2018

Application of remote sensing and machine learning modeling to post-wildfire debris flow risks

Priscilla Addison
Michigan Technological University, peaddiso@mtu.edu

APPLICATION OF REMOTE SENSING AND MACHINE LEARNING MODELING TO POST-WILDFIRE DEBRIS FLOW RISKS

By

Priscilla E. Addison

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Geological Engineering

MICHIGAN TECHNOLOGICAL UNIVERSITY

2018

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Geological Engineering.

Department of Geological and Mining Engineering and Sciences

Dissertation Advisor:	*Dr. Thomas Oommen*
Committee Member:	*Dr. Ann Maclean*
Committee Member:	*Dr. Qiuying Sha*
Committee Member:	*Dr. Stanley Vitton*
Department Chair:	*Dr. John Gierke*

Table of Contents

Author Contribution Statement

Chapters 2 and 3 are composed of published material that can be found in peer-reviewed academic journals, and Chapter 4 is work that has been submitted for publication review. Find details of these as follows:

Chapter 2: Addison, P. and Oommen, T., 2018. Utilizing satellite radar remote sensing for burn severity estimation. International Journal of Applied Earth Observation and Geoinformation, 73, pp.292-299

> **Author Contributions**: Addison and Oommen conceptualized the manuscript, and provided revision throughout the study. Addison processed the data and drafted the manuscript.
>
> **Permission**: Reproduction of this article falls under non-commercial use: "*Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: www.elsevier.com/about/our-business/policies/copyright#Author-rights.*"
>
> **Article Link**: www.sciencedirect.com/science/article/pii/S030324341830432X

Chapter 3: Addison, P., Oommen T., and Sha, Q., 2018. Assessment of Post-Wildfire Debris Flow Occurrence Using Classifier Tree. Geomatics, Natural Hazards and Risk.

> **Author Contributions**: Addison and Oommen conceptualized the manuscript, and provided revision throughout the study. Sha provided statistical modeling expertise, as well as manuscript review. Addison processed the data and drafted the manuscript.
>
> **Permission**: This is an Accepted Manuscript of an article published by Taylor & Francis Group in Geomatics, Natural Hazards and Risk journal. It has been published as an open access article and hence can be reprinted in a dissertation without requesting further permission. More information can be found here: https://authorservices.taylorandfrancis.com/sharing-your-work/

Chapter 4: Addison, P., and Oommen T., 2018. Post-Fire Debris Flow Modeling Analyses: Case Study of the Post Thomas Fire Event in California. Natural Hazards.

> **Author Contributions**: Addison and Oommen conceptualized the manuscript, and provided revision throughout the study. Addison processed the data and drafted the manuscript.

Abstract

Historically, post-fire debris flows (DFs) have been mostly more deadly than the fires that preceded them. Fires can transform a location that had no history of DFs to one that is primed for it. Studies have found that the higher the severity of the fire, the higher the probability of DF occurrence. Due to high fatalities associated with these events, several statistical models have been developed for use as emergency decision support tools. These previous models used linear modeling approaches that produced subpar results. Our study therefore investigated the application of nonlinear machine learning modeling as an alternative. Existing models identified the burn severity of wildfires as an important input in their development. Currently, the most widespread approach to obtaining this input is the use of the differenced normalized burn ratio (dNBR) index, which is determined using data from optical sensors on satellites. However, progress of this existing protocol is mostly hampered by the presence of cloud coverage during data acquisition since optical sensors cannot penetrate clouds. Radar sensors on the other hand can penetrate clouds and smoke. This study therefore developed a radar based algorithm to be used as an alternative to the dNBR metric. The results showed the SAR metric to perform even better than the dNBR, with an overall accuracy (OA) of ~60% and Kappa of 0.35 in comparison to an OA of ~35% and a kappa of 0.1 from the dNBR approach. Next we developed a nonlinear machine learning model to predict the likelihood of post-wildfire debris flow occurrences. This produced improved results over the linear modeling approach with an average sensitivity of 77%, depicting increased ability to predict ~8 out of 10 DF producing basins. Finally, we performed a case study to validate our DF model that showed the model's robustness in isolating especially high hazard locations. Having these improved models will furnish emergency responders with an increased ability to better assess the associated risks of potential debris flow producing basins and make informed decisions on mitigation and/ or prevention measures.

1 Introduction

About 350 million ha of land worldwide burn annually as a result of wildfires (van der Werf et al., 2006). This projected coverage is potentially larger in recent years since wildfire frequencies have increased due to the onset of drier climates (Bond-Lamberty et al., 2007; Berman, 2017; Orosco, 2017; Dolan, 2018). Starting from ignition, wildfires leave devastations in their wake, which continue even years after the fire is ended. In fact, history shows most post-fire hazards to be more deadly than the fire itself. An example is the Thomas Fire, in Southern California in December 2017. This fire resulted in two fatalities, whereas a consequent debris flow event in January caused 21 deaths (Berman, 2017; Orosco, 2017; Dolan, 2018).

Post-fire hazards include, but are not limited to emission of greenhouse gases, erosion, flash floods, sediment-laden floods, debris flows, and rock falls (Dixon and Krankina, 1993; Moody and Martin, 2001; Cannon et al., 2009; 2010). Our study focuses on debris flows because historically they are the most fatal of these hazards and their frequencies in western USA are increasing (Cannon and DeGraff, 2009; Eaton, 1935; Bailey et al., 1947; Wells, 1987). A debris flow is a fast-moving, high-density slurry of water, sediments and debris that travels under gravity with enormous destructive power. It usually occurs after periods of intense, short duration precipitation (rainfall, snow melt, etc.) on steep hillsides covered with loose erodible material (Cannon et al., 2010). Debris flows are not exclusive to fire affected areas but a location's vulnerability increases significantly after it experiences wildfires (Cannon and Gartner 2005; Canon et al., 2010). The process by which this occurs are outlined as follows. Burning off of the rainfall intercepting vegetation (Kinner and Moody, 2008; Wondzell and King, 2003; Cannon et al., 2010), sealing of the soil pores from the generated ash (Cannon and Gartner, 2005; Larsen et al., 2009), and predominantly, the condensation of organic compounds produce an aftermath of a water repellent soil (DeBano, 2000; Doerr et al., 2000; Cannon and Gartner, 2005; Moody and Ebel, 2012). These processes result in the formation a non-cohesive, water-repellent, bare soil, which is primed for the movement of large volumes of sediment within the burned area and its vicinity.

Behaviors of debris flows are extremely difficult to predict due to their complex physical structure, trigger mechanisms, and regional variability (Cannon et al., 2010). There are records of studies dating to the 1980s that began the task of detailing the physical behavior and impact of debris flows (Hungr et al., 1984; Johnson et al., 1991; Cannon et al. 2010; Staley et al, 2017; Prochaska et al. 2008; Santi et al. 2011). Recognition and study of the influence of wildfires on these events began in 1991 when Johnson et al. discovered an increase in volumes of eroded material at locations that had experienced wildfires in their recent past (Johnson et al., 1991). Currently, researchers are focusing on developing statistical models that attempt to predict the probability of post-fire debris flow occurrences. These models are developed by utilizing a number of different descriptive characteristics of a location such as: basin morphology, rainfall characteristics, burn severity, and soil properties (e.g., Hungr et al. 1984; Bovis and Jakob 1999; Gartner 2005; Gartner et al. 2008; Cannon et al. 2010).

Unsurprisingly, development of these models showed the burn severity of wildfires to be a very important input (Johnson et al., 1991; Cannon et al., 2010; Staley et al., 2017). Generally, a high burn severity basin has a higher likelihood to produce debris flows of larger volumes than one with low severity (Cannon et al., 2010). This makes the burn severity input especially critical to this study. Currently the common approach to assessing this parameter is by employing satellite remote sensing. In particular, data from optical sensors data have been used extensively to develop algorithms that have proved to be instrumental in mapping burned areas (Koutsias et al., 2000; Justice et al., 2002; Roy et al., 2002; Mitri and Gitas., 2004; Chuvieco et al., 2006; Polychronaki et. al, 2013; Kalogirou et al., 2014; Stroppiana, 2015). The differenced normalized burn ratio (dNBR) is the most common of these algorithms. These optical based approaches, however, have a major disadvantage associated with them, in that their progress can be hampered or even halted by smoke and/or cloud coverage during data acquisition. Substantial smoke and cloud coverage can corrupt an optical image and render it unusable for any analyses to be done. A user who encounters this problem is forced to wait for at least an entire temporal cycle of the satellite of interest for the chance to obtain usable data. For

emergency situations, this delay can be life threatening and/ or costly. Past research have also shown that the dNBR method of assessing fire severity works well for unburned and highly burned areas but becomes inconsistent in the intermediary severity levels. Some studies found it problematic to clearly delineate intermediate severity burns (Chuvieco et al., 2006; Allen and Sorbel, 2008; Hoy et al., 2008; Murphy et. al., 2008). Radar sensors, on the other hand, can penetrate smoke and clouds and have been used by several researchers for fire studies (Bourgeau-Chavez, 1997; Rignot et al., 1999; Hoekman et al., 2010; Polyvhronaki, 2013; Kalogirou, 2014), but limited work has been done in the particular area of burn severity determination, especially in the fire prone western United States. With these problem statements in mind, our study set out to complete the following three objectives:

- Develop a radar-based burn severity estimate to be used as an alternative in emergency situations where there is cloud and/ or smoke coverage.
- Develop a nonlinear machine learning model to predict the likelihood of post-wildfire debris flow occurrences in the western USA.
- Test robustness of developed debris flow model with a case study of recent event occurrence.

We employed machine learning modeling in all three aspects of the study because its algorithms offer flexibility through data driven predictions made by iteratively learning from the input data, as opposed to the strictly static algorithms of some linear models. Machine learning is a type of artificial intelligence that was originally developed to provide computers with the ability to learn without being explicitly programmed (Samuel, 1959). The algorithm was initially developed for the computer science discipline but is now gaining notoriety in other fields due to its robustness. Several models have been developed that use this algorithm to learn from previous computations by searching through data to look for patterns, iteratively re-adjusting until a final robust pattern is found (Samuel, 1959; Fogel et al., 1966; Kohavi and Provost, 1998; Kuhn and Johnson, 2013). Hence, the more data you have to learn from, the better your output. Machine learning modeling can be applied even when the theory behind the data in

question is not fully understood. Models that use these algorithms have become widespread in recent years because they have proved to be better at teasing out complex relationships than simple linear models (Kern et al., 2017). Some examples of the success of machine learning modeling include their use in the environmental science field to detect oil spills from satellite imagery (Kubat et al., 1998); geoscience and remote sensing data (Oommen et al., 2007; 2008; Samui et al., 2012; Gowda et al. 2015); as well as in the medical field to detect cancer tissue samples using microarray expression data (Furey, 2000; Stalin and Kalaimagal, 2016). The surveillance field has used them to develop algorithms for facial recognition (Rowley et al., 1998; Dolecki et al., 2016), whereas the financial world has used them for predicting bankruptcy as well as credit rating (Odom et al., 1990; Wu et al., 2016). There are many other disciplines that have explored machine learning modeling for various complex scenarios that depict the extreme usefulness of this approach and hence, touts its use as an emergency response tool for post-fire debris flow hazards.

To tackle the first aspect of our project, we developed a radar-based metric using the C5.0 decision tree algorithm. This furnished us with an alternative model to use in place of, or in the absence of the optical models. Specifically, we developed a synthetic aperture radar (SAR) based metric aimed at classifying the burn severity into three categories: low severity, moderate severity, and high severity. SAR is a technique that artificially lengthens a radar sensor's antenna by capitalizing on the flight movement to provide high resolution imagery. The way SAR works is that it transmits microwave energy to a target object, after hitting the target the wave scatters, part of it are lost but others are transmitted back to the SAR receiver. This is measured as the *backscatter* value. The amount of waves that are lost depends on the composition and nature of the target object. Its application to burned area studies is that after a surface experiences burn, the loss in vegetative cover causes backscatter variations. These changes are directly proportional to the degree of burn and therefore make it possible to map the burn severities (Tanase et al., 2015a; Polychronaki et al., 2013). The SAR data that was used was the Japan Aerospace Exploration Agency's Advanced Land Observing Satellite

Phased Array type L-band Synthetic Aperture Radar (ALOS PALSAR) obtained from the Alaska Satellite Facility. Details of this study have been provided in the Chapter 2 of this dissertation.

Moving on to the issue of debris flows prediction, we again employed the machine learning-based decision tree algorithm to develop a nonlinear model that predicts the probability of debris flow generation after wildfires. Previous studies, mostly by researchers at the United States Geological Survey (USGS), had predominantly employed linear models, specifically the logistic regression model (Cannon and Gartner 2005, Cannon et al., 2010, Staley et al., 2017). The logistic regression approach is advantageous because it considers simple linear classification boundaries, making model development simple as well as easy to interpret. However, its simplicity is a detriment to it in debris flow modeling due to its complex triggering and flow mechanics. Up until 2017, the best model reported a sensitivity of 44% (Cannon et al., 2010) — this translates to an approximate 4 out of 10 debris flow producing basins being correctly isolated. This was not ideal and hence, Staley et al., 2017 used a more data rich sample to develop an updated logistic regression model with an improved sensitivity of 83%. It, however, had a specificity of 58%, which means that ~6 out of 10 "debris flow safe" locations will be correctly predicted. Although this updated model provides an increased ability to isolate more vulnerable areas, it held the risk of desensitizing the public, due to the likelihood of producing high false positive predictions. We therefore proposed the use of the C5.0 decision tree algorithm to investigate if higher predictive capabilities could be harnessed from this nonlinear machine learning approach. Details of this have been provided in Chapter 3 of this dissertation.

Finally, the fourth chapter of this dissertation sought to validate the developed debris flow prediction model by considering a case study. We applied both the C5.0 tree and USGS' logistic regression models to a recent fire that happened in Southern California, the Thomas Fire. This fire, which was the largest in California's recent past at time of its ignition, occurred from December 4, 2017 to January 12, 2018 and consumed ~114,000 ha of land. Before the fire could be fully contained a major storm occurred on January 9,

2018, which triggered debris flows within the burned area, inundating several communities downstream in the Santa Barbara and Ventura Counties. Using this location as a case study, we validated our C5.0 tree model as well as compare its predictive strength with that of the logistic regression model. Details of this have also been provided in Chapter 4 of this dissertation.

2 Utilizing Satellite Radar Remote Sensing for Burn Severity Estimation[1]

Abstract: The increasing knowledge in the capabilities of satellite imagery to hazard applications is especially useful in emergency situations where timing and ability to cover large areas is of the essence. For optical imagery, cloud coverage can corrupt an image rendering it unusable for intended emergency analyses. This study proposes the use of Synthetic Aperture Radar (SAR) imagery for burn severity analysis for western United States sites, as an alternative to its optical based counterpart, differenced normalized burn ratio (dNBR). Unlike optical sensors, the radar sensor is an active sensor that is able to penetrate clouds and smoke, an attribute that is crucial in emergency situations where immediate burn severity data are needed to assess the vulnerability of fire affected areas to post-fire hazards. Using C5 decision tree algorithm we developed a SAR-based metric that attempts to classify burn severities of fire affected locations in the western USA. We then compared the performance of this developed metric to that obtained by the existing dNBR metric, to determine if there is any merit to its adoption as an alternative for the western USA landscape. The results showed the SAR approach to produce higher validation metrics in comparison to the dNBR. It had an overall accuracy and kappa of 60% and 0.35, respectively, in comparison to the 35% and 0.1 of the dNBR approach. This shows an improved ability to quickly obtain burn severity data and make better informed decision in emergency situations.

[1]This material has been published as an Open Access article in the *International Journal of Applied Earth Observation and Geoinformation*. Check page v for original article link.

2.1 Introduction

The hazards associated with a wildfire continue even after it is contained. A wildfire's aftermath usually results in the loss of vegetative cover, leaving the ground exposed and vulnerable to a plethora of post-fire hazards. These hazards include erosion, sediment flows, rock falls, flash floods, debris flows, and release of greenhouse gases into the atmosphere (Dixon and Krankina, 1993; Moody and Martin, 2001; Cannon et al., 2009; 2010). The intensity of such post-fire hazards are usually exacerbated by the severity of the fire. Emergency response teams therefore need immediate access to burn severity maps to enable them to assess the vulnerability of fire affected areas to post-fire hazards (Cannon et al., 2010; Staley et al., 2016; Kern et al, 2017). The United States Department of Agriculture (USDA), for example, has formed a dedicated team— Burned Area Emergency Response (BAER)— whose mission is to rapidly evaluate severity of fires and their implications on Federal lands and prescribe emergency stabilization treatments. The BAER team works on a short turnaround time because their activities have to be completed before the next major storm. Their work therefore begins even while the wildfire is still ongoing because it is imperative that stabilization measures (reseeding with quick-growing species, mulching, building slope breakers, etc.) are applied as soon as possible (Witt, 1999)

Currently, the common approach to determining burn severity of an extensive area is by employing satellite imagery since they provide objective estimates and cover wider investigative areas (Key and Benson, 2006; Miller and Thode, 2007; Parks et al., 2014). In particular, optical satellite data have been used extensively and proven to be useful for mapping burned areas (Koutsias et al., 2000; Justice et al., 2002; Roy et al., 2002; Mitri and Gitas., 2004; Chuvieco et al., 2006; Stroppiana, 2015). The optical approach, however, has a major disadvantage of being hindered by cloud coverage and smoke from ongoing wildfires (Schroeder et al., 2008; Wooster et al., 2013; Allison et al., 2016). Substantial cloud coverage and/or smoke can mask an optical image and render it unusable for the intended burn severity analysis. A user who encounters this misfortune is forced to wait for at least an entire temporal cycle of the satellite—16 days, in the case of

Landsat—for the chance to obtain usable data. For potential emergency hazards, this delay can be life threatening and/or costly. Also, studies have shown that the most widely used optical approach, the differenced normalized burn ratio (dNBR), shows inconsistencies in the intermediary burn severity levels. Some studies found it problematic to clearly delineate intermediate severity burns (Chuvieco et al., 2006; Allen and Sorbel, 2008; Hoy et al., 2008; Murphy et. al., 2008). Finally, the dNBR index is an absolute measure of landscape (vegetation) change, hence it fails to take into account the heterogeneous nature of landscapes. That is, for the same intensity of burn, a pixel with sparse pre-fire vegetation will measure a small change (lower dNBR) even if it experienced a high severity burn; whereas its counterpart with dense pre-fire vegetation will register a high change (higher dNBR).This becomes problematic when a user encounters a landscape with different vegetation types and densities and try to rank them on the same burn severity scale (Miller and Thode, 2007; Parks et al., 2014).

We are therefore proposing an alternate approach in this study: radar burn severity estimation. Unlike optical sensors, the radar sensor is an active sensor that is able to penetrate clouds and smoke, an attribute that is crucial in emergency situations. Radar satellite sensors use a technique known as synthetic aperture radar (SAR), which is basically the artificial lengthening of a sensor's antenna by capitalizing on the flight movement to provide high resolution imagery. SAR uses microwave energy to quantify and discern between ecological processes by measuring the differences in scattering based on the roughness of the target surface. Variation in the dielectric constant of target objects plays a central role in determining the intensity (backscatter) of the microwave energy that is received and processed into the resulting SAR image (Kasischke et al., 1997). Its application to burned area studies is that after a surface experiences burn, the loss in vegetative cover causes scattering variations. These changes are directly proportional to the degree of burn and therefore make it possible to map the burn severities (Tanase et al., 2015a; Polychronaki et al., 2013). The use of SAR to determine burn severity is in no way novel, however, it is a fairly new application area with limited literature currently available (Tanase et al., 2010a, 2010b, 2015a, 2015b). Further, the

existing studies focus predominantly on Mediterranean forests. The signature of SAR's measurement is unique to different vegetation types, hence with the landscape of the western USA being different from the Mediterranean landscape, an extensive study like this one with unique concentration on the western USA is warranted. Finally, noting the difficulty of the dNBR approach to account for the heterogeneous nature of landscapes we will consider relative measures of SAR burn severity by averaging over the pre-burn condition of each landscape to account for its relative change.

