Machine Learning and Satellite Remote Sensing for Monitoring Invasive Cactus in Laikipia County, Kenya

SRIKANTNAG A N

GRADUATE RESEARCH ASSISTANT | ECOSYSTEM SERVICES AI LAB| UF/IFAS

MASTER OF SCIENCE | Department of Electrical and Computer Engineering

University of Florida

INVASIVE SCIENCE RESEARCH Symposium, May 6-7, 2024

Australian Prickly Pears Story

1. By the 1920s, prickly pear(Opuntia Stricta) infestations had rendered 58 million acres of Australian land useless for agriculture, prompting large-scale land abandonment due to the cacti's rapid and uncontrollable spread.
2. The prickly pear was initially introduced to Australia in 1788 to establish a cochineal dye industry, exploiting the insect that feeds on the cactus and produces a scarlet dye used in British military uniforms.
3. The biological control of prickly pear using the cactoblastis moth proved to be the world's most monumental success in pest plant repression, dramatically reducing the cacti population and reclaiming millions of hectares for agriculture by 1932.

https://www.daf.qld.gov.au/__data/assets/pdf_file/0014/55301/prickly-pear-story.pdf

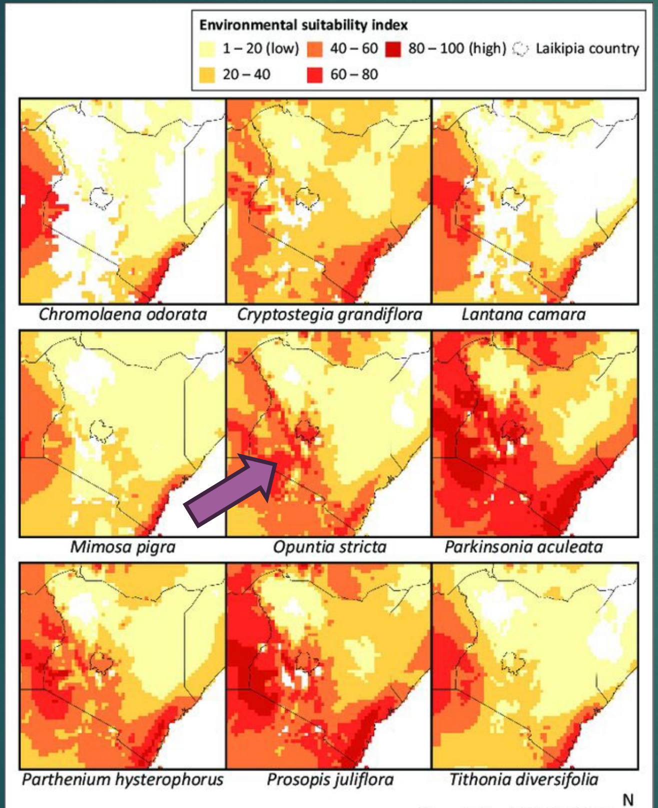
Opuntia Stricta

1. Opuntia stricta, a thorny cactus species in the Cactaceae family, has aggressively invaded the northern regions of Laikipia County, Kenya, significantly reducing prime grazing land and limiting access to grass for livestock.
2. The consumption of its spine-laden fruits by goats and other livestock during dry seasons leads to the attachment of glochids on their skin and gastrointestinal tracts, causing injuries, reduced milk production, and even death.

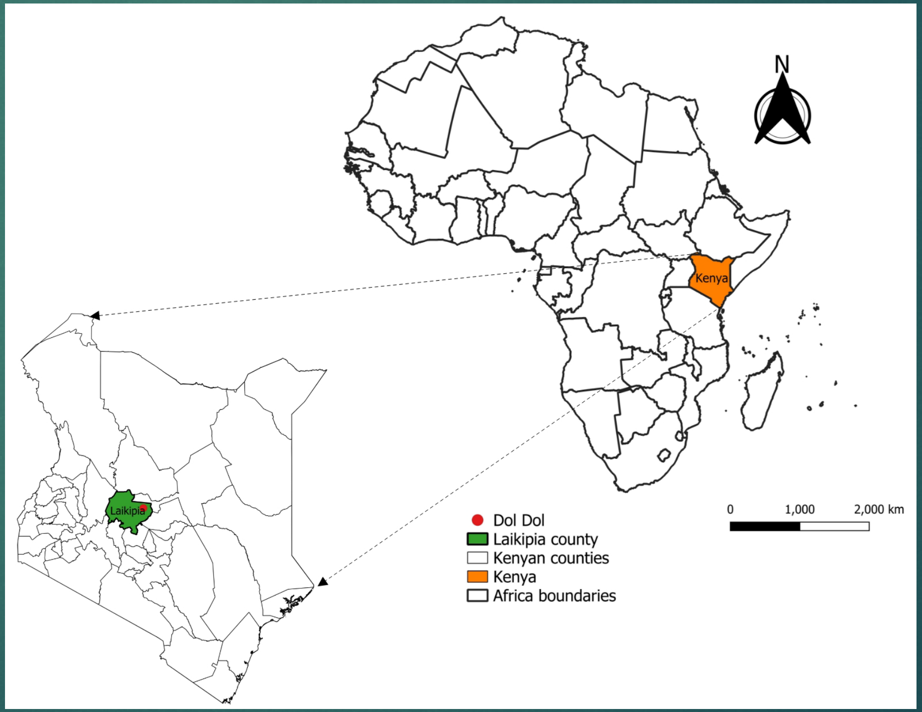
Opuntia stricta invasions in the arid and semi-arid lands of Kenya

- Extreme threats to biodiversity, ecosystem services and human well-being (Githae 2018)
  - Reduced foraging
  - Reduced livestock yields
  - Negatively impact wildlife habitats
  - Deplete soil and water resources and reduce the diversity of plants and animals

The Hotspots

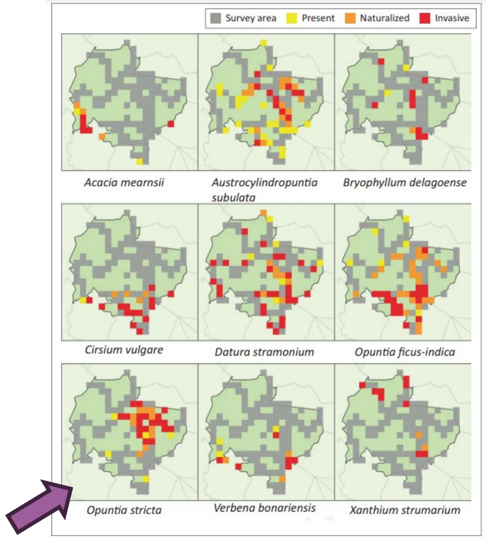


(Arne, et al. 2020)

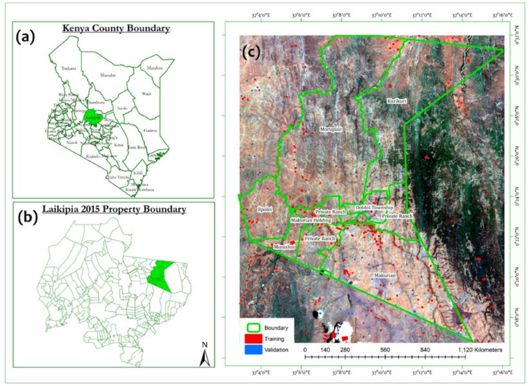


Previous Studies

- Broad-scale roadside surveys (Arne, et al. 2020)
- Demonstrated the ability of Sentinel-2 spectral properties to detect Opuntia stricta based on ensemble machine learning classifiers (Muthoka, J.M. 2021)



