

2024 FECM / NETL Carbon Management Research Project Review Meeting



A Data-driven Multiscale Phytotechnology Framework for Identification and Remediation of Leached-Metals-Contaminated Soil Near Coal Ash Impoundments

Principal Investigator:

Bahareh Nojabaei

Mining and Minerals Engineering, Virginia Tech

Email: baharehn@vt.edu



Our team

- ❑ **Bahareh Nojabaei (PI)**

Mining and Minerals Engineering, data analyses, fluid transport in porous media, underground multiphase multicomponent flow simulation

- ❑ **Hosein Foroutan (Co-PI)**

Civil and Environmental Engineering, Atmospheric science and remote sensing

- ❑ **Hoda Eldardiry (Co-PI)**

Computer Science, Artificial Intelligence and Machine Learning

- ❑ **Ryan Stewart (Co-PI)**

School of Plant and Environmental Sciences, Quantifying interactions between water, soil, and plant communities

- ❑ **Lee Daniels (senior personnel)**

School of Plant and Environmental Sciences , Rehabilitation of drastically disturbed soils

- ❑ **Mining Engineering Graduate Student: Salman Karim**

- ❑ **Computer Science Graduate Students**

Background: Coal Combustion Product Impoundments

- Many older impoundments are not lined.

Contamination

BORON, CHLORIDE, CHROMIUM, IRON, MANGANESE, SELENIUM, TOTAL DISSOLVED SOLIDS, NITRATE, SULFATE AND THALLIUM, ARSENIC, LEAD, AND PH.

Unlined Impoundment	
Impoundment Name	1964 Ash Pond
Hazard Rating	High
Capacity	72,664,773 gallons
Ash	2,535,600 tons
Wastewater	9,776,000 gallons
Dam Height	100
Active	Yes

Unlined Impoundment	
Impoundment Name	1982 Ash Pond
Hazard Rating	High
Capacity	290,007,390 gallons
Ash	686,000 tons
Dam Height	95
Active	Yes

Unlined Impoundment	
Impoundment Name	Stilling Pond
Hazard Rating	Not Rated
Active	Yes

Asheville Power Station

Arden, NC 28704

Owned and Operated by Duke Energy Corporation

Hazard Rating: **High**



Office of Fossil Energy and Carbon Management (FECM)
UNIVERSITY TRAINING AND RESEARCH FOR FOSSIL ENERGY AND CARBON
MANAGEMENT - UCR

Phytotechnology Development for Identification and/or Remediation of Sites Exhibiting Soil Contamination via Groundwater Transport of Metals from Coal Combustion Product Impoundments

- Identification of sites that need more aggressive monitoring and remediation
- Low-cost technology is needed to screen sites
- Facilitate a treatment development and deployment
- No removal of the affected soils or groundwater

Coal combustion product impoundments in the region

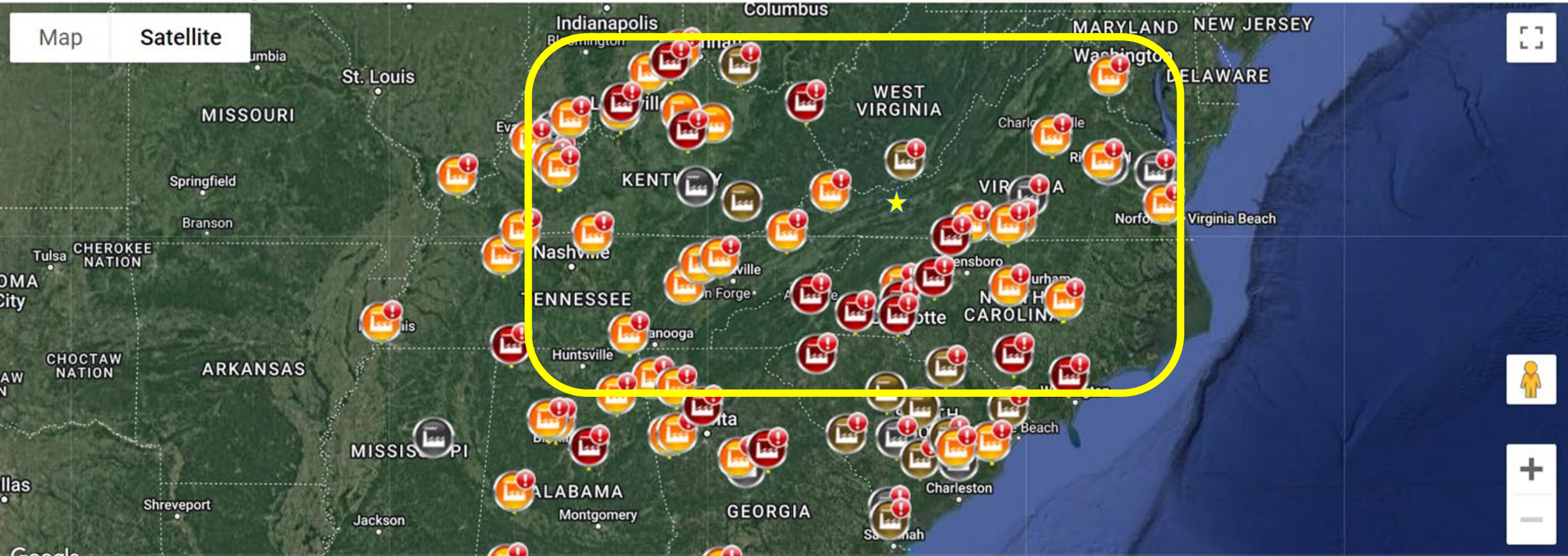


Map View

Take Action!

About Coal Ash

Learn More



Primary goals

- Using **phytotechnology** to identify leached-metal contamination
- **Not being limited to a few sites** (site identification should be the outcome, screening is important)
- **Not being limited to a few number of metal-leached contaminants**
- Facilitate (future) **development of phytoremediation**

Remote sensing: Satellite imagery

Benefits:

- Low-cost technology
- Searching for visual cues
- Extended coverage over multiple states
- No site access needed.

Challenges:

- What kind of visual cues we are searching for?
- Low resolution
- What type of metal-leached contamination we are looking for?

Image processing: searching for abnormalities

We don't know what the exact visual cues are!

Examples of contamination bio-indicators:

- Cadmium-induced stress for rice crops, **indicator:** color change (Liu et al., 2018)
- Cement dust pollution, **indicator:** effect on plant species density and diversity (Bayouli et al., 2021)
- Mercury in fly ash, **indicator:** presence in soil and vegetation samples (Huang et al., 2017)



Search for both temporal and spatial abnormalities

Factoring out other influencers

A Data-driven Smart Search!

- Climate and environmental conditions
- Soil type and properties, pH, moisture
- Structural conditions: underground geological features

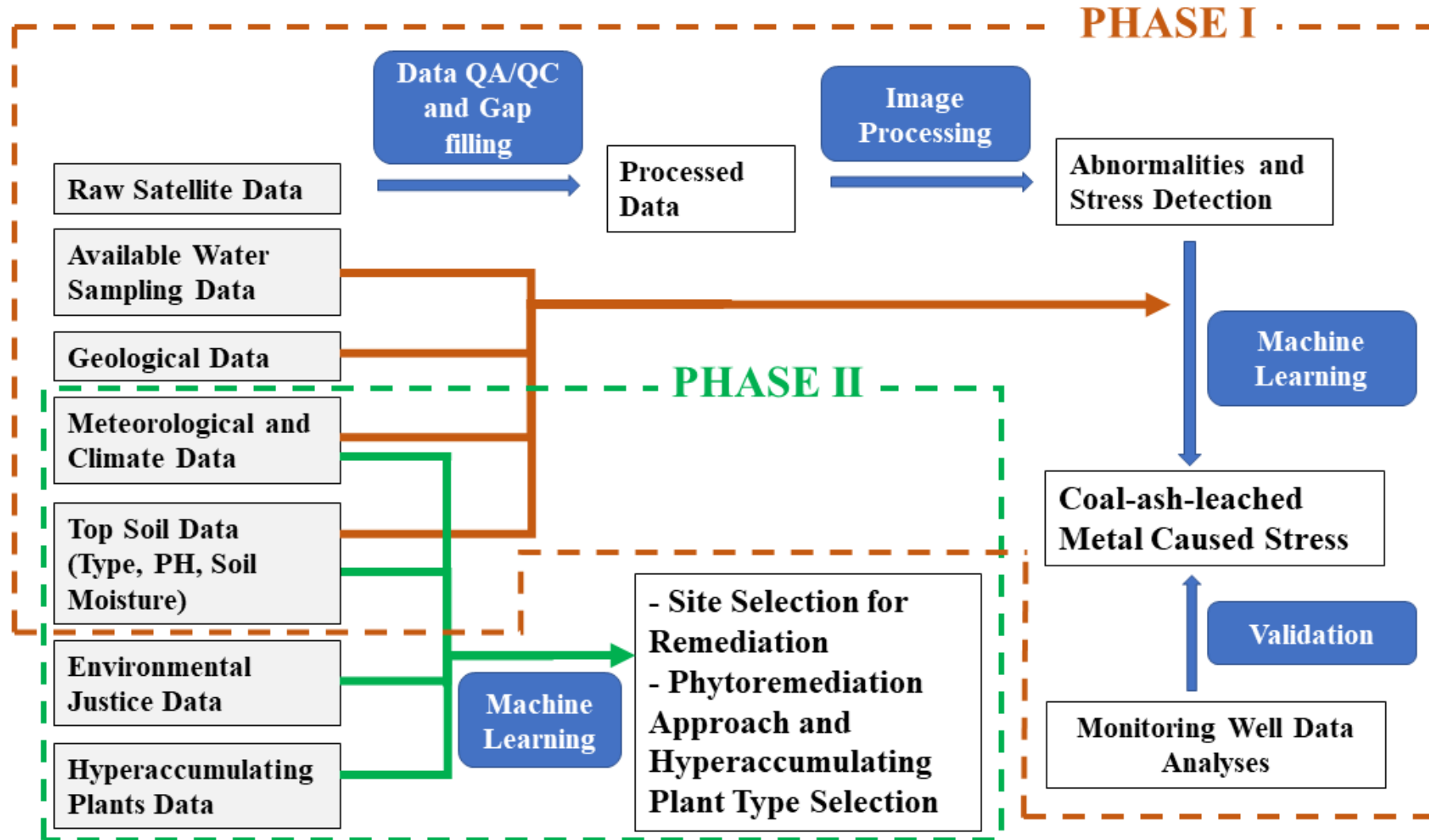
How to validate our model?

- Underground contamination data from monitoring wells
- Geological modeling

Phytoremediation

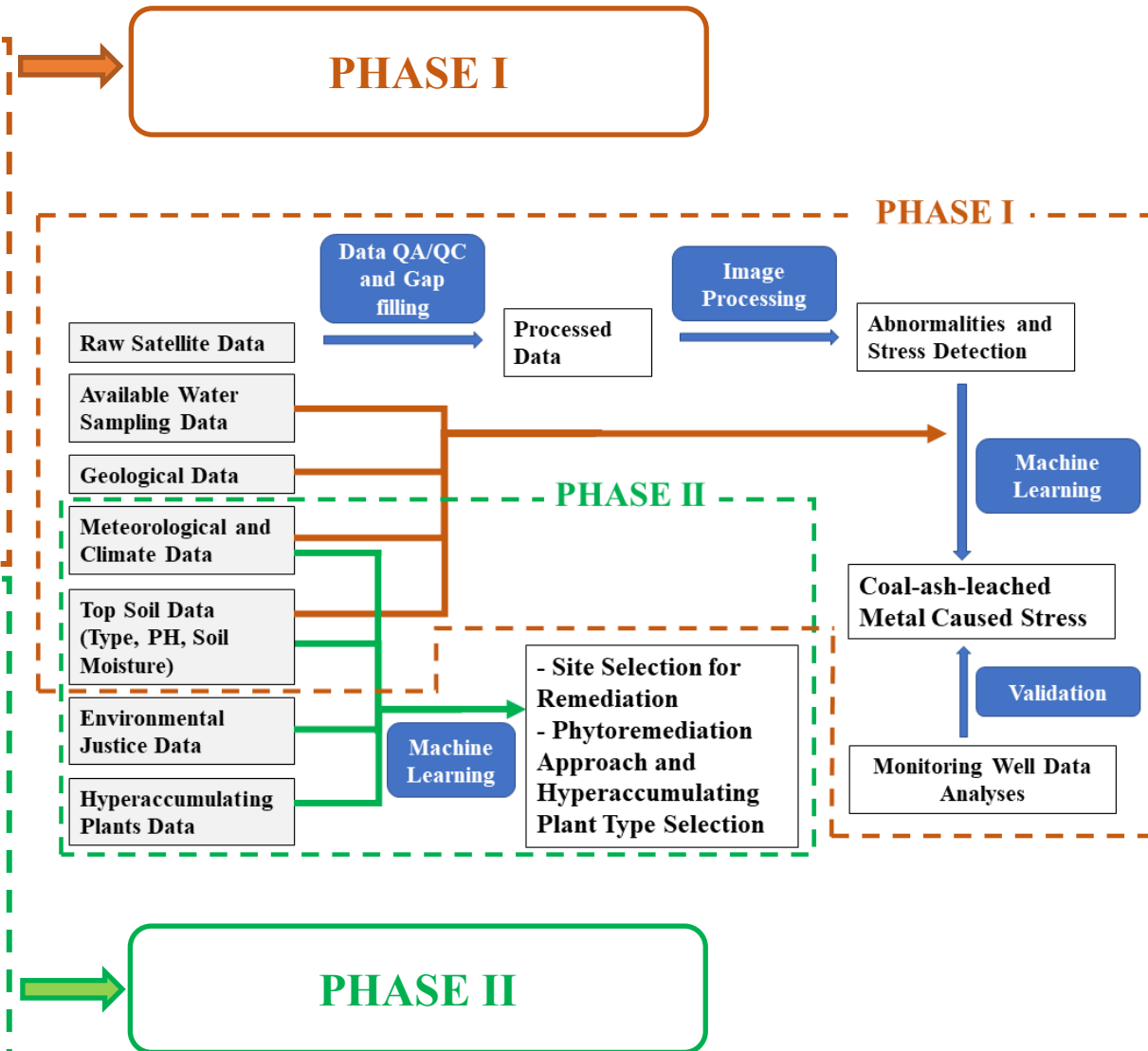
- Database of hyperaccumulating plants is available.
- How to select a site as a phytoremediation candidate? How can we rank them?
- What are the factors we should consider in order to select plant species?

