

UTILIZING LANDSAT IMAGERY TO DETERMINE LOCATIONS OF TREE REGENERATION AFTER HIGH SEVERITY WILDFIRES IN NORTHERN ARIZONA

By Kimberly Istok

A Practicum Project Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in Applied Geospatial Science

Northern Arizona University

May 2017

To be approved by:

Mark Manone, M.A.

Pamela E. Foti, Ph.D.

Jose Iniguez, Ph.D.

Abstract:

In this research project ground collected tree point data from 2014 to 2016 was tested for correlation with normalized difference vegetation index, NDVI. Based on the NDVI values associated with the ground collected data an NDVI range for regenerated trees will be used to estimate the amount of acres of specific tree regeneration in high burn severity areas. By being able to estimate the amount of trees returning, management organizations can plan for forest restoration techniques as well as prevention techniques for high severity wildfires. The following methodology was tested on two of Arizona's most severe wildfires. Landsat 7 ETM+ and Landsat 8 OLI were used to identify NDVI. In order to determine the NDVI value of the trees, a Landsat derived NDVI value taken directly after the fire (2002 -2005) was subtracted from NDVI value derived from NDVI value from a 2016 Landsat image. These delta NDVI pixel values were plotted against canopy cover area and a delta NDVI range was determine for tree regeneration. All pixel values that fell in the range were classified as locations of tree regeneration. This research found a low correlation between the number of trees and the delta NDVI value of the pixel. There was also a low to moderate correlation between the estimated area the trees take up and the delta NDVI. Low correlation was found to be significant however it does not solely determine the presence of trees. This low level of correlation could be contributed to the presence of grasses that also reflect during NDVI analysis causing noise. It could also be contributed to the large pixel size of the Landsat imagery in comparison to the small size of the seedlings. More research is required for NDVI ranges of specific tree species.

Acknowledgements:

The author would like to thank Pamela Foti, Mark Manone and Jose Iniguez for serving on the review committee. Their feedback and support was instrumental in the development and process of this project. The author would also like to the team at the Rocky Mountain Research Station for permitting the use of the collected stem map data for this practicum project. Thank you Caroline Seig, and Suzi Owen. Also a big thank you to all my family and friends for their support during this two years masters process

Table of Content:

List of Maps:	4
List of Figures:	4
Introduction:	5
Project Objectives:	8
Study Sites:	10
<i>Pumpkin Fire: 2000</i>	<i>11</i>
<i>Rodeo-Chediski Fire: 2002</i>	<i>13</i>
Methodology:	15
<i>Literature Review:</i>	<i>15</i>
<i>Data Utilized:</i>	<i>18</i>
Landsat imagery:	18
Tree point data:	23
Monitoring Trends in Burn Severity: (MTBS.gov)	25
<i>Methodology Performed:</i>	<i>28</i>
Preprocessing of Landsat Imagery:	28
Clean up of field collected ground data:	29
Creation of a Grid Layer for Reference:	30
Finding Delta NDVI:	33
Validating Regeneration:	37
Pulling Delta NDVI Ranges:	41
Selecting within High Severity Area:	42
Statistical test for Significance:	43
Results:	45
<i>Pumpkin Fire:</i>	<i>45</i>
<i>Rodeo-Chediski Fire:</i>	<i>47</i>
<i>Pumpkin Fire Conifer Test:</i>	<i>49</i>
Discussion:	51
<i>Future Research:</i>	<i>54</i>
Conclusion:	55

Citations:	57
Appendix 1: Time Logged	59
Appendix 2: Processing Notes	60

List of Maps:

Map 1: Fire Locations	11
Map 2: Pumpkin Fire Plot Location	13
Map 3: Rodeo-Chediski Plot Locations	15
Map 4: Pumpkin Fire Burn Severity	26
Map 5: Rodeo-Chediski Fire Burn Severity	27
Map 6: Pixel layer digitization example	32
Map 7: Pumpkin Fire NDVI Change Detection	35
Map 8: Rodeo-Chediski Fire NDVI Change Detection	36
Map 9: Pumpkin Fire Conifer Test Change Detection	37
Map 10: Example of Regeneration Selection	43
Map 11: Pumpkin Fire Areas of Regeneration	46
Map 12: Rodeo-Chediski Fire Areas of Regeneration	48
Map 13: Pumpkin Fire Conifer Test areas of Regeneration	50

List of Figures:

Figure 1: Landsat 7 ETM+ Band and Wavelength List	20
Figure 2: Landsat 8 OLI TIRS Band and Wavelength List	21
Figure 3: Pumpkin Fire Scatterplot	39
Figure 4: Rodeo-Chediski Fire Scatterplot	40

Figure 5: Pumpkin Fire Conifer Test Scatterplot	41
Figure 6: Pumpkin Fire Scatterplot with Regression line	47
Figure 7: Rodeo Chediski Fire Scatterplot with regression line	49
Figure 8: Pumpkin Fire Conifer Test Scatterplot with Regression line	51

Introduction:

Severe wildfires are becoming a growing concern in western forests of the United States. Over the past three decades wildfires have become much larger and more severe than fires in the past. For example the Wallow Fire May 2011 which burned 538,049 acres (Kramer 2012) and the Rodeo-Chediski in June 2002 which burned 468,000 acres (Strom 2005). These two fires are some of the largest wildfires recorded in the state of Arizona. This increase in severity can be correlated with an increase of fuels resulting from a century of selective logging of larger trees, livestock grazing and fire suppression (Strom & Fule 2007). Research on forest management techniques has been steadily increasing over the years because of the rate at which high severity wildfires are killing trees, is causing major concern for forest managers (Strom & Fule 2007). The ponderosa pines are of significant concern because of a shift of environment for seed regeneration, productions and seedling establishment due to the change in the climate within the southwest (Mast et. al. 1999, Strom & Fule 2007). Different species are being seen regenerating within the high burn severity areas that were not present before the burn. These new species are causing there to be a shift in the way in which the forests can be managed. Currently are two sides to the forest management argument: 1) the forest should be left alone to allow to natural

regeneration (Kramer 2012), or 2) the high severity areas should be replanted in an effort to accelerate the return of the pre-fire forest. (Kramer 2012).

Land managers and the research community have indicated that fire suppression a reactive technique and practice by the fire crews should be transitioned into more proactive techniques such as prescribed burns and thinning of dense areas (Strom & Fule 2007). Not only should the goal be to replant the forest but have an overall goal of returning the forest stands to pre-settlement forest conditions after wildfires (Strom & Fule 2007). Pre-settlement forest conditions refers to before human use and interaction within the forest. In pre-settlement times ponderosa pine forests in Arizona were not as dense as currently seen and created various wildlife habitats (Mast et al. 1999). The forest management technique of thinning an area consists of selectively cutting down seedlings allowing for seedlings and trees to have more resources to grow (Mast et al. 1999). There are locations that are seen within the Rodeo-Chediski Fire from 2002 that were thinned before the fire and were found to have a significantly lower burn ratio than areas that were not (Strom & Fule 2007, Kramer 2012). Currently, there is significant testing of forest management methods. All the techniques being tested are working to prevent massive catastrophic fires. The use of satellite imagery is one tool that has been used for forest management and determination of fire boundaries and severities.

Monitoring trends in burn severity (MTBS.gov) is a website that provides one location in which wildfire researchers can go for all fire mapping and analysis needs. This website is managed by the United States Forest Service and the US Geological Survey. For each wildfire after 1984 and later spatially located information can be downloaded. The main use is fire boundaries and burn severity of each fire. This is an excellent resource for any forest

management researcher and it is a free service. However MTBS.gov is not the only location for satellite imagery. It is available to everyone via open source websites that are managed by US Geological Survey who collects, preprocess and maintains excellent metadata. Unlike MTBS.gov the following websites are of raw satellite data. They provide a catalog of historical data, going back as far as 1972 depending on the image and sensor that is being used (Flood 2014, Chen et al. 2011). There are several different types of satellite data available through GloVis (glovis.usgs.gov) or Earth Explorer (EarthExplorer.usgs.gov). Other locations for download can be found on the internet however these are the most user friendly and have updates regularly for ease of use.

With the growing concern based on the increasing severity of wildfires, different techniques have been tested with satellite imagery. With satellite imagery being easily available, there have been several studies utilizing all different types and resolution size of imagery. These studies cover all types of forest topics and disciplines. It is not a new technique to utilize satellite imagery for measuring wildfire regeneration, fire severity or fire boundaries (Cocke, et al. 2005). But it is becoming a crucial piece to forest management studies and understanding the after effects of wildfires on the ecosystem of the area (Veraverbeke et al. 2010).

By using satellite imagery from MTBS.gov management bodies are able to identify the areas quickly that may need management intervention. It allows for easier calculation and budgeting for the departments and for what the forest needs. Forest management techniques such as thinning can be costly and require a lot of personnel to power the work as well as make it an effective and safe process. Prescribed burns need even more people and planning, because not only do forest crews go in and collect all of the down debris, fire crews are needed to light

and contain the fire. With prescribed burns there is also current climate and weather taken into account before the areas are burned, there lot of coordination that needs to be done to insure a smooth prescribed burn. Satellite imagery is being used as a tool in the forestry management and research community for regeneration over time, forest changes, severity levels within the fires, estimation of regeneration amount after wildfire, and it can even be used to determine what the present species are. By being able to assess the recovery of the vegetation after a wildfire, researchers are able to assess the ecosystem as a whole. Understanding the fire severity allows for a key hole into what the environment will be like after (R. Diaz-Delgado et al. 2013). High severity fires effect the physical, chemical and biological aspects that are seen after the fire (R. Diaz-Delgado et al. 2013). Even though wildfires can have so many negative effects on the human environment, wildfires are important to forest ecosystems, allowing for fresh beginnings (Yi et al. 2013).

This project was created with the intensions of producing and testing a process that can be recreated by forest management teams that will allow for tree regenerated areas after high severity wildfires to be identified. By providing areas that have the presents or absence of trees, forest management can be streamlined to address these areas. In order to make this process accessible to everyone free satellite imagery was used as well as tools and techniques that have several resources available for reference. The author has included the processing notes in appendix 2 to help with processing questions and techniques.

Project Objectives:

In an effort to understand forest resiliency after these large high severity wildfires the Rocky Mountain Research Station (RMRS) has been conducting several forest regeneration

research projects. The RMRS which is part of the United States Forest Service has worked closely with Northern Arizona University and top researchers all over the country to try and better understand forest recovery after large wildfires. To better understand the patterns seen in the tree regeneration 18 - 200 meter by 200 meter plots of regenerated areas were used as a sample of regeneration after two high severity wildfires in northern Arizona. This data was collected during 2014 – 2016 summers. Within each plot all seedlings and trees were mapped by field crews with a Trimble unit and range finder. There were two different types of plots recorded, high severity and moderate severity. There were 6 moderate severity plots collected however these plots were not used for this analysis, because there were surviving trees within the plots and there is a potential that the seedlings will not be visible by the satellite imagery because of the large existing trees canopy. This analysis focused on the remaining 12 high severity plots. The plots were split between two northern Arizona fires. These high severity plots did not contain any surviving trees after the wildfire.

The two high severity fires in northern Arizona that are included in this research are the Rodeo-Chediski and the Pumpkin Fire (Map 1). The larger of the two the Rodeo-Chediski wildfire began on June 18, 2002 as two separate fires. But because of the landscape of the White Mountain Apache Reservation and the Sitgreaves National Forest, and the weather conditions this fire became one of the largest fires in Arizona's history burning close to 468,000 acres (Strom 2005). This fire had a mixed burn severity but 169,042 acres classified as a high burn severity (MBTS.com). The second fire that was included in this research is the Pumpkin Fire, which began on May 24, 2000, burned a total of 15,800 acres with 3,536 acres classified as a

high burn severity. This burned area is located 1 mile southwest of the Kendrick Mountain Wilderness Area near Flagstaff, Arizona (Farnsworth 2001).

The first objective of this project is to test if field collected tree ground data can be utilized as training data for regeneration identification with Landsat satellite imagery. The ground data will allow for correlation analysis between tree count and delta NDVI, allowing for a delta NDVI value range to be established as regeneration. The second objective will be to estimate the overall acreage of regeneration only within the high burn severity areas of each wildfire. This estimation will utilize the range established to select all of the areas within the high burn severity boundary. The third objective is to determine if species specific delta NDVI range can be established for regenerating conifer trees. Conifer trees includes both ponderosa pines and pinon-junipers. Because conifers are green year round, the NDVI value should only contain them because the other greenery is dormant or dead during the winter. Being able to estimate acreage of this specific type of tree would be excellent help in creating management plans.

Study Sites:

The two study areas are within northern portion of Arizona. Northern Arizona as well as the entire southwestern region of the United States has a high frequency of wildfires (Dickson et. al 2005). Where 60 to 70 % of the wildfires are ignited by lightning (Dickson et. al 2005). This high percentage is contributed to the dry climate and the high intensity thunderstorms that are seen during summer monsoons in this region (Dickson et. al 2005). The summer monsoon season begins roughly around July 1st and continue into the end of August. Both study sites are located in Arizona's largest contiguous ponderosa pine forest. These ponderosa pine stands are forests

that are mainly dominated by ponderosa pine (*pinus ponderosa*) trees with little other species present (Dickson et. al 2005).

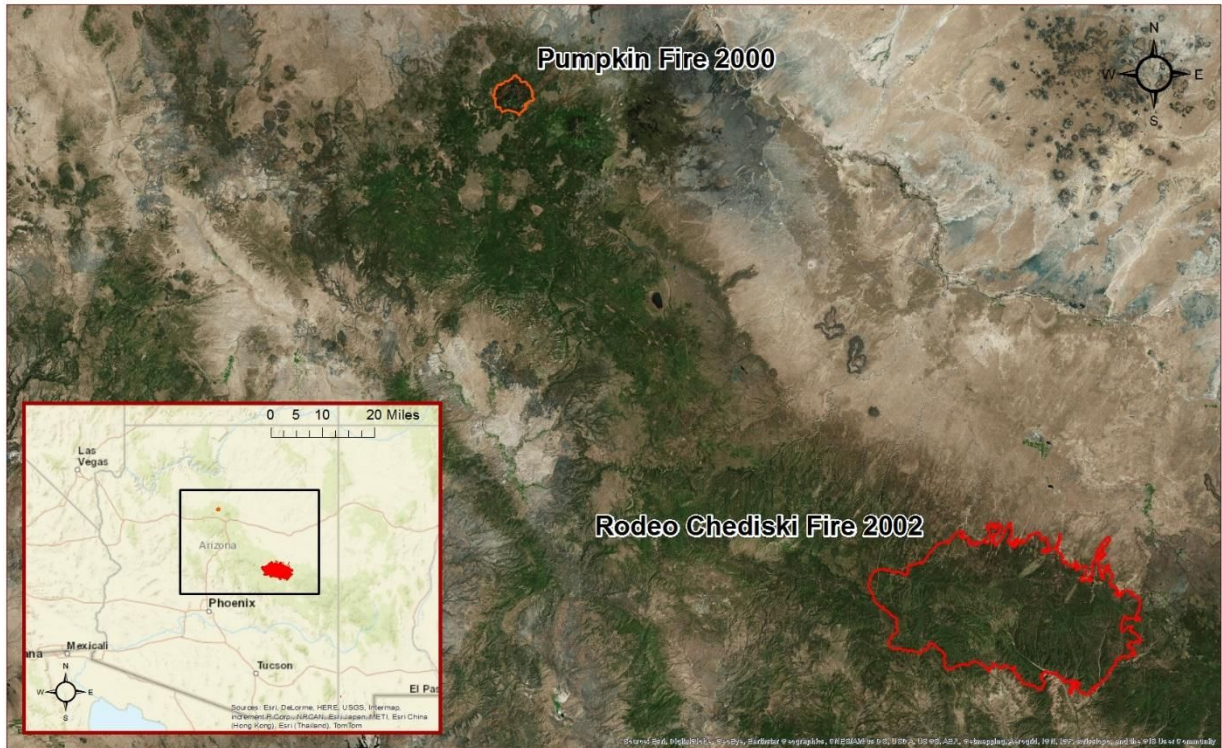
For this research project 12 - 200-m by 200-m plots were collected and will be used as the training data for delta NDVI values. Six high burn severity plots were collected from each fire. These plot locations were selected based on the location to alive trees and the high burn severity boundary from the Monitoring Trends in Burn Severity website (MTBS.gov). Three plots from each location were places along the edge of the high severity burn while the second was placed 200 meters from the nearest live tree but completely contained within the high burn severity area.

