Acid Mine Drainage (AMD) Detection Using Aerial Four-Band Imagery

Project Team: Scott Miller, Steve Porter, Jennifer Bowman

Voinovich School of Leadership and Public Affairs

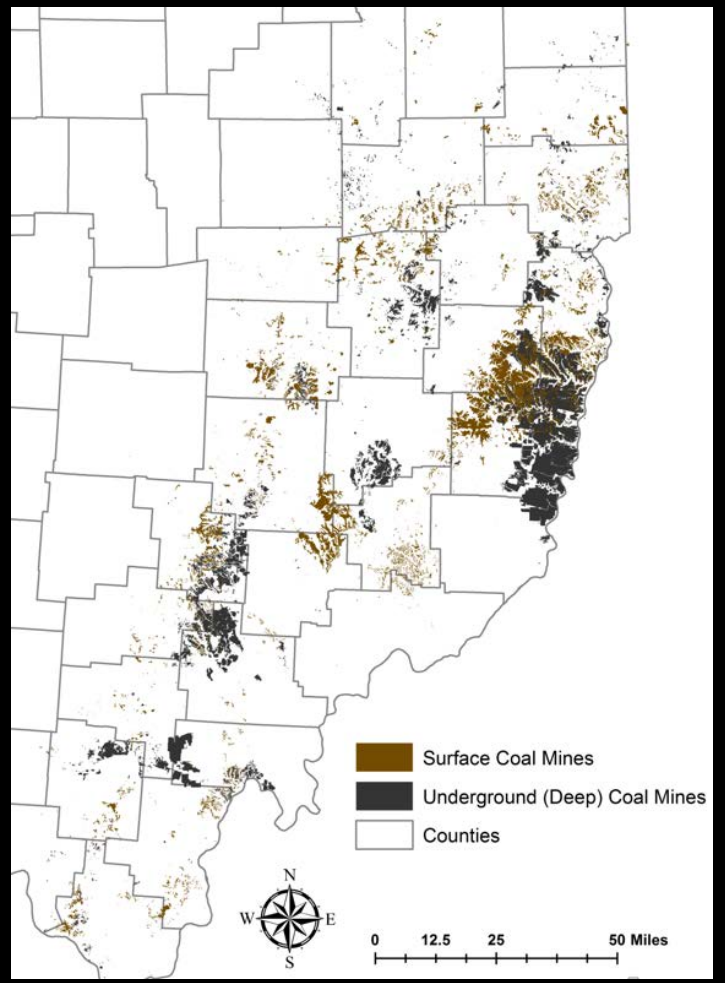
Woolpert, US Air Force Research Laboratory

OHIO UNIVERSITY

Voinovich School of Leadership and Public Affairs

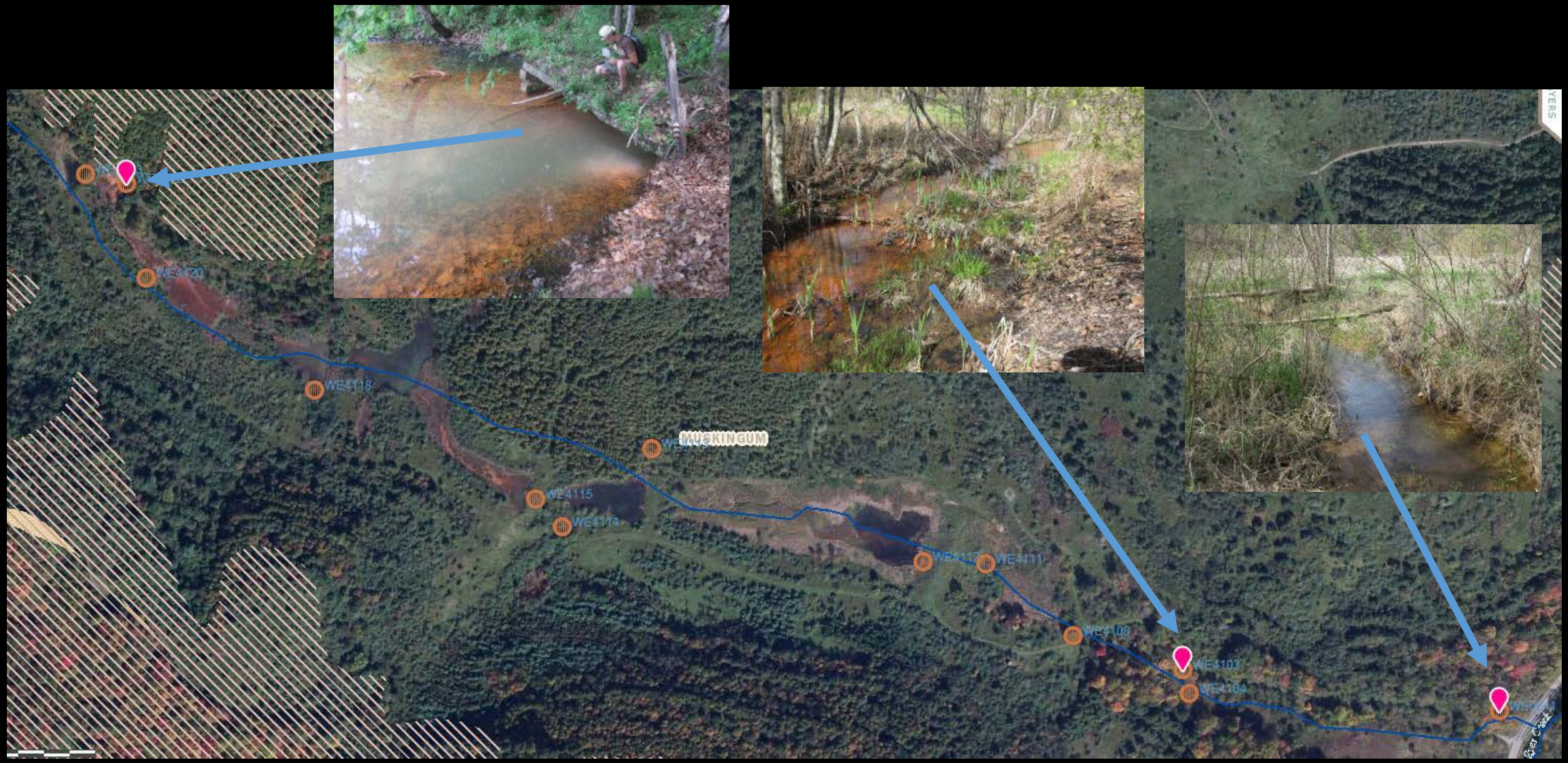
Acid Mine Drainage in the coal bearing region of Ohio