(Arne, et al. 2020)

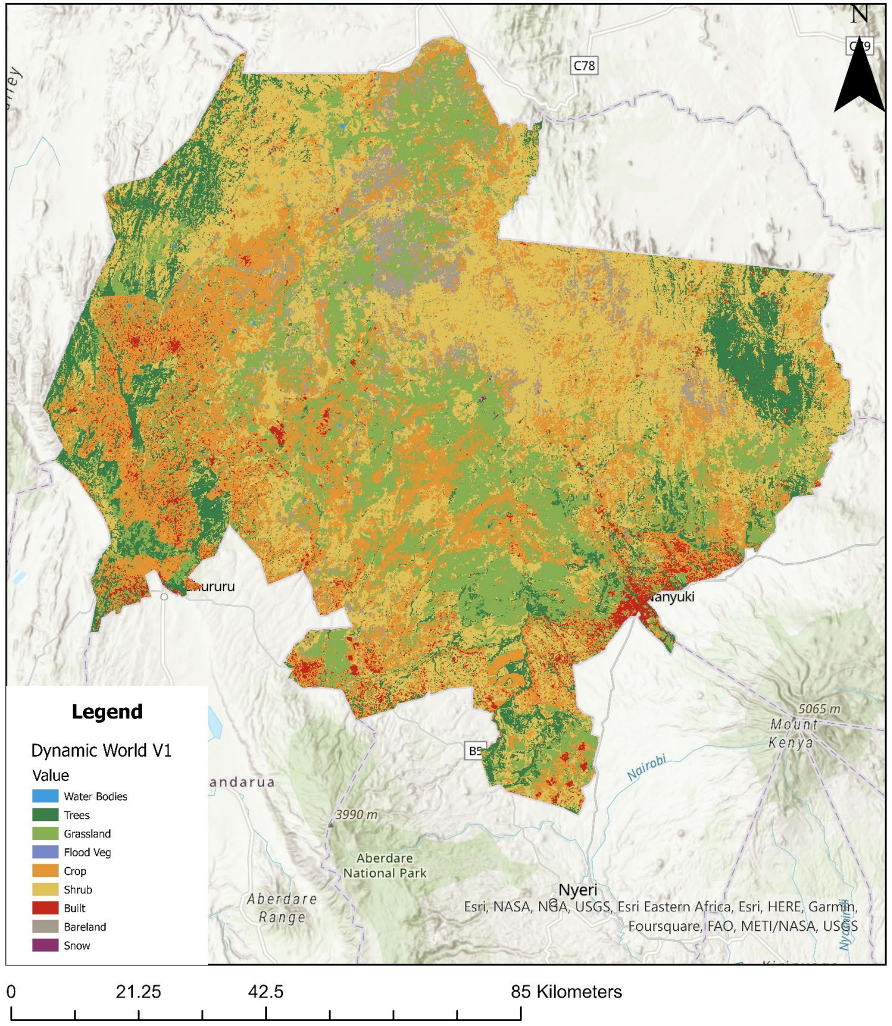


(Muthoka, J.M. 2021)

Research Gap

- Comprehensive information on the spatial distribution of invasion at fine-resolution across the Laikipia county.
- Accounting for geographical context, anthropogenic factors and temporal aspects of plant phenology in spatial prediction.
- Monitoring the spread of the species over time to properly manage Opuntia in the county.

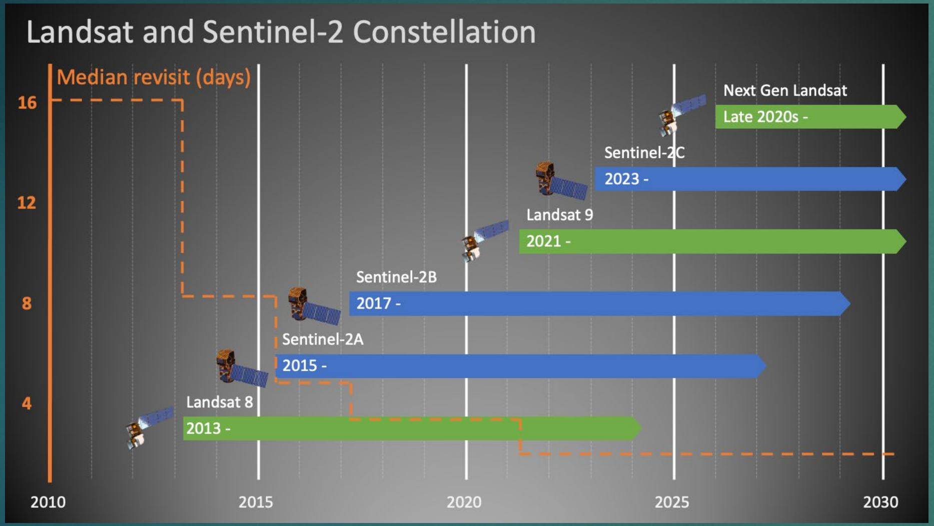
Dynamic LULC V1



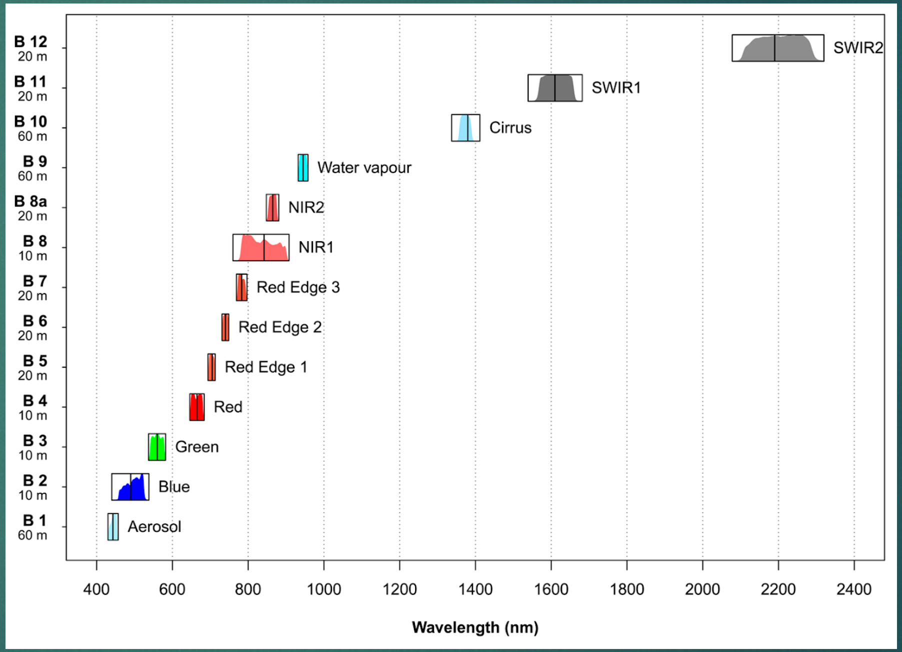
Research Objectives

1. Develop an accurate machine learning model for spatial prediction of O. stricta cactus across Laikipia County, Kenya at 10-m spatial resolution using freely available Sentinel 2 MIS data and open geospatial data.
2. To assess the importance of environmental and anthropogenic factors in spatial prediction of cactus using machine learning.
   - Climate, terrain, landscape proximity, population, building and livestock grazing density

