INSERT TITLE

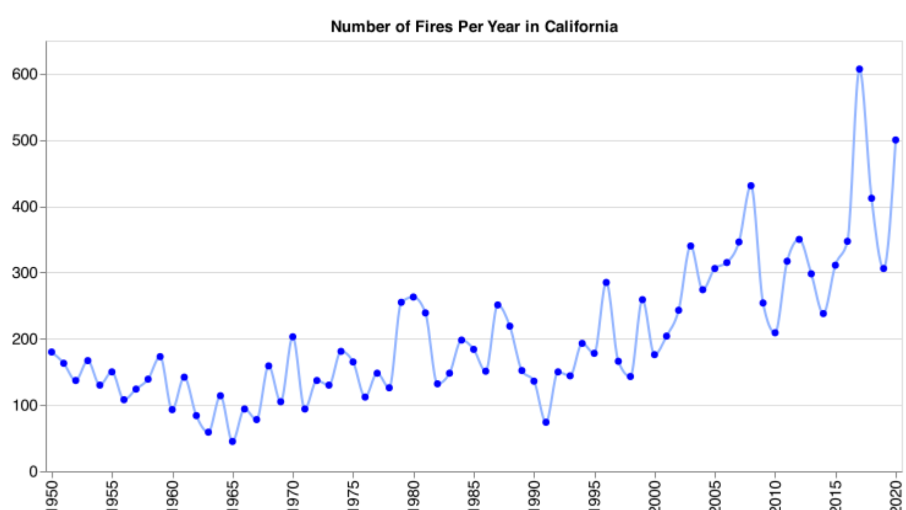
Anthony Chi **Jaskaranpal Singh** **Alice Lu** **Oscar Jimenez**

Abstract (revise)

Fire severity maps are an important tool for understanding fire damage and managing forest recovery. They are mainly produced by several federal fire mapping groups: MTBS, RAVG, and BAER. We have identified several key issues with the completeness, consistency, and efficiency of their mapping methods that reduce the effectiveness of their burn severity maps. In order to address these issues, we propose the use of machine learning in conjunction with remote sensing data as an alternative to traditional methods of producing severity maps, which rely on in-situ data and spectral indices derived from image algebra. Several supervised classifiers will be trained on sample data collected from 17 wildfires across Northern California and their performances will be evaluated to judge their accuracy at mapping fire severity.

1 Introduction

In recent years, wildfires have been growing into a much larger environmental and public safety threat. Fire seasons are larger, more destructive, and burning longer than ever before such that the US Forest Service has coined the term “fire year”. The exact causes for this behavior are not known, but scientists point to climate change, increased human activity from expansion into rural areas, and over-zealous fire prevention policies that have created environments ripe for wildfires with large buildups of combustible fire fuels. [1] This phenomenon is happening across the world, but is especially apparent in Northern California, which has historically been a global hotspot for wildfires. For example the 2018 fire season, the worst in California’s history, caused over $140 billion in damages and killed over 100 people. Our preliminary analysis also confirms that wildfires are more frequent and destructive. From 1990-2020 there has been a 267% increase in the number of fires annually and in the past 20 years there has been a 520% increase in the number of fire seasons that exceed 500,000 burned acres relative to the 50 years prior.



1.1 Burn Severity Maps

There is some ambiguity around the meaning and usage of fire intensity, fire severity, and burn severity. Fire intensity is strictly used to describe the total amount of energy released by a fire and not the effect on the ecosystem. Fire severity and burn severity describe the effect of fire on aboveground and belowground biomass. This includes measures like canopy cover, crown volume, surface litter, and soil hydrophobicity. These terms are often used interchangeably, but a minor distinction is that in certain applications burn severity can specifically refer to fire effects on soil. [2]
Our analysis uses remote sensing data to specifically focus on the effect of wildfire on above and belowground biomass, which we refer to interchangeably as fire or burn severity. This definition is widely used by federal fire mapping groups and past research on wildfires.

Burn severity maps are widely used by federal agencies and forest managers to map fire damage and extent and help manage forest recovery efforts, update vegetation and land cover maps, and monitor ecosystem health. Fire mapping responsibilities are shared by several federal interagency groups, mainly Monitoring Trends in Burn Severity (MTBS), Rapid Assessment of Vegetation Condition after Wildfire (RAVG), and Burned Area Emergency Response (BAER).

Traditional methods of producing these maps are expensive and time-consuming since they require teams of surveyors and ecologists to gather in-situ data. For many fires, this is infeasible due to harsh weather and inaccessible terrain. These methods are still used for certain fires, but have been largely phased out with the introduction of Earth observing satellites that provide remotely sensed data. Fires can be mapped at a much faster and larger scale at a fraction of the cost relative to field surveys, while still maintaining high accuracy. Remote sensing data is widely used in many other applications, such as agriculture, climate change, and natural disasters, since they cover long time spans and are continuously updated with high resolution, multi-spectral data

The most widespread spectral index for identifying burned areas and fire severity levels is the Normalized Burn Ratio, $NBR = \frac{(NIR-SWIR)}{(NIR+SWIR)}$ and with Landsat this is $NBR = \frac{(Band\ 5 - Band\ 7)}{(Band\ 5 + Band\ 7)}$.
The near-infrared band (*NIR*) is sensitive to chlorophyll present in live vegetation, while the short-wave infrared band (*SWIR*) is primarily sensitive to water content in soil and vegetation. It has also been shown to be capable of discerning dead wood from burned soil, ash, and charred wood. As a result, NBR is sensitive to absolute changes in live, photosynthetically active vegetation, moisture content, and certain post-fire surface conditions. [3] Computing NBR for select pre, post fire images and differencing the result provides a measure of absolute change caused by fire (dNBR). However in regions that are less vegetated (grassland, shrubs) or have heterogeneous land cover types, dNBR can underestimate burn severity. [3]

Federal fire-mapping groups mainly use this image differencing method with dNBR, but with slight differences based on their organizational needs.

1.2 Federal Fire Mapping Groups

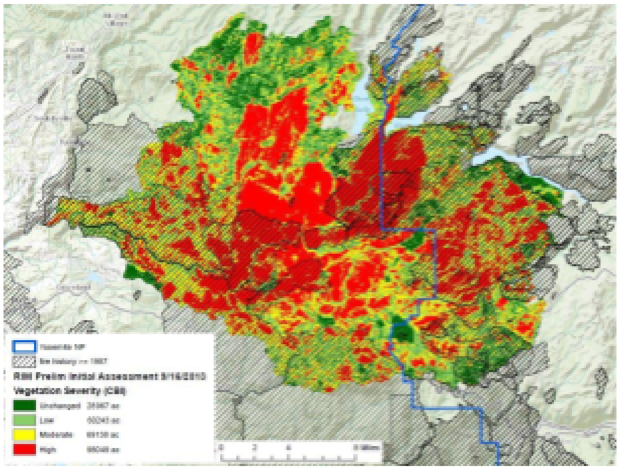
The main groups responsible for mapping fires in the US are MTBS, RAVG, and BAER. MTBS is the largest and most active federal mapping group and in California it maps fires larger than 5000 acres. RAVG maps fires that occur on at least 1000 acres of National Forest System (NFS) land and produces results usually within 60 days of fire containment. It specifically focuses on changes in canopy cover and basal area. BAER is slightly different since its main goal is to assess soil burn severity and identify and prescribe treatments for any hazards caused by fire, like water runoff from hydrophobic soil. Within a week of a fire's containment it provides satellite imagery and preliminary burn severity data to field teams, made up of ecologists, soil scientists, and engineers, that work in the field to stabilize a region. [7]

Several issues currently hinder the effectiveness of burn severity maps produced by these groups.