Map 1: Fire Locations

Location of the fires within Arizona

Created by Kimberly Istok

0 12.5 25 50 Miles



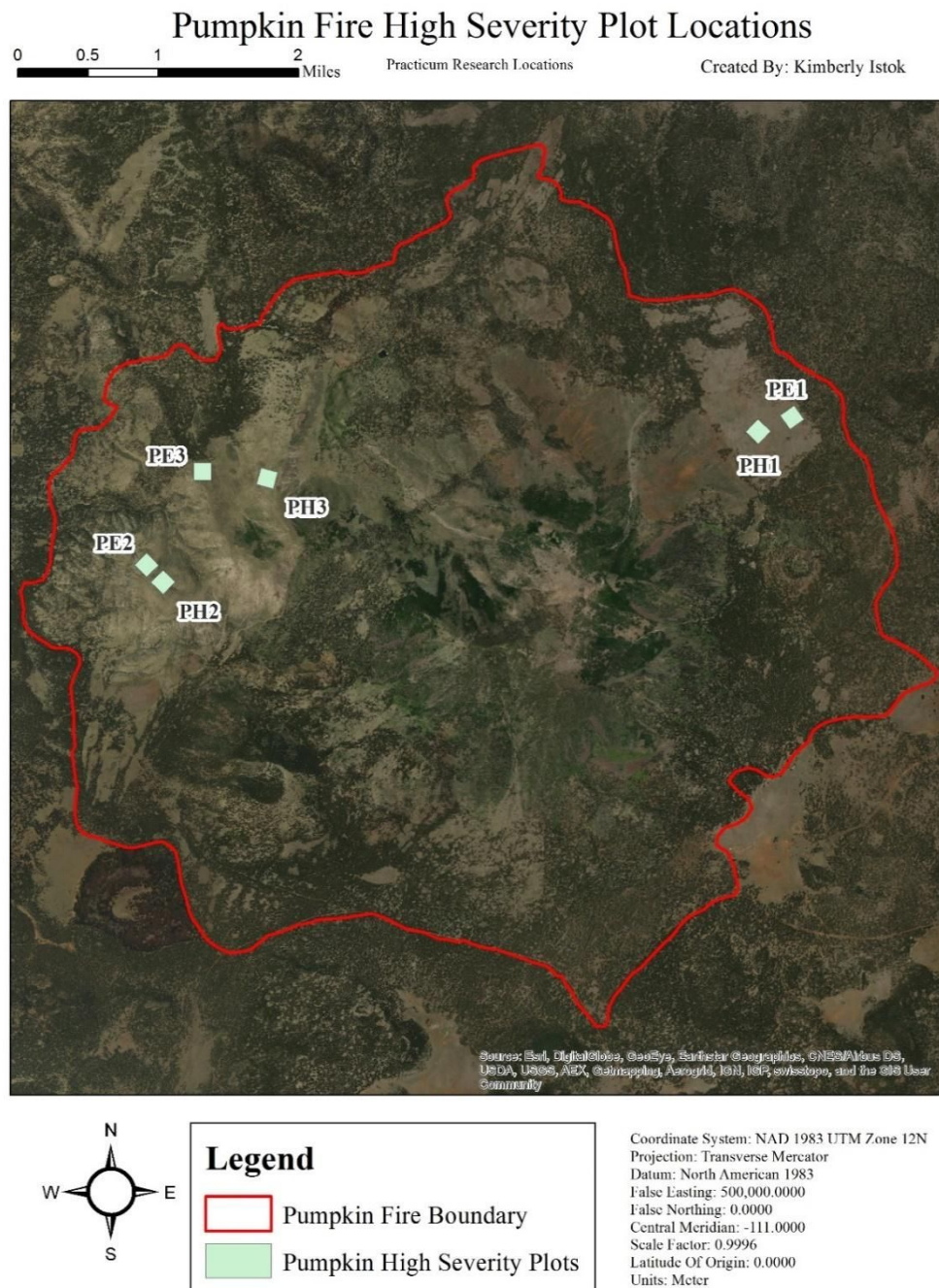
Pumpkin Fire: 2000

The Pumpkin Fire is the oldest wildfire in this research project, it started on May 24, 2000. Beginning by a lightning strike on the southwestern side of Kendrick Mountain Park (Jenkins et al. 2011). Kendrick Mountain is the second largest peak at 10,423 ft (3175 meters) within the San Francisco Peak mountain range, located north northwest of Flagstaff, Arizona (Jenkins et al. 2011). The forest that covers Kendrick Mountain is mainly composed of ponderosa pine and mixed conifers (Jenkins et al. 2011). Mixed conifer forests include ponderosa pine along with Douglas-fir (*pseudotsuga mensizii*), Quaking Aspen (*polulus tremuloidis*) and White fir (*Abis concolor*) (Jenkins et al. 2011). The burn was fully contained by June however it was not completely extinguished until August 2000 (Jenkins et al. 2011). The

burn encompassed 16,086 acres (6509.77 ha.), which is roughly 93 % of the Kendrick mountain (Jenkins et al. 2011).

Within the Pumpkin Fire six high severity plots were stem mapped. They are all located on the northern side of the mountain. There are three separate sites with two plots each. The plots are located on the western, central and eastern portions of the northern side of the mountain. These plots were selected based on the high severity burn areas. Three of the plots are along the edge of the high severity area, the others were placed 200 meters from the edge of the high severity. But all of the plots contained no surviving trees. The elevation within the plots ranged from 7901 to 8211 ft, plots located at an average elevation of 8028 ft (2244.6 m).

Map 2: Pumpkin Fire Plot Location

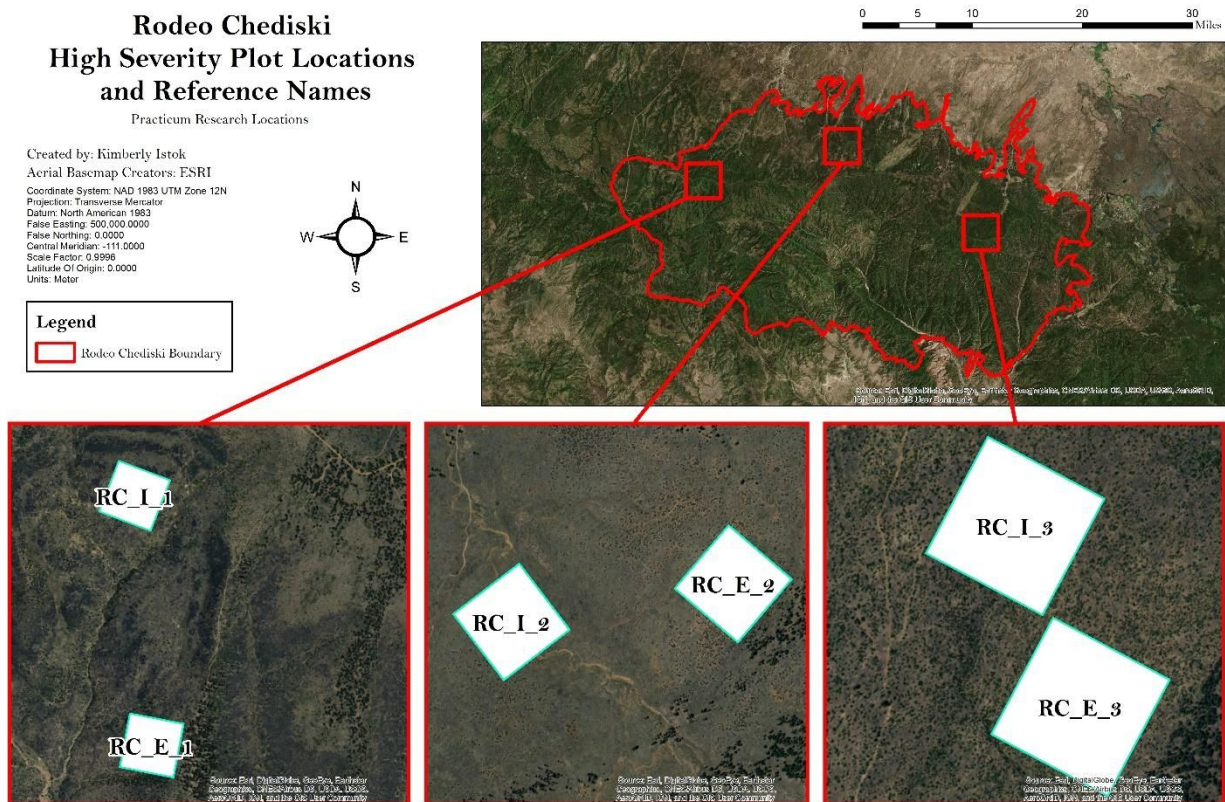


Rodeo-Chediski Fire: 2002

The Rodeo-Chediski Fire began as two separate fires. The first was set 2 miles northeast of Cibecue on the White Mountains Apache Reservation on June 18th 2002, this was named the Rodeo Fire (Kramer 2012). The conditions were perfect for the fire to spread. Arizona was in the middle of a drought at that time, and the fire soon grew uncontrollable (Kramer 2012). The Chediski Fire was started in the Sitgreaves National Forest, on June 20th, by a lost hiker (Kramer 2012). By June 22nd the two fires had combined to form the largest wildfire in Arizona history (Kramer 2012). This fire also spread to the Tonto National Forest and private land (Ffolliott et al. 2008), burning a total of 468,000 acres (189,392 ha.), and 450 homes (Kramer 2012) on four different land management types. Roughly 50% of the area burned by the fire was classified as high burn severity (Ffolliott et al. 2008).

The Apache – Sitgreaves National Forest which spans the Mogollon Rim is where the three sites for the plots were selected (Strom & Fule 2007). The plots elevations range from 6708 to 7455 ft with an average of 7033.8 ft (2143.9 m). The Rodeo-Chediski Fire has three sites, western, central and eastern areas of the fire. Each site included two plots, one along the edge of the high burn severity boundary and the other that was 200 meters from the high burn severity boundary and any surviving trees.

Map 3: Rodeo-Chediski Plot Locations



Methodology:

Literature Review:

For this project several resources and other satellite research projects were consulted. Based on the objectives of the project Landsat imagery was selected. This project utilizes only Landsat imagery. Landsat is not the only imagery that is available from EarthExplore.gov,

however it was chosen for this project because of its availability, this project requires dates as far back as May 2000 (Pumpkin Fire burn). Landsat imagery can be found as far back as 1982 depending on the satellite needed. The Landsat satellites pass over the same spot every 16 days, which allows for comparisons to be made over time on specified locations (Chen et al. 2010, Roder et al. 2006). Landsat imagery is also free to the user. An account must be created with both the Glovis or EarthExplorer websites but there is no other requirements.

The first objective for this project is to utilizing ground data in correlation with satellite imagery, however this is a practice that is not as common in the literature. In Cocke et al. 2005, tested ground tree mortality data to the Differenced Normalized Burn Ratio (NBR). The NBR is a combination of near-infrared band and mid-infrared bands of Landsat imagery (Chen et al. 2011, Escuin et. al 2008). A change detection was done between Landsat 7 ETM+ imagery that was pre the 2001 Leroux fire and one year after the burn to determine the NBR. The NBR value was then compared to Ninety-two - 20 by 50 meters plots that were measured before the fire and after the fire. The attribute collected for each of the plots, tree species, height, and diameter at breast height (DBH). These plots were also measured for condition of tree as well as debris, snags and understory information before the fire and after the fire (Cocke et al. 2005). This study determined that the medium severity plots had the greatest mortality occur. The authors equate this to the species that were present within the medium severity plots verse the high severity plots. However the authors state that there was a strong association seen between the satellite severity selections and the ground/understory debris severity seen in the 92 plots. It was determined that NBR was a reliable method for mapping the burned areas that need post fire management techniques based on burn severity (Cocke et al. 2005).

For burn severity this project will be using the high burn severity areas calculated by MTBS.gov. However for the regeneration analysis NDVI has been selected for this analysis. Normalized difference vegetation index or NDVI will be used because NDVI derived from Landsat data is the most widely used method when assessing wildfire vegetation recovery (Veraverbeke et al. 2010). NDVI is the reflectance of photosynthetic activity (Fernandez et al. 1997). The NDVI technique uses red and near-infrared band reflectance from the vegetation's chlorophyll (Yi et al. 2013). There is a long history in the literature that NDVI is utilized for measuring post fire vegetation recovery because after a high severity fire the NDVI value drop significantly because of the absence of photosynthetic material (Chen et al. 2011, Veraverbeke et al. 2010). NDVI has also been found to have more similarities with field estimates than any other index that can be used (Veraverbeke et al. 2010). The formula for NDVI is:

$$NDVI = \frac{\text{Near Infrared band} - \text{Red band}}{\text{Near Infrared band} + \text{Red band}}$$

NDVI is the best index for vegetation because it allows for there to be natural variation, meaning that even though all the plants are not the same it still can produce a reliable estimate (Veraverbeke et al. 2010).

Since the NDVI values decrease after high severity wildfires (absence of green reflectance) it is an excellent base for a change detection comparison. The change detection method involves an equation that is used to describe the relationship between the two scenes in comparison (Chen et al. 2011, Escuin et al. 2008, Fernandez et. al 1997, R. Diaz-Delgado et al. 2013, Yi et al. 2013,). These equations can be just between two different year scenes or they can be calculated between multiple years. Chen et. al 2011 compared the differences between six different Landsat derived NDVI values over 7 year after the Jasper Fire in Black Hills National

Forest, South Dakota. One of the most common change detection equations is subtracting the older date NDVI from the more recent (Chen et al. 2011, Cocke et al. 2005, Escuin et al 2008). Below is the equation that will be used for this analysis:

$$NDVI_{newest} - NDVI_{oldest} = NDVI_{change}$$

The third objective of this project is to utilize ground data and delta NDVI values to estimate the amount of conifers that are present within the regeneration of the Pumpkin Fire. The Pumpkin Fire was only utilized for this because of the available satellite data and because this fire had the largest amount of conifers recorded in the regeneration. Utilizing NDVI and Landsat imagery to identify species is limited in the research. However an experiment that was similar to this has been done in Phoenix's Active Management area (Zheng et al. 2015). NDVI was utilized to see if certain crops had specific NDVI ranges. The NDVI values were derived from a Landsat satellite, meaning the resolution was the same. This research started by tracking the growth cycle of the farm crops, in order to establish NDVI value ranges for each crop type. However they found that the same crops varied from farm to farm and even field to field in the same farm (Zheng et al. 2015). The authors contribute this difference to dates of when the seeds were sowed, weather, soil conditions and crop management (Zheng et al.2015).

Data Utilized:

There were three major types of data that were utilized for this project. The first was free Landsat 7 ETM+ and Landsat 8 OLI Satellite scenes downloaded from EarthExplorer (<https://earthexplorer.usgs.gov>). A total of six scenes were processed for this analysis. The second type of data that was used was collected by the Rocky Mountain Research Station field

teams over the 2014-2016 summers. There was a total of 18 plots collected but only 12 were used. This data was not free to obtain. It was funded through grant monies. The third type of data that was used was raster data downloaded from the Monitoring Trends in Burn Severity website. This website provided both fire boundaries as well as the burn severity designations for each fire.

Landsat imagery:

A little background on Landsat Satellites imagery. There have been 8 satellites sent up into space collecting imagery since 1972 (Wulder et al. 2012, Chen et al 2011). This information has been used to shows how the Earth has changed over almost 50 years of climate change and human interaction (Wulder et al. 2012). In 2008, all new and previously collected Landsat data was made public by the United States Geological Survey through Glovis and Earthexplore websites (Wulder et al. 2012).

Landsat imagery is collected on a polar orbit of the Earth allowing for the 16 day cycles (Irons 2017). The satellite measures the amount of light that is reflected off of the Earth's surface from the sun (Irons 2017). The Landsat satellites collect wavelengths from visible light spectrum (Red, Orange, Yellow, Green, Blue, Indigo and Violet) and infrared light (Irons 2017). The standard resolution for Landsat imagery is 30 meter by 30 meter pixels (Irons 2017). The scene size that is available for download are 106 miles by 115 miles.

