Spring 1-1-2017

A Cross-Scale Assessment of Historical Fire Severity Patterns, Landscape Dynamics, and Methodological Challenges in Mixed-Severity Fire Regimes of the Northern U.S. Rockies

Cameron Edwards Naficy

University of Colorado at Boulder, cameron.naficy@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/geog_gradetds

Part of the Environmental Microbiology and Microbial Ecology Commons, Forest Sciences Commons, and the Geography Commons

Recommended Citation


This Dissertation is brought to you for free and open access by Geography at CU Scholar. It has been accepted for inclusion in Geography Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
A CROSS-SCALE ASSESSMENT OF HISTORICAL FIRE SEVERITY PATTERNS, LANDSCAPE DYNAMICS, AND METHODOLOGICAL CHALLENGES IN MIXED-SEVERITY FIRE REGIMES OF THE NORTHERN U.S. ROCKIES

by

Cameron E. Naficy

B.A., Rice University, 2003
M.Sc., University of Montana, 2008

A thesis submitted to the faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirement for the degree of Doctor of Philosophy Department of Geography

2016
This thesis entitled:

A cross-scale assessment of historical fire severity patterns, landscape dynamics, and methodological challenges in mixed-severity fire regimes of the northern U.S. Rockies

written by Cameron E. Naficy

has been approved for the Department of Geography

__________________________
Thomas T. Veblen

__________________________
Waleed Abdalati

__________________________
Stefan Leyk

__________________________
Tania Schoennagel

Date____________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
ABSTRACT

Naficy, Cameron E. (Ph.D., Geography)

A cross-scale assessment of historical fire severity patterns, landscape dynamics, and methodological challenges in mixed-severity fire regimes of the northern U.S. Rockies

Thesis directed by Professor Thomas T. Veblen

Forests characterized by mixed-severity fire regimes (MSFR) exhibit high spatio-temporal variability of fire frequency and severity. These forests comprise much of western North America, but their ecological dynamics, spatial ecology, resilience mechanisms, and biophysical drivers are poorly understood. MSFR forests provide rich spatio-temporal data that predate landscape changes and alterations to disturbance regimes that have occurred since Euro-American colonization. However, the relative scarcity of large historical datasets and some persistent methodological limitations have impeded robust reconstructions of historical MSFRs, creating substantial debate and confusion. We address these problems by developing an extensive dataset of dendroecological records and geospatial maps of historical forest conditions for MSFR forests in seven watersheds across two distinct study regions in the northern U.S. Rockies. We develop a framework for scaling dendroecological reconstruction methods along the fire regime gradient that accounts for the increasing loss and spatial bias of historical data that occurs for MSFRs. We use this approach to reconstruct spatio-temporal patterns of fire severity-mediated dynamics for a subset of forest patches throughout each watershed. These data reveal novel landscape dynamics and a resilience mechanism
to high severity fire that has not been previously documented. Finally, we use the
dendroecological records as a validation dataset to develop a calibrated structure-
based fire severity model based on the photo-interpreted structural attributes and
patch boundaries in our geospatial maps. The resulting model reveals complex
relationships between forest structures and fire severity history, but provides a
relatively strong method for reconstructing spatial patterns of fire-severity-
mediated dynamics across a range of scales, from plots to patches to landscapes.
This model represents one of the most robust and unique tools for examining spatio-
temporal patterns of historical MSFR forests to date.
## Contents

**Chapter 1**  
Introduction ........................................................................................................................................... 1

**Chapter 2**  
Historical dynamics of mixed-severity fire regime forests: Results from the northern U.S. Rockies and a framework for scaling dendroecological methods along the fire regime gradient .............................................................................. 8  
  Introduction ........................................................................................................................................... 8  
  Methods.................................................................................................................................................. 15  
    Study area & design ............................................................................................................................... 15  
    Photogrammetric mapping of vegetation patches .............................................................................. 17  
    Dendroecological data collection and processing .............................................................................. 18  
    Characterization of patch age structure and identification of age cohorts .................................. 21  
    Drivers of cohort establishment and attribution of age cohorts to fire events ............................ 22  
    Development of the fire history record ............................................................................................... 24  
    Estimation of Fire Severity .............................................................................................................. 25  
    Spatial patterns of fire history information along a fire regime gradient .................................. 28  
    Examining bias due to compositing fire history information within patches .......................... 30  
    Compositing age structure data within patches .......................................................................... 31  
  Results .................................................................................................................................................. 33  
    Photogrammetric and dendroecological records ........................................................................... 33  
    Drivers of tree establishment ........................................................................................................... 34  
    Reconstructed fire regimes ............................................................................................................ 34  
    Spatial patterns of fire history information along a fire regime gradient .................................. 35  
    Fire history bias due to compositing .............................................................................................. 38  
    Characterization of patch-level age structure .......................................................................... 39  
  Discussion ............................................................................................................................................ 40  
    Age structure is closely linked to patch-level fire effects ........................................................... 40  
    Dendroecological methods should be scaled with the fire regime gradient ................................ 43
Paired photogrammetric and dendroecological records allow robust reconstruction of historical fire-mediated stand dynamics, spatial pattern, and spatial scale of fire effects.......................................................... 48

Spatio-temporal intermixing of fire frequency and severity created landscapes dominated by younger single- and multi-cohort patches .......... 49

Possible evidence for nonlinear relationships between fire frequency and severity .......................................................... 51

Conclusions .......................................................................................................................................................... 53

Acknowledgements ............................................................................................................................................. 54

Appendix A2: Calibration of relative dendroecological severity......................... 69

Biases associated with the use of raw basal area or density for estimating fire severity .................................................................................. 69

Calibration of the fire-event severity metric, $\Delta S(n_i)$ ........................................ 72

Chapter 3
A cross-scale method for modeling spatio-temporal patterns of fire severity in historical landscapes using dendroecology and photogrammetry.............................................................. 76

Introduction .......................................................................................................................................................... 76

Reconciling limitations to dendro-calibrated, structure-based fire severity models.................................................................................. 82

Methods............................................................................................................................................................. 85

Study Area........................................................................................................................................................ 85

Study Design ...................................................................................................................................................... 86

Photogrammetric mapping of patch boundaries and vegetation structure .... 87

Field sampling and laboratory methods ........................................................................................................... 89

Reconstruction of patch-level fire history and severity.............................................. 90

Characterization of patch-level age structure .............................................................. 93

Derivation of patch-level topographic variables and climatic setting .......... 93

Construction and validation of structure-based models........................................ 94

Accounting for temporal disparities between dendroecological and structural records ................................................................................. 98

Variable reduction & selection of the final fire severity model......................... 100

Rule-based structural fire severity models.............................................................. 102
Multivariate structural properties and relationships with fire severity history .......................................................... 102

Results .......................................................................................................................................................... 104

Photogrammetric analysis & dendroecological fire severity reconstructions ......................................................... 104
Structure-based models of fire severity ................................................................. 106
Variable reduction & selection of the final fire severity model.............. 108
Multivariate structural characteristics and fire severity history ............ 110
Relationships between top predictor variables and cumulative fire severity .................................................................................................................. 111

Discussion .............................................................................................................................................. 113

Limitations to dendroecologically-calibrated, structure-based models of fire severity ................................................. 115
Considerations for structure-based models of historical fire severity ...... 116
Comparison with other fire severity reconstruction methods.............. 119

Conclusions .............................................................................................................................................. 123

Acknowledgements ................................................................................................................................. 124

Appendix A3: Ancillary material from the photogrammetric analysis of historical aerial photos ................................................................. 139
Appendix B3: Dendroecological supplementary information ................. 144
Determination of age thresholds used in age structure classes ............. 144
Appendix C3: Structure-based model ancillary outputs ......................... 149

Chapter 4
Summary and conclusions ......................................................................................................................... 155

Landscape dynamics and resilience mechanisms in MSFRs .................... 155
Management implications of novel resilience mechanisms and dynamics in MSFRs ......................................................... 156
Dendro-calibrated, structure-based fire severity models: A novel cross-scale reconstruction method ................................................................. 158
Understanding methodological challenges and fire-mediated dynamics in the context of the fire regime gradient ......................... 159
Next steps: Application of structure-based models to evaluate the spatial ecology and biogeography of MSFRs across the northern Rockies ......................................................... 160

References .............................................................................................................................................. 162
TABLES

Table 1.1. Data sources and characteristics of the principal methods used to study historical fire severity. .................................................................................................................................................. 7

Table 2.1. Table outlining the range of scenarios of availability of fire history information along the fire regime gradient. ................................................................................................................. 55

Table 2.2. Summary information of the forest types, dendroecological data, and fire regime properties of each of the four study regions used to construct the fire regime gradient. ............................................................................................................................................................................................ 56

Table 2.3. Results from the cohort richness analysis stratified by cohort and fire severity classes.................................................................................................................................................................................. 58

Table 3.1. Description of the justification and reclassification procedure for each of the four reclassification hypotheses.......................................................................................................................................................................................................................................................... 125

Table 3.2. Summary information for each of the five study watersheds in the NCDE. .......................................................................................................................................................................................................................................................... 127

Table 3.3. Classification rules for the age structure classes. ...................................................................................................................... 128

Table 3.4. Detailed summaries of the structure- and rule-based model parameters and performance statistics. .......................................................................................................................................................................................................................................................... 129

Table 3.5. Descriptions of the structural groups identified by cluster analysis and their relation to severity classes. .......................................................................................................................................................................................................................................................... 132

Table A3.1. Description of all photo-interpreted vegetation attributes and topographic variables .......................................................................................................................................................................................................................................................... 139

Table A3.2. Rule-based key used to classify the severity of the most recent fire at each patch. .......................................................................................................................................................................................................................................................... 143

Table C3.2. Confusion matrix for the severity of the most recent fire Random Forest model predictions. .......................................................................................................................................................................................................................................................... 151

Table C3.3. Confusion matrix for the age structure class Random Forest model predictions. .......................................................................................................................................................................................................................................................... 152

Table C3.4. Confusion matrix for the severity of the most recent fire rule-based model predictions. .......................................................................................................................................................................................................................................................... 153
Table C3.5. Confusion matrix for the cumulative fire severity Random Forest model predictions. 

............................................................ 154
FIGURES

Figure 2.1. Diagram of the theoretical influence of the fire regime gradient on fire history reconstruction uncertainty and sampling methodology ........................................ 59
Figure 2.2. Map of the study areas for chapter 2.............................................................. 60
Figure 2.3. Graphics from the analysis of the drivers of cohort recruitment.......... 61
Figure 2.4. Graphics summarizing the fire regime properties of the NCDE and GYE. ................................................................................................................................. 63
Figure 2.5. Graphics showing the fire regime characteristics and the spatial bias of fire scar information along the fire regime gradient............................................... 64
Figure 2.6. Graphics from Six Mile Creek in the GYE, depicting the challenges to reconstructing fire history in MSFRs with high spatial bias of fire history records. .......................................................................................................................... 65
Figure 2.7. Fire interval distributions for different compositing methods. ........ 66
Figure 2.8. Rarefaction curves of cohort richness and extrapolated cohort richness as a function of patch size. ............................................................ 67
Figure 2.9. Boxplots of patch-level median fire intervals and the percent of patches with right-censored intervals stratified by severity class................................. 68
Figure A2.1. Influence of different thresholding values on classification agreement between $BA(n_i)_{surviving}$ and $\Delta S(n_i)$ severity metrics. ........................................ 75
Figure 3.1. Map of the study areas for chapter 3......................................................... 133
Figure 3.2. Hypothesized and observed relationships between time since high severity fire and model performance. ......................................................... 134
Figure 3.3. Relativized variable importance scores and results from the variable reduction procedure for the final cumulative fire severity models .................... 135
Figure 3.4. Nonmetric multidimensional scaling plots depicting the multivariate structural characteristics of classes produced by the hierarchical cluster analysis, the original fire severity classes, and the reclassified fire severity classes........... 136
Figure 3.5. Partial dependence plots for the top vegetation predictors used in the final random forest model of cumulative fire severity. .................................. 137
Figure 3.6. Partial dependence plots for the top topographic predictors used in the final random forest model of cumulative fire severity. .......................................................... 138

Figure B3.1. Frequency distribution histograms for the number of years to canopy closure and stand age. ........................................................................................................ 146

Figure B3.2. Examples of the fire history and age structure of sites from different severity classes........................................................................................................ 147

Figure B3.3. Examples of the fire history and age structure of sites affected by each of the four reclassification hypotheses. ................................................................. 148

Figure C3.1. Examples of a conditional tree and an evolutionary tree for the cumulative fire severity class................................................................................................. 149
Chapter 1

Introduction

Critical questions exist about the socio-ecological consequences of and potential responses to changing fire regime dynamics, ecosystem function and landscape resiliency resulting from past land management combined with increased climate-forcing of wildfire activity. Studies of historical fire regimes have played a key role in addressing these questions (Swetnam et al. 1999, Keane et al. 2009, Whitlock et al. 2010) because (1) they are able to capture broad temporal variation that isn’t possible using documentary records, including extremes in fire activity, shifting fire regimes, and changes in the relative strength of the biophysical drivers of fire dynamics over time, and (2) they predate Euro-American settlement and associated changes in landscape conditions that confound evaluation of the dynamics of active fire regimes that are likely to characterize future landscapes in many parts of the world. Most historical fire research has focused on reconstructing fire frequency (Kitzberger et al. 2007, Falk et al. 2011, Marlon et al. 2012) or size (Heinselman 1973, Johnson et al. 1998, Hessl et al. 2007). These types of studies have provided important empirical datasets used to describe the landscape dynamics of active fire regimes (Heinselman 1973, Swetnam and Baisan 1996, Fulé et al. 2002, Sherriff et al. 2014) and to derive statistical relationships (Williams et al. 2010), bound simulation model exercises (Lertzman et al. 1998), or validate assumptions of process-based models (Pechony and Shindell 2010) that in turn are used to model the ecological outcomes of future fire activity under alternative climate and land-use scenarios (Williams and Abatzoglou 2016). However, relatively few historical ecology studies have reconstructed broad-scale patterns of historical fire severity despite the important role it plays in moderating landscape dynamics, ecosystem function, post-fire recovery, feedbacks and tipping points (Peterson 2002, Johnstone et al. 2010, Turner 2010, Coop et al. 2016).
In part, interest in fire severity has been spurred by the recognition that wildfire in many ecosystems within western North America (Taylor and Skinner 1998, Schoennagel et al. 2004, Perry et al. 2011, Sherriff et al. 2014), and also more broadly (Holz and Veblen 2009, Shorohova et al. 2009), exhibit a broad range of fire severities, either over multiple fire events in time or spatially within individual fire events. These mixed-severity fire regime (MSFR) ecosystems exhibit unique spatio-temporal dynamics, resilience mechanisms, and biophysical controls than more well-studied ecosystems with low- or high-severity fire regimes that are poorly understood. For instance, in low-severity fire regime systems, high frequency (5-20 years) low severity fires limited tree density and understory fuels, conditions which created strong stabilizing feedbacks that perpetuated low severity fires (Moore et al. 1999, Allen et al. 2002). Fire exclusion over the last century is thought to have disrupted this feedback by facilitating fuel accumulation, increasing the risk of widespread high-severity fire, and removing the forest structural conditions that maintained this feedback (Covington 2000, Savage and Mast 2005, Savage et al. 2013). In high-severity fire regimes, forests are resilient to large, severe fires, due to favorable climatic conditions that promote tree regeneration coupled with long fire return intervals (> 100 years) (Kipfmuller and Kupfer 2005, Sibold et al. 2006). In contrast, MSFRs exhibit variable median fire frequencies (15-60 years) and a wide range of severities, ranging from regimes dominated by non lethal fires with only infrequent and small-scale high severity fire (Heyerdahl et al. 2001, Margolis and Balmat 2009, Heyerdahl et al. 2014) to those where high-severity fire is more frequent, intermixes with non lethal fire across a broader portion of the landscape or occurs in much larger patches (Bekker and Taylor 2001, Amoroso et al. 2011, Schoennagel et al. 2011, Heyerdahl et al. 2014, Sherriff et al. 2014).

As a result of this variability, neither high frequency, low severity fire nor infrequent, high-severity fire sufficiently describe the landscape dynamics and resilience mechanisms that operate in MSFRs. Most existing studies in MSFRs encompass small areas (Taylor and Skinner 1998, Bekker and Taylor 2001,
Margolis and Balmat 2009, Scholl and Taylor 2010, Heyerdahl et al. 2012, Heyerdahl et al. 2014, Tepley and Veblen 2015) and studies are lacking for MSFRs in many regions, limiting the data needed to inform hypotheses about the ecological dynamics and biophysical drivers of variability in MSFRs. The uncertainty created by this lack of data has generated substantial debate in the scientific community (Baker et al. 2007, Franklin and Johnson 2012, DellaSala et al. 2013, Fulé et al. 2013, Hagmann et al. 2013, Hagmann et al. 2014, Odion et al. 2014, Williams and Baker 2014, Stevens et al. 2016) and confusion regarding the influence of historical land management on the ecological integrity of MSFRs, their likely response to climate change, and the outcomes of alternative land management scenarios.

Many of these questions stem from uncertainties about the ecological role of high-severity fire in MSFRs and how resilience mechanisms change in response to the scale or frequency of high-severity fire. High-severity fire has received this special interest because it consumes most vegetation at a site, thereby weakening forest structural controls on subsequent fire behavior that can create negative feedbacks on future high-severity fire (Agee 1993). By altering these feedbacks, high-severity fire can create prolonged changes to landscape dynamics and ecological function, influence landscape resilience to future disturbance and trigger alternative landscape states (Larson et al. 2013, Johnstone et al. 2016). For instance, limited seed sources in the interior of large high-severity patches can lead to prolonged early seral conditions (Donato et al. 2009, Haire and McGarigal 2010, Kemp et al. 2015, Harvey et al. 2016) that may reduce recovery rates and resilience to future disturbances. If high-severity fire occurs over large areas or recurs over short time scales, it may even trigger switches from forests dominated by multi-aged, fire-resistant species (e.g. ponderosa pine, Jeffrey pine, or western larch) to non forest vegetation (Barton 2002, Savage and Mast 2005, Collins and Roller 2013, Falk 2013, Savage et al. 2013) or alternative forest types (Larson et al. 2013) that promote future high-severity fire. These ideas have motivated interest in the spatio-
temporal dynamics of high-severity fire in MSFRs and the ecological mechanisms that confer resilience in the face of high severity fire.

Despite great interest in historical reconstructions of fire severity to address these questions, significant methodological challenges exist that hinder reconstruction of fire severity over the spatio-temporal scales necessary to characterize the landscape dynamics and consequences of fire in MSFRs. Fire severity reconstructions (Table 1.1) have primarily been based on point- or stand-level dendroecological data (Margolis and Balmat 2009, Heyerdahl et al. 2012, Sherriff et al. 2014, Tepley and Veblen 2015, Yocom-Kent et al. 2015), or in a more limited set of cases, on inferences from forest structural characteristics derived from aerial photos (Taylor and Skinner 1998, Minnich et al. 2000, Bekker and Taylor 2001, Hessburg et al. 2007) or historical survey data (Collins et al. 2011, Vankat 2011, Williams and Baker 2012, Hagmann et al. 2013). Although dendroecological reconstructions of fire severity are subject to some biases and uncertainty (Johnson et al. 1994, Baker and Ehle 2001, Hessburg et al. 2007), they provide relatively robust temporal records of fire history and tree establishment that can be used to make qualitative or quantitative estimates of fire severity for specific locations (Table 1.1). However, they generally cover only small areas relative to the landscape (Baker and Ehle 2001, Moritz et al. 2009), are a labor-intensive data source for geostatistical modeling techniques (Farris et al. 2010, Scholl and Taylor 2010, O’Connor et al. 2014, Yocom-Kent et al. 2015), are often designed to produce only aspatial inferences about fire severity (e.g., proportion of a landscape or relative area) that don’t address scale or pattern (Hessburg et al. 2007, Moritz et al. 2009), and can suffer from biased sampling design (e.g., selective placement in fire scar-rich sites or exclusion of shrubfields and grasslands caused by high severity fire) (Johnson and Gutsell 1994, Baker and Ehle 2001, Van Horne and Fule 2006, Baker 2009, Farris et al. 2010). In contrast, aerial photographs and historical surveys can provide detailed records of forest conditions and inferred fire severity across a range of scales, from individual plots or patches to large landscapes, for
specific points in time (Hessburg et al. 2007, Williams and Baker 2012). Compared to dendroecological data, these structure-based reconstructions may therefore more effectively capture the spatial variation, scale and pattern of fire severity within a landscape (Table 1.1), although they also present a greater degree of ambiguity about the dynamics and fire history leading to specific forest structures that has limited their widespread use as a reconstruction method.

Notably few studies have incorporated multi-proxy historical data sources to reconstruct fire severity patterns, despite the clear potential benefits to such an approach (but see Taylor & Skinner 1998, Bekker & Taylor 2001, Williams & Baker 2012). Incorporating multiple data sources with unique spatio-temporal properties could facilitate the validation of interpretations made by individual methods (a task which is notoriously difficult for historical reconstructions), help to quantify biases or limitations inherent to different methodologies, and improve spatio-temporal fire severity inferences by leveraging the unique spatial or temporal properties of different methods. Here, we use a unique multi-proxy dataset of dendroecological data and geospatial maps of historical forest structural attributes and patch boundaries derived from photogrammetric interpretation of historical aerial photographs to address three objectives:

1. Develop a framework for scaling historical fire severity reconstruction methods along the fire regime gradient, with emphasis on the challenges to fire severity reconstructions in MSFRs.

2. Provide new insights about the spatio-temporal dynamics of MSFRs using dendroecological records from two distinct study regions in the northern U.S. Rockies, the Northern Continental Divide Ecosystem (NCDE) and the Greater Yellowstone Ecosystem (GYE).

3. Develop a dendroecologically-calibrated, structure-based fire severity model that incorporates both dendroecological records and photogrammetrically-
derived forest structural attributes to predict patch-level fire severity-mediated dynamics.

To accommodate these objectives, we present data from dry and moist mixed-conifer forests in seven watersheds located in the northern U.S. Rockies, two in the GYE and five in the NCDE. These forest types are representative of much of the lower-middle elevation forests within the northern U.S. Rockies. Previous research in both regions has documented MSFRs for small portions of the landscape (Barrett et al. 1991, Arno et al. 1995, Littell 2002), but there is insufficient understanding of the landscape dynamics of both systems. In chapter two we develop methods for quantifying the fire regime gradient using dendroecological data from both regions as well as previously published ancillary data, we examine how methodological challenges and fire severity reconstruction methods vary along this fire regime gradient, and we evaluate the spatio-temporal dynamics of the GYE and NCDE. In chapter three, we develop a dendroecologically-calibrated, structure-based model that predicts patch-level fire severity from aerial photo-interpreted forest structural attributes and we describe the structural characteristics that best predict distinct fire severity histories.
Table 1.1. Data sources, metrics used to infer fire severity, sample unit characteristics and spatio-temporal constraints for the principal methods used to study historical fire severity.

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Data source</th>
<th>Metrics of Inference</th>
<th>Sample unit</th>
<th>Effort/sample unit</th>
<th>Sample density</th>
<th>Spatial extent</th>
<th>Temporal resolution</th>
<th>Temporal extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dendroecology</td>
<td>tree establishment dates, fire scar dates, growth response, tree diameters/counts, spacial arrangement and species</td>
<td>age/cohort structure, species composition, forest structure, stem maps</td>
<td>plots (10^2 - 10^2 ha)</td>
<td>High</td>
<td>Low-high</td>
<td>Low-moderate</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Historical aerial photography</td>
<td>panchromatic imagery &amp; stereo photogrammetric visualization</td>
<td>horizontal and vertical (where stereo imagery is available) vegetation structure, species composition</td>
<td>pixels (1 - 10 m²), vector objects (&lt; 10^2 - 10^2 ha)</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Historical field surveys</td>
<td>tree diameters/counts, species, and spatial arrangement</td>
<td>species composition, forest structure, stem maps</td>
<td>plots (10^2 - 10^2 ha)</td>
<td>Moderate</td>
<td>Low-high</td>
<td>Low-high</td>
<td>Low</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

1. describes the principal data types employed by each method.
2. attributes that are used make inferences about fire severity.
3. unit (and scale) at which inferences are made for each method.
4. relative effort required per sample unit
5. density of sample units within a study area
6. areal extent encompassed by each method.
7. density of the time series information achievable by each data source.
8. timespan covered by each data source.

# Note that where the effort per sample unit is high, an inverse relationship exists between sample density and spatial extent.
Chapter 2

Introduction

Dendroecological studies of historical fire regimes have generally focused on reconstructing fire occurrence and the temporal (e.g. interannual to multi decadal climate variation) or spatial (e.g. topography, moisture gradients) drivers of variability in fire occurrence (Kitzberger et al. 2007, McKenzie and Kennedy 2012). Distinct methods for characterizing fire occurrence have been developed in forests that lie at opposite ends of the fire regime continuum. In forests where fires were historically frequent and primarily of low severity, most studies sample trees with multiple fire scars to estimate fire frequency (Swetnam and Baisan 1996, Falk et al. 2011), whereas studies in forests dominated by high severity fires with long return intervals typically map historical fire perimeters to estimate burn rates or fire rotations (Johnson 1979, Romme 1982, Sibold et al. 2006). In these cases, fire-caused tree mortality, here referred to as fire severity, is often not explicitly quantified because it is treated as an assumed property of the study system.

An important shift in dendroecology to explicitly reconstruct fire severity has been motivated by several factors: (1) many ecosystems across western North America (Schoennagel et al. 2004, Perry et al. 2011) and elsewhere (Heinselman 1973, Bergeron and Brisson 1990, Gonzalez et al. 2005, Holz and Veblen 2009, Shorohova et al. 2009) are now known to have experienced a complex mixed-severity fire regime (MSFR) characterized by high spatiotemporal variability of fire frequency and severity, (2) in MSFR systems, fire severity and frequency may
exhibit complex, nonlinear relationships (Pennanen 2002) so each must be independently quantified if spatiotemporal dynamics are to be understood, and (3) the spatial pattern, grain, and variability of fire severity strongly influences post-fire successional trajectories, vegetation feedbacks on subsequent disturbance, and a variety of ecosystem functions (Agee 1998, Turner 2010, Halofsky et al. 2011).

A number of methodological challenges make historical fire severity reconstructions in MSFRs uniquely difficult. The first stems from two countervailing effects; spatially-distributed dendroecological evidence is needed to characterize the high spatial heterogeneity of the patch mosaic formed in MSFRs, yet the prevalence of moderate-high severity fires in MSFRs can result in high loss rates and patchiness or spatial bias of fire history evidence. A number of analytical approaches exist for estimating historical fire severity from dendroecological records (Taylor and Skinner 1998, Ehle and Baker 2003, Amoroso et al. 2011, Schoennagel et al. 2011, Heyerdahl et al. 2012, Marcoux et al. 2015, Tepley and Veblen 2015, Yocom-Kent et al. 2015), all of which fundamentally rely on association between forest age structure and fire events. Age structure is vital to reconstructing fire severity because it may help date fire events where fire scars or other fire history evidence is lacking (Kipfmueller and Baker 1998, Heyerdahl et al. 2014) and because tree establishment and survival serve as the main proxies for fire severity (Sherriff and Veblen 2006, Schoennagel et al. 2011, Heyerdahl et al. 2012, O'Connor et al. 2014, Tepley and Veblen 2015). Although age structure can be determined wherever trees produce annual rings, the widespread influence of high- or repeated
moderate-severity fires in many MSFRs consumes fire history evidence (e.g. fire scars, growth anomalies, wood anatomical features), leaving sparse or spatially-biased (e.g. to ridgelines) records of past fire boundaries and effects. It is reasonable to expect that the distance between age structure plots and fire scars \((d_{AF})\) will therefore vary as a function of the high severity burn rate for a given area. High \(d_{AF}\) values could create significant uncertainties in estimates of fire frequency and the interpretation of age structure (Lorimer 1985) because without precise fire history information, attribution of cohorts to specific fire events or even to fire, exclusively, may be subject to error since a number of exogenous (pluvials, windthrow, insect outbreaks) and endogenous (succession, senescence) processes can trigger recruitment pulses (Lorimer 1980, Aplet et al. 1988, Ehle and Baker 2003, League and Veblen 2006, Dugan and Baker 2015).

If the \((d_{AF})\) increases reliably along the fire regime gradient, a general methodological framework can be outlined for dendroecological sampling design that accounts for varying scarcity and spatial bias in fire history records across the fire regime spectrum (Fig. 2.1). Under this hypothesis, the uncertainty in fire severity and frequency estimates and the reliance on age structure increases along the fire regime gradient as high severity fire becomes more prevalent and the loss rate of fire history information increases. This uncertainty peaks for MSFR forests, and declines for high-severity fire regimes even though \(d_{AF}\) continues to increase because reduced fire frequencies and large, high severity fires create a coarser, simpler patch mosaic that can be more easily reconstructed with age structure and
sparse fire history information (Romme 1982, Veblen et al. 1994, Johnson et al. 1998, Kulakowski et al. 2003, Sibold et al. 2006). Although this hypothesis is generally consistent with existing literature and dendroecological sampling practice (Agee 1998, Lertzman et al. 1998, Agee 2005, Lentile et al. 2005, Hessburg et al. 2007, Parsons et al. 2007), it has not been empirically tested or formalized in a framework that addresses the unique methodological challenges to reconstructing fire regime parameters along the fire regime gradient and for MSFRs, in particular.