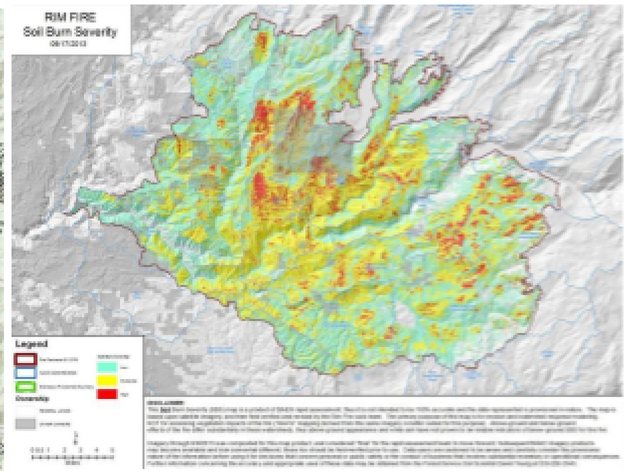
The first is that the completeness of wildfire mapping methods is insufficient. Due to various agency requirements, lack of resources, and the immense number of wildfires every year, federal agencies are only able to map a fraction of wildfires. This leads to lacking fire documentation and coverage, which could limit the work of groups that depend on fire severity maps. In the ten year span from the 2010-2019 fire seasons RAVG mapped 142 fires, MTBS mapped 424 fires, and BAER mapped 174 fires. In addition there is a lack of "completeness" in the data used to produce fire severity maps. Only two spectral bands, *NIR* and *SWIR*, are used from Landsat to calculate dNBR and contextual data like land cover or weather is not used. This additional data could contain relevant information that can be uncovered with machine learning.

The second issue relates to the consistency of severity maps. Maps produced by MTBS rely on analysts to subjectively determine dNBR thresholds to produce severity classifications. In addition, these thresholds are not validated with field data or ecologically quantified so the consistency of their maps is questionable. [7]
Another source of inconsistency is the use of different pre and post fire images since these agencies mostly operate independently and on different timelines. Ideally, the selected pre and post fire images are as close to a fire as possible because using images that are further apart can influence results. For example, selecting a post fire image from a later date allows vegetation regrowth from fire or seasonal changes in vegetation to occur. Or if a fire occurs in November but a pre-fire image from spring is used, this can increase a fire's dNBR value since the absolute decrease in vegetation is greater. [8] For these reasons, agencies often come up with conflicting results. The Rim Fire burned 250,000 acres near Yosemite and was one of the most destructive wildfires in state

history in 2013. The fire severity map produced by RAVG showed that 88% of the area was burned, while the soil burn severity map by BAER showed 56% of the land was unchanged or had a low severity burn.



RAVG Burn Severity Map of Rim Fire (2013)



BAER Burn Severity Map of Rim Fire (2013)

The third issue, which only affects MTBS, is the speed at which severity maps are produced. They release maps on a two year lag and as of today still have not released any for fires from the 2020 and 2021 fire seasons. This delay is likely due to the large number of fires they are responsible for and the amount of human influence required.

1.3 Related Research (need better title)

Fire severity is a well researched topic and common approaches revolve around in-situ sampling and spectral indices derived from image algebra, similar to other change detection applications.[9] In recent years machine learning applications for remote sensing have been growing in popularity, but are still fairly limited. This has been attributed to the limited support of machine learning methods in traditional remote sensing software, confusion on how to apply ML models, contradictory model performance, and parametric methods still being extremely popular even though they have been shown to perform worse overall. [4][10] Machine learning models perform well at modeling complex relationships between features and benefits from large, high-dimensional datasets, which are common with remote sensing datasets.

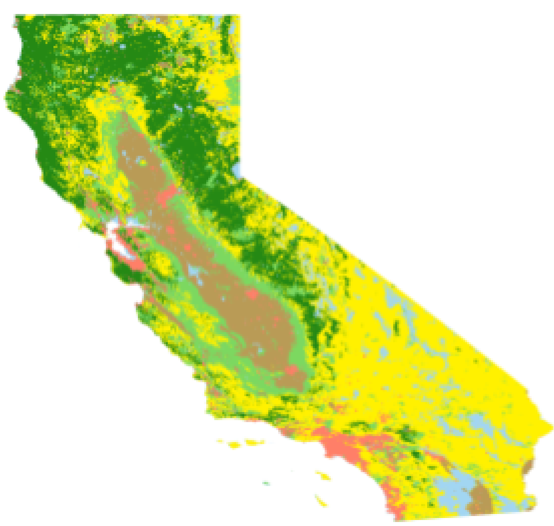
Given these benefits, we propose the use of machine learning methods with remote sensing data to map fire severity. This would address the issues present with current mapping methods by reducing the influence of human subjectivity, allow maps to be produced at a faster and larger scale, and allow us to incorporate additional Landsat 8 spectral bands along with contextual data including weather, land cover, and terrain. In the process we hope to gain insight on what contextual features are indicative of fire risk and severity. We plan to train and compare the performance of several supervised classifiers across 17 different fires in Northern California and our results should be applicable to other wildfires in the general region

A study similar in scope to ours was conducted by training a random forest classifier (RF) over in-situ and remote sensing data from 8 fires in Victoria, Australia. They found that RF generally produced more accurate burn severity results than a traditional spectral indexing approach. [5]

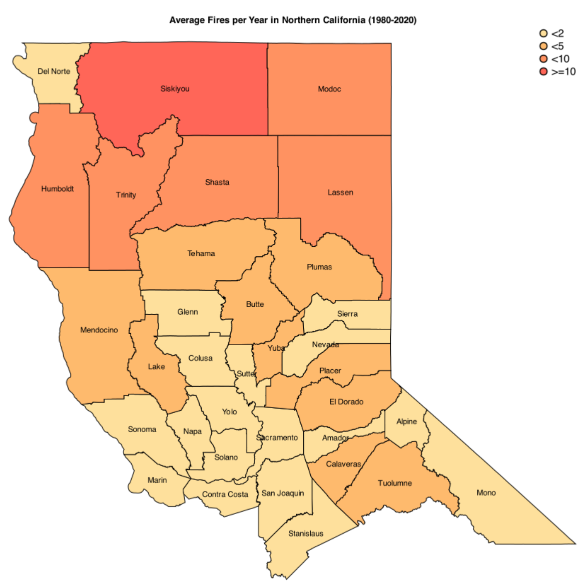
1.4 Region of Interest

Our study will be focused on Northern California since it is a global hotspot for wildfires that has affected a majority of students at UCSD. In addition, its infamous wildfires are well documented by CALFIRE, have been researched significantly in the past, and there are many remote sensing datasets that cover this region.

