Fisher’s Peak Ecological Forecasting

Mapping Biomass to Inform Conservation Planning of a Future State Park in Southern Colorado

 **Technical Report**

Final Draft – August 6th, 2020

Lauren Lad (Project Lead)

Scott Cunningham

Laura Krauser

Darby Levin

**Advisors:**

Dr. Paul Evangelista (Colorado State University, Natural Resource Ecology Laboratory)

Dr. Catherine Jarnevich (United States Geological Survey, Fort Collins Science Center)

Peder Engelstad (Colorado State University, Natural Resource Ecology Laboratory)

Nicholas Young (Colorado State University, Natural Resource Ecology Laboratory)

Dr. Tony Vorster (Colorado State University, Natural Resource Ecology Laboratory)

# 1. Abstract

Fisher’s Peak is a 77.5 km2 property southeast of Trinidad, Colorado that is planned to become Colorado’s newest state park. The area has experienced limited anthropogenic disturbance and is home to an abundance of unique habitats and species. A rapid, approximately 900 m change in elevation over the extent of the area nurtures a variety of plants and animals, including the endangered New Mexico meadow jumping mouse. In 2019, the State of Colorado obtained Fisher’s Peak with plans to make it Colorado’s second largest state park. A diverse group of collaborators, including the Colorado State Forest Service and The Nature Conservancy, worked closely to design the state park to maximize recreation opportunity while conserving the property’s rich habitats and biodiversity. The Fisher’s Peak Ecological Forecasting Team utilized Light Detection and Ranging (LiDAR) surveys, *in situ* forest inventory data, and Earth observations from Landsat 8 Operational Land Imager (OLI), Sentinel-1 C-band Synthetic Aperture Radar (C-SAR), Sentinel-2 Multispectral Instrument (MSI), Advanced Land Observing Satellite 2 (ALOS-2) Phased Array type L-band Synthetic Aperture Radar (PALSAR-2) and the Shuttle Radar Topography Mission (SRTM) to quantify and map biomass over the extent of the study area. The results from modeling biomass had an out-of-bag root mean square error of 55 Mg/ha and an R2 of 12. The resulting map indicates areas where carbon storage on the property is high, informing decision-making processes for future park development. While more *in situ* training data may improve modeling capacity for biomass in the Fisher’s Peak area, this work represents a feasible attempt to better understand biomass distribution using earth observation.

**Key Terms**

LiDAR, biomass, optical remote sensing, Landsat 8 OLI, Sentinel-1, carbon sequestration

# 2. Introduction

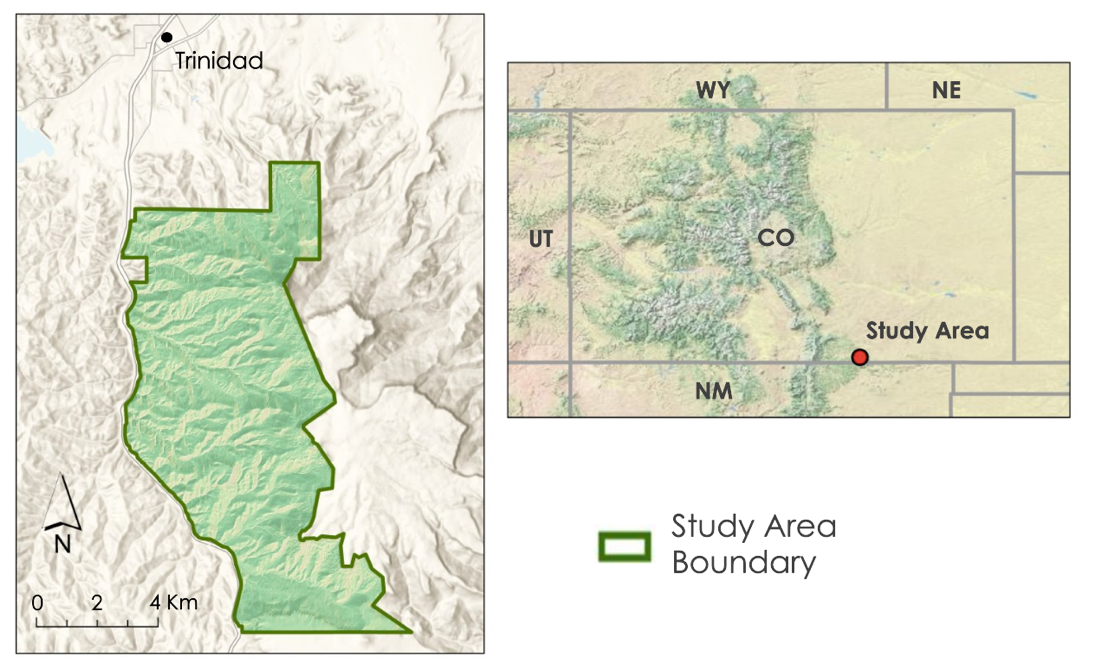
***2.1 Background Information***

Monitoring vegetation has been a core application of remotely-sensed imagery for decades. In particular, quantifying above-ground biomass (AGB) via optical and active remote sensing data has many advantages over traditional field-based AGB estimation methods, including the capability of estimating biomass at various spatial scales (Timothy et al., 2015). For the purpose of this paper, subsequent references to biomass will refer to live above-ground forest biomass. Traditional methods of estimating biomass require time and labor-intensive efforts that estimate the carbon stock of individual sites of vegetation. The use of remotely-sensed data to streamline biomass estimation offers an efficient way to support carbon sequestration efforts. Accurate, regional representation of biomass distribution is increasingly valuable as the importance of understanding carbon emission and sequestration creates potential revenue generation through the carbon market (Foody et al., 2003). The carbon market works on a system of credits and balances emissions standards for industry with financial support for conservation in the form of carbon sequestration.

A number of machine learning methods have been developed to estimate biomass using algorithms that create flexible, nonparametric comparisons of remotely-sensed data and biomass measurements (Gleason & Im, 2012; Wang et al., 2016). The Random Forest (RF) algorithm is a precise predictive method for classification and regression that has proven useful for ecological studies (Cutler et al., 2007). RF has been used to estimate biomass across a variety of vegetation types and applications, from agriculture (Wang et al., 2016) to tropical forests (Dang et al., 2019; Gleason & Im, 2012) to montane regions (Nandy et al., 2019). RF regression operates as an ensemble, nonparametric modeling approach using an aggregate of regression trees. Predictions result from a combination of independent, “bootstrap” samples of data filtered through a “forest” of decision trees (Breiman, 2001). RF has proven to be an effective tool because it automates and randomizes variable selection, requires few model input parameters, and generates results that can be used to estimate variable importance (Breiman, 2001; Stevens et al., 2015). Our study supports the literature with the application and discussion of RF predictive power for a smaller area with a unique landscape on which to test the limits of the established methods.

