Rebecca Risch and Elizabeth Pogue
EARS 77 - Environmental Applications of GIS
Professor Jonathan Chipman
13 February 2024

Lab 4 Report: Kuiseb River Riparian Zone Fractional Vegetation Cover

Introduction:

The objective of this lab is to model the Fractional Vegetation Coverage (FVC) in the Riparian Zone of the Kuiseb River, an ephemeral river that flows through the Namibian desert. To do so, we were given packages of UAV (drone) and Landsat-8 OLI imagery from 11 sites along the river and developed 4 models, one for the "Golf", "Foxtrot", and "Hotel" sites, and one for the three specified sites combined to model FVC for the whole study region. For each site, we first found the FVC from the UAV imagery by finding the NDVI (Normalized Differential Vegetation Index). We then manually set a binary threshold between soil and vegetation content per pixel based on the histogram generated by the NDVI layer, noting the minimum between the two peaks of NDVI-type frequencies to differentiate between pixels primarily containing vegetation vs soil. To translate this binary split between vegetation and soil from the 9x9 cm UAV imagery to the 30x30m landsat imagery, we created a fishnet grid of 30x30 cells restricted to the extent of the UAV image which we then used to store the fraction of binary values in each 30x30 cell. We then found the NDVI from the Landsat imagery, to use in later regression. From each imagery set (the UAV FVC and Landsat NDVI), we used the "Zonal Statistics as Table" tool to create tables relating the pixels to their mean statistic (UAV FVC or Landsat NDVI). We then used python packages to run a linear regression, modeling UAV FVC (y) vs Landsat NDVI (x), and applied that model to the Landsat pixels, based on their NDVIs, to predict their FVCs. After having run regressions at three separate UAV sites, we combined the data from each site to run a larger regression to model FVC over the whole study area. Finally, we masked out data from outside the riparian zone, as it is largely desert. The last step in our analysis included applying our FVC model to landsat images from both the oldest images in the dataset (1984-90) and the newest (2020-2021) to observe the change in vegetation, and also masked out data from outside the riparian zone to better visualize the change over time in the critical area.

Methods:

Overall Steps:

1. Create a regression model to predict FVC (0-100% vegetation cover) from Landsat imagery (here).
   a. Used Golf as test site to develop the methodology in ModelBuilder for relating the UAV and Landsat imagery
2. Create and run (in ModelBuilder) new regression models (repeat steps from above) for our other two sites. (here) - same as above

    a. Foxtrot
    b. Hotel
3. Combine the data from all three sites and produce a final regression model to predict FVC from Landsat data, based on all three sites together. (here)
4. Apply this combined regression model from all three sites to the entire Landsat image (the 30x30 meter grid cells along the entire length of the Kuiseb River within the study region) to produce the whole-region fractional cover map

Additional Work:
1. Constrain Study area FVC prediction to just the riparian zone
2. Time-series estimation of FVC for the whole study area

INDIVIDUAL SITE REGRESSION

**Because we completed the same process on all three of our UAV sites after developing the methodology for creating a regression model on the Golf test site, the documentation for each of these sites is identical in everything apart from the layer names. To show this and simplify the report, we walk through the steps for creating single-site regression models with a blank space ("____") to indicate where a layer name would be site-specific (either "golf," "foxtrot," or "hotel").
- The ModelBuilder used for this can be seen below, in figure 15.

Calculating UAV FVC (ModelBuilder): golf_ndvi - use this to get vegetation

1. **Calculate UAV NDVI (in order to find UAV FVC).** The Normalized Difference Vegetation Index is a standardized index that uses NIR and Red reflectance bands to display relative 'greenness,' or vegetation, on a landscape. We can use this to help us distinguish between bare soil and green vegetation.
    a. Tool: "Raster Calculator"
    b. Notes: Use the NIR (band 3 / "b3.img") and Red band (Band 1 / "b1.img") from the drone image
    c. Input: (NIR - R) / (NIR + R)
    d. Output: New Raster layer, named "___ndvi"
        i. "___ndvi" contains just 1 band, named band_1
2. **Look at the distribution (histogram) of golf_ndvi (not in ModelBuilder)**. Because the UAV images are such high resolution (9x9 cm), we are able to create a binary split between vegetated pixels and non-vegetated pixels to determine what percent of vegetation in a pixel is deemed "vegetation" and what's deemed "soil." These histograms and their respective vegetation/soil reflectance peaks can be seen below.
        i. Golf *(Figure 1)*, Foxtrot *(Figure 2)*, Hotel *(Figure 3)*

   b. Manually choose binary split between land and vegetation based on the low point between vegetation peak and soil peak reflectance in the histogram.
      i. Golf Binary Split: 0.065
      ii. Foxtrot Binary split: 0.055
      iii. Hotel Binary Split: 0.095
   c. (greater than = vegetation, less than = non-vegetated land, soil/sand)
3. **Binary Split between land and vegetation to create a layer representing FVC (titled golf_vegetation).** We used the Raster Calculator tool to apply an algebraic expression over the UAV raster layer (golf_ndvi) to create a new layer that contains the fractional vegetation cover for each 9x9 cm pixel.
   a. Tool: "Raser Calculator"
   b. Input: "____ndvi" > "Binary Split Value"
      i. Golf Binary Split: 0.065
      ii. Foxtrot Binary split: 0.055
      iii. Hotel Binary Split: 0.095
   c. Output: " ___vegeation"
4. **Input site mask to mask out all non-UAV areas from the landsat image.** This will allow us to estimate FVC of the landsat image over just the UAV area.
   a. Tool: "Raster to Polygon" (changed raster mask layer to a vector feature)
   b. Input: "____mask.img"
   c. Output: "____mask_polygon"
      i. This will be input for select layer by location
5. **Fishnet tool**. Allows us to calculate the fraction of vegetation vs. non-vegetation (1 vs 0 binary values) and store it in a 30x30 grid. We can assign the value in each individual fishnet cell and then correspond with the landsat grid cells.
6. (for later use in calculating fractional vegetation cover in each grid cell). In order
   a. Tool: "Fishnet"
   b. Template extent: golf_landsat.img (so that it matches the UAV and landsat area)
   c. Geometry type: polygon
   d. Cell size: 30x30m (to match Landsat.img cell size)
   e. Output: golft_fishnet
      i. Fishnet is not a layer in the project until we add it : we need to make feature layer
7. **Export feature layer before zonal stats as table.** This will make the fishnet layer into a raster feature so that we can use it for the following tool, to select
   a. Tool: "Make Feature Layer"
   b. Output: "_____fishnet_layer'
8. **Select layer by location.** This will allow us to limit our fishnet to just the UAV extent by selecting the fishnet cells that are spatially related to another dataset. In this case, this would be the cells that are completely within the UAV polygon ("___mask_polygon")

