

LINEAR MODELING AND EVALUATION OF CONTROLS ON FLOW

RESPONSE IN WESTERN POST-FIRE WATERSHEDS

by

Samuel Saxe

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Hydrology).

Golden, Colorado

Date _____

Signed: _____
Samuel Saxe

Signed: _____
Dr. Terri Hogue
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____
Dr. Terri Hogue
Hydrologic Science and Engineering Program

Signed: _____
Dr. John McCray
Department of Civil and Environmental Engineering

ABSTRACT

This research investigates the impact of wildfires on watershed flow regimes throughout the western United States, specifically focusing on evaluation of fire events within specified subregions and determination of the impact of climate and geophysical variables in post-fire flow response. Fire events were collected through federal and state-level databases and streamflow data were collected from U.S. Geological Survey stream gages. 82 watersheds were identified with at least 10 years of continuous pre-fire daily streamflow records and 5 years of continuous post-fire daily flow records. For each watershed, percent changes in annual runoff ratio (RO), low-flows (LF), high-flows (HF), peak flows (PF), number of zero flow days (Nzeros), baseflow index (BFI), and Richards-Baker flashiness index (RB) were calculated from pre- to post-fire. Numerous independent variables were identified for each watershed and fire event, including topographic, vegetation, climate, burn severity, and soils data. The national watersheds were divided into five regions through k-means clustering and LASSO linear regression models were calculated for each region. Regression models were also produced for watersheds grouped by total area burned. The coefficient of determination (R^2) was used to determine the accuracy of the resulting models. Model accuracy was highly variable, both by group and by response variable. Resulting coefficient values demonstrate that, of the watershed parameters applied in this study as explanatory variables, watershed area and burn severity parameters explain the greatest amount of the post-fire flow change variability. Burn area slope and soil erodibility factor (K_{fact}) also contribute significantly to post-fire response. Watershed area and K_{fact} are generally negatively correlated with response variables, while slope and percent moderate burn severity (BS_M) are positively correlated.

TABLE OF CONTENTS

ABSTRACT.....	iii
LIST OF FIGURES.....	vi
LIST OF TABLES.....	vii
Chapter 1 INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Study area.....	4
CHAPTER 2 METHODS.....	7
2.1 Data collection.....	7
2.2 Response variables.....	12
2.3 <i>k</i> -means clustering.....	14
2.4 LASSO regression.....	15
CHAPTER 3 REGRESSION RESULTS.....	17
3.1 Clustering.....	17
3.2 Response variable distribution and analyses.....	18
3.3 Regression results.....	24
CHAPTER 4 SUMMARY AND CONCLUSION.....	35

REFERENCES CITED.....36

LIST OF FIGURES

Figure 1.1	CONUS map of the locations of the 82 watersheds utilized in this study.....	4
Figure 1.2	Boxplots summarizing the range of watershed areas, elevations, percent area burned, percent of precipitation that falls as snow, and aridity index for studied watersheds.....	5
Figure 1.3	Distribution of NLCD land cover types.....	6
Figure 3.1	CONUS map of the results of <i>k</i> -means clustering of the GAGES-II watershed set.....	17
Figure 3.2	CONUS map of the results of watershed clustering.....	17
Figure 3.3	Variations on explanatory variables by cluster.....	18
Figure 3.4	Boxplots of the response variables utilized in this study.....	19
Figure 3.5	Scatterplots of 1 st and 2 nd year response variables versus 5 year mean response variables.....	20
Figure 3.6	Scatterplots of all response variables versus respective watershed percent area burned. Simple regression lines provided to show correlations.....	21
Figure 3.7	Boxplots of relative response variables by cluster.....	22
Figure 3.8	Scatterplots of R ² and p-value results of simple linear regression modeling of response variables versus as increasing percent burn limitation. Squares and R ² and circles are p-values. Filled squares are R ² >= 0.50, filled circles are p <= 0.05. The blue dashed line is R ² = 0.50, the red dotted line is p = 0.05.....	23
Figure 3.9	Error statistics of the results of LASSO regression modeling of response variables. Filled squares are models with R ² >= 0.50.....	25
Figure 3.10	β_w values for all response variables.....	26
Figure 3.11	β_w in logistic format (if $\beta_w > 0$, $\beta_w = 1$. If $\beta_w < 0$, $\beta_w = -1$)	27

LIST OF TABLES

Table 2.1	Summary of the explanatory variables used in this study.....	7
-----------	--	---

CHAPTER 1

INTRODUCTION

Identifying the effects of wildfires on watershed hydrologic systems and the watershed qualities that have the greatest influence on flow response are imperative studies in regards to understanding the consequences overall increasing rate of wildfires and the lack of large-scale research within the United States.

1.1 Introduction

The rate of wildfires in the western United States (hereby, western) is increasing annually, on average, costing federal agencies billions of dollars a year in suppression efforts (Whitlock, 2004) and causing increases in flood events destructive to both life and infrastructure in many parts of the world (Daniel G Neary, 2003; Juli G. Pausas, 2008). Westerling et. al. (2006) showed that the western fire regime exhibited a significant transition from infrequent and short-duration events to a higher frequency, longer duration regime during the mid-1980's. The greatest increases in fire frequency were found to occur in mid-elevation forests, most commonly in the Northern Rockies, Sierra Nevada, southern Cascades, and western Coast Ranges in northern California and southern Oregon. This marked change is strongly correlated with significant climate change events, such as warmer springs and longer dry seasons, commonly in occurrence with reduced winter precipitation rates and earlier spring snowmelt (REF). Overall, Westerling et. al. 2006 determine that, though land-use history may be a significant factor in the spatial distribution of wildfires within specific forest types, changes in fire regimes in the western US can most likely be attributable to recent changes in climate. Other notable research has also provided significant correlatory evidence between climate change and wildfire occurrences (Littell et al., 2009; Moritz et al., 2010)

Though wildfires are a part of the natural process of vegetation dynamics, they cause wide-ranging changes to ecosystems (Daniel G Neary, 2003; Santos et al., 2015) depending on numerous factors, including burn severity. Studies examining the effects of wildfires on a small-scale such as in plot-sized and laboratory experiments, show fire temperatures can result in the combustion of organic matter within soils and cause permanent alteration to the chemical structure of local clays, decreasing soil stability (Shakesby and Doerr, 2006). Water-repellent soil layers can be created in a discrete layer on or below the soil surface through chemical bonding of the combusted organic matter to mineral particles, potentially increasing overall topsoil erosion rates in burned regions (Wilkinson et al., 2009), though this hydrophobicity is highly variable depending on fire behavior, burn severity and soil properties (DeBano, 2000).

At larger scales, such as entire watersheds or multiple watershed systems, studies of post-fire erosion rates have shown incompatible conclusions (Moody and Martin, 2001; Owens et al., 2013; Smith et al., 2011), though this is most likely due to the variability of precipitation events and general climate patterns (Moody et al., 2013). In terms of water quality, contaminant levels can be dramatically increased for many years after a wildfire in both soil (Burke et al., 2010) and stream systems (Emelko et al., 2011; Stein et al., 2012), increasing the workload on source water protection organizations in communities reliant upon burned watersheds for drinking and farm water. Furthermore, wildfires are readily attributed as the cause of substantial increases in debris flows (Benavides-Solorio and MacDonald, 2001; Cannon et al., 2001; Meyer et al., 2001).

Studies evaluating post-fire water yield change are highly disparate owing to the transient nature of climate patterns, variations in basin geomorphology, and vegetation recovery patterns, and the resulting complex interactions (Moody et al., 2013). For example, studies in rangeland regions of the United States found moderate increases in flow, infiltration, and erosion rates after major wildfires, with trends continuing for as long as 15 years (Emmerich and Cox, 1994; Frederick B. Pierson, 2009;

Hester et al., 1997). Fires in chaparral environments, such as in southern California, exhibited increased flows up to as much as two orders of magnitude (Coombs and Melack, 2013; Kinoshita and Hogue, 2015; Loáiciga et al., 2001). Fires in other chaparral environments were found to also yield flow increases, such as in South Africa (Lindley et al., 1988; Scott, 1993), Cyprus (Hessling, 1999), and France (Lavabre et al., 1993). Additional increases to post-fire flow regimes were found in temperate, forested catchments as well (Neary et al., 2005; Watson et al., 2001). A concise summary of historic changes in U.S. post-fire stream systems is found in Neary et al. (2005), documenting changes in 1st year runoffs and peak flows, encompassing a range of ecological regions. Conversely, several studies found limited or no significant changes to hydrologic systems post fire, or attributed fluctuations to natural annual variability (Aronica et al., 2002; Bart and Hope, 2010; Britton, 1991; Townsend and Douglas, 2000).

These discrepancies lead to the question of which watershed characteristics have the greatest influence over post-fire flow response? Moody et al. (2013) provide a succinct summary of soil-related theories, such as reduced infiltration due to increases in soil-water repellency, increased overland flow velocities due to increased bare ground, and reduced infiltration caused by soil-sealing. Theories commonly found in literature attribute flow changes to a wider range of factors, including reduction in interception and evapotranspiration (Lavabre et al., 1993; Scott, 1993) and increased hydrophobicity of soils (Neary et al., 2005). In regards to altered peak flows, conflicting evidence is found regarding the importance of burned watershed area with some studies finding an inverse correlation between peak flows and watershed size (Biggio and Cannon, 2001; Neary et al., 2005) and others finding no relationship at all (Bart and Hope, 2010).

The current study undertakes a comprehensive assessment of post-fire streamflow changes in the United States by examining burned watersheds that encompass a wide spectrum of climatological and geophysical parameters. A variety of flow parameters are examined which describe changes to flow regimes at several levels. Furthermore, the variability in response by distinct regions is investigated,

anticipating distinct differences influenced by regional climate. Because of the discrepancies in correlations between post-fire flow changes and watershed parameters discussed above, we identify which geophysical parameters are positively and negatively correlated with various aspects of flow regime change, as well as summarize which of the parameters account for greatest variability in response. With downstream communities at risk for flooding, and also relying on catchment runoff for water supply, investigating alterations in post-fire discharge over large scales will provide critical information for regional managers on post-fire runoff mitigation. In addition, understanding factors controlling discharge response will help inform development and calibration of surface water models used for post-fire streamflow predictions.