Multiscale data-driven phytoindication and phytoremediation framework



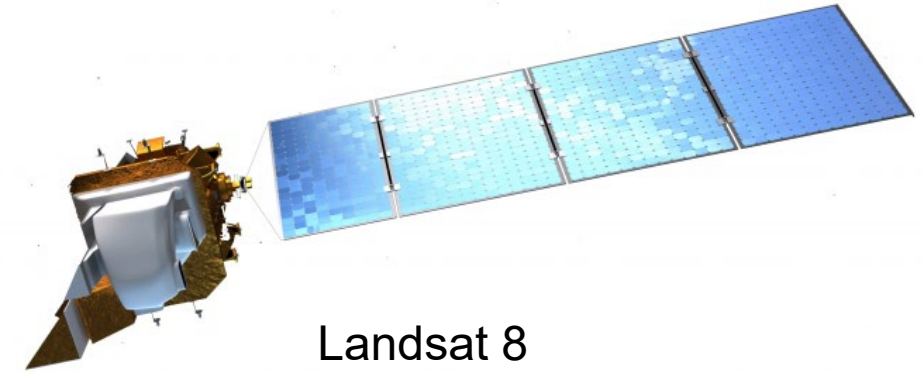
Main project objectives

- Process and analyze multi-year satellite retrievals to find vegetation-related abnormalities.
- Develop a model correlating the leached-metal contamination and the above-mentioned abnormalities.
- Validate our data-driven model for a limited number of sites.
- Propose a phytoremediation approach based on the specifics of hyperaccumulator plants, sites, and environmental and geological conditions.
- Select and rank site candidates for phytoremediation based on environmental and meteorological conditions, soil type and condition, severity of contamination, and environmental justice metrics.

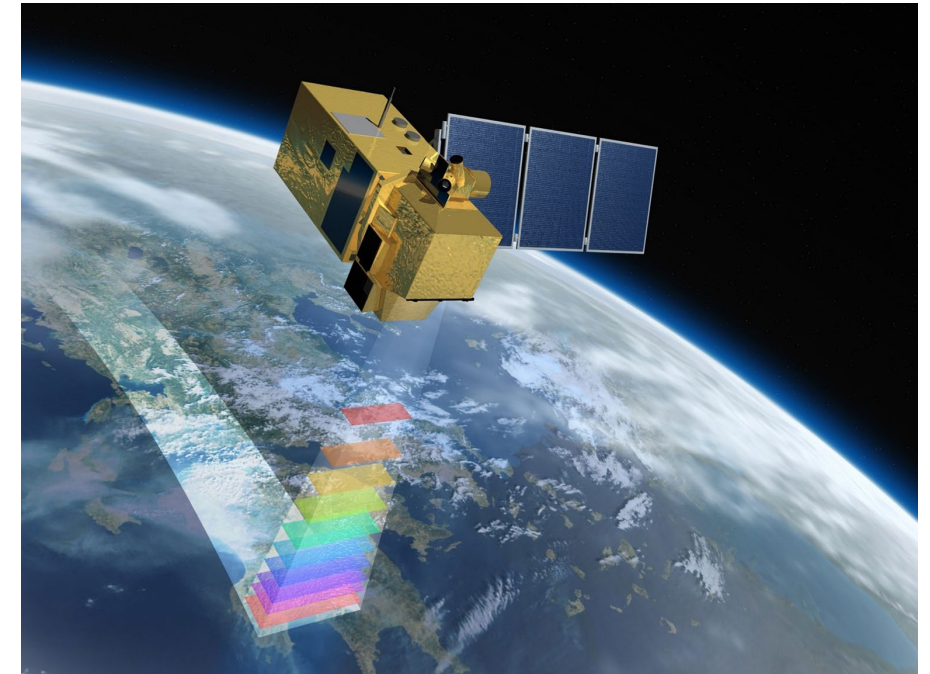


Multiplatform satellite observations

- **Landsat 8 OLI (Operational Land Imager)**
 - Launched in 2013
 - Resolution : 30 m
 - Revisit cycle : 16 days
 - Number of bands : 9
- **Landsat 7 ETM+ (Enhanced Thematic Mapper Plus)**
 - Launched in 1999
 - Resolution : 30 m
 - Revisit cycle : 16 days
 - Number of bands: 8
- **Sentinel-2**
 - Launched in 2015
 - Resolution : 10-20 m
 - Revisit cycle: 10 days
 - Number of bands: 13



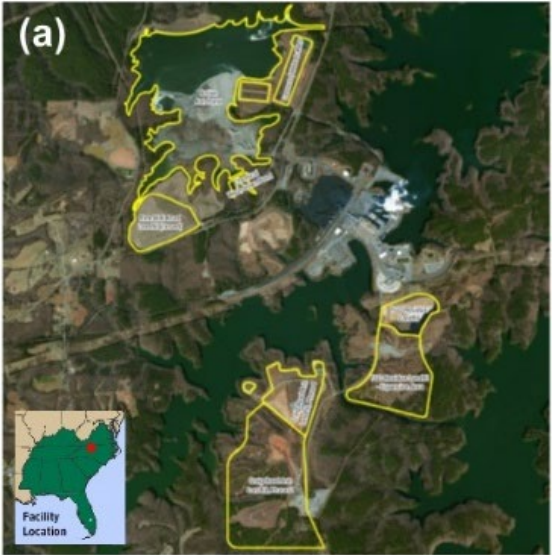
Landsat 8



Sentinel-2

Multiband Capability of Landsat 8 OLI

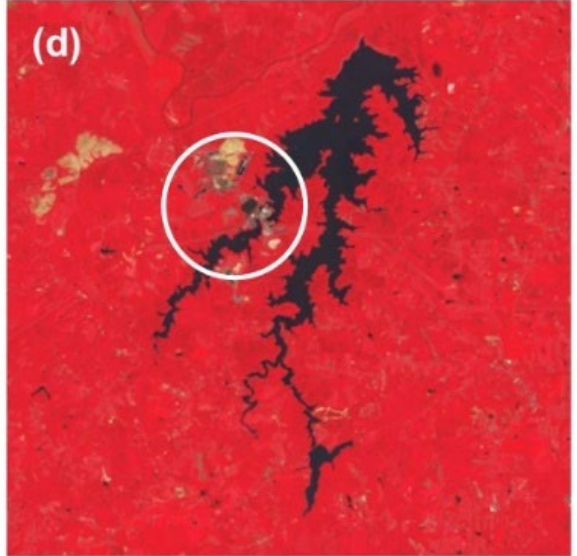
Belews Creek Stream Station in Walnut Cove, NC that has a "high hazard" rating



Satellite view with natural color



Satellite view with false color (combination of bands 4, 5, and 6)



Satellite view with color infrared (combination of bands 3, 4, and 5)



Satellite data post-processing

Raw Satellite Data

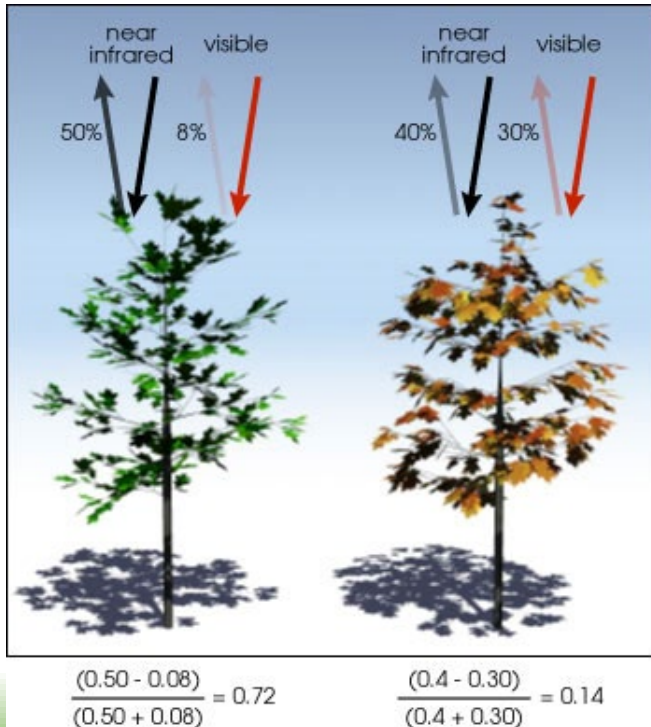
- Landsat 7 & 8
- Sentinel-2

Data Post-Processing

- QA/QC
- Filter out clouds
- Gap-fill
- Atmospheric correction

Vegetation Indices

- NDVI
- EVI
- NDWI1 & NDWI2
- CIRE
- REP
- MTCI
- NDRE1 & NDRE2



NDVI: Normalized Difference Vegetation Index
REP: Red Edge Position

“NDVI is calculated from the visible and near-infrared light reflected by vegetation. Healthy vegetation (left) absorbs most of the visible light that hits it, and reflects a large portion of the near-infrared light. Unhealthy or sparse vegetation (right) reflects more visible light and less near-infrared light.”

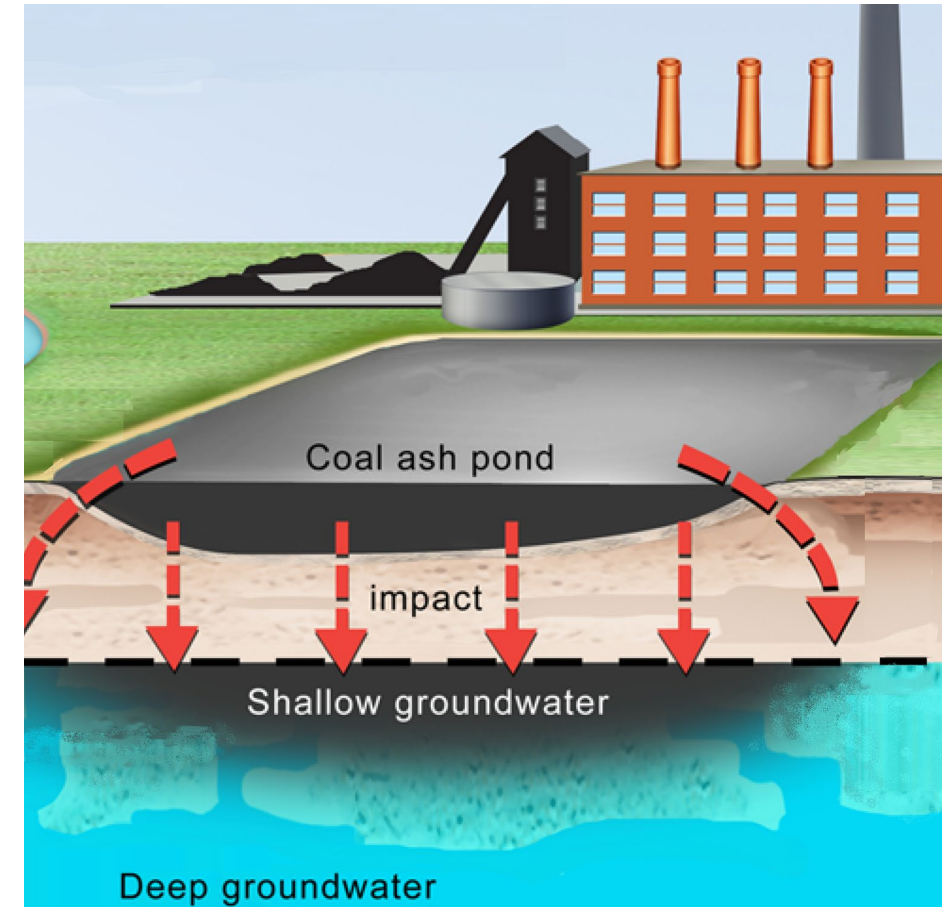
Contamination transport and leachability of contaminants

- Potential contaminants:

Cadmium, Lithium, Molybdenum, Boron, Arsenic, Manganese, Aluminum, Chromium, Selenium, and Mercury are present in the coal combustion products.

