Mapping Breeding Sites:

Leveraging GIS to Reduce Vector-Borne Disease

MAMCA 2020 Annual Conference

Melissa Nolan, PhD, MPH
Assistant Professor
University of South Carolina

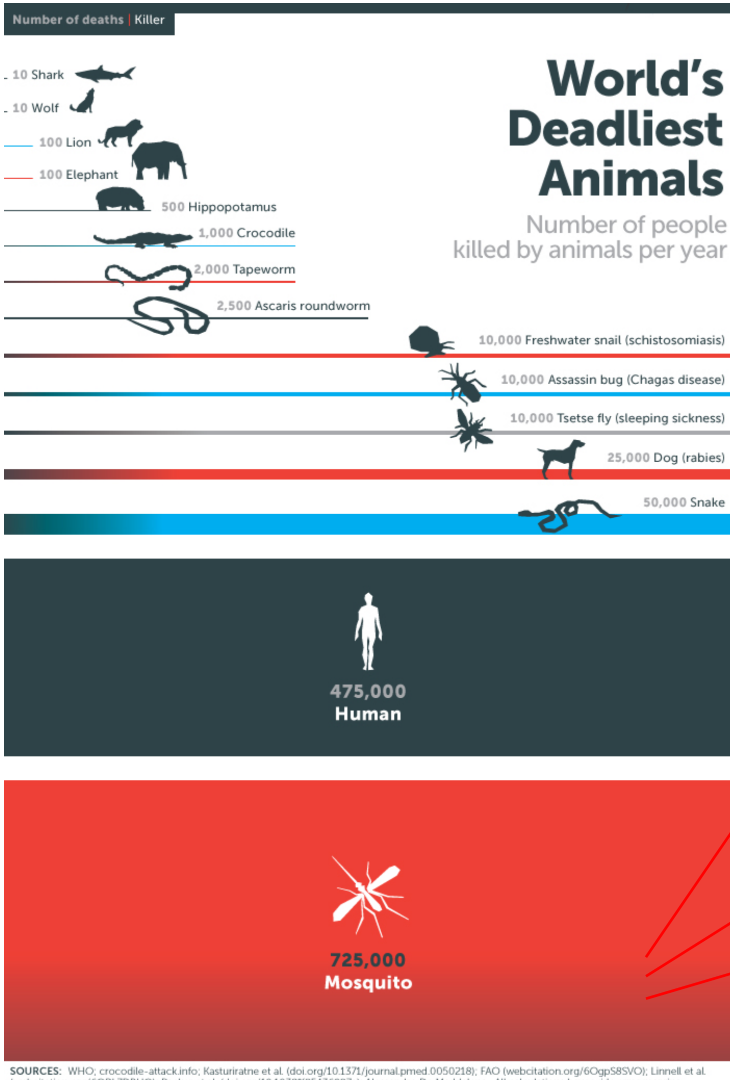
UNIVERSITY OF SOUTH CAROLINA

Baylor College of Medicine

ExxonMobil Energy lives here™

Harris County Public Health Building a Healthy Community

Public Health Impact of Mosquito-Borne Diseases



- ***Aedes spp.***
  - Chikungunya
  - Dengue fever
  - Lymphatic filariasis
  - Rift Valley fever
  - Yellow fever
  - Zika
- ***Anopheles***
  - Malaria
  - Lymphatic filariasis
- ***Culex***
  - Japanese encephalitis
  - Lymphatic filariasis
  - West Nile fever

UNIVERSITY OF SOUTH CAROLINA

https://www.gatesnotes.com/Health/Most-Lethal-Animal-Mosquito-Week

Challenges of Mosquito-Borne Disease Prevention

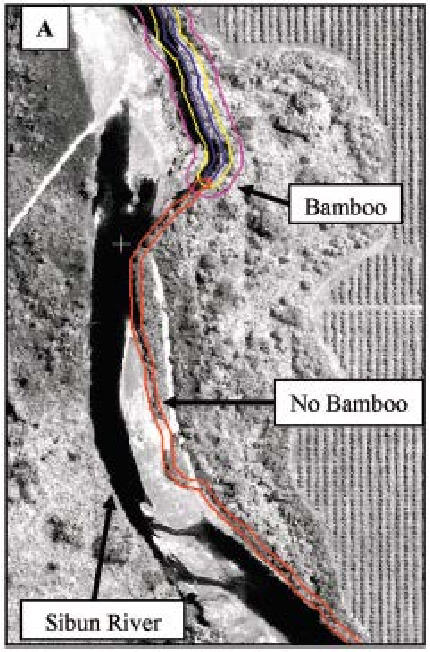
1. Majority of these diseases originate in infrastructure-poor, resource-limited countries
   I. Hard to predict spread of new Mosquito-Borne Diseases
      a. Arboviral mutations
         i. Unpredictable jump to new mosquito species-animal hosts
      b. Lack of surveillance
         i. Can't identify new epidemics
         ii. Can't track spread
         iii. Unaware of highest-risk populations

2. Limited funding for local mosquito control organizations

3. Paucity of available diagnostics, vaccines, and therapeutics

Development of BCM-ExxonMobil Collaboration:

Applying Remote Sensing Technologies to Enhance Mosquito Abatement



MODELING/GIS, RISK ASSESSMENT, ECONOMIC IMPACT

Use of Remote Sensing and Geographic Information Systems to Predict Locations of *Anopheles darlingi*-Positive Breeding Sites Within the Sibun River in Belize, Central America

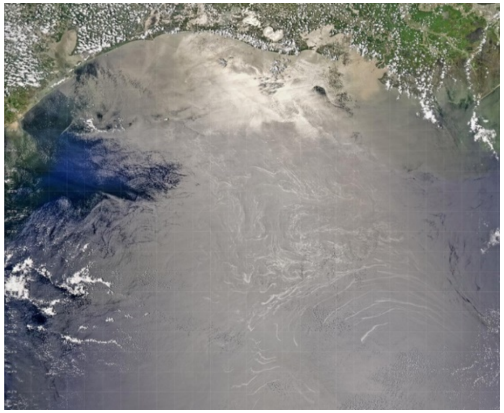
NICOLE L. ACHEE,[1] JOHN P. GRIECO,[1] PENNY MASUOKA,[1] RICHARD G. ANDRE,[1] DONALD R. ROBERTS,[1] JAMES THOMAS,[1] IRENEO BRICENO,[2] RUSSELL KING,[2] AND ELISKA REJMANKOVA[3]

J. Med. Entomol. 43(2): 382–392 (2006)

to be made up of forest land cover (Table 5). Evaluation of the confusion matrix indicated a 75.9% accuracy rate by which all land cover categories were classified. Bare ground, forest, and pasture/low grass land cover categories had the highest accuracy rates with 98.8, 97.0, and 94.9% of the pixels being correctly classified, respectively. The orchard and sandbar land cover classes suffered from the worse classification confusion, with 58.8 and 58.6% of the pixels, respec-

ExxonMobil: Ideal Remote Sensing Partner

- Environmental applications of remote sensing technologies
  - Assess environmental impact
    - Baseline survey of vegetation cover & health (*chlorophyll count*)
    - Post-Oil exploration and drilling survey of vegetation
  - Assess environmental recovery post-spill clean-up
  - Search for geographic features that indicate oil reserves
    - Surface oil slicks, phytoplankton



NASA's MODIS Aqua sensor; https://www.boem.gov/BOEM-2016-082/; Ian McDonald

Why Focus Project in Harris County, TX?

- Everything is bigger in Texas
  - Most populous county in Texas
  - 3rd most populous in the US
  - Home to 4.6 million people

- Home to vulnerable populations
  - 15.9% living in poverty
  - ~4,000 homeless residence
    - 1,515 living unsheltered

- One of the oldest vector control organizations in the country

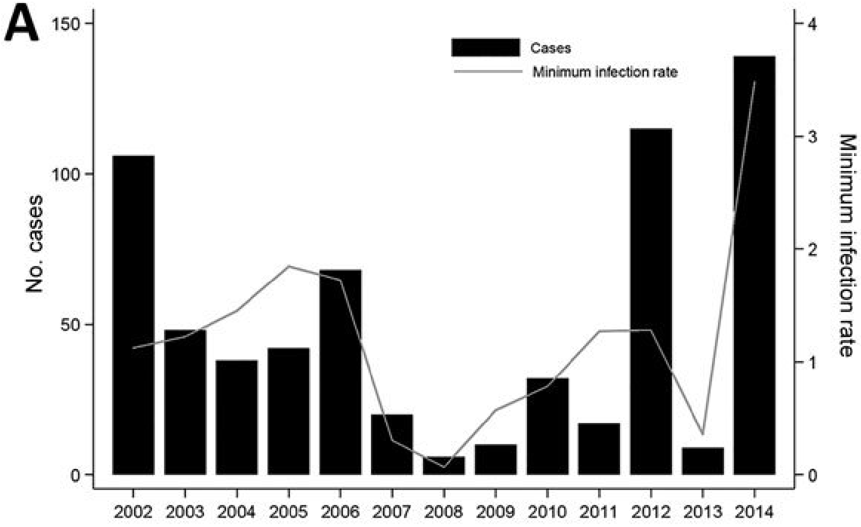
- 100+ resident mosquito species
  - *Aedes aegypti*, *Aedes albopictus* & *Culex quinquefasciatus*

