Hannah Siegel
December 12, 2024

Post-Fire Remote Sensing Analysis of Vegetation and Land Cover Transformation:

Using classification and NDVI to track trends in Eagle Creek's vegetation recovery

---

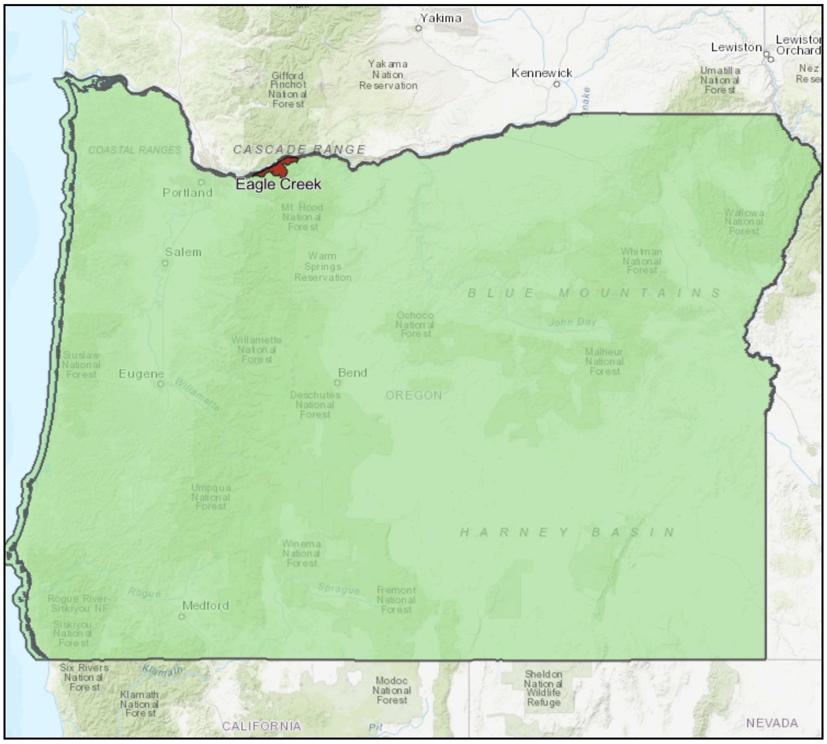
INTRODUCTION

The Eagle Creek Fire began on September 2, 2017, and burned nearly 50,000 acres across the Columbia River Gorge, causing widespread destruction to plant and animal habitats and significantly altering the beautiful landscape. Originally sparked by a firework thrown into vegetation during a dry season, the fire was rapidly spread by strong winds, burning forests on both sides of the Columbia in both Oregon and Washington. On November 30, the fire was declared 100 percent contained by the U.S. Forest Service (1), and despite its destruction, the affected areas have shown significant ecological recovery since then, demonstrating the Gorge's resilience and ability to heal.

This remote sensing study focuses on the fire's impact on vegetation loss and the processes of regrowth, offering insights into the Gorge's long-term ecological recovery and environmental changes. The driving question for my research was: How has vegetation health and vegetation land cover changed in the Eagle Creek study area over time, in response to the September 2017 fire event?

STUDY SITE AND CONTEXT

Below are some maps showing the size and location of the Eagle Creek burn area in the Columbia Gorge. The burn area (my study area) highlighted in red below covers approximately 309 square kilometers (2) and is situated on the Columbia River, right by the Bonneville Dam and just northwest of Mount Hood.



