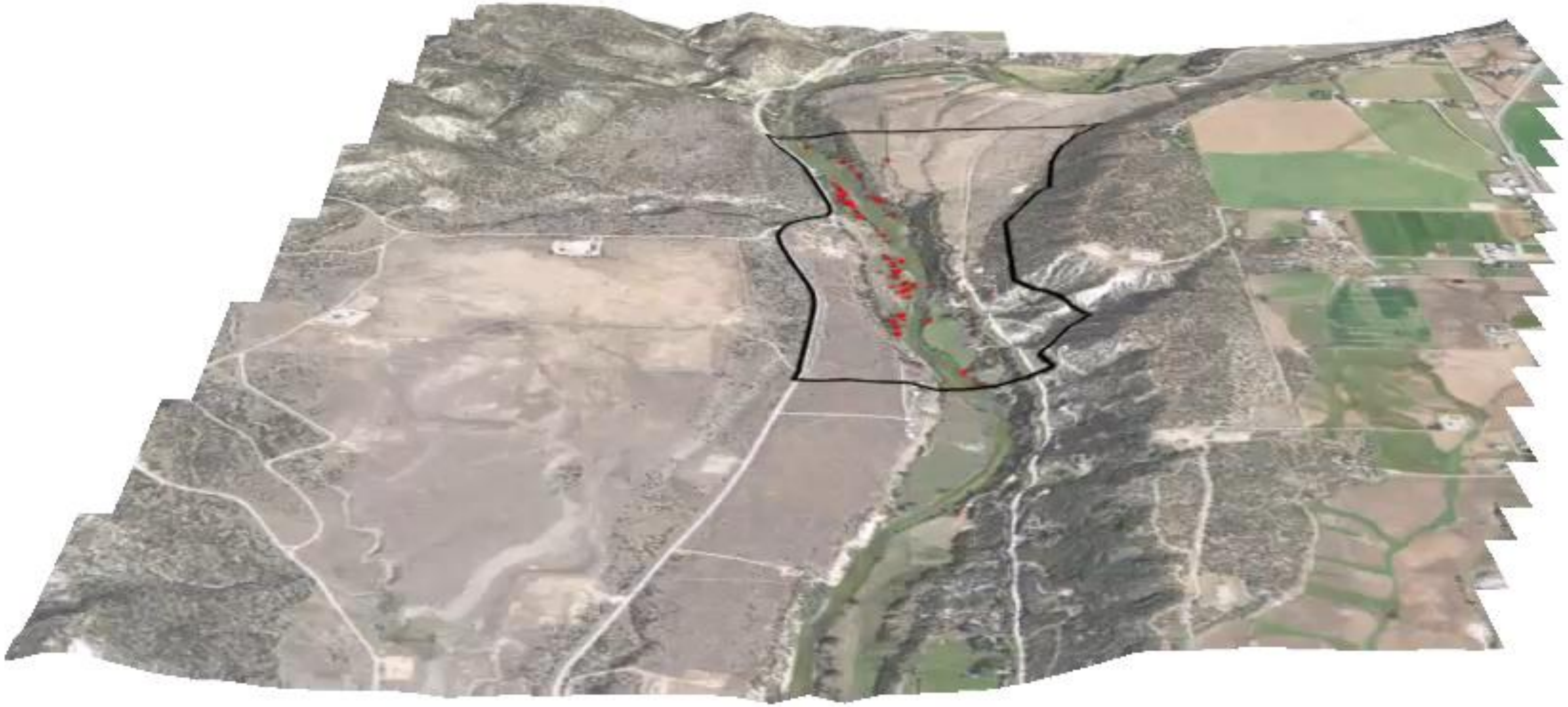


Supervised Classification of Russian Olive in the Animas Valley with NAIP Imagery and Object-Based Image Analysis

Anna Riling, MS, PG

2020 Riparian Restoration Conference

February 4, 2020

Object-Based Image Analysis (OBIA): uses spectral, textural, and spatial elements to classify an image

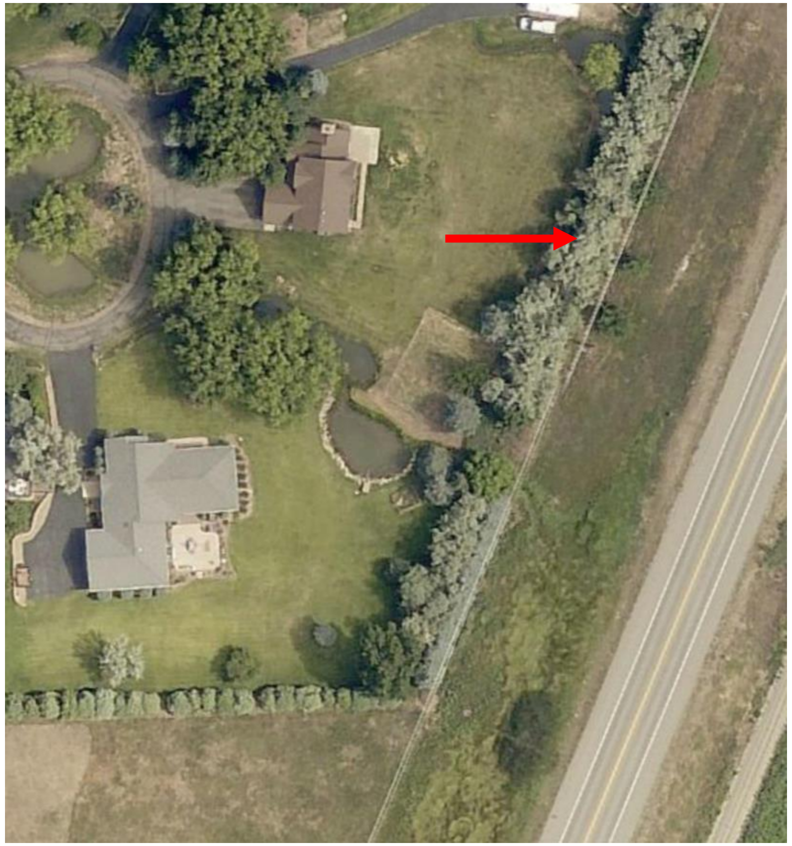
Russian olive is easily distinguished in aerial imagery due to silvery-green canopy.

NAIP: 1-meter, 4-band imagery used to classify Russian olive in a study area on the Animas River, with a user's accuracy of 91.3 percent

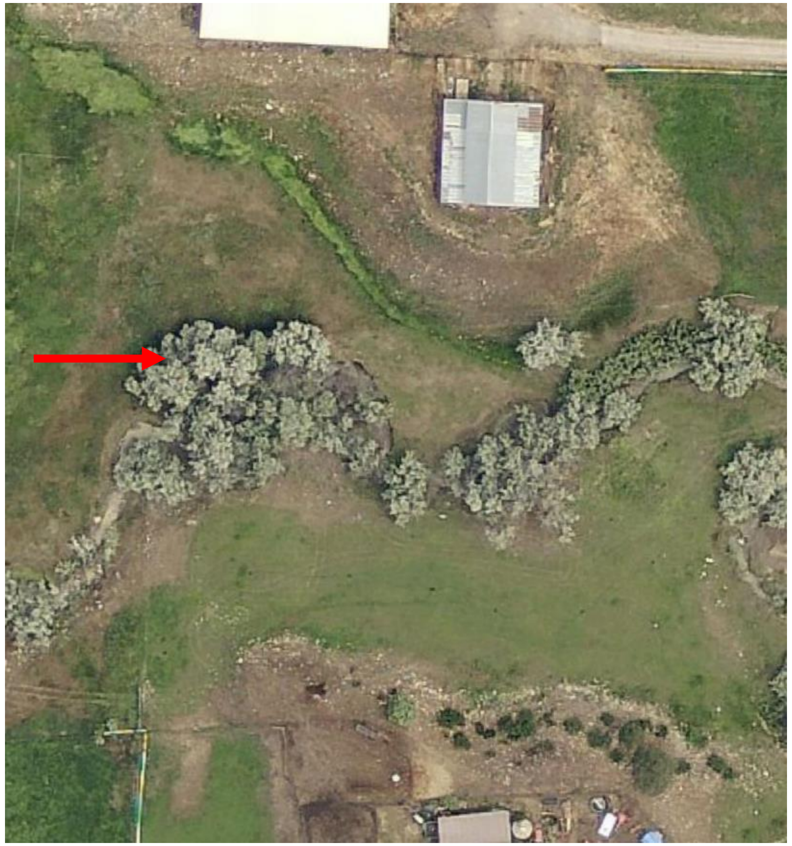
Methodology and parameters may be used in future efforts for mapping Russian olive on a regional scale.

Background

Landscaping



Windrow



Riparian

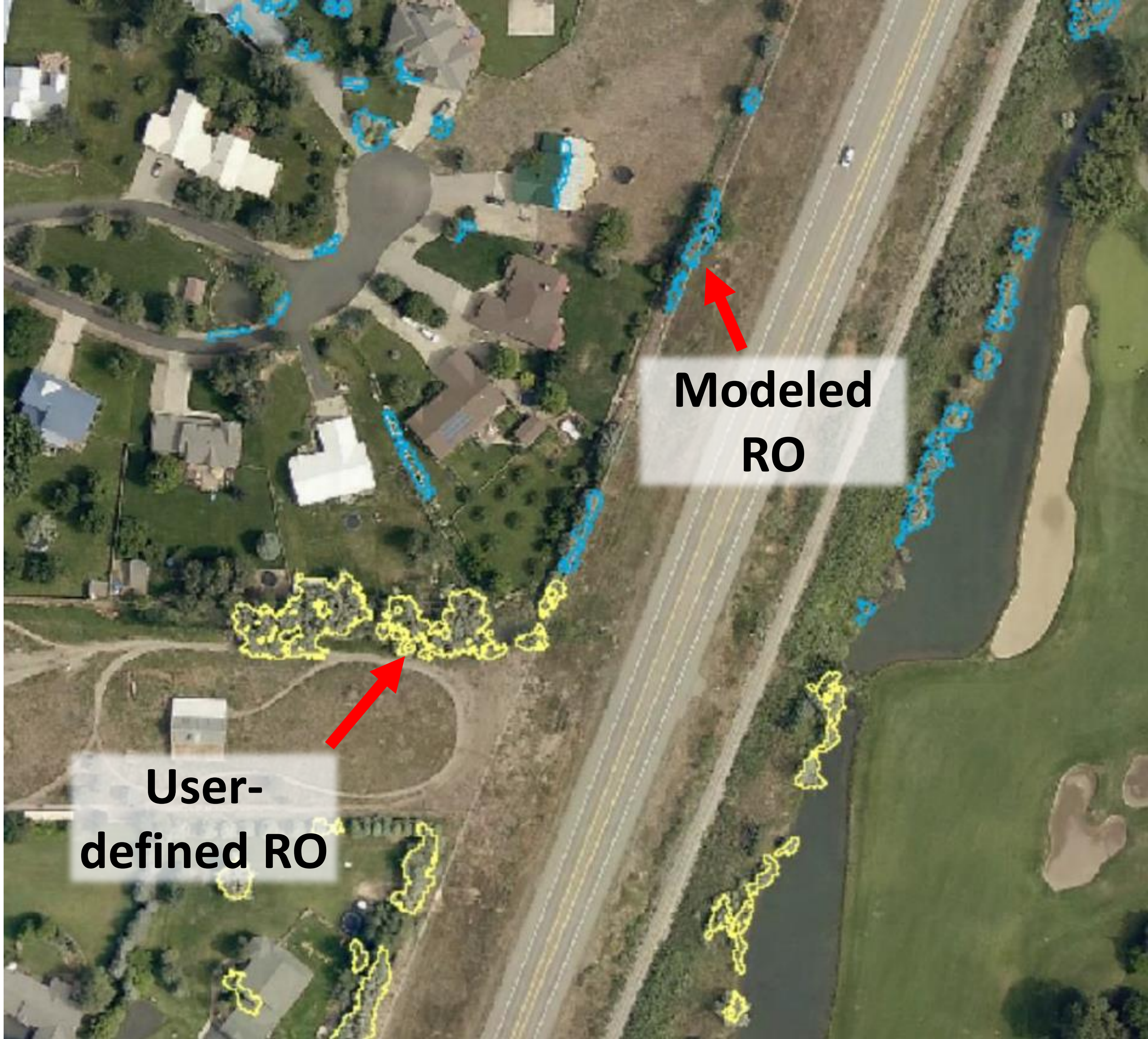
Distribution Patterns

Mitigation and Monitoring

- Mountain Studies Institute (MSI) has state funding to remove RO
- Quantitative, comprehensive documentation of RO distribution does not exist in the Animas Valley
- No means of evaluating mitigation efforts
  - *Photo: MSI*

