Python in Earth Science Applications

Ali L. Yagci

Post-Doctoral Fellow

Hydrological Sciences Laboratory (617)

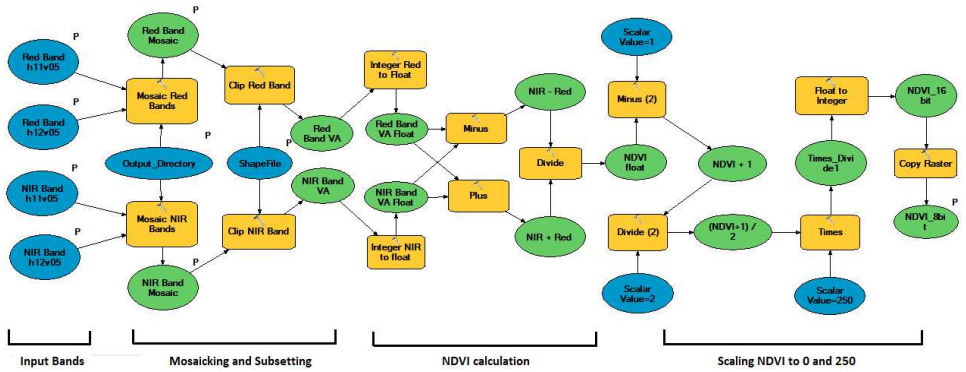
First Acquaintance

- Background
  - Geographical Information Systems (GIS)
  - Remote Sensing
- Introduction to GIS programming back in the graduate School
  - Introduction to Python
  - Implement a method to read shape file and visualize it on Tkinter.
    - Python's de-facto standard GUI (Graphical User Interface) package.
  - ESRI shape file specifications
- Introduction to Object-Oriented Programming
  - JAVA
- MATLAB
  - Heavy and hungry for memory

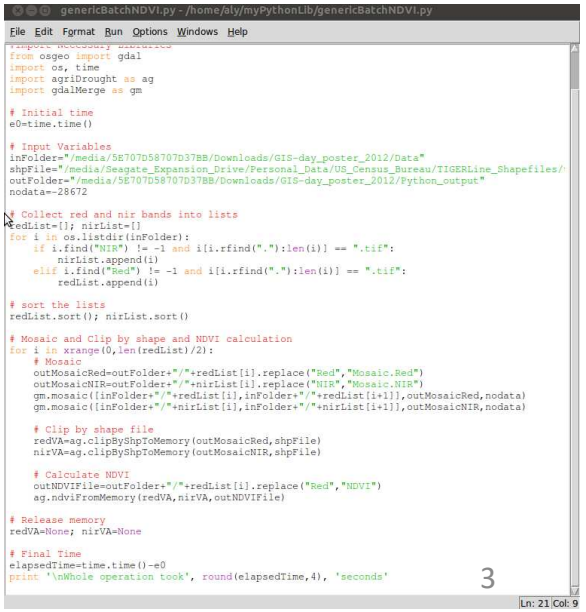
Simple Applications

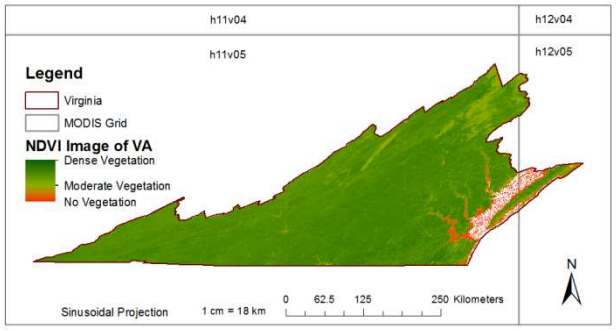
- Simple NDVI calculation for Virginia.
  - Two MODIS sensor tiles
  - Calculate NDVI,
  - Mosaic two tiles,
  - Clip the results by VA
- Per run
  - 61 seconds (ArcGIS per run)
  - 12 second with Python
- Why Python?
  - No Intermediate file clutter
  - Passing results as an array to functions
  - Fast
  - Customizable
  - Easily automated for time series analysis

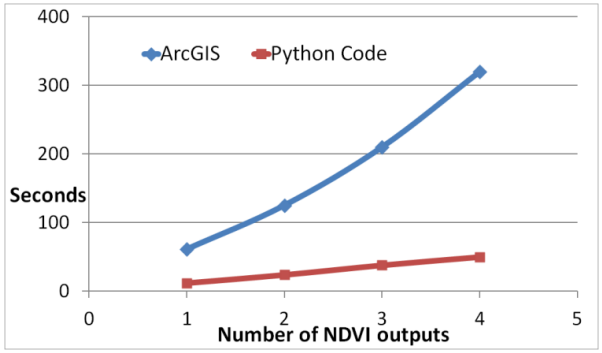
ArcGIS model builder



Python Script







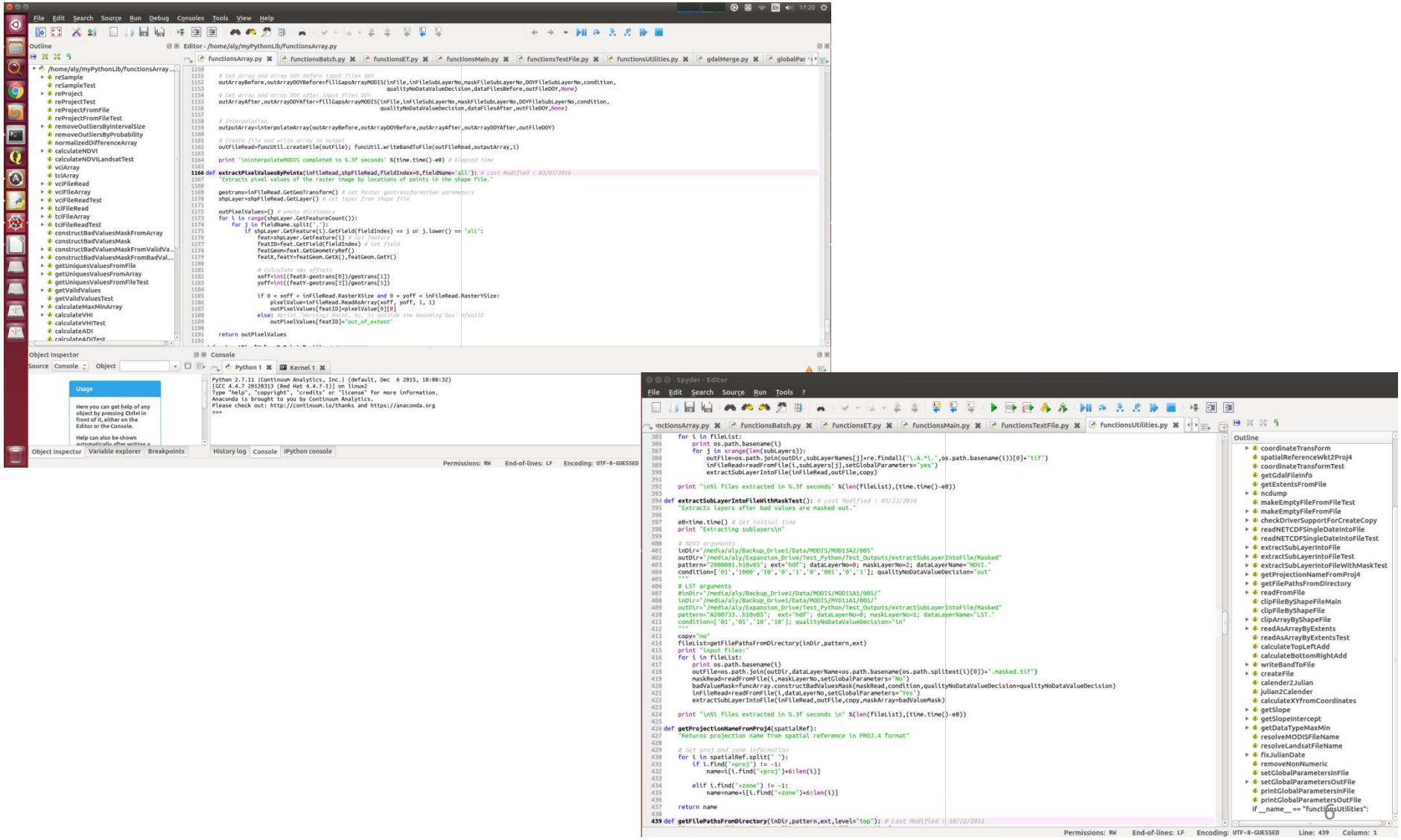
Applications

- Drought Monitoring
  - Based on the Earth observing satellite products
  - Vegetation Indices
    - Normalized Vegetation Difference Index (NDVI)
    - Enhanced Vegetation Index (EVI)
  - Land Surface Temperature (LST)
- Evapotranspiration Modeling
  - NDVI, LST and maximum/minimum air temperature
- Time series extraction from satellite data for validation and analysis.

How I do it?

- Ubuntu operating system
  - 14.04.4 LTS
- Libraries and their dependencies handled by Anaconda.
- Integrated Development Environment
  - Spyder and IDLE
- Old way downloading libraries from Ubuntu GIS stable/unstable repository
  - https://launchpad.net/~ubuntugis/+archive/ubuntu/ppa
  - Not convenient, dependency problems.
  - All libraries not available.
  - Not latest release libraries.

