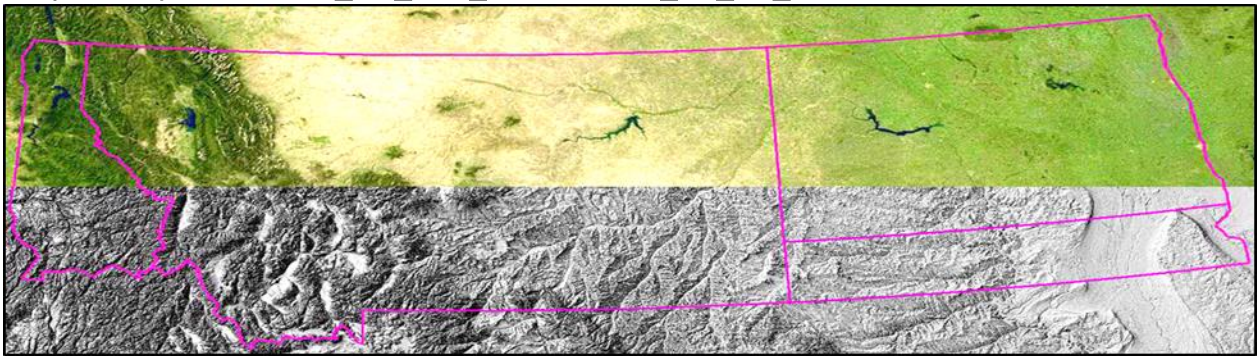
Project Report: NRGG_PR_B-D_VMAP2018_09_15_2018



The Beaverhead-Deerlodge National Forest Region 1 Existing Vegetation Database (VMap) Revision of 2018

A USDA Forest Service, Northern Region Geospatial Group Project Report
By Robert Ahl, Jed Gregory, Steve Brown, Kelsey David, Jeff Kaiden, and Fred Kellner. For further information contact:
Steve Brown at 406-329-3514, stevebrown@fs.fed.us

Table of Contents

List of Tables

List of Figures

1.0 Overview

Consistent, continuous, contemporary and accurate vegetation data are essential for effective ecosystem assessment and land management planning. The Northern Region Existing Vegetation Mapping Program (VMap) (USDA 2018) addresses this information need by providing a database of existing vegetation and associated map products that are constructed with an analytical methodology based on the Existing Vegetation Classification and Mapping Technical Guide (Brohman and Bryant, 2005) to support the Region 1 Multi-level Classification, Mapping, Inventory, and Analysis System, R1-CMIA (Berglund et al. 2009).

A VMap database has been published for every Forest in the USDA Forest Service Northern Region, and updated in a cyclical manner since 2003 (Brown and Barber 2012, Brown et al. 2012). In 2018, an updated VMap database was produced for the Beaverhead-Deerlodge National Forest (B-D). The VMap database consists of four primary spatially explicit attributes that include descriptions of 1) lifeform, 2) tree canopy cover class, 3) tree size class, and 4) tree dominance type. These attributes can be mapped and used to support mid and base-level analysis and planning. VMap uses the Region 1 Existing Vegetation Classification System (R1-ExVeg) (Barber et al. 2009) in its map unit design. This system defines the logic for grouping entities by similarities in their floristic characteristics. VMap products are derived using remote sensing technology, and are based on a combination of airborne imagery and a nationally available digital topographic and climatic data.

With a foundation of contemporary aerial imagery, a clear view of the project area is essential. The B-D mapping area of interest was obscured by forest fire smoke in 2015 and in 2017. For these reasons, continuous high resolution NAIP imagery (USDA Farm Service Agency 2015, 2017) was not available for the full extent of the B-D update area, and NAIP imagery from 2013 was utilized instead. Furthermore, to obtain contemporary and full coverage of the mapping area, Sentinel-2 satellite data (ESA 2015) was sourced between July 18, 2016 and August 8, 2017 to capture existing vegetation patterns during peak greenness conditions. The imagery was delivered with 10 meter pixel resolution, and thirteen spectral bands of radiometric resolution, including red, green, blue, and infrared components. However, even with a custom collection of image data, cloud cover was still present.