A majority of counties in Northern California are very rural, have sparse populations, and are mostly undeveloped. Their land covers are largely dominated by forests, low-lying shrubland, annual grassland, and mixed chaparral vegetation. These counties account for a majority of wildfires and related damages. Counties located in Central California near Sacramento, like Yolo, Sutter, and San Joaquin, are more developed and revolve around agriculture and livestock. On average these counties experience less than 2 wildfires per year, usually under 1000 acres.



California Land Cover (2016)



Northern California is historically prone to wildfire since it doesn't experience much rainfall and has dry, hot summers that lead to large accumulations of combustible fire fuels in the fall. Environmental factors, like strong downslope winds and lightning strikes, and human activity are common wildfire ignition sources. [1] California is especially susceptible to long droughts and often experiences consecutive dry years, which are characteristic of regions with Mediterranean type climates. As the effects of climate change become more apparent, droughts and wildfires in California will be a greater environmental and public safety hazard.

2 Data

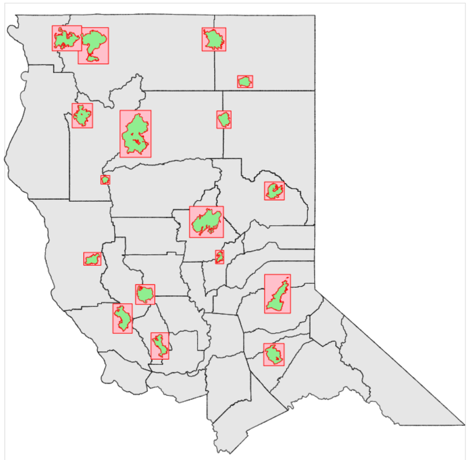
Google Earth Engine (GEE) is a cloud-based distributed computing environment that greatly reduces the technical barriers to entry for large scale geospatial analysis and hosts a large catalog of data including satellite imagery, climate forecasts, and geophysical data.[6] We used the GEE platform to access and run computations on remote sensing data from Landsat 8, NASA SRTM, NLCD 2016, and GRIDMET.

Data	Provider	Bands
Landsat 8 (Level 2, Collection 2, Tier 1)	USGS	7
NASA SRTM Digital Elevation	NASA / USGS / JPL-Caltech	1
NLCD: USGS National Land Cover Database (2016)	USGS	14
GRIDMET: University of Idaho Gridded Surface Meteorological Dataset	University of California: Merced	16

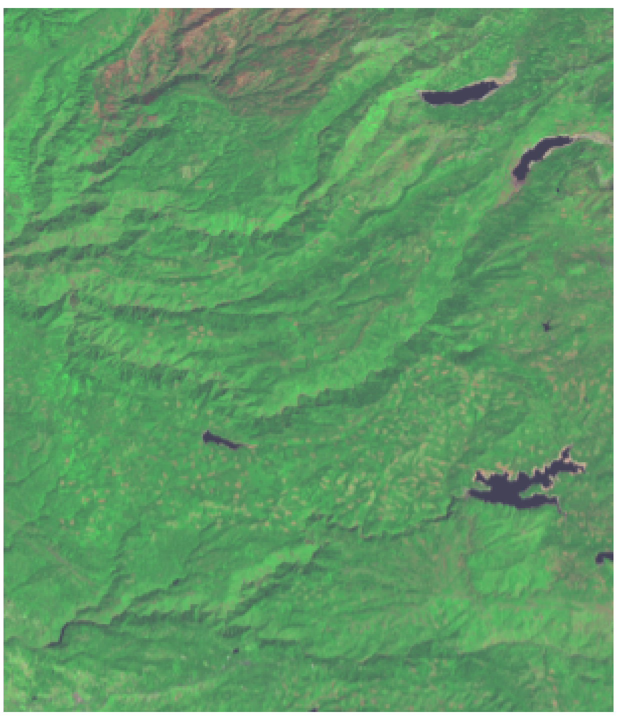
Data on California wildfire seasons from 1950-2020 is provided by CALFIRE and includes information on a fire's location, geometry, size, and duration.

2.1 Fire and Image Selection

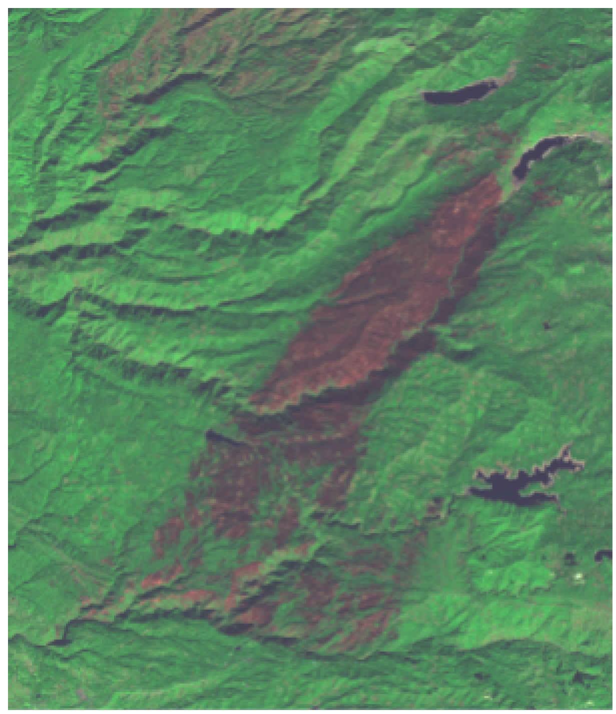
In total, 17 fires were selected from a candidate set of 79 fires. The fires occurred across Northern California between 2013-2020 because this coincides with the launch of Landsat 8 (February 2013) and the California wildfire dataset hasn't been updated to include any fires past the 2020 fire season. All selected fires are at least 10,000 acres in size because fires of this size are better documented and have more pixels to sample.
To get optimal pre and post fire images from Landsat 8, we considered all images that occurred 60 days before and after a fire. Images were selected based on their proximity to a fire's start or end and the presence of environmental factors that reduce image quality, like clouds, smoke, and snow. A majority of pre-fire images are within 14 days of a fire's ignition, but some post-fire images occur much later due to poor image quality.



