NASA DEVELOP National Program
Idaho – Pocatello
Summer 2020

Mark Twain National Forest Ecological Forecasting

Utilizing NASA Earth Observations to Classify Ground Cover Types in the Mark Twain National Forest

DEVELOP Technical Report
Final Draft – August 6th, 2020

Kaitlyn Bretz (Project Lead)
Madison Bradley
Sarah Hafer
Grant Verhulst

Advisor
Keith Weber, Idaho State University, GIS Training and Research Center (Science Advisor)

1. Abstract

The Mark Twain National Forest (MTNF) encompasses 1.5 million acres of public land in the Ozarks region of southeastern Missouri. The industrial boom between the 1880s and 1920s had devastating effects on the shortleaf pine (*Pinus echinata*), Missouri's only native pine species. The combination of fire suppression on this fire-dependent species and timber harvest of mature pine stands inhibited the development of pine seedlings and promoted the establishment of hardwood stands. Partners at the US Forest Service's MTNF are currently involved in restoration efforts in two ranger districts, which includes removing invasive eastern red cedar (*Juniperus virginiana*) and prescribed burning. To expand spatial coverage for the MTNF beyond *in situ* observation sites, the NASA DEVELOP team analyzed land cover change from 1986 through 2019 and forecasted changes based on a 'business-as-usual' scenario out to 2040. The team incorporated remotely sensed data from Landsat 5 Thematic Mapper (TM) and Landsat 8 Operational Land Imager (OLI) into the random trees supervised classification tool in ArcGIS Pro. This tool spectrally separated pixels into five distinct land cover classes and produced classifications for 1986 and 2019, with kappa statistics of 0.87 and 0.81, respectively. Overall, there was a net decrease in conifer and meadow land cover between 1986 and 2019 along with a net increase in water, developed, and deciduous land cover. The team used TerrSet's Land Change Modeler to forecast land cover through 2040. Results showed an increase in coniferous land cover and a decrease in deciduous cover, indicating a high probability that current restoration efforts will produce the intended effect.

Key Terms
supervised classification, ecological forecasting, eastern red cedar, shortleaf pine, forest restoration, Landsat 5 TM, Landsat 8 OLI, Missouri

2. Introduction

2.1 Background Information

Missouri's lumber boom between the 1880s and 1920s had devastating effects on the shortleaf pine (*Pinus echinata*), Missouri's only native pine species. This species typically occurs in upland areas above glades but is also planted in plantations in valley bottoms (K. Steele, personal communication, July 15, 2020). Early settlers engaged in the unregulated exploitation of local forest resources for timber harvest, railroad construction, and agriculture, and disrupted the natural fire regime. The combination of fire suppression on this fire-dependent species and timber harvest of mature pine stands inhibited the development of pine seedlings and promoted the establishment of hardwood stands (Cunningham, 2007). These factors also enabled the establishment of out-competing hardwood stands (Cunningham, 2007). The Missouri Ozarks were once covered by approximately 6 million acres of shortleaf pine-oak woodlands dominated by longer-lived pine, white oaks, and post oaks. Today, only 600,000 acres of these woodlands remain within their native range, as short-lived scarlet and black oaks and the invasive eastern red cedar (*Juniperus virginiana*) have come to dominate the area (Cunningham, 2007).

While the eastern red cedar tree species is native in parts of eastern North America, it is considered invasive in the Ozarks. This species typically occurs on

overgrown glades, dolomite cliffs, and old fields (K. Steele, personal communication, July 15, 2020). The presence of these trees can have long-lasting effects on native species, such as warm-season grasses and drought-adapted wildflowers in Missouri's glades, as well as the shortleaf pine. The eastern red cedar has four factors that enable it to quickly dominate grasslands and abandoned land, such as old and unused agricultural fields. These factors are (*1*) high seed production and the ability to disperse seeds rapidly; (*2*) actively pulling calcium into soils, which elevates the pH, making the soil inhospitable to native species; (*3*) falling bark, needles, and cones contribute to increases in the soil pH and prevent water from reaching the seedlings of native grasses; and (*4*) the deep shade produced by red cedar canopies hinders the survival of other plant species, causing bare earth to develop around cedar tree crowns (Corbett & Lashley, 2017).

Significant advancements in remote-sensing capabilities over the last four decades allow for the remote classification of tree species and forest types. Many published studies use methodologies with various remote sensing techniques, but few of these cover large-scale geographical areas (Fassnacht et al., 2016). Remote sensing can provide a deeper understanding of forest structure and biochemistry, along with land use and land cover change. Machine learning models can segment images, classify pixels by land cover classes, and simulate forest change with varying climate factors using robust remotely sensed data (Shugart, 2015). The random trees (also called random forest) classifier is one such model that uses supervised classification methods coupled with advanced machine learning to segment multispectral images into different classifications based on input training sample data (Pal, 2005). The use of high-resolution, multispectral imagery provides spatially explicit information as a basis for large-scale mapping of forest composition based on land cover classification and can prove useful for forest administrations (Fassnacht et al., 2016). The Landsat program provides the longest visual record of land use data. Landsat 5 Thematic Mapper (TM) launched in 1984 provides data through 2013, and Landsat 8 Operational Land Imager (OLI) launched in 2013 is still operational. Because of their reliability and expansive timeframe, Landsat 5 TM and Landsat 8 OLI are commonly used in forest classification studies.

2.2 Study Area & Period

This study focuses on Mark Twain National Forest (MTNF), which encompasses 1.5 million acres of public land, mostly within the Missouri Ozarks *(Figure 1)*. The forest is composed of nine distinct geographic areas with six ranger districts. The heterogeneous landscape, characterized by a mosaic of ecologically complex open grassy woodlands and savannas, supports nearly 750 native animal species and over 2,000 plant species. Of these, 500 native plant species are found in glade habitats, as many of Missouri's rarest plant species are glade-dependent and cannot grow elsewhere. Management objectives of the forest seek to restore natural communities to ensure the long-term sustainability of this globally distinct landscape. These restoration efforts are currently active on two fronts: glade restoration and shortleaf pine restoration.

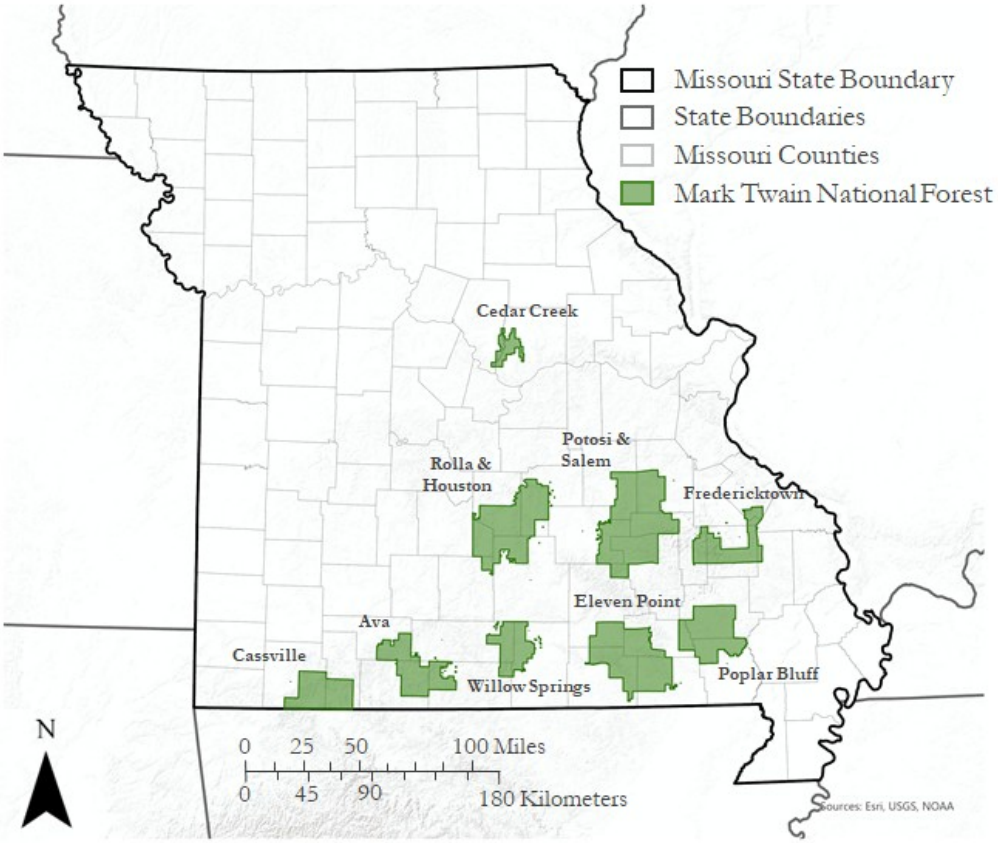

Figure 1. Nine geographically distinct regions of Mark Twain National Forest in Southern Missouri.

