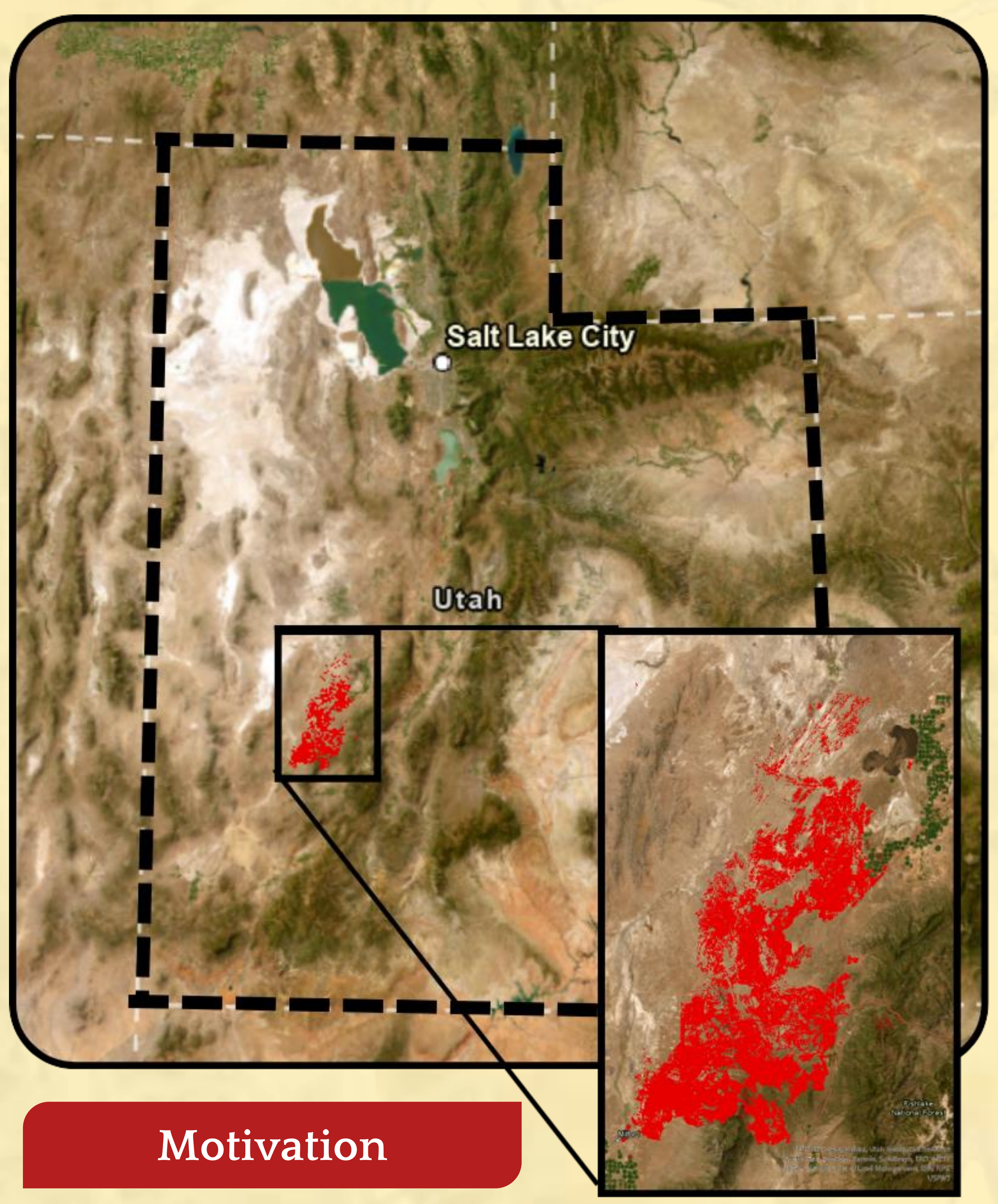
Pixels to Predictions: Mapping and Analyzing Vegetation Rebound After the 2007 Milford Flat Fire