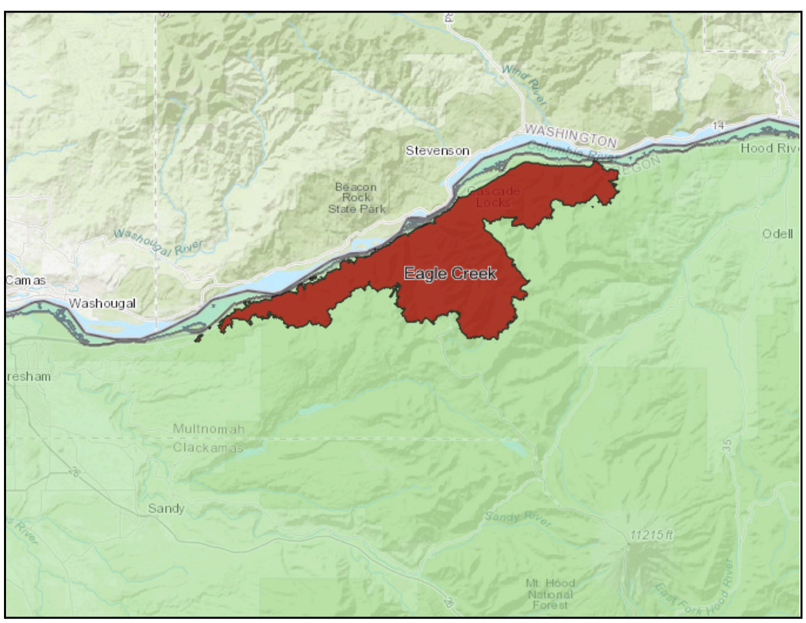
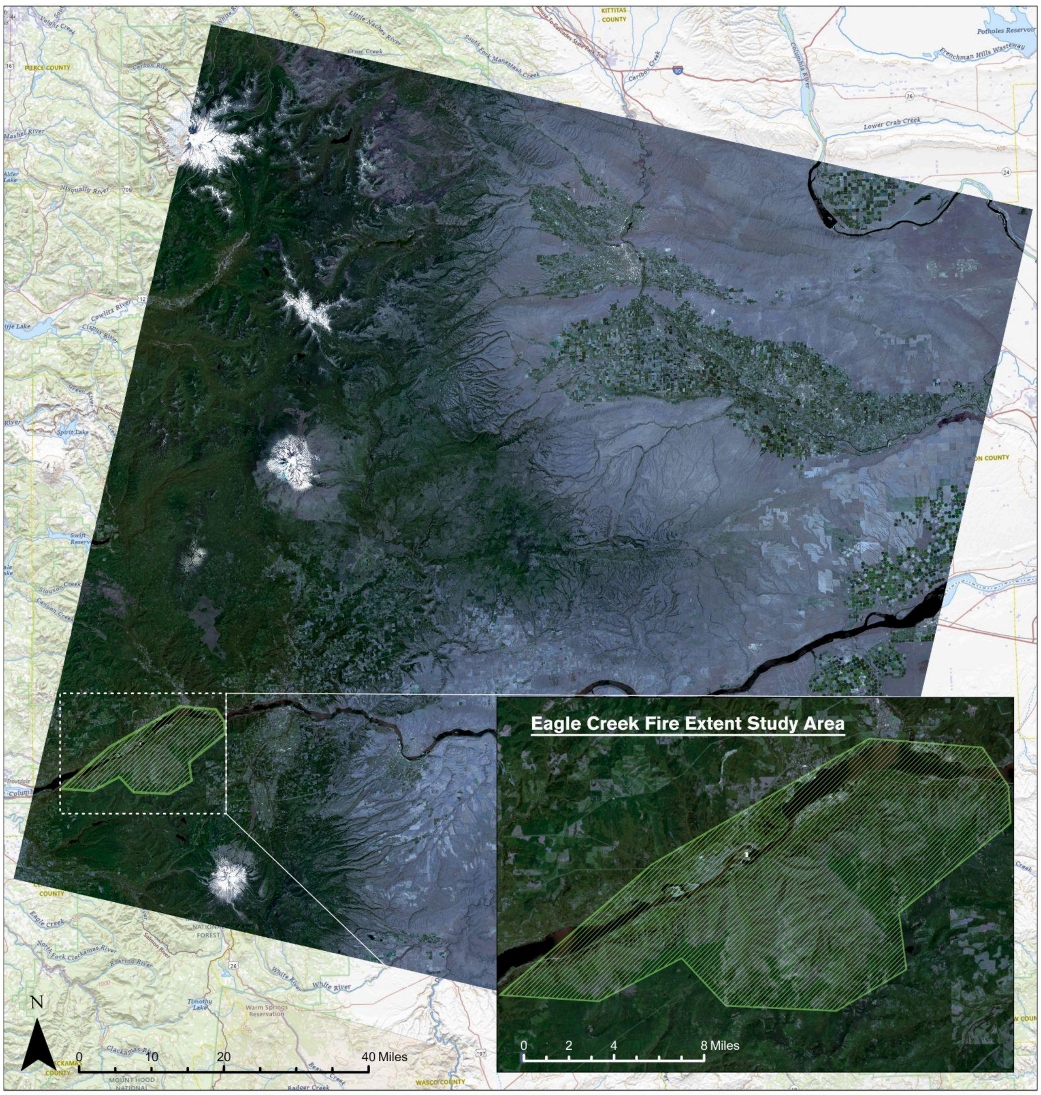


Figure 1. (Maps created in ArcGis Online)

METHODS

Data Acquisition:

To conduct my remote sensing analysis, I downloaded satellite imagery from USGS Earth Explorer (3), using Landsat 8's freely-available images of my region of interest. Below is a more detailed (true-color) picture of my study area in the context of the much larger, full-sized downloaded satellite image.



Satellite Imagery Extent,
Enhanced Eagle Creek Fire Study Area

Figure 2.

For my analysis, I downloaded Landsat 8 bands 1, 2, 3, 4, and 5 for the area, and I calculated an NDVI (Normalized Difference Vegetation Index) as well.

- Band 1 is useful for mapping coastal and aerosol studies (4), and I included it to better differentiate between water (the river in my study area) and vegetation/soil.
- Band 2 helps distinguish soil from vegetation, important for my land cover analysis, and also distinguishes deciduous from coniferous vegetation.
- Band 3 (green) emphasizes peak vegetation, which is useful for assessing plant health (useful for my study)
- Bands 4 and 5 (red and NIR) are also useful in terms of mapping vegetation, but I mainly downloaded them to create my NDVI "band".
- Normalized difference vegetation index (NDVI) is a standardized index that is used to quantify vegetation greenness, also known as relative biomass, and is useful for monitoring vegetation density and health (5).
    - It is calculated with Landsat 8's bands 4 and 5: NDVI = ((5 - 4) / (5 + 4))

I combined the bands 1, 2, and 3, along with NDVI, to create one multi-band raster to use for my remote-sensing land cover classification. These 4 inputs served as the best bands to monitor the changes in vegetation size and health across the region, over 3 separate years of growth, burning, and regrowth.