The degree of leachability and diffusion into the ground water system is different!

- Metal type and concentration,
- The aqueous phase pH,
- Soil and underground formation porosity and permeability
- Presence of fractures, faults, or any pathway from the impoundment to the underground water system



Underground hydrological modeling for the sites with abnormalities

Input data:

- Contamination report from monitoring wells
- Geological maps of the flagged impoundments
- Available soil type and properties
- Local hydrological data

Approach:

Mass conservation equation will be solved along with, Darcy-type flow, diffusion and solubility of metal contaminants.

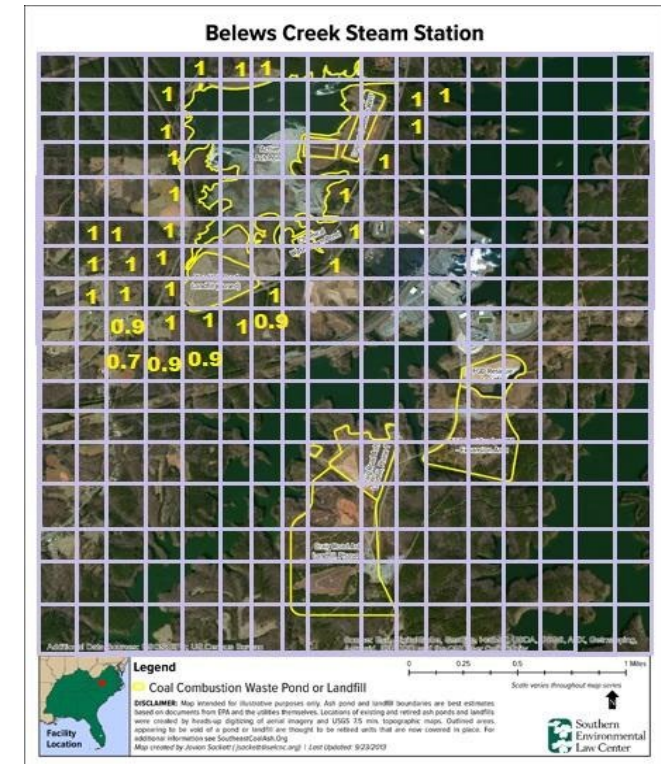
Goal:

Map a probability function to the studied area

One additional assessment tool to identify:

- 1) whether the observed abnormality is because of leached metal contamination,
- 2) the type of leached metal(s) potentially contaminating the soil.

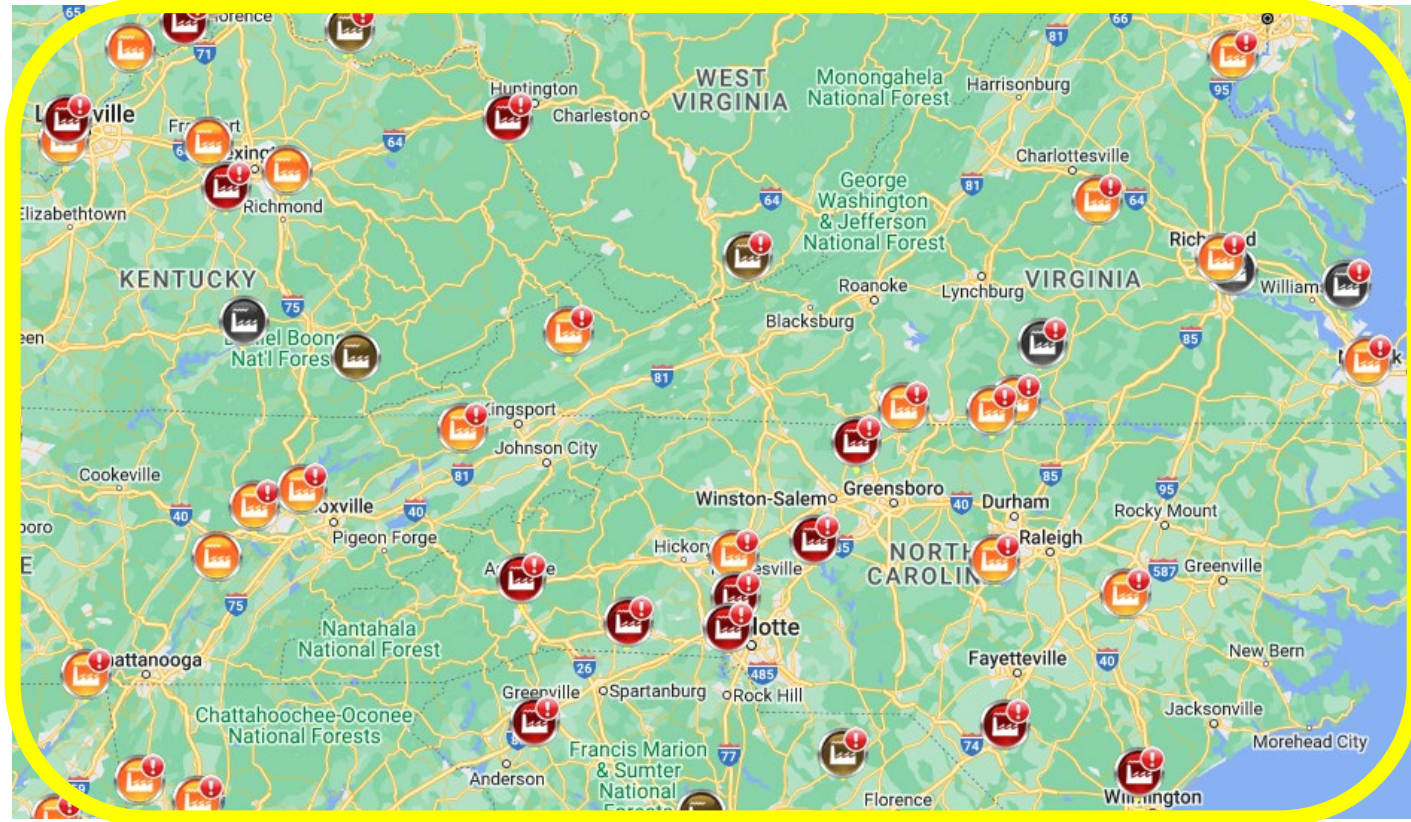
The probability function is a 2D map and provides values from zero to one for each pixel/location, zero being the lowest and one being the highest probability of any specific metal leached from the impoundment to that location's soil.



Selenium release from this plant caused local extinction of 80% all fish species in Belews Lake.

Overview of study area

- 43 coal powerplants spanning in 7 states and owned & operated by 11 different companies.
- Built between 1940 to 1992 and 70% of powerplants are active.
- Generate 36663 MW in total and operate with 133 boilers.
- 31 different pollutants have been found (no contamination data is available for 8 plants).
- All powerplants have one or more associated contamination above safe threshold.
- Total number of known impoundments are 216.

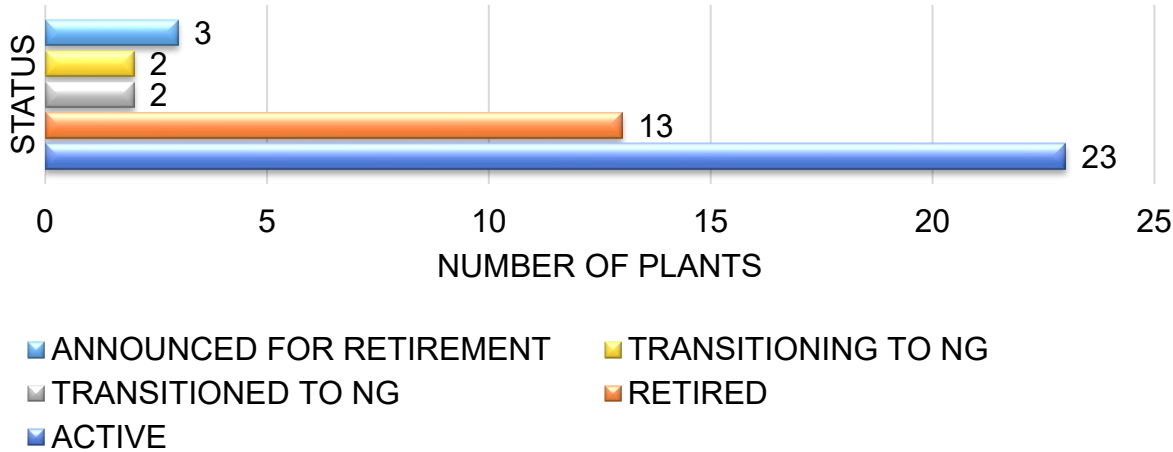


ASHTRACKER.ORG
SOUTHEASTCOALASH.ORG

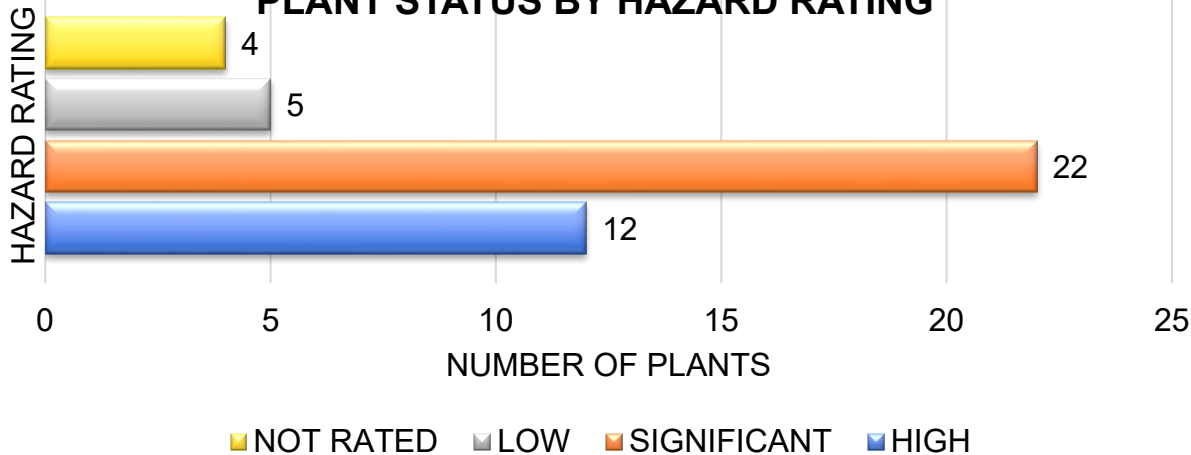
Available data

Some Statistics

PLANT ACTIVITY STATUS BY NUMBERS



PLANT STATUS BY HAZARD RATING



Constituents	Threshold	No of Plants above threshold
Arsenic	0.01 mg/L	27
Cobalt	0.006 mg/L	25
Boron	3.0 mg/L	23
Molybdenum	0.04 mg/L	23
Sulfate	500.0 mg/L	23
Manganese	0.3 mg/L	22
Lithium	0.04 mg/L	19
Lead	0.015 mg/L	16
Radium	0.005 mg/L	15
Beryllium	0.004 mg/L	11
Cadmium	0.005 mg/L	11
Selenium	0.05 mg/L	9
Thallium	0.002 mg/L	9
Chromium	0.1 mg/L	7
Antimony	0.006 mg/L	6
Barium	2.0 mg/L	5
Fluoride	4.0 mg/L	3
Nickel	0.1 mg/L	3
Strontium	4.0 mg/L	3
Mercury	0.002 mg/L	2
Ammonia	30.0 mg/L	1
Nitrate	10.0 mg/L	1

Google Earth Engine Workflow

- Open google earth engine code editor.
- Select region of interest and set map center.
- Load satellite-2 image collection from google cloud.
- Filter image collection with bands, datetime, cloud cover and region of interest.
- Add visualization parameters to visualize better.
- Add image collection to map.
- Define functions to calculate vegetation indices and biophysical parameters.
- Add calculated VI/BPI to image collection and add to map.
- Chart VI/BPI.

A simple code snippet is added here.

