Project Summary

The California Air Resources Board (CARB) creates a natural and working lands (NWL) carbon inventory that tracks the amount of carbon in vegetation biomass and soils in order to understand how much carbon is on the Californian landscape, where the carbon stocks are increasing/decreasing, and the reasons behind those stock changes. Stocks are quantified in 6 land cover categories: 1) Forestland, 2) Grassland, 3) Cropland, 4) Settlements, 5) Wetlands, and 6) Other Lands. Total carbon stocks in vegetative biomass and soils is quantified for each of these 6 land cover categories during a year, as well as the amount of carbon and land acreage that transitioned from one land cover category to another.

Identification of the urban forest extent (biomass carbon stocks within the Settlements category was identified as an area needing improvement. The current methodology uses the US Census TIGER layer of Incorporated Places to identify the urban forest biomass carbon stocks. This leads to carbon accounting issues arising from the exclusion of unincorporated areas within the built environment (e.g. rural communities and ski resorts) being excluded from the urban forest inventory, even though those land cover types would fall within the Settlements category.

This project investigated two potential avenues for improving upon the current methods. The first is to use a supervised land cover classification to improve upon remote sensing of rural communities that are missed in the currently utilized spatial data product, and thus more thoroughly include rural communities and other currently excluded land cover areas. The second is to compare the known areas of excluded Settlements to the Defense Meteorological Program (DMSP) Operational Line-Scan System (OLS) (1992 – 2013) and Visible Infrared Imaging Radiometer Suite (VIIRS) Stray Light Corrected Nighttime Composite bands, Version 4 (2014 - present) datasets to determine whether nighttime lights would be an efficient improvement on determining boundaries of the built environment.

Santa Barbara County was chosen as the study area because it is completely within one ecoregion, has large areas of wildland urban interface (WUI) that are excluded as Settlements in the current NWL inventory, and because the analyst has extensive knowledge of the study area.

The results of the analysis show that the current execution of the supervised classification option was ineffective in improving upon current methods in a time-efficient manner; this method resulted in extensive confusion between land cover categories and did not improve on picking up known areas of rural development. The second method was determined to improve upon picking areas of rural development, but only when the older, DMSP OLS dataset was used. The newer VIIRS dataset didn’t pick up additional developed areas when compared to the currently used LandfireC dataset.

Project Purpose

This project was designed and performed in order to investigate potential avenues for improving the current urban forest biomass carbon inventory produced by CARB. Two potential avenues for improvement were investigated: 1) performing an ecoregion specific, supervised land cover classification, and 2) using the VIIRS Stay Light Corrected Nighttime Composite Bands datasets to identify unincorporated developed areas. The goal is to identify a more accurate and time efficient method for delineating what is considered urban forest ecosystem carbon and what is ecosystem carbon contained within the Settlement land cover category.

Methods

Data Description

Six datasets were used for this project:

1. DMSP OLS: Nighttime Lights Time Series Version 4 (NTL 1)[1]
2. VIIRS Stray Light Corrected Nighttime Composite Bands, Version 1 (NTL 2)[2]
3. LANDSAT 8 analysis ready, surface reflectance data for the year 2014 (LS8)[3]
4. LandfireC (LFc)
5. National Aerial Imagery Program imagery for the year 2014 (NAIP)
6. Santa Barbara County boundary, US Census TIGER layer[4]

The NTL 1 dataset "contains the lights from cities, towns, and other sites with persistent lighting, including gas flares. Ephemeral events, such as fires, have been discarded. The background noise was identified and replaced with values of zero."[1] The pixel size is 30 arc seconds. The units are not listed, but the values represent the average visible band digital number for each pixel in the year 2013. Since this data isn't being used for arithmetic calculations, but rather to create categories that represent levels of urban density, the units were deemed unimportant.

The NTL 2 dataset is similar to the NTL 1 dataset, but the resolution is 15 arc seconds and the units are in nanoWatts/cm2/sr. Descriptions of the methods used to download the NTL 1 & 2 datasets from Google Earth Engine are located in *Appendix I*.

The LS8 dataset was downloaded from the United States Geological Survey (USGS) Earth Explorer website. Two tiles were downloaded for 2014:

1. LC08_CU_002011_20191012_20191022_C01_V01 in row 11, column 2, and
2. LC08_CU_003011_20191012_20191022_C01_V01 in row 11, column 3.

These tiles contain 8 band imagery (Table 1) at 30 meter resolution.

Table 1. Description of bands in the LANDSAT 8 dataset.

Band	Description	Wavelength Range (ɥm)
1	Coastal	0.433 - 0.453
2	Visible blue	0.450 - 0.515
3	Visible green	0.525 - 0.600
4	Visible red	0.630 - 0.680
5	Near-infrared	0.845 - 0.885
6	Short wavelength infrared	1.56 - 1.66
7	Short wavelength infrared	2.10 - 2.30
8	Cirrus	1.36 - 1.39

The LandfireC dataset is a derived product of the LANDFIRE public dataset produced by a collaboration between the United State Forest Service (USFS) and The Nature Conservancy (TNC)[6]. The derived LandfireC dataset contains the IPCC categories for 2014 that were created in-house by a model that uses the existing vegetation types, heights, and disturbance layers from LANDFIRE to determine the amount of biomass carbon per pixel, disturbance type, carbon gain/loss, and IPCC land cover category.

The 2014 NAIP imagery was acquired from the in-house data library.

The boundary of Santa Barbara was downloaded as a TIGER shapefile from the United States Census Bureau for the year 2014.

A 6th dataset, LANDSAT 8 greenest pixel data for the year 2014, was acquired and used in preliminary analysis, but discarded after poor results were generated. Methods for acquiring this data are located in *Appendix I*.

Data Organization

Multiple steps were taken to prepare the data for analysis. First, all data were projected into the NAD 1983 California Teale Albers projection. Next, the eight LANDSAT 8 band layers for both tiles were clipped to the Santa Barbara County boundary. Once the band layers were clipped, they were subsequently compiled into two composites, one for each tile. The two composites were then mosaicked together to produce an 8-band image with wall-to-wall coverage for Santa Barbara County.

Supervised Classification

Supervised classification of land cover was performed on the LANDSAT 8 imagery under two scenarios. The first was using the composited imagery without further data transformation. The display chosen was color infrared, which displays the red band as near-infrared (band 5), the green band as visible red (band 4), and the blue band as green (band 3) (Figure 1). This color combination was chosen because the goal of this project is to differentiate urban and built environments in the WUI, and the analyst decided that the color infrared color display would provide the greatest differentiation between vegetation and built environments.