A second challenge to reconstructing historical fire severity is the difficulty of scaling point, or small-area (e.g. < 1 ha), dendroecological records to larger spatial scales or spatially explicit patterns. Reconstructing the patch sizes (i.e. grain) and pattern of the mosaic of different burn severities in MSFRs is critical because these properties are thought to constitute primary feedback mechanisms that regulate future fire patterns and severity, landscape dynamics and successional trajectories, and ecosystem resilience (Pennanen 2002, Hessburg et al. 2005, Halofsky et al. 2011, Perry et al. 2011, Miquelajauregui et al. 2016). The geographic distribution of fire regime classes defined by historical fire severity has been modeled from biophysical variables without explicitly reconstructing spatial patterns caused by MSFRs (Sherriff et al. 2014). Geostatistical modeling methods have been used to interpolate fire perimeters from point dendroecological records in low-severity fire regime systems where fire history information is abundant (Hessl et al. 2007, Farris et al. 2010) and these techniques can also be used to map fire severity in MSFRs (O’Connor et al. 2014, Yocom-Kent et al. 2015). However, the scale at which
geostatistical methods can be applied is generally limited by the high costs associated with collection of the high-density, spatially-distributed dendroecological records needed to interpolate fire severity in heterogeneous landscapes. As a result, most dendroecological studies either describe plot-scale results that lack spatial information (Sherriff and Veblen 2006, Iniguez et al. 2009, Margolis and Balmat 2009, Amoroso et al. 2011, Heyerdahl et al. 2012) or reconstruct spatial pattern over small areas ≤ 10³ ha (Morrison and Swanson 1990, O’Connor et al. 2014, Tepley and Veblen 2015, Yocom-Kent et al. 2015). Only a few examples exist of spatially-explicit dendroecological fire severity reconstructions that encompass study areas of sufficient size (e.g. 10⁴ – 10⁶ ha) to adequately describe historical fire severity-driven dynamics over large landscapes (Sherriff and Veblen 2007, Sherriff et al. 2014).

An alternative approach to reconstructing historical fire severity that we explore here is to combine dendroecological records with photogrammetric delineation of patch boundaries based on photo-interpretation of historical forest structures. Patch boundaries delineated from historical aerial photographs provide direct observations of the spatial patterns of forest structure created by previous disturbances and biophysical conditions (Avery and Berlin 1992, Hessburg et al. 1999b, Morgan et al. 2010). If delineated patch boundaries optimize the ratio of within-patch to between-patch variation of age structure and disturbance histories, as they are intended, pairing of dendroecological fire severity reconstructions with photogrammetrically interpreted patch boundaries would provide robust empirical
information about the temporal history of fire severity-mediated dynamics and the spatial scale (i.e. patch size) and pattern (i.e. patch shape, intermixing of patch types) of historical fire severity. It would also confer numerous methodological advantages compared to random, stratified random or gridded plot designs that are commonly used in dendroecology (Fulé et al. 2002, Schoennagel et al. 2011, Heyerdahl et al. 2012). Stratifying dendroecological sampling grids by patches could improve sample design, sampling efficiency, and the areal coverage of dendroecological studies because it provides direct knowledge of the spatial grain of the patch mosaic that allows precise control of within- and between-patch sample intensity and avoidance of patch boundaries. Where patch boundaries are not used to stratify sampling, the appropriate plot spacing and arrangement is determined by relatively arbitrary spacing criteria rather than the specific scale and pattern of the landscape mosaic being studied.

Conceptually, if patches represent ecologically meaningful units, then both fire history and age structure data collected from small plots within a patch could be composited to produce more robust estimates of the age structure, fire history, and fire severity of a patch. However, if photointerpreted boundaries do not accurately delineate stands with unique disturbance histories, compositing dendroecological data could bias interpretations towards shorter fire intervals and lower fire severity. Fire intervals are known to decrease as a function of compositing area because small fires are cumulatively detected and included in fire frequency estimates for the entire area over which data is composited (Arno and Petersen
Although compositing of age structure data has not been investigated previously, a bias analogous to that which impacts fire frequency estimates may exist. For instance, if poorly delineated patch boundaries caused compositing of cohorts from plots with distinct fire history or effects, the aggregate patch age structure could become increasingly multi-modal, thereby biasing interpretation of age structure towards lower severity. To address this problem, systematic study of the effect of compositing both fire history and age structure within photogrammetrically interpreted patch boundaries is needed.

To address these questions, here we present new dendroecological data collected from dry and moist mixed-conifer forests across seven watersheds in two distinct study regions within the northern U.S. Rockies where patch boundaries have been mapped using intensive photogrammetric techniques. To examine how dendroecological record quality and methodological challenges vary along a fire regime gradient, we compare these data with previously published dendroecological records from other regions. We address the following study objectives and research questions:

1. Reconstruct historical fire frequency and severity patterns for mixed-severity fire regime forests in two distinct regions of the northern U.S. Rockies.

2. Highlight the methodological challenges unique to dendroecological study of MSFRs, evaluate our proposed framework of the spatial quality of dendroecological records along a fire regime gradient, and provide a practical set
of recommendations for how dendroecological sample design should scale with the fire regime gradient.

3. Evaluate drivers of cohort initiation and develop methods that improve attribution of age cohorts to fire events where fire scar evidence is sparse.

4. Evaluate whether patch boundaries are appropriate units for binning dendroecological data.

4.1. Does compositing fire history and age structure data at the patch scale bias fire frequency and severity interpretations?

4.2. Do patch boundaries represent ecologically meaningful units that can be used to quantify the scale and pattern of reconstructed fire severity?

**Methods**

*Study area & design*

Our study was conducted in two ecoregions within the northern U.S. Rocky Mountains in Montana, the northern range of the Greater Yellowstone Ecosystem (GYE) and the Northern Continental Divide Ecosystem (NCDE) (Fig. 2.2). These regions are characterized by contrasting climatic settings, with cooler, drier conditions in the GYE (mean annual Temperature = 1°C, mean annual precipitation = 71 cm) compared to the NCDE (mean annual Temperature = 6°C, mean annual precipitation = 109 cm). Vegetation of the northern GYE consists of shrubfields and montane grasslands interspersed with Douglas-fir (Pseudotsuga menziesii), Rocky Mountain Juniper (Juniperus scopulorum), and scattered limber pine (Pinus flexilis) at low to moderate elevations and spruce-fir (Picea engelmannii - Abies
lasiocarpa) or lodgepole pine (Pinus contorta) forests at higher elevation. In the NCDE, mixed-conifer forests of ponderosa pine (Pinus ponderosa), western larch (Larix occidentalis), Douglas-fir, and lodgepole pine, and minor components of grand fir (Abies grandis) or trembling aspen (Populus tremuloides) occur at low-moderate elevations. Mesic sites often have components of Engelmann spruce, subalpine fir, western white pine (Pinus monticola), western red cedar (Thuja plicata), or paper birch (Betula papyrifera). Low- and mixed-severity fire regimes have been documented in the NCDE (Freedman and Habeck 1985, Barrett et al. 1991, Arno et al. 1997, Arno et al. 2000, Naficy et al. 2015), but the prevalence and geography of each regime is poorly understood. Fire regime studies in the GYE have mostly documented forest expansion into ecotonal grasslands (Houston 1973, Arno and Gruell 1983, Barrett 1994, Heyerdahl et al. 2006) and there is evidence of mixed-severity fire effects from a limited number of studies (Littell 2002, Korb 2005, Naficy et al. 2015). However, relatively little is known about the historical fire ecology and dynamics of montane Douglas-fir forests.

We collected dendroecological data and mapped patch boundaries and vegetation structures using historical aerial photographs for two watersheds in the northern GYE and five in the NCDE (Fig. 2.2). To maximize the range of historical fire effects captured in our sample, we selected study watersheds (5,500-18,000 ha in size) representing a broad range of biophysical conditions and forest types. Forest structures and patch boundaries were mapped across the full extent of each watershed, but dendroecological sampling was restricted to dry-moist mixed-conifer
and lodgepole pine cover types in the NCDE and Douglas-fir types in the GYE. To
ensure collection of dendroecological records across a range of structural types and
disturbance histories, maximize field sampling efficiency, and ensure a spatially-
distributed network of sample patches within each watershed, clusters of 2-3
neighboring patches were randomly selected from groups stratified by topographic
and vegetation characteristics. Groups used for stratification included two broad
aspect classes (defined along the SW-NE axis), two elevation bands (defined as the
upper and lower 50% of DEM pixel values), and up to seven forest structural types
(Hessburg et al. 1999a) derived from the photointerpreted structural variables.
Watersheds with more complex forest structures and apparent disturbance histories
were sampled more intensively.

Photogrammetric mapping of vegetation patches

We used photogrammetric techniques to delineate vegetation patches with
internally consistent characteristics (i.e. tree composition and structure) across each
study watershed. For each watershed we obtained high resolution (1: 15840 -
1:26,000 scale), panchromatic stereo aerial photographs from local Forest Service
offices, the National Archives or the USDA Aerial Photography Field Office. To
ensure that forest structures and patch boundaries closely reflected closely historic
fire regimes, we selected the earliest available photo series for each study
watershed (GYE: 1956-1962, NCDE 1934-1955). These dates are generally
coincident with the onset of effective fire exclusion in our study regions around the
1930s (Arno 1980, McKay 1994, Barrett et al. 1997), although extensive grazing was
underway in the GYE by the late 1800s that may have begun to reduce fire frequency at the time of the aerial photographs (Hansen et al. 1995, Sankey et al. 2006).

The stereo aerial photos were used to interpret a number of forest structural and compositional variables (see detailed variable descriptions in Hessburg et al. 1999) that were assessed separately for understory and overstory layers, including: canopy cover, tree size class, species composition and life form, snag abundance, and a number of textural variables describing fine-scale tree cover patchiness (e.g. clumpiness, crown differentiation). Patches were delineated using an iterative process that refined patch boundaries as structural interpretations were made, down to a minimum patch size of 4 ha. Delineated patches for all mapped watersheds in the NCDE were digitized, orthorectified and converted into a vector layer. At the time of publication, patch boundaries for watersheds in the GYE were not yet orthorectified, so no spatial information associated with these patches is presented here.

Dendroecological data collection and processing

To characterize the age structure of each sample patch, a variable-length transect was initiated at a random point and azimuth within the interior of each patch (> 300m from patch boundaries). Circular, fixed area (0.04 ha) plots were placed along each transect at 30-100 m spacing (average spacing = 60 m), depending on patch size, dimensions and age structure complexity. To balance the competing goals of intensive age structure sampling within a patch and maximal number of
patches sampled within each watershed, we employed two plot sampling procedures. In the first procedure, our goal was to intensively sample the within-patch heterogeneity of tree ages from a large number of plots in each patch (range = 5-22, median = 10 plots). For plots within intensively sampled (IS) patches, we recorded the species and measured the diameter of all live and dead trees ≥ 4 cm diameter at breast height (DBH) until a minimum of > 300 live trees had been measured. Increment cores were sampled from a minimum of 15-20 trees of each shade intolerant species and 5-10 shade tolerant species, in proportion to their diameter distribution within the patch. To convert the age distribution in the subsample of trees of each species into the absolute density of the transect, the size distribution was multiplied by the relative density of that species documented along the transect.

In the second procedure, dendroecological data were sampled using a rapid assessment (RA) method that employed a lower intensity sample in a greater number of patches. For RA patches, the number of plots per transect varied from 1-5 (median = 3) depending on the size and apparent complexity of the patch fire severity history. In our study regions, where fire frequency and effects are highly variable, stands can vary from dense, even-aged lodgepole pine stands with > 5,000 trees/ha to open dry mixed-conifer stands with < 200 trees/ha, necessitating this adaptive sampling strategy. Fire history complexity was visually assessed in the field by inspection of fire frequency from initial fire scar samples and field counts of tree ages. In patches with higher fire frequency and apparent number of cohorts, a
higher number of plots were sampled. A 0.02-ha subplot was centered within each 0.04-ha RA plot and increment cores were collected from all trees within the nested subplot while only trees ≥ 35 cm DBH were sampled in the full plot.

To minimize the loss of growth rings with increasing coring height (Veblen 1992), increment cores were sampled at 15-30 cm height above the root collar in all but some of the largest trees with interior rot. In most cases where rot occurred, replacement cores were sampled from similarly-sized trees of the same species nearby the plot. Increment cores were resampled until they were judged to be less than 10 years from the pith date. We estimate the number of years missed at our average coring height by sampling two cross sections from 10 seedlings of each species from each watershed, one at the root collar and one at 25 cm, and calculating the average age difference. All sampled saplings were grown in open (closed) environments for shade intolerant (tolerant) species to represent the conditions under which they naturally regenerate.

An extensive survey of fire scars was performed in each sample patch, adjacent patches, and throughout many unsampled patches. Most fire scars found on dead trees were sampled, unless wood deterioration was severe. A few live trees (< 10% of total samples) were sampled in sites where deadwood was unavailable. In patches where fire scars were abundant, 5-10 scars with the best records were sampled depending on the apparent complexity of site fire history. To increase our sample density and the spatial representation of our fire scar network, we sampled
additional fire scars opportunistically as they were encountered and by systematically searching stumps in previous timber harvest units.

All increment cores and fire scars were surfaced and sanded with successively finer sandpaper (to a maximum 600 grit) until the wood cellular structure was clearly visible. Master chronologies were created for each study region using 45-60 cores from the most abundant species in each study region, including ponderosa pine, Douglas-fir, and western larch for the NCDE and only Douglas-fir for the GYE. Cores were measured to 0.001 mm accuracy using a Velmex manual sliding stage micrometer. All cores used in each master chronology were statistically crossdated using COFECHA (Holmes 1983). Using each master chronology, all remaining cores and fire scars were visually crossdated or, when visual crossdating was not effective, statistically crossdated using COFECHA. For cores that did not contain the pith, the number of missing rings was estimated using methods in Duncan (1989). Any cores missing > 30 years were discarded.

Characterization of patch age structure and identification of age cohorts

We identified cohorts in order to test whether fire, pluvials, or other factors significantly influence patch age structure and to aid with reconstruction of fire history where fire scars or other evidence was lacking. Age structure was summarized by binning tree establishment dates from all plots within a patch into 10-year intervals. To best represent forest conditions prior to fire exclusion and for consistency between aerial photo interpreted forest structures and dendroecological data, we excluded tree establishment after 1955, the date of the most recent aerial
photo. Similar to Heyerdahl et al. (2012), we identified cohorts independent of fire history as a minimum of 50 trees/ha with establishment in a 30 year window preceded by at least one decade with no establishment. The 50 trees/ha threshold was empirically determined to filter out most background variation while preserving the more distinct recruitment peaks. To minimize inclusion of establishment pulses that were not likely caused by fire, ages from shade tolerant species were excluded from this analysis. For each cohort, we identified its initiation and peak establishment year separately because the timing of these responses may result from different processes (Dugan and Baker 2015). We defined the cohort initiation and peak as the modal year of establishment within the first decadal bin and the bin with maximum establishment, respectively.

Drivers of cohort establishment and attribution of age cohorts to fire events

To attribute cohorts to specific fire events, we implemented a three-part strategy: 1) we tested for statistical dependence between tree establishment, fire and climate, 2) we quantify the post-fire time lags over which fire-caused establishment is consistently observed, and 3) we used this window as an upper limit for attribution of age cohorts to fires. To address the confounding influence that climate-triggered cohort initiation could have on interpretation of fire history from age structure (Dugan and Baker 2015), we used bivariate event analysis (BEA) from the K1D software package (Gavin et al. 2006). We tested for statistical dependence between two phases of tree establishment, 1) cohort initiation and 2) peak cohort establishment, and two potential drivers, 1) fire and 2) pluvials. BEA is
a one-dimensional form of the bivariate Ripley’s K-function, a spatial point pattern analyses modified to evaluate the temporal scale of dependence between two types of events over a range of lags (t). For this analysis, we filtered out localized recruitment and fire events by excluding cohorts with less than 150 trees/ha and fires recorded in less than three patches. Pluvials were reconstructed using a gridded tree-ring reconstruction of summer Palmer Drought Severity Index (PDSI) (Grid point_{NCDE}= 68, Grid point_{GYE}= 100, Cook et al. 2004). We used PDSI because it is correlated with summer drought and peak soil moisture water deficit (Dai et al. 2004), conditions that are critical determinants of tree recruitment success. Similar to Dugan & Baker (2015), we defined pluvials as periods with 5-year mean PDSI > 1 that also had at least 1 year with PDSI ≥ 2. We used a bidirectional model to test for dependence between pluvials and establishment because wet years can favor recruitment by either creating favorable conditions preceding cohort establishment that increase germination success or augmenting seedling survival through wet years following germination. Because fires stimulate establishment post-fire through creation of canopy opening and soil disturbance, we used a forward selection model to test for synchrony of tree establishment events with fires.

We graphed the BEA results using the L-function, \( \hat{L}_{AB}(t) \), which is a conversion of the K-function output that stabilizes its mean and variance. We calculated 95% confidence envelopes around \( \hat{L}_{AB}(t) \) using 1,000 randomizations based on the circular shift simulation method available in K1D, which wraps the temporal series from beginning to end and shifts it a random number of years for
each simulation. Compared to random simulation methods, the circular shift method preserves the inherent frequency properties of each event series. Due to the broad and highly variable recruitment windows that defined establishment in our study systems, we found this method improved interpretability of the results over random methods. Time lags for which $\hat{L}_{AB}(t)$ exceeds the upper (lower) confidence envelope indicate statistically significant synchrony (asynchrony) between events, whereas time lags where $\hat{L}_{AB}(t)$ does not exceed confidence envelopes indicate temporal independence.

Finally, we used the time lag identified by BEA over which cohort initiation was statistically dependent on fire ($t_{\text{crit}}$) to define the upper limit of the post-fire window in which a cohort could be attributed to fire with relatively high confidence. Cohorts that initiated in less than or equal to $t_{\text{crit}}$ years from a fire were considered fire-initiated while those initiating after $t_{\text{crit}}$ years were attributed to non fire agents (e.g. insect outbreaks, stand dynamics). Fire-initiated cohorts identified by this data-driven process were used as an important supplementary piece of evidence for the fire history reconstructions.

*Development of the fire history record*

To address the challenges of reconstructing fire history where fire scar evidence is patchy, we incorporated multiple lines of evidence, in addition to fire scars, into our fire history reconstructions including: 1) growth anomalies, 2) wood anatomy (e.g. traumatic resin ducts, latewood discoloration), 3) buried scars found
on increment cores and stumps, 4) fire-initiated establishment cohorts, and 5) approximate fire perimeters reconstructed by calculating the convex hull from all fire scar and age structure plots recording each fire date. These auxiliary data were only used to date fires if they corroborated fire scar dates from within a patch or adjacent patches. If no fire scars or evidence of fires other than an age cohort was documented in a patch, a number of criteria were used to attribute specific fire dates to it (Table 2.1). If dated fire scars from adjacent patches preceded the cohort initiation date by $\leq t_{\text{crit}}$ years, the fire date that most closely preceded the cohort initiation year was assigned to the patch. If no fire scars were sampled within adjacent patches, but the patch bordered or was within a reconstructed fire perimeter and had a cohort pulse whose initiation date was $\leq t_{\text{crit}}$ years from the fire date, the fire date that most closely preceded the cohort initiation year was assigned to the patch. If a patch with an unassigned cohort bordered a fire perimeter and had a weak recruitment response, defined as a visually evident peak that did not meet the 50 trees/ha minimum density threshold required to be considered a cohort by our algorithm, or no recruitment at all, we used indirect information, including species life history traits and topographic features to judge whether or not to assign the fire perimeter date to the patch (Table 2.1).

Estimation of Fire Severity

To reconstruct fire severity for each patch, we used a tree density decay function, $d_i(n_t)$, that uses age structure and fire dates to quantify the rate of loss in stand density ($d$) with increasing numbers of fires ($n$) for the $i^{\text{th}}$ patch (Tepley and
Patches with a higher $d_i(n_i)$ experience greater cumulative tree mortality for a given number of fires. To quantify the relative variation in fire severity among the stands of each study region, we calculated the average density of trees that predate the $n^{th}$ most recent fire per region, $\bar{d}(n)$. We then assessed the relative cumulative severity over a given number of fires by comparing the density of trees that predate the $n^{th}$ most recent fire per stand to the average density of trees that predate the same number of fires per stand across the study region:

$$S(n_i) = \frac{d_i(n_i)}{\bar{d}(n)}$$  \hspace{1cm} \text{[Eq. 1]}

The proportional change to $S(n_i)$ with each fire added when extending the fire record back in time was then used as a proxy for the relative severity of each event:

$$\Delta S(n_i) = -1 \times \frac{S(n_i) - S(n_{i-1})}{S(n_{i-1})}$$  \hspace{1cm} \text{[Eq. 2]}

$\Delta S(n_i)$ values mostly range between -1 and 1, with positive (negative) values representing higher (lower) fire severity relative to other fires pooled across all sites in each region. To allow severity calculations for the largest number of fires possible in this study we used a minimum density cutoff of 15 trees/ha. Below this threshold we truncated calculation of $\Delta S(n_i)$ because we viewed age structure data as insufficient for characterizing fire effects.

We used the event-level severity metric, $\Delta S(n_i)$, to define three severity classes (low, moderate, high) for each fire event and we then assigned a cumulative fire severity class (low, non stand-replacing, mixed, or high) to each patch based on the mix of fire severities that influenced it. We set event-level class breakpoints by finding the $\Delta S(n_i)$ values that produced the best class agreement between $\Delta S(n_i)$
and severity estimates calculated directly from the percentage of stand basal area (BA) surviving a fire, \( BA(n_i)_{surviving} \) (Appendix A). Although event-level tree mortality, or survival, metrics based on absolute BA and density can bias fire severity interpretations for sites with multiple fires and are therefore not useful as a general method for estimating event-level fire severity, this analysis used only a subset of fires for which these biases are minimal (Appendix A). For this analysis, fire severity classes based on \( BA(n_i)_{surviving} \) were defined as high = \( \leq 20\% \), moderate = 21-69\% and low = \( \geq 70\% \). These threshold values are similar to the values used in other studies (Agee 1993, Hessburg et al. 2007, Schoennagel et al. 2011) but were modified to slightly broaden the low severity class and narrow the high severity class. Our calibration procedure iteratively classified fire severity from \( \Delta S(n_i) \) using a factorial range of thresholding values until the threshold combination that maximized agreement with \( BA(n_i)_{surviving} \) was found. These optimal breakpoints were then applied to \( \Delta S(n_i) \) for all fires to assign a calibrated severity class to each fire. Once fire severity classes had been assigned to each fire event, the cumulative fire severity history for each patch was defined as: (1) low severity if all fires were of low severity; (2) non stand-replacing (NSR) if fires were a mix of low and moderate severity; (3) mixed severity if at least one high and one low or moderate severity fire occurred within a patch or if all fires were moderate severity; or (4) high severity if all fires within a patch were high severity.
Spatial patterns of fire history information along a fire regime gradient

To expand the fire regime gradient represented by our dataset we combined previously published dendroecological data from central Oregon (Heyerdahl et al. 2014, Merschel et al. 2014) and southwestern Colorado (Tepley & Veblen 2015) with our data from the northern U.S. Rockies. These studies were selected because: (1) they represent sites with differing fire regimes along the low to mixed-severity gradient, (2) they used similar sampling strategies and recorded the area over which increment cores were collected from trees, allowing conversion of frequency counts of tree ages to a density distribution that forms the basis of the fire severity metric used here, and (3) each dataset contained paired tree establishment dates and fire scars. The central Oregon data were obtained from the International Multiproxy Paleofire Database (www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/fire-history, accessed online 9/23/2016). It included four study sites (Potholes, Lytle Creek, McKay Creek and Green Ridge), each dominated by dry mixed-conifer forest (Table 2.2). The southwestern Colorado dataset included samples from two watersheds in the San Juan Mountains dominated by dry- and moist-mixed conifer forests (Table 2.2).

We compared fire severity along the fire regime gradient by comparing the slope of linear regressions fit to the log-transformed $\bar{d}(n)$ values for each region (Tepley & Veblen 2015), where $\lambda$ represents the absolute value of the slope of the regressions. The higher the value of $\lambda$ for a region, the more rapid the decrease in the average density of surviving trees with increasing number of fires when
extending the record back in time, and therefore the higher the predominant fire severity for the region. We also evaluated the patchiness and spatial bias of dendroecological data along this gradient by plotting the cumulative $d_{AF}$ curve, calculated as the distance between each age structure plot in a region and the nearest fire scar that contributed to its fire history. To assess the spatial bias of the fire history evidence, we calculated the topographic position index (TPI) class from a 10m DEM obtained from the National Elevation Dataset (Gesch et al. 2002) for each pixel containing a fire scar and compared the distribution of these values to the expected distribution for each region. The expected distribution of TPI classes was calculated as the proportion of total area encompassed by each TPI class for each watershed. The TPI is a scale-dependent measure of the elevation of a focal pixel relative to the mean elevation of surrounding pixels. We calculated TPI using an annulus of 300m, which represents relatively fine-scale topographic position. TPI classes that group pixels of similar relative terrain shape were constructed by thresholding TPI into 6 groups: 1=flat areas, 2=valleys, 3=lower, 4=middle, 5=upper slopes and 6=ridgelines. TPI classes were derived by standardizing relative to the mean TPI value from all pixels across all regions and placing thresholds at each 0.5 SD interval around the mean (Weiss 2001).

We tested for statistical differences between observed and expected TPI of fire scars using a $\chi^2$ test. To evaluate the strength of the relationship between the fire regime gradient and (1) the $d_{AF}$ and (2) the spatial bias of fire scars, we used linear regression models with the $\lambda$ (i.e., the slope of the log-transformed $\bar{d}(n)$ curve)
for each region as the predictor variable and the mean $d_{AF}$ and mean TPI class of
pixels with fire scars in each region as the response variables. Prior to modeling, the
response variables were $z$-score transformed to standardized effect sizes.

**Examining bias due to compositing fire history information within patches**

To evaluate the influence that compositing fire scar samples within patches
has on fire history interpretations, we compared the distribution of fire intervals
from composite and point (i.e. individual trees) samples. Two types of composite fire
records were constructed, one with all unique fire years found on all fire-scarred
cross-sections within a patch (referred to as the composite record) and the other
from these same records plus any unique fire dates inferred from age structure
(referred to as the full composite record). The median composite fire interval (MCFI)
was calculated as the median interval of fires in these records. Point fire intervals
may under or overestimate the true fire interval by either not recording scars for
every fire or by recording fires that affected only a small portion of the stand.
Although they may represent an incomplete fire history record (Swetnam and
Baisan 1996), point records are not biased by sample size as is the MCFI (Baker
2009). The median point fire interval ($PFI_{\text{median}}$) was calculated as the median of all
fire intervals recorded in these point records. We also calculated the median point
fire interval for the fire scar sample (or the average of multiple samples in the case
of ties) recording the greatest number of fires in a patch ($PFI_{\text{max}}$). Compared to the
$PFI_{\text{median}}$, the $PFI_{\text{max}}$ should better approximate the complete fire history of a patch
without introducing the bias created by composite records. We tested for statistical
differences in the distribution of fire intervals obtained using each method using the k-sample Anderson-Darling (AD) test. To test whether the magnitude of bias was related to sample intensity, we evaluated the difference between composite and point intervals as a function of the number of samples collected within a patch using linear models.

*Compositing age structure data within patches*

To evaluate whether our dendroecological methods effectively captured patch-level age structure and whether photo-interpreted boundaries represented patches with internally consistent fire histories, we examined accumulation curves of age structure as a function of the number of plots sampled and the variation of age structure between plots within a patch. For this analysis we used the incidence (e.g. presence/absence) of unique fire-initiated cohorts in place of species as the basis of all calculations. Sample-based rarefaction was used to interpolate smoothed accumulation curves with increasing sample size for each patch using the Coleman method (Coleman et al. 1982) in the Vegan R package (Oksanen et al. 2016). We used sample-based, rather than individual-based rarefaction, because the former accounts for the underlying spatial structure of the data whereas the latter does not (Gotelli and Colwell 2001).

We used the Chao estimator (Chao 1987), which calculates the number of missing species based on the incidence of rare species (Oksanen et al. 2016), to estimate the total number of cohorts (i.e. extrapolated richness) within each patch. Richness estimators generally produce a lower bound richness estimate (Gotelli and
Colwell 2001) and the bias towards underprediction of total richness increases with smaller sample sizes (Colwell and Coddington 1994, Chiu et al. 2014), as is the case in this study. However, the Chao estimator has been found to perform better than other estimators in these cases (Colwell and Coddington 1994). We calculated the richness difference (extrapolated - observed) to estimate the magnitude of undersampling, if any, that occurred for each patch. We used linear regression models stratified by severity and cohort class to test whether patch size, rather than fire history, influenced extrapolated cohort richness. As a final check on the adequacy of our sample design, we also calculated the minimum number of sample plots required to reach extrapolated richness and compared it to the number of plots actually sampled for each cohort and severity class.