Ellie Meyer

Tufts University, Civil and Environmental Engineering Department | Spatial Statistics Spring 2025

Abstract

The 2007 Milford Flat fire was the largest in Utah history, burning an area of 1,469 square km in southwestern Utah. **This study aims to explore the way vegetation bounced back in this area over the course of three summers post-fire**. **The main goal of this work is to identify regions where additional rehabilitation efforts would have been useful (areas that had more difficulty rebounding) and examine whether there are commonalities between these regions spatially**. This study uses spectral data from Landsat 4-5 TM Level-2. This is surface reflectance, which gets rid of atmospheric effects. First, the extent of the fire was analyzed using the difference in Normalized Burn Ratio (NBR), immediately pre and post fire. K-means clustering of a four year time-stacked image of NBR is done to identify trends within pixels of rebound and examine whether rebound differed spatially during these 3 years. Then, **geospatial data such as elevation, slope, and distance to water as well as spectral data, like burn severity and pre-fire NDVI, is used in an autologistic regression to predict areas of good rebound.**



Motivation

This area is within the Basin and Range province, which is characterized by arid climate, with alternating higher elevation and lower valleys. The flora are typically specialized high-desert plants: sagebrush, saltbush, scrub oak, Utah juniper, etc, with different species adapted to the various elevations and conditions.

This varied study area offers a unique opportunity at analyzing post-fire response by examining different geospatial variables. The goal of this project is to use spectral images immediately pre and post fire, as well as readily available geospatial data to predict where rebound will be best. This can help identify and pinpoint the locations to focus rehabilitation efforts.

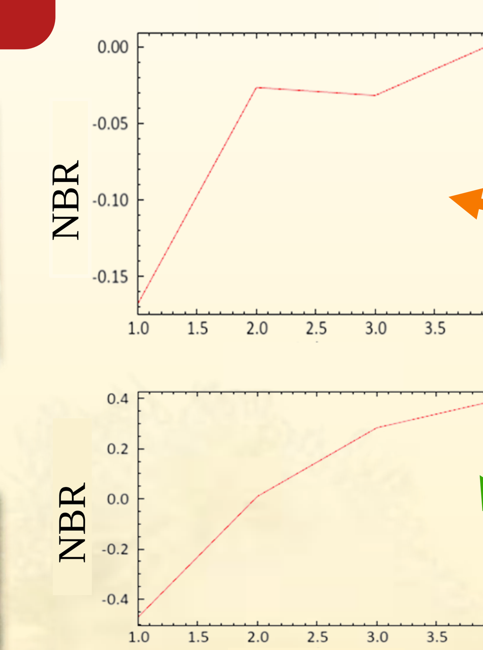
Photo from AP

K-Means Clustering to Identify Areas with Best Rebound

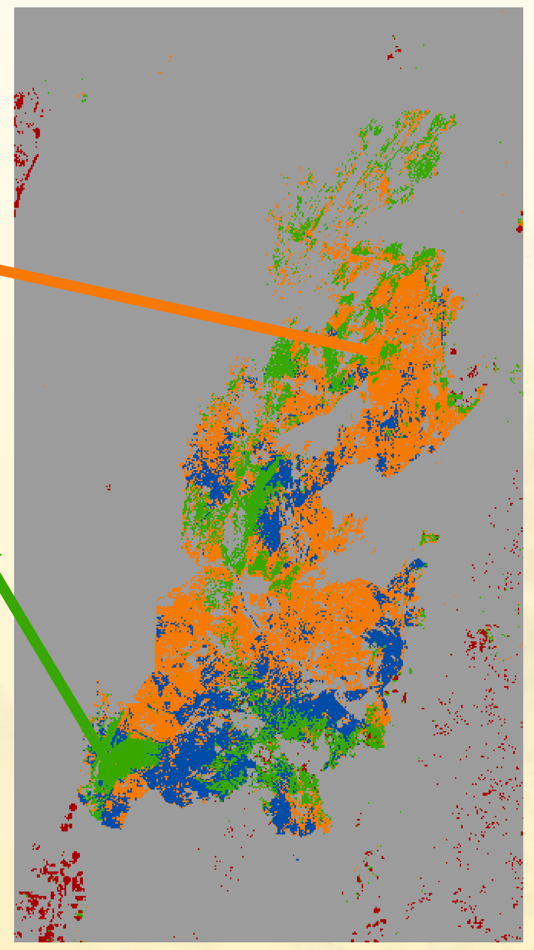
Analysis Steps:

1. Gathered Landsat 4-5 TM Level-2 spectral reflectance data from immediately before the fire and after the fire, as well as spectral data from the same season for three years following the Milford Flat fire, totaling 5 images.
2. Resized spectral images to fire extent
3. Calculated Normalized Burn Ratio (NBR) for each image.
   NBR = (Band 4 – Band 7) / (Band 4 + Band 7)
4. Created a time stack of each NBR following the fire (four timestamps) to proxy vegetation rebound over the years pixel by pixel.
5. Used K-Means clustering algorithm on the time-stacked image to examine trends in vegetation rebound
6. Randomly sampled 1000 pixels from each cluster and plotted them to identify which clusters rebounded the best.
7. Identified cluster 4 as unburned areas that bypassed the masks, clusters 1 and 3 as having the best vegetation rebound, and cluster 2 as having less successful recovery after 3 years.

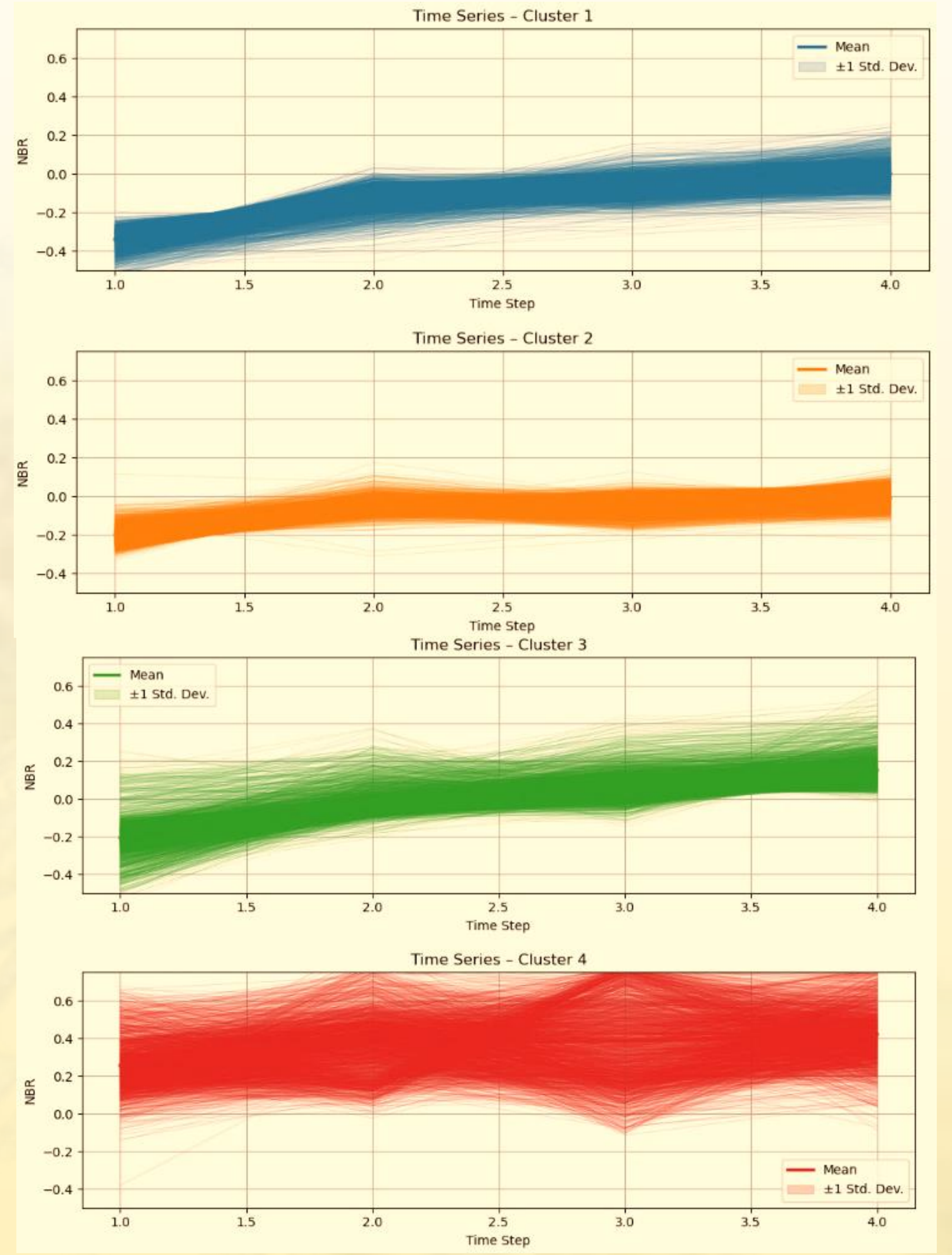
Visualization of raw spectral images from Landsat 4-5 collected from USGS EarthExplorer. All resizing, indices, and time stacking was done using ENVI 6.0 software