1.2 Study Area

A total of 82 burned watersheds were utilized in this study (Fig. 1.1), encompassing a wide range of spatial, temporal, climatological, and topographic factors (Fig. 1.2), exclusively limited to watersheds with significant wildfires (burned > 5%) and adequate (continuous 15 years daily flow) discharge records available in the USGS streamflow database (U.S. Geological Survey, 2014) identified through the GAGES-II database (Falcone, 2011). The majority of available watersheds are overwhelmingly found in the western United States, predominantly in California, Oregon, and Idaho, though several are located in the

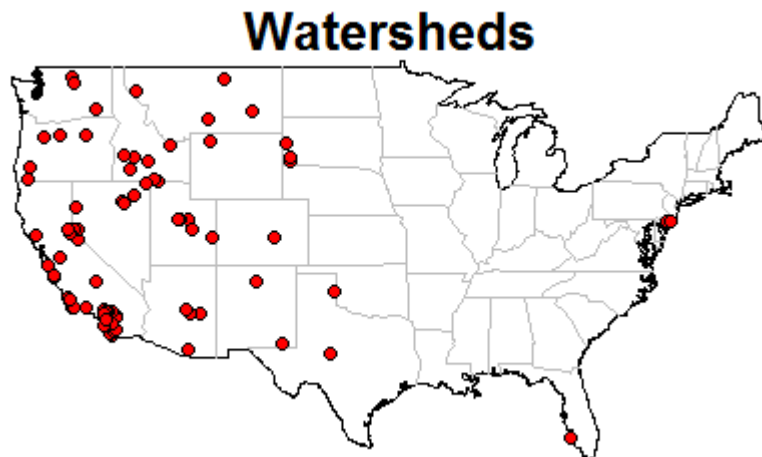


Figure 1.1. CONUS map of the locations of the 82 watersheds utilized in this study.

North East, Florida, and Kentucky. Due to discharge record and burn severity data limitations, the fires in this study cover a temporal range from water years 1984 through 2010. Average percent area burned ranges from 5-97%, with a mean of 25%, over a range of watershed areas from 4.6-9209 km². The wide spatial distribution of the studied watersheds results in mean elevations and burn-area slopes varying from 13-2760 m and 0.11-16% respectively (Fig. 1.2).

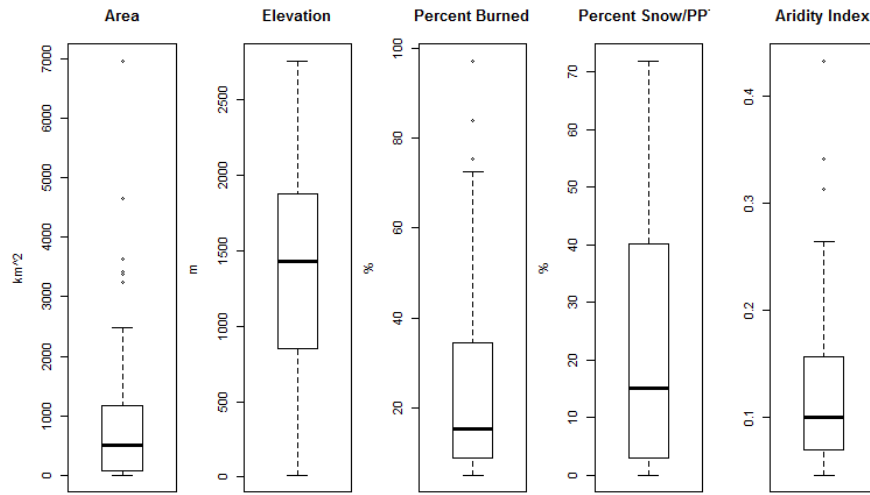


Figure 1.2. Boxplots summarizing the range of watershed areas, elevations, percent area burned, percent of precipitation that falls as snow, and aridity index.

The most important difference between many of these watersheds is the variation in climate, the values of which were collected from the GAGES-II dataset (Falcone, 2011), which catalogs all watersheds in the United States monitored by the USGS and have at least one twenty year period with continuous daily flow records. Average basin precipitation ranges from 29-220 mm/yr, with a mean of 72 mm/yr, and average temperature ranges from 1.4 -23 °C, with a mean of 10 °C. Important for identifying snow dominated regions is the percent of precipitation (PPT) that falls as snow (%Snow/PPT), which ranges from 0-72%. Relative humidity ranges from 39-73%, with a mean of 55, and potential evapotranspiration ranges from 400-1200 mm/yr, with a mean of 633 mm/yr.

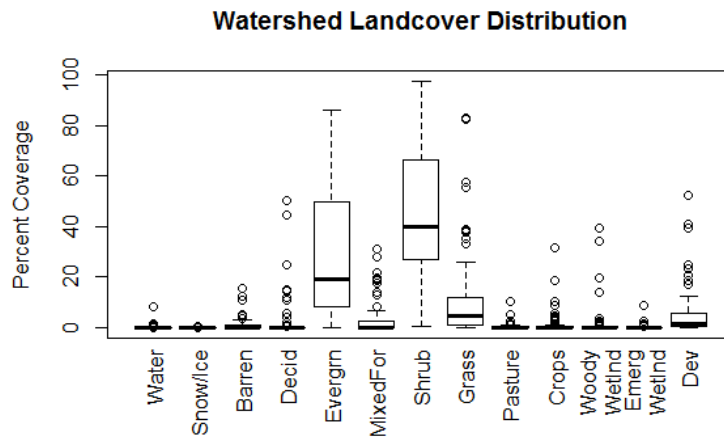


Figure 1.3. Distribution of NLCD land cover types.

Watershed vegetation types vary as well. Evergreen forest and shrub vegetation are the overwhelmingly dominant land cover type over all watersheds used in this study (Fig. 1.3). The high proportion of evergreen is due to the dominance of high elevation fires in mountainous regions and shrub prevalence is due to the abundance of watersheds found in the chaparral regions of Southern California. Grassland, mixed forest, and developed land account for a smaller proportion of land cover types. Barren land and wetland account for only a small percentage of land cover types throughout the watersheds in this study.

CHAPTER 2

METHODS

Methodology for this study includes collection and filtration of ground- and satellite-based physical and geospatial measurements, simple hydrostatistics to evaluate response, *k*-means clustering to regionalize watersheds, and LASSO regression to identify the most influential watershed qualities in regards to post-fire flow response.

2.1 Data Collection

The current study focuses on watersheds with significant historical fires, collected through the MTBS (2009) and GAGES-II (Falcone, 2011) databases. Watershed parameters (Table 2.1) utilized in this study were chosen to encompass the variability in geophysical parameters found throughout the watersheds used in this study. To identify spatial trends in post-fire response, watersheds were grouped through *k*-means clustering based on geographic and climatological data.

Table 2.1. Summary of the explanatory variables used in this study

Name	Abbreviation	Description
Burn Severity: Low	BS_L	Percent of the burn area with burn severity categorized as low
Burn Severity: Moderate	BS_M	Percent of the burn area with burn severity categorized as moderate
Burn Severity: High	BS_H	Percent of the burn area with burn severity categorized as high
KFACT	KFACT	Mean soil erodibility factor of the burn area
Area	Area	Mean area of the watershed
Slope	Slope	Mean slope of the burn area
Aspect	Aspect	Mean aspect of the burn area
Percent Watershed Burned	Percent Burned	Total percent area of the watershed burned
NDVI	NDVI	Mean NDVI of the burn area for the four years preceding the fire

2.1.1 – Watershed selection

Watersheds were selected based upon the continuity of USGS mean daily flow records preceding and succeeding the fire event. Watersheds were required to have continuous daily flow records (>95% of daily flow records accounted for in each year) for a minimum of 10 years pre-fire and 5 years post-fire. Using the approximately 9,000 watersheds in the GAGES-II dataset (Falcone, 2011), watersheds delineations were spatially cross-referenced with the MTBS database of historic wildfires (2009) . The results were again cross-referenced with USGS daily flow records (U.S. Geological Survey, 2014) to identify watersheds with the required flow records, resulting in 263 unique watersheds in the United States with greater than 5% total burn area in a single water year. Of these, 23 contained 2-3 wildfires within the same year burning over 5% of the total area. The remainder contained only a single significant fire in the year of interest. Further exclusion of watersheds was based on the presence of major dams within the watershed flow regimes extracted from the GAGES-II database (Falcone, 2011), resulting in a final collection of 82 watersheds.

2.1.2 – Hydrologic and Precipitation Data

Daily flow and peak flow data were obtained from the USGS National Water Information System (U.S. Geological Survey, 2014), pared down to a range of 10 years pre-fire and 5 years post-fire. Monthly precipitation data was collected from the PRISM database (PRISM Climate Group, 2004), a nation-wide 4 km resolution gridded monthly dataset that extrapolates station climate measurements over unmonitored areas using a complex topographic- and climate-based algorithm. Monthly national precipitation rasters were averaged for each watershed for all months within the flow record period.

2.1.3 – Topographic data

Topographic data collected included watershed area and the elevation, slope, and aspect of burn areas. Watershed area was calculated in ArcGIS while elevation, slope, and aspect were calculated

through a 30 meter resolution national Digital Elevation Model (DEM). Burn areas of watersheds were obtained by clipping each watershed delineation to the relevant fire perimeters available through MTBS. Slope and aspect functions within ArcGIS were applied to the DEM to produce unique rasters, which were then iteratively clipped to each burn area to collect average elevation, slope, and aspect.

2.1.4 – Soils Data

Soils data were collected through an adapted version of the State Soil Geographic (STATSGO) database, a national collection of over 78,000 polygons containing a host of soil characteristics (Schwartz and Alexander, 1995). Initially, data were to be collected through the 10 meter resolution gridded gSSURGO-10 database (National Cooperative Soil Survey, 2014) but large spatial gaps required the use of STATSGO. The soil erodibility factor (K_{fact}) was utilized to numerically represent average soil types, as it provides a quantitative description of a soil's erodibility:

$$K_{fact} = (1.292)[2.1E - 6 * f_p^{1.14}(12 - P_{om}) + 0.0325(S_{struc} - 2) + 0.025(f_{perm} - 3)] \quad [1]$$

$$f_p = P_{silt}(100 - P_{clay}) \quad [2]$$

where f_p is the particle size parameter, P_{om} is percent organic matter, P_{clay} is the percent clay, S_{struc} is the soil structure index, and f_{perm} is the profile-permeability class factor (Goldman et al., 1986). K_{fact} increases as the potential erodibility of a soil increases. A national raster of KFACT values was produced in ArcGIS and iteratively clipped and averaged for each burn area.

2.1.5 – Vegetation data

Vegetation data were collected for each burn area prior to the fire event. Initially, data were obtained through the 30 meter resolution National Land Cover Database (Homer et al., 2004), though several issues arose. First, NLCD data were only available for the years of 1992, 2001, 2006, and 2011, meaning that fire events occurring prior to 1993 would have to be thrown out. Additionally, the large temporal gaps between measurements would have significantly decreased the overall accuracy of any statistical methodology used. Second, due to the overall size of the watershed set being used in this study, as well as further subdivisions discussed later, a single value for vegetation would prove more accurate for statistical methods, as opposed to the many values resulting from NLCD analysis.