2.2 Methods

2.2.1 Study Area

Our study considered 15 fires that occurred across six States in the fire-prone western USA (Figure 2.1, Table 2.1) from 2008 to 2010. These fires were chosen out of the 37 fires in the FIRESEV (FIRE SEVerity mapping system project) database (Dillon et al, 2011a; Sikkink et al., 2013) because they were the ones with documented containment dates. FIRESEV was a project funded by a Joint Fire Sciences Program "geared toward providing fire managers across the western United States with critical information for dealing with and planning for the ecological effects of wildfire at multiple levels of thematic, spatial, and temporal detail" (Dillon et al, 2009). This project collected ground-based burn severity data on 37 fires that occurred in the western USA from 2008-2010 (Sikkink et al., 2013).

2.2.2 Field Data

The field protocol used in the FIRESEV project was the Composite Burn Index (CBI), which assesses the burn severity of a landscape on a continuous scale by visually examining vegetation conditions of 90-meter diameter plots in the aftermath of a fire with respect to the condition of vegetation before the fire. Values range from 0 (unburned) to 3 (high severity) for any given plot (Key and Benson, 2006).

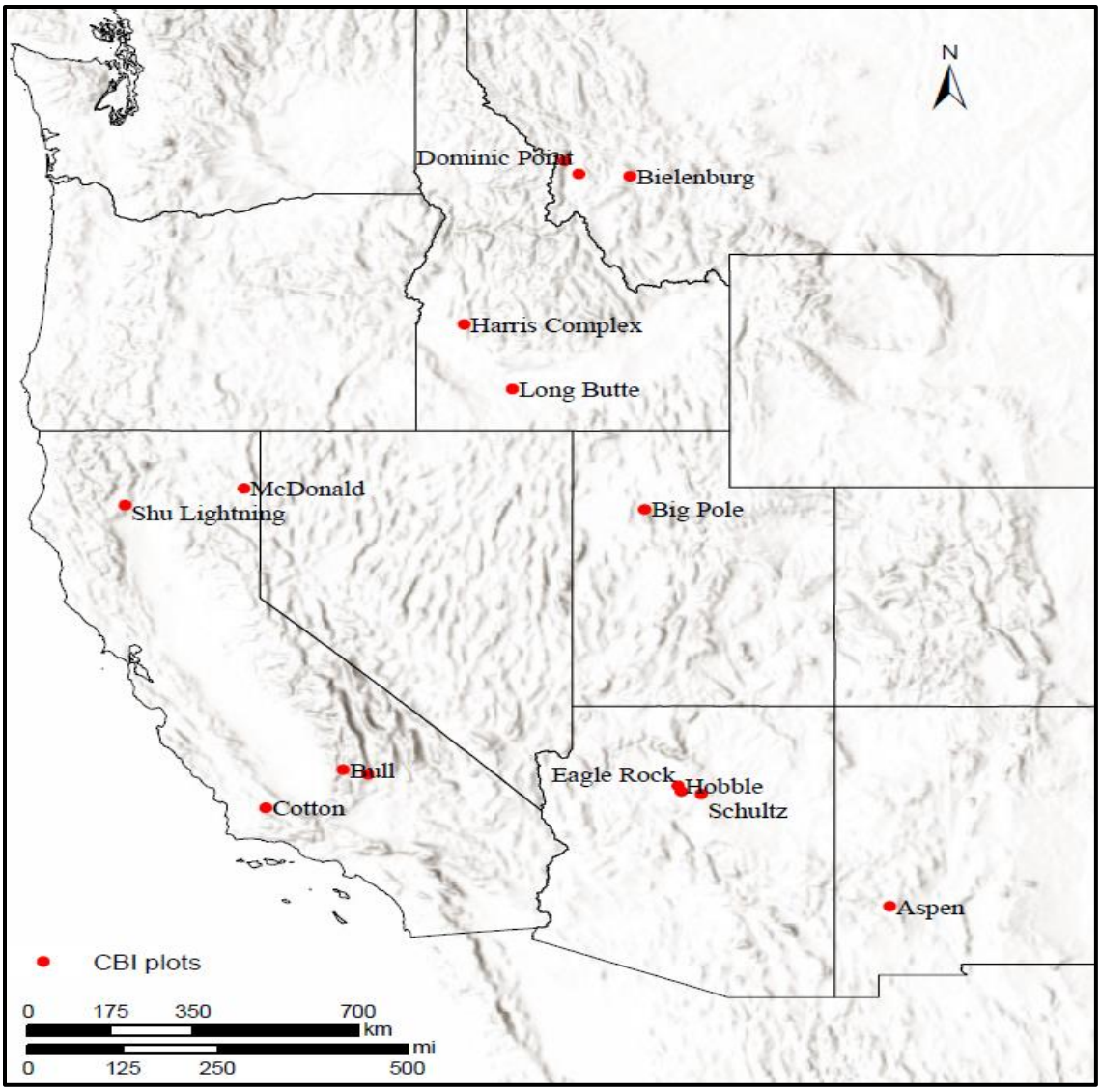

Figure 2.1 Study area of western USA showing the 15 fires from 2008-2010 from the FIRESEV database.

Setting CBI thresholds is arguably as much an art as science, so depending on study location and study focus, threshold values could differ for different studies. After doing some preliminary explorations of different thresholds, we decided on the ranges proposed by Key and Benson, 2005 (Table 2.2). The only exception was the unburned class that was set at a threshold of 0.1 because preliminary analysis showed the noted range limit of 0.5 encapsulated most of the low severity samples. Documentation on the FIRESEV project noted that it was predominantly geared towards isolating high severity burn

locations (Dillon et al, 2011b) therefore there was an intentional sample bias towards this class, which was evident even in the subset of data we obtained (Figure 2.2); this further gave confidence to the class thresholds set. The unburned class comprised ~2% of the sample size, so we removed them from further analyses to prevent this extreme class imbalance.

Table 2.1. List of fires used in the study

Fire Name	**State**	**Ignition Date**	**Containment Date**
Eagle Rock	Arizona	6/11/10	6/28/10
Hobble	Arizona	8/30/10	10/15/10
Schultz	Arizona	6/20/10	7/7/10
Bull	California	7/26/10	8/10/10
Cotton	California	5/15/10	5/17/10
Indian	California	7/18/10	7/24/10
McDonald	California	7/27/10	8/10/10
Shu Lightning	California	6/21/08	7/20/08
Harris Complex	Idaho	8/27/10	9/4/10
Long Butte	Idaho	8/21/10	8/31/10
Bielenburg	Montana	7/12/09	10/1/09
Dominic Point	Montana	7/25/10	8/7/10
Kootenai Creek	Montana	7/12/09	8/5/09
Aspen	New Mexico	6/6/10	6/26/10
Big Pole	Utah	8/6/09	8/16/09

Table 2.2 CBI and dNBR severity category definitions. Bilinear interpolation was used to cross-reference CBI values with dNBR values

Class	CBI	dNBR	Description
Unburned	0.00 – 0.10	0.00 – 0.10	Location experienced no fire. This may also include a location that recovers quickly after fires.
Low	0.10 – 1.24	0.10 – 0.27	Minimal vegetation consumption; patches of scorched foliage.
Moderate	1.25 – 2.24	0.27 – 0.66	The landscape exhibits transitional conditions between low and high severity characteristics described.
High	2.25 – 3.00	> 0.66	~ 90% to total consumption of vegetation. Sites normally exhibit over 50% cover of newly exposed mineral soil or rock fragments

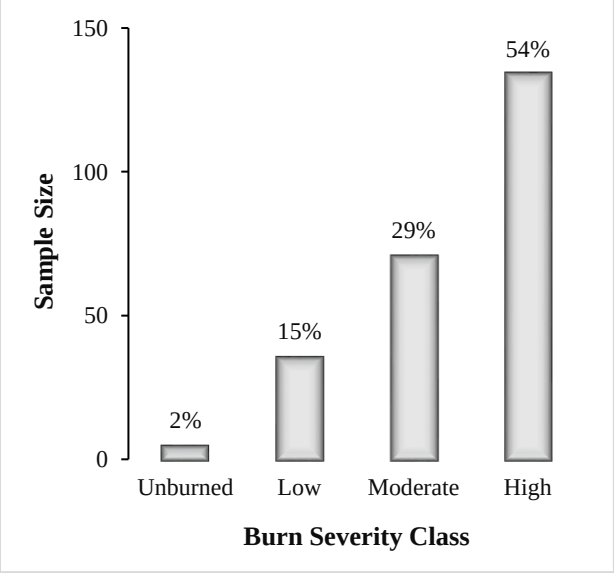

Figure 2.2 Percentage of the different CBI burn severity classes from the FIRESEV database

2.2.3 Remote Sensing Data

2.2.3.1 Optical Burn Severity Determination

The dNBR index determines burn severities by harnessing the spectral values of the near infrared (NIR) and short wave infrared (SWIR) wavelengths. NIR records high reflectance values for healthy vegetation and low values for burned vegetation; whereas the converse is true for SWIR; low reflectance in healthy vegetation and high reflectance in burned areas. With these intrinsic signatures, the normalized burn ratio (NBR) (equation 1) was developed to map burn areas. A high NBR value generally indicates healthy vegetation whereas a low value indicates low or no vegetation, such as a result of fire. To go a step further to quantify different severity classes, the dNBR (equation 2) was developed as a temporal change detection metric. This metric runs from -2 to +2, with high positive values corresponding to severely burned locations.

$$NBR = \frac{NIR\ band - SWIR\ band}{NIR\ band + SWIR\ band} \quad (1)$$

$$dNBR = NBR_{prefire} - NBR_{postfire} \quad (2)$$

Cloud-free pre- and post- images were downloaded from Landsat 5 Thematic Mapper (TM). These images were downloaded as Level 1 products (geometrically and radiometrically corrected) from USGS' Earth Explorer website. Burn severity for each scene was determined using equations (1) and (2) with bands 4 and 7 corresponding to the NIR and SWIR infrared wavelengths, respectively. We then extracted the dNBR values to correspond to each CBI plot location using bilinear interpolation, as a past study showed this approach to give representative values corresponding to the CBI plots (Parks et al., 2014). Bilinear interpolation is a resampling approach that uses the four closest pixels of an input raster, with a user defined statistic metric, to determine the value of the output raster. The statistic metric used for our study was the mean, which was adjusted to account for the distance of each of the four closest rasters to the centroid of the output cell. This resampling step was necessary to account for the extent of the CBI plot, which is unlikely to fall within a single raster pixel. The continuous dNBR values were then

classified into their respective severity classes of low, moderate, and high (Table 2.2). This four class categorization is known in the fire community as BARC-4 (burned area reflectance classification, using 4 classes).

2.2.3.2 Radar Burn Severity Determination

SAR scattering is sensitive to vegetation structure and biomass. Removal of leaves and branches from trees after wildfires alters the scattering mechanisms causing temporal variations of the backscatter coefficient (Polychronaki et al., 2013). This temporal alteration makes it possible to map burn severity. The wavelength and polarization of the SAR sensor strongly influence the accuracy of the output. The higher the wavelength the more penetration the SAR signals are able to achieve. The higher wavelength of the L-band (24 cm) therefore allows it to penetrate vegetation canopy and interact with large branches, stems, and the forest floor (Le Toan et al. 1992), which makes it possible to discern burned structures from unburned ones. Conversely, the X- and C- bands, by virtue of their lower wavelengths (5.6 cm and 3 cm, respectively), have lower penetration capabilities; scattering occurs only in the upper few centimeters of the forest canopy making them less favorable in burn severity applications (Tanase et al., 2010a, 2010b, 2015a, 2015b). In terms of polarizations, the cross-polarized state (HV) has been shown to be sensitive to volume changes hence, the removal of leaves and branches by fire and consequent thinner, dryer vegetation results in a decreased backscatter. These processes translate to the needed contrast in SAR image to discern changes.

The Advanced Land Observing Satellite Phased Array type L-band Synthetic Aperture Radar (ALOS PALSAR) was the sensor employed in our studies. This was because it is the only L-band sensor that was in operation during our investigative period and also that its data are readily available at no cost. Pre- and post- fire data were downloaded through the Alaska Satellite Facility (ASF) as fine beam dual polarization (FBD) products, which meant that they had both HH and HV polarization information. A preliminary assessment of the mean backscatter values for the study locations pre- and post- fire conditions showed that unburned locations recorded higher backscatter coefficients in comparison to

their burned counterparts for both HH and HV polarizations (Figure 2.3). This supports the theory found in earlier literature.

The ALOS PALSAR data were also downloaded as radiometrically terrain-corrected (RTC) products from ASF. This meant that single look complex (SLC) data had already been converted to radiometrically and terrain geo-coded data by ASF. By this, co-registering was done for scenes obtained from the same sensor and track. Multi-looking had also been performed to obtain representative pixel sizes as well as reduce the characteristic speckle noise associated with SAR data. For images with fully developed speckle noise, further filtering had been done by applying a sensor-suitable adaptive or non-adaptive filter; keeping in mind to preserve good radiometric as well as spatial resolution (Bernhard et al., 2011; Gimeno et al., 2004). Geocoding had finally been carried out using the best available digital elevation model (DEM). Two sets of data were downloaded for each fire for pre- and post- fire scenarios, respectively.

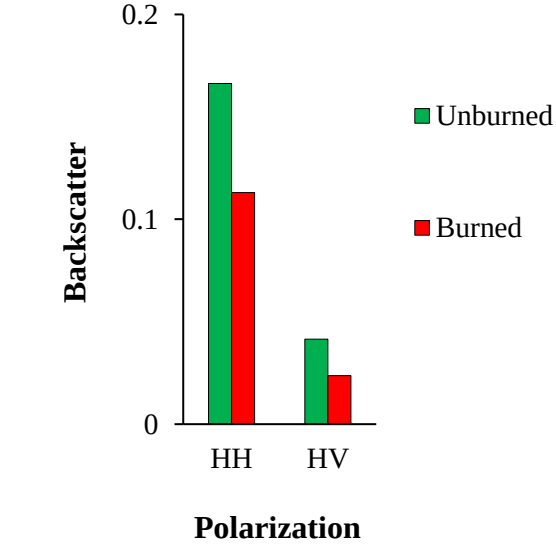

Figure 2.3 A plot of mean backscatter values for the fifteen burned areas considered in this study for before (green) and after (red) the fires.

Next, we developed the following absolute and relative predictors from the HH and HV backscatter data to quantify the landscape changes due to wildfire:

$$Abs\_HH = HH_{pre} - HH_{post} \quad (3)$$

$$Abs\_HV = HV_{pre} - HV_{post} \quad (4)$$

$$Rel\_HH\_1 = \frac{HH_{pre} - HH_{post}}{HH_{pre}} \quad (5)$$

$$Rel\_HV\_1 = \frac{HV_{pre} - HV_{post}}{HV_{pre}} \quad (6)$$

$$Rel\_HH\_2 = \frac{HH_{pre} - HH_{post}}{\sqrt{HH_{pre}}} \quad (7)$$

$$Rel\_HV\_2 = \frac{HV_{pre} - HV_{post}}{\sqrt{HV_{pre}}} \quad (8)$$

Equations (3) and (4) to give the absolute changes of the landscape after fires, whereas equations (5) to (8) give the relative changes with respect to the existing condition on the ground before the fire. We extracted each of these predictors to correspond to the CBI plots using bilinear interpolation, as was done for the dNBR data. Land Cover Data

We observed the existence of different vegetation types across the six States that were investigated. Past studies indicate that the differences in the moisture content and individual components of different vegetation types results in unique dielectric constants (Rignot et al, 1994; Kasischke et al., 1997; Wegmuller and Werner, 1997). The different dielectric constants cause different vegetation groups to have different scattering and attenuating responses to the microwave energy transmitted by radars (Kasischke et al., 1997). Therefore, we obtained land cover data from the National Land Cover Database (NLCD) from the United States Department of Agriculture's geospatial data gateway to quantify the broad land cover types. Preliminary analyses showed improved results when land cover data was integrated into model development. Our sites comprised the following three dominant land cover types: evergreen forest, herbaceous, and shrub.

These were coded as three separate dummy variables to be used as predictors and were extracted for each CBI plot using nearest neighbor resampling method. With these three dummy variables, together with the six predictors from equations (3) – (8), we had a total of nine candidate predictors for developing our model.

2.2.4 Data Analyses

2.2.4.1 C5.0 Algorithm

The algorithm used for this study was the C5.0 decision tree (Quinlan, 1993). It is a simple method that uses inductive inference to approximate discrete-valued functions. It is robust to noisy data and capable of learning mutually exclusive expressions. Its algorithm works by sorting the data from a base decision into smaller, more homogeneous groups. It does this by developing a general decision called the root, then branching out from there with other more specific decisions known as branches, till it gets to a decision that produces homogeneous samples, which are then assigned a classification known as the leaf (Kuhn and Johnson, 2013). This represents a typical tree structure, hence, the name decision tree. Figure 2.4 gives a schematic of a simplified tree with one root, two branches and four leaves.

At each node, each predictor is tested to assess how well it alone separates the training data according to the target classification. The process is repeated on each new subset until a subset contains only samples of a single class, or the partitioning tree has reached a predetermined maximum depth. Trees are grown to their maximum size and then a pruning step is usually applied by removing a decision's precondition if the accuracy of the decision improves without it. This is done to avoid overfitting to the training data. The C5.0 tree algorithm uses boosting in its model training process, which works by combining average model decisions together to improve overall model performance of the final output.

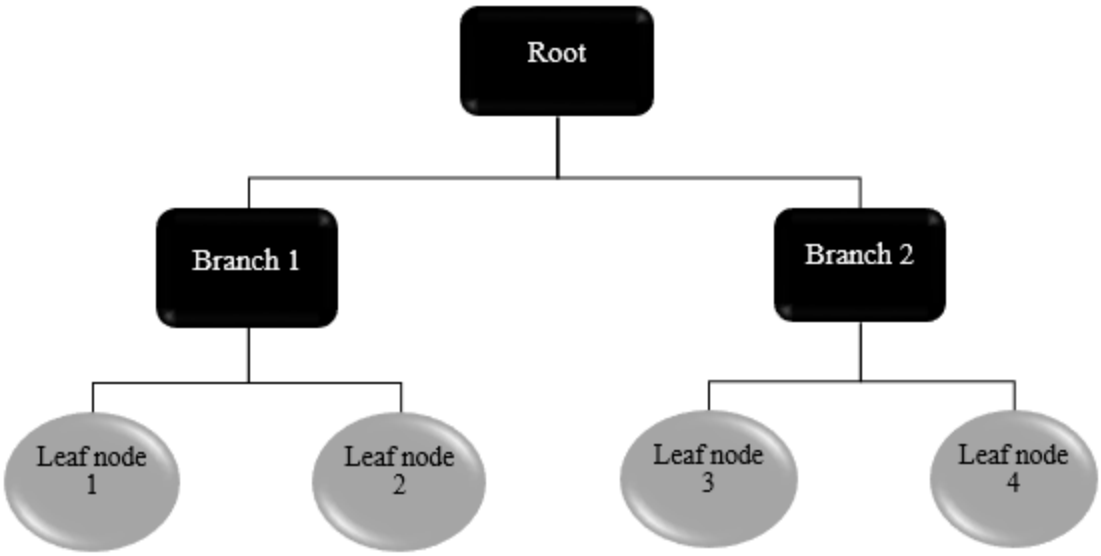

Figure 2.4 Schematic of a simplified decision tree with a root, two branches and four leaf nodes

2.2.4.2 Model Development

Figure 2.5 shows boxplots of the six continuous predictors (excluding the three land cover predictors) as they relate to the different burn severity classes from the CBI data. A look at this shows the low and moderate severity classes to have mostly similar median values. This likely due to the fact that there were not as many samples for these two as there were for the high severity class. Preliminary data pre-processing steps taken to ensure optimal model performances included omitting observations with missing values and testing for predictor degeneracy.

