

# Geospatial analysis of the 2020 LNU Lightning Complex fires

**Alex Marden**  
**12/6/2020**

The LNU Lightning Complex fires started on August 17, 2020 and burned until October 2, 2020. These fires—sparked by a series of lightning strikes— burned 363,220 acres, destroyed over 1,000 structures, and caused 6 deaths (Cal Fire, 2020). The fires occurred within a context of increasing regional and global fire activity that has been linked to climate change and land use (e.g. Bowman et al., 2009). Fires events like the LNU Lightning Complex fires are predicted to become a new normal, so it is imperative to identify factors that influence fire intensity to inform future fire management. The identification of contributing factors is inherently spatial because fire burns in heterogenous patterns across landscapes that is affected by fuel, topography, and human infrastructure (Krasnow et al., 2017). Here I measured fire intensity across Napa County (one of the counties affected by the LNU fires), quantified the relationship between fire intensity and land cover, and measured the influence of topographical and fuel variables across space.

## **Research questions:**

- 1) How did fire severity vary over space across Napa County?  
Hypothesis: Fire severity was likely highest in areas with continuous fuel loads and a low amount of human infrastructure. This is likely to occur in the mountainous zones of the county.
- 2) What factors contributed to high fire severity?  
Hypothesis: High fuel loads and steep terrain are likely to contribute significantly to fire severity. These factors likely change over space depending on context (e.g. different vegetation communities)

## **Methods:**

Part 1—Data Acquisition

Part 2—Vegetation type and fire intensity using zonal statistics

Part 3—OLS Regression and Geographically Weighted Regression to identify important factors affecting fire intensity

## Part 1—Data acquisition

**Step 1.** NBR: Normalized Burn Ratio:  $NBR = \frac{NIR-SWIR}{NIR+SWIR}$

The Normalized Burn Ratio (NBR) provides a numerical measurement of estimated fire severity. The formula utilizes near-infrared (NIR) and shortwave-infrared (SWIR) wavelengths. It is designed to be used with [Landsat](#) imagery (Cocke et al., 2005).

Unburned vegetation has very high NIR reflectance and low SWIR reflectance whereas burned vegetation has low NIR reflectance and high SWIR reflectance. A high NBR denotes healthy unburned vegetation while a low value indicates a lack of vegetation. dNBR ( $prefireNBR - postfireNBR$ ) tracks change from pre fire to post fire conditions so a pixel with a high dNBR value indicates a change from healthy vegetation to burnt vegetation.

Most of the data were processed and acquired using Google Earth Engine (GEE)— a geospatial processing service equipped with Google’s cloud computation (Gorelick et al., 2017). GEE runs on JavaScript, here is the annotated code:

This portion of code defines the image collection from which to get data—In this case Landsat 8 (Line 1). Here I am interested in imagery from when the fires were burning, so the date range reflects that of recorded fire activity (Line 2). Line 3 defines the area of the analysis—in this case Napa Valley (table). Lines 5-7 define the bands (5 & 6) to use for the NBR calculation. Lines 10-15 contain the formula for Normalized Burn Ratio. Lines 18-20 apply the formula to the image collection. Line 23 selects the median NBR value from the collection.