- Abandoned underground mines (1850-1970's)
- High acidity, conductivity, SO_4^{2+} , Fe, Al, Mn
- Hundreds of small/diffuse and large AMD seeps throughout coal bearing region of Ohio



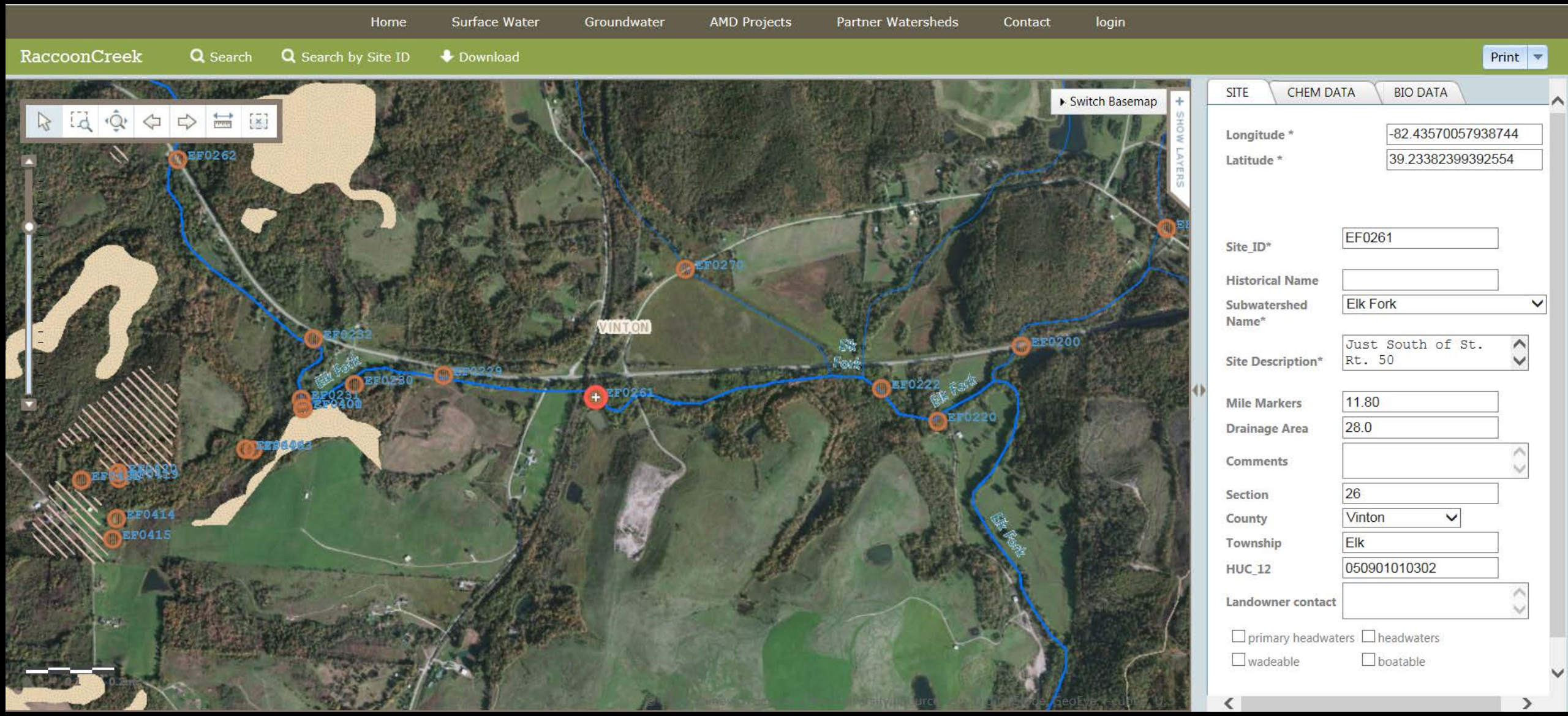
Acid Mine Drainage

Reconnaissance – finding mine sources



Chemical Water Samples

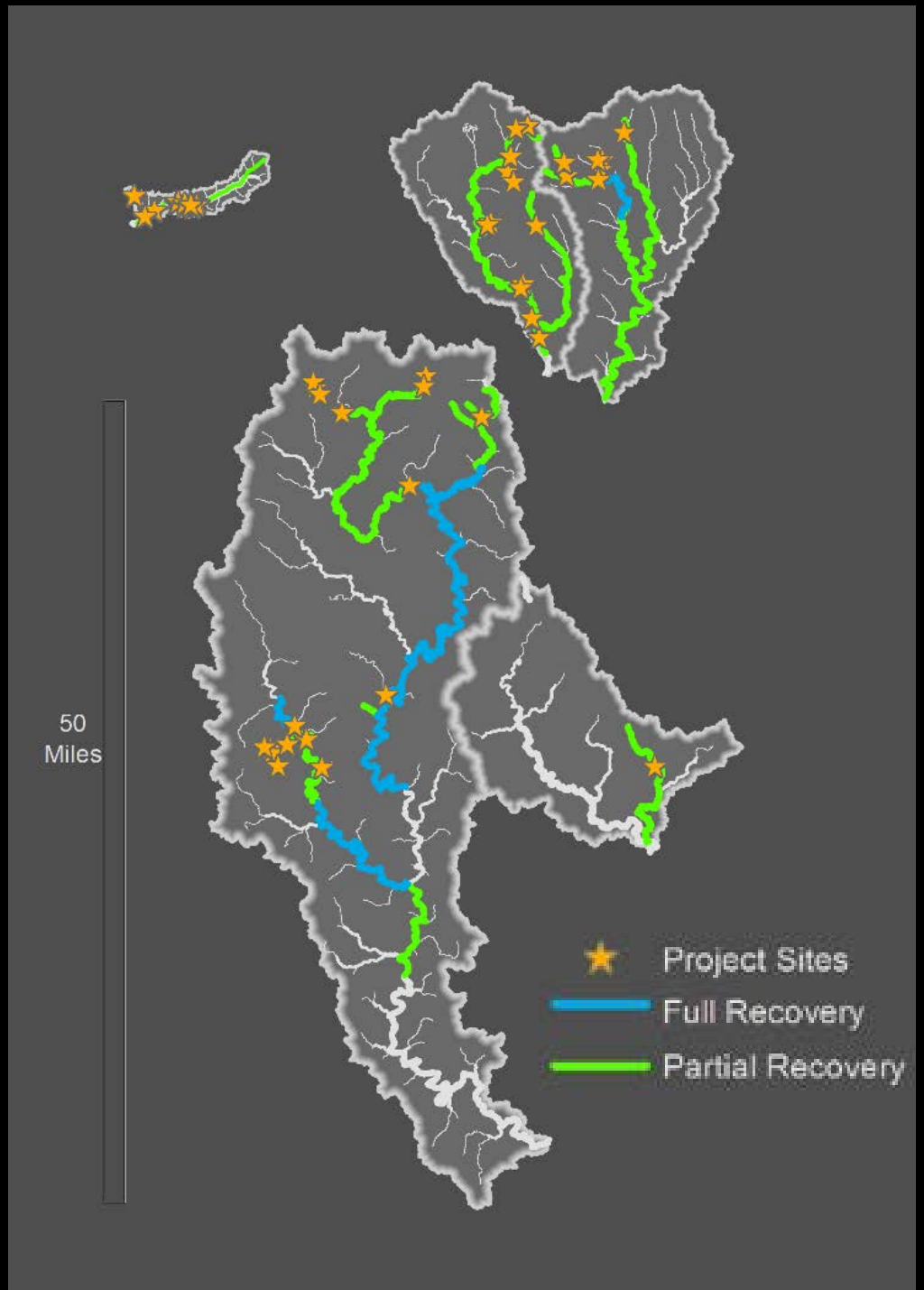
Watersheddata.com



Active Treatment

Biological Recovery

Biological Recovery



Acid Mine Drainage in Appalachia Ohio

- Abandoned underground mines (1900-1970's)
- Acidity, high conductivity, SO_4^{2+} , Fe, Al, Mn

AMD restoration projects complete: 66

Money spent: $30M

Acid Load Reductions: 10,000 lbs / day

Partial recovery: 172 stream miles

Full recovery: 47 stream miles

Acid Mine Drainage (AMD) Detection Study – Developing a new tool to streamline field work