The Sentinel-2 data was used in combination with a biophysical characterization layer in a segmentation procedure to create a vector-based layer of polygons that represent a delineation of patches of similar vegetation types across the study area (Haralick and Shapiro 1985, Zaitoun and Aqel 2015). The resulting set of polygons are the elements attributed by the VMap process and issued in the final database.

Field-based reference information was collected and used as training data to make spatial predictions of the vegetation attributes contained in the database. Predicted raster surfaces of the attributes were then summarized to the delineated polygons.

As draft map products were created by the labeling algorithms, they were reviewed and appropriate changes were made in the database. Upon a satisfactory conclusion, the final products were used to populate the VMap database.

After draft products were inspected and adjusted, an accuracy assessment was conducted to provide a quantitative validation of the database. Estimates of overall map accuracy and confidence measures of individual map classes can be inferred from the error matrix derived from the comparison of known reference sites to mapped data, for each attribute. The stated accuracy assessment results are applicable to the entire B-D database, and ranged from 63-93%, depending on the attribute in question.

2.0 Source Data

A combination of field reference, recent image, and biophysical data are needed to produce the VMap database. Once collected, ground reference data is used to build relationships between the observed phenomena and the spectral and biophysical information derived from remotely sensed and ancillary data.

2.1 Field Data Collection

Collectively, ground and other reference data are known as "training data" because they are used to construct algorithms that relate observations to quantified variables and are used to interpret and label areas that have not been sampled within a study area. Thus, they "train" algorithms to distinguish between and label the un-sampled areas. For the development of the VMap database, training data was specifically collected in the field with plot-based measurements that are used in conjunction with image interpretation to identify and distinguish lifeform, tree canopy cover, tree size, and vegetation dominance type classes. For a more detailed explanation of the field data collection process and the findings of the field season please see the Story Map at https://arcg.is/0vSSnX.

2.2 Image Data Collection and Pre-Processing

Two distinct types of spectral image data were used in the production of the VMap database. They include NAIP imagery from 2013, and a 2016-2017 composite of Sentinel-2 data. NAIP data have the finest grain size, with 1 meter pixel resolution (USDA Farm Service Agency 2016), but, due to persistent cloud and wildland fire smoke cover, contemporary coverage over the mapping area was not available for this project in 2017. While current NAIP image data was not available, image data from 2013 was used, in part, for algorithm development, and visual inspection. Sentinel-2 image data was specifically obtained for the B-D project, and was delivered with 10 meter pixel resolution, and 13 bands of radiometric resolution (ESA 2015). The Sentinel-2 data were extracted from Google Earth Engine using standard methods (Gorelick et al 2017). For algorithm development in this project, NAIP imagery was used to model the fine grain texture associated with tree canopy cover and tree size. It was also used in combination with Sentinel-2 data to model lifeform. Sentinel-2 data are both contemporary and spectrally consistent across the project area and have a higher number of spectral bands with which to better discriminate between individual tree species for species distribution and tree dominance type modeling.

2.3 Creation of Image Derivatives

Image derivatives are transformations of raw image data that provide spectral and texture-based information useful for land cover mapping. Regardless of the native format, all derivatives used in the B-D mapping process were converted to a 10 meter pixel resolution to enhance processing speed and reduce variability in the dataset.

The derivatives used in the B-D process were based on NAIP and Sentinel-2 data. As a first step in the derivative creation process, the NAIP imagery was degraded from 1 meter spatial resolution to 5 meter spatial resolution. Then, a principal component analysis (PCA) (Jolliffe 2002) of the four bands of 5 m NAIP image data was conducted. The first three components were retained and stacked to yield a three band principal component raster with 5 meter pixel resolution. From this raster, a focal mean, focal standard deviation, and contrast gray level co-occurrence matrix were created, using a seven pixel by seven pixel moving window. The results of the focal and gray level co-occurrence matrix computations were then degraded to 10 meter pixel resolution for the final application. In a similar fashion, the 10 meter native resolution Sentinel-2 image data were also transformed into a three band principal component raster. Focal and contrast derivatives were also created for the Sentinel-2 data, using a three by three pixel window, to maintain a close tie to both the NAIP imagery and general size of plots used to acquire ground-based measurements. Finally, the NAIP and Sentinel-2 focal and contrast derivatives, based on the top three principal components, at 10 meter spatial resolution, were used as spectral inputs for lifeform, tree canopy, tree size, and species distribution and tree dominance type modeling.