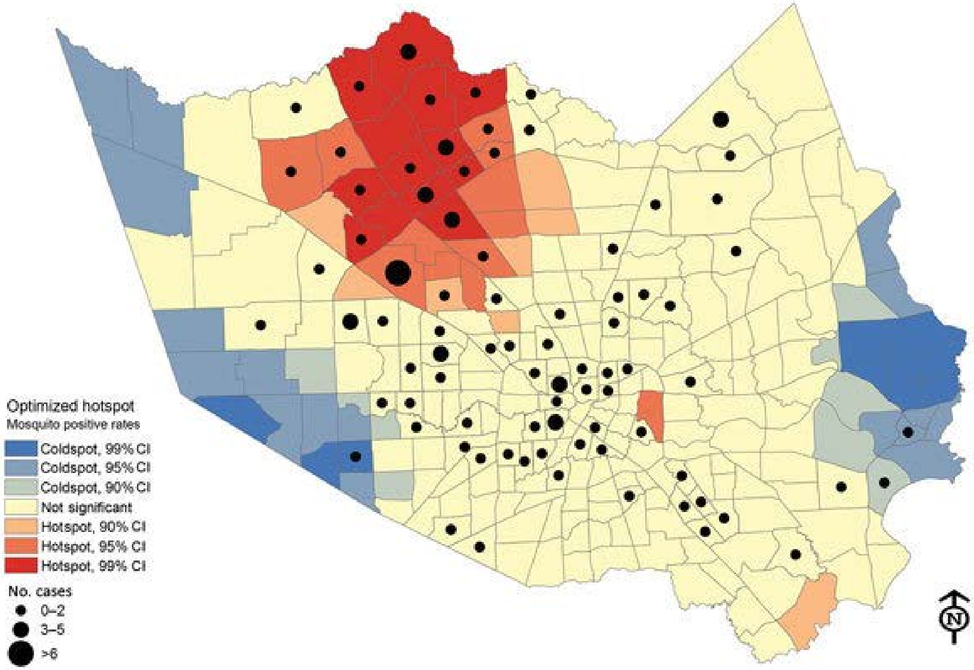
Harris County
Public Health
Building a Healthy Community

UNIVERSITY OF SOUTH CAROLINA

https://www.census.gov/quickfacts/harriscountytexas
http://www.homelesshouston.org/wp-content/uploads/2018/05/Final_2018_PIT_FactSheet_Digital_3.pdf

Why *Culex quinquefasciatus* in Harris County?





- High disease burden
- Sustained transmission
- Geographic dispersion of infected mosquitos & human cases

Emerg Infect Dis. 2013 Jan;19(1):137-9.
Emerg Infect Dis. 2013 Nov; 19(11): 1836–1838.
Emerg Infect Dis. 2017 Aug;23(8):1372-1376

Collaborative Project Goals

Overall goal: Create a reproducible workflow to identify priority areas for mosquito abatement

1) Identify relevant land data from imagery collected over the course of our study period

2) Build a predictive model using statistical land data correlations with mosquito trapping data from across Harris county




Workflow: Mosquito Data Collection

- *Culex quinquefasciatus* mosquito collections were conducted from October 2017 to September 2018
  - Comprised of 10,767 trap nights
  - 934 unique trapping locations

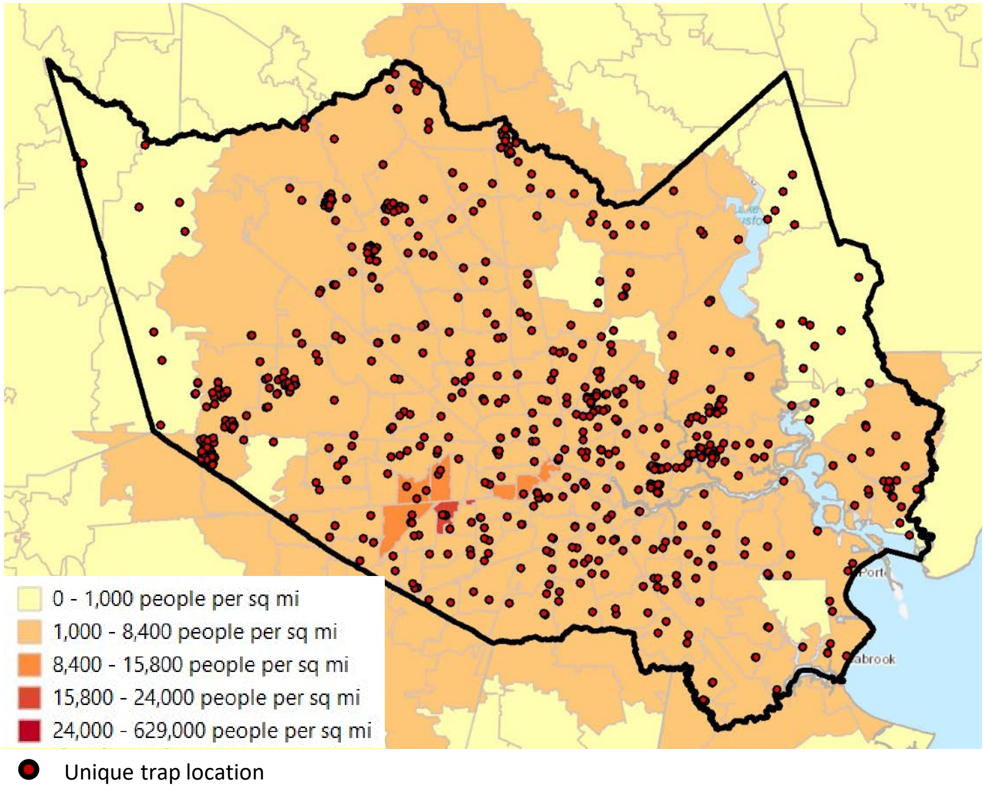
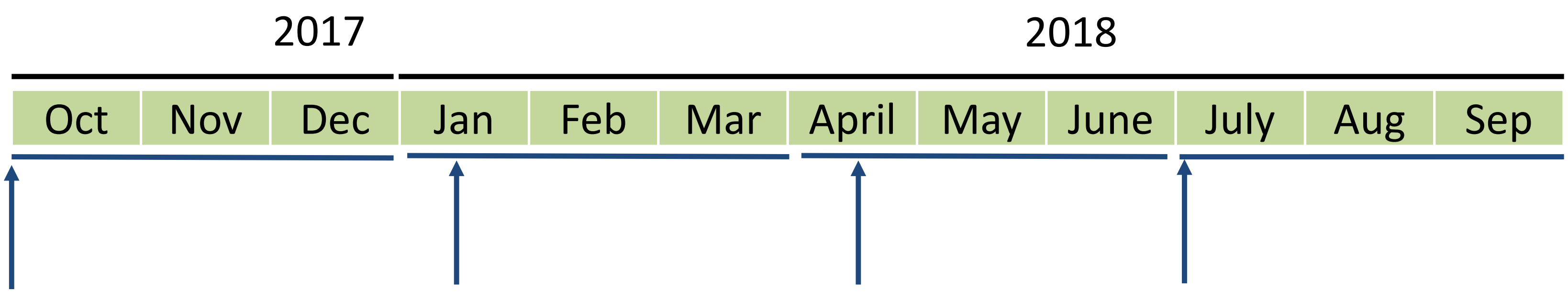


Image Source: Landsat - 8

- Reasons for selection:
  1. Ease of access
  2. Reasonable resolution (30m) for analysis
  3. Advance products created for and made available to non-experts including:
     - Surface Reflectance
     - Normalized Difference Vegetation Index (**NDVI**)
     - Normalized Difference Moisture Index (**NDMI**)
     - Enhanced Vegetation Index (EVI)
     - Soil Adjusted Vegetation Index (SAVI)
     - Modified Soil Adjusted Vegetation Index (MSAVI)
     - Normalized Burn Ratio (NBR)
     - Normalized Burn Ratio 2 (NBR2)