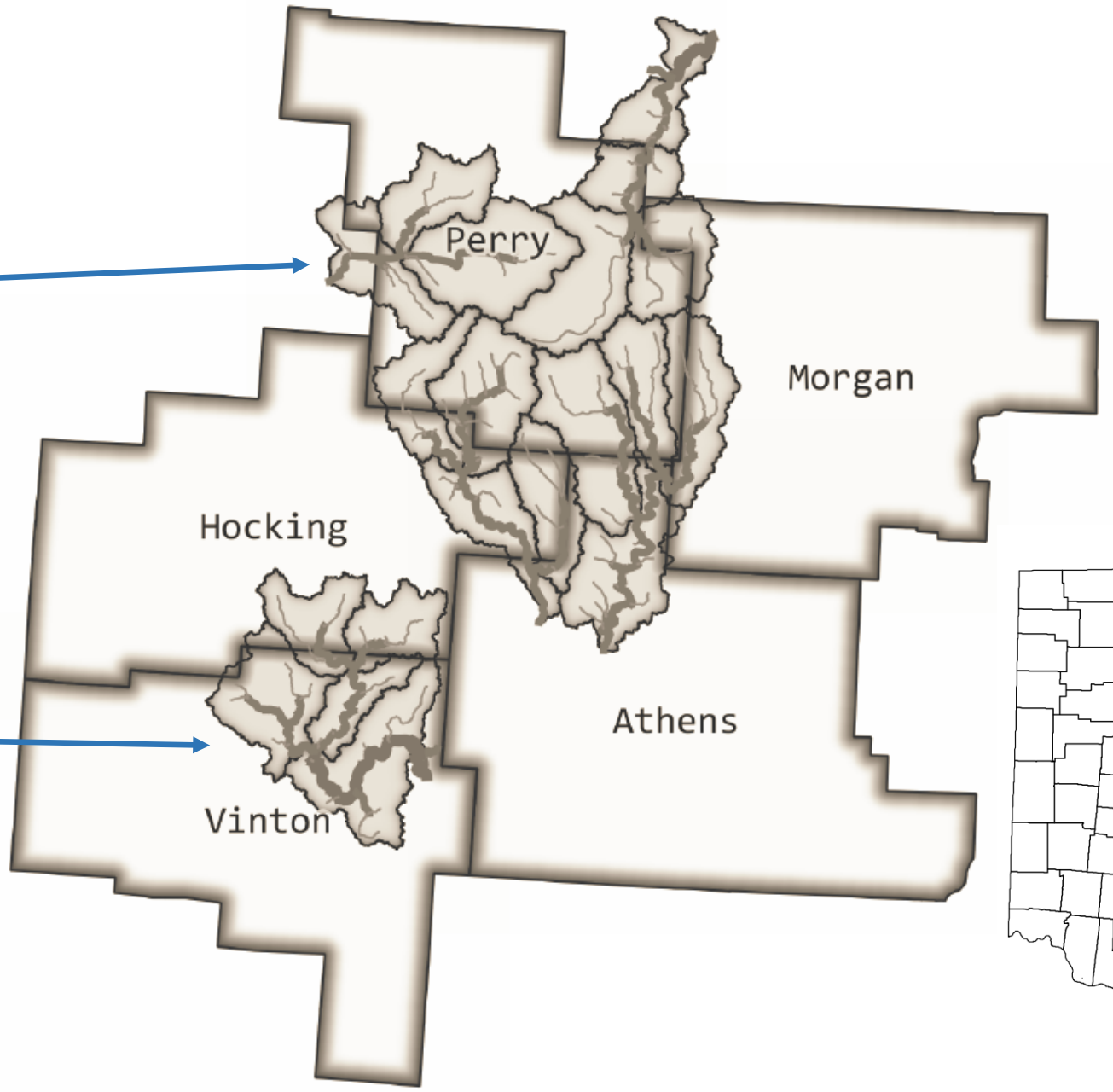
<u>Project Goal:</u> Is it possible to classify streams as AMD-impacted or non-impacted from aerial imagery?

- Focused on using machine learning techniques on unclassified 16-bit aerial four-band (R, G, B, IR) imagery collected by Woolpert as part of Ohio's Statewide Imagery program (OSIP).
  - Athens, Hocking, Vinton 6 inch pixel resolution
  - Morgan, Perry 12 inch pixel resolution

5 County Study Area

Hocking Valley and Muskingum River Coal Field

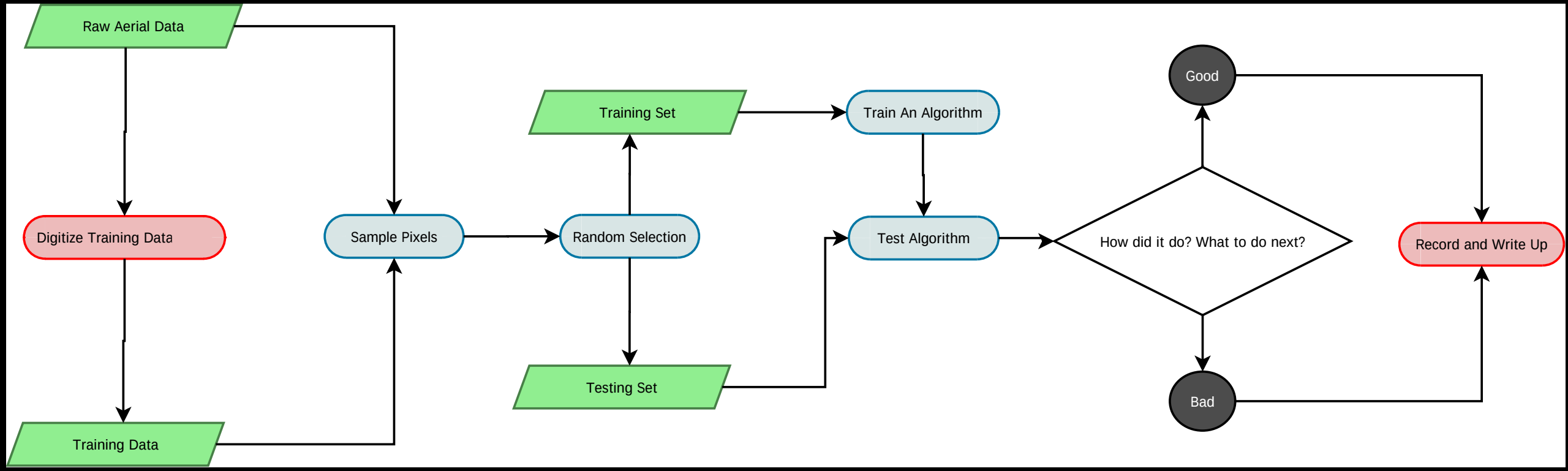
Raccoon Creek



Data Collection

- Aerial Data Collection
  - 16-bit 4-band aerial sensors R, G, B, IR county-wide RAW image files