Fire	Start	End	Acres	Size Code	Pre-Fire Image Date	Post-Fire Image Date
ATLAS	2017-10-08	2017-11-01	51624	I	2017-10-04	2017-11-05
BALD	2014-07-30	2014-08-15	39752	H	2014-07-24	2014-09-10
BUCK	2017-09-12	2017-10-30	13357	H	2017-08-24	2017-12-14
BUTTE	2015-09-10	2015-09-28	70846	I	2015-09-06	2015-10-24
CALDWELL	2020-07-22	2020-09-01	81224	I	2020-07-15	2020-09-01
CAMP	2018-11-08	2018-11-26	153335	J	2018-10-07	2018-12-26
CARR	2018-07-23	2018-09-01	229651	J	2018-07-10	2018-10-14
CASCADE	2017-10-08	2017-10-26	16140	H	2017-10-04	2017-11-05
COVE	2017-07-25	2017-09-06	30890	H	2017-07-16	2017-10-04
FRYING PAN	2014-08-11	2014-09-30	133177	J	2014-07-15	2014-10-03
HAPPY	2015-07-30	2015-08-08	68095	I	2015-07-25	2015-09-20
KINCADE	2019-10-23	2019-11-10	77762	I	2019-10-01	2019-11-18
KING	2014-09-13	2014-10-10	97684	I	2014-09-03	2014-10-21
OAK	2017-08-11	2017-10-10	91125	I	2017-07-14	2017-10-18
REDWOOD VALLEY	2017-10-08	2017-10-25	36522	H	2017-09-25	2017-10-27
ROCKY	2015-07-29	2015-08-14	69438	I	2015-07-27	2015-08-19
WALKER	2019-09-04	2020-01-15	54614	I	2019-09-01	2020-02-24



King Fire (9/3/14)



King Fire (10/21/14)

2.2 Data Extraction

In addition to surface reflectance data from Landsat 8, we also used land cover, elevation, and weather data from NLCD, NASA SRTM, and gridMET respectively. These images are clipped over each fire's bounding box and their bands are merged into a single image in GEE. The selected Landsat 8 images are pre-orthorectified to account for terrain and we used the standard image differencing method to calculate dNBR.

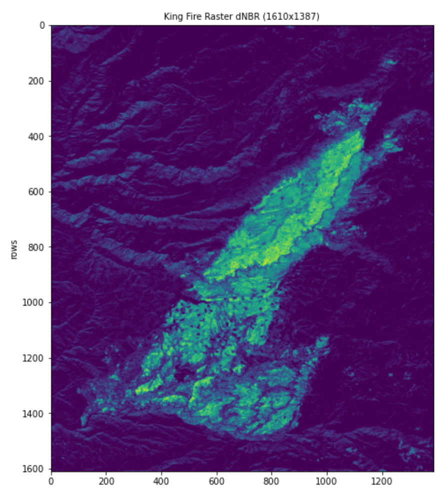
Using proposed burn severity values from the USGS, which we simplified from seven classes to five, we classified each pixel as either vegetation growth, unburned, low severity burn, moderate severity burn, and high severity burn. [2] ~~Images of land cover from NLCD, elevation from NASA SRTM, and weather from GRIDMET are also clipped to each fire's buffered bounding box and merged with the Landsat 8 image to produce a single image for each fire.~~

Rasters for each fire are extracted from GEE and points are sampled from every fire by projecting a uniformly buffered grid over a fire's geometry and randomly sampling a point within every grid box. This data is used to train and validate our supervised classifiers.

2.3 Data Cleaning (maybe blank for now)

~~2.4 Raster Manipulations (rewrite)~~

~~The raster data format was used to facilitate data consolidation and determine data granularity. Since much of the data was found using GEE, which natively supports raster exports, features were more easily added across datasets. The output raster has a spatial cell size of 30 x 30 meters which was consistent amongst most datasets. Polygon areas were used to target fire bounds directly. This was particularly important since most datasets spanned the entire country or even the globe. Even with these reductions, There were still over 400,000 data points for each fire which caused large file sizes.~~



~~To alleviate this issue, each of the fires analyzed receives their own raster which scales with fire size. Figure K shows the representation of the King Fire raster representation on the dNBR feature, the x and y axis represent the data's columns and rows respectively. While cell sizes are consistent among rasters, the size of the rasters themselves are not uniform. In practice this means some fires have more data points than others. However, by scaling each raster with fire size, we can keep a similar proportion of burned and~~

~~unburned cells across fires. Since the fires chosen were all very large to begin with, we accepted the tradeoff that came with scaling on fire size.~~

3 Methods (update to new models)

3.1 Feature Selection and Engineering
(write about sampling strategy)

To limit the amount of data being fed to our machine learning models, we decided to limit the number of features to the most relevant ones. We only kept the features with a **(specify score)** score of 0.02 or higher. All features were standardized to prevent one feature's variance from dominating the other features in the dataset.

Based on our baseline models, we determined that NDVI, elevation, Band 3 (green) surface reflectance, Band 6 (shortwave infrared 1) surface reflectance, and tree cover percentage are the most important features to our model.

The tabular format of the data allows us to preserve the geospatial awareness for each point such that each point considers its neighbor's features as well as their own. We used a point's connected components, which are made up of the 8 closest points, because they are a good measure of the immediate surrounding area while not creating too many features. Edge cases where points are missing neighbors were removed to make sure our data was dense. Even though some observations are lost, these edge points only make up a small fraction of the points in the spatial buffer we use for each fire.

3.2 Model X

3.2 Baseline Models (probably add into each model's section)
Our target feature, burn severity, is ordinally encoded into five classes that are: vegetation growth, unburned, low severity burn, moderate severity burn, and high severity meaning that we have a classification problem. As a result, our selected baseline models are a Logistic Regression model, MLP Classifier and a Random Forest classifier. Of the three baseline models, we found that the Random Forest Classifier provided the most accurate predictions. The Random Forest Classifier consistently has a R^2 score of about 0.85 without any parameter tuning. This performance is expected since previous studies have shown that Random Forests perform well without any parameter tuning and do not overfit to training data as the number of trees and leaves increase. [11] [12]

3.N app section maybe

~~3.N Model #N + Tuning~~

4 Results (revise)

(rough notes based on app)

- Larger fires
- Fires that don't occur in winter months
- Shrubland/Grassland can lead to mixed performance
- Very good at identifying unburned pixels / fire scar
- Not that good at identifying high severity
- Unburned - low severity is kinda bad
- Image quality is very important
- Seasonal vegetation loss adds a lot of noise
- Accuracy is shitty metric (Grant Fire example)
- Models (logistic+mlp+trees) perform about the same
-

To evaluate how generalizable our models are to other Northern California wildfires, we benchmarked their performances on fires of varying sizes, times of year, and land cover types. These factors can affect fire behavior and produce inaccurate or skewed model results. We selected __ fires that represent these factors to demonstrate the strength and weaknesses of our models.

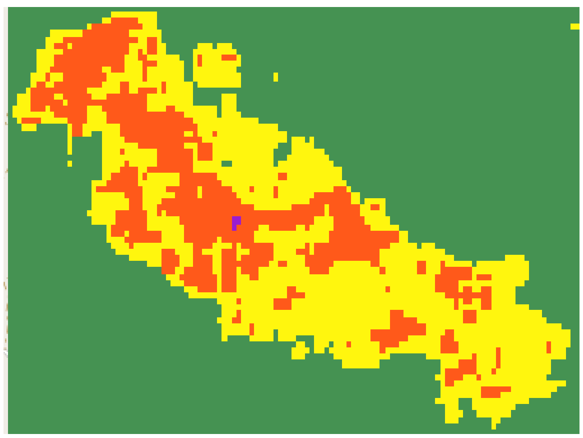
Fire Name	Start Date	End Date	Acres	County	Land Cover
Atlas	2017-10-08	2017-11-01	51624	Napa	Mixed
Badger	2020-07-18	2020-07-19	557	Siskiyou	Shrub
Day	2014-07-30	2014-08-13	13150	Modoc	
Hennessey	2020-08-17	2020-09-16	305351	Napa	Shrub
~~North Complex~~	~~2020-08-17~~	~~2020-12-03~~	~~318776~~	~~Plumas~~	~~Forest~~
Steele	2017-07-26	2017-08-13	45704	Modoc	Shrub
Wallow	2017-08-11	2017-11-13	63785	Siskiyou	Forest

We tested our models on four different fires: Badger (E), Day (H), Hennessey (J+), and Natchez (H), ~~Gold~~, …, ….. Overall, all our models had trouble predicting vegetation growth and high severity burned areas. Our models tend to overestimate the burn severity of each area, particularly they have trouble with predicting Moderate severity areas on actually high severity areas. This is to be expected since we have limited data on high severity fires. Unburned and Moderate areas were predicted most accurately since we have significantly more training data on these areas. While our models have a moderate accuracy rate, our models predict the overall shape of the fire really well.

