

Final Report: ColoradoView Project

Project Theme: Invasive species habitat suitability modeling for the Wheat Stem Sawfly in Colorado.

Project Team:

- Jordan Lestina, Forest and Rangeland Stewardship, jlestin@rams.colostate.edu
- Maxwell Cook, Forest and Rangeland Stewardship, mccook3@rams.colostate.edu

Faculty Advisor: Sunil Kumar, Natural Resource Ecology Laboratory (NREL), sunil.kumar@colostate.edu

Project Description:

The purpose of this project is to use remote sensing and GIS to create input layers for a species distribution model to determine if the model is improved with the use of satellite imagery. We are using MODIS 250 meter spatial resolution vegetation indices. The vegetation indices that we are using are Normalized Difference Vegetation Index (NDVI), which is a measure of greenness of vegetation, and Enhanced Vegetation Index (EVI) which is an improvement of NDVI that is more sensitive to densely vegetated areas as well as having better correction for atmospheric haze. This imagery will be stacked with monthly PRISM climate data that describes maximum/minimum temperature, precipitation, and mean temperature as well as layers for soils, hydrology, land use/land type, elevation, presence data for the Wheat stem sawfly, and possibly others. These data layers will be inputs to Maxent, a software based on the maximum-entropy approach for species habitat modeling that will produce a model of the potential range of the given species of interest.

Objectives/Research Question:

Wheat is an economically important crop in Colorado, with approximately 2.2 million acres planted across the state and a revenue of \$580 million based on 2012 harvest information. Wheat stem sawfly is an invasive species observed in eastern Colorado and in other wheat growing areas of the United States. This invasive can have serious impacts on wheat production and thus economic gains

from that production. This study will derive bioclimatic variables from PRISM Climate data that are significantly correlated with wheat stem sawfly infestations and, along with other layers, will determine potential future distribution of this species in Colorado. This project, while focused on remote sensing and GIS, also involves ecological theory and statistics and also could have implications for management of this invasive fly.

Progress:

Up to this point, Jordan and my-self have acquired most of the relevant datasets needed for input into Maxent. My part of the project thus far has involved downloading PRISM monthly climate datasets and re-projecting/clipping them to the extent of Colorado. However, this climate data is at 4 km spatial resolution when we would prefer to start with a 1km dataset. After being in contact with the Colorado Climate Center here at CSU it seems that this dataset may be available but would come at a cost. The alternative to a 1km dataset is to resample the 4km datasets to the pixel size of the MODIS NDVI and EVI imagery. NDVI and EVI were collected from the USGS Modis Reprojection Tool (MRTWeb) which processes HDF files and outputs GeoTiff, mosaicked MODIS scenes. In addition to the extracted NDVI and EVI layers, I have the MODIS raw data in .hdf format. To process the MODIS and PRISM datasets I built models in ModelBuilder which re-projected and extracted data to the Colorado boundary. Using the Cell Statistics tool in ArcMap, I ran statistics on the NDVI and EVI to get 3-year monthly means and standard deviations. It has been a good experience working with MODIS data as my prior remote sensing experience had been with Landsat imagery. Jordan and I are almost ready to compile our datasets and begin running our model in Maxent. We will run many iterations of the model until we figure out the right methods and parameters. Thus far I have received great support from my teammate Jordan, the Geospatial Centroid, and from Sunil Kumar. I believe that Jordan and I are on track at this point and with a final push to get the climate data prepared, we should be running the model at the beginning of next week.