While biomass estimation has been explored through remote sensing analyses at various scales for its role in carbon stocks, estimating biomass in heterogeneously vegetated areas presents many challenges. Fisher’s Peak State Park will be Colorado’s second largest state park and is envisioned to grow the recreation economy of the nearby City of Trinidad and provide jobs for residents (Figure 1). The state park will cover over 77 km2 of diverse ecological communities, with the varying elevations and tree species providing important habitat for animal, plant, and insect populations in the prairies and mountains. The protection of Fisher’s Peak’s diverse habitats and the species they support is a major priority of the park’s development plan. The Nature Conservancy (TNC), the Colorado State Forest Service (CSFS), and Colorado Parks and Wildlife (CPW) hope to understand the feasibility of using remotely-sensed biomass estimations in conjunction with previously collected forest inventory data from the field and high-resolution radar for areas like Fisher's Peak State Park. Project partners conducted a forest inventory assessment and LiDAR analysis of the area in the summer of 2019. This project’s study period ran from June through August of 2019. The study area is undergoing a comprehensive inventory evaluation, so the full extent of the property’s biomass, fuel loads, and carbon storage is not yet quantified.

The area has evidence of historic logging and grazing disturbance over the past five decades, indicating a distinct land management history, oscillating between economic productivity and conservation. The property’s namesake comes from the iconic Fisher’s Peak, a nearby flat-topped mountain. This montane formation is located on traditional lands of Jicarilla Apache, Pueblo, Comanche, and Ute nations, and has long been a notable landmark of southeast Colorado. A steep 3,000-foot elevation gradient from west to east facilitates high levels of ecological diversity: tree species transition between piñon pine and junipers at low elevations, Gambel oak and ponderosa pine as elevation increases, and a mix of fir and pines as the elevation peaks.

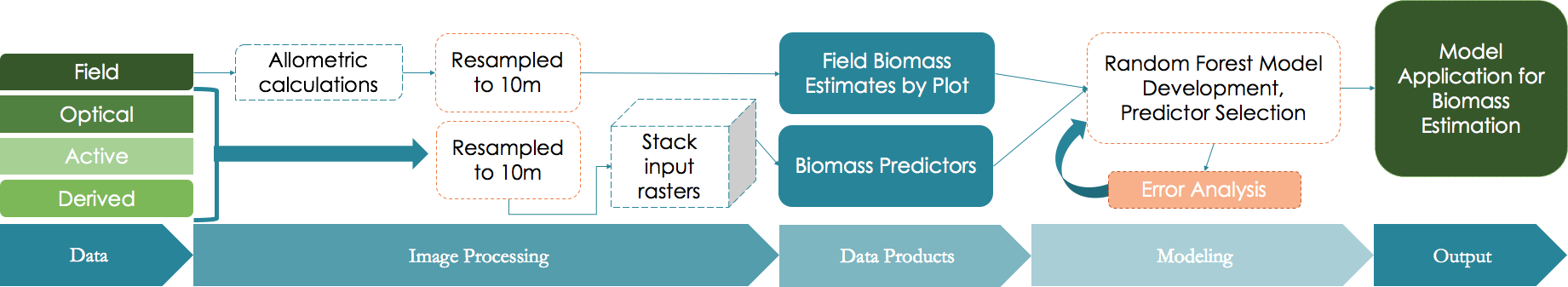


*Figure 1.*The location of Fisher’s Peak State Park in relation to Trinidad, CO.

***2.2 Project Partners & Objectives***

Conducted in collaboration with the TNC and the CSFS, this project employed remote sensing-based biomass estimation methodologies to quantify biomass of the heterogeneous landscape that will make up Fisher’s Peak State Park. The TNC furthers its mission of conserving lands and waters by purchasing high conservation-value land. The TNC and the CSFS partnered with NASA DEVELOP to conduct a comprehensive inventory of the property’s biomass assets. This inventory will add to the park’s planning and development resources, helping to strike a balance between recreation and conservation by highlighting areas of high biomass and carbon storage. The TNC is also exploring entering other properties into the Colorado carbon market. The carbon market offers financial support by facilitating the sale of emission offsets in the form of stored carbon stocks (biomass) to companies seeking to comply with regulatory emission standards. Additionally, *in situ* field methods for estimating biomass are expensive and time-consuming. A remote sensing-based methodology has the potential to make estimating biomass (and therefore carbon) more economically feasible while aiding TNC’s conservation mission.

The primary objective of this project was to map biomass across the Fisher’s Peak property to inform park development. To do this, the team developed a RF regression model that combined *in situ* forest inventory data, high resolution LiDAR, and passive remote sensing imagery to quantify biomass at Fisher’s Peak. The team used the RF model to create a map of biomass estimates over the extent of the study area. This map helped identify areas of high carbon storage and fuel loads as well as the priority for preservation. Lastly, the methodology used by the team provided a framework for future biomass evaluation efforts (Figure 2).



*Figure 2.* Flow-chart depicting the methods used to generate training and predictor data for the random forest (RF) regression model of biomass estimations.

# 3. Methodology

***3.1 Data Acquisition***

Project partners at the CSFS collected forest inventory data from June 11 to July 11, 2019 across over 80 plots (Appendix A). These forest inventory data were used to derive ground truth biomass estimates for our training and testing data. The CSFS also provided a LiDAR canopy height map, created by differencing a Digital Terrain Model (DTM) and a Digital Surface Model (DSM). Landsat 8 Operational Land Imager (OLI) WRS (Path/Row 33/34) and Sentinel-2 Multispectral Instrument (MSI) (MGRS Tile 13SEB) data were downloaded through the Google Earth Engine (GEE) API (Table 1).

**Table 1**

*The table below showcases the satellite data used in this project. All data were accessed and processed in GEE.*

|  |  |  |
| --- | --- | --- |
| **Platform and Sensor** | **Data Product** | **Dates** |
| ALOS-2 PALSAR-2 | L-band | 2017 |
| Landsat 8 OLI | Top of Atmosphere Tier 1 | July 9, 2019 |
| Sentinel-1 SAR | C-band | June – August 2019 |
| Sentinel-2 MSI | Level 1C – Top of Atmosphere | June 8, 2019 |
| LiDAR | Canopy height | June 2019 |
| Shuttle Radar Topography Mission (SRTM) | Slope, Elevation, Northness, Eastness | February 11-22, 2000 |

***3.2 Data Processing***

*3.2.1 In situ data*

The forest inventory *in situ* data included 84 plots, the species of each tree in a plot, and a measurement of tree diameter at breast height (DBH) for each tree. To develop training data values, a Python script was used to automate the calculation of biomass based on the field-collected DBH. The allometric equation coefficients for calculating biomass based on tree species were taken from Jenkins et al., 2003 (Appendix A). First, the team calculated the biomass of each tree within a prism-based variable radius plot. Each variable radius plot had a basal area factor (BAF) of 10. Then, the team summed the values of each individual tree biomass. Biomass for saplings (any tree with a DBH ≥ 5 in) was also calculated on a plot by plot basis. Biomass values for mature trees and saplings were resampled separately, as sapling samples were taken using a fixed radius plot of 13.5m. To assign biomass values to a 10 m cell (spatial resolution of the predictor variables), each plot’s radius factor was multiplied by the DBH of the largest tree in the plot to find the horizontal limiting distance (HLD) (Powell 2014). HLD is equivalent to the radius of the plot and was used to find the area of each plot. This area was used to derive a resampled mature tree biomass value for each 10 m training pixel, with the cell focal points matched to coordinates of the forest inventory plots. Sapling biomass values were separately resampled to this standard 10m training pixel size then added to the mature tree biomass value. This summed value was the total biomass content for each training pixel. An example biomass calculation for a Douglas fir dominated plot can be seen in Appendix A. A summary of training data distribution can be seen in Table 2 below. A breakdown of biomass distribution by dominant tree species can be found in Appendix C.