To evaluate within-patch age structure heterogeneity, we used the AD and Kruskal-Wallis (KW) tests to determine the number of patches that exhibited a statistically significant difference in the distribution of establishment dates between plots. Both methods are non-parametric tests of the null hypothesis that age structures from plots within a patch originate from the same distribution. The KW tests for differences in the central tendency of the rank order and assumes identical scale and shape of the sample distributions whereas AD makes no assumptions about distribution properties and is more sensitive to the shape of the sample distributions, including outliers. Within-patch heterogeneity is presented here as the percent of sites within a severity class where the null hypothesis is
rejected, suggesting differences in the distribution of establishment dates between plots within a patch.

**Results**

*Photogrammetric and dendroecological records*

We mapped patch boundaries and interpreted vegetation attributes on 738 patches, ranging in size from 2-674 ha (mean = 61.15 ha, median = 33.99 ha), in the NCDE. Patch boundaries were delineated for the 44 patches where dendroecological samples were collected, but spatial information for these patches could not be determined without orthorectification of the base imagery. Across the five watersheds in NCDE, we sampled and successfully crossdated 2,572 increment cores from 255 plots in 70 patches and 146 fire-scarred cross-sections containing 292 scars (Table 2.2). Tree establishment dates from these samples ranged from 1462-1999 A.D. and cross-sections yielded 40 unique fire years during the period 1634-1932 A.D. From our two study watersheds in the GYE, we sampled and crossdated 1,419 increment cores from 221 plots in 44 patches and 82 fire-scarred cross-sections containing 122 scars (Table 2.2). Tree establishment dates in the GYE spanned the 1495-1996 period and cross-sections yielded 34 unique fire years during the period 1607-1938. Most fire-scarred cross-sections contained only 1-2 scars (Fig. S1a, c), although cross-sections from ponderosa pine (median=3 scars/sample) and individual samples from other species sometimes recorded ≥ 3 fires. Western larch, Douglas-fir and ponderosa pine were the most important
recorders of fire in the NCDE, whereas Douglas-fir comprised most of the crosssections sampled in the GYE (Fig. S1b, d).

Drivers of tree establishment

The cohort detection algorithm identified 106 cohorts in the NCDE and 83 cohorts in the GYE (Fig. 2.3c, g). Over the period from 1600-1950, we documented 73 and 99 pluvial events lasting 1-15 years in the NCDE and GYE, respectively, with slightly longer average pluvials in the GYE (Fig. 2.3c, g). The BEA analysis indicated a lack of synchrony between pluvials and cohort initiation or peak years (Fig. 2.3d, h). Similarly, no relationship was found between cohort peaks and either fire or pluvials. However, synchrony between fire and cohort initiation was near the upper limit of the confidence envelope at multiple time lags over the initial 30 years following fire and was statistically dependent at lags of 10-23 years in NCDE and 10-13 years at GYE. Based on this analysis, we set t_{crit} to 20 years and only considered cohorts that initiated within a 20-year post-fire window to be due to fire.

Reconstructed fire regimes

Patch composite fire intervals for each region exhibited high variation around median values of 35.5 and 49 years for the NCDE and GYE regions, respectively (Fig. 2.4a). Variation in watershed-scale median fire intervals was high for the NCDE and low for the GYE, reflecting, at least in part, the greater geographic spread of our study areas in the NCDE compared to the GYE (Fig. 2.1, Table 2.2). The calibration procedure we used to set class thresholds for $\Delta S(n_i)$ showed strong patterns of agreement between $BA(n_i)_{surviving}$ and $\Delta S(n_i)$ for all three performance
statistics (Appendix A1). Although a number of thresholding values in each region would have produced similarly high classification accuracy, we selected the 30/80 and 20/70 $\Delta S(n_l)$ percentile thresholds for the NCDE and GYE, respectively. Of the 292 and 122 fire events documented in the NCDE and GYE, respectively, severity could not be calculated for 112 and 11 fires due to the absence of age structure data preceding high or repeated moderate severity fires. Median $\Delta S(n_l)$ values for fires where severity could be calculated were similar between regions but also exhibited high variation (Fig. 2.4b). The cumulative fire severity history of patches was classified as low, non stand-replacing, mixed or high for 0%, 17%, 46% and 37% in the NCDE and 0%, 14%, 47% and 40% in the GYE, respectively (Fig. 2.4c). The prevalence of moderate and high severity fire in both regions resulted in primarily young- (< 100 years) to intermediate-aged (100-200 years old) patches (Fig. 2.4d-e) with a mix of simple and multi-cohort age structures and successional conditions (Fig. S2-3). However, different landscape dynamics between regions are revealed by the higher proportion of young forest (< 100 years old) in the NCDE and the dominance of intermediate-aged forest patches in the GYE (Fig. 2.4d-f).

Spatial patterns of fire history information along a fire regime gradient

The four study areas (OR, NCDE, GYE, and SJ) we used to represent the fire regime gradient were successful in capturing a broad range of conditions consistent with low- and mixed-severity fire regimes. The fire regime of the central OR sites was characterized by frequent and predominantly low severity fires, as illustrated by the lowest $\lambda$ of the four study regions (Table 2.2, Fig. 2.5a). The San Juans
experienced relatively frequent fire, and exhibited intermediate density decay that was more similar to the northern Rockies region than central OR. The NCDE and GYE had longer and more variable fire frequency, steeper density decay slopes that reflect an important role of repeated moderate or high severity in these regions, and low tree survivorship beyond the most recent 5-7 fires per stand.

There was marked difference in the spatial patchiness and bias of fire scar records along the fire regime gradient (Fig. 2.5b-c). Central OR had very low d_{AF} values, with almost all age structure plots located within 500 m of fire scars evidence (Fig 5b). In contrast, a much greater number of plots were located 500-1,000 m from fire scar evidence in the other regions, with some plots located up to 1,500 m from the nearest fire scar. Chi-squared tests confirmed a bias in the topographic position of fire scar evidence in all regions (p < 0.01), but the magnitude and nature of this bias varied among regions (Fig. 2.5c). Fire scars were found more frequently than expected on middle slopes in all regions and most of the fire scar evidence in all four regions was found in this topographic position class. However, fire-scar evidence in central OR was more proportionally distributed throughout the landscape. The NCDE and GYE had a much greater than expected occurrence of fire scars on upper slopes and ridgetops and much less than expected in lower slopes and valley bottoms. The San Juan Mountains showed a similar pattern to the northern Rockies regions, but with a less pronounced bias towards upper slopes and ridgetops. Linear models of the relationships between λ (i.e., the absolute value of the slope of log-transformed \( \tilde{d}(n) \)) and the mean TPI class of pixels containing fire
scars was not significant \( (p > 0.10, R^2 = 0.52) \), whereas the model for \( d_{AF} \) was highly significant \( (p > 0.05, R^2 = 0.95) \).

The large \( d_{AF} \) and spatial bias that characterized portions of some study watersheds made interpretation of age structure difficult for some patches (Fig. 2.6). This is exemplified in Fig. 2.6a-c, which details a portion of Six Mile creek in the GYE that contained a region with large \( d_{AF} \). Some patches in this area had high quality fire history records that documented five fires years (Fig. 2.6d, i), 1741, 1771, 1803, 1820, and 1870. However, extensive field surveys within and between five neighboring patches (Fig. 2.6e-h) revealed no further fire- scarred trees. Using the approach outlined in Table 2.1, we were able to attribute the majority of cohorts identified by our algorithm to fire dates from adjacent patches with relatively high confidence (Fig. 2.6e-i). In one patch (Fig. 2.6g,) we assigned the 1820 fire date based on a weak establishment pulse, the patch location entirely within the 1820 fire perimeter (Fig. 2.6b), and widespread evidence of the 1820 fire in all other surrounding patches. There were a small number of cases where a patch was partially overlapped by a fire perimeter but no age structure information or other direct evidence was available to inform the decision (e.g. 1741 fire, Fig. 2.6e). In these cases, we considered the evidence too weak to assign the fire perimeter date to the patch. A small number of cohorts identified by our algorithm were attributed to non fire causes, either due to their species composition (Fig. 2.6d) or because the cohort initiation date fell outside the \( t_{\text{crit}} \) threshold identified by BEA (Fig. 2.6e). Because fires may occur without initiation of distinct cohorts or if subsequent fires
erase evidence of these cohorts (1803 and 1870 fire, Fig. 2.6d) this approach may underestimate fire frequency for cases where both fire scar and cohort evidence is lacking. However, these cases are most likely for sites where multiple non stand-replacing fires have occurred and the predominant fire severity was classified as low or NSR (Fig. 2.6d, g). As a result, omission of fires due to a lack of fire history evidence is likely to affect fire frequency estimates more than severity classifications.

_Fire history bias due to compositing_

Compositing fire history data at the patch level did not significantly bias the distribution of fire intervals (Fig. 2.7a, d). There were no statistically significant differences between any of the point and composite intervals (Anderson-Darling test, p > 0.1), although addition of unique fire dates evidenced by age cohorts in the GYE noticeably modified the probability distribution of fire intervals. We interpret this to be a result of the critical role that age cohorts played in reconstructing fire history in the GYE, where dAf was high and a disproportionate percentage of fire scars were found on high ridges or upper slopes (Fig. 2.5b-c), fire free intervals were the longest (Table 2.2), and individual fire scar samples recorded few fires (Fig. S1c). Linear models showed no significant trends (p > 0.1, R² < 0.1 for all tests) in the differences between the MCFI and the PFI_connected or the PFI_max with increasing sample size in either region (Fig. 2.7b-c, e-f).
Characterization of patch-level age structure

Accumulation curves of within-patch cohort richness (i.e. number of cohorts per patch) demonstrate that most double- and some multi-cohort stands in both regions reached, or nearly reached, asymptotic richness using our sample methodology (Fig. 2.8a, c). Some multi-cohort sites in both mixed and NSR sites did not reach an asymptote (Fig. 2.8a-d), although most of these sites showed a decreasing rate of cohort accumulation that indicates the beginning of saturation. Differences between observed and extrapolated patch cohort richness were small (e.g. < 1 cohort) (Table 2.3), indicating that few distinct aged cohorts likely would have been detected with further sampling even in most sites where asymptotic cohort richness was not reached. In the NCDE, where patches were orthorectified, no statistically significant relationship was found between extrapolated cohort richness and patch size (Fig. 2.8e-f) for any cohort or severity class (p > 0.05, adjusted R² < 0.2). Calculation of the number of plots required for cohort richness to reach saturation in mixed and NSR patches corroborated that our sample sizes were likely sufficient to capture patch-level age structure for most sites (Table 2.3). Comparison of the distribution of establishment dates between plots within a patch using the AD- and KW-tests showed that within-patch heterogeneity of age structure is generally low for high and NSR patches. Within-patch heterogeneity was highest for mixed-severity patches characterized by double- and multi-cohort age structures (Table 2.3). This suggests that within-patch heterogeneity of age structure in our study area varies with disturbance history but over small spatial...
scales (e.g. < 100 m that separated age plots within patches) and can be accurately characterized with a relatively small number of dendroecological sample plots (average 3-5 plots) for most patches.

**Discussion**

*Age structure is closely linked to patch-level fire effects*

Our use of age structure to identify fire events where fire-scar evidence was lacking relies on accurate attribution of cohorts to specific fire events and exclusion of alternative, non fire causes of cohort initiation (e.g. stand dynamics or pluvials). We present strong evidence that age cohorts in our study regions are initiated primarily by fire or other canopy-opening disturbances (Fig. 2.3). Using the time lags over which fire and tree establishment are significantly synchronous, we were able to attribute fire dates from neighboring patches to age structure in patches that lacked sufficient fire history information with high confidence in most cases (Fig. 2.6). Our data contribute to a growing body of evidence across a range of ponderosa pine, dry and moist mixed-conifer forests that finds little support for a direct role of favorable climate (Brown and Wu 2005) on the initiation of cohorts in the absence of canopy disturbances (Heyerdahl et al. 2006, Meunier et al. 2013, Dugan and Baker 2015, Tepley and Veblen 2015, Iniguez et al. 2016). Instead, they corroborate a general pattern that has emerged from recent studies in which climate may initiate tree establishment indirectly by reducing fire frequency in drier forests where high fire frequency generally limits tree survival and maintains open canopy conditions. In more mesic forests with lower fire frequency, such as
those of our study regions, tree establishment in the fire-free decades following fire create a strong negative feedback on subsequent establishment of shade intolerant trees. This is evident in many patches within both regions that experienced long periods with little or no recruitment of shade intolerant species in the absence of fire (Fig. 2.3e, f, Fig. 2.6f, h, Fig. S2b, d, e, f, Fig. S3a, c-e), despite periods of favorable climate (Fig. 2.3c, g).

The BEA results and inspection of recruitment patterns for individual patches (Figs. S2-3) also suggests that post-fire recruitment is more protracted than would be expected for mixed-conifer forests of our study region that are characterized by intermediate moisture conditions. Many studies document post-fire recruitment initiation immediately following fire or up to < 10 years post-fire, especially following low-moderate severity fire or in forests with intermediate moisture conditions (Chappell and Agee 1996, Ehle and Baker 2003, Sherriff and Veblen 2006, Kemp et al. 2015, Tepley and Veblen 2015). In our study regions, BEA revealed that recruitment initiation did not occur until 10-20 years following fire, while peak recruitment often occurred decades later and was not consistently associated with fire (Fig. 2.3d, h). We interpret the cause of the latter result to be the high variability in the timing of peak post-fire recruitment in our study areas, which varied between 10-60 years (Fig. 2.3a-b, e-f, Figs. S2-3). Such broad recruitment windows are notable because the climate of our study regions is more favorable for recruitment than in many ponderosa pine or dry mixed-conifer forests across the West.
Based on these observations, we speculate that protracted recruitment periods may be a historical feature of patches in MSFR forests where high severity fire occurred in large enough patches to limit abundant seed dispersal to patch interiors. Studies from modern wildfires suggest that seed dispersal in forests similar to those we studied declines sharply after 100-400 m from seed source (Donato et al. 2009, Kemp et al. 2015, Harvey et al. 2016b). In these cases, protracted tree recruitment occurs because trees infill burned patches slowly through rare long-distance dispersal events and gradual progression from the nearest seed sources (Haire and McGarigal 2010). Protracted historical tree recruitment lasting ~30-75 years post-fire has been documented in coastal Douglas-fir forests following high- and some mixed-severity fire (Freund et al. 2014, Tepley et al. 2014) and in some ponderosa pine and mixed-conifer forests (Nagel and Taylor 2005, Sherriff and Veblen 2006). In contrast, in dry ponderosa pine and many mixed-conifer forests, limited tree regeneration in the interior of high severity patches has been interpreted as a signal of recruitment failure (but see Haire and McGarigal 2010) and possible vegetation type conversions from forest to shrub- or grass-dominated cover types (Savage and Mast 2005, Strom and Fule 2007, Roccaforte et al. 2012, Collins and Roller 2013). If protracted tree establishment is a common feature of some MSFRs, then low levels of tree recruitment in the first several decades following fire may not be strong indicators of long-term recovery trends or vegetation conversion to alternative states. Protracted recruitment may, in fact, have been an important process that maintained compositional, structural
and successional diversity within historical landscapes (Swanson et al. 2011, Hessburg et al. 2016). Further research is needed to better understand the spatio-temporal patterns of forest recovery in relation to fire severity, resilience mechanisms in MSFRs, and the geographic variation in the ecological role played by high-severity fire across the broad distribution of MSFRs.

**Dendroecological methods should be scaled with the fire regime gradient**

We present evidence from a diverse range of MSFR forests that the scarcity, spatial patchiness and bias of fire history records scales with the local fire regime (Figs. 5, S1), consistent with our predictions (Fig. 2.1). Although our dataset does not span the entire fire regime gradient, previous research in low- and high-severity fire regime forests suggests that they generally fit the pattern of increasingly sparse fire history evidence with increasing prevalence of high severity fire (Parsons et al. 2007, Farris et al. 2010, Cyr et al. 2016). This pattern presents significant challenges and uncertainty and requires greater reliance on stand-level age structure patterns to reconstruct fire history and severity as fire scar evidence is lost. Although we highlight the challenges presented by high $d_{AF}$ here (Fig. 2.6), the increasing spatial bias of fire history evidence towards upper slopes and ridges for MSFRs with greater amounts of high severity fire that we document can also be problematic because these features form natural topographic barriers to fire spread and therefore often coincide with fire perimeters and patch boundaries (Swanson 1998, Dorner et al. 2002). If tree establishment patterns in a patch that coincide with a fire date from a ridgetop fire scar are somewhat ambiguous, it can be difficult
to determine whether the ridgetop fire scar recorded a low severity fire that affected patches on lower slopes and produced only a weak establishment response or if the fire stopped at the ridgetop and the weak establishment pulse in patches around the recorded fire date is stochastic variation.

Despite these challenges, the approach we present here for attributing fire dates to patches based on the presence of fire-initiated age cohorts (Table 2.1) was effective in limiting uncertainty in all but a small number of age cohorts (Fig. 2.6). We propose that despite large $d_{AF}$ and spatial bias in some MSFRs, the strategy we propose here will be generally effective in reconstructing fire regime properties because both the grain of the patch mosaic and the rate of evidence loss are a function of the high-severity burn rate and scale in a system. Thus, where high fire frequency creates fine-scale heterogeneity of age structures and fire severity effects, more fire scar evidence will exist on the landscape and uncertainty in fire history reconstructions will be relatively low. As fire history evidence is lost along the fire regime gradient due to increasing prevalence of high severity fire, less fire history evidence is necessary to describe the coarsening landscape patch mosaic and age structure will be increasingly relied upon. We do not explicitly test changes in the spatial grain of the patch mosaic along the fire regime gradient in this paper because patch boundaries were not evaluated in the auxiliary data we used from central OR and the San Juans. However, the combined dendroecological and photogrammetric methods we present here could be applied in other areas to evaluate this question.
The fire regime gradient we constructed to test our theoretical framework consisted of regionally-aggregated dendroecological data that demonstrated broad-scale differences in the fire regimes of each region but ignored internal variation of fire regime properties within regions. For example, the local fire regime for BC1 in the NCDE and POT in central OR had substantially longer fire return intervals (Table 2.2) and higher severity than the other study sites within their respective regions. The fire severity metric and spatial metrics of fire scar evidence we employed are all derived from fine-scale plot and point data that can be aggregated at multiple scales (i.e. plot, patch, watershed, or region). Therefore, our framework should be easily transferable across multiple scales and to other study systems where mixed-severity fire effects are important.

In MSFRs, multi-cohort patches may exist in close proximity to patches with much simpler disturbance history and age structure that can be accurately characterized with much lower sample effort. We show that the sample effort needed to characterize patch-level age structure is related to the number of cohorts within a patch and its fire severity history (8a-d, Table 2.3). Single- and double-cohort patches can be effectively characterized with two or fewer plots (Table 2.3). Even in multi-cohort patches with more complex fire severity histories, a moderate number of spatially distributed sample plots (i.e. 3-4 plots) was generally adequate to characterize age structure and fire severity history in our study regions, although some sites required sampling of up to five plots (Fig. 2.8a-d, Table 2.3). Therefore, in MSFRs the most efficient sample strategy will incorporate a flexible design, with
the sample intensity and spatial distribution of plots within patches contingent on the apparent complexity of patch-level fire history. Preliminary assessments of disturbance history can be made with detailed aerial photo interpretation prior to field sampling, as we have done here, or by field assessment of fire history (e.g. qualitative estimation of fire frequency from visible fire scars within stands) and age structure complexity (e.g. field counts of tree ages from multiple plots within a patch). The risk of under-sampling using this scheme can be moderated by conservative interpretation of disturbance history that is biased towards more complex effects or by a phased sample design wherein initial sample collection is supplemented by subsequent data collection for patches where initial age structure data is deemed insufficient.

We present a calibration method for the average density loss function, \( \bar{d}(n) \) (Tepley & Veblen 2015), and demonstrate that it is a useful metric for comparing the predominant fire severity among regions (Fig. 2.5a). We attribute the strong agreement between \( BA(n_t)_{surviving} \) and \( dS \) to the fact that our study areas were characterized by a broad mix of fire severities that spanned the absolute range of possible fire effects. Thus, \( \Delta S(n_t) \) in our study regions may closely approximate absolute fire severity. This is likely not to be the case for many areas where fire regimes are dominated by a narrower range of fire effects, e.g. the OR study region where fires were frequent and mostly low to moderate severity (Fig. 2.5a). An alternative to the calibration procedure we used here that should be explored further is cross-regional calculation of \( d(n) \). The \( \Delta S(n_t) \) metric is essentially a
measure of the change in the ranking of a patch relative to other patches in the region (or whatever pool of sites is used to calculate $\tilde{d}(n)$) caused by a specific fire event (Tepley & Veblen 2015), so it therefore should be minimally biased by biophysical factors (e.g. edaphic or topographic factors) that affect absolute stand density. This same logic could theoretically be applied to the extension of $\Delta S(n_i)$ calculations across a broader range of conditions than found within regions or small study areas. The benefit of this approach is that existing dendroecological datasets from public archives, such as those we used here to examine regional fire regime gradients, could be compiled into regional or continental scale networks and used to calculate $\tilde{d}(n)$. As a greater number of data from more diverse sites is incorporated into calculations of $\tilde{d}(n)$ and $\Delta S(n_i)$, this relative metric should more closely approximate absolute fire severity. This is an area of needed future research.

Although MSFRs are primarily defined by variation in fire severity, they also present notable analytical challenges to estimating fire frequency. For example, it is likely that our sample estimate of the median fire return interval and variation around the MCFI, is low, due to the presence of some doubly- and many right-censored fire intervals that were excluded from our fire frequency estimates. Doubly- and right censored-intervals occur when no fires or only one fire is recorded at a site, respectively (Polakow & Dunne 1999, Cyr et al. 2016). Both situations occurred in both of our study regions (Fig. 2.9), although right-censored intervals documented in old forest patches recovering from a single, old high severity fire were much more common in both regions. As is standard practice in fire history
reconstructions (Reed & Johnson 2004, Grissino-Mayer 1999), we included only scar-to-scar intervals in our fire interval estimates, so these double- and right-censored intervals are excluded. As a result, our summary statistics are likely biased towards shorter intervals and reduced variability. Even in frequent fire systems where longer fire history records exist (Swetnam & Baisan 1996), fire-free intervals are generally short, and the timing of fire exclusion is easy to detect, censored data can result in biased fire interval estimates (Polakow & Dunne 1999, Moritz et al. 2009). In MSFRs, censorship of fire intervals may produce stronger biases in fire frequency estimates, in part because fire history statistics are known to be sensitive to accurate sampling of long fire intervals that define the tail of the distribution (Finney 1995). Analytical techniques have been developed to account for incomplete intervals (Polakow & Dunne 1999), but they are not widely used and should be considered in future studies of MSFRs.

Paired photogrammetric and dendroecological records allow robust reconstruction of historical fire-mediated stand dynamics, spatial pattern, and spatial scale of fire effects

A fundamental paradigm in landscape ecology is that landscape processes (e.g. fire) and patterns occur at a series of nested hierarchical scales (Wu & Loucks, Urban et al. 1987, Hessburg et al. 2015, Turner 1989). Ecological phenomena are often driven by non linear, cross-scale interactions and threshold behaviors that can only be understood, therefore, through evaluation of spatio-temporal dynamics across a range of scales (Peters et al. 2004, Pascual & Guichard). Dendroecological
data provide robust temporal proxy records of historical information (e.g. climate, disturbance history) for particular points or small plots in the landscape that have been used to effectively describe small-scale ecological dynamics and coarse-scale spatial phenomena, i.e. synoptic forcing of fire frequency (Kitzberger et al. 2007), but they have been much less effectively used to describe meso-scale landscape dynamics. We show that point and small-plot dendroecological data can be used to accurately characterize patch-level age structure and fire severity history across a range of patch sizes, forest types and fire effects (Fig. 2.8a-b, Table 2.3). This validates a key, but poorly tested, assumption of photogrammetric analysis – that patch boundaries circumscribe areas with internally consistent heterogeneity and are therefore ecologically meaningful units for landscape analysis. The critical benefit, of our approach is that dendroecological interpretation of fire-mediated dynamics can be scaled to the patch-level, where photogrammetric analysis can be used to document the scale and pattern of the patch mosaic. Because patches are the fundamental units used in most landscape-scale analyses (Gustafson 1998), our approach can be efficiently applied across scales from individual patches to larger landscapes.

*Spatio-temporal intermixing of fire frequency and severity created landscapes dominated by younger single- and multi-cohort patches*

Patch-level reconstructions of historical fire frequency and severity in our study regions depict landscapes driven by highly variable fire frequencies and non-equilibrium patch-level dynamics caused by a temporal mix of different fire effects
and frequencies. High severity fire affected > 80% of the forested landscape in both regions at least once during the study period, much of it within the last 100-200 years (Fig. 2.4d-e). This estimate does not include any grasslands or shrubfields that may have been created by high severity fire but were not sampled in this study. The broad influence of high severity fires is reflected in the aggregate age distributions (Fig. 2.3c, g), which show low survival of synchronous tree recruitment pulses prior to the early 1800s in the NCDE and 1700s in the GYE. Approximately 40% of the forested landscape in each region was dominated by early seral and even-aged forest of different ages (Fig. 2.4d-e). However, the remaining area burned by high severity fires subsequently reburned by one or multiple low-moderate severity fires (mixed severity patches in Fig. 2.4c-f), producing young- to intermediate-aged, double- and multi-cohort patches (Fig. S2-3) across ~45% of the forested area in each region. The remaining forested area was mostly older forests (Fig. 2.4d-e) shaped either by relatively frequent low-moderate severity fires, infrequent low-moderate severity fires, or frequent low-moderate severity fires following an old high severity fire. Frequently burned patches were characterized by open-canopy, multi-aged forest structures and stand dynamics similar to those documented for many dry-mixed conifer forests with low severity fire regimes (Figs. S2a-b, S3a), but were classified as NSR or old mixed-severity, rather than low-severity, patches due to a significant number of moderate severity fires or an old high severity fire 250-450 years ago. The most striking feature of both landscapes is the broad mix of non forest, even-aged patches and patches with two or more
cohorts that occurred in relatively close proximity. These findings are consistent with the idea that MSFRs exhibited higher beta diversity (i.e. differences in compositional and structural complexity between patch) than alpha diversity (i.e. within-patch structural complexity) (Perry et al. 2011, Agee 2005).

Possible evidence for nonlinear relationships between fire frequency and severity

Theory and some empirical observations suggest a negative relationship between fire frequency and severity that should lead to a greater representation of old, multi-cohort forests in landscapes with higher fire frequency (Lertzman et al. 1998, Lutz et al. 2011, Steel et al. 2015). In contrast to this expectation, higher median fire frequency in the NCDE (MCFI_{median} =35.5) compared to the GYE (MCFI_{median} =49) actually resulted in a greater prevalence of young forest recovering from high severity fires (Fig. 2.3c, g, Fig. 2.4d-e). Moreover, for both regions, median fire frequency of NSR and mixed-severity patches was similar (Fig. 2.9), despite very different disturbance dynamics and structural properties. This appears to contradict the generally accepted frequency-severity relationship and suggests that mean or median fire frequency estimates are poor indicators of the local fire regime.

We suggest that these observations can be reconciled with theory if (1) the fire frequency-severity relationship is characterized by a strong nonlinear threshold behavior and (2) this threshold behavior is related as much to the variability of fire frequency as it is to measures of central tendency (e.g. mean, median). Nonlinear relationships between fire frequency and severity could explain the inconsistent support it has received in empirical studies (Lutz et al, van Wagendonk et al. 2012,
Steel et al.), and why the relationship does not appear to hold for our study regions, which are characterized by intermediate fire return intervals. Consistent with this idea, empirical results from resumed fire regimes in portions of the western U.S., including our study regions, suggest that negative severity feedbacks may occur primarily when fire intervals are quite low, < 10-25 years (Parks et al. 2014, Harvey et al. 2016a). Above this mean threshold, as is the case in our study areas, variability around the MCFI increases resulting in much more mixed severity fire effects. Evidence of this can be seen in the increasing variability in median fire intervals of NSR versus mixed-severity patches, despite the similar median values (Fig. 2.9). This pattern suggests that consistently frequent fire is needed to maintain a strong negative feedback of previous fire on subsequent fire severity.

This is important for two reasons: (1) the assumption that mean or median fire interval summary statistics are a good indicator of the overall fire regime, and fire severity specifically, is not supported by our data. (2) variability around the mean or median fire interval may be as important and informative as central tendencies, and (3) there may be a fire frequency threshold below which negative fire-fire feedbacks maintain low severity fire, whereas above this threshold variable-severity fire is more likely. Identifying the fire frequency characteristics that define this threshold behavior is critical to understanding MSFRs, especially in relation to climate-forced or management-induced changes in fire frequency. It also suggests that proxies of ecological departure based on mean fire interval, such as fire return
interval departure (Hann and Bunnell 2001, van Wagendonk et al. 2012), may be uninformative and potentially misleading, at least for MSFRs.