Check the app developed for all the calculation:

[Vegetation Indices Calculator](#)

```
New Script *
Get Link Save Run Reset Apps
Imports (1 entry)
var roi: Polygon, 4 vertices
1 //Select Region of Interest nad add to map
2 Map.addLayer(roi, {}, 'ROI', false);
3 Map.setCenter(-80.0641, 36.2922, 14);
4 //Load Image collection
5 var imColl = ee.ImageCollection('COPERNICUS/S2_SR_HARMONIZED')
6     .filterBounds(roi)
7     .filterDate('2017-03-28', '2023-03-28')
8     .select('B1','B2','B3','B4','B5','B6','B7','B8','B8A','B9','B11','B12','QA60')
9     .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE',30));
10 // add image collection to map
11 var visParam={ bands: ["B4","B3","B2"], gamma: 1, max: 2994.6, min: 198.25714285714287, opacity: 1};
12 Map.addLayer(imColl.median().clip(roi), visParam, ' True Color Image', true);
13 // Calculate EVI and the calculatd EVI band to image collection
14 var addEVI = function(image) {
15     var calEVI = image.expression(
16         '(2.5 * float((B8-B4)/( B8 + (6 * B4)-(7.5 * B2) + 1)))',
17         {
18             'B8': image.select('B8'),
19             'B4': image.select('B4'),
20             'B2': image.select('B2'),
21         }).rename('EVI');
22     return image.addBands(calEVI);
23 };
24 var evi = imColl.map(addEVI).select('EVI');
25 // Chart EVI
26 var chart_evi = ui.Chart.image.series({
27     imageCollection: evi,
28     region: roi,
29     reducer: ee.Reducer.mean(),
30     scale: 30
31 })
32 .setOptions({
33     interpolateNulls: true,
34     lineWidth: 1,
35     pointSize: 3,
36     title: 'Mean EVI over Time',
37     vAxis: {title: 'EVI',titleTextStyle: {italic: false, bold: true}},
38     hAxis: {title: 'Date', format: 'YYYY-MMM', gridlines: {count: 20},titleTextStyle: {italic: false, bold: true}}
39 });
40 Map.addLayer(evi.median(), {}, 'EVI', false);
42
```

Vegetation indices calculation

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

$$NDWI1 = \frac{NIR - SWIR1}{NIR + SWIR1}$$

$$NDWI2 = \frac{NIR - SWIR2}{NIR + SWIR2}$$

$$NDRE11 = \frac{NIR - Red\ Edge\ 1}{NIR + Red\ Edge\ 1}$$

$$NDRE2 = \frac{NIR - Red\ Edge\ 2}{NIR + Red\ Edge\ 2}$$

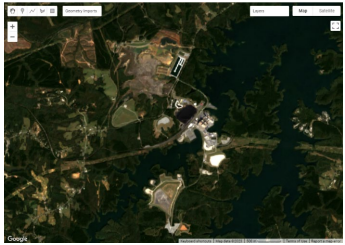
$$EVI = 2.5 * \frac{(NIR - Red)}{(NIR + 6 * Red - 7.5 * Blue + 1)}$$

$$SAVI = 1.5 * \frac{(NIR - Red)}{(NIR + Red + 0.5)}$$

$$REP = 705 + 35 * \frac{(\frac{Red + Red\ Edge\ 3}{2 - Red\ Edge\ 1})}{(Red\ Edge\ 2 - Red\ Edge\ 1)}$$

$$MTCI = \frac{Red\ Edge\ 2 - Red\ Edge\ 1}{Red\ Edge\ 1 - Red}$$

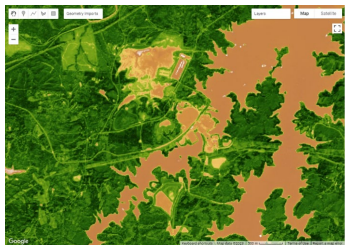
$$CI_{RE} = \frac{Red\ Edge\ 3}{Red\ Edge\ 1} - 1$$



TRUE COLOR



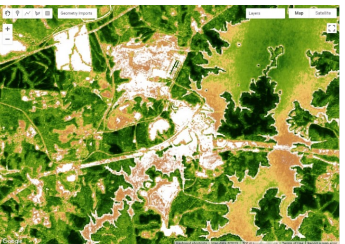
FALSE COLOR



EVI



NDVI



NDWI1



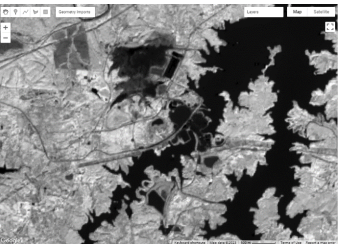
NDWI2



NDRE1



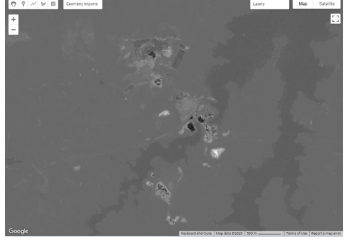
NDRE2



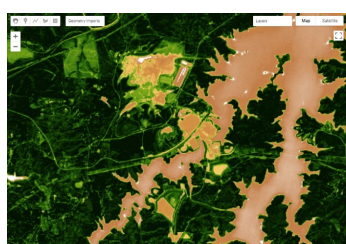
CERI



MTCI



REP



SAVI

Biophysical parameter index

- ❑ Quantitative measurements that relate to the physical properties of plants.
- ❑ Used in remote sensing to estimate parameters related to vegetation status.

$$LAI = (NIR - 1.35 * Red) * 11.728$$

Leaf Area Index

$$LCI = \frac{NIR - Red\ Edge\ 1}{NIR - Red}$$

Leaf Chlorophyll Index

$$LCC = \left(\frac{NIR - Red\ Edge\ 1}{NIR - Red} * 0.1134 \right) + 0.14$$

Leaf Chlorophyll Content

$$CCC = \left(-1.3231 + 0.8364 * \left(\frac{NIR}{Green} - 1 \right) \right)$$

Canopy Chlorophyll Content

$$CVI = \frac{\left(\frac{NIR}{Green} \right)}{Green/Red}$$

Chlorophyll Vegetation Index

Sensitivity analysis

- Which vegetation indices are suitable and sensitive to high stress level for our study area?
- R-squared between VI and BPI have been calculated.
- for each calculation, linear regression, exponential regression, logarithmic regression, and polynomial regression have been used to establish correlations between the spectral indices.

Coefficients of spatiotemporal variation (CSTV)

A standard deviation from mean calculated for the satellite images with space and time **factoring out all other influencers**

A higher CSTV value referred to abrupt stress, whereas a lower CSTV value indicated no stress or stable stress with a uniform changing trend during growth.

$$CSTV = \frac{|Q_i - \bar{Q}|}{\delta}$$

Q_i : the VI variations over time,

\bar{Q} : area-averaged mean value,

δ : standard deviation of Q.

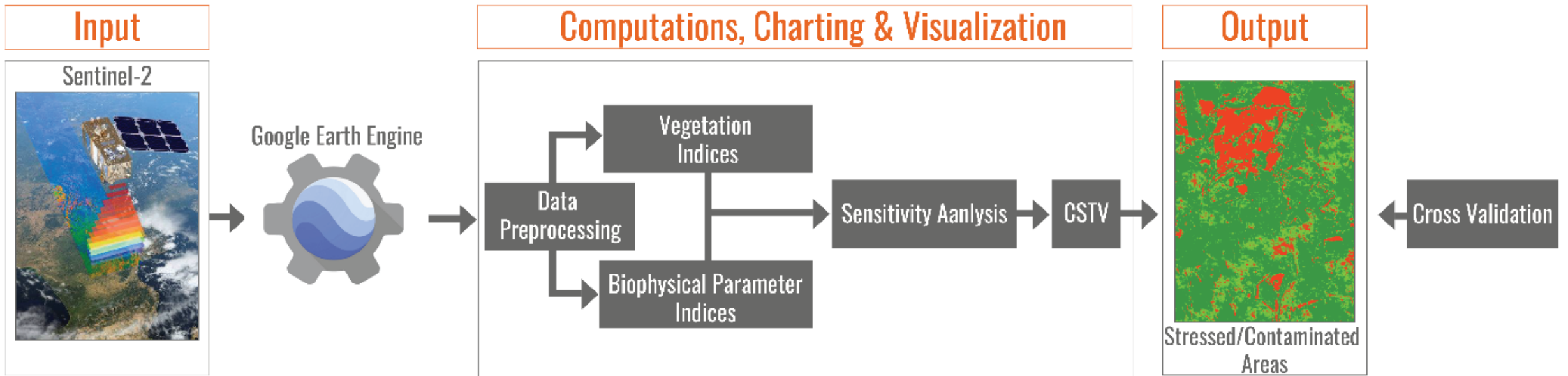
$$Q_i = VI_{m+1,i} - VI_{m,i}$$

$$\bar{Q} = \sum_i^j \frac{Q_i}{j}$$

$$\delta = \sqrt{\sum_i^j (Q_i - \bar{Q})^2 / j}$$

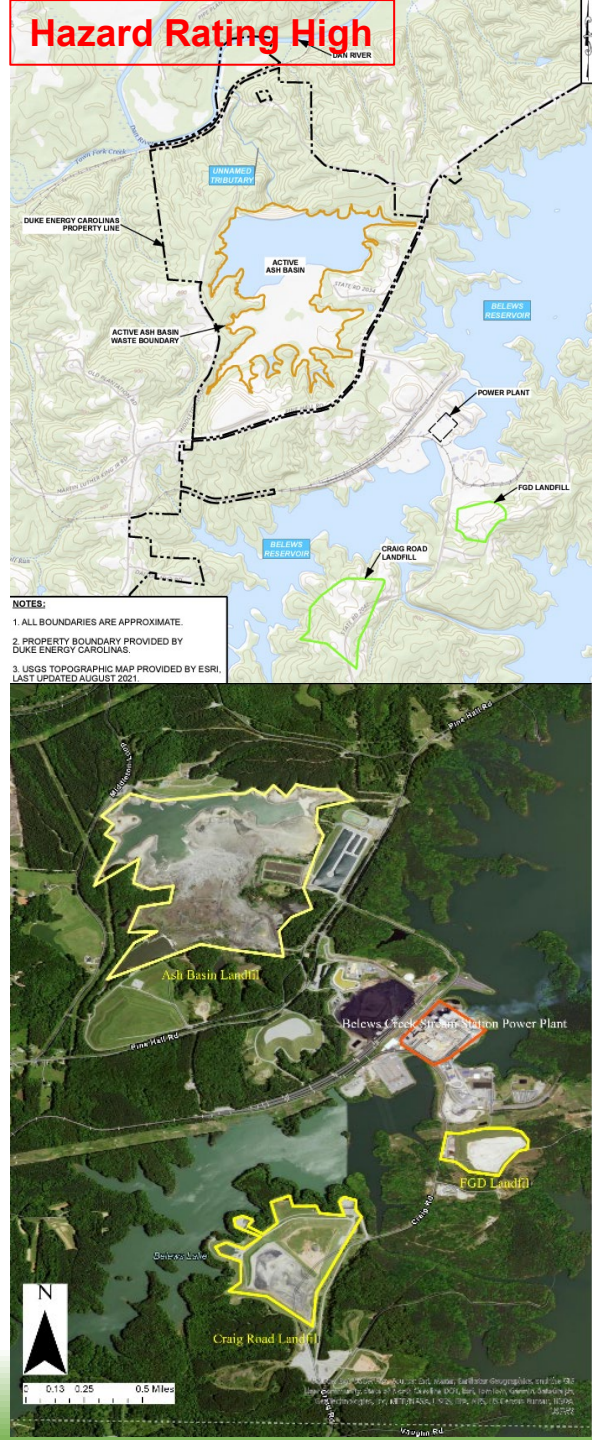
Vegetation stress detection workflow

- (i) Collecting multispectral satellite imagery data of sentinel-2 from Google cloud, preprocessing, and filtering the data with the required bands reducing cloud cover and atmospheric noise and other interference with area of interest for any specified time duration for further analysis,
- (ii) Calculating different vegetation indices (VI) and biophysical parameter indices (BPI),
- (iii) Performing sensitivity analysis to find sensitive VI and BPI to leached metal induced vegetation stress,
- (iv) Calculating coefficient of spatiotemporal variation (CSTV) for the better performing VI,
- (v) Visualizing the calculated CSTV on map to identify the potential contaminated area with high vegetation stress,
- (vi) Cross validation and comparison of the results with the ones from monitoring well data.