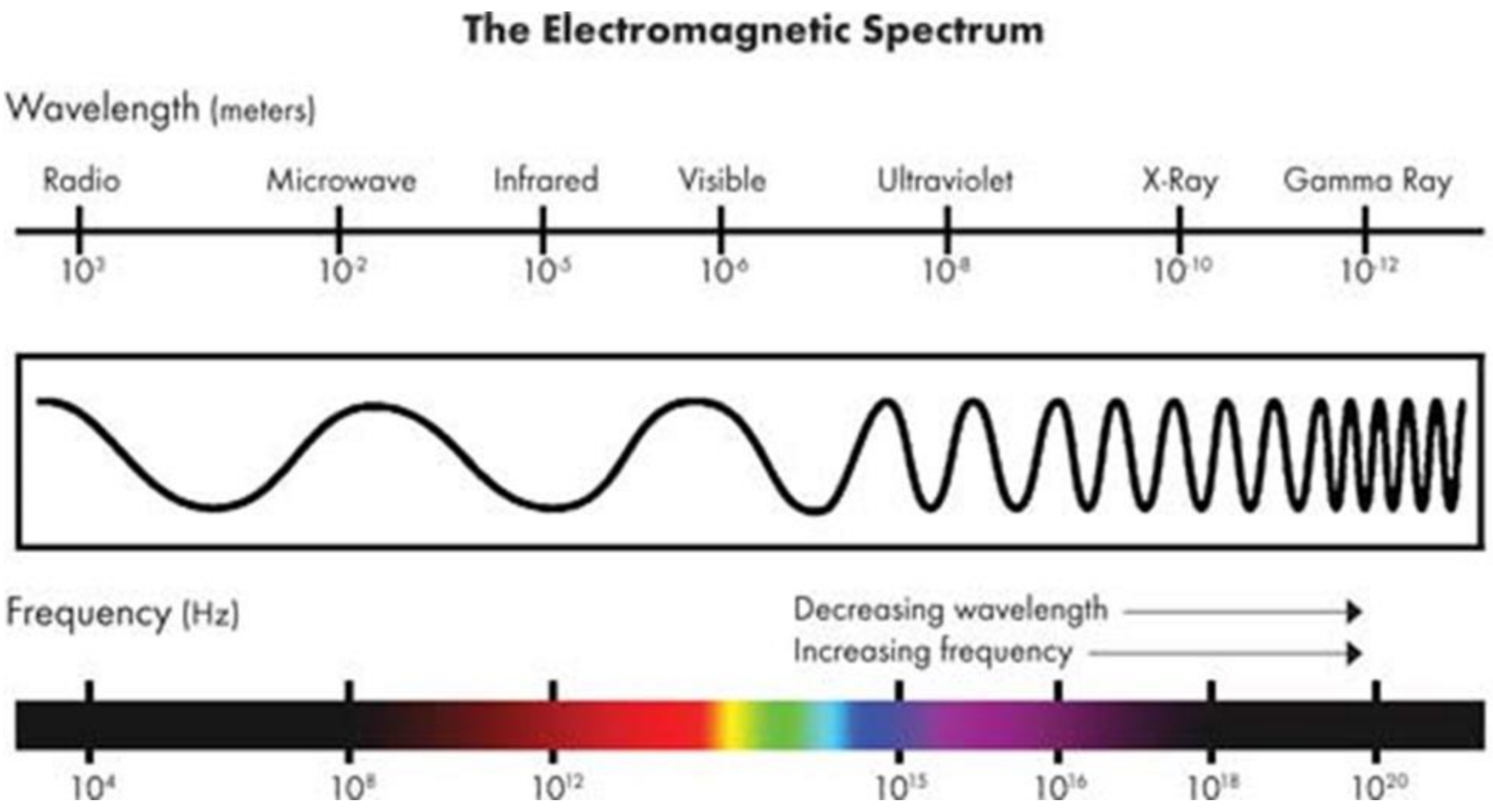
Mitigation and Monitoring

- MSI has used Feature Analyst to classify RO with 3-band (visible spectrum) aerial imagery
- Class confusion with sidewalks, roofs
  - *Source: Giggy, Kuenzi, and Roberts (2018)*



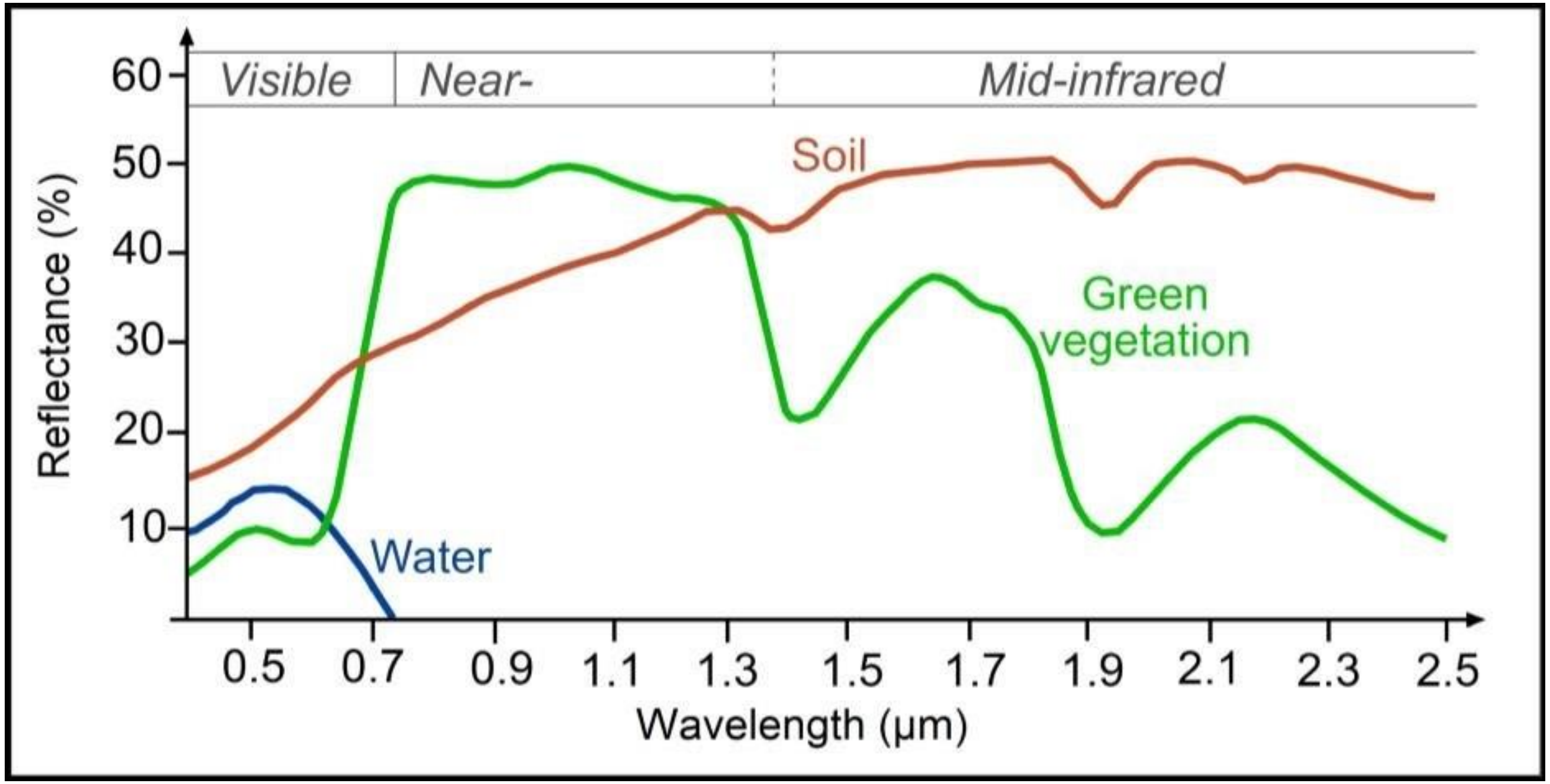
Multispectral Imagery

- Collected across several spectral ranges in several bands
- NIR is a portion of the EM (700 to 1,400 nm) just beyond the visible light spectrum
  - *Source: [Western Reserve Public Media 2009]()*



Multispectral Imagery

- Vegetation strongly reflects in the NIR spectrum
- Different plants have different spectral signatures
  - *Source:* [*GrindGIS 2017*]



OBIA Workflow

SOFTWARE
- ArcGIS for Desktop Advanced[1]
- ENVI
- TPI Extension for ArcGIS[2]

DATA
- DEM[3]
- NAIP[4]
- BLM PLSS[5]
- La Plata County[6]
- LiDAR[7]

STUDY AREA
- Delineate valley bottom with TPI
- Optionally select PLSS Section(s)
- Mosaic to New Raster Tool to mosaic NAIP tiles (if necessary)
- Digitize study area by clipping to sections and valley bottom, other data as needed
- Clip NAIP and LPC imagery to study area

MASK AND ANCILLARY DATA
- Create NDVI from clipped NAIP and inspect for RO values[8]
- Create binary raster for veg/nonveg; convert polygon and export “nonveg”
- Merge with additional data (If necessary); clip to study area.
- Create LiDAR processing extent mask (polygon) that covers the study area
- Create DTM from “ground” points and DSM from “first return” points
- Use Minus tool to subtract DTM from DSM

SAMPLE SET
- Inspect imagery for class types[8]
- Use 50 samples in testing sample set and 3x that for training sample set
- Use Random Points tool to create Sample Set within non-masked region
- Assign class to each point by visual inspection of imagery and height band
- Use Sample Parsing Model to separate sample set into training and testing points, with 2/3 as training samples

SEGMENTATION
- Stack layers[9]:
  - Input = Clipped NAIP
  - Mask = NDVI (use inverse)
  - Ancillary data = Height
  - Custom bands = NDVI, Hue, Saturation, Intensity
- Scale = 20; Merge = 50, accept default settings