The Landsat satellite series has multiple satellites that have imagery available however the band and wavelength technology has changed over time. Landsat 1 launched in 1972 contained two camera systems, the return beam vidicon (RBV) and the multispectral scanner (MSS) (Irons 2017). Originally the RBV was to be the prime focus but the information collected

by the MSS was found to be better (Irons 2017). Four spectral bands were recorded by the MSS sensor, green, red and two near infrared bands (Irons 2017). Landsat 1 operated until 1978 when it was decommissioned (Irons 2017). Landsat 2 (launched in 1975 and decommissioned in 1983) and Landsat 3 (launched in 1978 and decommissioned in 1983) also carried RBV and MSS sensors based on the sensor on Landsat 1 (Irons 2017). Landsat 4 launched in 1982 carried the MSS and the thematic mapper (TM) sensor. TM was able to see a wide electromagnetic spectrum which improved the spectral and spatial resolution with 7 spectral bands (Irons 2017). Within a year the two main transmitting ports were no longer working and Landsat 5 was sent up in 1984 with the same sensors TM, MSS. Landsat 4 was continued to be used with limited data collection and transmission capabilities until its decommission in 2001 (Irons 2017).

Landsat 6 was sent up in 1993 but this satellite was lost at launch (Irons 2017). It contained a new sensor called Enhanced Thematic Mapper (ETM), this sensor included the same 7 bands as were previously seen on Landsat 4 and 5 but it collected an additional band, band 8 (Irons 2017). Band 8 is the panchromatic band that has a spatial resolution of 15 meters. Landsat 5 was continued to be used until 2013 when it was decommissioned (Irons 2017). In 1999, Landsat 7 was sent up with the ETM+, Enhanced Thematic Mapper Plus sensor (Irons 2017). Similar to Landsat 5 ETM sensor but also with other features included: on-board, full aperture, 5% absolute radiometric calibration, thermal infrared channel with 60 meter spatial resolution, on-board data recorder as well as the panchromatic band with 15 meter resolution (Irons 2017). In May 2003, 6 weeks into the life of Landsat 7 satellite the scan line collector failed causing striping over 75% of the collected scenes (Irons 2017). USGS is working to find the best ways to merge the data products for use, however the best method to do this has not been

determined (Irons 2017). It is still currently in use collecting striped imagery. Landsat 8 was launched in 2013 and contains two sensors, Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) (Irons 2017). Spatial resolution for the Landsat 8 is 30 meters in the visible spectrum, near infrared and short infrared, 100 meter for thermal, and 15 meter resolution for the panchromatic band (Irons 2017).

Figure 1: Landsat 7 ETM+ Band and Wavelength List

Landsat 7 ETM+ (Enhanced Thematic Mapper Plus)	Bands	Wavelength (Micrometers)	Resolution (meters)
	Band 1 – Blue	0.45-0.52	30
	Band 2 - Green	0.52-0.60	30
	Band 3 – Red	0.63-0.69	30
	Band 4 – Near Infrared (NIR)	0.77-0.90	30
	Band 5 – Shortwave Infrared (SWIR) 1	1.55-1.75	30
	Band 6 – Thermal	10.40-12.50	60* (30)
	Band 7 – Shortwave Infrared (SWIR) 2	2.09-2.35	30
	Band 8 – Panchromatic	.52-.90	15

Figure 2: Landsat 8 OLI TIRS Band and Wavelength List

Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS)	Bands	Wavelength (Micrometers)	Resolution (meters)
	Band 1 – Ultra Blue (coastal/aerosol)	0.43-0.45	30
	Band 2 – Blue	0.45-0.51	30
	Band 3 - Green	0.53-0.59	30
	Band 4 – Red	0.64-0.67	30
	Band 5 – Near Infrared (NIR)	0.85-0.88	30

	Band 6 – Shortwave Infrared (SWIR) 1	1.57-1.65	30
	Band 7 – Shortwave Infrared (SWIR) 2	2.11-2.29	30
	Band 8 – Panchromatic	0.50-0.68	15
	Band 9 – Cirrus	1.36-1.38	30
	Band 10 – Thermal Infrared (TIRS) 1	10.60-11.19	100*(30)
	Band 11 – Thermal Infrared (TIRS) 2	11.50-12.51	100*(30)

<https://landsat.usgs.gov/what-are-band-designations-landsat-satellites>

There were six scenes that were utilized for this analysis. See chart below for Landsat satellite and scene information:

Fire Name & Scene	Satellite	Date
Pumpkin Fire Scene 37,35	Landsat 7 ETM+ Landsat 8 OLI	06-12-2002 07-12-2016
Rodeo-Chediski Fire Scene 36,36	Landsat 7 ETM+ Landsat 8 OLI	07-7-2002 06-03-2016
Conifer Pumpkin Fire Scene 37,35	Landsat 7 ETM+ Landsat 8 OLI	11-19-2002 12-19-2016

The three groups of Landsat scenes that were selected were based mainly on the data that was available. The original idea for the satellite imagery was to have the Pumpkin and the Rodeo-Chediski tests have the same seasonality and time gap from the start of the fire. However because of the data that was available two different time gaps were used. The data scenes that were selected for this analysis were collected by two different satellites, Landsat 7 ETM+ and Landsat 8 OLI TIRS. These satellites were chosen mainly off the scene clarity from cloud obstructions and the seasonality. The goal was to find previous image and current images from the same season. June to September month images were considered for the Pumpkin Fire and Rodeo-Chediski Fire. There is some correction that can be made for clouds however the best

scenes to use are free from cloud cover. However other factors for choosing Landsat 7 and Landsat 8 did come into play. Both Landsat 7 and 8 have the same resolution for the visible and near infrared bands 30 –meter by 30 –meter. The satellites also capture the same scene location every 16 days (Flood 2014, Irons 2017). Another important similarity between Landsat 7 and Landsat 8 is that they collect the same visible and near infrared wavelengths needed for NDVI change detection.

For the Pumpkin Fire, the earlier image from 2002 is two years post-fire and the current image is from 2016 – 16 years post fire. The theory behind this analysis starts with the grasses seen after the fire. During the 2 years after the fire grasses are the first to come back, hence the pixel NDVI values will be reflections from grass types in the area. In the 2016 image the NDVI value is a function of both grasses and the regenerated trees. By using the change detection equation $B2(2016) - B1(2002)$ in theory it should remove the grass NDVI value from the 2016 scene, the remaining NDVI value would only be the regenerated tree value.

For the Rodeo-Chediski Fire the Landsat scenes were selected differently based on availability. Rodeo-Chediski scenes were not able to be selected a few years after the fire because of the down sensor on the Landsat 7 ETM+ satellite from 2003 to current each scene has striping over the 75% of the scene. There are methods that could help correct striping however it enters in a high level of error into the scene, because the missing pixel values are being estimated off of the established pixels. The current Landsat image is from 2016 making the two comparative scenes are 14 years apart. The 2002 scene was taken directly after the fire had been extinguished. It shows the area is burned with very little or no greenery. This change detection

will show both the grass and the trees. However with the use of the training ground data this should show some correlation between the number of trees and the NDVI value.

Landsat 5 satellite images were checked for cloud obstructions during the months of June – August from years 2002 -2005 for the Rodeo-Chediski however there was a lot of cloud disturbance and the scenes could not be used.

For the third test, determining the acreage of regeneration of conifers on the Pumpkin Fire winter scenes were found in the November to February month range, reviewed for cloud cover and snow obstruction. The earlier scene was taken 2 years after the 2000 Pumpkin Fire, while the second scene was taken in 2016. Neither of the scenes have visible snow cover within them. The thought process for this analysis was to select clear scenes immediately after the fire and compare it to a more recent scene. Because both grass and the Gambel oak trees are dormant or dead in the winter, and the NDVI pixel values will only be reflecting the greenery of the junipers and the ponderosa pine seedlings. This will allow for an estimate of the amount of conifers that are returning.

Tree point data:

This data was collected by the Rocky Mountain Research Station (RMRS) in Flagstaff, Arizona. This layer was collected, created and processed by the field technicians. The collection was done over three summers, 2014 – 2016. A team of four technicians were used for a majority of the summers. The field crews utilized a Trimble unit along with a range finder to map the spatial location of each tree or seedling in the 200-m by 200-m plots within each fire. The process that was used for mapping began with placing a reference locations with sub meter satellite accuracy. Then utilizing the range finder to map each tree by obtaining the offsetting

distance for each tree location. The field technicians would then relay the information of each of the trees or seedlings to the Trimble unit operator. Each location was of one individual tree or seedling. Clumps were used to help expedite the process of mapping, however this clumping could only be used for trees of the same species or seedlings of the same species. The difference between a tree and a seedling was determined by the height of the sample. Seedlings were determined as species less than or equal to 1.4 meters in height. For seedlings only height, species type, clump radius, and clump count (if there was a clump) were recorded. For trees (taller than 1.4 meters in height) in addition to the already listed attribute for the seedlings, diameter at breast height (DBH) was collected. Diameter at breast height means the diameter of the tree at 1.4 meters. The attributes: height, DBH and clump radius were measured with a combination of estimation and exact measurement in meters by the field crews. The author has worked with this data since summer 2015.

The overall goal of this project is to create a process that can be re-creatable. Something that needs to be mentioned is that this point data was not free. The technicians and tools were all paid for with grant monies that were obtained by the Rocky Mountain Research Station. This data's primary function was to fulfill other research requirements. However this data has been used in several other research projects other than the intended use. Each use of this data has been working towards tracking forest change after wildfires. It is very valuable research data and is worth finding or producing the funding needed for ground data collection. This data is a crucial piece to this analysis.

Monitoring Trends in Burn Severity: (MTBS.gov)

Monitoring Trends in Burn Severity website was created in order to establish a standardization of all the burn severities across all the wildfires after 1984. This website is managed by a multi-agency oversight group, the US Geological Survey National Center for Earth Resources Observation and Science and the USDA Forest Service Remote Sensing Applications Center (RSAC). This data has been used to provide information for measuring the National Fire Plan and Healthy Forest Restoration Act (MTBS.gov). All of the Landsat imagery that is provided in the download is 30-m resolution and is free to download. This information comes in as package of data. The metadata that is provided shows the calculations of how the severity levels were classified for each fire. The package also includes pre-fire and post-fire Landsat imagery. Fire boundary and the raster imagery for the severity levels.

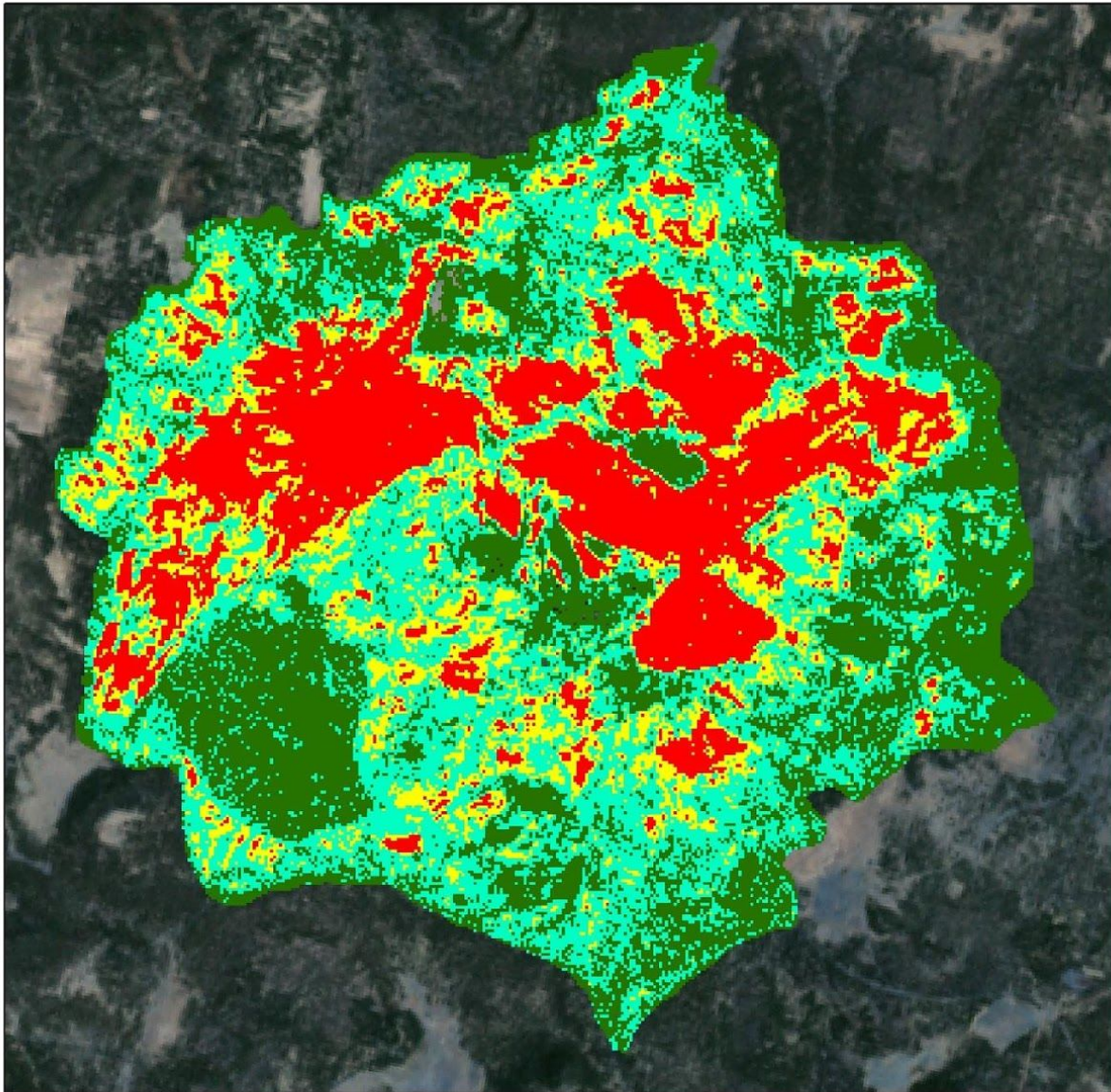
The raster imagery of the severity levels was the important part of the package for this project. This raster image was converted into a polygon to allow for selection of the final delta NDVI range in order to calculate the tree regeneration within high burn severity areas. Because it is in raster form this layer had to be converted into a vector data form, polygon. The raster image was first clipped by the fire boundary because the raster images are so large that is often bogs down the processing speed. The next step was to convert the raster clip to a polygon, not simplified. The not simplified selection creates blocky polygon boundaries that follow the pixel boundaries. This is needed because the Landsat imagery with the selected delta NDVI range will be following the pixel boundaries. Only code 4 (High Burn Severity) was kept for the polygons, and they were merged together to form one record.