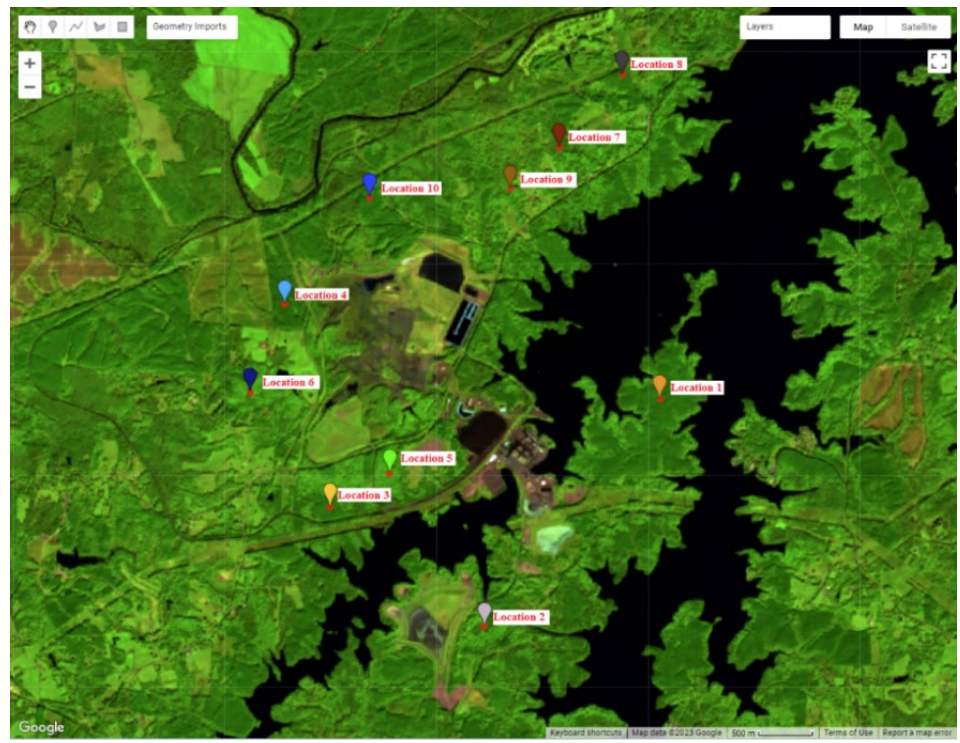
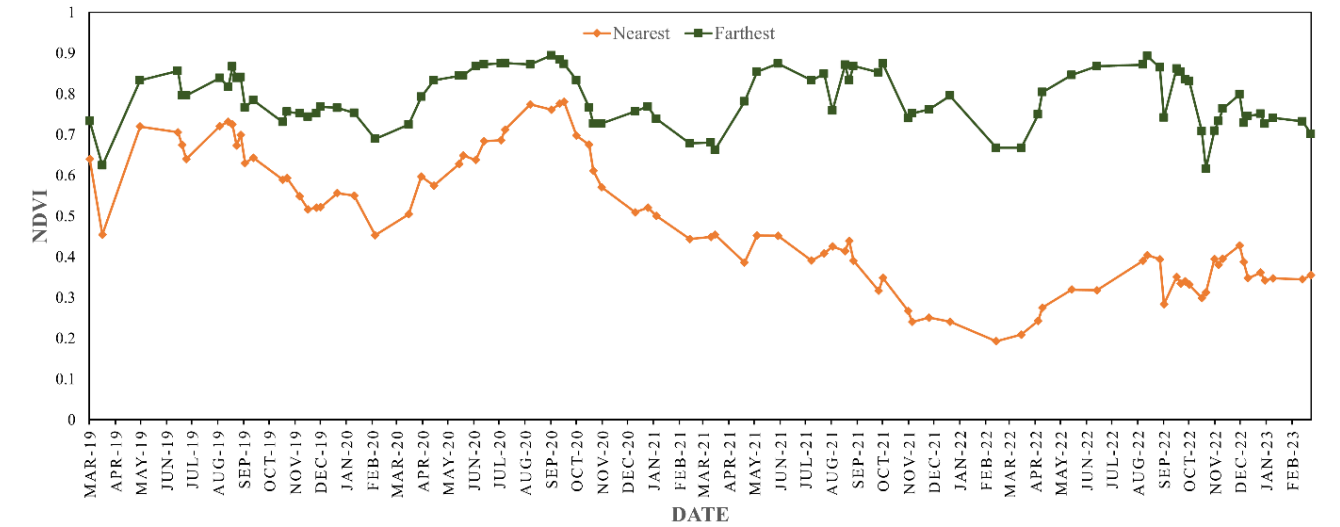
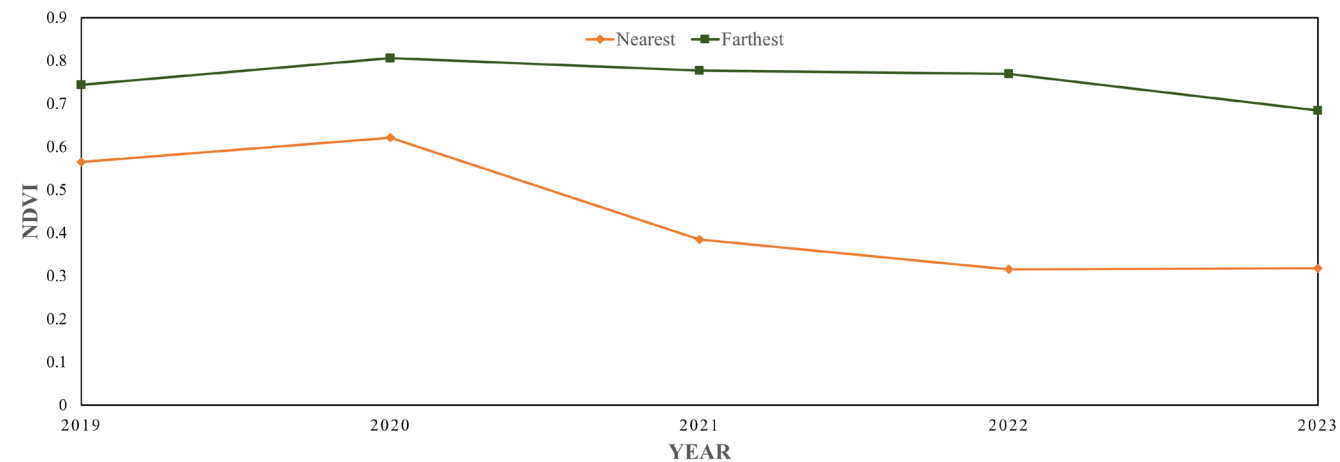


Case study 1: Belews Creek steam station

- Owned and operated by Duke Energy.
- Commercial operations began in 1974 with Unit 1 (1,120 MW), followed by Unit 2 (1,120 MW) in 1975.
- 3 impoundments: FGD Landfill, Craig Road Landfill, and Active Ash Basin.
- **104 groundwater monitoring wells**, 54 of which have been polluted above federal advisory levels based on samples collected between January 06, 2011, and April 15, 2019.
- **Groundwater at this site contains unsafe levels of arsenic, beryllium, boron, chromium, cobalt, lithium, manganese, mercury, molybdenum, radium, selenium, and sulfate.**
- Pollutants monitored and found below guidelines are antimony, barium, cadmium, copper, fluoride, lead, nickel, nitrate, and thallium.
- Duke Energy announced in 2022 that the facility would be shifted to natural gas by 2035.

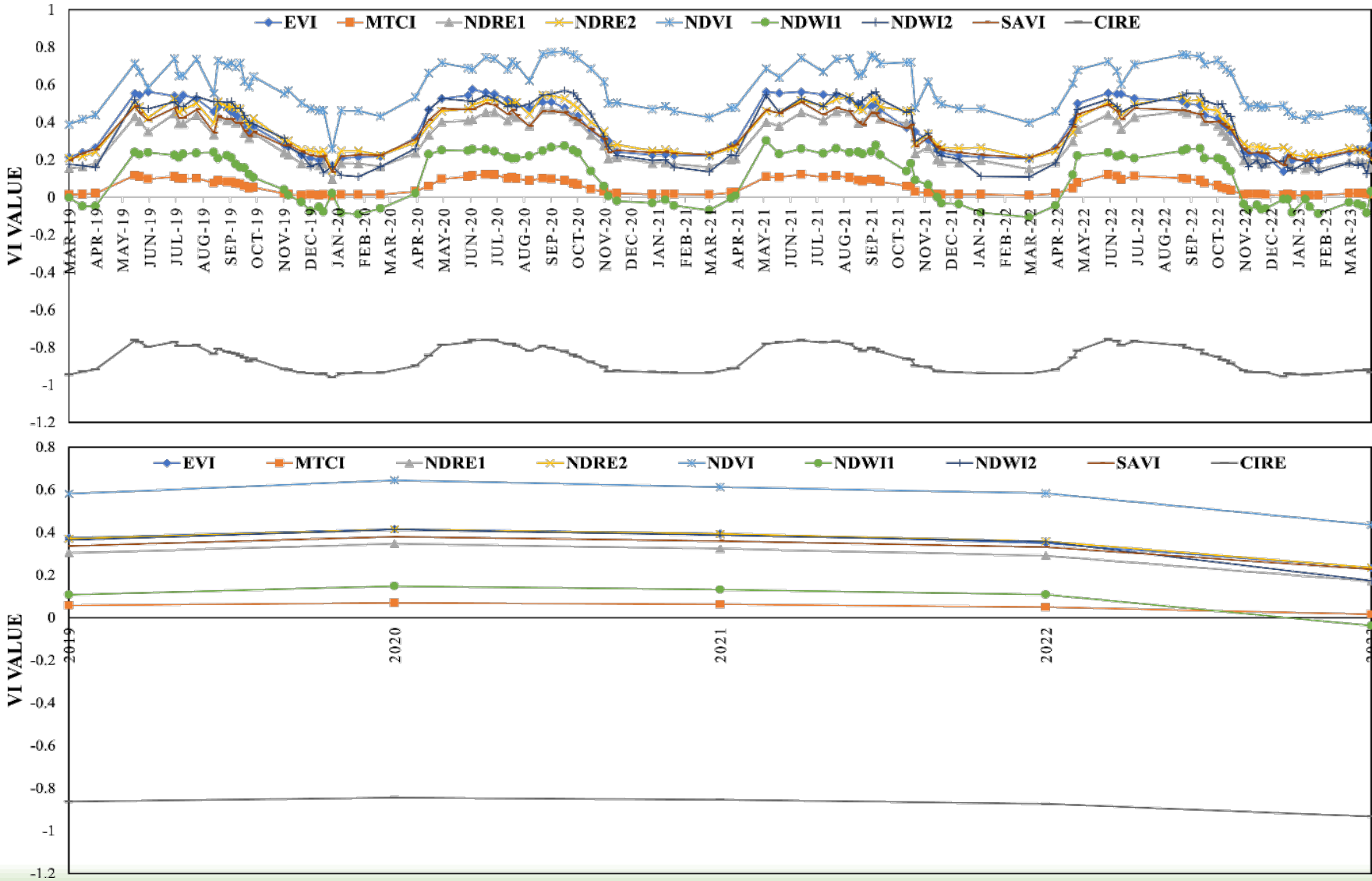


NDVI time series

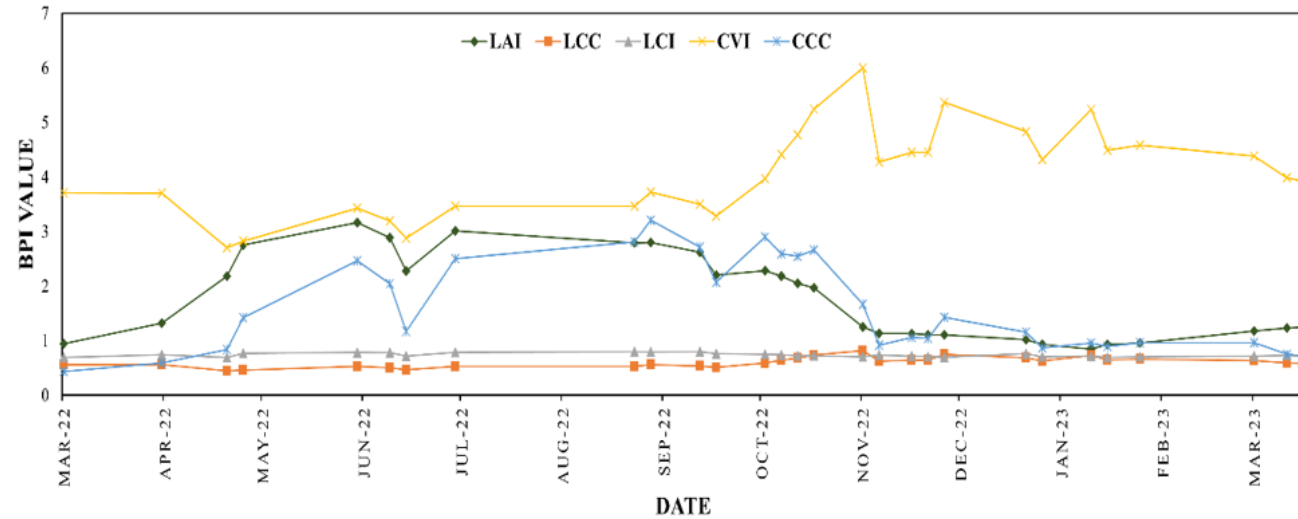
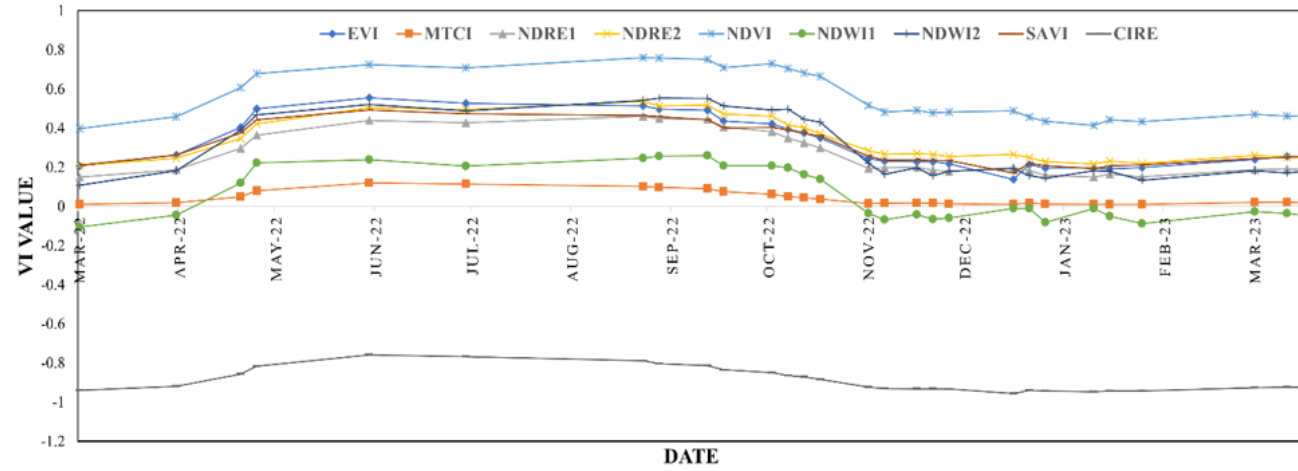