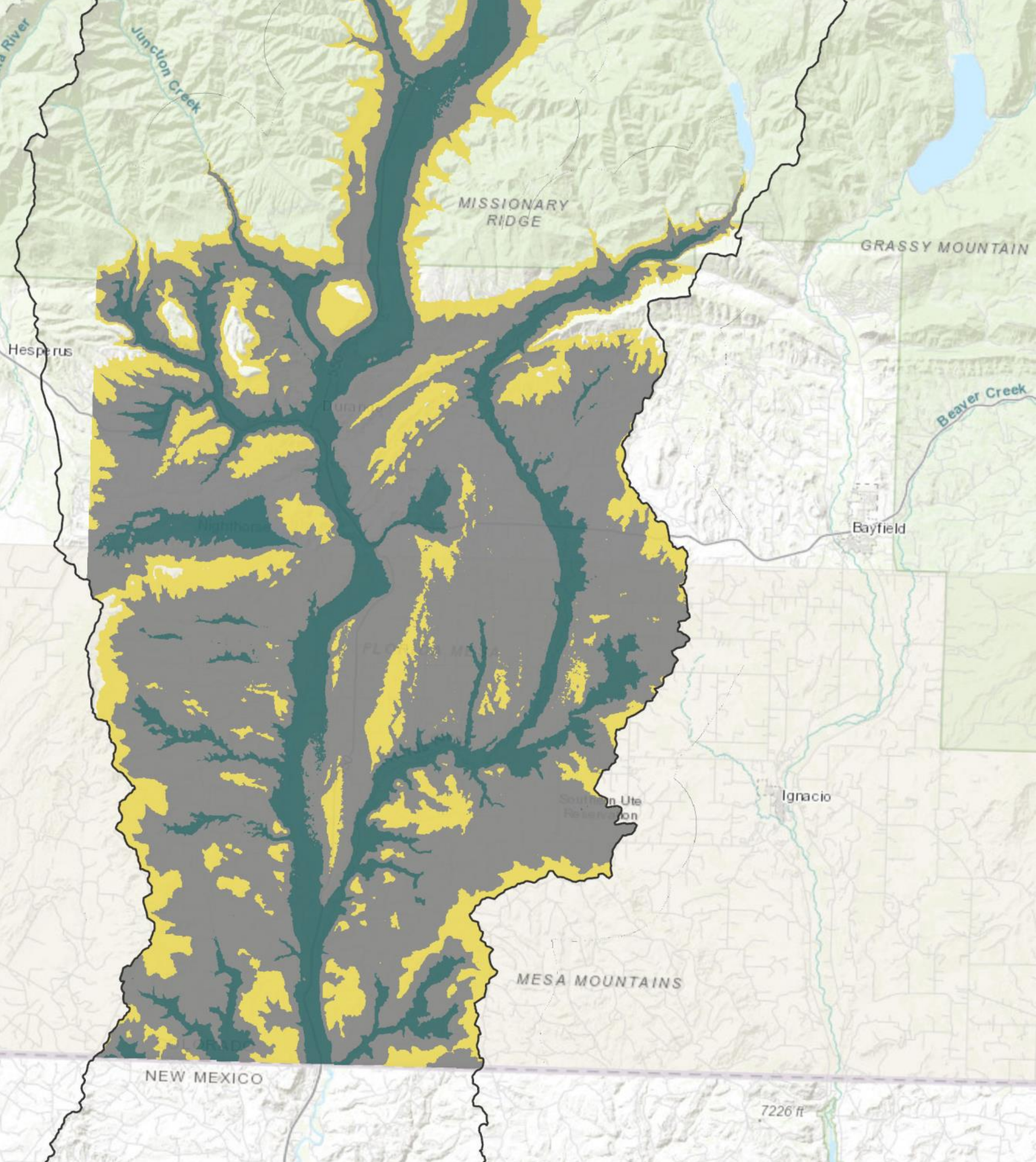
CLASSIFICATION
- Import training data, refine if necessary[9]
- Allow ENVI to select best predictor attributes
- Use KNN with K value of 3, do not merge adjacent segments
- Export all results[10]

VALIDATION
- Generate Confusion Matrix[9]
- Revise classification parameters as necessary

Study Area

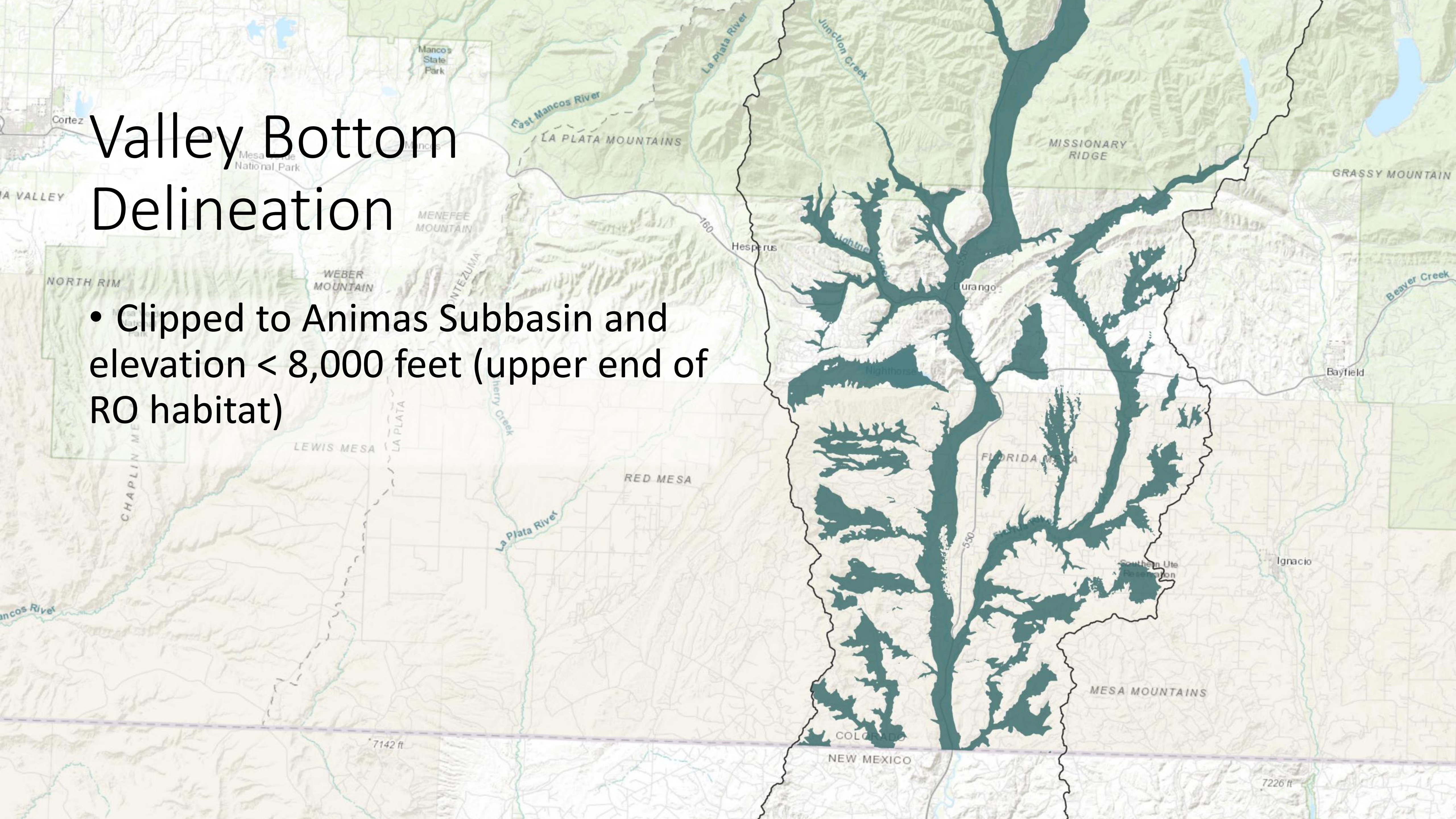
Valley Bottom Delineation

- Confines analysis to riparian zones
- Approximated using topographic position index (TPI) tool [(Jenness, et al 2013)](#)
- TPI = difference between a cell elevation value and the average elevation of the neighborhood around that cell
- Combined with slope, classifies DEM into ridge, valley bottom, slope



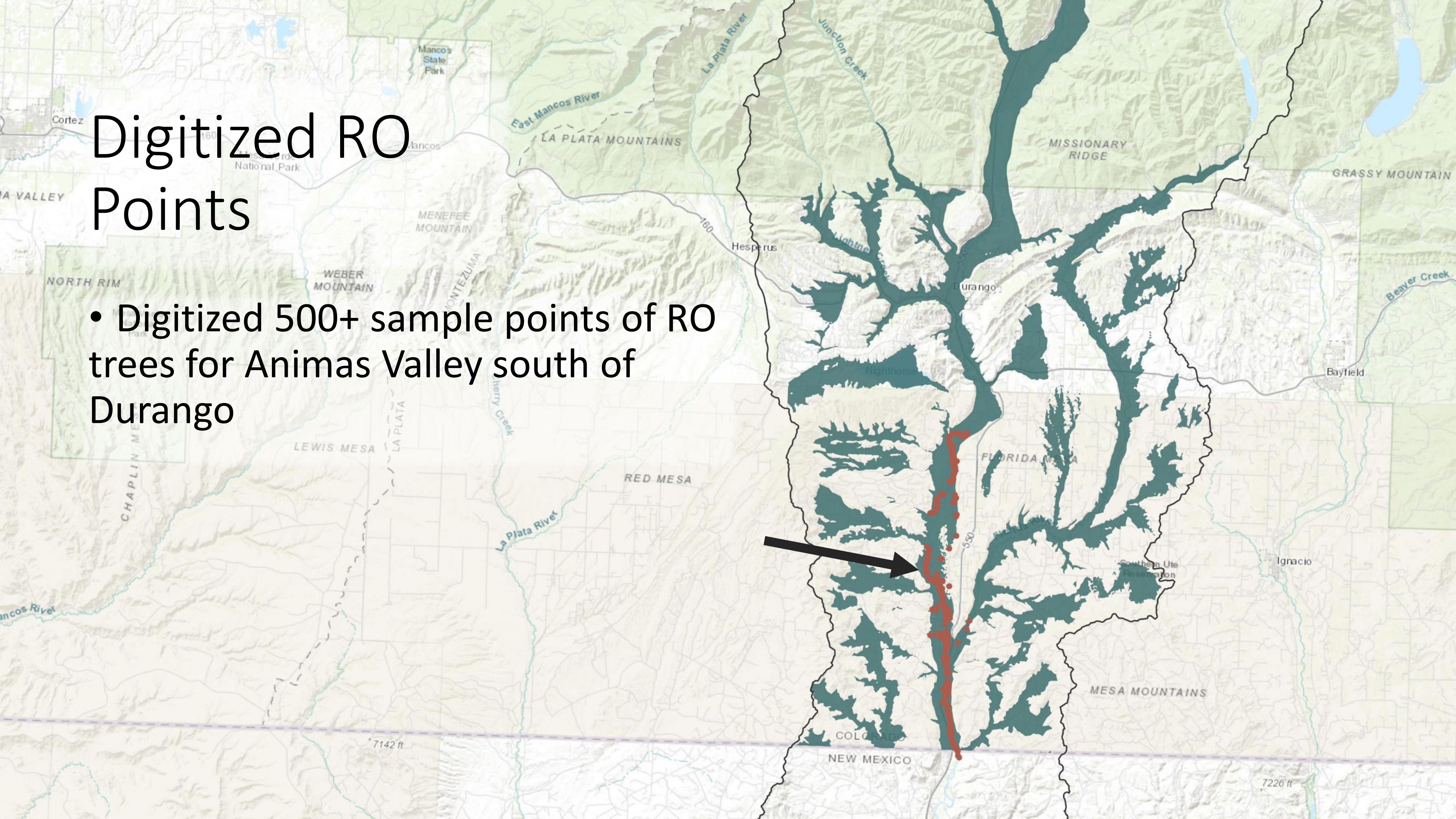
Valley Bottom Delineation

- Clipped to Animas Subbasin and elevation < 8,000 feet (upper end of RO habitat)



Digitized RO Points

- Digitized 500+ sample points of RO trees for Animas Valley south of Durango



Digitized RO Points

La Plata County 9" (2017)