My goal was to investigate these vegetation changes over three years: June 2017 (before the fire), June 2018 (immediately after the fire), and June 2023 (some years after the fire, with more time for regrowth). I ensured that I downloaded the necessary bands from the same month/season each year, to maintain consistency in vegetation levels and provide a reliable comparison of vegetation changes over time. Below are some true-color composites of my study area in these three years for context (Figure 3), as well as the multi-band combination raster for each year (created as described above, to optimize for vegetation detection) (Figure 4).

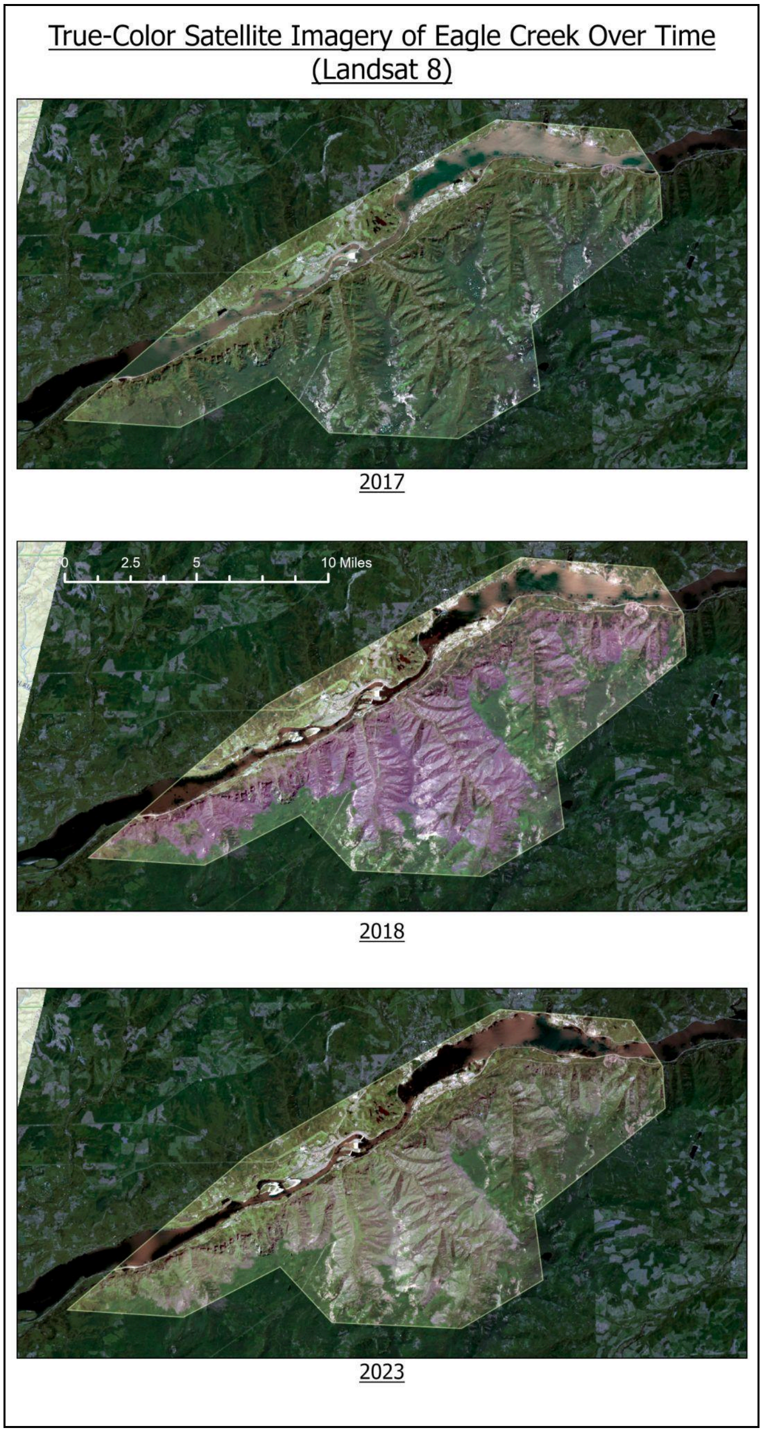


Figure 3.