2.4 Long term Site Characterization

Vegetation indices provide another useful metric for describing and distinguishing various vegetation characteristics. The normalized difference vegetation index (NDVI) is commonly used and yields a measure of photosynthetic activity in plants, using information related to the wavelengths of light that are captured by image sensors (Rouse et al. 1974, Lillesand et al. 2015). In its original format, NDVI quantifies photosynthetic activity at the instantaneous time of image collection, and while this is useful, it does not provide information about long term processes or trajectories over time. By summing individually collected NDVI values over each time period the data are collected, seasonal patterns of green-up to senescence can be interpreted by the magnitude of accumulated values. For example, an area of deciduous shrubs that is very active photosynthetically will have very high individual NDVI values during the active growing season, and those values will accumulate to be higher than its evergreen counterparts that rely on lower levels of sustained photosynthesis over longer periods of time. An index that captures the accumulated values of NDVI is called Time-Integrated NDVI (TINDVI) (Reed et al. 1994). A TINDVI for a 30-year period record for the growing season months (July to September) was calculated from the Landsat data record for use in modelling long term site photosynthetic activity as a surrogate for productivity.

2.5 Biophysical Characterization Data

In the arid West moisture availability is often the limiting factor in vegetative growth/productivity and species distribution. As such, biophysical setting can be a useful piece of information when characterizing vegetation. To address this information gap a raster derivative that integrates precipitation, solar radiation, and topography was used to quantify the physical environment. This provides a physical foundation for processes that are associated with the availability of water. Because it integrates precipitation, heat load from the sun, and water routing by topographic elements, it is called PHEAT (Precipitation Heat & Elevation Adjusted Topography). PHEAT is used to help inform the delineation of polygons in the segmentation process and the derivation of vegetation characteristics in modeling processes.

2.6 Image Segmentation

Image segmentation is the process of combining unique picture elements, or pixels, within digital images into spatially cohesive regions. These individual regions are called image objects and represent distinct areas within the image that generally correspond to patches of similar vegetation type/conditions (Haralick and Shapiro 1985, Zaitoun and Aqel 2015). Ultimately, the raster-based image objects are converted to vector-based polygons. These image objects depict patterns of vegetation and other elements on the landscape, and all VMap attributes are associated with the polygons derived from the segmentation process.

For this database, polygon delineation was based on Sentinel-2 imagery, and biophysical characterization supplied by the PHEAT derivative. Both datasets had 10 meter pixel resolution. Figure 1 provides an example of how imagery has been segmented to yield the delineation of distinct vegetation patches. In this particular display, the simplified polygons of the Beaverhead-Deerlodge National Forest VMap dataset are shown in combination with the color infrared components of the 2013 NAIP imagery.