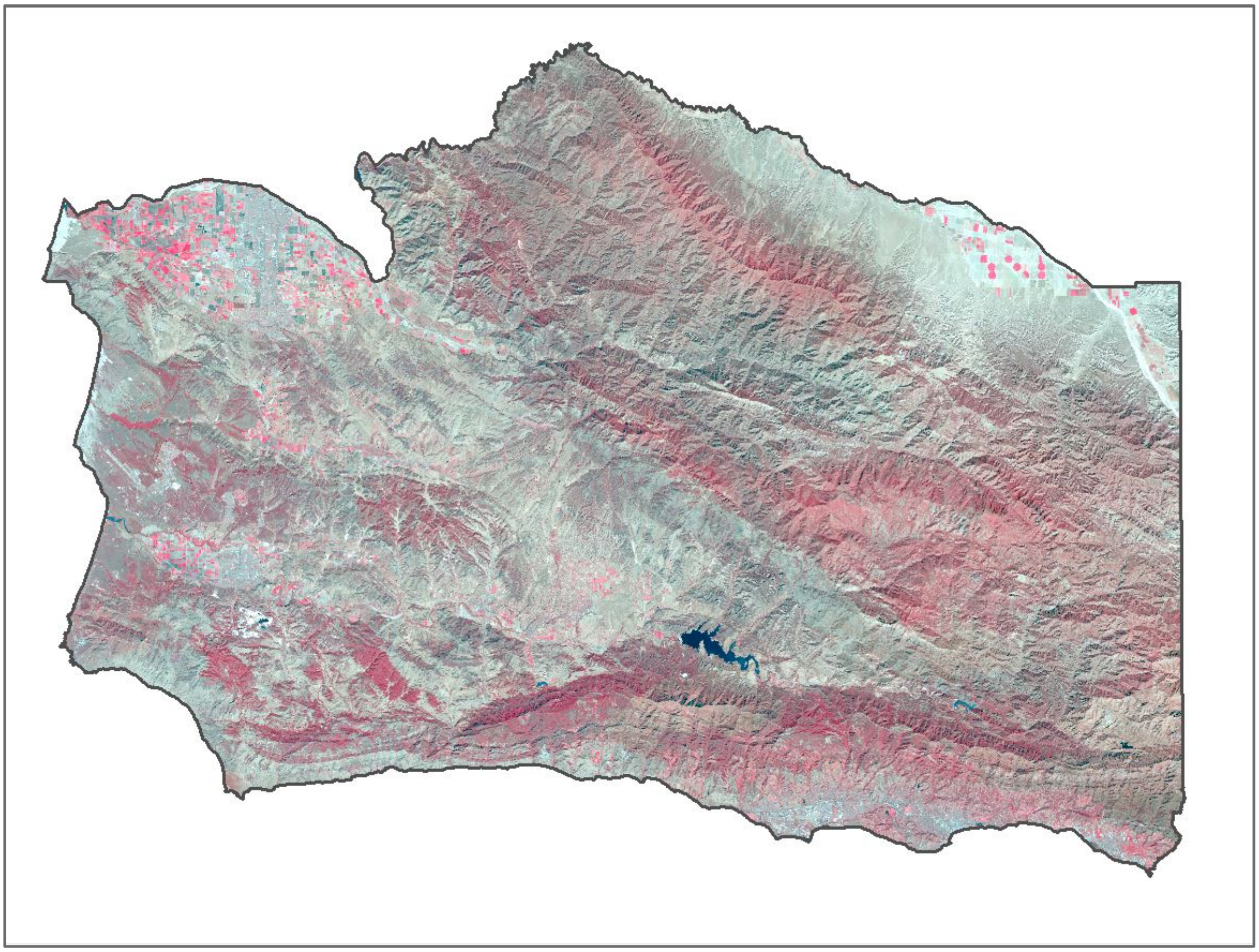


Figure 1. Color infrared display of the 2014 LANDSAT 8 composite data for the year 2014.

The second scenario required a band transformation to normalized differentiation vegetation index (NDVI). The LANDSAT 8 composite was loaded into ArcGIS Pro to perform this analysis and the methods offered in the “Band Rations and Transformations” write up for this course were used. First, the NDVI was created using the raster function “NDVI” to produce a layer that depicts the NDVI of the LANDSAT 8 composite image for Santa Barbara County. The NDVI layer was then used to produce a colorized NDVI layer using the “NDVI Colorized” raster function (Figure 2) for data visualization.

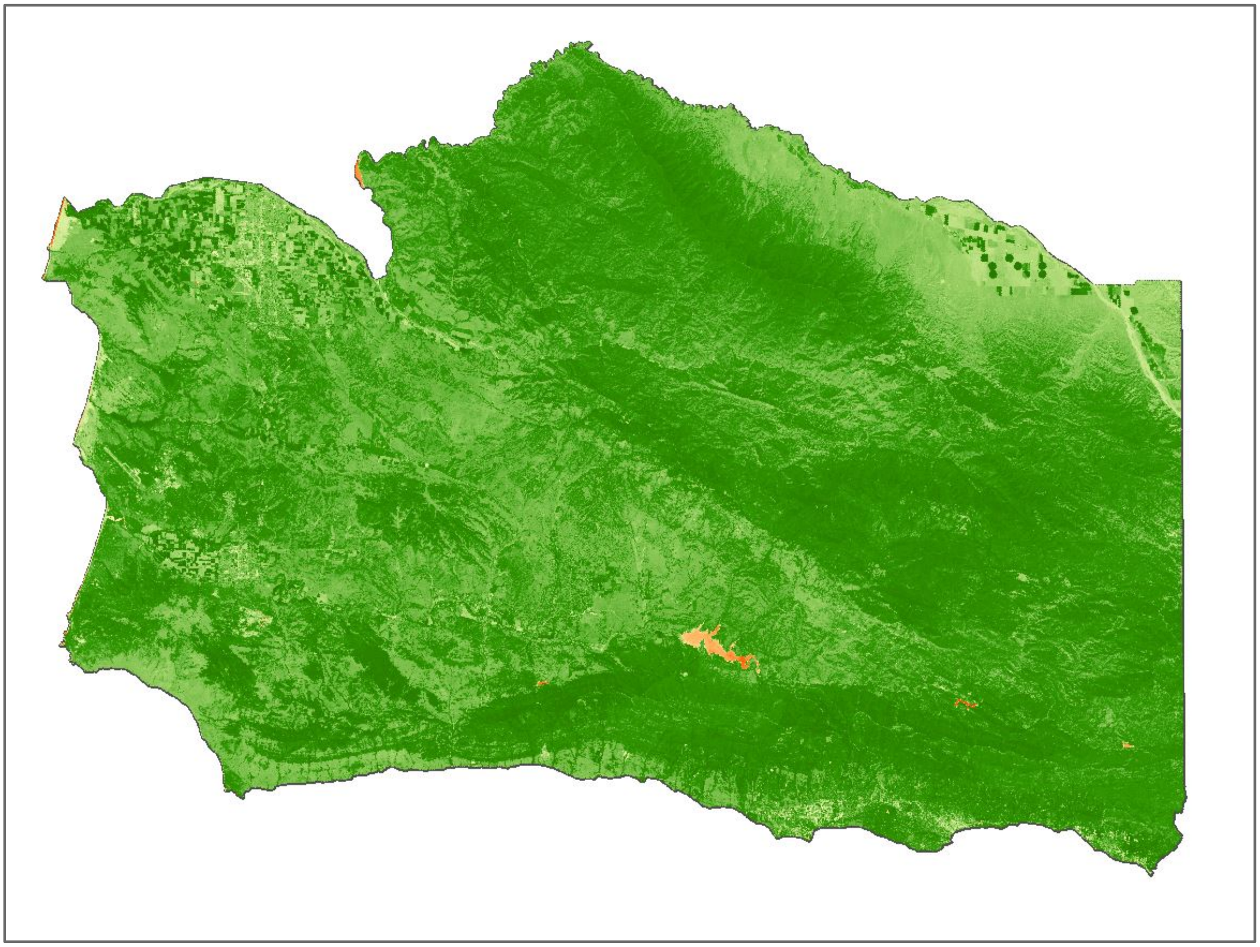


Figure 2. Normalized Difference Vegetation Index of the 2014 LANDSAT 8 composite data for the year 2014.

The raster function “Band Arithmetic” was then used to create two new layers that only displayed the green band and the near-infrared (NIR) bands. The band arithmetic was “Band = 5” for the near-infrared and “Band = 3” for the green band. These two images were then composited with the NDVI image in the order of the green band, the NDVI band, then the NIR band to produce a single image. The color displays were set in the same order as the composite sequence (Figure 3).