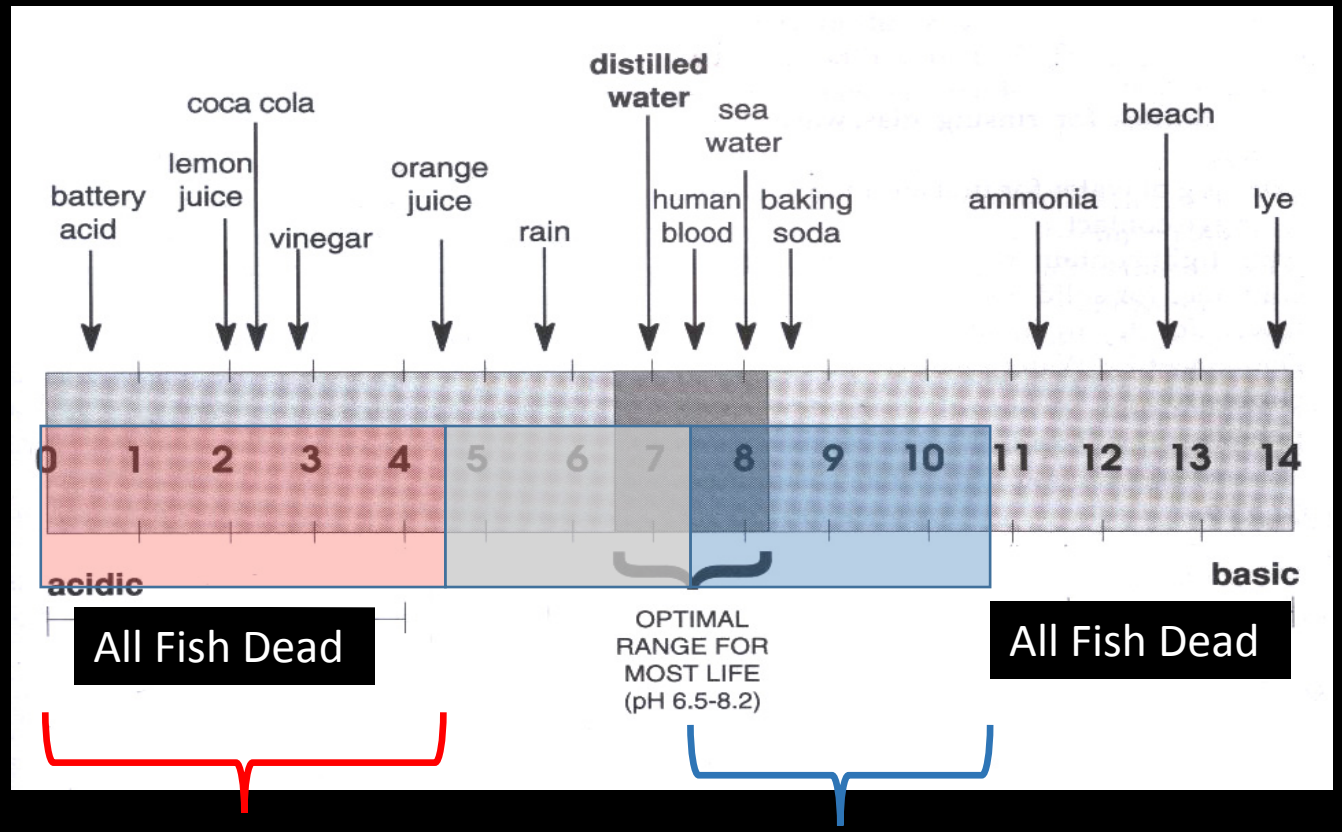
- Ground verification based upon 20 years of lab-analyzed water quality measurements along with lat/long coordinates stored in Ohio University-developed database - watersheddata.com


Raw Aerial Data
Digitize Training Data
Training Data
Sample Pixels
Random Selection
Training Set
Testing Set
Train An Algorithm
Test Algorithm
How did it do? What to do next?
Good
Bad
Record and Write Up

Training Data

- Areas with known AMD from previous samples were used as training data

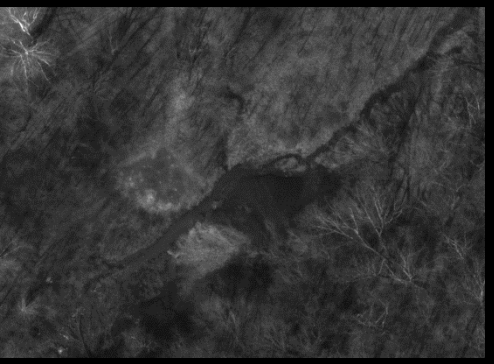
- Assumptions:
  - AMD-impacted streams:
    - pH < 4.5
    - Fe > 1.5 mg/l
    - Al > 1 mg/l
  - Non-AMD-Impacted Streams:
    - pH > 7.5
    - Fe < 1 mg/l
    - Al < 0.75 mg/l