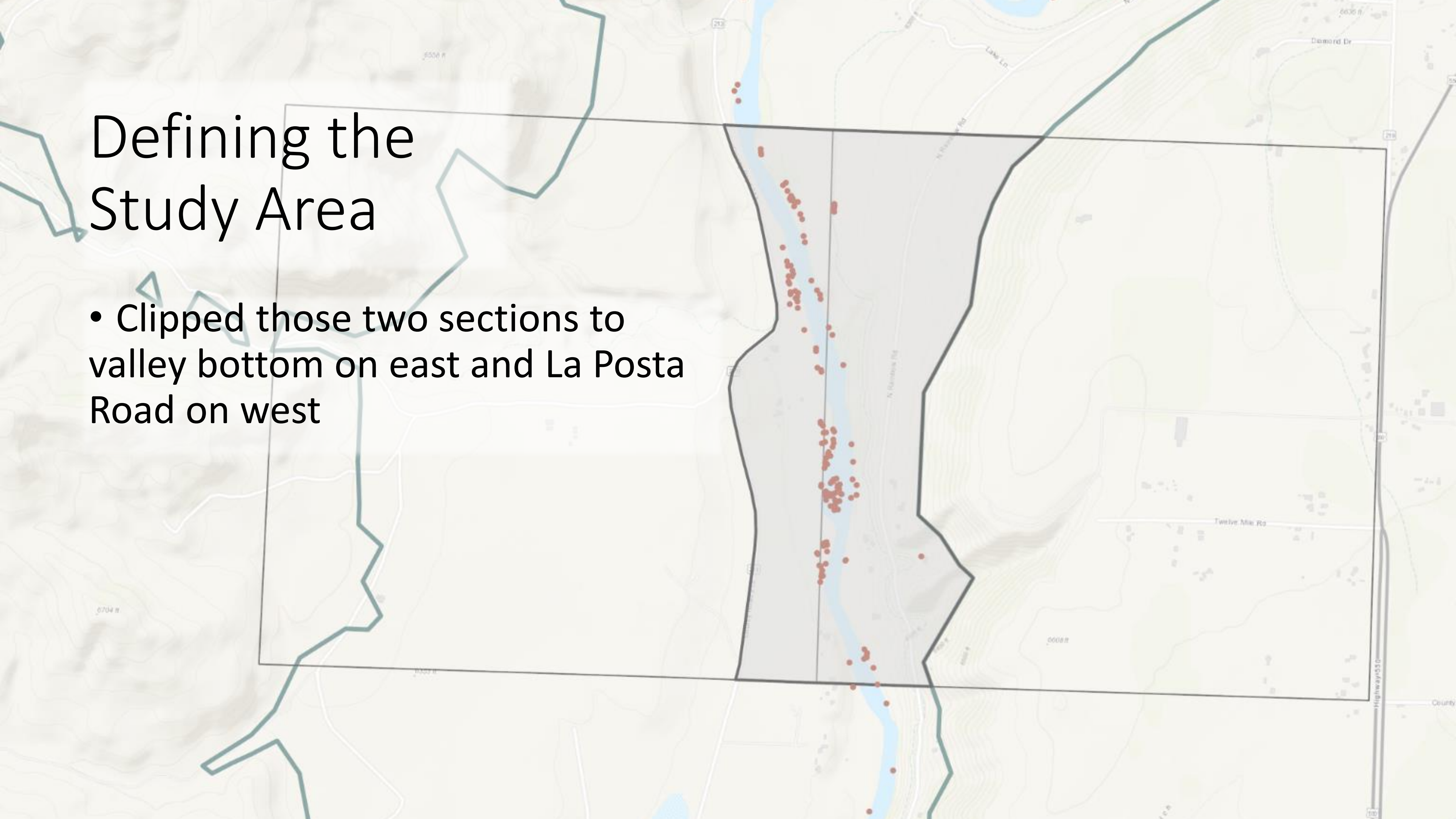
NAIP 1-meter (September 2017)

Defining the Study Area

- I selected the two PLSS Sections with the highest number of RO samples

Defining the Study Area

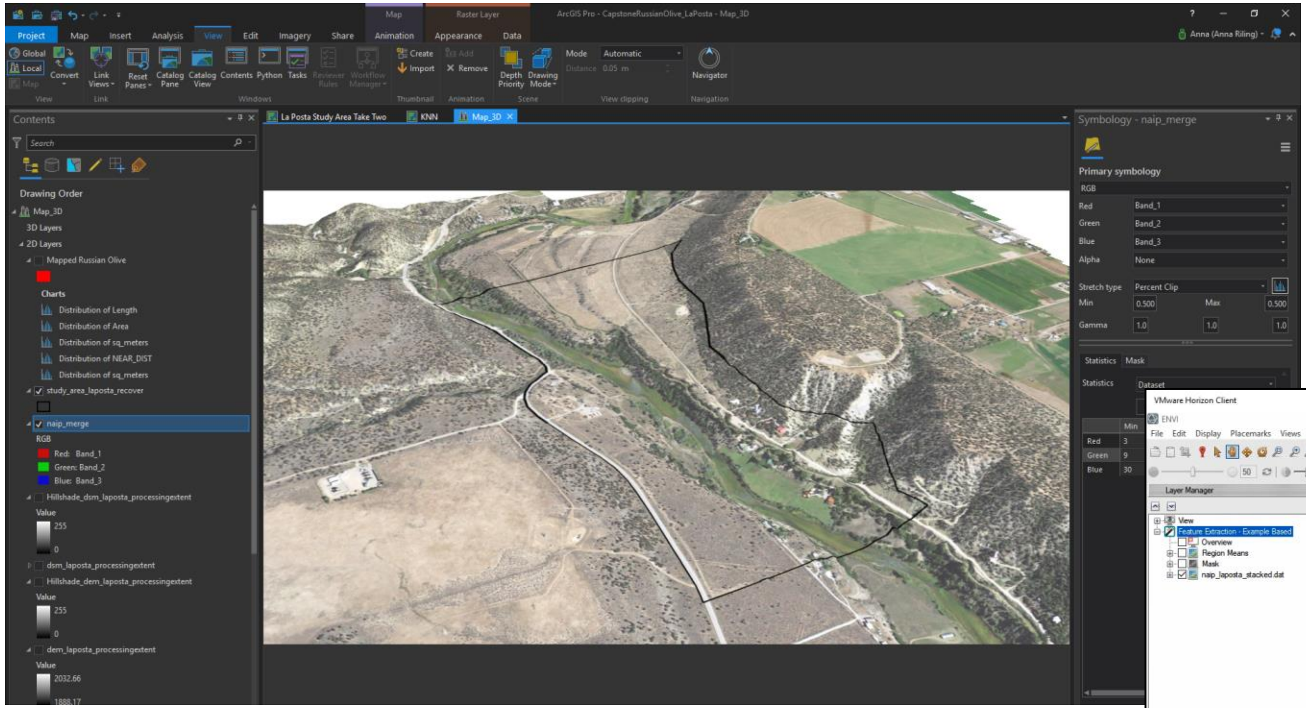
- Clipped those two sections to valley bottom on east and La Posta Road on west



Defining the Study Area

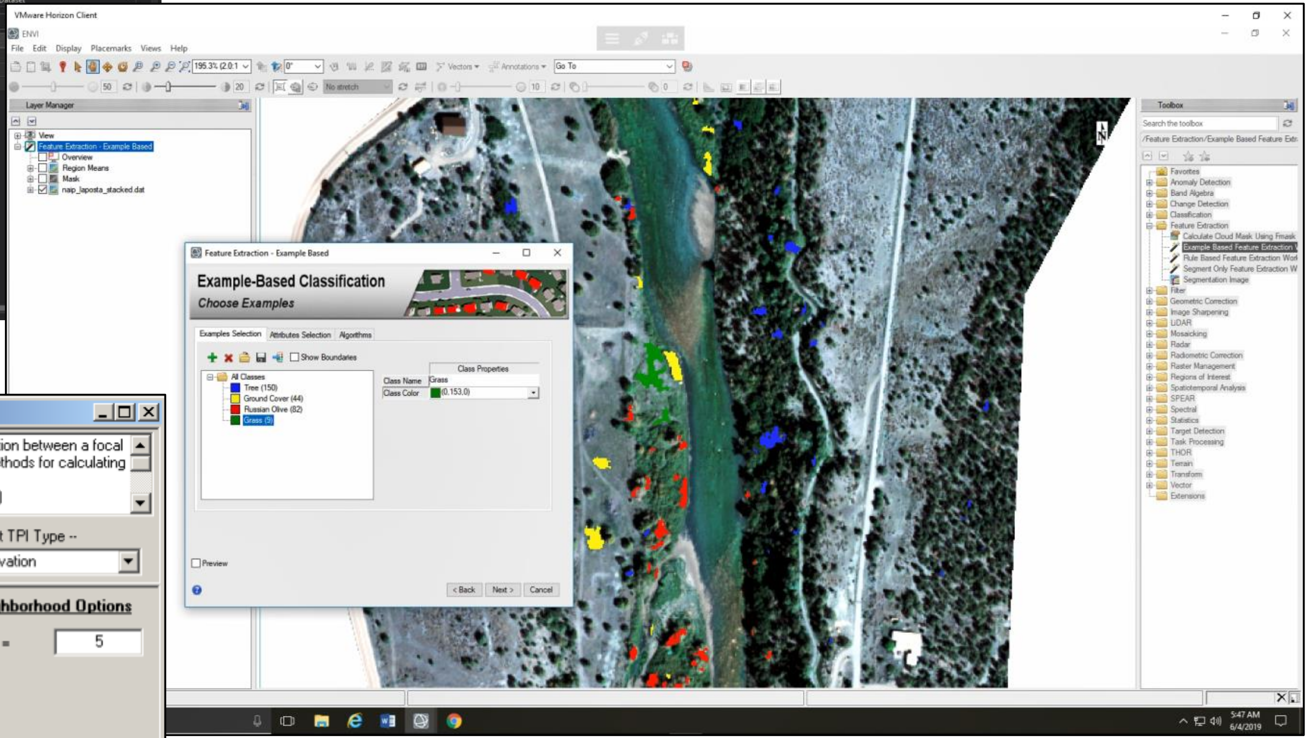
- Area ≈ 1 sq km

Data and Methods

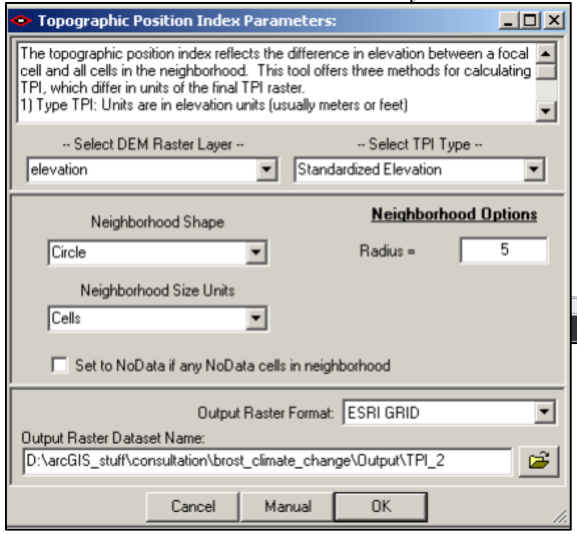


ArcGIS Pro 2.3.2

Harris Geospatial ENVI 5.5



Land Facet
Corridor Designer



NAIP Imagery

- Multispectral
- Free
- Flown every two years = Repeatable
- “High” resolution 1-meter
- Previous studies have used it to successfully classify RO with OBIA