   a. Tool: "Select layer by location"
   b. Input feature = "____ fishnet_layer" (from step 6)
   c. Selecting feature: "____mask_polygon" (from step 4)
   d. Relationship: Completely within
   e. Output:
      i. A layer with selection (masked fishnet)
      ii. The output layer names
      iii. Count
9. **Exporting Feature**
   a. Tool: "Export Feature" (changes shapefile to a feature class - vector polygon)
      i. Input: layer with selection
      ii. Output: "____ finshnet_ExportFeature"
10. **Zonal Stats as table: (for UAV FVC).** This tool/step extracts the stats (min, max, mean, etc) from a 'zone' of one raster layer. The zone from which these values are extracted is determined by another raster or vector layer (the input raster or feature). This will allow us to extract data from each of the cells to use for the regression model.
    a. Tool: "Zonal Statistics as table"
    b. Input raster or feature zone data: "____ExportFeature" (from step 8)
    c. Input value raster: golf_fvc
    d. Zone field: objectID (manually write this in)
    e. Stats type: Mean
    f. Output: "____fvc_table"
11. **Alter field of above stats table to rename MEAN column to fvc_mean.**
    a. Tool: "Alter Field"
    b. Input: "____fvc_table"
    c. Field name: MEAN
    d. New field: fvc_mean

Finding NDVI for Site Landsat Image (ModelBuilder). Because we want to predict FVC for the whole *landsat* image, we need landsat NDVI to use as our independent variable in the regression. This process follows the same steps as before, when we found the NDVI for the UAV image.

1. **Calculate NDVI from NIR and Red bands in landsat image**
   a. Tool: "Raster Calculator:
   b. Input: (Layer 5(NIR) - Layer 4 (red)) / (Layer 5 + Layer 4)
   c. Output: "____ndvi"
2. **Input this NDVI layer to Zonal stats to get a table.** This will allow us to collect stats (based on the UAV image extent) from the NDVI raster layer into a csv table, to use in regression.
   a. Tool: Zonal Statistics as Table

b. Input value raster: "____ndvi"
c. Input raster or feature zone data: Output: "____finshnet_ExportFeature"
d. Stats type: Mean
e. Zone field: object id (manually write this in)
f. Output: ndvi_table

3. **Alter field of above stats table to rename MEAN column to ndvi_mean**
   a. Tool: "Alter Field"
   b. Input: "___ndvi_table"
   c. Field name: MEAN
   d. New field: ndvi_mean

Joining / Exporting FVC and NDVI output tables for regression model (ModelBuilder). This step will join the two data tables we created for the landsat NDVI and the UAV FVC into one table, using the objectID as the key.

1. Joining the UAV FVC and Landsat NDVI tables together to prepare for regression model:
   a. Tool: "Join Field"
   a. Input table: "_____ ndvi_table"
   b. Input join field: objectID
   c. Join table: "_____ fvc_table"
   d. Join table field: objectID
   e. Output table: "_____joint_table.csv"
2. Export table as CSV, save in folder
   a. Tool: Export Table
   b. Output: "___fvc_table"

Python Portion (running regression model) - further explanation of steps in python notebook/code comments

1. Import libraries
2. Import csv file from folder
3. Assign x and y variables
   a. X = landsat_ndvi mean (independant)
   b. Y = uav_FVC mean (dependant)
4. Run regression on x and y
   a. "___model" = stats.linregress(__x,__y)
   b. Viewing plot
      i. Golf *(figure 10)*, Foxtrot (*figure 11*), Hotel *(figure 12)*
5. Note slope and y intercept
   a. GOLF SLOPE: 1.58667
   b. GOLF INTERCEPT: -0.014846196

c. FOXTROT SLOPE: 1.9768
d. FOXTROT INTERCEPT: -0.1037
e. HOTEL SLOPE: 1.683441
f. HOTEL INTERCEPT: -0.138237600

Landsat FVC prediction. Apply the regression model to the landsat image to predict FVC over each site on the landsat level (30x30 m grid). Results pictured below: Golf, (*Figure 4*), Foxtrot (*Figure 5*), Hotel (*Figure 6*).

1. **Apply regression model to landsat image (step)** to estimate FVC for landsat based on landsat NDVI. We used a linear regression model.
    a. Tool: "Raster Calculator"
    b. Input: ($b_0 + b_1X$)
        i. b_1 (slope) = varies per site / per regression model used
        ii. b_0 (y-intercept) = varies per site / per regression model used
        iii. X = ___landsat.img
        iv. (SLOPE*(____ndvi)) - Y-INT
    c. Output: "____fvc_pred"
2. **Add layer to map**

WHOLE SITE REGRESSION MODEL

Combine the data from all three sites and produce a final regression model to predict FVC from Landsat data, based on all three sites together.

1. In Excel, we joined the three tables containing UAV FVC and landsat NDVI into one large table, which we then ran the same regression on in python as we did on each site individually to calculate a regression model for the whole study region. This gave us a new slope and intercept to apply to the whole landsat image.

Whole Python portion (running regression model).

6. Import libraries
7. Import csv file from folder (that contains data from ALL three UAV sites)
8. Assign x and y variables
    a. wx = whole_joint_table (landsat) NDVI mean (independant)
    b. wy = (UAV) FVC mean (dependant)
9. Run regression on x and y
    a. View plot, as seen in figure 13
    b. whole_model= stats.linregress(wx, wy)
10. Note slope and y intercept
    a. WHOLE SLOPE: 1.892557

    b. WHOLE Y-INT: -0.12228208

WHOLE STUDY REGION APPLICATION

Applied our combined regression model from all three sites to the entire Landsat image (the 30x30 meter grid cells along the entire length of the Kuiseb River within the study region) to produce the whole-region fractional cover map

Applying Regression model (Done in ModelBuilder, as seen in figure 14).

1. Calculating NDVI for the whole landsat image to use in regression, using the NIR (layer 4) and Red (layer 1) bands
    a. Tool: Raster Calculator
    b. Input: (Layer 4 - Layer 1) / (Layer 4 + Layer 1)
    c. Output: whole_ndvi
2. Calculating whole landsat FVC prediction
    a. Tool: Raster Calculator
    b. Input: (1.892557 * whole_ndvi) - 0.122282)
    c. Output: whole_fvc_pred
        i. Fvc range: 1.80446 - -3.705

MASKING RIPARIAN AREA

Masking out non-riparian areas (in ModelBuilder, after calculating the whole area FVC prediction). This will allow us to show predicted FVC for just the river's riparian zone, as most study region consists of desert, with little to no vegetation.