- Data between these boundaries were not used for training data

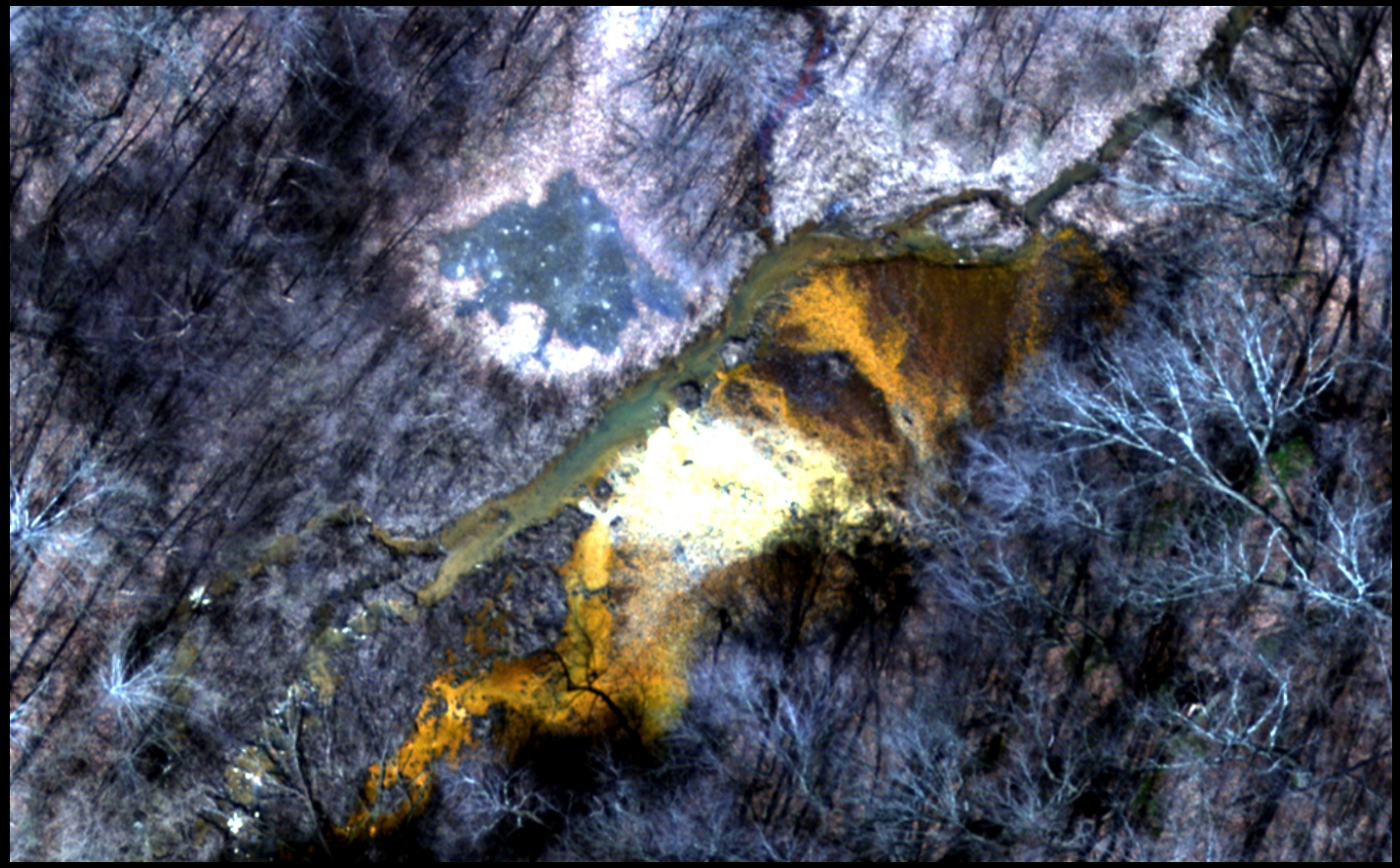
R

G

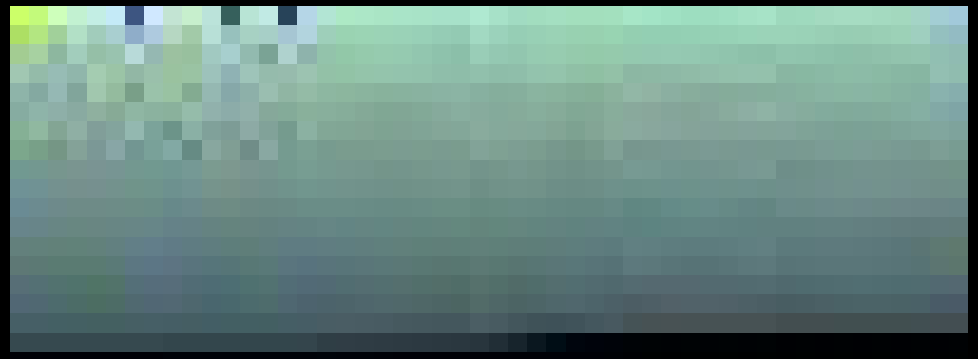


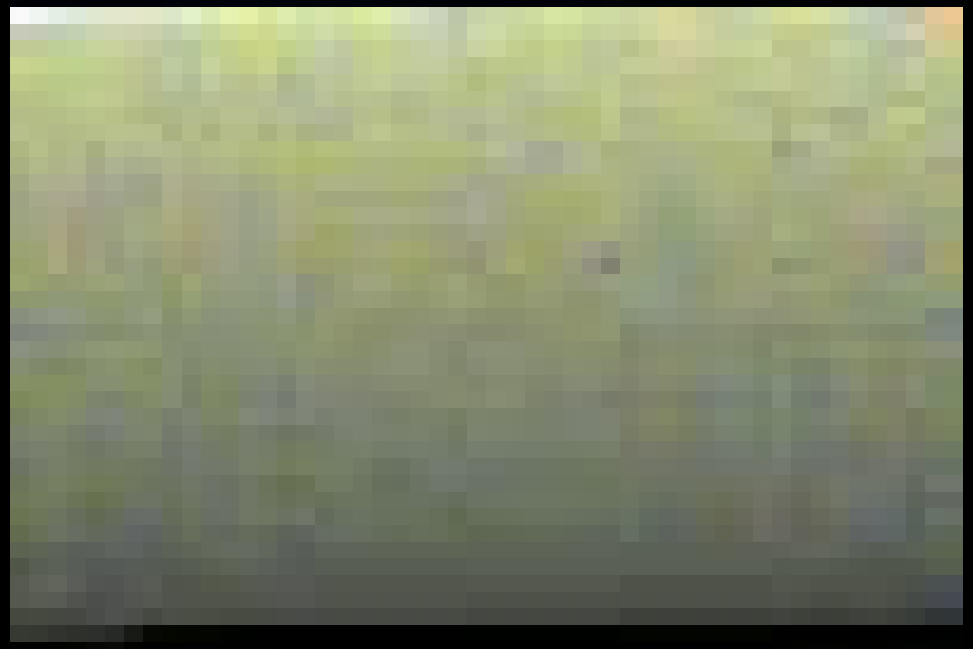
B

I



Training Data





Physical properties of Fe

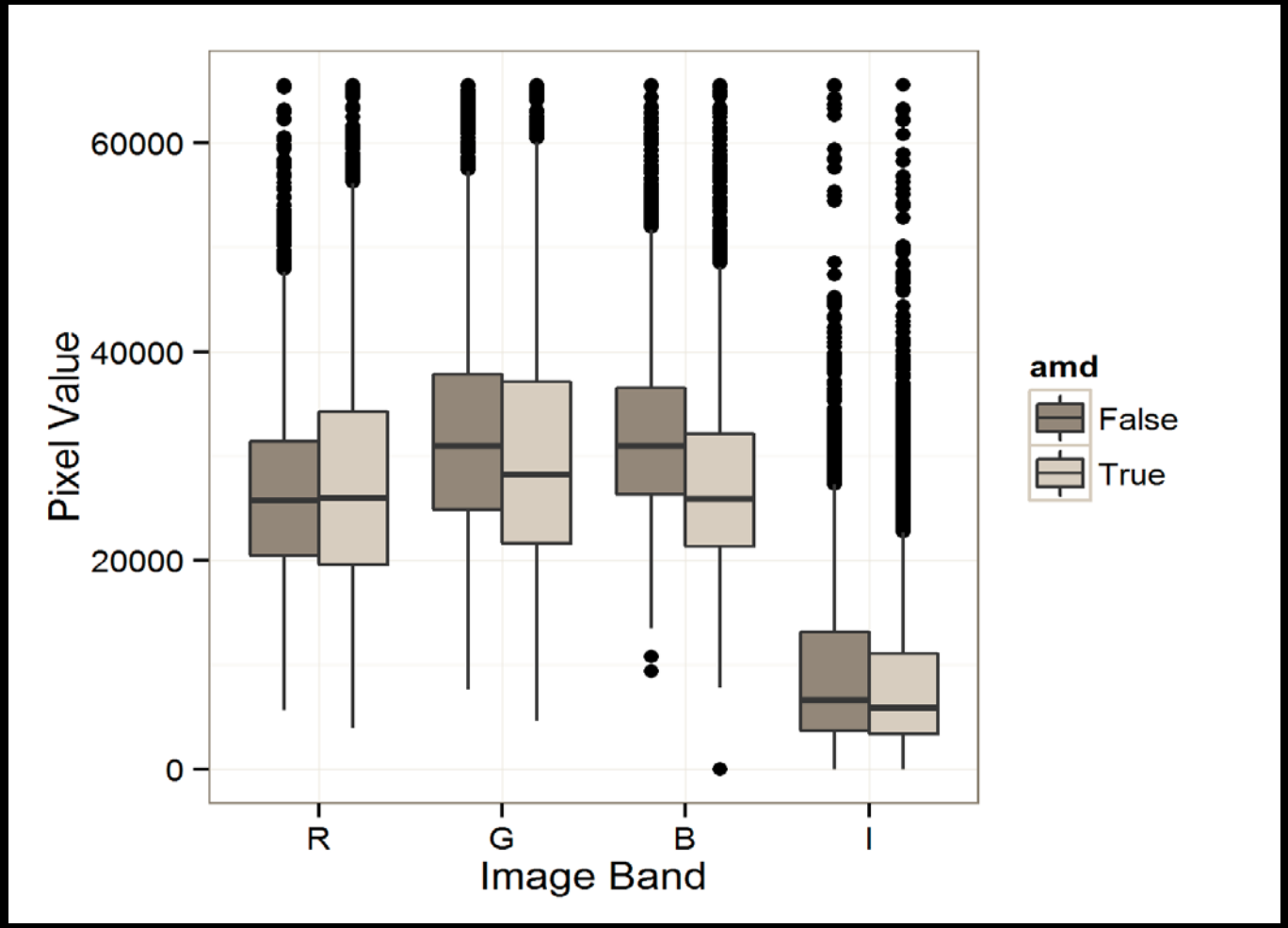
Very Acidic pH<3.5
Water appears healthy and clear, but it is not

pH >3.5 and <6.5
Iron precipitates out of the water
giving AMD classic orange look.

pH>7.5 Non-AMD impacted streams

Training Results

- Training data resulted in 11,812 sampled pixels
  - 8,065 pixels represented AMD-impacted streams
  - 3,747 pixels represented non AMD-impacted streams
- AMD pixels have lower average Green, Blue and Infrared
- “Yellow Boy” – AMD Staining reflects Red = imaging sensor absorbs Blue
- T-tests seem to indicate Blue is the most important band for differentiating between AMD and non-AMD-impacted streams