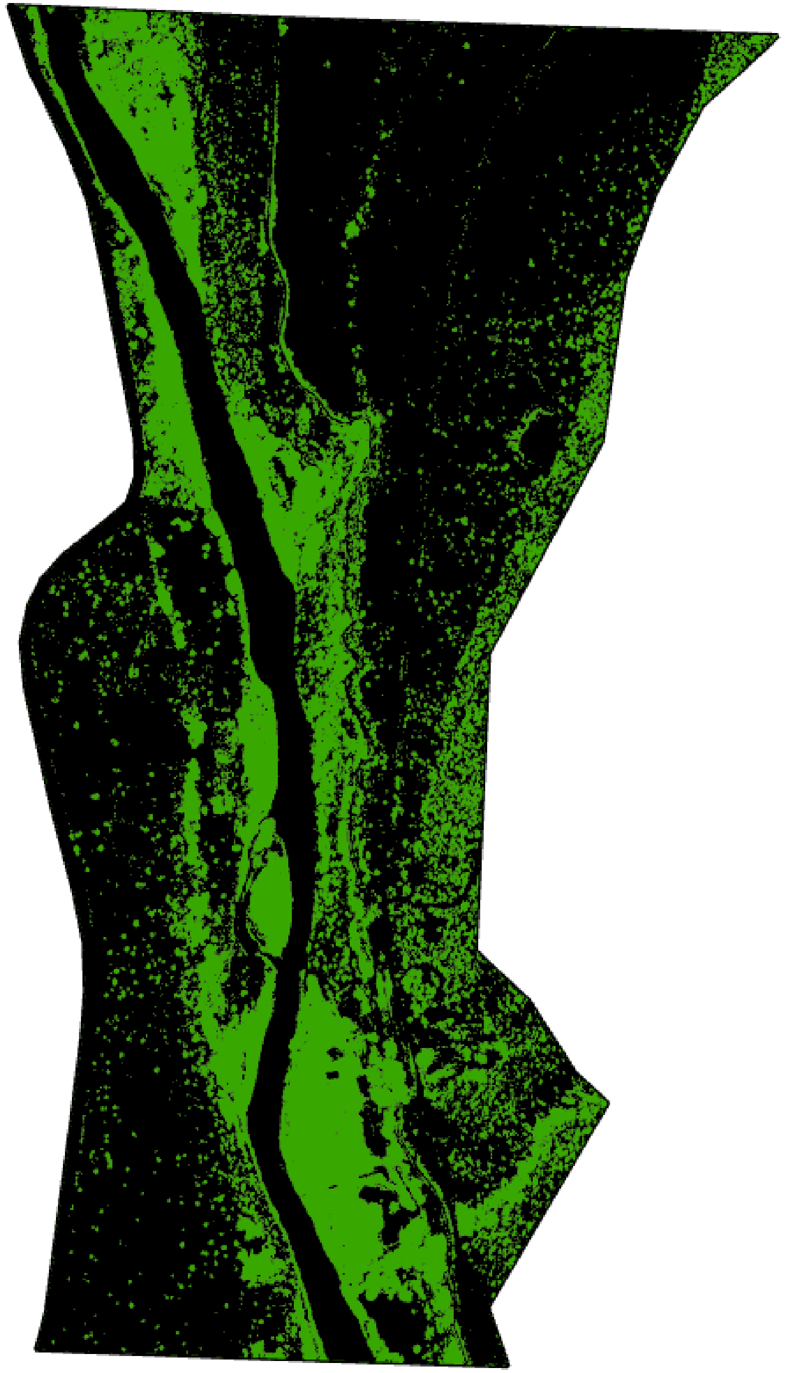
(Hamilton et al. 2006; X. Li and Shao 2014; Tobalske and Vance 2017)

Masking: NDVI

- Masking eliminates areas of non-interest from analysis to speed up and simplify classification
- Normalized Difference Vegetation Index (NDVI) distinguishes between vegetation and non-veg.
- NDVI = (NIR – Red) / (NIR + Red)

Masking: NDVI

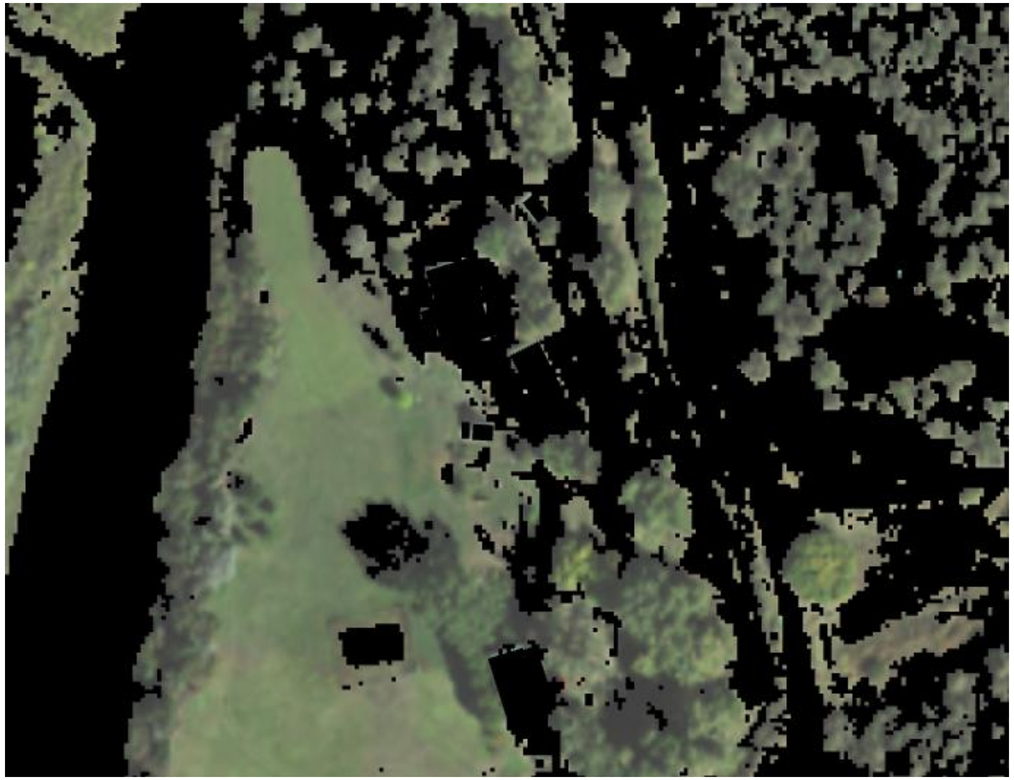
- NDVI RO values ranged from 0.1 to 0.6.
- NDVI reclassified into binary raster:
  - Non-veg < 0.1
  - Veg ≥ 0.1
- Some rooftops were still unmasked



Masking: NDVI + Structures Layer

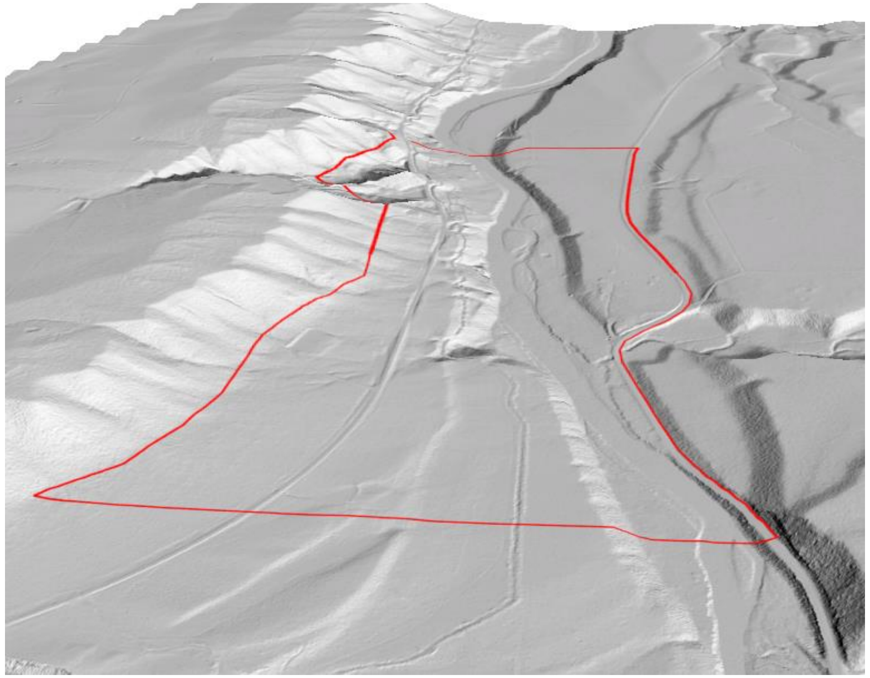
NAIP:

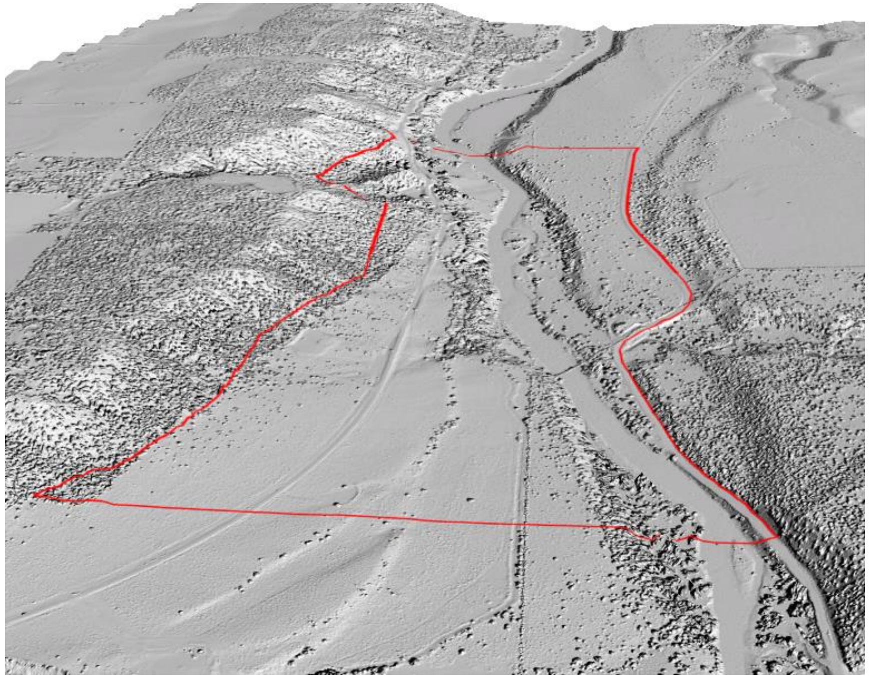
Final mask:



Ancillary Data: LiDAR

- Ancillary data like LiDAR increases classification accuracy in most cases
- I created a Height Band by subtracting the DEM from the DSM