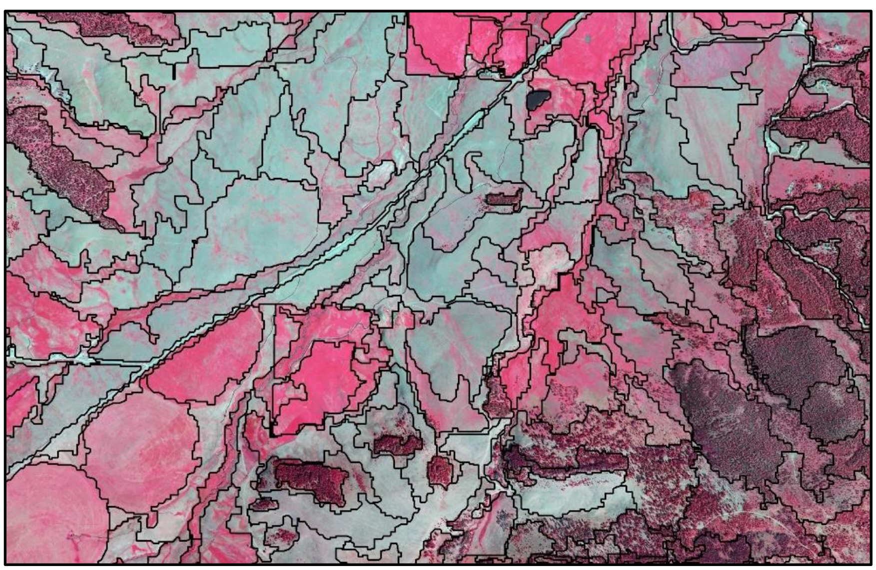


Figure 1 - *Delineation of distinct vegetation and terrain features using an image segmentation routine. Illustrated are the simplified polygons of the Beaverhead-Deerlodge National Forest VMap dataset are shown in combination with the color infrared components of the 2013 NAIP imagery.*

3.0 Mapping Process

3.1 Lifeform Classification

At the level of individual polygons, the lifeform class is attributed through a combined process of image object classification and refined with manual image interpretation and editing, following the rules established by the R1 Existing Vegetation Classification document. Labeling of the lifeform groups is accomplished with the Random Forest classification algorithm (Breiman 2001, Liaw and Wiener 2002, Liaw 2015) using field collected reference information and summarized image derivative, biophysical derivative, and vegetation index derivative statistics associated with the polygons obtained from the segmentation process. In this process, spectral inputs from both NAIP and Sentinel-2 data were used. Mapped lifeform classes include Coniferous Tree, Deciduous Tree, Shrub, Herbaceous, Sparsely Vegetated, Urban, and Water. A polygon-based accuracy assessment is given in the Accuracy Assessment section of this document.

3.2 Tree Canopy Cover Classification

For polygons where a Coniferous Tree lifeform is assigned, tree canopy cover values are estimated. Traditionally the tree canopy cover values in the VMap database were only available in four classes: low (10-24.9% Cover), moderate low (25-39.9% Cover), moderate high (40-59.9% Cover) and high (60%+ Cover). In this VMap update, however, canopy cover estimates were produced as continuous variables that were distributed into the stated classes. Providing continuous canopy cover percentage values and categorical groupings of canopy cover enable increased precision for model and decision support.

Canopy cover models were based on reference data obtained through analyst-based image interpretation, and a Random Forests regression algorithm (Breiman2001, Liaw and Wiener 2002, Liaw 2015).

Using a 70 meter by 70 meter grid, which resembles the dimensions of an FIA plot (Bechtold and Patterson 2005), an image analyst randomly selected 400 grid cells across the mapping area and then used high resolution imagery to assign a percent canopy cover estimate to each cell. A full range of canopy cover values, ranging from a minimum of 10% to 80%, were generated and used as training data in the modeling process. A selection of reference sites were used in combination with the NAIP 2013 image derivatives, and biophysical setting variables that include 30-year composite vegetation index and a moisture availability model in a Random Forest regression model to estimate the full range of canopy cover values across the mapping area. Another subset of the reference data were used for model evaluation. Table 1 provides a description of sites used for model development and model evaluation.

Table 1 - *Canopy cover classes, and numbers of reference sites used for model development and model evaluation*.

Canopy cover quantity (%)	Number of reference sites for model development	Number of reference sites for model evaluation
0	27	14
10	30	14
20	30	16
30	30	8
40	30	11
50	30	47
60	30	23
70	30	4
80	10	6
sum	247	143

The output from this regression was evaluated against the training data and residuals were computed. Those residuals were then used to generate a model of the regression errors. The modeled errors were then used to adjust (added to) the initial canopy cover regression model output. By doing so the strength of the relationship between measured and modeled canopy cover values was enhanced. The initial canopy cover model explained roughly 60% of the variation in the measured values while the residual adjusted model accounted for nearly 80% of that variation (R^2 0.58 vs R^2 0.81). The graphical expression of these models is given in Figure 2 and Figure 3 - *Residual-adjusted regression on measured versus modeled canopy cover.*.