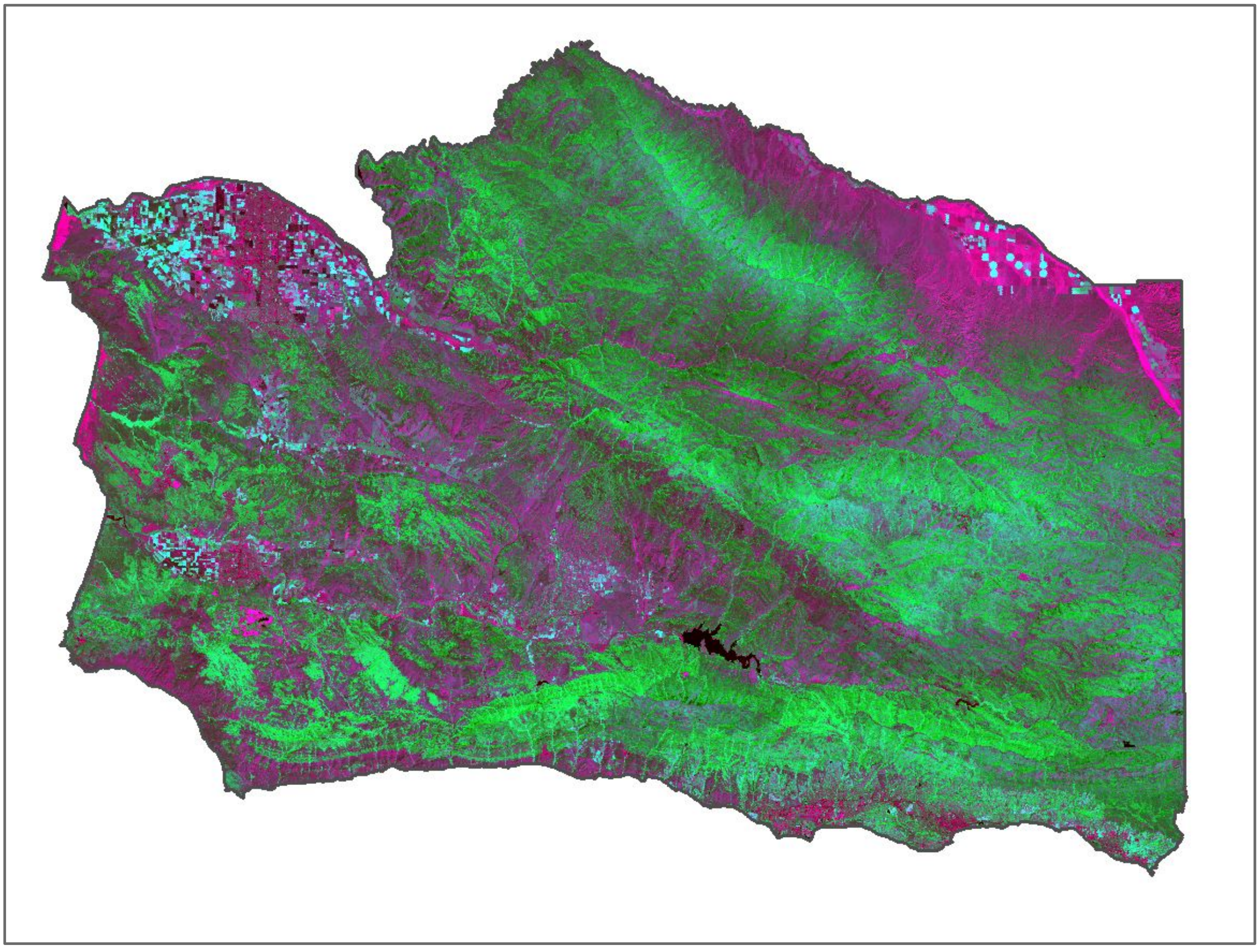


Figure 3. Green band, NDVI, and NIR band composite image of the 2014 LANDSAT 8 data for the year 2014.

The green, NDVI, NIR composite image was then loaded back into ArcGIS Desktop for the supervised classification process. Supervised classification was performed twice; once on the color infrared display of the untransformed LANDSAT 8 composite, and once on the NDVI transformed composite that was created in ArcGIS Pro.

Landcover classification was divided into 8 land cover types: 1) agricultural fields, 2) agricultural orchards, 3) forestland, 4) grassland, 5) industrial development, 6) residential development, 7) wetlands, and 8) open water. When possible, a statistically significant sample size of 40 training sites and 40 test sites was used (Table 2).

Table 2. Number of training and test sites per cover type.

Cover Type	**Number of Training Sites**	**Number of Test Sites**
Agriculture - Fields	40	40
Agriculture - Orchards	40	40
Forestland	40	40
Grassland	40	40
Development - Industrial	10	10
Development - Residential	20	20
Wetland	7	7
Open Water	20	20

The 2014 NAIP imagery was loaded into ArcGIS Desktop and set beneath the LANDSAT 8 imagery. The LANDSAT imagery was set to 50% transparency in order to use the higher resolution (3m) NAIP imagery as a guide in creating training samples. A dendogram and spectral signature was produced for each cover type training sample using the untransformed LANDSAT imagery to ensure that the maximum distance between in-group samples remained below the analyst designated threshold of 10. Then an overall dendogram and spectral signature was created and analyzed for both the untransformed and transformed LANDSAT 8 training samples.

Next, a Maximum Likelihood Classification was performed for both the untransformed and transformed LANDSAT 8 data (Figures 4 and 5).

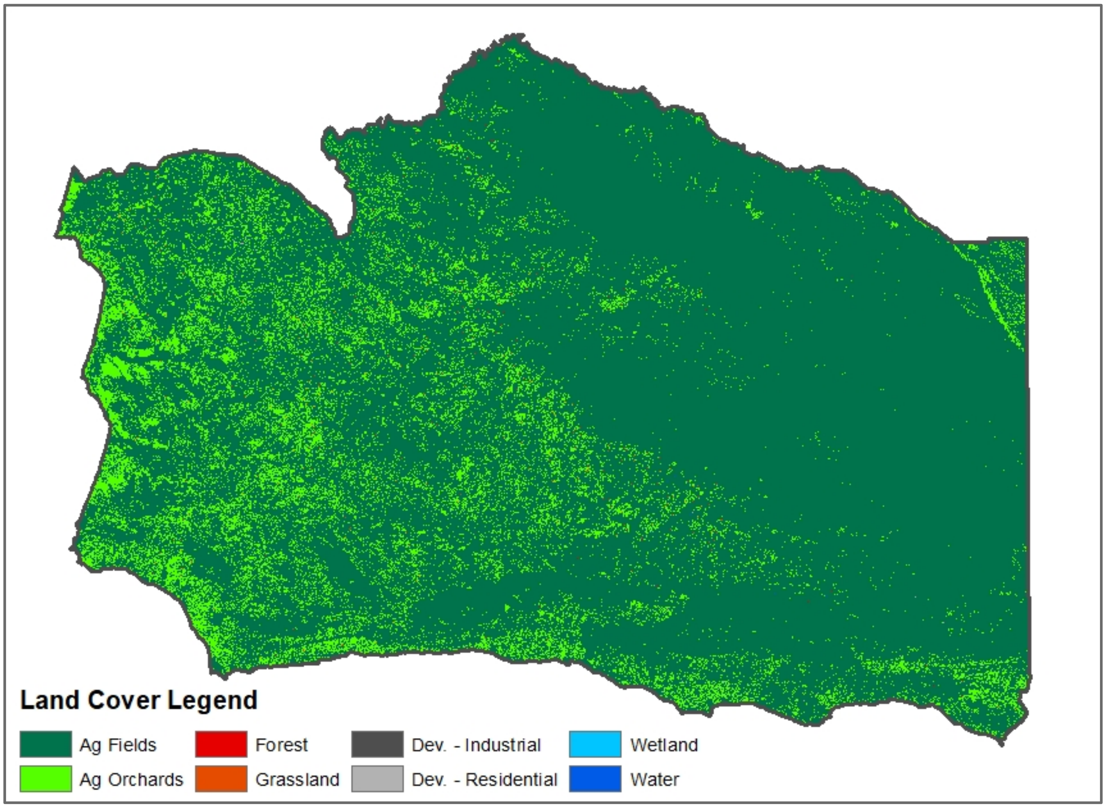


Figure 4. Maximum Likelihood Classification for the untransformed LANDSAT 8 data.

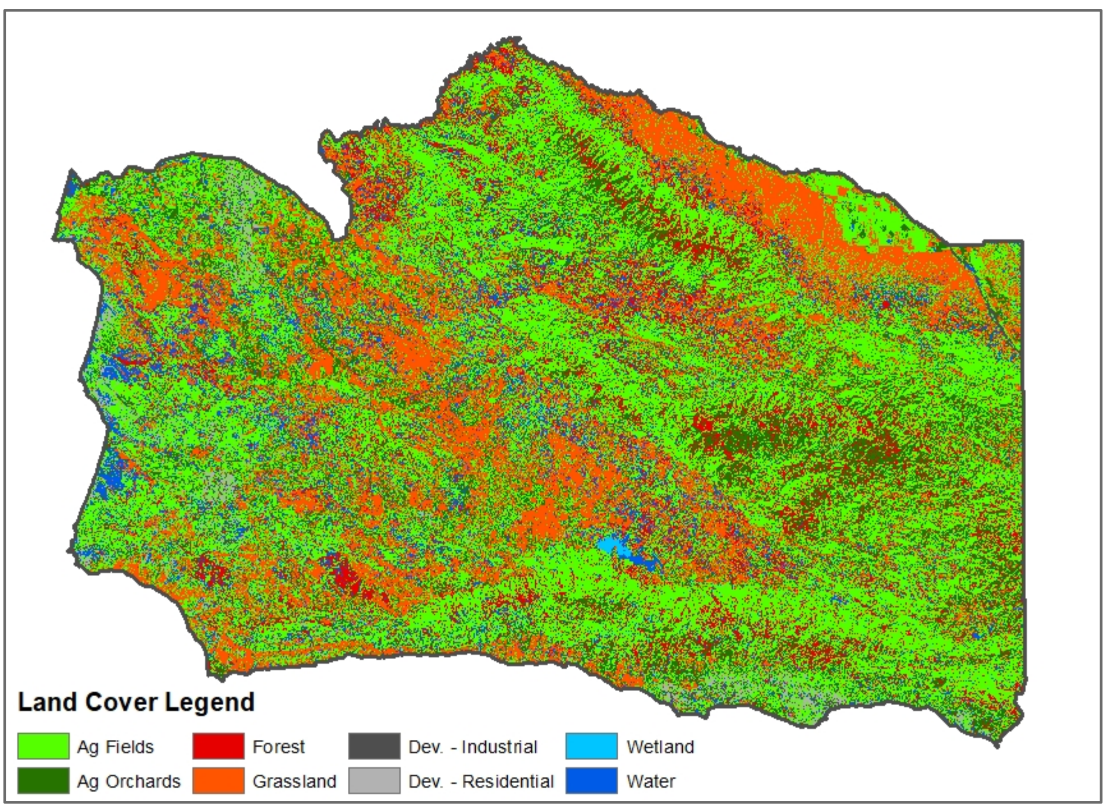


Figure 5. Maximum Likelihood Classification for the transformed LANDSAT 8 data.