Due to the temporal gaps in the National Land Cover Database (Homer et al., 2004), an averaged normalized difference vegetation index (NDVI) was collected for pre-fire burn areas to quantitatively summarize vegetation, similar to previous remote sensing fire disturbance studies (Barbosa et al., 1999; Kinoshita and Hogue, 2011; Lee and Chow, 2015). NDVI is defined as:

$$NDVI = \frac{(a_{nir} - a_{vis})}{(a_{nir} + a_{vis})} \quad [3]$$

where a_{nir} and a_{vis} are surface reflectances averaged over the ranges of wavelengths in the near infrared and visible spectrums, respectively. Despite NDVI having been shown to have accuracy issues related to atmospheric interference and variations in soil brightness (Carlson and Ripley, 1997), the extended timespan over which values were being averaged may have muted any such error responses.

Average values were collected for each watershed through national 32-day NDVI rasters hosted on Google Earth Engine (GEE) (Google, Inc.), in turn calculated from Landsat5 composite satellite data freely available through the U.S. Landsat archive at the USGS Earth Resources Observation and Science (EROS) Center (C.E. Woodcock et al., 2008). 10 years of monthly NDVI data were not available for

approximately 10% of the fire events in this study, as Landsat5 imagery is only available beginning in 1984. For each watershed, mean NDVI values were calculated for 1, 2, 5, 8, and 10 years pre-fire, when available. Mean differences in values between each of the sets of results were examined and the 4 year mean NDVI was found to have limited differences (< 20%) from the ten year period values. In order to include more watersheds, mean NDVI values were produced for the burn areas of all watersheds for the 4 years pre-fire.

2.1.6 – Burn Severity data

Burn severity is the classification of burn areas relating visible changes in living and non-living biomass, fire byproducts, and soil exposure within one growing season (2009), including low, moderate, and high severity categories (Eidenshink et al., 2007). Though categorization varies by region, some generalizations can be made. Typical high severity burns result in complete kills of canopy trees and almost complete consumption of surface litter and organic soil layers (Neary et al., 2005).

Characteristics of moderate burn severity include partial canopy cover kill, completely charred or consumed understory vegetation, and widespread destruction of the soil organic layer. Low severity burns lightly scorch trees, char or consume surface litter, and produce no to little charring of the soil organic layer. Wildfires are almost always a patchwork of varying degrees of burn severity. More specifically, burn severity is the qualitative assessment of the heat pulse directed toward the ground during a fire, relating soil heating, fuel consumption, and mortality of buried plant parts (National Wildfire Coordinating Group).

Burn severity data were obtained for each unique fire in raster format through MTBS (2009), wherein pixels are designated class descriptions: 0 – background, 1 – unburned to low, 2 – low severity, 3 – moderate severity, 4 – high severity, 5 – increased greenness, 6 – non-processing area. Burn severity values were calculated as the percent coverage of total watershed area. To maintain a low total

parameter list, percentages of burn severity were limited to severity categories of Low (BS_L), Moderate (BS_M), and High (BS_H).

2.1.7 – Climatological data

Watershed climatological parameters used in this study included percent of precipitation that falls as snow (%Snow/PPT) and the aridity index (AI). The %Snow/PPT for each watershed was available through the GAGES-II dataset (Falcone, 2011) and the aridity index was calculated for each watershed as:

$$AI = \frac{P_{avg}}{PET_{avg}} \quad [4]$$

where P_{avg} is average precipitation and PET_{avg} is average potential evapotranspiration, both of which were available in the GAGES-II dataset.

2.2 Response Variables

Response values provide the means for quantifying post-fire flow changes across a variety of regimes, such as flows relating to dry seasons (low flows, base flows) and wet seasons (high flows, peaks flows).

2.2.1 – Low, high, and peak flows

Pre-fire low-flow (LF) and high-flow (HF) metrics were calculated for each of the ten years prior to the fire water year and averaged to produce a single value, using the daily flow data collected for each watershed from the USGS (U.S. Geological Survey, 2014).

Low flows (LFs) were defined as the average of mean daily flows with a 90% exceedance within a single water year, excluding zero flow days. High flows (HFs) were defined similarly, with a 10% exceedance to isolate larger volume flows (Kinoshita and Hogue, 2015). To reduce calculation bias due

to zero flow days commonly found in ephemeral stream systems, zero flow days were eliminated from exceedance value calculations. Changes in LFs and HF.s were calculated as the post-fire percent change from the average 10 water years pre-fire. Post-fire values were calculated for the 1st year (LF.one, HF.one), the 2nd year (LF.two, HF.two), and the 5 year mean (LF.five, HF.five).

Peak flows (PFs) were defined as the largest mean daily flow measurement each water year. Post-fire changes in PF were calculated as the percent change of the 1st year (PF.one), 2nd year (PF.two), and 5 year mean (PF.five) peak flow measurements from the pre-fire ten year mean. Percent changes in the number of zero flow days were calculated similarly (Nzero.one, Nzero.two, Nzero.five).

2.2.2 – Runoff Ratios

The runoff ratio (RO) of a watershed is the fraction of total annual runoff depth over total annual precipitation:

$$RO = \frac{Q_{tot}/A_{ws}}{P_{tot}} \quad [5]$$

where P_{tot} is total annual precipitation, Q_{tot} is total annual runoff depth, and A_{ws} is watershed area. RO was calculated for the ten years pre-fire and 5 years post-fire using PRISM precipitation and USGS mean daily flow data. Post-fire RO response was calculated as the percent change between the 1st year, 2nd year, and average 5 year values post-fire (RO.one, RO.two, RO.five) and the pre-fire 10 year mean.

2.2.3 – Base flow and Richards-Baker indices

Base flow index (BFI), the fraction of total streamflow that is baseflow (Baker et al., 2004), was calculated for each water year through the R package ‘hydrostats’ (Bond, Nick, 2015), that applies the Lyne-Hollick filter (V. D. Lyne, 1979). BFI response was calculated as the percent change of the 1st, 2nd, and 5 years BFI post-fire from the mean of the 10 years pre-fire BFI (BFI.one, BFI.two, BFI.five).

The Richards-Baker index (RB) quantifies the frequency and rapidity of short-term changes in streamflow (flashiness) based on daily flow data through the equation:

$$R - B \text{ Index} = \frac{\sum_{i=1}^n |q_i - q_{i-1}|}{\sum_{i=1}^n q_i} \quad [6]$$

where q is mean daily flow, t is time, and q is daily flow (Baker et al., 2004). RB response was calculated as the percent change from the average RB over the ten years pre-fire to the RB of 1 year, 2 years, and average of 5 years post-fire (RB.one, RB.two, RB.five).

2.3 *k*-means Clustering

In order to regionalize differences in post-fire flow response and create region-specific regression models, watersheds were classified into unique regions through *k*-means clustering (MacQueen, 1967). This method partitions an N-dimensional population of observations into clusters with minimal variation, allowing for relatively simple similarity grouping. For the current study, the ideal ensemble of clusters was one that produced easily recognizable regions with unique climatological characteristics. Large-scale clustering methods have been applied in prior watershed classification studies, but utilized more complex streamflow and ecological indices as parameters (McManamay et al., 2014; Poff, 1996).

Wildfires in this study are typically found in western evergreen and shrub environments, so clustering by only these watersheds would likely produce regions biased by fire occurrence. To limit this, we applied the *mclust* package in R (Fraley et al., 2012) to cluster over 9,000 GAGES-II watersheds to produce national regions. The *mclust* package was chosen over the standard *kmeans* function in R due to its inclusion of numerous model-based approaches and application of the Bayes Information Criteria (BIC) to determine the most accurate model and cluster count (Schwarz, 1978). Various

groupings of simple parameters were used for clustering including watershed latitude and longitude, elevation, AI, %Snow/PPT, and mean monthly and seasonal flow statistics.

2.4 LASSO Regression

Linear regression models were fit for each grouping of watersheds through the least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996). This method limits the number of standardized explanatory variables through penalization of the absolute value of the regression coefficients by:

$$\hat{\alpha}, \hat{\beta} = \arg \min \left\{ \sum_{i=1}^N (y_i - \alpha - \sum_j \beta_j x_{ij})^2 \right\} \quad \text{subject to } \sum_j |\beta_j| \leq t \quad [7]$$

$$\min_{\beta_0, \beta} \frac{1}{N} \sum_{i=1}^N w_i l(y_i, \beta_0 + \beta^T x_i) + \lambda \left[\frac{(1-\alpha) \|\beta\|_2^2}{2} + \alpha \|\beta\|_1 \right] \quad [7]$$

where x_i are the predictor variables, y_i are the explanatory variables, β_i are the coefficients, and λ is a tuning parameter that controls the overall strength of the coefficient penalty, applied in R through the ‘glmnet’ package (Hastie et al., 2013). As λ increases, select β s are driven to zero. This method differs from ridge regression in that rather than shrinking the β 's of correlated explanatory variables towards each other, the LASSO picks one and discards the others (Condon et al., 2015). Thus, identification and exclusion of correlated predictors is not a required step in this regression process, significantly decreasing user workload. This has the added benefit of producing a sparse matrix of β values, producing $\beta=0$ for non-contributing x 's. The optimal values of λ and β_i are determined through minimization of the mean squared error (MSE) for this study. Prior to modeling, explanatory and predictor (response) variables were standardized by:

$$\frac{(X_i - \mu_X)}{\sigma_X} \quad [8]$$

where X_i is the observation, μ is the mean of all X observations, and σ is the X standard deviation. This study utilized the explanatory variables shown in Table 1. Accuracy of LASSO regression models is then determined through the coefficient of determination (R^2) by:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad [9]$$

where SS_{res} is the sum of squared errors and SS_{tot} is the total sum of squares. This fraction quantitatively represents the response variable variation explained by the model. Linear regression models utilized in the trend analysis discussion section will use adjusted R^2 , which alters the standard R^2 by penalizing increases in total number of explanatory variables relative to sample size through:

$$Adj R^2 = 1 - \frac{SS_{res}/(n-K)}{SS_{tot}/(n-1)} \quad [10]$$

where n is sample size and K is parameter count. This common statistic may not be applicable to LASSO model results as the algorithm is inherently biased in favor of only significant explanatory variables. LASSO regression was applied through a Monte Carlo simulation with $n = 500$ for each group and response variable. The optimal model was chosen by highest R^2 value.

CHAPTER 3
RESULTS AND DISCUSSION

Results demonstrate that flow response is correlated to percent area burned, but is significantly influenced by several other watershed qualities. Grouping watersheds by climate parameters shows that response is much more variable and greater in certain regions. Watershed area, slope, and the soil erodibility factor appear to have significant control over post-fire flow response.