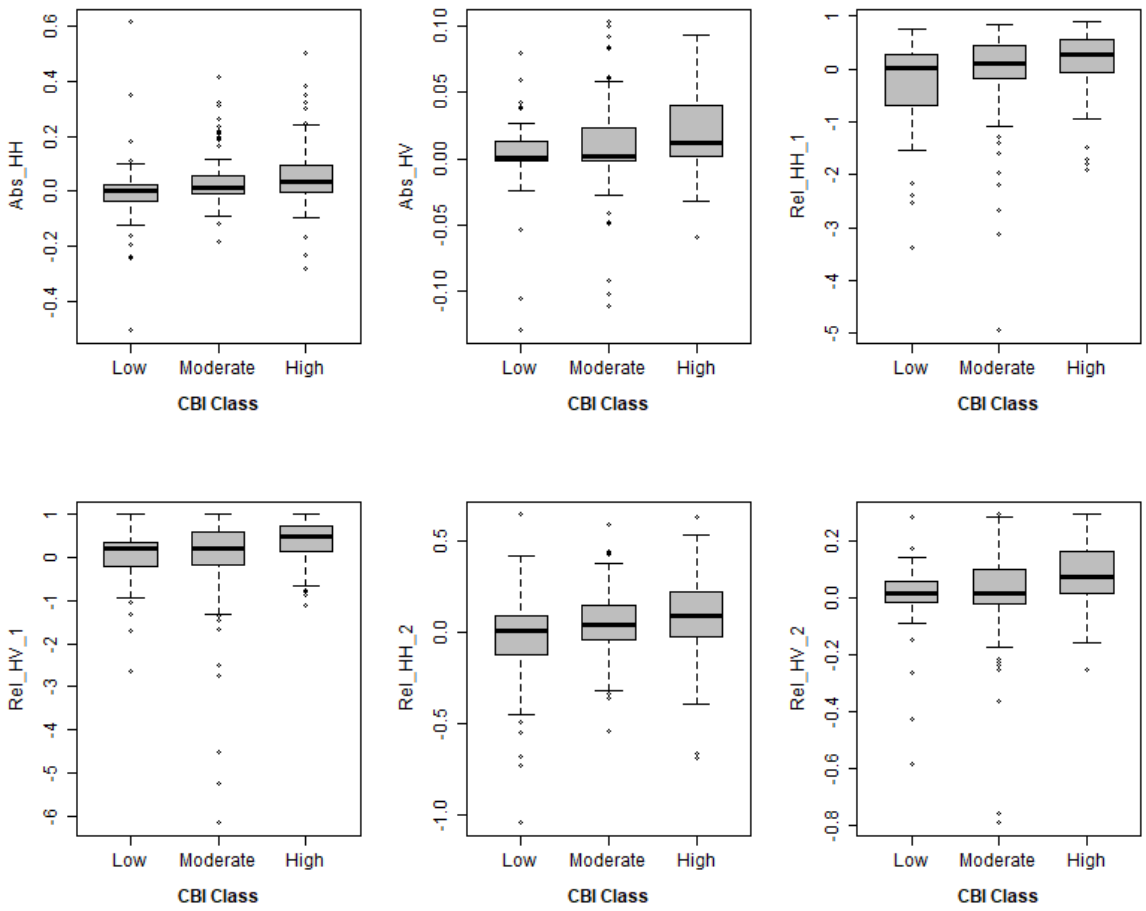

Figure 2.5 Boxplots showing how the six continuous predictor variables vary with the CBI classes of low, moderate, and high burn severity.

We then run a predictor selection routine that tested the performance of candidate models with successively fewer predictors (Dillon et al., 2011b and Birch et al., 2015). By this, we assessed the influence of each individual predictor on the model as a whole. We examined the variable importance from the C5.0 tree modeling process. With each model run, C5.0 tree calculates variable importance by randomly permuting the values of each variable, one at a time, and calculating the change in overall model performance (mean decrease in accuracy) as a result. We determined the rankings of stable predictor importance by running 10 reproducible C5.0 tree models with all the nine predictors. From here, we determined a single value of importance for each predictor from the mean variable importance of all 10 candidate models and ranked them (1= highest importance;

9 = least importance). The results identified REL_HV_1, evergreen forest, and Abs_HV to be most informative, having an aggregated ranking of less than a threshold of 50 (Figure 2.6).

Using the three final predictors, REL_HV_1, evergreen forest, and Abs_HV, an 80-20 data sampling split was made without replacement. 80% of the data was used to train the model and 20% was used to validate it. This split was done using stratified random sampling to ensure that representative distributions of the response variable were sampled in each set since our class sizes were imbalanced. A leave-one-out cross validation resampling was applied to determine the number of trials needed to achieve optimal model performance. An interval of 1 – 30 was set as the candidate for this process. The results from the resampling were then aggregated into a performance profile which revealed optimal number of trials to be 6. Setting this as the optimum number of trials, the model was developed a final time using the entire training data.

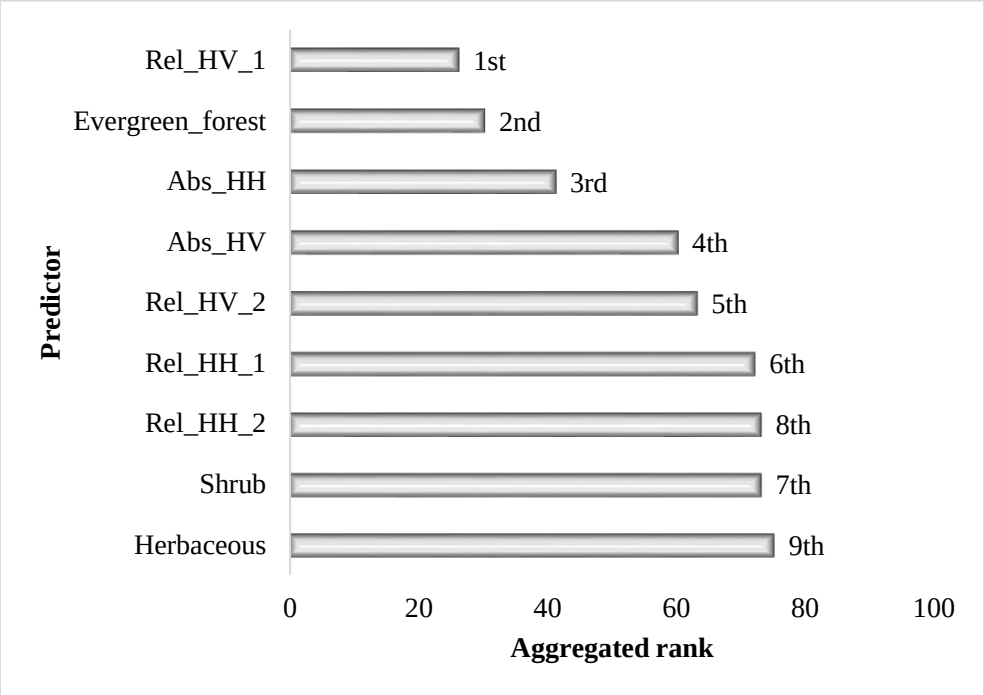

Figure 2.6 Aggregated ranking of variable importance for 10 candidate model runs for all nine predictors

2.2.4.3 Model Validation

To first establish the baseline to compare our SAR analyses to, we determined a confusion matrix (Table 2.3) using the CBI values for the 15 fires as the reference data and the BARC-4 data as the predictor. We then tested our developed SAR model on the initial 20% validation set that was retained after preprocessing.

From this we determined the overall accuracy (Eq. 9), which measures the overall performance of the model in correctly identifying the three different burn severity classes. It ranges from 0 to 1, with 0 representing a model with no predictive capability and 1 representing a perfect model. Cohen's Kappa (Eq. 10), also provides a measure of the overall performance of the model by measuring the precision between predictions and observations while correcting for chance agreements (Cohen, 1960). This statistic takes into account the possibility of chance predictions hence it tends to be more conservative. Typical values from 0.30 – 0.50 on a scale of -1 to 1 represent a model with reasonable agreement (Kuhn and Johnson, 2013).

Table 2.3 A three class confusion matrix, setting the "high" class as the true positive

		Predicted Class		
		Low	Moderate	High
Actual Class	Low	True Negative (TN)	False Negatives (FN)	False Positives (FP)
	Moderate	False Negatives (FN)	True Negative (TN)	False Positives (FP)
	High	False Negatives (FN)	False Negatives (FN)	True Positives (TP)

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (9)$$

$$Kappa = \frac{p_o - p_e}{1-p_e} \quad (10)$$

where p_o is the total accuracy given as: $p_o = \frac{TP+TN}{TP+FP+TN+FN}$, and p_e is the random accuracy given as: $p_e = \frac{(TN+FP)+(TN+FN)+(FN+TP)+(FP+TP)}{(TP+FP+TN+FN)^2}$

2.3 Results and Discussion

Values of the continuous dNBR index ranged from -0.20 to 0.84; these corresponded to continuous CBI values from 0 to 3 across the 15 burned areas. A plot of the continuous values of the CBI confirmed the challenges of the dNBR having difficulty in differentiating between moderate and high severity burns, as there did not seem to be much separation between these two classes (Figure 2.7).

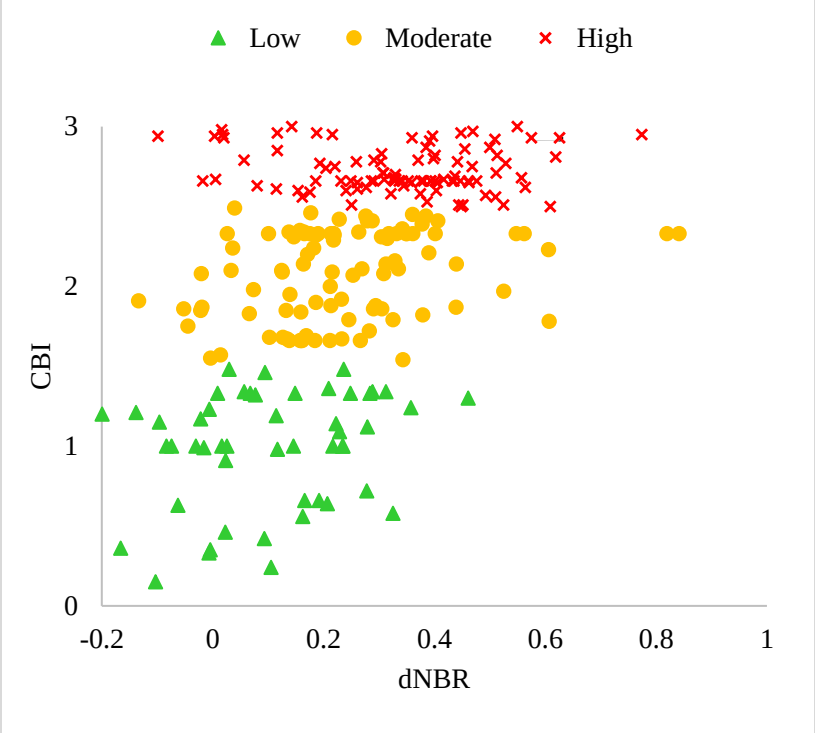

Figure 2.7 A plot of continuous values of the CBI against the continuous values of dNBR

A baseline confusion matrix between CBI and the BARC-4 data also confirmed this difficulty, as only ~21% and ~26% of the moderate and high severity classes, respectively, were correctly identified (Table 2.4). The confusion matrix shows the low severity class to have the highest producer's accuracy of 80%, which goes to show that the reflectance-based BARC-4 does a good job of correctly predicting low severity locations, however, its ~28% user's accuracy is problematic because here, a large count of moderate and high severity locations are being incorrectly classified as low severity.

With this baseline established, we developed our SAR model with the three final predictors to check against. Figure 2.8 and Table 2.5 show the breakdown of the tree that was developed from our C5.0 tree algorithm. This tree had one root node, four branches and six leaves. The algorithm identified the Rel_HV_1 predictor as its root. This makes the Rel_HV_1 predictor the most important, as it forms the base of every decision made. It therefore means that relating the response of the landscape to its condition before the fire occurrence makes it possible to better discern the different severity classes.

Table 2.4 Confusion matrix of test data CBI (columns) vs. BARC-4 data (kappa = 0.10)

		CBI			User's Accuracy (%)
		Low	Moderate	High	
BARC-4	Low	**8**	14	7	27.6
	Moderate	1	**4**	7	33.3
	High	1	1	**5**	71.4
Producer's Accuracy (%)		80.0	21.1	26.3	35.4

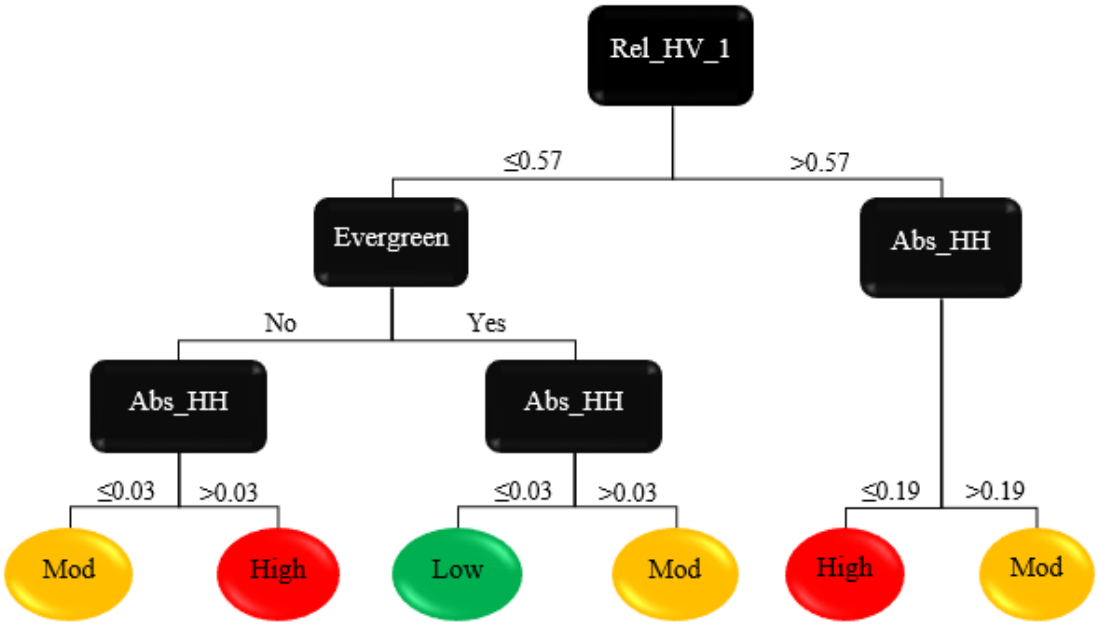

Figure 2.8 Break down of the decisions for developing the final SAR burn severity model

The Evergreen decision branch generally classified locations where this land cover was present as low to moderate severity and locations without evergreen forests as moderate to high severity. This means that the likelihood of having a high severity burn is low when evergreen forests are present. A critical look at this predictor confirmed that the low severity locations had the majority of this land cover type whereas the high severity had the least (Figure 2.9).

Table 2.5 Summary of decision paths for each leaf node

Leaf	Decisions	Burn Severity
1	$Rel\_HV_1 \leq 0.57$ Evergreen = No $Abs\_HH \leq 0.03$	Moderate
2	$Rel\_HV_1 \leq 0.57$ Evergreen = No $Abs\_HH > 0.03$	High
3	$Rel\_HV_1 \leq 0.57$ Evergreen = Yes $Abs\_HH \leq 0.03$	Low
4	$Rel\_HV_1 \leq 0.57$ Evergreen = Yes $Abs\_HH > 0.03$	Moderate
5	$Rel\_HV_1 > 0.57$ $Abs\_HH \leq 0.19$	High
6	$Rel\_HV_1 \leq 0.57$ $Abs\_HH > 0.19$	Moderate

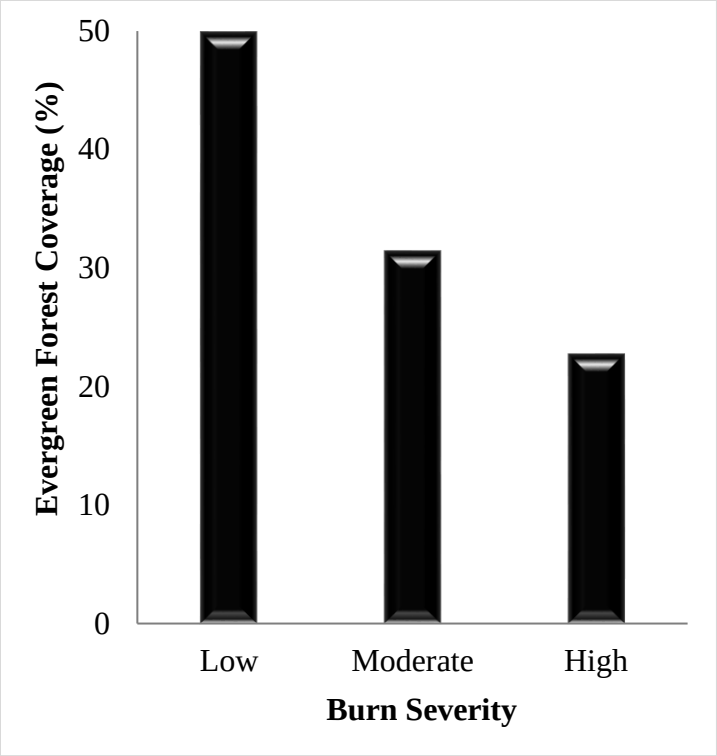

Figure 2.9 Plot showing percent coverage of evergreen vegetation in the different burn severity classes.

The HH polarization also played an integral role in the developed model, as it was the final determinant in all the six decisions. This was unexpected because past studies showed the cross-polarized HV to be a much better discerner of burn severity classes than the co-polarized HH polarization. The argument was that the strength of the backscatter from the HV polarization was much higher since by virtue of the transmitted waves going in horizontally and received vertically translates to volumetric scattering. However, we see here that the HH polarization, though not as important as the HV polarization, is crucial in discerning between burn classes.

Some other observations were that generally a low Abs_HH corresponded to low severity classification whereas higher Abs_HH values corresponded to the higher severity classes. This is intuitive because from equation 3, low Abs_HH is obtained when pre- and post-backscatter values are close to each other in magnitude. This is what is to be expected in

a low severity burn because the landscape would not have experienced that much change to translate to a large difference in recorded values. The converse is also true, as higher Abs_HH corresponded to higher severity burns.

Testing the 20% validation data set that was retained during model development on our final model gave the results presented in Table 2.6. An overall accuracy of ~60% and a kappa of 0.35 for a three-class model show good predictive performance. The model performed better in identifying moderate and high severity locations but had difficulty with the low severity locations. The fact that the low severity had the least sample size (15%) likely contributed to this underperformance, hence we are hopeful that improvement will be possible with a data rich model. Comparing this to the results obtained from the dNBR approach (Table 2.5) clearly shows the SAR approach to perform much better.

Table 2.6 Confusion matrix of test data CBI vs. SAR classified data (kappa = 0.35)

		CBI			User's Accuracy (%)
		Low	Moderate	High	
SAR	Low	**1**	1	0	50.0
	Moderate	7	**15**	6	53.6
	High	2	3	**13**	72.2
Producer's Accuracy (%)		10.0	78.9	68.4	60.4

Finally to test the importance of the land cover input, the model was tested without the Evergreen predictor. This resulted in an accuracy of 48% and kappa of 0.14. This is a clearly mediocre model in comparison and proves that knowledge of the land cover type is necessary for radar burn severity estimation. It is even possible that dNBR approach might be improved with a land cover input as well and future studies can look into it.

2.4 Conclusions and Future Work

The main focus of this study has been to develop a SAR based metric aimed at classifying the burn severity of locations in the western USA that have experienced wildfires as an alternative to the commonly used dNBR metric. We consequently compared the performance of SAR approach to the accepted dNBR approach to determine if there is any merit to its adoption as an alternative in the western USA landscape. The results showed the SAR approach to produce higher validation metrics in comparison to the dNBR. It had an overall accuracy and kappa of ~60% and 0.35 respectively in comparison to the ~35% and 0.1 of the BARC-4 data.

Generally, smaller differences in backscatter values for pre- and post- fire data translated to lower burn severities and vice versa. This was intuitive because for low severity, the landscape would not have experienced much change, thereby translating to lower differences in scattering of the radar signals and consequently, lower differences in backscatter values. Another noteworthy discovery was that the presence of evergreen forests seemed to limit the severity of fires.