Example of two different pixel profiles of the immediate post-fire to three-year post-fire time-stacked NBR image showing different rebound patterns



K-means Cluster of time-stacked NBR image showing distinct spatial clusters with different rebound patterns. Clusters 1 and 3 had the best vegetation rebound

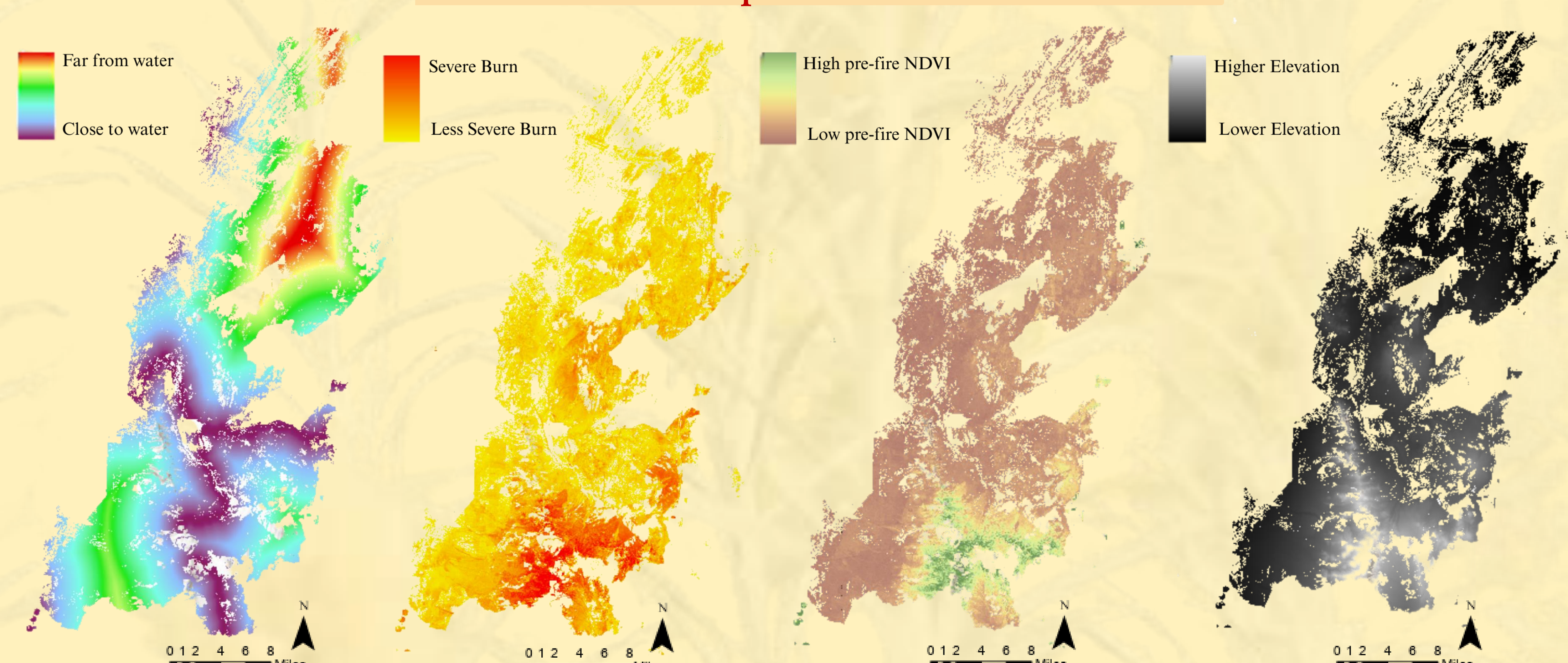


Logistic Regression and Autologistic Regression

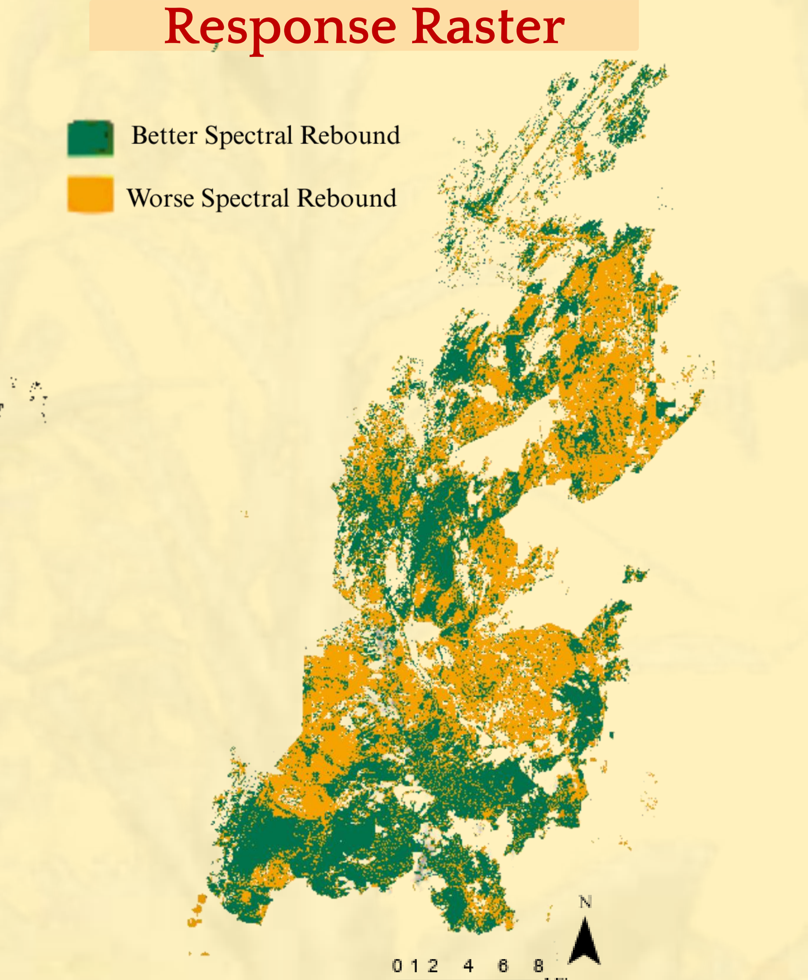
Analysis Steps:

1. After identifying the cluster with the best rebound, we converted the clustered map to a binary map where clusters 1 and 3 are coded as (1) and clusters 2 and 4 are coded as (0).
2. Gathered geospatial predictors and performed ESDA, finding that all of these are spatially correlated, but have no issues with multicollinearity
3. Use the geospatial predictors in a logistic regression with the binary raster as the response variable to get probabilities of good rebound by pixel Trained on 80% of data, tested on 20%. Predicted using the training data on the full raster after testing to get a raster of probabilities
4. Use these probabilities to create an autocovariate term to account for the spatial correlation of the response (Rooks 2nd order). Do autologistic regression with the autocovariate using the same 80/20 split
5. Compare the accuracy metrics
6. Train on full datasets and produce probability rasters to compare spatially

