**NASA DEVELOP National Program**



USGS at Colorado State University *Fall 2015*

Colorado Agriculture

Mapping Forest Species Composition at the Colorado State Forest State Park using Landsat 8 with Integrative Spatial Modeling

**Technical Report** 

Rough Draft – October 8, 2015

Brian Woodward (Project Lead)

Sarah Carroll

Nicholas Kotlinski

Eric Rounds

Dr. Paul Evangelista, Natural Resources Ecology Lab, Colorado State University (Science Advisor)

Tony Vorster, Bioenergy Alliance Network of the Rockies (Mentor)

# I. Abstract

[Placeholder - do not put anything here until the final draft submission. The abstract in the project summary is where the working draft of the abstract should “live”]

**Keywords**

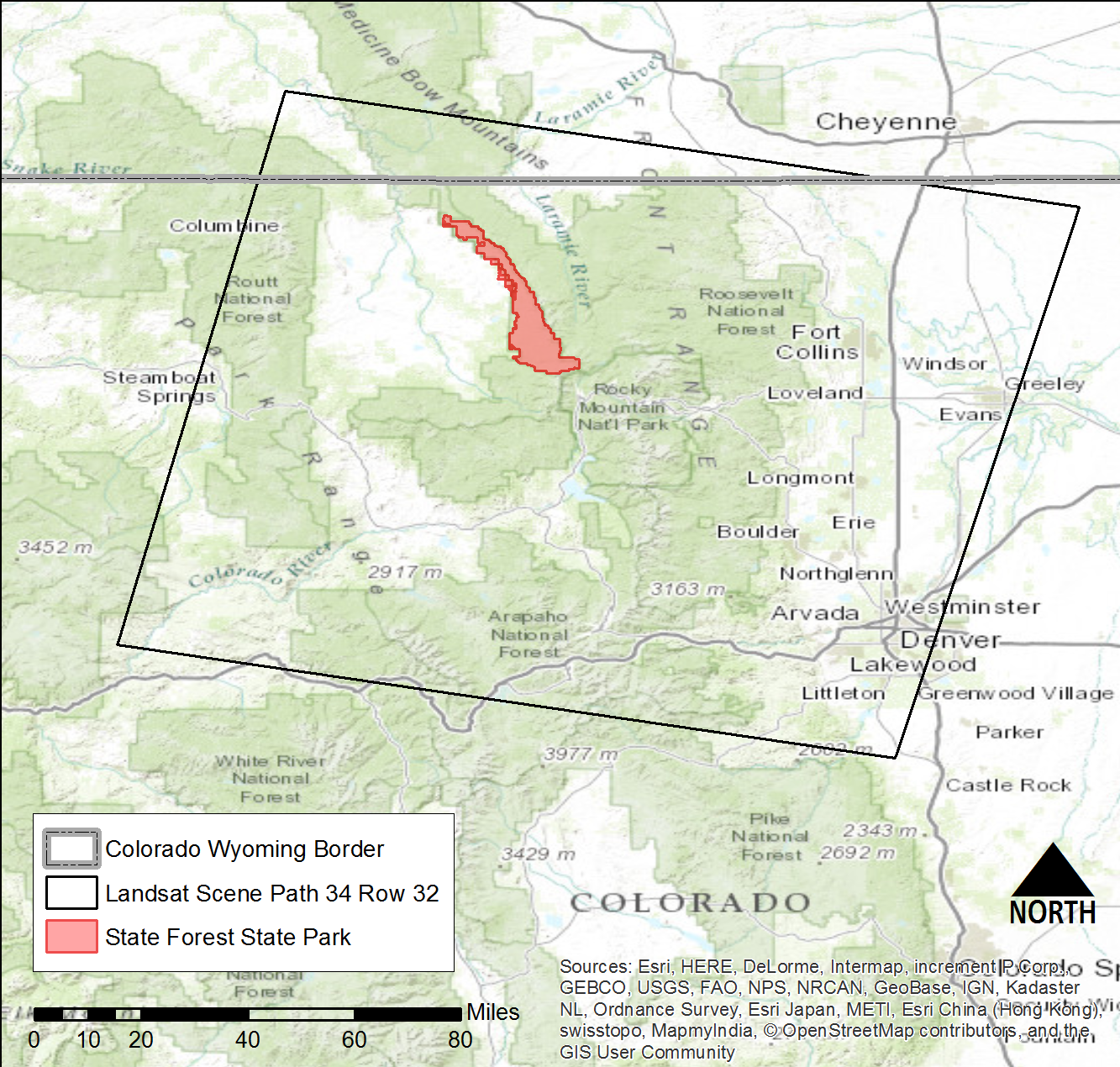
Remote Sensing, Species Composition, Colorado, Forest Management, Agriculture, Spatial Modeling