Conclusions

The recognition that MSFRs are much more widespread than historically acknowledged has sparked great interest in understanding their distribution, landscape dynamics, and drivers. The relative dearth of large-scale, spatio-temporally robust historical datasets available to characterize the landscape variability, dynamics, geographic distribution and spatial ecology of historical MSFRs is principally due to the methodological limitations enumerated here. These methodological barriers have led to significant debate about the nature of historical fire regimes, the implications of MSFR dynamics for management, resilience mechanisms in MSRFs, and likely outcomes of increased climate-driven increases in fire activity in MSFRs (Williams and Baker 2012, DellaSala et al. 2013, Fulé et al. 2013, Franklin et al. 2014, Odion et al. 2014, Stevens et al. 2016). We present new insights about historical fire regime dynamics from two MSFR forests that contrast with those described for low- and high-severity forests and contribute to our expanding understanding of the diverse landscape dynamics that historically characterized MSFR forests. Rather than debate aspects of a single monolithic MSFR model, we place our results in the context of the fire regime gradient and we provide a quantitative framework for scaling dendroecological methods along this gradient that can be expanded and adapted as more data are incorporated.
**Acknowledgements**

This research was supported by a National Science Foundation Partnerships for International Research and Education (PIRE) grant (award no. 0966472), an NSF Doctoral Dissertation Research Improvement grant (award no. 1302233), the Jerry O’Neal National Park Service Student Fellowship, and a Dissertation Completion Grant from the Graduate School at CU Boulder. We thank the CU Undergraduate Research Opportunities Program for generous support of the many undergraduate students involved with this project. For field and laboratory assistance, we thank: S. Diaz, A. Simler, P. Wickey, J. Scharff, K. Anderson, and B. Heitshusen. We also thank E. Heyerdahl (Fire Lab) and E. Sutherland (RMRS) for the generous use of their tree ring facilities, D. Divoky (Glacier National Park), T. Carolin (Glacier National Park) and T. Wehunt (Flathead National Forest) for logistical support, and P. Ohlson (Okanogan-Wenatchee NF) for photogrammetric training.
Table 2.1. Table outlining a range of scenarios of availability of fire history information encountered in mixed-severity fire regime forests and the distinct attribution methods that we used in each scenario to maximize the accuracy of assigning specific fire dates to age cohorts. Availability of fire history information decreases from Scenario 1-5, with increasingly sparse, indirect or uncertain evidence.

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>cohort present</th>
<th>fire documented within age structure plot</th>
<th>fire documented within patch</th>
<th>fire documented in adjacent patch</th>
<th>patch is mostly within reconstructed fire perimeter</th>
<th>patch borders a reconstructed fire perimeter</th>
<th>Attribution method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Cohort assigned fire date from within age structure plots or the larger patch.</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Cohort assigned to fire date from adjacent patch.</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Cohort assigned to fire date from fire perimeter.</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>If patch exhibits strong tree establishment pulse within 20 years of bordering fire perimeter, assign fire date from fire perimeter.</td>
</tr>
<tr>
<td>Scenario 5</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>Do not assign fire date or assign fire date only if indirect evidence supports extension of fire throughout patch. Auxiliary evidence includes:</td>
</tr>
</tbody>
</table>

- **Species traits** — If patch age structure shows significant establishment of fire-sensitive species closely following, but not preceding, the fire perimeter date, assign the fire perimeter date to patch.

- **Topographic influences on fire spread** — If there is evidence of a small recruitment event and the patch is located on a topographic position where fire was likely to spread into patch (i.e. immediately uphill of fire perimeter), assign the fire perimeter date to patch.
Table 2.2. Summary information describing the forest types, dendroecological data, and fire regime properties of each of the four study regions. Study area codes are: GRR = Green Ridge, LYT = Lytle Creek, KAY = McKay Creek, POT = Potholes, SM = Squaretop Mountain, WC = Williams Creek, BC1 = Big Creek (NCDE), EC = Emery Creek, MF = Middle Fork, SB = Spotted Bear, SV = Swan Valley, SM = Six Mile, BC2 = Big Creek (GYE). Forest type abbreviations are: PIPO = Pinus ponderosa, PICO = Pinus contorta, PIMO = Pinus monticola, PIFL = Pinus flexilis, PSME = Pseudotsuga menziesii, LAOC = Larix occidentalis, ABGR = Abies grandis, ABCO = Abies concolor, ABLA = Abies lasiocarpa, CADE = Calocedrus decurrens, JUOC = Juniperus occidentalis, JUSC = Juniperus scopulorum, PIEN = Picea engelmannii, POTR = Populus tremuloides, and BEPA = Betula papyrifera.
<table>
<thead>
<tr>
<th>Region</th>
<th>Study area</th>
<th>Forest types‡</th>
<th># age plots</th>
<th># cores</th>
<th># fire-scarred cross-sections</th>
<th>Median fire interval (years)</th>
<th>slope of d(n)</th>
<th>Severity regime description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR¹</td>
<td>GRR</td>
<td>PIPO, PSME, ABGR, LAOC, CADE</td>
<td>30</td>
<td>902</td>
<td>103</td>
<td>23</td>
<td>-0.13</td>
<td>Mixed-severity fire regimes characterized by frequent low-moderate severity fires, with small-scale high severity fire effects. POT experienced a higher frequency and spatial scale of high severity than the other study areas.</td>
</tr>
<tr>
<td></td>
<td>LYT</td>
<td>PIPO, ABGR, PSME, JUOC, LAOC</td>
<td>30</td>
<td>901</td>
<td>155</td>
<td>15</td>
<td>-0.07</td>
<td>Mixed-severity fire regimes characterized by frequent low-moderate severity fires, with small-scale high severity fire effects. POT experienced a higher frequency and spatial scale of high severity than the other study areas.</td>
</tr>
<tr>
<td></td>
<td>KAY</td>
<td>PIPO, ABGR, PSME, LAOC</td>
<td>30</td>
<td>899</td>
<td>124</td>
<td>16</td>
<td>-0.08</td>
<td>Mixed-severity fire regimes characterized by frequent low-moderate severity fires, with small-scale high severity fire effects. POT experienced a higher frequency and spatial scale of high severity than the other study areas.</td>
</tr>
<tr>
<td></td>
<td>POT</td>
<td>PICO, PIPO</td>
<td>30</td>
<td>909</td>
<td>69</td>
<td>58</td>
<td>-0.56</td>
<td>Mixed-severity fire regimes characterized by frequent low-moderate severity fires, with small-scale high severity fire effects. POT experienced a higher frequency and spatial scale of high severity than the other study areas.</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>120</td>
<td>3,611</td>
<td>451</td>
<td></td>
<td>-0.1</td>
<td>Mixed-severity fire regimes characterized by frequent low-moderate severity fires, with small-scale high severity fire effects. POT experienced a higher frequency and spatial scale of high severity than the other study areas.</td>
</tr>
<tr>
<td>SJ²</td>
<td>SM</td>
<td>PIPO, PSME, POTR, ABCO</td>
<td>235</td>
<td>2,620</td>
<td>45</td>
<td>21</td>
<td>-0.31</td>
<td>Mixed-severity fire regime characterized by fairly frequent fires and variable severity. Many fires were low-moderate severity, but occasional high severity fires also occurred in portions of each study area.</td>
</tr>
<tr>
<td></td>
<td>WC</td>
<td>PIPO, PSME, POTR, ABCO</td>
<td>165</td>
<td>1,715</td>
<td>46</td>
<td>24</td>
<td>-0.26</td>
<td>Mixed-severity fire regime characterized by fairly frequent fires and variable severity. Many fires were low-moderate severity, but occasional high severity fires also occurred in portions of each study area.</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>400</td>
<td>4,335</td>
<td>91</td>
<td></td>
<td>-0.27</td>
<td>Mixed-severity fire regime characterized by fairly frequent fires and variable severity. Many fires were low-moderate severity, but occasional high severity fires also occurred in portions of each study area.</td>
</tr>
</tbody>
</table>

† Mixed-severity fire regime characterized by intermediate fire frequency and highly variable severity. Most patches experienced high severity fires at variable intervals, with more frequent low-moderate severity fires in between high severity fires.

‡ Mixed-severity fire regime characterized by intermediate fire frequency and highly variable severity. Most patches experienced high severity fires at variable intervals, with more frequent low-moderate severity fires in between high severity fires.

¹ Data sourced from: Heyerdahl et al. 2014, Merschel et al. 2014
² Data sourced from: Tepley & Veblen 2015
³ Data sourced from: this study
‡ Forest types are listed in order of abundance within each study area
! Fire frequency statistics were not calculated for EC because only two patches were sampled within this watershed.
Table 2.3. Observed and extrapolated cohort richness (e.g. # cohorts per plot), richness difference (Extrapolated - Observed), # of plots sampled, # of plots required to reach saturation, and the within-patch heterogeneity for all sites stratified by cohort class and cumulative fire severity class for both the NCDE and GYE.

Observed richness is the cumulative number of fire-initiated cohorts comprising at least 10% of stand basal area at a site. The estimated cohort richness was derived using the Chao estimator (Chao 1987, Chiu et al. 2014). The number of plots required to reach the estimated saturation point (i.e. the extrapolated cohort richness) was derived from the species accumulation curve for each site (Ugland et al. 2003). The percent of all sites within a row that exhibited statistically significant (α = 0.95) differences in the distribution of establishment dates between plots within a patch (i.e. the within-patch age structure heterogeneity) was determined using the non-parametric k-sample Anderson-Darling (AD) and Kruskal-Wallis (KW) tests. Correlation between AD and KW was relatively high (r = 0.72), indicating good agreement between the tests.

<table>
<thead>
<tr>
<th>Cohort class</th>
<th>Observed cohort richness</th>
<th>Extrapolated cohort richness</th>
<th>Richness Difference</th>
<th># plots sampled</th>
<th># plots to saturation</th>
<th>Within-patch heterogeneity (% sites)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Med</td>
<td>Mean</td>
<td>Max</td>
<td>Min</td>
<td>Med</td>
</tr>
<tr>
<td>NCDE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Double</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Multiple</td>
<td>3</td>
<td>4</td>
<td>3.88</td>
<td>5</td>
<td>0</td>
<td>0.08</td>
</tr>
<tr>
<td>GYE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Double</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Multiple</td>
<td>3</td>
<td>4</td>
<td>4.08</td>
<td>7</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Severity class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCDE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>2</td>
<td>2</td>
<td>2.52</td>
<td>5</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>NSR</td>
<td>3</td>
<td>3</td>
<td>3.08</td>
<td>5</td>
<td>-1</td>
<td>0.17</td>
</tr>
<tr>
<td>GYE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>1</td>
<td>1.12</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>1</td>
<td>2</td>
<td>2.35</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NSR</td>
<td>3</td>
<td>5</td>
<td>5.50</td>
<td>7</td>
<td>0</td>
<td>0.50</td>
</tr>
</tbody>
</table>

58
Figure 2.1. Diagram of the theoretical influence of the fire regime gradient on $d_{AF}$ (solid line) and scaling of dendroecogical sampling methodology (dashed line) needed to accommodate increasing $d_{AF}$ and spatial scale of the patch mosaic. $d_{AF}$ increases along the fire regime gradient because the spatial scale or temporal frequency, or both, of high severity fire increases along the gradient and results in higher loss rates and spatial bias of fire history information (e.g. fire scars, fire-caused growth anomalies or wood anatomical features in surviving trees). As $d_{AF}$ increases, therefore, so does the uncertainty and difficulty of reconstructing fire history. Fire history reconstructions may have to shift from primarily point fire scar records to increasing reliance on age cohorts.
Figure 2.2. (a) Map of the western U.S. showing locations of the NCDE (5 sites), GYE (2 sites), central OR (4 sites), and SJ (2 sites) study regions. Insets from (a) show the geographic distribution of study watersheds from the NCDE and GYE. Panels (c-h) show the distribution of fire scars and age structure plots for each watershed. Photogrammetrically derived patch boundaries are shown for the NCDE watersheds in panels (c-f) but not for the GYE watersheds in panels (g-h) because patch boundaries for the latter were not digitized and orthorectified.
Figure 2.3. Graphics of (a-b, e-f) the raw dendroecological data used to determine cohort initiation and recruitment events for a subset of sites that demonstrate the diversity of establishment patterns in our study regions, (c, g) occurrence of fires, pluvials and cohort pulses in relation to the Palmer Drought Severity Index (PDSI), and (d, h) results from the bivariate event analysis showing synchrony of cohort initiations and peaks with fire and pluvials over a range of temporal lags. Results are shown for the NCDE (a-d) and GYE (e-h) regions separately. Asterisks in (a-b) and (e-f) indicate the 10-year establishment bins where initiation (*) and peak (⋆) establishment occurred for each example site. Species codes in (a-b) and (e-f) are listed in Table 2.2. In (c) and (g), smoothed PDSI was calculated using a 5-year low pass filter. Vertical light grey shading indicates years classified as pluvial. Panels (d) and (h) show BEA. The upper and lower confidence intervals are depicted by the grey shaded area. Areas where \( \hat{L}_{AB}(t) \) is above (below) the confidence interval represent synchrony (asynchrony) between variables.
Figure 2.4. Graphs of some properties of the reconstructed fire regimes for the NCDE and GYE, including, (a) fire-free intervals, (b) event-level fire severity ($\Delta S(n_i)$) for all fires, (c) percent of sites in each cumulative fire severity class, (d-e) percent of sites in different age-cumulative fire severity class groupings for the NCDE (d) and GYE (e), and (f) cumulative distribution of patch ages. In panels (a-b) horizontal lines are medians, boxes represent the interquartile range (IQR) of values, whiskers represent up to 1.5 times the upper and lower IQR, and points are outliers. In panels (d-e), abbreviations represent: Y = Young (< 100 years), I = Intermediate (100-200 years), Old (> 200 years), NSR = non stand-replacing, M = mixed, H = High. Stand age was defined as the median age of the oldest cohort comprising $\geq 10\%$ of total stand BA.
Figure 2.5. Graphs showing (a) mean density decay, $\bar{d}(n)$, (b) the cumulative distribution of all pairwise distances between age plots and the nearest fire scar, $d_{AF}$, (c) expected and observed topographic position index classes of pixels containing sampled fire scars, and (d) slopes from linear regression models fit to the log-transformed $\bar{d}(n)$ values ($\lambda$) versus standardized $z$-scores of the mean TPI of pixels where fire scars were found and mean $d_{AF}$ (y-axis) for each of four regions. In panel (c), TPI class codes represent: 1 = flat slopes, 2 = valleys, 3 = lower slopes, 4 = middle slopes, 5 = upper slopes, 6 = ridgetops. Expected values were derived as the proportion of pixels in each TPI class within each study area. For ease of interpretation, the x-axis in panel (d) is plotted as the absolute value of the $d(n)$ slopes so that lower (higher) values represent lower (higher) severity. Lines in panel (d) are fitted curves from linear regressions. The linear model for TPI class was not significant ($p > 0.10$, $R^2 = 0.52$) whereas the model for $d_{AF}$ was highly significant ($p > 0.05$, $R^2 = 0.95$).
Figure 2.6. Graphics demonstrating the challenges to reconstructing historical fire severity where fire history records are patchy and spatially biased. Data shown here are from Six Mile Creek in the GYE. Graphics depict (a) the spatial distribution and distance between age plots and fire scars, $d_{AF}$, (b) fire perimeters reconstructed from convex hulls in relation to age plots and fire scars, (c) a closeup view of age plots and fire scars overlaid on a 1961 aerial photograph showing the mosaic of forests (dark grey) and grassland/shrubland (light grey), and (d-i) detailed dendroecological data for a subset of sites depicted in (a-c). In panels (d-i), vertical dashed lines represent fires documented by fire scars (red), age cohorts (orange), and fire perimeters (pink) or fires documented in neighboring patches that were not attributed to a patch due to lack of evidence (grey). Asterisks above an establishment bin indicates that it was identified as a fire-initiated age cohort pulse by our algorithm. Grey filled circles above an establishment bin indicates cohorts identified by our algorithm that were not attributed to fire, either due to lack of association with fire, species composition, or repetitiveness with already identified cohorts. Species codes in (d-f) are listed in Table 2.2.
Figure 2.7. Influence of compositing fire history data at multiple-scales on fire interval data used to calculate fire history statistics. Graphics are shown for the NCDE (a-c) and GYE (d-f) regions separately. In panels (a, d), the full composite fire interval (FI) represents the distribution of median patch-scale FIs calculated based on the composite record of fire dates from all sampled cross-sections and cohort pulses detected within a patch. The composite FI was calculated similarly, but only included fire dates evidenced by fire scars (i.e. it excluded fire dates evidenced solely by cohort pulses). The point FI distribution represents median patch-level fire intervals calculated from all intervals recorded on individual cross-sections. The max point FI was calculated similar to the point FI, but only included intervals from the cross-section(s) containing the maximum number of scars within each patch. Panels (b-c, e-f) show the difference between the composite and point (b, e) or maximum point (c, f) FIs as a function of the number of cross-sections sampled in a patch. Negative (positive) values indicate longer (shorter) point FIs compared to the composite FI. Large difference between fire interval distributions in panels (a, d) or a negative trend in the data in panels (b-c, e-f) would suggest that fire interval statistics are sensitive to compositing at the patch-scale. Anderson-Darling tests showed no statistically significant differences between FI distributions using different compositing methods (p > 0.1 for all tests). Linear models fit to the data in panels (b-c, e-f) found no statistically significant differences between composited and point FIs as a function of sample size (p > 0.1, R² < 0.1 for all tests).
Figure 2.8. Accumulation curves (a-d) and the extrapolated cohort richness as a function of patch size (e-f) displayed for cohort (a, c, e) and severity (b, d, f) classes. Accumulation curves are presented for the GYE (a-b) and the NCDE (c-d). Extrapolated cohort richness (e-f) is only presented for the NCDE, because patch boundaries for the GYE were not orthorectified and do not provide accurate area estimates. Accumulation curves show the cumulative number of age cohorts detected within a patch (within-patch cohort richness) as a function of the number of dendroecological plots sampled. Accumulation curves were not displayed for single cohort patches, including most high severity patches. Extrapolated richness is the estimated total number of cohorts in a patch calculated using the Chao estimator (Chao 1987). Curves from linear regression models were fit to the data and displayed in (e-f) to display possible trends in the data, but were not statistically significant (p >0.05, adjusted $R^2 < 0.2$).
Figure 2.9. Boxplots showing the distribution of patch-level median composite fire intervals binned by cumulative fire severity classes for the a) NCDE and b) GYE. NSR stands for “non stand-replacing” cumulative fire severity class. Black solid lines with hollow squares show the percent of patches with right-censored fire intervals within each cumulative severity class. Right-censored fire intervals occurred mostly in sites with long time since fire and are excluded from boxplots because they represent incomplete fire intervals. Where right-censored fire intervals comprise a significant portion of the data, boxplots likely underestimate the actual distribution of fire intervals.
Appendix A2: Calibration of relative dendroecological severity

Biases associated with the use of raw basal area or density for estimating fire severity

Many field and remote sensing techniques for estimating fire severity are based on detecting the degree of change caused by fire (Key & Benson 2006, DeSantis & Chuvieco 2009, van Wagendonk et al 2004, Escuin & Fernandez 2009), often expressed as a percent of pre-fire condition. This approach is effective for recent fires, when pre and post-disturbance measurements can be made. For historical reconstructions of fire severity, however, the use of tree density- or basal area-weighted age structure to estimate historical severity may suffer from multiple biases. Some of these biases may act counter-directionally to each other but the net effect is likely to result in a bias towards higher severity fire that increases with the number of fires occurring at a site. These biases are caused by effects related to both the numerator (e.g. BA or density surviving a fire) and denominator (e.g. total stand BA or density).

When only live trees are available for inclusion in the age-structure data and fire severity is estimated as the proportion of present tree density or BA that predates the fire, one of the strongest biases occurs because the denominator (total stand density or basal area) is fixed and the numerator (the density or BA of trees that predate the fire) decreases monotonically as the number of fires experienced increases going back in time (Tepley & Veblen 2015). For instance, the density or BA of trees that established before the second most recent fire in a stand and survived to the present can only be less than or equal to the density or BA of living
trees that predate the most recent fire. This will result in a tendency for higher severity estimates as the number of fires at a site increases, biasing interpretations of historical fire effects.

Another bias related to the numerator exists if BA, rather than density, is used to estimate severity when BA calculations for surviving trees are based on current and not reconstructed bole diameter at the time of each fire. If current bole diameters are used, the BA surviving a fire will be an overestimate, therefore biasing severity estimates towards lower severity. To account for this, ring widths for each core used to reconstruct tree ages would have to be measured and the diameter at the time of each fire would need to be used in calculations of BA surviving. For large sample sizes this represents an immense task. The last source of error related to the numerator is the inability to account for lost evidence resulting from fire, other disturbances and stand dynamics between a given fire date and sampling. Any post-fire BA lost due to these or other factors will be included in severity estimates based on current stand structure and will lead towards higher severity estimates.

Two primary factors influence the denominator. First, the use of current BA as the denominator presumes that this value represents a maximum site-specific BA that, when compared with the BA of trees that survived a specific fire, could provide a useful measure of fire severity. Although fire exclusion has likely moved current stand BA closer to maximum potential values, some sites may not have reached their maximum BA. A related, but much stronger bias, however, is caused
by the fact that stand BA at the time of any historical fire likely was not near its maximum BA value unless preceded by a long fire-free interval. For dry pine and many mixed-conifer forests where median fire intervals ranged from 10-60 years (Schoennagel et al. 2004, McKenzie et al. 2000, Swetnam & Baisan 1996) it is likely that most stands on the landscape would not have reached their maximum BA in the intervals between fires. If actual BA at the time of a fire was lower than the current BA, the actual percent of BA surviving a fire would be higher (i.e. fire severity would be lower) than estimates using current BA. The magnitude of this bias for any given fire is unknown and would vary as a function of the ratio between the time-since-fire and the biomass recovery rate.

As is evident from this discussion, multiple biases influence the accuracy of raw BA and density estimates of historical fire severity. These factors have motivated development of alternative severity metrics (Tepley & Veblen 2015, Marcoux et al. 2015, Heyerdahl et al. 2012). However, it is also important to recognize that many of these biases can be minimized by filtering the dataset to events meeting the following criteria: (1) the most recent fire (n=1) at a site, (2) high severity fires occurring subsequent to the most recent fire (e.g. n=2) if the most recent fire was non stand-replacing, or (3) fires from anywhere in the record that are classified as low severity according to the BA criteria. In the first case, where severity is only estimated for the most recent fire in a patch, the bias towards high severity fire resulting from an increasing amount of lost evidence with subsequent fires, one of the strongest biases, is minimized. Tree survival and establishment
patterns measured in current field campaigns should largely represent the effects of the last fire, barring any major subsequent disturbance. The reasoning for the second case is similar to that for the first. In these cases, we submit that abrupt losses in tree biomass more likely indicate severe fire rather than cumulative fire effects of two non stand-replacing fires. In the case of fires classified as low severity, even if they occur following multiple previous low severity fires, the bias towards higher severity fire that can occur with increasing numbers of fires is sufficiently weak that it does not biased fire severity interpretations. This subset of fires comprises the largest possible dataset that could be used to calculate absolute severity with minimal bias. We use this dataset as a calibration dataset for the relative severity fire event metric, $\Delta S(n_i)$, employed in this paper.

*Calibration of the fire-event severity metric, $\Delta S(n_i)$*

The cumulative $S(n_i)$, and event-level, $\Delta S(n_i)$, severity metrics developed by Tepley & Veblen (2015) are standardized by $\bar{d}(n)$ and therefore represent severity relative to all other sites used to calculate $\bar{d}(n)$. Discrete severity classes for individual fires can be derived from $\Delta S(n_i)$ by converting $\Delta S(n_i)$ values from all fires pooled across a study area to percentile scores and placing thresholds at user-set percentiles. In this study, we calculated $\bar{d}(n)$ separately for each region, so the patch-level severity metrics, $S(n_i)$ and $\Delta S(n_i)$, can only be compared to other patches within regions. In order to make direct comparisons of fire severity estimates between study regions, we developed a method to calibrate $\Delta S(n_i)$ percentile scores to an absolute fire severity scale.
The calibration method we developed involved calculating fire severity using both the relative fire severity metric, \( \Delta S(n_i) \), and an absolute estimate of fire severity based on basal area, \( BA(n_i)_{surviving} \). We calculated \( BA(n_i)_{surviving} \) as:

\[
BA(n_i)_{surviving} = 100 * \frac{\sum_{i=1}^{n_{max_i}} BA_i(n_i)}{BA_{i,total}}
\]  

[Eq. 3]

where \( BA_{i,total} \) is the present stand BA for the \( i \)th patch, \( BA_i(n_i) \) is the BA of trees that established before the \( n \)th fire and survived to the present, and \( n_{max_i} \) is the total number of fires in the \( i \)th patch that meet the three filtering criteria detailed above. We performed an iterative search for the threshold values of \( \Delta S(n_i) \) percentile scores that maximized classification agreement with \( BA(n_i)_{surviving} \). We used a basal area metric to calibrate \( \Delta S(n_i) \) rather than density because basal area measures of fire severity are more widely used in assessments of fire effects in modern field studies (Miller et al. 2009) and are more closely related to site biomass or canopy cover estimates than density (Mitchell and Popovich 1997). This analysis was only conducted for the subset of fires described in the previous section that experience minimal bias when calculating severity using raw BA. Thresholds for the low-moderate and moderate-high classes for \( BA(n_i)_{surviving} \) were fixed at 70% and 20%, respectively. These threshold values are similar to the values used in other studies (Agee 1993, Hessburg et al. 2007, Schoennagel et al. 2011) but were modified to slightly broaden the low severity class and narrow the high severity class.

For each fire in the calibration dataset, classification agreement was evaluated for the fixed \( BA(n_i)_{surviving} \) classes and \( \Delta S(n_i) \) classes thresholded across
a full factorial range of percentile rankings from 10-40% and 60-90%, for the low-moderate and moderate-high classes, respectively. Classification agreement was evaluated using the Kappa statistic, classification accuracy and Pearson’s correlation. The relative $\Delta S(n_i)$ values were then classified into absolute event-level severity classes (low, moderate, high) for each fire based on the threshold values that maximized classification agreement for all three statistics.

The calibration procedure showed a clear pattern of variation across severity thresholds that was quite consistent between the three statistical measures of agreement for most threshold combinations in both regions (Fig. A2.1). Thresholding of the low-moderate severity classes at 10% or 40% percentile values for both regions consistently produced poor classification agreement compared to the 20-30% thresholds. Overall classification agreement in the GYE declined significantly when the moderate-high severity threshold was placed above 70% (Fig. A2.1b), whereas classification agreement in the NCDE was relatively insensitive to the moderate-high severity threshold value (Fig. A2.1a). Although a number of thresholding values in each region would have produced similarly high classification accuracy, we selected the 20/80 and 20/70% threshold points for the NCDE and GYE, respectively. These thresholds were used to partition $\Delta S(n_i)$ values for all fires into severity classes.
Figure A2.1. Influence of different thresholding values on classification agreement between $BA(n_i)_{surviving}$ and $\Delta S(n_i)$ severity metrics for the (a) NCDE and (b) GYE. Classification agreement between $BA(n_i)_{surviving}$ and $\Delta S(n_i)$ metrics was estimated using the kappa statistic, accuracy, and Pearson’s correlation. Agreement statistic values on the y-axis range from 0 to 1 for accuracy and correlation and from -1 to 1 for the kappa statistic. Values on the x-axis represent a full factorial combination of threshold values from 10-40% for the low-moderate threshold and 60-90% for the moderate-high threshold.
Chapter 3

Introduction

Many lower and middle elevation forests of western North America were historically characterized by mixed-severity fire regimes (Taylor & Skinner 1998, Arno et al. 2000, Schoennagel et al. 2004, Amoroso et al. 2011, Heyerdahl et al. 2011, Perry et al. 2011, Sherriff et al. 2014). Definitions of MSFRs vary in their details (Agee 1993, Arno et al. 2000, Schoennagel et al. 2004, Agee 2005, Sugihara et al. 2006), but there is general agreement that they caused variable levels of tree mortality (here referred to as fire severity) resulting from a mix of crown and surface fire behavior that manifest over a range of spatial and temporal scales. This process resulted in a landscape mosaic of even-aged and uneven-aged patches, whose relative proportions, spatial pattern and scale were driven by a combination of top-down (e.g. weather, ignition patterns, climatic variation) and bottom-up (e.g. vegetation characteristics, topography) influences (Agee 1998). Robust spatio-temporal records are needed to quantify the complex mosaic of MSFRs and to understand the landscape dynamics that created them, but few historical data sources provide both the temporal depth and spatial attributes needed to make these robust inferences. For instance, dendroecological records (e.g. fire history, age structure) have contributed significantly to our current understanding of fire severity-mediated stand dynamics and the geographic distribution of MSFRs over relatively small areas ($\leq 10^3$ ha) (Taylor and Skinner 1998, Heyerdahl et al. 2012, Heyerdahl et al. 2014, Tepley and Veblen 2015). But a variety of practical and methodological barriers significantly limit their uses for reconstructing the spatial
pattern and scale of historical fire severity over areas of sufficient size to capture
the variability that defined MSFR forests historically (Hessburg et al. 2007). As a
result, there is a notable lack of understanding and debate (Williams and Baker
2014, Odion et al. 2016, Stevens et al. 2016) about the spatial ecology, landscape
dynamics and resilience mechanisms of the diverse set of forests across western
North America that historically experienced MSFRs.