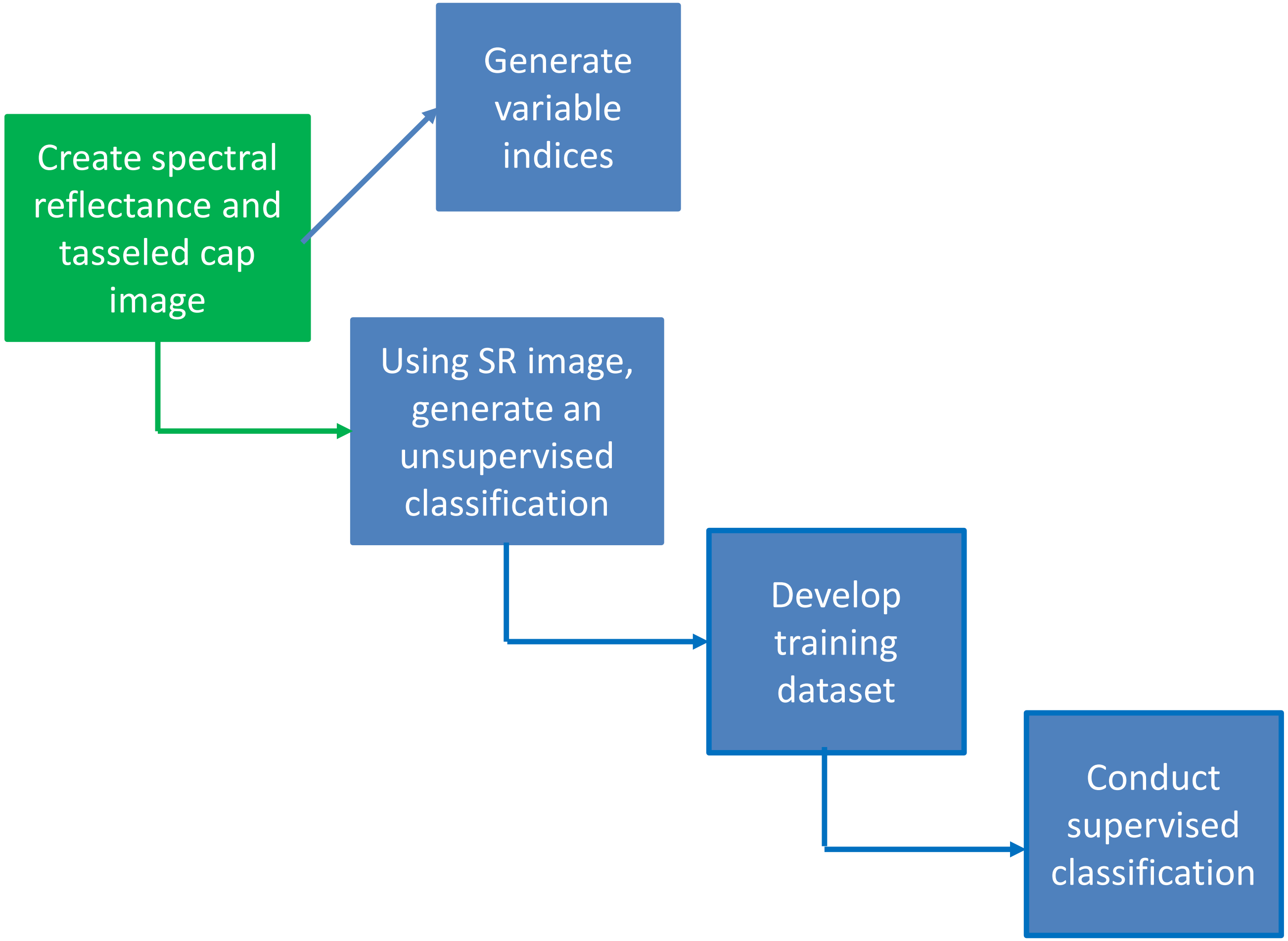
Used for this project

Landsat-8 Workflow: Image Selection

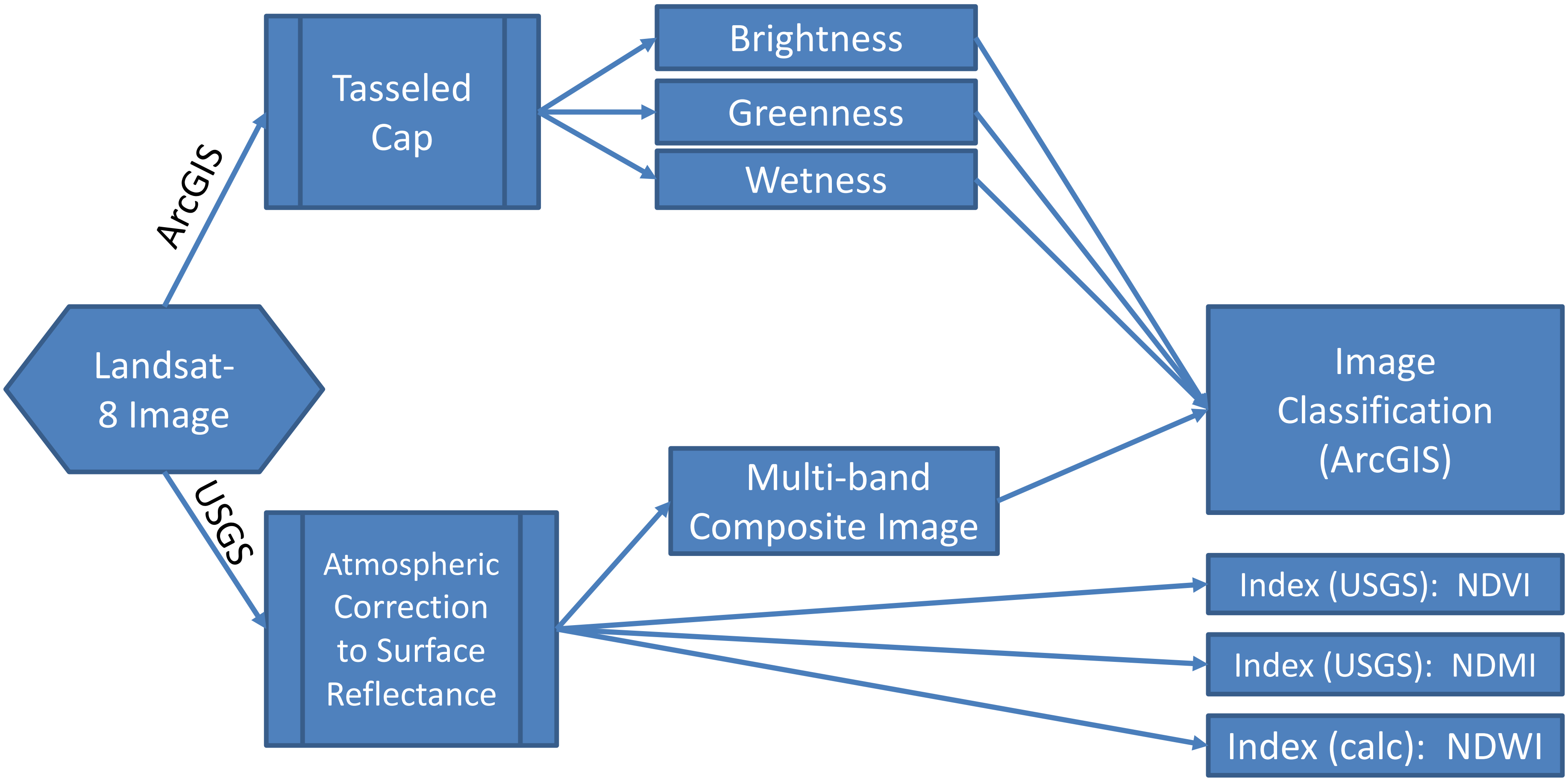


- 4 images chosen to be representative of the study period
- Images were chosen because of their even spacing through the study period, low cloud cover (0.27-5.37%)
- WRS path 025 WRS Row 039, was used to cover our catchment area in Harris County, Texas

Image Analysis Workflow

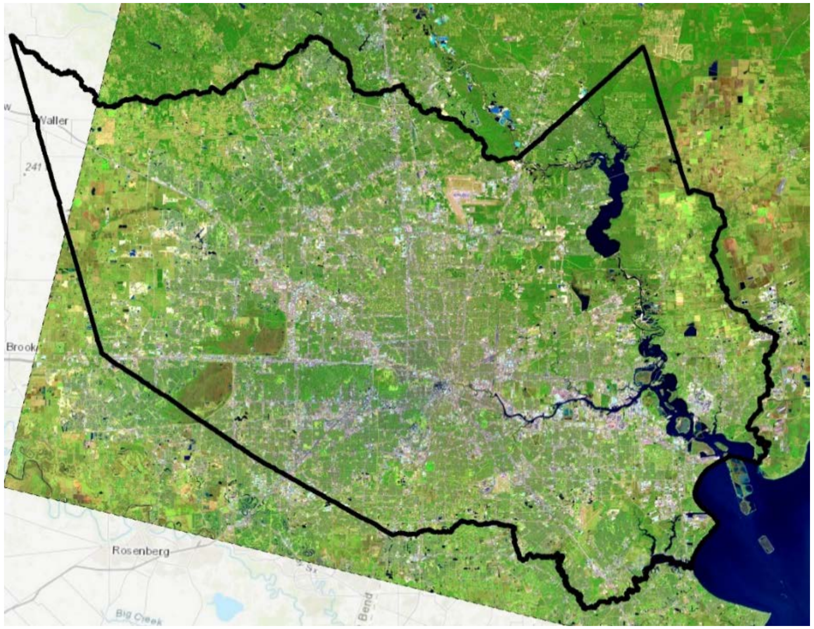


Landsat-8 Workflow: Technical Image Corrections



Landsat-8 Workflow: Technical Image Corrections

Tassled cap image:
- Greenness (green)
- Brightness (blue)
- Wetness (red)