Spyder IDE



Libraries

- Using existing libraries, rather than implementing from scratch.
  - Handling file input/output. Read files and passing it to libraries.
- GDAL/OGR
  - GDAL: Geospatial Data Abstraction Library
  - OGR: OpenGIS Simple Features Reference Implementation.
  - Reading geospatial data in raster and vector format.
  - inFile=gdal.Open(raster format like HDF, GeoTiff and NetCDF)
  - inFile=ogr.Open(raster format like HDF, GeoTiff and NetCDF)

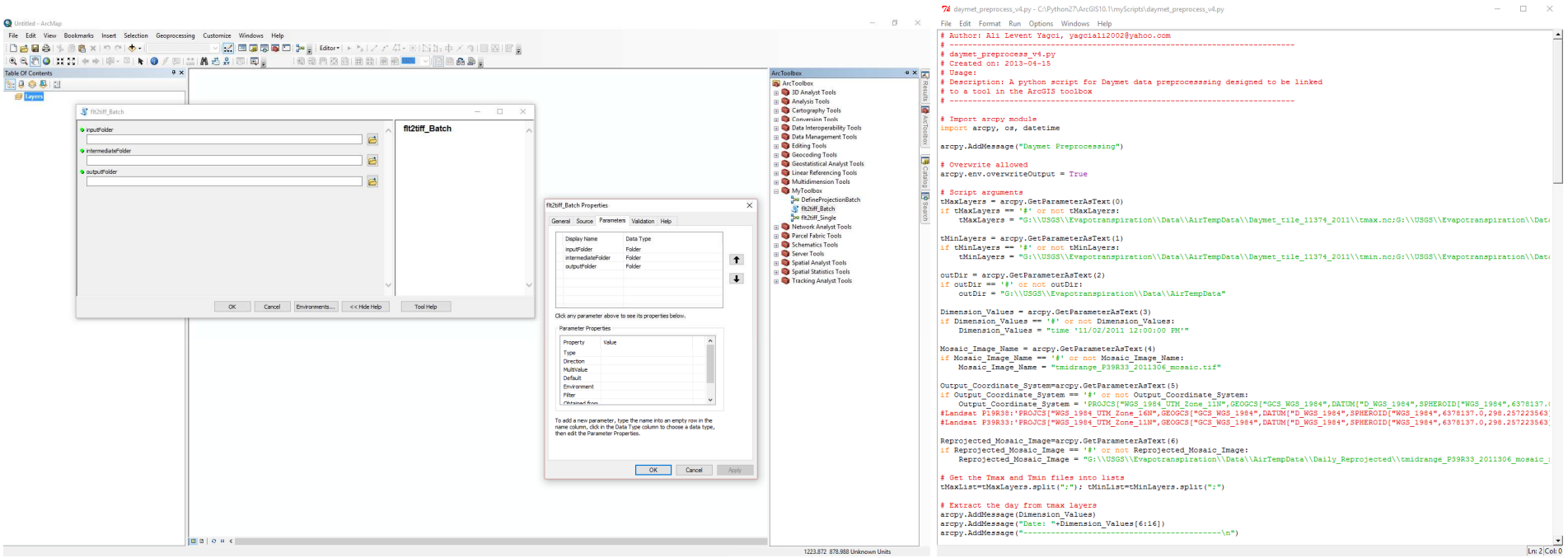
Libraries

- Scientific Libraries for Array computation
  - NumPy : Reading 2-D data into numpy arrays
  - SciPy: some utility functions for arrays
  - Matplotlib: plotting

matplotlib

Libraries from Commercial GIS Software

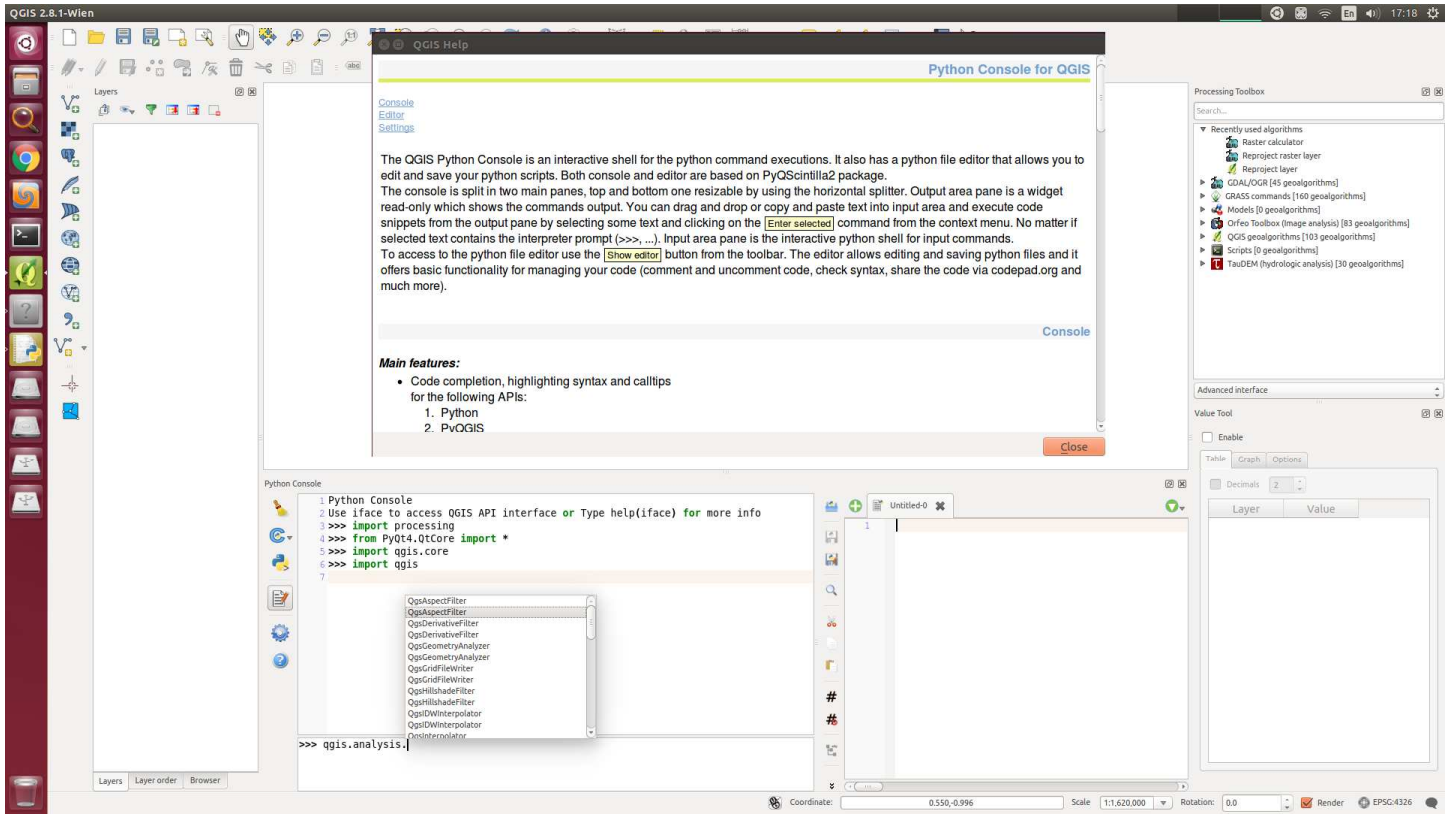
- ArcGIS
  - Built-in python console or,
  - Stand-alone python script
- ArcPy
  - Associate Python script to a tool and save it in toolbox.



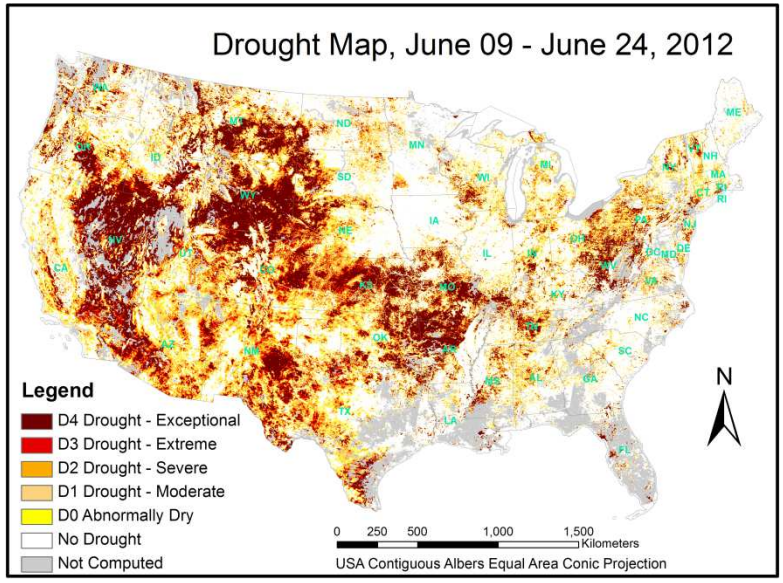
- inputFolder = arcpy.GetParameterAsText(0)

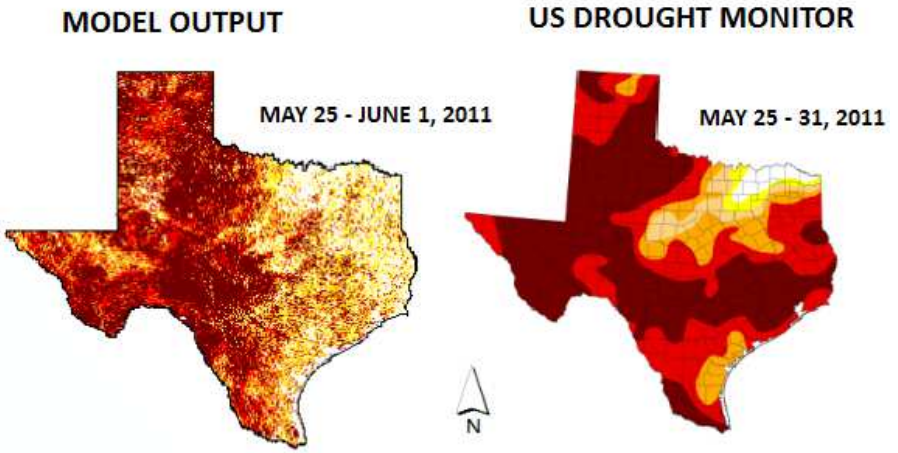
Libraries from Free GIS Software

- Quantum GIS (QGIS)
  - No license issues
- Built-in Python console and editor
- Stand-alone python scripts
- Associate Python script to a tool and save it in toolbox.