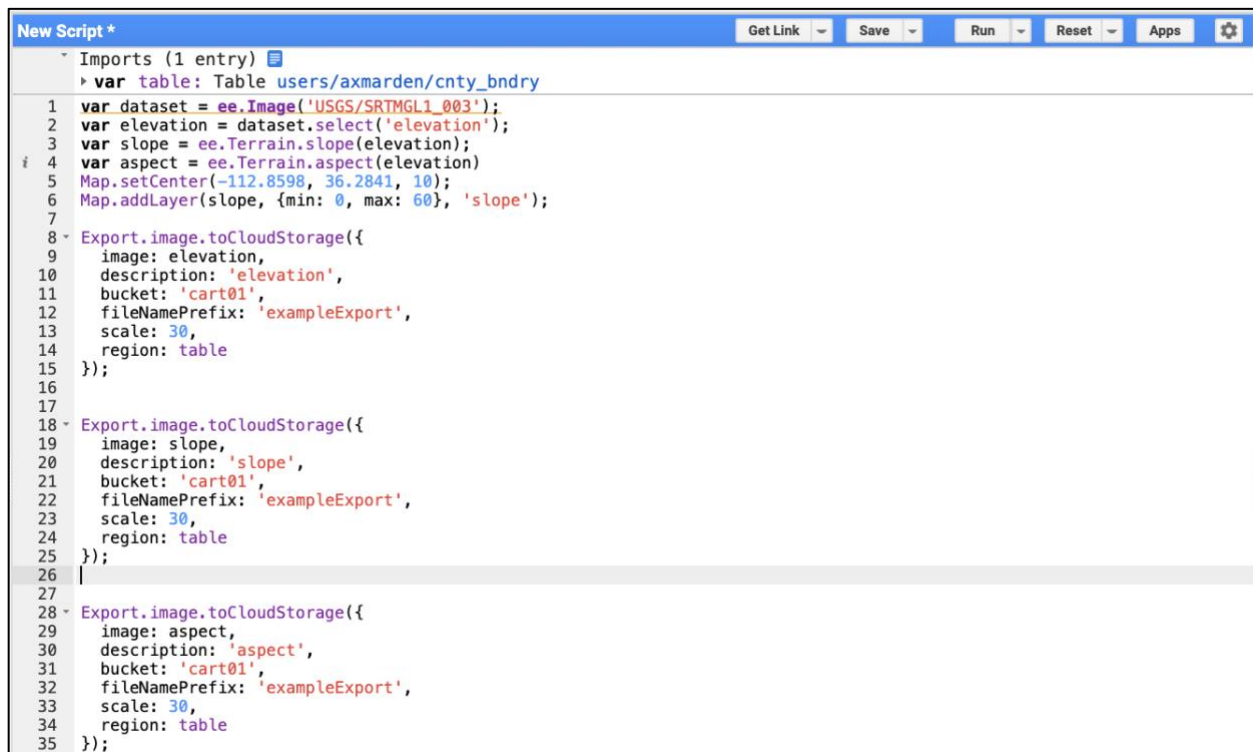
```
1 var dataset_burned = ee.ImageCollection('LANDSAT/LC08/C01/T1_TOA')
2   .filterDate('2020-08-17', '2020-10-02')
3   .filterBounds(table);
4
5 print(dataset_burned)
6 var nbr_burned =
7   dataset_burned.select(['B5', 'B6']);
8
9
10 var nbr2_burned = function(image) {
11   return image.expression(
12     '(NIR - swir) / (NIR + swir)', {
13     'NIR': image.select('B5'),
14     'swir': image.select('B6'),
15     }).rename('nbr').set('system:time_start', image.get('system:time_start'));
16 };
17
18 var dataset2_burned = dataset_burned
19   .select(['B5', 'B6'])
20   .map(nbr2_burned).select('nbr');
21
22 //print(ui.Chart.image.series(dataset2_burned, geometry, ee.Reducer.mean(), 30));
23
24 var max_burned = ee.Image(dataset2_burned.median().clip(table))
25
26 Map.addLayer(max_burned, {min: -1, max: 1, palette: ['blue', 'white', 'green']}, 'Burned');
27
28 //-----
29
```

Figure 1—GEE NBR code



## Part 1 Step 2: Terrain data Acquisition

This step was also performed using GEE, but is relatively simple.



```
Imports (1 entry)
  var table: Table users/axmarden/cnty_bndry
1 var dataset = ee.Image('USGS/SRTMGL1_003');
2 var elevation = dataset.select('elevation');
3 var slope = ee.Terrain.slope(elevation);
4 var aspect = ee.Terrain.aspect(elevation)
5 Map.setCenter(-112.8598, 36.2841, 10);
6 Map.addLayer(slope, {min: 0, max: 60}, 'slope');
7
8 Export.image.toCloudStorage({
9   image: elevation,
10  description: 'elevation',
11  bucket: 'cart01',
12  fileNamePrefix: 'exampleExport',
13  scale: 30,
14  region: table
15 });
16
17
18 Export.image.toCloudStorage({
19   image: slope,
20   description: 'slope',
21   bucket: 'cart01',
22   fileNamePrefix: 'exampleExport',
23   scale: 30,
24   region: table
25 });
26
27
28 Export.image.toCloudStorage({
29   image: aspect,
30   description: 'aspect',
31   bucket: 'cart01',
32   fileNamePrefix: 'exampleExport',
33   scale: 30,
34   region: table
35 });
```

Figure 5—GEE code for Acquisition of terrain data

Elevation data (DEM) were acquired from the [NASA shuttle radar topography mission \(SRTM\)](#). This dataset was selected because it is 30m resolution which matches the resolution of Landsat 8. The GEE functions `ee.Terrain.slope()` and `ee.Terrain.aspect()` were used to produce slope and aspect datasets from the DEM. For more information on these functions see: <https://developers.google.com/earth-engine/apidocs>. Each dataset—DEM, Aspect, and Slope—were exported to cloud storage for use with ArcMap.

