**NMDS**

I have run numerous variations of NMDS on various combinations of the data. Some examples include:

* LPI with soils as environmental factor
* Soils with LPI as environmental factor
* LPI with shrub density as environmental factor
* Shrub density with soils as environmental factor
* Using just the seven categories we came up with
* Using height classes, and live/dead in height classes

I have also run various combinations of specific soil variables, as well as doing all of these on just my data, just the USGS data, and on both combined. I have tried several distance measures including Euclidean and Bray-Curtis.

**Random Forests**

Classification Forest on binary presence absence of live sagebrush and of live and dead sagebrush combined on soils (Also with ATCA2 & KRLA2).

Regression Forest on ARTR2, ATCA2, KRLA2, BOGR2, SPCR, & HECO26.

These have been run with several attempts at tuning parameters such as cp, minsplit, and using testing and training datasets, as well as different random forest packages.

These were run on all applicable soil variables as well as subgroups suggested by Janis and the Variable Importance Plots. These were also run on both the LPI and shrub density data.

The Boruta package was also used with Random Forests to attempt variable selection. It was run repeatedly on all of the soils data, as well as multitudinous combinations.

This was done using just our data from North and South Plains, just USGS data, and both our datasets combined.

Results were not consistent and had extremely low % variance explained. No amount of tuning or variable selection changed this.

**PCA**

On Soils for variable reduction

On both LPI and Shrub Density to see if would result in shrubs and grass pulling apart.

Done at various scales as well as on correlation matrices.

**Poster Analysis in SAS**

Looked at two models: Live and dead, presence and absence. This was suggestive, but certainly not a strong analysis, especially considering the use of stepwise selection, and the fact that the live and dead analysis had even fewer data points.

**Various Univariate and multivariate prescreening, and changes to soil variables**

Prescreening included 1 way ANOVA, correlation and multicollinearity reduction, and lots of combining categorical soil data classes. New soils variables were intermittently added including depth weighted average (DWA) of the entire pedon, and of specific sections of a pedon, variables at specific depths, maximum & minimum values, surface and subsurface values, and many more. This meant rerunning several analyses each time the soils variables were changed to see if it made a noticeable difference to the statistics.

**K-Means Cluster Analysis With NMDS**

Run on LPI not soils or shrub density

Run with and without PCA and on all LPI as well as subsets. Still need to talk to Susan about whether some of the things I did make sense. From what I can tell, the brown is mostly PJ communities, the blue is primarily shrubless communities( SALS0 and OPPO), and the green and red seem to be the shrubs(ARTR2, ATCA2, KRLA2) and the grasses (BOGR2, SPCR, HECO26) and their associated forbs. I was also able to overlay the soil environmental variables. It gives us some ideas about the communities, but they are roughly coming out of where ARTR is, so my best guess is that ARTR prefers lower values of these variables.