Sentinel 2 MSI Data



https://www.earthdata.nasa.gov/esds/harmonized-landsat-sentinel-2



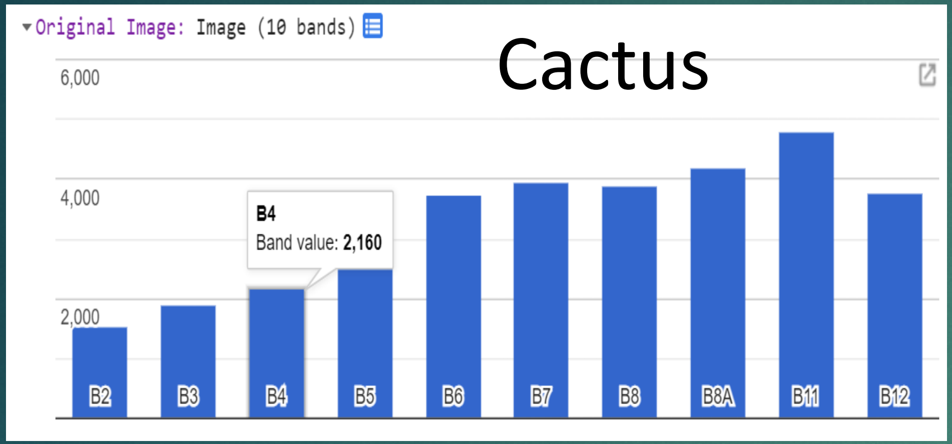
Multispectral instrument (MSI) on board Sentinel-2

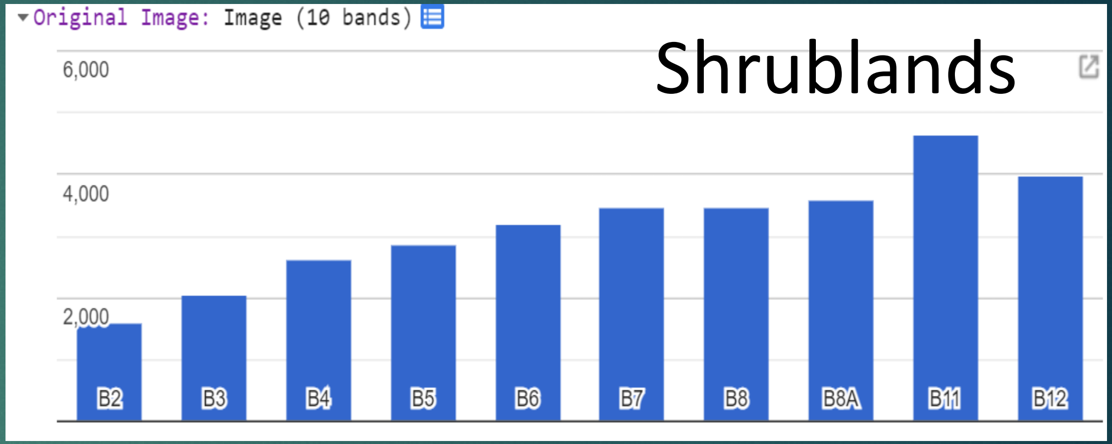
(Immitzer, M., et al 2016: https://www.mdpi.com/128566)

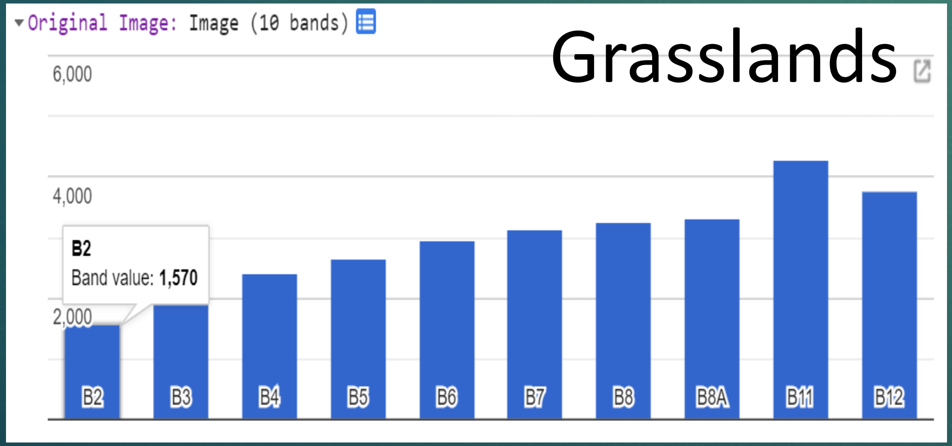
Sentinel 2 MSI Data Preprocessing

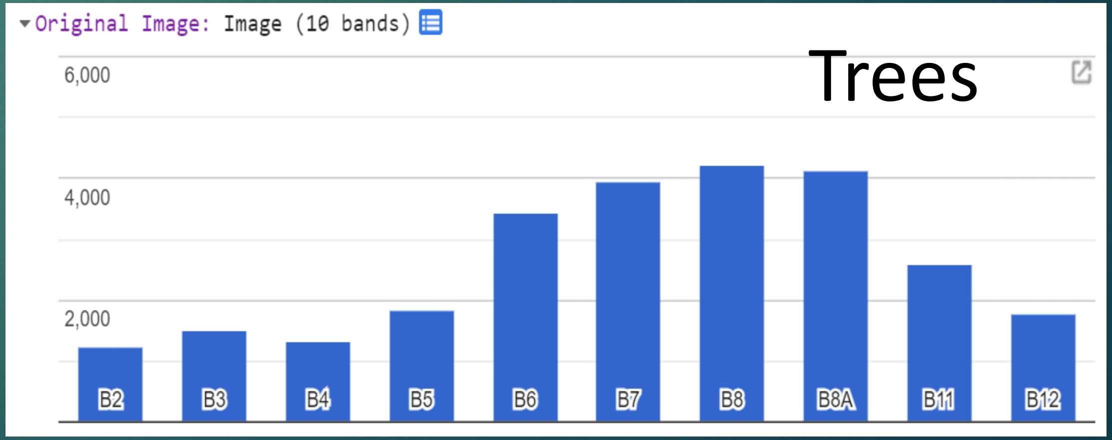
- Retrieve Monthly Images from Google Earth Engine
- Cloud Coverage Optimum Criteria: < 10% ~ 30%
  - Only available for 10 months from Sept 2022 to Aug 2023
- Clouding Masking and Illumination Correction
  - Converted clouds and shadow pixels to No Data, ensured consistent lighting across datasets
- Temporal Stacking Criteria to generate 10 monthly images:
  - Quality Mosaic: images from different dates within the 1-month window that have the lowest cloud percentage were spatially mosaiced together to cover the whole county