## Part 1 Step 3: Fuels data acquisition

Here data were collected that reflect fuel conditions.

First, tree cover percentage was acquired from the [2016 USGS National Land Cover Database \(resolution 30m\)](#).



Second, NDVI (a measure of vegetation productivity) was acquired from [Landsat 8 Collection 1 Tier 1 32-Day NDVI Composite](#) (30m) for the month prior to when the fires started (July, 2020).

## Part 1 Step 4: Vegetation for Zonal Statistics

For zonal statistics, detailed vegetation data in a vector format was acquired from a map produced by a University of California, Davis (UCD) group ([Thorne et al., 2020](#)). The shapefile contains a scheme with 71 different vegetation types and a simplified scheme with 11 different types (Figure 8).

## Part 2: Zonal Statistics

All of the acquired data were added to ArcMap and projected in NAD83/ UTM Zone 10.

The tool “Zonal Statistics as Table (Spatial Analysis)” was used to produce statistics from the dNBR raster for each vegetation type from the Napa County vegetation shapefile. These statistics show Min/Max/Mean of dNBR values for each vegetation type.

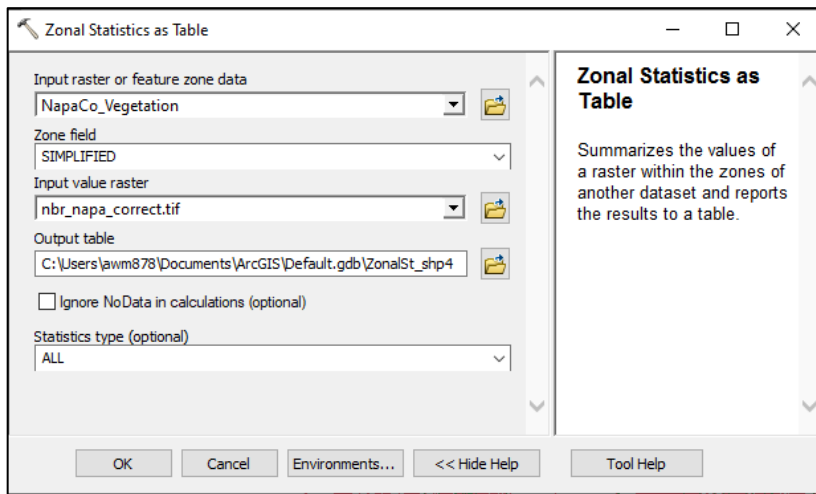


Figure 6— ArcGIS Zonal Statistics GUI

### **Part 3 Step 1: Prepare data for regression and Geographically Weighted Regression**

Ordinary least squares regression is a method of estimating the relationship between a dependent variable and independent variable(s). In this case focus is on the relationship between fire severity (dependent variable) and topographical/vegetation characteristics (independent variables). To prepare raster data for OLS regression I produced XYZ tables where each pixel is represented by an X and Y value with a series of Z values—NBR, NDVI, Percent Tree Cover, Slope, Aspect, and Elevation. To produce each table it was necessary to stack each raster image as separate bands within an image file so that all of the pixels were exactly aligned. The ArcMap tool [“Composite Bands”](#) was used to stack bands. Each raster layer was exported from GEE at the same resolution (30m) and clipped to the same extent (Napa County) so no further processing was necessary prior to producing the composite band image.

After the composite bands image was produced, the tool [“Raster to ASCII”](#) was used to produce a table with X, Y values and corresponding Z values. The ASCII was exported to a CSV using Excel. The software [STATA](#) was used to run the Ordinary Least Squares Regression.