In 2005, the United States Department of Agriculture (USDA) US Forest Service (USFS) at MTNF published the Land and Resource Management Plan (hereafter the '2005 Forest Plan'), a forest-wide plan that guides natural resource management activities for a 10- to 15-year period. The 2005 Forest Plan directed individual ranger districts to determine restoration priorities from 2005 through approximately 2020 (Moore, 2005). In 2009, the US Congress passed the Collaborative Forest Landscape Restoration Program (CFLRP). The CFLRP is a federal initiative that encourages the restoration of natural forest landscapes across the United States. This catalyzed a collaborative project between private landowners and the USFS MTNF focused on the restoration of shortleaf pine-oak woodlands. The CFLRP aligns with the goals of the MTNF 2005 Forest Plan. Both management strategies support reforestation to restore historic shortleaf pine-oak woodland community composition and structure. However, the 2005 Forest Plan has highlighted 29% of MTNF's 1.5 million acres as potentially eligible for restoration efforts, whereas the CFLRP highlights eligible land in just two of the six ranger districts, totaling only 8% of MTNF (approximately 100,000 acres in Eleven Point and Poplar Bluff districts, USDA Forest Service, 2020a). Glade restoration efforts within the Ava district include the thinning of eastern red cedar and prescribed fires along the Glade Top Trail. The Ava ranger district holds 26,708 acres of glades within MTNF's total of 44,000+ acres, with 8,709 of those acres under conservation management (USDA Forest Service, 2020b). Using NASA Earth observations (EO), conifer forest can be delineated from other forest types on a landscape-level throughout the forest, with the potential to also spectrally

separate shortleaf pine and red cedar tree stands, for efficient, targeted restoration.

2.3 Project Partners & Objectives

The team partnered with the USFS MTNF and the USFS Geospatial Technology and Applications Center (GTAC) to complete this project. Management decisions regarding restoration efforts are determined on a case-by-case basis for individual ranger districts (e.g., Eleven Point, Poplar Bluff, and Ava districts), rather than at the landscape-scale for the entire MTNF. The use of remotely sensed and geospatial data is not a requirement within the USFS's 2005 Forest Plan or the CFLRP. Though the USFS GTAC in Utah serves as a resource, they are not required to be involved with a project unless managers reach out and invite them. This means that remote sensing data resources are generally used on a limited basis and are not comprehensive across MTNF.

This project had two primary objectives: (1) create a land cover type analysis of MTNF that can be used to complete species-level classifications; and (2) forecast changes in land cover composition out to the year 2040 based on current USFS MTNF land management practices. Eastern red cedar and shortleaf pine trees are the species of focus. The USFS MTNF uses eastern red cedar trees as markers for shallow or exposed soil in the national forest, which helps create general maps of areas in need of restoration. The shortleaf pine is the only native pine species in the Missouri Ozarks and under the CFLRP. Shortleaf pine-oak woodland areas are targeted for restoration to their original extents (USDA Forest Service, 2020a). In addition to their historical significance, native shortleaf pines also store carbon and influence groundwater stocks and slope erosion. If missing from an area, these factors can run unchecked and greater targeted restoration will be necessary.

These concerns are the chief motivation for delineating between red cedar stands, pine stands, and other forest types on a landscape-scale in MTNF. The USFS MTNF has yet to complete a comprehensive, forest-wide inventory of forest and glade ecological sites in MTNF. A cover type analysis of the forest canopy forecasted out to the year 2040 will assist forest managers with future species-level classification attempts and make evaluating management practices more efficient.

3. Methodology

3.1 Data Acquisition

Landsat 5 TM and Landsat 8 OLI surface reflectance (SR) imagery was acquired through Google Earth Engine (GEE; Gorelick et al., 2017). These images are pre-processed to remove atmospheric effects and generate bands for pixel masking. To evaluate pixels for masking, the US Geological Survey (USGS) uses CFMask, an algorithm that populates cloud, cloud confidence, cloud shadow, and snow/ice pixels within the quality assessment band (Foga et al., 2017). Landsat 5 TM and Landsat 8 OLI imagery was atmospherically corrected using the Landsat Ecosystems Disturbance Adaptive Processing System (LEDAPS) and the Landsat Surface Reflectance Code (LaSRC), respectively. These two algorithms process top-of-atmosphere data to remove atmospheric effects, which involves characterizing the absorption and scattering of electromagnetic radiation signals from gases and aerosols in the Earth's atmosphere (Song et al., 2001).

Atmospheric correction via the LEDAPS and the LaSRC are comparable; however, the LaSRC algorithm for Landsat 8 has been found to have slightly higher accuracy (Dwyer et al., 2018).

The goal was to generate two composite images for each year: one image representing leaf-off (winter) conditions, generally created from scenes captured between December and February, and one image representing leaf-on (summer) conditions, created from scenes captured from June through August. Image quality and availability were different between each year, so imagery was selected at approximately 5-year intervals to cover the project timeline: 1986-87, 1991-92, 1996-97, 2001-02, 2006-07, 2010-11, 2016-17, 2019-20. Several of the selected years required different filter values to cover the full study area (see Appendix A for a detailed table). These images were processed within GEE prior to exporting them into ArcGIS Pro for analysis.

Elevation data were collected from the National Elevation Dataset (NED) hosted by USGS' The National Map (TNM) online downloader. 1/3 arc-second digital elevation models (DEM) were collected for analysis. Data were mosaicked in Esri ArcGIS Pro and masked to the MTNF border extent. Slope and aspect were calculated from the DEM using the built-in raster functions toolbox in ArcGIS Pro.

High-resolution aerial imagery was collected from the National Agriculture Imagery Program (NAIP) and the Missouri County-Extent Digital Orthophoto Quarter Quads (DOQQ) collection via the State of Missouri's spatial data information service. 1990 DOQQ and 2018 NAIP imagery was imported into ArcGIS Pro via image servers and used to develop training sample data. DOQQ imagery was gathered by Missouri in 1990, 1995-1997, 2008, and 2015 and provide high-resolution black-and-white images for each county. NAIP was first contracted to collect imagery on a five-year cycle beginning in 2003. In 2009, the contracts transitioned to a 3-year cycle. The default spectral resolution is 'natural color,' so there are three bands: red, green, and blue. Beginning in 2007, some states delivered 4-band aerial imagery that included near infrared (NIR); however, this imagery was not available within the Missouri state lines. All NAIP images are required to have a one-meter ground-sample distance, 6-meter horizontal accuracy, and no greater than 10% cloud cover.

3.2 Data Processing

Landsat imagery from GEE was filtered by location, date, and cloud cover. Images with pre-calculated cloud cover greater than 10% were removed and all scenes were clipped to the MTNF boundaries. The team applied a cloud mask to the selected scenes and removed the pixels where the quality band indicated a high confidence of cloud cover. A snow mask was created by calculating the Normalized Difference Snow Index (NDSI) and masking out the pixels where NDSI > 0.3. This value was selected by trial-and-error methods involving a high-level visual review of the masked pixels. The NDSI uses a ratio between the green and shortwave infrared (SWIR) spectral bands to identify areas with snow cover (*Equation B1*). Index values range from –1 to 1, with negative values indicating little to no snow cover and positive values indicating a high probability of snow cover.

After processing within GEE, the selected Landsat imagery was subsequently composited into two images for each year: leaf-off and leaf-on. The team used these composite images to calculate several image derivatives within GEE. Normalized Difference Vegetation Index (NDVI) is a commonly used vegetation index that quantifies vegetation greenness (*Equation B2*). This index is often used to monitor droughts, croplands, and overall vegetation density and health. The Enhanced Vegetation Index (EVI) is similar to NDVI, as it quantifies vegetation greenness, but it corrects for several atmospheric variables and canopy background noise, and is less prone to oversaturation (i.e. is more sensitive to areas with dense vegetation) *(Equation B3)*. Normalized Difference Water Index (NDWI) is a ratio between the green and NIR spectral bands. NDWI values help differentiate between open waterbodies and any surrounding soil and vegetation features; however, this index can be sensitive to turbidity (measure of suspended solids, such as soil, within water features; McFeeters, 1996 [*Equation B4*]). These indices were selected based on their common application in previous land cover classification studies and utility in spectrally separating seasonally affected vegetation. They are important for classification as they show seasonal changes, record differences in vegetation, and isolate bodies of water (see Appendix B).

Composited Landsat images and their derivatives were exported from GEE and were mosaicked and resampled in ArcGIS Pro in preparation for supervised classification. Two mosaics were created for each select year: leaf on (summer months) and leaf off (winter months). These mosaicked composite images simply combined all the bands from each scene into a single raster. ArcGIS Pro can classify a single raster at a time, but that raster can have any number of bands. In anticipation of this, the team further composited all of the processed data into a single raster. The resulting images slated for classification had 32 bands, including all the bands from both original composite images (winter and summer), image derivatives, and elevation derivatives.

The Training Samples Manager tool in ArcGIS Pro was used to create sample training and testing data from Landsat 5 (1986 - 2012) and Landsat 8 (2013 - 2019) products in preparation for pixel-based supervised classification. The land cover analysis consisted of 5 classes – coniferous forest, deciduous forest, meadow, water, and developed – each with approximately 100 training/testing polygons drawn around visually-classified pixels. The team validated these training/testing samples through visual comparison with high-resolution (1m) DOQQ imagery (1990, 1995 - 1997) and NAIP imagery (2002, 2006, 2012, 2016, and 2018). Training and testing polygons were subset into two groups on a 60:40 split, respectively.