Few alternative data sources to dendroecological records are available for the
study of historical fire severity and dynamics, but inferences have been made from
forest structural conditions derived from historical aerial photographs (Minnich et
al. 2000, Hessburg et al. 2007) and land surveys (Baker 2012, Williams and Baker
Stephens et al. 2015). Compared to dendroecological data, these structure-based
reconstructions can span a greater range of scales, from individual plots or patches
to large landscapes, for specific points in time, and may therefore more effectively
capture the spatial variation, scale and pattern of fire severity within a landscape.
Structural interpretations of historical stand dynamics and fire severity are based
on an assumed relationship between tree size and age that is used to interpret the
distribution of tree biomass (e.g. canopy cover, basal area or density) and species
composition among tree size classes or canopy strata (e.g. overstory vs. understory).
So whereas dendroecological reconstructions are based on direct evidence of age
structure in relation to fire events, only indirect interpretations of fire severity and
stand dynamics can be made from forest structural conditions. Although size-age relationships can be relatively strong at the scale of individual stands, especially for shade intolerant, and some shade tolerant, tree species that colonize following disturbance (Lorimer and Krug 1983, Lorimer 1985, Antos and Parish 2002, Abella 2008) they become weaker with increasing sample scale as sites with different biophysical characteristics are incorporated (Lorimer 1980, Veblen 1992). This reliance on tree age-size relationships creates ambiguity for structure-based interpretations of fire severity that is compounded by the myriad influences on forest structure and composition resulting from biophysical gradients, endogenous processes (e.g. stand development, competition, senescence) and exogenous disturbances other than fire (e.g. insect outbreaks, climatic events, windstorms) (Oliver 1980, Spies 1998).

Due to this ambiguity, many studies make qualitative inferences about historical fire severity and stand dynamics (Hagmann et al. 2013, Hagmann et al. 2014, Collins et al. 2015, Stephens et al. 2015) without quantitatively or mechanistically linking forest structure, fire severity and stand dynamics. Only a few studies have explicitly reconstructed fire severity from forest structures (Hessburg et al. 2007, Baker 2012, Williams and Baker 2012, Baker 2014) and despite some general similarities between them, differences in their analytical approaches demonstrate two critical points that have not been sufficiently addressed by previous research. First, ecologically distinct target variables can be derived from structural conditions, but no previous research has directly tested
which aspects of disturbance history are best represented by structural conditions. For instance, Williams & Baker (2012) reconstructed cumulative, or time-integrated, fire severity classes that accounted for the influence of multiple fire events over time. In contrast, Hessburg et al. (2007) postulated that forest conditions best reflect the effects of the most recent fire and proposed that a space-for-time substitution could effectively capture temporal variability of fire severity effects in a system if reconstructed over sufficiently large areas. Second, rigorous validation procedures have generally not been used to identify the multivariate structural characteristics and variable thresholds that maximize classification performance for distinct target variables. Hessburg et al. (2007) used a dichotomized hierarchical logic structure and user-set canopy cover and tree size thresholds to assign patch severity classes, but did not use an independent dataset to validate this method. Williams & Baker (2012) used mean values documented in published dendroecological studies to place stand density and tree diameter thresholds for their classification scheme. As such, their study represents one of the only calibrated structure-based fire severity reconstructions, but even in this case no sensitivity analysis was used to evaluate whether their multivariate criteria produced optimal severity classifications.

In MSFRs, similar stand structures may result from multiple pathways (Tepley et al. 2013), so rigorous validation of structure-based fire severity inferences is essential for a clear understanding of the spatio-temporal dynamics in these systems. Multiple pathways in MSFRs result from the complex interplay between
fire frequency and severity that is less apparent in other fire regime models (Bekker and Taylor 2001, Wimberly and Kennedy 2008, Halofsky et al. 2011). For instance, old forest structures in low- and high-severity regimes can be clearly related to two contrasting processes: recurrent low severity fires (Abella et al. 2007) in the former and long time-since-fire in the latter (Spies 2004). In MSFRs, these alternative scenarios may both exist, potentially in stands within relative proximity, or old forest structures may result from distinct processes such as intermediate fire frequency and severity. The presence or abundance of large trees may not clearly discriminate between these alternative scenarios although other factors, or combinations of factors (e.g. species composition, density, canopy cover, number of canopy strata), might. Similarly, patches dominated by intermediate-sized trees could represent even-aged cohorts recovering after intermediate time-since-high-severity-fire, multiple moderate severity fires that result in a high cumulative mortality rate and dominance of intermediate-sized trees, or old, multi-aged stands influenced by repeated low severity fires on low productivity sites that limit tree size. While discrimination among multiple pathways is feasible in most cases using dendroecological methods and may also be feasible using stand structure it is not clear which combinations of structural variables, other ancillary variables (e.g. topography), or variable thresholds will maximize discrimination among these alternative scenarios and yield severity classifications that are closely linked to the disturbance dynamics that generated them.
In this paper, we address the need for rigorous evaluation of structure-based reconstructions of historical fire severity and fire-mediated landscape dynamics in MSFRs. We use a unique multi-proxy dataset from the northern U.S. Rockies consisting of dendroecological data, a digital elevation model, photogrammetrically-derived patch boundaries, and photo-interpreted (PI) patch-level forest structural attributes derived from historical aerial photographs to develop a calibrated structure-based model of patch-scale fire severity that is closely linked to stand dynamics. This research builds on previous work which demonstrated that dendroecological data can effectively characterize the internal heterogeneity of age structure and fire history within patches, allowing the scaling of dendroecological interpretations of fire severity-mediated dynamics from sample plots to the patch-level (Naficy 2016 Ch. 2 Dissertation). Thus, our multi-proxy dataset permits direct comparisons between patch-scale fire severity history, fire-mediated stand dynamics and the structural attributes of patches. We capitalize on this dataset to address the following objectives: (1) model three response variables (severity of the most recent fire, cumulative fire severity history, and age structure) that are strongly related to fire severity and stand dynamics as a function of patch-level structural variables and topographic features, (2) evaluate multiple machine learning model frameworks and the rule-based classification method previously used by Hessburg et al. (2007) to determine the optimal structure-based modeling method, and (3) identify the top predictor variables and describe their relationships to the response variable of the top performing model. Our analysis represents the
first attempt to combine dendroecology and detailed photogrammetric analysis of historical aerial photos to produce dendroecologically calibrated structure-based models of patch-level fire severity.

We suspected that inherent differences in the temporal resolution of dendroecological records and forest structural attributes might reduce model performance if not controlled for appropriately. Therefore, as a fourth and final goal we empirically evaluate four hypothesized temporal disparities between the two methods and we examine the temporal thresholds over which each of these hypotheses influence model results (detailed in the following section).

Reconciling limitations to dendro-calibrated, structure-based fire severity models

In MSFRs with a significant component of high severity fire, the elapsed time-since-high-severity-fire (TSHSF) will be a key driver of successional and structural properties that strongly influences structural interpretations of fire severity. High severity fire consumes evidence of previous fire effects and creates even-aged stands of small trees that are easily detected in both aerial photographs and age structure. These stand structures may last for decades or more, but will change over time due to autogenic stand developmental processes and subsequent non stand-replacing disturbances (Agee 1993, Oliver and Larson 1996). After sufficient time, high-severity effects may no longer be apparent in stand structure although they are clearly preserved in age structure (Arno et al. 1995). For example, a stand initiated by high severity fire over 400 years ago that has experienced multiple subsequent non stand-replacing fires would likely be dominated by mixed-
sized trees including many large old trees, similar to a 400 year old stand shaped exclusively by recurrent non stand-replacing fires. In this scenario, no differences in the fire history of these stands would be detected based on PI forest structures and both would be classed as low-severity patches whereas dendroecological records would class the former as mixed-severity and the latter as low-severity. At the other end of the TSHSF gradient, non stand-replacing disturbances that occur shortly after high-severity fire may not be detectable in PI structural attributes due to the difficulty of distinguishing between closely-aged cohorts of predominantly small trees.

Although age structure may preserve the high severity signal over a broader temporal range than PI vegetation structure, there may be cases in which evidence of high severity fire is lost in dendroecological records but preserved in forest structural attributes. For example, in sites where multiple non stand-replacing fires occur following high severity fire, the high-amplitude pulse of establishment that is apparent in the age structure of most stands initiated by high severity fire may be partially lost. In these cases, it is difficult to determine whether age truncation is due to a high severity fire or cumulative mortality resulting from repeated fires. Vegetation structure in these sites, however, would likely still reflect the previous effects of high severity fire, as evidenced by the predominance of smaller trees and a lack of large trees that generally indicate a long TSHSF.

To test whether these factors could be assessed empirically, determine the temporal constraints of these hypotheses, and evaluate if their incorporation into
our modeling approach would improve model performance we evaluated four hypotheses relating TSHSF to model performance metrics (Table 3.1). Hypothesis 1 (H1) states that young stands influenced only by non stand-replacing fires (NSR) have experienced different stand dynamics and fire severity history than old NSR stands and may be more appropriately classed as mixed-severity sites. The truncated age distributions in young NSR stands either results from a moderately old high severity fire (e.g. 100-250 years old) whose effects have been masked by subsequent non stand-replacing fires or these sites have experienced significantly greater cumulative tree mortality in non stand-replacing disturbances than those which have an old tree component. In either scenario both age and forest structural attributes indicate higher severity effects than is generally thought to occur in stands with a history of NSR fires. Therefore, we hypothesized that reclassifying these sites as mixed-severity, although no high severity fire was determined quantitatively, was reasonable and would improve model performance. Hypothesis 2 (H2) states that non stand-replacing effects in mixed-severity sites with short TSHSF are unlikely to be detected using PI vegetation attributes due to the difficulty in distinguishing between two young cohorts. Reclassifying these mixed-severity sites as high severity recognizes this lower TSHSF constraint on PI structure and the dominant influence of recent high-severity fire at these sites. Hypothesis 3 (H3) concerns sites that were initiated by high-severity fire centuries ago and have experienced recurrent non stand-replacing fires since. As described above, the old high severity fire in these sites likely falls outside the upper TSHSF
boundary within which PI structure can detect high severity fire effects, so these sites would likely be classified as NSR severity by structure-based models. This upper TSHSF limit can be accounted for by reclassifying mixed-severity sites with long TSHSF as NSR. Hypothesis 4 (H4) recognizes that young and old high severity sites are characterized by unique stand structures and composition that have developed with different TSHSF. Under this hypothesis, young and old high severity sites should be classed separately.

Methods

Study Area

The study area is the Northern Continental Divide Ecosystem (NCDE), an approximately 2 million ha area in western Montana centered around Glacier National Park and the Bob Marshall Wilderness (Fig. 3.1). Low to middle elevation forests in the NCDE are dominated by mixed-conifer forests of ponderosa pine (*Pinus ponderosa*), western larch (*Larix occidentalis*), Douglas-fir (*Pseudotsuga menziesii*), lodgepole pine (*Pinus contorta*), with occasional Rocky Mountain Juniper (*Juniperus scopulorum*), grand fir (*Abies grandis*) or patches of trembling aspen (*Populus tremuloides*). Mesic sites may also have components of Engelmann spruce (*Picea engelmannii*), subalpine fir (*Abies lasiocarpa*), western white pine (*Pinus monticola*), western red cedar (*Thuja plicata*), and paper birch (*Betula papyrifera*) (Habeck 1987). Xeric grasslands and shrublands are found primarily in large valley bottoms, dry aspects, or in recently burned areas. Understory vegetation may consist of grasses and forbs in drier sites but often includes a significant shrub
component (Pfister et al. 1977, Fischer and Bradley 1987). Elevations range from 738 to 3,188 m (average = 1,766 m). Annual average temperature and precipitation are approximately 6°C and 109 cm, respectively (Daly et al. 2002), with most precipitation occurring as winter snow or spring rain. Topography is varied, with broad river valleys framed by steep, mountainous terrain. Lower to middle elevations in the region are dominated by mixed-severity fire regimes that grade into high-severity regimes at higher elevations (Freedman & Habeck 1985, Arno et al. 1995, Barrett et al. 1991, Naficy 2016 Ch. 2 Dissertation).

Study Design

We collected dendroecological samples and mapped forest structures using historical aerial photographs from 1934-1955 for five study subwatersheds (area$_{\text{mean}}$ = 8,345 ha, area$_{\text{min}}$ = 5,558 ha, area$_{\text{max}}$ = 12,216 ha) situated along a latitudinal gradient (Fig. 3.1). Study areas were selected to include a range of forest types, vegetation characteristics, climatological settings, topographic features and fire regime characteristics (Table 3.2). We excluded the highest elevation forest types (e.g. whitebark pine, subalpine larch, and high elevation spruce-fir) from sampling, focusing instead on the dry mixed-conifer, mesic mixed-conifer and lodgepole pine forests that comprised the lower and middle elevations.

Prior to dendroecological sampling, patches of distinct forest structure and composition were delineated and structural attributes were interpreted from aerial photographs. To identify a subset of patches in each watershed for dendroecological sampling, we employed a spatially balanced, stratified sampling scheme that
ensured a representative sample across topographic gradients and vegetation types. Clusters of 2-3 patches were randomly placed across each watershed in proportion to the representation of two broad aspect classes (defined along the SW-NE axis), two elevation bands (defined as the upper and lower 50% of DEM pixel values), and seven forest structural types (Hessburg et al. 1999) derived from the PI data. Where large gaps in the coverage of sample plots occurred within a watershed, random locations for new sample clusters were generated and this procedure was repeated until a spatially balanced sample network was achieved. To minimize the confounding influence of management on age structure any sites with evidence of previous timber harvest were excluded.

**Photogrammetric mapping of patch boundaries and vegetation structure**

For each of the 5 study watersheds, we acquired full coverage series of high-resolution (1:15840 - 1:26,000 scale), panchromatic stereo aerial photographs from the National Archives or local Forest Service offices. All photo-interpretation was performed by expert photogrammetrists following methods developed by Hessburg et al. (1999). To delineate patch boundaries, stereo photo pairs were visually interpreted using a stereoscope in an iterative process. Initial patch boundaries were defined along zones with clear changes in vegetation composition or structure. As patch attributes were photo-interpreted, boundaries were revisited and redrawn if any attribute varied by one class level, down to a minimum patch size of 4 ha. Photo-interpreted attributes used to delineate patches and describe vegetation characteristics of each patch include (Table A3.1): total canopy cover, canopy cover
of over and understory layers, diameter size class of over and understory canopy layers, the number of canopy layers, species composition of the over and understory layers, the abundance of snags, and a series of textural variables (e.g. crown differentiation, canopy clumpiness). Patch boundaries were digitized, orthorectified and joined with the tabular database of photo-interpreted attributes in a vector layer that was used in all further analysis.

From the raw attributes, we also derived several multivariate variables. First, we calculated the canopy cover of trees in each diameter class for each patch, regardless of over or understory status. To account for the possibility of bias in attribute values between watersheds mapped by different photo-interpreters, especially for tree size class which we hypothesized to be a critical predictor variable, we created a canopy cover variable with relative tree size class rather than the absolute size classes in the raw dataset. Relative tree size classes (denoted $\text{Size}_{\text{std}}$) for each patch were standardized by dividing the over and understory size class by the median size class of overstory canopy trees within each watershed. Canopy cover values for all relative size classes were then pooled into three categories: SMALL CC ($\text{Size}_{\text{std}} < \text{Size}_{\text{median}}$), MED CC ($\text{Size}_{\text{std}} = \text{Size}_{\text{median}}$), or LARGE CC ($\text{Size}_{\text{std}} > \text{Size}_{\text{median}}$). We reasoned that if absolute size classes were better predictors than relative size classes then tree size likely reflected actual fire effects whereas if the reverse was true, some photo-interpreter bias likely existed in the dataset and relative size classes would better reflect fire effects. Finally, eight multivariate structural classes were derived following methods in Hessburg et al.
(1999) from the raw canopy cover and size class attributes to describe patches with distinct forest structural properties.

Field sampling and laboratory methods

To characterize the age structure of each sample patch, we placed a randomly oriented, variable-length transect within each patch interior (> 300m from patch boundaries). Along each transect, we placed fixed area (0.04 ha) dendroecological sample plots at intervals of 30-100 m (average spacing = 60 m), depending on patch size, topography, dimensions and age structure complexity. The transect length, and number of plots sampled within each patch, increased with the apparent complexity of the fire history of patch (range=1-22 plots, median=3 plots). Within each plot, we recorded the species and measured the diameter of all live and dead trees ≥ 4 cm diameter at breast height (DBH). To balance the competing goals of intensive age structure sampling within a patch and maximal number of patches sampled within each watershed, we employed two different plot sampling procedures. In the first, increment cores were sampled from a representative subset of trees in each plot but in a greater number of plots per patch, while in the second procedure all trees in each plot were sampled, resulting in a more intensive sample within fewer plots. See (Naficy 2016 Ch. 2 Dissertation) for detailed sampling methods.

Following extensive field surveys within each sample patch and adjacent patches, partial or full cross-sections were sampled from up to 5-10 trees with the best records, primarily from dead trees. To increase the spatial distribution of
sampled cross-sections, document buried fire scars, and ensure the most complete fire history records, we opportunistically sampled additional cross-sections from non-sample patches and areas with previous timber harvest.

All increment cores and fire scars were surfaced and sanded with successively finer sandpaper (to a maximum 600 grit) until the wood cellular structure was clearly visible. All increment cores and fire scars were visually crossdated or, if visual crossdating was not possible, statistically crossdated with COFECHA (Holmes 1983) using species-specific master chronologies developed from within our study area (Naficy 2016 Ch. 2 Dissertation). For cores that did not contain the pith, the number of missing rings was estimated using methods presented in Duncan (1989). Corrections for rings missed as a function of coring height were made using the mean number of years required for 10 samples of each species in each watershed to reach 25 cm DBH, the mean coring height.

Reconstruction of patch-level fire history and severity

The fire history of each patch was documented using a combination of fire scars and age cohorts, primarily, with growth anomalies, buried fire scar tips in increment cores and wood anatomical features used as secondary evidence in a small number of cases. Tree establishment dates were binned into 10-year intervals and discrete cohorts were identified as continuous tree establishment pulses of at least 50 trees/ha within a 30-year moving window that were preceded by at least one decadal bin with no establishment. Previous research showed strong statistical dependency between cohorts and fire for temporal lags of up to 23 years (Naficy
so we considered a cohort to be fire-initiated if it occurred within \( \leq 20 \) years following a fire date documented within a patch, an adjacent patch or a reconstructed fire perimeter based on surrounding fire scars and cohorts. Fire-initiated cohorts were attributed to and assigned the date of the fire that most closely preceded cohort initiation. For consistency between our dendroecological data and the PI attributes observed in the historical aerial photos and to minimize the influence of fire exclusion, which became effective in our study area in the 1930s, we excluded all trees from this analysis that established after 1955, the date of the most recent aerial photo. The fire history of each patch was created by compositing all fire dates recorded on cross-sections or by fire-initiated cohorts. Previous analysis determined that this method produced the most accurate and unbiased patch-scale fire histories and age structures Naficy 2016 Ch. 2 Dissertation).

To estimate the severity of each fire and the cumulative, time-integrated, fire severity history of each patch, we used a relative severity metric developed by Tepley & Veblen (2015). This metric allows quantitative, cumulative- and event-level fire severity estimates that are not biased towards higher severity fire as the number of fires at a site increases. It is based on a monotonically decreasing function, \( d_i(n_i) \), that describes the rate of stand density loss (\( d \)) with increasing numbers of fires (\( n \)) for the \( i^{th} \) patch. Based on the average density of all patches in the study area after \( n \) fires, \( \bar{d}(n) \), a relative estimate of cumulative severity, \( S(n_i) \), and event-level severity, \( \Delta S(n_i) \), can be calculated as:
\[
S(n_i) = \frac{d_i(n_i)}{\bar{d}(n)} \quad [\text{Eq. 1}]
\]
\[
\Delta S(n_i) = -1 \times \frac{S(n_i) - S(n_i-1)}{S(n_i-1)} \quad [\text{Eq. 2}]
\]

\(\Delta S(n_i)\) values mostly range between -1 and 1, with positive (negative) values representing higher (lower) fire severity relative to other fires pooled across all sites in the study domain. We truncated fire severity calculations for fires that occurred when patches had < 15 surviving trees/ha due to lack of sufficient age structure data to characterize fire severity with confidence. To create discrete severity classes each fire closely approximate absolute fire effects, we pooled \(\Delta S(n_i)\) values of all fires across the entire study area, converted them to percentile scores, and thresholded these values at the 30\(^{th}\) and 80\(^{th}\) percentiles to produce three event-level fire severity classes (low, moderate, and high). These threshold values were derived from an iterative calibration procedure (Naficy 2016 Ch. 2 Dissertation) that searched for the \(\Delta S(n_i)\) percentile score thresholds that maximized agreement with basal area (BA) estimates of fire severity classes defined as low when \(\geq 70\%\) BA survived a fire, moderate when 21-69% BA survived, or high when \(\leq 20\%\) BA survived. These threshold values are similar to the values used in other studies (Agee 1993, Hessburg et al. 2007, Schoennagel et al. 2011) but were modified to slightly broaden the low severity class and narrow the high severity class. Based on the severity of all fires affecting a patch, we defined the cumulative fire severity history as: (1) low severity if all fires were of low severity, (2) non stand-replacing (NSR) if fires were a mix of low and moderate severity, (3) mixed severity if at least one high and one low or moderate severity fire occurred within a patch or if all fires
were moderate severity, or (4) high severity if all fires within a patch were high severity.

**Characterization of patch-level age structure**

To determine whether structure-based models could predict key features of detailed age structure that are closely linked with fire severity history, we defined six age structure classes based on stand age and the number of age cohorts (Table 3.3). These classes are similar to widely-used stand structural classes (O’Hara et al. 1996, Oliver and Larson 1996) that we modified to more closely reflect past fire effects by excluding structural components that are unrelated to fire, e.g. shade tolerant cohorts, cohorts initiated by non fire-disturbances. Stand age was defined as the median age of the oldest cohort comprising > 10% of total stand basal area. Based on stand age, we grouped sites into one of three developmental stages: stand initiation (stand age < 50 years), young forest (stand age 50-199 years) and old forests (stand age ≥ 200 years) (see Appendix B for derivation of threshold values). The number of cohorts for each stand was defined as the number of fire-initiated cohorts comprised of shade intolerant species with greater than 10% of total stand BA. The minimum BA criterion was used to exclude cohorts that represent a minor portion of age structure and would be difficult to detect in PI forest structure.

**Derivation of patch-level topographic variables and climatic setting**

We characterized the topographic setting of each patch using a suite of variables (Table A3.1) derived from a 10m digital elevation model (Gesch et al. 2002). The median elevation, mean slope and mean transformed aspect were
calculated from all pixels in each patch. Aspect was cosine-transformed into a northness index and a heat load index was calculated from the mean aspect, slope and latitude of each patch using equation 3 of McCune & Keon (2002). To characterize the relative landform position for each patch, we calculated the topographic position index (TPI) for each pixel in the DEM (Weiss 2001). The TPI describes the position of any pixel relative to its neighboring pixels, with higher TPI values indicating higher slope positions. The TPI is scale-sensitive, so we performed calculations at two scales, 300 and 1500 pixels, to represent fine- and broad-scale topographic settings. Based on the mean and standard deviations of the raw TPI values (Weiss 2001), we partitioned TPI into 6 broad classes describing their relative landform positions, including: flat slopes, valley bottom, lower slopes, middle slopes, upper slopes and ridgetops. The mean TPI value and the median TPI class were assigned to each patch. All raster processing was carried out using Python v2.7.

*Construction and validation of structure-based models*

Our goal was to use the combined dataset of dendroecological reconstructions, topographic variables, and patch-level PI forest structures to: determine which severity-related response variable is best predicted by structure, develop a modeling approach that optimizes structural interpretations of historical fire severity and dynamics, and describe relationships, including non linearities, between PI structure, topographic variables and fire severity for the top performing model. To determine which aspects of disturbance history can best be predicted from PI forest
structure, we evaluated models for three separate patch-level response variables: cumulative fire severity class, severity of the most recent fire, and age structural class. To accommodate the high dimensionality of forest structure-based interpretations of fire severity, we employed nonparametric machine learning techniques to build and tune models of PI vegetation attributes for each response variable (Table 3.3), using the dendroecological data as a validation dataset. A variety of machine learning methods have been shown to skillfully capture high-dimensional relationships even with relatively small sample sizes (Hastie et al. 2001, Olden et al. 2008, Kuhn and Johnson 2013), making them an ideal modeling tool for this study. To evaluate the influence of different learning methods on model performance for our dataset, we employed two supervised machine learning classification tree algorithms. First, we used Random Forests (Breiman 2001) and a modified implementation of Random Forests that uses a conditional inference framework to reduce variable selection bias (Hothorn et al. 2006, Strobl et al. 2008). Random Forests (RF) is an ensemble machine learning technique that minimizes overfitting and instability associated with individual classification trees by constructing models based on large numbers of trees derived from subsamples of the data (De'ath and Fabricius 2000, Breiman 2001, Prasad et al. 2006, Cutler et al. 2007). The RF procedure also provides a measure of variable importance based on the difference in model accuracy using original and permuted variables. This method for assessing variable importance can be biased, however, in cases where correlated predictors are used or where the type (e.g. categorical vs. continuous
variables) and range of predictor values vary greatly (Strobl et al. 2007, Strobl et al. 2008). In these cases, unbiased variable importance measures can be derived using a conditional framework (Strobl et al. 2008). Although most of the PI vegetation predictors used in this study are categorical variables with a relatively low range of factor levels (Table A3.1), some continuous variables with a broad range of values (e.g. many topographic variables) were incorporated in models, thereby necessitating evaluation of both conditional and traditional RF procedures.

Second, we implemented an evolutionary algorithm which differs substantially from RF in the mechanics used for model optimization. Whereas RF searches recursively for locally optimized solutions that maximize the homogeneity of groups resulting from each split, evolutionary algorithms search for globally optimal solutions (Olden et al. 2008). For high-dimensional modeling problems, such as structure-based models of historical fire severity, we reasoned that globally-optimized solutions might outperform the RF method. Evolutionary algorithms apply principles of natural selection to achieve optimization by using a fitness function that evaluates populations of solutions (i.e. subsets of predictor variables) which are mutated or recombined stochastically to produce variation through subsequent generations (Eiben and Schoenauer 2002). We used an evolutionary algorithm implemented by the evtree package in R (Grubinger et al. 2014) to evaluate whether a globally optimized solution might improve performance over RF models.
For each model, we used an iterative process to evaluate model performance, identify important predictors and interpret relationships between predictor and response variables. Prior to initial model runs, we assessed correlations between predictor variables and removed any variables with correlations higher than $r = 0.6$. We reasoned that vegetation attributes more closely reflected the direct influence of fire whereas topographic factors were more likely to be indirect indicators of fire severity. Accordingly, vegetation attributes were preferentially retained where vegetation and topographic variables were highly correlated. We used the Caret package in R (Kuhn 2008) to optimize model parameters and evaluate model performance. Each RF model was constructed from a relatively high number of individual trees ($\text{ntree} = 1501$) and a standard number of variables for each node ($\text{mtry} = \sqrt{\# \text{ of predictors}}$). For the evolutionary tree, we used an intermediate complexity parameter value ($\alpha = 0.5$), which was found to be near optimal in model tuning. For each model run we utilized 70% of the dendroecological data as a training data set and withheld the remaining 30% for model validation. We used repeated k-fold cross-validation during model tuning to determine the optimal parameter settings and the full training and test datasets to construct the model and estimate its performance. Cross-validation is a data-efficient method for improving estimates of model performance over the built-in out-of-bag (OOB) error estimate provided by RF, which are based on data used in model training and are therefore not strictly independent. To minimize both bias and variability in performance estimates, we used repeated 5-fold cross-validation with 10 repeats.
For model tuning, we designated the kappa statistic (Cohen 1960) as the optimization parameter because it performs well for datasets with unbalanced class design (Kuhn 2008).

Initial model runs exhibited some instability, likely resulting from the use of k-fold cross-validation with a small sample dataset (Kuhn and Johnson 2013) combined with the high-dimensional relationships that exists between vegetation attributes, topography, and severity. To improve estimation of model performance, we ran all models 500 times using unique data partitions (i.e. unique seeds) to select training and test datasets for each run. The median value of the kappa statistic and prediction accuracy for each model was calculated from all runs and used as the basis for identifying the best modeling approach.