Final Report: Lessons Learned, Methods, Discussion and Issues

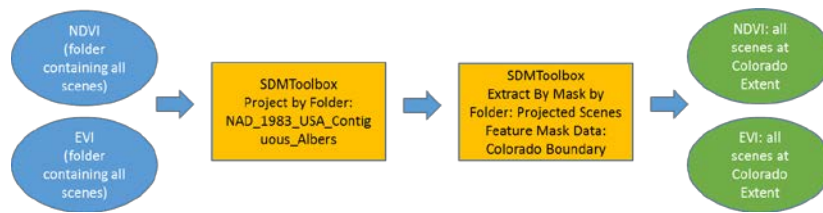
Lessons Learned

Preparing NDVI/EVI for use in MaxEnt

Part of this project is to determine the effectiveness of remote sensing in ecological niche modeling. We incorporated the MODIS products NDVI and EVI (Normalized Difference Vegetation Index and Enhanced Vegetation Index respectively) to help us determine suitable habitat for the Wheat Stem Sawfly (WSS). MODIS products such as these are available for download via USGS MRTWeb Application. There are a couple other places to get this data but we found that MRTWeb was the most efficient and time effective method to download the data. It is a standalone web application that allows you to select a MODIS time series and download your selected products (there is a whole suite of products to choose from in the MODIS image). We used MODIS NDVI and EVI for the months of April, May, June, and July which included two scenes per month. The downloaded file is in ESRI GRID format, making easy to bring in to ArcMap and analyze. In addition, Colorado is the intersection for 4 MODIS flight paths so in order to get a complete scene we needed to stitch together the 4 tiles before extracting the Colorado boundary. MRTWeb automatically stitches together the scenes that fall within your area of interest. Once downloaded, the scenes were brought into ArcMap to be re-projected and clipped to Colorado extent. Here I used the SDM (Species Distribution Model) Toolbox to re-project and extract by mask to get coverage for our study area. SDMToolbox is designed to facilitate many complicated processing steps commonly required for species distribution modeling and other analysis. The SDM toolbox allows for these tools to be run on entire folders, eliminating the hassle of doing them individually, running the tools in batch mode, or through a python script. After we re-projected and clipped to Colorado extent, we ran cell statistics on the images to get the mean and standard deviation of all the data by month. The workflow for that procedure is below. The major issues I encountered during this process were finding

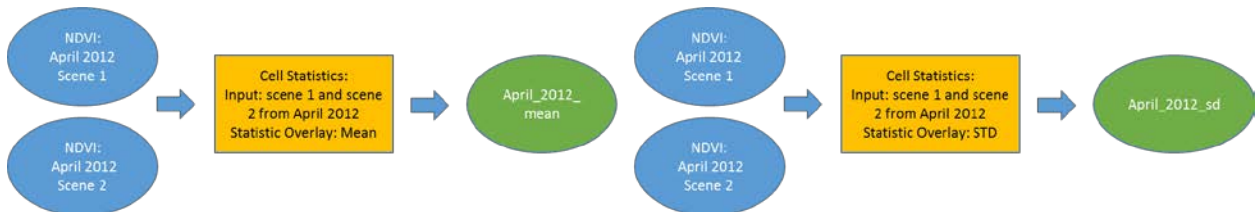
the correct way to download and extract MODIS NDVI and EVI. At first, I used the USGS Global Visualization Viewer (GloVis) which allowed me to download all my required scenes in .hdf format. These .hdf files contained spatial subsets of MODIS products (9 of them) and were in the conventional MODIS projection. I ran into problems trying to extract the correct spatial subsets that I needed as well as with the processing time it would have taken to do so. The use of ENVI and certain add-ons that are meant to handle .hdf files could have been the answer. However, I found MRTWeb to be more time efficient and straightforward.

Workflow Diagrams for MODIS:

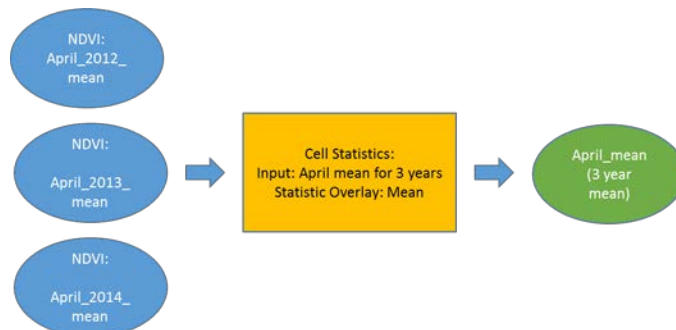


a. Preparing NDVI/EVI to run statistics

b. Running statistics for NDVI/EVI (example of just one year out of the 3)



*same process for all months for all years and then...



*Final Products: NDVI 3-year mean for April, May June, July; NDVI 3-year standard deviation for all months; EVI 3-year mean and standard deviation for all months.

All Final rasters in GRID format were then converted to ASCII using the Raster to Ascii (Conversion) Tool in ArcMap 10.2.2

Preparing PRISM data for use in MaxEnt:

The PRISM Climate Group gathers climate data from a wide range of monitoring networks and develops spatial climate datasets that reflect current and future climate patterns. The Climate Group offers historical data in 30-yr Normals, current data (monthly or annually), and others. The data consists of precipitation, maximum temperature, minimum temperature, and mean temperature. When I first began to look into PRISM data, we wanted layers for the years of 2011-2014, which were available in the “Recent Years” section of the PRISM download portal at a resolution of 4 km. after attempting to resample the climate data, we realized that it was not efficient to downscale from 4km to 250 meter resolution. Because of this issue we decided to use the 30-year Normals, available at 1 km resolution. This data covered the time period of 1981-2010 and represented the average monthly and annual conditions over the most recent three full decades and covered the extent of the United States. We reprojected this data into our Nad83 Contiguous USA Albers projection and extracted the data to the Colorado extent. These layers were then used to generate 19 bio-climatic variables using an AML script produced by WorldClim, a climate group that gathers global data from various monitoring networks. The bioclimatic variables represent annual trends, seasonality, and extreme or limiting environmental factors which offers more biologically meaningful variables for ecological niche modeling.