1. Use the Extract by Mask tool to extract the cells of a raster (whole_fvc_pred) that correspond to the areas defined by a mask (riperian_polygon)
    a. Tool: Extract by mask
    b. Input Raster: whole_fvc_pred
    c. Input raster or feature mask data: riperian_polygon
    d. Output: riperian_fvc_pred
2. Results: See figure 9

APPLYING WHOLE FVC MODEL TO PAST LANDSAT IMAGES (TIMESERIES)

Applied our combined regression model from all three sites to entire Landsat Images from 1984, 1990, 2020, 2021 to produce the whole-region fractional vegetation coverage map over these specified years.
In ModelBuilder (see Model here, figure 16)

1. For each year's landsat image, Make Raster Layer out of bands of the Landsat image to isolate single bands for NDVI calculation (repeat for Band 1 (RED) and Band 4 (NIR)
    a. Tool: Make Raster Layer
    b. Input Raster: (ex:) L5_1984_09_09.img
    c. Output Raster Layer Name: (ex:) RED_1984
    d. Bands: (ex:) 1
2. Calculate NDVI using Raster Layers of Landsat Imagery Bands
    a. Tool: Raster Calculator
    b. Input: (ex:) ("%NIR_1984%" - "%RED_1984%") / ("%NIR_1984%" + "%RED_1984%")
    c. Output Raster: (ex:) ndvi_1984
3. Predict FVC using Whole model's slope and intercept
    a. Tool: Raster Calculator
    b. Input: (ex:) (1.892557 * "%ndvi_1984%") - 0.12228208
    c. Output Raster: (ex:) fvc_pred_1984
4. Average the 1984 & 1990 FVC prediction layers, and also separately, the 2020 & 2021 FVC prediction layers
    a. Tool: Raster Calculator
    b. Input: (ex:) ("%fvc_pred_1984%" + "%fvc_pred_1990%") / 2
    c. Output Raster: (ex:) mean_84_90
5. Subtract the average layers to find the change in FVC between the modern and antique periods
    a. Tool: Raster Calculator
    b. Input: "%mean_20_21%" - "%mean_84_90%"
    c. Output Raster: fvc_pred_diff
6. Extract by mask to mask data out outside of Riparian Zone (gives better visualization)
    a. Tool: Extract by Mask
    b. Input raster: fvc_pred_diff
    c. Input raster or feature mask data: riparian_polygon.shp
    d. Output raster: fvc_pred_diff_rip
    e. Extraction Area: Inside
7. Results: See figure 17.

Results:

	Binary Threshold (UAV NDVI)	Model	Slope	Intercept	R-Squared	Std Error
Golf	0.065	Linear Regression	1.58667	-0.014846196	0.5795	
Foxtrot	0.055	Linear Regression	1.9768	-0.1037	0.8519	
Hotel	0.095	Linear Regression	1.683440795	-0.138237600	0.8182	
Whole	—	Linear Regression	1.892557	-0.12228208	0.7925	

Analysis:

Across our four regression models (Golf, Foxtrot, Hotel, and the Whole), the Golf site seemed to be somewhat of an outlier from the others, whose values across the board were more similar to each other then they were to Golf, as you can see in the above table. The slopes for our regression models are relatively similar, apart from the Golf site, which is the lowest of our slope values. As a higher slope indicates a stronger positive relationship between NDVI and FVC, it makes sense that as the model with the lowest slope, Golf is also the model with the lowest R^2 value.

Notably, the R-squared value of Golf (0.5795) is much lower than the other models' R-squared values (29.16% lower than the closest site, Hotel (0.8182)). As the R^2 value indicated the independent variable (NDVI)'s ability to explain all the variability in the dependent variable, in this case the R^2 of 0.5795 means that the NDVI explains approximately 57.95% of the variability in the Golf FVC. In the other 3 regression models, this R^2 value is higher (closer to 1), suggesting a better fit of the model to the data at these sites. Our best fit was with the Foxtrot model, (which also had the highest slope) and who's R^2 value indicated that NDVI predicts 85.19% of the variability in the Foxtrot FVC. For these models, in particular the Golf site model, the fact that NDVI while NDVI is a predictor of FVC (a more significant predictor, depending on the model), there are probably other factors influencing FVC that are not captured by the model. We think these could include things like climate/precipitation, vegetation types, etc. When looking at the histograms for our three independently modeled sites (Golf, Foxtrot, Hotel; Figures 1-3 respectively) one can see that in the Golf_ndvi histogram, there is a much higher rate of reflectance from vegetation than the other two. This might explain some of the difference between Golf and the other two sites. One can also see, in figures 4-6, that there is a much starker difference in FVC at the riparian zone and further away from the river at the Foxtrot and Hotel site compared to the Golf site. The regression

We noticed that slope and R^2 value seem to positively correlate, which came as a surprise to us because we also noticed, while working on the maps, that so too did slope and range of FVC. FVC should ideally range between 0 and 1, but the models, being imperfect, often ranged a bit outside of those values. Seemingly, the higher the slope, the wider the range of FVC. As R^2 is the coefficient of determination, or proportion of the variation in the dependent variable that is explainable from the independent variable, we'd expect predictions with wider ranges (erroneously outside of 0 and 100% covered in vegetation) to be less accurate, where UAV FVC is less explainable from Landsat NDVI than if the range was closer to plausible values of FVC (between 0 and 1).

In an effort to understand the change in FVC over time, we averaged the landsat image FVC predictions (based off of our Whole model) between 1984 and 1990, and also between 2020 and 2021. We subtracted these averages (modern versus antique) giving us percent change in FVC between these two time periods. The percent change in FVC ranged between 20 and -100 percent, meaning the fractional vegetation of the pixels in the riparian zone changed somewhere between 100% loss and 20% gain. Because we used the Whole model, the R^2 value remains 0.7925 for these predictions.

FIGURES:

Figure 1: Golf Site Histogram

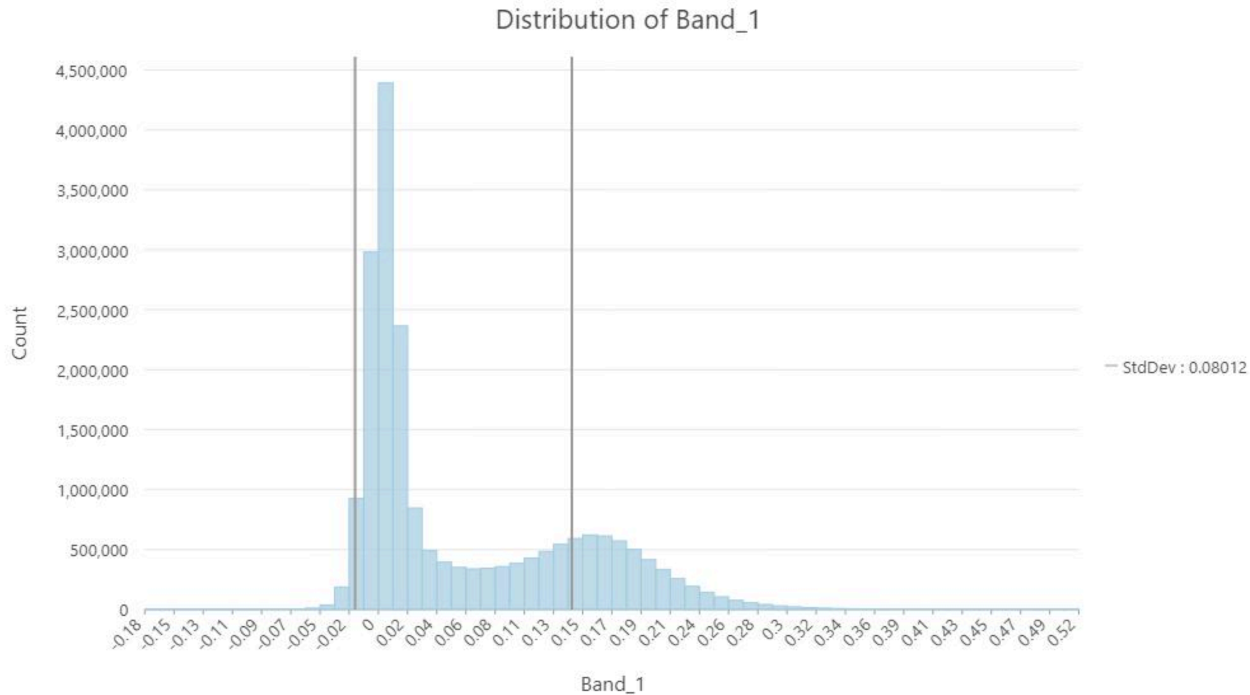

Figure 2: Foxtrot Histogram

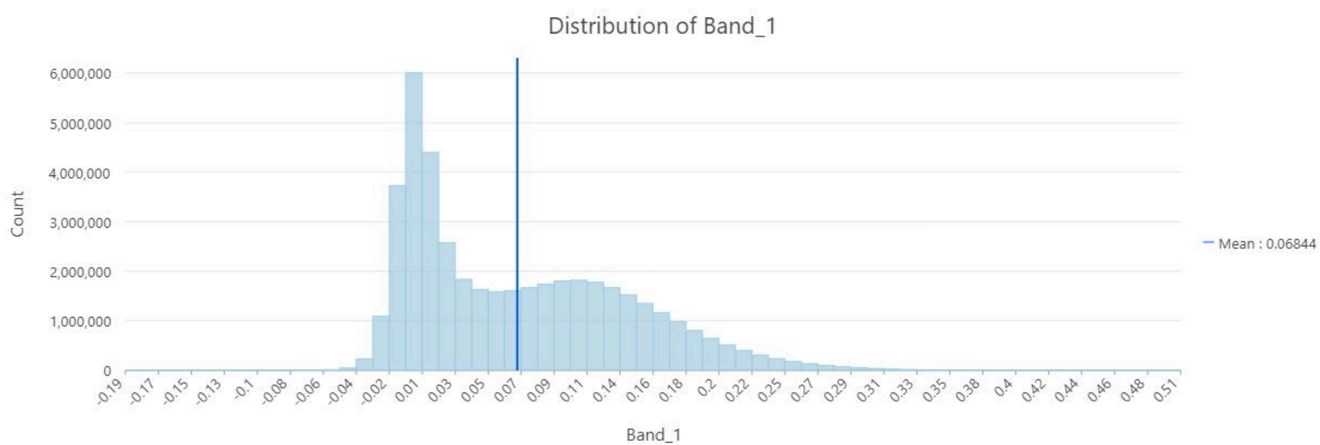

Figure 3: Hotel Histogram

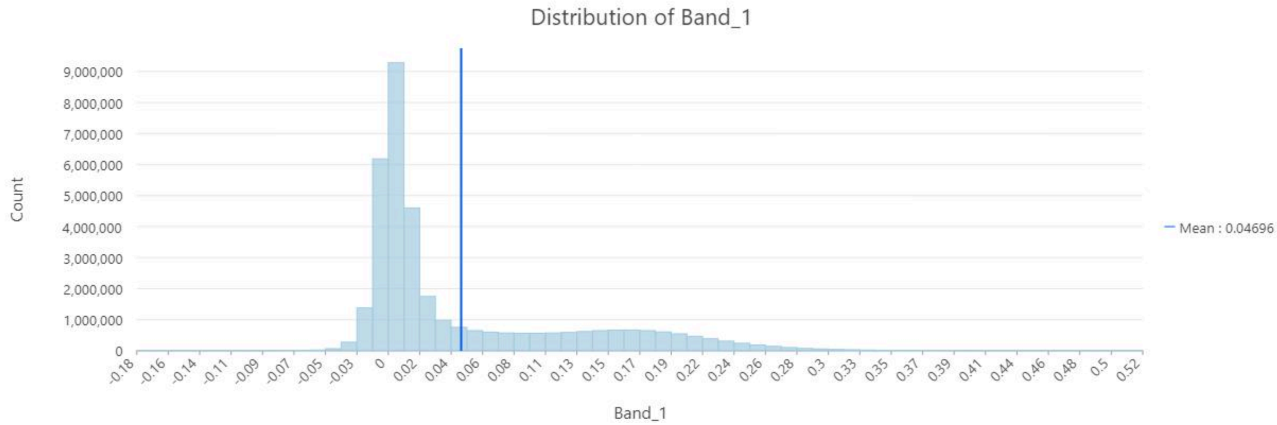

Figure 4: Golf Site Predicted FVC

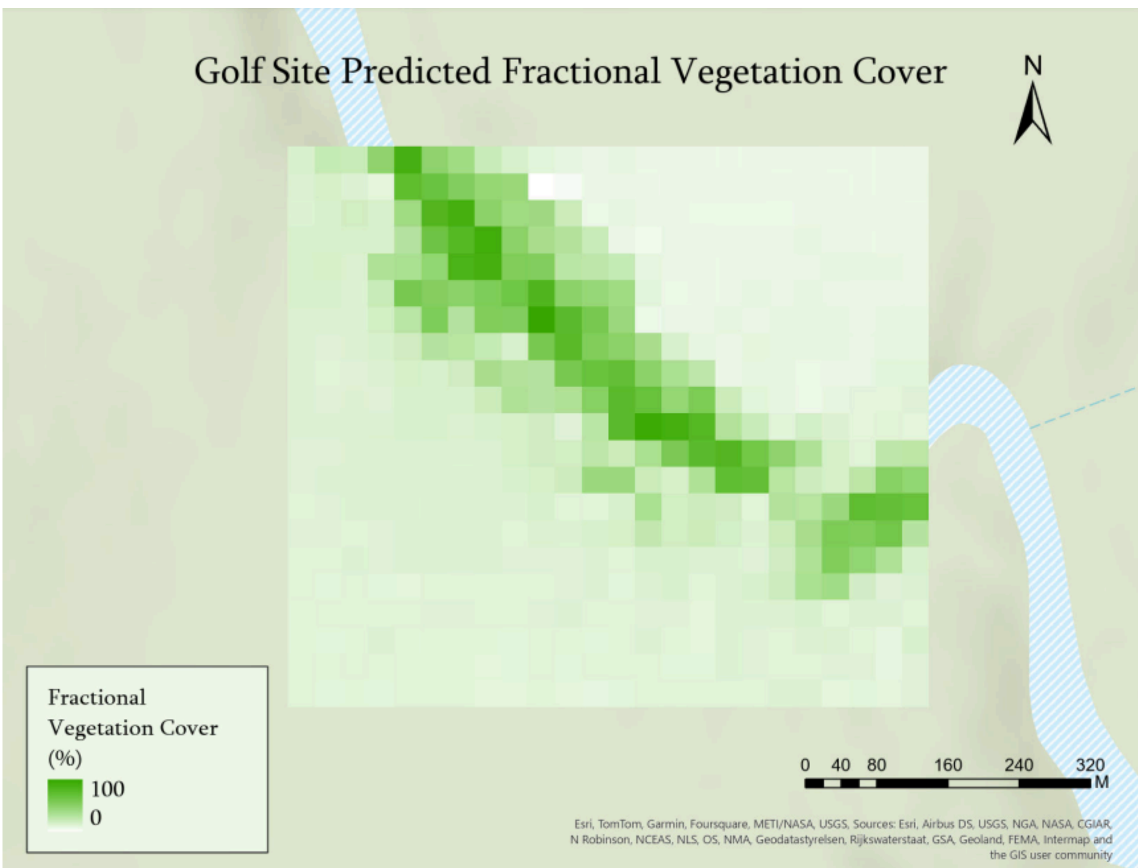

Figure 5: Foxtrot Site Predicted FVC

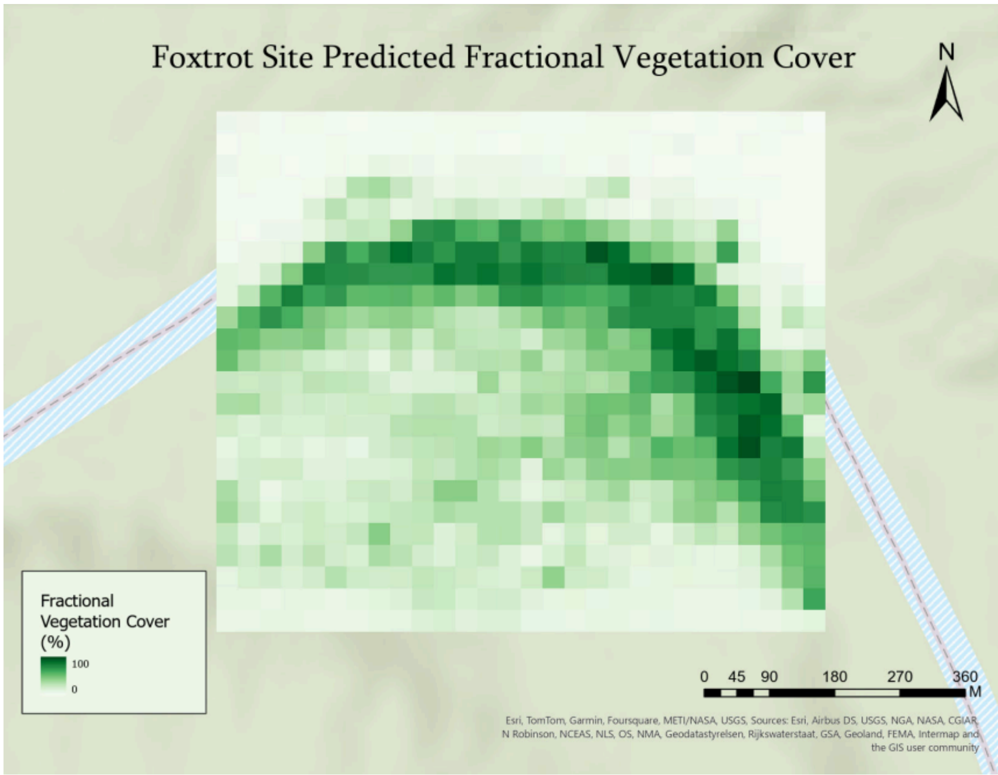