# II. Introduction

The Colorado State Forest (CSF) is managed for multi-use objectives including timber harvest, recreation, and wildlife. Forest management decisions are based on an adaptive and active management strategy that requires accurate and current data to inform management decisions. Presently, these decisions are based primarily on historical records, field observations, and National Agricultural Imagery Program (NAIP) imagery. Although forest managers do have a species composition map of the State Forest, the map is inaccurate due to the lack of revisions over time and the outdated methods used to map tree species. Moreover, the need for an up-to-date species composition map has been exacerbated by drastic changes in forest health caused by a pine-beetle epidemic in the early 2000s. As a result, an accurate map detailing forest species composition throughout the CSF would provide a baseline for ecological studies and timber harvest planning, while at the same time providing a reference for future studies of beetle-kill.

As a part of the Agriculture category of the DEVELOP National Application Area, this project utilizes NASA Landsat 8 (OLI) imagery in addition to ancillary datasets from fieldwork as inputs into a novel regression tree model to map forest species composition at a fine scale. Previous studies have proven rather successful for mapping percent canopy cover (PCC) by utilizing moderate resolution imagery like Landsat along with other datasets such as LiDAR (Ahmed, 2014; Carreiras 2006). The primary objective is to bridge the gaps in knowledge of forest species composition in the CSF and other regional forests by creating a more thorough and comprehensive record of species distribution. The resulting maps produced by this project will better inform forest management decisions and assist in the study of harvesting bioenergy feedstocks. Additionally, the results should provide useful data for ecological studies, as species composition information has important impacts on ecological functionality and resilience (Peterson et al., 1998). This type of study is critical; many studies that have attempted to characterize species composition of forest canopy cover in the past have done so without coinciding field data or a large study area (Alistair et al., 2009). As such, the combination of timely and accurate field data, spectral remote sensing products (Landsat 8 OLI), and zero-inflated modelling processes, should allow for a highly accurate view of species composition in north-central Colorado with implications for larger-scale studies.

The study area focuses on the Colorado State Forest in north-central Colorado (Figure 1), which stretches across 70,000 acres through Jackson and Larimer counties. The area was established as a state park in 1970 and is now actively managed by Colorado Parks and Wildlife, Colorado State Forest Service, and other partners. Elevation in the park ranges from approximately 2500 to nearly 4000 meters and more than seventy percent of the acreage is covered with coniferous forests (Colorado Parks & Wildlife). These montane forests are dominated by lodgepole pine *(Pinus contorta*) and fir-spruce communities in the subalpine zone, also including Douglas fir (*Pseudotsuga menziesii)*, Engelmann Spruce (*Picea engelmannii*), subalpine fir (*Abies lasiocarpa)*, and aspen (*Populus tremula*) (Wiken et al., 2011). In order to capture current species composition in the CSF, Landsat imagery from 2014 and 2015 was utilized. The data covers summer (June and July) and a fall month (September) in order to provide maximum sun cover along with subtle changes in canopy character that may inform the study.