Downloading and Preparing a 90m DEM for Colorado using SRTM data:

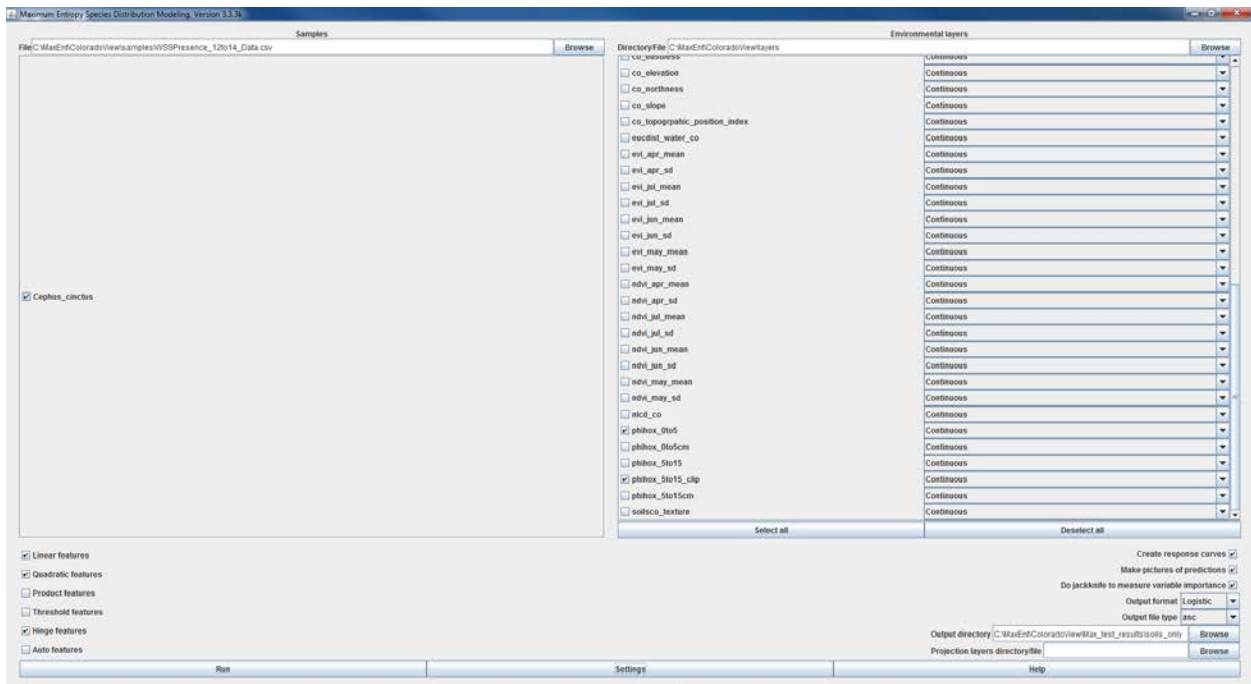
Shuttle Radar Topography Mission data was used for elevation and other DEM related layers including aspect, slope, and eastness/northness. The original SRTM data was downloaded from the CGIAR Consortium for Spatial Information download portal webpage (URL = <http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp>). We downloaded the data in 4 tiles to get full coverage for Colorado. These tiles were then mosaicked to a new raster layer in ArcMap, reprojected to Nad83 Contiguous USA Albers, and extracted to the Colorado extent. Once at the Colorado Extent, we resized the layer using the Aggregate tool in ArcMap from a 90 meter resolution to 250 meter resolution to match the rest of our data. Next, aspect was derived from the DEM using the Aspect tool in ArcMap. Eastness and Northness (a continuous measure of how “east” a pixel is and how “north” a pixel is) was derived using Raster Calculator on the aspect layer; Eastness = $\text{Sin}(\text{Aspect Layer (in radians)})$, Northness = $\text{Cos}(\text{Aspect Layer})$. Slope was derived from the DEM using the Slope tool in ArcMap. A Solar Radiation layer was generated as well using the Area Solar Radiation tool in ArcMap. However, due to the lack in variation of incoming solar radiation in the eastern plains of Colorado, we did not use the layer in our final model.