3.3 Data Analysis

The team used the Random Trees classifier (also referred to as the Random Forest classifier) to complete a pixel-based supervised classification of the composite images. The training points (60% of the total samples collected) were used to train the model on what classes the team needed to identify. After preliminary runs, the team decided that the most accurate classification was achieved by using a maximum of 150 trees, with a maximum tree depth of 30, and a maximum of 1000 samples per class. After reviewing results from the supervised classification for the originally selected years, the team elected to focus on the supervised classification

results from 1986 and 2019 due to missing data from pixel masks and low accuracy values of the intermediate years during preliminary runs.

To quantify the error associated with the supervised classifications, the team created a confusion matrix within ArcGIS Pro and calculated the associated Cohen's Kappa Statistic. The confusion matrix is created by using randomly generated accuracy assessment points within ArcGIS Pro to compare with the classification results. The accuracy assessment points (n = 500) are randomly placed within the testing polygons created earlier (40% of the total samples collected). The confusion matrix calculates both user's accuracy (false positives) and producer's accuracy (false negatives), both of which are captured by a percent, with 100% being perfectly accurate. An error confusion matrix with Kappa statistics was created to validate data and check the accuracy of the model for each supervised classification (Appendix B). Cohen's Kappa Statistic measures consistency and agreement of ratings for categorical items. The equation is as follows (*Equation 1*) where p_o is the observed agreement and p_e is the expected agreement. A kappa value of $\kappa=1$ indicates full agreement. This statistic is calculated automatically by ArcGIS Pro.

(1)

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

A land cover change assessment and forecasting were completed in Idrisi TerrSet's Land Change Modeler (LCM). Forecasting with LCM is a step-by-step process that requires three elements: (1) Change Analysis; (2) Transition Potential Modeling; and (3) Change Prediction. First, the team ran a land cover change analysis between 1986 and 2019 using the classified composite images from each year. After a quick analysis of the net change between these years, the team opted to focus on six potential land cover transitions – conifer to deciduous, meadow to deciduous, conifer to meadow, deciduous to meadow, deciduous to conifer, and meadow to conifer – as these transitions were predicted to be the most relevant within MTNF and were important to the project partners at USFS. From the resulting land cover change maps, all potential transitions between the selected land cover classes were modeled from transition-driving variables using a Multi-Layer Perceptron (MLP) neural network with the default parameters. The transition-driving variables used in this analysis included DEM, slope, and aspect layers derived from the NED. Next, the transition potential maps were used to quantify the change in each transition between the five land cover classes, resulting in a forecasted land cover change map out to the year 2040 based on trends between 1986 and 2019. Finally, the 2019 classified composite image and the 2040 forecasted land cover map were imported into the Change Analysis to quantify the predicted net change out to 2040.

4. Results & Discussion

4.1 Analysis of Results

4.1.1 Land Cover Change

Supervised classifications for 1986 *(Figure 2)* and 2019 *(Figure 3)* were completed with Kappa statistic values of 0.86 and 0.81, respectively. This indicates a good agreement with validation data and an accurate classification model (*Appendix C*).

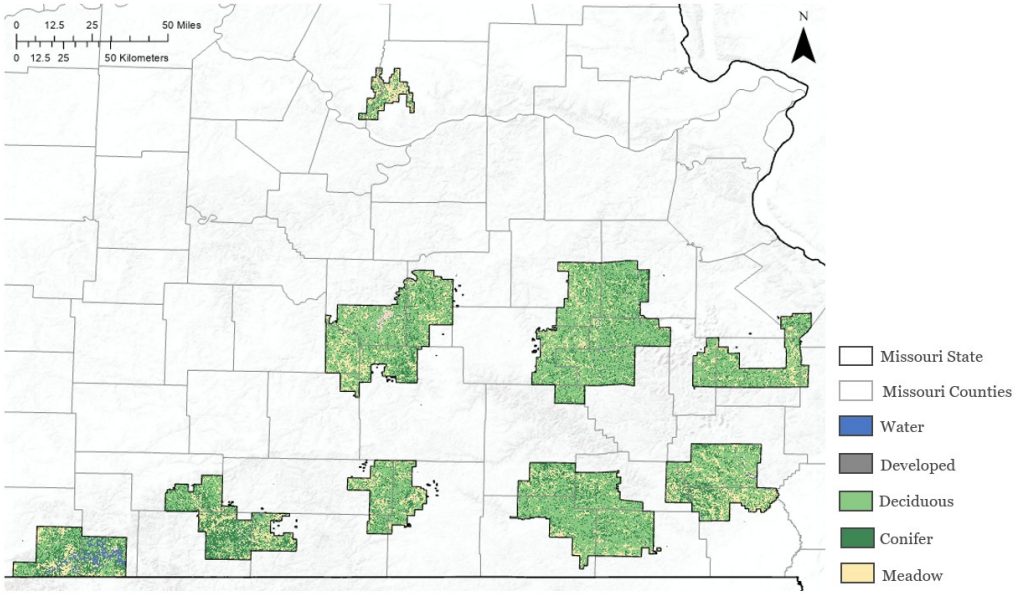

Figure 2. Land cover classification for 1986

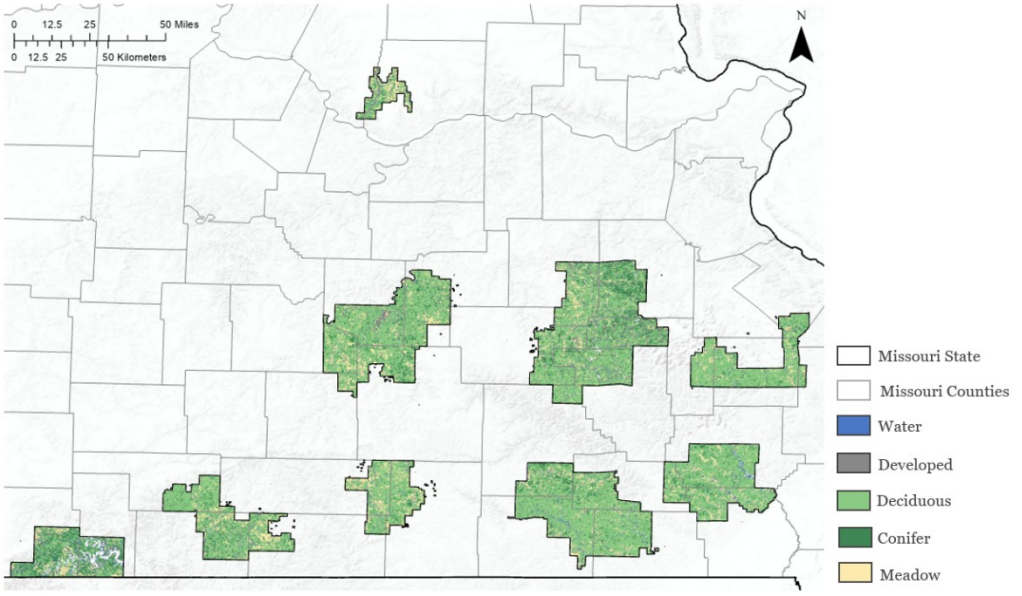

Figure 3. Land cover classification for 2019

A land cover change analysis for the entirety of MTNF between the classified years 1986 and 2019 is shown in *Figure 4*. This analysis indicates a loss in conifer forest and meadow land and an increase in water, developed, and deciduous forest *(Figure 5)*. It is important to note that 2019 was a record-breaking year for flooding in Missouri, which may account for the increase in water between the two classifications. A more detailed land cover change analysis of the Cassville ranger district is shown in *Figure D1*. This land cover change map for the Cassville ranger district accounts for all potential transitions between the five land cover classes used in the analysis. One of the most prevalent transitions is from conifer forest and meadows to deciduous forest (light green). There is also a large area that

transitions from vegetation to water or developed areas, seen in the blue/purple pixels in the upper right side of this ranger district. Land cover changes, especially in the loss of conifers, can be seen more closely in *Figure D2*, which compares the 1986 and 2019 classifications for the Eleven Point and Ava ranger districts.