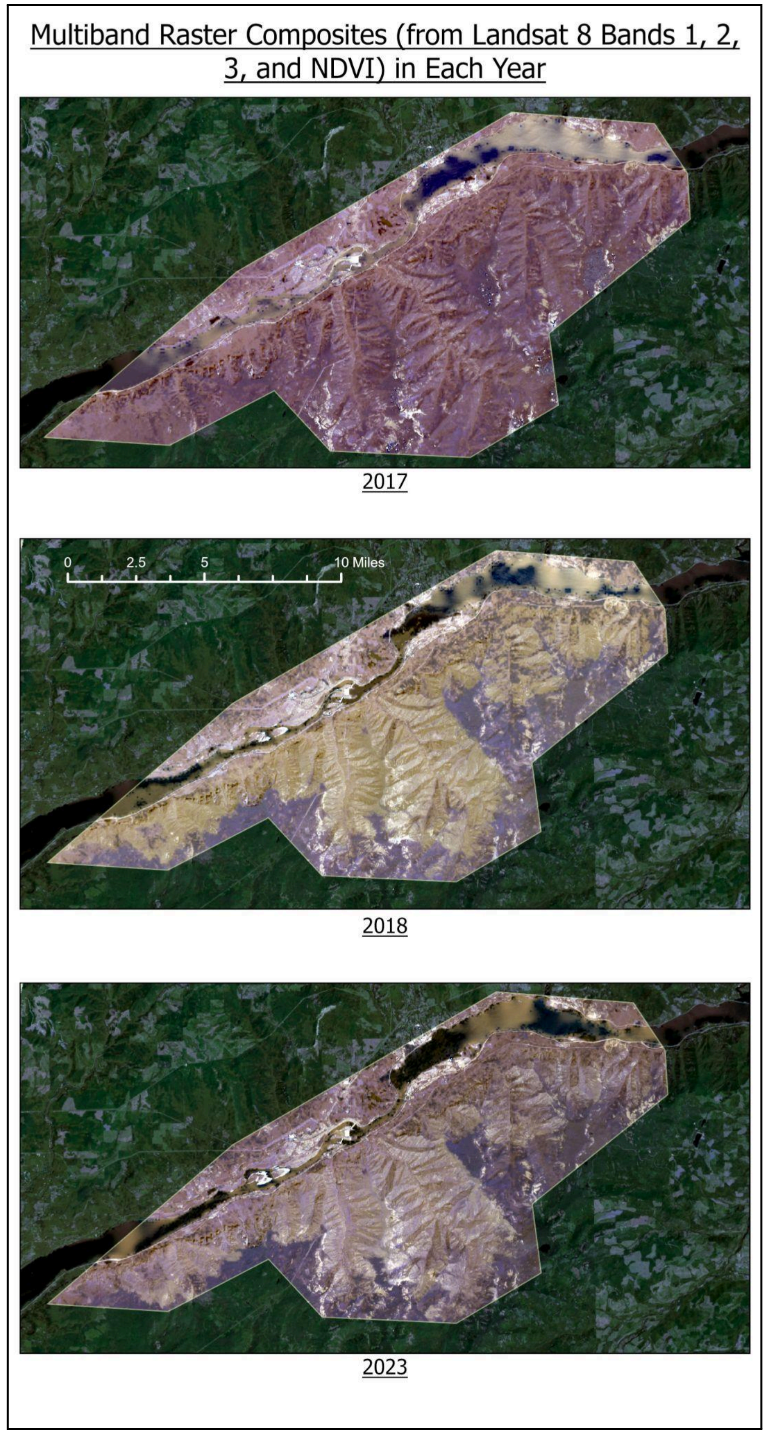


Figure 4.

Already, in both the true-color composites and the multiband composites we can see the stark differences between the years. In 2017 the area was largely forested, in 2018 (immediately after the fire) we can see the contrast between the burned area/soil and the surrounding vegetation, and in 2023 we can still see the burn area but it seems to be regrowing.

<u>Supervised Classification:</u>

To study the fire's impact on vegetation loss and regrowth, I performed a random forest supervised classification on my study area raster for each year. First, I generated 100 randomly placed points within the study area to serve as training data for the model. Then, using each year's true-color composite as ground truth, I classified the points into 5 land cover classes (for each year). The points' locations remained identical across the years, ensuring comparability, but their classifications varied because of significant changes in land cover year-to-year. This resulted in 3 sets of spatially identical random points, each with different classifications (example shown in Figure 5 below).

MY 5 CLASSES:

- Healthy Vegetation (value = 1)
- Burned Vegetation (value = 2)
- Soil (value = 3)
- Water (value = 4)
- Urban (value = 5)

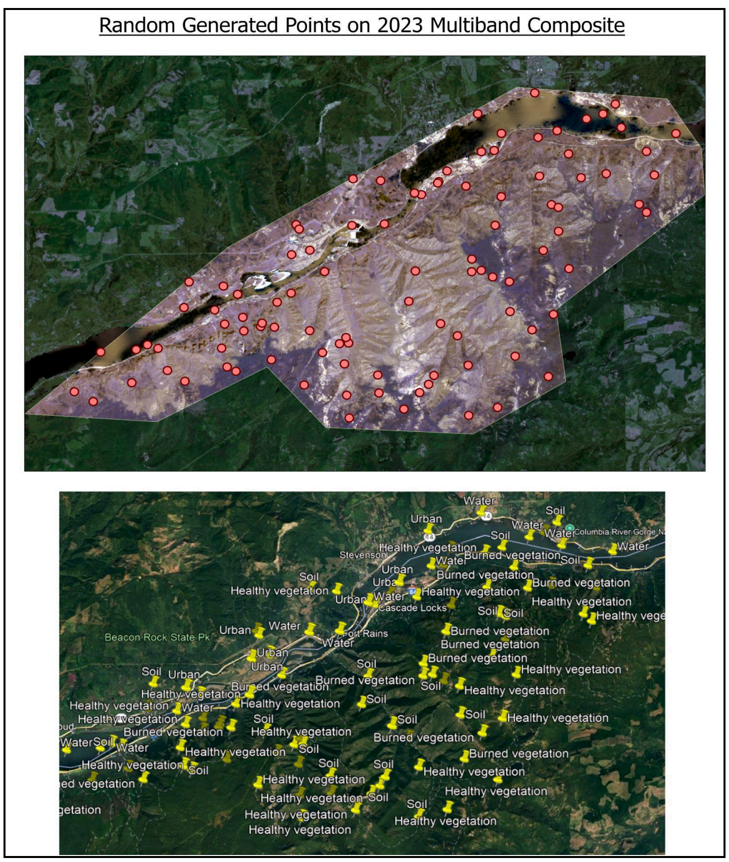


Figure 5.