Surface reflectance corrected image

UNIVERSITY OF SOUTH CAROLINA

Image Analysis Workflow

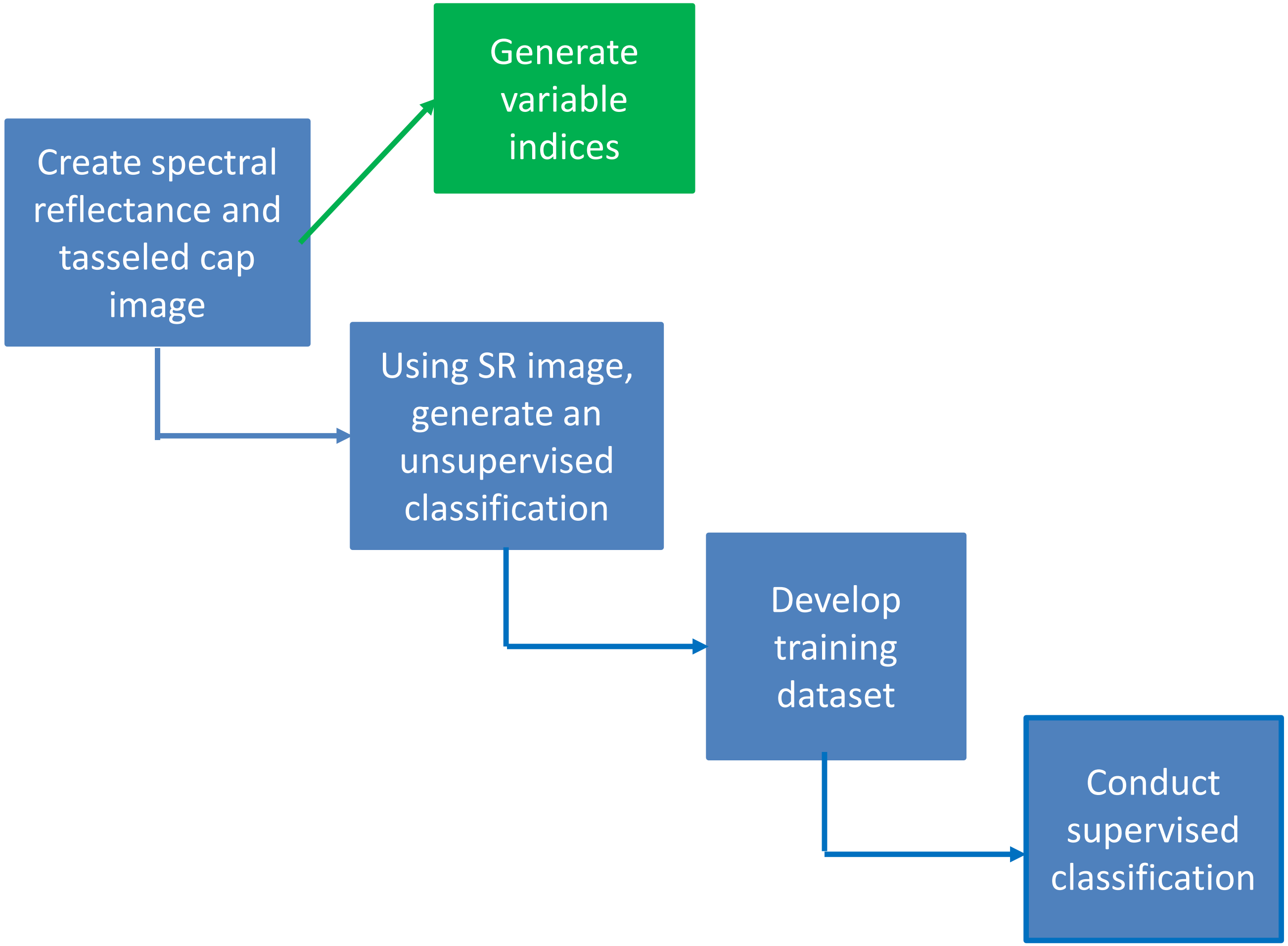


Image Analysis Workflow: Indices

- Allows for the detection of vegetation greenness (chlorophyll)
- Generated using ArcGIS geoprocessing tool
  - Recategorized into 20 even categories
- Generated for all 4 images

NDVI

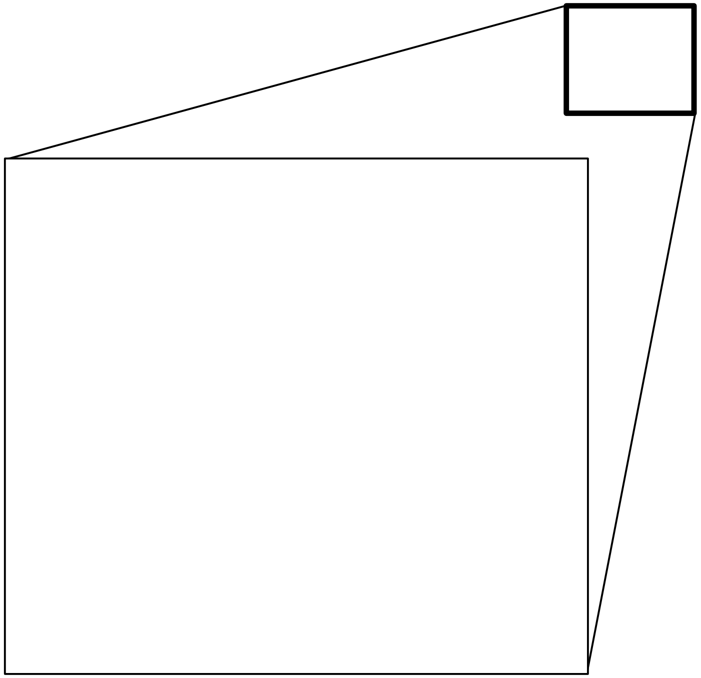


Image Analysis Workflow: Indices

- Allows for the detection of moisture in vegetation
- Generated using ArcGIS geoprocessing tool
  - Recategorized into 20 even categories
- Generated for all 4 images

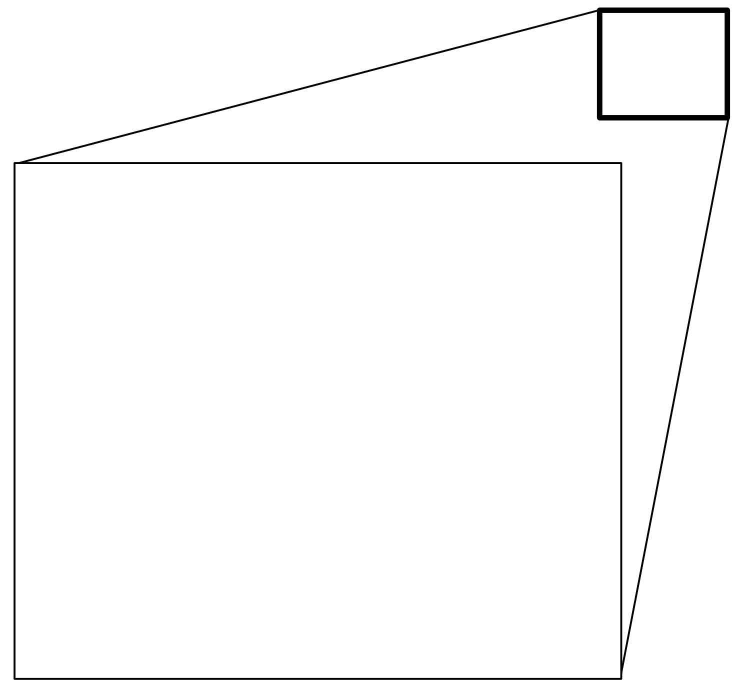


Image Analysis Workflow: Indices

- Allows for the detection of standing water
- Generated using standard formulas
  - Recategorized into binary variable
- Generated for all 4 images

