Burning Area Estimates on the Flint Hills using MODIS, NLCD, and R

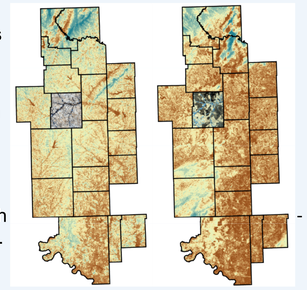
Andy Hawkins[1], Yao Tang[2], Douglas Watson[2]

[1]US EPA, Region 7, Lenexa, KS 66219; [2]Kansas Department of Health and Environment, Topeka KS 66612

Andy Hawkins | hawkins.andy@epa.gov | (913) 551-7179

Introduction

The Flint Hills region in Kansas and Oklahoma is an economically and ecologically important area encompassing the largest (~12,000 square miles) remaining tallgrass prairie ecosystem in North America. A portion of this area is frequently burned in the spring for management purposes. This burning can contribute to air quality problems in the form of both $PM_{2.5}$ and ozone. Estimating emissions from this burning first requires accurate estimates of the acreage burned, while forecasting potential impacts requires knowing how much is left to potentially burn. This poster looks at calculating burn estimates using MODIS satellite data with a high resolution land use mask (30m NLCD) using the freely available R statistical package to do the analysis. Comparison of the R burn area estimate is compared with other estimating methods currently being used for determining burning in the Flint Hills.



Flint Hills Study Area Chase County, KS highlighted, Left—pre burn, Right—post burn

Approach

Currently KDHE does burn scar analysis using a supervised classification process using commercial software. EPA Region 7 wanted to investigate the use of using the R statistical computing package which has some potential advantages over the KDHE approach such as being freely available and highly scriptable. Similar to the KDHE approach, EPA acquired the MODIS MOD09GQ and MYD09GQ datasets which consists of the Daily Surface Reflectance, Bands 1 and 2 from Terra and Aqua. These data sets are level 2 data providing an estimate of the surface spectral reflectance as it would be measured at ground level in the absence of atmospheric scattering or absorption. The data are provided at an approximate 250 meter resolution in a binary hdf format. Each MODIS dataset is then processed into a raster format (geotiff) using the MODIS Reprojection Tool.

The R software has libraries for both statistical computing and graphics. Several libraries were used in this analysis, including raster, rgdal, maps, maptools, and cluster. For this project the main libraries used were raster and rgdal. Raster allows for the creation and manipulation of raster objects, including masking, cropping, resamping, etc. Rgdal allows for reprojection of spatial objects. In addition to the add in libraries the base R software package has statistical functions such as “kmeans” which performs k-means clustering on a data matrix. One advantage of R is the ability to script the code for a repetitive/repeatable process. In this case the study area contains 20 counties, and multiple dates of MODIS imagery that potentially requires processing for multiple years. The goal was to investigate R as a viable alternative solution to the current process and to develop code to assist KDHE in future burn scar analysis of prescribed burning in the Flint Hills.

Processing of the MODIS data was done in a stepwise process with the main steps summarized below:

Step 1—Load and process MODIS image, crop and mask to each of the 20 counties of interest.

Step 2—Evaluate cloud cover and discard those images with cloud cover or cloud shadows.

Step 3—For the remaining cloudless days during the period evaluate burn scars, estimate the burn area.

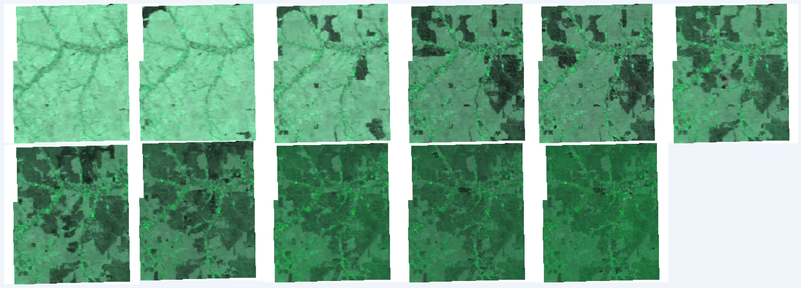
Step 4—Mask areas that likely did not burn (e.g. roads, cultivated land, etc.)

Step 5—Evaluate area burned estimate from R approach with KDHE estimate both in magnitude

Results

Step 1 and 2: Process MODIS images, crop and mask, select cloud free days.

In these two steps R code was written to read each MODIS geotiff and crop and mask to each county keeping only the clearest days for further analysis. It was observed that in general an image with clouds contains higher reflectance values, and a larger standard deviation of reflectance values over a county so those days were excluded from further analysis in R using a very simple algorithm. Additional cloud analysis using a better algorithm, including cloud shadow masking, could be performed but the authors instead focused on trying to develop the burn scar analysis code. At some point further cloud analysis may be warranted as many days in the spring period are partially cloudy but these additional days may provide additional temporal resolution of when burns occur.