The limitation of OLS regression is that there is a loss of spatial information. Spatial patterns are exchanged for single values. Geographically Weighted Regression allows for a combination of regression and spatial patterns by calculating local effects for each geographical unit in a dataset (Fotheringham et al., 2003). These local coefficients are calculated by comparing the dependent and explanatory variables of features with their neighbors. [More can be read here](#). This analysis requires building a computationally taxing neighborhood matrix which limits the number of features that can be included. Because the data used is 30 meter resolution raster data, there are way too many datapoints to use the full Napa County study area in the GWR analysis. One option to deal with this would be to aggregate the cells to a coarser resolution. This option would allow for an analysis to take place over a larger area, but spatial nuance pertaining to the variables in question would be lost. For this analysis, I decided that it was more interesting to gain a fine resolution perspective of the variables over a small space rather than a coarse perspective over a larger space. To retain a high resolution, a subset was taken of the study area in an area with variation in dNBR, topography, and vegetation(see figure 9).The clip function was used to clip the multiband image to the smaller area. The ArcMap GWR function does not support raster data so the the tool [“Raster to ASCII”](#) was used again to create an ASCII file which was reformatted as a CSV using Excel. The application [GWR 4.0](#) which is run by the Spatial Analysis Research Center at Arizona State University was used to perform the GWR analysis because it supports raster data.