Spectral Analysis for Pixel Annotation







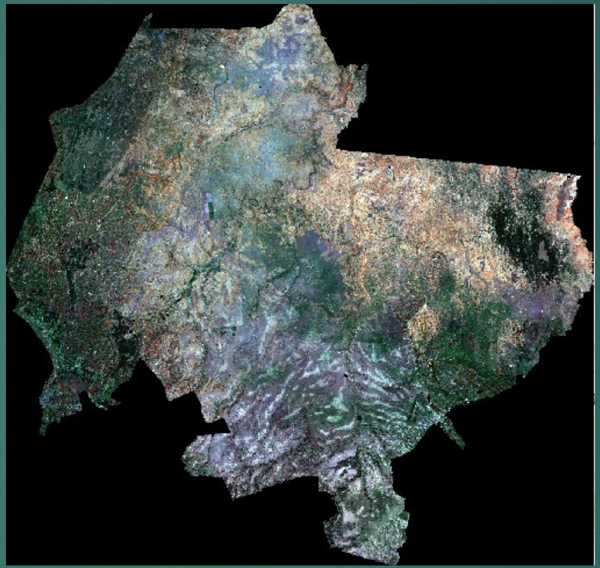


Training Data Annotation

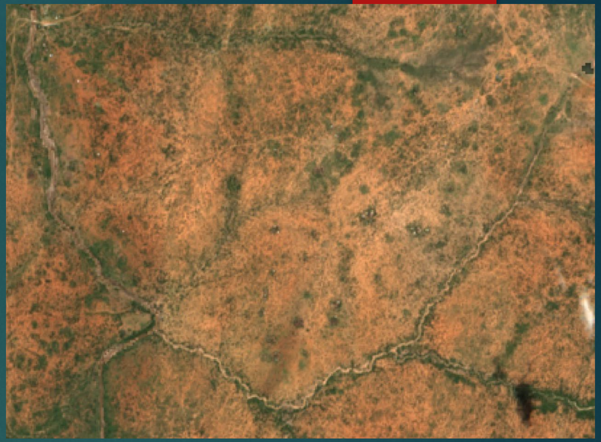
Person 1

Cross-Verification

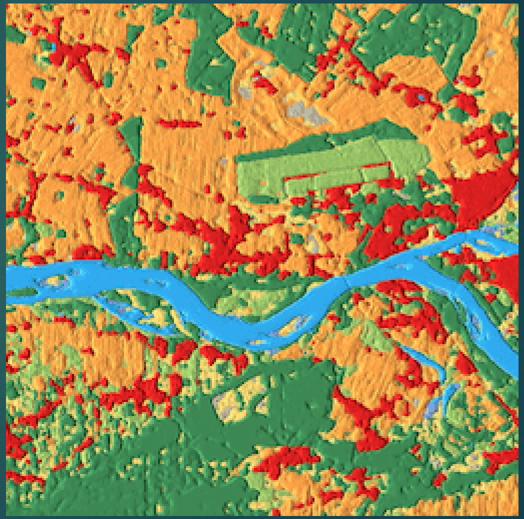
Person 2



Best-pixel mosaic from the Sentinel-2 L2A (Surface Reflectance) collection during the period from **'2023-03-01' to '2023-06-17'**, using the quality mosaic method