NDWI

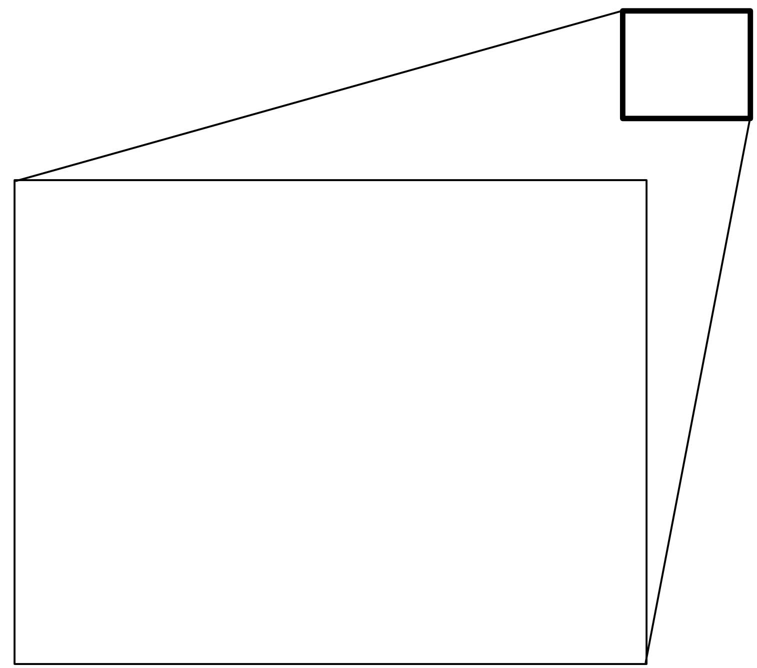


Image Analysis Workflow

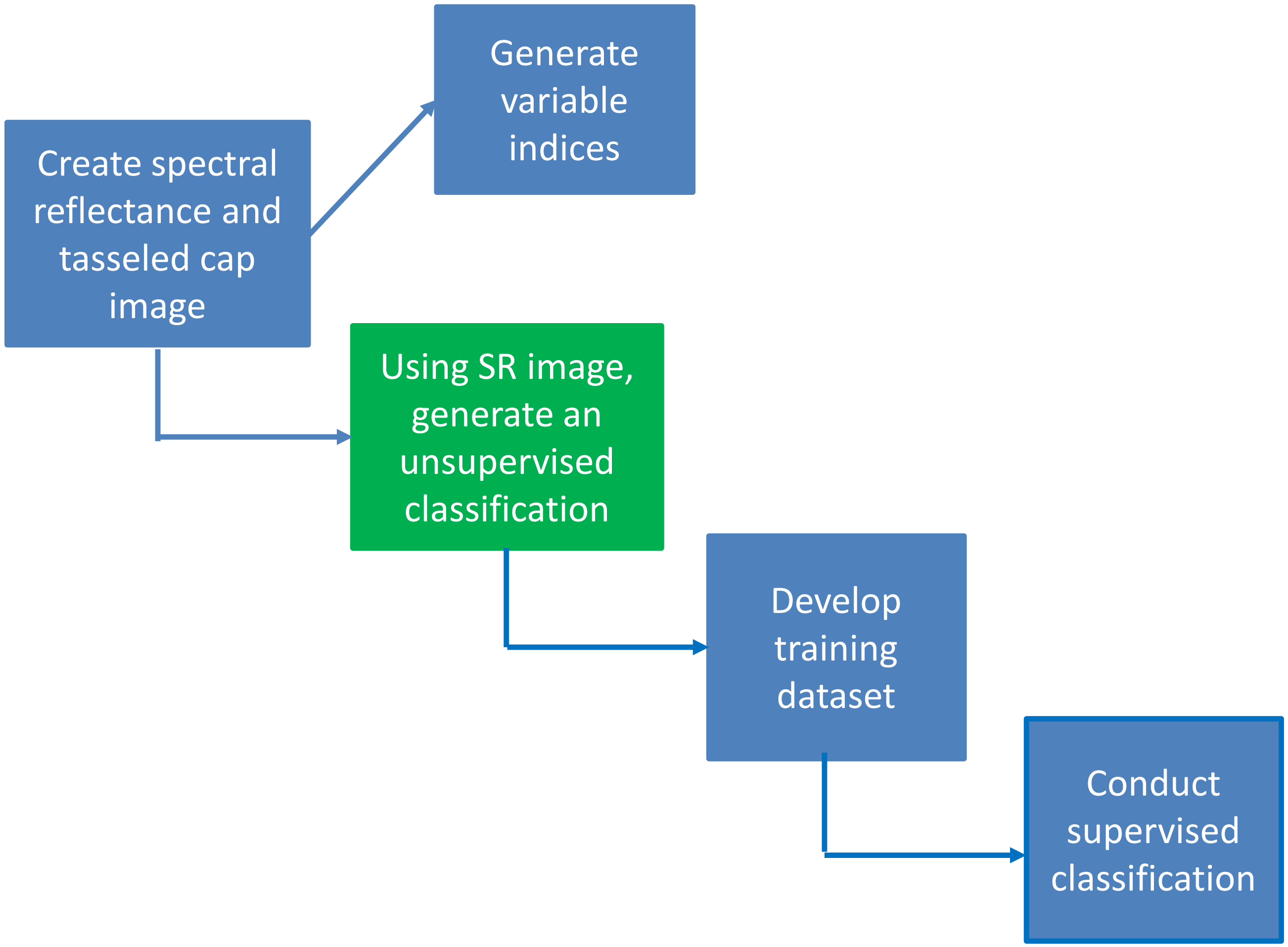


Image Analysis Workflow: Unsupervised Classification

Unsupervised classification

01/18/2018 image

Image Analysis Workflow: Unsupervised Classification

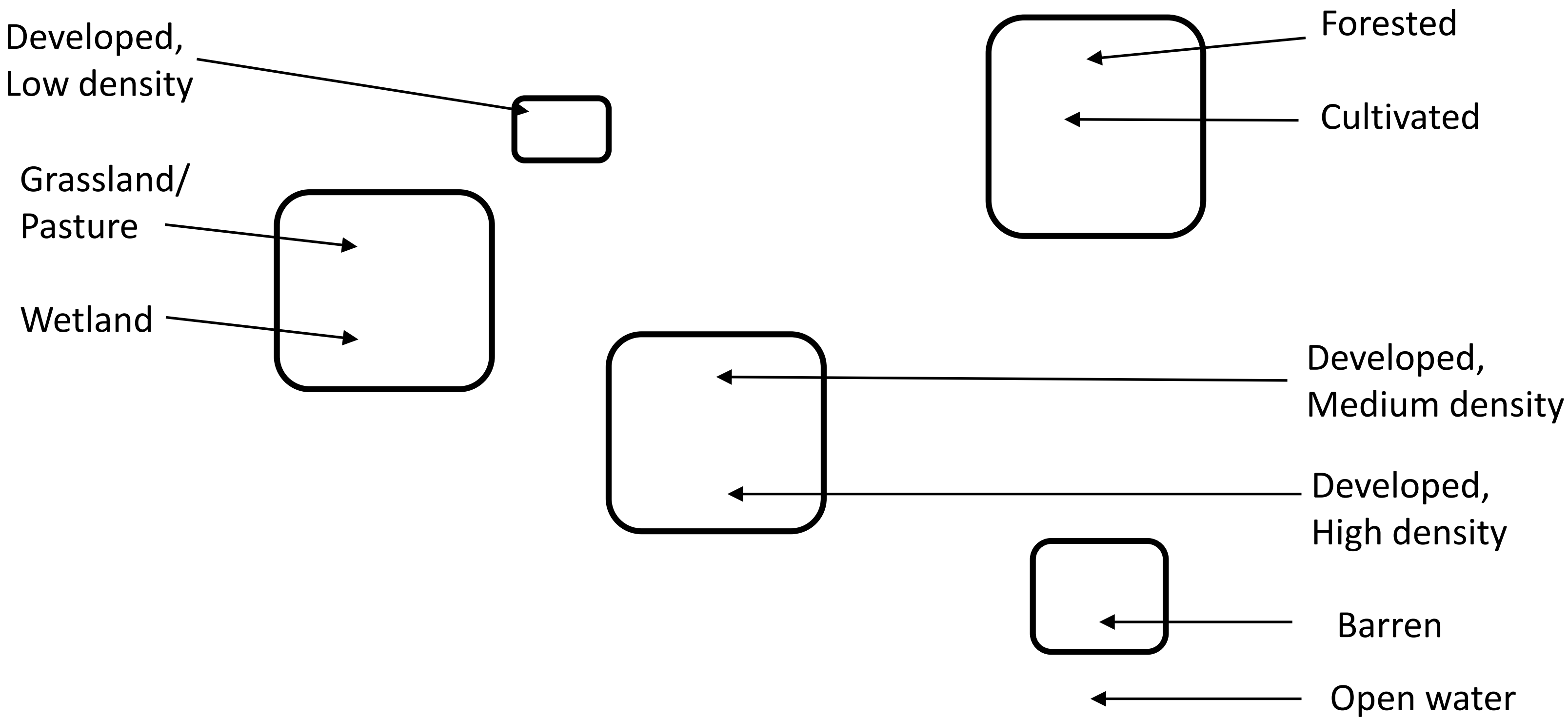


Image Analysis Workflow

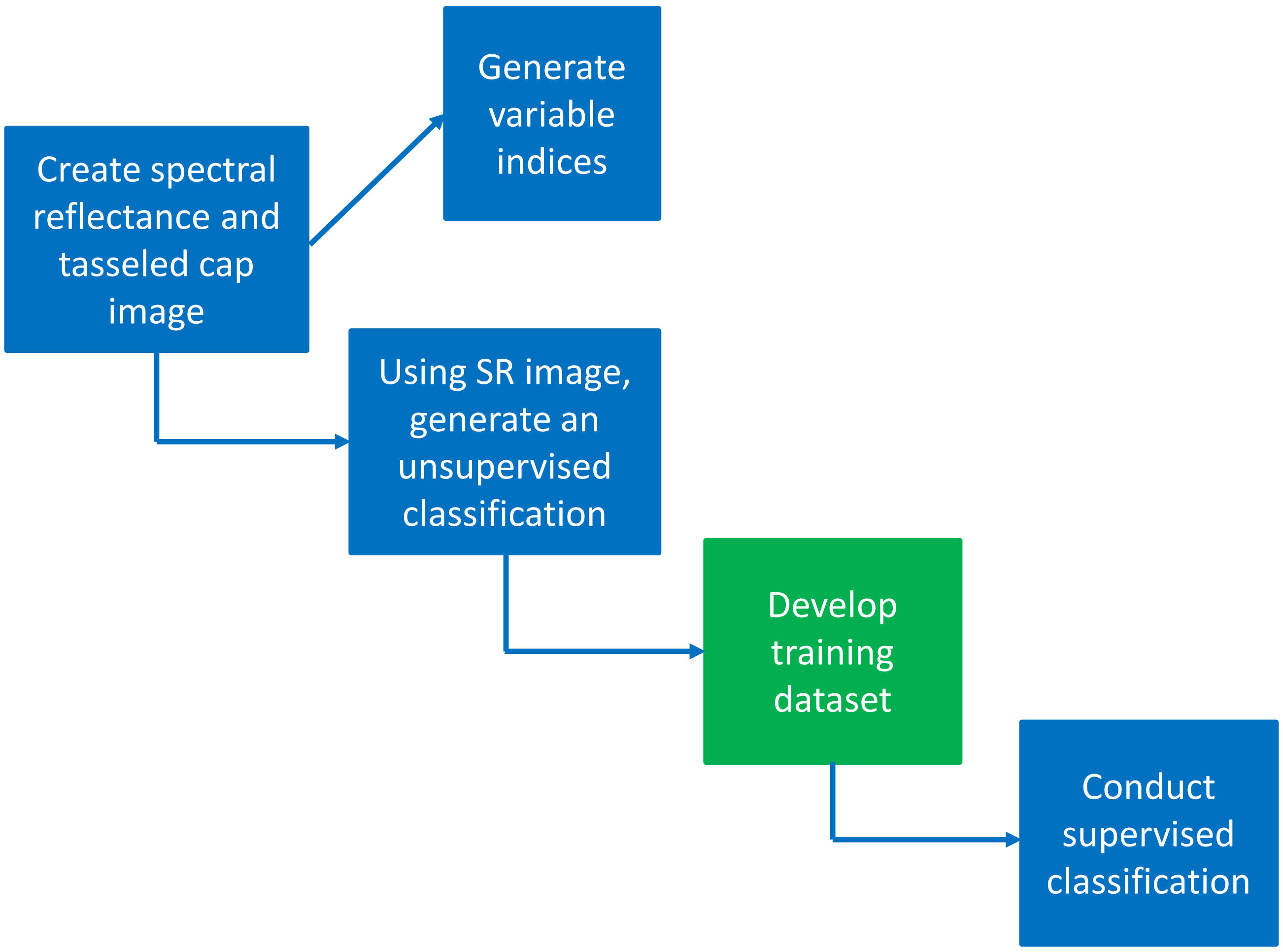


Image Analysis Workflow: Supervised Classification

Developed		
	open space	Mixture of some constructed materials, mostly vegetation from lawns. Includes: large lot single-family homes, parks, golf courses, erosion control areas.
	low intensity	Mixture of constructed materials and vegetation. Includes: single family homes, rural residential areas.
	Medium intensity	Mixture of constructed materials and vegetation. Includes: single family homes, suburban residential areas.
	High intensity	Highly developed densely populated area. Includes: urban centers, urban residential areas, shopping centers.
Barren		
	Barren or Unconsolidated shore	Areas of bedrock, desert pavement, gravel pits OR silt or sand that is subject to inundation by water. Minimal vegetation is present in the area.
Forest		
	Forest	Area with vegetation cover consisting predominately of trees including deciduous, evergreen, and/or scrub species
Pasture/Grassland		
	Grassland	Dominated by herbaceous vegetation. Areas are not utilized for intensive management such as tilling or farming but can be used for livestock grazing or the production of seed or hay crop, typically on a perennial cycle.
Cultivated		
	Cultivated Crop	Area of crop vegetation including annual crops, soybeans, vegetables, orchards, vineyards, etc.
Wetland		
	Wetland	Includes estuarine and palustrine wetland with forested, scrub, or emergent vegetation
Water		
	Water	Non-flowing, and non-flowing bodies of water