Map 4: Pumpkin Fire Burn Severity

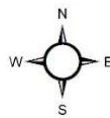
Pumpkin Fire Burn Severity

MTBS.com provided the burn severity levels

Created by: Kimberly Istok



Coordinate System: Albers Conical Equal Area
Projection: Albers
Datum: North American 1983
false easting: 0.0000
false northing: 0.0000
central meridian: -96.0000
standard parallel 1: 29.5000
standard parallel 2: 45.5000
latitude of origin: 23.0000
Units: Meter



Natural Colors, Landsat 7 +ETM, 37,35
June 2002

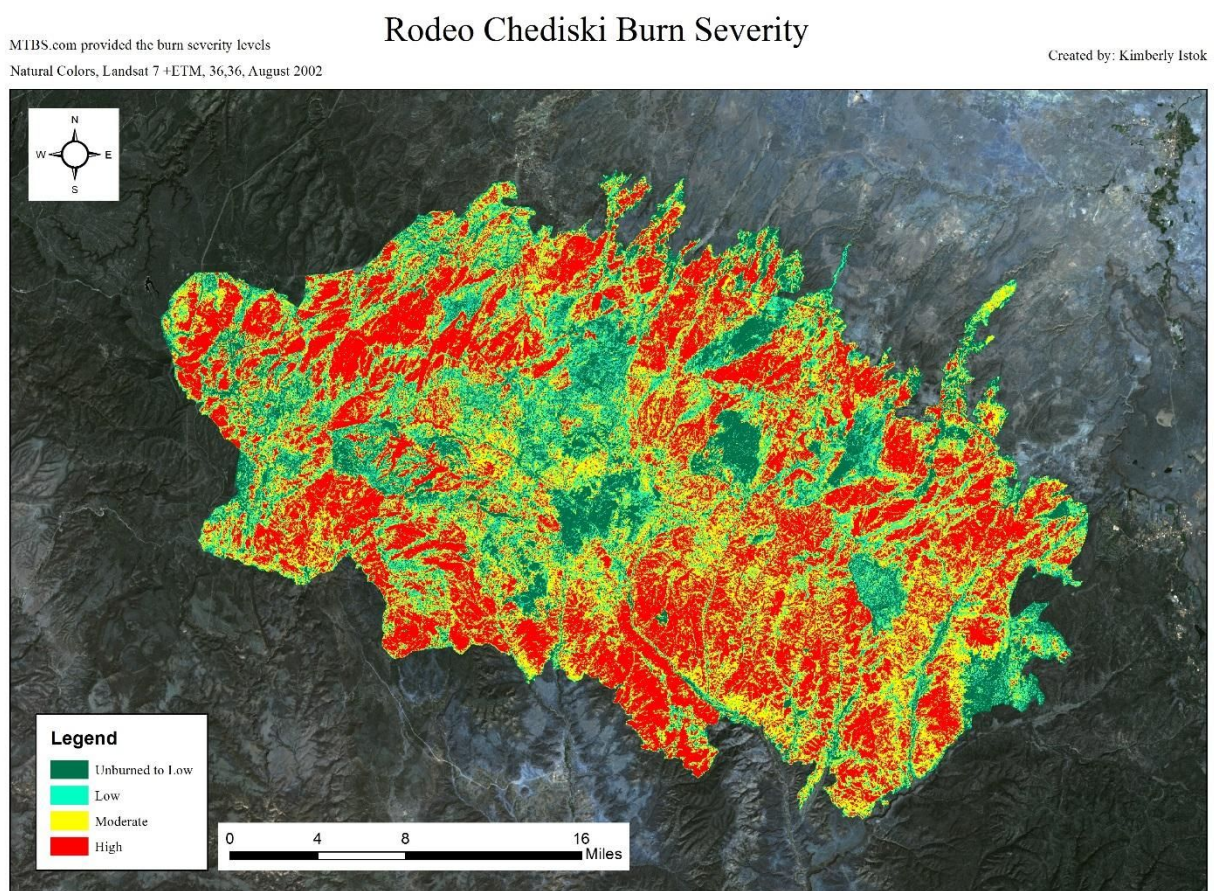
0 0.5 1 2 Miles

Legend

- Unburned to Low
- Low
- Moderate
- High

The Pumpkin Fire was found to have 4,595 acres in unburned to low severity, 5,363 acres in low burn severity, 2,572 acres in moderate burn severity, and 3,536 acres with high burn severity for a grand total of 16,086 acres effected. This analysis done by the MTBS.gov team was conducted with pre-fire image from Landsat 5 – June 25, 1998 and post-fire image Landsat 7 – June 09, 2001.

Map 5: Rodeo-Chediski Fire Burn Severity



The Rodeo-Chediski fire was found to have 69,616 acres in unburned or low severity, 76,754 acres in low burn severity, 146,296 acres in moderate burn severity, 169,042 acres in high burn severity for a total of 461,753 acres effected. This analysis done by the MTBS.gov team

was conducted with pre-fire image from Landsat 7 – May 01, 2001 and the post-fire image from Landsat 5 – April 29, 2003.

Methodology Performed:

Preprocessing of Landsat Imagery:

The preprocessing of the Landsat imagery is a crucial step to ensure that the analysis is correct. There are several different methods and programs that can be used for preprocessing. These Landsat images were preprocessed in ENVI 4.8 and 5.3 classic. Because two different sensors were used two different method had to be applied. Landsat 7 ETM+ was preprocessed in Envi 4.8. Bands 1-5 and 7 were preprocessed. First each band was converted to radiance from the digital number format that it was delivered in. The 6 bands were then stack and the band names and wavelengths were then edits to ensure proper identification for the analysis portion. The stacked bands were then converted from band sequential BSQ raster file to band interleaved by line (BIL) raster file. Once the BIL conversion is complete a good check within the preprocessing is to see if the pixels fall within the proper bit range. Landsat 7 data is 8 bit data and the pixel values after the BIL conversion should fall between 0-256 (2^8).

The final step in the preprocessing is the use of the FLAASH tool. FLAASH stands for the fast line-of-sight atmospheric analysis of spectral hypercubes. This tool performs an atmospheric correction taking the influence of the atmosphere out of the values that are used for analysis. The atmospheric influence is seen in all of raw data from satellites because the sun wavelengths travel from the earth to the satellite. The FLAASH tool corrects wavelengths in the visible spectrum, near infrared and short-wave infrared regions. For multispectral data like we are using here the FLAASH tool takes into account the water vapor, time of day, season, scene

elevation, scene visibility and sun angle when correcting. Landsat 7 – 8 bit data after the FLAASH tool will have pixel ranges form 0-10,000.

Landsat 8 OLI TIRS preprocessing is a little different, because Landsat 8 comes with a multispectral layer (MTL layer) with all of the visible bands, near infrared and short infrared wavelengths already layered together and ready for conversion to radiance. Radiometric calibration in ENVI 5.3 was used to convert the Landsat 8 images to radiance. With the radiometric calibration tool the user is able to specify the exported raster type and the product was selected to BIL. This extra step allowed for faster processing time with Landsat 8 than Landsat 7. Once the Radiometric calibration was done the pixel values were checked and fell within 0 – 65535. This number is very different from the Landsat 7 because the Landsat 8 data is 16 bit. (2^{16}). The FLAASH tool was used in ENVI 5.3 because ENVI 4.8 does not have Landsat 8 sensor within the drop down for selection. After the FLAASH tool was complete the pixel values for Landsat 8 pixels should be 0 – 4096. By double checking the pixel values between each step human error can be removed from the results. This is an important step for preprocessing that is commonly missed. A total of 8 Landsat Images were preprocessed for analysis and map creation.

Clean up of field collected ground data:

Originally when the field crews were sent out into the field to begin collecting, protocol for entering data was established. Domains were used to help increase the collection speed and reduce spelling errors. However, the second summer a different Trimble unit and different operators/field crew were doing the collection. Whenever you have multiple people collecting data, everyone recorded the data a little different. The point layer was cleaned up to match the

established and approved data variable names. For example ponderosa pine was sometimes entered as PIPO, Pipo, pipo and pine. When doing queries the numbers would all vary, so to avoid any confusion. The data was cleaned up to have uniform species entries and uniform plot names. Another check that was done for this data consisted of making sure all the records with 1.4 or lower were marked as seedling and anything higher than 1.4 labeled as tree. Metadata was created for this data because this data belongs to the RMRS and will be returned there after the completion of the project.

When working with a range finder if the locations of the trees are difficult to see or are blowing in the wind points can be over shot. In order to eliminate this from the data set all the points that had delete in the comments field or where located outside of the plot were removed from the point dataset. Another removal from the dataset that was made was the reference points. Each plot had roughly 10-20 reference points for where the range finder offset started. In an effort to make sure that these points were not accidentally counted as trees, a definition query was used to remove them. Definition queries are not permanent, they just only show the desired records.

Creation of a Grid Layer for Reference:

The next step in the processing was to create a layer that brought together the information from the ground point layer and the location of each tree within the Landsat pixels. This reference layer was completely created by the author and was designed to incorporate all of the information into one layer. The attribute fields that were created for this layer are listed below:

TotalTreecount; (Short) = Total number of trees present in the pixel square.

QuagaCount: (Short) = Total number of Gamble oak.

PipoCount: (Short) = Total number of Ponderosa Pine.

JuniCount: (Short) = Total number of Juniper.

ManzCount: (Short) = Total number of Manzanita.

OtherCount: (Short) = Total number of other species, Aspen was one of these.

Majority: (Text) = Species name that had the most trees within each grid

PercentMajority: (Short) = Percent of the majority =

$$\text{Majority Species Count} / \text{Total trees} * 100$$

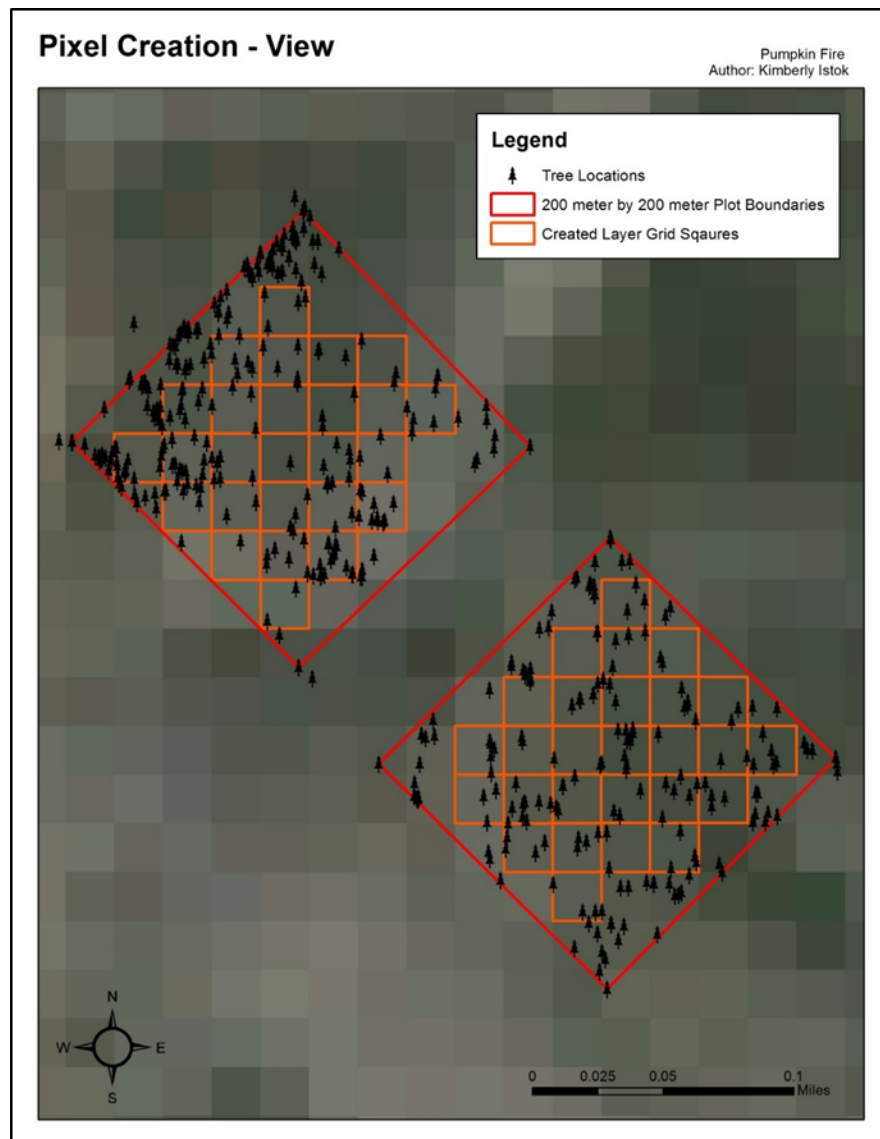
Delta NDVIValue (Float) = The NDVI Value of the pixel square after NDVI tool and Change detection equation.

Winter Delta NDVIValue (Float) = The NDVI Value of the pixel square after NDVI tool and Change detection with the Winter scene and original scene.

CanopyPercent: (Float) = The percentage of area that is covered by trees. Calculated from stem map data.

The fields that are not listed include fire plot number and recordID which was the unique identify for this layer. The attributes listed above are color coded to help with processing explanation. Because the pixel layer is only a reference layer to combine the pixel location and tree point data the pixel boundaries were manually digitized. Digitization of each pixel that fell completely within the 12 fire plots boundary. Because the pixels are all 30-m by 30-m the advance editing tools length and perpendicular were used to create 30-m by 30-m squares that followed along the pixel boundaries. See map 6 below for clarification.

Map 6: Pixel layer digitization example



The attributes listed above in green were manually counted using the select by location tool. The attribute field Majority was entered based on the species that had the majority count in each pixel square. Once Majority was identified, PercentMajority (listed in purple) was calculated. Based on Majority of each species per pixel. This was queried for and Field Calculator was used with the following equation to calculate the PercentMajority of the species within each pixel. The equation that was used:

$$\frac{\text{Majority Species Count}}{\text{Total trees}} * 100$$

PercentMajority attribute will be used to identify which pixels have 75% species majority of junipers and ponderosa pines for the conifer test.

NDVI Calculation:

Once the grid layer was completely digitized and all of the attribute fields were populated, NDVI was calculated for each of the 6 Landsat scenes. Envi 5.3 has a NVDI tool that performs the NDVI equation. The following equation was used for the three Landsat 7 images:

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

Because Landsat 7 and Landsat 8 sensors both collect different bands Landsat 8 image equation is:

$$\frac{NIR_{Band\ 5} - R_{Band\ 4}}{NIR_{Band\ 5} + R_{Band\ 4}} = NDVI$$

Band location is something that must be tracked throughout the entire preprocessing step for Landsat imagery. During both methods for Landsat 7 and Landsat 8 there can be band position mix ups resulting in odd results for NDVI calculations. This is something that the author did experience, but caught the error and corrected it for the final analysis. After the

NDVI tool all of the pixel values should be between -1 and 1. If this is not the case, this is another indication of preprocessing error or band positioning error.

Finding Delta NDVI:

In order to figure out the amount of growth after the wildfires a change detection will need to be done for identifying the areas of new tree growth. With the three tests that were established the first the Pumpkin Fire equation will look at 2 years after the fire with the NDVI value of just grasses subtracted from the 2016 NDVI value with trees and grasses, to provide just the trees NDVI value. The second test on the Rodeo-Chediksi Fire looks directly after the fire and subtracts any immediate 2002 greenery NDVI value from 2016 NDVI, this will show potentially grasses and new trees. The Pumpkin Fire Conifer test will be looking at winter NDVI values 2 years after the fire and winter NDVI values from 2016.

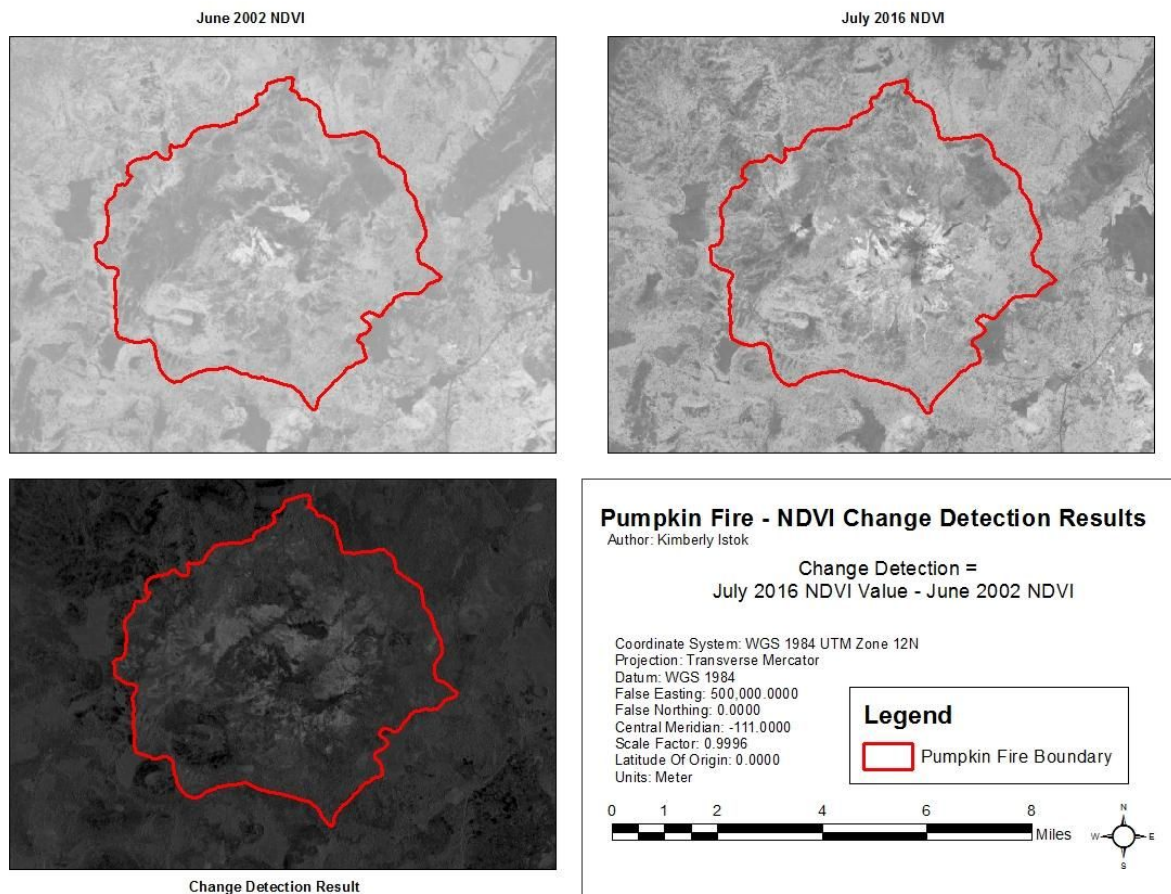
When analyzing the difference between two separate years a change detection is an important relationship to be able to quantify. By utilizing an equation a value allows for a deeper understand of the relationship. The higher the value the bigger the difference between the years. The change detection equation that was used for change over time for each of the tests was:

$$NDVI_{2016} - NDVI_{2002} = \text{delta NDVI}$$

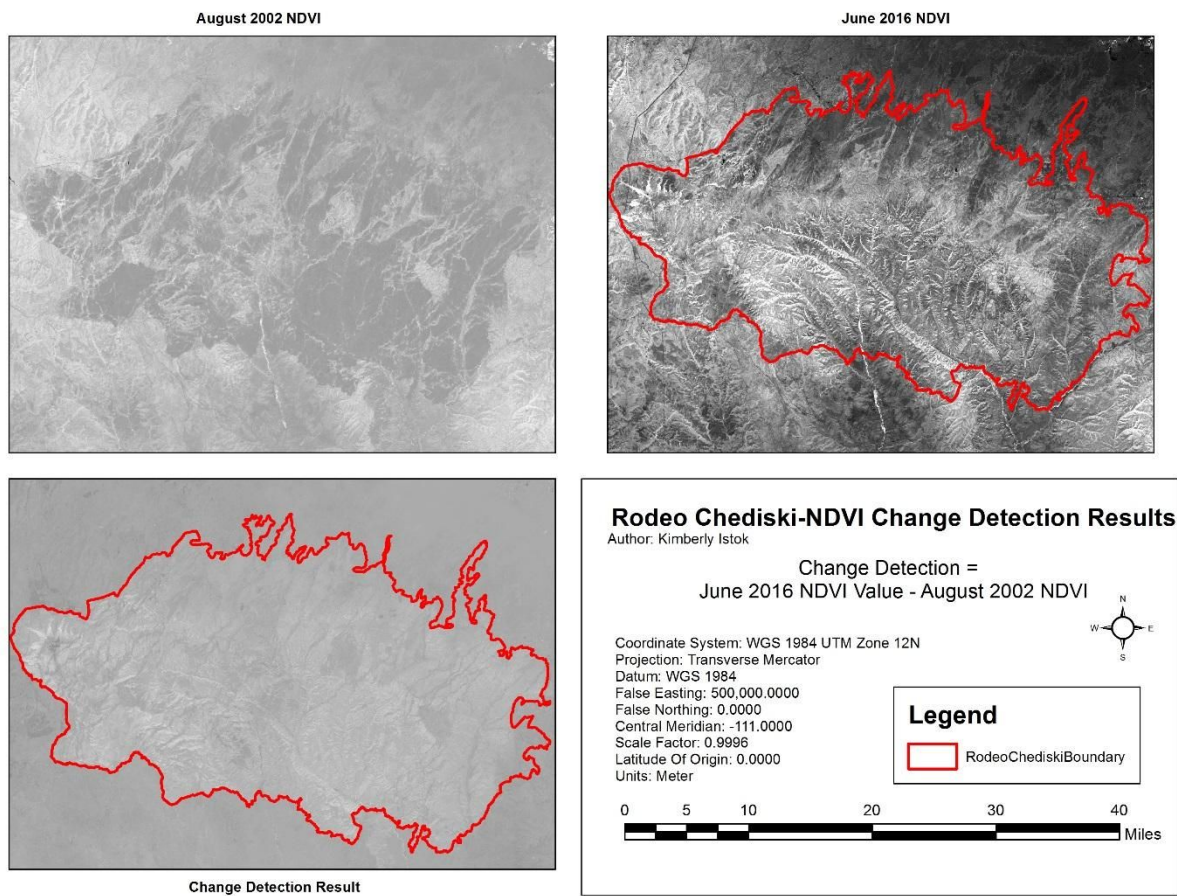
All of these equations were calculated the same way by using Band Math tool in Envi 5.3. Before the Band Math tool could be used each pair of NDVI Landsat Images had to be Layer Stacked together. Keeping track of band position during this step is also important because once they are stacked together it is hard to tell which one is which. The equation that was entered in the Band Math tool was B2-B1, meaning the more recent image was B2 and the

earlier image was B1. Once the NDVI and delta NDVI were completely calculated the attribute fields in blue were entered for each of the pixels. The author did this manually however it could be scripted in python to increase processing time.

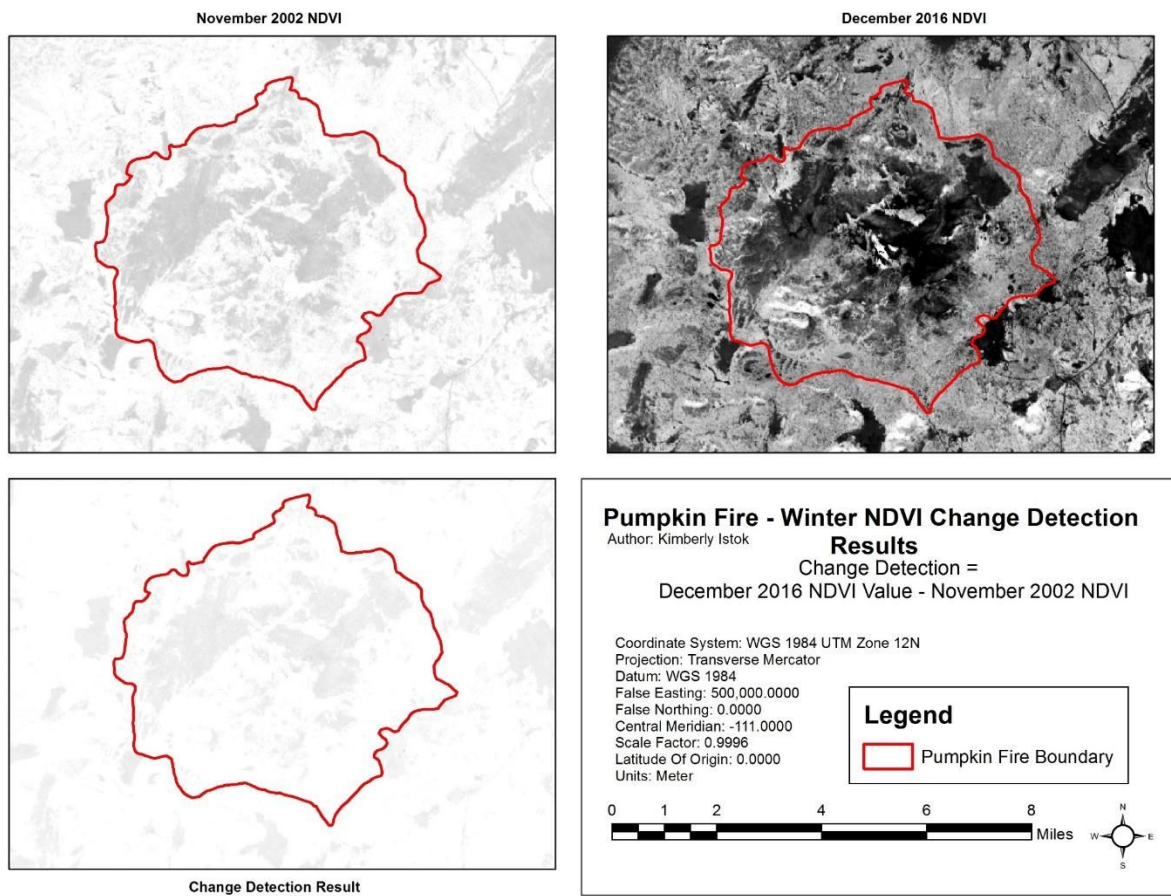
Map 7: Pumpkin Fire NDVI Change Detection



Map 8:Rodeo-Chediski Fire NDVI Change Detection



Map 9: Pumpkin Fire Conifer Test Change Detection



Validating Regeneration:

NDVI values are based on the reflection of the photosynthetic material in plants. Because Landsat imagery resolution is so large (30-m by 30-m) in comparison to seedlings determining what exactly is causing the greenery reflection is important in validating this method for accuracy. Greenery could be grasses, shrubs, and trees. By defining canopy cover percent for each pixel utilizing the ground collected point data, a correlation can be tested between the delta NDVI value and the ground tree data. That way a delta NDVI range can be identified as regeneration.

The canopy cover percent per pixel was calculated by first buffering each of the tree species with an average meter buffer to account for the potential greenery shown by the trees from above. The species were buffered individually by the following distance: ponderosa pine 1 meter, Gamble oak 2 meters, juniper 1 meter, manzanita 2 meters, and other species 1 meter. The distance in which each tree species was arbitrary and standardized to make it easier to calculate. The individual buffer layers were then merged together to form one layer with all the species buffered. A spatial join based on location was then used for calculating the area of the buffers in each pixel. Canopypercent (in orange) attribute field was then able to be calculated by:

$$\frac{\text{Area Sum of Canopy Cover}}{900 (\text{Total Area of Landsat pixel})} * 100$$

This value was then plotted on the x axis with the delta NDVI plotted on the y axis for each pixel. These scatterplots will be used to identify areas of clumping allowing for a range of regeneration to be estimated. The following scatter plots were created in ArcMap.

Figure 3: Pumpkin Fire Scatterplot

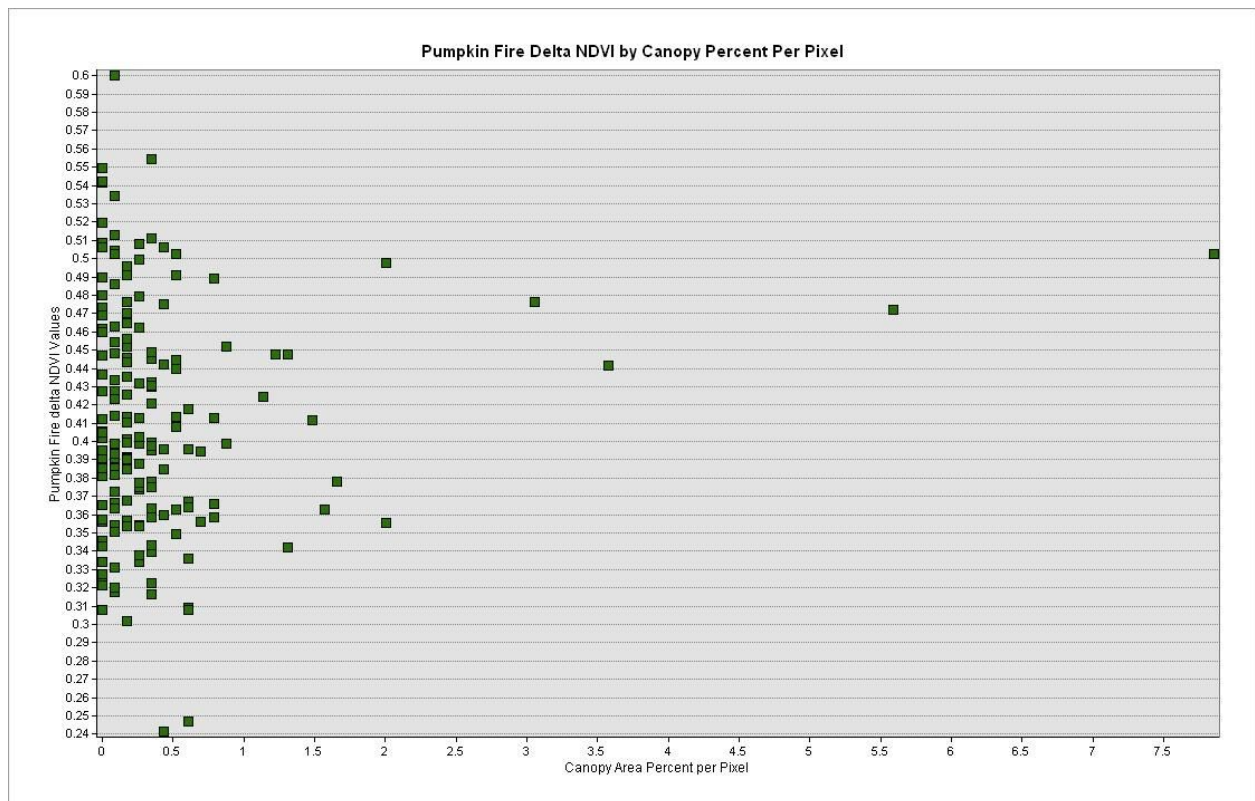


Figure 4: Rodeo-Chediski Fire Scatterplot

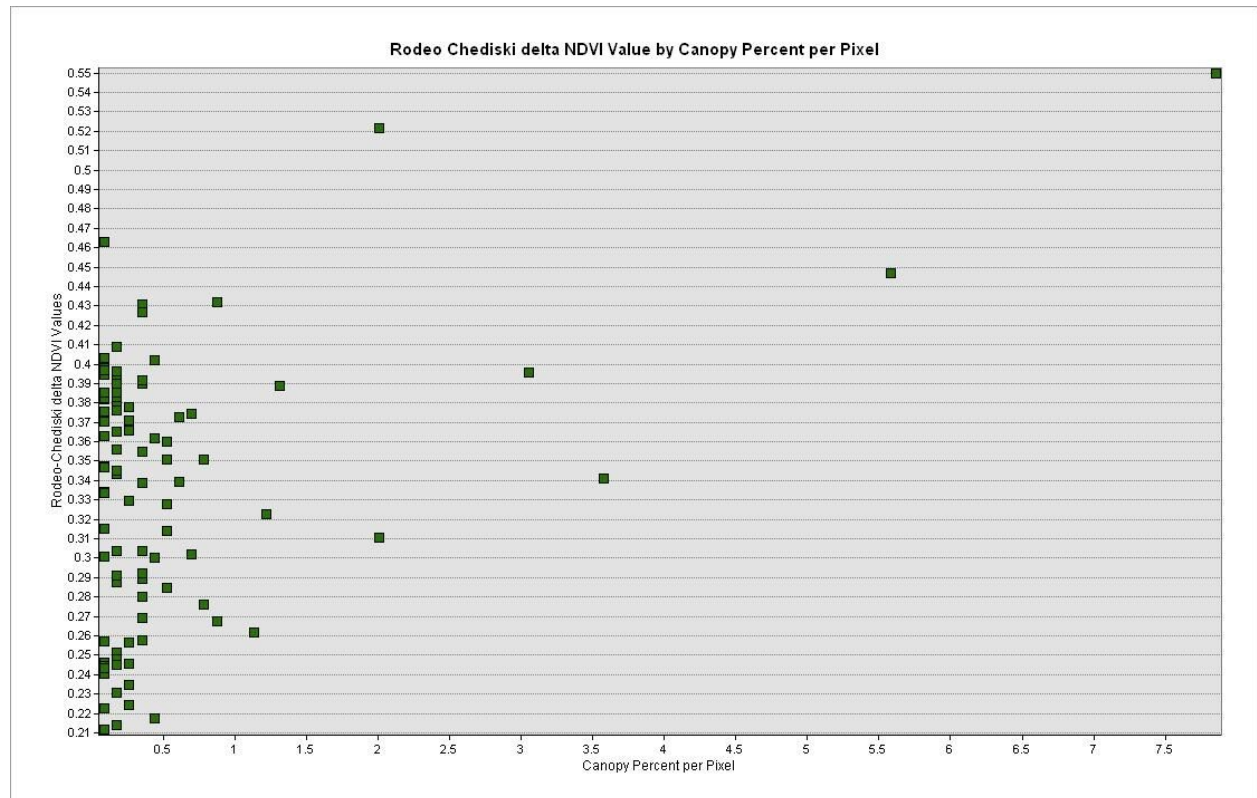
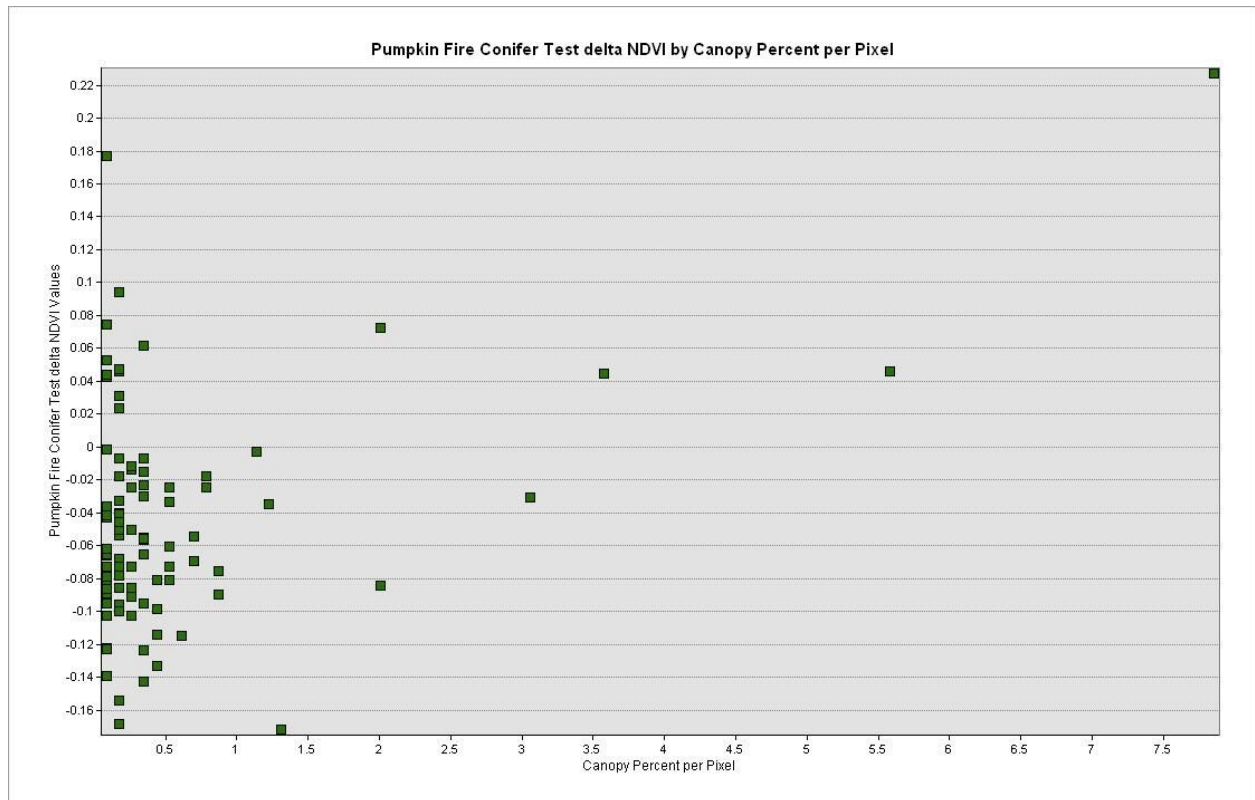