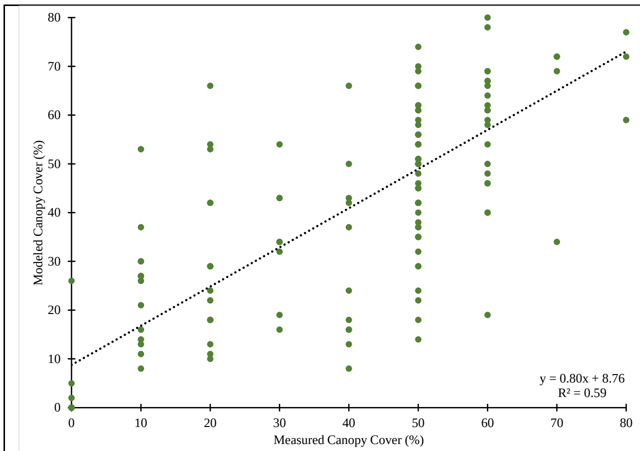


Figure 2 - *Scatter plot of regression on measured versus modeled canopy cover.*

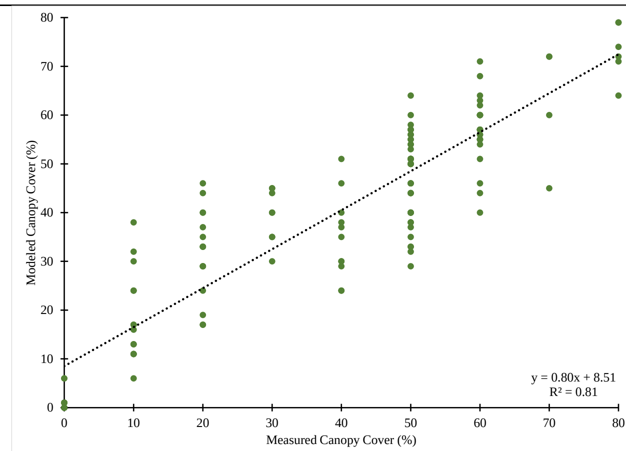


Figure 3 - *Residual-adjusted regression on measured versus modeled canopy cover.*

Although the residual adjusted model fit the measured values well, the slope of this relationship indicates that the lowest model values tend to be slightly higher than measured values, and conversely high measured values tend to have slightly lower modeled values. To further refine the estimated canopy cover values, a stretch was applied to the residual adjusted output. This procedure maintained the strength of the relationship but increased the steepness of the slope of the regression line between measured and modeled values more, thus low values became lower and high values became higher. The final relationship between measured and modeled canopy cover values is given in Figure 4 - *Scatter plot of final regression on measured versus modeled canopy cover.*.