Google High Resolution Image

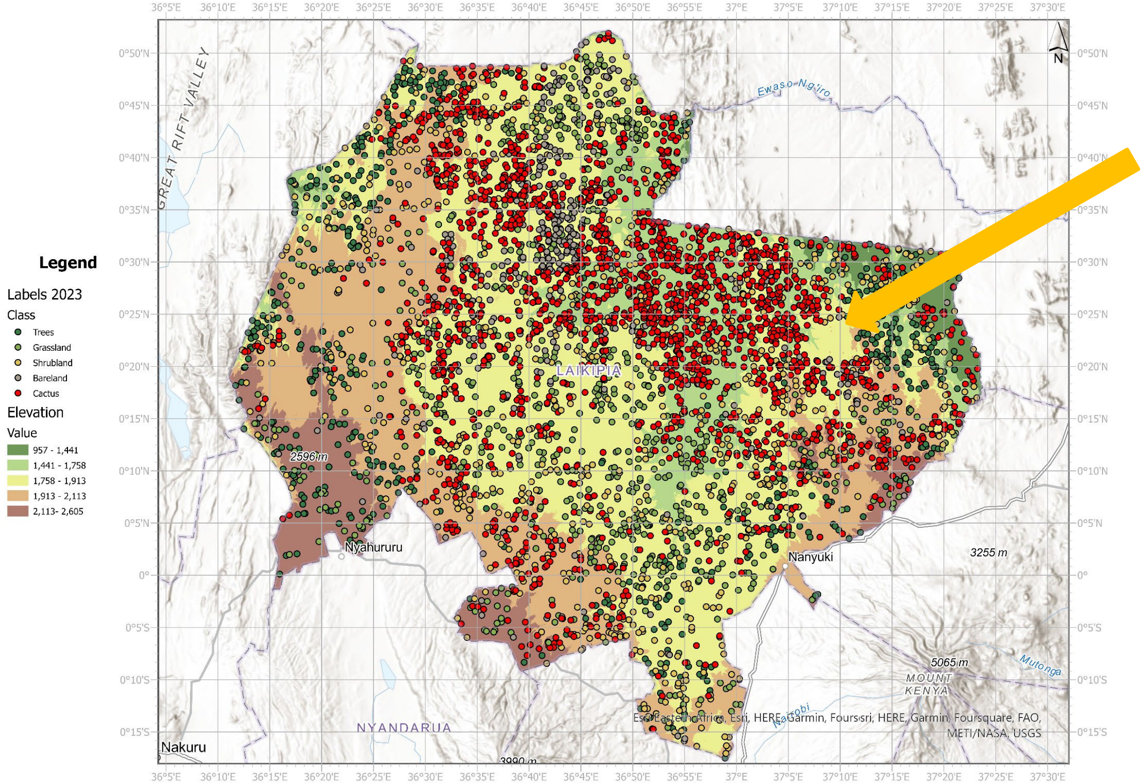


Google Dynamic World LULC

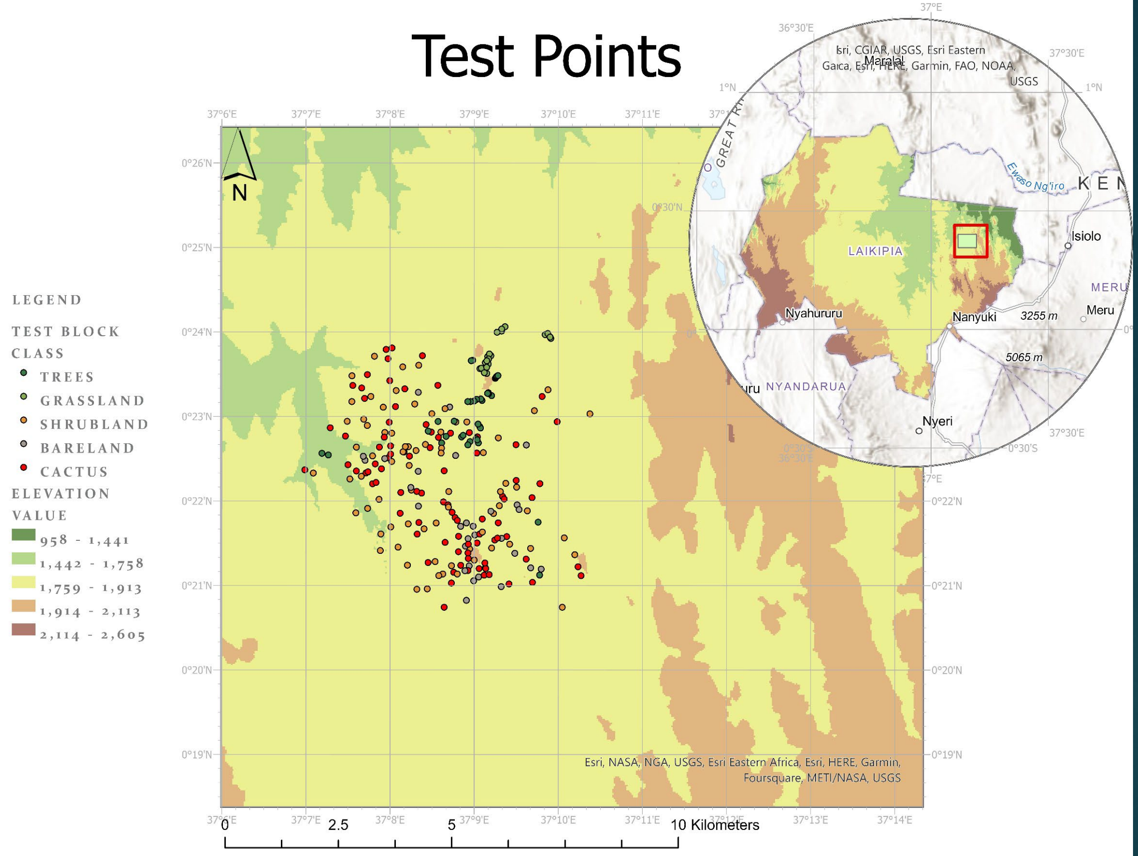
Table 1. Number of Labels by Land Cover Class

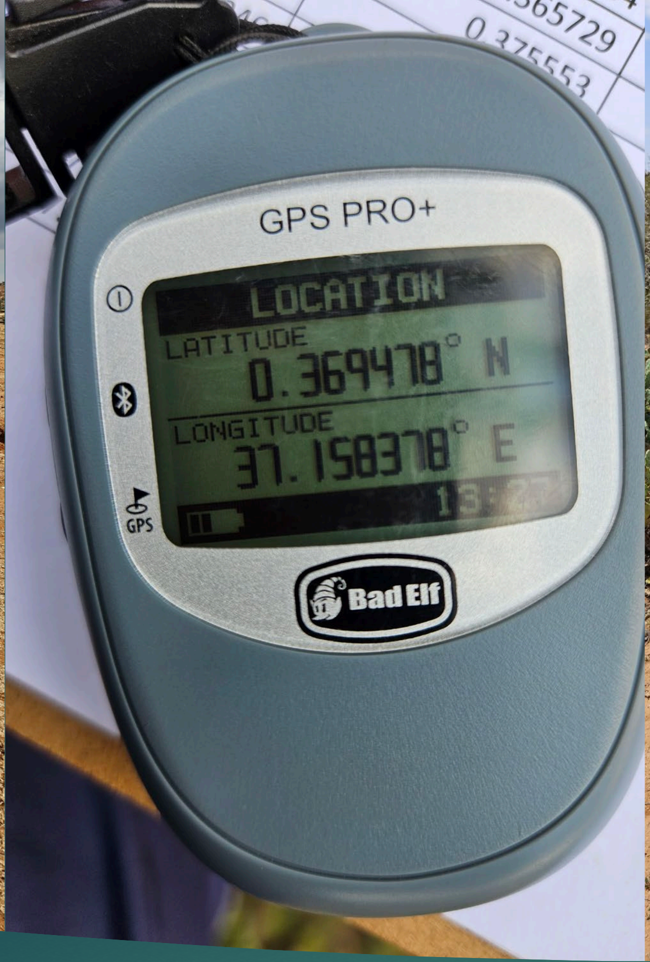
Classes	Class ID	Description	Samples	
			Training	Validation
Trees	1	High-density woody vegetation	470	117
Grassland	2	Low lying vegetation	686	171
Shrubland	5	Open and high-density low vegetation	766	192
Bare Soils	7	Dry exposed soils	355	89
Cactus	9	Opuntia stricta	1636	409
Total			3913	978
				4891