## Results

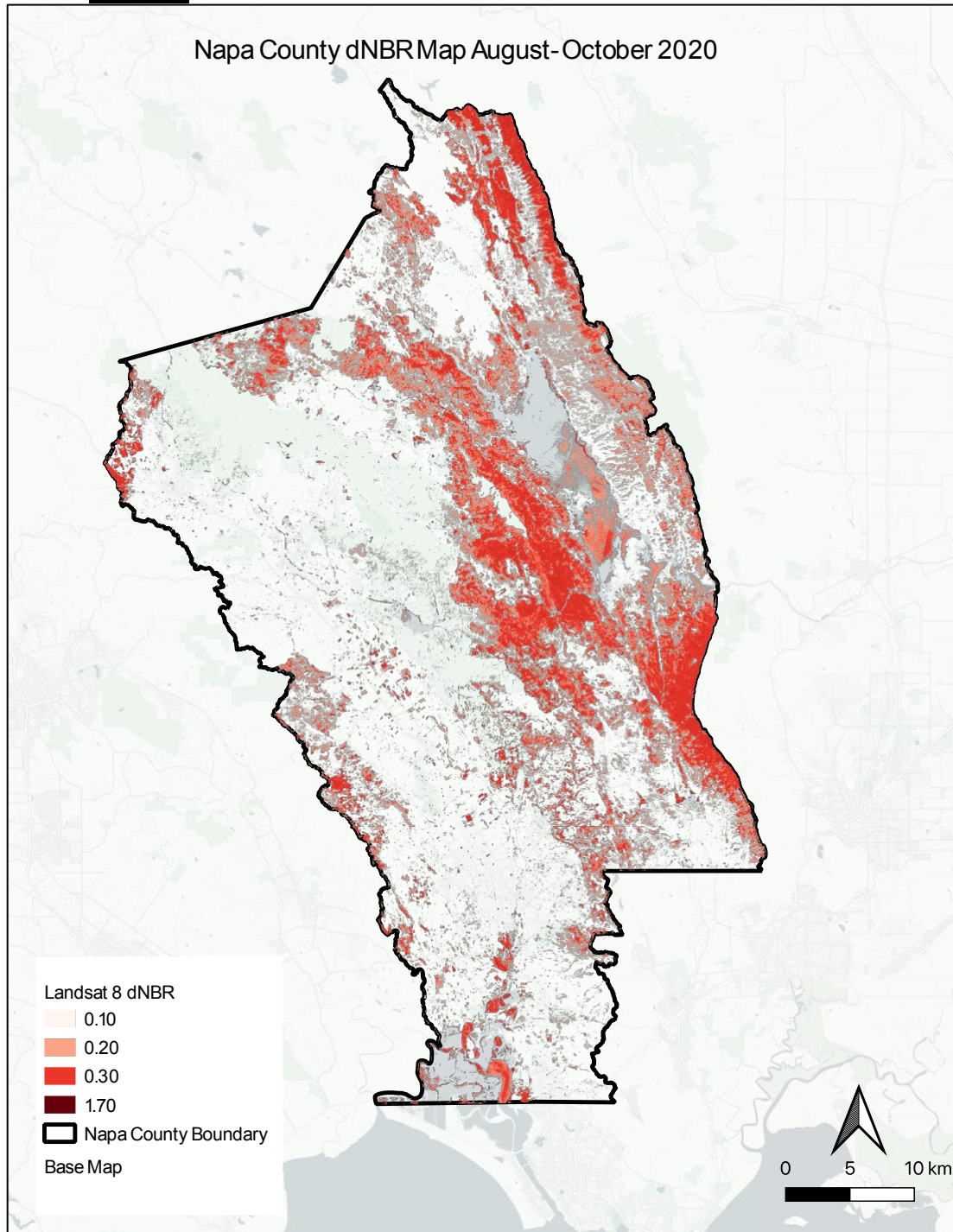


Figure 7—Classified dNBR image for August-October 2020. dNBR values under 0.10 are transparent leaving only pixels classified as burnt shown in graduated red where the darker the red the higher the fire severity.



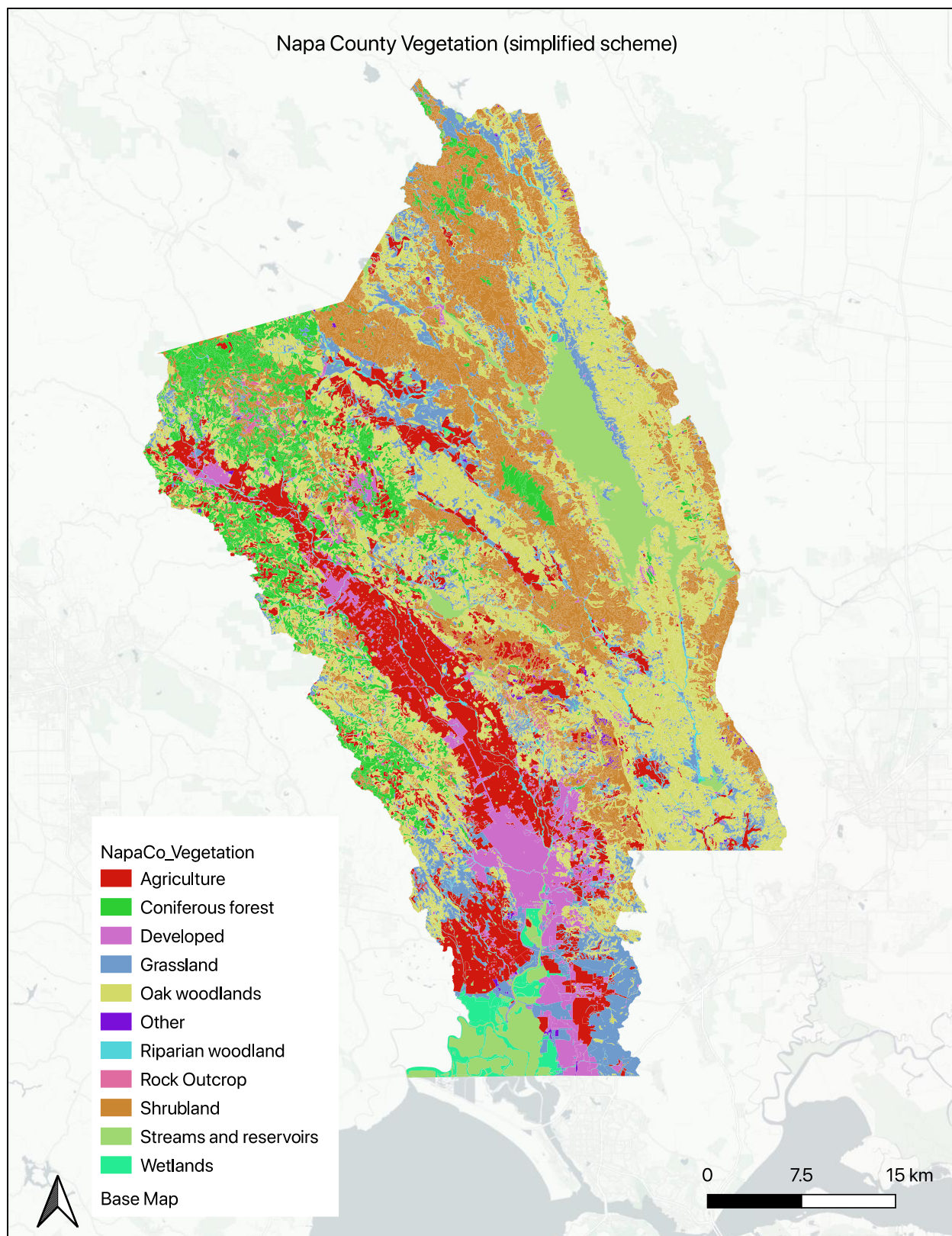


Figure 8—Simplified vegetation scheme from [\(Thorne et al., 2020\)](#). Basemap from [Carto.com](#)