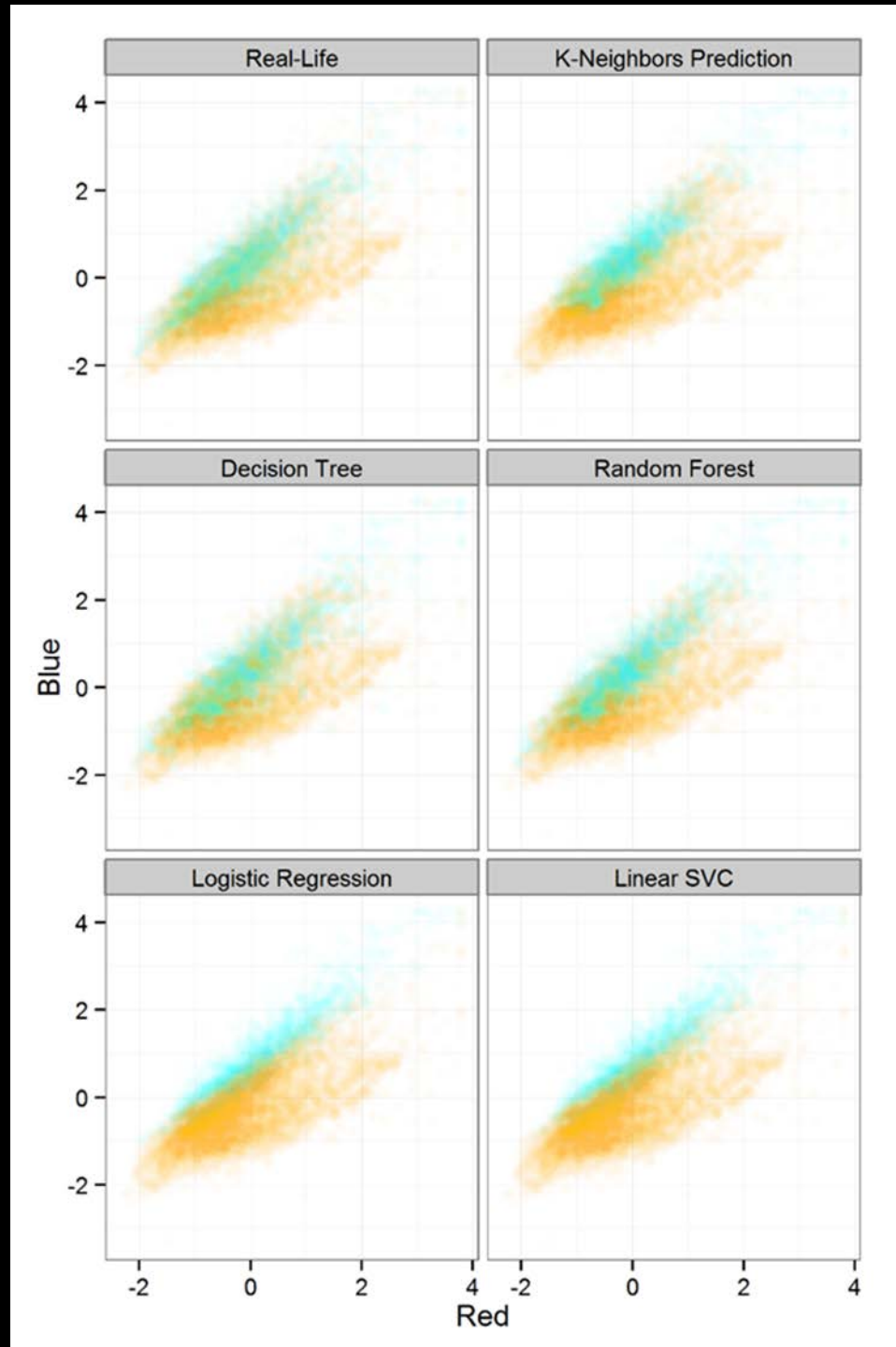
Classification

- 1/3 random selection of the training data was used to train models
- Resulting models were then tested against the other 2/3 of the data
- Five algorithms were used for this study
  - Nearest Neighbor
  - Decision Tree
  - Random Forest
  - Logistic Regression
  - Linear Support Vector Machine

(More Python! Classification done using Scipy, Sklearn, and Pandas)

Results

	Correct Prediction	False Positive	False Negative	% Correctly Predicted
Nearest neighbor	6,907	1,210	1,077	75%
Decision tree	6,448	1,350	1,396	72%
Random forest	6,793	1,086	1,315	74%
Logistic regression	6,650	1,742	802	72%
Linear support vector machine	6,640	1,793	757	72%



Classification Discussion

- Overall, machine learning algorithms correctly classified roughly 72-75% of all sampled pixels with many misclassifications attributed to the same few bodies of water
- Similar numbers of false positives and false negatives were found in all of the models
- Many of the “bad” pixels were from many of the same streams
- False positives, non-AMD streams classified as AMD, seem to be correlated with other phenomena such as turbidity or sedimentation.
- False negatives, AMD streams classified as Non-AMD (i.e. Snow Fork, Brush Fork) where pH of streams were less than 3.5 and clear

Further Discussion

- Agricultural impacts and heavy recreational use (horse and ATV trails) seem to mimic AMD in certain areas with orange sediment loadings
  - Honey Fork, Little Monday Creek, Salt Run
- Extremely low pH streams (pH < 3.5 i.e. Snow Fork) where iron flocculates dissolve back into solution making water seem clear, without the classic orange AMD look
- Additional field work is needed in areas with little or no recent data
  - Indian Run, Little Greens Run
- Training data clearly shows the blue wavelengths being captured by the sensor in lower intensities in red/orange-colored AMD-impacted streams

Recommendations

- Strong performance of statistical classification warrants additional study
- Conduct a larger watershed-level study
- Improved data input – Down select into blue and possibly green wavelengths may enhance predictive ability
- Night time cold weather infrared/thermal data may help in mine portal/ source detection
- Expanded training data needed to better distinguish sediment and agricultural runoff from AMD

No substitute for field work, just another a tool

Voinovich School of Leadership and Public Affairs:
Steve Porter - Geospatial Software Development Engineer

Jen Bowman – Acting Director of Environmental Programs,
bowmanj2@ohio.edu

OHIO UNIVERSITY
Voinovich School of Leadership and Public Affairs