Image Analysis Workflow

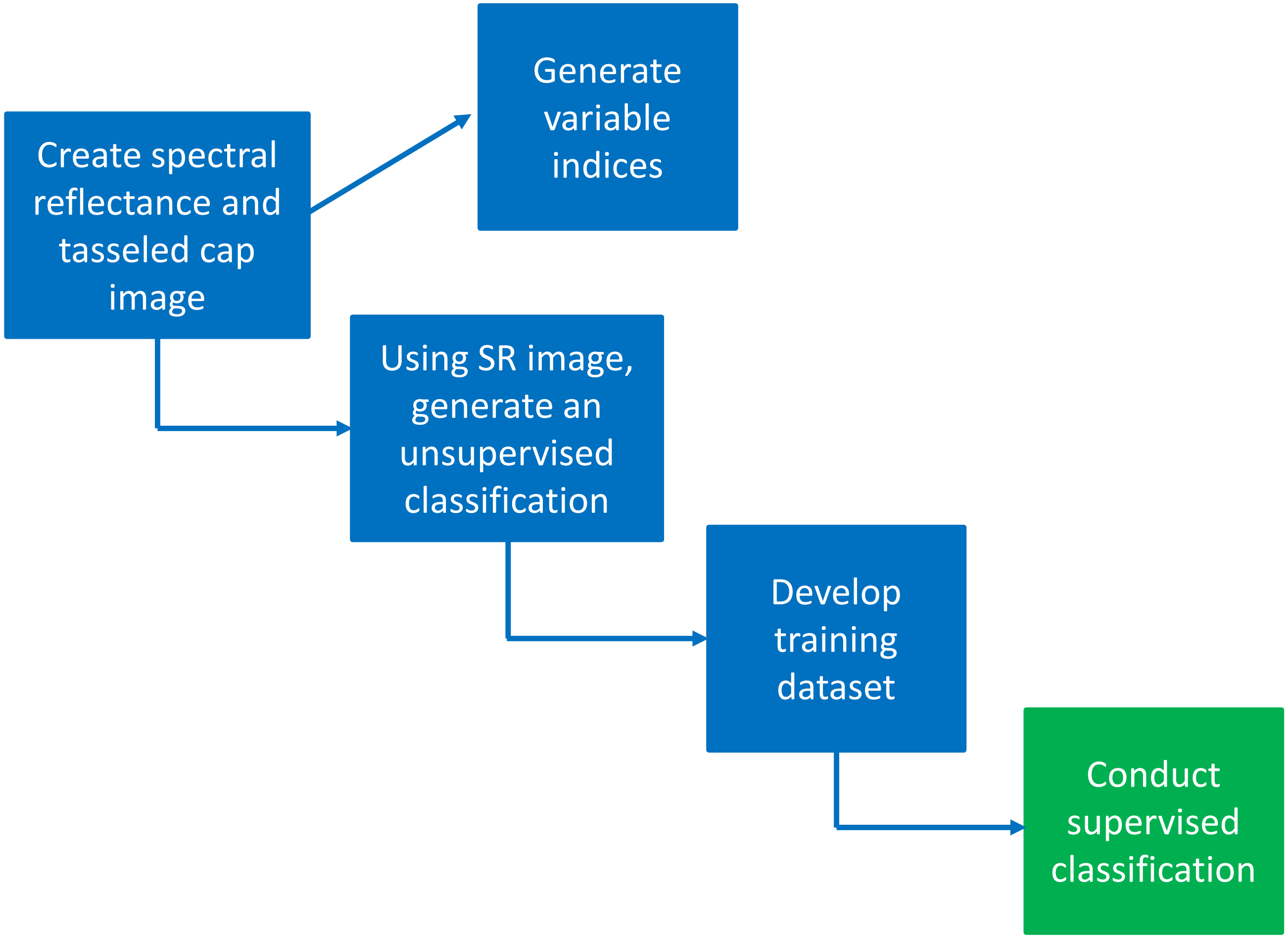
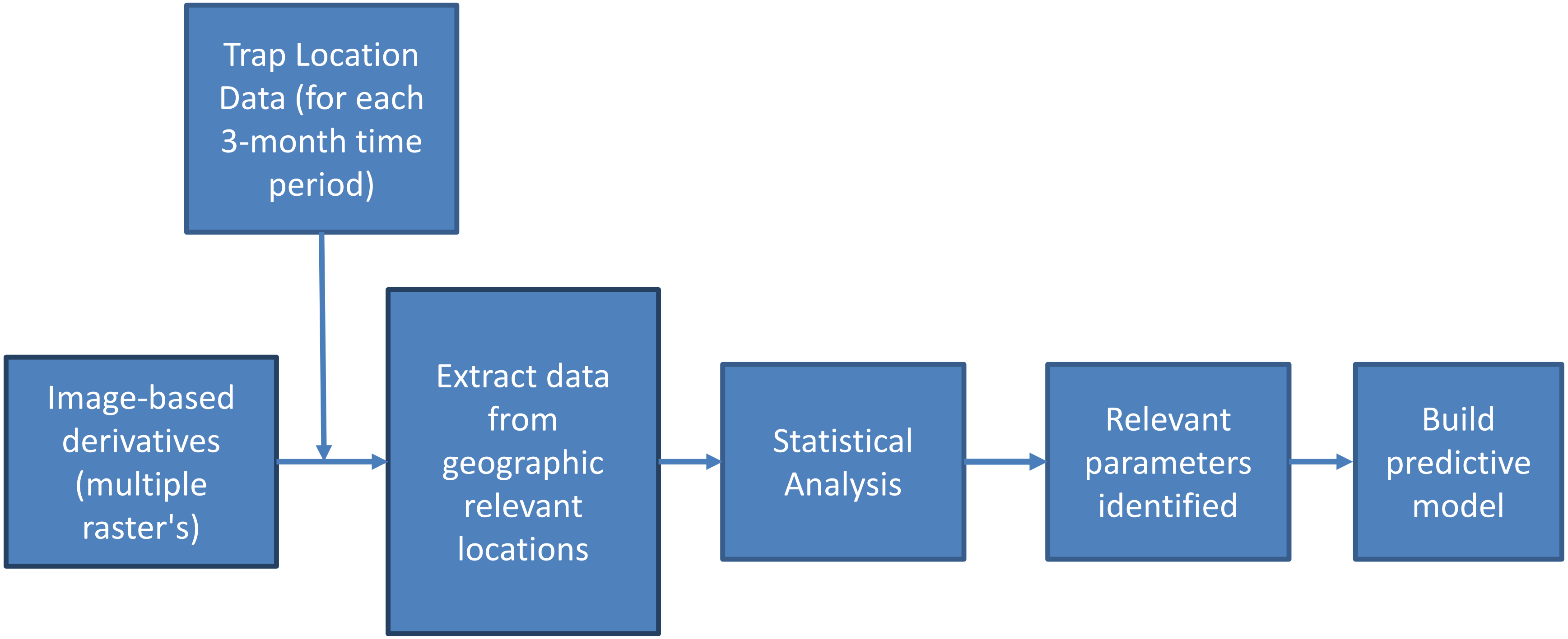