4.2 Different Sized Fires

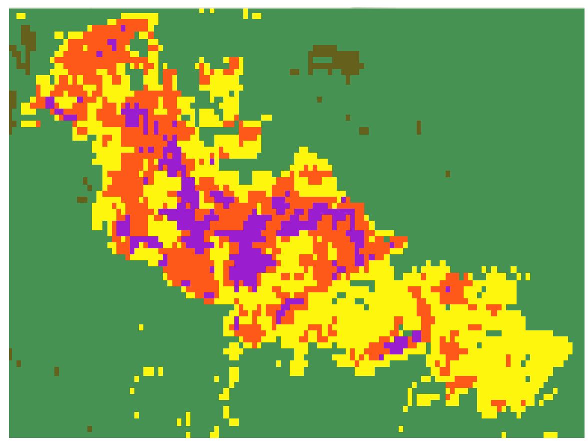
Badger Fire

Actual Fire



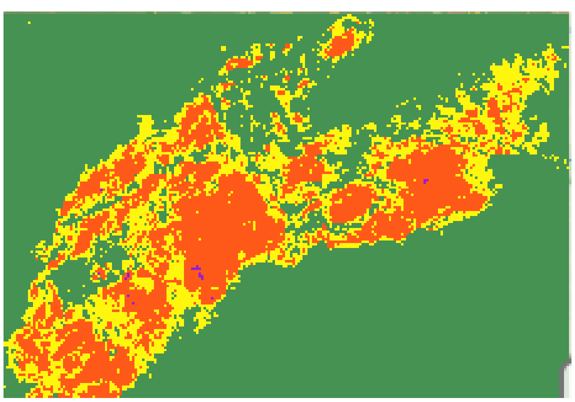
Accuracy: 81.02% (Log)

Predicted Fire



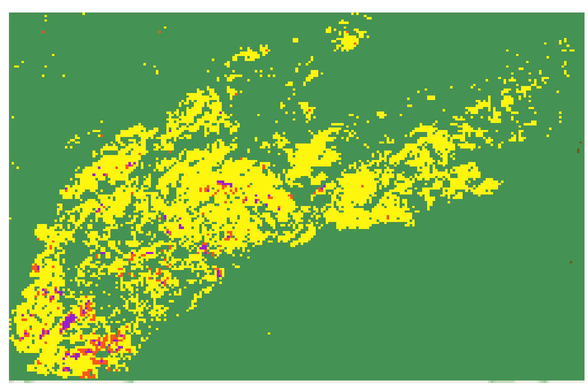
Day Fire

Actual Fire



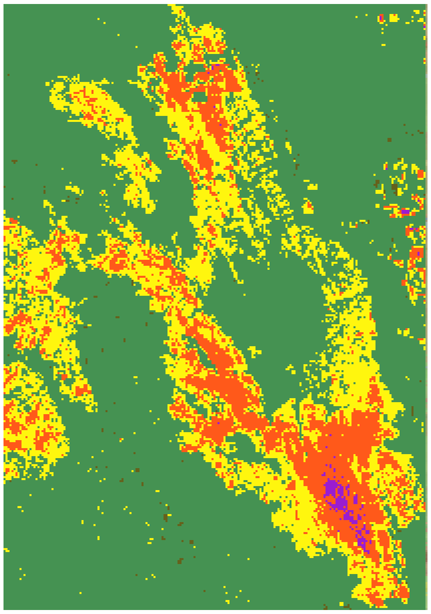
Accuracy: 72.75%

Predicted Fire



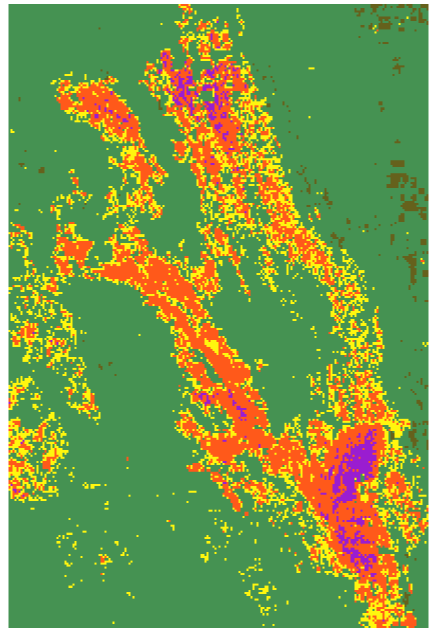
Hennessey Fire

Actual Fire



Accuracy: 78.7%

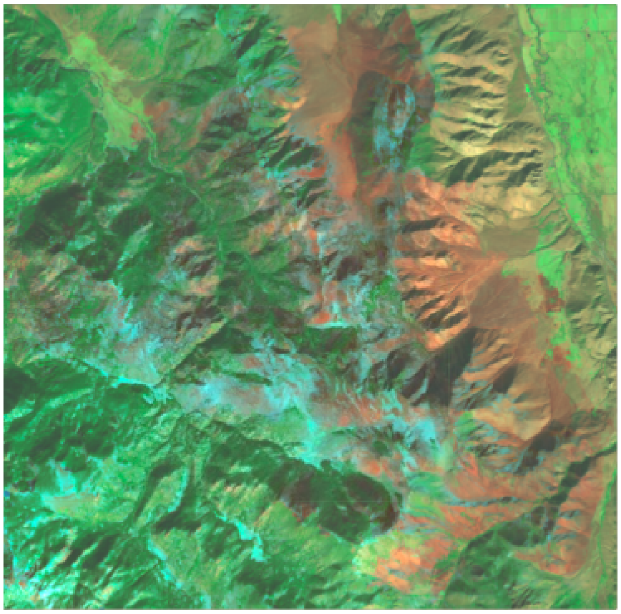
Predicted Fire



4.3 Season

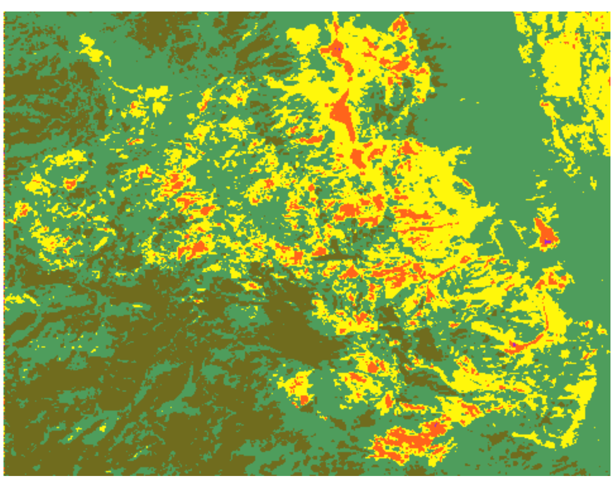
We found that a major roadblock to producing burn severity maps with either method for fires in winter months is that seasonal changes in vegetation and the presence of snow can produce inaccurate or misleading results.