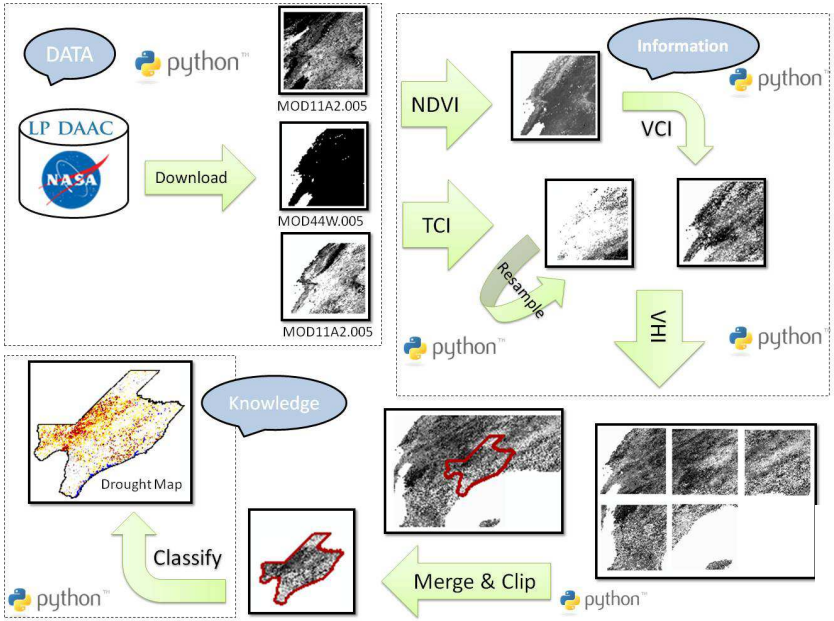


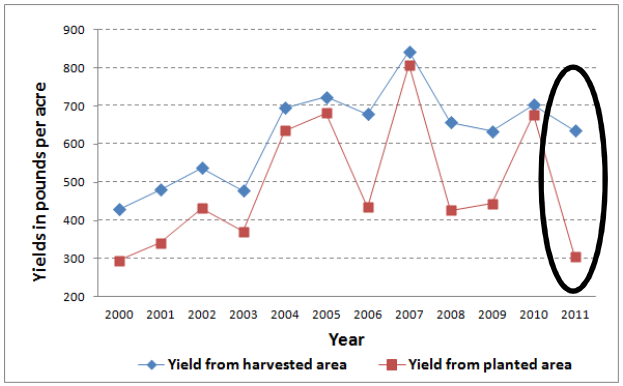
Drought Monitoring

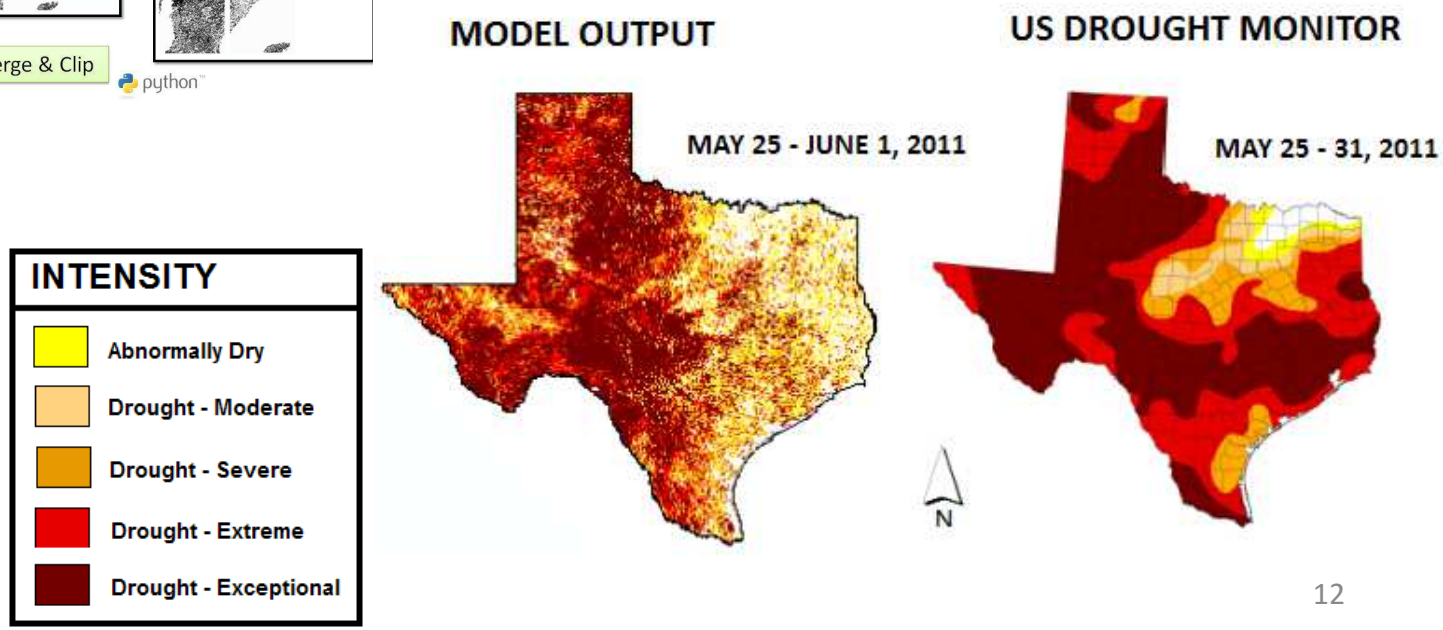




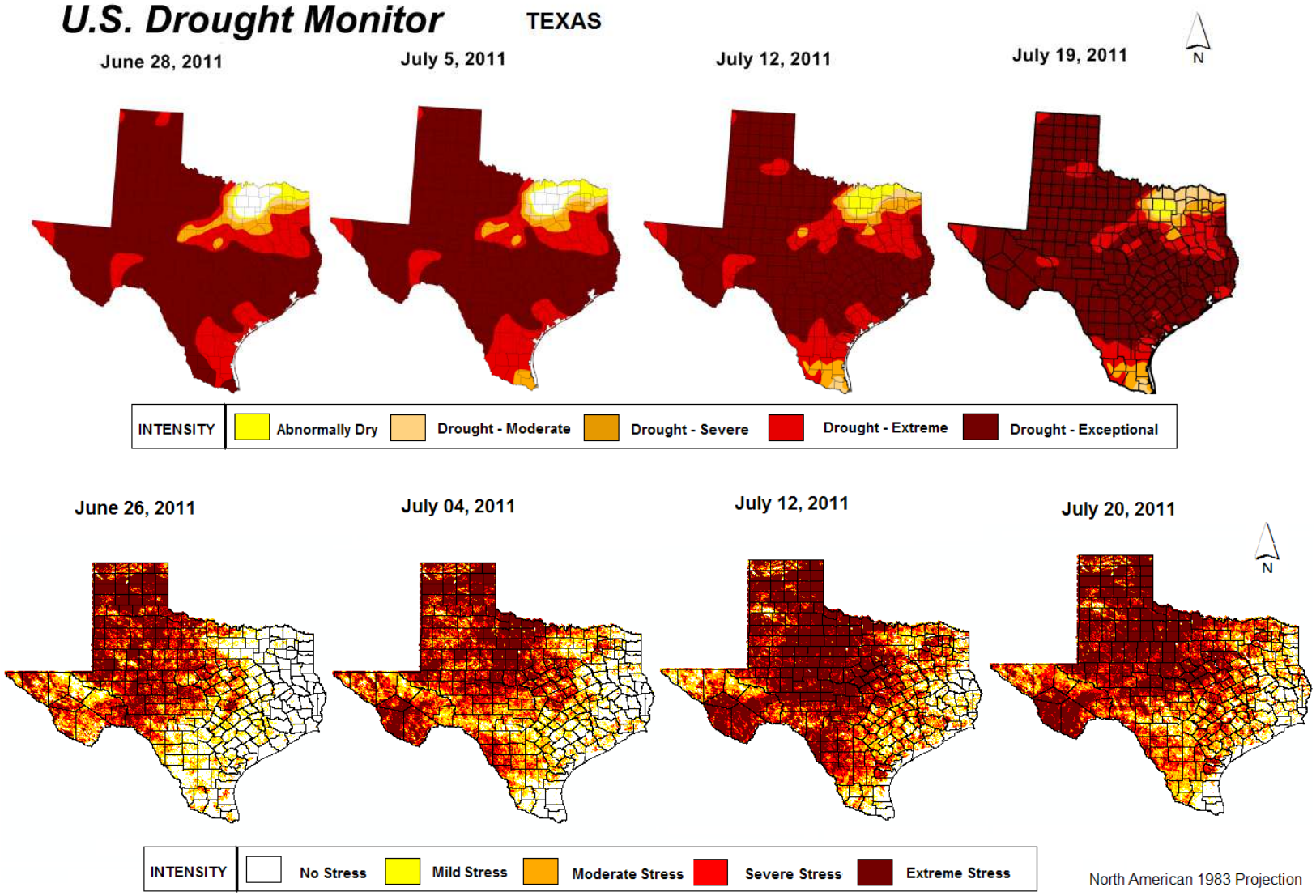
Texas Drought in 2011



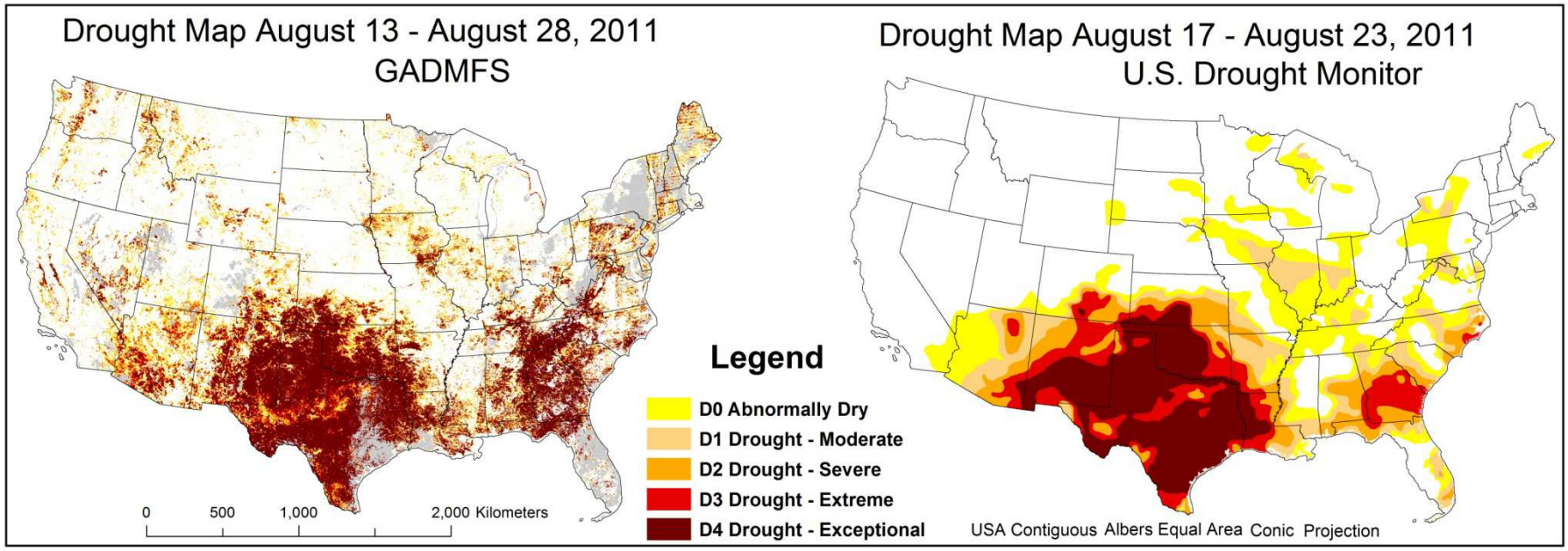




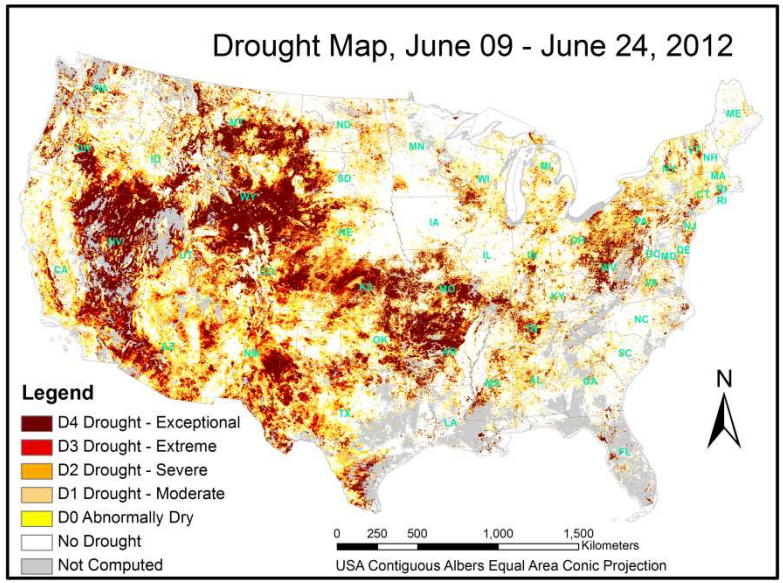
Texas Drought in 2011

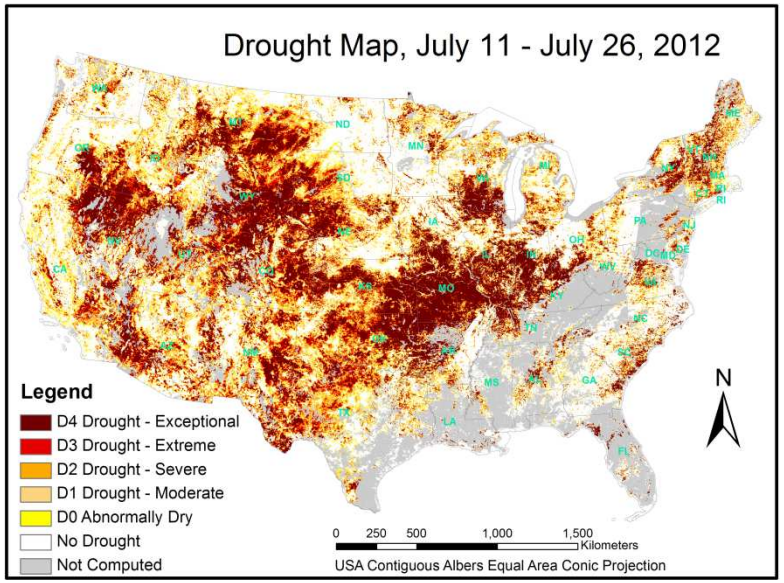


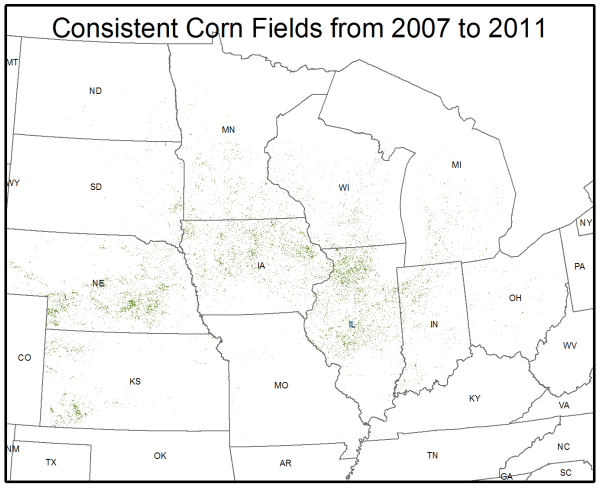
Drought in 2011

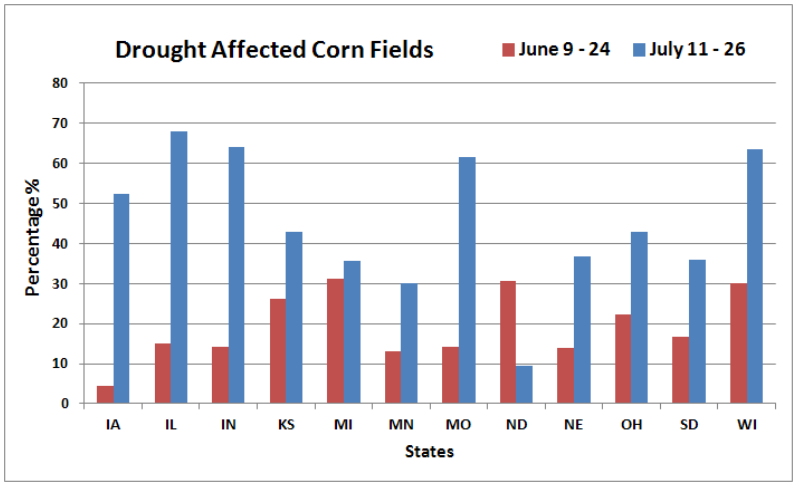