Finally, using each *multi-band combination* raster I created above (to optimize vegetation detection), I trained a different random forest model on each year's data, using the ground truth point classifications and the raster's spectral signature at each point. These random forest models came from QGIS's [Dzetsaka plugin](), which leverages Python Scikit-learn for predictive modelling. I used each model to classify each year's land cover data, creating three deliverable raster maps (see Figure 6 in *Results* section).

Change Detection:

This study's goal was to explore vegetation loss and regrowth, which means understanding the *change* in vegetation health and land cover over this three-year time period. One way we can do this is by visually analyzing the differences in size, shape, and location of the 5 land cover classes created above. Another way is through assessing *numeric change* using NDVI to map vegetation greenness/health.

I executed this second option by using exclusively NDVI and no other bands to map my region of interest, so I could identify areas with the highest values (healthiest vegetation) and lowest values (unhealthy vegetation or bare soil) from each year. To analyze burn and regrowth trends, I calculated the year-to-year changes in NDVI by subtracting each year's NDVI raster from the subsequent year's raster, which created a raster product (one for the change from 2017-2018, one for 2018-2023). I then calculated the mean of each of these two raster products, to understand the average change. This approach accounts for spatial variability in changes between areas, and lets us know the average *numeric* change in vegetation cover/health between years.

(See Figure 8 in *Results* section for NDVI change between years.)

RESULTS

Classification Results:

When I used a random forest model to classify each year's land-cover data (from the *Supervised Classification* process above), I was able to create these three final classification maps, shown below (Figure 6). The confidence raster for each classification map is also shown (Figure 7).

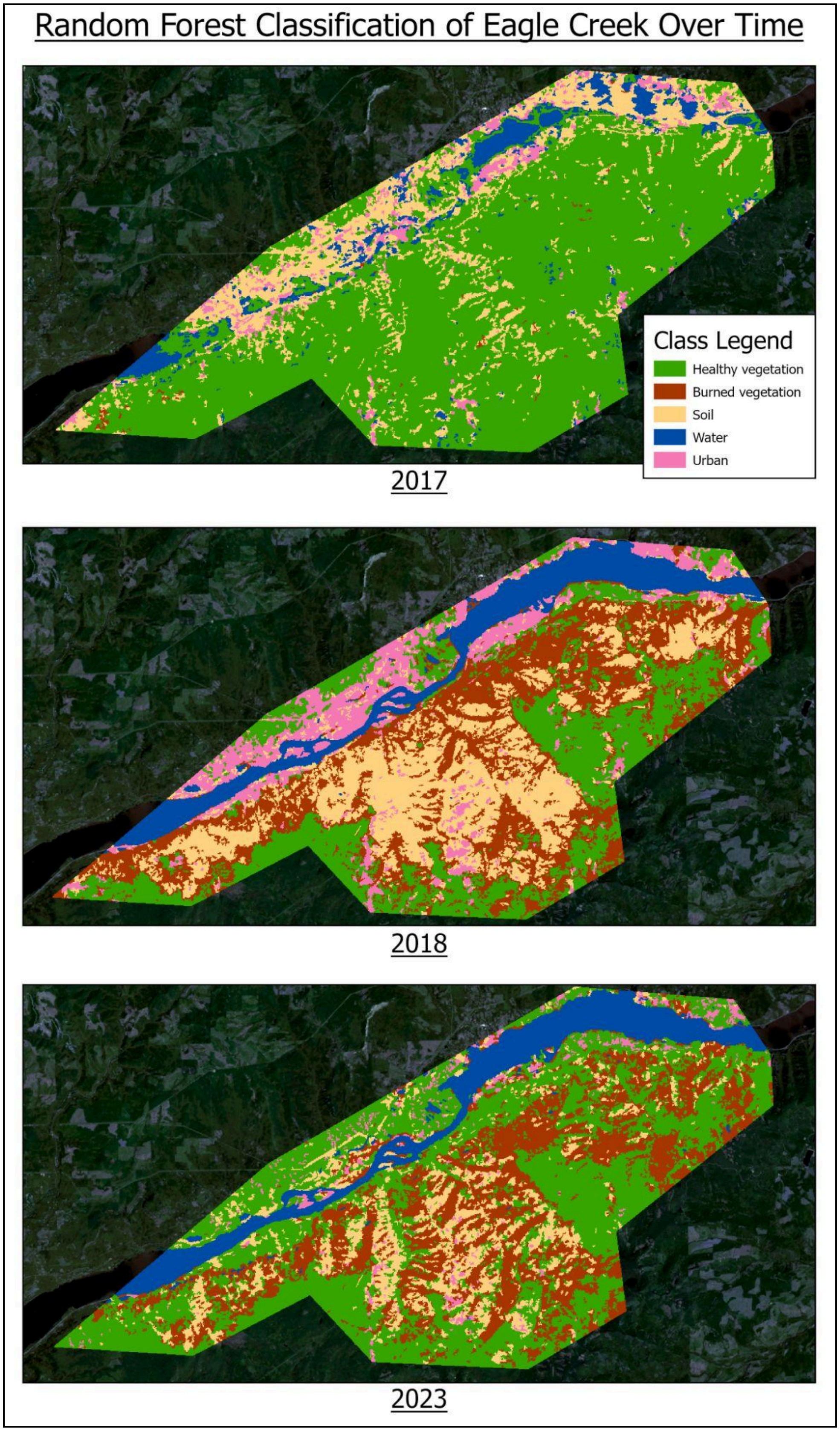


Figure 6. (Classification maps have been run through the *sieve* tool to smooth/minimize noise (6))