Overall, this developed SAR based metric showed improved results and can therefore be recommended as an alternative to the optical sensor approach of the dNBR. Also, because the radar signals of SAR can penetrate clouds, this will be even more beneficial in emergency situations where burn severity information is needed right away to implement stabilization measures on unstable slopes.

FUNDING: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

3 Assessment of Post-Wildfire Debris Flow Occurrence Using Classifier Tree[2]

Abstract: Besides the dangers of an actively burning wildfire, a plethora of other hazardous consequences can occur afterwards. Debris flows are among the most hazardous of these, being known to cause fatalities and extensive damage to infrastructure. Although debris flows are not exclusive to fire affected areas, a wildfire can increase a location's susceptibility by stripping its protective covers like vegetation and introducing destabilizing factors such as ash filling soil pores to increase runoff potential. Due to the associated dangers, researchers are developing statistical models to isolate susceptible locations. Existing models predominantly employ the logistic regression algorithm, however, previous studies have shown that the relationship between the predictors and response is likely better predicted using nonlinear modeling. We therefore propose the use of nonlinear C5.0 decision tree algorithm, a simple yet robust algorithm that uses inductive inference for categorical data modeling. It employs a tree-like decision making system that makes conditional statements to split data into homogeneous classes. Our results showed the C5.0 approach to produce stable and higher validation metrics in comparison to the logistic regression. A sensitivity of 81% and specificity of 78% depicts improved predictive capability and gives credence to the hypothesis that data relationships are likely nonlinear.

[2]This material has been published as an Open Access article in the Geomatics, Natural Hazards and Risk journal. Check page v to see the citation for this.

3.1 Introduction

An average of 350 million hectares of land are affected annually by wildfires worldwide (van der Werf et al., 2006). There are predictions of even further increase in these occurrences with increasing trends in temperature (Westerling et al., 2006; Bond-Lamberty et al., 2007; Miller et al., 2009; Flannigan et al., 2009). The hazards associated with a wildfire continues even after it is contained. Its aftermath can yield itself to a spectrum of post-effects, of which debris flows are at the most disastrous end (Brock et al., 2007; Cannon et al., 2010). Debris flows are fast-moving, high-density slurry of water, sediments and debris that travels under gravity and are endued with enormous destructive power (Cannon and Gartner 2005, Cannon et al., 2010, Staley et al., 2017). They usually occur after periods of intense, short duration precipitation on steep hillslopes covered with loose erodible material (Cannon et al., 2010). They are very destructive to anything in their paths and are usually associated with fatalities. A recent event occurring on January 09, 2018 in Montecito, California resulted in 21 confirmed deaths, hundreds of injuries, and extensive damage to infrastructure including roadways, causing major freeways to be closed for days (CNN, 2018; Fox News, 2018). This debris flow event occurred following the largest fire in California's recent past, Thomas Fire, which had been ignited a month prior and consumed a little over 280,000 acres of land.

Although debris flows are not exclusive to fire affected areas, wildfires have been known to increase the susceptibility of otherwise stable locations (Cannon et al., 2010). With the upward trend in wildfire frequencies, it is likely these events will also be on the rise. The threat of debris flow occurrence can persists for years after wildfires. A 2015 study showed that a location that experiences a wildfire event can be at risk of producing debris flows for an average of up to two years after the fire after which period the risk reduces due to repopulation of vegetation (DeGraff et al., 2015). A wildfire increases the susceptibility of a basin to debris flow occurrences by several mechanisms, which all work to decrease infiltration and increase the runoff generated. Some of these factors include the burning off of the rainfall intercepting vegetation (Kinner and Moody, 2008; Wondzell and King, 2003; Cannon et al., 2010), sealing of the soil pores from the

generated ash (Cannon and Gartner, 2005; Larsen et al., 2009), and predominantly, the condensation of water repellent organic compounds which leave in their wake a water repellent soil that is primed for the movement of large volumes of erodible material (DeBano, 2000; Doerr et al., 2000; Cannon and Gartner, 2005; Moody and Ebel, 2012).

Accurately predicting the behavior of debris flow is decidedly difficult due to their complex physical structure, triggering mechanisms, and regional variability. There are records of studies dating as far back as the 1980s that began the onerous task of detailing the physical behavior and impact of debris flows (Hungr et al., 1984; Bovis and Jakob, 1999; Gartner et al., 2008; Cannon et al., 2010; Staley et al., 2017). With the knowledge that wildfires increase a location's susceptibility, scientists in the past have utilized this information to develop statistical models to isolate these potentially hazardous locations. Using information on the severity of the wildfire together with other pertinent descriptors of the affected location such as basin morphology, rainfall characteristics, and soil properties have been utilized in developing these predictive models (Hungr et al., 1984; Bovis and Jakob, 1999; Gartner et al., 2008; Cannon et al., 2010; Staley et al., 2017).

Generally, there are two different approaches for predicting the likelihood of post-wildfire debris flow occurrences: deterministic (Hungr et al., 1984; Johnson et al., 1991) and probabilistic (Cannon et al., 2010; Staley et al, 2017). Researchers at the United States Geological Survey (USGS) have spearheaded work using the probabilistic approach since it provides objective results even with scanty or low quality datasets (Hammond et al., 1992; Donovan and Santi, 2017). They have developed probability models using datasets on past debris flow events such as basin morphology, burn severity, rainfall characteristics, and soil properties to build logistic regression models that predict the statistical likelihood of post-fire debris flow occurrence in western United States (Canon et al., 2010, Staley et al., 2013; Staley et al., 2017). This work began in 2005 but there have been several refinements to the models over the years (Cannon and Gartner 2005, Cannon et al., 2010, Staley et al., 2017). The logistic regression approach is advantageous mostly because it considers simple linear relationships which are computationally faster and easy to interpret. Up until 2017, the best available logistic

regression model developed by USGS researchers for the intermountain west United States reported a sensitivity of 44% (Cannon et al., 2010) — this translates to an approximate of 4 out of 10 potential hazardous debris flow events being correctly predicted. This classifier had each of the input predictors modelled to influence the response variable independently, as such, probabilities greater than the cutoff points occurred even when the rainfall input was zero (Cannon et al., 2010). This was problematic because it is impossible for debris flows to occur in the absence of a driving high intensity, short duration rainfall (Cannon et al., 2010; Gartner et al, 2005). In a bid to improve upon this, in 2016, USGS researchers added more samples to the initial 2010 dataset to investigate if the now data-rich database could improve the initial model (Staley et al., 2016). The data size was increased from 608 samples in 2005 (Gartner 2005 et al., 2005) to 1550 samples in 2016 (Staley et al., 2016). Also, this new study introduced a link function whereby the critical inputs of the basin characteristics were multiplied by the rainfall inputs to ensure that the response probability was close to zero when there was no rainfall event. The best of these updated models had an improved sensitivity of 83% as compared to the previous 44%, with a corresponding specificity of 58% (Staley et al., 2017).

Other researchers have also looked into nonlinear probability modeling approach to investigate if more of the complex relationships between basin predictors and debris flow occurrence, which might not be discernible to linear models, can be captured with the nonlinear approach. Kern et al., 2017 explored the use of machine learning algorithms to model debris flow response. Their study explored both linear and nonlinear relationships between the predictors and response variable. They compared the accuracies offered by different linear and nonlinear models using the same dataset in Cannon et al. 2010's study. Their results showed the nonlinear models outperformed the linear ones by as much as ~64% giving credence to their hypothesis that the relationship between basin predictors and the debris flows occurrence might be a nonlinear one. The top model identified from the Kern et al. (2017) study was one that was built using the Naïve Bayes algorithm. This model resulted in a sensitivity of 72%, an improvement on the 44% that

was initially obtained from Cannon et al., 2010's study, and a corresponding specificity of 90% showing improved ability to predict these debris flow locations with the nonlinear model.

We are therefore proposing the application of nonlinear modeling to the 2016 dataset as well to further improve the debris flow prediction. Preliminary analyses done with the Naïve Bayes algorithm resulted in a sensitivity 75% of and a specificity of 81%, showing improved results. However, in this current study, we propose the use of the nonlinear C5.0 tree algorithm (Quinlan, 1993) as opposed to the Naïve Bayes algorithm because the Naïve Bayes model is a black box model whose inner workings are unknown. It does not offer any insight into the relationships of the predictors as they relate to the response. The C5.0 tree algorithm, was therefore chosen in particular because it is one of the simplest nonlinear algorithms with transparent outputs. It affords a nonlinear approach by identifying unique cutoffs in the different distributions of the predictors as they relate to the response. The algorithm works by splitting the data into smaller, more homogeneous groups. Stepwise decisions are made on predictors at different levels to iteratively determine unique breakpoints as they relate to the different classes of the response variable. C5.0 builds trees from a set of training data using concepts from information theory. The algorithm makes different decisions at different nodes in an attempt to sort the response variable into its homogenous classes. To determine which predictor to choose to ask which question at each node, it uses a concept called *information entropy* (H), a statistic measure for the average rate at which information is produced by a stochastic source of data (Shannon, 2001). Essentially, H calculates the uncertainty in any particular decision at each node. Shannon defined H of a discrete random variable, X, with possible values (x1, x2,… xn) and probability mass function P(x) as:

$$H = -\sum_{i=1}^{n} P(x_i) \log_b P(x_i) \quad (1)$$

where b is usually taken as base 2. For a binary response like in this project's case, "no" debris flow and "yes" debris flow, the entropy distribution looks like Figure 3.1 below.

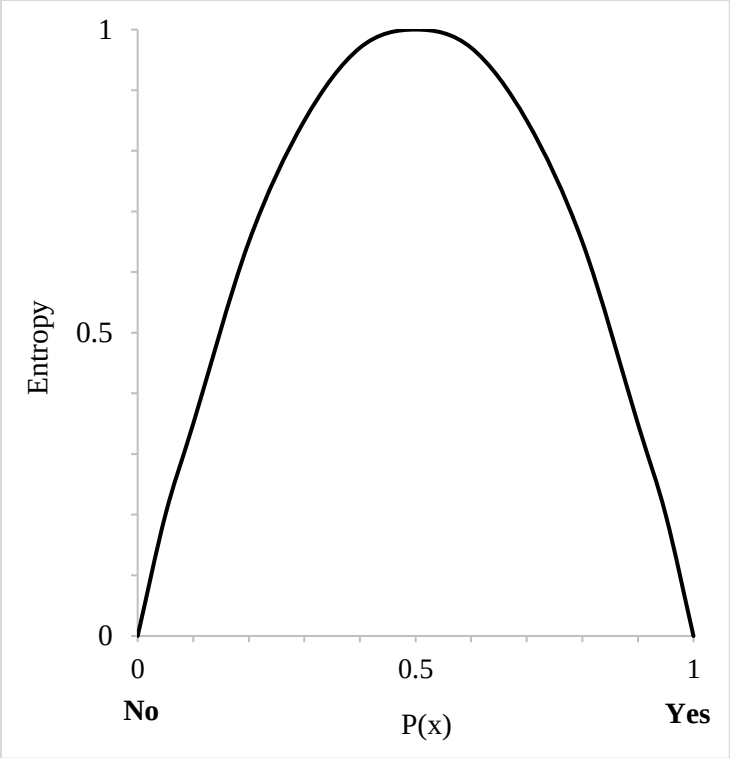

Figure 3.1 A plot of a binary entropy function showing the distribution of entropy (uncertainty) with changes in the probability of response classes. Uncertainty is lowest when the probability approaches 0 or 1, and reaches maximum when probability is 0.5.

Entropy reaches a maximum at the halfway point when the probability is 0.5; i.e. there is a 50-50 chance that it could go either way ("no" or "yes"), uncertainty is at its maximum. It is lowest when the probability approaches 0 or 1. The goal is to choose the predictor which gives us the lowest entropy. Moving on from there the process is repeated for the node below, however, this time the *gain*, measure of entropy change, is also determined. This is to assess the magnitude of information increase in comparison to a prior node (Mitchell, 1997; Shannon, 2001).

$$Gain = Entropy_{prior} - Entropy_{current} \quad (2)$$

All candidate predictors are considered and the one with the largest information gain is chosen for the decision using a greedy system. This process is applied recursively from the root node down until a subset contains only samples of a single class, or the partitioning tree has reached a predetermined maximum depth. The C5.0 tree algorithm uses boosting in its model training process, which works by combining average model decisions together to improve overall model performance of the final output. The particular objective of this study was to investigate the applicability of the nonlinear C5.0 tree algorithm in predicting the likelihood of post-wildfire debris flow occurrences in the western USA and to determine if there is any advantage to be obtained over the linear logistic regression approach with this new dataset.

3.2 Method

3.2.1 Data

Data for this study was obtained from USGS' open file report 2016-1106 (Staley et al., 2016). It comprises descriptive information on 1550 burned basins in western United States (Figure 3.2). Besides the response variable, there were a total of 16 predictors, which included information on the basin morphology, burn severity, soil properties, rainfall characteristics, and other ancillary data. Brief summaries of predictors have been provided in Table 3.1.

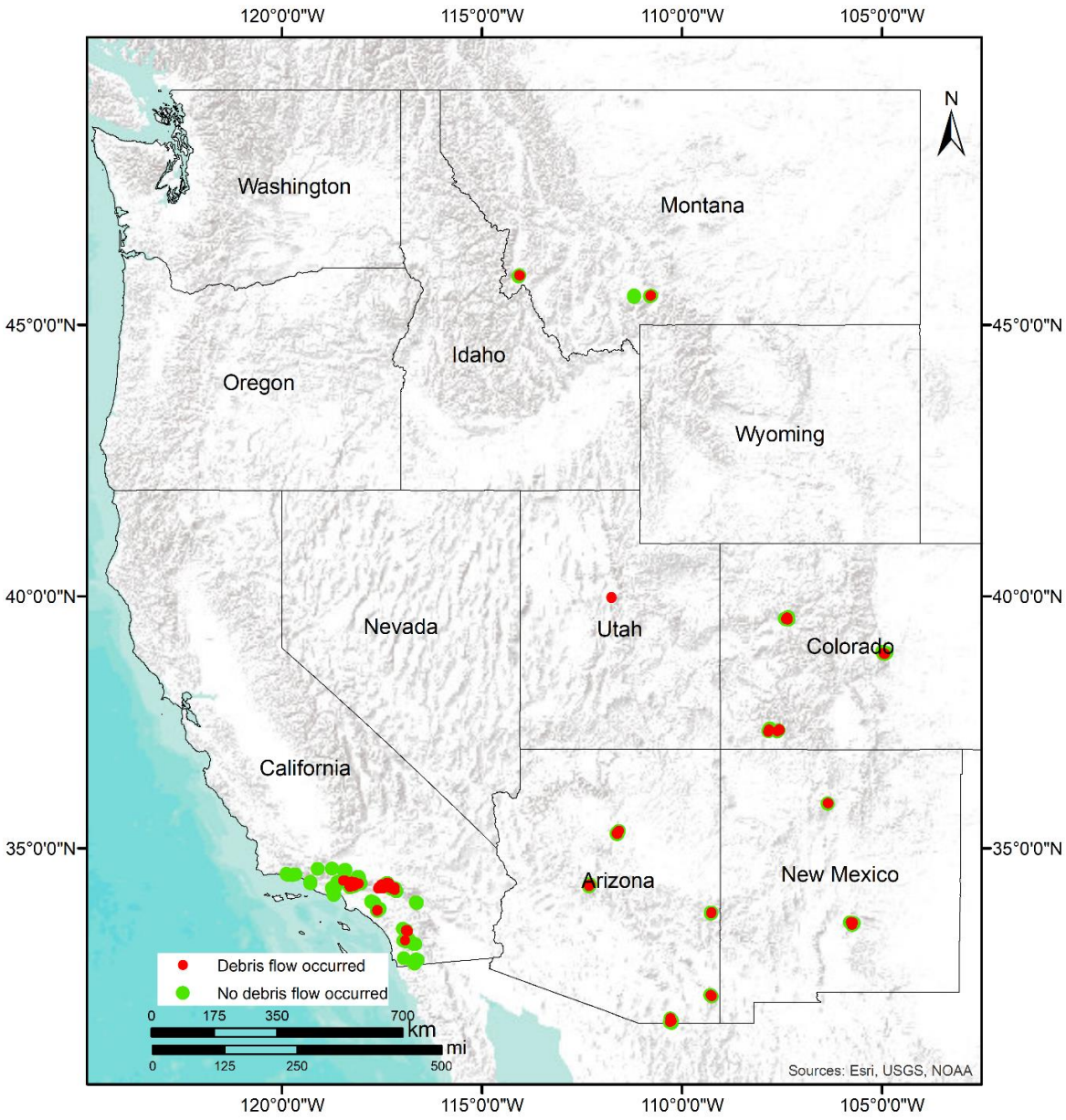

Figure 3.2 Map showing the 1550 burned basins considered in this study. Red dots show locations that experienced debris flows and the green dots show locations that experienced none

Table 3.1 Brief descriptions of model predictors

Predictor	Description
Acc015	Peak 15-minute rainfall accumulation of storm, in millimeters
Acc030	Peak 30-minute rainfall accumulation of storm, in millimeters
Acc060	Peak 60-minute rainfall accumulation of storm, in millimeters
Area	Contributing area of observation location, in square kilometers
dNBR	Average differenced normalized burn ratio of watershed
GaugeDist	Distance from rain gauge to documented response location, in meters
Iave	Average storm intensity, in millimeters per hour
KF	Average KF-factor a basin. Also known as the *erodibility* factor. It is the susceptibility of soil particle to detached and get transported by rainfall.
Peak_I15	Peak 15-minute rainfall intensity of storm, in millimeters per hour
Peak_I30	Peak 30-minute rainfall intensity of storm, in millimeters per hour
Peak_I60	Peak 60-minute rainfall intensity of storm, in millimeters per hour
PropHM23	Proportion of watershed burned at high or moderate severity and with slope>23º
Soil	Thickness of soil to the closest “restrictive layer” that significantly impede the movement of water and air through the soil.
StormAccum	Total rainfall accumulation of storm, in millimeters
StormDate	Date of storm that produced the debris-flow response
StormDur	Total duration of storm, in hours

3.2.2 Model Development

The data from Staley et al., 2016 were all given as independent predictors. As was done in Staley et al., 2017, we introduced a link function by multiplying the morphological and burn properties predictors by rainfall predictors, since debris flows cannot occur in the absence of a driving storm. A total of 35 compound predictors were obtained after this. Preliminary data pre-processing steps taken to ensure optimal model performance included omitting observations with missing values and assessing predictor degeneracy. We also observed the existence of correlations between predictors so we performed a pairwise collinearity test. The results showed that 15 predictors had 99% or more correlations with other predictors. These were regarded as redundant information and were thus deleted, leaving 20 predictors for further analyses. We run a predictor selection routine that tested the performance of candidate models with successively fewer predictors (Dillon et al., 2011 and Birch et al., 2015). By this, we assessed the influence of each individual predictor on the model as a whole. We examined the variable importance from the C5.0 tree modeling process. With each model run, C5.0 tree calculates variable importance by randomly permuting the values of each variable, one at a time, and assessing the overall improvement in the optimization criteria (accuracy, in this case). We determined the rankings of stable predictor importance by running 10 reproducible C5.0 tree models with all the 20 remaining predictors. From here, we determined a single value of importance for each predictor from the mean variable importance of all 10 candidate models and ranked them (1=highest importance; 20 = least importance). A threshold of 50 was applied, narrowing down the predictor size to five most informative. Finally, to ensure that predictors were truly independent we performed a pairwise collinearity test once more and applied a cutoff of 60%. This resulted in these three final predictors for modelling: Soil* peakI30, PropHM23*peakI30, and KF*peakI30.

Using the final three predictors, we split the data into 70% training and 30% validation sets using stratified random sampling to ensure that representative distributions of the response variable were represented in each set, since the data was skewed towards the

"no debris flow" response at about 3:1. A repeated ten-fold cross validation resampling was applied to determine the number of trials needed to achieve optimal model performance. An interval of 1 – 30 was set as the candidates for this process. The results from the resampling were then aggregated into a performance profile which revealed optimal number of trials to be 11. Setting this as the optimum number of trials, the model was developed a final time using the entire training data.

3.2.3 Model Evaluation

The developed model was then tested on the initial 30% validation set that was retained after preprocessing. To define the performance profile, we first generated a confusion matrix (Table 3.2) to summarize how the model's predictions performed against the actual data.