Figure 6: Hotel Site Predicted FVC

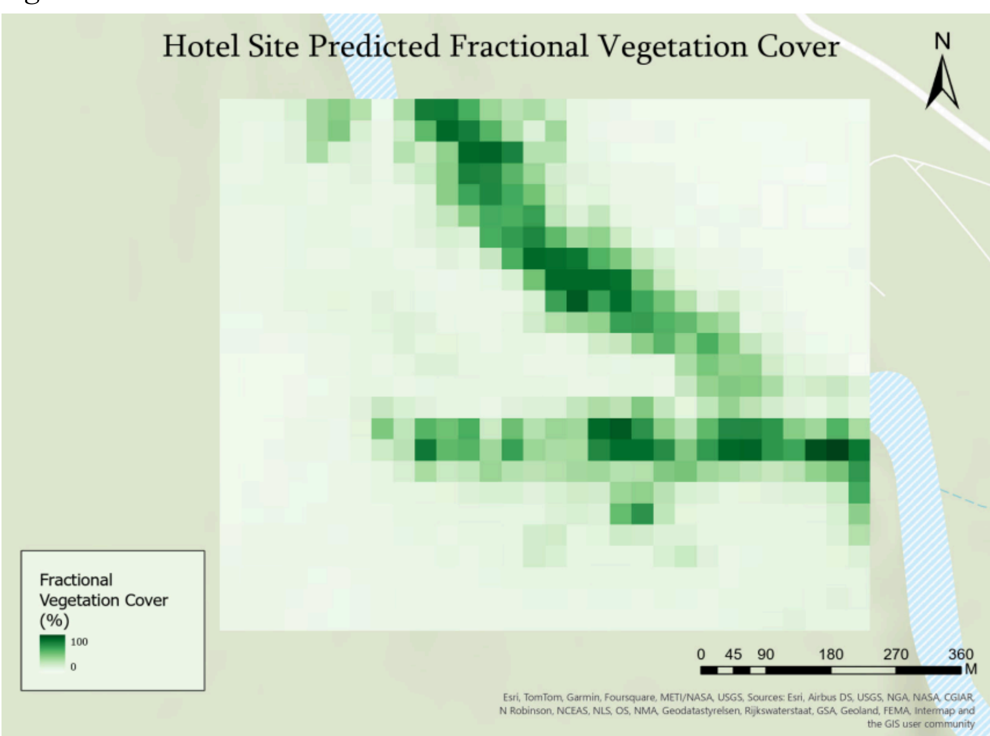

Figure 7: Whole Study Region Predicted FVC

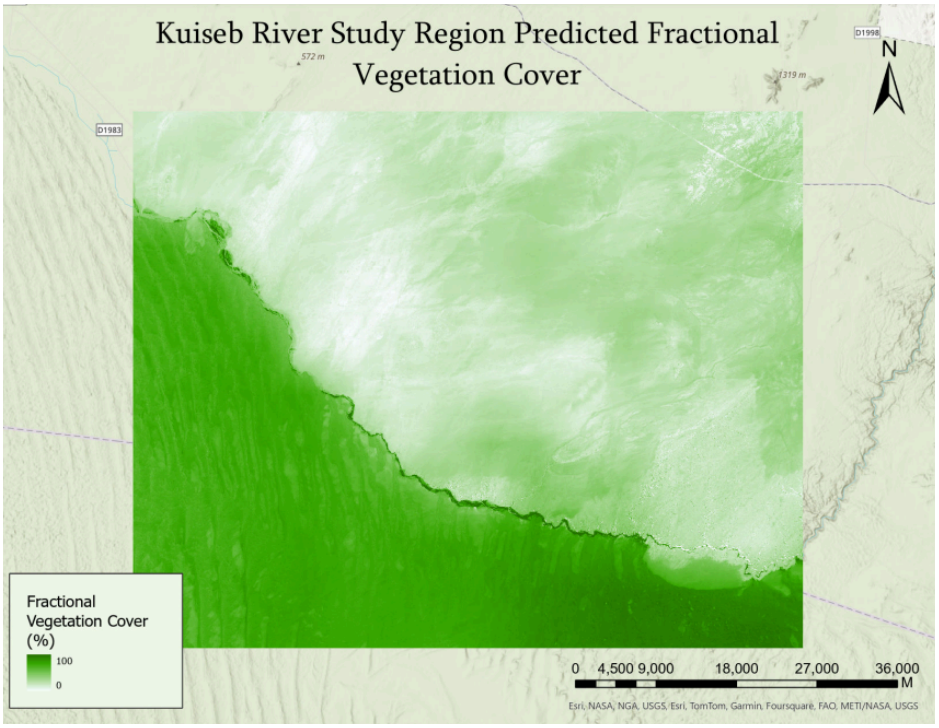

Figure 8: Enlarged Foxtrot Site; Predicted FVC

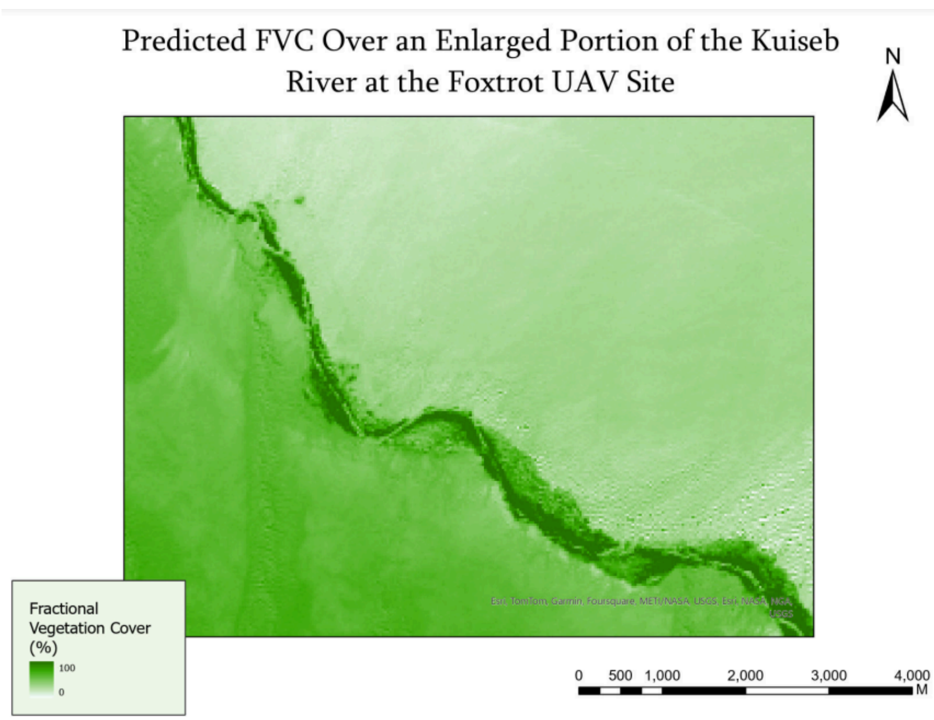

Figure 9: Riparian Zone Over Enlarged Foxtrot Site; Predicted FVC

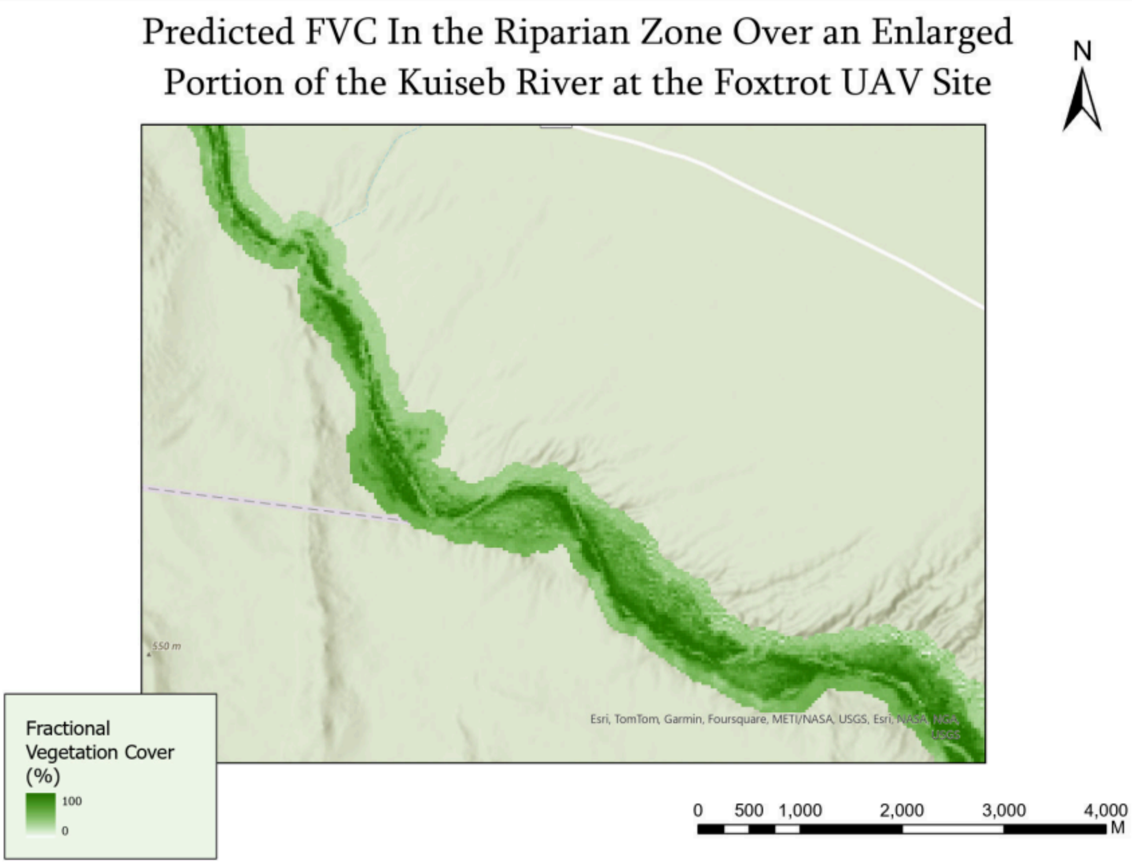