*Figure 1. Colorado Agriculture study area in north-central Colorado*

Our project partners and end-users include Tony Vorster and John Twitchell. Tony Vorster is part of the feedstock supply team with the Bioenergy Alliance Network of the Rockies (BANR). This large, collaborative project is assessing the feasibility of using beetle-kill wood as a feedstock for biofuels. A map of species composition could be used by BANR to map live and dead biomass as they assess the location and quantity of potential feedstock for biofuel in beetle-kill areas. John Twitchell is a district manager for the Colorado State Forest Service. The forest service is in need of a current, highly accurate, and comprehensive species composition map to aid in ecological knowledge and management of the CSF. Species composition data provides baseline information on forest structure and health. This information is crucial for forest managers in areas across the state of Colorado where high impact disturbance events, including wildfire and beetle-kill outbreaks, are driving forest structure and recovery trajectories. In 2014, over 890,000 acres of Colorado forests were impacted by insect or disease activity. Species composition data are becoming an increasingly informative management tool in areas of bark beetle outbreak, as forest recovery trajectories are often dependent upon pre-outbreak species composition. These same data can aide foresters in identifying areas susceptible to severe outbreaks as well as with predicting future spread.

# III. Methodology

**Data Acquisition**

Three Landsat 8 OLI scenes for path 34 row 32 were downloaded from USGS Earth Explorer, including two summer scenes, July 2014 and June 2015, as well as one fall scene, September 2014. The scenes were downloaded as a Level 1 data product with atmospheric correction provided by EROS data center. Shuttle Radar Topography Mission (SRTM) data was downloaded from USGS Earth Explorer, providing the team with a 30 meter digital elevation model (DEM) of the study area. In addition, a National Agriculture Imagery Program (NAIP) image containing the study area was downloaded from the USDA.

Reference data on the CSF was collected during fieldwork undertaken by BANR and their associates. The reference data provides forest composition information (e.g. species, status, diameter at breast height [DBH], and densiometer readings for over ninety 30x30 plots within the CSF. The geographic center of these plots corresponds with the center of Landsat pixels. Other ancillary data utilized in this project include a CSF administrative boundaries shapefile and a point shapefile with the locations of plots surveyed during the BANR fieldwork.

**Data Pre-processing**

Several pre-processing steps with the field data were taken to prepare it for input into the model created by Rick Lawrence and Shannon Savage at Montana State University. The species composition reference data was used to create relative canopy cover percentages for the four dominant species (ABLA, PICO, PSME, and PEIN) within all ninety-two plots surveyed by BANR. Percent canopy cover was determined for each species by deriving a total basal area sum for each tree species within a plot, converting those to percentages, and then adjusting those percentages to account for open sky readings from densiometer data.

Once the reference data was properly converted, imagery data had to be formatted to the specifications provided by Savage et al. (2015). Initially all imagery acquired for the model was converted to ERDAS Imagine (.img) file format, and then subsequently all imagery was clipped to the CSF study area using the extract by mask tool in ArcMap. Using the same tools, slope and aspect images were generated in ArcMap using a 30 meter DEM for additional bands.

A principal component analysis was then performed on 1 meter NAIP imagery in ENVI. The resulting standard deviation of the first principal component was used to create texture mean and minimum values by utilizing the focal statistics tool in ArcMap. To make these data consistent for analysis these images were resampled to a 30 meter pixel size.

With all of the imagery formatted properly, the next step was to create a single stack composite, as detailed by Savage et al. (2015) in ArcMap. The stack (see Table 1. Appendix) consisted of select raw Landsat bands for each year’s scene, a DEM, slope and aspect raster’s, and NAIP texture values. Finally, a point shapefile containing the pixel locations of the corresponding reference data was used to extract the values in Spatial Analyst (Extract Multi Values to Points) from each band in the stack composite for all ninety-two pixels surveyed. The resulting values were inserted into a table displaying every value derived from the images included in the 21-band composite for each plot.