From this we determined the sensitivity (Eq. 1), which measures the fraction (or percentage) of debris flow producing locations that were correctly predicted by the model. This metric was given the highest priority because it gives a direct measure of the objective of the study. It runs from 0 to 1, with 1 representing a perfect model that correctly classifies all the debris flow locations. The specificity (Eq. 2) was also determined as the measure of the fraction (or percentage) of locations that did not produce any debris flows that were correctly predicted by the model. This metric also runs from 0 to 1, with 1 representing a perfect model that correctly classifies all the no debris flow locations. We were also interested in this metric since it assesses the model's robustness in preventing false positives.

The third metric was the overall accuracy (Eq. 3), which measures the overall performance of the model in correctly distinguishing between debris flow and no debris flow locations. A score of 1 indicates a perfect model and 0 indicates a model with no predictive capability. Threat score (Eq. 4), also known as the critical success index (Schaefer, 1990), was also considered. It is another metric that measures a model's overall performance. It is especially used when the class distributions of the response variable are as skewed toward one class, as is present in our data. This metric, however,

does not consider true negative (TN) events. Cohen's Kappa (Eq. 5), also provides a measure of the overall performance of the model by measuring the similarity between predictions and observations while correcting for agreement that occurs by chance (Cohen, 1960). The kappa statistic tends to be quite conservative but it is a more robust measure than the overall accuracy since it takes into account the possibility of chance predictions. Therefore typically, values within 0.30 to 0.50 on a scale of -1 to 1 indicate reasonable agreement (Kuhn and Johnson, 2013).

$$Sensitivity = \frac{TP}{TP+FN} \quad (1)$$

$$Specificity = \frac{TN}{TN+FP} \quad (2)$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

$$Threat\ score = \frac{TP}{TP+FP+FN} \quad (4)$$

$$Kappa = \frac{p_0-p_e}{1-p_e} \quad (5)$$

where p_o is the total accuracy given as: $p_o = \frac{TP+TN}{TP+FP+TN+FN}$, and p_e is the random accuracy given as: $p_e = \frac{(TN+FP)+(TN+FN)+(FN+TP)+(FP+TP)}{(TP+FP+TN+FN)^2}$

Table 3.2 Confusion matrix

		Predicted Class	
		Yes	No
Actual Class	Yes	True Positives (TP)	False Negatives (FN)
	No	False Positives (FP)	True Negative (TN)

We evaluated the validation data on all these five metrics together because each of these metrics has its own biases, hence, using them together gave a better picture of the model's performance. For example, by virtue of the skewed nature of the data, if we consider the threat score metric, it will afford us the ability to prioritize the low frequency debris flow locations, since it ignores the TNs in its computation. However, considering this metric alone would have meant that we would not have been able to fully assess the influence of the false positives in relation to how many no-debris flow locations were present. That information is necessary to assess whether or not the model holds the risk of desensitizing the public, showing the need for considering all these metrics. To further test robustness and stability, the entire modeling process from resampling, training, and validation was repeated for ten different combinations of data samples and the best amongst these was selected.

3.3 Results and Discussion

In this study, we used C5.0 decision tree algorithm to develop predictive models with the aim of isolating locations in western USA that will likely produce post-wildfire debris flows. Ten candidate models were built in order to investigate the robustness and stability of the algorithm. The average of validation metrics for the candidate models have been presented in Table 3.3. The results gave stable as well as high metrics for all ten candidates with an average sensitivity of 77%, which is an improvement on the 66% from the logistic regression method (Table 3.4), thereby giving confidence in the use of the nonlinear C5.0 algorithm for our study. This algorithm was robust in distinguishing between the two classes of the response, even though the data distribution was skewed towards the "no debris flow" class. The 2017 study done on this same dataset using logistic regression approach (Staley et al., 2017) reported validation outputs as TPs, TNs, FPs, and FNs hence, we have computed the corresponding sensitivities, specificities, kappas, and accuracies (Table 3.4), to allow for direct comparison of the two different approaches.

Table 3.3 Validation metrics of the 10 candidate models developed with the C5.0 tree algorithm

Model	Validation Metrics				
	Accuracy	**Sensitivity**	**Specificity**	**Kappa**	**Threat Score**
1	0.77	0.82	0.75	0.47	0.45
2	0.80	0.71	0.83	0.49	0.46
3	0.78	0.81	0.78	0.49	0.47
4	0.81	0.74	0.84	0.53	0.49
5	0.76	0.78	0.75	0.45	0.44
6	0.78	0.78	0.78	0.49	0.46
7	0.80	0.77	0.81	0.52	0.48
8	0.80	0.76	0.81	0.51	0.47
9	0.79	0.75	0.80	0.49	0.46
10	0.76	0.74	0.77	0.44	0.43
Average	**0.79**	**0.77**	**0.79**	**0.49**	**0.46**

Table 3.4 Validation metrics computed from the results of the logistic regression models developed by Staley et al., 2017

Model	Validation Metric				
	Accuracy	**Sensitivity**	**Specificity**	**Kappa**	**Threat Score**
M1_15	0.64	0.81	0.58	0.30	0.39
M1_30	0.67	0.71	0.66	0.29	0.34
M1_60	0.73	0.46	0.81	0.26	0.27
M2_15	0.64	0.81	0.57	0.30	0.38
M2_30	0.67	0.74	0.65	0.30	0.35
M2_60	0.74	0.46	0.82	0.27	0.28
M3_15	0.60	0.84	0.52	0.26	0.36
M3_30	0.62	0.82	0.56	0.27	0.34
M3_60	0.71	0.48	0.76	0.20	0.24
M4_15	0.61	0.79	0.54	0.25	0.36
M4_30	0.61	0.67	0.58	0.20	0.31
M4_60	0.69	0.37	0.79	0.16	0.23
Average	**0.66**	**0.66**	**0.65**	**0.26**	**0.32**

Comparing Table 3.3 to Table 3.4, we can see that the logistic regression algorithm produces generally lower metrics than the C5.0 approach. Its highest sensitivity of 84% corresponded to a specificity of 52%, which goes with the general theme of its results, where high sensitivities corresponded to low specificities, and vice versa. In other words, the model had a harder time isolating actual debris flow locations without lumping some of the no debris flow locations (false positives) with them. This is likely due to the linear nature of the logistic regression algorithm being unable to capture much of the complex relationships to discern between little nuances as they relate to the likelihood of debris flow occurrence.

The C5.0 tree, on the other hand, affords a nonlinear approach by identifying unique cutoffs in the different distributions of the predictors as they relate to the response. Model 3 from Table 3.3 was chosen to be the overall best model since it had a high sensitivity with an equally high specificity, as well as a simple tree structure. A confusion matrix of the resulting classification of the test data for this model is presented in Table 3.5. The resulting sensitivity is 81% and the specificity is 78%, which shows an increased capacity to correctly identify ~8 out of 10 of these hazardous debris flow locations with a very low risk (~22%) of collating numerous false positives in the process.

Table 3.5 Confusion matrix of the top model using the C5.0 tree algorithm

		Predicted Class	
		Yes	No
Actual Class	Yes	75	18
	No	67	234

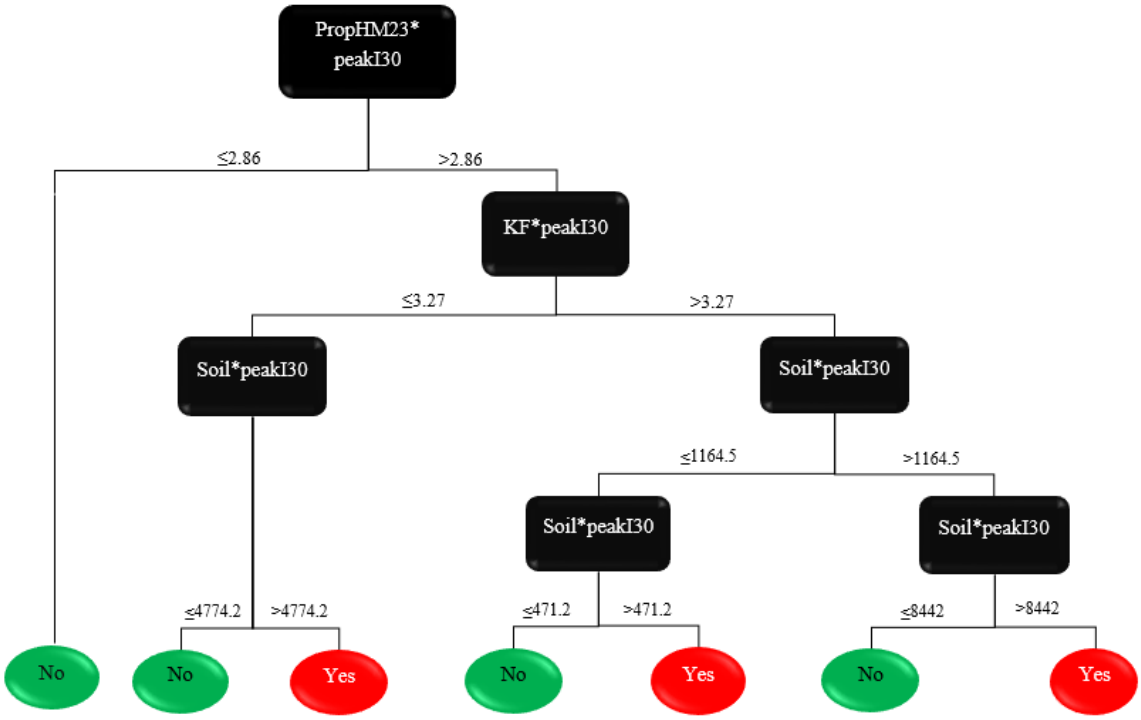

Figure 3.3 Graphical representation of the adopted C5.0 tree model

The comprehensive list of rainfall predictors investigated in this study include rainfall accumulations in 15, 30, and 60 minutes, respectively, peak rainfall intensities in 15, 30, and 60 minutes, respectively, total rainfall accumulation as well as average rainfall intensity. A look at the decision tree plot (Figure 3.3), however, reveals that it was only the peakI30 rainfall predictor that the model found to be most informative. This suggests that it does not take much time (at most 30 minutes) for a post-wildfire debris flows to be triggered once an intense storm (peak) starts, which agrees with what is found in literature (Cannon et al., 2010).

As was discussed earlier, the rainfall predictors are very essential because debris flows cannot occur in the absence of a triggering storm event. However, since the peakI30 predictor was multiplied to each of the other predictors used in the study, it will be considered as a constant in the following discussion of the tree decisions. The generated simple tree comprised a root, five branch nodes and seven leaf nodes; therefore there were seven different decision paths. The accuracy of the decisions were individually

~60% or better. The most important predictors were the PropHM23, which indicates the proportion of moderate-high burn severity locations that have slopes that are 23% or higher, as well as the Soil predictor, which measures the overburden thickness over a "restrictive layer" such as bedrock, cemented layers, frozen layers etc. The PropHM23 was the root of the tree and thus formed the basis of every decision made, whereas the Soil was the "tie-breaker" for ~86% of the decisions.

A quick overview of the seven decision nodes showed that the model was flexible enough to discourage the training data from overfitting to it. Taking the decision path from the root to leaf node 1, shows that fire affected location that are on higher elevations (slope>23%) but sustained mostly lower severity burns, i.e. lower moderate-high burn severity burn (PropHM23*peak_I30 < 2.9) will likely not produce any debris flows. This decision was especially reassuring because it agrees with the main premise of this study that wildfires increases the susceptibility of a location to debris flow occurrences. A look at decisions 2 through to 7 seem to suggest that the thicker the soil overburden over a "restrictive layer" (Soil*peak_I30), the higher a location's chance of generating post-wildfire debris flows. A speculation can be drawn that this is likely due to the fact that the basin will now have a greater supply of loose sediment primed for movement by a triggering rainfall. Decisions 4 and 6 further buttress this point by suggesting that even with a higher erodibility factor (KF), a basin might not be susceptible to post fire debris flows, if it has little supply of loose material to move. In summary, the comprehensive tree seems to suggest that:

- Post-wildfire debris flows are triggered by high intensity storms that occur over short periods.
- Any location with an ample steepness in terrain (slope>23) that experiences moderate to high severity burns (PropHM23) has an increased capacity to produce debris flows.
- Coupled with this, the thicker the soil overburden over a "restrictive layer", the higher the chances of generating these events.

These are in no way new discoveries as past studies have reported or alluded to them. In fact, the recent logistic regression modeling identified these similar predictors as well (Staley et al., 2017). The agreement confirms the importance of these predictors and gives rise to the recommendation to focus future studies on the cutoff points identified by the C5.0 tree algorithm. This will further our understanding of the triggering and driving forces to better prepare and mitigate future hazardous events.

3.4 Conclusions and Future Work

The main focus of this study has been to investigate the use of nonlinear C5.0 tree algorithm in predicting the likelihood of post-wildfire debris flow occurrences in the western USA and compare it to the results obtained from the logistic regression approach. The results showed the C5.0 approach to produce stable and higher validation metrics in comparison to the logistic regression. It had an average sensitivity of 77%, which is an improvement on the 66% from the logistic regression approach. Analyzing the resulting tree of the adopted C5.0 tree model buttressed the intuitive hypotheses that a location with a steep terrain (slope>23) that experiences moderate to high burn severity fires and laden with thick overburden soil is highly susceptible to such events. All these give rise to the conclusion that the relationship between the predictors and the response is nonlinear as opposed to the linear one offered by the logistic regression.

ACKNOWLEDGEMENT: This project was funded by the US Department of Transportation (USDOT) through the Office of the Assistant Secretary for Research and Technology. The authors would also like to thank the following individual and institution for their contributions and roles in making this work possible: Caesar Singh, USDOT program manager and the United States Geological Survey for making the data freely available online.

DISCLAIMER: The views, opinions, findings, and conclusions reflected in this paper are the responsibility of the authors only and do not represent the official policy or position of the USDOT/OST-R or any State or other entity.

4 Post-Fire Debris Flow Modeling Analyses: Case Study of the Post Thomas Fire Event in California[3]

Abstract: There is an increased risk in post-fire debris flow (DF) occurrences in the western USA with a recent increase in wildfire frequencies. DFs are destructive, causing high loss to lives and infrastructure. A lot of effort is going into possible preventive and/ or mitigation measures. Recently, researchers have been especially working on developing statistical models that assist emergency response efforts in isolating vulnerable locations after fires. There are two general approaches to this statistical modeling: linear and nonlinear. This study has looked into applying a linear-based logistic regression model and a nonlinear-based C5.0 decision tree model to assess the strength of each in predicting the locations within the Thomas Fire boundary that produced DFs. To do this, DF scars were delineated by running a change detection protocol known as delta normalized difference vegetation index (dNDVI), using high resolution data from Planet Labs. These scars were further validated with data gathered by the Santa Barbara County on affected areas. Results from the two statistical models were then overlain on the delineated DF scars and compared against each other to assess predictive strengths. The results showed both models to perform well in predicting high probabilities for locations that were evidenced to produce DFs. The logistic regression model predicted an overall ~44,800 ha (49%) more high hazard coverage compared to the C5.0 tree, and therefore showed greater urgency. However, a closer look at the basin predictions with the delineated DF scars showed that most of these high hazard basins by the logistic regression did not to have any discernible scars. It was projected that most of these locations were likely false positives and the model was recommended for fine-tuning to avoid desensitizing the public against this hazard. Further recommendation was made concerning the development of models to predict the potential inundation DF paths after these statistical models have isolated their origination points upstream since the most associated danger with DFs is not necessarily where they start but rather further downstream where communities and infrastructure are located.

[3]This chapter has been submitted for peer review to the *Natural Hazards* journal.

4.1 Introduction

Wildfire influenced hazards can be even more deadly after their containment. Factors such as the consumption of rainfall intercepting vegetation, sealing off of soil pores from generated ash, together with the creation of water repellent due to the condensation of organic compounds can result in decreased rainfall infiltration and subsequent significant increase in runoff (Cannon and Gartner 2005, Cannon et al., 2010). All these factors work together to increase a location's susceptibility to debris flow occurrences after it experiences a wildfire. A debris flow (DF) is a geological phenomenon whereby a soil-laden slurry of mud and rock fragments, triggered by runoff, flow under gravity and pick up any loose erodible material on its way (Cannon et al., 2010). They garner appreciable speeds which cause them to be very deadly and destructive to anything in their paths, in several cases being more destructive than the wildfires that preceded them.

The focus of this study is the Thomas Fire, which occurred in southern California (Figure 4.1). This fire was ignited on December 4, 2017, burned for about a month, consuming ~114,000 ha of land before being officially contained on January 12, 2018. Before the fire could be fully contained, a major storm occurred on January 9, 2018, which triggered DFs within the burned area, inundating several communities downstream in the Santa Barbara and Ventura Counties. The actively burning fire resulted in two fatalities whereas the ensuing DF event caused 21 fatalities (Berman, 2017; Orosco, 2017; Dolan, 2018). Not only that but many homes and roads were severely damaged, causing ~$207 million in infrastructure damage (RDN, 2018; Magnoli, 2018). It is therefore imperative that any mitigating and/or preventive measures that can be applied to lessen the effect or prevent such events altogether are done.

In a bid to support emergency response efforts researchers have developed probability models that aim at isolating especially vulnerable locations within a burned area immediately after wildfires. There exist two general probability modeling approaches: linear (Rupert et al., 2008; Cannon et al., 2010; Staley et al., 2017), and nonlinear

machine learning modeling (Kern et al., 2017, Addison et al., 2018), both of which have inherent advantages and disadvantages.

The objective of this paper is to apply representative models from these two approaches to the DF event that occurred after the Thomas Fire and assess their predictive capabilities. The models considered are the linear-based logistic regression model from Staley et al., (2017) and the nonlinear-based C5.0 decision tree model from Addison et al., (2018). To do this, DF scars were delineated by running a change detection protocol with the Normalized Difference Vegetation Index (NDVI) algorithm, using high-resolution data from Planet Labs Inc. The mapped scars were further validated with data gathered by Santa Barbara County officials on affected structures. DF probability results from the two models were then overlain on the delineated DF scars and compared against each other to assess predictive strength.

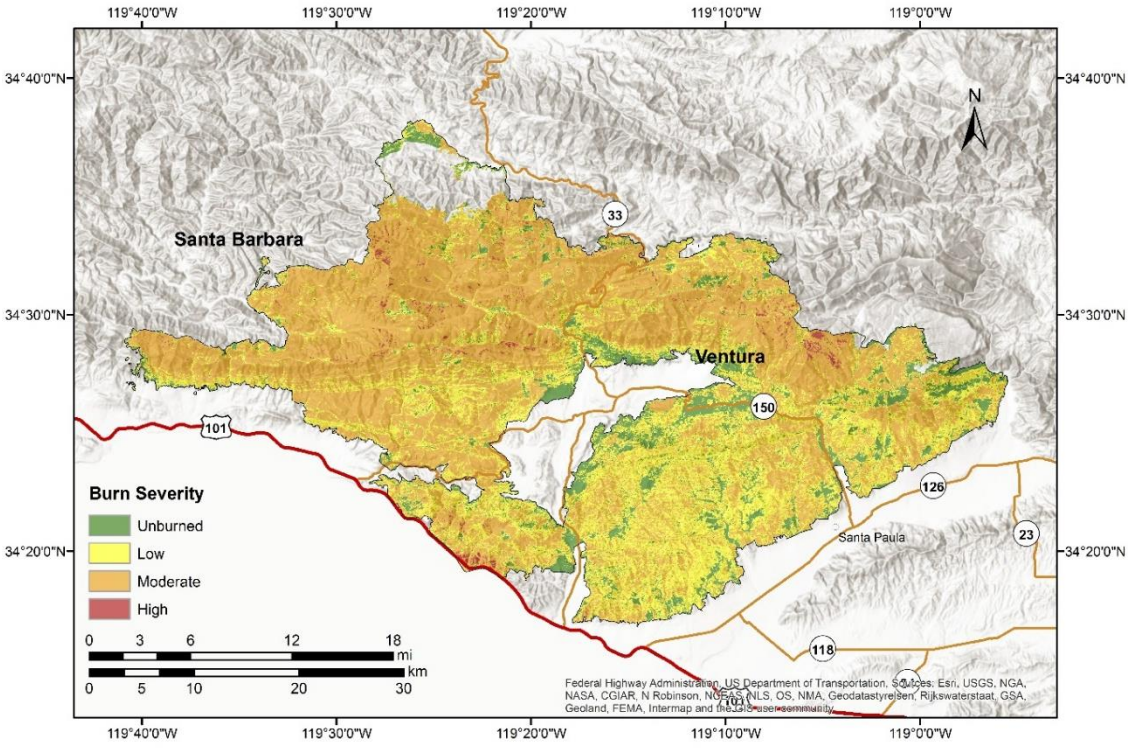

Figure 4.1 Location of the Thomas Fire in California.