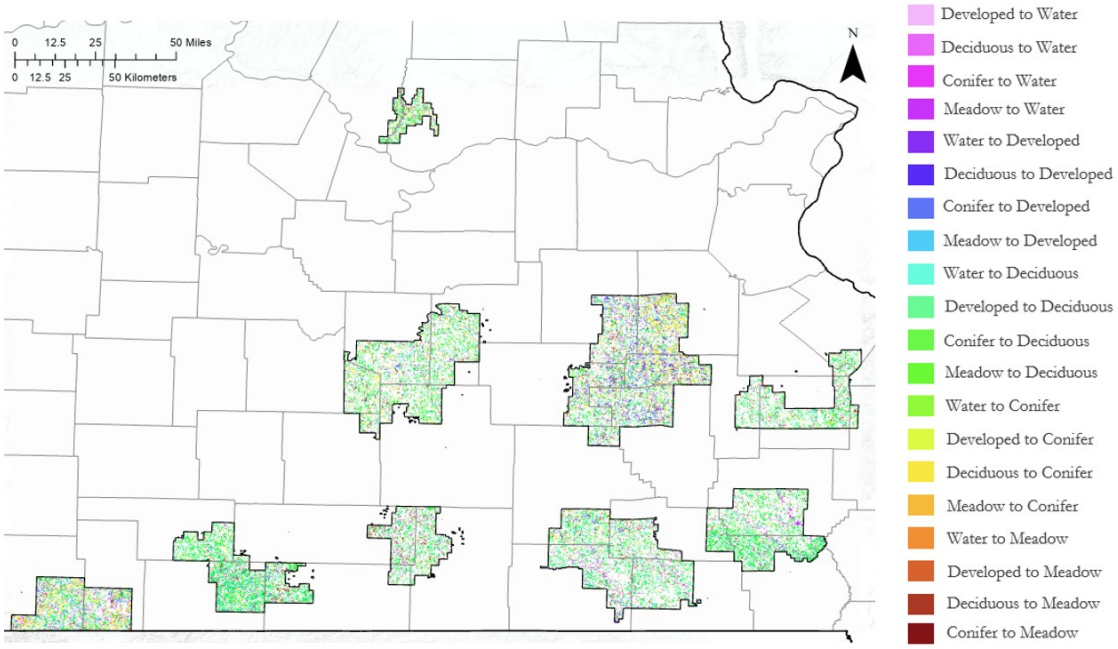

Figure 4. Land cover change analysis from TerrSet between 1986 and 2019

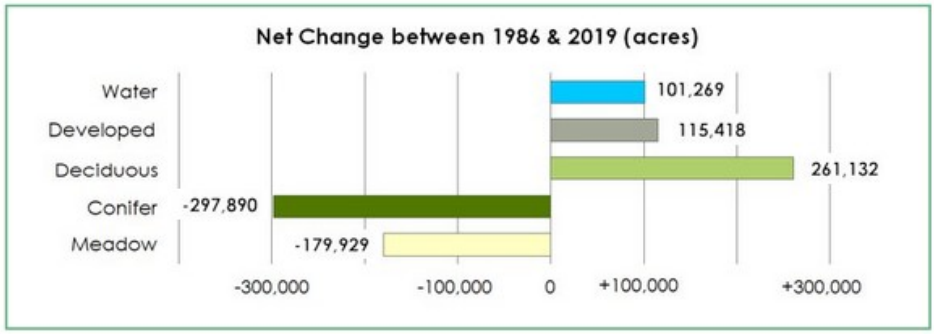

Figure 5. Net land cover change analysis for MTNF between 1986 and 2019

The 2040 forecasting map *(Figure 6)* estimates forest composition in the year 2040 under business-as-usual practices based on current land management strategies. This map is again classified into coniferous forest, deciduous forest, water, developed, and meadow. Transitions were focused between the conifer, deciduous, and meadow classes in the analysis of net change from 2019 and 2040. There is a projected overall increase in coniferous forest and meadows, with a projected decrease in deciduous forest between the years 2019 and 2040 as seen in *Figure 7*. A land change analysis map for 2040 *(Figure 8)* shows the predicted land cover transitions in 2040. A more detailed look at the Cassville district can be seen in *Figure D4*. The forecast for the Cassville district differs from the majority of the forest in terms of conifer regrowth.

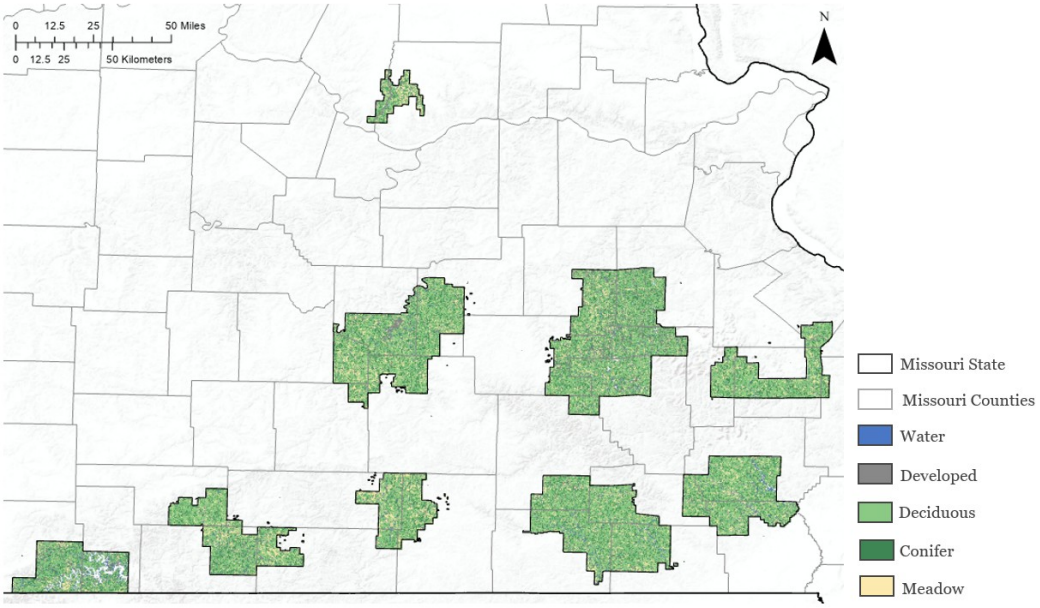

Figure 6. Forecasted land cover classification for 2040

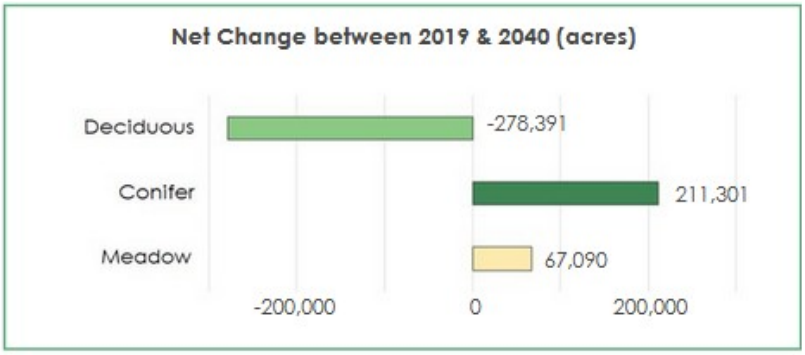

Figure 7. Net land cover change analysis for MTNF between 2019 and 2040

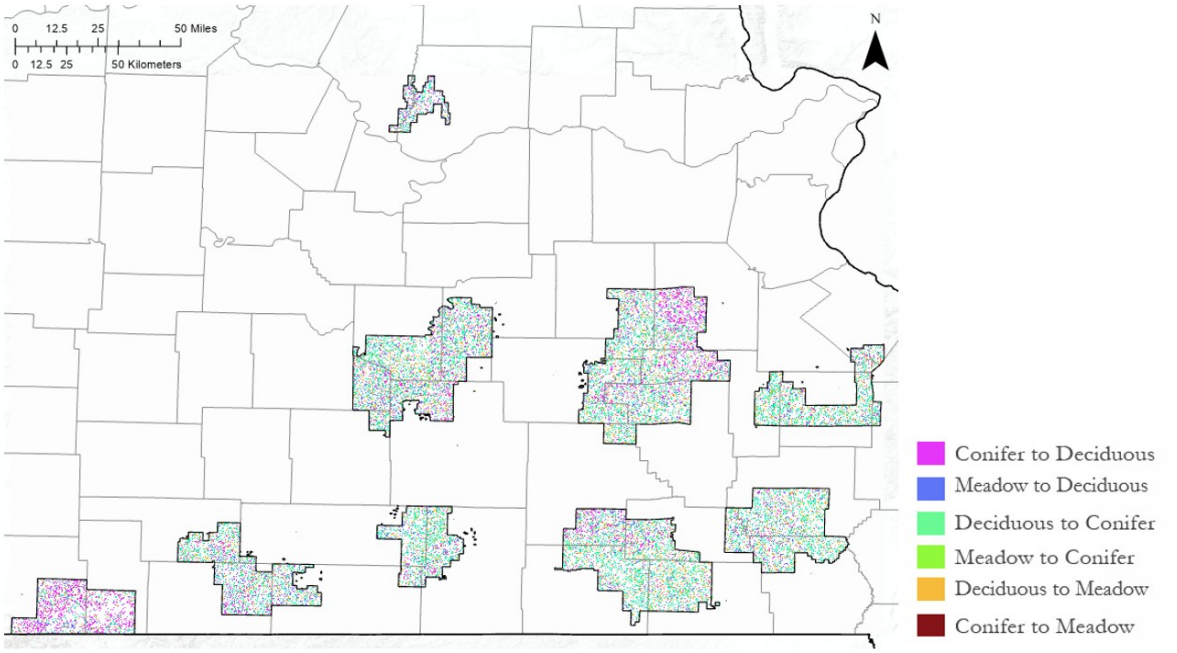

Figure 8. Forecasted land cover change out to 2040

A comparison of land cover change in the Eleven Point and Ava ranger districts, which are actively engaged in restoration projects, are shown in *Figure D5*. This highlights the potential for widespread, scattered conifer regrowth in 2040 if restoration efforts are successful. The results indicate that the current restoration efforts in MTNF have a strong potential to succeed in restoring natural shortleaf pine and glade communities if they are continued into the future. They also indicate the potential that the shortleaf pine could be restored to near historic levels.