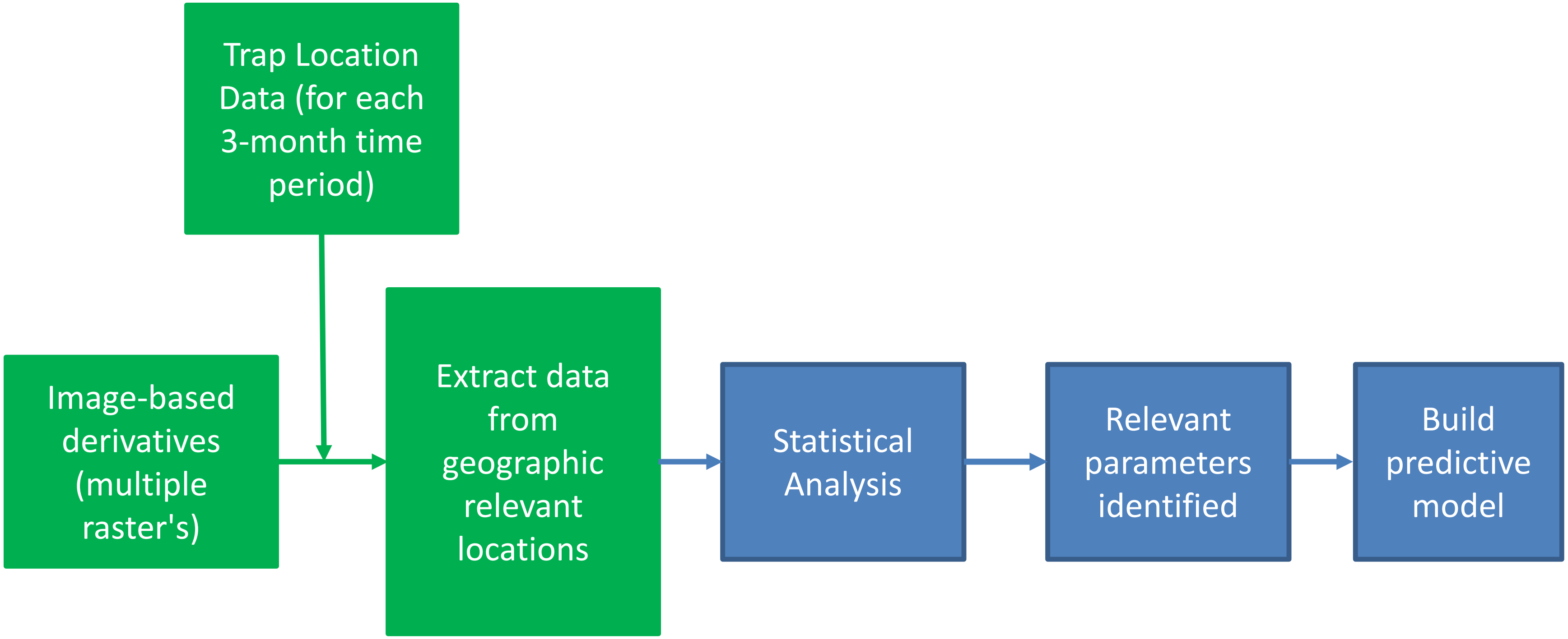
Image Analysis Workflow: Supervised Classification

- Barren
- Cultivated
- Forrest
- Grassland
- High Dev
- Low dev
- Med Dev
- Dev- open
- Water
- Wetland

Data Analysis Workflow



Data Analysis Workflow



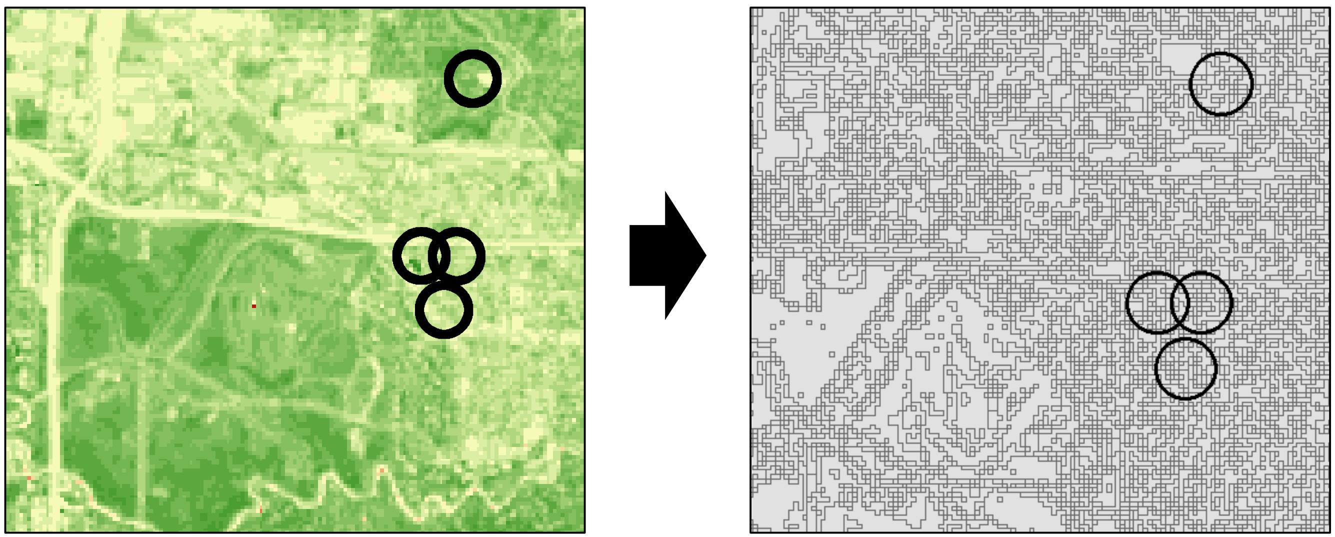
Data Analysis Workflow: Extracting data

- To capture relevant data around each mosquito trap a 200m buffer was generated

Nolan et al. 2012 J Biomed Biotechnol

UNIVERSITY OF SOUTH CAROLINA

Data Analysis Workflow: Extracting data



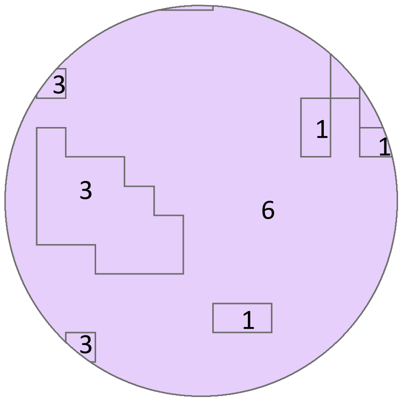
- The raster data (NDVI, NDMI, NDWI, and land cover) generated from our imagery analysis was converted to vector data
  - Raster to polygon conversion

Data Analysis Workflow: Extracting data

- Only data within buffer around the traps was extracted
- *Intersect and Clip functions*
    - *Buffer*
    - *Polygon analysis layers*

Data Analysis Workflow: Extracting data

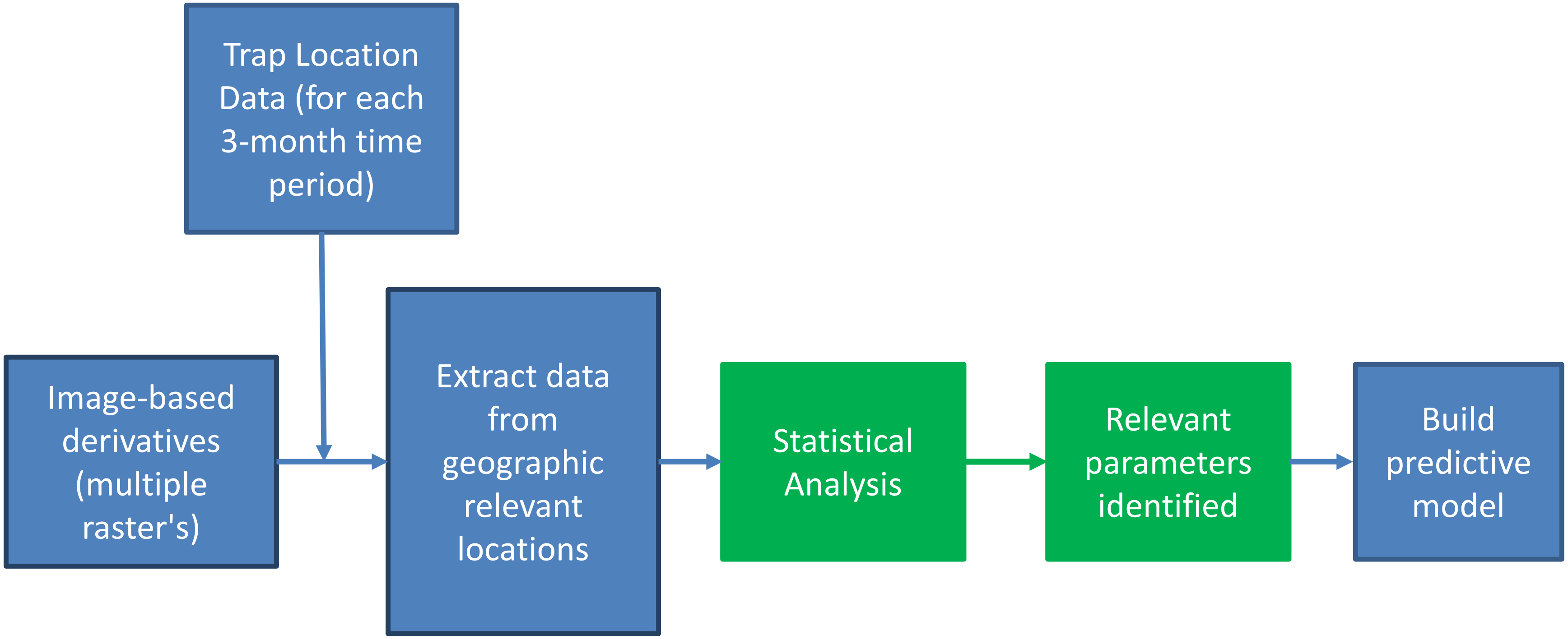
Land cover classification



	Class	Area	%
1	Water	100	5%
3	Medium Developed	400	20%
6	High Developed	1500	75%

Study period	Trap location	Median Mosquito count	Percent water	Percent Med. develop	Percent high develop	Median NDVI
10/17	79	56	5	20	75	15
4/18	223	102	0	45	26	9
7/18	106	28	2	40	15	20