Figure 5: Pumpkin Fire Conifer Test Scatterplot



Pulling Delta NDVI Ranges:

In each of the tests there was some clumping seen. The clump edges were estimated and ranges for each test were defined. Please see table below for each test ranges.

Figure : Ranges and Equations established for each test

Pumpkin Fire	Rodeo Chediski	Conifer Test
Range: .3 - .55 Equation 1: B1 gt .3 Equation 2: B1 lt .55 Equation 3: B1*B2	Range: .26 - .6 Equation 1: B1 gt .26 Equation 2: B1 lt .6 Equation 3: B1*B2	Range: -.15 - .05 Equation 1: B1 gt -.15 Equation 2: B1 lt .05 Equation 3: B1*B2

For each test, to be able to pull these established pixel ranges three equations were entered in the Band Math tool in ENVI 5.3. The three equations are listed above. In this step it is very important to keep track of the equation that was used to create each of the layers. The first two

equations actually create individual layers of all the values above or below the specified value. All of the pixels that fall within the parameters of the first two equations receive a 1 for the pixel value while the other pixels get a 0. The third equation brings together the two equations with multiplication. All of the pixels that fall within the parameters of both equations will have a 1 ($1*1$) while all others have a 0 ($1*0, 0*0, 0*1$). The pixels that have a 1 after the last equation show the areas of regeneration.

Selecting within High Severity Area:

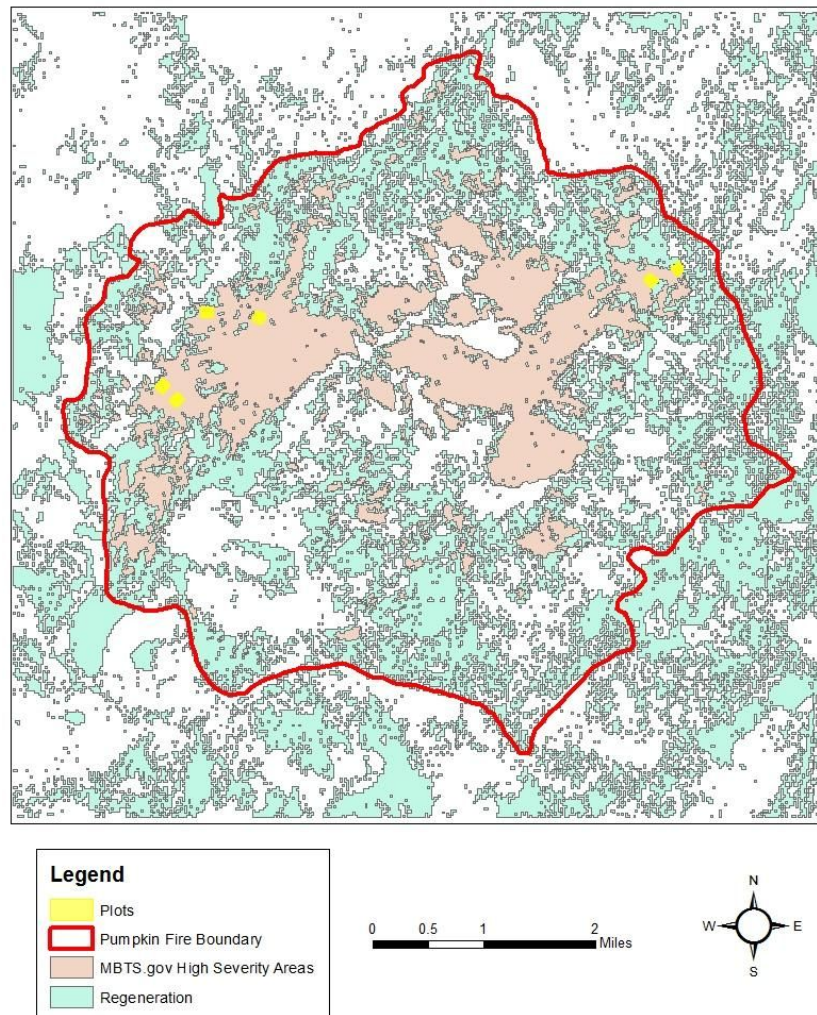
After identifying the regeneration pixels for the entire Landsat scene, identifying the ones that are within high burn severity areas of each fire needs to be done. This range only works for identifying regeneration in high burn severity because it was developed from areas within the high burn areas. It does not include NDVI values from trees that survived the fire only for trees that have returned after the fires. The trees that survived should have a different NDVI value based on what Zheng et al. 2015 found with field crops and being able to see a difference between different germination times of the same species.

The process that was used for calculating the areas of regeneration within the high burn severity areas of each fire began with clipping the raster image of regeneration by the fire boundary. This step significantly increases the processing speed. Then the regeneration raster image was converted to a polygon layer with the Raster to Polygon tool. No simplification was used to keep the pixel edges. The records coded 0 were deleted and the records with 1 were merged together. Then the regeneration layers were stacked underneath the MTBS.gov polygon (see Map 10). The clip tool was used to select the areas that fell within the high burn severity

boundaries. The map below shows how it was stacked to help describe the relationships that were set up for the use of the clip tool.

Map 10: Example of Regeneration Selection

How Regeneration was selected for High Severity Areas



Statistical test for Significance:

There are several different ways in which the ranges can be tested. Field checking can be done to estimate the percentage of correctly identified areas or different statistical testing can be

used. The method that was selected for this project was the Pearson's Correlation Coefficient combined with hypothesis testing for significance and regression line creation for visual inspection. The Pearson's Correlation Coefficient allows for two different attributes to be compared determining the relationship (Wong and Lee 2005). In this case the comparison correlation was done between the number of trees and the delta NDVI value and between the percent canopy cover and the delta NDVI. The factor being calculated is r . The equation for r is:

$$\frac{\sum(x-\bar{x})(y-\bar{y})}{(n-1)S_xS_y}$$

The S_x and S_y are the standard deviation of each attribute (Wong and Lee 2005). This equation was done in Excel by using the tool CORREL. It automatically calculates r removing the human error from the math. The r value provides a number between -1 and 1. The values less than 0 show a negative correlation, close to 0 there is no correlation between the values and close to 1 there is a positive correlation (Wong and Lee 2005). In assessing the significance of this value hypothesis testing was used (Wong and Lee 2005). The following hypothesis test and t statistic were used for each of the three tests in order to determine significance.

$H_0 : r = 0$ (Meaning there is no significant relationship between Canopy percent or Number of trees to delta NDVI)

$H_a : r \neq 0$

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

A two tailed p-value was then determined from a chart for the significance. This chart tool can be found here (<http://oak.ucc.nau.edu/rh83/Statistics/normal/>). With p-value, the lower the value the more significant the relationship. For this analysis 95% confidence was utilized,

meaning the p-value had to be smaller than .05 in order to reject the null hypothesis indicating significance.

The regression line was calculated and created in Excel. The data analysis extension “Data Analysis” was utilized. The charts were created within Excel to show the regression line. Excel calls this line the Trend line in the chart creation tool bar.

Results:

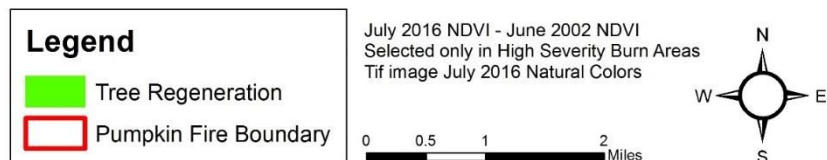
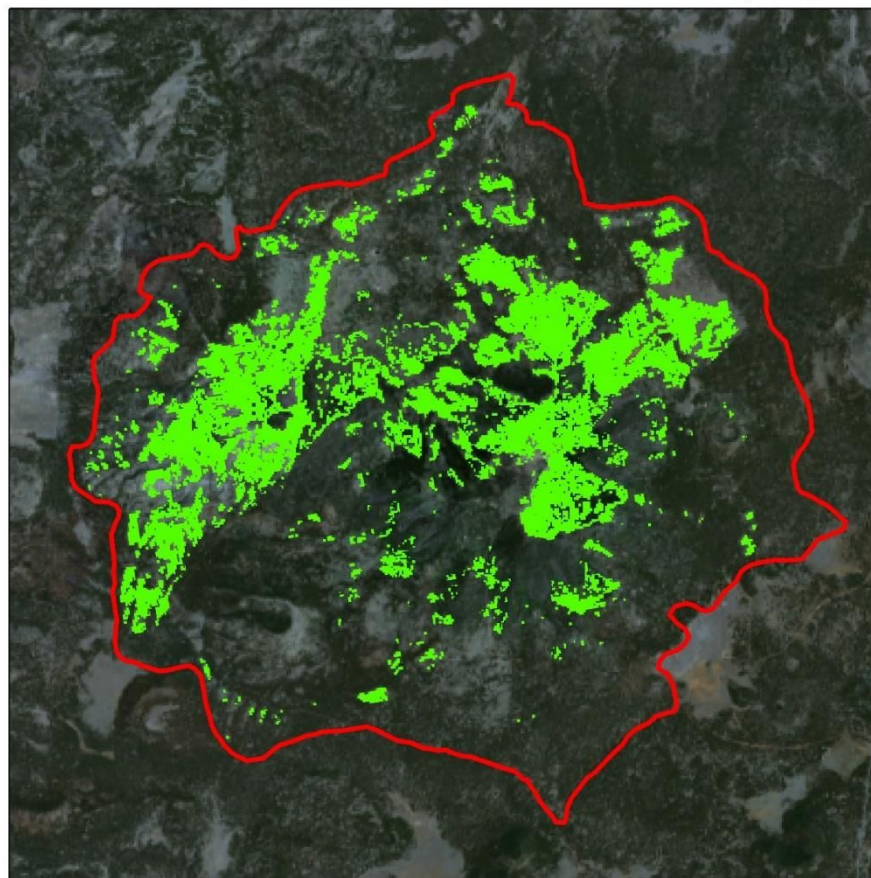
Pumpkin Fire:

The Pumpkin fire had six plots with a total of 1,558 seedlings and trees in 170 grid squares. The species found within the plots were: Gamble oak (*quercus gambelii*), pinon-juniper (*juniperus osteosperma*) and ponderosa pine (*pinus ponderosa*). The range that was identified as regeneration based on the scatterplots was .3 -.55. There were 2659.49 acres (1076.26 ha) identified by this range. Originally there were 3,536 acres identified as areas of high burn severity (MBTS.gov). This is roughly 75% of the high severity areas has some type of greenery regenerating.

Map 11: Pumpkin Fire Areas of Regeneration

Tree Regeneration after the 2000 Pumpkin Fire

By Kimberly Istok

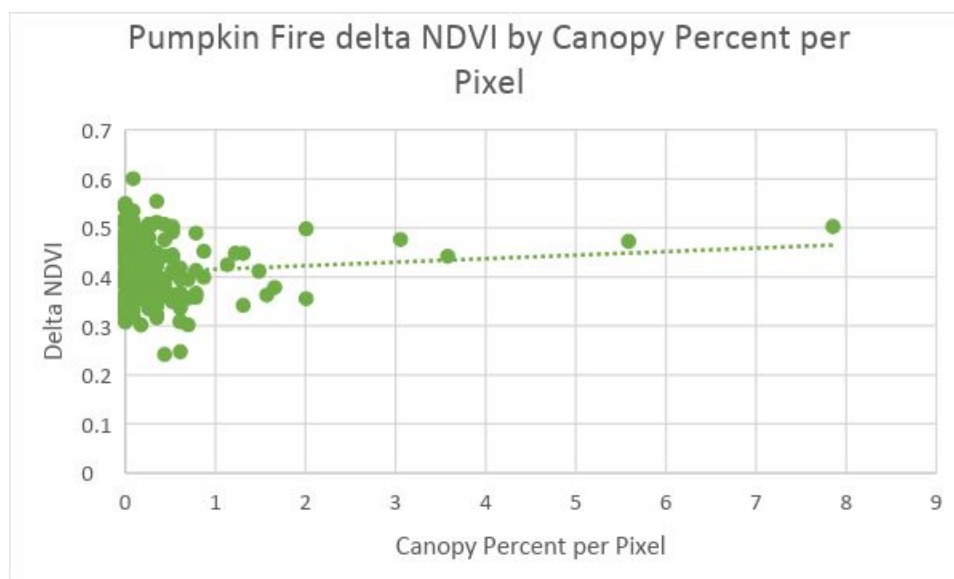


The correlation coefficient between the number of trees per pixel and the delta NDVI value was $r = .09$. The t calculated was 1.17 giving a p -value of .2436 causing an acceptance of the null hypothesis because it is larger than .05, indicating no correlation between the two attributes.

The correlation coefficient between delta NDVI value and the canopy area percent of the recorded trees is $r = .103$. The t calculated was 1.34 giving a p -value of .1821 causing an acceptance of the null hypothesis indicating no correlation between the two attributes. The hypothesis test in this case was not necessary because the r values and the regression line show that there is no correlation.

The r value alone shows that there is little to no relationship between the NDVI value and the presences of trees or percent canopy area in the grid squares. The regression line charted also shows a very straight line supporting the results of the hypothesis test, no correlation. See figure below for regression line angle.