Drought-affected corn fields in 2012

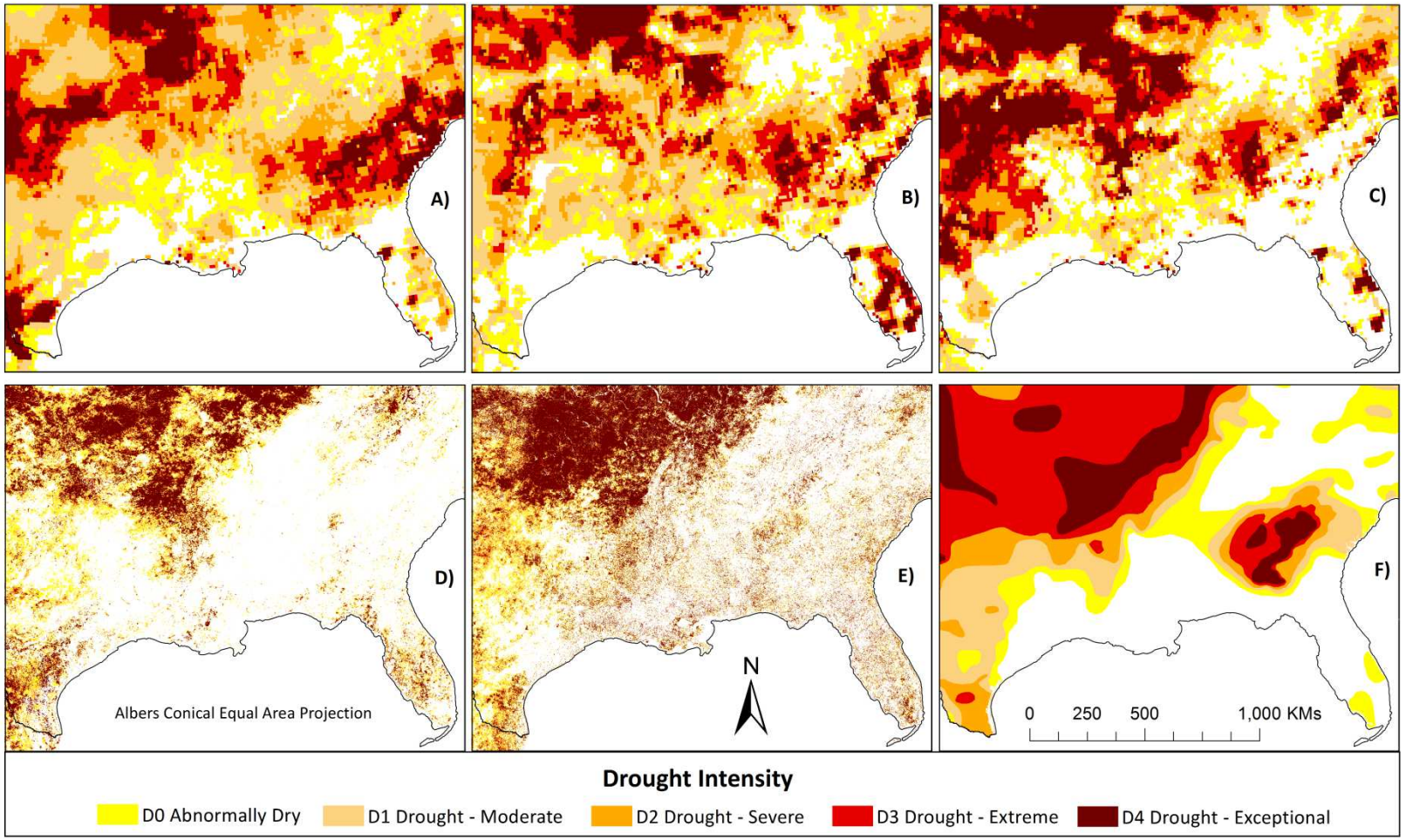






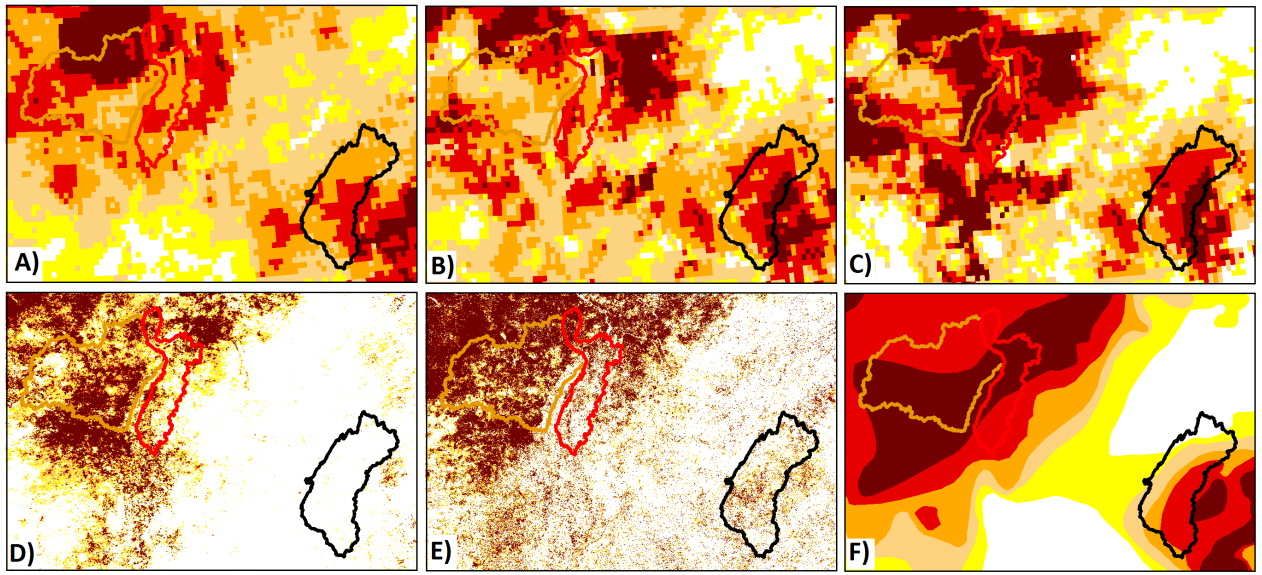


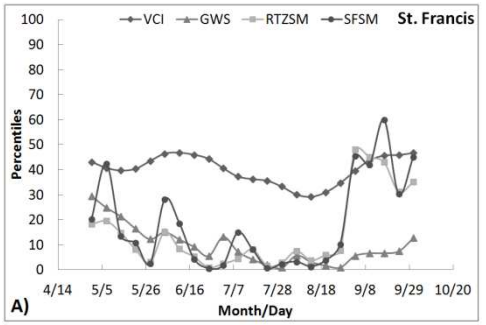
2012 Drought in Southeast US

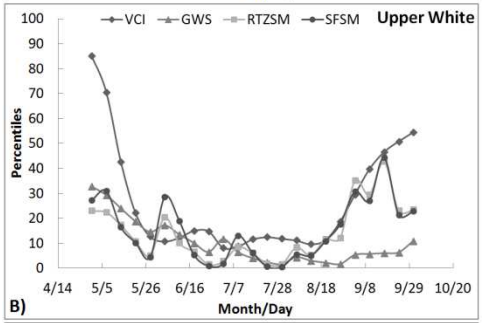


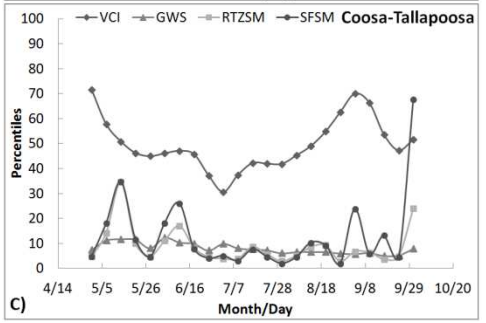
2012 Drought in Southeast US

- Creating time series of basin averages of drought indices by boundary files of Basins and writing them into text file.

