t$:
  \[
  f(t) = 1 - (1 - p)^t
  \] (see proof at end)
- If the function $f(t)$ were fit to the empirical data, $f_t$, we could solve for $p$. However, the empirical quantity $f_t$ cannot be derived directly because $N$ is unknown.

To solve this dilemma, a quantity, $G_t$, that depends only on the available empirical data ($C_t$, $R_t$, and $M_t$) is required, along with the theoretical function, $G(t)$, that estimates it.

- One such quantity is $G_t = M_t / \sum C_t$, the proportion of all “captured” fire dates that are unique (i.e., the proportion of $U_t$ in $S_t$).
- Since $M_t \equiv f(t) \times N$ and $\sum C_t \equiv p \times q \times N$, we get the theoretical function $G(t) = f(t) / p^t$.

For one value of $t$, on a single site, over a single time period, $p_t$ could be derived directly from the empirical data, by solving $G(t) = G_t$ for $p$ (although I have not found an analytic solution for this equation). However, $p_t$ can also be estimated by fitting the function $G(t)$ to a plot of all values of $G_t$ computed for various values of $t$ at a number of sites and/or time periods. This yields an estimate of the average value of $p_t$ over all trees and time periods included in the computation.
Assumptions:

Both of the methods described above make two key assumptions that are most likely not borne out in nature. Empirical studies will be required to determine how adversely departures from these assumptions in nature affect the results:

1. Each tree is assumed to act as an independent recorder. This assumption will not hold if, for example, a fire burns in such a way that it scars more trees in one area than it does in some other area (i.e., the rate of fire scarring is spatially autocorrelated).

2. Each fire is assumed to have an equal probability of being recorded. This assumption will not hold if, for example, one fire burns in such a way that it scars fewer trees than another fire (i.e., the rate of fire scarring varies among fire years).

A third assumption that must be carefully controlled for during data sampling is that recorder trees grouped together on the same site are assumed to have experienced the exact same history of fire — if a fire burned one tree on a site, then it burned all the trees on a site. Thus, sample sites should be as small and homogeneous as possible.

Note that the methods do not assume that each tree has an equal probability of recording a fire (just as readers A and B had different probabilities of detecting errors). Thus, the rate of fire scarring may vary between trees and so trees of different species or with different bark thickness may be included in a single computation.
Proof: \( f(t) = 1 - (1-p)^t \)

We require a function \( f(t) \equiv f_i \), where \( f_i = M_i / N \), the proportion of fires found after the \( t^{th} \) tree is examined.

Let \( p = \) the probability that a tree will record any given fire year.

Then \( f_{i+1} \equiv f_i + p(1-f_i) \)

If \( f(t) \equiv f_i \) then \( f(0) = 0 \), and \( f(t+1) = f(t) + p(1-f(t)) \)

Which yields: \( f(t) = p \sum_{i=0}^{t-1} (1-p)^i \) for \( t \geq 1 \), and 0 otherwise.

Proof by induction:

\[
\begin{aligned}
f(1) &= f(0) + p(1-f(0)) = p = p(1-p)^0 = p \sum_{i=0}^{0} (1-p)^i \\
\text{Assume } f(t) &= p \sum_{i=0}^{t-1} (1-p)^i ; \text{ then } \\
f(t+1) &= f(t) + p(1-f(t)) \\
&= f(t)(1-p) + p \\
&= \left(p \sum_{i=0}^{t-1} (1-p)^i\right)(1-p) + p \\
&= \left(p \sum_{i=1}^{t} (1-p)^i\right) + p \\
&= p \sum_{i=0}^{t} (1-p)^i
\end{aligned}
\]

To obtain the closed form for \( f(t) \), we solve:

\[
(1-p) \cdot f(t) - f(t) = p \sum_{i=1}^{t} (1-p)^i - p \sum_{j=0}^{t-1} (1-p)^j
\]

\[
f(t) - p \cdot f(t) - f(t) = p(1-p)^t - p
\]

\[
f(t) \cdot (-p) = p\left((1-p)^t - 1\right)
\]

\[
f(t) = 1 - (1-p)^t.
\]
Literature Cited


Riccius, E. 1998. Scale issues in the Fire History of a fine grained landscape. MRM Thesis Simon Fraser University, Vancouver, B.C.