3.1 Clustering

The *k*-means clustering performed on the GAGES-II watershed set yielded 9 clusters or regions (Fig. 3.1). The most important clusters for the current study are 6 through 9, which assemble 77 of the 82 watersheds into unique regions (Fig. 3.2). These four significant clusters have unique characteristics (Fig. 3.3). Watersheds in cluster 6, on average, have the highest K_{fact} , though almost all are burned less

Clustered GAGES-II Watersheds

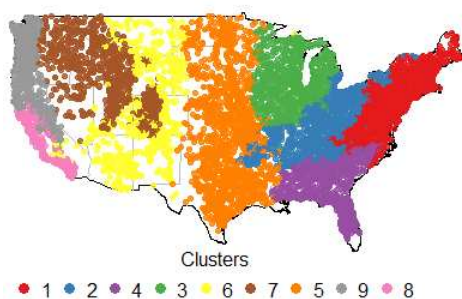


Figure 3.1. CONUS map of the results of *k*-means clustering of the GAGES-II watershed set.

Study Watersheds by Cluster

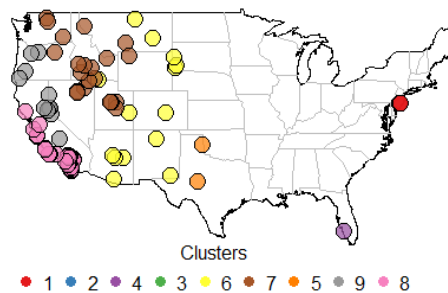


Figure 3.2. CONUS map of the results of watershed clustering.

than 20%. They have relatively moderate %Snow/PPT values and the lowest AI values. Watersheds in cluster 7 have the highest average area and elevation, and accordingly the highest %Snow/PPT and the lowest NDVI. Cluster 8 contains watersheds with the widest range of relative fire sizes, including watersheds burned from as little as 10% to as great as 97%. Watersheds in cluster 8 also have the

lowest average elevations and areas, as well as the smallest %Snow/PPT and low AI. The percent of the burn area rated as high burn severity is also the greatest on average in cluster 8 watersheds. Cluster 9 watersheds have the lowest KFACT and highest elevations. These watersheds also have high %Snow/PPT and AI.

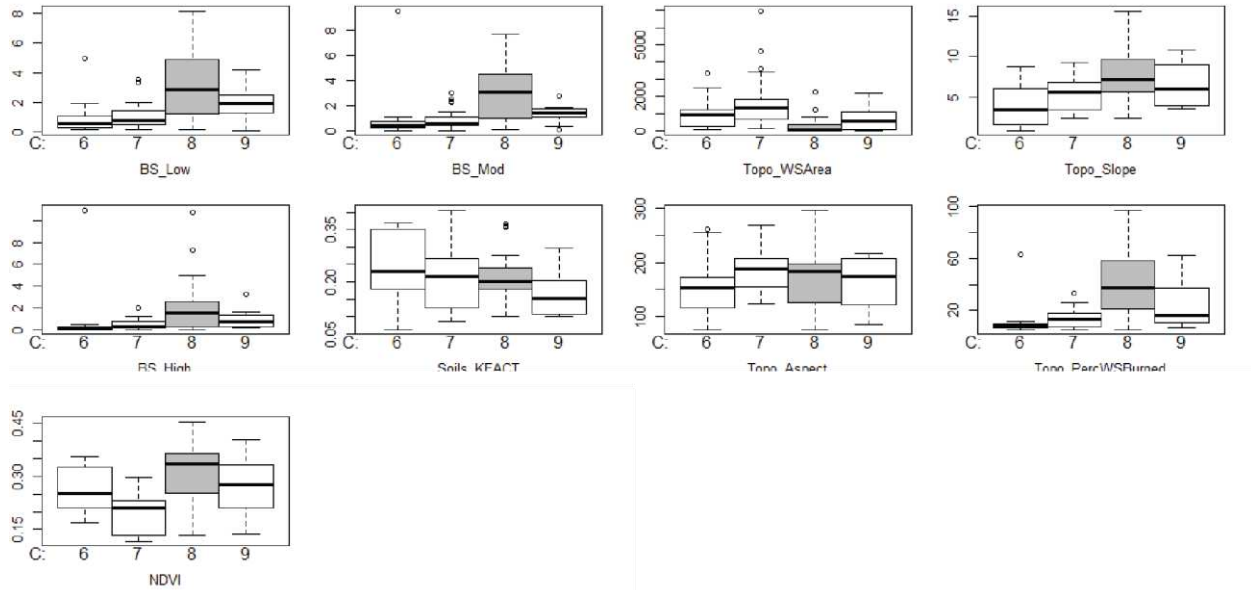


Figure 3.3. Variations in explanatory variables by cluster.

3.2 Response Variable Distribution and Analyses

Calculated flow response variables indicate an extremely wide range of post-fire system responses (Fig. 3.4). The greatest ranges occur within variables representing changes in low flows, such as Nzero.one (st. dev = 243%), BFI.one (236%) and Nzero.five (201%). The tightest ranges are typically found within mean five year variables where extreme changes are muted, such as RB.five (28%), RO.five (39%), and BFI.five (48%). Response variable means range from as low as 0.92% (RB.five) to as great as

115% (Nzero.one). Due to the nature of response calculations (section 2.1), Nzero values were limited. Nzero.one and Nzero.two were found for 22 watersheds and Nzero.five was found for 27 watersheds.

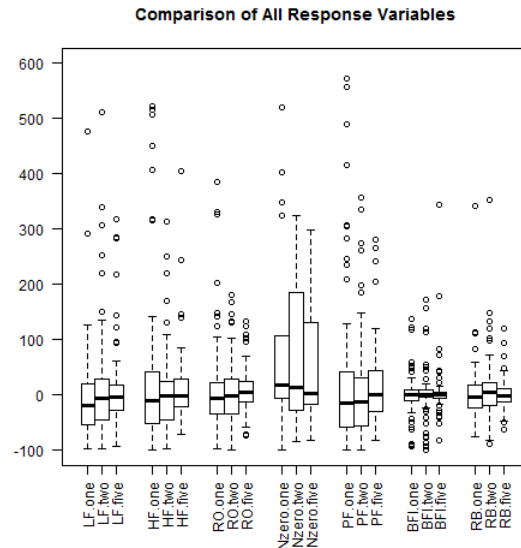


Figure 3.4. Boxplots of the response variables utilized in this study.

3.2.1 - Trend analysis

Comparison of 1st year response variables to 5 year mean variables generally produce trendline slopes greater than one, with a mean slope of 1.9 (Fig. 3.5). The greatest trendline slopes are found in LF.one, HF.one, and BFI.one. BFI.one is unique in the magnitude of its slope versus BFI.five, which is twice as large as that of LF.one. Only Nzero.one versus Nzero.five yields a slope less than one.

Trendlines of 2nd year response variables versus 5 year mean variables produce significantly lower slopes, with a mean of 0.78. Only LF.two and RB.two yield greater values. We can infer from this that the greatest increases in these response variables in the five water years following a fire are found in the first year. Additionally, only LFs exhibit greater response in the second year than in the following three years.

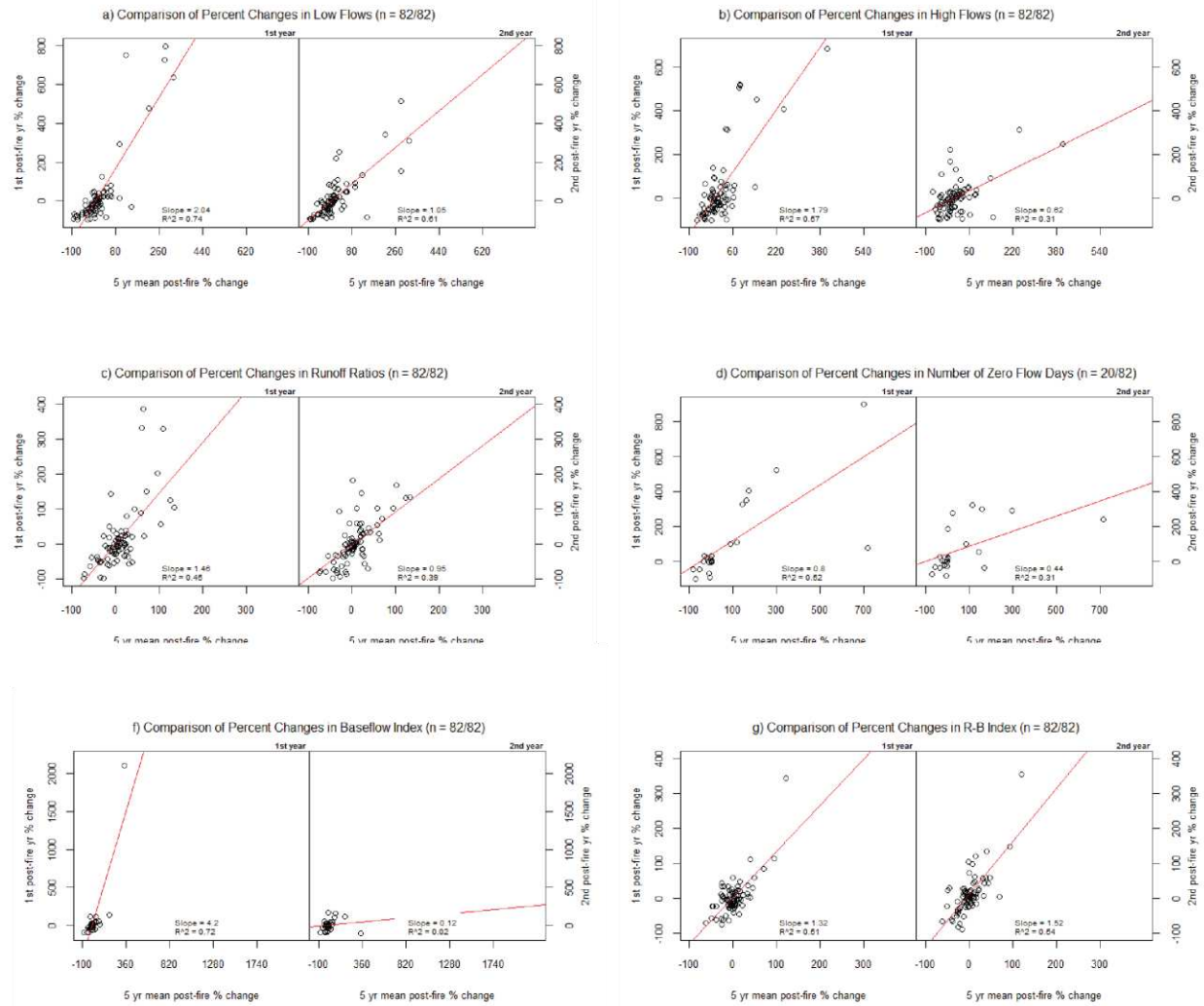


Figure 3.5. Scatterplots of 1st and 2nd year response variables versus 5 year mean response variables.