4.1.2. Species-level Classification

Preliminary results for the species-level classification and modeling were found through the Forest-Based Classification and Regression tool in ArcGIS Pro. The results for the shortleaf pine species are shown in *Figure 9* and are indicated in green for each figure.

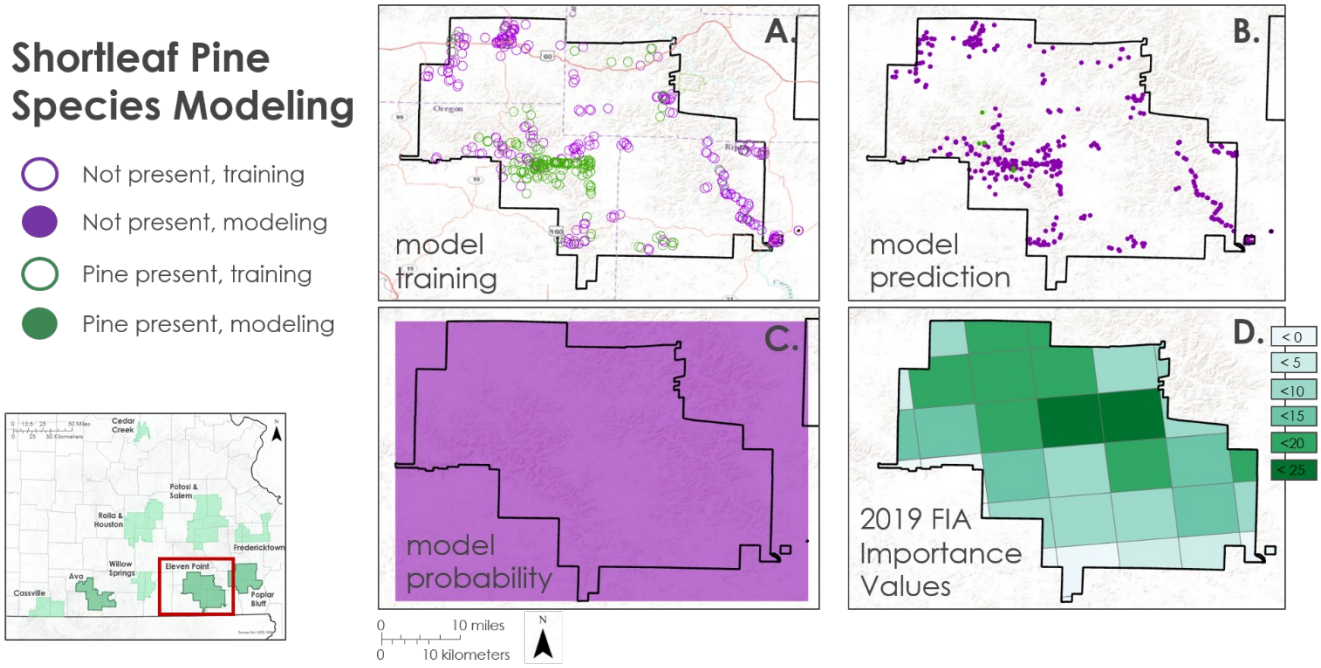

Figure 9. Species-level modeling for shortleaf pine
A) training the model; B) running the model; C) probability from model; D) 2019 FIA Importance Values

To train the model, a binary shapefile was included to mark the presence or absence of shortleaf pine trees: '1' indicated the tree was present and '0' indicated absence (*Figure 9A*). Next, several raster layers were imported to help explain the species' distribution, including 2018 NAIP imagery, conifer distribution from the supervised classification, elevation data, and elevation derivatives. It should be noted that there was an attempt to include 'distance from' rasters (e.g., distance from streams, roads, and glades rasters) and soil-layers, but the rasters did not provide enough variation for the model to incorporate and apply these data.

The results are preliminary and are based on the training points ArcGIS Pro provided during the model training. For shortleaf pine, the model predicted that trees would be found only in the southwest region of the Eleven Point ranger district (*Figure 9B)*, and the overall probability of the trees was '0' for the district (*Figure 9C)*. Again, this can be attributed to the lack of ground-truthed data and limited variability amongst the explanatory variables. *Figure 9D* is data taken from the USFS Forest Inventory Analysis, a data-collection effort that estimates the probability of tree species distribution across the United States. The values within the grids are 'importance values', which in forestry, equates to how dominant a certain tree species type can be in a certain area. These values range from 0 to 300 and are a sum of three variables (relative frequency, relative density, and relative basal area, all expressed as percentages and then added together). Overall, this model indicates that few shortleaf pines will be found in this ranger district. However, this is a weak model and will need greater refinement before moving forward with these results.

Results for the eastern red cedar tree *(Figure 10)* have a similar outcome as the species-level classification for shortleaf pine *(Figure9)*. The cedar trees are marked as orange-red within each of these figures. The results indicated here mirror those we discussed previously; however, the cedar species' model shows an even smaller probability of finding these trees within the Eleven Point district. There are much fewer cedar trees indicated by red points in the 'model prediction' figure than those for the shortleaf pine species (*Figure 10B*).

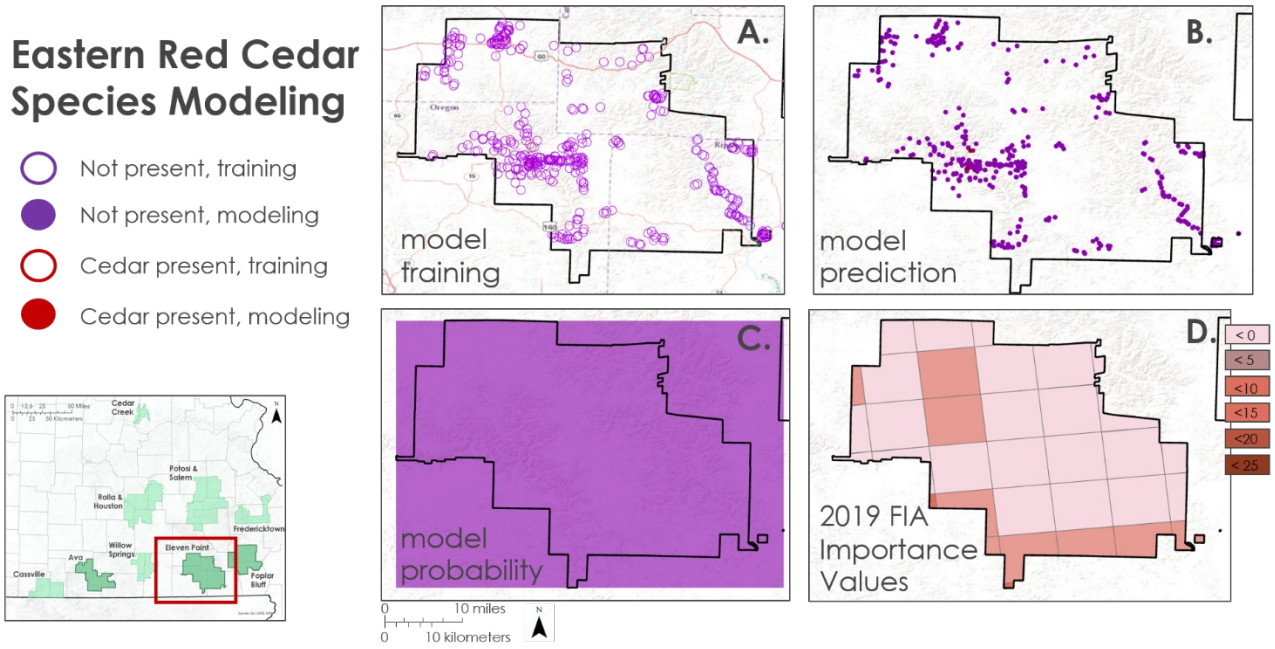

Figure 10. Species-level modeling for eastern red cedar
A) training the model; B) running the model; C) probability from model; D) 2019 FIA Importance Values

Again, this model indicates that few eastern red cedar trees will be found in this ranger district, but it also is a weak model and will need greater refinement *(Figure 10C)*. These results for both shortleaf pine and eastern red cedar were also compared to the USFS's Forest Inventory Analysis (FIA) data (*Figure 10D*). Values were taken from FIA records and used to verify classification and land cover change assessments *(Figure D3)*.

4.2 Future Work

Future work could include a shorter study period for which higher resolution data are available. NAIP imagery is available at 1m spatial resolution starting in 2013. Focusing on the years in which this higher resolution data are available could be highly beneficial and lead to more accurate results. LiDAR has been flown over approximately a quarter of MTNF in recent years and first-return data has been used to study the canopy in select ranger districts. Unfortunately, LiDAR coverage is not comprehensive across MTNF and is only available for smaller-scale studies

within select districts. While LiDAR is not widely available, incorporating it where possible could also improve results for certain areas of the forest.

When completing the supervised classification only five training classes were used. It is possible to expand upon this and break down classes even further. For example, the meadow class could be split further into categories such as grasslands, glades, active agriculture fields, and abandoned agriculture fields. Additional classes could give a more descriptive land cover classification, and could aid in identifying areas in need of restoration.