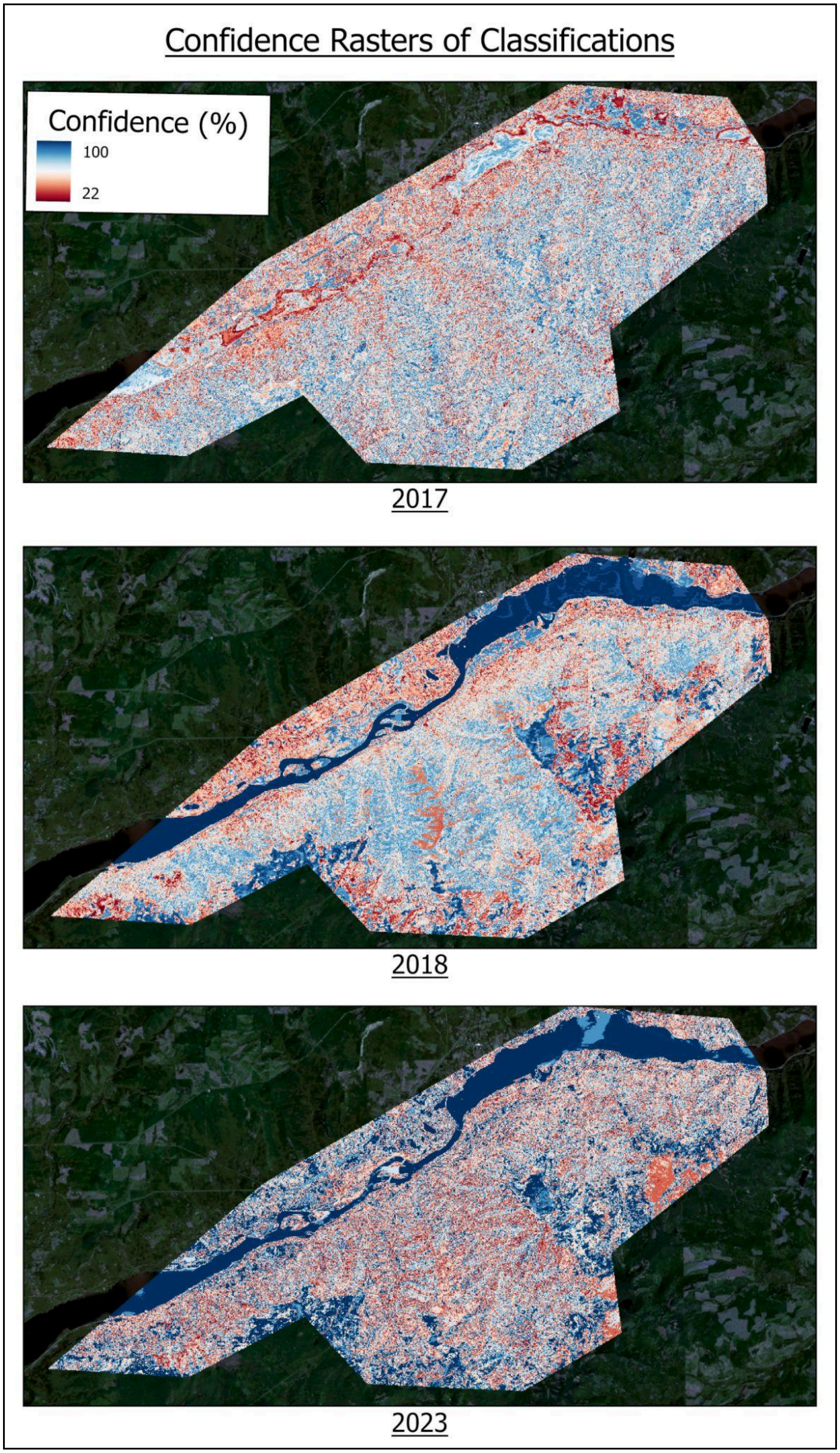

Confidence Rasters of Classifications
Confidence (%)
100
22
2017
2018
2023

Figure 7.

Overall, we can see that the fire had a devastating impact on vegetation land cover, resulting in over half the land being classified as either Burned Vegetation or Soil (after the trees had been burned) in 2018, compared to nearly 100% Healthy Vegetation in 2017. A few years later in 2023, we can see that the proportion of Healthy Vegetation is somewhat higher, but there are still large patches of Burned Vegetation and Soil. We can see from these land cover maps that areas nearer to the Columbia river and on the mountain tops/ridges seem to be recovering more quickly, already becoming reforested by 2023. However, reforestation on the sides of the mountains is taking longer, as large patches of these areas are still Soil and Burned Vegetation classes in 2023.