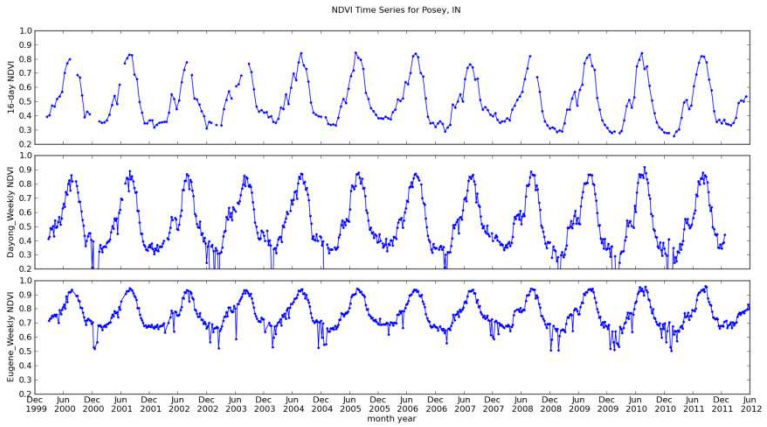


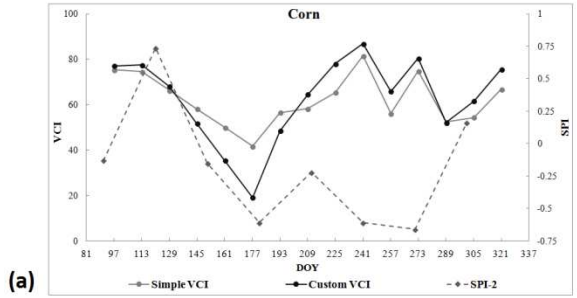




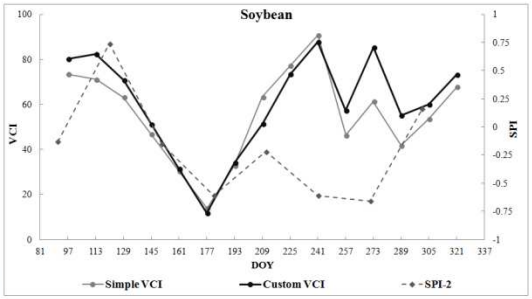


Time Series Generation





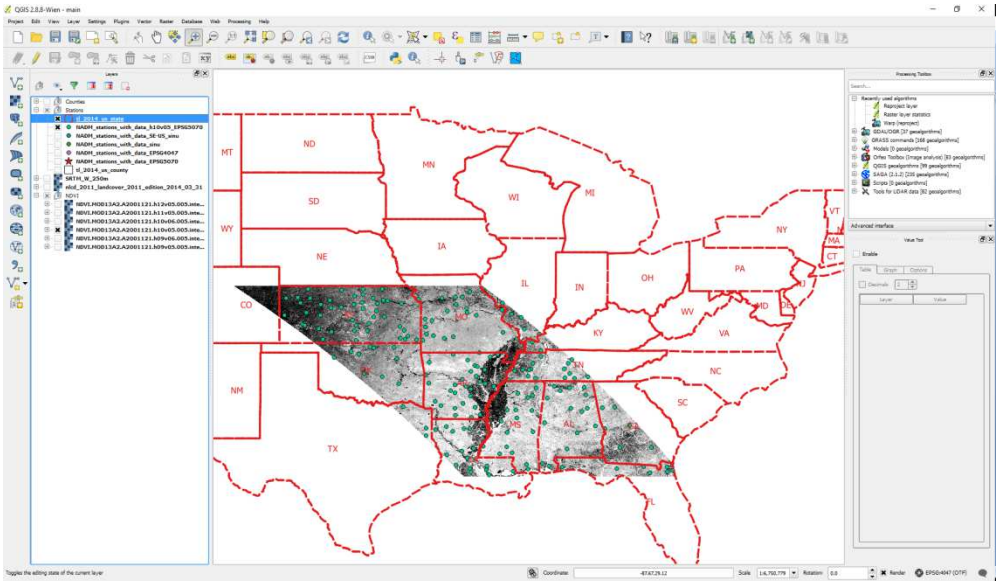
(a)

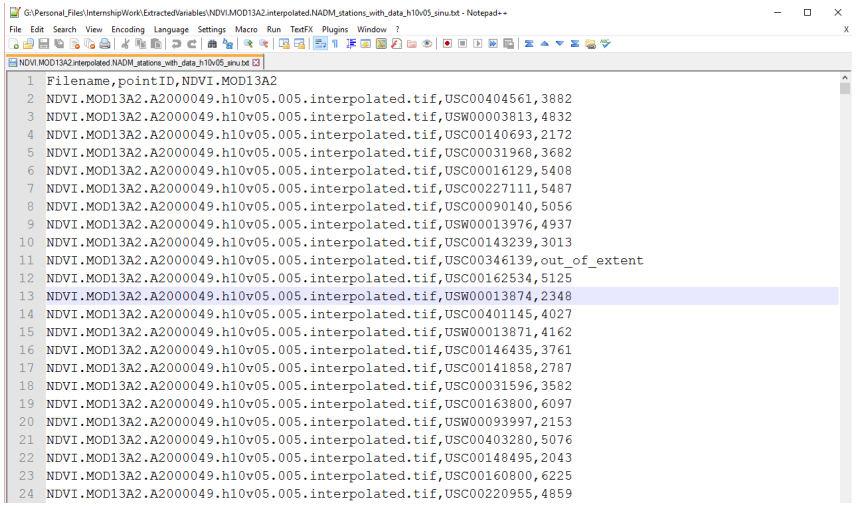


(b)

Time series extraction by points

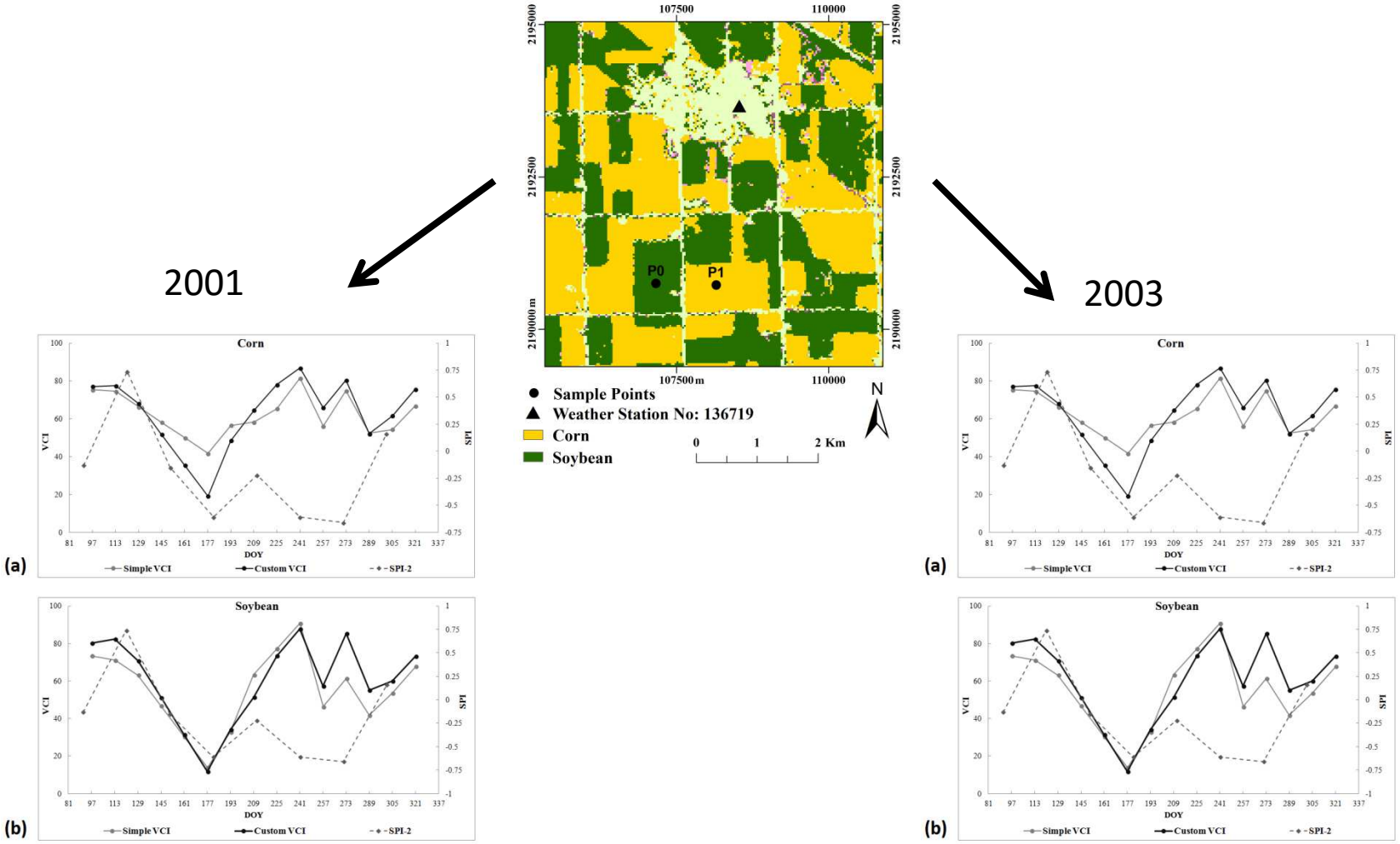
- 196 weather stations in a MODIS tile.
- Extracting overlapping NDVI and EVI pixels with stations between 2000 and 2016 (391 MODIS layers).