Finally, a manual creation of error matrices was used in order for the analyst to gain direct experience in the arithmetic process behind error matrix creation. The test samples were used as masks for the untransformed LANDSAT 8 data (Table 3), the transformed LANDSAT 8 data (Table 4), and on the LandfireC classifications currently used by CARB (Figure 6) (Table 5). The attribute tables of the extracted classifications were then exported to Excel and organized to create both the error matrices and accuracy assessments. Producer's accuracy was calculated according to Equation 1:

$$\textbf{Equation 1}: Producer's\ Acuracy\ (\%) = \frac{\#\ of\ correctly\ categorized\ pixels\ in\ test\ data}{\#\ of\ total\ pixels\ in\ category\ column} \times 100$$

User's Accuracy was calculated according to Equation 2:

$$\textbf{Equation 2}: User's\ Acuracy\ (\%) = \frac{\#\ of\ correctly\ categorized\ pixels\ in\ test\ data}{\#\ of\ total\ pixels\ in\ category\ row} \times 100$$

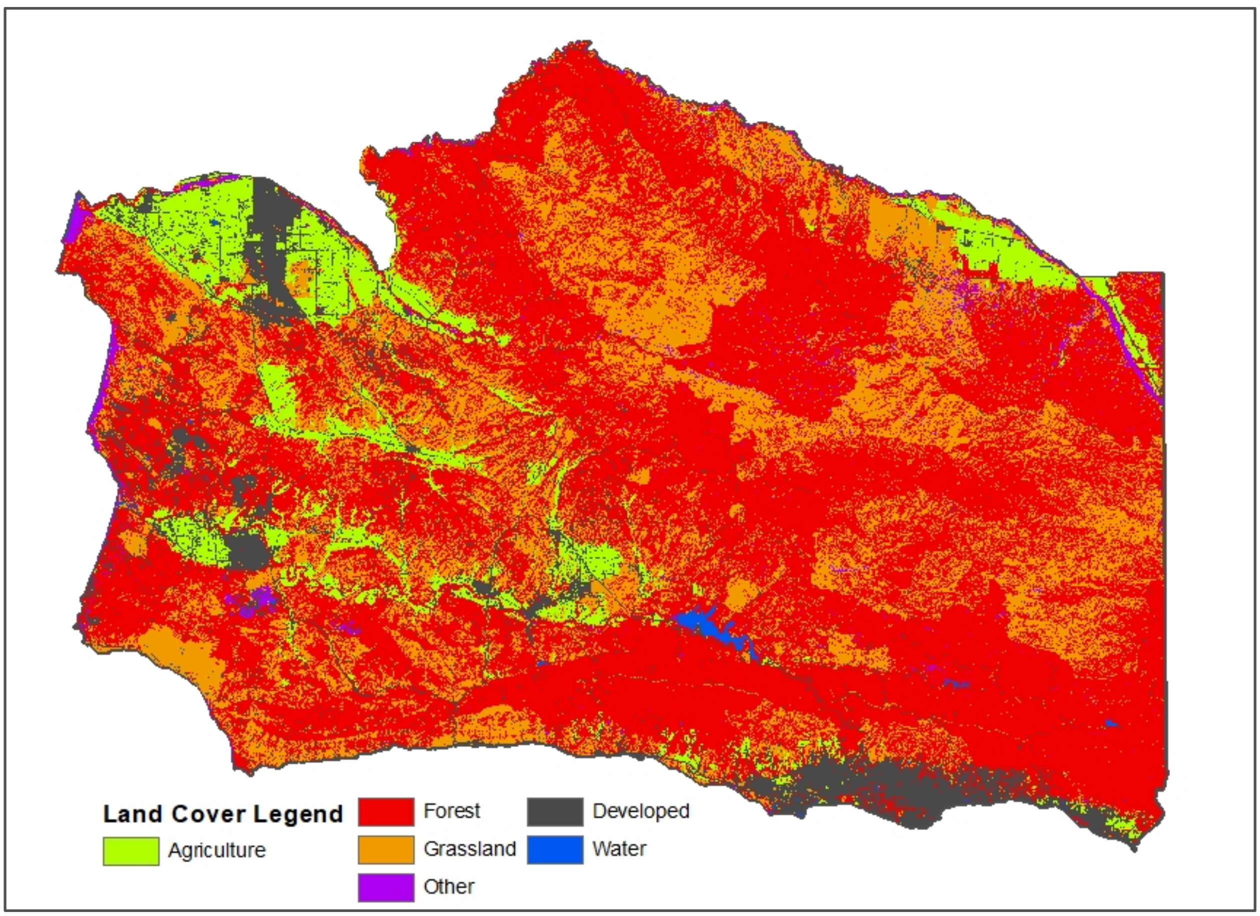


Figure 6. CARB in-house classification of LANDFIRE data to IPCC land cover categories.

Table 3. Error matrix of the maximum likelihood classification performed on the untransformed LANDSAT 8 data.

	Field	Forest	Grass	Indust.	Orch.	Res.	Water	Wet.	Total
Field	14,971	15,353	13,583	614	6,657	8,749	4,393	5,188	69,388
Forest	1,841	106	2,524	449	774	467	31	204	6,396
Grass	0	1	0	0	0	0	0	0	1
Indust.	4	0	12	0	3	0	0	0	19
Orch.	1	0	0	17	6	1	0	1	26
Res.	0	0	0	0	1	14	1	2	18
Water	0	0	0	0	0	0	0	1	1
Wet.	0	0	0	0	0	0	0	1	2
Total	16,818	15,430	16,119	1,080	7,351	9,231	4,425	5,397	75,851

Table 4. Error matrix of the maximum likelihood classification performed on the transformed LANDSAT 8 data.

	Field	Forest	Grass	Indust.	Orch.	Res.	Water	Wet.	Total
Field	9,556	10,072	2,961	308	2,462	4,279	63	1,316	31,017
Forest	3,737	594	334	15	3,188	1,391	12	541	9,812
Grass	44	4,572	20	0	319	253	9	2,228	7,445
Indust.	3,182	8	12,314	23	975	63	6	42	16,613
Orch.	51	0	373	597	9	292	53	361	1,736
Res.	95	1	18	73	237	2,511	24	46	3,005
Water	0	0	0	0	0	0	2,402	0	2,402
Wet.	153	183	99	64	161	442	1,856	863	3,821
Total	16,818	15,430	16,119	1,080	7,351	9,231	4,425	5,397	75,851

Table 5. Error matrix created using the training samples that were used to perform the maximum likelihood classifications in Figures 6 and 7 and the LandfireC data.

	Ag.	Forest	Grass	Other	Dev.	Wetland	Water	Total
Ag.	20,734	0	762	0	0	62	12	21,570
Forest	357	15,245	3,769	0	471	3,160	55	23,057
Grass	204	65	10,481	0	48	512	15	11,325
Other	0	0	0	0	0	65	91	156
Dev.	2,848	122	471	0	9,791	1,140	54	14,426
Wetland	23	0	639	0	0	0	0	662
Water	2	0	0	0	0	459	4,198	4,659
Total	24,168	15,432	16,122	0	10,310	5,398	4,425	75,855

Both the producers and users accuracy was significantly lower for the maximum likelihood classification on the untransformed dataset (Table 6) and transformed dataset (Table 7) as opposed to the LandfireC data (Table 8).

Table 6. Producers and users accuracy for the classification performed on the untransformed LANDSAT 8 dataset.

Land Cover Category	Producer's Accuracy	User's Accuracy
Agriculture Fields	89%	22%
Forest	1%	2%
Grassland	0%	0%
Developed – Industrial	0%	0%
Agriculture Orchards	0%	23%
Developed – Residential	0%	78%
Water	0%	0%
Wetlands	0%	50%

Table 7. Producers and users accuracy for the classification performed on the transformed LANDSAT 8 dataset.

Land Cover Category	Producer's Accuracy	User's Accuracy
Agriculture Fields	57%	31%
Forest	4%	6%
Grassland	0%	0%
Developed – Industrial	2%	0%
Agriculture Orchards	0%	1%
Developed – Residential	27%	84%
Water	54%	100%
Wetlands	16%	23%

Table 8. Producers and users accuracy for the classification performed on the LandfireC dataset.

Land Cover Category	Producer's Accuracy	User's Accuracy
Agriculture	86%	96%
Forest	99%	66%
Grassland	65%	93%
Other Land	0%	0%
Settlement	95%	68%
Wetland	0%	0%
Waterbody	95%	90%

Use of the Nighttime Lights Datasets

The DMPS OLS and VIIRS datasets did not have data for the same year and had differing units, so the initial analysis required the creation of like categories. The original DMPS OLS data had a DBN range of 0 – 63 (Figure 7); the final publication year of 2013 was used for this analysis. The original VIIRS data had a band range of 0 – 270 (Figure 8); the first publication year of 2014 was used for this analysis.