Business-as-usual modeling based on current land management practices and restoration was the only forecasting model used. It is possible for future work to explore and incorporate other models. This could include modeling for different management strategies or taking into account climate change and disturbances such as flooding, fire, forest clearing, or mining. Comparing forecasts for various management strategies would be very useful in decision-making practices, and accounting for climate change and disturbance could help generate more accurate results.

5. Conclusions

General trends in land cover change between 1986 and 2013 show a decrease in meadow/grasslands, as well as a decrease in conifer forest. In this time period, the MTNF increased in deciduous forest, water, and developed area. The year 2019 was a record-breaking year for floods in Missouri, which may account for the increase in area classified as water. Urban development in the Ozarks is generally increasing, which could account for the increase in developed areas. The increase in deciduous and hardwood forest, accompanied by a decrease in conifers and pine, is consistent with literature regarding forest habitat in the Ozarks.

Forecasting to 2040 shows the highest potential for conifer regrowth in the MTNF given physiographical limitations. This implies that if current management practices continue, conifers in the forest will recover per the goals of the CFLRP. The forecast also shows potential for meadowland revival. In order to conduct reasonably accurate, species-level modeling, imagery with higher temporal resolution and finer spatial resolution would be required.

6. Acknowledgments

The Mark Twain Ecological Forecasting team would like to thank the following project partners and advisors for their invaluable time and assistance to this project:

- US Forest Service, Mark Twain National Forest
  - Kyle Steele, Forest Ecologist
  - Kevin Godsey, Soil Scientist
- US Forest Service, Geospatial Technology and Applications Center
  - Nick Klein-Baer, Remote Sensing Specialist
- Idaho State University, GIS Training & Research Center (GIS TReC)
  - Keith T. Weber, GIS Director
- NASA DEVELOP National Program Fellow
  - Mason Bull

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 NNL16AA05C.

7. Glossary

CFLRP – Collaborative Forest Landscape Restoration Program
DEM – Digital Elevation Model
DOQQ – Digital Orthophoto Quad Quarters
Earth observations – Satellites and sensors that collect information about the Earth's physical, chemical, and biological systems over space and time
EVI – Enhanced Vegetation Index
EROS – Earth Resources Observation and Science
MTNF – Mark Twain National Forest
NAIP – National Agriculture Imagery Program
NDSI – Normalized Difference Snow Index
NDVI – Normalized Difference Vegetation Index
NDWI - Normalized Difference Water Index
OLI – Operational Land Imager
TM – Thematic Mapper
USFS – United States Forest Service
USGS – United States Geological Survey

8. References

Corbett, E. A., & Lashley, A. (2017). Laboratory studies of allelopathic effects of *Juniperus virginiana L.* on five species of native plants. *Oklahoma Native Plant Record*, *17*(1). DOI: 10.22488/okstate.18.100005

Cunningham, R. J. (2007). Historical and social factors affecting pine management in the Ozarks during the late 1800s through 1940. In *Shortleaf pine restoration and ecology in the Ozarks: proceedings of a symposium. USDA Forest Service General Technical Report NRS-P-15* (pp. 1-7).

Dwyer, J. L., Roy, D. P., Sauer, B., Jenkerson, C. B., Zhang, H. K., & Lymburner, L. (2018). Analysis ready data: enabling analysis of the Landsat archive. *Remote Sensing*, 10(9), 1363. DOI: 10.3390/rs10091363

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., ... & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, *186*, 64-87. DOI: 10.1016/j.rse.2016.08.013

Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Hughes, M.J., Laue, B. (2017). Cloud detection algorithm

comparison and validation for operational Landsat data products. *Remote Sensing of Environment, 194*, 379-390. DOI: 10.1016/j.rse.2017.03.026.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, *202*, 18-27. DOI: 10.1016/j.rse.2017.06.031

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425-1432. DOI: 10.1080/01431169608948714

Missouri Spatial Data Information Service (1990). MSDIS/Digital Orthographic Quarter Quad (DOQQ) 1990 [image server]. Retrieved from: http://moimagery.missouri.edu/arcgis/rest/services/MSDIS/DOQQ1990/ImageServer. Accessed August 6, 2020. DOI: 10.5066/F7125QVD

Missouri Spatial Data Information Service (2019). MSDIS/National Agriculture Imagery Program (NAIP) [image server]. Retrieved from: http://moimagery.missouri.edu/arcgis/rest/services/MSDIS/NAIP2018/ImageServer. Accessed August 6, 2020. DOI: 10.5066/F7QN651G

Moore, R. (2005). *2005 Land and Resource Management Plan (2005 Forest Plan)*. Retrieved from: https://www.fs.usda.gov/main/mtnf/landmanagement/planning

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*(1), 217-222. DOI: 10.1080/01431160412331269698

Shugart, H. H., Asner, G. P., Fischer, R., Huth, A., Knapp, N., Le Toan, T., & Shuman, J. K. (2015). Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models. *Frontiers in Ecology and the Environment*, *13*(9), 503-511. DOI: 10.1890/140327

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., & Macomber, S. A. (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects?. *Remote sensing of Environment*, 75(2), 230-244. DOI: 10.1016/S0034-4257(00)00169-3

USDA Forest Service (2020a). *Forest's Collaborative Forest Landscape Restoration Project 2012 – 2022*. Retrieved from https://www.fs.usda.gov/detail/mtnf/landmanagement/?cid=stelprdb5423048. Accessed August 4, 2020.

USDA Forest Service (2020b). *Glade Top Trail National Scenic Byway*. Retrieved from https://www.fs.usda.gov/detail/mtnf/specialplaces/?cid=fseprd514266. Accessed August 4, 2020.

United States Geologic Survey (USGS) (1999). National Elevation Dataset [data]. Retrieved from: https://viewer.nationalmap.gov/basic/. Accessed August 6, 2020. DOI: 10.3133/fs14899

USGS (2020a). USGS EROS Archive – Landsat Archives – Landsat 4-5 TM level-1 Data Products [data]. Retrieved from https://www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-archives-landsat-4-5-thematic-mapper-tm-level-1-data?qt-science_center_objects=0#qt-science_center_objects. Accessed August 6, 2020. DOI: 10.5066/F7N015TQ

USGS (2020b). USGS EROS Archive – Landsat Archives – Landsat 8 OLI and TIRS Level-1 Data Products [data]. Retrieved from https://www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-archives-landsat-8-oli-operational-land-imager-and?qt-science_center_objects=0#qt-science_center_objects. Accessed August 6, 2020. DOI: 10.5066/F71835S6

USGS. National Elevation Data [data].

9. Appendices

Appendix A

Table A1

Detailed description of filters used to acquire leaf-off Landsat data from Google Earth Engine

Years	Winter Start	Winter End	Winter Cloud Cover Filter
1986-1987	12/1/1986	2/28/1987	30%
1991-1992	10/1/1991	2/28/1992	40%
1996-1997	10/1/1996	2/28/1997	10%
2001-2002	12/1/2001	2/28/2002	25%
2006-2007	12/1/2006	2/28/2007	10%
2010-2011	10/1/2010	2/28/2011	10%
2016-2017	12/1/2016	2/28/2017	10%
2019-2020	10/1/2019	2/28/2020	10%

Table A2

Detailed description of filters used to acquire leaf-on Landsat data from Google Earth Engine

Year	Summer Start	Summer End	Summer Cloud Cover Filter
1986	06/01/1986	08/30/1986	10%
1991	06/01/1991	08/30/1991	10%
1996	06/01/1996	08/30/1996	25%
2001	06/01/2001	08/30/2001	15%
2006	06/01/2006	08/30/2006	10%
2010	06/01/2010	08/30/2010	10%
2016	06/01/2016	08/30/2016	10%
2019	05/01/2019	08/30/2019	25%

Appendix B

These indices were selected based on their common application in previous land cover classification studies and utility in spectrally separating seasonally affected vegetation.

NDSI is the ratio between the green and SWIR spectral bands and is used to identify areas with snow cover. Index values range from –1 to 1, with negative values indicating little to no snow cover and positive values indicating a high probability of snow cover.

(1)

$$NDSI = \frac{(¿ - SWIR)}{(¿ + SWIR)}$$

Landsat 5:

(1a)

$$NDSI = \frac{TM_{band\,2} - TM_{band\,5}}{TM_{band\,2} + TM_{band\,5}}$$

Landsat 8:

(1b)

$$NDSI = \frac{OLI_{band\,3} - OLI_{band\,6}}{OLI_{band\,3} + OLI_{band\,6}}$$

NDVI is the ratio between the red and NIR values, which are found on different bands between Landsat 5 TM and Landsat 8 OLI (*Equation B2*). The range of index values is –1 to 1, with negative values indicating less green vegetation and positive values indicating more green vegetation.

(2)

$$NDVI = \frac{(NIR - ¿)}{(NIR + ¿)}$$

Landsat 5:

(2a)

$$NDVI = \frac{TM_{band\,4} - TM_{band\,3}}{TM_{band\,4} + TM_{band\,3}}$$

Landsat 8:

(2b)

$$NDVI = \frac{OLI_{band\,5} - OLI_{band\,4}}{OLI_{band\,5} + OLI_{band\,4}}$$

Akin to NDVI, EVI also uses a ratio of the R and NIR bands to assess vegetation greenness but incorporates several other values, listed with *Equation 3*. The range

of index values is –1 to 1, with negative values indicating less green vegetation and positive values indicating more green vegetation (*Equation 3*).