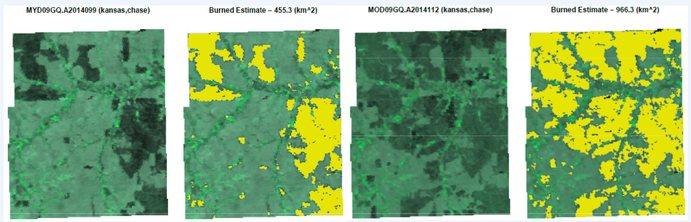


Example RGB images using MODIS Surface Reflectance Bands 1 and 2 and sqrt(band1*band2) for RGB channels for Chase County, KS from Julian Day 068 to day 129 in 2014. Darker areas represent recent burn scars. Regrowth is evident in later images.

Step 3: Evaluate burn scars, estimate the burn area.

Step 3 was the main focus of the initial study with the goal of creating an inexpensive easy to use tool that could automate the process of calculating burned area during the burning season using only the MODIS surface reflectance bands. Because the authors were new to both processing Raster data in R and remote sensing analysis in general much time was spent on learning the basic steps of manipulation of the data within R and graphical analysis and output. At the time of this conference (April 2015) the best estimating solution found was to perform a k-means clustering technique to identify the burn scars using common burn indices. Because each day of MODIS data has a different spectral range varying by county throughout the growing season the automation of the analysis was challenging and no single approach tried worked across all counties across multiple years. However, because of the scripting ability in R it was possible to loop through various techniques and visualize the output of each. From this visualization is was clear which technique worked best on any given day and the user could then use that answer. The authors believe additional refinement of the methods could improve the tool and estimation accuracy.

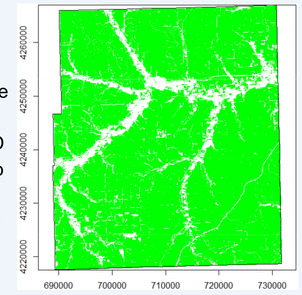
The current solution calculates three indices, NDVI, GEMI, BAI (Mohler, 2011) and uses those indices along with the raw wavelengths from the Red and NIR MODIS bands to calculate k-means clusters for each day. The clusters identify like areas, such as recent burn scars, and the software provides a visual estimate of the burns identified along with a calculated area burned.



Results

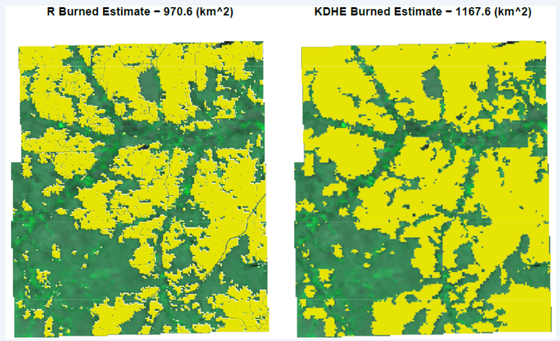
Step 4: Mask areas that likely did not burn.

For this step we used the 2011 National Land Cover Database available at a 30m resolution and excluded land cover types not likely to burn in a prescribed burning scenario. The NLCD dataset has 20 classifications. Of these 20 classifications two were retained for our study, the grassland/herbaceous and pasture/hay classifications. Other classifications such as water, developed areas, etc. were excluded.



Step 5: Evaluate area burned estimate from R approach with KDHE estimate.

In this final step we compare the estimate from R after resampling and masking to the higher resolution NLCD data set. Overall the magnitude and spatial burn scar areas identified between the two approaches are similar, generally less than a 20% area difference on any given day. Using the higher resolution mask removes some areas, such as roadways, that are not excluded in the KDHE estimate that uses a 250 meter resolution.



Conclusions

The freely available R statistical package is a tool that can be used to identify burn scars in the Flint Hills of Kansas and Oklahoma. The available R libraries contain all the tools needed to manipulate MODIS raster datasets and tools are available to do various statistics on the raster data including k-means clustering. Because of the ease in which R can be scripted the approach outlined in this poster could be applied to analyze multiple years and counties of burns. While the R statistical software does not contain all the features of commercial GIS packages, it is able to adequately estimate burn scars using unsupervised classification techniques giving results similar to other estimation techniques.

Future Work

Future work will include refining the R approach to better identify burn scars, including the potential of using other types of cluster approaches (neural net, random forest, etc.), and the investigation of additional remote sensing data including using additional MODIS bands. Ground-Truth of this R approach is also needed and has not yet been performed other than comparing R results with the KDHE estimates.

References

Mohler, L. Rhett, 2011: MULTI-SCALE BURNED AREA MAPPING IN TALLGRASS PRAIRIE USING IN SITU SPECTROMETRY AND SATELLITE IMAGERY, https://krex.k-state.edu/dspace/bitstream/handle/2097/11986/RhettMohler2011.pdf?sequence=1

MODIS Technical Specifications: http://modis.gsfc.nasa.gov/about/specifications.php

R Statistical Package: http://www.r-project.org/

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., and Xian, G. 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sensing of Environment, 132: 159 – 175.