4.1.1 Fire Background

The Thomas Fire was ignited north of Santa Paula, California, close to a college called Thomas Aquinas College, from which the fire derived its name. Aided by strong winds, it grew to a total of 114,000 ha, breaching the county boundaries of Santa Barbara and Ventura Counties. Analyzing burn severity data showed that 11% of the area within the burn boundary were unburned, 31% burned with low severity, 56% moderately burned, and a final 1% burned with high severity. It devastated ~1,300 structures and caused a projected ~$2.2 billion in damages (Ding, 2018; Hersko, 2018; Santa Cruz, 2018; NIFC, 2018). By the time it was contained on January 12, 2018, the Thomas Fire was the largest fire in California's recent past, after the 2003 Cedar fire, which consumed ~110, 600 ha of land. However, in the August 2018, it was surpassed by the Mendocino Complex Fire, which consumed ~63% more land, with a ~185,800 ha total coverage (Phillips et al., 2018). This was not surprising because scientists have been predicting increase in the frequencies of big fires with the onslaught of drier climates (Westerling et al., 2006; Bond-Lamberty et al., 2007; Miller et al., 2009; Flannigan et al., 2009), however, its fulfillment increases our urgency in fine-tuning emergency response tools.

4.1.2 Candidate Models

The two models considered in this study are the linear-based logistic regression model from Staley et al., (2017) and the nonlinear-based C5.0 decision tree model from Addison et al., (2018). It is important to note that these models predict the origination points of DFs and not their inundation paths downstream. The following sections provide brief overviews of each algorithm, together with their associated validation metrics.

4.1.2.1 Logistic Regression Model

Logistic regression modeling was the approach adopted by Staley et al., (2017) in their study. This is a linear modeling approach that assesses the probability (from 0-100%) of a location's potential to generate DFs after experiencing wildfires. The data utilized for model development was obtained from a database developed by researchers at the United

States Geological Survey (USGS), which included descriptive information on 1550 basins from 34 fires across the western United States (Staley et al., 2016). There were a total of 26 predictors and ancillary data, which included information on fire location, year of occurrence, burn severity, rainfall rates, soil properties, as well as topographical properties. The data was split into training and validation sets. Running multiple data iterations, four final variables were identified as the important predictors for model development (Table 4.1).

Table 4.1 Predictors used for the two candidate models: logistic regression (Staley et al., 2017) and C5.0 decision tree (Addison et al., 2018)

Predictor	**Description**	**Staley et al., 2017**	**Addison et al., 2018**
R_{15}	Peak 15-minute rainfall intensity of storm, in millimeters per hour ($PeakI_{15}$)	✓	
R_{30}	Peak 30-minute rainfall intensity of storm, in millimeters per hour ($PeakI_{30}$)		✓
X_1	Proportion of watershed burned at high or moderate severity and with gradients in excess of 23º (PropHM23)	✓	✓
X_2	average differenced normalized burn ratio of the basin (dNBR)	✓	
X_3	Average erodibility index, also known k-factor, of the fine fragments of the soil of the watershed (KF)	✓	✓
X_4	Average soil thickness to a restrictive layer, like bedrock ($Soil_{thick}$)		✓

Using the training dataset, a final logistic regression model was developed as:

$$P = \frac{e^x}{(1 + e^x)} \quad (1)$$

where, P = the probability of debris-flow occurrence,

e = the exponential function, and

$$x = (0.41 \times X_1 R_{15}) + (0.67 \times X_2 R_{15}) + (0.7 \times X_3 R_{15}) - 3.63 \quad (2)$$

where, X_1 = proportion of basin that had moderate–high burn severity with gradients ≥ 23°,

X_2 = average differenced normalized burn ratio (dNBR) of the basin,

X_3 = soil erodibility index of the basin, and

R_{15} = peak 15-minute rainfall accumulation of the design storm.

The model was then tested on the validation set, which yielded a sensitivity of 81% and a specificity of 58% (Staley et al., 2017; Addison et al., 2018). A sensitivity of 81% means that the model has the capability of correctly predicting ~8 out of 10 DF producing locations, whereas a specificity of 58% means that ~6 out of 10 "DF safe" locations within a burned area will be correctly identified. The model had an overall accuracy of 64%.

4.1.2.2 C5.0 Tree Model

Using the same dataset on the 1550 burned basins in western USA from USGS' databases (Staley et al., 2016), Addison et al., (2018) employed the C5.0 decision tree algorithm to develop a model for post-wildfire DF prediction. This is a nonlinear machine learning algorithm which uses inductive inference to split data into smaller, homogeneous classes. Stepwise decisions are made on predictors at different levels to iteratively determine unique breakpoints as they relate to the different classes of the response variable. Specifically, a general decision called the root is first developed, then branching out from there more specific decisions, known as branches, are made, till it gets to a decision that produces homogeneous samples, which are then assigned a classification known as the

leaf (Kuhn and Johnson, 2013). Figure 4.2 gives a schematic of a simplified tree with one root, two branches and four leaves. To determine which predictor to use at each node, each predictor is tested to assess how well it alone separates the training data according to the target classification. The process is repeated on each new subset until a subset contains samples of only a single class, or the partitioning tree has reached a predetermined maximum depth (Kuhn and Johnson, 2013).

The dataset was split into 70:30 training and testing sets respectively. A predictor selection routine was ran that identified the four predictors presented in Table 4.1 to be the most important for model development. Using these predictors the C5.0 tree algorithm was ran to develop the final tree structure presented in Addison et al., (2018). The tree had seven decision paths, with three belonging to the "yes" classification and four belonging to the "no". Summary of this has been presented in Table 4.2.

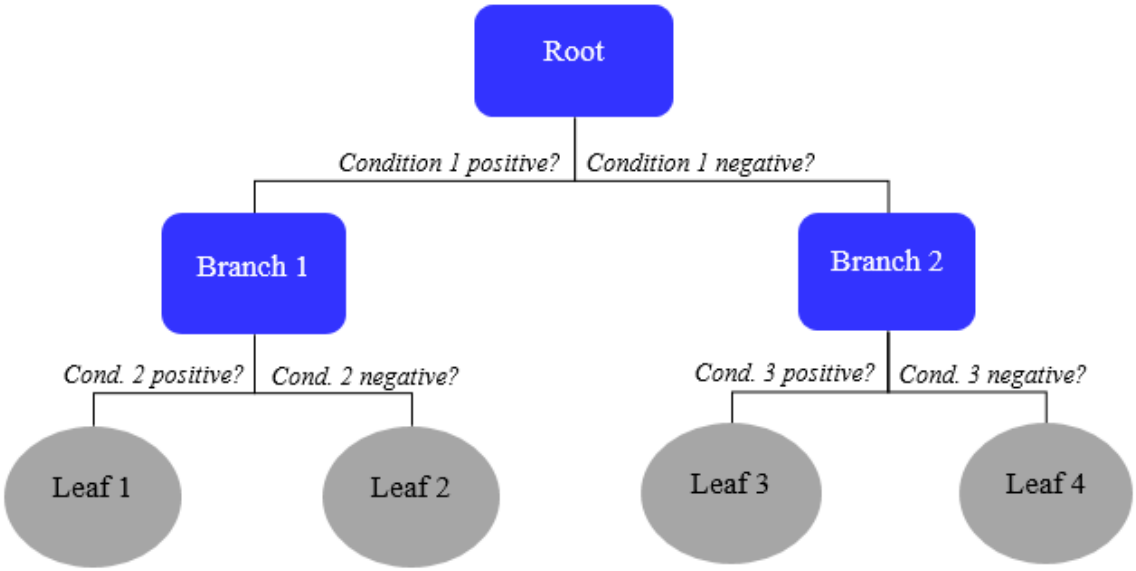

Figure 4.2 Graphical schematic of a simplified decision tree

Table 4.2. Summary of decision paths for each leaf node

Leaf	Conditions	Debris Flow
1	$(X_1 * R_{30}) > 2.86$ $(X_3 * R_{30}) \leq 3.27$ $(X_4 * R_{30}) > 4774.2$	Yes
2	$(X_1 * R_{30}) > 2.86$ $(X_3 * R_{30}) > 3.27$ $(X_4 * R_{30}) > 8442$	Yes
3	$(X_1 * R_{30}) > 2.86$ $(X_3 * R_{30}) > 3.27$ $471.2 < (X_4 * R_{30}) \leq 1164.5$	Yes
4	$(X_1 * R_{30}) \leq 2.86$	No
5	$(X_1 * R_{30}) > 2.86$ $(X_3 * R_{30}) \leq 3.27$ $(X_4 * R_{30}) \leq 4774.2$	No
6	$(X_1 * R_{30}) > 2.86$ $(X_3 * R_{30}) > 3.27$ $(X_4 * R_{30}) \leq 471.2$	No
7	$(X_1 * R_{30}) > 2.86$ $(X_3 * R_{30}) > 3.27$ $(X_4 * R_{30}) \leq 8442$	No

General summaries of the tree was that a location on steep slope that experiences moderate to high burn (high X_1 predictor) had a higher probability to produce DFs, if it was also laden with thick overburden material (high X_4 predictor), with high erodibility index (high X_3 predictor). Finally, the DF had to be triggered by a high intensity, short duration storm (predictor R30). Validation of this model with the test set gave a sensitivity of 81% and a specificity of 78%, resulting in an overall accuracy of 78%. This means that the model can identify both DF stable and unstable locations within a burned area ~8 out of 10 times.

4.2 Methodology and Approach

4.2.1 Delineating Debris Flow Paths

Although the probability models discussed above only isolate DF origination points upstream, we postulated that by isolating the DF inundation paths, we can back-trace them to the origination points upstream to use as model validation data. To do this, we performed a change detection analysis to map the DF paths. Two sets of high-resolution multispectral data were obtained from Planet Labs Inc. These were from their Planet Scope product, which comes with four bands: blue, green, red, and near infrared (nIR), at a 3m spatial resolution. The pre- image acquisition date was December 29, 2017, and the post image acquisition date was January 11, 2018 (Planet Team, 2018). To run the change detection analyses, we determined the NDVI for these two time stamps using the following equation (Rouse et al., 1973):

$$NDVI = \frac{Band\ 4 - Band\ 3}{Band\ 4 + Band\ 3} \tag{3}$$

where, Band 3 = reflectance of the red band at a spectral resolution of 590 – 670 nm

Band 4 = reflectance of the nIR band at a spectral resolution of 780 – 860 nm

NDVI is a metric that is usually used to determine vegetation health; it runs from -1 to +1. Values close to +1 indicate healthy vegetation whereas values close to -1 give indication of poor to no vegetation. The rationale behind it is that leaves of healthy vegetation absorb red light for use in photosynthesis (low reflectance value) whereas its cell structure strongly reflects nIR (high reflectance value), thereby resulting in high NDVI values. For poor to no vegetation, the converse is true— that is, the red band has a higher reflectance and the nIR has lower reflectance, resulting in lower NDVIs. We therefore expected the DF path to have lower NDVI values. To make them even more discernible, we generated a resultant metric by subtracting the NDVI after the DF event (NDVIpost) from the NDVI before the event (NDVIpre). This metric shall be called *delta NDVI (dNDVI)*:

$$dNDVI = NDVI_{pre} - NDVI_{post} \quad (4)$$

Locations with DF scars had larger magnitudes of negative NDVI, thereby subtracting them from the pre- image resulted in relatively higher values than their surroundings, making the contrast even sharper (Figure 4.3). To validate these scars as being truly from the DFs we overlaid ground truth data obtained from Santa Barbara County (Santa Barbary County, Finalized Damage Inspections, 2018). These data consisted of point locations of infrastructure downstream to the burned area that were affected by the runoff.

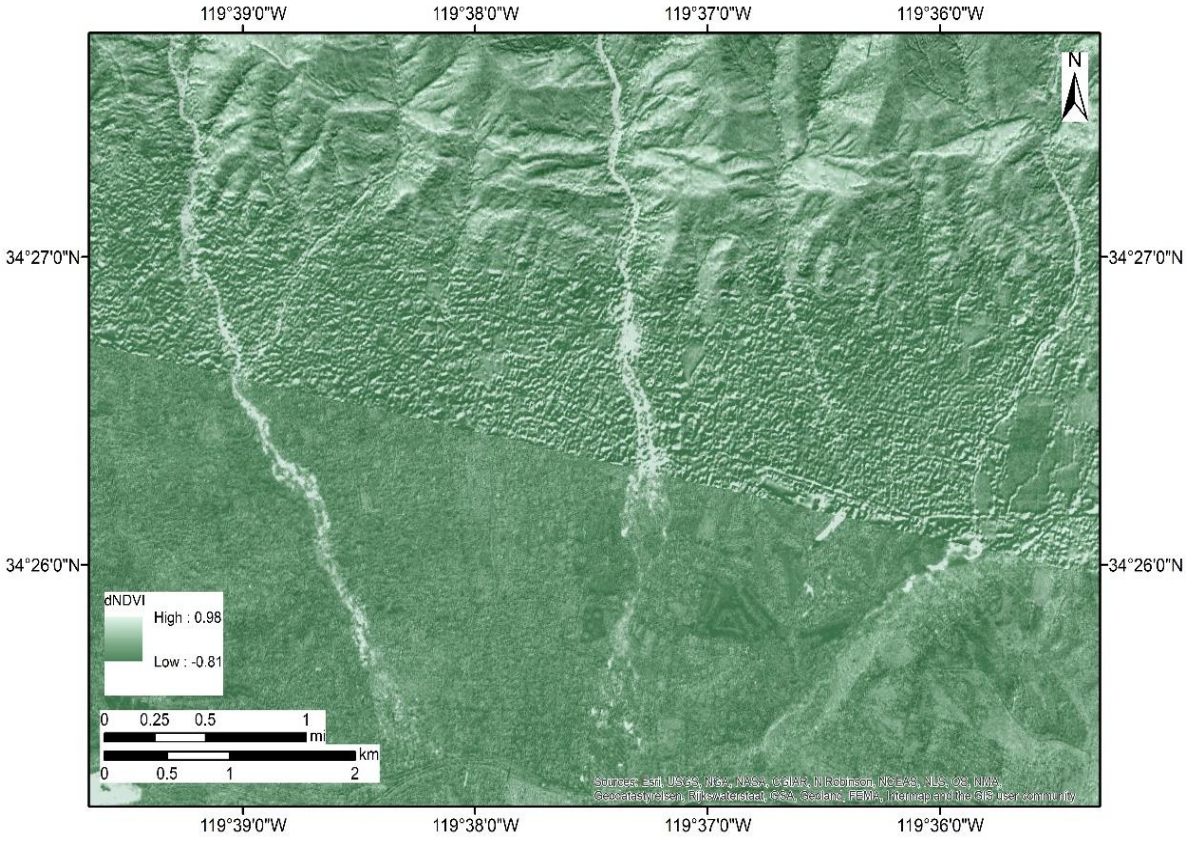

Figure 4.3 dNDVI analyses output of with Planet Lab data. Locations with lighter green to white colors show DF scars, with high dNDVI values.

4.2.2 Assessing Debris Flow Response

Analyses of the DF response of the Thomas fire basin using the Staley et al., (2017) logistic regression approach was ran by researchers from the USGS. These analyses were done for different rainfall scenarios and uploaded on the web for public access (https://landslides.usgs.gov/hazards/postfire_debrisflow/detail.php?objectid=178). The output for a 40mm/h rainfall was downloaded for use as the candidate logistic model output for use in later assessments.

To run similar prediction analyses with C5.0 algorithm, we first downloaded all needed data. A 10 m digital elevation model (DEM) was downloaded from the United States Department of Agriculture (USDA) Geospatial Gateway. Peak 30-minute rainfall intensity (PeakI30) input was obtained from the Santa Barbara County Public Works website. Burn severity input was obtained from the Burned Area Emergency Response (BAER) team, who field validated Burned Area Reflectance Classification (BARC) imagery derived from the dNBR data (Key and Benson, 2006). Finally, the soil properties, KF and soil thickness, were obtained from the Soil Survey Geographic (SSURGO) database (USDA, 1995).

The DEM was used to delineate the area within the Thomas Fire boundary into sub-basins using the hydrology suite offered by TauDEM toolbox add-on in ArcMap. This add-on was developed by researchers at Utah State University (Tarboton, 2005). The basin delineation analyses resulted in the creation of 6,611 sub-basins. For the PeakI30 input, we analyzed historical per minute rainfall data from the Santa Barbara County for 13 rainfall stations within, and around the burned area on the day of the event, January 09, 2018 (Table 4.3). An inverse distance weighted (IDW) interpolation was done to display the spatial distribution of the rainfall (Figure 4.4). These showed variable intensities across the stations with the higher intensities concentrated within the burned boundary. For representative rainfall scenario, we took an average of the five stations within the Thomas Fire boundary to obtain a PeakI30 input of 49.4 mm/h, which was used for further analyses.

Table 4.3 Historical rainfall data for January 09, 2018 obtained from 13 stations within the Thomas Fire boundary.

Station	Lat.	Long.	Accumulation (mm)		Peak Intensity (mm/h)	
			15 mins	30 mins	15 mins	30 mins
Carpinteria Fire	34.39697	-119.518	21.6	28.2	86.4	56.4
Don Victor	34.6375	-119.464	6.9	9.4	27.4	18.8
Doulton Tunnel	34.45778	-119.566	25.7	31.0	102.6	62.0
Edison Trail	34.4428	-119.508	19.6	26.7	78.2	53.3
Jameson Reservoir	34.49083	-119.507	24.9	31.8	99.6	63.5
KTYD	34.47111	-119.677	19.6	22.1	78.2	44.2
La Conchita	34.36683	-119.442	16.3	20.3	65.0	40.6
Matilija Canyon	34.54131	-119.372	12.2	23.4	48.8	46.7
Montecito	34.42752	-119.64	18.3	19.6	73.2	39.1
Old Man Mountain	34.50431	-119.44	11.2	15.2	44.7	30.5
Seacliff	34.34544	-119.42	6.4	9.1	25.4	18.3
Stanwood Fire	34.4439	-119.691	14.0	15.0	55.9	30.0
Summerland	34.4153	-119.581	18.3	21.6	73.2	43.2

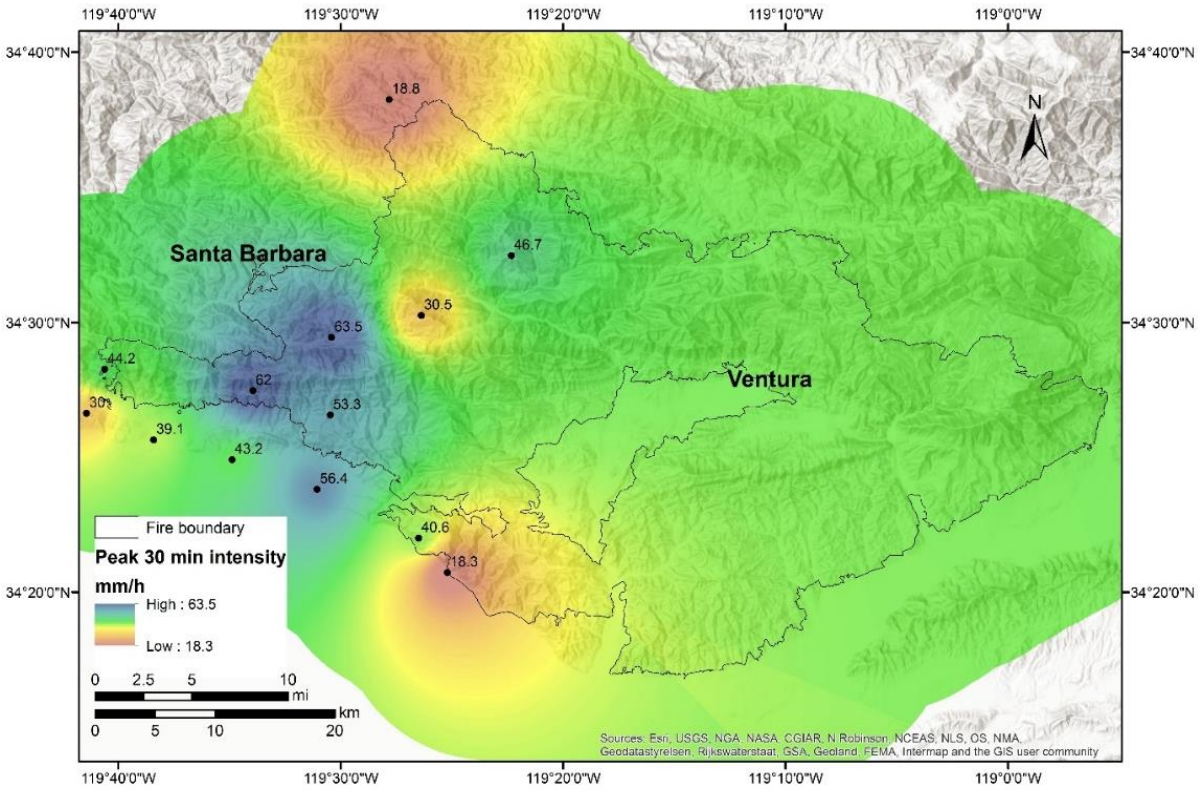

Figure 4.4 Inverse distance weighted interpolation showing spatial distribution of the January 09, 2018 rainfall data within the Thomas Fire Boundary in Santa Barbara County

We then compiled values for each of the other three needed predictors; if more than one value for any one predictor occurred in a basin, we calculated a single spatially weighted average for it. Finally, using the C5.0 tree model we calculated the probability of DF occurrence. The calculated values were then proportioned into probability classes and assigned to their respective sub-basins.