Comparing response variables to percent area burned, trendlines demonstrate the greatest slopes in 1st year responses for LF, HF, and PF variables (Fig. 3.6). In the case of RO, trendlines exhibit similar increases of RO.one and RO.two. Typically, 2nd year trendlines tend to be steeper in slope than 5 year mean values. The exception to this is Nzero.two, where 2nd year values decrease significantly with increased fire size. Overall, Nzero is found to increase post-fire, though due to both a small sample size and a short time period, results are most likely uncertain. BFI increases with increasing burn size,

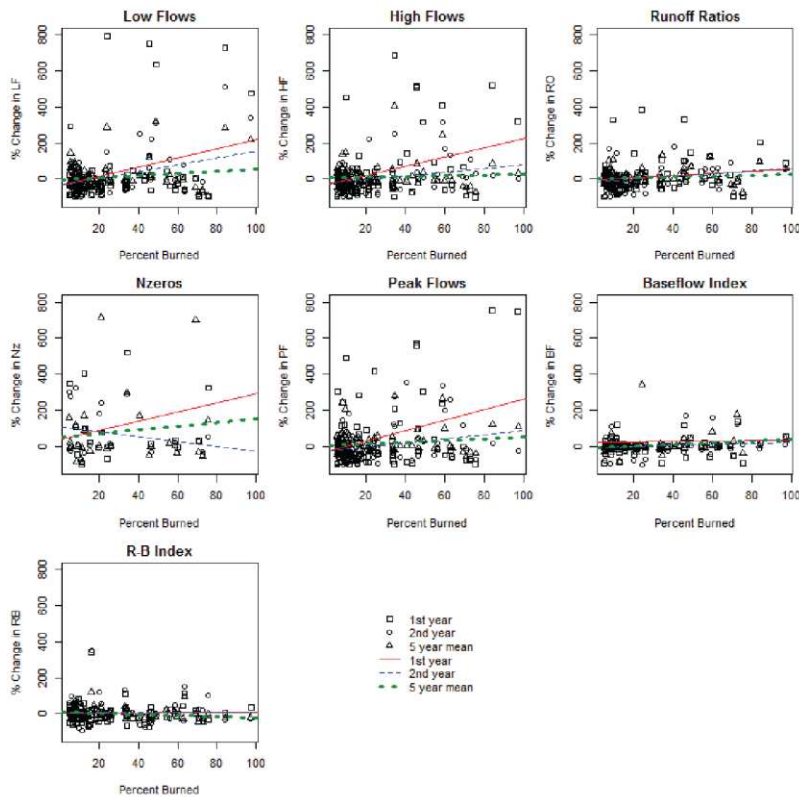


Figure 3.6. Scatterplots of all response variables versus respective watershed percent area burned. Simple regression lines provided to show correlations.

though to a lesser extent than the previously mentioned variables. Only RB indicates little linear correlation to burn area, with marginal 1st and 2nd year slopes. In fact, RB.five decreases with increasing percent area burned. Overall, these findings confirm previous smaller-scale studies in which the greatest flow responses occur immediately after fire events and decrease with time.

3.2.2 - Response variability by cluster

CONUS plots of the spatial distribution of post-fire watershed response are difficult to understand and do not provide adequate detail of results. To simplify an analysis of response variables, boxplots are provided comparing responses across the four clusters (Fig. 3.7). In this instance and that of the CONUS plots, all variables are scaled by dividing the variable by the percent burn area of the watershed in order to show relative response.

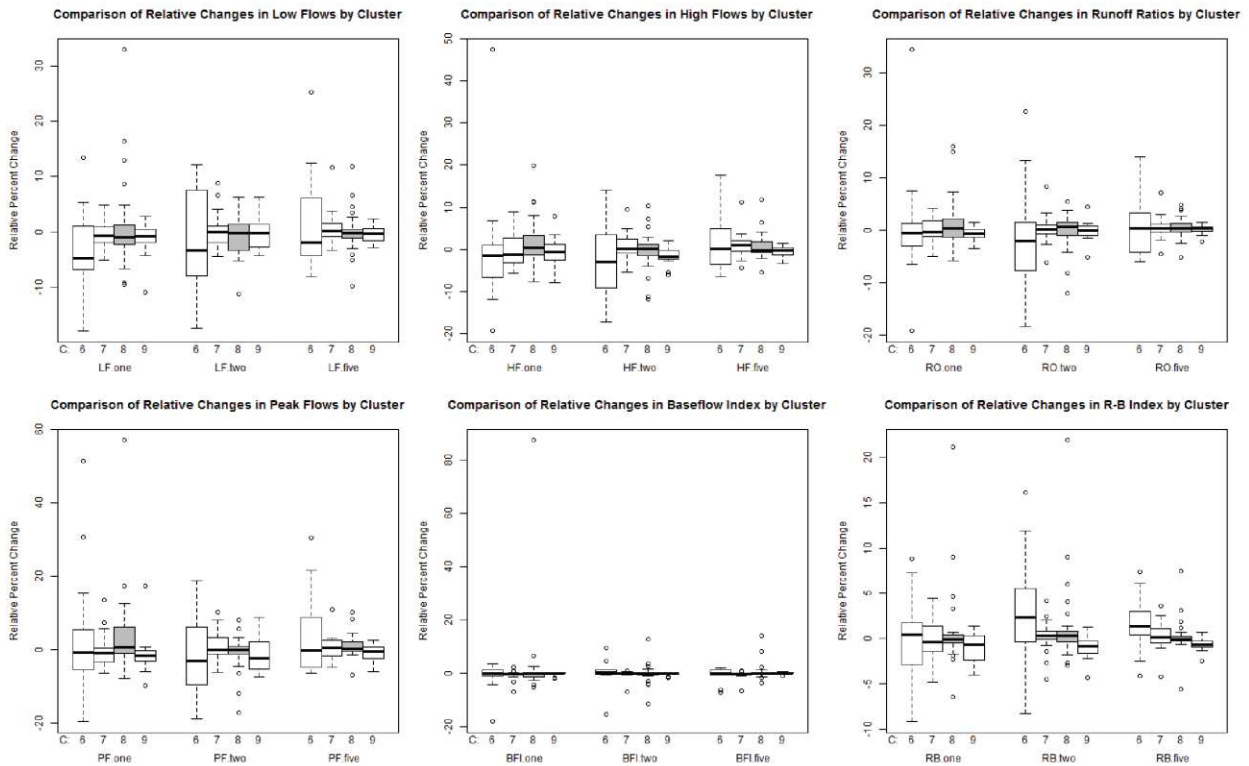


Figure 3.7. Boxplots of relative response variables by cluster.

Cluster 6 demonstrates the greatest overall variability in relative response variables. In the case of LFs, standard deviation of the three time periods (1st year, 2nd year, and 5 year mean) averages 8.6%. The magnitude of this value is pronounced when compared to the standard deviation of other cluster LF ranges, the next largest of which is 5.7% (cluster 8). Similarly high variability of cluster 6 values in the remaining response variables, especially in HFs (mean st. dev = 10.5%), ROs (mean st.dev = 9.3%), and

PFs (mean st.dev = 13.4%). Despite the high variability, cluster 6 typically demonstrates the overall lowest mean responses, such as in the cases of LFs, HFs, and BFIs. In regards to the remaining response variables, mean responses rarely deviate from the range of the other clusters.

Variability in response variables is similar between clusters 7 and 9. Standard deviations of each response variable is typically very similar and in almost all cases, lower than those found in clusters 6 and 8. Mean responses, on the other hand, are very different. Mean relative responses in cluster 9 are generally negative, while those in cluster 7 are positive.

Watersheds in cluster 8 differ in overall magnitude with respect to both the standard deviations and the mean values of the other clusters. Standard deviation within relative response variables is significantly higher than that of clusters 7 and 9, though lower than that of cluster 6. Notable instances of this are observed in the cases of PF.one (st. dev = 11.7%) and BFI.one (16.4%). Mean values are the second largest among the clusters, close in magnitude to cluster 6.

Variability in response variables generally decreases with increasing percent watershed burned.

Linear regression modeling of response variables by percent watershed burned yields the error statistics

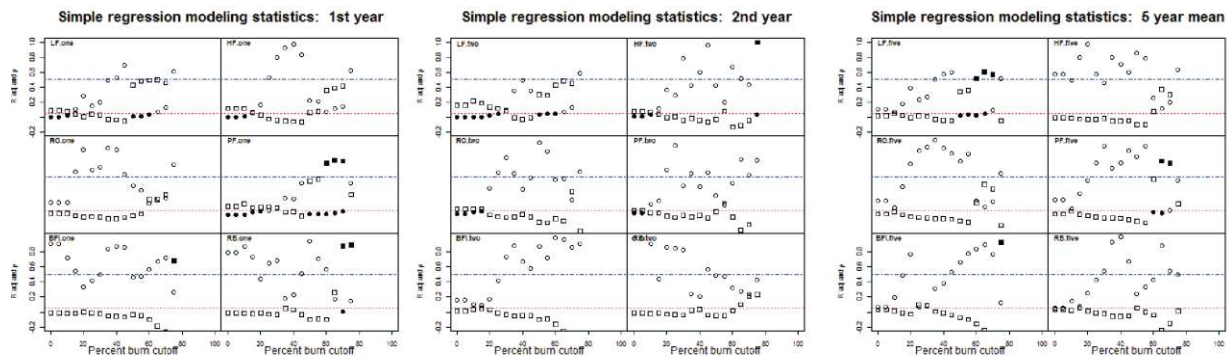


Figure 3.8. Scatterplots of the R^2 and p-value results of simple linear regression modeling of response variables versus an increasing percent burn limitation. Squares are R^2 and circles are p-values. Filled squares are $R^2 > 0.50$, filled circles are $p \leq 0.05$. The blue dashed line is $R^2 = 0.50$, the red dotted line is $p = 0.05$.

found in Figure 3.8 (i.e. decreasing sample size with increasing percent area burned). Figures include the adjusted R^2 and p-value significance tests ($\alpha = 0.05$).

Generally, linear modeling of 1st year response variables increases in accuracy as included values are limited by increasing percent burn area. LF.one, HF.one, PF.one, and RB.one show substantial increases in adjusted R^2 once included watersheds are limited to those exceeding a 50% burn area ($n=12$). Included p-values indicate that several of the LF.one and PF.one models are statistically significant.