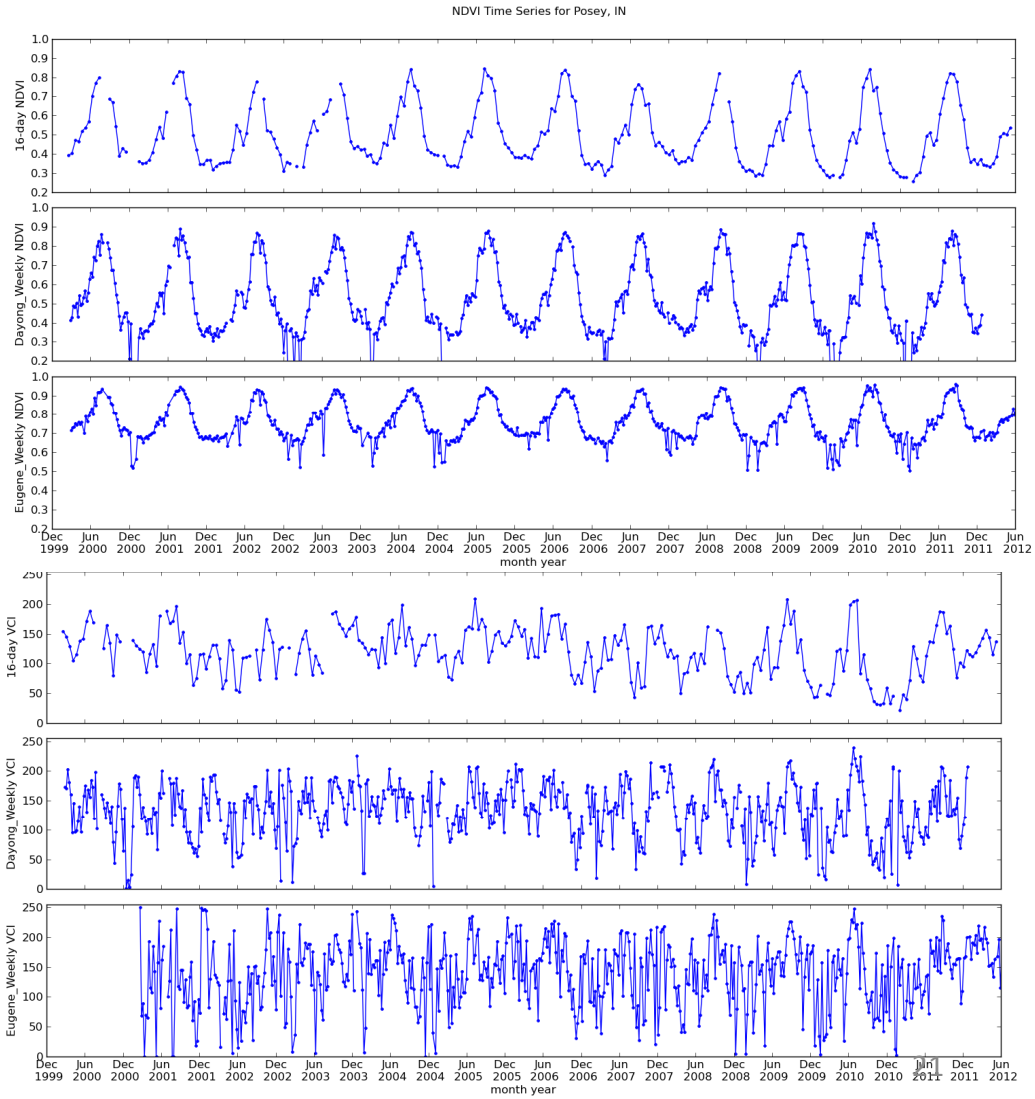
```
Filename,pointID,NDVI.MOD13A2
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00404561,3882
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USW00003813,4832
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00140693,2172
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00031968,3682
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00016129,5408
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00227111,5487
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00090140,5056
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USW00013976,4937
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00143239,3013
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00346139,out_of_extent
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00162534,5125
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USW00013874,2348
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00401145,4027
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USW00013871,4162
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00146435,3761
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00141858,2787
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00031596,3582
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00163800,6097
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USW00093997,2153
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00403280,5076
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00148495,2043
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00160800,6225
NDVI.MOD13A2.A2000049.h10v05.005.interpolated.tif,USC00220955,4859
```

Time series extraction by points



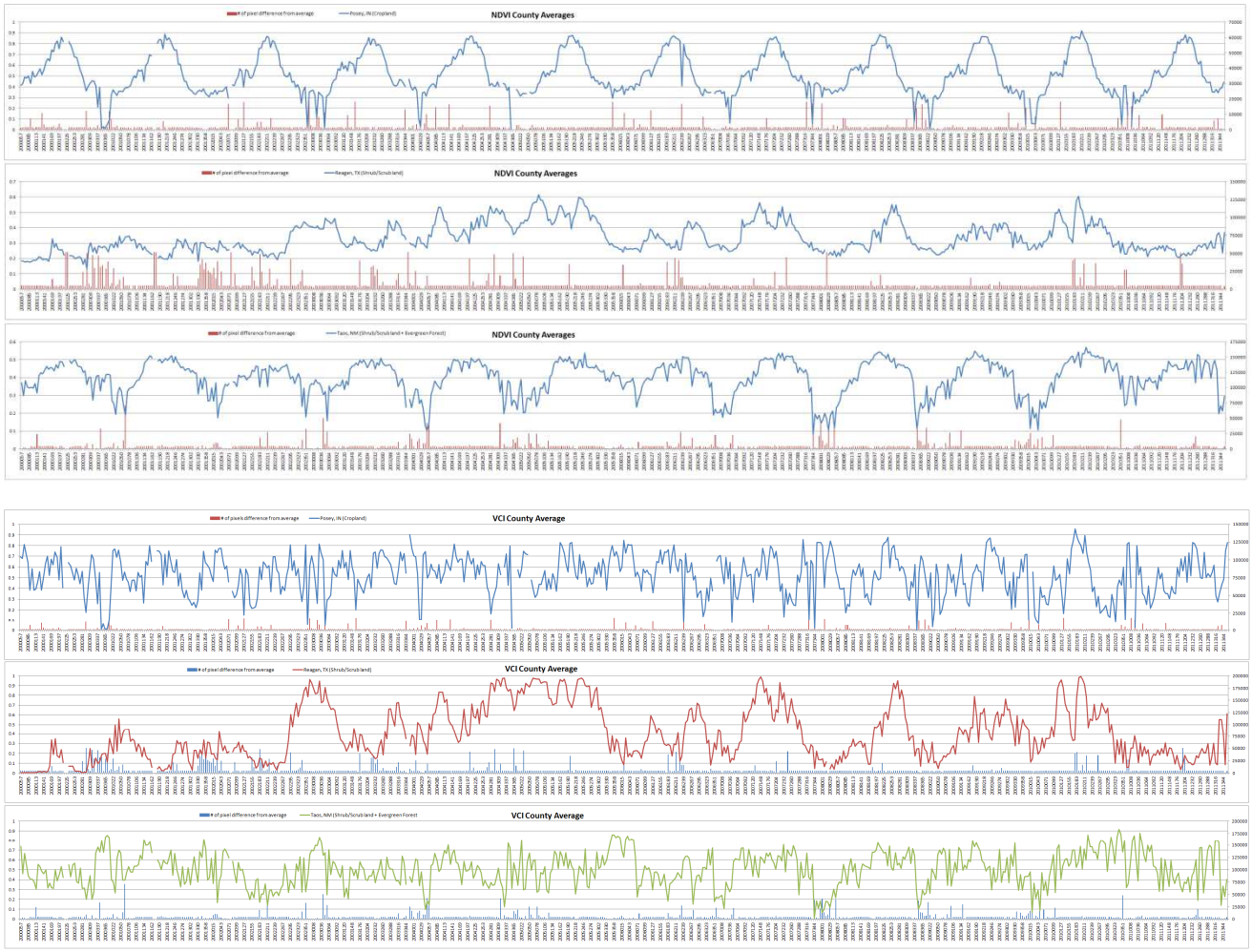
Time Series generation

- Generating Time series of NDVI and VCI by county boundaries.
- No intermediate files, everything done on memory
- Reusable functions
- Plotted in Matplotlib

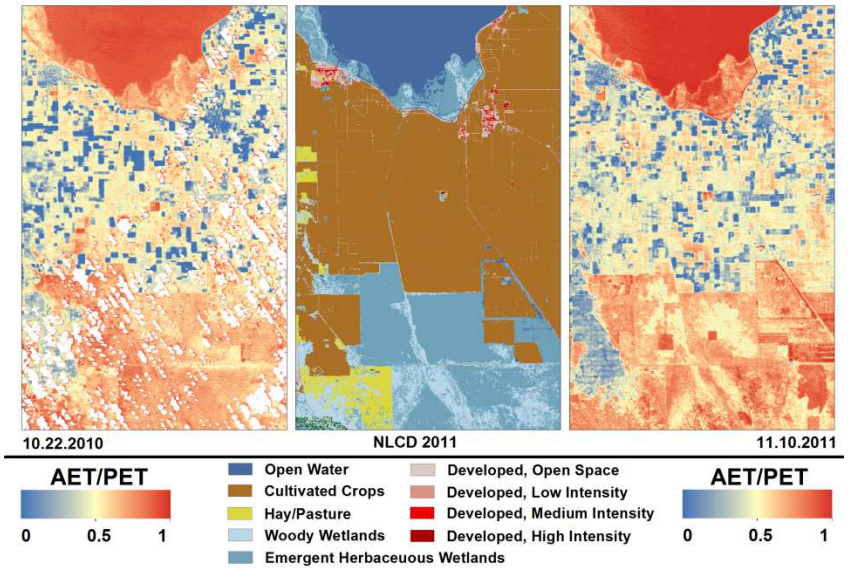


Time Series generation 2. way

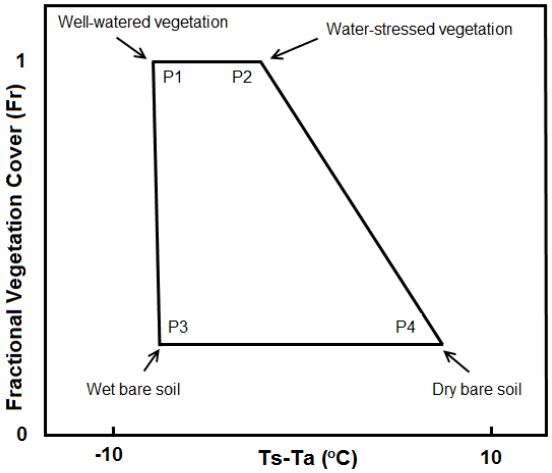
- Extract time series of indices with counts into text files.
- Import the text files into excel and plots them.
- More interactive