Vegetation indices time series

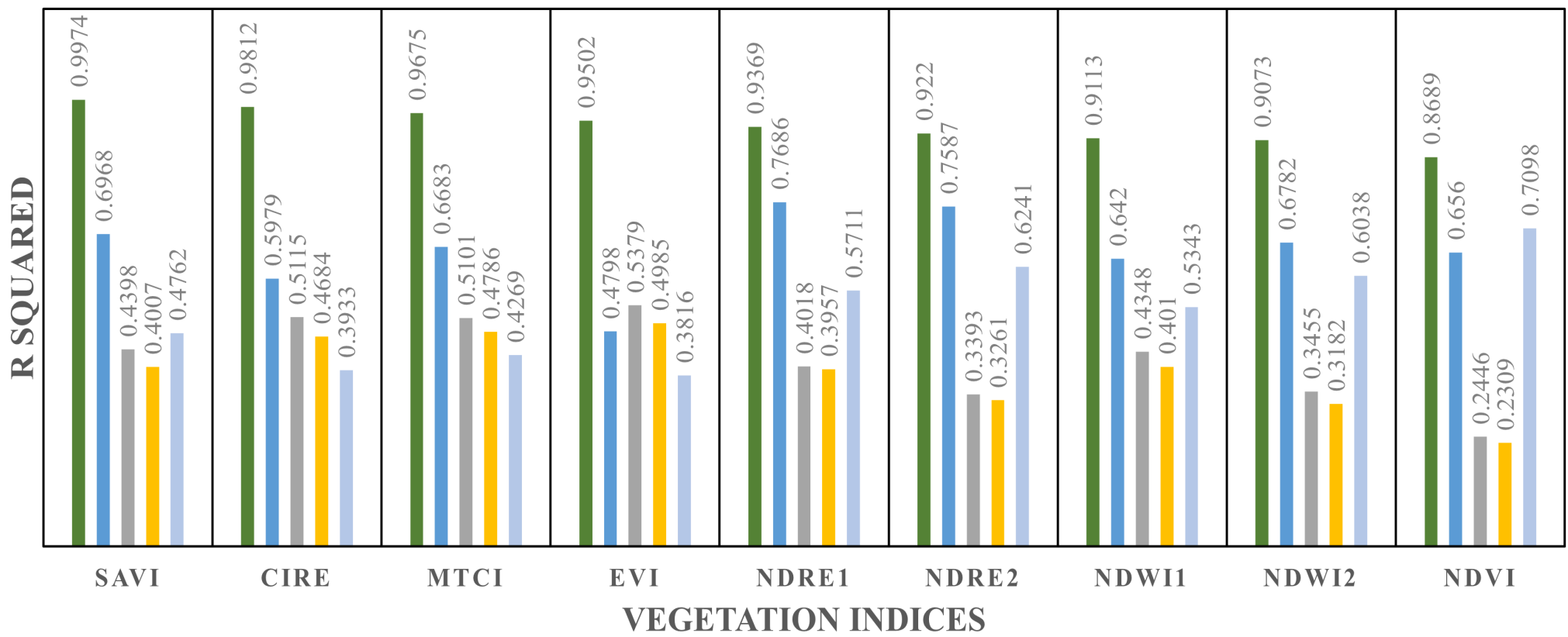
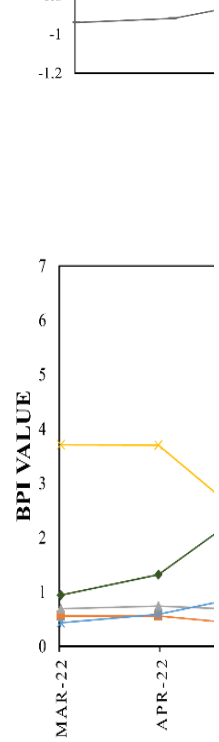
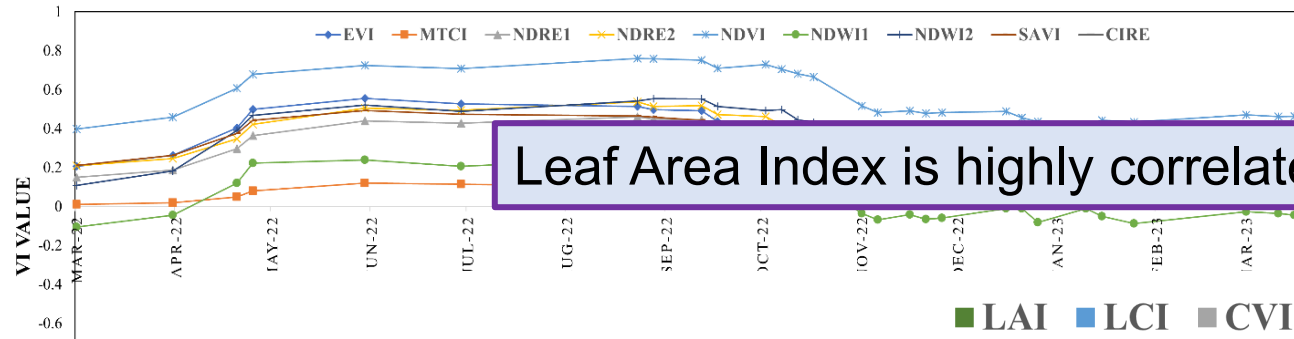
- Various vegetation indices has been calculated from Sentinel-2 satellite with less than 30% cloud cover.



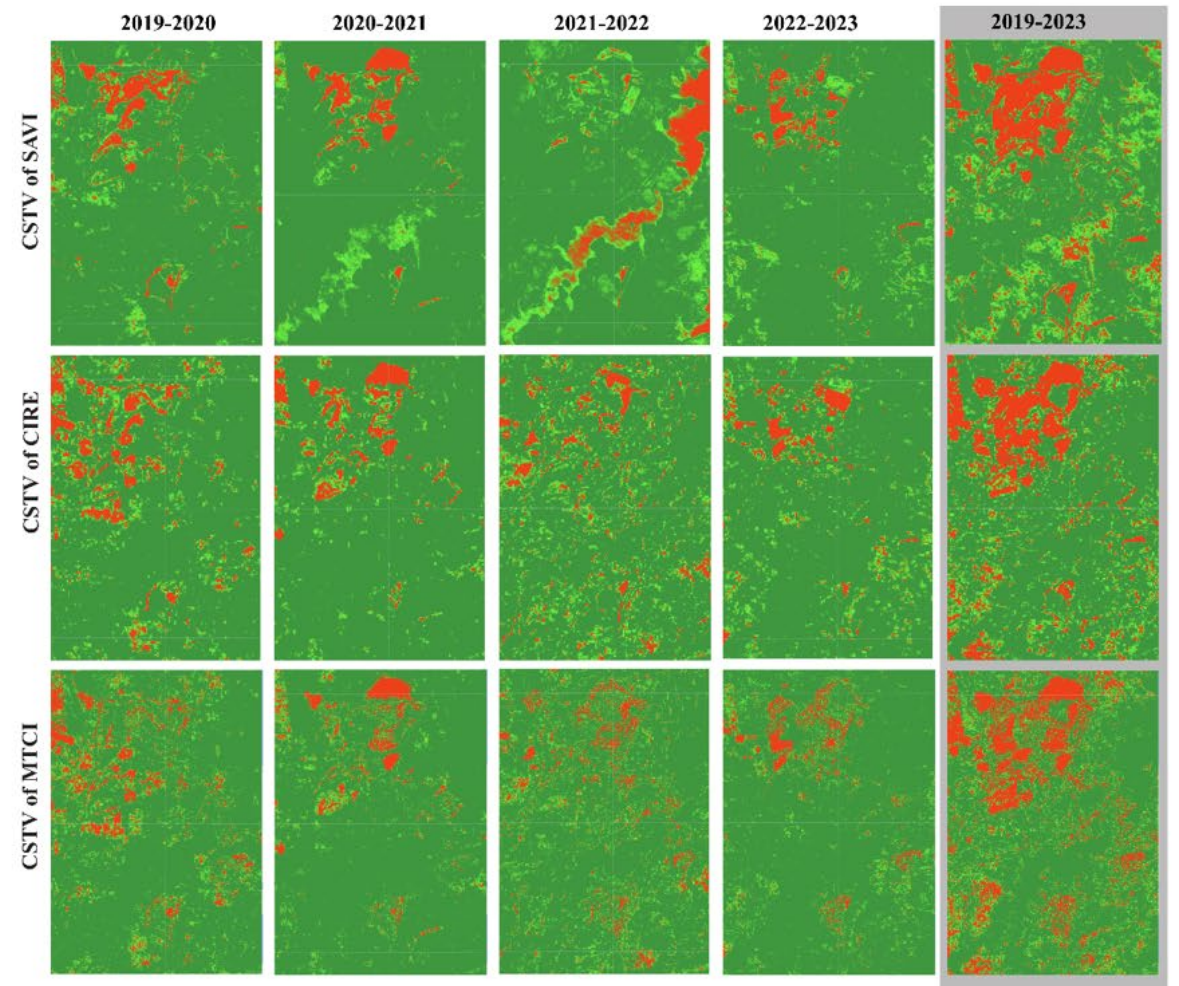
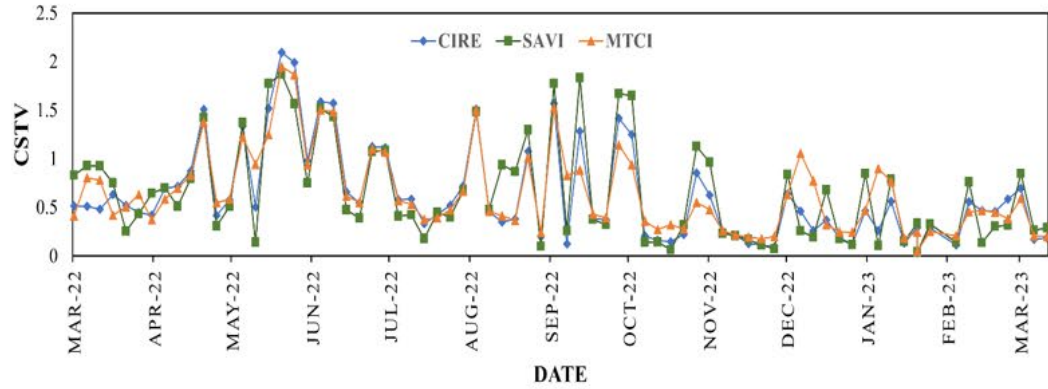
VI/BPI sensitivity analyses