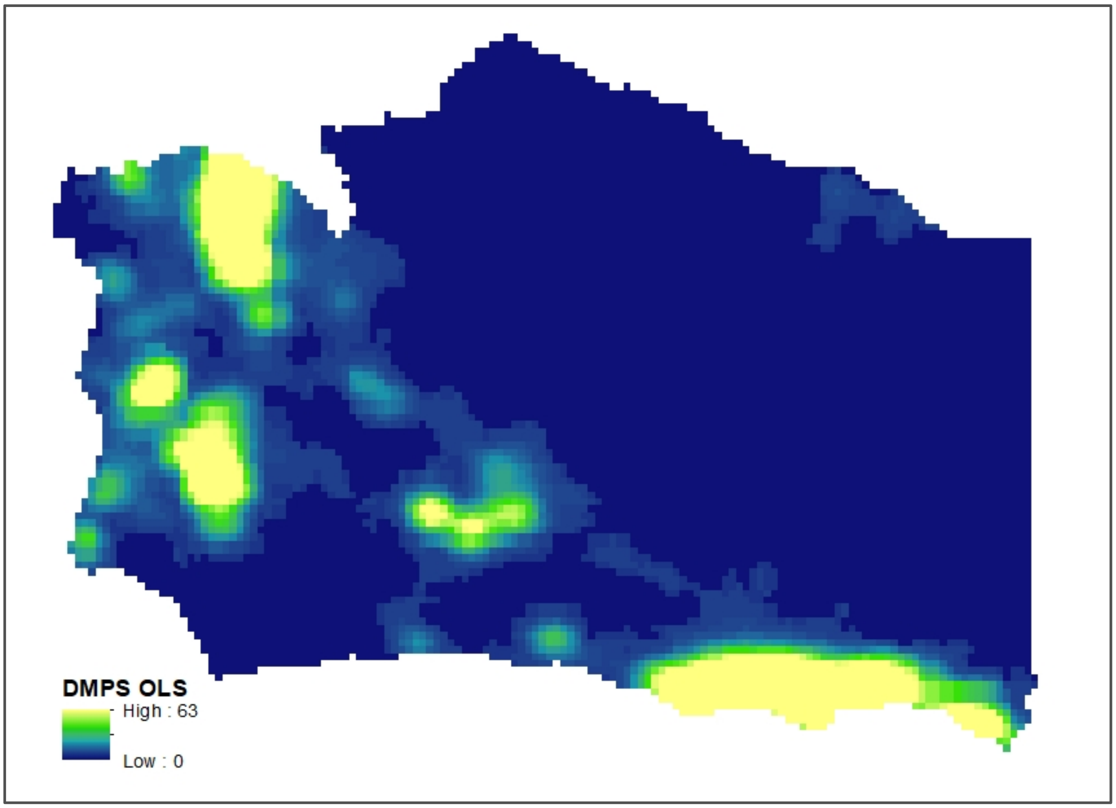


Figure 7. Uncategorized, raw data import of the DMPS OLS data for the year 2013

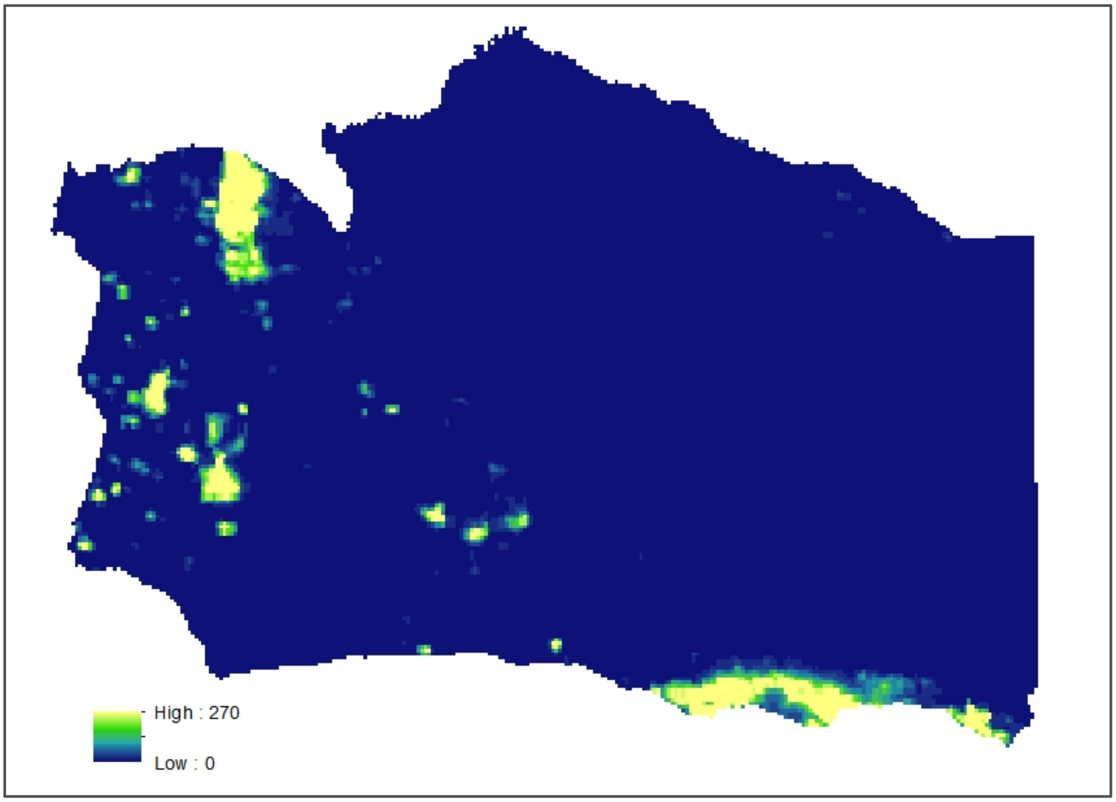


Figure 8. Uncategorized, raw data import of the VIIRS data for the year 2014.

Although an overlapping year wasn't available, the analyst determined that there wouldn't be a signficant amount of additional development between 2013 and 2014 after the financial crisis. Hence, the categorization was performed for the 2013 and 2014 data while assuming that the developed area would remain constant between years.

The digital band numbers for both the DMPS OLS (Figure 9) and VIIRS (Figure 10) datasets were broken out into 4 categories: 1) No Development, 2) Low Development, 3) Mid Development, and 4) High Development. Low development was defined as small rural communites, Mid Development as suburban communities or the fringes of High Development areas, and High Development was defined as the urban core. The categorization was performed by classifying both datasets in the "Properties" section, "Symbology" tab into 4 categories (Table 9).

Table 9. Classification of digital band number values for the 2013 DMPS OLS and 2014 VIIRS datasets.

	DMPS OLS	VIIRS
No Development	0 – 5	0
Low Development	6 – 10	1 – 75
Mid Development	11 – 25	76 – 100
High Development	26 – 63	101 - 270