Data Annotation



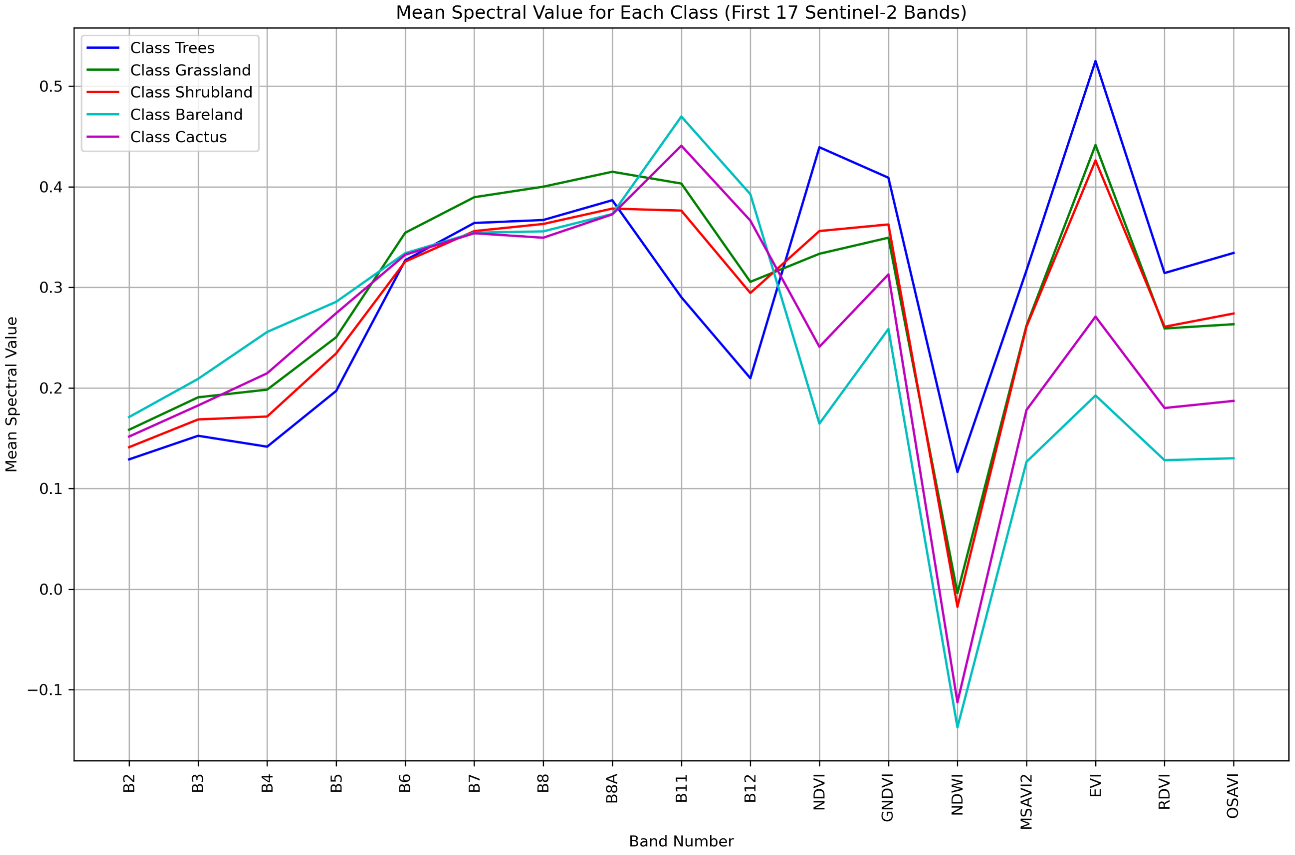
Test Points



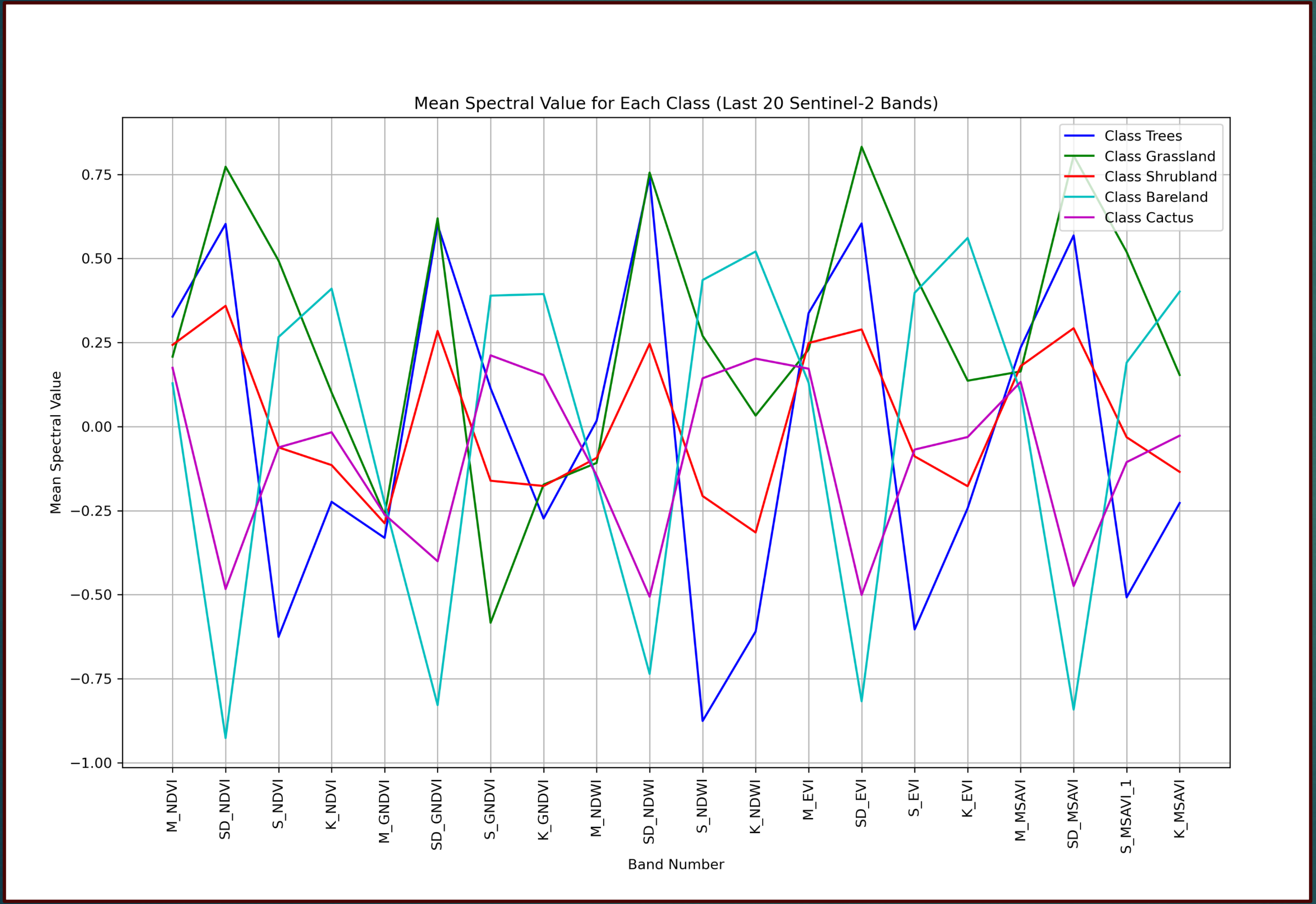



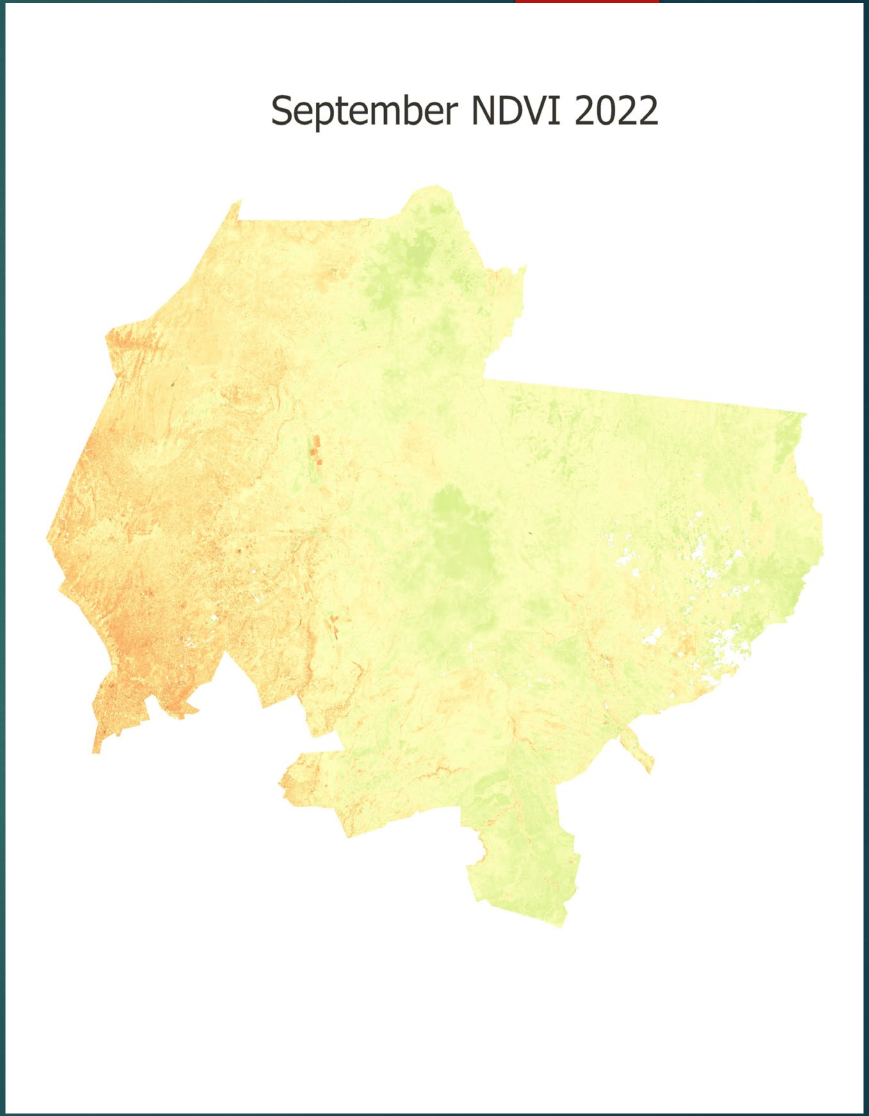

Field Observations

Spectral Profile of S2 bands by LUC class



Spectral Profile of Time Series VIs by LUC class





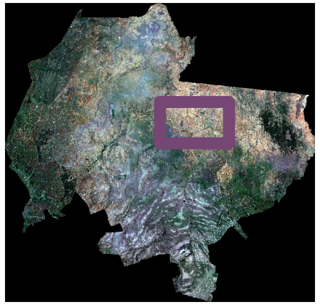
Environmental and Socioeconomic Data

CLIMATE, TOPOGRAPHY, HYDROLOGY, LANDSCAPE PROXIMITY, POPULATION AND GRAZING ACTIVITIES

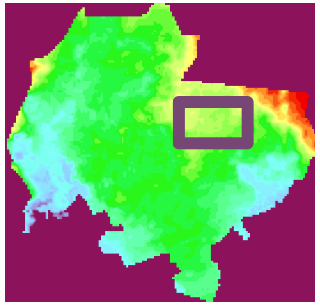
Environmental and Socioeconomic Data

- Time-variant Data:
  - Monthly Climate Data: Daylight and Night Land Surface Temperature, Precipitation and Evapotranspiration
- Time-invariant Data:
  - Terrain: DEM, Slope, Aspect, TPI
  - Landscape Proximity: Distance to rivers, roads, protected areas, wetlands, and major waterbodies
  - Population and Grazing Intensity: Population density, human settlement density, and livestock density