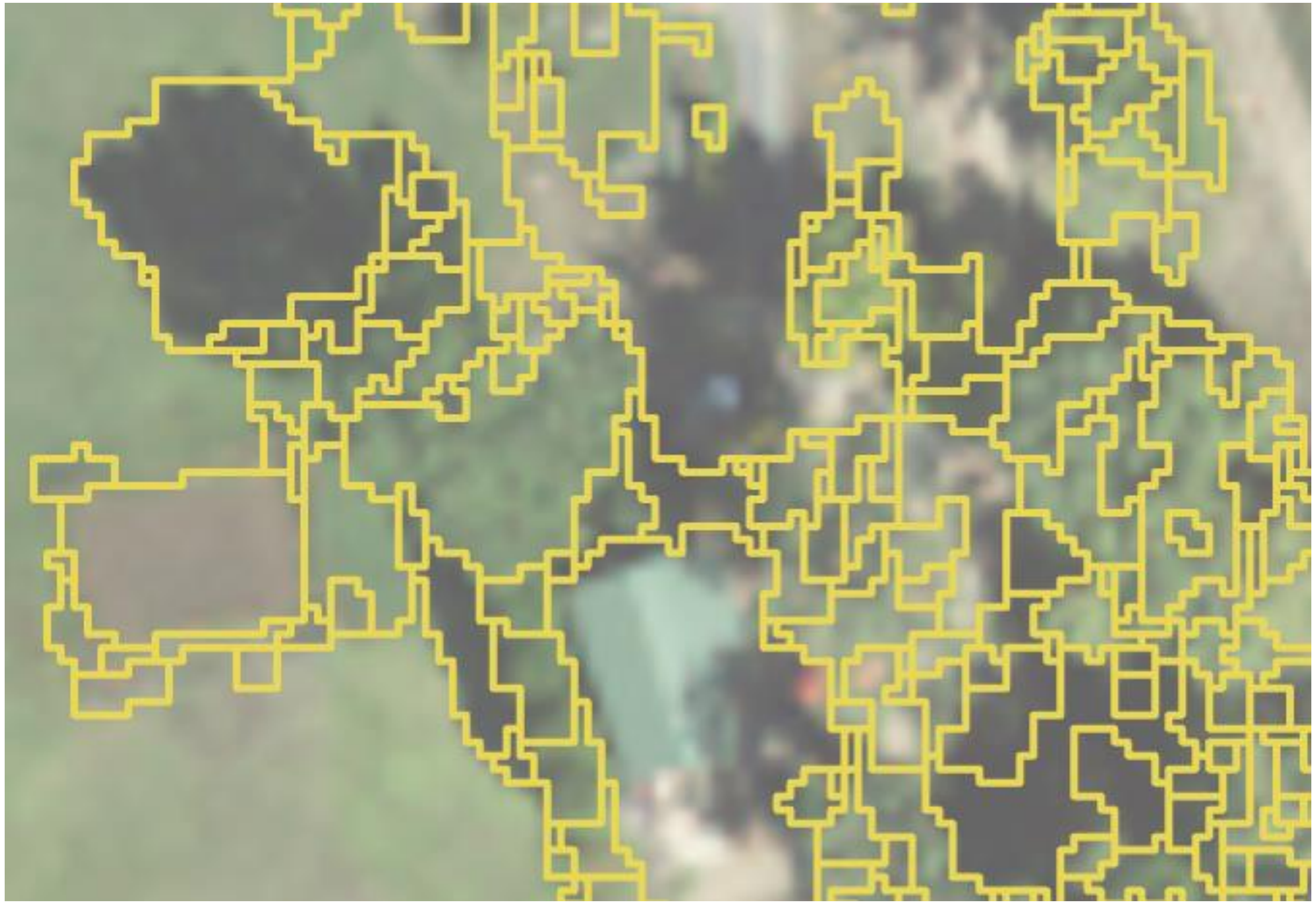


Training and Testing Samples

- Training points used to train the classification
- Testing points used to see how well the classification did
- Classes = Grass, Ground Cover (Shrubs), Russian Olive, Tree
- Generate 800 random points within the non-masked area, each assigned to a class

Segmentation

- Partitions the image into regions of similar spectral intensities
- Each segment is assigned the mean spectral value of the pixels within that segment



Segmentation

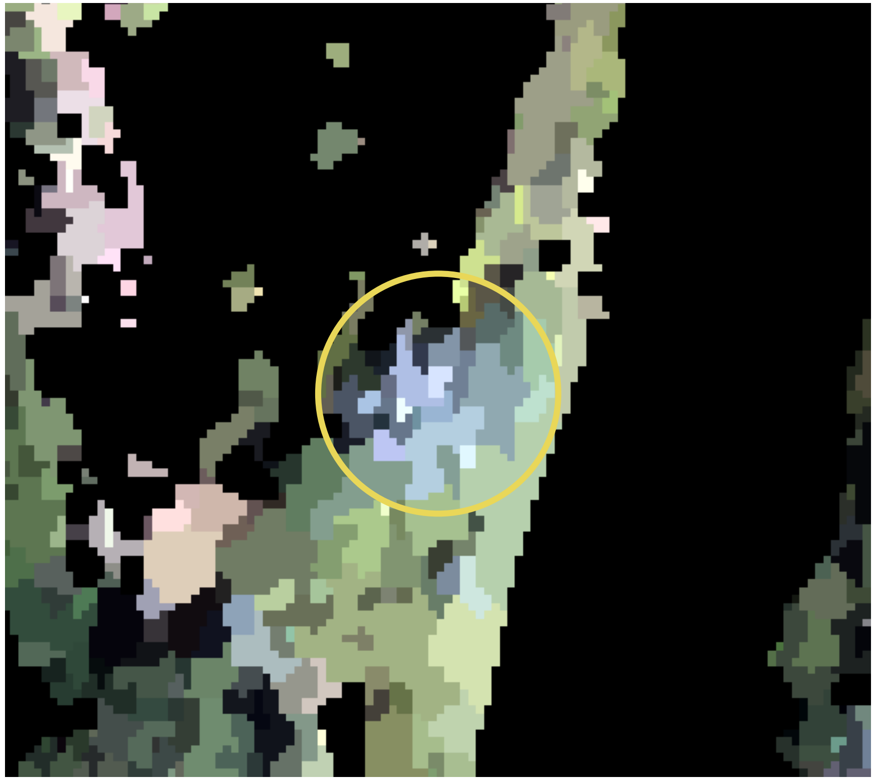
Output is a 9-band raster with the following bands:

1. NAIP Red Band
2. NAIP Green Band
3. NAIP Blue Band
4. NAIP NIR Band
5. Height
6. NDVI
7. Hue
8. Saturation
9. Intensity

Segmentation

NAIP:

Segmented version (RGB values):



Attributes

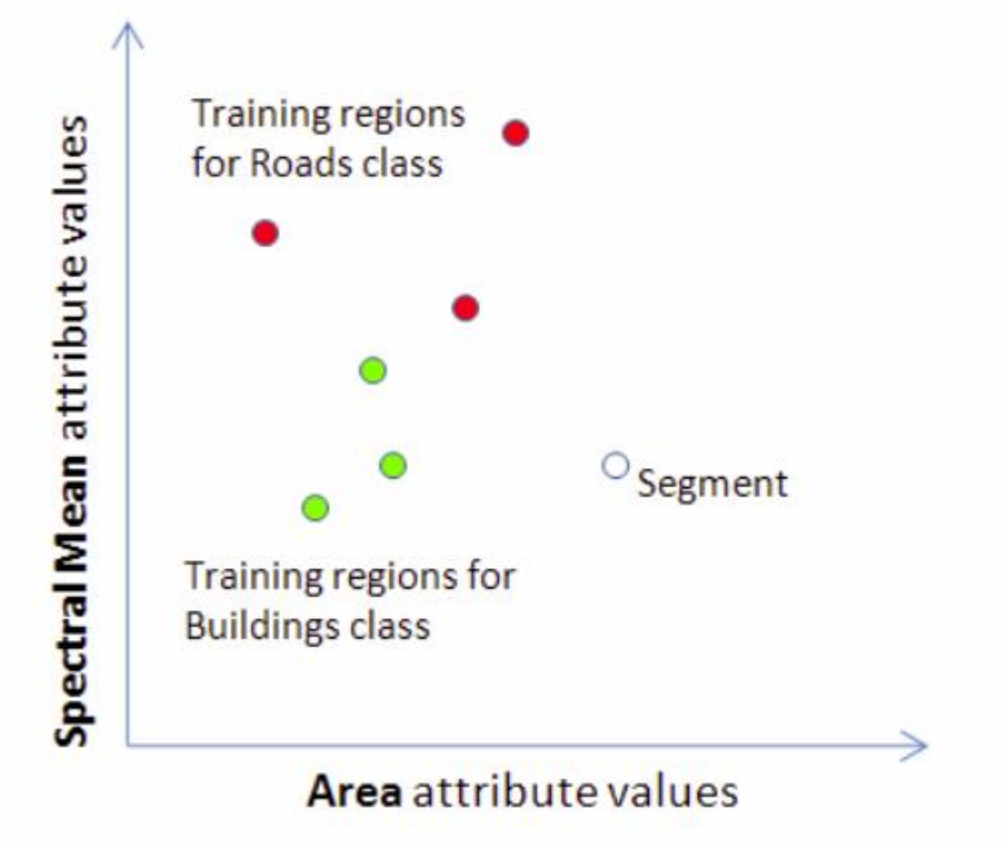
- ENVI calculates 22 **spatial**, **spectral**, and **textural** attributes for each band in the segmentation raster and uses them to classify each segment

Classification Algorithms

- Two commonly used algorithms were compared:
  - K Nearest Neighbors (KNN)
  - Support Vector Machine (SVM)
- There are many, many others
- KNN was ultimately used to derive results.

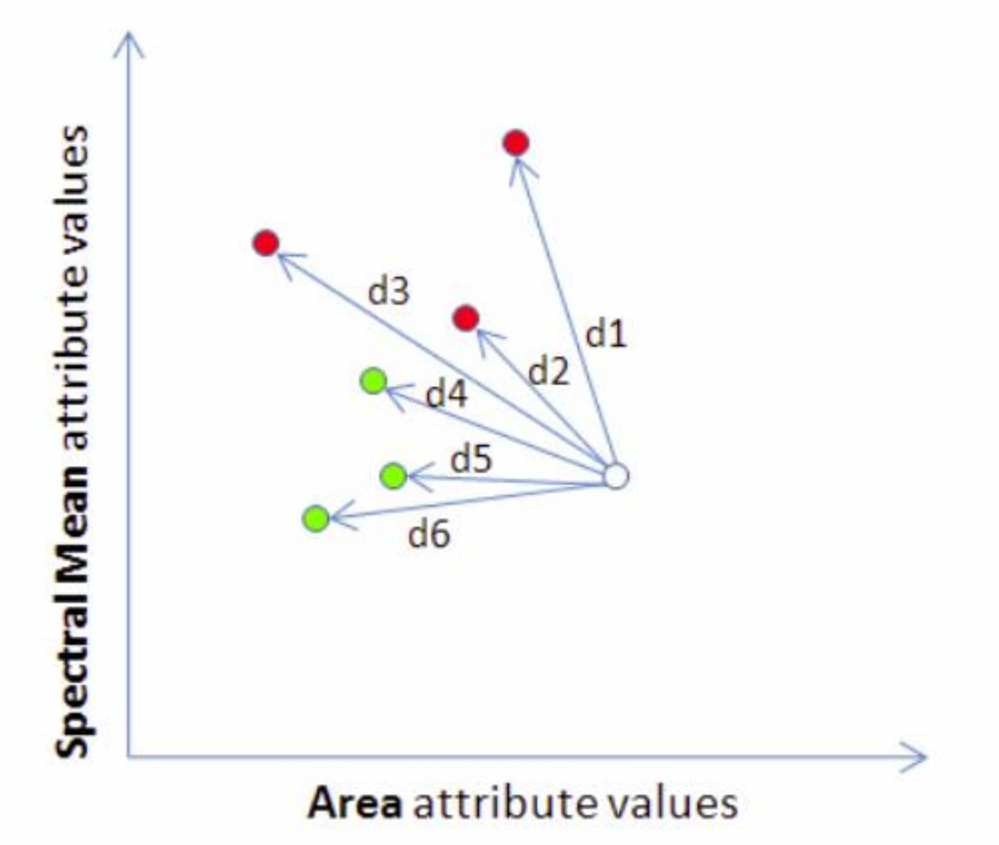
K Nearest Neighbor

- Measures Euclidean distance from each unknown segment to each training segment
  - *Source: [Harris Geospatial Solutions 2019]()*



K Nearest Neighbor

- Segments are assigned the most common class of the *k* number of training samples that are nearest to it
- The higher the *k* value, the greater the generalization
- This study used a *k* value of 3
- *Source: [Harris Geospatial Solutions 2019]*