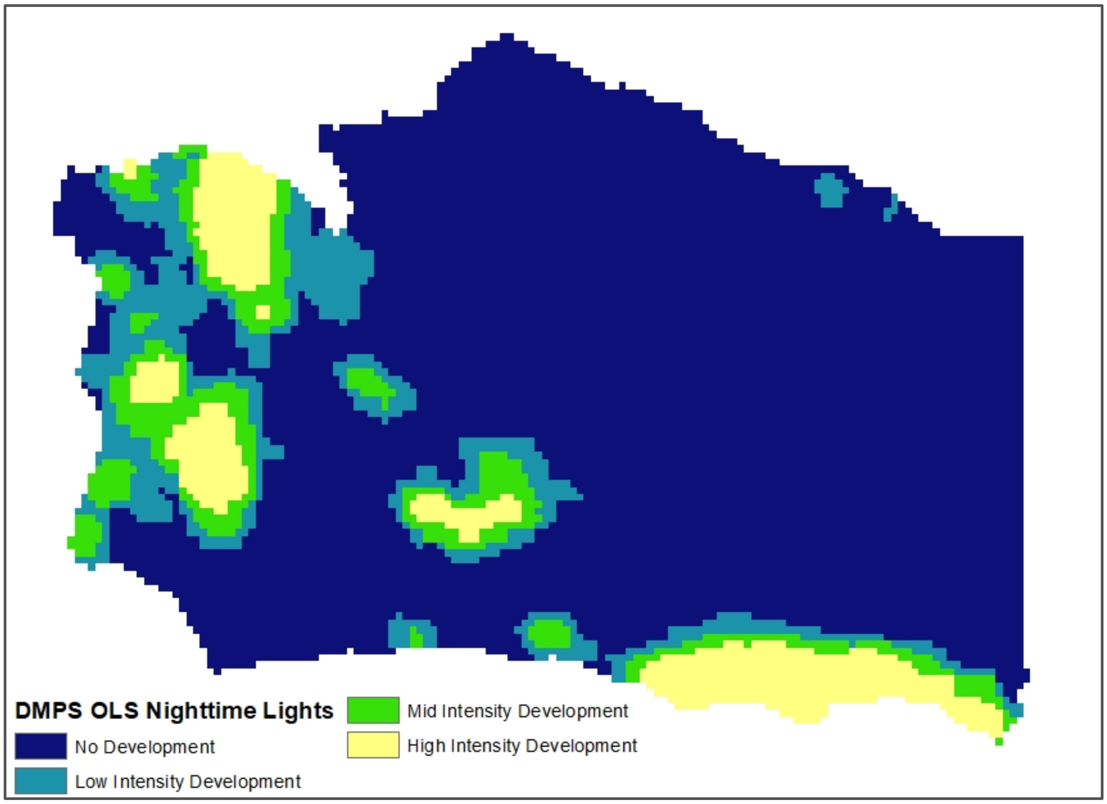


Figure 9. Classified DMPS OLS dataset for the year 2013.

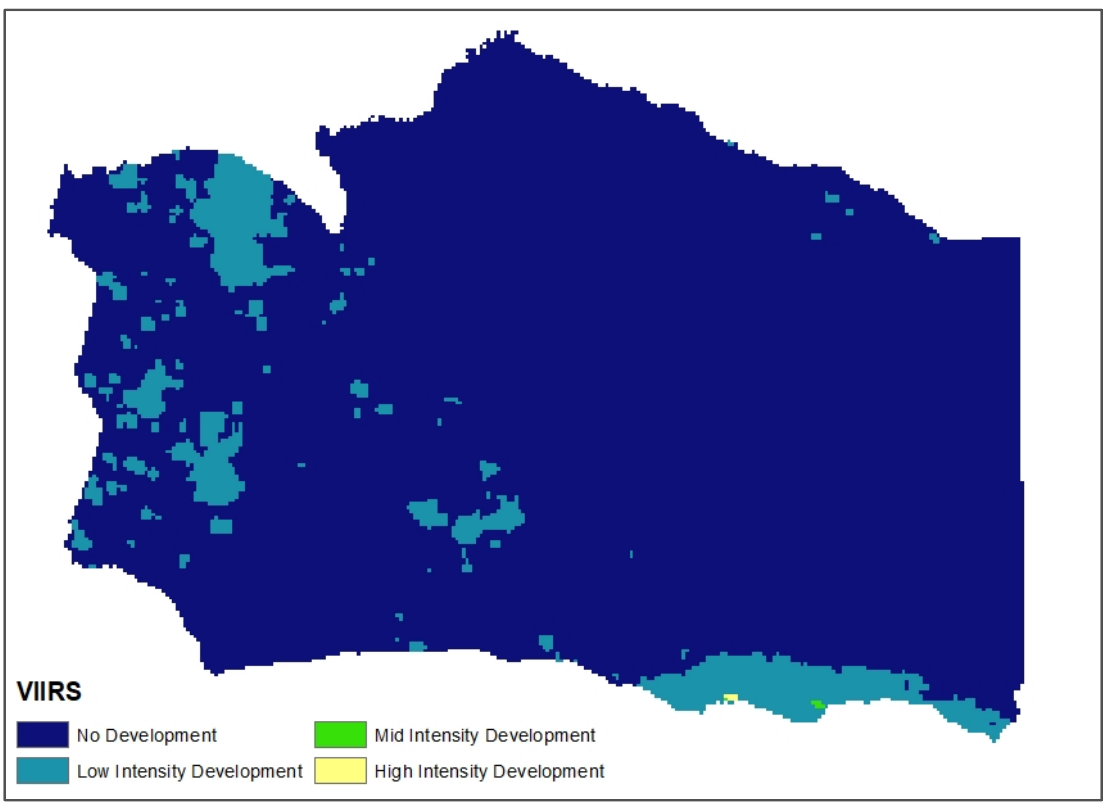


Figure 10. Classified VIIRS dataset for the year 2014.

Analysis Issues

Supervised Classification

Multiple analysis issues arose during the classification process. First, the original LANDSAT 8 greenest pixel dataset chosen for this analysis produced poor results, and was replaced with the analysis ready surface reflectance dataset used in this report. The greenest pixel data was originally chosen because 2014 was a severe drought year, and the analyst initially expected that differentiation between water stressed vegetation and developed areas would be more accurate using the greenest pixel data. It was found that using the analysis ready results provided better differentiation between agriculture/developed areas that received irrigation and other, non-irrigated vegetated areas.

Second, when creating the training samples for the LANDSAT 8 Analysis Ready data, the image classification extension would not produce the desired distribution curves; only the scatter plots were available to analyze the quality of training samples produced. The lack of distribution visualizations likely led to poor training samples. The analysis ready dataset did not produce acceptable results, with the untransformed data performing more poorly than the transformed dataset. Both classifications performed more poorly than the currently used LandfireC dataset.

The results produced via the supervised classification on the untransformed LANDSAT 8 imagery produced unusable classifications, as illustrated by the producer's and user's accuracy for this classification (Table 6). Conversely, the supervised classification on the transformed LANDSAT 8 imagery also produced results that are currently unusable, but the producer's and user's accuracy for this initial classification (Table 7) suggest that this method could produce usable results with further refinement.

An issue that was identified as a potential source of between-class confusion in the supervised classification of the transformed data was that I used the same training and test sample polygons that were created using the untransformed data. When revisiting the data after analysis had been completed, I noticed that the polygons that had appeared to be uniform while using the untransformed data now had a rainbow of colors within a single training polygon. New training and test polygons should be created for the tranformed data – this is expected to produce improved classification results.

Nighttime Lights Categorization

The primary analysis issue is that one, continuous dataset of nighttime lights was not found for this project. The data used was at different resolutions, during different years, and used different sensors. This resulted in the inability to produce like classifications. Most notably, the VIIRS data has a large area of DBN = 0 that the DMPS OLS data identifies as having non-background levels of nighttime illumination. At this point in the analysis, no solution to this has been identified.

Discussion

Supervised Land Cover Category Classification

The analysis was very informative and will guide further refinement of the NWL carbon inventory outside the scope of this course. The supervised classification of land cover using the untransformed LANDSAT 8 imagery produced results that classified nearly the whole county as agricultural. This classification is wholly false and signifies an extreme level of confusion between land cover categories. This method will not be used in future analyses.

The classification performed on the transformed data produced much more promising results. While the confusion of agricultural lands with other vegetated cover types (most notably forest and grassland) still exist to a significant degree, this method could be used in conjunction with an initial land ownership segmentation to differentiate forests from orchards and fields from grassland. This could be achieved by downloading parcel data and segmenting it first by private versus public ownership, then further segmenting the private ownership by zoning to reduce confusion between cover categories. Focus would be on improving the accuracy of the Developed – Industrial category, as there is considerable confusion.

A possible source of between category confusion was identified after the classifications and error matrices had been produced. The training and test data polygons used to create the spectral signatures for supervised classification on the transformed data were created using the untransformed data as a visual guide. When the polygons were reinvestigated as an overlay on the transformed data, the analyst noted that the spectral signatures of each polygon were no longer consistent, and classifications will likely improve drastically once training and test polygons are drawn specifically for the transformed dataset.

Another promising finding that is outside the focus of this project, but relevant to my broader work, is the improved accuracy of wetland identification as compared to the accuracy of the LandfireC data. LANDFIRE is known to have great difficulty in identifying wetlands, and this analysis shows that it is possible for the analyst to improve upon current wetland mapping via remote sensing with the LANDSAT 8 aerial imagery.

Nighttime Lights Categorization

The Nighttime lights data does provide a time efficient method for quickly identifying developed rural communities and unincorporated areas. The most notable drawback would be the inconsistency between the two datasets; developed areas measured using the DMPS OLS dataset are significantly over-estimated, and developed areas measured using the VIIRS dataset are significantly under-estimated. Unfortunately, this method was unsuccessful in identifying rural developments in the WUI. A possible use of these two datasets would be to identify areas of dense development that are outside of US Census Incorporated Areas (e.g. prisons and ski resorts) by producing a two category classification of background and dense development. This route will be investigated outside of this project.