**Table 2**

*An assessment of the distribution of training data, excluding the removed plot 285. This table was used to evaluate the efficacy of our training data, and as a reference for the RF model RMSE.*

|  |  |
| --- | --- |
| **Training Data Breakdown (Mg/ha)** | |
| **Average Biomass Value** | 79.17 |
| **Maximum Value** | 265.31 |
| **Minimum Value** | 9.86 |
| **Standard Deviation** | 57.67 |

*3.2.2 Passive Remote Sensing Data*

Using Google Earth Engine (GEE), Landsat 8 OLI (Path/Row 33/34) and Sentinel-2 MSI (MGRS Tile 13SEB) top of atmosphere (TOA) data were filtered to the study area for June and July of 2019. A single image was selected for both Landsat 8 OLI (July 9, 2019) and Sentinel-2 MSI (June 8, 2019). The images were chosen based on their lack of clouds, complete coverage of the study area, and proximity to the dates of field data collection. Several spectral indices shown to be useful for predicting biomass were calculated from each image (Adame-Campos et al., 2019; Powell et al., 2009; Hyde et al., 2006). Normalized Difference Vegetation Index (NDVI) and Tasseled cap indices were chosen because of their demonstrated usefulness in previous studies (Powell et al., 2009; Hyde et al., 2006). Three tasseled cap transformations (brightness, greenness, and wetness) were computed for each image using empirically-derived coefficients specific to each sensor (Baig et al., 2014; Nedkov, 2017) (Appendix B Table 1, 2). Remote sensing of ecological phenomena regularly incorporates these indices for summarizing and distinguishing vegetation qualities on the landscape. Since Sentinel-2 has bands with varying pixel sizes, these rasters were resampled to 10m spatial resolution before calculating the indices. For sensors with uniform spatial resolutions, indices were calculated prior to resampling. All bands were resampled by exporting them from GEE as assets, while specifying the scale as 10m and the projection as WGS84/UTM Zone 13. All bands were resampled to Sentinel-2's 10m resolution in order to account for fine-scale shifts in biomass across a highly heterogenous landscape.

*3.2.3 Active Remote Sensing Data*

All active remote sensing datasets, except for the aerial LiDAR canopy height, were acquired using GEE. Aspect northness and eastness, elevation, and slope layers were calculated from the SRTM digital elevation model. The Advanced Land Observing Satellite 2 (ALOS-2) Global Yearly Phased Array type L-band Synthetic Aperture Radar (PALSAR)/PALSAR-2 image collection was filtered to an image from 2017, the most recent year available. The HH and HV bands were selected, and a ratio of the HH/HV band was calculated. Additionally, the Sentinel-1 C-band SAR dataset was filtered for the summer of 2019, 10m resolution, and Interferometric Wide Swath image mode. VH and VV bands were selected for use as predictor variables. The LiDAR-derived canopy height layer was acquired from the CSFS and imported to GEE as an asset. All active remote sensing layers were clipped to the Fisher’s Peak study region and resampled as necessary to 10m resolution, using the same methodology as for the passive remote sensing data.

*3.3 Model Development*

The team used GEE to extract the values of all predictor variables at the location of each forest inventory plot. These values, as well as those from the pixel level biomass values, were used as training data in a RF regression model in RStudio using the randomForest library. For all model runs, the team utilized 5,000 decision trees and 2 variables for consideration at each node. Determining the number of decision trees requires an analysis of the asymptotic relationship between the number of trees and mean-squared error associated with those numbers. The number of variables for consideration at each node was optimized in RF model by generating out-of-bag error metrics for varying numbers of variable predictors.

In order to evaluate the best variables for model selection, the team conducted three model comparisons. The aim of these comparisons was to determine both the best combination of variables, as well as the best combination of variable types. These types were optical, topographic, radar, and LiDAR. Prior to the development of each model, we conducted a predictor variable correlation analysis. First, we examined the correlations between our *in situ* forest inventory data. Next, we ranked predictor variables based on the highest correlation with *in situ* data. Lastly, we examined the correlations between predictor variables to eliminate redundancy. The correlation threshold for eliminating variables was |0.7|. Using this predictor variable correlation analysis, the team narrowed each model comparison run to eight variables. A run was not conducted using Sentinel-2 MSI optical data because all variable correlations were above the allowed threshold of |0.7|. The descriptions of each model used for comparison, as well as the out-of-bag (OOB) accuracy metrics, are described in Table 3. Each model used several Landsat 8 OLI variables and several topographic variables, while some also used radar and/or LiDAR variables. After narrowing to eight predictor variables, we utilized the highest performing variables for a final model. The best performing model included four predictor variables including LS Wetness, SRTM Northness, LS Greenness, SRTM Elevation.

# 4. Results & Discussion

***4.1 Analysis of Results***

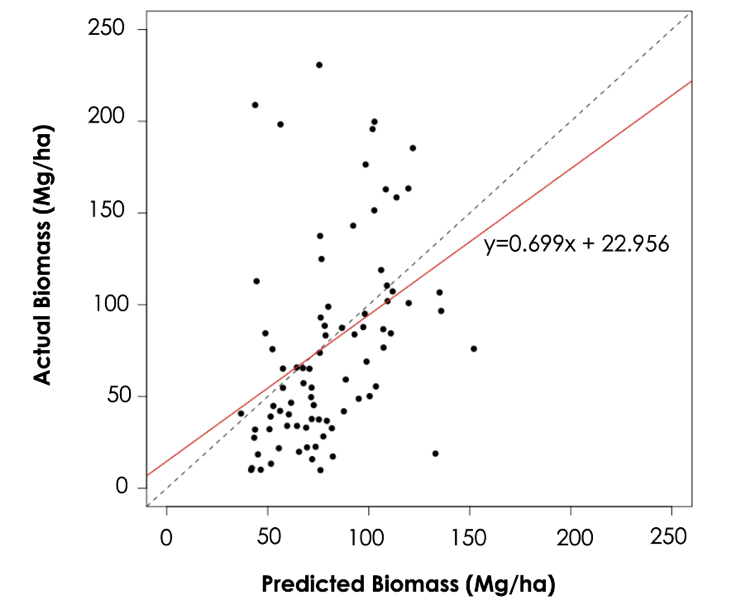
The highest performing model, after iterative and comparative analyses, retained predictor variables of Landsat Wetness, SRTM Northness, Landsat Greenness, SRTM elevation, Landsat Brightness, Sentinel VH, SRTM Eastness, and ALOS-2 PALSAR-2 HH band, listed in order of importance. This list was narrowed to the four most important variables for our final run. This final run utilized Landsat 8 Wetness, SRTM Northness, Landsat Greenness, and SRTM elevation.