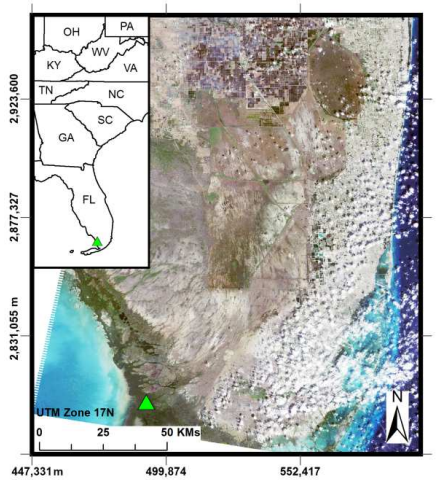
Evapotranspiration

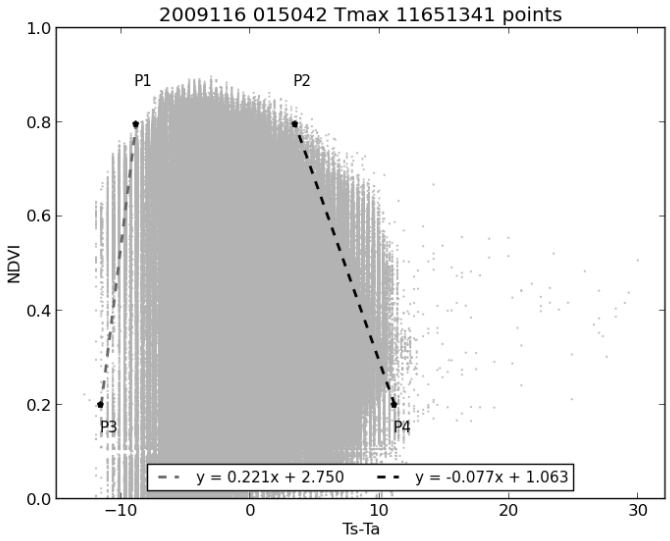


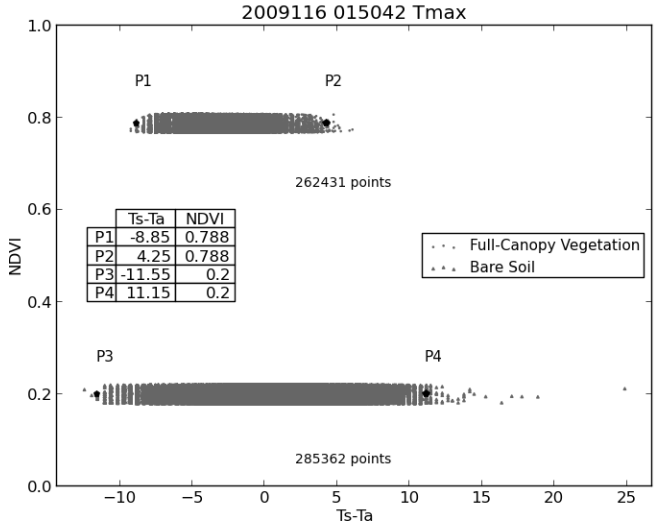
Extracting end-members from Landsat data



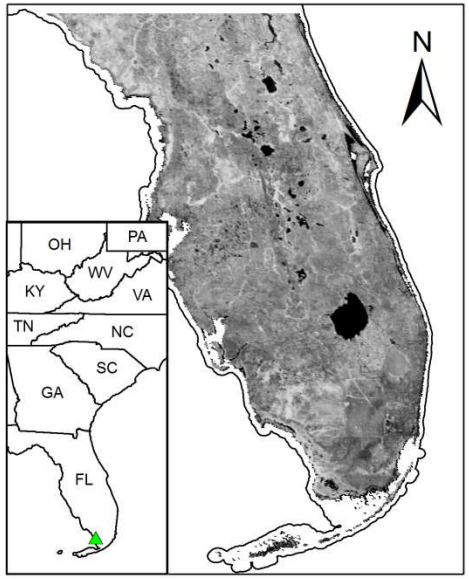
Flux tower located at Everglades National Park in Florida, US