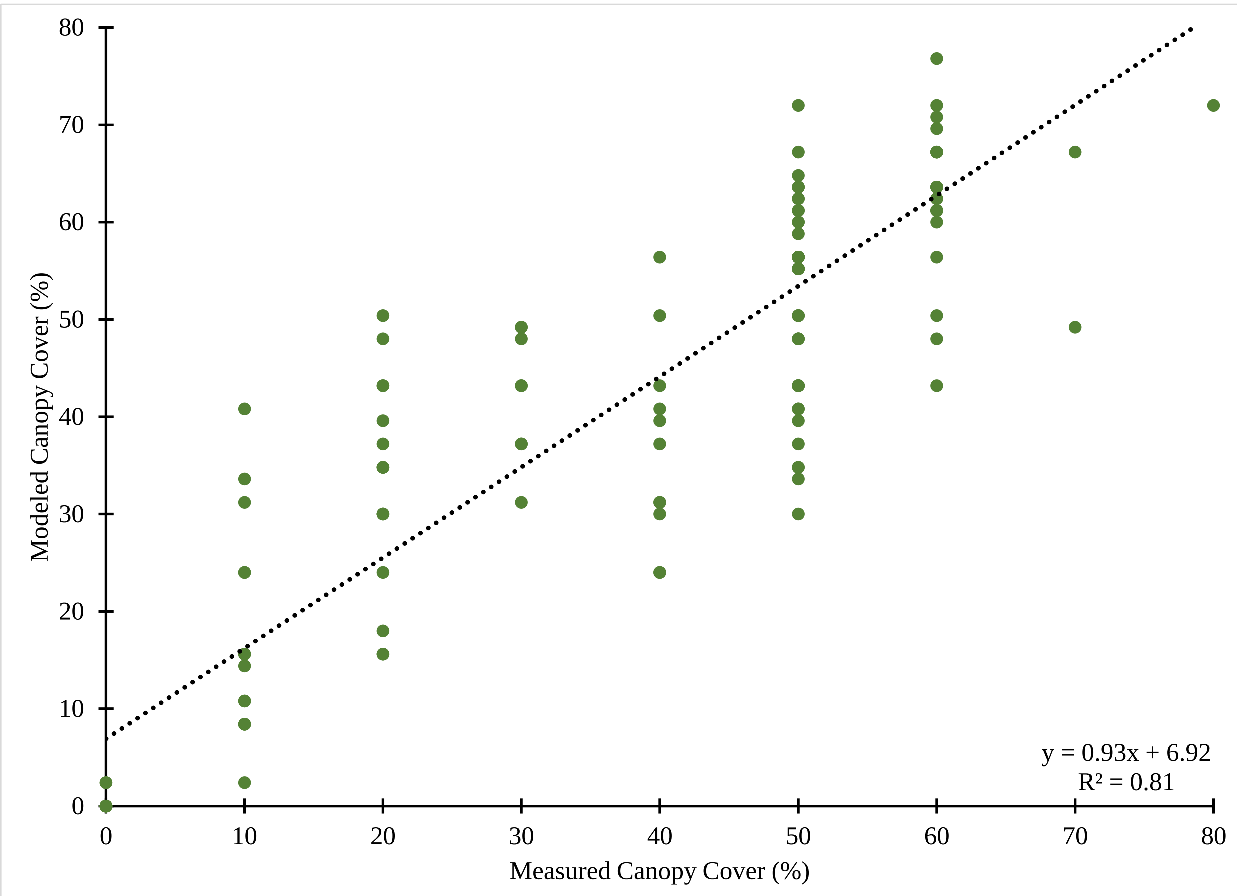


Figure 4 - *Scatter plot of final regression on measured versus modeled canopy cover.*

This final continuous canopy cover model has a mean absolute error of 8% when compared to measured values. The raster containing these final values was summarized to the image segmentation and then grouped into canopy cover classes based on the specifications of the Region 1 Existing Vegetation Classification System, as described above.

Within the designated Regional canopy cover classes, the mean class absolute error is:

Table 2 - *Mean canopy cover class absolute error***.**

Regional Canopy Cover Classes	Percent
Mean Absolute Error Overall	8
Mean Absolute Error 10-24.9%	12
Mean Absolute Error 25-39.9%	12
Mean Absolute Error 40-59.9%	9
Mean Absolute Error >= 60%	7

A polygon-based accuracy assessment is given in the Accuracy Assessment section of this document.

3.3 Tree Size Class

For polygons where a Coniferous Tree lifeform is assigned, tree size values are estimated. Traditionally the tree size values in the VMap database were only available in five classes: 1 (0-4.9”), 2 (5-9.9”), 3h (10-14.9”) and 4 (15-19.9”), and 5 (>= 20”). In this VMap update, however, tree size estimates were produced as continuous variables that were distributed into the stated classes. Providing the flexibility to assess continuous tree size values or categorical groupings of canopy cover enable increased precision for model and decision support.