Figure 10: UAV FVC vs Landsat NDVI plot, Golf model

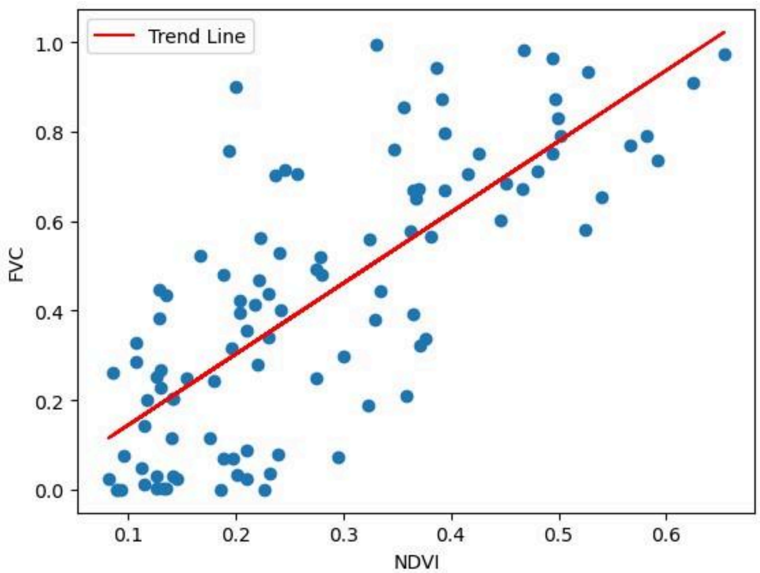

Figure 11: UAV FVC vs Landsat NDVI plot, Hotel model

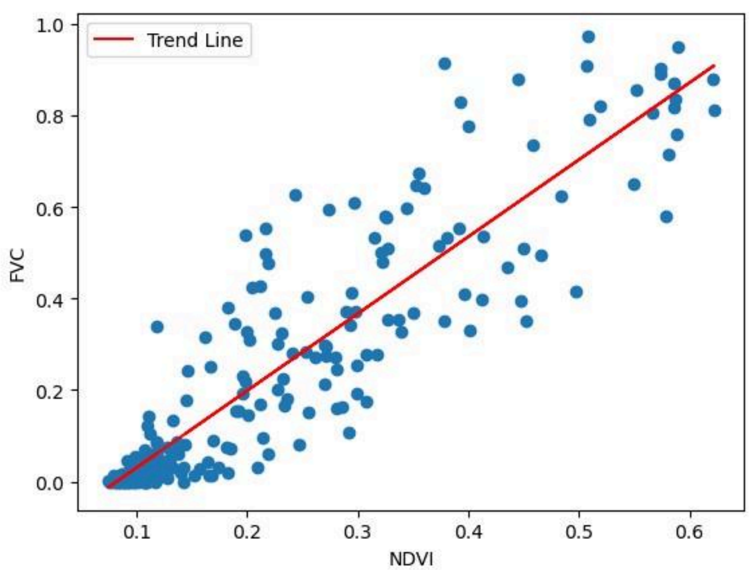

Figure 12: UAV FVC vs Landsat NDVI plot, Foxtrot model

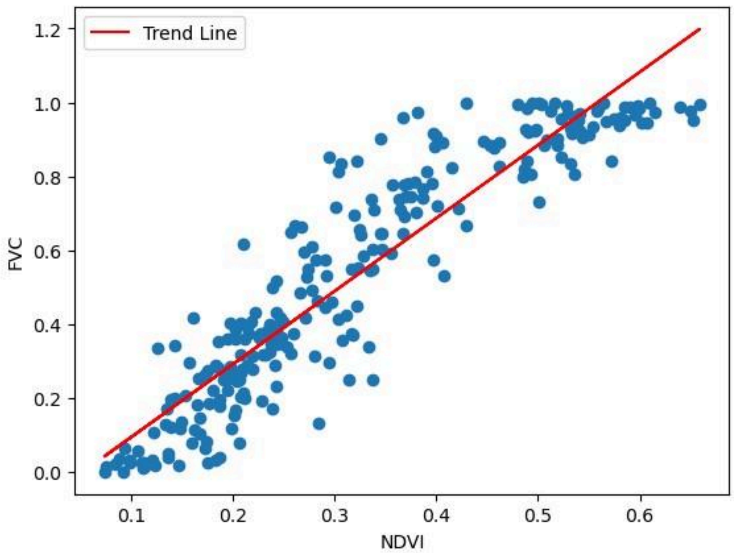

Figure 13: UAV FVC vs Landsat NDVI plot, Whole (combined) model