Applying the same methods to 2nd year values produces dissimilar results, with adjusted R^2 exceeding 0.5 in only a single instance (HF,two), and few significant p-values. However, in the cases of LF.two and RB.two, R^2 values increase with increasing percent burn threshold. Unsurprisingly, simple regression of 5 year mean values versus percent area burned produce mixed results with only LF.five and PF.five allowing for adjusted R^2 values greater than 0.5, few of which are statistically significant. BFI.five shows a single instance of a high adjusted R^2 value but is statistically insignificant. Simple regression modeling was also performed on response variables by limiting included watersheds by decreasing percent area burned (i.e. decreasing sample size with decreasing percent area burned) and results demonstrated zero significant adjusted R^2 values.

In the next section we investigate which watershed parameters have the greatest statistical control over watershed response variables. Findings so far suggest that for this study, 1st year values will most likely yield LASSO regression models with the highest R^2 .

3.3 Regression results

LASSO regression was performed on 8 groups of watersheds for all response variables, for two categories: 1) all of the watersheds progressively limited by their percent area burned, hereby referred to as burn-area limited groups. Group 1 contains all watersheds in the study and Groups 2 through 4

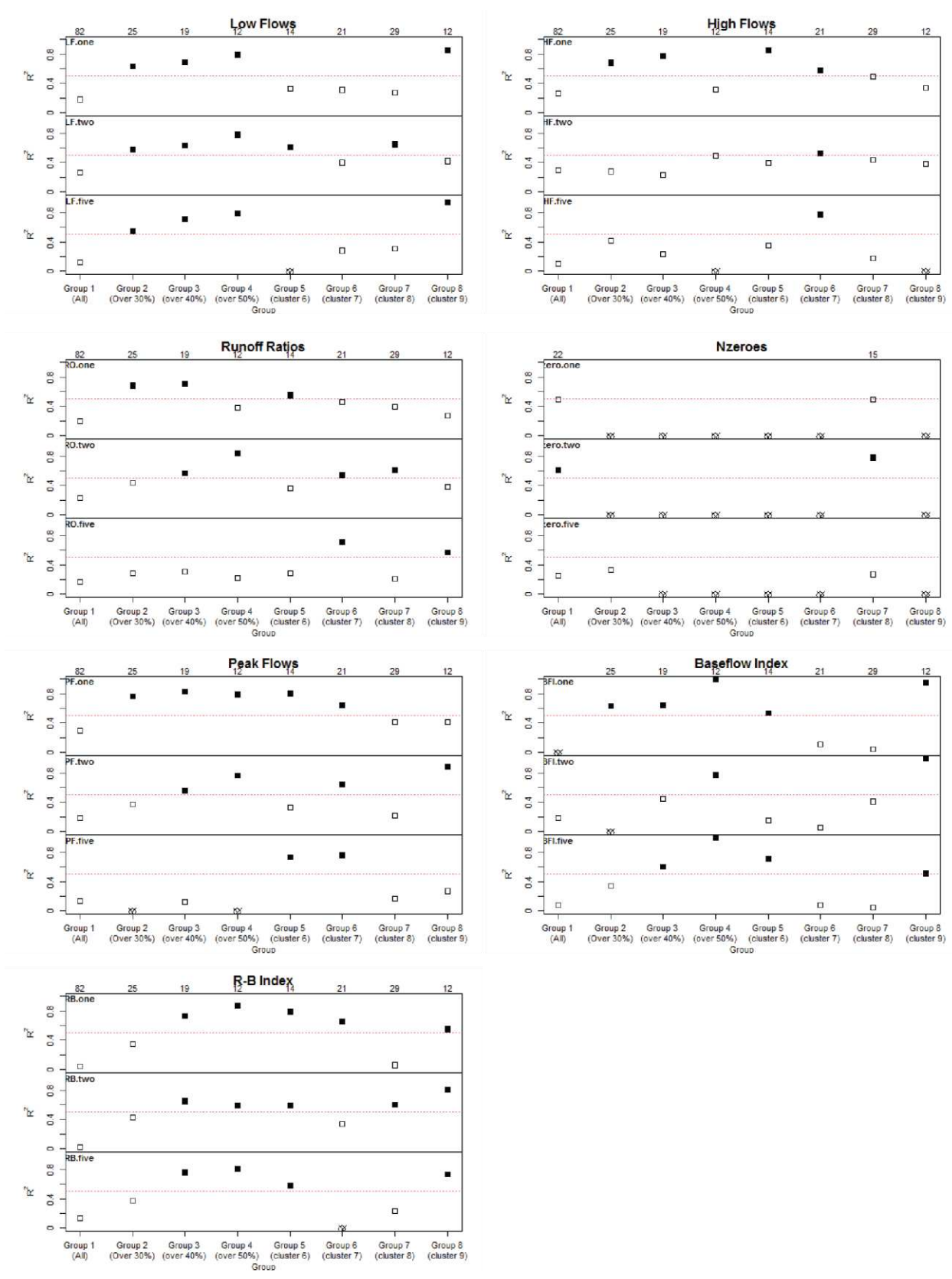


Figure 3.9. Error statistics of the results of LASSO regression modeling of response variables. Filled squares are models with $R^2 \geq 0.50$.

contain watersheds limited to burn areas exceeding 30, 40, and 50%; and 2) groups 5 through 8 correspond to clusters 6 through 9, hereby referred to as clustered groups. Figures summarizing the error statistics can be found in the Figure 3.9. To relate coefficient (β) results, heatmaps were produced of both the β values (β_V) and β ratios (β_R). β_R 's are the absolute values of each coefficient in a model divided by the sum of the absolute value of all coefficients in the same model, effectively representing how explanatory variables account for the variability in response (Condon et al., 2015). A full discussion of each model's results would be overly lengthy and time-consuming. For the sake of brevity, this discussion will cover the error statistics of significant regression models. Because of the large number of groups and response variables, mention of coefficients of individual models will be brief.

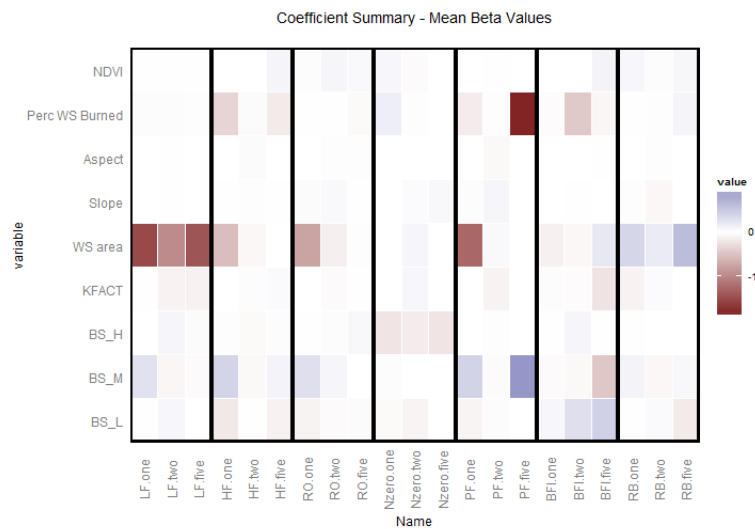


Figure 3.10. β_w values for all response variables.

To summarize, weighted mean β_R 's (β_w) are calculated by multiplying the logistic value (β_L) for each β_V ($\beta_V > 0$ { $\beta_L = 1$ }, $\beta_V < 0$ { $\beta_L = -1$ }) by the β_R for all models, again in turn multiplied by the model's ratio R^2 value, essentially using the model accuracy statistic as a means to weight the β_R results. Resulting values are then summed across all models for each response variable (Fig. 3.10). This helps to condense and simplify the complicated results into a single plot. Converting the resulting β_w 's to a logistic format further simplifies results (Fig. 3.11). Due to a small sample size ($n < 30$ for most groups),

resulting β values trend towards high variability. Results and discussion will focus on general trends and abnormal outcomes.

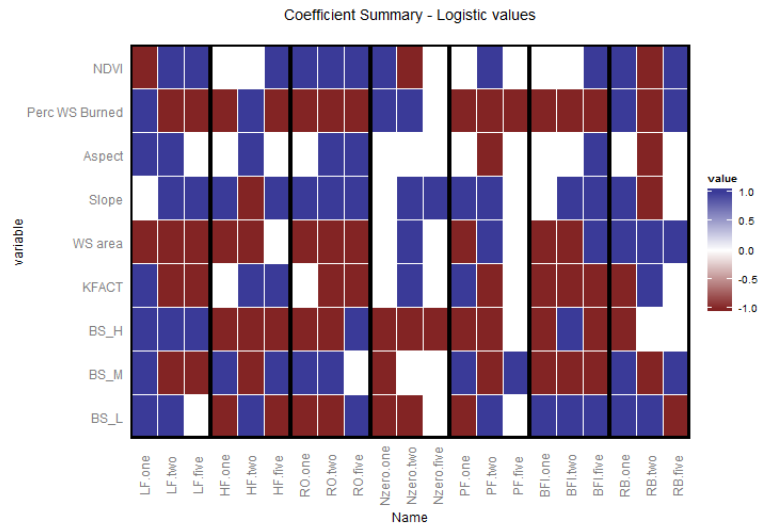


Figure 3.11. β_w in logistic format (if $\beta_w > 0$, $\beta_w = 1$. if $\beta_w < 0$, $\beta_w = -1$).

3.3.1 - Low flows

R^2 values greater than 0.50 were produced for burn-area limited groups 2 through 4 for the LF.one, LF.two, and LF.five periods. Regression of group 1 consistently resulted in unsatisfactory R^2 values much less than 0.50. LASSO regression results of clustered groups 5 through 8, corresponding to clusters 6 through 9, were more variable. Only the regression model of cluster 9 yielded a significant R^2 value for LF.one. Satisfactory models were produced for both clusters 6 and 8 for LF.two, and only for cluster 9 for LF.five. Overall, the most accurate models were produced for the LF.one period.

β_v 's show a strong negative correlation between watershed area and LF.one and a strong positive correlation between BS_M and LF.one. K_{fact} is typically weakly negatively correlated with LF.one. Interestingly, BS_H is somewhat weakly negatively correlated as well. β_r 's show that the most significant coefficient for LF.one is, by far, watershed area. Burn severities, K_{fact} , and total percent area burned also account for significant proportion of LF.one variability.

β_V 's for LF.two demonstrate similar results as with LF.one, with a negative correlation for watershed area and positive correlation for BS_M, as well as a increases in positive correlation for BS_L and BS_H. Again, K_{fact} is weakly negatively correlated with LF.two. Slope is shown to demonstrate a weak positive correlation with LF.two. β_R 's for burn-area limited groups show results comparable to those of LF.one, with a propensity towards watershed area dominating LF.two variability and moderate proportions of burn severities and K_{fact} . β_R 's for cluster groups are much more highly skewed towards increased ratios of burn severities and K_{fact} , as well as slope and NDVI in cluster 8. In LF.two results also show a small increase in relative importance of slope coefficients not found in LF.one.