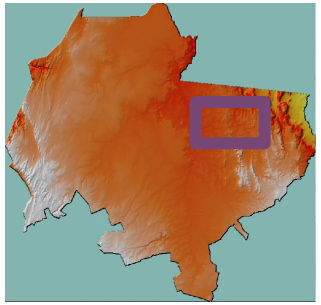
- Allows a GeoAI model to extract rich spatio-contextual information
- To gain stronger confidence in high-precision object detection and recognition



Remote Sensing Images



Climate Maps

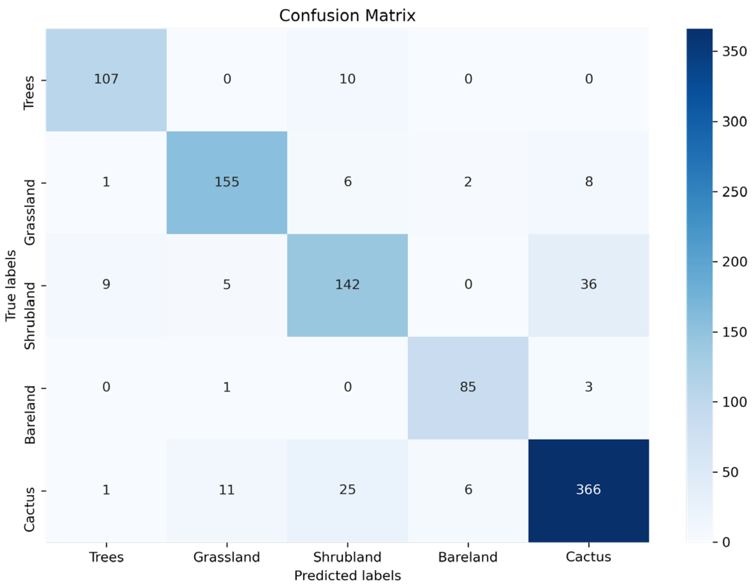


Digital Elevation Model (DEM) data

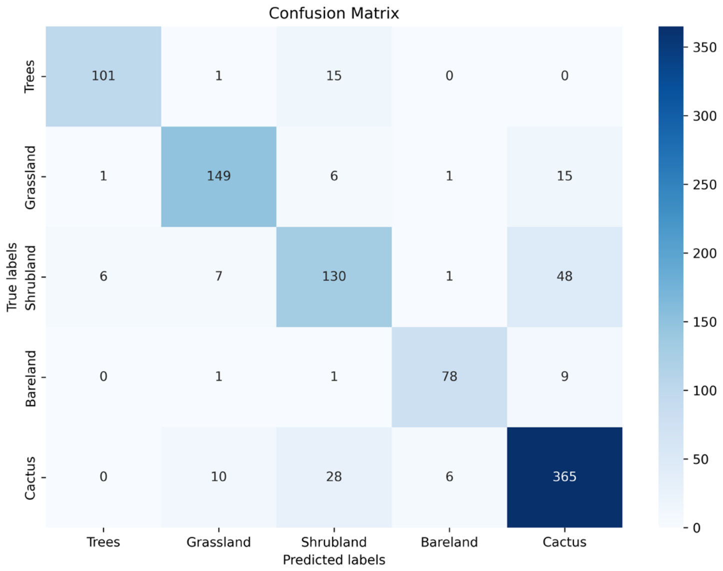
Multi-source GeoAI Framework

Multiclass Classification

		Trees	Grassland	Shrubland	Bareland	Cactus	O.A.
MLP	Precision	0.95	0.86	0.75	0.98	0.91	0.88
	Recall	0.90	0.95	0.73	0.97	0.89	
	F1-Score	0.92	0.90	0.74	0.91	0.90	
RF	Precision	0.93	0.86	0.71	0.89	0.84	0.84
	Recall	0.87	0.88	0.65	0.90	0.88	
	F1-Score	0.90	0.87	0.68	0.89	0.86	
XGBoost	Precision	0.93	0.90	0.71	0.90	0.84	0.84
	Recall	0.86	0.86	0.69	0.89	0.88	
	F1-Score	0.89	0.88	0.70	0.89	0.86	
AdaBoost	Precision	0.81	0.82	0.57	0.88	0.71	0.73
	Recall	0.86	0.65	0.47	0.58	0.87	
	F1-Score	0.83	0.72	0.52	0.70	0.78	