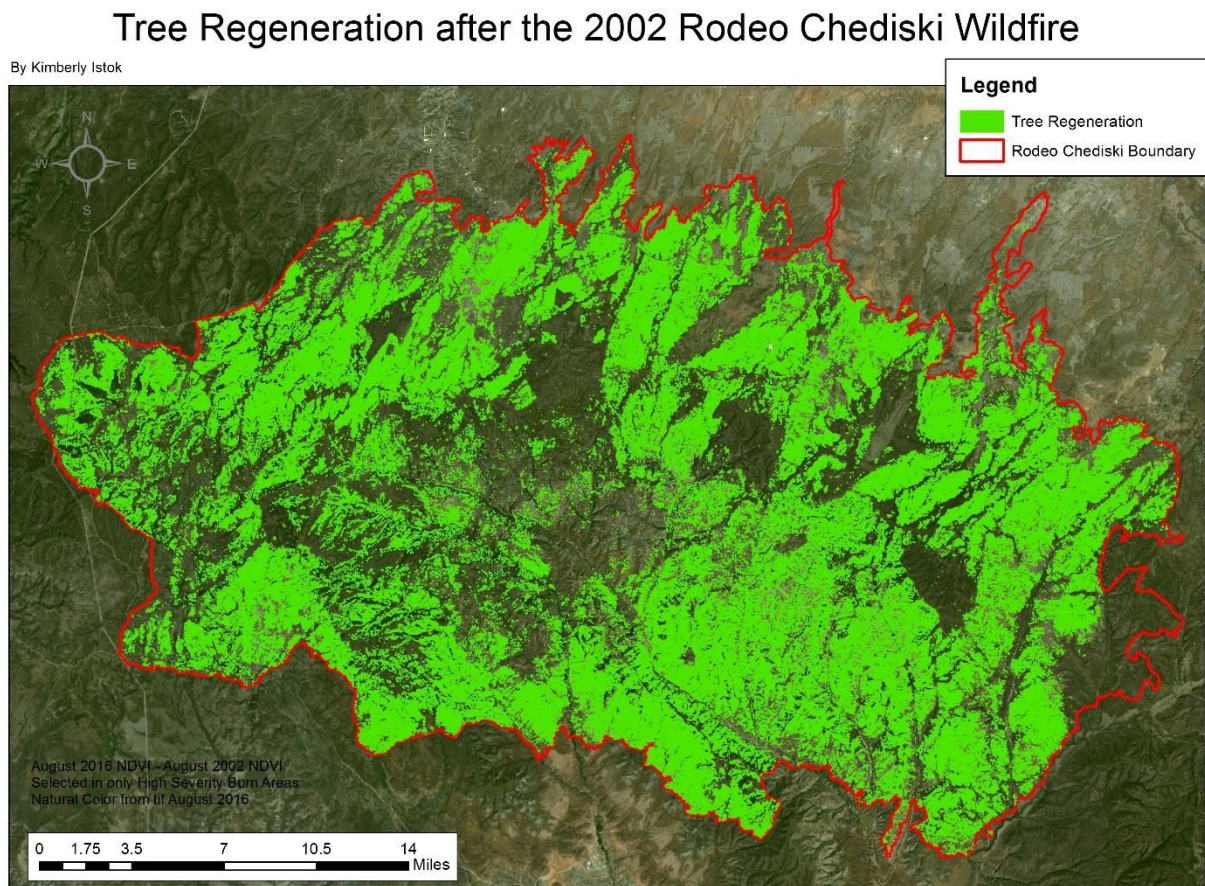
Figure 6: Pumpkin Fire Scatterplot with Regression line



Rodeo-Chediski Fire:

The 6 Rodeo-Chediski plots 171 grids were created. The species found within these grids were Gamble oak (*quercus gambelii*), pinon-juniper (*juniperus osteosperma*) and ponderosa pine (*pinus ponderosa*) seedlings and trees. In addition, Manzanitas (*arctostaphylos pungens*) were also recorded. There was a total 35,188 trees and seedlings measured in the 6 plots. The range that was selected for regeneration was .26 - .6. There were 144,052.3 acres (58,295.88 ha) identified to have some type of greenery within the Rodeo-Chediski high burn severity areas. There were originally 169,042 acres recorded as locations of high burn severity (MBTS.gov). This is 85% of the areas with high burn severity have some type of greenery regenerating.

Map 12: Rodeo-Chediski Fire Areas of Regeneration

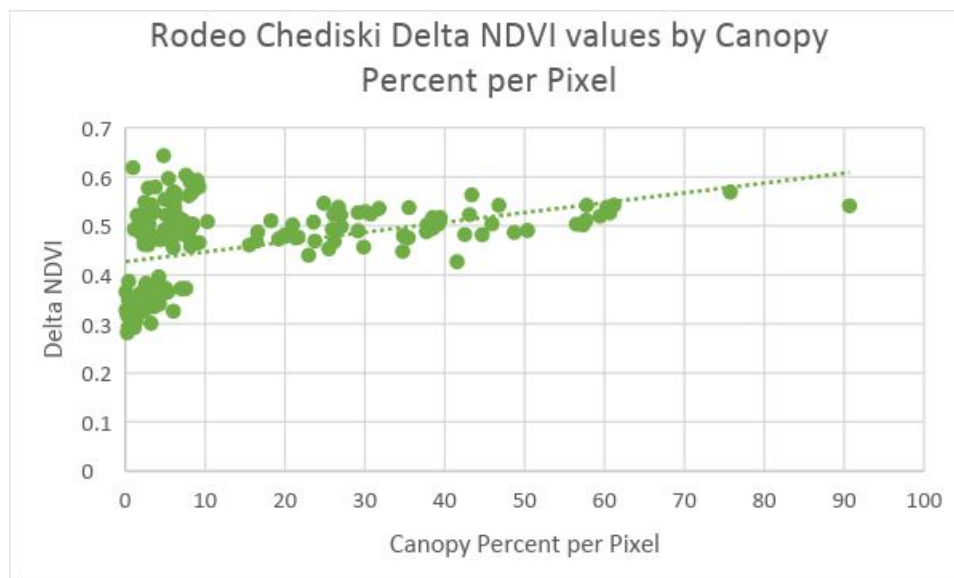


The correlation coefficient between the number of trees per pixel and the NDVI value $r = .33$. The t calculated was 4.54 giving a p -value of 0.000 allowing for safe rejection of the null hypothesis indicating that this value is significant and there is a correlation between the number of trees per pixel and delta NDVI.

The correlation coefficient between the delta NDVI value and the canopy area percent of the recorded trees is $r = .424$. The t calculated 6.08 giving a p -value of 0.000 allowing for a safe rejection of the null hypothesis indicating that this value is significant and there is a correlation between the canopy area percent and delta NDVI. This correlation is higher than what is seen with the Pumpkin fire.

Both r values above are farther from 0 indicating a positive correlation. This positive correlation is also represented in the charting of the regression line, as there is an angle in the line running from southwest to northeast. See figure below for regression line.

Figure 7: Rodeo Chediski Fire Scatterplot with Regression line



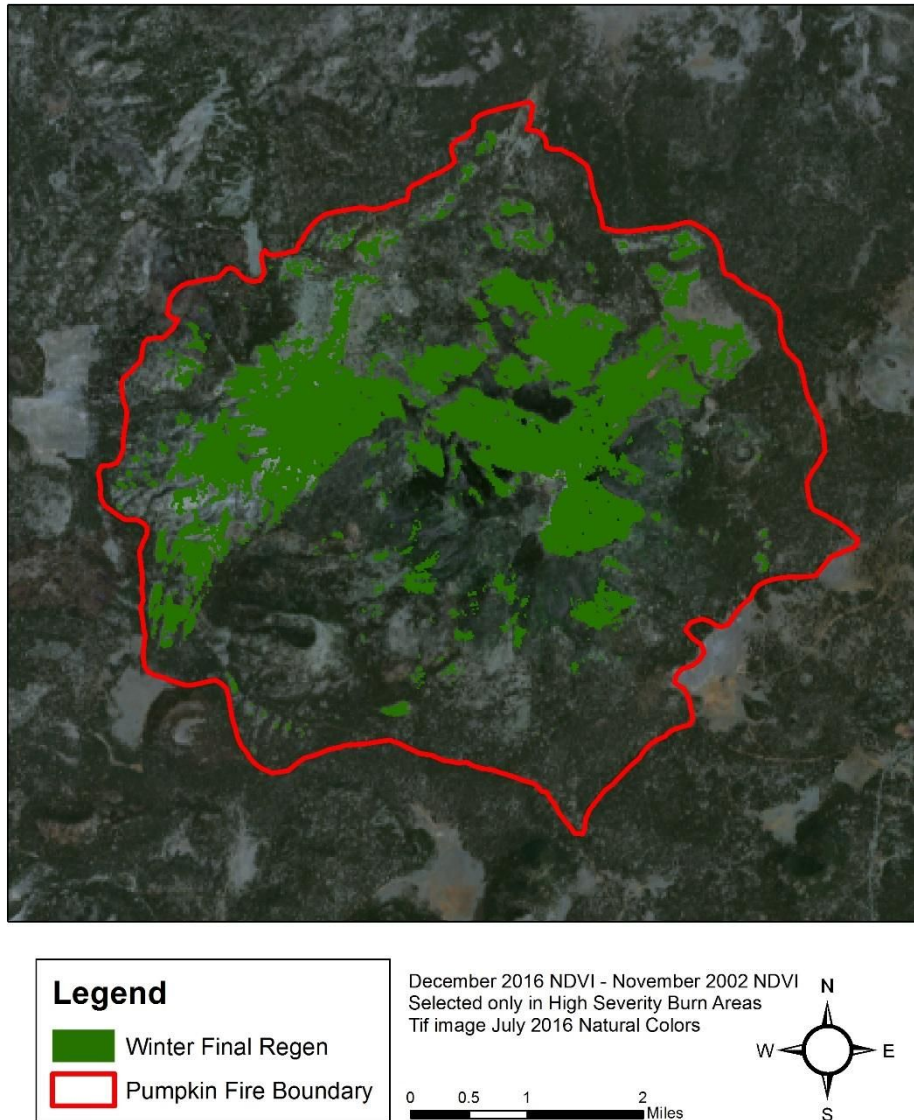
Pumpkin Fire Conifer Test:

For the conifer change detection only pixel grid squares with a 75% majority or above of juniper or ponderosa pines seedlings were utilized. There were a total of 92 pixel grid squares with 535 trees and seedlings counted. The range that was utilized to locate areas with ponderosa pines or junipers was $-.15-.05$. There were 2200 acres (890.31 ha) identifies as locations for conifers.

Map 13: Pumpkin Fire Conifer Test areas of Regeneration

Conifer Tree Regeneration after the 2000 Pumpkin Fire

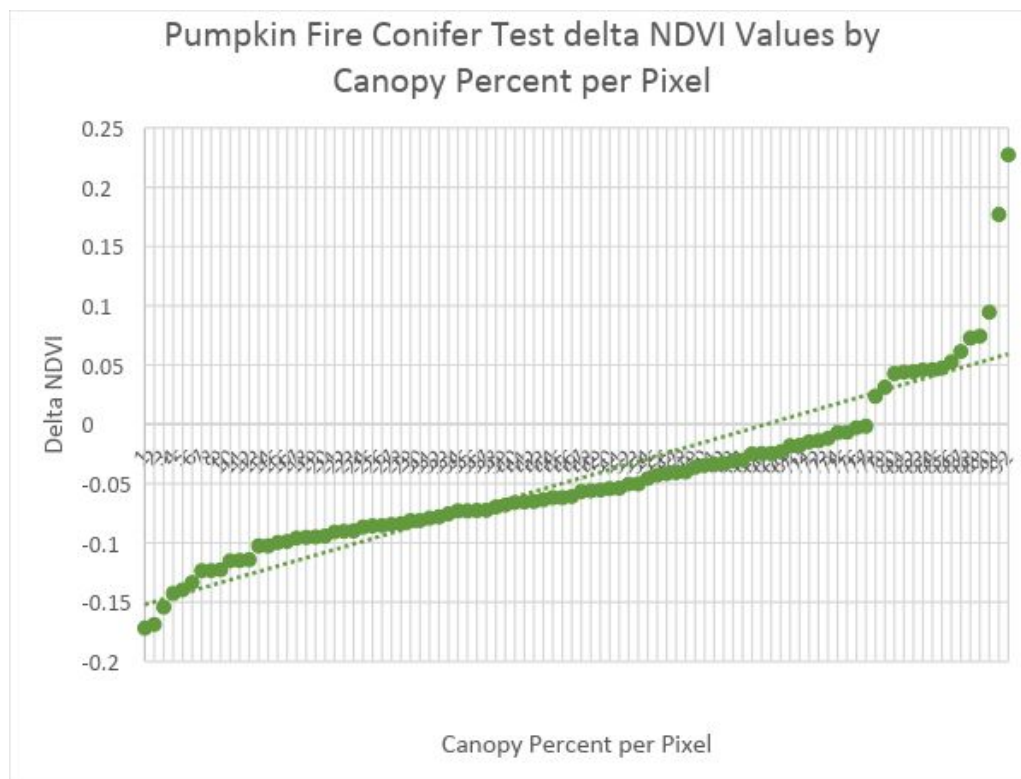
By Kimberly Istok



The correlation coefficient between the number of trees per pixel and the delta NDVI value is $r = .41$. The t calculated was 4.26 giving a p -value of .0001 allowing for the null hypothesis to be rejected indicating that this value is significant and there is a correlation between the number of conifer trees and delta NDVI.

The correlation coefficient between the pixel square delta NDVI value and the area percent of the recorded trees is $r = .417$. The t calculated was 4.35 giving a p -value of 0.000, allowing for the null hypothesis to be safely rejected indicating that this value is significant and there is a correlation between the area percent of recorded trees and the delta NDVI. The regression line also shows the positive correlation. See figure below for regression line.

Figure 8: Pumpkin Fire Conifer Test Scatterplot with Regression line



Discussion:

During the three tests, two of the test were found to have a positive correlation between number of trees per pixel compared to delta NDVI and canopy area percent compared to delta NDVI. The two test that were found to be positive were the Rodeo-Chediski greenery change detection and the Pumpkin Fire Conifer test. The Rodeo-Chediski was a test that performed a change detection between a NDVI values derived from Landsat images taken directly after the

fire and subtracted from NDVI values derived from Landsat images taken 14 years after the fire. The Pumpkin Fire Conifer test was done with NDVI values derived from winter Landsat scene two years after the fire and subtracted from NDVI values derived from a winter Landsat scene 14 years after the fire.

The third test that was performed on the Pumpkin Fire did not have a positive correlation found between the number of trees or percent canopy and delta NDVI. This test consisted of NDVI values derived from a Landsat scene taken 2 years during the summer after the fire and subtracted from NDVI values derived from a Landsat scene taken 14 years during the summer after the fire. There are several reasons why the Pumpkin Fire analysis did not show as strong of a correlation as the Rodeo-Chediski or Pumpkin Fire Conifer test. The thought process behind the 2002 NDVI values being subtracted from 2016 NDVI values was intended to pull out the grass NDVI values leaving the remaining NDVI values just the regenerated trees.

The first major reason for such a low correlation could be the Landsat scene selection for this fire. The earlier image was taken two years after the fire. After a wildfire the first plants to come back are the grasses, but small seedlings could be present as well. Gambel oak and juniper have easier times regenerating after a fire compared to the ponderosa pines, these species could be part of the NDVI value seen in the 2002 Landsat image. By there being seedlings potentially part of the NDVI value in 2002 this would produce a skewed delta NDVI value. Making the delta NDVI only show some of the regenerated trees.

Another reason the Pumpkin Fire did not have a correlation seen could be the number of trees that were actually recorded in the 170 grids. 1,558 seedlings were recorded in 170 grids,

compared to the Rodeo-Chediski fire which had 171 grids with 35,188 seedlings recorded. For this size resolution of imagery (30-m by 30-m) the seedlings might not have been dense enough to make enough of a difference in the NDVI value for there to be a correlation seen.

Overall the r values of the three tests were all lower than expected. If there is a very strong correlation between the two values the r value is .8 or above. However all of the values seen were smaller than .5. The values that were seen in the Rodeo-Chediski Fire and the Pumpkin Fire Conifer test were positive and found to have a significant correlation but were still very low. There are a few factors that might contribute to these low numbers. The first factor could be contributed to the resolution of the imagery. Just like in the Pumpkin Fire test the Landsat imagery is 30-m by 30-m and the seedling are very small. Making it difficult to determine the NDVI tree value only. The NDVI value will include grass, shrubs as well as the trees. These numbers could be lower because of the background noise seen from the grasses and shrubs.