Input Rasters

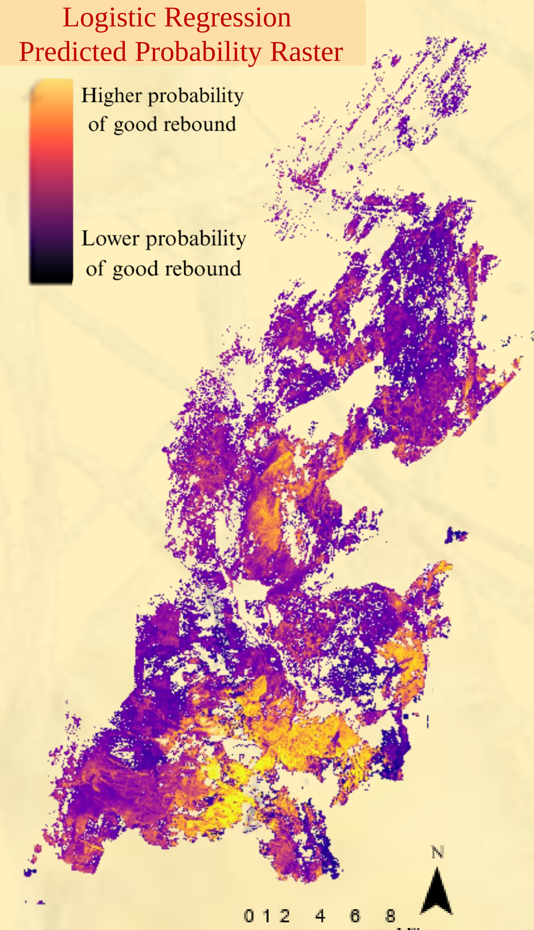


Response Raster



Results

Logistic Regression
Predicted Probability Raster



Logistic Regression

Model Performance on Test Pixels:

Autologistic Regression Summary	
Accuracy	0.739
AUC	0.756
Precision Class 1(Better rebound)	0.74
Recall Class 1(Better rebound)	0.42
Precision Class 0 (Worse rebound)	0.74
Recall Class 0 (Worse rebound)	0.92

Model summary after training on all pixels:

Autologistic Regression Summary	
Pseudo R-squared	0.1481
Autologistic Coefficients (All p-values significant at 0.005)	
Elevation	-0.0022
Distance to Water	-6.122
Pre-fire NDVI	–0.0153
Burn Severity	12.007

Autologistic Regression

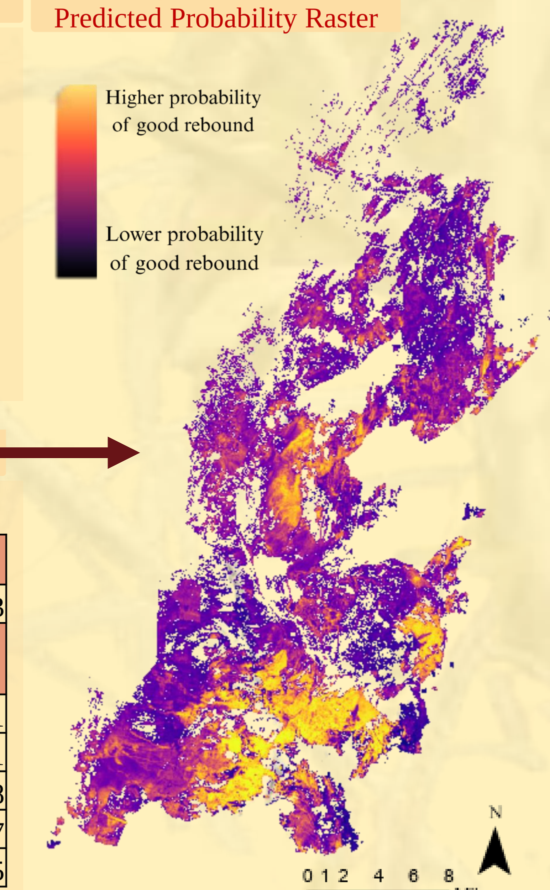
Model Performance on Test Pixels:

Autologistic Regression Summary	
Accuracy	0.742
AUC	0.756
Precision Class 1(Better rebound)	0.74
Recall Class 1(Better rebound)	0.44
Precision Class 0 (Worse rebound)	0.74
Recall Class 0 (Worse rebound)	0.91

Model summary after training on all pixels:

Autologistic Regression Summary	
Pseudo R-squared	0.1563
Autologistic Coefficients (All p-values significant at 0.005)	
Elevation	-0.0011
Distance to Water	-2.631
Pre-fire NDVI	–0.0048
Burn Severity	3.877
Autocovariate	0.3085

Autologistic Regression
Predicted Probability Raster



Conclusions:

- **The models are extremely similar, with the autologistic probability map smoothing out the logistic probability map.**
- **All the input variables were significant in predicting three-year spectral response with a significance level of .005 for both models**
- Between the logistic and autologistic model, the **coefficients kept the same signs but decreased in magnitude,** likely meaning this effect shifted to the autocovariate vs raster predictors.
- In the autologistic model, **the autocovariate was significant and the pseudo r-squared value increased,** meaning the spatial model improved prediction.
- The **computational expense of the autologistic model may not be worthwhile if the goal is a simple, usable model** to help local agencies pick areas for intervention after a fire.

References

Balch, J. K. B. A. Bradley, C. M. D'Antonio, J. Gómez-Dans. 2013. Introduced annual grass increases regional fire activity across the arid western USA (1980–2009). Global Change Biology 19: 173–183.

Bright, B.C., Hudak, A.T., Kennedy, R.E. et al. Examining post-fire vegetation recovery with Landsat time series analysis in three western North American forest types. fire ecol 15, 8 (2019). https://doi.org/10.1186/s42408-018-0021-9

DEM models courtesy of the Utah Geospatial Resource Center. Accessed April 2025

Distance to water pulled from 2023 ESRI data from Tufts M-drive, accessed April 2025

Duniway, M. C., E. Palmquist, and M. E. Miller. 2015. Evaluating rehabilitation efforts following the Milford Flat Fire: successes, failures, and controlling factors. Ecosphere 6(5):80. http://dx.doi.org/10.1890/ES14-00318.1

Khairani, Nabila & Sutoyo, Edi. (2020). Application of K-Means Clustering Algorithm for Determination of Fire-Prone Areas Utilizing Hotspots in West Kalimantan Province. International Journal of Advances in Data and Information Systems. 1. 9-16. 10.25008/ijadis.v1i1.13.

Landsat-5-5 images courtesy of the U.S. Geological Survey, accessed April 2025

Miller, Mark & Bowker, Matthew & Reynolds, Richard & Goldstein, Harland. (2012). Post-fire land treatments and wind erosion – Lessons from the Milford Flat Fire, UT, USA. Aeolian Research. 7. 29–44. 10.1016/j.aeolia.2012.04.001.

USDI BLM [United States Department of Interior, Bureau of Land Management]. 2007. Burned area emergency stabilization and rehabilitation handbook. BLM Handbook H-1742-1. USDI BLM, Washington, D.C., USA.

Xuexia Chen, James E. Vogelmann, Matthew Rollins, Donald Ohlen, Carl H. Key, Limin Yang, Chengquan Huang & Hua Shi (2011) Detecting post-fire burn severity and vegetation recovery using multitemporal remote sensing spectral indices and field-collected composite burn index data in a ponderosa pine forest, International Journal of Remote Sensing, 32:23, 7905-7927, DOI: 10.1080/01431161.2010.524678

Zerbe K, Cook T, Vulcano A. Multivariate spatial regressions help explain wildfire hot spot intensities in Washington, USA. Nat Hazards Res. 2024;4(2):288–294. doi:10.1016/j.nhres.2023.11.006