OBJECTID*	SIMPLIFIED	ZONE_CODE	COUNT	AREA	MIN	MAX	RANGE	MEAN
1	Shrubland	1	616331	0.044762	-0.247957	0.744694	0.992651	0.175567
2	Oak woodlands	2	927630	0.067371	-0.365204	0.789322	1.154526	0.131958
8	Wetlands	8	35215	0.002558	-0.85878	0.78399	1.642769	0.130977
5	Other	5	9706	0.000705	-0.479241	0.602683	1.081924	0.127696
6	Riparian woodland	6	46886	0.003405	-0.325018	0.575931	0.90095	0.096828
11	Coniferous forest	11	249555	0.018125	-0.263563	0.658919	0.922482	0.08674
4	Grassland	4	301269	0.02188	-0.510595	0.631244	1.141839	0.0635
9	Streams and reservoirs	9	184710	0.013415	-7.086919	3.793003	10.879922	0.050456
7	Agriculture	7	376159	0.027319	-0.487333	0.574803	1.062135	0.029284
10	Rock Outcrop	10	10172	0.000739	-0.182304	0.582529	0.764833	0.028343
3	Developed	3	162630	0.011811	-0.551394	0.565173	1.116567	0.028218

Table 1—Zonal result for the simplified vegetation scheme—ordered descending based on the mean NBR value

Rowid	NAME	ZONE-CODE	COUNT	AREA	MIN	MAX	RANGE	MEAN
55	California Buckeye	55	509	0.000037	0.034501	0.647516	0.613015	0.373751
53	California Bay - Interior Live Oak	53	243	0.000018	0.040562	0.66504	0.624478	0.354502
9	Scrub Interior Live Oak - Scrub Oak - (California Bay - California Ash - Birch Leaf Mountain Mahogany -	9	63070	0.004581	-0.162229	0.744694	0.906923	0.311635
38	Sparse California Juniper - Canyon Live Oak - California Bay - California Buckeye / Steep Rock Outcrop	38	3005	0.000218	0.007192	0.672937	0.665744	0.263131
69	Arroyo Willow	69	28	0.000002	0.12269	0.355301	0.232611	0.262967
54	Yerba Santa Alliance	54	71	0.000005	0.163521	0.449917	0.286396	0.253558
40	Interior Live Oak	40	29955	0.002176	-0.217796	0.789322	1.007117	0.252226
20	Chamise - Wedgeleaf Ceanothus	20	41121	0.002987	-0.156062	0.703094	0.859157	0.235828
27	(Bulrush - Cattail) Fresh Water Marsh	27	4525	0.000329	-0.272542	0.782229	1.054771	0.234685
52	Occasionally Flooded Grassland & Forbs	52	1118	0.000081	-0.161773	0.670706	0.832479	0.233518

Table 2—Zonal result for the full vegetation scheme—the top 10 vegetation types based on mean dNBR value are included

Source	SS	df	MS	Number of obs	=	80,676
				F(4, 80671)	=	11003.53
Model	576.072137	4	144.018034	Prob > F	=	0.0000
Residual	1055.85032	80,671	.01308835	R-squared	=	0.3530
				Adj R-squared	=	0.3530
Total	1631.92246	80,675	.020228354	Root MSE	=	.1144
nbr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
slope	.0038083	.0000602	63.24	0.000	.0036903	.0039263
elevation	-.0000178	2.93e-06	-6.10	0.000	-.0000236	-.0000121
percent_tree	-.0005732	.0000238	-24.04	0.000	-.0006199	-.0005264
ndvi	.6196549	.0040511	152.96	0.000	.6117148	.6275951
_cons	-.1302716	.0020401	-63.85	0.000	-.1342703	-.126273

Table 3— OLS Regression output from STATA. Aspect was taken out of the analysis because it was insignificant.