Using MaxEnt:

After all of our data layers were prepared (projected the same, with the same pixel dimensions, and in ASCII format) we began the process of producing our species distribution model. We

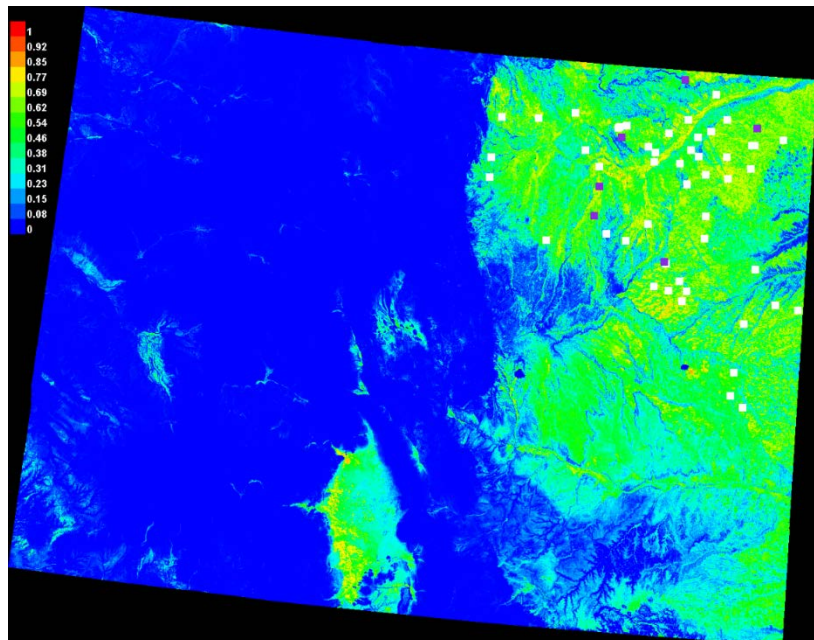
used the Maximum Entropy (MaxEnt) method which assumes that species spread out to fill areas with suitable conditions in the absence of external influences. It is a general technique for estimating a probability distribution from incomplete information and uses presence-only data (presence of the species in a given area). The goal for producing a satisfactory model was to eliminate the variables that did not contribute to the distribution of our species or were highly correlated with other variables being used. MaxEnt has a very user friendly graphic user interface, making it relatively intuitive to run models (see below). The hard task is in determining which variables to use and which to drop based on the

statistics that the program runs on the data. In addition, computer modeling may find correlations between data but it does not necessarily interpret them as important for the species from a biological



perspective. Thus, as the modelers, we had to distinguish between what was statistically significant and what was biologically relevant. We started the process with 47 data layers and, after 8 iterations of the model, narrowed it down to 5 variables that make the most sense statistically and biologically. Using

these 5 environmental variable, MaxEnt produced a heat map for potential distribution of the species that matched what you would expect given the nature of WSS. A Picture of the image is below. The probability scale is 0-1, with blue being zero and red being 1 (100% chance of occurrence).



The squares are the presence data we used for the WSS.

Reflection:

This project is a great opportunity to create a product that has implications for land managers in Colorado, especially those who are in the wheat production profession. The Wheat Stem Sawfly can pose a major threat to wheat

production when outbreaks are at epidemic levels. The photo to the right shows the effect of Sawfly infestation on wheat crops. The insect over winters in the stems of the wheat plants, causing the plants



to fall over at the base and die off. The diameter of the wheat stem is directly related to the suitability for WSS. The thicker the stem, the better the habitat for the fly. Inevitably, the modeling of the WSS is almost a prediction of where wheat can occur because it is the main host for this insect. One issue we faced with running the model is how to separate the modeling of wheat from the modeling of the fly itself. This involved choosing environmental variables that made sense for the biology of the fly more so than the biology and habitat for wheat. Our goal for the model was to determine the potential distribution of the fly and not of wheat crops (although they are so closely related). By using remote sensing products such as NDVI and EVI, we were able to get some information on how biomass or “greenness” affects the distribution of the fly. Because we wanted to avoid modeling wheat distribution, we had to be careful about which months to use for the NDVI and EVI. Based on the concept that thicker stems (especially early in the growing season) provide better habitat, we found that April mean EVI provided a decent indicator for suitable habitat. This tells us that the greener it is (the more biomass) in April at the start of the growing season, the more suitable the habitat for WSS is. This makes sense

intuitively because this means that there are thicker stems in the wheat crops which increase host suitability. Implications for land managers of this information could be that on especially productive years, there is a better chance for WSS outbreaks. I believe that this project was a good primer for more research on using NDVI and EVI to model invasive insect species and suitable habitat for outbreaks of those species. In addition, the use of the 19 bioclimatic variables proved to be very useful in modeling the insect's potential distribution. They give a better idea of the specific climate interactions that matter for suitable habitat of the WSS. The climate variables give a great indication of where the fly can occur across the state especially because the climate in Colorado varies greatly across the elevation gradient.