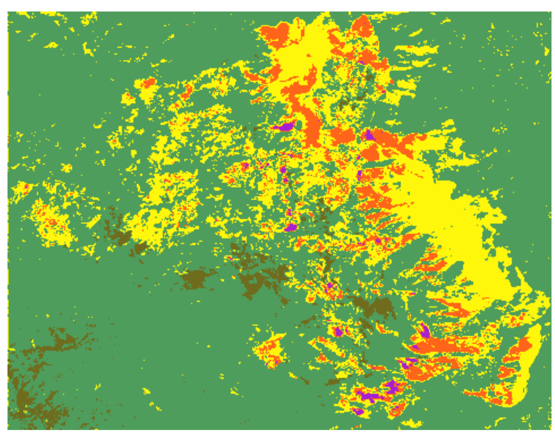
~~For example the North Complex Fire (2020) is one of the largest fires in California's history and was finally contained on December 3, 2020~~ after burning for nearly three months. Since it burned so late into the winter, the candidate set of post-fire images used to produce a burn severity map are strongly affected by snow and seasonal vegetation loss. In the false-color image below, the fire scar is fairly visible and there is some snow in the central part of the fire and in the surrounding region.



An issue with both severity maps is that there are many pixels classified as vegetation growth, mostly near regions with snow. Another is that pixels with seasonal vegetation loss, in the bottom right and top center of the images, are misclassified as low severity burns in both images. These issues contribute a lot of noise to the burn severity maps and can make it more difficult to interpret a fire's burn severity. As demonstrated with the MLP classifier, our models are not as sensitive to these issues and produce a more concise burn severity map. Since none of our training fires contain snow, this could be a direction for further model tuning and improvement.

(A) (B, mlp)

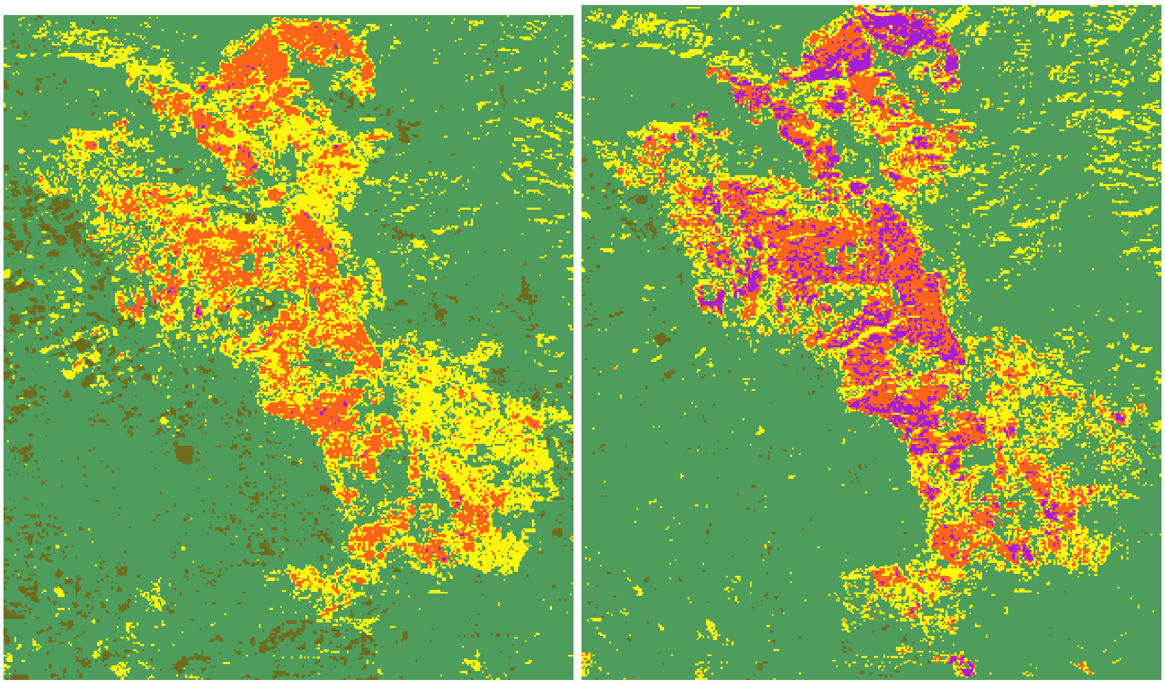




4.4 Land Cover

Our model performed well on fires of mixed land covers that occur most frequently in Napa and Sonoma counties. These fires can be difficult to produce burn severity maps for because discrete severity thresholds may not accurately represent how fires behave in different land covers. We tested our models on the Atlas Fire (2017) as it is adjacent to Napa, CA and has a very mixed land cover composition with lots of agricultural and urban areas.

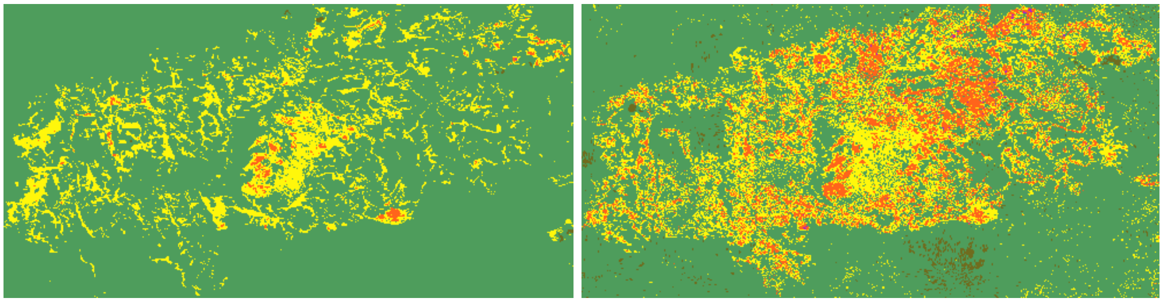
(mlp)



Using our MLP classifier we are able to produce a map that accurately identifies burned and unburned regions and shows the shape of the Atlas Fire. A key difference is that the map produced by the MLP classifier shows the Atlas Fire as having a significantly more severe burn compared to the linear threshold method. This is likely a more accurate assessment of the Atlas Fire as dNBR thresholding is known to underestimate burn severities in shrub and grassland. [3] A strength of using our models in environments with mixed land covers is that they are robust to changes in agricultural regions from crop sowing and harvesting. The linear threshold picks up on these changes and classifies many pixels on unburned farmland as having a low severity burn or vegetation growth, which adds a lot more noise.