As for our confidence in these predictions, we can see from the confidence rasters (Figure 7) that the most accurate prediction of the model seems to be the shape of the Columbia river in 2018 and 2023 (although in 2017, there might not have been enough "Water" class points to train the model on). Apart from the river, when we look to the mountainous, forested part of the study area, the model had its highest overall confidence in 2018 when the hills were mostly burned vegetation and soil. As the vegetation started to regrow in 2023, the model had less confidence in its predictions of whether the mountainous areas would be Healthy Vegetation, Burned Vegetation, or Soil. This is likely because of the variance of reforestation patterns, as well as the "mosaic" burn pattern of the fire, meaning burned areas and regrowth areas might be changing at different or unpredictable rates, and the training points might have been unreliable for places still in the early stages of regrowth.

<u>NDVI Change Detection Results:</u>

Below are two maps of the change in NDVI (representing vegetation health change) as described in the *Methods – Change Detection* section. One map shows the change from 2017-2018 and another shows the change from 2018-2023 (Figure 8).

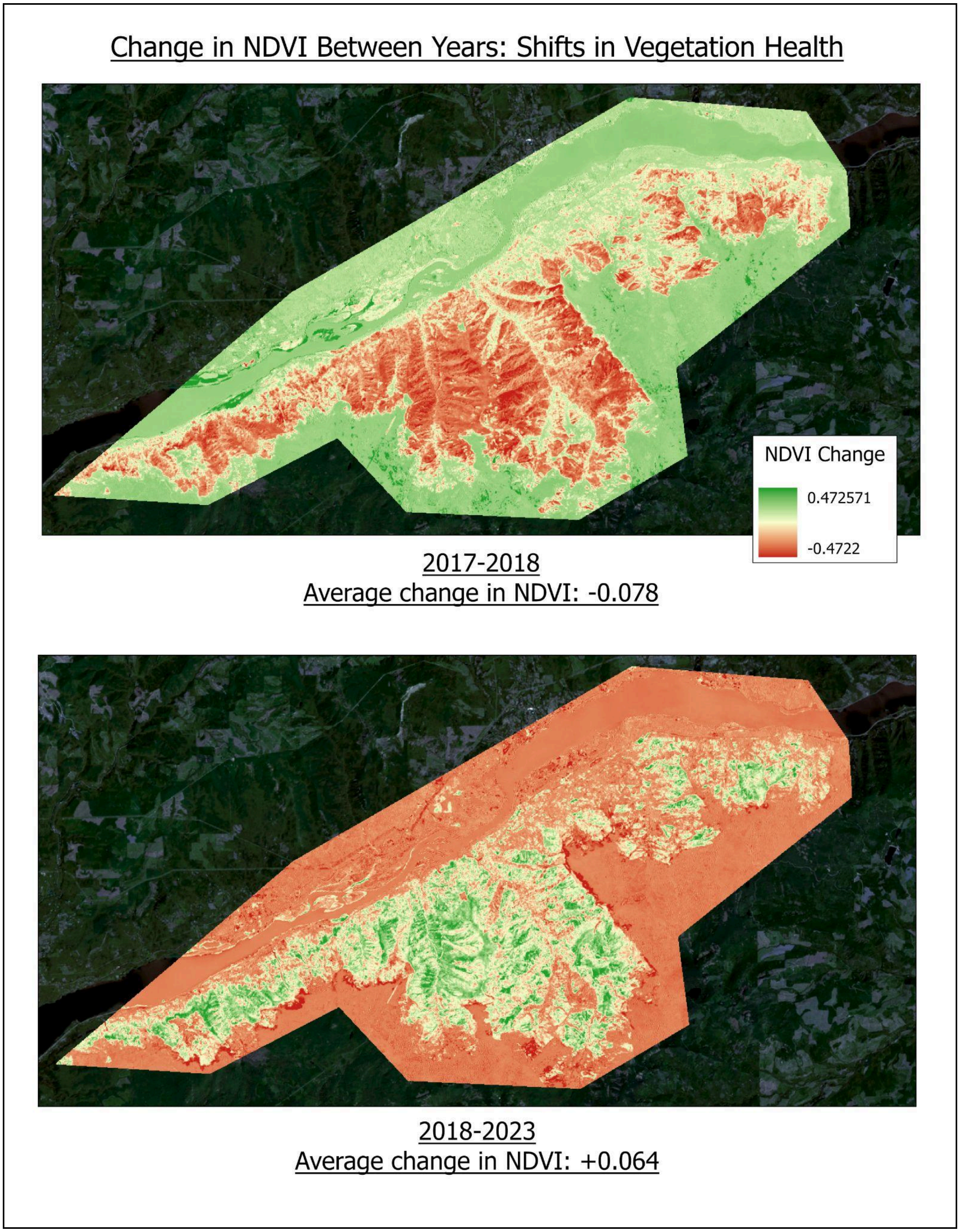

Change in NDVI Between Years: Shifts in Vegetation Health
NDVI Change
0.472571
-0.4722
2017-2018
Average change in NDVI: -0.078
2018-2023
Average change in NDVI: +0.064

Figure 8.

Since NDVI is one of the most useful indices for monitoring vegetation *density and health*, mapping change in NDVI provides us with these powerful visualizations of the shifts in vegetation during different time periods. The shift in NDVI from 2017-2018 reveals the destruction of vast amounts of vegetation in the fire, as nearly the entire mountainous area is red, indicating the most drastic decrease in NDVI of -0.47. However, from 2018-2023, we can see the regrowth in those hills, with green pixels indicating a strong positive change of 0.47 in small locations and yellow pixels indicating a slight positive change in the rest of the burned area.