(3)

$$EVI = G * \left(\frac{(NIR - ¿)}{(NIR + C_1 * ¿ - C_2 * ¿ + L)} \right)$$

$Gain\ Factor\ (G) = 2.5$
$Coefficient\ 1\ of\ aerosol\ resistance\ term\ (C_1) = 6$
$Coefficient\ 1\ of\ aerosol\ resistance\ term\ (C_2) = 7.5$
$Canopy\ Background\ (L) = 1$

Landsat 5:

(3a)

$$EVI = 2.5 * \left(\frac{(TM_{band\,4} - TM_{band\,3})}{(TM_{band\,4} + 6 * TM_{band\,3} - 7.5 * TM_{band\,1} + 1)} \right)$$

Landsat 8:

(3b)

$$EVI = 2.5 * \left(\frac{(OLI_{band\,5} - OLI_{band\,4})}{(OLI_{band\,5} + 6 * OLI_{band\,4} - 7.5 * OLI_{band\,2} + 1)} \right)$$

NDWI ranges from –1 to 1 and indicates the presence of water bodies within a scene. Note that there is another index termed the Normalized Difference Water Index, but is a ratio of the NIR and SWIR spectral bands and looks at the water content of leaves (*Equation 4*).

(4)

$$NDWI = \frac{(¿ - NIR)}{(¿ + NIR)}$$

Landsat 5:

(4a)

$$NDWI = \frac{TM_{band\,2} - TM_{band\,4}}{TM_{band\,2} + TM_{band\,4}}$$

Landsat 8:

(4b)

$$NDWI = \frac{OLI_{band\,3} - OLI_{band\,5}}{OLI_{band\,3} + OLI_{band\,5}}$$

Appendix C

Table C1
Confusion matrix and Kappa Statistic for 1986

		Validation Data						
		Water	**Developed**	**Deciduous Forest**	**Conifer Forest**	**Meadow**	**Total**	
Classified Data	**Water**	72	0	0	0	0	72	
	Developed	3	55	0	0	6	64	
	Deciduous Forest	1	0	94	5	2	102	
	Conifer Forest	1	0	20	89	0	110	
	Meadow	1	11	3	0	137	152	
	Total	78	66	117	94	145	500	
								K = 0.86469

Table C2
Confusion matrix and Kappa Statistic for 2019

		Validation Data						
		Water	**Developed**	**Deciduous Forest**	**Conifer Forest**	**Meadow**	**Total**	
Classified Data	**Water**	40	2	1	1	2	46	
	Developed	3	37	4	2	17	63	
	Deciduous Forest	0	2	62	13	6	83	
	Conifer Forest	0	0	0	132	0	132	
	Meadow	2	7	4	2	135	150	
	Total	45	48	71	150	160	474	
								K = 0.810276

Appendix D

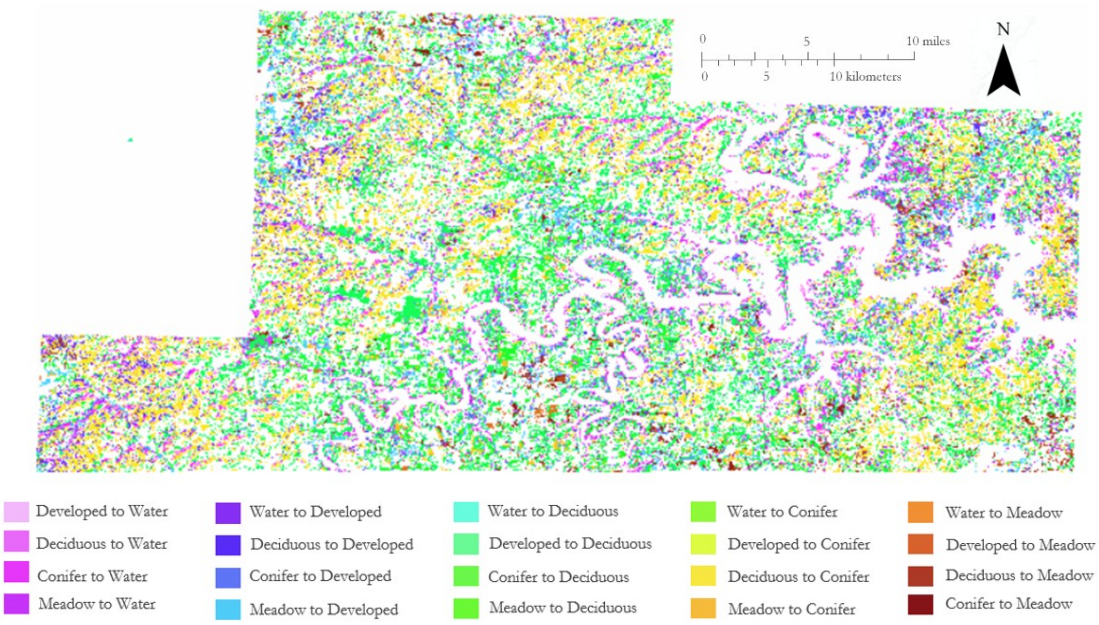

Figure D1. Land cover change analysis for Cassville ranger district, 2006-2019

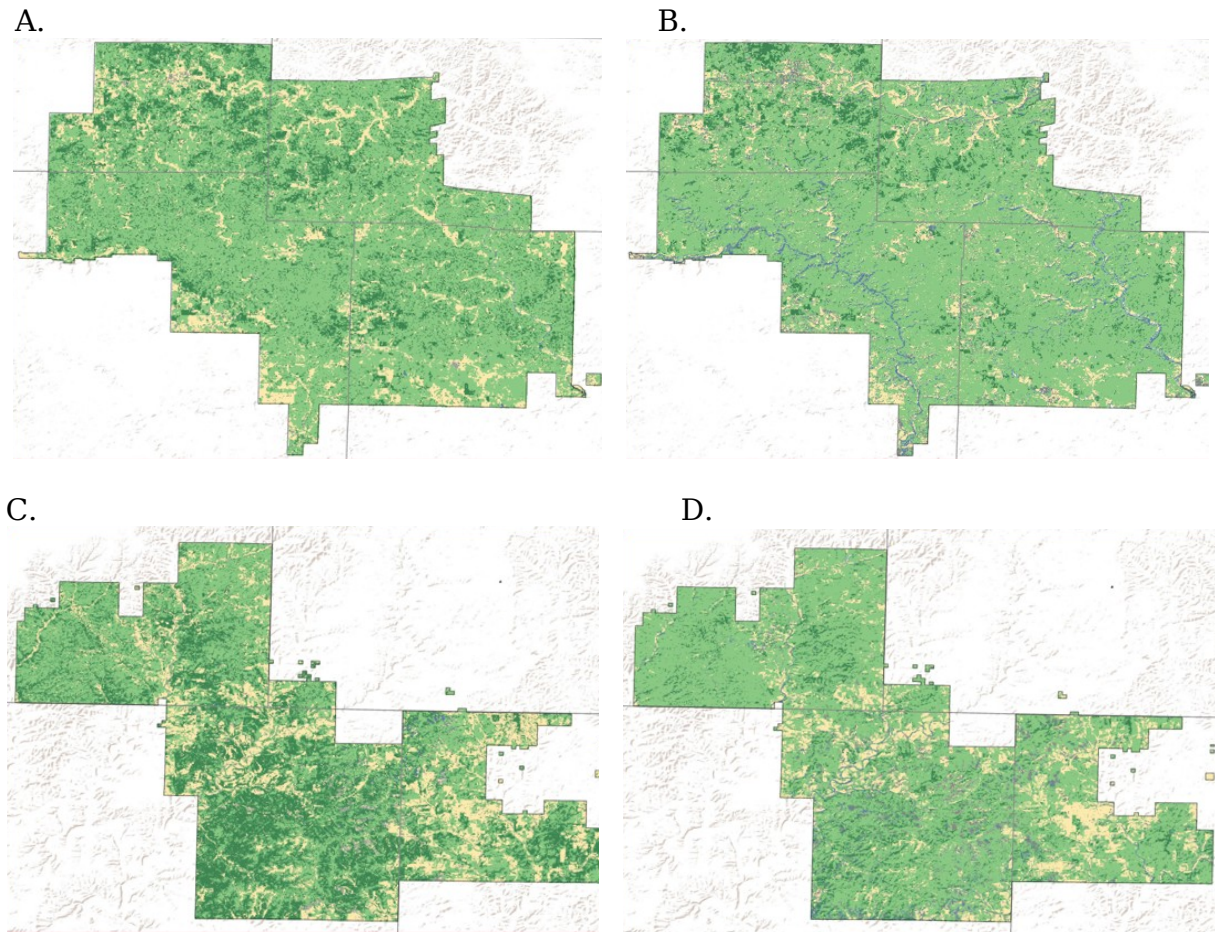

Figure D2. Land cover change for two ranger districts involved in restoration efforts
A) Eleven Point, 1986; B) Eleven Point, 2019; C) Ava, 1986; D) Ava, 2019