Figure 9—GWR subset is shown as red box. A detailed close up shows aspect overlaid with partially transparent classified dNBR. Basemap from Carto.com.



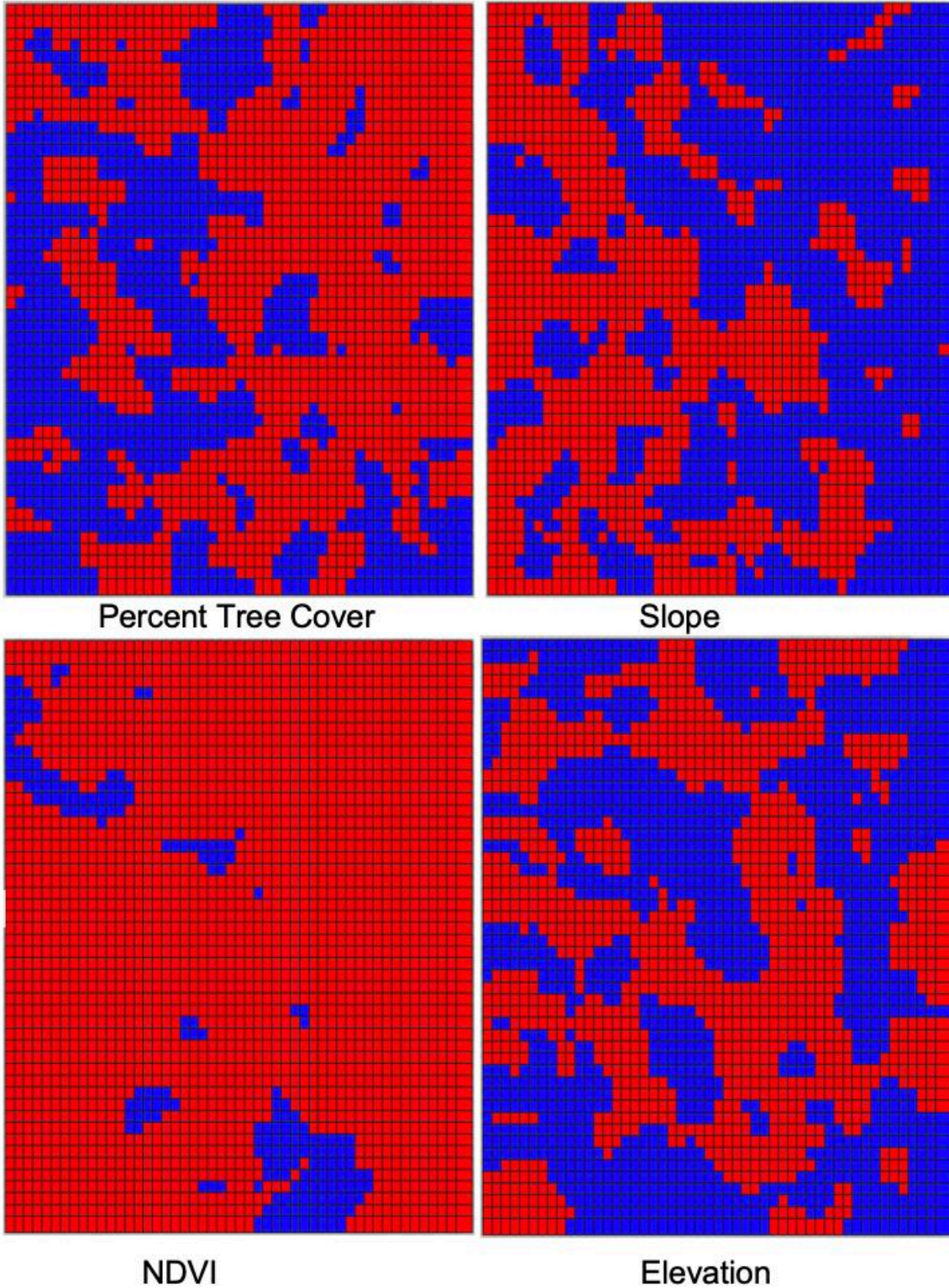


Figure 10—Local Coefficients for each variable are plotted across the GWR subset area. Positive local coefficients are represented by red pixels while negative local coefficients are represented by blue pixels