Tree size models were based on reference data obtained through field-based plot measurements, and a Random Forests regression algorithm (Breiman2001, Liaw and Wiener 2002, Liaw 2015). In the development process, the suite of image derivatives, the vegetation index derivative, and the biophysical derivative, described in the above sections, were incorporated with the plot data.

Over the course of the 2017 field season, analysists collected plot-based measurements of tree size at 1,164 locations across the Beaverhead-Deerlodge National Forest. At each location, variable radius plots were installed and the diameter at breast height (DBH) of all “in” trees was measured. This information was then used to calculate the basal area-weighted diameter at breast height (BAWD) for each plot. The assemblage of tree size information at all plot locations was divided into 1) reference sites for model development, and 2) reference sites for model evaluation. Of all the data collected, 70% were used for model development and 30% were used for model evaluation. Figure 5 - *Relative proportions of model development and model evaluation reference sites used in the tree size modeling process.* provides a graphic display of the relative proportion of model development and model evaluation data used in the tree size modeling process.

The full range of tree size values represented in the model development set, ranging from a minimum of 1 inch to a maximum of 36 inches BAWD, including a set of reference locations that represent zero tree size, were used in combination with the NAIP 2013 image derivatives, and biophysical setting variables that include 30-year composite vegetation index and a moisture availability model in a Random Forest regression model to estimate tree size values, expressed as BAWD, across the mapping area.