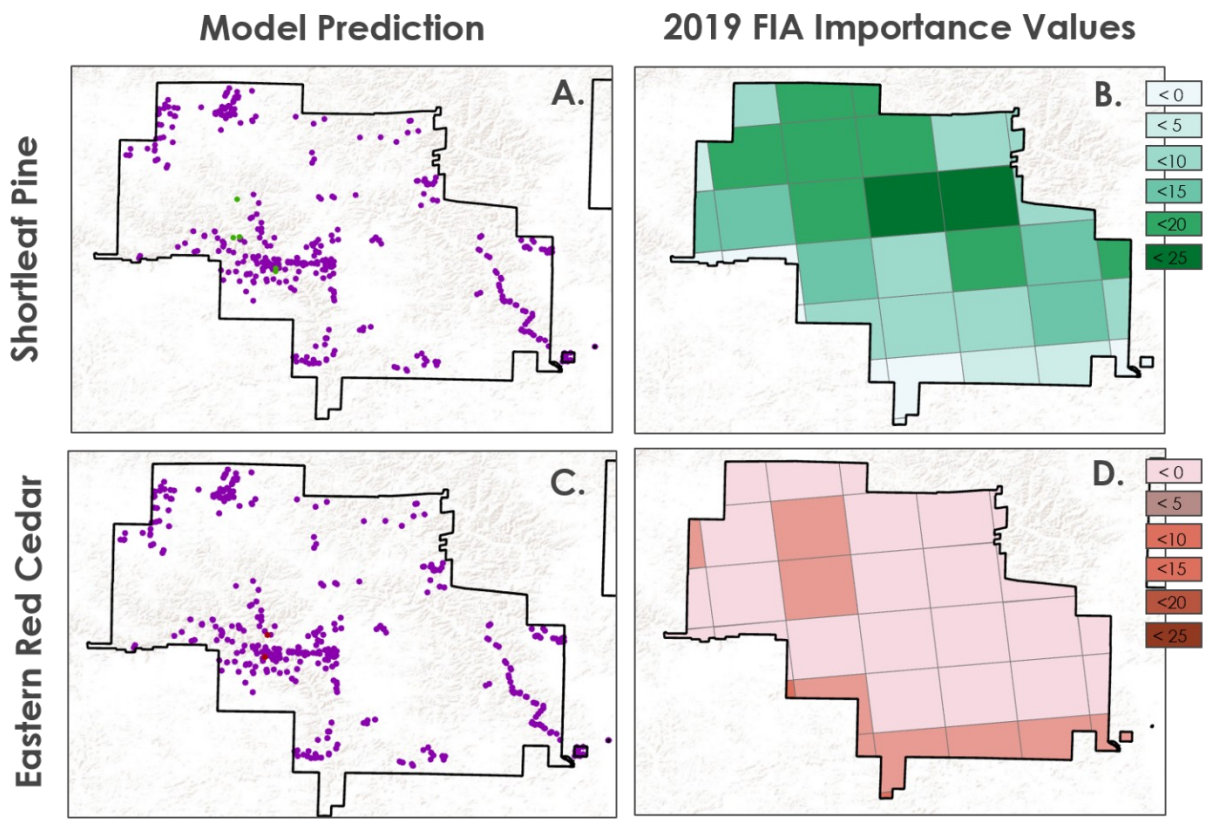

Figure D3. Species-level modeling for eastern red cedar and short leaf pine, compared to 2019 USFS FIA data
A) Predicted shortleaf pine; B) FIA predicted shortleaf pine; C) predicted eastern red cedar; D) FIA predicted red cedar

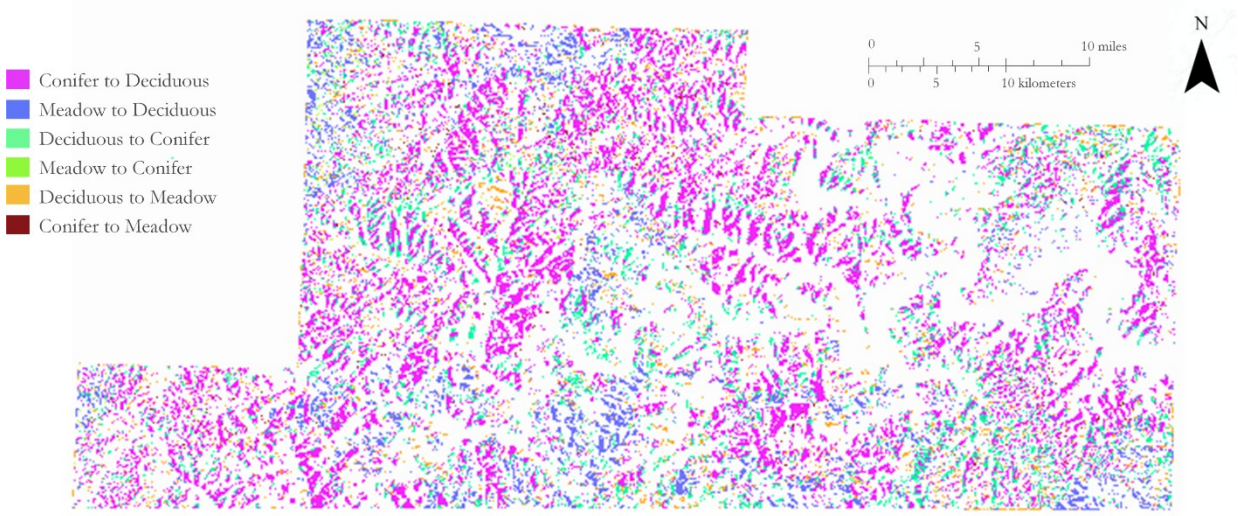

Figure D4. Forecasted land cover change analysis for Cassville ranger district, 2019-2040

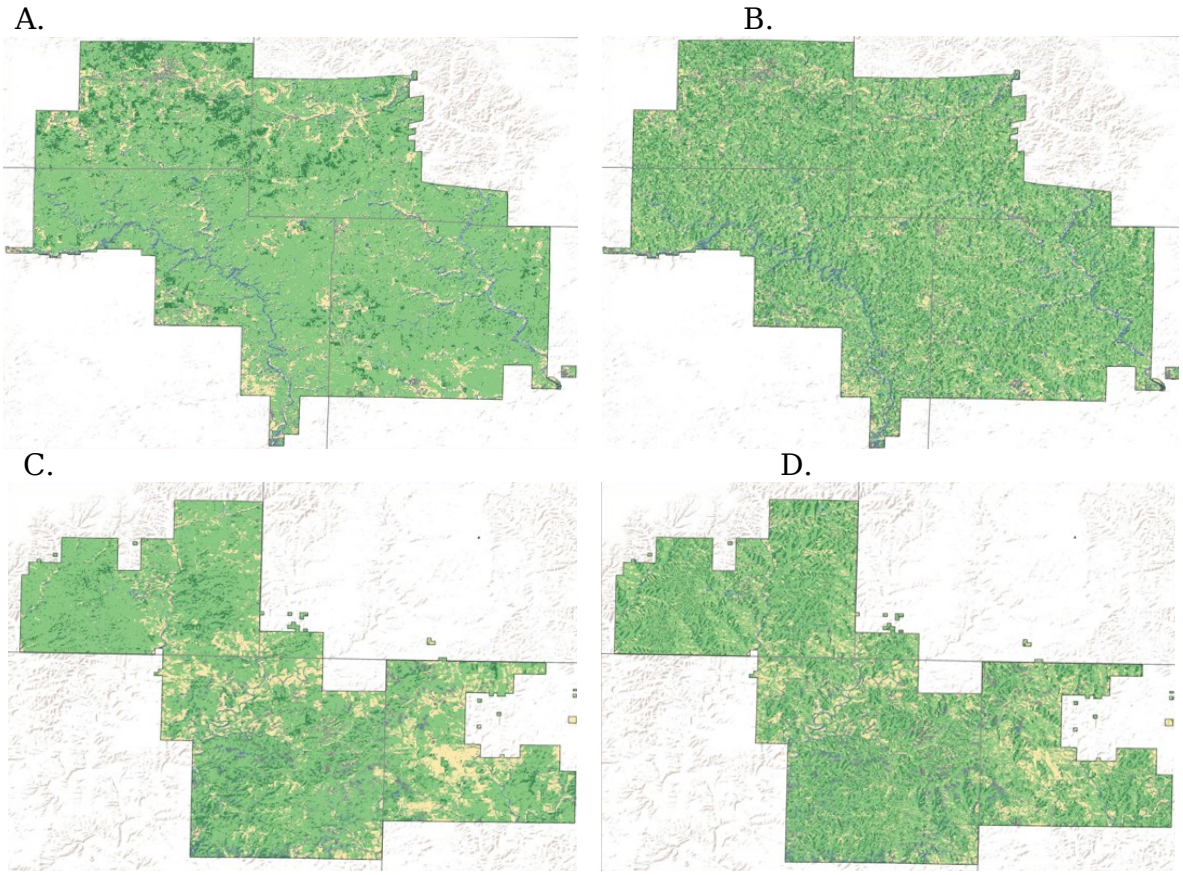

Figure D5. Forecasted land cover change for two ranger districts involved in restoration efforts
A) Eleven Point, 2019; B) Eleven Point, 2040; C) Ava, 2019; D) Ava, 2040