**Accounting for temporal disparities between dendroecological and structural records**

To test our reclassification hypotheses (H1-H4) and to determine which TSHSF thresholds improved model performance, we applied each hypothesis and ran conditional forest models of fire severity over a range of TSHSF intervals (0-300 years, at 30 year intervals). For H1, which addressed NSR severity sites where no evidence of high severity fire exists, we used stand age as a close proxy for TSHSF. At each TSHSF interval we modified the severity class of each site that met the reclassification criteria for each hypothesis (Table 3.5). For example, for TSHSF = 60 years under H2 all mixed severity sites with TSHSF > 60 years would be reclassed as high severity prior to model training and evaluation. Following reclassification, model training and evaluation was performed just as it was for the
original model except that the number of model iterations at each TSHSF interval was reduced from 500 to 50 to reduce the high computational costs of this procedure. This procedure allowed us to track model performance over a range of TSHSF values for each hypothesis to determine if each hypothesis improved performance relative to the null and, if so, over which temporal lags.

To evaluate the importance of each individual hypothesis on model performance and determine the optimal TSHSF threshold, we initially analyzed each hypothesis in a one-by-one fashion (data not shown). After this initial run, all analyses were re-run, this time in an all-but-one fashion using the best performing thresholds found in the one-by-one analysis for each hypothesis. The all-but-one analysis was found to improve model stability and overall performance between runs, likely due to improvements made by incorporating some of the hypotheses, so only results from the all-but-one analysis are presented in the results. Using this approach we tested which hypothesis improved overall model performance, how tradeoffs between specificity and sensitivity varied with TSHSF and which TSHSF thresholds for each hypothesis maximized model performance.

In addition to reporting the median kappa statistic and accuracy from all runs for each TSHSF interval, we calculated the class-specific specificity (true negative rate), sensitivity (true positive rate), and multi-class area under the receiver operating curve (AUC) (Hand and Till 2001). Receiver operating curves (ROC) depict changes in the true positive and false positive classification rates and are useful for evaluating tradeoffs between specificity and sensitivity as well as
assessing classification accuracy. The AUC is a robust measure of classification accuracy used in binary classification problems (Fawcett 2006), with values near 1 indicating a perfect predictor and values of 0.5 (the minimum possible value) indicating a random or uninformative predictor. The AUC can be expanded to multi-class problems by calculating the mean AUC from ROC curves built for all pairwise comparisons (Landgrebe and Duin 2007), thereby providing an overall assessment of model performance that complements the model kappa and accuracy.

For hypotheses that improved model performance, the optimal TSHSF thresholds were applied to the original data to produce two reclassified response variables, Sev Reclass 4 and Sev Reclass 3. These variables were identical except that all old high severity class patches (produced by application of H4) were excluded from Sev Reclass 4 to produce Sev Reclass 3. This step was deemed necessary because the old high class comprised a small sample size (n=9) that may not be sufficient for effective model construction. The modeling process used to evaluate the original data was then re-run using the new reclassed severity response variables with vegetation predictor variables alone, with topographic and vegetation variables included, or topographic variables alone.

**Variable reduction & selection of the final fire severity model**

To produce the most parsimonious model with optimal performance and interpretability, we ran the top performing machine learning method through a feature reduction routine. In addition to maximizing parsimony, we expected that this process would increase mean model performance by reducing variation between
model iterations as the number of predictor variables decreased and individual
trees share a more similar structure. This procedure was only performed for fire
severity models, since this was our primary response variable of interest. We used a
two-step variable reduction process (Guyon and Elisseeff 2003, Díaz-Uriarte and
Alvarez de Andrés 2006, Genuer et al. 2010), wherein poor predictor variables used
in the initial model runs were first culled. The remaining predictor variables were
then recursively discarded in step two through a univariate backward elimination
procedure that tracked model performance as the number of variables was reduced.
To cull variables that consistently showed low variable importance in the initial 500
model runs (step 1), we calculated the median standardized variable importance
values from the upper 50th percentile of the best performing models. We used
natural breakpoints in the median variable importance scores to conservatively
eliminate only variables with consistently low importance values. This set of
retained variables was useful for interpretation of relationships, but required
further refinement to produce a parsimonious final model.

The stepwise backward elimination procedure (step 2) that we implemented
to further reduce variable numbers and select a final model began with the retained
set of variables from step one. For each iteration of the selection procedure, 50
model runs were performed using unique data partitions for training and test data
sets but the same set of predictor variables. Importance values for each variable
were converted to ranks and the median variable ranking and Kappa statistic were
calculated from all 50 runs for each iteration. The variable with the lowest median
ranking was removed from the set of predictors and this process was repeated until only the top three variables remained. The final model was selected using the 1-SE rule (Breiman et al. 1984, De'ath and Fabricius 2000, Genuer et al. 2010), which states that the optimal, most parsimonious model is the one with the fewest variables whose error lies within one standard error of the minimum.

**Rule-based structural fire severity models**

For comparison with the three machine learning modeling approaches described above, we applied the ruleset developed by Hessburg et al. (2007), which was designed to classify the severity of the most recent fire affecting a patch (Table A3.2). As with the machine learning models, we evaluated the rule-based model performance using confusion matrices with the dendroecological records as a validation dataset. To test whether this rule-based classification scheme was a more robust predictor of the severity of the most recent fire, as it was intended, or cumulative fire severity effects, we also evaluated its performance using the original and reclassified cumulative severity classes as response variables.

**Multivariate structural properties and relationships with fire severity history**

To depict the multivariate structural properties of patches with distinct fire severity history, we used a combination of cluster analysis and non-metric multidimensional scaling (NMDS) to identify distinct PI forest structural groups and determine their relationship to fire severity history. Prior to analysis, PI variables with high correlations (>0.6) were removed and remaining variables were standardized by subtracting the observed minimum and dividing by the range of
observed values, a method which has been found to result in optimal clustering (Milligan and Cooper 1988). To define unique structural groups, hierarchical agglomerative cluster analysis was implemented on the PI structural variables using the hclust function from the Vegan package in R (Oksanen et al. 2016) using Ward’s linkage method (Murtagh and Legendre 2014). None of the identified structural groups were closely associated with a fire severity class (see Results), so we describe the range of structural conditions associated with different fire severity histories by calculating the proportion of patches within each severity class assigned to each of the identified structural groups. We used NMDS based on Sørensen distances to visualize gradients in PI forest structural properties that defined patches assigned to (a) structural groups resulting from the cluster analysis, (b) the original severity classes (ML$_{Sev$ Original$}$), and (c) the reclassified severity classes (ML$_{Sev Reclass 4}$). To aid with interpretation, we calculated and plotted the mean (median in the case of CRWN DIFF) NMDS axis scores and standard deviations for each grouping scheme. Multiresponse permutation procedures (MRPPs) (Mielke and Berry 2007) were used to test for statistical differences in the multivariate dissimilarity matrix of classes within each grouping scheme.

We also used partial dependence plots from the final severity model (only shown for ML$_{Sev Reclass 4 Final}$), to illustrate relationships between the top predictors and the conditional probability of each severity class. Partial dependence plots show the contingency of the prediction probability of a specific response variable level (i.e., severity class) on the range of values of a specific predictor (Hastie et al. 2001,
Cutler et al. 2007). They are therefore useful for evaluating the nature of the relationship (i.e., linear vs. nonlinear) and threshold behaviors which exist between response variable classes and specific predictors.

**Results**

*Photogrammetric analysis & dendroecological fire severity reconstructions*

Across all five study watersheds (41,723 ha in total), we mapped patch boundaries and interpreted vegetation attributes on 738 patches, ranging in size from 2-674 ha (mean = 61.15 ha, median = 33.99 ha) (Table 3.2). All of our study watersheds were historically characterized by pole- to small-sized trees (~13-40 cm), with few patches dominated by large trees (Table 3.2). Dendroecological samples were collected from 255 plots in 70 patches comprising a large cumulative proportion of the total area of mixed-conifer forest in each watershed (mean=52%, min=37%, max=72%), excluding Emery Creek where only two patches were sampled (Fig. 3.1, Table 3.2). In total, 2,579 increment cores and 188 cross-sections were successfully crossdated, although 42 of these cross-sections (mostly from lodgepole pine, subalpine fir or spruce) were determined to be due to agents other than fire (i.e. insects or mechanical damage). Establishment dates covered the period 1462-1949. The 146 crossdated fire-scar samples contained 292 fire scars (average=1.9, maximum=5 scars/cross-section) from 40 unique fire years over the period 1471-1955 (oldest fire=1635, most recent fire =1932). Fire return intervals in patches were highly variable between and within all watersheds (FI$_{\text{median}}$=36-120 years, FI$_{\text{min}}$=8-64 years, FI$_{\text{max}}$=108-255 years, Table 3.2). Severity could not be calculated.
for 112 of the 292 crossdated fire scars in our fire history record due to the lack of age structure information. Of the 180 fire scars for which severity could be calculated, applying the optimal severity thresholds from the calibration procedure resulted in 63 (35%), 62 (34%), and 55 (31%) fires of low, moderate and high severity, respectively. Based on the cumulative fire severity history of each patch, 0 (0%), 12 (17%), 32 (46%), and 25 (36%) of patches were classified as low, NSR, mixed and high severity, respectively. Because no sites were classified as low severity, this class was not included in subsequent modeling.

Consistent with studies in low severity fire regimes, many NSR sites in our study area (Fig. B3.2a-d) had recurrent non stand-replacing fires and multiple cohorts with at least one older cohort (i.e. pre-1750). However, many of these patches had two or more moderate severity fires that caused them to be classified as NSR rather than low severity. A small number of NSR sites exhibited different fire histories and stand dynamics than those traditionally ascribed to a low-severity fire effects (Fig. B3.2a). These stands were characterized by an old dominant cohort but had fairly low fire frequencies, large fire-free gaps, and a small number of subdominant cohorts (generally only one or two) caused by infrequent non stand-replacing fires. Mixed-severity patches generally exhibited intermediate ages and had two or more cohorts, but there was high variability of stand ages, numbers of cohorts and fire frequency between sites within this class (Fig. B3.2e-j). High severity patches were mostly younger even-aged patches caused by the most recent
fire (Fig. B3.2k-l), although some patches resulting from older high severity fire also were documented (Fig. B3.3g-h).

**Structure-based models of fire severity**

For the cumulative fire severity models, all three machine learning algorithms exhibited moderate predictive power in the initial model runs, although RF performed consistently better than CF or EV (Table 3.3). The reclassification procedure improved the cumulative fire severity models and showed support for a number of hypotheses (Fig. 3.2). H1 resulted in the least improvement of any hypothesis, with model performance varying marginally over TSHSF, with the highest statistics for TSHSF values >180 years (Fig. 3.2b). The best balance between specificity and sensitivity was achieved with the initial model or TSHSF values of 150-180 years (Fig. 3.2c-d). Based on these results, we determined that initial models or models with a TSHSF threshold of < 200 years would be optimal. For H2, model performance also varied marginally, with slight peaks around TSHSF values of 90 and 180 years (Fig. 3.2f). Class-specific error rates were similar to or slightly improved, relative to initial models, for TSHSF of 60 and 60-120 years for mixed- and high-severity classes, respectively, and diverged strongly after that (Fig. 3.2f-h). For H2, we concluded that a TSHSF threshold of 60 years might improve model performance. Model performance for H3 showed minimal improvement over all TSHSF thresholds, with the notable exception of improved kappa values at the 240 threshold (Fig. 3.2j) that was also clearly reflected in improved class-specific error rates for mixed-severity patches (Fig. 3.2k) and only
minimal change for NSR patches (Fig. 3.2l). For H4, model performance and class-specific error rates indicated optimal model performance at around 60 years TSHSF (Fig. 3.2n-p). In general, the degree of change in model performance resulting from the reclassification procedure appeared to be related to the number of sites affected by each hypothesis, with weaker support for H1-H2 (< 5% of sites affected) than H3-H4 (9-11% of sites affected). Figure B3 depicts the age structure and fire history of some example sites affected by each hypothesis.

Applying all hypotheses with the identified thresholds changed the severity class for 27% of sites (Table 3.1) and resulted in significantly improved model performance over initial PI models (Table 3.3, ML$_{Sev}$ Reclass 4 Kappa$_{RF}$ = 0.42, ML$_{Sev}$ Reclass 3 Kappa$_{RF}$ = 0.47, ML$_{Sev}$ Original Kappa$_{RF}$ = 0.35) or PI models with the weaker performing hypotheses removed (H1$_{removed}$ Kappa$_{RF}$ = 0.38, H2$_{removed}$ Kappa$_{RF}$ = 0.38). Downsizing the dendroecological training dataset to produce a balanced sample size across all severity classes reduced model performance significantly (data not shown) and was not implemented further. Despite general similarities in the structure of different machine learning models (Fig. C1), the RF models consistently outperformed the CF and EV algorithms, with EV models generally performing second best (Table 3.3). Models based solely on topographic variables performed poorly (Table 3.3, Kappa$_{RF}$ ML$_{Sev}$ Reclass 3 Topo Only = 0.11), but inclusion of topographic variables in PI severity models resulted in similar or slightly improved performance (Table 3.3, ML$_{Sev}$ Reclass 4 topo Kappa$_{RF}$ = 0.40, and ML$_{Sev}$ Reclass 3 topo Kappa$_{RF}$ = 0.49).
Cumulative fire severity models outperformed models using other response variables (Table 3.3, ML\text{Sev Reclass 3} Kappa\text{RF} = 0.47 with ML\text{Sev Most Recent} Kappa\text{RF} = 0.29, ML\text{Age Structure} Kappa\text{RF} = 0.32), suggesting that PI vegetation structure is a better predictor of cumulative fire severity history than either the severity of the most recent fire alone or detailed age structure. Confusion matrices for ML\text{Sev Most Recent} and ML\text{Age Structure} show that class-specific accuracy was generally highest for high severity and stand initiation structures, which are closely associated with high severity, and was notably lower for most other classes (Table C3.2-3.3). Compared to all three machine learning models, the rule-based severity model performed poorly at predicting the severity class of the most recent fire (Table 3.3, Rule\text{Sev Most Recent} Kappa = 0.12) or cumulative fire severity class (Rule\text{Sev Original} Kappa = 0.09), although use of the reclassified severity response variable improved this method marginally (Rule\text{Sev Reclass 3} Kappa = 0.15). Similar to ML\text{Sev Most Recent}, the Rule\text{Sev Most Recent} model had the highest prediction accuracies for high severity fire (Table C3.4), although accuracy was lower than the machine learning models for all classes.

**Variable reduction & selection of the final fire severity model**

Three vegetation and three topographic variables were removed from the pool of candidate predictors due to high correlation (Table A3.1). Importance values from the initial model runs were consistently low for four variables (DEAD SNAG, SIZE US, CLUMP SIZE, LARGE CC), so these variables were also culled prior to running the final model selection procedure, leaving 16 predictor variables in total. Results from the variable reduction procedure (Fig. 3.3b, d) show slight variation of model
performance around an average of 0.42 and 0.48 Kappa ($\text{Kappa}_{\text{max}} = 0.44$ and 0.51) for ML\text{Sev Reclass 4 final} and ML\text{Sev Reclass 3 final}, respectively, until fewer than 10-12 predictor variables were retained, after which point performance declines notably. All models with 12 or more predictor variables for ML\text{Sev Reclass 3 final} and 10 predictors for ML\text{Sev Reclass 4 final} performed similar to or better than the initial models (Table 3.3) and were within 0.027 Kappa of the 1-SE rule threshold (Fig. 3.3b, d), although only models with 12-13 predictors for ML\text{Sev Reclass 3 final} and >13 for ML\text{Sev Reclass 4 final} were strictly within one standard error of the optimal model. Based on this, we selected the model with the top 12 predictors for ML\text{Sev Reclass 3 final} and the top 10 predictors for ML\text{Sev Reclass 4 final} as the final models. The relative ranking of predictors by importance scores were identical for the top five predictor variables for the ML\text{Sev Reclass 4 final} and ML\text{Sev Reclass 3 final} models and similar for the top 10-12 predictors retained in both final models (Fig. 3.3a, c). Both topographic and PI structure predictors ranked among the top five predictors. All topographic variables except the median TPI$_{300}$ Class were retained in the final models.

Class-specific accuracy for the final model (ML\text{Sev Reclass 4 final}) was lowest for old high-severity classes (Table C3.5), likely due to the small sample size of this severity class in our dataset and the similar abundance of medium-sized trees compared to NSR or some mixed-severity sites (see following section). Mixed-severity sites exhibited intermediate accuracy, likely attributable to the broad range of forest structures and fire severity effects in mixed-severity sites that can resemble NSR or high-severity sites at either extreme.
Multivariate structural characteristics and fire severity history

Hierarchical cluster analysis of the PI structural variables produced four groups (Fig. 3.4a) with distinct multivariate structural properties (MRPP test, p < 0.01), which we named: stand initiation, open canopy multi-story, closed canopy small trees and closed canopy medium trees (Table 3.5). All severity classes were comprised of a mix of two or three forest structural groups, suggesting that broadly-defined structural conditions may arise from multiple fire-mediated pathways. NSR patches were predominantly characterized by open canopy, multi-story forest structures, with a smaller, but significant, proportion of sites in closed canopy medium tree structural groups (Table 3.5). Mixed-severity patches were also comprised of a large percentage of open canopy multi-story structural groups, but with a greater representation of closed canopy forest structures than NSR sites. Young high-severity patches were either characterized by stand initiation structures or a mix of open canopy multistory and closed canopy small trees, likely reflecting the diversity of remnant structures, post-fire regeneration and times-since-high-severity fire contained within this class. Old high-severity patches were characterized by a broad mix of forest structures, other than stand initiation, that likely reflect variation in the underlying productivity gradients and autogenic successional processes that influence stand structure with TSHSF.

The multivariate space resulting from the NMDS analysis was defined by overstory canopy cover (OS CC r = -0.61), tree size (SIZE OS r = 0.89) and crown differentiation (CRWN DIFF r = 0.76) for Axis 1 and the canopy cover of intermediate-sized trees (SIZE 2 CC r = 0.86, SIZE 3 CC r = 0.80) for Axis 2 (Table
3.5). Structural groups showed good separation in this multivariate space (Fig. 3.4a). However, for the cumulative severity classes the within-group dissimilarity (δ) was substantially greater and the chance-corrected within-group agreement was smaller (data not shown) compared to the structural groups, resulting in poor group separation between the original cumulative severity classes (Fig. 3.4b). The reclassification procedure improved group separation, providing further support for H1-H4, although substantial structural variation and marginal separation of the reclassified severity classes persists (Fig. 3.4c, Table 3.5).

**Relationships between top predictor variables and cumulative fire severity**

Partial dependence plots show a mix of linear and nonlinear relationships between cumulative fire severity class (shown for the MLSevReclass 4 final model only) and the top vegetation (Figs. 5) and topographic (Figs. 6) predictor variables. The probability of NSR classification was higher in patches with low to intermediate overstory canopy cover (OS CC = 10-50%) of small to medium overstory trees (SIZE OS = 3-4) and a variable cover of the understory stratum (SIZE 1 CC = 0-40%, SIZE 2 CC = 0-70%) that resulted in high crown differentiation values (CRWN DIFF = 3). The probability of the mixed-severity class was highest for patches with pole-sized overstory trees (SIZE OS = 2) and moderate crown differentiation (CRWN DIFF = 2). Overstory canopy closure (OS CC) and the cover of intermediate tree sizes (SIZE 2 CC, SIZE 3 CC) were also important predictors for the mixed-severity class, but they exhibited broad variation that likely reflects the diverse forest structural groups that comprised these patches. The probability of young high-severity sites
was highest where crown differentiation (CRWN DIFF) and the cover of small to medium trees (SIZE 3 CC, SIZE 4 CC) was low and canopy cover of saplings was moderate to high (SIZE 1 CC = 40-80%). Variation of the probability associated with overstory tree size (SIZE OS) and the bimodal peaks in overstory canopy cover (OS CC=10% and 70%) probably result from variation in the presence of residual surviving small and medium overstory trees at some young high severity sites. The probability of old high severity sites was highest in patches with intermediate to large overstory tree size (SIZE OS = 2-4), high canopy cover (OS CC > 70%) and intermediate crown differentiation values (CRWN DIFF = 2).

Partial dependence plots for topographic variables (Fig. 3.6) showed the highest probability of NSR severity on flat or moderate slopes (MEAN SLOPE < 30°), low elevation (MEAN ELEVATION < 1,400m) areas with higher heat load (HEAT LOAD > 0.9), and lower slope positions (MEAN TPI < 0) suggesting a strong association with large valley bottoms or areas of moderate terrain (Fig. 3.6). The high partial dependence values observed at higher elevations (MEAN ELEVATION > 1,600m) may result from the occurrence of some NSR patches near the upper elevation limits of mixed-conifer forest characterized by generally long fire free intervals punctuated by 1-2 NSR fires. Mixed-severity showed contrasting patterns, with higher partial dependence on intermediate values of all topographic predictors. Partial dependencies for young high-severity sites were highest on steep slopes (MEAN SLOPE > 30°), intermediate elevations (MEAN ELEVATION = 1,300-1,800m) and high or low slope positions (30 > MEAN TPI < 0). Old high-
severity sites exhibited the highest partial dependencies for low-moderate elevations (MEAN ELEVATION < 1,600m), sites with flat terrain to upper slope positions (MEAN TPI_{300} > 0) and steep terrain (MEAN SLOPE > 30°).

**Discussion**

We present a novel combination of dendroecological methods and photogrammetric analysis of historical aerial photographs to address some long-standing methodological challenges to reconstructing broad-scale, spatio-temporal patterns of historical fire severity. This research constitutes the first rigorous, data-driven evaluation of the utility, performance and limitations of PI forest structure-based models of historical fire severity and stand dynamics. Our analysis reveals nonlinear relationships between PI forest conditions and fire severity history (Figs. 5, 4b-c, Table 3.5), but we also provide a framework for incorporating dendroecological records into structure-based models that strengthens the link between forest structural conditions and fire effects. We show that although structure-based models were not effective at predicting detailed age structure (e.g. ML_{Age Structure}) or some aspects of disturbance history (e.g. ML_{Sev Most Recent}), if calibrated, they can successfully reconstruct spatio-temporal patterns of cumulative fire severity. Our modeling approach builds upon previous work (Naficy 2016 Ch. 2 Dissertation) that addressed challenges stemming from the spatial scale-disparity of sample units that comprise dendroecological records (e.g. plots < 1 ha) and photogrammetric maps (e.g. patches ≥ 4 ha) and developed a calibration procedure that maximized the accuracy of thresholds used to create dendroecological fire
severity classes. Here, we complement these analyses by identifying and accounting for temporal disparities between dendroecological and structural data types in our models. Together, the combined dataset and the modeling framework we have developed allow cross-scale reconstruction of spatio-temporal fire severity dynamics, from individual dendroecological plots to forest patches and, ultimately, to landscapes.

Combining both data types is labor intensive, requiring detailed photogrammetric analysis and collection of extensive dendroecological networks of point fire scar records and age structure plots across a large number of patches with differing disturbance histories, biophysical settings and stand conditions. It is important to note that the PI patch boundaries and the large number of structural variables used as predictors in our models requires detailed photogrammetric techniques applied systematically across patches. These methods eclipse traditional uses of aerial photographs in most historical landscape reconstructions which have generally used them to delineate high severity patches (Johnson and Larsen 1991, Kipfmueller and Baker 2000) or distinguish between severity classes using a small number of structural characteristics and a priori rules (Taylor and Skinner 1998, Bekker and Taylor 2001). A notable benefit of using these more intensive photogrammetric techniques is that, once calibrated, our model is highly scalable and can be applied to predict cumulative fire severity across large landscapes, where patch boundaries and PI forest structures have been mapped, without further dendroecological sampling. For example, if applied to just the mixed-conifer
forest within the five subwatersheds studied here (Table 3.2), cumulative fire severity could be modeled for 738 patches over 27,000 ha, making it one of the larger reconstructions to date and the only validated structural reconstruction of its size to use empirical, not modeled, spatially-explicit patch boundaries. As such, our reconstruction approach has unique potential to address critical questions about the spatial ecology and geographic variability of fire severity-mediated dynamics in historical landscapes.

**Limitations to dendroecologically-calibrated, structure-based models of fire severity**

The structure-based fire severity models we develop here perform fairly well (ML\textsubscript{Sev Reclass 4 Final} Kappa\textsubscript{RF} = 0.41, Accuracy = 0.61; ML\textsubscript{Sev Reclass 3 Final} Kappa\textsubscript{RF} = 0.49, Accuracy = 0.71), but exhibit some instability that requires a computationally expensive model iteration procedure and complicates model evaluation and prediction. We suspect that the observed variability in model performance resulted from three principal factors. First, relationships between PI forest structures and fire severity history are fundamentally complex (Fig. 3.4b-c), so structure-based interpretations are likely to assume a greater degree of error than dendroecological methods. Second, although we used a large dendroecological dataset and robust analytical techniques to derive our dendroecological severity estimates, our validation dataset (n=70 patches) constitutes a relatively small sample size for the k-fold cross-validated machine learning models (Kuhn and Johnson 2013) that we used. An increased sample size, especially of NSR and old high severity sites would balance our dataset, reduce some of the current variability, and likely increase
overall model performance. Third, there are limitations to dendroecological and photogrammetric methods that can influence model performance, but are difficult to assess quantitatively due to their historical nature. We used dendroecological records to validate structural interpretations, but fire severity reconstructions in MSFRs are inherently challenging (Naficy 2016 Ch. 2 Dissertation) and any error in these reconstructions is propagated through the machine learning process to ultimately influence model performance and predictions. Similarly, the quality of patch boundaries and structural predictor variables depends on multiple factors, including the experience-level of photogrammetrists and aerial photo image quality relative to the detail of attributes being interpreted (Avery and Berlin 1992). We used expert photogrammetrists and a systematic photointerpretation method to ensure record quality, but there are inherent limitations to aerial photo interpretations (e.g. limited visibility of understory layers in dense forests) and some level of interpreter error that is not easily quantified for historical studies. Although the influence of some of these errors on model performance may not be directly quantifiable, our approach makes an important contribution to the use of structure-based fire severity models by presenting a transparent and rigorous analysis of many of the quantifiable sources of error that have not been previously addressed.

**Considerations for structure-based models of historical fire severity**

In our study system, we identified TSHSF as a critical variable that caused divergence between the dendroecological and structure-based severity metrics when
TSHSF was long (H3-H4, Fig. B3.2e-h) or when non stand-replacing disturbance occurred following short TSHSF (H2, Fig. B3.2c-d). TSHSF is likely not the only factor that influences the accuracy of structure-base methods, but because high-severity fire leaves an enduring imprint on forest conditions, we suspect that TSHSF will be a key influence on structure-based severity estimates in many other ecosystems as well. TSHSF is likely to be of particular relevance for MSFR forests that are influenced by interacting non lethal and lethal fires, but the factors that limit structure-based interpretations of fire severity and their specific bounds are likely to be highly specific to particular regions and should be assessed locally.

The model performance improvements resulting from the reclassification procedure (Table 3.3) demonstrate the importance of identifying the temporal limits to structure-based disturbance history reconstructions but they also highlight the tradeoffs between temporal resolution and spatial robustness. The reclassification procedure we implemented improved model performance by incorporating the temporal constraints on structure-based fire severity reconstructions, but at the expense of some accuracy in interpreting fire dynamics. For instance, the limits we document for structure-based models at both low (H2) and high (H3) TSHSF suggest that our models would not be effective at detecting some types of extreme, i.e. short- or long-interval, high-severity fire events which can be important triggers of critical transitions or state-shifts in landscape dynamics (Donato et al. 2009, Enright et al. 2015, Fitch and Meyer 2016). Because our goal was to build a model that could be used to quantify the spatial pattern and variability of fire severity
over large landscapes and because only a relatively small portion of patches in our study area were subject to these constraints, we considered this an acceptable tradeoff. However, we stress that tradeoffs inherent to structure-based reconstructions should be explicitly addressed and carefully considered in the context of specific research applications.

Both machine learning and rule-based methods were better predictors of cumulative severity class than severity of the most recent fire (Table 3.3), suggesting that forest structural conditions strongly reflect a time-integrated disturbance legacy spanning multiple centuries. This is a direct violation of a critical assumption underlying the rule-based severity classification system we used here – that forest structures most strongly reflect the influence of the most recent fire – which explains the poor predictive power of the rule-based models (Table 3.3). Because the rule-based system attributes all tree mortality to the most recent fire it discriminated true high severity fire impacts relatively well, but it exaggerated the degree of fire-caused tree mortality for patches where the most recent fire was of low-moderate severity (Table C3.4). However, our model results successfully identified key structural features associated with cumulative fire severity history (Figs. 5-6) and could be used as the basis for classification rules in areas where no dendroecological data exist. Although the relative importance of specific variables and placement of variable thresholds that best separate severity classes may vary between ecosystems, the relationships we document here provide the first calibrated
description of the PI vegetation attributes associated with different disturbance histories.

Although there were some qualitative similarities between the three machine learning algorithms we evaluated (Fig. C3.1), we document significant differences in their performance that contrasts with our expectations. We expected that the unbiased variable importance values generated by the CF method or the globally-optimized EV algorithm would outperform RF, but the RF models were consistently ranked first (Table 3.3). Although we do not speculate about the causes or generality of this finding, it suggests that multiple modeling approaches should be explored for new applications, especially for cases of high-dimensional problems with relatively sparse data.