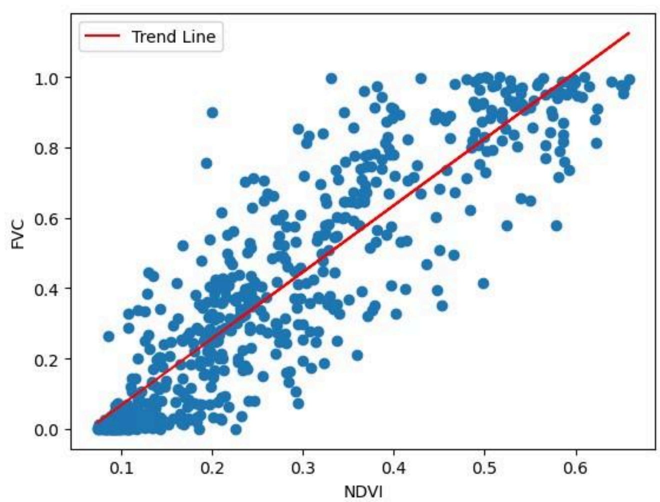

Figure 14: Whole Study Area Regression and Riparian Zone Mask

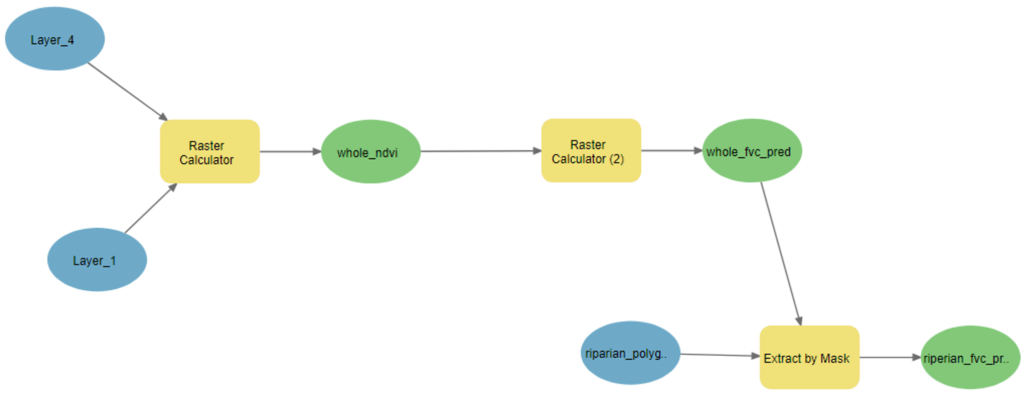

Figure 15: Individual site methodology for building FVC regression model (Foxtrot site pictured, but was the same for Golf, Foxtrot, and Hotel).

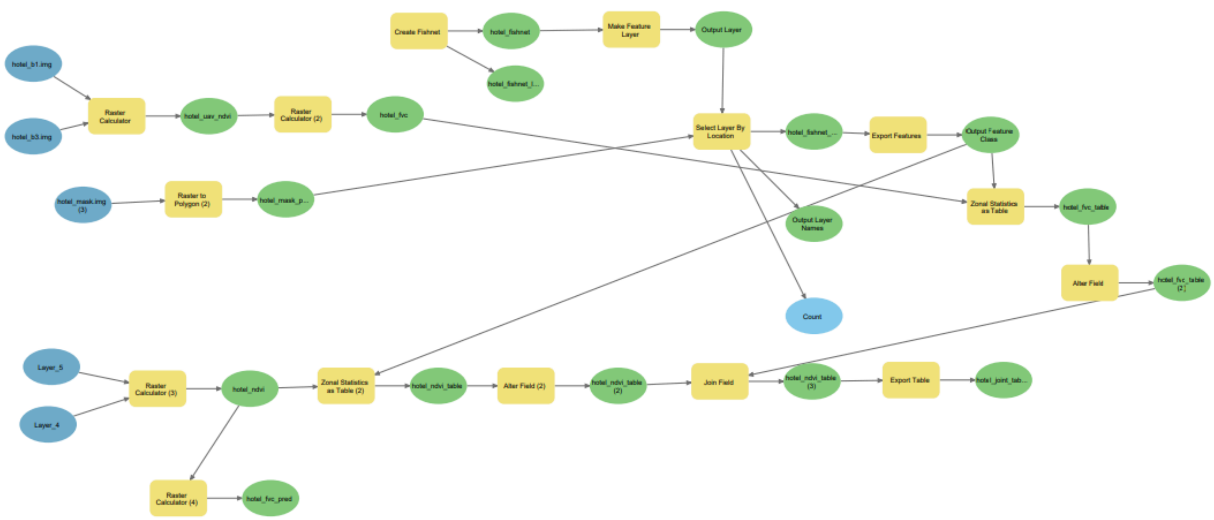