We also tested our models on the Steele Fire (2017) which occurs in Modoc County, a part of California that is mostly covered by shrubland. As mentioned with the Atlas Fire, dNBR thresholding struggles to produce accurate burn severity classifications for fires in shrubland and other less vegetated regions. A quick solution around this issue that is employed by MTBS is to have analysts subjectively determine severity thresholds. However this approach introduces a lot of human influence, produces inconsistent results, and is time-consuming.

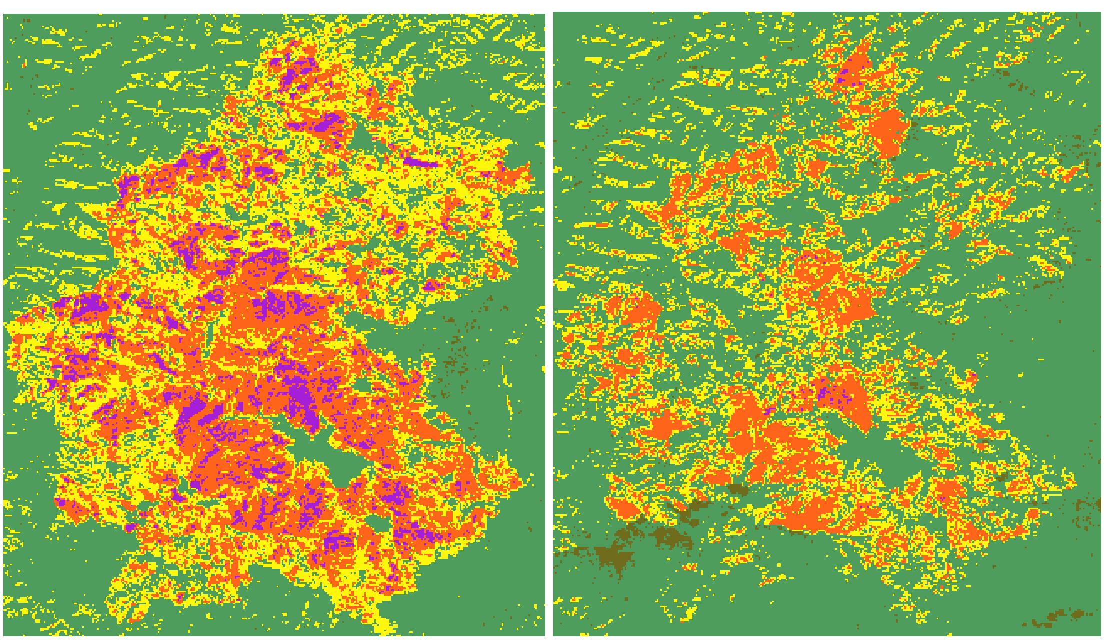
(rf)



The burn severity map produced by thresholding demonstrates this issue as a large majority of pixels are classified as unburned. In contrast, the map produced with a Random Forest classifier shows that it is much more capable of identifying burned areas and depicting the fire's outline.

In general our models all perform similarly and are able to identify fire scars very well, but struggle to identify pixels with vegetation growth and high severity burns. This is likely due to pixels with these classes not occurring frequently in our training data.

A weakness of our models is that they underestimate burn severities by misidentifying pixels with low severity burns as being unburned. This leads to predicted burn severity maps that are very sparse and discontinuous.
(carr fire)



5 Analysis / Further work

(section on dNBR vs rdNBR maybe)
(l8 vs sentinel 2)

5.3 Environmental Features
(weather, elevation, landcover)

6 Conclusion

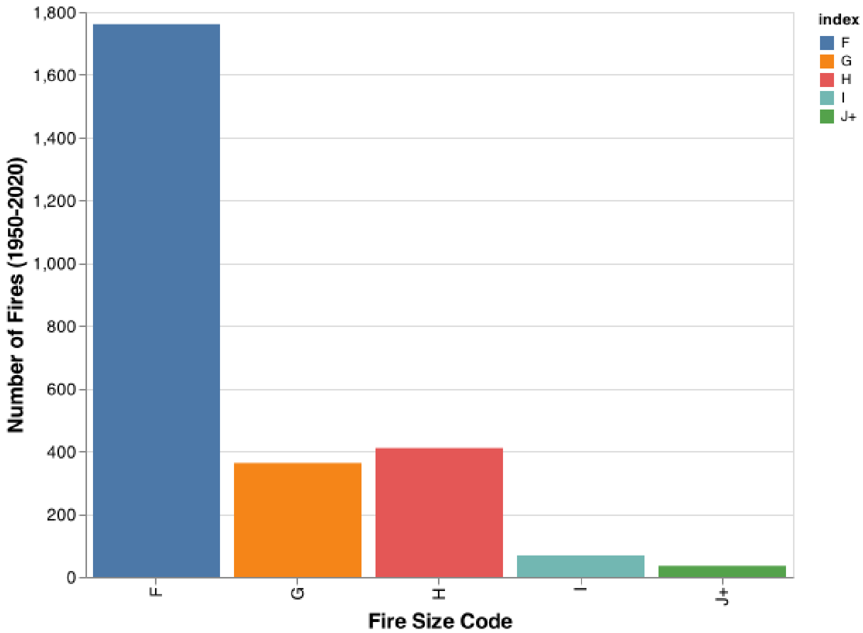
References (format citations)

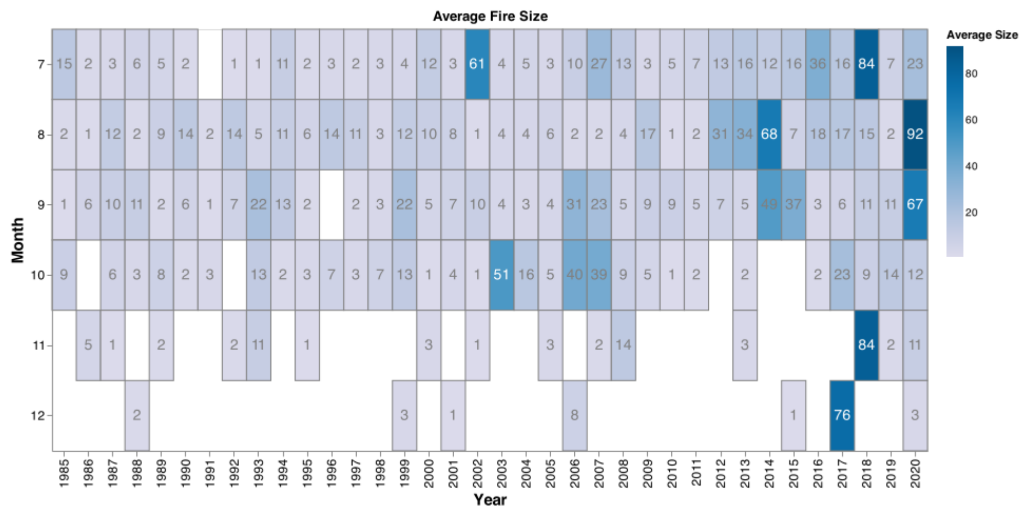
1. **Observed Impacts of Anthropogenic Climate Change on Wildfire in California (2019)**
2. **Fire intensity, fire severity and burn severity: a brief review and suggested usage (2009)**
3. **Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR) (2006)**
4. **Implementation of machine-learning classification in remote sensing: an applied review (2017)**
5. **Fire-severity classification across temperate Australian forests: random forests versus spectral index thresholding (2019)**
6. **Google Earth Engine: Planetary-scale geospatial analysis for everyone (2017)**
7. **Limitations and utilisation of Monitoring Trends in Burn Severity products for assessing wildfire severity in the USA (2015)**
8. **Digital change detection techniques using remotely-sensed data (1989)**
9. **Classifying and Mapping Wildfire Severity (2005)**
10. **Meta-discoveries from a synthesis of satellite-based land-cover mapping research (2014)**
11. **Random forest classifier for remote sensing classification (2005)**
12. **Random forest in remote sensing: A review of applications and future directions (2016)**