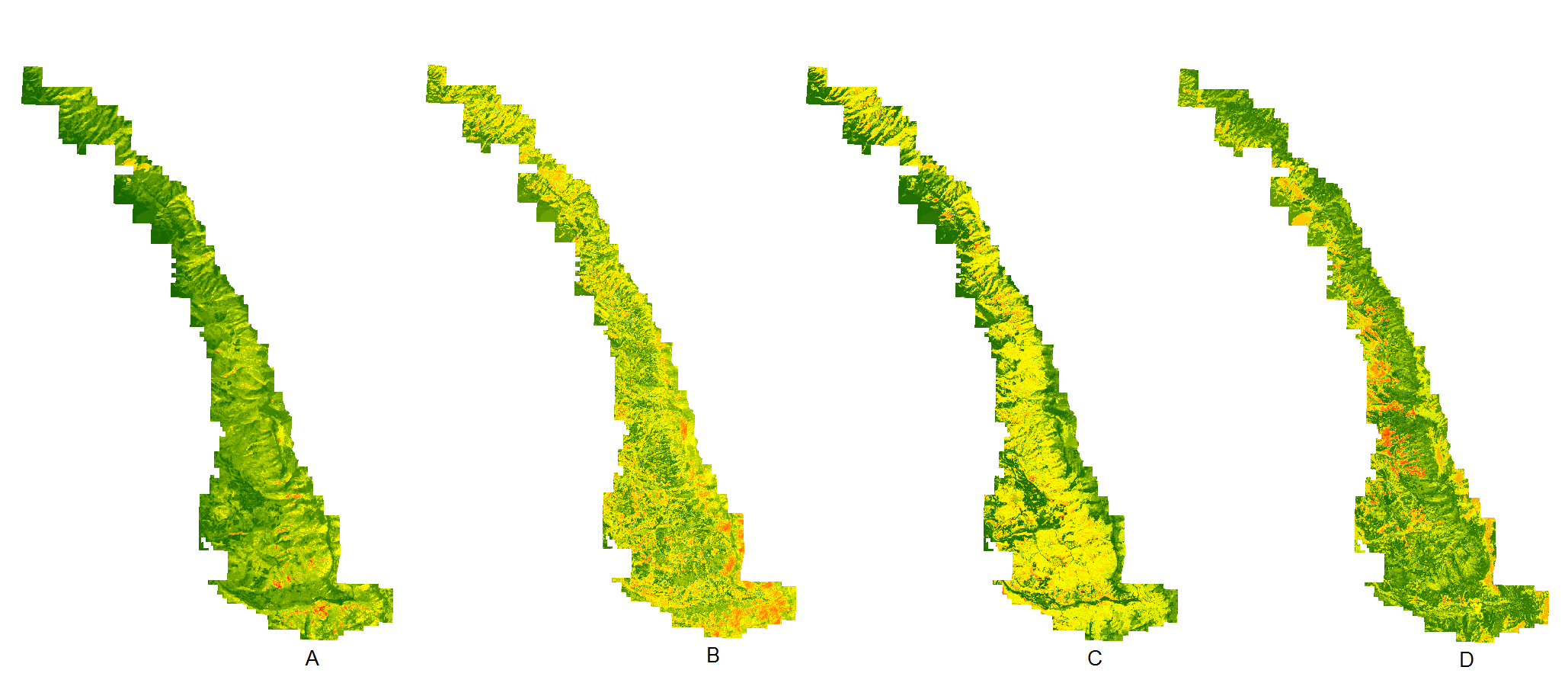
**Data Analysis**

The stacked image composite, species presence field-data, and associated digital number (DN) values at each field-level point for individual stacked bands were run through a two-step Zero-Inflated Model (generic\_ZImodeling \_rinput) in RStudio. Step one involved the input of all reference data into a Generalized Linear Model (GLM), resulting in a binary (BIN) map of species presence (1) or absence (0). For step two, all non-zero (species present) data were used in a Support Vector Machine (SVM), creating a continuous (CON) map of species distribution in the study area. The two associated maps (BIN & CON) for each species are then combined to only include the continuous output data (1) where the binary map is. Any pixel marked as (0) will include no data for that particular species.

# IV. Results & Discussion

**Analysis of Results**

The output of the Zero-Inflated Model provided a reasonable visual output (see Figure 2), with location confidence representing varying spatial patterns for each species with high probability present in specific pockets of forest area.



*Figure 2. Preliminary percent canopy cover composition map outputs per species from R zero-inflated model based on individual species data. Reds and oranges represent high canopy cover, yellows moderate cover and green minimal to no canopy cover. Mapped species include: (A) Subalpine Fir [ABLA] (B) Lodgepole Pine [PICO] (C) Engelmann Spruce [PIEN] (D) Aspen [POTR]*

**Error and Uncertainty**

Problems with running the model included issues with constructing an image composite that mirrored Savage et al. (2015) due to unclear documentation in the literature. Additionally, model outputs at this stage have not been verified and the statistics remain unclear. A major issue with the model was in acquiring a BIN map output from the SVM, which prohibits combining CON and BIN results for a finished species composition product. As such, more time is required to analyze and understand the impacts of preliminary composition maps (Figure 2), as well as to troubleshoot and clean up the binary model output before continuing to compile a final map of specific species locations in CSF.