Figure 16: ModelBuilder to apply FVC to past years of landsat images, average the early years and the late years together, and calculate the change between the two (early vs late)

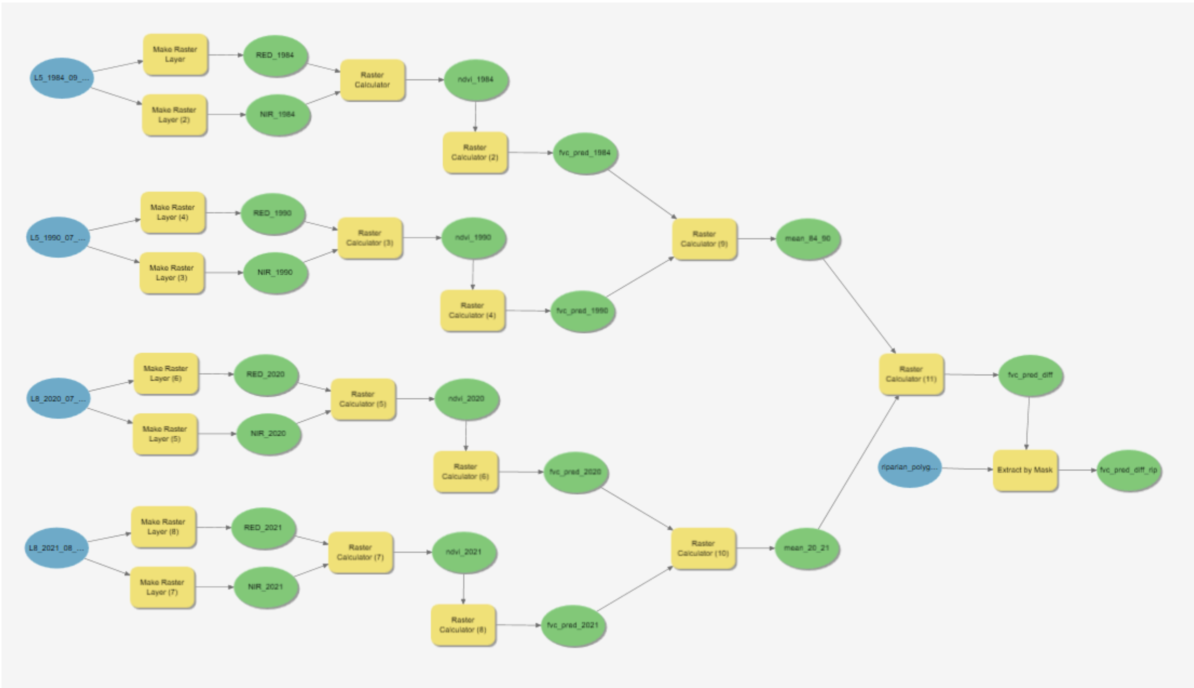

Figure 17: Average Change of Predicted FVC in the Riparian Zone between 1984/1990-2020/2021 Over an Enlarged Portion of the Kuiseb River at the Foxtrot UAV Site