A significant takeaway from these maps is that the red area (indicating a -0.47 decrease in NDVI) in the 2017-2018 map is much larger than the green area (indicating a +0.47 increase in NDVI) in the 2018-2023 map. This shows that the vegetation loss caused by the fire has not been fully offset or "undone" even after five years; for full regrowth, the green area in 2018-2023 would need to match the spatial extent of the red area in 2017-2018. Instead, much of the area remains in mid-range yellow, suggesting a slower recovery process.

Finally, looking at the average values of NDVI change from each map, we can see that the average change from 2017 to 2018 was -0.078, whereas the average change from 2018 to 2023 was only +0.064. This further reinforces our analysis that the regrowth in the second time period, although significant, has still not fully replenished the extent of the fire's destruction from the first time period. Our first time period is only one year long and our second (post-fire) time period is 5 years long, meaning that regrowing the forest has taken 5 times as long as it did to destroy it, and will likely take even longer to fully heal.

REFLECTION

Both the visual analysis of land cover classification across years and the numeric comparison of NDVI change demonstrate the Eagle Creek Fire's overall destruction and the Gorge's slow pace of regrowth and recovery. The random forest model identifies for us some areas that might see regrowth first, like the riverbanks and the ridges/tops of mountains. Our NDVI change detection maps reinforce this, indicating that although growth is slow, the most growth areas are indeed in the mountainous regions.

As a reflection on my remote sensing analysis process, we can understand how well our classification model performed by looking at the *confusion matrices* it provided (Fig. 9-11).

Fig. 9: 2017 confusion matrix

	predicted healthy	predicted burned	predicted soil	predicted water	predicted urban
actual healthy	6	2	2	0	0
actual burned	0	0	1	0	0
actual soil	2	0	1	0	0
actual water	0	0	0	3	1
actual urban	1	0	1	0	1

Notes:

- The most misclassified class seems to be burned vegetation, which makes sense because in the 2017 data most of the forest was, in fact, healthy vegetation and not burned.
- The confusion matrix indicates water as being accurately classified with no omission or commission error, however looking at the classification maps we see a lot of river area being incorrectly classified as forest.

Fig. 10: 2018 confusion matrix

	Predicted healthy	Predicted burned	Predicted soil	Predicted Water	Predicted urban
Actual healthy	0	0	0	0	0
Actual burned	2	4	3	0	0
Actual soil	0	3	1	0	1
Actual water	0	0	0	3	0
Actual urban	2	0	1	0	2

Notes:

- There seems to be some mixup between Burned vegetation and Soil, with each being predicted as the other. However, in our classification map they both serve a similar purpose of indicating areas affected by the fire.
- Areas predicted to be healthy forest were actually often burned forest, indicating that there may have been even *less* healthy vegetation in 2018 than what we see on the 2018 classification map.

Fig. 11: 2023 confusion matrix

	Predicted healthy	Predicted burned	Predicted soil	Predicted Water	Predicted urban
Actual healthy	5	1	2	0	0
Actual burned	0	2	3	0	1
Actual soil	2	1	0	0	1
Actual water	0	0	0	3	0
Actual urban	1	0	0	0	0

Notes:

- Areas predicted to be soil are actually often either burned or healthy forest, indicating that 5 years after the fire even, there isn't much bare ground; it is actually either burned forest or new growth.
- Urban areas aren't picked up very well by the classifier / are often missed during classification.

Overall, from both a visual assessment of our classification maps and from our confusion matrices, we can see that our classifier:

- Overclassified burned vegetation in 2017 (understandable, since there wasn't a major fire yet)
- Classified water very accurately in 2018 and 2023 but not in 2017
- Often confused burned vegetation and soil in 2018, although this didn't hinder our analysis
- Overpredicted urban areas in 2018 but underpredicted them in 2017 and 2023

These classification errors could be addressed with a more robust training dataset, with more points and a better distribution of spectral signatures within each class. However, by analyzing other data alongside our classification maps (like NDVI or our confusion matrices), we can get a better picture of the whole data and gain deeper insights. For example, the burned areas in 2018 may have been even *more* extensive than our model initially predicted, and areas near the Columbia river and on the mountain ridges appear to be experiencing the fastest regrowth.

A last suggestion for improving this project's scope could be changing the NDVI change detection section slightly, to encompass greater changes over many years. It would be interesting to calculate the average NDVI value for an entire raster of each individual year, and plot/graph the trends in average yearly NDVI over a larger range of time. This could reveal interesting patterns in the timeframe of regrowth.

CONCLUSION

The results of this study highlight the devastating impact of the Eagle Creek Fire on vegetation health and land cover. While some recovery has occurred over the subsequent five years, large portions of the landscape remain slow to regenerate. By combining supervised classification and NDVI change detection, this research has provided valuable insights into areas of the most significant change, building understanding of complex post-fire recovery patterns and the challenges of ecosystem restoration.

SOURCES

(1) https://gorgefriends.org/home/eagle-creek-fire--frequently-asked-questions.html
(2) https://portal.opentopography.org/raster?opentopoID=OTSDEM.112018.26910.1
(3) https://earthexplorer.usgs.gov/
(4) https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research
(5) https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/ndvi-function.htm
(6) https://daveparr.com/2024/03/01/machine-learning-classification-in-qgis/#:~:text=Vectorizing%20the%20Results-,Smoothing%20the%20Results,View%20%3E%20Panels%20%3E%20Processing%20Toolbox.