Comparison with other fire severity reconstruction methods

Few other methods present comparable potential for reconstructing spatio-temporal patterns of historical fire severity and its biogeographic variation across large regions. Geospatial modeling based on point dendroecological data has been used to reconstruct fire perimeters and fire severity patterns over scales of hundreds to thousands of hectares (O’Connor et al. 2014, Yocom-Kent et al. 2015). The spatial extent and quality of interpolated surfaces is generally restricted by the intensive labor required for each dendroecological plot that in turn limits the scale or sample intensity of dendroecological grids used in modeling. Relative to geostatistical modeling, the method we present here has two important advantages: (1) it is based on empirical, not modeled, patch boundaries that more effectively
describe patch shapes and the location of abrupt boundaries and (2) it is scalable beyond the area where dendroecological data are available once models have been calibrated.

A small number of other methods have reconstructed fire severity over larger landscapes (≥ 10⁶ ha) than those presented here based on dendroecological data (Sherriff et al. 2014) or forest structural conditions (Hessburg et al. 2007, Williams and Baker 2012). Compared to these efforts the methods we present here are unique for their integrated use of multiple data sources (but see Williams & Baker 2012), a flexible modeling framework that can incorporate ancillary data sources, and the fine-scale, empirical, and spatially-explicit patch boundaries used to define the landscape (but see Hessburg et al. 2007). For instance, multiple studies (Sherriff and Veblen 2007, Sherriff et al. 2014) have successfully modeled the spatial distribution of low- versus mixed-severity fire regimes in ponderosa pine forests of Colorado using a suite of topographic variables (primarily elevation). However, our analysis showed poor performance of topographic models based on a similar set of predictors (Table 3.3, MLSev Reclass 3 Topo). These contrasting results can be explained by clear geographic differences between our study areas. The Colorado Front Range is defined by a unidirectional, east-west elevational gradient that entrains moisture availability, vegetation characteristics and fire dynamics and contributes to the success of topographic fire regime models. Topographic variables were also important in our final models (Table 3.3, Fig. 3.3a, c) but the poor performance of models based on topography alone (MLSev Reclass 3 Topo) suggests that
in areas with greater topographic diversity, such as the NCDE, other factors must be accounted for. In the NCDE, direct measurement of forest vegetation condition was necessary to capture the patchy nature of fire effects. Because fire severity patterns are generally shaped by multiple factors, we suspect that our modeling approach, which is based primarily on vegetation characteristics, has greater transferability across diverse study areas.

Historical vegetation surveys have increasingly been used to reconstruct early 20th century forest conditions and make inferences about historical fire severity over intermediate to large areas \((10^3-10^6 \text{ ha})\) (Vankat 2011, Baker 2012, Williams and Baker 2012, Hagmann et al. 2013, Baker 2014, Dolanc et al. 2014, Hagmann et al. 2014, Collins et al. 2015, Stephens et al. 2015). Our analysis corroborates that fire severity can be successfully reconstructed from forest structures, but only if rigorous analysis is undertaken to calibrate models and validate interpretations. Although our structural variables are not directly comparable with historical surveys, there are qualitative similarities between variables measured by each method, our model predictions, and criteria commonly used to infer fire severity from historical survey data. For instance, the probability of NSR severity in our models (Fig. 3.5) was highest for patches with low-intermediate overstory canopy cover \((\text{OS CC} = 10-60\%)\), larger overstory trees \((\text{SIZE OS} \geq 3)\), and a sparse understory of small trees \((\text{SIZE 1 CC, SIZE 2 CC} < 40\%)\), conditions which are strikingly similar to those used to depict historical stands influenced by a low severity fire regime (Baker 2014, Williams & Baker 2012,
Hagmann et al. 2013, Stephens et al. 2015). Similar congruency is evident between our model predictions for other severity classes and criteria used in other studies (Fig. 3.5). However, we stress that the apparent consistency of a classification procedure with a logical set of descriptors does not ensure robust predictive power of the resulting classification system. Many vegetation and topographic predictors in our models exhibited nonlinear relationships and threshold behaviors (Figs. 5-6) that would be difficult to detect without rigorous analysis using an independent data source, as we have done here. Therefore, numerically similar class thresholds, which may be qualitatively equivalent, can result in divergent model predictions and performance. These thresholds are likely to vary significantly between ecosystems or studies with unique designs and should be assessed locally.

Notable discrepancies have emerged in the literature between some reconstructions of historical forest structural conditions, inferred fire severity dynamics, and their implications for forest management and restoration (Baker 2012, Hagmann et al. 2013, Baker 2014, Collins et al. 2015, Stephens et al. 2015). The methodological sources of these discrepancies are unclear but a number of possibilities have been posited, including: variable record quality (Frayer and Furnival 1999), biased plot locations or sample designs that may over represent older trees or forests (Frayer and Furnival 1999, Bouldin 2009), and exaggerated density values resulting from aggregation of sparse tree data in General Land Survey records to larger scales (Hagmann et al. 2013, Hagmann et al. 2014, Collins et al. 2015, Stephens et al. 2015, but see Williams & Baker 2011). This highlights
another important advantage of the methods used here—we mapped forest structures using full coverage aerial photographs of our study areas, so there is no bias in our reconstructions due to spatial sampling design and no need to model forest structural variables or spatial patterns from plot data. Direct comparisons of forest conditions and inferred fire dynamics between the methods presented here and historical survey data would be a fruitful area of future research.

Conclusions
Methodological limitations to historical fire regime reconstructions impose strong tradeoffs between capturing variation over broad spatial extents, quantifying spatial pattern, and revealing temporal dynamics. We address a critical need in fire ecology by presenting a novel cross-scale method that improves integration of these goals and facilitates more robust analysis of the spatial ecology of fire severity in historical landscapes. Our multi-proxy approach to modeling historical fire severity is labor-intensive and reveals complex challenges presented by structure-based model of fire severity. However, it also presents multiple benefits including scalability, a flexible modeling framework, empirical documentation of spatial patterns, and a data-driven validation procedure that uses dendroecological records to identify model limitations and quantify performance. These features make it a unique and useful tool for reconstructing historical fire severity patterns, quantifying their biogeographical variation and evaluating their biophysical drivers.
Acknowledgements

Support for this project was provided by a National Science Foundation Partnerships for International Research and Education (PIRE) grant (award no. 0966472), an NSF Doctoral Dissertation Research Improvement grant (award no. 1302233), the Jerry O’Neal National Park Service Student Fellowship, and a Dissertation Completion Grant from the Graduate School at CU Boulder. We thank the CU Undergraduate Research Opportunities Program for generous support of the many undergraduate students involved with this project. Special thanks are due to those who contributed to the intensive field and laboratory work involved in this research, including: S. Diaz, A. Simler, P. Wickey, J. Scharff, K. Anderson, and B. Heitshusen. We also thank E. Heyerdahl (Fire Lab) and E. Sutherland (RMRS) for the generous use of their tree ring facilities, D. Divoky (Glacier National Park), T. Carolin (Glacier National Park) and T. Wehunt (Flathead National Forest) for logistical support, P. Ohlson (Okanogan-Wenatchee NF) for photogrammetric training, and A. Tepley for discussions regarding dendroecological estimation of fire severity.
Table 3.1. Descriptions of the four hypotheses proposed to explain disagreement between dendroecological and PI structure-based severity classification and to detect the limits to detection of the target variable by PI structure. For each hypothesis, the justification, proposed reclassification action, and the hypothesized relationship between the target variable and model performance is provided. The optimal reclassification criteria identified in our analysis and used to reclassify cumulative fire severity class is listed for each hypothesis along with the number of reclassified sites. The main reclass variable for H2-4 is time since high severity fire (TSHSF). For H1, which reclasses sites within the NSR severity class that, by definition, have not experienced high severity fire, stand age was used in place of TSHSF as a measure of minimum time since stand initiation. The number of sites affected by each reclassification procedure is listed as a fraction (and percent) of the total number of sites in our dataset (n=70).

<table>
<thead>
<tr>
<th>Explanation of hypotheses</th>
<th>Target variable(s)</th>
<th>Reclass criteria</th>
<th># Sites affected</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H1</strong>: Stand structures and dynamics differ between young and old non lethal severity sites</td>
<td>1. Stand age</td>
<td>Stand age (&gt; 200) years</td>
<td>3/70 (4%)</td>
</tr>
<tr>
<td><strong>Justification</strong>: Old and young non lethal severity sites are characterized by different stand dynamics and structures. Old non lethal sites have not experienced high severity fire in the known record (usually (&gt; 300-400) years) and cumulative fire severity has been low enough to permit significant numbers of older trees and multiple age cohorts to survive within a stand. These stand structures differ substantially from young non lethal sites that have 1) either experienced a high severity fire in the past 150-250 years which is not detectable in the stand age structure due to lost evidence caused by subsequent non stand-replacing fires or 2) they have experienced significantly higher cumulative tree mortality on average than old non lethal sites. Young non lethal stands generally have multiple cohorts and mixed age (size) classes, but they are dominated by a narrower range of tree ages (sizes) and specifically lack the large, old trees that dominate old non lethal sites.</td>
<td>True: non lethal → mixed</td>
<td>False: non lethal → non lethal</td>
<td></td>
</tr>
<tr>
<td><strong>Proposed action</strong>: Reclass young non lethal severity sites as mixed severity to better reflect the younger stand structures in these sites.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hypothesized influence of the reclass variable on model performance</strong>: Initial model performance changes little over the youngest stand age thresholds since few very young non lethal severity sites exist. Performance increases slightly over intermediate stand age thresholds, as these stands are reclassified to mixed severity. Model performance decreases below initial models at higher stand age thresholds as older non lethal severity sites are incorrectly reclassified as mixed.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H2</strong>: Non stand-replacing fire effects in mixed severity sites are not detectable with short time since high severity fire</td>
<td>1. TSHSF</td>
<td>TSHSF (&lt; 60) years</td>
<td>2/70 (3%)</td>
</tr>
<tr>
<td>Mixed severity sites with a non stand-replacing fire following a relatively recent high severity fire would be dominated by small trees and might not be distinguishable from a young stand initiating after a single recent high severity fire.</td>
<td>True: mixed → high</td>
<td>False: mixed → mixed</td>
<td></td>
</tr>
</tbody>
</table>
Proposed action: Reclassify mixed severity sites with short time since high severity fire as high severity.

Hypothesized influence of the reclass variable on model performance: Relative to the initial model, performance will increase over short TSHSF thresholds as these sites are reclassified to high severity. Subsequently, model performance will decrease below initial model performance as young and old mixed severity sites are misclassified as high severity.

H3: Detection of high severity fire effects in mixed severity sites decreases with time since high severity fire

- Old mixed severity sites initiated by an older high severity fire (i.e. > 200 years) followed by one or multiple non stand-replacing fires would likely be dominated by large, old trees and, mixed age classes similar to non lethal severity sites. Older high severity fires may therefore not be detectable in stand structure.

Proposed action: Mixed severity sites with an old high severity fire and a significant large/old tree component should be reclassified as non lethal severity.

Hypothesized influence of the reclass variable on model performance: With low to intermediate TSHSF thresholds many mixed severity sites will be incorrectly classified as non lethal and models will underperform relative to initial models. At higher TSHSF thresholds (i.e. > 200 years), model performance will increase as only the oldest mixed severity sites with a significant component of large, old trees are reclassified to non lethal.

H4: Stand structures differ between young and old high severity sites

- Old and young high severity sites will have contrasting stand structure as a function of time since high severity fire. Tree size, canopy cover, vertical canopy diversity (e.g. number of canopy strata) and species composition will diverge with time since fire.

Proposed action: Old and young high severity sites should be classed separately.

Hypothesized influence of the reclass variable on model performance: Model performance will be poor at very short TSHSF thresholds, as most sites, including recent post-high severity fire are classed as old high severity sites. At intermediate TSHSF thresholds, model performance will increase as recent post-high severity sites are classed separately from intermediate to older post-high severity stands. At high TSHSF, model performance will decrease to the initial model performance as increasingly fewer stands are classified as old high severity.

In non stand-replacing sites, there is no evidence of high severity fire, so we use stand age as a close proxy for time since high severity fire. Stand age is defined as the median age of the oldest cohort comprising > 10% of total stand basal area.
Table 3.2. Summary information including aerial imagery dates, watershed characteristics, topographic features, summary information for dendroecologically sampled patches in each of the five study watersheds.

<table>
<thead>
<tr>
<th></th>
<th>Swan Valley (SV)</th>
<th>Spotted Bear (SB)</th>
<th>Emery Creek (EC)</th>
<th>Middle Fork (MF)</th>
<th>Big Creek (BC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watershed characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hydrologic subbasin</td>
<td>Swan</td>
<td>South Fork Flathead</td>
<td>South Fork Flathead</td>
<td>Middle Fork Flathead</td>
<td>North Fork Flathead</td>
</tr>
<tr>
<td>area (ha)</td>
<td>12,216</td>
<td>6,956</td>
<td>5,558</td>
<td>10,511</td>
<td>6,482</td>
</tr>
<tr>
<td># of patches</td>
<td>236</td>
<td>122</td>
<td>127</td>
<td>134</td>
<td>119</td>
</tr>
<tr>
<td>Patch size (ha)</td>
<td>51.76 (2.21-650.34)</td>
<td>57.02 (5.69-403.16)</td>
<td>43.75 (5.61-673.82)</td>
<td>78.43 (6.74-430.19)</td>
<td>54.47 (3.77-242.66)</td>
</tr>
<tr>
<td>mixed-conifer forest area (ha)</td>
<td>10,038 (82%)</td>
<td>5,495 (79%)</td>
<td>4,464 (80%)</td>
<td>5,186 (49%)</td>
<td>1,865 (29%)</td>
</tr>
<tr>
<td># of mixed-conifer forest patches</td>
<td>167 (71%)</td>
<td>85 (70%)</td>
<td>87 (69%)</td>
<td>64 (48%)</td>
<td>40 (34%)</td>
</tr>
<tr>
<td>Topographic characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>slope (degrees)</td>
<td>11 (0-78)</td>
<td>18 (0-63)</td>
<td>22 (0-63)</td>
<td>26 (0-71)</td>
<td>21 (0-60)</td>
</tr>
<tr>
<td>elevation (m)</td>
<td>1,460 (1,118-2,842)</td>
<td>1,482 (1,118-2,235)</td>
<td>1,490 (1,086-2,331)</td>
<td>1,597 (1,082-2445)</td>
<td>1,726 (1,206-2,187)</td>
</tr>
<tr>
<td>Dendroecological sample summary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mixed-conifer area sampled (ha)</td>
<td>2,434 (57%)</td>
<td>3,945 (72%)</td>
<td>169 (4%)</td>
<td>2,145 (42%)</td>
<td>693 (37%)</td>
</tr>
<tr>
<td># of mixed-conifer patches sampled</td>
<td>15 (25%)</td>
<td>31 (37%)</td>
<td>2 (2%)</td>
<td>16 (25%)</td>
<td>6 (15%)</td>
</tr>
<tr>
<td>Dendroecological fire severity class (# patches)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non stand-replacing</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>mixed</td>
<td>9</td>
<td>18</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>high</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

$ - calculations based only on mixed-conifer cover types within each watershed.
† - percent calculated based on total watershed area.
‡ - percent calculated based only on the area of mixed-conifer forest on public lands that was available for sampling within each watershed.
Table 3.3. Classification rules for each of six age structure classes. Age classes included stand initiation (< 50 years old), young forest (50-200 years) or old forest > 200 years old. See Appendix B for explanation of age threshold placement. The number of cohorts was defined as the number of cohorts comprised of shade intolerant trees within a patch that were initiated within 20 years of a fire.

<table>
<thead>
<tr>
<th>Age structure class</th>
<th>Age Class</th>
<th># Cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand Initiation</td>
<td>Stand initiation</td>
<td>1-2</td>
</tr>
<tr>
<td>Stem exclusion</td>
<td>Young forest</td>
<td>1</td>
</tr>
<tr>
<td>Understory reinitiation</td>
<td>Young forest</td>
<td>2</td>
</tr>
<tr>
<td>Young forest multi-cohort</td>
<td>Young forest</td>
<td>≥ 3</td>
</tr>
<tr>
<td>Old forest single cohort</td>
<td>Old forest</td>
<td>1</td>
</tr>
<tr>
<td>Old forest multi-cohort</td>
<td>Old forest</td>
<td>≥ 2</td>
</tr>
</tbody>
</table>
Table 3.4. Detailed summaries of model parameters, including: model name, response variable, response variable levels, predictor variables used, modeling method, and the accuracy and Kappa statistic. For the model names, ML= machine learning, Rule = rule-based classification. Where "vegetation" or "topography" are listed for the predictor variables, all variables of the listed type were included in the models (see Table C1 for full list of variables). If only a subset of the potential predictor variables of either type were used, they are listed individually. Three machine learning model frameworks were used; RF= random forests, CF= conditional forests, and EV=evolutionary tree. Accuracy and Kappa values in bold represent the highest performing model method.

<table>
<thead>
<tr>
<th>Model name</th>
<th>Response variable</th>
<th>Response variable levels</th>
<th>Predictor variables</th>
<th>Method</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML Sev Original</td>
<td>Severity$_c$umulative</td>
<td>non lethal, mixed, high</td>
<td>vegetation</td>
<td>RF</td>
<td>0.63</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CF</td>
<td>0.63</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EV</td>
<td>0.53</td>
<td>0.23</td>
</tr>
<tr>
<td>ML Sev Reclass 4</td>
<td>Severity$_c$umulative</td>
<td>non lethal, mixed, high, old high</td>
<td>vegetation</td>
<td>RF</td>
<td>0.58</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CF</td>
<td>0.53</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EV</td>
<td>0.53</td>
<td>0.32</td>
</tr>
<tr>
<td>ML Sev Reclass 3</td>
<td>Severity$_c$umulative</td>
<td>non lethal, mixed, high</td>
<td>vegetation</td>
<td>RF</td>
<td>0.65</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CF</td>
<td>0.59</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EV</td>
<td>0.53</td>
<td>0.27</td>
</tr>
<tr>
<td>ML Sev Reclass 4 Topo</td>
<td>Severity$_c$umulative</td>
<td>non lethal, mixed, high, old high</td>
<td>vegetation, topography</td>
<td>RF</td>
<td>0.58</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CF</td>
<td>0.53</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EV</td>
<td>0.53</td>
<td>0.34</td>
</tr>
<tr>
<td>ML Sev Reclass 3 Topo</td>
<td>Severity$_c$umulative</td>
<td>non lethal, mixed, high</td>
<td>vegetation, topography</td>
<td>RF</td>
<td>0.65</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CF</td>
<td>0.59</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EV</td>
<td>0.65</td>
<td>0.43</td>
</tr>
<tr>
<td>Model Name</td>
<td>Response Variable</td>
<td>Response Variable Levels</td>
<td>Predictor Variables</td>
<td>Method</td>
<td>Accuracy</td>
<td>Kappa</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------------</td>
<td>--------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>ML Sev Reclass 3 Topo Only</td>
<td>Severity&lt;sub&gt;cumulative&lt;/sub&gt;</td>
<td>non lethal, mixed, high  topography</td>
<td>RF</td>
<td>0.39</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CF</td>
<td>0.42</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EV</td>
<td>0.37</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>ML Sev Reclass 4 Final</td>
<td>Severity&lt;sub&gt;cumulative&lt;/sub&gt;</td>
<td>non lethal, mixed, high, old high</td>
<td>MEAN SLOPE, MEDIAN ELEVATION SIZE 1 CC, SIZE OS, MEAN TPI&lt;sub&gt;300&lt;/sub&gt;, HEAT LOAD INDEX MEDIAN CC OS CC SIZE 2 CC CRWN DIFF SIZE 4 CC SIZE 3 CC</td>
<td>RF</td>
<td>0.61</td>
<td>0.41</td>
</tr>
<tr>
<td>ML Sev Reclass 3 Final</td>
<td>Severity&lt;sub&gt;cumulative&lt;/sub&gt;</td>
<td>non lethal, mixed, high</td>
<td>MEAN SLOPE, MEDIAN ELEVATION SIZE 1 CC, SIZE OS, MEAN TPI&lt;sub&gt;300&lt;/sub&gt;, HEAT LOAD INDEX MEDIAN CC OS CC SIZE 2 CC CRWN DIFF SIZE 4 CC SIZE 3 CC</td>
<td>RF</td>
<td>0.71</td>
<td>0.49</td>
</tr>
<tr>
<td>Method</td>
<td>Response variable</td>
<td>Response variable levels</td>
<td>Predictor variables</td>
<td>RF</td>
<td>Kappa</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>--------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-----</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>ML Sev Most Recent</td>
<td>Severity$_{most recent fire}$</td>
<td>low, moderate, high</td>
<td>vegetation</td>
<td>0.55</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.50</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.50</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>ML Age Structure</td>
<td>Age structural class</td>
<td>stand initiation,</td>
<td>vegetation</td>
<td>0.47</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>stem exclusion,</td>
<td></td>
<td>0.32</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>understory reinitiation,</td>
<td></td>
<td>0.32</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>young forest multi-story,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>old forest single-story,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>old forest multi-story,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule Sev Original</td>
<td>Severity$_{cumulative}$</td>
<td>non lethal, mixed, high</td>
<td>vegetation</td>
<td>0.40</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Rule Sev Reclass 3</td>
<td>Severity$_{cumulative}$</td>
<td>non lethal, mixed, high</td>
<td>vegetation</td>
<td>0.43</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Rule Sev Most Recent</td>
<td>Severity$_{most recent fire}$</td>
<td>low, moderate, high</td>
<td>vegetation</td>
<td>0.43</td>
<td>0.12</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.5. Summary of the total number of sites and percent of sites in each cumulative severity class assigned to each of the four groups identified by the cluster analysis, as well as the mean (median for CRWN DIFF) value of each structural variable found to correlate significantly with the NMDS axes.

<table>
<thead>
<tr>
<th>Total # sites</th>
<th>Stand initiation</th>
<th>Open canopy multi-story</th>
<th>Closed canopy small trees</th>
<th>Closed canopy medium trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of sites</td>
<td>Non lethal</td>
<td>Mixed</td>
<td>High</td>
<td>Old high</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>0%</td>
<td>42%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>64%</td>
<td>48%</td>
<td>26%</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>31%</td>
<td>26%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>21%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>OS CC</td>
<td>75</td>
<td>34</td>
<td>72</td>
<td>75</td>
</tr>
<tr>
<td>SIZE OS</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>SIZE 1 CC</td>
<td>75</td>
<td>15</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>SIZE 2 CC</td>
<td>0</td>
<td>21</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>SIZE 3 CC</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>SIZE 4 CC</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CRWN DIFF</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Figure 3.1. Map showing the location of Montana within the U.S. (a), the location of the NCDE within the state of Montana, locations of the five study watersheds (c), detailed maps depicting the PI patch boundaries and cover types of each study watershed (d-h), and the location of each watershed within climate space (i). Diagonal and cross-hatched areas within (c) represent Glacier National Park and federally designated wilderness areas, respectively. Grey stippling in panels (d-h) indicate patches where dendroecological samples were collected. All watersheds in (d-h) are displayed at the same scale and share the same scale bar. Points in panel (i) represent mean annual maximum temperature ($T_{\text{max}}$) and precipitation values for all pixels in each study watershed derived from gridded climate normals (Daly et al. 2006).
Figure 3.2. Hypothesized and observed relationships between time since high severity fire (TSHSF), or stand age in the case of low severity sites, and model performance. The curves in the first column (panels a, e, i, m) graphically depict the hypothesized influence of TSHSF on model performance (see Table 3.5 for detailed explanation). Columns 2-4 present different model performance metrics as a function of TSHSF. Column 2 (panels b, f, j, n) shows TSHSF influences on model performance metrics including accuracy, Kappa statistic, and the multi-class area under the curve (AUC) of the receiver operating characteristic curve. Columns 3 (panels c, g, k, o) and 4 (panels d, h, l, p) show TSHSF influences on the specificity (true negative rate) and sensitivity (true positive rate) of each severity class affected by the reclassification. Desirable model performance occurs where sensitivity and specificity are maximized, which for this dataset occurs where sensitivity and specificity curves approach each other or cross. Unfilled points show the performance of the original model, prior to any reclassification. These points serve as a baseline for assessing the utility of each reclassification hypothesis and identifying the TSHSF breakpoints where model performance is optimized. Variation in the range of TSHSF values where model performance is evaluated is a function of the maximum and minimum TSHSF where sufficient sample size exists to run models. In the legend, subscripts refer to non stand-replacing (NSR), mixed (m), high (h), and high old (H) severity classes.
Figure 3.3. Relativized variable importance scores (a, c) and results from the variable reduction procedure (b, d) for the final cumulative fire severity models, \textit{MLSev Reclass 4 Final} (a, b) and \textit{MLSev Reclass 3 Final} (c, d). Results of the variable reduction procedure depict changes in model performance, determined by the Kappa statistic, as poorly performing predictor variables are eliminated. Filled points and error bars represent the median and standard error of the Kappa statistic values, respectively, from all 50 model runs for each iteration. Missing data values, e.g. for 12 predictor variables in (b), occurred where two variables in the previous iteration were tied for the minimum importance rank and were removed jointly. The dashed line shows the 1-SE rule threshold (Breiman 1984), which provides a useful benchmark for selection of the optimal, most parsimonious model.
Figure 3.4. Nonmetric multidimensional scaling (NMDS) plots depicting the relative position in multivariate space for a) structural classes produced by hierarchical cluster analysis, b) the original fire severity classes, and c) reclassified fire severity classes for all 70 sample patches. The multivariate space was defined using only the raw PI attributes selected as top predictors in the model selection routine. Each point in the NMDS plot represents the mean and 2-dimensional standard deviations for all patches within a class. NMDS axis 1 is correlated with SIZE OS (r=0.89), OS CC (r=-0.61), SIZE 4 CC (r=0.54), and CRWN DIFF (r=0.76). NMDS axis 2 is correlated with SIZE 2 CC (r=0.86) and SIZE 3 CC (r=0.80). Symbols in panel (a) depict black=stand initiation, dark grey=open canopy multi-story, light grey=closed canopy small trees, white=closed canopy medium trees and for panels (b-c) red=high, dark red=old high, orange=mixed, green=NSR.
Figure 3.5. Partial dependence plots for a subset of the top PI vegetation predictor variables used in the final random forest model (ML\textsubscript{Sev Reclass 4 Final}) of the four classes of fire severity. For classification, partial dependence is a measure of the influence of a specific predictor on the probability of a specific factor level when all other predictors are held constant. See Table A3.1 for detailed descriptions of each predictor variable.
Figure 3.6. Partial dependence plots for a subset of the top topographic predictor variables used in the final random forest model (ML\text{Sev Reclass 4 Final}) of the four classes of fire severity. For classification, partial dependence is a measure of the influence of a specific predictor on the probability of a specific factor level when all other predictors are held constant. See Table A3.1 for detailed descriptions of each predictor variable.
Appendix A3: Ancillary material from the photogrammetric analysis of historical aerial photos

Table A3.1. Description of all photo-interpreted vegetation attributes and topographic variables used as candidate predictors in the modeling exercises. For each variable, the variable type, its description and the range of values or factor levels, for continuous and categorical variables, respectively, is listed.