**Table 3**

*A model comparison was conducted to decide on the final list of predictor variables for our biomass model. Sentinel-2 MSI data were not used for model comparisons as the variables were too highly correlated. The optimal set of predictor variables was a combination of topographic variables and Landsat 8 optical data.*

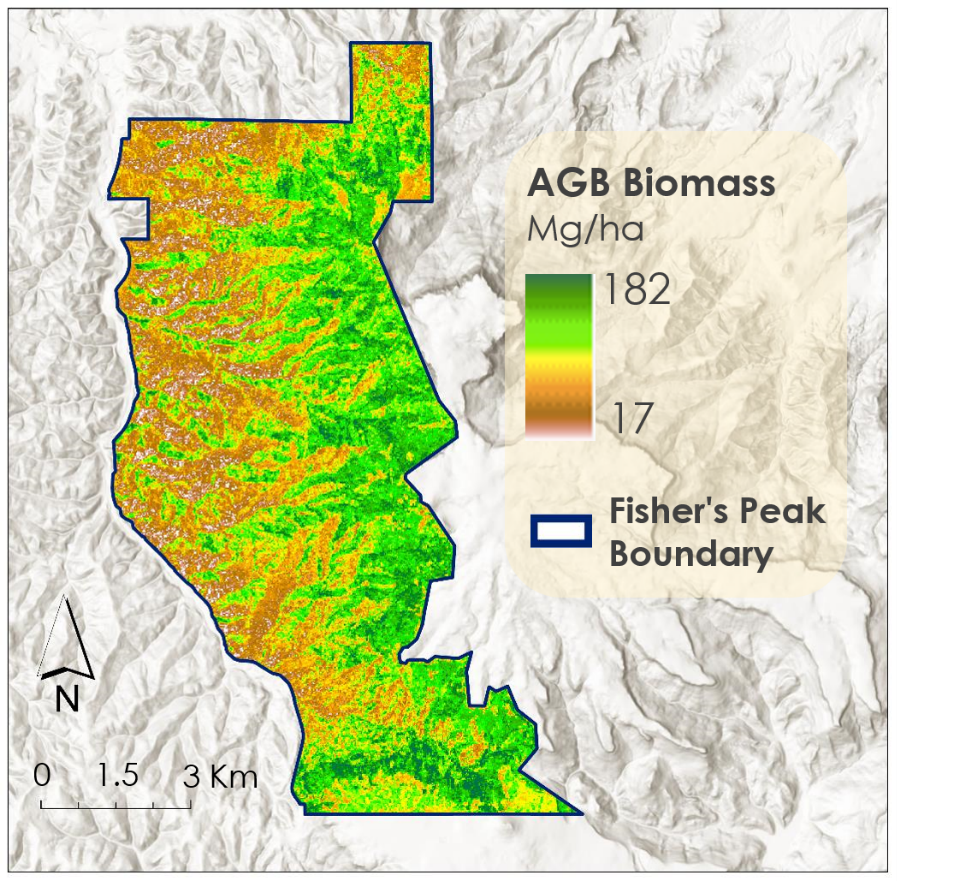
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model Name** | **Satellite/Sensors** | **Predictor Variables** | **OOB RMSE**  **(Mg/ha)** | **OOB R2** | **Final Model? (Y/N)** |
| M1 | Landsat 8 OLI (LS), SRTM | LS Wetness, SRTM Northness, LS Greenness, SRTM Elevation, LS Brightness, SRTM Eastness, LS NIR, SRTM Slope | 55.24 | 7.01 | N |
| M2 | LS 8 OLI, SRTM, LiDAR | LS Wetness, SRTM Northness, LS Greenness, LiDAR Canopy Height, SRTM Elevation, LS Brightness, SRTM Eastness and Slope | 55.77 | 5.23 | N |
| M3 | LS 8 OLI, SRTM, ALOS-2 PALSAR-2, Sentinel-1 C-band | LS Wetness, SRTM Northness, LS Greenness, SRTM Elevation, LS Brightness, Sentinel VH, SRTM Eastness ALOS-2 PALSAR-2 HH | 54.98 | 7.46 | N |
| M4 | LS 8 OLI, SRTM | LS Wetness, SRTM Northness, LS Greenness, SRTM Elevation | 54.48 | 9.56 | Y |

M4 was selected as our final model based on its OOB RMSE (54.48 Mg/ha) and R2 (9.56) values. An additional evaluation was conducted through a comparison of the model’s predicted and actual values, shown in Figure 5.

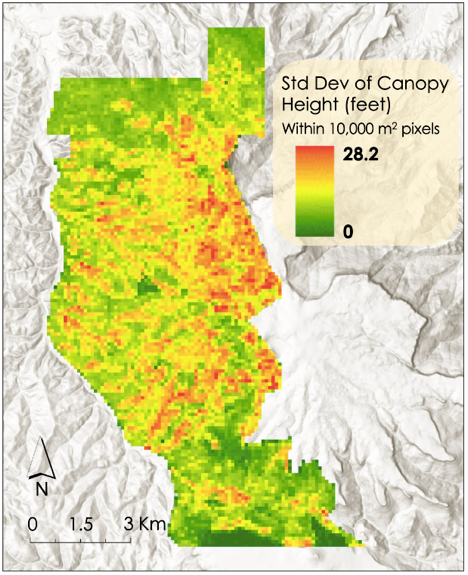
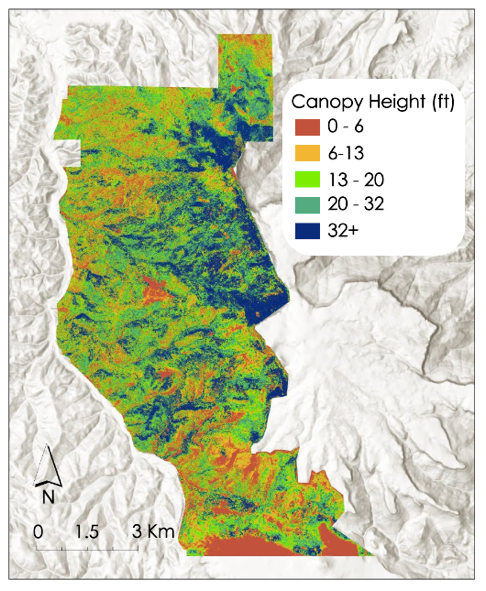


*Figure 5.* The graph illustrates the predicted versus actual biomass values from our predictive model. The dashed line represents the ideal 1:1 relationship of predicted to actual values. The solid red line represents the actual best-fit of predicted to actual biomass. The saturation point of predicted biomass is approximately 150Mg/ha.