## **Discussion/Conclusion**

The analyses here provide some important takeaways regarding the LNU Lightning Complex Fire and methodological insights regarding NBR and regression analyses. First, it is clear that the most severe fire occurred in the eastern portion of the study area as evident in Figure 7. Zonal statistics showed that among the simplified vegetation scheme, Shrublands had the highest mean dNBR while Oak Woodlands had the highest single dNBR value (0.789). It is important to note that while streams and reservoirs had the highest maximum value at 3.79, this value is almost certainly a result of water features and not fire given the outlier nature of the figure and the fact that it is a water feature. What is questionable is the wetlands class which has the third highest mean and second highest maximum value (discounting the water features). Wetlands can and do burn, particularly during droughts, but they are also prone to the same fire detection complications caused by water. Further ground and remote truthing is necessary to explore this dynamic in this instance. The detailed vegetation classification with 71 classes identified several vegetation types with high mean dNBR values (e.g. California Buckeye & Interior Live Oak). These vegetation types are likely limited enough in area that would facilitate intensive management regarding post fire erosion, ecological management, and future vegetation specific fire management strategies.

The OLS regression analysis showed that both topographic and vegetation characteristics have relationships with fire severity. However, the topographic variables alone (elevation, slope, and aspect) only had a small  $r^2$  value of .09. While this is a low explanatory value, elevation and slope were highly significant. It is understandable with all of the complexity and stochasticity of where fire occurs that there is a low  $r^2$  value. The addition of vegetation variables (percent tree cover and NDVI in July 2020) to the two significant topographical variables improved the  $r^2$  value to 0.35 (Table 3). NDVI proved to be especially influential, although this relationship is somewhat problematic because NDVI is spectrally related to dNBR measurements. A spectrally independent variable that accounts for fine fuel characteristics would be ideal for further analyses.

The GWR analysis showed that beside NDVI, the other variables varied greatly over space. Percent tree cover and slope have especially interesting relationships where percent tree cover is mainly positively related fire severity on the eastern side of the mountain (Figures 9 & 10) and negatively related on the western side. Slope has an opposite pattern where it has a negative association on the eastern side and positive on the western side. This pattern is likely caused by the relationship between trees and slope where at high slopes vegetation growth is limited and therefore fire severity is also limited. Slopes that are steep and can still support vegetation tend to encourage burning compared to flat vegetated slopes. Furthermore, the GWR analysis seems to show that aspect is an important factor regarding these dynamics. Aspect was insignificant in the OLS regression because the 360 measurement system doesn't work well in a regression analysis. Alternatively a better option may be to classify aspect (N, S, E, W) and treat them as dummy variables. These complexities that are illustrated by GWR show the importance of spatial analysis and why local spatial statistics are important for diagnosing issues within non-spatial statistics like OLS regression.



## Data sources and references.

Bowman, David MJS, Jennifer K. Balch, Paulo Artaxo, William J. Bond, Jean M. Carlson, Mark A. Cochrane, Carla M. D'Antonio et al. "Fire in the Earth system." *science* 324, no. 5926 (2009): 481-484.

Cal Fire, 2020:<https://www.fire.ca.gov/incidents/2020/8/17/lnu-lightning-complex-includes-hennessey-gamble-15-10-spanish-markley-13-4-11-16-walbridge/>

Cocke, Allison E., Peter Z. Fulé, and Joseph E. Crouse. "Comparison of burn severity assessments using Differenced Normalized Burn Ratio and ground data." *International Journal of Wildland Fire* 14, no. 2 (2005): 189-198.

Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons, 2003.

Gorelick, Noel, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. "Google Earth Engine: Planetary-scale geospatial analysis for everyone." *Remote sensing of Environment* 202 (2017): 18-27.

Krasnow, Kevin D., Danny L. Fry, and Scott L. Stephens. "Spatial, temporal and latitudinal components of historical fire regimes in mixed conifer forests, California." *Journal of Biogeography* 44, no. 6 (2017): 1239-1253.

[2016 USGS National Land Cover Database \(resolution 30m\).](#)

[Thorne et al., 2020 Vegetation Map](#)

[Landsat 8 Collection 1 Tier 1 32-Day NDVI Composite](#)

[Landsat 8 Collection 1 Tier 1 Top of Atmosphere Reflectance](#)

[NASA shuttle radar topography mission \(SRTM\)](#)