# V. Conclusions

Coming Soon

# VI. Acknowledgments

This project was made possible through support and mentorship of select individuals and organizations. We extend a kind thanks to the following:

* Dr. Paul Evangelista, Colorado State University (CSU), Natural Resource Ecology Laboratory (NREL)
* Tony Vorster, CSU, NREL, Bioenergy Alliance Network of the Rockies (BANR)
* Shannon Savage, Montana State University (MSU)

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract NNL11AA00B and cooperative agreement NNX14AB60A.

# VII. References

Ahmed, O.S., Franklin, S.E., and Wulder, M.A. 2014. Integration of Lidar and Landsat

data to estimate forest canopy cover in coastal British Columbia.

*Photogrammetric Engineering & Remote Sensing*: Volume 80, pp. 953-961.

Alistair, M.S., Falkowski, M.J., Hudak, A.T., Evans, J.S., Robinson, A.P., and Caiti M. Steele.

2009. A cross-comparison of field, spectral, and lidar estimates of forest canopy

cover. *Canadian Journal of Remote Sensing*: Vol. 35, No. 5. pp. 447-459.

Carreiras, J.M.B., Pereira, J.M.C., and Pereira, J.S. 2006. Estimation of tree canopy cover

in evergreen oak woodlands using remote sensing. *Forest Ecology and*

*Management*: Volume 223, pp. 45-53.

Colorado Parks & Wildlife. Colorado Parks & Wildlife. http://cpw.state.co.us/placestogo/

parks/ StateForest/Pages/default.aspx.N.p.,n.d. Web. 01 Oct. 2015.

Peterson, G., Allen, C.R., and Holling, C.S. 1998. Ecological Resilience, biodiversity, and

scale. *Ecosystems,* 1, pp. 6-18.

Savage, S.L., Lawrence, R.L., and John R. Squires. 2015. Predicting Relative Species

Composition within Mixed Conifer Forest Pixels Using Zero-Inflated Models and Landsat Imagery. Unpublished Manuscript (*In Review*).

Wiken, Ed, Francisco Jiménez Nava, and Glenn Griffith. 2011. North American Terrestrial

Ecoregions—Level III. Commission for Environmental Cooperation, Montreal, Canada. pp. 88-92.

# VIII. Content Innovation

Coming Soon

# IV. Appendices

Insert here

# IV. Appendices

|  |  |
| --- | --- |
| Model Image Stack Band Number | Landsat 8 (Year - Band) / Raster Data Source |
| 1 | July 2014 - Blue |
| 2 | July 2014 - Green |
| 3 | July 2014 - Red |
| 4 | July 2014 - NIR |
| 5 | July 2014 - MIR1 |
| 6 | July 2014 - MIR2 |
| 7 | July 2014 - TIR1 (Thermal) |
| 8 | July 2014 - TIR2 (Thermal) |
| 9 | September 2014 - Blue |
| 10 | September 2014 - Green |
| 11 | September 2014 - Red |
| 12 | September 2014 - NIR |
| 13 | September 2014 - MIR1 |
| 14 | September 2014 - MIR2 |
| 15 | September 2014 - TIR1 (Thermal) |
| 16 | September 2014 - TIR2 (Thermal) |
| 17 | June 2015 - Blue |
| 18 | June 2015 - Green |
| 19 | June 2015 - Red |
| 20 | June 2015 - NIR |
| 21 | June 2015 - MIR1 |
| 22 | June 2015 - MIR2 |
| 23 | June 2015 - TIR1 (Thermal) |
| 24 | June 2015 - TIR2 (Thermal) |
| 25 | DEM, 30-meter (SRTM) |
| 26 | Slope |
| 27 | Aspect |
| 28 | NAIP Mean (Texture) |
| 29 | NAIP Minimum (Texture) |

*Table 1. A 21 band image combination based on specifications from Savage et al. (2015) used as model input*