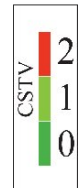
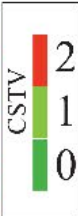
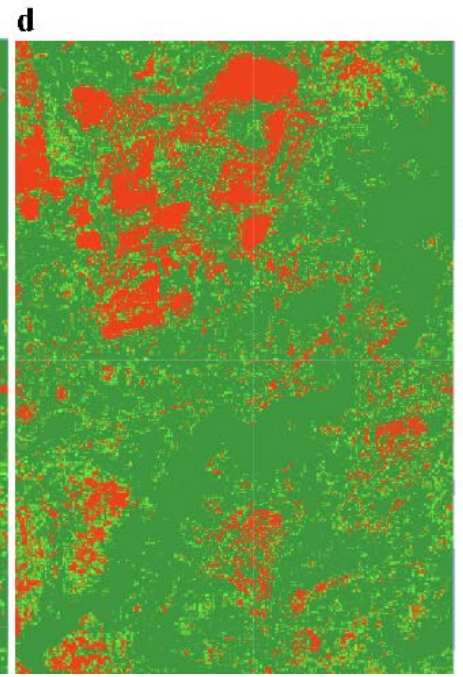
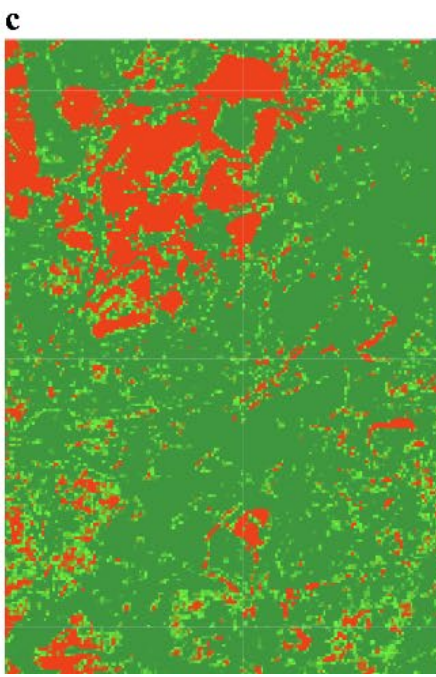
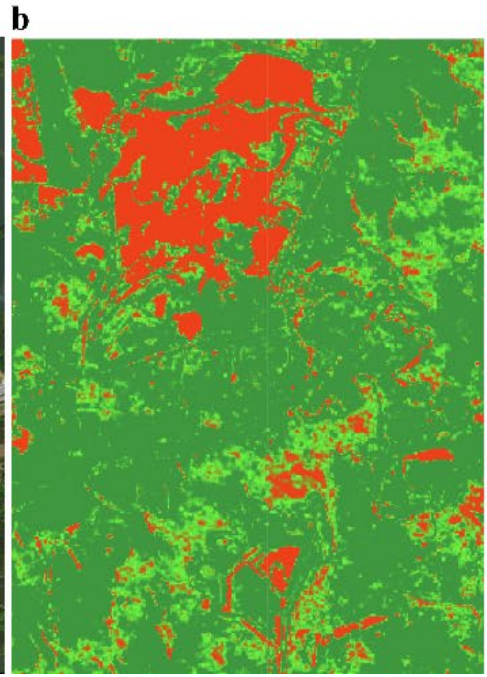
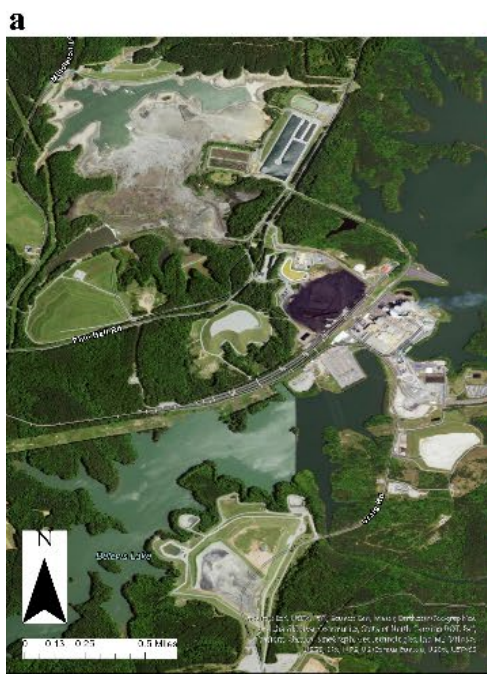
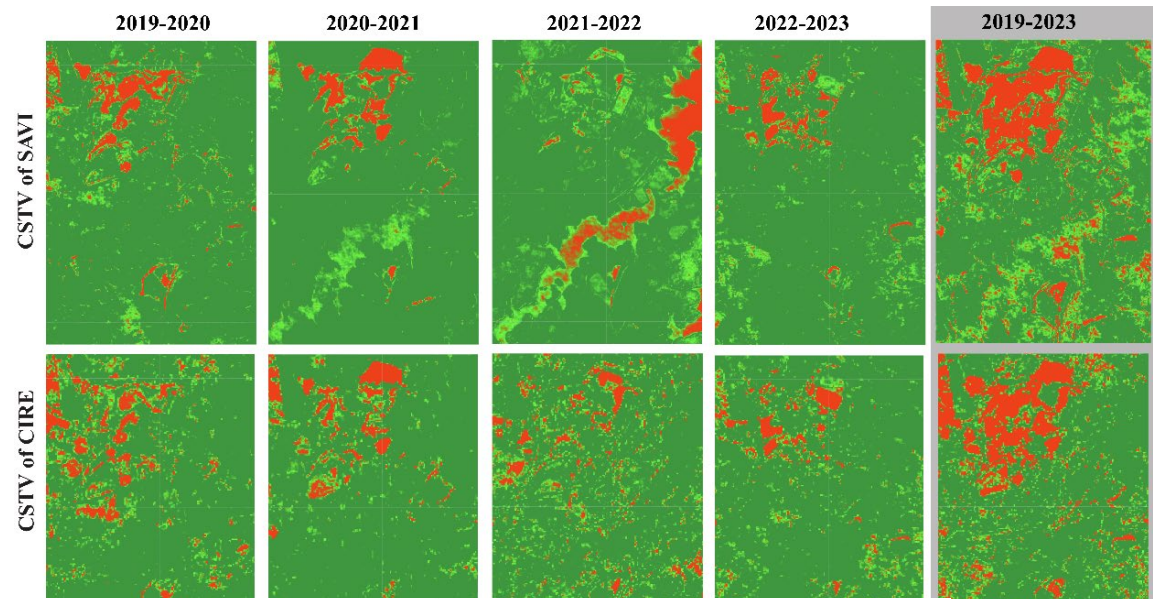
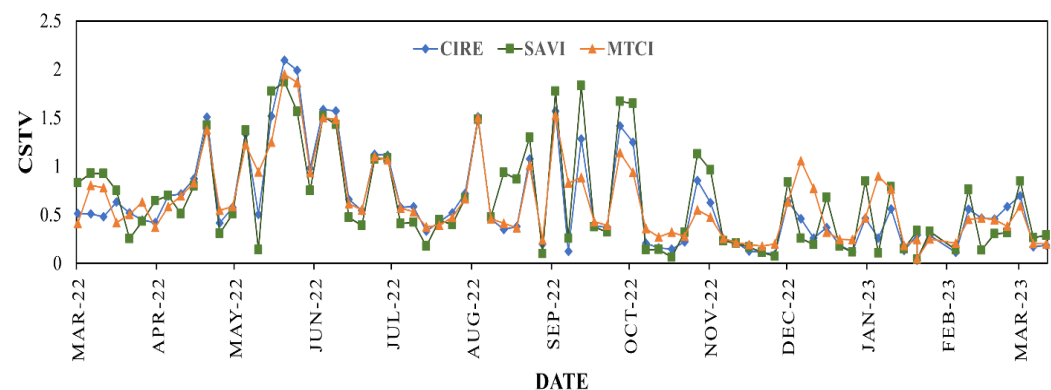
VI/BPI sensitivity analyses



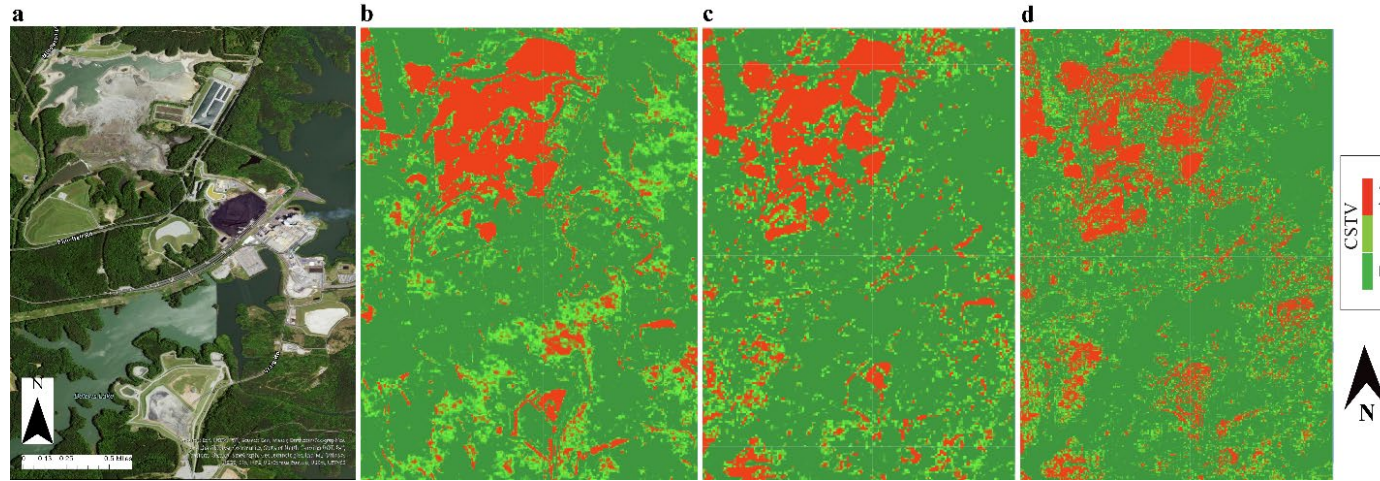
CSTV calculation



CSTV calculation



Cross validation



Ongoing work:

- searching for potential correlations between vegetation stress and type of leached metal

Case study 2: Four Virginian power stations

Chesterfield

- Active plant with significant hazard owned by Dominion Virginia Power in Chester, VA
- Capacity of 1353 MW built in 1952
- 4 impoundments, and 61 monitoring wells, 46 of which are polluted
- 12 contaminating elements (ammonia, arsenic, beryllium, boron, chromium, cobalt, lead, lithium, manganese, molybdenum, radium, sulfate).

Bremo

- Active plant with significant hazard by Dominion Virginia Power in Bremo Bluff, VA
- Capacity of 254 MW, built in 1951,
- 3 impoundments, and 32 monitoring wells, 28 polluted
- 9 contaminants (arsenic, boron, cobalt, lead, lithium, manganese, molybdenum, radium, sulfate).

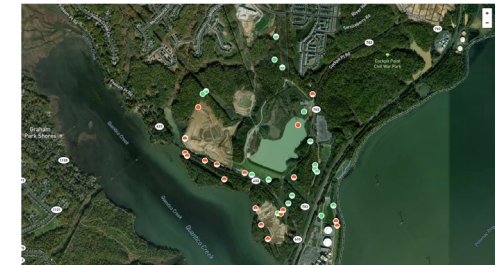
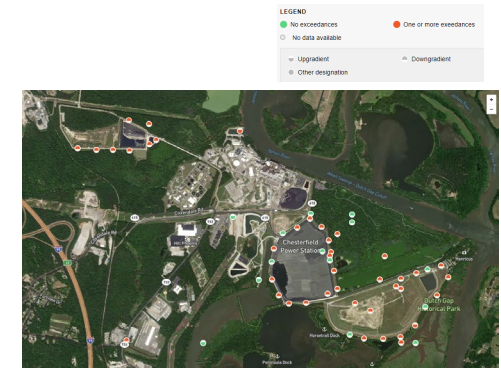
Possum Point

- Active with significant hazard plant in Dumfries, VA, also owned by Dominion Virginia.
- Capacity of 1235 MW, built in 1956
- 5 impoundments, and 36 monitoring wells, 18 polluted
- 7 known contaminants (arsenic, boron, cadmium, cobalt, manganese, nickel, radium).

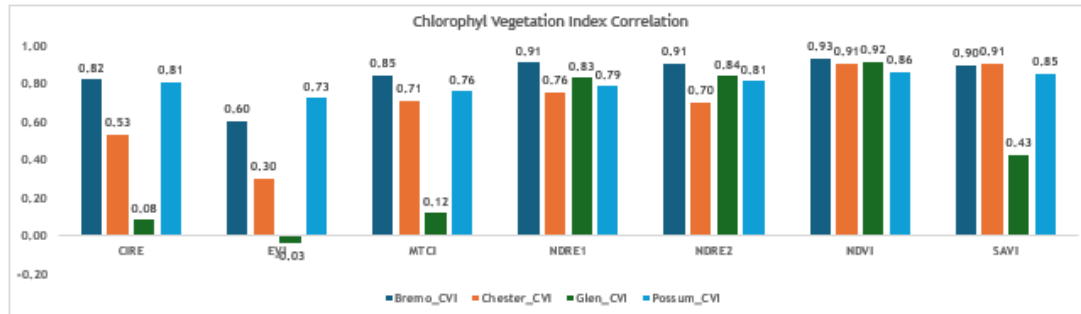
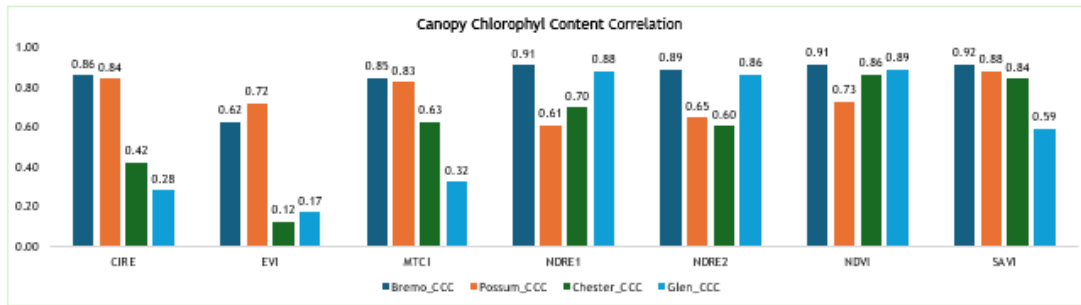
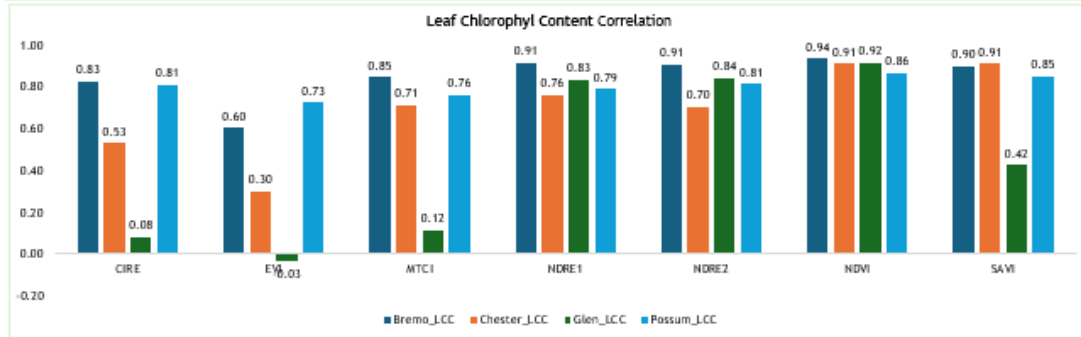
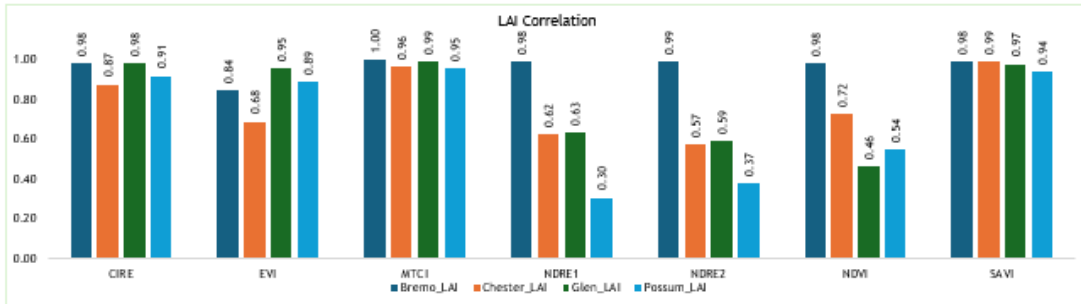
Glen Lyn