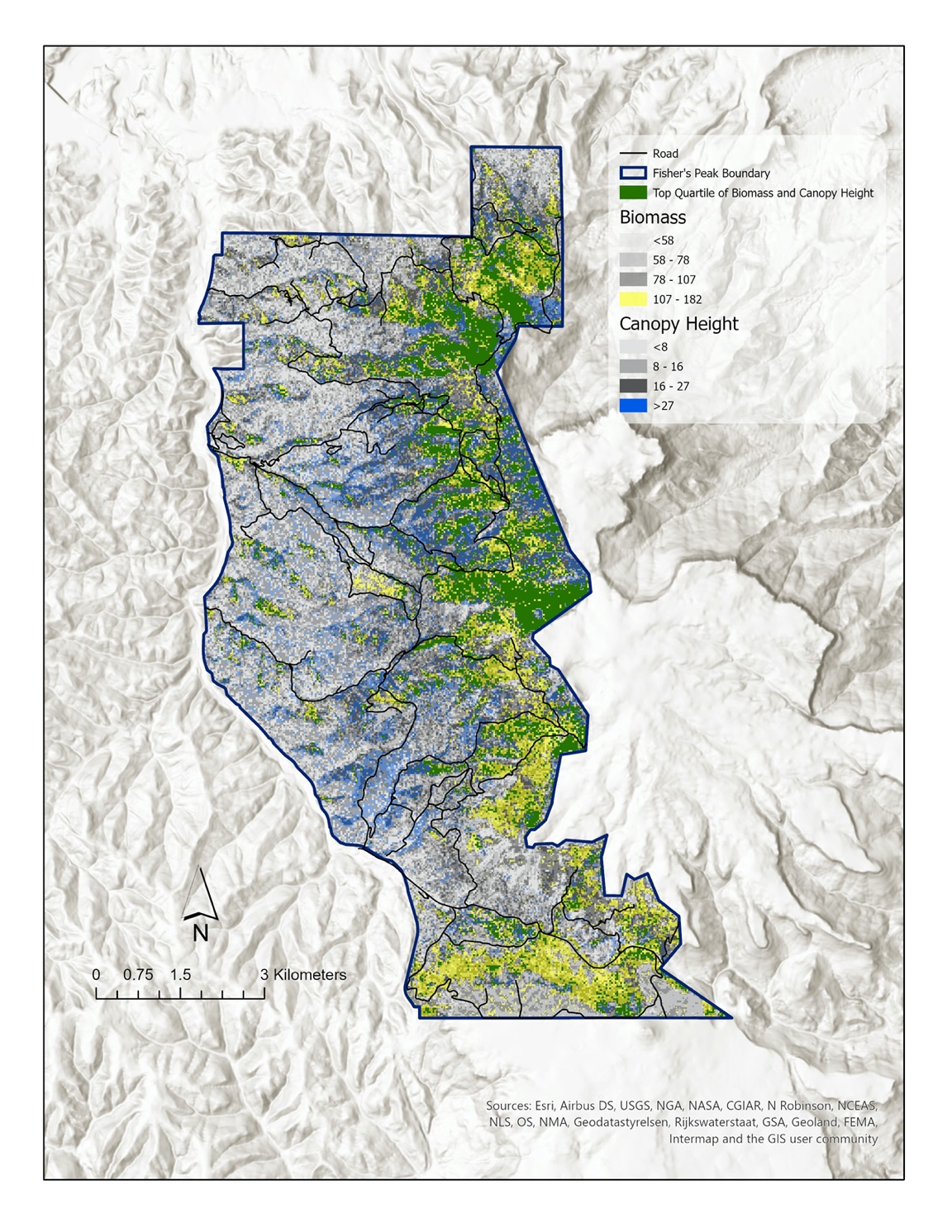
Using our predictive biomass model, we generated a biomass map (Figure 6) for project partners to incorporate into informed decision making regarding future park development. The biomass map quantified the property’s total biomass at 640,366 Mg, which is equal to 320,183 metric tons of carbon. The team also utilized LiDAR canopy height data to develop a canopy height summary and canopy variability map (Figure 7). Additionally, the team developed a map of high carbon storage locations using the top quantile of our predictive biomass map (Figure 8 and Appendix C).



*Figure 6.* The map of biomass generated by the random forest model M4, with remotely-sensed predictor variables: LS Wetness, SRTM Northness, LS Greenness, and SRTM elevation.



*Figure 7.* The canopy height map (left) is derived from LiDAR canopy height data received from partners. The red section on the southern border of the property is a mesa grassland. The canopy variability map (right) shows the standard deviation from the average canopy height among measurements within 100m pixels, which may relate to sub-pixel variability.



*Figure 8.* Using the predictive biomass map from Figure 6, this map isolates the top quantile of biomass in yellow, and the top quantile of canopy height in blue. The green areas demonstrate the overlap of tall stand species and high biomass storage.

***4.2 Discussion***

This project’s predictive biomass model, although low in accuracy (i.e. RMSE=54.48 Mg/ha, R2=9.56), provides insight into optimal predictor variable combinations. Based on predictive map comparisons, a combination of optical, topographic, and radar data produces a realistic approximation of AGB patterns in Fisher’s Peak State Park. However, the saturation point of predicted biomass (Figure 5), is approximately 150Mg/ha, while our field data record biomass values of over 250Mg/ha. The saturation point describes the biomass level at which a model or sensor can no longer distinguish further increases in biomass. As the forest canopy closes, optical imagery cannot distinguish taller canopies from shorter ones. This causes the model to identify areas of closed canopy as containing uniform amounts of biomass (Joshi et al., 2017; Rodríguez et al., 2017). Because optical remote sensors have lower saturation points than SAR or LiDAR (Zhao et al., 2016), the importance of optical predictors in our model may have contributed to a lower saturation point. Additionally, the majority of our training points had relatively low biomass, with only 8 plots containing more than 150Mg/ha of biomass. This corroborates field observations that the Fisher’s Peak property contains many low-biomass stands.

The biomass map (Figure 6) largely aligns spatially with the team’s expectations and knowledge of the study area. The property generally increases in elevation from west to east, accompanied by a transition to tall stand species like ponderosa pine. A corresponding increase in biomass can be seen from west to east on the map and the areas of the highest elevation have the highest modeled biomass. An exception to this trend is the grasslands on a mesa near the park’s southern boundary. These grasslands are at some of the highest elevations in the study area and the map indicates relatively high biomass values. However, we know this area to be dominated by grasses and to contain less biomass than the surrounding areas. Unsurprisingly, our model does not accurately capture most grasslands as our model was specific to live, aboveground, forested biomass, and our training data did not contain non-forested plots. Future studies to quantify more types of biomass in the study area must take this into account in their sampling method design.