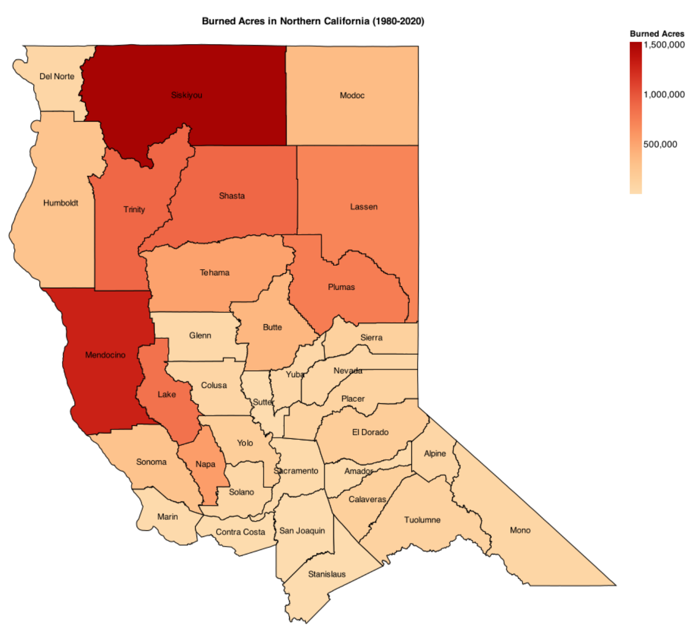
Appendix Figures

Fire	Agriculture (%)	Developed (%)	Forest (%)	Grassland (%)	Other (%)	Shrub (%)
ATLAS	10.6	12.0	26.1	17.5	2.8	30.9
BALD	0.4	0.7	34.4	36.9	1.0	26.6
BUCK	0.0	5.1	65.0	6.6	0.0	23.3
BUTTE	0.0	2.7	39.9	35.3	0.2	21.8
CALDWELL	8.7	1.4	18.7	11.5	2.3	57.4
CAMP	5.5	4.7	46.4	13.8	2.0	27.5
CARR	0.0	7.3	53.8	4.1	4.8	30.0
CASCADE	0.0	3.2	51.6	25.1	1.0	19.0
COVE	4.6	0.4	50.0	14.7	0.3	30.0
FRYING PAN	0.3	5.4	49.5	26.2	0.9	17.8
HAPPY	0.0	2.7	51.3	27.6	0.1	18.3
KINCADE	9.3	7.4	31.2	23.4	1.5	27.2
KING	0.0	1.9	61.0	23.9	1.5	11.7
OAK	0.1	3.7	71.1	3.0	0.3	21.8
REDWOOD VALLEY	6.4	5.3	41.2	9.7	0.4	37.0
ROCKY	0.0	2.4	3.0	50.6	0.5	43.5
WALKER	0.1	1.1	55.6	1.0	1.0	41.3

	County	Agriculture (%)	Developed (%)	Forest (%)	Grassland (%)	Other (%)	Shrub (%)
0	Alpine	0.1	0.6	40.8	3.5	2.3	52.6
1	Amador	1.3	3.4	42.5	20.9	1.7	30.2
2	Butte	26.2	5.3	37.0	13.5	3.9	14.1
3	Calaveras	0.2	3.1	42.5	26.9	1.7	25.5
4	Colusa	46.3	3.1	3.3	22.3	3.5	21.4
5	Contra Costa	10.6	31.8	13.7	20.3	12.9	10.7
6	Del Norte	1.1	4.7	61.5	3.3	19.6	9.8
7	El Dorado	0.0	4.7	61.0	9.9	5.7	18.7
8	Glenn	33.3	3.8	13.4	24.7	3.2	21.5
9	Humboldt	1.6	5.3	63.4	6.4	13.5	9.8
10	Lake	0.9	6.1	27.6	17.6	6.0	41.9
11	Lassen	3.1	1.0	27.8	15.6	4.6	47.9
12	Marin	0.7	8.7	19.6	9.6	39.8	21.6
13	Mendocino	1.1	4.6	55.9	5.3	10.2	22.9
14	Modoc	5.3	0.9	25.1	14.4	8.6	45.8
15	Mono	0.8	0.8	18.3	5.6	9.6	64.9
16	Napa	6.6	6.3	25.7	14.8	7.1	39.5
17	Nevada	0.0	5.2	68.8	4.2	1.7	20.1
18	Placer	5.2	8.8	51.4	12.8	6.6	15.1
19	Plumas	0.8	1.4	65.1	4.5	4.0	24.2
20	Sacramento	31.9	31.7	0.1	27.6	6.0	2.7
21	San Joaquin	65.2	13.0	0.6	16.1	3.3	1.8
22	Shasta	2.3	3.6	54.4	11.7	2.6	25.4
23	Sierra	0.2	1.5	64.9	3.0	2.8	27.6
24	Siskiyou	5.6	3.4	55.2	10.3	2.9	22.7
25	Solano	32.6	12.2	2.5	23.8	21.7	7.2
26	Sonoma	6.2	9.2	35.7	17.3	12.6	19.0
27	Stanislaus	43.5	8.6	3.9	30.4	2.5	11.1
28	Sutter	79.0	7.1	0.4	5.7	4.2	3.7
29	Tehama	5.1	3.0	23.6	26.6	1.8	39.8
30	Trinity	0.0	4.5	62.3	8.4	1.4	23.4
31	Tuolumne	0.0	1.4	46.2	16.0	4.2	32.1
32	Yolo	58.0	6.3	2.1	15.4	2.5	15.6
33	Yuba	27.7	5.4	33.9	18.0	4.4	10.7







(maybe remove)
We decided to use surface reflectance data from Landsat 8 over alternative options, such as Sentinel-2, because it is used by federal fire mapping agencies, has the same spatial resolution as our other data, and has been operating longer, especially if previous Landsat satellites are considered.

(maybe)
Wildfire Burn Severity Prediction using Machine Learning

Similar study aimed at predicting burn severity classes using random forest classification, but on a much smaller scale. This study looked only at 3 specific fires in British Columbia from 2017. The variables used in the model for this study also differed from ours, with these researchers including variables such as “fuel types” and “bark beetle infestation”.