**Next steps:**

Set up RStudio workflow:

For environmental data:

Bring in data:

If uploading straight to R - separate for locations of each sample

or upload already extracted info from QGIS

make single data set

remove NA’s

Correlation analysis

Do I approach it the same way as the budworm?

Combination of correlation factor, clustering with dendrogram and VIF?

Explore data structure

Which model? Likely poisson?

Carbon ~ poisson(?)

Carbon = logit? (µ)

µ = ∑ßenvx\*Envx

fit to carbon estimates

predict() function for new environmental variables

what spatial extent?

\* for spatial analysis see Heckford et al 2021 Statistical analysis:

1. Extracted the value of each remotely sensed covariate
   1. Elevation (and landcover type?)
      1. combined DEM images together to create a seamless raster
      2. ‘Clip’ function we limited our DEM raster to our AOI (what is R equivalent?)
      3. normalized our aspect raster by replacing any value > 180 by subtracting -180
      4. ‘subs’ to normalize using legend of corresponding values
      5. ‘raster.transformation’ from ‘spatialEco’ to standardize
   2. Forest inventory – stand info
      1. Using unique forest polygon identifiers, we attributed spatial covariates to the FRI datasets
      2. Unoverlap using clip, erase, and merge in ArcGIS (R equivalent?)
      3. Subset to variables of interest
      4. Remove ‘white-space’ (non intentional ones)
      5. Text values to integers
      6. Create raster ‘rasterize’ in ‘raster’ for each covariate
   3. Tree type
      1. Create codes (represent different dominant tree species) for vector mask of polygons?\*\* confused by this part
2. General Linear Models (GLM) with the response variables against explanatory spatial predictors
3. ranked models based on Akaike Information
4. removed models with uninformative parameters
5. If more than one model was within a\2 DAICc we averaged model coefficients and extracted full coefficient estimates for use in the construction of distribution models
6. Constructed spatial surfaces of foliar ESP traits by summing top model or the averaged coefficients estimates for top competing models, multiplied against their corresponding spatial covariate -> create predictive spatial surfaces of plant trait variability

Start deciding on sample sites:

Look at map compared to those selected by Rachael

What were the ones she didn’t get to/back up

Will we be going to the more remote sample sites?

What distribution at bw GM and TN are Jeremy and Elizabeth looking

Am I including fire - yes

\*\* could I use some of Heckford’s carbon data to validate or expand the data set?

Does what we sample include parts of what he sampled for carbon and could our data be subset to match his? Is the sampling procedure similar?

To decide which sites to visit:

Decide which metrics I want to get full distribution:

* Forest age and height (height is good so probably just age)
* Elevation for sure!
* We will see about bioclim – just annual averages I think \*will take a while to get info
* Species