- Retired with low hazard and operated by Appalachian Power Company in Glen Lyn, VA
- Capacity of 338 MW, built in 1945
- 6 impoundments, and 19 monitoring wells, 17 polluted
- 8 contaminants (arsenic, boron, cobalt, lead, manganese, molybdenum, strontium, sulfate).

Arsenic, boron, cobalt and manganese are detected for all 4.

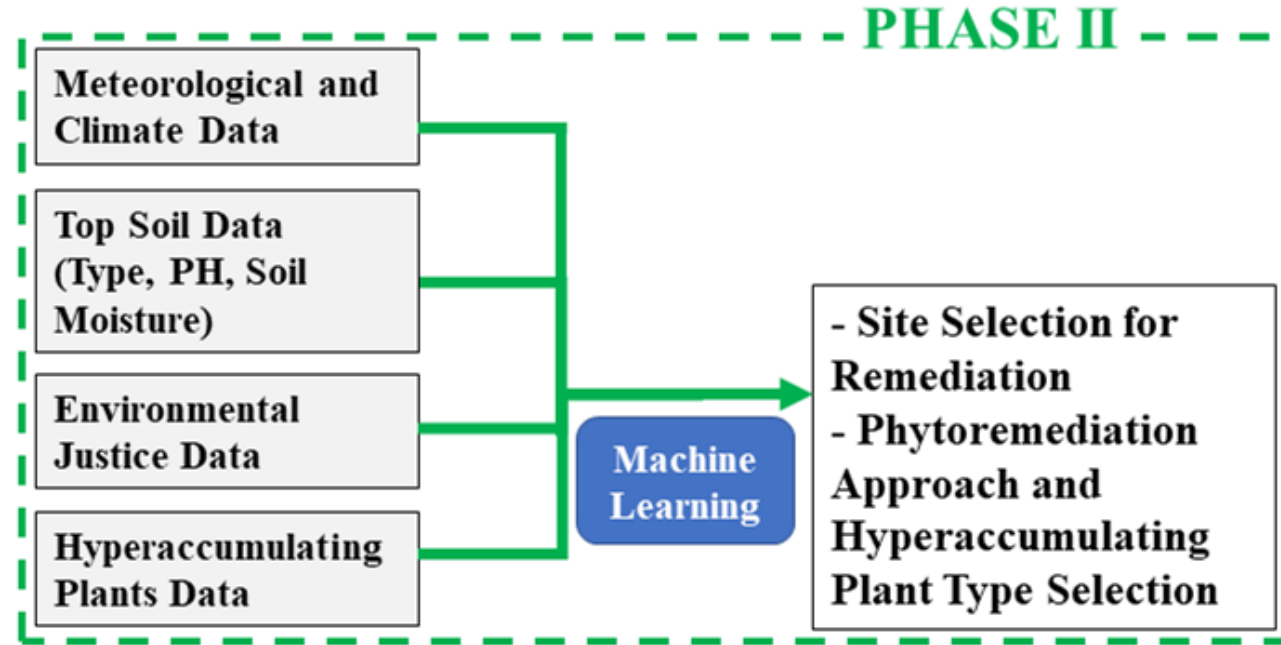


Correlations between seven VI and four BPI of 4 sites



- Sentinel-2 image data from 2018 to 2023 were analyzed.
 - Sensitivity analysis and correlation studies showed that NDVI performed the best across three of the four sites in measuring vegetation health/stress and strongly correlated with BPI indices: LCC, CCC, and CVI, while MTC1 & SAVI were more correlated with LAI.
 - NDVI, MTC1 and SAVI were utilized to calculate CSTV and visualize the stressed areas.
 - GEE based applications are developed for calculating and visualizing Vegetation Indices, Biophysical Parameter Indices and Co-efficient of Spatiotemporal Variations (CSTV) for further analysis of vegetation stress and contamination surrounding the power stations.
- Bremono: [VI](#), [BPI](#), [CSTV](#)
 - Chesterfield: [VI](#), [BPI](#), [CSTV](#)
 - Possum Point: [VI](#), [BPI](#), [CSTV](#)
 - Glen lyn: [VI](#), [BPI](#), [CSTV](#)

Phase II: propose locally adaptable phytoextraction approaches to remediate contaminated regions



- ❑ A content-based filtering ranking algorithm will be developed for effective and environmentally justified case-specific phytoremediation

Output:

- Ranking phytoremediation candidates
- Suggested plant species to be used for the phytoremediation of candidate sites

Data Acquisition and Classification based on soil type and properties, hyperaccumulating plant types, and metal contamination types

Which plants are a good fit for different sites?

- Plant type
- Soil type
- Climate
- Native or non-native



Letters

A global database for plants that hyperaccumulate metal and metalloid trace elements

Hyperaccumulator plants and the need for a database

Table 1 Hyperaccumulator species in the Global Database (as of September 2017) with the global records that are the highest concentrations reported to date

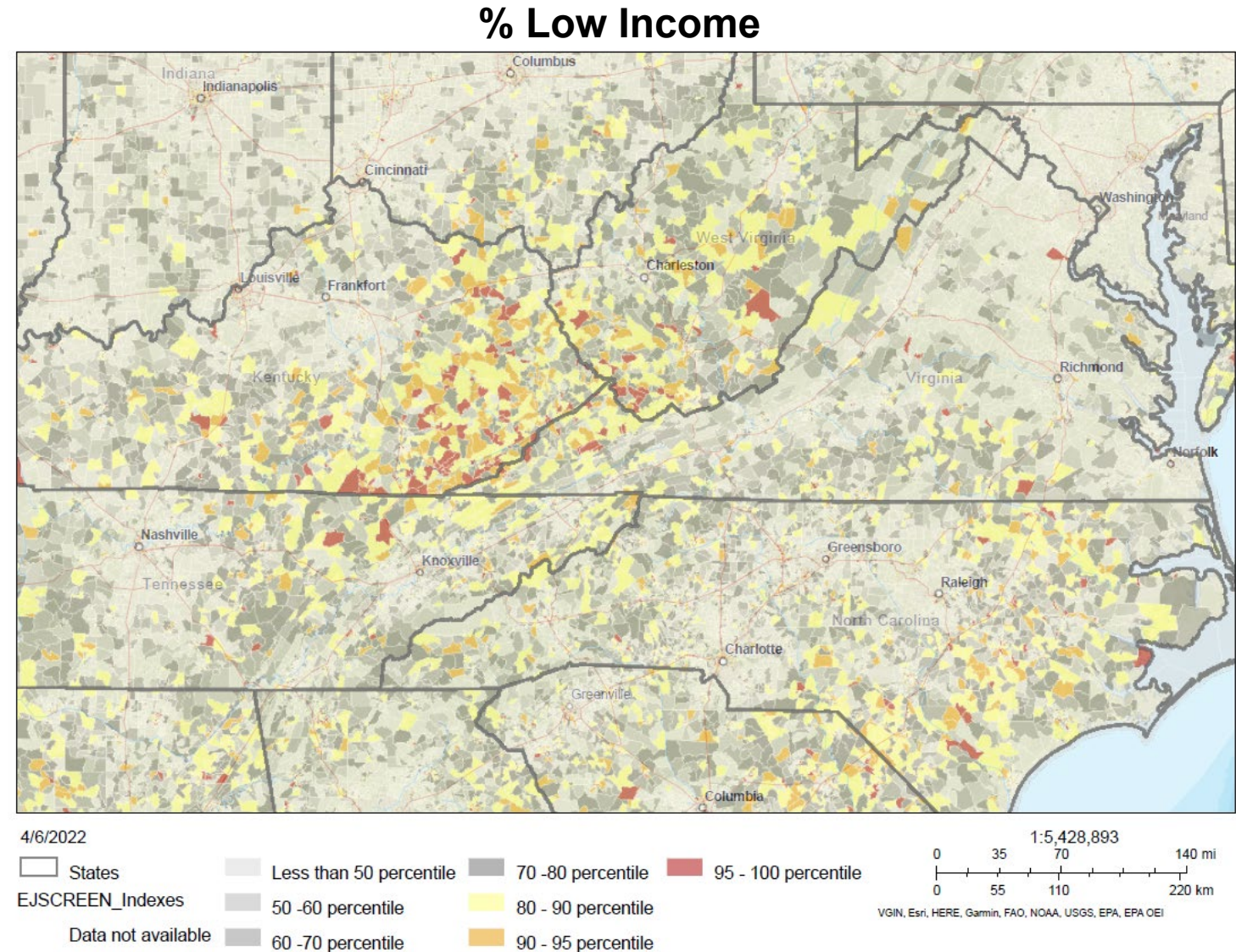
Element	Threshold ($\mu\text{g g}^{-1}$)	Families	Genera	Species	Global records
Arsenic (As)	> 1000	1	2	5	<i>Pteris vittata</i> ¹ (2.3%)
Cadmium (Cd)	> 100	6	7	7	<i>Arabidopsis halleri</i> ² (0.36%)
Copper (Cu)	> 300	20	43	53	<i>Aeolanthus biformifolius</i> ³ (1.4%)
Cobalt (Co)	> 300	18	34	42	<i>Haumaniastrum robertii</i> ⁴ (1%)
Manganese (Mn)	> 10 000	16	24	42	<i>Virotia neurophylla</i> ⁵ (5.5%)
Nickel (Ni)	> 1000	52	130	532	<i>Berkheya coddii</i> ⁶ (7.6%)
Lead (Pb)	> 1000	6	8	8	<i>Noccaea rotundifolia</i> subsp. <i>cepaefolia</i> ⁷ (0.8%)
Rare earth elements (lanthanum, La; cerium, Ce)	> 1000	2	2	2	<i>Dicranopteris linearis</i> ⁸ (0.7%)
Selenium (Se)	> 100	7	15	41	<i>Astragalus bisulcatus</i> ⁹ (1.5%)
Thallium (Tl)	> 100	1	2	2	<i>Biscutella laevigata</i> ¹⁰ (1.9%)
Zinc (Zn)	> 3000	9	12	20	<i>Noccaea caerulescens</i> ¹¹ (5.4%)

¹Ma *et al.* (2001); ²Stein *et al.* (2017); ³Malaisse *et al.* (1978); ⁴Brooks (1977); ⁵Jaffré (1979); ⁶Mesjasz-Przybyłowicz *et al.* (2004); ⁷Reeves & Brooks (1983); ⁸Shan *et al.* (2003); ⁹Galeas *et al.* (2006); ¹⁰LaCoste *et al.* (1999); ¹¹Reeves *et al.* (2001).

Data acquisition and classification based on environmental justice

Identify priority regions based on the data from **US EPA's Environmental Justice Screening and Mapping Tool (EJSCREEN)**

- % Low Income
- % Unemployed
- % Limited English Speaking
- % Less than High School Education
- Low Life Expectancy



Project benefits

- A scalable, low-cost technology to screen a wide range of locations and to identify coal ash-sourced or mine tailings- sourced leached metal-contaminated areas.
- Train and educate students in a multi-disciplinary setting
- Potential application of phytotechnology in other fields (e.g. phyto-mining)
- Phytoremediation pilot design and implementation
- Post phytoremediation monitoring

Bahareh Nojabaei
baharehn@vt.edu

Virginia Tech
Mining and Minerals Engineering