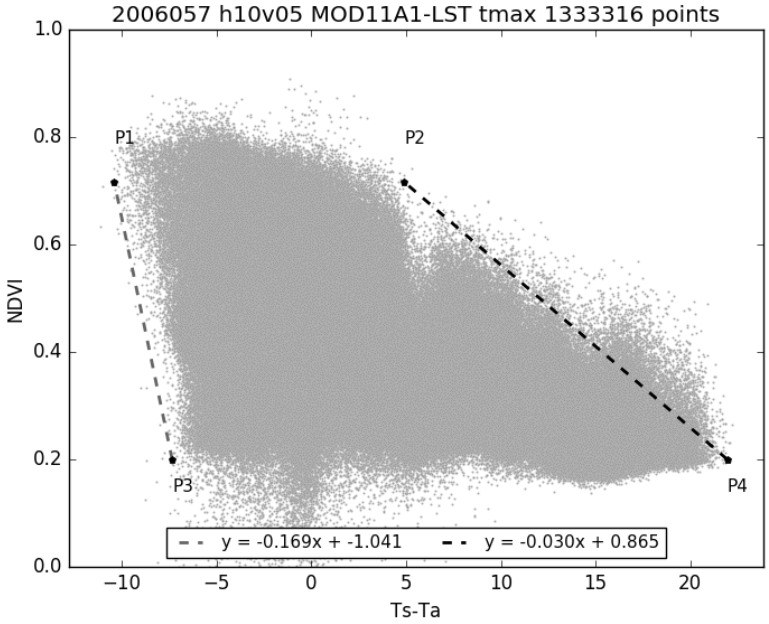


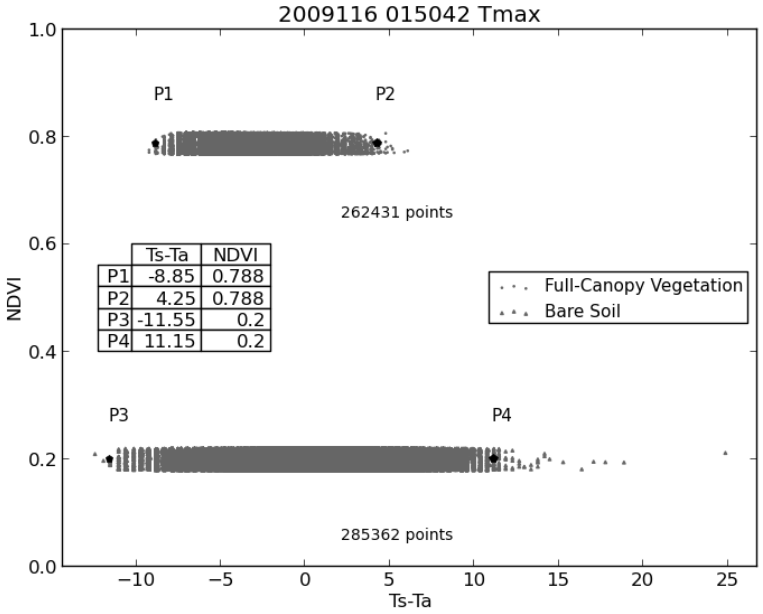




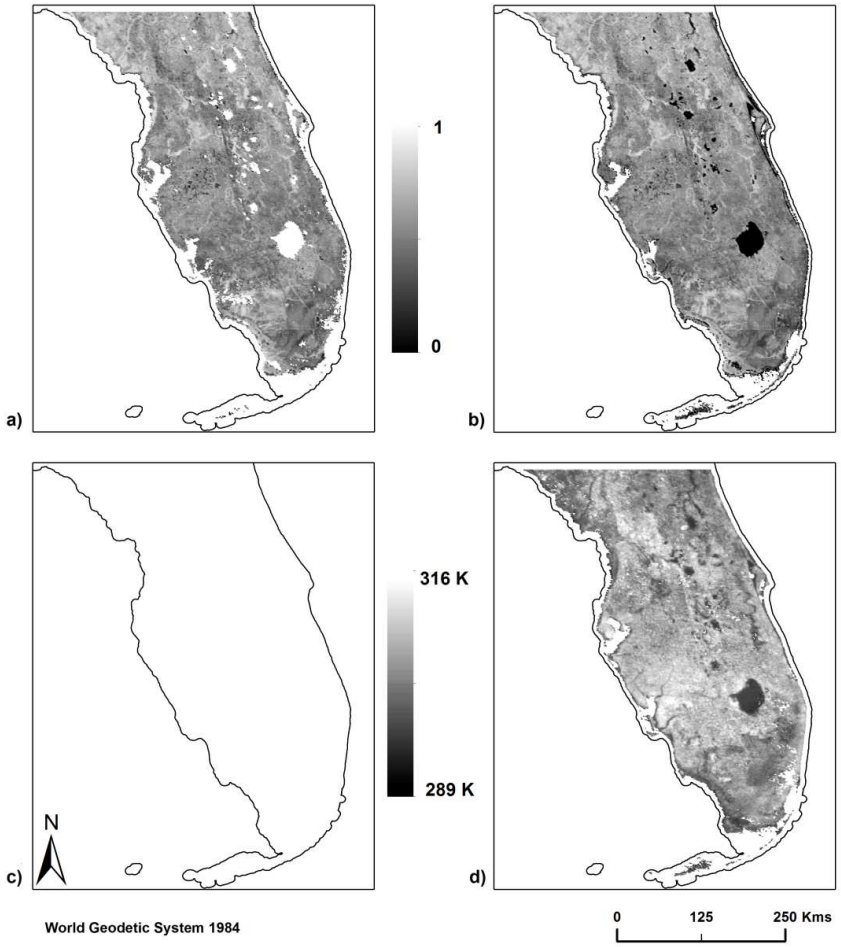
Extracting end-members from MODIS data





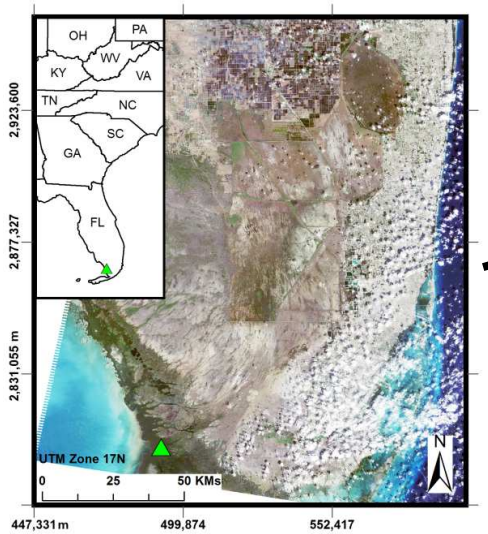


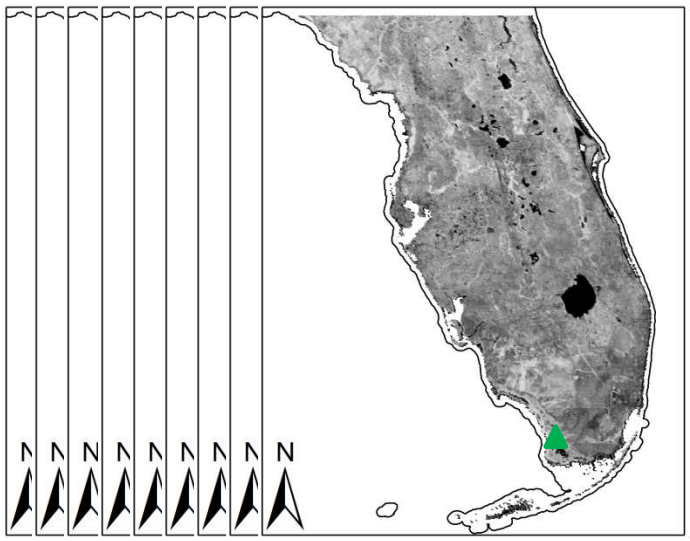
Interpolation

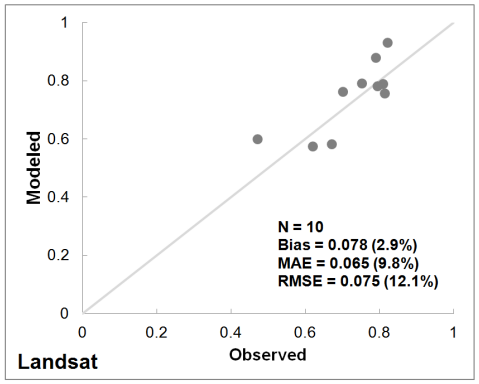


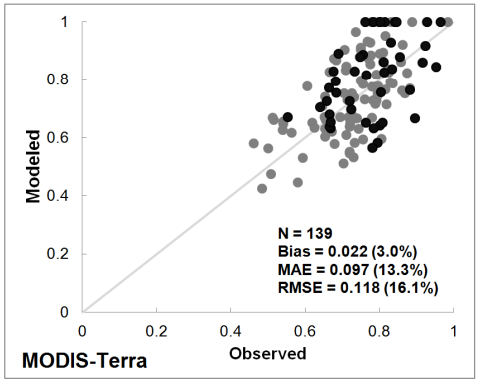
- NDVI gap-filling
  - 32 days before/after
  - Spatially auto-correlated
- LST gap-filling
  - 4 days before/after
- Analysis of Quality Bands
  - Cloudy pixels
  - Cloud adjacent pixels
  - Low-quality
  - High errors

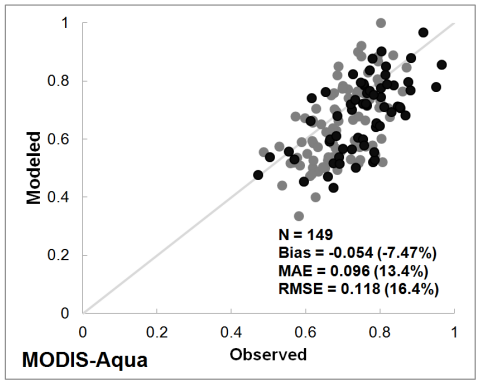
Validation of model inputs



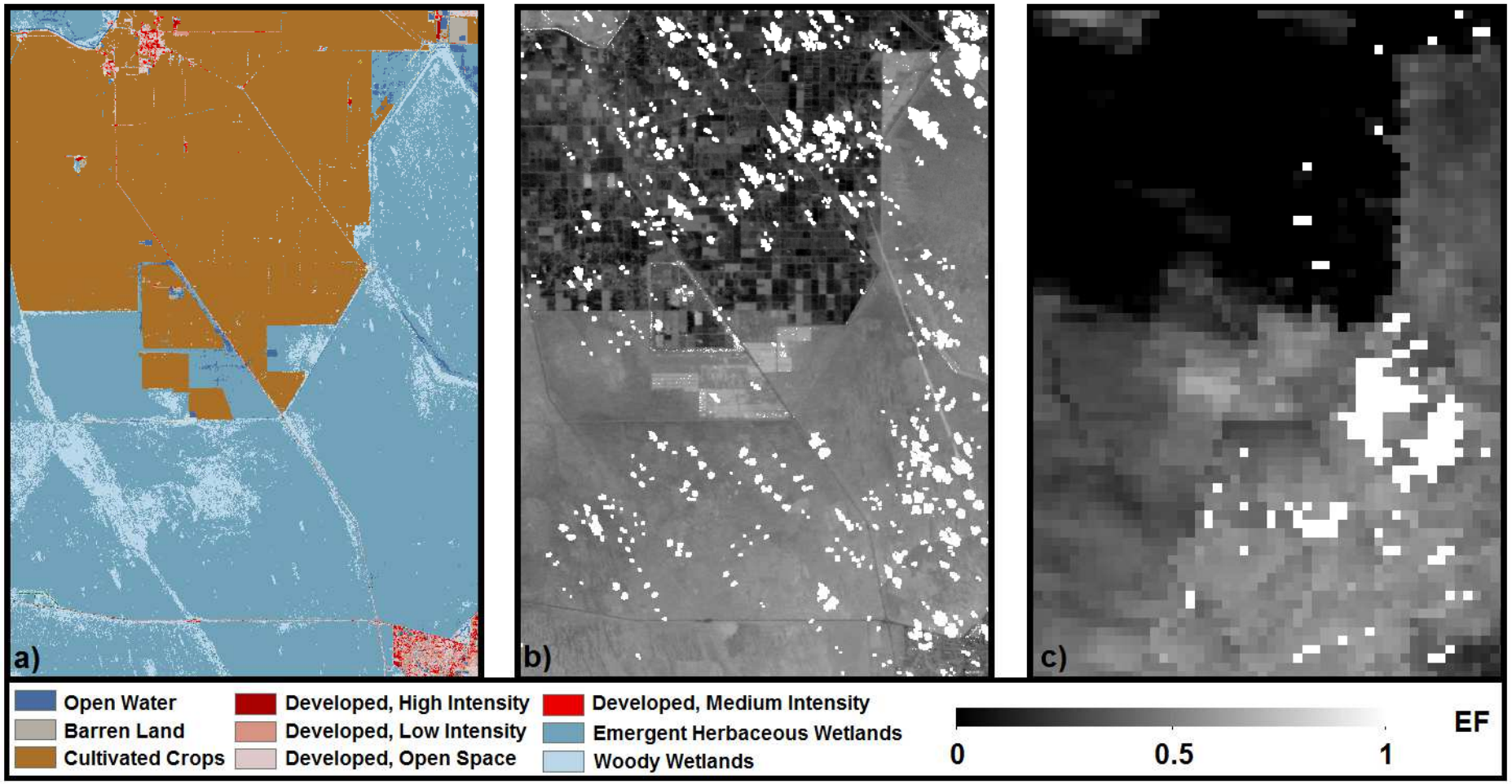




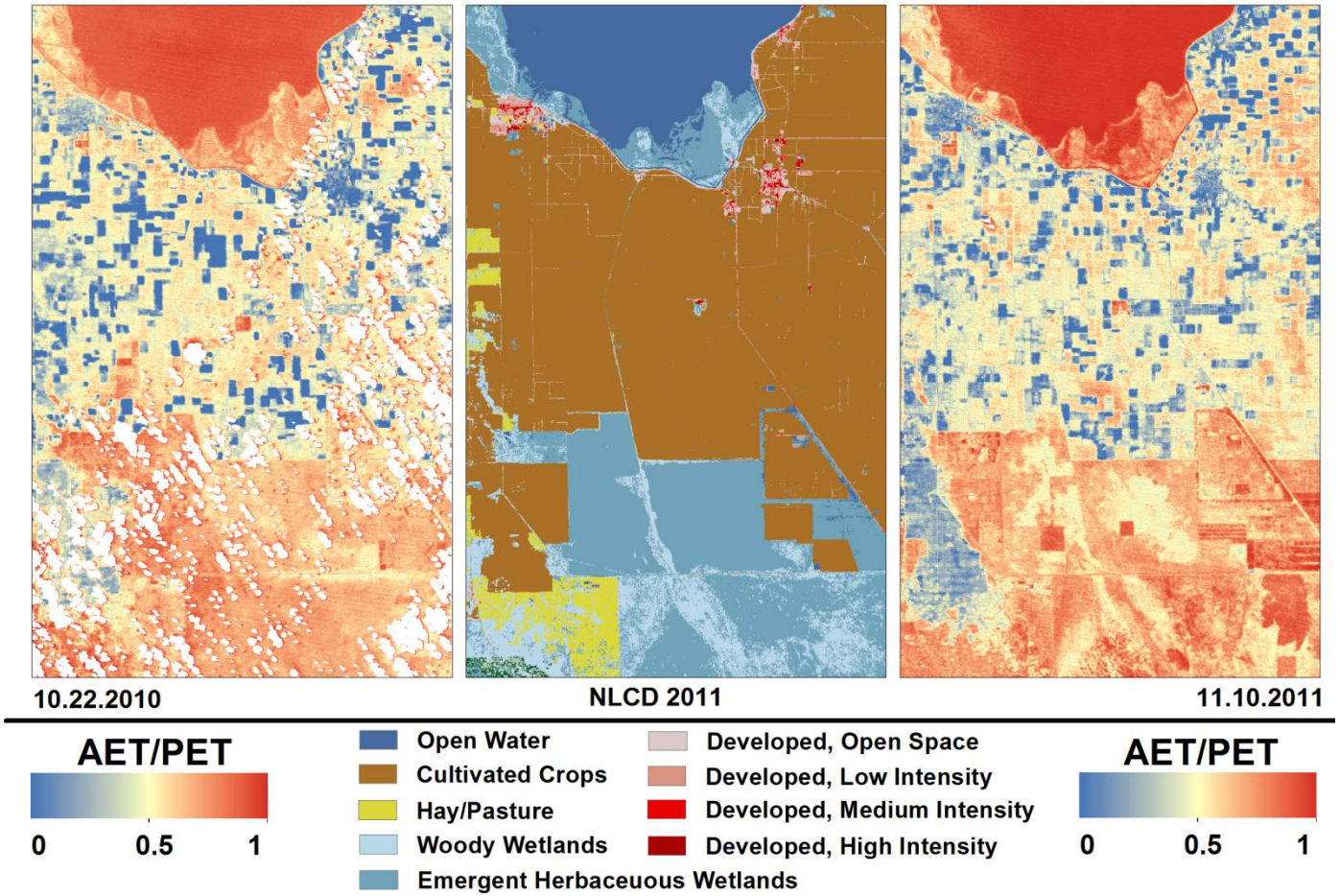




ET maps of Everglades National Park



ET maps of Florida



Important Points

- Reusable functions
  - NDVI functions takes arrays as inputs
  - NDVI for MODIS, NDVI for Landsat or any sensor product
    - Handle which bands you need to pull
  - Drought index for a compositing period 171st day of 2012.
- Setting unused variables (e.g., arrays) to none to release memory.
  - ndvi=None

Questions?