β_V 's resulting from LF.five modeling yield similar results to LF.one and LF.two, with watershed area and K_{fact} showing negative correlation and BS_M showing positive correlation. β_R 's for LF.five yield results more consistent with those of LF.one than LF.two, with watershed area, burn severity, and K_{fact} parameters dominating variation.

β_W summaries show the overwhelming negative dominance of watershed area to LFs and the positive correlation between BS_M and LFs. The heatmap of logistic β_W 's show that watershed area, K_{fact} , and, surprisingly percent area burned, are typically negatively correlated with LFs. Burn severities, slope, and aspect are generally positively correlated with LFs.

3.3.2 - High flows

Modeling results of HFs are similar in variability to those of LFs. Regression models of burn-area limited groups for HF.one were satisfactory for groups 2 and 3. Unlike the LF.one models, group 4 did not produce significant HF.one results, despite having the strictest burn-area limitations placed on it. However, of the burn-area limited groups, only modeling of group 4 produced an R^2 closeto 0.50 for HF.two. None of these groups yielded statistically significant regression models for HF.five. Cluster models for clusters 6 and 7 were significant for HF.one, with cluster 8 yielding an R^2 of 0.49. Only cluster

7 yielded significant results for HF.two and HF.five, though all clusters resulted in $R^2 > 0.38$. We can infer from these results that post-fire high flows are highly variable and, likely to sample size, difficult to accurately model through multiple regression. However, it is clear that the most accurate models are those produced for HF.one, similar to the results of LF.one.

β_V 's show that BS_M and slope are positively correlated with HF.one. BS_L, BS_H, and watershed area are negatively correlated with HF.one. β_R 's show the variation in respective importance of watershed parameters, though burn severities almost universally account for greater than 50% of HF.one variation for models exceeding the R^2 threshold of 0.50. Other significant parameters include watershed area and percent area burned.

β_V 's for HF.two are more difficult to interpret, given the lack of significant models. By including results of models with $R^2 > 0.40$, we can see that watershed area and BS_H are negatively correlated with HF.two, while slope and aspect are positively correlated. Similarly, β_R values are highly variable, though generally burn severities and slope dominate watershed parameters, with an increase in burn area aspect importance not previously found in LFs and HFs.

β_V 's show positive correlation between HF.five and BS_M, slope, aspect, and NDVI, and negative correlation with BS_H and percent area burned. β_R again show a propensity towards high importance of burn severity variables, as well as percent area burned, slope, and K_{fact} .

β_W summaries indicate that, overall, watershed area, percent area burned, and BS_H are negatively correlated with HFs, while BS_M and slope are positively correlated.

3.3.3 - Runoff ratios

Results of regression modeling of the RO response variables followed trends similar to LFs and HFs. For instance, significant models were produced for groups 2 and 3 for both RO.one, and groups 3

and 4 for RO.two, typically exceeding R^2 values of 0.60. However, none of these groups produced significant results for RO.five. Modeling results of the clustered watersheds were more inconsistent. Only the regression model of cluster 6 was found to be significant for RO.one, though clusters 7 and 8 showed R^2 results > 0.40 . RO.two and RO.five models were significant for clusters 7 and 8 and clusters 6 and 9, respectively. The overall increased accuracy of regression models during the RO.two period indicates that the most predictable changes in RO may occur later in the post-fire period than those of HF and LF.

β_V 's for RO.one indicate a strong negative correlation with watershed area, BS_L, and BS_H, and a strong positive correlation with BS_M and a weak positive correlation with slope and NDVI. β_R 's show that burn severity variables account for greater than 50% of the RO.one variability, with watershed area and slope accounting for the majority of the rest.

β_V 's for RO.two indicate significant negative correlation with watershed area, K_{fact} , BS_H, and BS_L. Positive correlations are found with slope, NDVI, BS_M, and aspect. β_R 's show that burn severities, K_{fact} , watershed area, slope, and NDVI account for significant portions of RO.two variability. However, results are highly variable between groups.

Regression results for RO.five are contradictory to those of earlier time periods, with β_V values showing positive correlations for watershed area, BS_L. NDVI, aspect, and slope are positive as well. β_R values reveal very different proportions of variable contribution by group, though burn severities and watershed area are the dominant variables.

3.3.4 - Number of zero flow days

Modeling results of Nzero was very poor, with the exception of groups 1 and 7, in which adequate models were produced for Nzero.one and Nzero.two. However, due to the relatively small sample size and overwhelmingly poor modeling results, these will be disregarded for future discussion.

3.3.5 - Peak flows

LASSO modeling of peak flows produced the greatest overall number of statistically significant regression models, especially in regards to the 1st and 2nd year response variables. Regression models with R^2 values greater than 0.50 were produced for PF.one for burn-area limited groups 2 through 4, and for groups 3 and 4 for PF.two. Group 1, containing all watersheds, yielded poor models. All models of these burn-area limited groups were insignificant for PF.five, indicating that the only predictable PF responses occur within 1-2 years of a wildfire. Modeling of PFs by cluster also yielded significant results. Clusters 6, and 7 yielded high R^2 values for PF.one, while clusters 8 and 9 values were close to 0.50. Clusters 7 and 9 yielded significant R^2 values for PF.two, and clusters 6 and 7 yielded significant results for PF.five.

β_V 's show that watershed area, BS_L, and BS_H are negatively correlated with PF.one, and BS_M is positively correlated. β_R 's show that burn severities, watershed area, and percent area burned overwhelmingly account for the majority of PF.one variability, with very limited contributions from K_{fact} , NDVI, and slope.

β_V 's for PF.two show negative correlations with all variables but slope. β_R 's indicate that burn severities, K_{fact} , slope, area, and aspect account for most of the PF.two variability. PF.five is negatively correlated with area burned and BS_H, and positively correlated with BS_M. Burn severities and area burned account for the majority of variability.

3.3.6 - Baseflow index

Regression modeling results of post-fire changes in BFI were overwhelmingly poor, with only a single significant model (Cluster9 in BFI.two), unsurprising given the trend analysis discussion in section 2.1. This strongly suggests that, in the case of these watersheds, there is little to no correlation between wildfires and baseflow within the five year period post-fire. Regression modeling results of post-fire

changes in BFI were exceptionally strong, especially in 1st year values. Burn-area limited groups 2 through 4 yielded high R^2 values, while only groups 3 and 4 were significant for BFI.two and BFI.five. Regression modeling of clusters 6 and 9 yielded significant R^2 values for BFI.one and BFI.five, and cluster 9 was significant for BFI.two. Outstanding models include those of cluster 9 for BFI.one and BFI.two, where R^2 values were greater than 0.95.

β_v 's for BFI.one are somewhat variable, though overall trends are similar to those of previously discussed response values. Generally, BS_M and BS_L are positively correlated with BFI.one, while watershed area and percent area burned are negatively correlated. β_R 's are much more consistent, showing that burn severities, watershed area, and percent area burned account for the majority of BFI.one variability. β_v 's for BFI.two show that BS_L and BS_H are positively correlated with response, while watershed area, percent area burned, K_{fact} , and BS_M are negatively correlated. β_R 's for BFI.two are similar to those of BFI.one. β_v 's for BFI.five are difficult to interpret due to the multitude of different results, but general trends indicate positive correlation between response and BS_L, BS_H, NDVI, and aspect. BS_M, slope, and K_{fact} are negatively correlated. Again, burn severity values account for a large portion of the variability, with K_{fact} , watershed area, slope, and aspect accounting for more of the rest.

3.3.7 - R-B flashiness index

LASSO regression of RB response variables yielded several significant models. Many of these models yielded accurate results, occasionally exceeding an R^2 of 0.75. Several other models, notably in the RB.two plot, yielded error statistics very close to the cutoff limit. Burn-area limited groups 3 and 4 allowed for the best modeling during RB.one, RB.two, and RB.five. Clusters 6 and 9 yielded statistically significant models for all RB periods. In addition, models for cluster 7 and cluster 8 were significant for RB.one and RB.two, respectively, though they demonstrated relatively poor results for the remaining time periods.

β_V 's for RBs are highly variable. For RB.one and RB.two, several variables represent both positive and negative correlations with response, depending on group. Similarly, β_R 's show little continuity in results between groups. The only consistency is found in RB.five, where percent area burned, watershed area, and BS_M are positively correlated with response, while BS_L and BS_H are typically negatively correlated. In groups where NDVI is significant, it is positively correlated with response. Overall, β_R 's indicate that burn severities are consistently strong indicators of response, while remaining variables fluctuate in relative strength by group.

3.4 Discussion

In summary, simple regression analysis shows that percent area burned is rarely the only significant indicator of post-fire flow changes. Watershed area and burn severity parameters generally account for the greatest proportion of flow response variability. Results further indicate that slope and K_{fact} also contribute significantly. Watershed area, frequently the strongest control over post-fire response, is strongly negatively correlated with LFs, HFs, and PFs, but strongly positively correlated with RBs. Slope, aspect, and BS_M are positively correlated with almost all response variables. K_{fact} is typically negatively correlated with most response variables.

Interestingly, BS_H and BS_L are generally negatively correlated with response variables. This may be due to the actual pattern of burning that occurred within the fire perimeters which were later lineated in MTBS. Within most fire perimeters used in this study, a very high percentage of the burn area is actually designated as background or unburned. Perhaps many of the fires with a relatively high proportion of burn severity rated "high" are in fact mostly unburned elsewhere within the perimeter. The common negative correlation of total percent area burned with many response variables may also be due to this, in that total actual burned area may be significantly lower than the percent area burned variable in this study indicates. The negative correlation of K_{fact} with many response variables is also

intriguing, implying that as the erodibility of a soil increases, post-fire response becomes more muted. This is possibly because less erodible soils maintain their post-fire compressed structures and hydrophobic layers for a longer period of time, thereby allowing for greater increases of several types of flow response.

CHAPTER 4

Conclusions

Many of the results summarized in the discussion complement prior smaller-scale studies, providing further evidence of the importance of watershed area, slope, and burn severity. However, the negative correlation of K_{fact} with flow response is contradictory to the authors prior assumptions. While the scale of this study is broader than most previous research on the subject, the sample size is in actuality quite small. Furthermore, the soils database used here is very low resolution, which may be a significant factor in the K_{fact} results. Though the regression models produced in this study indicate typically negative correlation of the soil erodibility factor with flow response, future research will help to clarify and substantiate these theories.

Results from this study may help water resources organizations develop post-fire water budgets, as well as highlight regions where emergency services may be required to increase their funding and preparedness following significant fires. The LASSO regression, which helped to identify the relative strengths of watershed parameters in affecting flows, will help to both calibrate and justify post-fire watershed models, most importantly the Precipitation Runoff Modeling System (PRMS) for which this project was first implemented.