***4.3 Caveats and Uncertainties***

In addition to our biomass map, we produced a canopy height summary and canopy variability maps (Figure 7). These maps can be used for identifying patterns in vegetation structure. For example, linking canopy variability and biomass estimates, we were able to identify areas of high biomass that contain both tall stand trees and saplings. This will allow our partners to demarcate areas that sequester large amounts of carbon, but will also increase their carbon storage as the saplings reach maturity. Additionally, this map can be used by the Natural Resource Ecology Laboratory at Colorado State University as they develop a land cover map of Fisher’s Peak using our canopy variability layer as a predictor variable in classification models.

The forest inventory data for saplings only included a range for tree DBH (ex. 1-1.9 in). Therefore, the median value for the given range was used for calculating sapling biomass. This could be problematic if the DBH values of the saplings differed from the median value significantly, as it would skew the training data values. Additionally, the species for each sapling DBH were formatted as an unassigned list for some plots, so DBH was assigned to a tree species based on the most dominant species in that plot, unless the individual species was noted. When converting the variable radius plots to uniform 10m pixels, assumptions were made that A) the resampled biomass values match the ground truth values, B) each training plot holds the same weight in the resampling, and C) that every tree species/mix of species can be resampled with the same methods. Additionally, mature trees were sampled using variable radius plots, while saplings were sampled with fixed radius plots. To maintain training data consistency and reduce sources of error, a future sampling method using exclusively fixed radius plots would be recommended.

Our work could be improved by validation data, fixed-radius or plot sampling scheme, or comparison with a map of detailed landcover. Better representation on the ground to correlate with remote-derived indices could also strengthen the RF model, as training data captures more diverse landcover represented in the area. With research continuing in the area, opportunities to further investigate biomass with higher accuracy assessment exist for project partners.

***4.3 Future Work***

This project developed a biomass map that will be best used in combination with other forest inventory efforts, including a comprehensive landcover map being developed by the Evangelista Lab at Colorado State University. These two products could be used to describe species level biomass and carbon sequestration estimates. Additionally, the biomass map should be compared to park development plans in order to avoid the development of areas with tall stand forests or high carbon storage potential.

This project would also benefit from future model runs using additional datasets. Since this project included sapling biomass, the typically strong predictor variable of LiDAR canopy height was much lower in variable importance than expected. The evaluation of LiDAR-derived minimum canopy height, maximum canopy height, and canopy density layers could provide increased accuracy (Gleason & Im, 2012). Finally, the Global Ecosystem Dynamics Investigation (GEDI) and the ESA’s Biomass missions may be useful given that they specifically target biomass-related metrics. These missions could provide high-resolution and highly correlated predictor variable datasets for future projects aiming to quantify forest biomass.

# 5. Conclusions

Employing random forest predictive modeling utilizing data collected in the field, remotely-sensed imagery, and derived indices, the team established a framework for the creation of a predictive biomass map using forest inventory and remotely sensed data. However, this model would benefit from a higher number of plots and/or sampling on a fixed radius. Additionally, while previous studies have noted the usefulness of canopy height for biomass estimates, this model found that radar data from the ALOS-2 PALSAR-2 L-band were more useful. This could be the result of our inclusion of saplings in our biomass training data. Since the L-band penetrates forest canopies, it could pick up on saplings better than a canopy height dataset that only looks at the taller trees in an area.

Investigating an ecologically diverse property such as Fisher’s Peak State Park proved to be complex and multifaceted. The work of estimating biomass in the Fisher’s Peak State Park through RF modeling using remotely-sensed data demonstrated potential workflows for future work on the property and with other state park projects. While carbon market entry requirements cannot be satisfied based on our work, project partners could use a more complex representation of biomass distribution to further explore the potential for carbon credits on the property. Partners can leverage our products to investigate areas of interest for priority conservation and designated infrastructure development. The work of the Fisher’s Peak Ecological Forecasting team generated highly useful resources for project partners’ to continue their work exploring development scenarios of park infrastructure and its impact on local ecology.

# 6. Acknowledgments

The Fisher’s Peak Ecological Forecasting team would like to express our thanks to the project partners and mentors that made this project possible.

Partners:

* Chris Pague, The Nature Conservancy, Senior Ecologist
* Dr. Amanda Fordham, Colorado State Forest Service, Science Information Manager

Mentors/Science Advisors:

* Dr. Paul Evangelista, Colorado State University, Natural Resource Ecology Laboratory
* Dr. Catherine Jarnevich, USGS, Fort Collins Science Center
* Peder Engelstad, Colorado State University, Natural Resource Ecology Laboratory
* Nicholas Young, Colorado State University, Natural Resource Ecology Laboratory
* Dr. Tony Vorster, Colorado State University, Natural Resource Ecology Laboratory
* Kristen Dennis, NASA DEVELOP Colorado Lead/Fellow

This material contains modified Copernicus Sentinel data (2019), processed by ESA.

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

**AGB** – Aboveground biomass

**BAF** – Basal area factor, common method for describing stand density, typically of merchantable timber

**DBH** – Diameter at breast height, a standard dendrometric measurement used to capture the diameter of a tree’s trunk/bole at 1.37m

**Earth observations** – Satellites and sensors that collect information about the Earth’s physical, chemical, and biological systems over space and time

**Landsat** – Jointly managed earth-observing NASA/USGS satellite mission

**LiDAR** – Light Detection and Ranging, an active remote sensing technique

**NDVI** – Normalized Difference Vegetation Index; a transformation that is a proxy for the “greenness” of a pixel, and enhances spectral differences on the basis of absorption and reflectance in the red and near-infrared bands.

**RF** – Random forest, an ensemble machine learning method that can be used for regression modeling

**Tasseled Cap Indices** – Tasseled cap indices compress spectral data into a few useful bands of data: brightness, greenness, wetness, and other uncorrelated data computed using a linear combination of bands and coefficients found using Principal Component Analysis (Crist and Cicone, 1984; Nedkov 2017). The bands of uncorrelated data were not included in the model development.