4.3 Results

The objective of this study was to apply the linear-based logistic regression model from Staley et al., (2017) and the nonlinear-based C5.0 tree model by Addison et al., (2018) to assess the strength of each in predicting the locations within the Thomas Fire boundary that produced DFs. To do this, DF scars were delineated by running a change detection protocol known as dNDVI, using high resolution data from Planet Labs. These scars were

further validated with data gathered by the Santa Barbara County on affected infrastructure. Results from the two models were overlain on the delineated DF scars and compared against each other to assess predictive strengths.

The dNDVI was instrumental in delineating DF scars. Overlaying the Santa Barbara point data of affected areas on it revealed great agreement between them and the delineated scars (Figure 4.5). We traced these scars upstream to points where they terminated and classified them as the DF origination points.

Figure 4.6 shows the model outputs for the logistic regression and the C5.0 tree algorithms, respectively. Table 4.4 also gives a breakdown of the prediction summary for different probability bins. The rainfall predictors that were used by these models were the peak 15-minute rainfall intensity (PeakI15) for the logistic regression, and the peak 30-minute rainfall intensity (PeakI30) for the C5.0 decision tree.

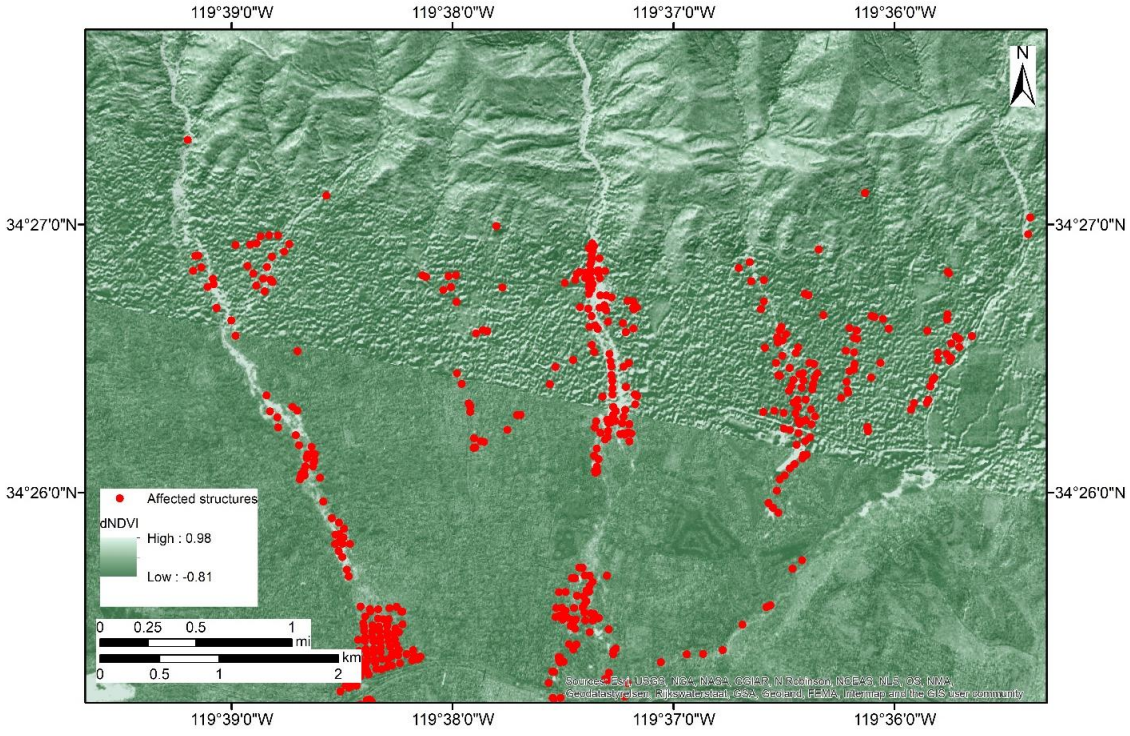

Figure 4.5 dNDVI image with locations of infrastructure that were affected in Santa Barbara County by the DF event overlaid.

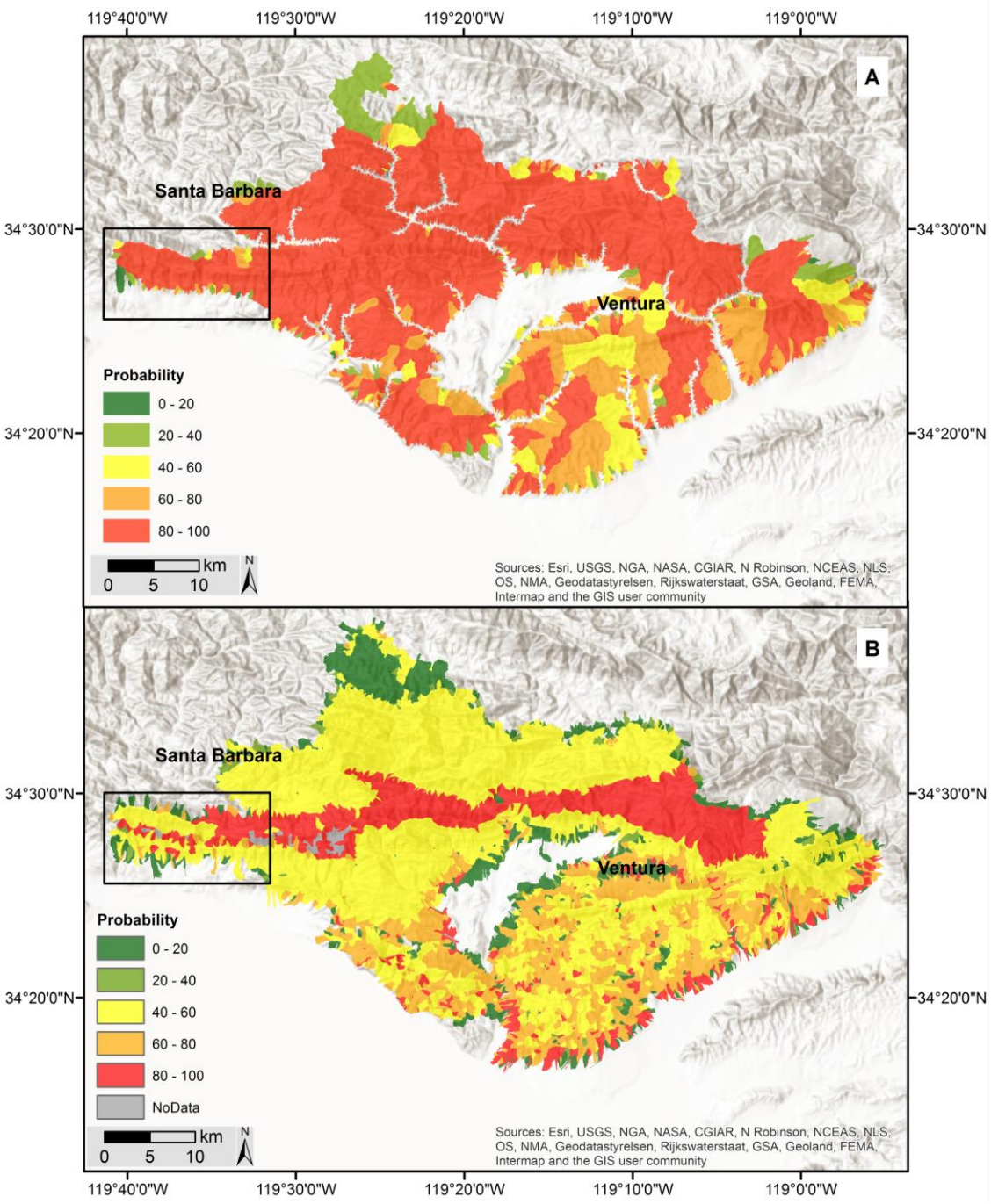

Figure 4.6 Model probability predictions of the logistic regression (A) and C5.0 tree (B) algorithms respectively. The map displays the probability of DF (in %) in 20% bins. The black bounding boxes in the Santa Barbara area shows the location that was zoomed in later for further analyses (see Figure 4.7).

Table 4.4 Summary of the area covered by the different probability bins of the two models

Probability (%)	Hazard	Area Covered (%)	
		Logistic Regression	C5 Tree
0-20	Low	0.3	10.5
20-40	Moderate	5.6	1.6
40-60		9.0	51.7
60-80	High	15.9	18.8
80-100		69.3	17.4

It is important to note that the average of the intensities of the actual rainfall that fell that day were 73.2 mm/h for the PeakI15 and 49.4 mm/h for the PeakI30. However, the logistic regression model showed in Figure 4.6A was made with a PeakI15 of 40 mm/h. The reason for using this instead of the actual 73.2 mm/h is that USGS researchers made these predictions before the actual DF event; they made them for different PeakI15 rainfall scenarios from lowest 12 mm/h to highest 40mm/h. The rainfall experienced on January 09, 2018 was a 50-year return event (Burns, 2018). Before the event, this seemed unlikely and hence, probably the reason why USGS researchers predicted to a highest rainfall scenario of PeakI15 of 40mm/h, which is a ~5-year return event, according to NOAA's precipitation frequency estimates (https://hdsc.nws.noaa.gov/hdsc/pfds/).

It was therefore this available maximum 40mm/h PeakI15 model output that was considered in our study. Even with the lower rainfall intensity, the logistic regression model predicted the majority of the Thomas Fire boundary, 84.9% to be in high hazard (probability of 60-100%). The C5.0 tree model, on the other hand, predicted 36.2% of the area to be high hazard, with its majority (53.3%) being a moderate hazard prediction

(probability of 20-60%). Overall, the logistic regression predicted ~44,800 ha (49%) more high hazard locations than the C5.0 tree.

Since we know that the DF did occur, in terms of emergency response, the logistic regression model shows better urgency than the C5.0 tree. However, with majority of the study area being classified as high hazard, it does bring a question to bear: is the model predicting more false positives for some of these areas, especially since it has a lower specificity (58%) than the C5 tree (78%)? From the dNDVI images, it does not look like all the basins upstream had DF origination points. Zooming into the Santa Barbara County, where we had DF scar validating data, we see that both models performed well in predicting high probabilities for the mountain ranges for the traced out debris flow scars from the dNDVI image (Figure 4.7). Going off from the delineated DF scars in the zoomed in image, the logistic regression model predicts a 100% of all terminating basins as high hazard, whereas the C5.0 tree model predicts 80% of the locations as high hazard and the remaining 20% as moderate hazard. We performed further analyses to isolate true positive predictions, as basins predicted as high hazards with visible DF scars, and their converse as false positives (basins predicted as high hazards with no visible DF scars). Logistic regression had ~72% of its area as false positives whereas the C5.0 tree had 29%. A perfect model is one with 100% true positives and 0% false positives; and a realistic robust model should be one that maximizes true positives and minimizes its false positives. Therefore from these analyses, the C5.0 tree performed better overall since it did not sacrifice its specificity for a high sensitivity as the logistic regression did.

All in all, we can conclude that both models performed well in isolating the high hazard basins, which is critical in emergency situations. The logistic regression showed better urgency but the C5.0 tree performed better overall with a high true positive and corresponding low false positive rate. The high false positive rate from the logistic regression waters its robustness down, as it holds the risk of desensitizing the public to when there is real danger. We therefore recommend that it be fine-tuned to improve its specificity to match its sensitivity.

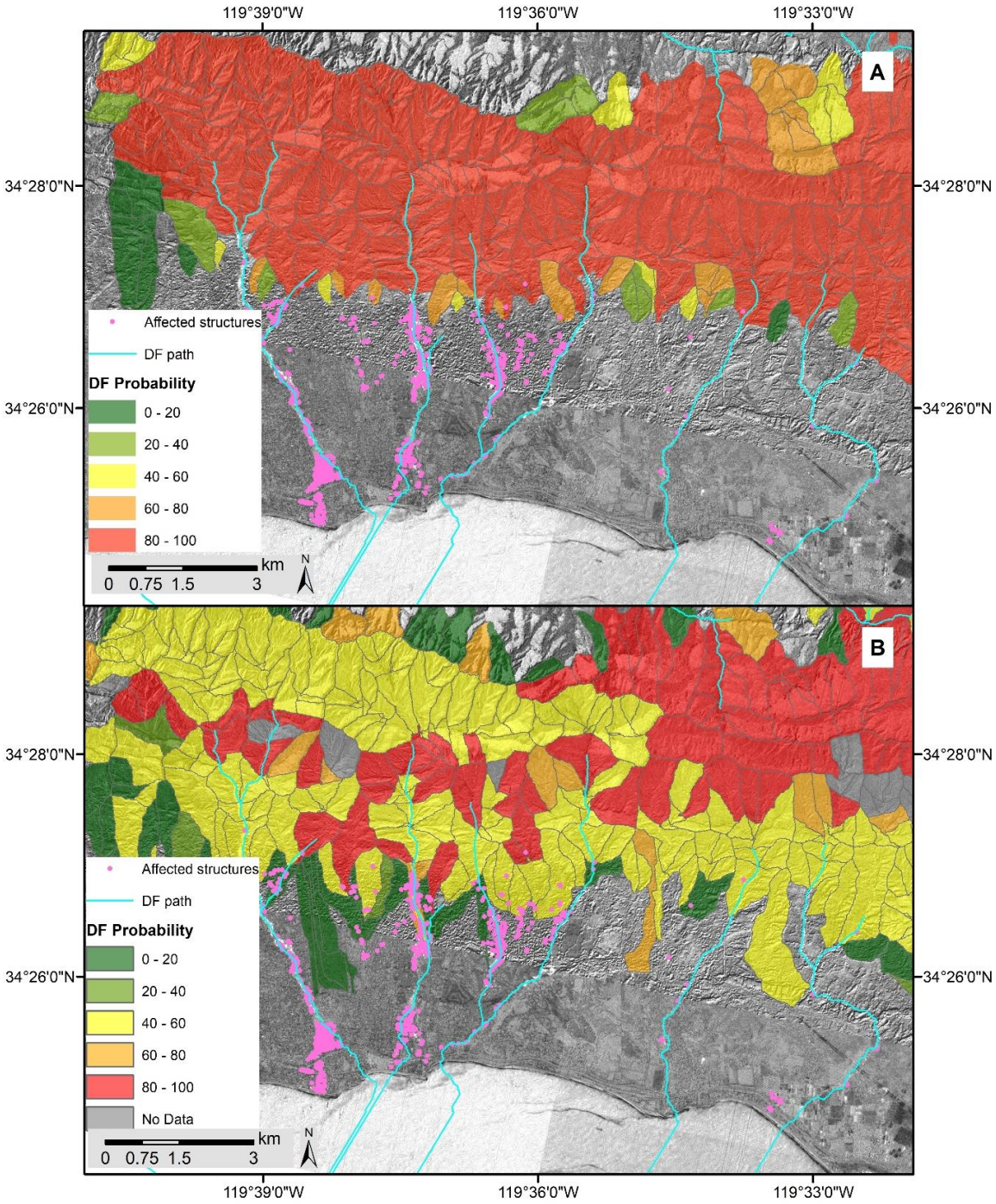

Figure 4.7 Model probability outputs in Santa Barbara County from the logistic regression algorithm (A) and the C5.0 tree algorithm (B), respectively. These outputs are overlaid on a greyscale dNDVI map with isolated DF paths traced out to their origination points upstream (cyan color). The pink show point data for locations downstream that were affected by the DF runoff.

4.4 Conclusion

This study has looked into applying a linear-based logistic regression model and a nonlinear-based C5.0 tree model to assess the strength of each in predicting the locations within the Thomas Fire boundary that produced DFs. The dNDVI metric was instrumental in delineating DF scars. Locations with higher dNDVI values showed good agreement with ground truthing data and were mapped as debris flow scars. The results from the statistical modeling approaches showed that both models performed well in isolating the high hazard basins, which is critical in emergency situations. The logistic regression model, however, predicted an overall ~44,800 ha (49%) more high hazard locations compared to the C5.0 tree, which with the knowledge that the debris flow event happened, shows better urgency. However, overall, the C5.0 tree performed better since it did not sacrifice its specificity for a high sensitivity. A closer look at the DF predictions by the logistic regression showed most of the basins predicted as high hazards to not have any discernible DF scars. ~72% of the area in the Santa Barbara County were classified as false positives, which is worrying since it holds the risk of desensitizing the public to when there is real danger. We, however, do not discount this model since it had a high sensitivity. We recommend fine-tuning to improve its specificity as well. Further recommendation include further studies to predict potential inundation DF paths after these statistical models have isolated their origination points upstream since their most associated danger is not necessarily where they start but rather further downstream where most communities and infrastructure are located.

Acknowledgments: We thank researchers at USGS for providing data together with model outputs for the analyses. We will also like to thank Santa Barbara County for providing ground truthing data; as well as the Planet Labs team for the high resolution data for debris flow delineation. Finally, we will like to thank Dr. McGuire from the University of Arizona for his assistance in obtaining rainfall data.

5 Overarching Conclusions

The main focus of this dissertation project has been to investigate the application of machine learning modeling as an emergency response tool for post-wildfire debris flow occurrences. The study was split into three sub-projects to meet the following objectives:

- Developing a radar-based burn severity estimate to be used as an alternative in emergency situations where there is cloud and/ or smoke coverage.
- Developing a nonlinear machine learning model to predict the likelihood of post-wildfire debris flow occurrences in the western USA.
- Testing robustness of developed debris flow model with a case study of recent event occurrence.

To meet the first objective, we developed a SAR based burn severity classification model as an alternative to the optical based approach, dNBR. During the modeling process, a significant discovery showed that evergreen vegetation seem to limit the severity of fires, as we observed locations with evergreen forests to generally experience low burn severity. The developed SAR metric produced higher validation metrics. It had an overall accuracy (OA) of ~60% and Kappa of 0.35 in comparison to an OA of ~35% and a kappa of 0.1 from the dNBR approach. This therefore proves the robustness of the SAR metric, however, we recommend further studies to investigate the influence of rainfall on this metric since weather effects are sometimes considered sources of noise for SAR data. Rainfall especially may result in differing water content of a target, affecting its relative permittivity and effectively altering its dielectric constant (alterations in backscatter values).