REFERENCES CITED

- Aronica, G., Candela, A., and Santoro, M. (2002). Changes in the hydrological response of two Sicilian basins affected by fire. (IAHS Press), pp. 163–169.
- Baker, D.B., Richards, R.P., Loftus, T.T., and Kramer, J.W. (2004). A New Flashiness Index: Characteristics and Applications to Midwestern Rivers and Streams. *JAWRA Journal of the American Water Resources Association* 40, 503–522.
- Barbosa, P.M., Stroppiana, D., Gregoire, J.M., and Pereira, J.M.C. (1999). An assessment of vegetation fire in Africa (1981-1991): Burned areas, burned biomass, and atmospheric emissions. *Glob. Biogeochem. Cycle* 13, 933–950.
- Bart, R., and Hope, A. (2010). Streamflow response to fire in large catchments of a Mediterranean-climate region using paired-catchment experiments. *Journal of Hydrology* 388, 370–378.
- Benavides-Solorio, J., and MacDonald, L.H. (2001). Post-fire runoff and erosion from simulated rainfall on small plots, Colorado Front Range. *Hydrol. Process.* 15, 2931–2952.
- Biggio, E.R., and Cannon, S.H. (2001). Compilation of post-wildfire runoff data from the Western United States (U.S. Geological Survey).
- Bond, Nick (2015). hydrostats: Hydrologic Indices for Daily Time Series Data.
- Britton, D.L. (1991). Fire and the chemistry of a South African mountain stream. *Hydrobiologia* 218, 177–192.
- Burke, M.P., Hogue, T.S., Ferreira, M., Mendez, C.B., Navarro, B., Lopez, S., and Jay, J.A. (2010). The Effect of Wildfire on Soil Mercury Concentrations in Southern California Watersheds. *Water Air Soil Pollut* 212, 369–385.
- Cannon, S.H., Kirkham, R.M., and Parise, M. (2001). Wildfire-related debris-flow initiation processes, Storm King Mountain, Colorado. *Geomorphology* 39, 18.
- Carlson, T.N., and Ripley, D.A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment* 62, 241–252.
- C.E. Woodcock, R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, et al. (2008). Free access to Landsat imagery. *Science* 320, 1011.
- Condon, L.E., Hering, A.S., and Maxwell, R.M. (2015). Quantitative assessment of groundwater controls across major US river basins using a multi-model regression algorithm. *Advances in Water Resources* 82, 106–123.
- Coombs, J.S., and Melack, J.M. (2013). Initial impacts of a wildfire on hydrology and suspended sediment and nutrient export in California chaparral watersheds. *Hydrol. Process.* 27, 3842–3851.
- Daniel G Neary, G.J.G. (2003). Post-wildfire watershed flood responses.

- DeBano, L.F. (2000). The role of fire and soil heating on water repellency in wildland environments: a review. *J. Hydrol.* 231, 195–206.
- Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.L., Quayle, B., and Howard, S. (2007). A project for monitoring trends in burn severity. *Fire Ecology* 3.
- Emelko, M.B., Silins, U., Bladon, K.D., and Stone, M. (2011). Implications of land disturbance on drinking water treatability in a changing climate: Demonstrating the need for “source water supply and protection” strategies. *Water Research* 45, 461–472.
- Emmerich, W.E., and Cox, J.R. (1994). Changes in Surface Runoff and Sediment Production after Repeated Rangeland Burns. *Soil Science Society of America Journal* 58, 199.
- Falcone, J. (2011). GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow (Reston, Virginia: U.S. Geological Survey).
- Fraley, C., Raftery, A.E., Murphy, T.B., and Scrucca, L. (2012). *mclust* Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation (Department of Statistics, University of Washington).
- Frederick B. Pierson, C.A.M. (2009). Prescribed-fire effects on rill and interrill runoff and erosion in a mountainous sagebrush landscape. *Earth Surface Processes and Landforms* 34.
- Goldman, S.J., Jackson, K., and Bursztynsky, T.A. (1986). *Erosion and sediment control handbook* (McGraw-Hill).
- Google, Inc. Google Earth Engine.
- Hastie, T., Tibshirani, R., and Friedman, J. (2013). *The Elements of Statistical Learning; Data Mining, Inference, and Prediction* (Springer).
- Hessling, M. (1999). Hydrological modelling and a pair basin study of mediterranean catchments. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* 24, 59–63.
- Hester, J.W., Thurow, T.L., and Charles A. Taylor, J. (1997). Hydrologic Characteristics of Vegetation Types as Affected by Prescribed Burning. *Journal of Range Management* 50, 199–204.
- Homer, C., Huang, C., Wylie, B., and Coan, M. (2004). Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing* 70, 829–840.
- Juli G. Pausas, J.L. (2008). Are wildfires a disaster in the Mediterranean basin? - A review. *Int J Wildland Fire*. *International Journal of Wildland Fire - INT J WILDLAND FIRE* 17.
- Kinoshita, A.M., and Hogue, T.S. (2011). Spatial and temporal controls on post-fire hydrologic recovery in Southern California watersheds. *Catena* 87, 240–252.
- Kinoshita, A.M., and Hogue, T.S. (2015). Increased dry season water yield in burned watersheds in Southern California. *Environ. Res. Lett.* 10, 014003.

- Lavabre, J., Torres, D.S., and Cernesson, F. (1993). Changes in the hydrological response of a small Mediterranean basin a year after a wildfire. *Journal of Hydrology* 142, 273–299.
- Lee, R.J., and Chow, T.E. (2015). Post-wildfire assessment of vegetation regeneration in Bastrop, Texas, using Landsat imagery. *GISci. Remote Sens.* 52, 609–626.
- Lindley, A.J., Bosch, J.M., and van Wyk, D.B. (1988). Changes in water yield after fire in fynbos catchments. *Water SA* 14, 7–12.
- Littell, J.S., McKenzie, D., Peterson, D.L., and Westerling, A.L. (2009). Climate and wildfire area burned in western U.S. ecoregions, 1916–2003. *Ecological Applications* 19, 1003–1021.
- Loáiciga, H.A., Pedreros, D., and Roberts, D. (2001). Wildfire-streamflow interactions in a chaparral watershed. *Advances in Environmental Research* 5, 295–305.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. (The Regents of the University of California),.
- McManamay, R.A., Bevelhimer, M.S., and Kao, S.-C. (2014). Updating the US hydrologic classification: an approach to clustering and stratifying ecohydrologic data. *Ecohydrol.* 7, 903–926.
- Meyer, G.A., Pierce, J.L., Wood, S.H., and Jull, A.J.T. (2001). Fire, storms, and erosional events in the Idaho batholith. *Hydrol. Process.* 15, 3025–3038.
- Moody, J.A., and Martin, D.A. (2001). Initial hydrologic and geomorphic response following a wildfire in the Colorado Front Range. *Earth Surface Processes and Landforms* 26, 1049–1070.
- Moody, J.A., Shakesby, R.A., Robichaud, P.R., Cannon, S.H., and Martin, D.A. (2013). Current research issues related to post-wildfire runoff and erosion processes. *Earth-Science Reviews* 122, 10–37.
- Moritz, M.A., Moody, T.J., Krawchuk, M.A., Hughes, M., and Hall, A. (2010). Spatial variation in extreme winds predicts large wildfire locations in chaparral ecosystems. *Geophys. Res. Lett.* 37, L04801.
- National Cooperative Soil Survey (2014). Gridded Soil Survey Geographic (gSSURGO-10) Database for the Conterminous United States - 10 meter.
- National Wildfire Coordinating Group Glossary of Wildland Fire Terms.
- Neary, D.G., Ryan, K.C., and DeBano, L.F.; (2005). *Wildland fire in ecosystems: effects of fire on soils and water.*
- Owens, P.N., Giles, T.R., Petticrew, E.L., Leggat, M.S., Moore, R.D., and Eaton, B.C. (2013). Muted responses of streamflow and suspended sediment flux in a wildfire-affected watershed. *Geomorphology* 202, 128–139.
- Poff, N. (1996). A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors. *Freshwater Biology* 36, 71–79.
- PRISM Climate Group (2004).

- Santos, R.M.B., Sanches Fernandes, L.F., Pereira, M.G., Cortes, R.M.V., and Pacheco, F.A.L. (2015). Water resources planning for a river basin with recurrent wildfires. *Science of The Total Environment* 526, 1–13.
- Schwartz, G.E., and Alexander, R.B. (1995). Soils data for the Conterminous United States Derived from the NRCS State Soil Geographic (STATSGO) Data Base.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Statist.* 6, 461–464.
- Scott, D.F. (1993). The hydrological effects of fire in South African mountain catchments. *J Hydrol. Journal of Hydrology* 150, 409–432.
- Shakesby, R.A., and Doerr, S.H. (2006). Wildfire as a hydrological and geomorphological agent. *Earth-Sci. Rev.* 74, 269–307.
- Smith, H.G., Sheridan, G.J., Lane, P.N.J., Nyman, P., and Haydon, S. (2011). Wildfire effects on water quality in forest catchments: A review with implications for water supply. *Journal of Hydrology* 396, 170–192.
- Stein, E.D., Brown, J.S., Hogue, T.S., Burke, M.P., and Kinoshita, A. (2012). Stormwater contaminant loading following southern California wildfires. *Environ. Toxicol. Chem.* 31, 2625–2638.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society (Series B)* 58, 267–288.
- Townsend, S.A., and Douglas, M.M. (2000). The effect of three fire regimes on stream water quality, water yield and export coefficients in a tropical savanna (northern Australia). *Journal of Hydrology* 229, 118–137.
- U.S. Geological Survey (2014). National Water Information System—NWISWeb: accessed on various dates in 2014 at <http://waterdata.usgs.gov>.
- V. D. Lyne, M.H. (1979). Stochastic Time-Variable Rainfall-Runoff Modeling. Australian National Conference Publication 79.
- Watson, F., Vertessy, R., McMahon, T., Rhodes, B., and Watson, I. (2001). Improved methods to assess water yield changes from paired-catchment studies: application to the Maroondah catchments. *Forest Ecology and Management* 143, 189–204.
- Westerling, A.L., Hidalgo, H.G., Cayan, D.R., and Swetnam, T.W. (2006). Warming and earlier spring increase western US forest wildfire activity. *Science* 313, 940–943.
- Whitlock, C. (2004). Land management: Forests, fires and climate. *Nature* 432, 28–29.
- Wilkinson, S.N., Wallbrink, P.J., Hancock, G.J., Blake, W.H., Shakesby, R.A., and Doerr, S.H. (2009). Fallout radionuclide tracers identify a switch in sediment sources and transport-limited sediment yield following wildfire in a eucalypt forest. *Geomorphology* 110, 140–151.
- (2009). Monitoring Trends in Burn Severity.