MLP

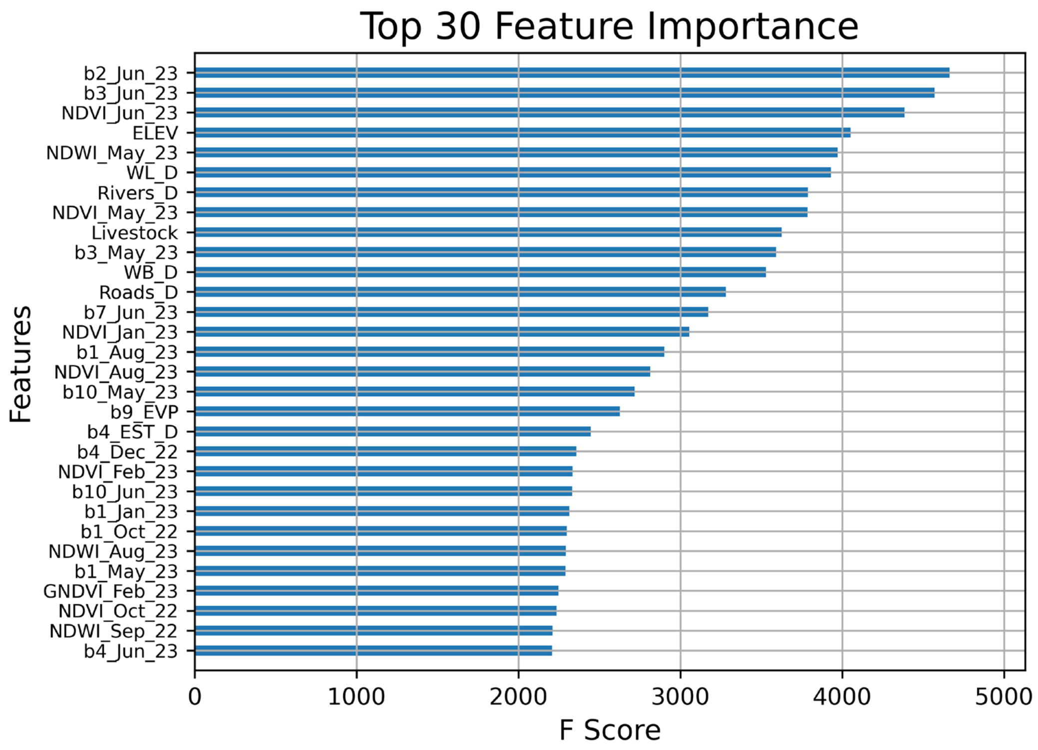


XGBoost

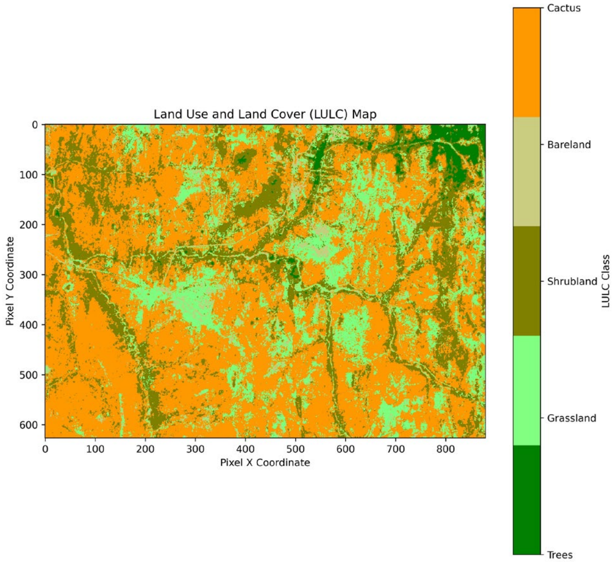
Multiclass Classification: Confusion Matrix

Multiclass Classification: Feature Importance

- XGBoost
  - Green and red bands in Jun
  - NDVI in Jun
  - Elevation
  - NDWI in May
  - Distance to wetlands and rivers
  - Livestock grazing density
  - Red band in May
  - Distance to waterbodies and roads
  - NIR, blue and shortwave Infrared
  - Evapotranspiration and surface temperature
  - Red edge band

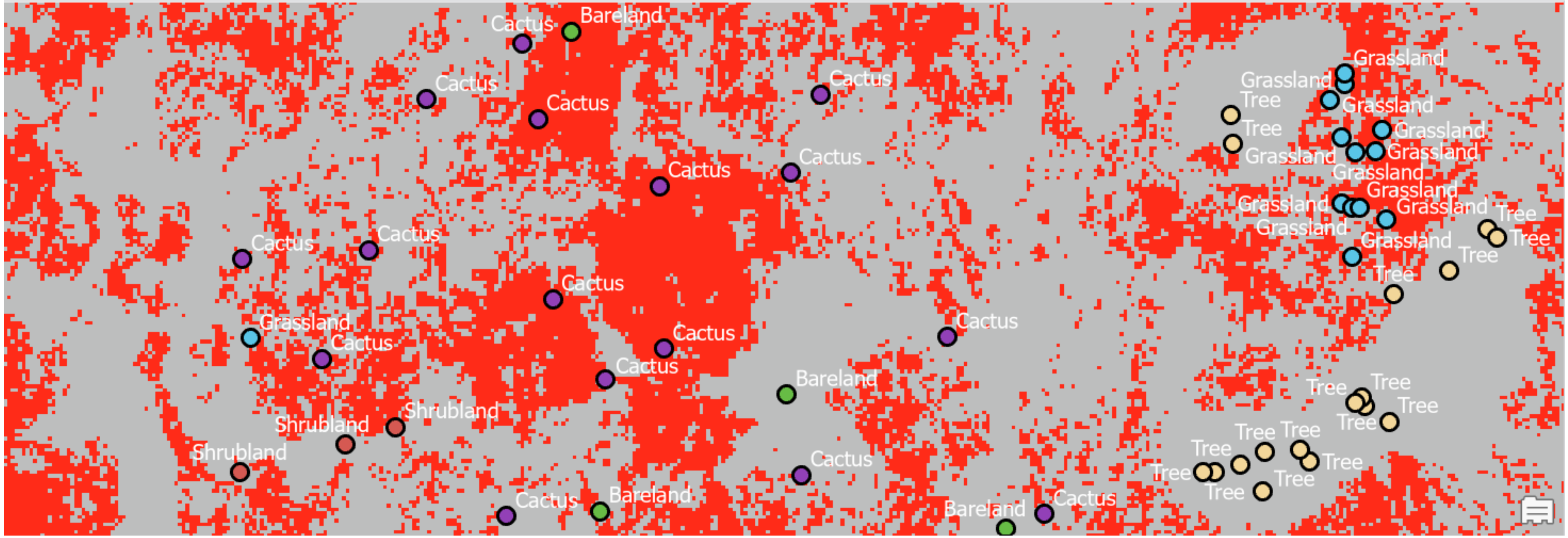


Sentinel 2 True Color Image



MLP Prediction

Prediction Maps for Independent Testing Regions



Ground Labels Vs Predictions

Discussion

- Examine the added-value of multi-source geographical data in predicting cactus distribution by including and excluding landscape and socioeconomic variables
- Examine the added-value of time in-variant variables using Deep Learning Models.
- Multiclass/mix-pixel issues

Acknowledgment

Jiayi Song
Master Student, ECE

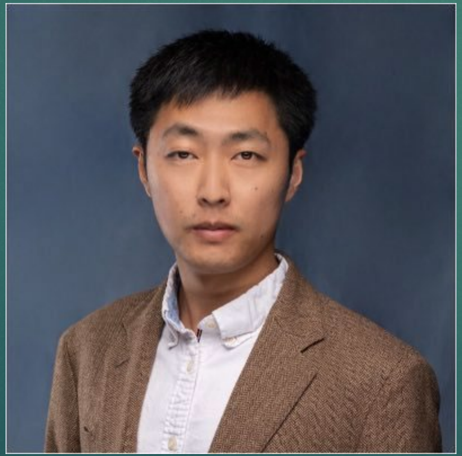
Kenneth Oduor
PhD Student, Agronomy

Dr. Jose Dubeux
Grassland Ecology, Agronomy

Dr. Chang Zhao,
Agronomy Department,
UF/IFAS

Student Volunteers:

- Prerana Venugopal (Master Student, CS)
- Dinesh Gogineni (Master Student, CS)



Dr. Wei Shao
Medical Imaging & AI, ECE

Dr. Zhou Tang
Postdoc, Agronomy

Dr. Igor Bretas
Postdoc, Agronomy

Thank You

Q&A: SRIKANTNANGONDAL@UFL.EDU