**USGS** – United States Geological Survey

# 8. References

Crist, E. P., & Cicone, R. C. (1984). A physically-based transformation of Thematic Mapper data---The TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing, (3*), 256-263. DOI: 10.1109/TGRS.1984.350619

Dang, A. T. N., Nandy, S., Srinet, R., Luong, N. V., Ghosh, S., & Kumar, A. S. (2019). Forest aboveground biomass estimation using machine learning regression algorithm in Yok Don National Park, Vietnam. *Ecological Informatics*, 50, 24-32. DOI: 10.1016/j.ecoinf.2018.12.010

Foody, G. M., Boyd, D. S., & Cutler, M. E. (2003). Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote Sensing of Environment,* *85*(4), 463-474. DOI: 10.1080/0143116031000103853

Gleason, C. J., & Im, J. (2012). Forest biomass estimation from airborne LiDAR data using machine learning approaches. *Remote Sensing of Environment*, *125*, 80-91. DOI: 10.1007/978-3-319-01854-6\_26

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-Scale Biomass Estimators for United States Tree Species. *Forest Science,* *49*(1). Retrieved from <https://www.fs.fed.us/ne/newtown_square/publications/other_publishers/OCR/ne_2003jenkins01.pdf>

Joshi, N. Mitchard, E.T.A., Brolly, M., Schumacher, J., Fernández-Landa, A., Johannsen, K.V., Marchamalo, M., & Fensholt, R. (2017). Understanding ‘saturation’ of radar signals over forests. *Scientific Reports*, 7(3505). DOI: https://doi.org/10.1038/s41598-017-03469-3

Nandy, S., Ghosh, S., Kushwaha, S. P. S., & Kumar, A. S. (2019). Remote sensing-based forest biomass assessment in northwest Himalayan landscape. In *Remote Sensing of Northwest Himalayan Ecosystems* (pp. 285-311). Springer, Singapore. DOI: 10.1007/978-981-13-2128-3\_13

Nedkov, R. (2017). Orthogonal transformation of segmented images from the satellite Sentinel-2. *Comptes rendus de l’Academie bulgare des Sciences, 70*(5), 687-692.

Rodríguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K., & Balzter, H. (2017). Quantifying forest biomass carbon stocks from space. *Current Forestry Reports,* *3*, 1–18. DOI: https://doi.org/10.1007/s40725-017-0052-5

Powell, D. C. (2014). *Using Mathematics in Forestry*. USDA Forest Service. <https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5413736.pdf>

Timothy, D.H..(2016). Remote sensing of aboveground forest biomass: A review. *Tropical Ecology*, *57*(2), 125-132. Retrieved from: <https://www.tropecol.com/pdf/open/PDF_57_2/1%20Timothy%20et%20al..pdf>

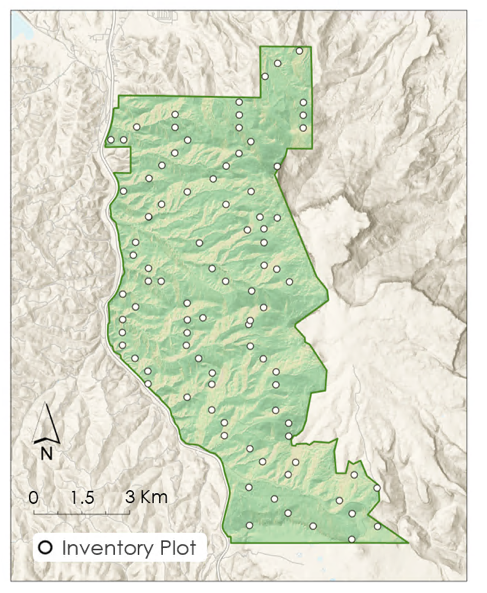
Zhao, P., Lu, D., Wang, G., Liu, L., Li, D., Zhu, J., & Yu, S. (2016). Forest aboveground biomass estimation in Zhejiang Province using the integration of Landsat TM and ALOS PALSAR data. *International Journal of Applied Earth Observation and Geoinformation*, *53*, 1-15. DOI: https://doi.org/10.1016/j.jag.2016.08.007

Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal*, *4*(3), 212-219. DOI: https://doi.org/10.1016/j.cj.2016.01.008

# 9. Appendices

**Appendix A:**

**Calculating *In-Situ* Biomass and Converting Variable to Fixed Radius Plots**



*Figure A1.*The locations of the forest area plots are shown above.

*Table A1*

*The allometric equation and coefficients used for calculating biomass for each variable radius plot within the study area (Jenkins et al. 2003).*

|  |  |  |
| --- | --- | --- |
| **Biomass (kg) = exp^(B0+(B1 \* ln(DBH))** | | |
| **Species** | **B0** | **B1** |
| Douglas Fir  *(Pseudotsuga menziesii)* | -2.2304 | 2.4435 |
| Gambel Oak  *(Quercus gambelii)* | -2.0127 | 2.4342 |
| New Mexico Locust  *(Robinia neomexicana)* | -2.48 | 2.4835 |
| One Seed Juniper  *(Juniperus monosperma)* | -0.7152 | 1.7029 |
| Pinon Pine  *(Pinus edulis)* | -2.5356 | 2.4349 |
| Ponderosa Pine  *(Pinus ponderosa)* | -2.5356 | 2.4349 |
| Rocky Mountain Juniper  *(Juniperus scopulorum)* | -0.7152 | 1.7029 |
| White Fir  *(Abies concolor*) | -2.5384 | 2.4814 |

*Table A2*

*Example biomass calculation and variable radius plot conversion to a uniform 10m pixel. This calculation example is for Fisher’s Peak State Park Forest Inventory plot 20. Plot 20 has 9 living adult Douglas fir trees and one Douglas fir sapling. In step 1, the DBH should be in centimeters. In step 3, the plot radius factor of 2.75 is derived from a BAF of 10 (Jenkins et al. 2003, Powell 2014). In this example the final value is given in kg/100m2, however, this value should be communicated in Mg/ha. The biomass value in Mg/ha is derived by dividing the value in kg/100m2 by 10.*

|  |  |  |
| --- | --- | --- |
| **Biomass (kg) = exp^(B0+(B1 \* ln(DBH))** | | |
| **Species: Douglas Fir** | **B0=-2.2304** | **B1=2.4435** |
| **Step 1: Calculate the biomass of adult trees on the plot.** | Biomass = exp^(-2.2304+(2.4435\*ln(Tree 1 DBH))  Use the DBH of the remaining 8 trees to calculate tree biomass, then add the 9 biomass values together to get the total for plot 20. | |
| **Step 2: Calculate the biomass of the saplings. Then resample to the 10m pixel size. B10=Biomass of 10m pixel** | Biomass = exp^(-2.2304+(2.4435\*ln(Sapling))  If there are multiple saplings, follow the method above and sum the biomasses of each tree to find the plot total.  B10 = (Area of desired cell size \* original plot area)/ Biomass  B10 = (100m2 \* 41.37kg)/13.50m2  B10 = 306.53kg | |
| **Step 3: Find the horizontal limiting distance for the plot. Then convert the HLD from feet to meters.** | HLD(ft) = Plot radius factor \* DBH (in) of the largest tree  HLD = 2.75 \* 28  HLD = 77 ft or 23.46 m | |
| **Step 4: Resample the HLD to a 10x10m pixel. B10=Biomass of 10m pixel** | B10 = (Area of desired cell size \* original prism plot area) / Biomass  B10 = (100m2 \* 9296.52 kg)/1729.4 m2  B10 = 537.56kg | |
| **Step 5: Add mature tree biomass and sapling biomass for 10x10m pixel value. Divide this value by 10 for communication purposes.** | Total B10 = B10 Mature Trees + B10 Saplings  Total B10 = 537.56kg + 306.53kg  Total B10 = 844.09kg  Biomass for plot 20 = 84.4 Mg/ha | |