<table>
<thead>
<tr>
<th>Variable type</th>
<th>Variable Name</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>TOTAL CC</td>
<td>Total canopy cover estimated to the nearest ten percent</td>
<td>Range: 0-100% in ten percentile classes</td>
</tr>
<tr>
<td></td>
<td>OS CC</td>
<td>Canopy cover of the overstory layer estimated to the nearest ten percent</td>
<td>Range: 0-100% in ten percentile classes</td>
</tr>
<tr>
<td></td>
<td>US CC</td>
<td>Canopy cover of the understory layer estimated to the nearest ten percent</td>
<td>Range: 0-90% in ten percentile classes</td>
</tr>
<tr>
<td></td>
<td>SIZE OS</td>
<td>Overstory layer tree size class</td>
<td>Range: 5 size classes</td>
</tr>
<tr>
<td></td>
<td>SIZE US</td>
<td>Understory layer tree size class</td>
<td>Range: 4 size classes</td>
</tr>
<tr>
<td></td>
<td>SIZE 1 CC</td>
<td>Canopy cover of size class 1 trees</td>
<td>Range: 0-100% in ten percentile classes</td>
</tr>
<tr>
<td></td>
<td>SIZE 2 CC</td>
<td>Canopy cover of size class 2 trees</td>
<td>Range: 0-100% in ten percentile classes</td>
</tr>
<tr>
<td></td>
<td>SIZE 3 CC</td>
<td>Canopy cover of size class 3 trees</td>
<td>Range: 0-100% in ten percentile classes</td>
</tr>
<tr>
<td></td>
<td>SIZE 4 CC</td>
<td>Canopy cover of size class 4 trees</td>
<td>Range: 0-100% in ten percentile classes</td>
</tr>
<tr>
<td></td>
<td>SIZE 5 CC</td>
<td>Canopy cover of size class 5 trees</td>
<td>Range: 0-100% in ten percentile classes</td>
</tr>
<tr>
<td>Variable Type</td>
<td>Description</td>
<td>Values</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Vegetation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL CC</strong></td>
<td>Total canopy cover estimated to the nearest ten percent</td>
<td>Range 0-100% in ten percentile classes</td>
<td></td>
</tr>
<tr>
<td><strong>OS CC</strong></td>
<td>Canopy cover of the overstory layer estimated to the nearest ten percent</td>
<td>Range 0-100% in ten percentile classes</td>
<td></td>
</tr>
<tr>
<td><strong>US CC</strong></td>
<td>Canopy cover of the understory layer estimated to the nearest ten percent</td>
<td>Range 0-100% in ten percentile classes</td>
<td></td>
</tr>
<tr>
<td><strong>SMALL CC</strong></td>
<td>Canopy cover of tree size classes smaller than the subwatershed median</td>
<td>Range 0-100% in ten percentile classes</td>
<td></td>
</tr>
<tr>
<td><strong>MEDIUM CC</strong></td>
<td>Canopy cover of the median tree size class within a subwatershed</td>
<td>Range 0-100% in ten percentile classes</td>
<td></td>
</tr>
<tr>
<td><strong>LARGE CC</strong></td>
<td>Canopy cover of tree size classes larger than the subwatershed median</td>
<td>Range 0-100% in ten percentile classes</td>
<td></td>
</tr>
<tr>
<td><strong>CANOPY LAYERS</strong></td>
<td>Number of canopy layers</td>
<td>Range: 3 canopy strata</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = Single canopy layer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = Two canopy layers</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = More than two layers</td>
<td></td>
</tr>
<tr>
<td><strong>DEAD SNAG</strong></td>
<td>Percent of dead tree canopy cover</td>
<td>Range: 5 snag abundance classes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = none</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = &lt; 10% dead tree canopy cover</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = 10-39% dead tree canopy cover</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 = 40-70% dead tree canopy cover</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 = &gt; 70% dead tree canopy cover</td>
<td></td>
</tr>
<tr>
<td><strong>CLMP DENS</strong></td>
<td>Degree of clumpiness of tree cover</td>
<td>Range: 3 clumpiness classes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = widely scattered clumps</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = moderately dense clumps</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = dense clumps</td>
<td></td>
</tr>
<tr>
<td><strong>CRWN DIFF</strong></td>
<td>Degree of crown size differentiation among overstory trees</td>
<td>Range: 3 crown differentiation classes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = low (&lt; 30% difference)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = moderate (30-100% difference)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = high (&gt; 100% difference)</td>
<td></td>
</tr>
<tr>
<td><strong>CLUMP SIZE</strong></td>
<td>Degree of clumpiness of tree cover</td>
<td>Range: 3 clump size classes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = small (&lt; 1 acre)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = medium (1-5 acres)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = large (&gt; 5 acres)</td>
<td></td>
</tr>
<tr>
<td><strong>STRUCTURE</strong></td>
<td>Multivariate structural classes describing forest structural properties derived from raw canopy cover and size class attributes. Structural classes may arise from different successional pathways or disturbance histories and are not meant to represent a fixed chronological developmental sequence.</td>
<td>Range: 7 structural classes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = stand initiation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = open canopy stem exclusion</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = closed canopy stem exclusion</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 = understory reinitiation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 = young forest multi-story</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 = old forest, single-story</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7 = old forest, multi-story</td>
<td></td>
</tr>
</tbody>
</table>
**STRUCTURE**

Multivariate structural classes describing forest structural properties derived from raw canopy cover and size class attributes. Structural classes may arise from different successional pathways or disturbance histories and are not meant to represent a fixed chronological developmental sequence.

**COVER**

The dominant cover type characterizing a patch. Cover type was assigned based on the composition of understory and overstory canopy components that comprised the majority of canopy cover.

---

**Topographic**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEDIAN ELEVATION</strong></td>
<td>Median elevation of all pixels within a patch</td>
</tr>
<tr>
<td><strong>MEAN SLOPE</strong></td>
<td>Average slope of all pixels within a patch</td>
</tr>
<tr>
<td><strong>MEAN TRANSFORMED ASPECT</strong></td>
<td>Average azimuth of all pixels within a patch. Aspect was cosine-transformed to a northness index, where higher (lower) values represent more northern (southern) aspects.</td>
</tr>
<tr>
<td><strong>HEAT LOAD INDEX</strong></td>
<td>Multi-variate potential heat index based on latitude, slope and aspect</td>
</tr>
<tr>
<td><strong>MEAN TPI&lt;sub&gt;300&lt;/sub&gt;</strong></td>
<td>Patch mean TPI index value calculated with a neighborhood scale of 300 pixels</td>
</tr>
<tr>
<td><strong>TPI&lt;sub&gt;300&lt;/sub&gt; CLASS</strong></td>
<td>Median patch TPI value classed as:</td>
</tr>
</tbody>
</table>

- < mean - 1 Std. dv.
- -0.5 Std. dv. < x ≥ -1.0 Std. dv.
- mean + 0.5 Std. dv. < x > mean - 0.5 Std. dv. & slope ≤ 5°
- Range: 6 TPI classes
  - 1 = valley bottom
  - 2 = lower slopes
  - 3 = flat slopes

---

**Values**

1 = small (< 1 acre)
2 = medium (1-5 acres)
3 = large (> 5 acres)

<table>
<thead>
<tr>
<th>Range: 7 structural classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = stand initiation</td>
</tr>
<tr>
<td>2 = open canopy stem exclusion</td>
</tr>
<tr>
<td>3 = closed canopy stem exclusion</td>
</tr>
<tr>
<td>4 = understory reinitiation</td>
</tr>
<tr>
<td>5 = young forest multi-story</td>
</tr>
<tr>
<td>6 = old forest, single-story</td>
</tr>
<tr>
<td>7 = old forest, multi-story</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Range: 12 cover classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = PIPO</td>
</tr>
<tr>
<td>2 = LAOC</td>
</tr>
<tr>
<td>3 = PSME</td>
</tr>
<tr>
<td>4 = ABGR</td>
</tr>
<tr>
<td>5 = PICO</td>
</tr>
<tr>
<td>6 = PIEN/ABLA</td>
</tr>
<tr>
<td>7 = TSHE/THPL</td>
</tr>
<tr>
<td>8 = POTR</td>
</tr>
<tr>
<td>9 = PIFL</td>
</tr>
<tr>
<td>10 = grassland</td>
</tr>
<tr>
<td>11 = shrubland</td>
</tr>
<tr>
<td>12 = non vegetated</td>
</tr>
</tbody>
</table>

---

**Size class**

- Size class 1 = seedlings and saplings (< 5.0" DBH)
- Size class 2 = poles (5 to 8.9" DBH)
- Size class 3 = small trees (9 to 15.9" DBH)
- Size class 4 = medium trees (16 to 25.0" DBH)
- Size class 5 = large trees (> 25.0" DBH)

---

14
Patch mean TPI index value calculated with a neighborhood scale of 1500 pixels

Median patch TPI value classed as:

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; mean - 1 Std. dv.</td>
<td>&lt; mean - 1 Std. dv.</td>
</tr>
<tr>
<td>-0.5 Std. dv. &lt; x ≥ -1.0 Std. dv.</td>
<td>-0.5 Std. dv. &lt; x ≥ -1.0 Std. dv.</td>
</tr>
<tr>
<td>mean + 0.5 Std. dv. &lt; x &gt; mean - 0.5 Std. dv. &amp; slope ≤ 5°</td>
<td>mean + 0.5 Std. dv. &lt; x &gt; mean - 0.5 Std. dv. &amp; slope ≤ 5°</td>
</tr>
<tr>
<td>mean + 0.5 Std. dv. &lt; x &gt; mean - 0.5 Std. dv. &amp; slope &gt; 5°</td>
<td>mean + 0.5 Std. dv. &lt; x &gt; mean - 0.5 Std. dv. &amp; slope &gt; 5°</td>
</tr>
<tr>
<td>mean + 1 Std. dv. ≤ x &gt; mean + 0.5 Std. dv.</td>
<td>mean + 1 Std. dv. ≤ x &gt; mean + 0.5 Std. dv.</td>
</tr>
<tr>
<td>&gt; mean + 1 Std. dv.</td>
<td>&gt; mean + 1 Std. dv.</td>
</tr>
</tbody>
</table>

Range: 6 TPI classes

1 = valley bottom
2 = lower slopes
3 = flat slopes
4 = middle slopes
5 = upper slopes
6 = Ridgetops

‡ - Overstory and understory canopy cover sum to TOTL CC

* - Cover type codes consist of 4 letters representing the first 2 letters of the genus and species. PIPO = Pinus ponderosa, LAOC = Larix occidentalis, PSME = Pseudotsuga menziesii, PICO = Pinus contorta, PIEN = Picea engelmannii, ABGR = Abies grandis, ABLA = Abies lasiocarpa, TSHE = Tsuga heterophylla, THPL = Thuja plicata, POTR = Populus tremuloides, and PIMO = Pinus monticola, PIFL = Pinus flexilis.

# - Heat load index was calculated using Eq. 3 in McCune & Keon 2002.

$ - TPI and TPI Class were calculated according to Weiss 2001

† - variable was removed from models due to high correlation
Table A3.2. Dichotomous key adapted from Hessburg et al. (2007) used to classify the severity of the most recent fire at each patch. Rangeland includes grassland and shrubland cover types, whereas non-rangeland includes non-vegetated types such as bare rock or mineral soil. Overstory size class = SIZE OS, understory size class = SIZE US, overstory canopy percent = OS CC, cover type = COVER. Shade intolerant cover types include PIPO, PSME, LAOC, PICO, PIMO, and JUSC.

<table>
<thead>
<tr>
<th>Decision Path</th>
<th>Description</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. Patch is not forested</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a. Patch is rangeland</td>
<td></td>
<td>High severity</td>
</tr>
<tr>
<td>2b. Patch is non-rangeland</td>
<td></td>
<td>No severity</td>
</tr>
<tr>
<td>1b. Patch is forested</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a. Overstory size class ≥ small trees and understory size class ≤ small trees</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4a. Overstory canopy percent ≥ 70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5a. Cover type is shade tolerant</td>
<td></td>
<td>High severity</td>
</tr>
<tr>
<td>5b. Cover type is shade intolerant</td>
<td></td>
<td>Low severity</td>
</tr>
<tr>
<td>4b. Overstory canopy percent &lt; 70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6a. Overstory canopy percent ≤ 20%</td>
<td></td>
<td>High severity</td>
</tr>
<tr>
<td>6b. Overstory canopy percent &gt; 20%</td>
<td></td>
<td>Moderate severity</td>
</tr>
<tr>
<td>3b. Overstory size class &lt; small trees or understory size class &gt; small trees</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7a. Overstory size class &lt; small trees</td>
<td></td>
<td>High severity</td>
</tr>
<tr>
<td>7b. Understory size class &gt; small trees</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8a. Overstory canopy percent ≤ 20%</td>
<td></td>
<td>High severity</td>
</tr>
<tr>
<td>8b. Overstory canopy percent &gt; 20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9a. Overstory canopy percent ≥ 70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10a. Cover type is shade tolerant</td>
<td></td>
<td>High severity</td>
</tr>
<tr>
<td>10b. Cover type is shade intolerant</td>
<td></td>
<td>Low severity</td>
</tr>
<tr>
<td>9b. Overstory canopy percent &lt; 70%</td>
<td></td>
<td>Moderate severity</td>
</tr>
</tbody>
</table>
Appendix B3: Dendroecological supplementary information

Determination of age thresholds used in age structure classes

To characterize key features of the age structure of each patch that are closely linked to disturbance history, we developed a set of age structure classes (Table 3.3) based on the number of cohorts observed in a patch and its age class. Determination of the number of cohorts within a patch is described in the main text (see Reconstruction of patch-level fire history and severity). To determine appropriate classification thresholds for the stand age class variable, we drew directly from patterns evident in our age structure data and consulted the literature. Conceptually, the stand initiation phase represents the first successional phase following high severity disturbance that is characterized by prolific regeneration and ends as increased crown closure limits subsequent regeneration of colonizer species (O’Hara et al. 1996, Oliver 1980). Although the length of the stand initiation phase likely varies between forest types and across climatic or edaphic gradients, a number of studies from forest types similar to ours have documented periods of active recruitment for 40-60 years following high severity fire (Oliver 1980, Ehle & Baker 2003, Haire & McGarigal 2010, Tepley et al. 2014). Post-fire tree recruitment in sites classified as high severity in this study corroborated this range (Fig. B3.1a). The mode of tree establishment following high severity fire in our sites, an approximation of the timing of canopy closure, occurred within < 50 years for 96% of our sites. Therefore, we used the 50-year threshold to separate the stand initiation and young forest phases.
Compared to the stand initiation phase, there is much greater variation in the range of stand ages used to define old forest, but minimum age thresholds of 150-300 years old are commonly used (Habeck 1988, Franklin & Spies 1991, Kaufman et al. 2007). Based on age data pooled across all five of our study watersheds, we found that a 200-year threshold closely matched the 75th percentile of stand ages (215 years, Fig. B3.1b) and the mean age of individual trees within the 75th percentile (192 years, data not shown). We therefore judged that an age threshold of 200 years was consistent with the literature as well as the aggregate tree and stand age structures in our study area.
Figure B3.1. Histograms displaying frequency distributions for (a) years to canopy closure for all sites classed as high severity and (b) stand age expressed as a percent of the total number of study sites. The number of years to canopy closure was estimated as the modal value of pith dates. Only high severity sites were used in this calculation since, by definition, the stand initiation stage only occurs following high severity fire. Stand age is defined as the median age of the oldest cohort comprising > 10% of total stand basal area.
Figure B3.2. Example graphics showing the fire history and age structure of NSR (a-d), mixed (e-j), and high (k-l) severity patches. Establishment dates are binned into 10 year intervals. Note that the y-axis scales differ between graphs. These sites were unaffected by the reclassification procedure. Species codes are: PIPO=Pinus ponderosa, PSME=Pseudotsuga menziesii, LAOC=Larix occidentalis, ABGR=Abies grandis, PIMO=Pinus monticola, THPL=Thuja plicata, PICO=Pinus contorta, PIEN=Picea engelmannii, ABLA=Abies lasiocarpa, BEPA=Betula papyrifera, JUSC=Juniperus scopulorum.
Figure B3.3. Example graphics showing the fire history and age structure of sites affected by hypothesis 1 (a-b), hypothesis 2 (c-d) hypothesis 3 (e-f), and hypothesis 4 (g-h). Under H1, young NSR sites were reclassed as mixed severity. Under H2, mixed severity sites with low time since-high-severity-fire (TSHSF) were reclassed as high severity. Under H3, mixed severity sites with long TSHSF were reclassed as NSR severity. Under H4, high severity sites were separated into young and old high severity. See Table 6 for details about the reclassification hypotheses. Establishment dates are binned into 10 year intervals. Note that the y-axis scales differ between graphs. Species codes are the same as in Fig. B3.2.
Appendix C3: Structure-based model ancillary outputs

Figure C3.1. Examples of (a) conditional tree and (b) evolutionary tree resulting from the conditional Random Forest and evolutionary tree algorithms using cumulative fire severity class (MLSev Reclass 4 Final) as the response variable and PI structural attributes as predictors. Each tree represents the output from a single model run and does not reflect the exact structure of the ensemble models. Note the qualitative similarity in the structure (e.g. selected variables and split points) of the models. Severity classes are coded as: nsr= non stand-replacing, m=mixed, h=high, and h.o=old high. See Table 3.3 for detailed descriptions of the predictor variables.
Table C3.2. Confusion matrix of the reference (dendroecological) and predicted (structure-based) classes for the Random Forest model of the severity of the most recent fire (MLSev Most Recent), the overall model performance (Accuracy and Kappa), and class-specific performance statistics. Values in the confusion matrix are the percent of all predictions. Sensitivity and specificity represent the true positive and true negative rates, respectively, of the predicted severity classes. Overall accuracy, Kappa, sensitivity, specificity, and balanced accuracy represent median values from 500 model iterations.

<table>
<thead>
<tr>
<th>Predicted (Structure-based)</th>
<th>Reference (Dendroecological)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Balanced accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>low</td>
<td>13</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>moderate</td>
<td>9</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>8</td>
<td>4</td>
<td>26</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td></td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappa</td>
<td></td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C3.3. Confusion matrix of the reference (dendroecological) and predicted (structure-based) classes for the age structure Random Forest model (ML_Age structure), the overall model performance (Accuracy and Kappa), and class-specific performance statistics. Values in the confusion matrix are the percent of all predictions. Sensitivity and specificity represent the true positive and true negative rates, respectively, of the predicted severity classes. Overall accuracy, Kappa, sensitivity, specificity, and balanced accuracy represent median values from 500 model iterations. Age structure class abbreviations are: SI = stand initiation, SE = stem exclusion, YFMS = young forest multi-story, UR = understory initiation, OFMS = old forest multi-story.

<table>
<thead>
<tr>
<th>Predicted (Structure-based)</th>
<th>Reference (Dendroecological)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Balanced accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI</td>
<td>10 3 0 3 0 0.64 0.92 0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>2 6 1 3 3 0.30 0.87 0.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YFMS</td>
<td>0 2 5 1 2 0.44 0.94 0.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UR</td>
<td>3 3 0 10 5 0.37 0.85 0.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OFMS</td>
<td>0 7 4 9 16 0.62 0.73 0.68</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall accuracy 0.47
Kappa 0.32
Table C3.4. Confusion matrix of the reference (dendroecological) and predicted (structure-based) classes for the rule-based model of the severity of the most recent fire (Rule\textsubscript{Sev Most Recent}), the overall model performance (Accuracy and Kappa), and class-specific performance statistics. Values in the confusion matrix are the percent of all predictions. Sensitivity and specificity represent the true positive and true negative rates, respectively, of the predicted severity classes. Overall accuracy, Kappa, sensitivity, specificity, and balanced accuracy represent median values from 500 model iterations.

<table>
<thead>
<tr>
<th>Predicted (Structure-based)</th>
<th>Reference (Dendroecological)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Balanced accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>7</td>
<td>0.23</td>
<td>0.77</td>
<td>0.50</td>
</tr>
<tr>
<td>moderate</td>
<td>10</td>
<td>0.33</td>
<td>0.80</td>
<td>0.56</td>
</tr>
<tr>
<td>high</td>
<td>14</td>
<td>0.67</td>
<td>0.56</td>
<td>0.61</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappa</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C3.5. Confusion matrix of the reference (dendroecological) and predicted (structure-based) classes for the cumulative fire severity Random Forest model (MLSev Reclass 4 Topo), the overall model performance (Accuracy and Kappa), and class-specific performance statistics. Values in the confusion matrix are the percent of all predictions. Sensitivity and specificity represent the true positive and true negative rates, respectively, of the predicted severity classes. Overall accuracy, Kappa, sensitivity, specificity, and balanced accuracy represent median values from 500 model iterations.

<table>
<thead>
<tr>
<th>Predicted (Structure-based)</th>
<th>Reference (Dendroecological)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Balanced accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSR</td>
<td>NSR 14 mixed 6 high 1 old high 4</td>
<td>0.66</td>
<td>0.86</td>
<td>0.76</td>
</tr>
<tr>
<td>mixed</td>
<td>NSR 5 mixed 27 high 10 old high 3</td>
<td>0.63</td>
<td>0.69</td>
<td>0.66</td>
</tr>
<tr>
<td>high</td>
<td>NSR 0 mixed 6 high 15 old high 1</td>
<td>0.58</td>
<td>0.90</td>
<td>0.74</td>
</tr>
<tr>
<td>old high</td>
<td>NSR 2 mixed 3 high 1 old high 3</td>
<td>0.28</td>
<td>0.94</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Overall accuracy 0.58
Kappa 0.40
Chapter 4

Summary and Conclusions

*Landscape dynamics and resilience mechanisms in MSFRs*

We document some differences in the landscape dynamics of the GYE and NCDE, but the most striking result of our analyses is the strongly contrasting landscape dynamics of both study regions compared with those of low- and high-severity fire regime forests. The mix of lethal and non lethal fire effects that influenced most patches in both study regions created non equilibrium dynamics at the patch-scale, resulting in landscapes dominated by younger forests in a variety of successional and structural conditions (Chapter 2). The prevalence of high severity fire in both of our study regions suggests that MSFR forests were quite resilient to high-severity fire over larger spatio-temporal scales than has traditionally been considered characteristic of mixed-conifer forests. High-severity fire may have created positive feedbacks (Larson et al. 2013) in small portions of our study watersheds, but the predominance of mixed-severity fire implies that most patches impacted by high-severity fire escaped this feedback loop.

Although our data do not reveal the exact mechanism that confers such resilience, they strongly indicate that an alternative resilience mechanism must exist for MSFRs that is not accounted for by current theories. We suspect that this mechanism is related to the spatio-temporal heterogeneity of fuel structures and the generally low fire hazard that persists for many decades following high-severity fire prior to the eventual dominance of closed canopy, even-aged forest that
homogenizes landscape pattern and promotes high-severity fire (Johnson and Gutsell 1994, McCarthy et al. 2001). Short-term (< 10 years) empirical studies and theoretical models have documented spatial variability in the patterns of vegetation regeneration and fuel profile development shortly after high-severity fire resulting from pre-fire conditions, variable retention of legacy structures, topoedaphic influences, the size and shape of high-severity patches, and post-fire climate (Turner et al. 1998, Swanson et al. 2011, Donato et al. 2012, Hudec and Peterson 2012, Harvey et al. 2016). Relatively few studies have quantified multi-decadal changes in the spatial patterns of forest and fuel succession following high-severity fire in mixed-conifer forests, but those that do generally corroborate our hypothesis that fuel accumulation, and presumably fire hazard, is relatively low over a 40-120 year post-fire window, depending on the fuel variable of interest and site productivity (Hall et al. 2006, Dunn and Bailey 2015). Fires that recur during this time of fuel discontinuity and low fire hazard are likely to exhibit spatially heterogeneous fire behavior and predominantly low-moderate severity effects. Given that median fire return intervals in our study areas ranged between 30-100 years (Chapter 2), the alternative resilience mechanism we propose for MSFRs seems feasible and warrants greater study in the future.

*Management implications of novel resilience mechanisms and dynamics in MSFRs*

There are important debates about the influence of past land management (e.g. timber harvest, grazing, fire exclusion) on stand and landscape structural conditions (Franklin and Johnson 2012, Williams and Baker 2012, DellaSala et al.
2013, Fulé et al. 2013, Odion et al. 2014, Stevens et al. 2016) and resilience of MSFRs to management- or climate-driven increases in fire activity. Although it is commonly asserted that fire exclusion has altered forest conditions in MSFRs (Brown 2010, Perry et al. 2011, Hessburg et al. 2016), previous studies have shown mixed results for this hypothesis at both stand- (Arno et al. 1995, Bekker and Taylor 2001, Scholl and Taylor 2010, Schoennagel et al. 2011, Heyerdahl et al. 2014, Tepley and Veblen 2015) and landscape-scales (Hessburg et al. 1999, Hessburg et al. 2005, Platt and Schoennagel 2009). This result is likely attributable to the influence of historical high-severity fire in portions of some MSFRs that created dense forest conditions and persistent negative feedbacks on subsequent tree recruitment prior to fire exclusion (Baker et al. 2007). Although we have not explicitly addressed fire exclusion impacts here, examination of age structure graphs from our study regions (Chapters 2-3) corroborates a mixed establishment response to fire exclusion that is consistent with this mechanism. Strong stand-level changes since the onset of fire exclusion in the young, moderately dense forest patches with simple forest structures (e.g. 1-2 cohorts) that characterized most of our study areas are therefore unlikely. Rather, the main influence of fire exclusion has more likely been a loss of landscape heterogeneity and greater landscape contagion that could facilitate larger high-severity burns. The resilience to high-severity fire exhibited by MSFR forests in our study region suggests that resilience and management-induced changes to forest structural conditions may be more weakly coupled in many MSFRs than in low-severity fire regime forests.
Based on these observations, application of widely-used stand-level fuels treatments designed to reduce stand density and create multi-aged, open-canopy conditions (Agee and Skinner 2005) would not restore historical conditions, are not consistent with historical fire-mediated dynamics or resulting stand structures, and may not be necessary or effective in imparting increased resilience. In place of traditional fuel management approaches, our results highlight a separate set of critical questions that must be addressed by future research to understand the management implications associated with MSFRs, including: (1) what is the relative importance of vegetation versus other bottom-up (e.g. topography) or top-down (e.g. climate variability) controls on fire severity patterns, (2) what characteristics of the vegetation mosaic (e.g. grain, composition of structural types, patch age) most strongly influence fire severity-mediated dynamics, especially the scale of high-severity fire, (3) how sensitive are landscape dynamics and resilience of MSFRs to changing scales of high-severity fire.

Dendro-calibrated, structure-based fire severity models: A novel cross-scale reconstruction method

In chapter 3, we demonstrate that dendro-calibrated structure-based models of cumulative fire severity can perform relatively well, but require rigorous validation and accounting of the temporal and spatial limitations that exist for both data types. The ability to characterize fire severity-mediated dynamics of entire patches using small-plot dendroecological data (chapter 2) is the keystone result that serves as the foundation for the cross-scale, spatio-temporal reconstructions of
historical fire severity that we develop in subsequent chapters. The structure-based fire severity reconstruction method we develop in chapter 3 represents a substantial improvement over previous aerial photo-based severity reconstructions (Hessburg, et al. 2007, Taylor & Skinner 1998, Beatty & Taylor 2008) because it is the only method based on a data-driven sensitivity analyses of the response variables that are best predicted by forest structure, predictor variable relationships to fire severity classes, and performance of different modeling methods. The performance of our final model compares relatively well to other calibrated fire severity models (Sherriff and Veblen 2007, Williams and Baker 2012, Sherriff et al. 2014), but it is maybe most notable because it provides empirically observed, spatially explicit patch sizes rather than coarsely modeled spatial extent of severity classes. For this reason, our model may be an especially important tool for evaluating the spatial patterns of fire severity in MSFRs that have been difficult to address using other reconstruction methods.

Understanding methodological challenges and fire-mediated dynamics in the context of the fire regime gradient

Debate about MSFRs often takes place within the context of a singular MSFR model akin to low-and high-severity models. We view this as a key weakness of the current scientific debate that has impeded research progress and generated confusion. In Chapter 2, we attempt to recast this debate by placing MSFRs in the context of a fire regime gradient. As we demonstrate, this context is valuable for understanding both the methodological challenges of historical reconstructions as
well as the variation in ecological dynamics exhibited by MSFRs. It may also lead to a stronger conceptual understanding of how and why landscape dynamics vary along the fire regime gradient that could improve spatially explicit modeling of fire regime dynamics and responses to alternative management or climate scenarios. For instance, geospatial statistical techniques have been developed to identify the environmental drivers of recent fire activity (Parisien and Moritz 2009, Krawchuk and Moritz 2011) and severity (Whitman et al. 2015) and model their spatial distribution. The broad-scale reconstruction of fire severity patterns that is possible with our structure-based models could be combined with these geospatial modeling techniques to map the distribution of historical fire regimes along environmental gradients.

**Next steps:** Application of structure-based models to evaluate the spatial ecology and biogeography of MSFRs across the northern Rockies

Some of the most pressing questions about MSFRs regard their spatial ecology, biogeography and biophysical drivers (Barton 2002, Savage and Mast 2005, Collins and Stephens 2010, Halofsky et al. 2011, Perry et al. 2011, Collins and Roller 2013, Mallek et al. 2013). Specifically, critical questions that have been identified in the literature are: (1) what scale of high-severity fire was experienced by MSFRs historically?, (2) how did the patch-level mix of fire severities and the scale of high-severity fire vary across the landscape?, and (3) what biophysical factors drove the variability fire severity across the landscape? To address these questions, we have used the structure-based model developed in chapter 3 to predict
fire severity across all five watersheds in the NCDE where our dendroecological dataset was collected. Using this dataset, we have constructed patch size distributions and determined the percent of each watershed influenced by each fire severity class. We are developing statistical methods to assess the relative importance and relationships between biophysical factors (fire frequency, topography, and climatic setting) and the percent of watershed area or mean patch size in each severity class. Once this work is completed, we will apply these analytical and statistical methods to a large dataset of almost 100 photo-interpreted watersheds of equivalent size to the five used here, covering > 1,000,000 ha and > 25,000 patches in total, to reconstruct spatio-temporal patterns and drivers of historical fire severity across western Montana, northern Idaho and northeastern Washington. For our GYE study areas, we aim to complete the aerial photo-interpretation for our two study areas so that these same methods and questions can be applied to that region.
References


Barrett, S. W., S. F. Arno, and J. P. Menakis. 1997. Fire Episodes in the Inland Northwest (1540-1940) Based on Fire History Data. USDA Forest Service, Intermountain Research Station, Ogden, UT.


Littell, J. S. 2002. Determinants of fire regime variability in lower elevation forests of the northern Greater Yellowstone Ecosystem. MS thesis. Montana State University, Bozeman, MT.


Pfister, R. D., B. L. Kovalchik, S. F. Arno, and R. C. Presby. 1977. Forest habitat types of Montana. INT GTR-34.