Data Analysis Workflow



Data Analysis Workflow: Statistical Analysis Variables

- **Outcome variable**
  - Median Total female *Culex sp*. mosquitos per trap over the 3-month period represented by each image
- **Color Band Ratios (all images):**
  - NDVI
    - Median NDVI value for the buffer
  - NDWI
    - Percentage of the buffer with water
  - NDMI
    - Median NDMI value for buffer
- **Supervised Image Classification**:
  - Percentage of buffer made up of land cover category
  - Each land cover type is a variable in the analysis

Data Analysis Workflow: Univariate Analysis

- Negative binomial regression analysis
- We used backwards stepwise model building to determine our final multivariate model
- Initially we conducted univariate analysis to determine which variables are significantly associated and would be included in model building
  - Initial p-value cut-off of 0.25

Variable	P-value total
NDVI	<0.001
NDMI	<0.001
NDWI	0.002
grassland	0.753
Open developed	0.557
forest	0.728
cultivated	0.078
wetland	0.013
barren	0.8
Low developed	0.35
High developed	0.067
Med developed	0.103
water	<0.001

Data Analysis Workflow: Multivariate Analysis

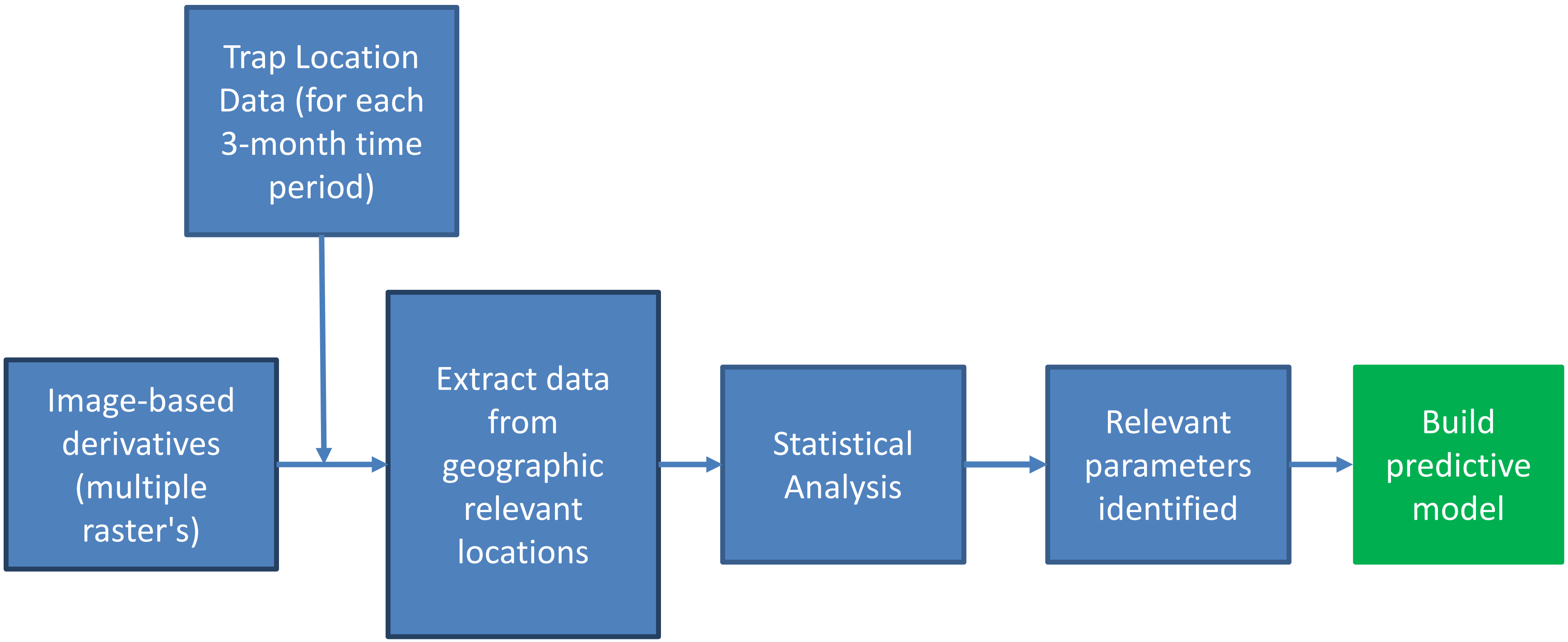
- Variables with a p-value of ≤0.1 were included in the final model
  - NDVI, Cultivated land, wetland, high development, and water were significant predictors of mosquito abundance

Variable	Coefficient	P-value total	95% CI
NDVI	0.113	<0.001	0.077-0.149
cultivated	0.864	0.014	0.178- 1.54
wetland	0.689	0.048	0.0074-1.370
high developed	0.222	0.024	0.030-0.416
water	-4.218	0.001	-6.78- -1.660

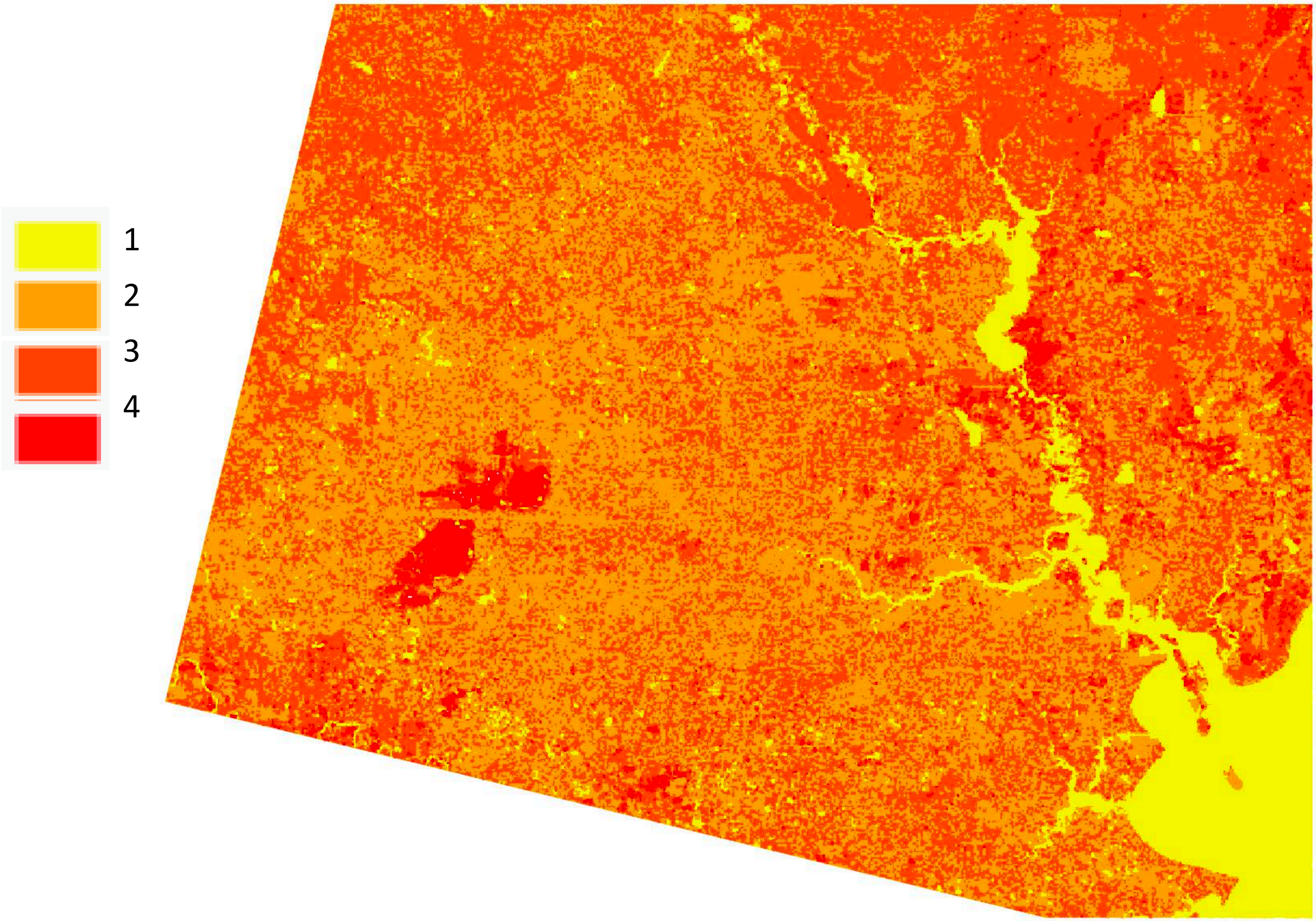
<u>Model Equation:</u>

Y= exp (NDVI*0.113 + Cultivated* 0.864+ wetland* 0.689+ High dev*0.222+ Water*-4.218 - 0.893)

Data Analysis Workflow



Remote Sensing Based Mosquito Risk Map



Future Directions

- **Improve model predictions using additional predictor variables**
  - Intermittent moisture
  - LiDAR- depressions in the earth that can hold water
  - Junk index-– need machine learning
- **Validate/rebuild model using higher resolution imagery**
  - Worldview 2 & 3 Imagery
- **Expand our efforts to new regions/new vectors**
  - Aedes spp. in South Carolina coastal regions
  - Tick-borne disease

Acknowledgements

- **Study Team**
  - ExxonMobil Upstream Division
    - Jerry Helfand (retired)
    - Tim Nedwed
  - Baylor College of Medicine
    - Dr. Peter Hotez
    - Dr. Abi Oluyomi
    - Dr. Sarah Gunter
  - University of South Carolina
    - Dr. Melissa Nolan
  - Harris County Public Health-Mosquito & Vector Control
    - Dr. Mustapha Debboun
    - Chris Fredregill
    - Kyndall Dye-Braumuller

- **Grant Support**
  - ExxonMobil Foundation
  - DigitalGlobe Foundation