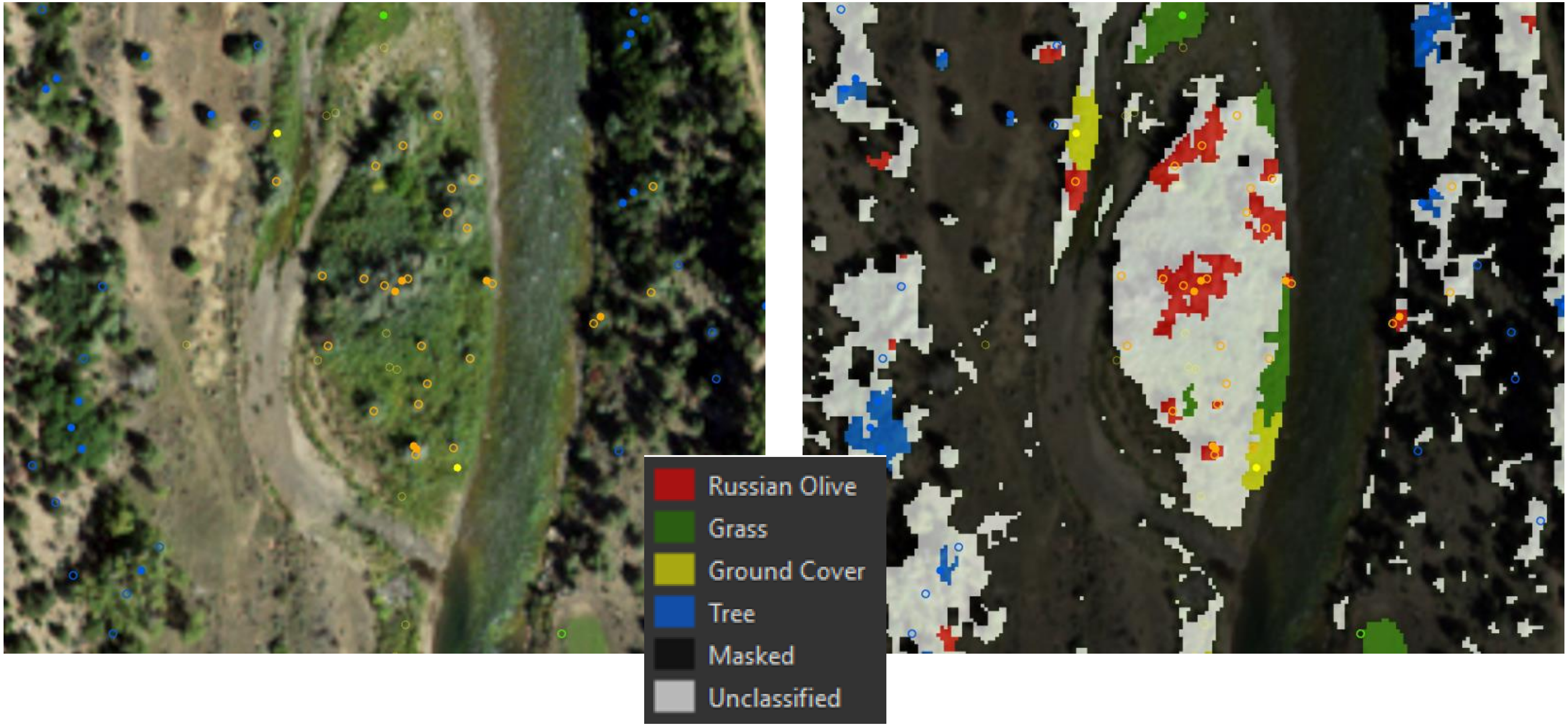


Results & Validation

Results

- 4,031 sq meters of mapped Russian olive
- In ENVI, I manually edited the KNN classification result to exclude segments that were misclassified or too small to determine if they were RO

KNN Results


Russian Olive
Grass
Ground Cover
Tree
Masked
Unclassified

Validation: Confusion Matrix

- Compares the relationship between test samples with the classification results
- Reports different metrics of accuracy: **Overall**, **Producer's**, **User's Accuracy** and **Kappa Coefficient**

User's Accuracy

- the number of correctly classified objects divided by the total number of objects that were classified in that class
- 91.30% of the areas identified as RO truly represent RO on the ground
- Represents a measure of commission: the likelihood of someone going to an an area on the ground that has been classified as RO, and actually finding a Russian olive tree at that spot in the field

Confusion Matrix (Count)

REFERENCE DATA

CLASSIFICATION DATA

Class	Grass	Ground Cover	Russian Olive	Tree	Total
Unclassified	2	47	36	148	233
Grass	**58**	0	0	1	59
Ground Cover	0	**5**	1	0	6
Russian Olive	0	2	**63**	4	69
Tree	0	0	**0**	12	12
Total	60	54	100	165	379

Class	**Commission Error (%)**	**Omission Error (%)**
Grass	1.69	3.33
Ground Cover	16.67	90.74
Russian Olive	8.70	37.00
Tree	0.00	92.73

Class	**Producer's Accuracy (%)**	**User's Accuracy (%)**
Grass	96.67	98.31
Ground Cover	9.26	83.33
Russian Olive	63.00	91.30
Tree	7.27	100.00

Overall Accuracy (%)	**Kappa Coefficient**
36.41	0.30

Producer's Accuracy

- number of correctly classified objects in a class divided by the number of reference samples used for that class
- 63.00% of RO reference data samples were correctly identified as RO
- Represents a measure of omission: if I am standing at a RO in the field, what is the probability that the classification will correctly identify that object as RO

Confusion Matrix (Count)

REFERENCE DATA

CLASSIFICATION DATA

Class	Grass	Ground Cover	Russian Olive	Tree	Total
Unclassified	2	47	36	148	233
Grass	**58**	0	0	1	59
Ground Cover	0	**5**	1	0	6
Russian Olive	0	2	**63**	4	69
Tree	0	0	**0**	12	12
Total	60	54	100	165	379

Class	**Commission Error (%)**	**Omission Error (%)**
Grass	1.69	3.33
Ground Cover	16.67	90.74
Russian Olive	8.70	37.00
Tree	0.00	92.73

Class	**Producer's Accuracy (%)**	**User's Accuracy (%)**
Grass	96.67	98.31
Ground Cover	9.26	83.33
Russian Olive	63.00	91.30
Tree	7.27	100.00

Overall Accuracy (%)	**Kappa Coefficient**
36.41	0.30

Overall Accuracy

- calculated by dividing the number of correctly classified objects by the total number of reference objects
- Overall: 36.41%
- Yikes…or not?

Confusion Matrix (Count)

REFERENCE DATA

CLASSIFICATION DATA

Class	Grass	Ground Cover	Russian Olive	Tree	Total
Unclassified	2	47	36	148	233
Grass	**58**	0	0	1	59
Ground Cover	0	**5**	1	0	6
Russian Olive	0	2	**63**	4	69
Tree	0	0	**0**	12	12
Total	60	54	100	165	379

Class	**Commission Error (%)**	**Omission Error (%)**
Grass	1.69	3.33
Ground Cover	16.67	90.74
Russian Olive	8.70	37.00
Tree	0.00	92.73

Class	**Producer's Accuracy (%)**	**User's Accuracy (%)**
Grass	96.67	98.31
Ground Cover	9.26	83.33
Russian Olive	63.00	91.30
Tree	7.27	100.00

Overall Accuracy (%)	**Kappa Coefficient**
36.41	0.30

Kappa Coefficient

- measure of the difference between actual and chance agreement between reference data and classified data
- Kappa: 0.30
- These results are 30% better than one resulting from chance

Confusion Matrix (Count)

REFERENCE DATA

CLASSIFICATION DATA

Class	Grass	Ground Cover	Russian Olive	Tree	Total
Unclassified	2	47	36	148	233
Grass	**58**	0	0	1	59
Ground Cover	0	**5**	1	0	6
Russian Olive	0	2	**63**	4	69
Tree	0	0	**0**	12	12
Total	60	54	100	165	379

Class	**Commission Error (%)**	**Omission Error (%)**
Grass	1.69	3.33
Ground Cover	16.67	90.74
Russian Olive	8.70	37.00
Tree	0.00	92.73

Class	**Producer's Accuracy (%)**	**User's Accuracy (%)**
Grass	96.67	98.31
Ground Cover	9.26	83.33
Russian Olive	63.00	91.30
Tree	7.27	100.00

Overall Accuracy (%)	**Kappa Coefficient**
36.41	0.30

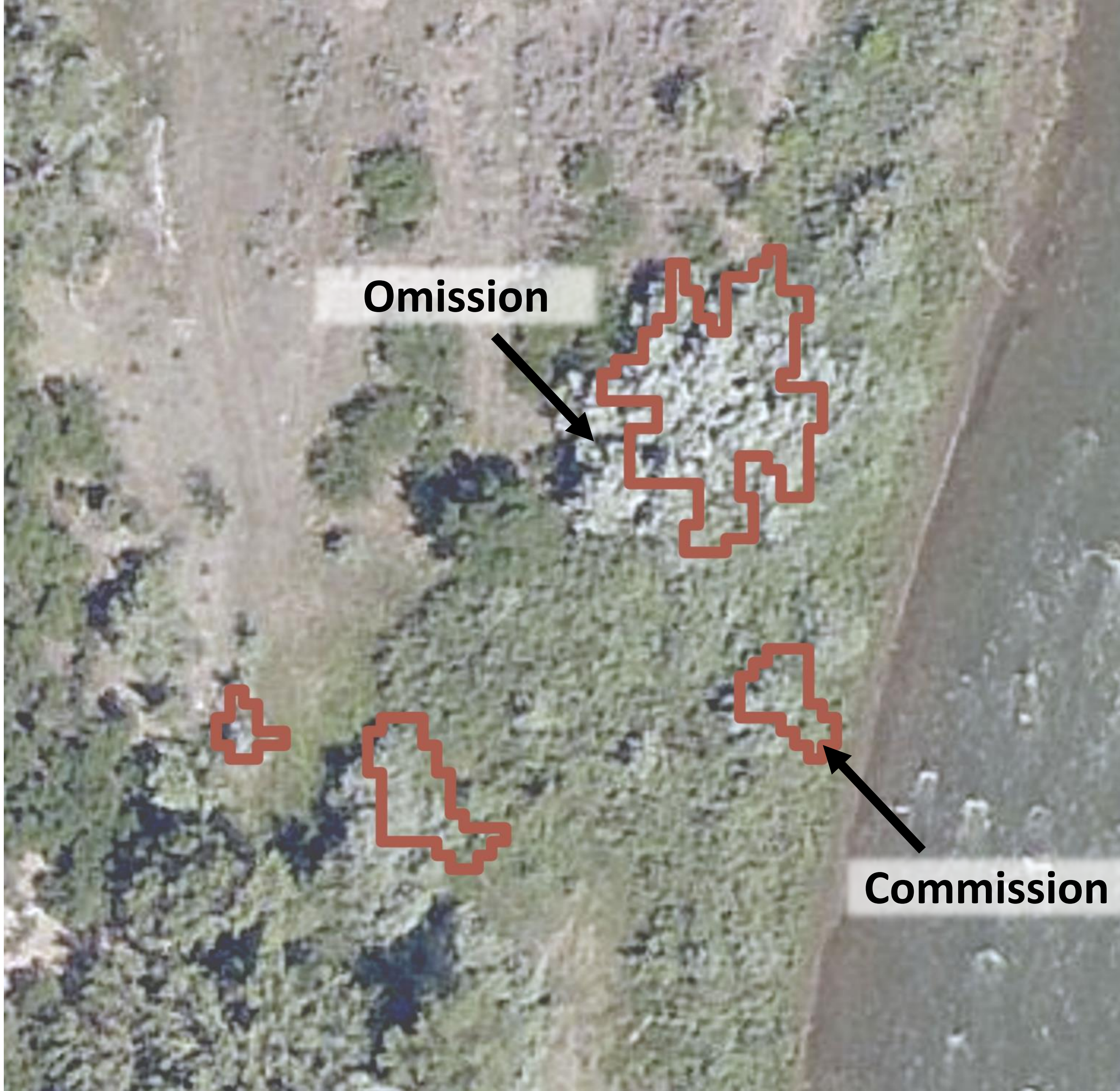
Discussion

The case against Overall Accuracy

- The most important quality of this classification is whether RO is there or not
- Accuracy of other classes is not that important
- Very low OA and Kappa would be unacceptable in a general land cover classification, but in this case they matter little
- This argument has been made by others in the classification of rare classes (Maxwell, Warner, and Fang 2018).