Bibliography

1. (NTL 1). *Defense Meteorological Program Operational Linescan System, National Oceanic and Atmospheric Administration 2014.* Available online at: https://developers.google.com/earth-engine/datasets/catalog/NOAA_DMSP-OLS_NIGHTTIME_LIGHTS
2. (NTL 2). *Visible Infrared Imaging Radiometer Suite Day/Night Band, National Oceanic and Atmospheric Administration 2014.* Available online at: https://developers.google.com/earth-engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMSLCFG
3. (LS8). *LANDSAT 8 OLI/TIRS C1 Level-1, tiles LC08_CU_002011_20191012_20191022_C01_V01 and LC08_CU_003011_20191012_20191022_C01_V01, United States Geological Survey 2014.* Available online at: https://earthexplorer.usgs.gov/
4. TIGER. *United States TIGER/Line Shapefiles: Counties (and equivalent), United States Census Bureau 2014. Available online at: https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2014&layergroup=Counties+%28and+equivalent%29*
5. LANDFIRE. *Landscape Fire and Resource Management Planning Tools, United States Forest Service 2014.* Available online at: https://www.landfire.gov/vegetation.php

APPENDIX I - Data Download Methodology for Obtaining Data from Google Earth Engine

Two datasets that measure the “brightness”, measured in average digital number band values from nighttime lights were used in this analysis. The first is the DMSP OLS: Nighttime Lights Time Series Version 4 produced by the Defense Meteorological Program Operational Linescan System hosted by the National Oceanic and Atmospheric Administration (NOAA). This dataset is produced for the years 1992 - 2013 and uses measures the brightness of visible and near infrared light emission sources at night. The cell size for this dataset is 30 arc seconds. This dataset is available on the Google Earth Engine platform at:

https://developers.google.com/earth-engine/datasets/catalog/NOAA_DMSP-OLS_NIGHTTIME_LIGHTS

The second dataset is the VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1 from the same NOAA program. This dataset is available for the years 2014 - present and has a cell size of 15 arc seconds. It is available on the Google Earth Engine platform at:

https://developers. google.com/earth-engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMSLCFG

The second dataset is an updated sister product of the first dataset. The following code chunk was used to extract the 2013 DMPS OLS year data from the Google Earth Engine platform and export it as a geoTIFF to my Google Drive:

```
// Create a new variable that contains the nighttime lights (NTL) data from 1992 - 2013, found at:
// https://developers.google.com/earth-engine/datasets/catalog/NOAA_DMSP-OLS_NIGHTTIME_LIGHTS

// Filter the data to only contain observations from 2013 to match as closely with the 2014 study period
var OldData = ee.ImageCollection('NOAA/DMSP-OLS/NIGHTTIME_LIGHTS')
                .filter(ee.Filter.date('2013-01-01', '2013-12-31'));

// Select the display varaible to be 'avg_rad', which displays the avg. digital number values
var NTL_avgvis = OldData.select('avg_vis');

// Set the visibility range to display values from 0 - 50 in GEE
var NTL_old_Vis = {
  min: 0.1,
  max: 50.0,
};
```

```
// Add the layer 'NTL_avgvis' with the visibility parameters 'NTL_old_Vis' and call it 'NTL Old
2013' in GEE
Map.addLayer(NTL_avgvis, NTL_old_Vis, 'NTL Old 2013');

// Create a new variable called 'CA' that is a subset of the FeatureCollection 'table' that is the
// California state boundary
var CA = (table).filter(ee.Filter.eq("NAME", "California"));

// Create a subset of 'NTL_avgvis' that is clipped to 'CA'
var NTL_avgvis_CA = NTL_avgvis.filterBounds(CA);

// Modify 'NTL_avgvis_CA' such that all of the tiles in "NTL_avgvis_CA' are mosaicked
together
var NTL_avgvis_CA = NTL_avgvis_CA.mosaic();

// Modify 'NTL_avgvis_CA' so that it only contains data within the bounds of 'CA'
var NTL_avgvis_CA = NTL_avgvis_CA.clip(CA);

// Add the layer 'NTL_avgvis_CA' to the map, give it the visibility specs of 'NTL_old_Vis', and
call it 'NTL Old 2013' in GEE
Map.addLayer(NTL_avgvis_CA, NTL_old_Vis, 'NTL Old 2013');


// Export 'NTL_avgvis_CA' to my Drive, call it 'NTL_Old_2013_CA', The scale is 30 m, the
export region is 'CA', and
// the maximum number of pixels that can be exported is 1e10
Export.image.toDrive({
  image: NTL_avgvis_CA,
  description: 'NTL_Old_2013_CA',
  scale: 30,
  region: CA,
  maxPixels: 1e10
});
```

This is a global dataset, and when first loaded into the Google Earth Engine platform, it looks like this:

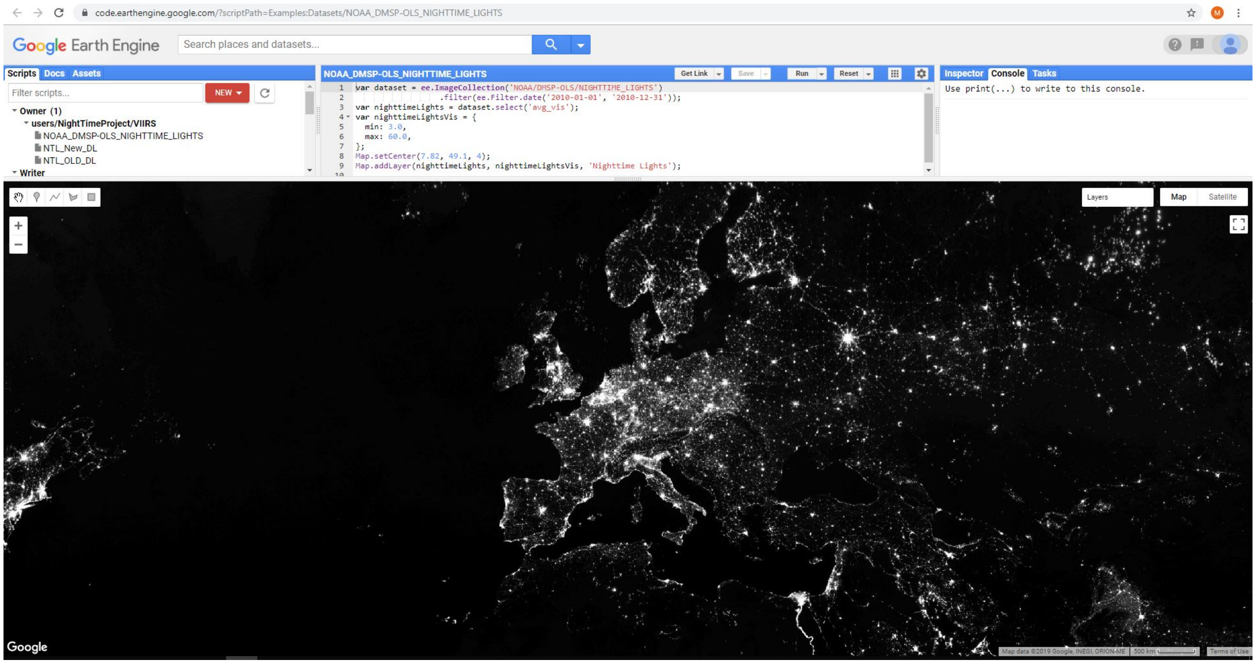


Figure 1: DMSP OLS: Nighttime Lights Time Series Version 4 before processing.

After I clipped the data to California and processed it for export, it looks like this:

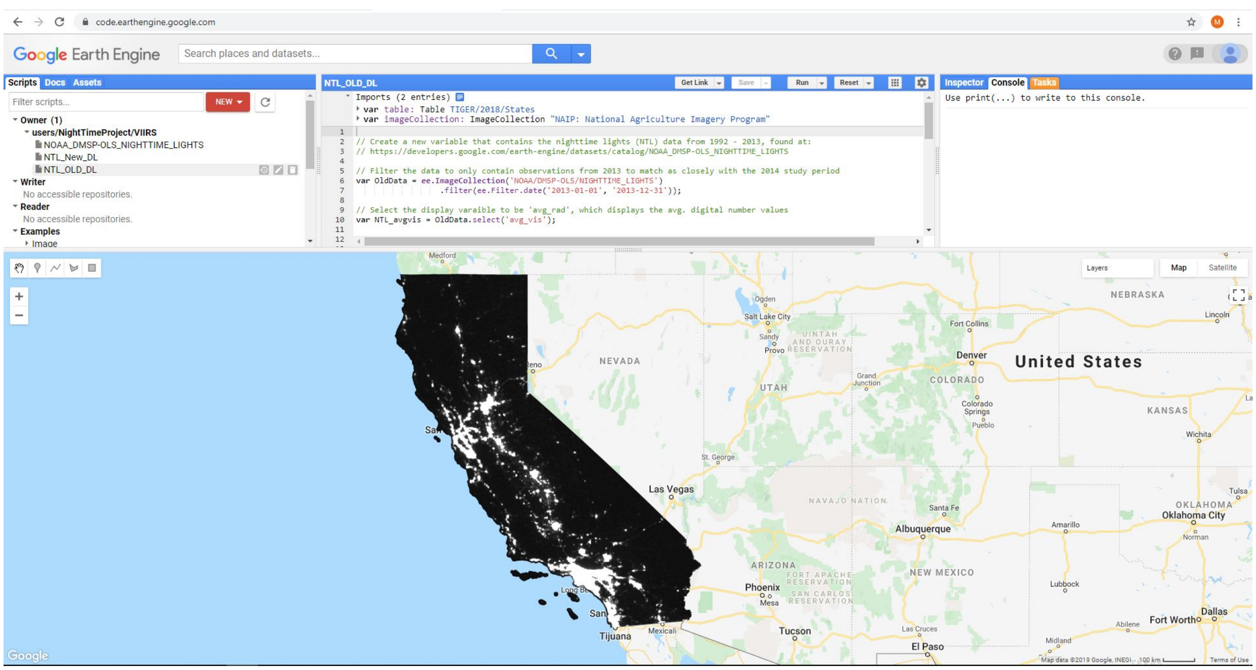


Figure 2: DMSP OLS: Nighttime Lights Time Series Version 4 after processing.

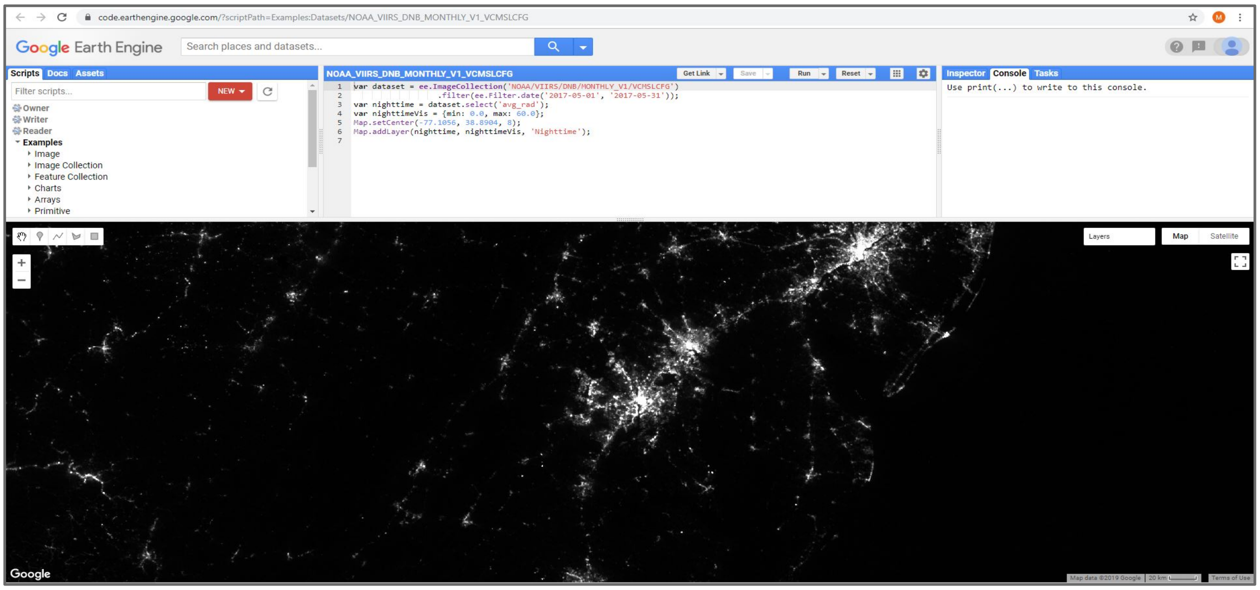
This process was repeated for the second dataset with the following code chunk:

```
// Create a new variable containing the nighttime lights (NTL) data from 2014 - present, found
at:
// https://developers.google.com/earth-
engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMSLCFG

// Filter the data to only contain observations from the 2014 study period
var NewData = ee.ImageCollection('NOAA/VIIRS/DNB/MONTHLY_V1/VCMSLCFG')
                  .filter(ee.Filter.date('2014-01-01', '2014-12-31'));

// Select the display variable to be 'avg_rad', which displays the avg. digital number values
var NTL_avgrad = NewData.select('avg_rad');

// Set the visibility range to display values from 0 - 50 in GEE
var NTL_new_Vis = {
  min: 0.0,
  max: 50.0,
};

// Add the layer 'NTL_avgrad' with the visibility parameters 'NTL_new_Vis' and call it 'NTL
New 2014' in GEE
Map.addLayer(NTL_avgrad, NTL_new_Vis, 'NTL New 2014');

// Create a new variable called 'CA' that is a subset of the FeatureCollection 'table' that is the
// California state boundary
var CA = (table).filter(ee.Filter.eq("NAME", "California"));

// Create a subset of 'NTL_avgrad' that is clipped to 'CA'
var NTL_avgrad_CA = NTL_avgrad.filterBounds(CA);

// Modify 'NTL_avgrad_CA' such that all of the tiles in "NTL_avgrad_CA' are mosaicked
together
var NTL_avgrad_CA = NTL_avgrad_CA.mosaic();

// Modify 'NTL_avgrad_CA' so that it only contains data within the bounds of 'CA'
var NTL_avgrad_CA = NTL_avgrad_CA.clip(CA);

// Add the layer 'NTL_avgrad_CA' to the map, give it the visibility specs of 'NTL_new_Vis', and
call it 'NTL New 2014' in GEE
Map.addLayer(NTL_avgrad_CA, NTL_new_Vis, 'NTL New 2014');


// Export 'NTL_avgrad_CA' to my Drive, call it 'NTL_New_2014_CA', The scale is 30 m, the
export region is 'CA', and
// the maximum number of pixels that can be exported is 1e10
Export.image.toDrive({
  image: NTL_avgrad_CA,
```

```
  description: 'NTL_New_2014_CA',
  scale: 30,
  region: CA,
  maxPixels: 1e10
});
```

Here are examples of the second dataset pre and post-processing

a)



b)

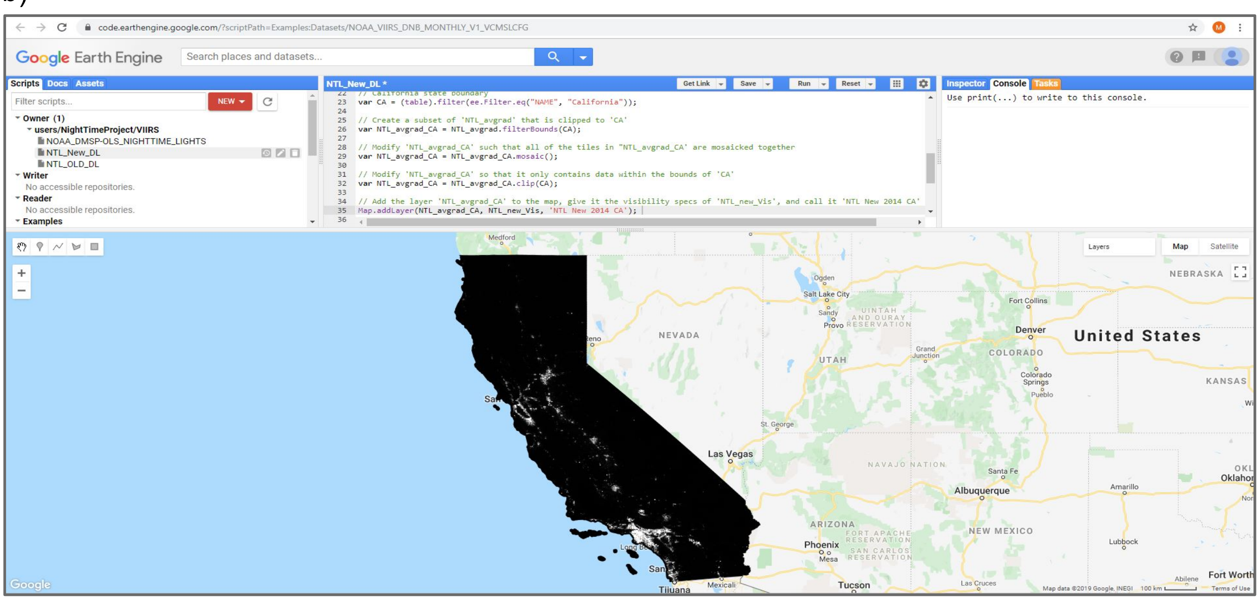


Figure 3. VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1; a) pre-processing and b) post-processing.

Both datasets were then downloaded to my local drive and stored in a file geodatabase where they could be accessed by ArcPro.