Another factor that could potentially contribute to the low correlation r values is the arbitrary establishment of area per species. The seedlings radius was arbitrarily determined by the author in relationship to a majority of the plants that were visually observed in the field. However a better way to test this would be to utilize the crown radius developed by Sanchez Meador (Tuten et al. 2015). The equations was developed to model canopies, which would work for establishing seedling canopies as well. The equations is $Crown\ Radius = 0.1387(DBH)$ (Tuten et al. 2015). This equation could have been used for just the records marked as tree (height 1.4m or higher) because the DBH was collected. This might provide a more accurate

canopy percentage per pixel for the trees, then the seedlings can be kept at an arbitrary radius because the DBH was not collected for them.

Seasonality of the Landsat scenes could have been a contributing factor to the NDVI values seen before the change detection equation. Because both Pumpkin Fire and the Rodeo-Chediski Fire scenes were taken during the summer June - August months. In the early months of the spring to summer April – June the color of the grasses is different from the July and August months. Grass growing season is during the monsoons and when comparing the April to June time period to July August time period the NDVI values will be different. Potentially by using Landsat scenes from the same growing seasons when comparing might help to minimize the greenery noise from the grasses seen in the final results.

Human error could also potentially be a reason for the low correlation values. Even though all of the values throughout the preprocessing and NDVI tool use, there still could be something that the author missed. The next time this process is produced python should be used to create a tool that does all of the preprocessing without any input from the researcher. This will allow for completely objective results.

Future Research:

There are several ways this research can be built on. First step would be to find or collect data with a smaller resolution. A smaller resolution will allow for the seedlings to be seen easier and potentially narrow down grass values and remove them from the analysis and test these areas again. Then utilizing the higher resolution data try and identify if the regeneration within the moderate severity plots has a different delta NDVI range from the high severity.

Another way this research could be built upon would be to review the NDVI values before the change detection. Maybe a stronger correlation was seen before the change detection equation. Ranges could be pulled from this data and test the same way without performing a change detection.

If higher resolution imagery was collected it would be interesting to test if different band combinations can be determined for each of the regenerated species in question and calculate the regeneration that way. Band combinations can be done with all aerial imagery that has been collected. Band combinations mean changing the order of the bands to make each of the species pop. For example the natural spectrum is created with Band 3,2,1 in that order. Bands 3,2,1 are the red, green, blue bands of an image. To find false color bands are ordered 4, 3, 2 and include the near infrared band. By manipulating the bands a potential combination can be determined to pull out all the areas of regenerated species. Another way of doing this would be to look up the wavelength at which each tree species reflects. Select by each wavelength for each species then combined all the selected areas together to get one layer with all the species areas selected. Then clip by the high severity boundaries to then calculate the amount of regeneration returning after the fire.

Conclusion:

The first objective of this project was to test if ground data can be utilized as training data for Landsat derived NDVI imagery. The second objective was to calculate the tree regeneration after two high severity wildfire. The third objective was to calculate the acreage of the conifers on one of the high severity wildfires. The acreage of the Rodeo-Chediski was calculated and showed a positive correlation between the number of tree and the delta NDVI found, meaning

that ground data can be used for training data. The Pumpkin Fire Conifer test also showed a correlation between the ground data and the delta NDVI determine. However the Pumpkin Fire test did not show any correlation, this could be contributed to the seasonality of the scenes, the resolution of the Landsat imagery as well as the amount of trees actually regenerating. The correlations for both the Rodeo-Chediski and the Pumpkin Fire Conifer test were low but they were both found to be significant.

One of the other goals of this project was to create a process that was from free satellite imagery. High resolution satellite imagery is expensive to acquire and because of this it is not available to everyone. Because this process utilizes the free imagery available through USGS and MTBS.gov, forest management agencies are able to replicate this test to potentially help identify forest areas that are in need of management techniques. These management techniques are costly to the agencies and require hours and hours of coordination to successfully apply them. These techniques are important because of the increase in high severity wildfire in the southwestern United States. Thinning and prescribed burns have shown time and time again in research and in use that these methods help to lessen the effects of wildfires on the landscape. By using Landsat imagery and geospatial science, management agencies are able to increase efficiency, in planning as well as potentially limit the severity of wildfires.

Citations:

Chen, X., Vogelmann, J., Rollins, M. Ohlen, D., Key, C., Yang, L., Huang, C., Shi, H., (2011). Detecting post-fire burn severity and vegetation recovery using multi-temporal remote sensing spectral indices and field-collected composite burn index data in a ponderosa pine forest. *International Journal of Remote Sensing*, 32(23), 7905-7927.

Cocke, A., Fule, P., and Crouse, J., (2005). Comparison of burn severity assessments using Differenced Normalized Burn Ratio and ground data. *International Journal of Wildland Fire*, 14, 189-198.

Delgado, R., Lloret, F., & Pons, X., (2010). Influence of fire severity on plant regeneration by means of remote sensing imagery. *International Journal of Remote Sensing*, 24(8), 1751-1763.

Dickson, B., Prather, J., XU, Y., Hampton, H., Aumack, E., Sisk, T., (2006). Mapping the probability of large fire occurrence in northern Arizona, USA. *Landscape Ecology*, 21, 747-761.

Escuin, S., Navarro, R., Fernandez, P., (2008). Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *International Journal of Remote Sensing*, 29(4), 1053-1073.

Farnsworth, A. (2001). Fighting the Pumpkin Fire - Indirect Attack and Aerial Ignition. *Fire Management Today*, 61(4), 34-34.

- Fernandez, A., Illera, P., Casanova, J. (1997). Automatic Mapping of Surfaces Affected by Forest Fires in Spain Using AVHRR NDVI Composite Image Data. *Remote Sensing Environment*, 60, 153-162.
- Ffolliott, P., Stropki, C., Neary, D., (2008). Historical Wildfire Impacts on Ponderosa Pine Tree Over stories: An Arizona Case Study. Research Paper RMRS-RP-75.
- Flood, N.,(2014). Continuity of Reflectance Data between Landsat-7ETM+ and Landsat-8 OLI, for Both Top-of-Atmosphere and Surface Reflectance: A Study in the Australian Landscape. *Remote Sensing*, 6, 7952-7970.
- Irons, J.R (2017, April 27). Landsat Science, Retrieved May 1, 2017, from <https://landsat.gsfc.nasa.gov/landsat-1/>
- Jenkins, S., Sieg, C., Anderson, D., Kaufman, D., Pearthree, P., (2011). Late Holocene geomorphic record of fire in ponderosa pine and mixed-conifer forests, Kendrick Mountain, northern Arizona, USA. *International Journal of Wildland Fire*, 20, 125-141.
- Kramer, K.,(June 2012). A Burning Issue. Arizona Highways. Pg 40-47.
- Mast, J, N, et al. 1999. Restoration of Presettlement Age Structure of an Arizona Ponderosa Pine Forest. *Ecological Applications*, 9(1), 288-239.
- Roder, A. , Hill, J., Duguy, B., Alloza, J.A, Vallejo, R. (2006). Using long term series of Landsat data to monitor fire events and post-fire dynamics and identify driving factors. A case study in the Ayora region (eastern Spain). *Remote Sensing of Environment*, 112, 259-273.
- Strom, B. A.,(2005). Pre-fire treatment effects and post-fire forest dynamics on the Rodeo-Chediski burn area, Arizona. Northern Arizona University, May 2005.
- Strom, B. A., & Fulé, P. Z. (2007). Pre-wildfire fuel treatments affect long-term ponderosa pine forest dynamics. *International Journal of Wildland Fire*, 16(1), 128. doi:10.1071/wf06051
- Tuten, M. C., Meador, A. S., Fule, P.,Z.,(2015). Ecological restoration and fine-scale forest structure regulation in southwestern ponderosa pine forests. *Forest Ecology and Management*. 348, 57-67.
- Veravebeke, S., Gitas, I., Katagis, T., Polychronaki, A., Somers, B., Goossens, R., (2012). Assessing post-fire vegetation recovery using red-near infrared vegetation indices: Accounting for background and vegetation variability. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68, 28-39.
- Wong, D. W., Lee, J., & Lee, J. (2005). *Statistical Analysis of Geographic Information with ArcView and ArcGIS*. John Wiley & Sons, Inc.
- Wulder, M., Masek, J., Cohen, W., Loveland, T., Woodcock, C.,(2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122, 2-10.

Yi, K., Tani, H., Zhang, J., Guo, M., Wang, X., Zhong, G., (2013). Long-term Satellite Detection of Post-Fire Vegetation Trends in Boreal Forest of China. *Remote Sensing*, 5, 6938-6957.

Zheng, B., Myint, S., Thenkabail, P., Aggarwal, R., (2015). A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. *International Journal of Applied Earth Observation and Geoinformation*, 34, 103-112.

Retrieved May 1, 2017, from

<https://www.harrisgeospatial.com/docs/ExampleMSSensorsFLAASH.html>

Appendix 1: Time Logged

This is the timeline of the preparation and classes completed by the student that lead to this final practicum project.

Summer 2015 – Worked full time for the RMRS collecting tree locations and species.

In Fall 2015 – GSP 529 – Remote Sensing Class: ENVI was taught to the author.

Data cleanup was done for summer 2015 collection, This continued until March 2016. ~80 hours of cleanup work.

The final project for GSP 529 With permission from the RMRS staff, the collected point data were utilized for the final project. This final project allowed for processing of the Landsat imagery, grid building and allowed the author to gain familiarity with ENVI as well as the RMRS tree point data

In Fall 2016 – Background research was started, sources were collected. Proposal for project was developed. Meetings with advisors. GSP 538 – Vector analysis and Statistics was taken.

November 2016 – All Landsat imagery was downloaded.

In Spring 2017: Landsat preprocessing was all completed by 2/15

Square grid was built by 2/20.

All NDVI and Change detection information processed and done by 3/1

Final write up began 3/1

½ Draft turned in 3/24, the other half of the write up to be turned in 3/29

3/30 Pumpkin noise test was completed.

4/4 revised full draft was sent

4/16 revised and full final draft sent

4/28 10 am paper presented

5/8 Final Submission

Appendix 2: Processing Notes

Point Data processing and clean up: 80 hours

Preprocessing for Landsat – Envi 4.8 and 5.3:

For the Landsat 7 data preprocessing was done in ENVI 4.8.

Each band was calibrated and turned into radiance.

Bands 1-5, 7 were stacked

Attributes were edited

Matching the correct wavelength

Band Name

Converted file from BSQ to BIL

Utilized FLAASH, Cell values falling between 0-10,000
For Landsat 8 data preprocessing was done in ENVI 5.3.
MTL bands 1-7 turned into a radiance image by Radiometric Calibration
Utilized FLAASH,

The NDVI tool was used on all 6 Landsat images:

NDVI Equation : $(\text{NIR}-\text{Red})/(\text{NIR}+\text{Red})$

Landsat 7: Band 3 Red and Band 4 Near Infrared.

Landsat 8: Band 4 Red and Band 5 Near Infrared.

Change Detection:

Pumpkin Fire:

B1 = PumpkinSept2000_NDVI

B2 = PumpkinJuly2016_NDVI

Layer stacked: (Pumpkin2000_2016_layerstacked_NDVI)

Band Math tool: $B2-B1 = \text{Change (Pumpkin2000_2016_Change)}$

Test 2: Using a Landsat image that is a few years after the burn, however before trees regeneration, to see if the noise from the grass can be removed from the analysis. However this will not be able to be tested using the Rodeo Chediski because Landsat 7 SLC sensor went down and created streaking over 75% of the scenes from 2004-2007.

B1 = PumpkinJune2002_NDVI (NDVI_test)

B2 = PumpkinJuly2016_NDVI

Layer stacked: Pumpkin2002_2016_Layer

Band Math tool: $B2-B1 = \text{Change (Pumpkin2002_2016_Change)}$

Field added to grid squares layer: PumpkinTest2NDVI

Rodeo-Chediski:

B1 = RCAug2002_NDVI

B2 = RC2016_NDVI

Layer Stack: (RC2002_2016_Layersstack_NDVI)

Band Math tool: $B2-B1 = \text{Change (RC2002_2016_Change)}$

NDVI Value field in grid squares will be manually populated with the two NDVI Change layers.

Scatter plots will be created for range selection, And correlation r^2 was found

Change detection for determining Range for only conifers (PIPO & JUNI)

Pumpkin Fire:

B1 = PumpkinSept2000_NDVI

B2 = Pumpkin2016_Winter_NDVI

Layer Stacked: (Pumpkin_2000_2016_Winter_Layerstack)

Band Math tool: $B2-B1 = \text{Change (WINTERPumpkin_2000_2016_NDVI)}$

Rodeo-Chediski: Removed from the project

B1 = RCAug2002_NDVI

B2 = RC2016_Winter_NDVI

Layer Stacked: (RC_2000_2016_Winter_Layerstack)

Band Math tool: $B2-B1 = \text{Change (WINTERRC2002_2016_NDVI)}$

NDVI values for WinterNDVI was manually entered for grid squares that have a majority percent of PIPO or JUNI 75% or above. PIPO: 89 grid squares, JUNI: 45 grid squares.

Create scatter plot with NDVI values and Area for identifying conifer ranges.
Grids with a majority of PIPO and JUNI and a percent majority that is larger than 75%.
Grid/Reference Layer creation: (Grid square 339 final number)
This layer was created by using Landsat imagery cells as a reference to created grid squares within each plot. These grid squares will include all of the pixels that are completely contained within the plot. Pixel squares were even included if they did not have a tree mapped in them.

Field in the Grid layer:

TotalTreecount; (Short)

QugaCount: (Short)

PipoCount: (Short)

JuniCount: (Short)

ManzCount: (Short)

OtherCount: (Short)

Majority: (Text) (Must remember the ClumpNumber field when calculating this)

PercentMajority: (Short)

NDVIValue (Float)

WinterNDVIValue(Float)

Fire Severity – High severity portions only

Information Layers were downloaded from MTBS.gov for the Pumpkin and RC

Dnbr6.tif was then turned into a polygon

In an edit session all of the polygons with a gridcode of 0 were merged together. This continued for 1-6. This layer created a way in which to isolate the High Severity Areas and run the select NDVI on just that area.

Canopy Cover Area – Created for Validation and Persons Correlation value and regression to show the relationship between the Canopy Area and NDVI value

Creation of layers for each species, QUGA, PIPO, JUNI, MANZ, OTHER

Appended point layers together. Point data originally came from two sources, hand held GPS and Trimble

Appended PumpkinMissingHighSeverity to HighSeverityPointsONLY

Total 15477 points (this does not include the clump counts)

Buffer: PIPO : 1 meter, .5 buffer entered

PIPO_BUFF_1M

Buffer: MANZ : 2 meter, 1 buffer entered

MANZ_BUFF_2m

Buffer: JUNI: 1 meter, .5 buffer entered

JUNI_BUFF_1m

Buffer: QUGA: 2 meter, 1 buffer entered

QUGA_BUFF_2m

Buffer: OTHER: 1 meter, .5 buffer entered

OTHER_BUFF_1m

Union was used to bring all the layers together

TotalCanopyArea

Spatial Join with Sum attribute summarized

GridsquaresWCanopyArea

AreaPercent Field was found by

$(\text{Sum_Shape_Length}/900(30 \text{ m} \times 30 \text{ m}))*100$

Pearson's Correlation Coefficient and regression analysis: CORREL in excel

Pulling ranges based on scatter plots - utilizing Band Math in ENVI:

Pumpkin Fire: Range NDVI Value

Equation 1: $B1 \geq .3$

Equation 2: $B1 \leq .55$

Equation 3: $B1 * B2$

Final Layer name: PumpkinRangeFinal

Rodeo Chediski: Range

Equation 1: $B1 \geq .26$

Equation 2: $B1 \leq .6$

Equation 3: $B1 * B2$

Final Layer name: RCRangeFinal

Conifer/Winter Pumpkin Fire: Range

Equation 1: $B1 \geq -.15$

Equation 2: $B1 \leq .05$

Equation 3: $B1 * B2$

Final Layer name: RCRangeFinal

1 and 0 were given to the values in each of the scenes with this. 1 is the presents of a value that falls within the range and 0 are the values that fall outside of it.

Converted into Tifs to utilize ArcMap

The rasters were converted to polygons without simplification because the pixel edges are needed.

0 and 1's were merged together

All range selections were then clipped by high severity boundaries from MTBS.gov

Final Maps were produced

Acreage and Ha was calculated by Geometry calculator