For the second project, we developed a nonlinear C5.0 tree model to predict the likelihood of post-wildfire debris flow occurrences in the western USA. This model was then compared to an existing logistic regression by USGS researchers. The C5.0 tree model gave an average sensitivity of 77% as compared to 66% from the logistic regression approach. Generally, the validation results from the C5.0 approach were higher and more stable. Analyzing the tree output confirmed the hypotheses that a steep

landscape that is laden with thick overburden soil, becomes highly susceptible to debris flow occurrences when it experiences moderate to high severity burns.

The final project was a case study to validate our C5.0 tree model from the previous project. We applied both the C5.0 tree and USGS' logistic regression models to a recent fire that happened in Southern California: the Thomas Fire. Using this burned location as a case study, we validated our C5.0 tree model as well as compare its predictive strength with that of the logistic regression. Both models were robust in isolating high hazard locations evidenced to have produced debris flows. The logistic regression model, however, predicted an overall ~44,800 ha (49%) more high hazard locations compared to the C5.0, showing better urgency. However, a closer look at the DF predictions showed that most of those extra basins had no discernible debris flow scars. It was therefore projected that these locations are likely false positives and caution was therefore recommended in using the logistic regression model to avoid desensitizing the public against debris flow hazards.

All these work to show the machine learning modeling as a superior tool in teasing out complex relationships. It also show utility of the approach for yet one more discipline for application.

6 References

Addison, P., Oommen T., and Sha Q., 2018. Assessment of Post-Wildfire Debris Flow Occurrence Using Classifier Tree. Geomatics, Natural Hazards and Risk.

Allen, J.L., and Sorbel, B., 2008. Assessing the differenced Normalized Burn Ratio's ability to map burn severity in the boreal forest and tundra ecosystems of Alaska's national parks. International Journal of Wildland Fire, 17(4), pp.463-475.

Allison, R.S., Johnston, J.M., Craig, G. and Jennings, S., 2016. Airborne optical and thermal remote sensing for wildfire detection and monitoring. Sensors, 16(8), p.1310.

Bailey, R.W., Craddock, G.W. and Croft, A.R., 1947. Watershed management for summer flood control in Utah (No. 639). US Dept. of Agriculture.

Bernhard, E.M., Twele, A. and Gähler, M., 2011. Rapid Mapping of Forest Fires in the European Mediterranean Region–a Change Detection Approach Using X-Band SAR-Data. Photogrammetrie-Fernerkundung-Geoinformation, 2011(4), pp.261-270.

Berman, M., 2017. California firefighter killed during response to historic Thomas Fire. Washington Post. ISSN 0190-8286. Retrieved December 14, 2017.

Birch, D.S., Morgan, P., Kolden, C.A., Abatzoglou, J.T., Dillon, G.K., Hudak, A.T. and Smith, A., 2015. Vegetation, topography and daily weather influenced burn severity in central Idaho and western Montana forests. Ecosphere, 6(1), pp.1-23.

Bond-Lamberty, B., Peckham, S.D., Ahl, D.E. and Gower, S.T., 2007. Fire as the dominant driver of central Canadian boreal forest carbon balance. Nature, 450(7166), pp.89-92.

Bovis, M.J. and Jakob, M., 1999. The role of debris supply conditions in predicting debris flow activity. Earth surface processes and landforms, 24(11), pp.1039-1054.

Brock, R. J., Cannon, S. H., Gartner, J., Santi, P. M., Higgins, J. D., and Bernard, D. R., 2007. An ordinary storm with an extraordinary response: Mapping the debris-flow response to the December 25, 2003 storm on the 2003 Old and Grand Prix fire areas in southern California. In 2007 GSA Denver Annual Meeting.

Cannon, S.H. and DeGraff, J., 2009. The increasing wildfire and post-fire debris-flow threat in western USA, and implications for consequences of climate change. In Landslides–disaster risk reduction (pp. 177-190). Springer Berlin Heidelberg

Cannon, S.H. and Gartner, J.E., 2005. Wildfire-related debris flow from a hazards perspective. In Debris-flow hazards and related phenomena (pp. 363-385). Springer Berlin Heidelberg.

Cannon, S.H., Gartner, J.E., Rupert, M.G., Michael, J.A., Rea, A.H. and Parrett, C., 2010. Predicting the probability and volume of post-wildfire debris flows in the intermountain western United States. Geological Society of America Bulletin, 122(1-2), pp.127-144.

Chuvieco, E., Riaño, D., Danson, F.M. and Martin, P., 2006. Use of a radiative transfer model to simulate the post-fire spectral response to burn severity. Journal of Geophysical Research: Biogeosciences, 111(G4).

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), pp.37-46.

CNN, 2018. California mudslides: Death toll rises to 20, 4 still missing. Retrieved January 15, 2018.

DeBano, L.F., 2000. The role of fire and soil heating on water repellency in wildland environments: a review. Journal of Hydrology, 231, pp.195-206.

DeGraff, J.V., Cannon, S.H. and Gartner, J.E., 2015. The timing of susceptibility to post-fire debris flows in the Western United States. Environmental & Engineering Geoscience, 21(4), pp.277-292.

Dillon, G.K., Holden, Z.A., Morgan, P. and Keane, B., 2009. A Fire Severity Mapping System (FSMS) for real-time management applications and long term planning: Developing a map of the landscape potential for severe fire in the western United States. Fourth Annual International Congress on Fire Ecology and Management: Fire as a Global Process.

Dillon, G.K., Morgan, P. and Holden, Z.A., 2011a. Mapping the potential for severe fire in the Western United States. Fire Management Today, 71, pp.1-28.

Dillon, G.K., Holden, Z.A., Morgan, P., Crimmins, M.A., Heyerdahl, E.K. and Luce, C.H., 2011b. Both topography and climate affected forest and woodland burn severity in two regions of the western US, 1984 to 2006. Ecosphere, 2(12), pp.1-33.

Ding, A., 2018. Charting the Financial Damage of the Thomas Fire. The Bottom Line. Retrieved May 26, 2018.

Dixon, R.K. and Krankina, O.N., 1993. Forest fires in Russia: carbon dioxide emissions to the atmosphere. Canadian Journal of Forest Research, 23(4), pp.700-705.

Doerr, S.H., Shakesby, R.A. and Walsh, R., 2000. Soil water repellency: its causes, characteristics and hydro-geomorphological significance. Earth-Science Reviews, 51(1-4), pp.33-65.

Dolan, J., 2018. “Search teams find 21st victim of Montecito mudslide”. Los Angeles Times. Retrieved January 21, 2018.

Donovan, I.P. and Santi, P.M., 2017. A probabilistic approach to post-wildfire debris-flow volume modeling. Landslides, 14(4), pp.1345-1360.

Eaton, E.C., 1935, Flood and erosion control problems and their solution: American Society of Civil Engineers Transactions, v. 101, p. 1302–1362.

Flannigan, M.D., Krawchuk, M.A., de Groot, W.J., Wotton, B.M. and Gowman, L.M., 2009. Implications of changing climate for global wildland fire. International journal of wildland fire, 18(5), pp.483-507.

Fox News, 2018. California mudslides: Where and why they happen. Retrieved February 09, 2018.

Gartner, J.E., Cannon, S.H., Bigio, E.R., Davis, N.K., Parrett, C., Pierce, K.L., Rupert, M.G., Thurston, B.L., Trebish, M.J., Garcia, S.P. and Rea, A.H., 2005. Compilation of data relating to the erosive response of 608 recently-burned basins in the western United States (No. 2005-1218).

Gartner, J.E., Cannon, S.H., Santi, P.M. and Dewolfe, V.G., 2008. Empirical models to predict the volumes of debris flows generated by recently burned basins in the western US. Geomorphology, 96(3), pp.339-354.

Gimeno, M., San-Miguel-Ayanz, J. and Schmuck, G., 2004. Identification of burnt areas in Mediterranean forest environments from ERS-2 SAR time series. International Journal of Remote Sensing, 25(22), pp.4873-4888.

Gowda, P., Oommen, T., Misra, D., Schwartz, R., Howell, T., & Wagle, P. (2015). Retrieving leaf area index from remotely sensed data using advanced statistical approaches. GIScience and Remote Sensing, 4, 156.

Hammond, C.J., Prellwitz, R.W. and Miller, S.M., 1992, February. Landslide hazard assessment using Monte Carlo simulation. In Proceedings of 6th international symposium on landslides, Christchurch, New Zealand, Balkema (Vol. 2, pp. 251-294).

Hersko, T., 2018. Ventura County agriculture suffers over $170 million in damages from Thomas Fire. VC Star. Retrieved January 25, 2018.

Hoy, E.E., French, N.H., Turetsky, M.R., Trigg, S.N. and Kasischke, E.S., 2008. Evaluating the potential of Landsat TM/ETM+ imagery for assessing fire severity in Alaskan black spruce forests. International Journal of Wildland Fire, 17(4), pp.500-514.

Hungr, O., Morgan, G.C. and Kellerhals, R., 1984. Quantitative analysis of debris torrent hazards for design of remedial measures. Canadian Geotechnical Journal, 21(4), pp.663-677.

Johnson, P.A., McCuen, R.H. and Hromadka, T.V., 1991. Magnitude and frequency of debris flows. Journal of hydrology, 123(1-2), pp.69-82.

Justice, C.O., Giglio, L., Korontzi, S., Owens, J., Morisette, J.T., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F. and Kaufman, Y., 2002. The MODIS fire products. Remote Sensing of Environment, 83(1), pp.244-262.

Kasischke, E.S., Melack, J.M. and Dobson, M.C., 1997. The use of imaging radars for ecological applications—a review. Remote Sensing of Environment, 59(2), pp.141-156.

Kern, A., Addison, P., Oommen T., Salazar S., and Coffman R., 2017. Machine learning based predictive modeling of debris flow probability following wildfire in the intermountain Western United States. Mathematical Geosciences.

Key, C.H. and Benson, N.C., 2005. Landscape assessment: remote sensing of severity, the normalized burn ratio and ground measure of severity, the composite burn index. FIREMON: Fire effects monitoring and inventory system Ogden, Utah: USDA Forest Service, Rocky Mountain Res. Station.

Key, C.H. and Benson, N.C., 2006. Landscape assessment (LA). FIREMON: Fire effects monitoring and inventory system. Gen. Tech. Rep. RMRS-GTR-164-CD, Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Kinner, D.A. and Moody, J.A., 2008. Infiltration and runoff measurements on steep burned hillslopes using a rainfall simulator with variable rain intensities. U. S. Geological Survey.

Koutsias, N., Karteris, M. and Chuvico, E., 2000. The use of intensity-hue-saturation transformation of Landsat-5 Thematic Mapper data for burned land mapping. Photogrammetric Engineering and Remote Sensing, 66(7), pp.829-840.

Kuhn, M., and Johnson, K., 2013. Applied Predictive Modeling. New York, NY: Springer, 2013.

Larsen, I.J., MacDonald, L.H., Brown, E., Rough, D., Welsh, M.J., Pietraszek, J.H., Libohova, Z., de Dios Benavides-Solorio, J. and Schaffrath, K., 2009. Causes of post-fire runoff & erosion: water repellency, cover, or soil sealing? Soil Science Society of America Journal, 73(4), pp.1393-1407.

Le Toan, T., Beaudoin, A., Riom, J. and Guyon, D., 1992. Relating forest biomass to SAR data. Geoscience and Remote Sensing, IEEE Transactions on, 30(2), pp.403-411.

Magnoli, Giana (27 February 2018). "County Estimates $46 Million Cost for Thomas Fire, Montecito Debris Flow Response, Repairs". Noozhawk. Retrieved 26 May 2018.

Martin, D.A. and Moody, J.A., 2001. Comparison of soil infiltration rates in burned and unburned mountainous watersheds. Hydrological Processes, 15(15), pp.2893-2903.

Miller, J.D. and Thode, A.E., 2007. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). Remote Sensing of Environment, 109(1), pp.66-80.

Miller, J.D., Safford, H.D., Crimmins, M. and Thode, A.E., 2009. Quantitative evidence for increasing forest fire severity in the Sierra Nevada and southern Cascade Mountains, California and Nevada, USA. Ecosystems, 12(1), pp.16-32.

Mitchell, T.M., 1997. Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 45(37), pp.870-877.

Mitri, G.H. and Gitas, I.Z., 2004. A semi-automated object-oriented model for burned area mapping in the Mediterranean region using Landsat-TM imagery. International Journal of Wildland Fire, 13(3), pp.367-376.

Moody, J.A. and Ebel, B.A., 2012. Hyper-dry conditions provide new insights into the cause of extreme floods after wildfire. Catena, 93, pp.58-63.

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(10), pp.1049-1070.

Murphy, K.A., Reynolds, J.H. and Koltun, J.M., 2008. Evaluating the ability of the differenced Normalized Burn Ratio (dNBR) to predict ecologically significant burn severity in Alaskan boreal forests. International Journal of Wildland Fire, 17(4), pp.490-499.

National Interagency Fire Center (NIFC), 2018. 2018 National Large Incident Year-to-Date Report. Retrieved August 30, 2018.

Neary, D.G., Ryan, K.C. and DeBano, L.F., 2005. Wildland fire in ecosystems: effects of fire on soils and water. Gen. Tech. Rep. RMRS-GTR-42-vol, 4, p.250.

Orozco, L., 2017. Authorities Release Details of Fatality in Thomas Fire; 70 Year Old Woman Dies during Evacuations. KCLU-FM. Retrieved August 01, 2018.

Parks, S.A., Dillon, G.K. and Miller, C., 2014. A new metric for quantifying burn severity: the Relativized Burn Ratio. Remote Sensing, 6(3), pp.1827-1844.

Phillips, K., Kaplan S., and McMillan K., 2018. A record 'you don't want to see': Mendocino Complex fire has become California's largest ever. The Washington Post. Retrieved October 22, 2018.

Planet Team, 2018. Planet application program interface: In space for life on Earth, San Francisco, CA. Retrieved from https://api.planet.com (Last accessed: August 01, 2018).

Polychronaki, A., Gitas, I.Z., Veraverbeke, S. and Debien, A., 2013. Evaluation of ALOS PALSAR imagery for burned area mapping in Greece using object-based classification. Remote Sensing, 5(11), pp.5680-5701.

Quinlan, J.R., 1993. Constructing decision tree. C4.5, pp.17-26.

Rignot, E.J., Williams, C.L., Way, J. and Viereck, L.A., 1994. Mapping of forest types in Alaskan boreal forests using SAR imagery. IEEE Transactions on Geoscience and Remote Sensing, 32(5), pp.1051-1059.

Robert D. Niehaus, Inc. (RDN), 2018. The Economic Impacts of the Montecito Mudslides: A Preliminary Assessment. Retrieved May 26, 2018.

Rouse Jr, J.W., Haas, R.H., Schell, J.A. and Deering, D.W., 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation.

Roy, D.P., Lewis, P.E. and Justice, C.O., 2002. Burned area mapping using multi-temporal moderate spatial resolution data—A bi-directional reflectance model-based expectation approach. Remote Sensing of Environment, 83(1), pp.263-286.

Rupert, M.G., Cannon, S.H., Gartner, J.E., Michael, J.A. and Helsel, D.R., 2008. Using logistic regression to predict the probability of debris flows in areas burned by wildfires, southern California, 2003-2006 (Open-File Report No. 2008-1370). US Geological Survey.

Santa Barbara County, 2018. Finalized damage inspections. Retrieved from https://sbc-gis.maps.arcgis.com/apps/webappviewer/index.html?id=ee848a57d8b2416eb2802da300df5b6e. Retrieved on June 01, 2018.

Santa Cruz, N., 2018. Trump approves disaster funds for Thomas fire victims. The Los Angeles Times. Retrieved January 25, 2018.

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. Weather and forecasting, 5(4), pp.570-575.

Schroeder, W., Csiszar, I. and Morisette, J., 2008. Quantifying the impact of cloud obscuration on remote sensing of active fires in the Brazilian Amazon. Remote Sensing of Environment, 112(2), pp.456-470.

Shakesby, R.A. and Doerr, S.H., 2006. Wildfire as a hydrological and geomorphological agent. Earth-Science Reviews, 74(3), pp.269-307.

Shannon, C.E., 2001. A mathematical theory of communication. ACM SIGMOBILE mobile computing and communications review, 5(1), pp.3-55.

Sikkink, P.G., Dillon, G.K., Keane, R.E., Morgan, P., Karau, E.C., Holden, Z.A. and Silverstein, R.P., 2013. Composite Burn Index (CBI) data and field photos collected for the FIRESEV project, western United States. Fort Collins, CO.

Staley, D.M., Kean, J.W., Cannon, S.H., Schmidt, K.M. and Laber, J.L., 2013. Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California. Landslides, 10(5), pp.547-562.

Staley, D.M., Negri, J.A., Kean, J.W., Laber, J.L., Tillery, A.C. and Youberg, A.M., 2016. Updated logistic regression equations for the calculation of post-fire debris-flow likelihood in the western United States (No. 2016-1106). US Geological Survey.

Staley, D.M., Negri, J.A., Kean, J.W., Laber, J.L., Tillery, A.C. and Youberg, A.M., 2017. Prediction of spatially explicit rainfall intensity–duration thresholds for post-fire debris-flow generation in the western United States. Geomorphology, 278, pp.149-162.

Stroppiana, D., Azar, R., Calò, F., Pepe, A., Imperatore, P., Boschetti, M., Silva, J., Brivio, P.A. and Lanari, R., 2015. Integration of optical and SAR data for burned area mapping in Mediterranean Regions. Remote Sensing, 7(2), pp.1320-1345.

Swanston, D.N., 1991, Natural processes, in, Meehan, W.L., ed., Infl uences of Forest and Rangeland Management on Salmonoid Fishes and their Habitat: American Fisheries Society Special Publication 19, p. 139–179.

Tanase,M. A., de la Riva, J., Santoro, M., Le Toan, T., & Perez-Cabello, F., 2010a. Sensitivity of X-, C- and L-band SAR backscatter to fire severity in Mediterranean pine forests. IEEE Transactions on Geoscience and Remote Sensing, 48, 3663–3675.

Tanase, M.A., Kennedy, R. and Aponte, C., 2015a. Radar Burn Ratio for fire severity estimation at canopy level: An example for temperate forests. Remote Sensing of Environment, 170, pp.14-31.

Tanase, M.A., Kennedy, R. and Aponte, C., 2015b. Fire severity estimation from space: a comparison of active and passive sensors and their synergy for different forest types. International Journal of Wildland Fire, 24(8), p.1062.

Tanase, M.A., Santoro, M., Wegmüller, U., de la Riva, J. and Pérez-Cabello, F., 2010b. Properties of X-, C-and L-band repeat-pass interferometric SAR coherence in Mediterranean pine forests affected by fires. Remote Sensing of Environment, 114(10), pp.2182-2194.

Tarboton, D.G., 2005. Terrain analysis using digital elevation models (TauDEM). Utah State University, Logan.

United States Department of Agriculture (USDA). Natural Resources Conservation Service, 1995. Soil Survey Geographic (SSURGO) Data Base: Data Use Information. National Cartography and GIS Center.

van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G.J., Kasibhatla, P.S. and Arellano Jr, A.F., 2006. Interannual variability in global biomass burning emissions from 1997 to 2004. Atmospheric Chemistry and Physics, 6(11), pp.3423-3441.

Wegmuller, U., and Werner, C., 1997. Retrieval of vegetation parameters with SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing, 35(1), pp.18-24.

Wells, W.G., 1987. The effects of fire on the generation of debris flows in southern California. Reviews in Engineering Geology, 7, pp.105-114.

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(5789), pp.940-943.

Witt, S., 1999. Burned Area Emergency Response, BAER. Available from: https://www.fs.fed.us/biology/watershed/burnareas/background.html, [accessed: January 2017]

Wooster, M.J., Roberts, G., Smith, A.M., Johnston, J., Freeborn, P., Amici, S. and Hudak, A.T., 2013. Thermal remote sensing of active vegetation fires and biomass burning events. In Thermal Infrared Remote Sensing (pp. 347-390). Springer, Dordrecht.

Wondzell, S.M. and King, J.G., 2003. Post-fire erosional processes in the Pacific Northwest and Rocky Mountain regions. Forest Ecology and Management, 178(1), pp.75-87.