**Appendix B:**

**Landsat 8 and Sentinel-2 Top of Atmosphere Tasseled Cap Coefficients**

*Table B1*

*Tasseled Cap Transformation coefficients for Landsat 8 OLI TOA imagery (Baig et al., 2014).*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Landsat 8 OLI Tasseled Cap Top of Atmosphere Band Coefficient** | | | | | | |
| **Component** | **Band 2 (Blue)** | **Band 3 (Green)** | **Band 4 (Red)** | **Band 5 (NIR)** | **Band 6 (SWIR1)** | **Band 7 (SWIR2)** |
| **Brightness** | 0.3029 | 0.2786 | 0.4733 | 0.5599 | 0.508 | 0.1872 |
| **Greenness** | -0.2941 | -0.243 | -0.5424 | 0.7276 | 0.0713 | -0.1608 |
| **Wetness** | 0.1511 | 0.1973 | 0.3283 | 0.3407 | -0.7117 | -0.4559 |

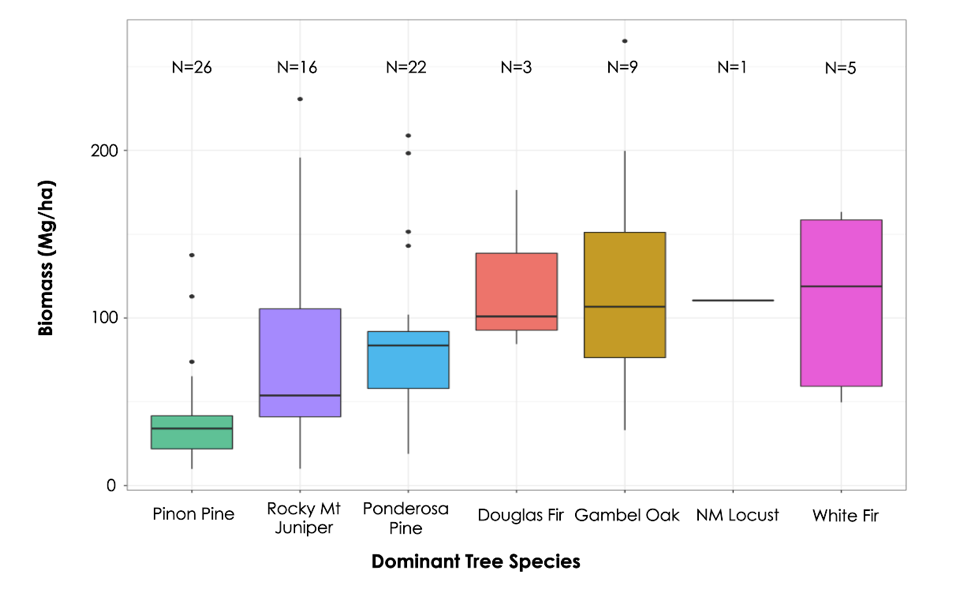
*Table B2*

*Tasseled Cap Transformation coefficients for Sentinel-2 MSI Level 1C (TOA) imagery (Nedkov, 2017).*

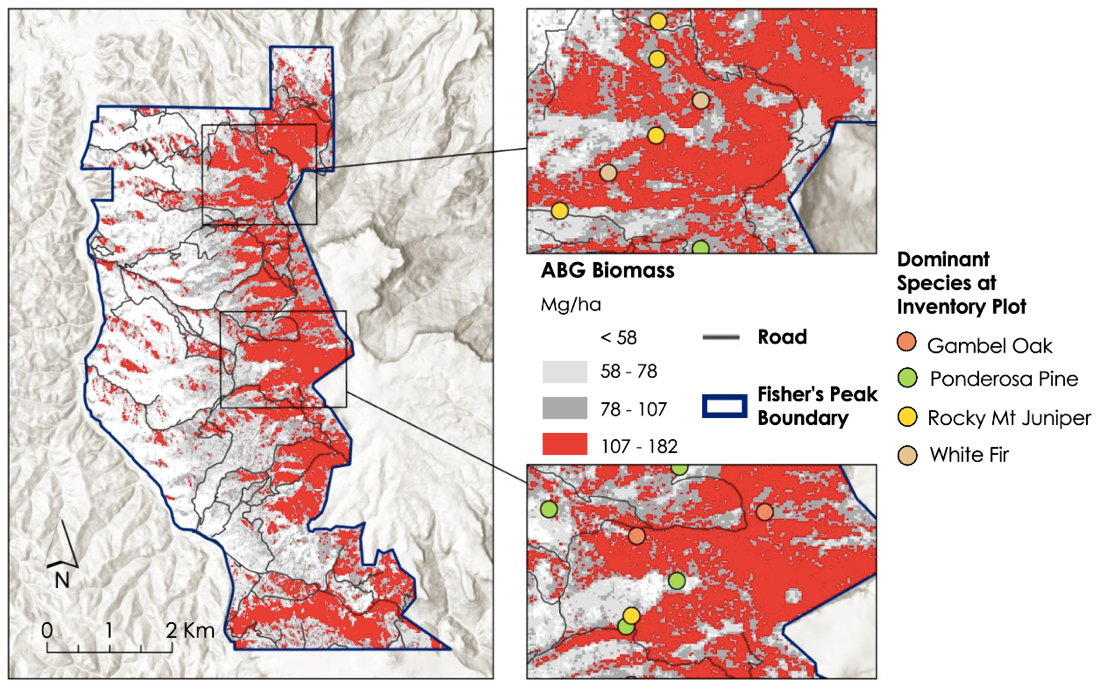
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sentinel-2 MSI Tasseled Cap Top of Atmosphere Band Coefficients** | | | | | | |
| **Component** | **Band 2 (Blue)** | **Band 3 (Green)** | **Band 4 (Red)** | **Band 8 (NIR)** | **Band 11 (SWIR1)** | **Band 12 (SWIR2)** |
| **Brightness** | 0.0822 | 0.1360 | 0.2611 | 0.3895 | 0.3882 | 0.1366 |
| **Greenness** | -0.1128 | -0.1680 | -0.3480 | 0.3165 | -0.4578 | -0.4064 |
| **Wetness** | 0.1363 | 0.2802 | 0.3072 | -0.0807 | -0.4064 | -0.5602 |

**Appendix C:**

**Biomass Distribution by Species**



*Figure C1.* This box plot shows forest inventory plots, sorted by the dominant tree species in each plot, in relation to the amount of biomass that plot contains.



**AGB Biomass**

Mg/ha

*Figure C2.* Using the predictive biomass map from Figure 6, this map isolates the top quantile of biomass in red. The outlined areas isolate regions in the top quantile of biomass that also features canopy peaks and high canopy variability.