KNN

- Results could be changed/improved by altering user-defined parameters
- For this study, KNN was used due to ease of post-classification editing and higher User's accuracy



Samples

- Classification dependent on quality and quantity of training samples
- Confusion matrix compares results to reference data (testing samples)
- All samples were identified from visual inspection of aerial imagery, and some could be wrong!
- Accuracy of other classes could also be improved by separating into subclasses (i.e coniferous and deciduous)

Other observations

- ¾ of segments removed through manual editing were < 17 m². **Classification confidence increases with object size**.
- Most RO located within 14 meters of river. **Mask areas from analysis not adjacent to water sources** (might exclude other distribution types).

Conclusions & Future Work

Conclusions

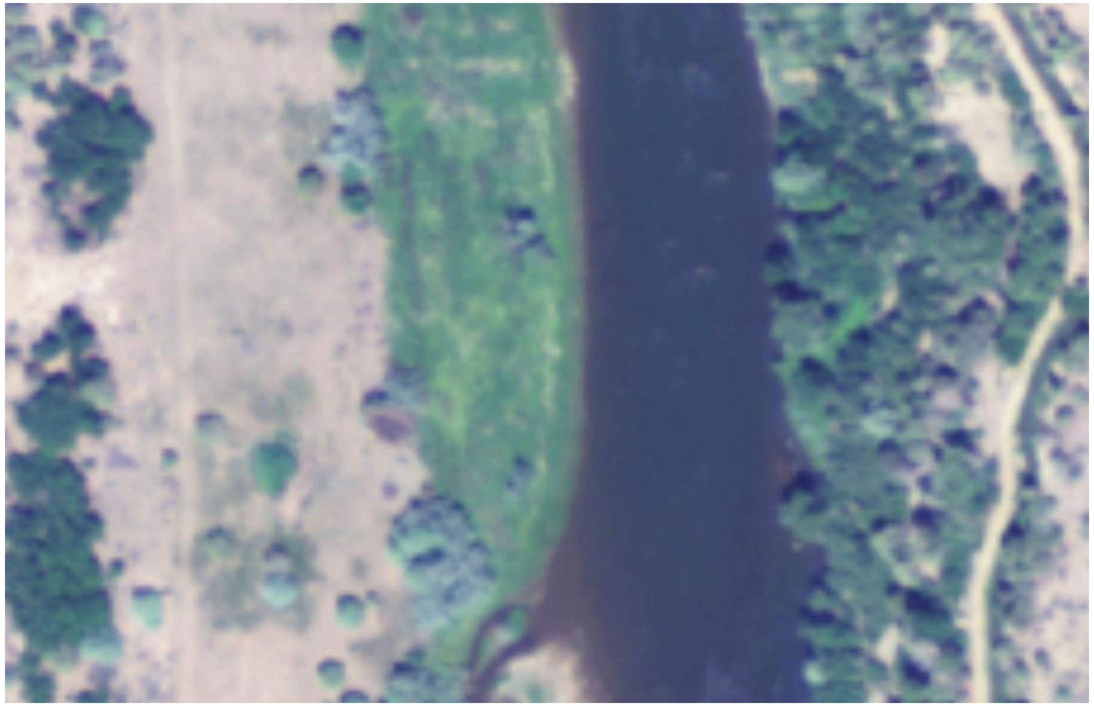
- 63.00% probability that the location of a Russian olive tree will be correctly predicted
- 91.30% probability that each object classified as Russian olive actually represents a Russian olive tree in reality
- Probability increases with object size

Future Work: Regional Distribution

- Use lessons learned to create distribution map of RO in the Animas Subbasin
- This is the ultimate goal of this pilot project!

Future Work: Change over time

2011

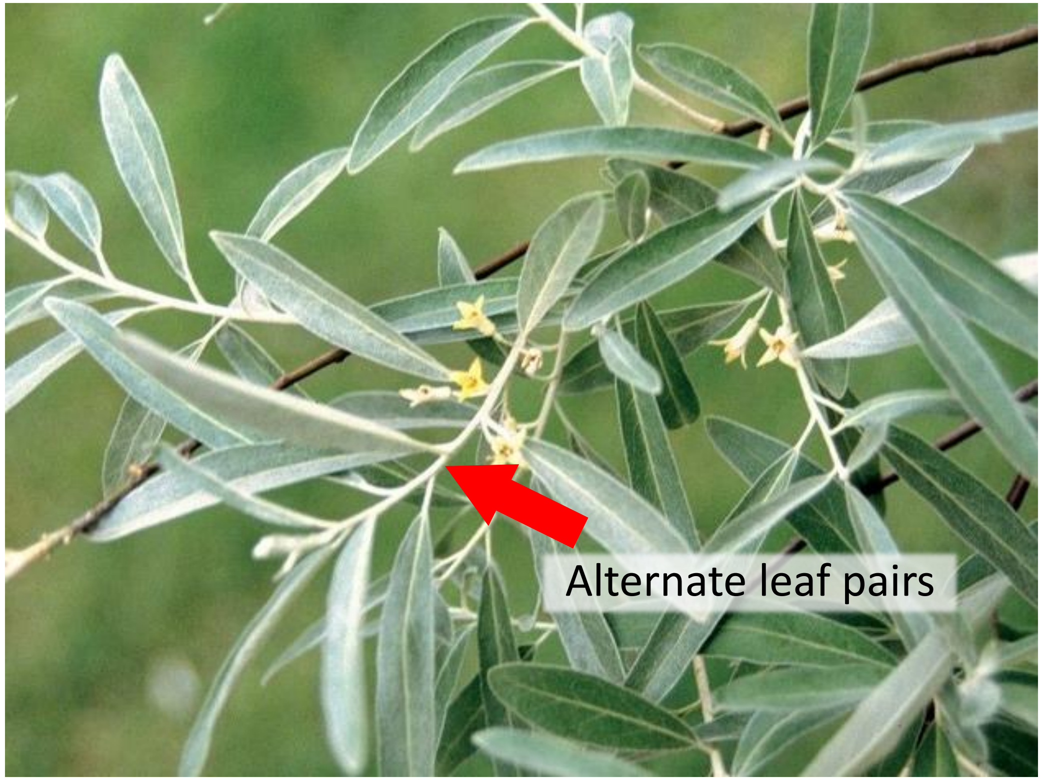


2017

Future Work: Distinguish from Native Silverleaf Buffaloberry

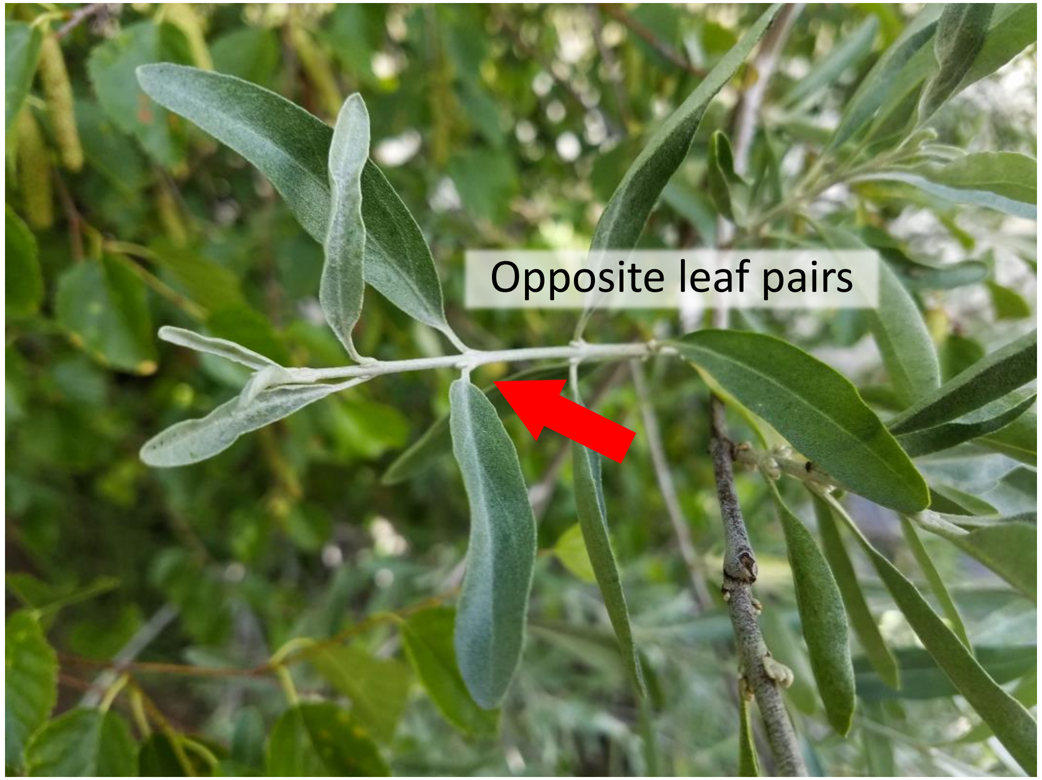
Russian olive

(*Elaeagnus angustifolia*):



Silverleaf Buffaloberry

(Sheperdia argentea)



Thank you!

Story Map
bit.do/russianolive

Anna Riling
annariling@gmail.com
(970) 903-8575

References

Aharoni, Udi. 2007. *SVM with Polynomial Kernel Visualization*. https://www.youtube.com/watch?v=3liCbRZPrZA.

Gandhi, Rohith. 2018. "Support Vector Machine — Introduction to Machine Learning Algorithms." Towards Data Science. June 7, 2018. https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47.

Giggy, Lauren, Amanda Kuenzi, and Scott Roberts. 2018. "Feature Analyst Tool for Mapping Russian Olive: Animas and Florida River Watershed." Mountain Studies Institute.

GrindGIS. 2017. "Vegetation Spectral Signature Cheat Sheet." May 16, 2017. https://grindgis.com/remote-sensing/vegetation-spectral-signature-cheat-sheet.

Hamilton, Randy, Kevin Megown, Henry Lachowski, and Robert Campbell. 2006. *Mapping Russian Olive: Using Remote Sensing to Map an Invasive Tree*.

Jenness, Jeff, Brian Brost, and Paul Beier. 2013. "Land Facet Corridor Designer." USFS Rocky Mountain Research Station. http://www.jennessent.com/downloads/Land_Facet_Tools.pdf.

"K Nearest Neighbor Background." n.d. Accessed June 5, 2019. https://www.harrisgeospatial.com/docs/BackgroundKNN.html.

Li, Xiaoxiao, and Guofan Shao. 2014. "Object-Based Land-Cover Mapping with High Resolution Aerial Photography at a County Scale in Midwestern USA." *Remote Sensing* 6 (11): 11372–90. http://dx.doi.org.du.idm.oclc.org/10.3390/rs61111372.

Tobalske, Claudine, and Linda Vance. 2017. "Predicting the Distribution of Russian Olive Stands in Eastern Montana Valley Bottoms Using NAIP Imagery." 95815401. University of Montana.

"Western Reserve Public Media." 2009. Ubiquitous Science. 2009. https://westernreservepublicmedia.org/ubiscience/electromagnetic.htm.