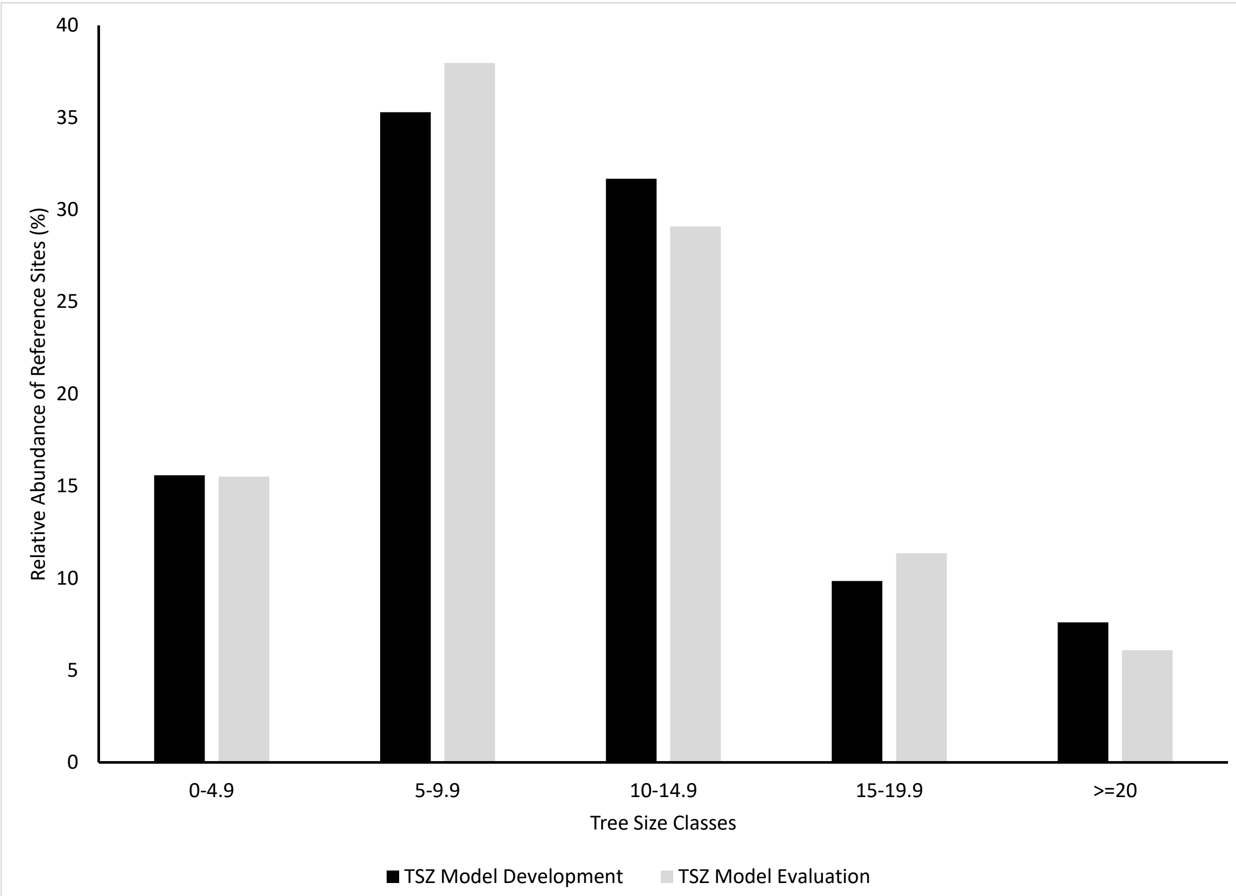


Figure 5 - *Relative proportions of model development and model evaluation reference sites used in the tree size modeling process.*

The output from this regression was evaluated against the model evaluation data and residuals were computed. Those residuals were then used to generate a model of the regression errors. The modeled errors were then used to adjust (added to) the initial tree size regression model output. By doing so the strength of the relationship between measured and modeled tree size values was enhanced. The initial tree size model explained roughly 74% of the variation in the measured values while the residual adjusted model accounted for 75% of that variation (R^2 0.74 vs R^2 0.75). The graphical expression of these models is given in Figure 6 - *Scatter plots of initial tree size regression model.* and Figure 7 – *Scatter plot of residual adjusted tree size regression model.*.

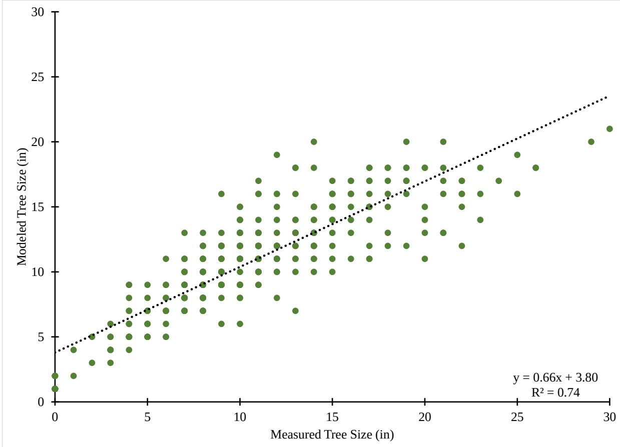


Figure 6 - *Scatter plots of initial tree size regression model.*

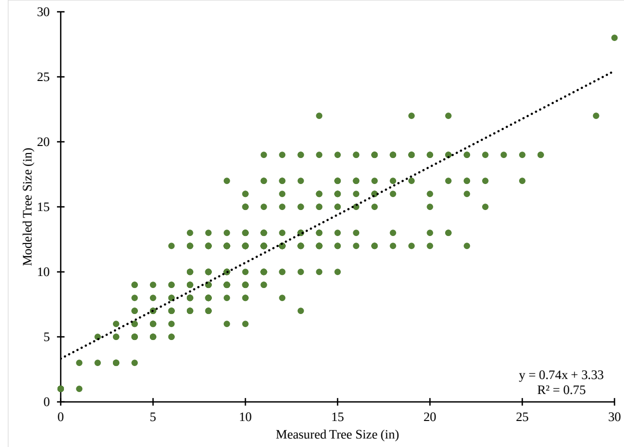


Figure 7 – *Scatter plot of residual adjusted tree size regression model.*

Although the residual adjusted model fit the measured values well, the slope of this relationship indicates that the lowest model values tend to be slightly higher than measured values, and conversely high measured values tend to have slightly lower modeled values. To further refine the estimated canopy cover values, a stretch was applied to the residual adjusted output. This procedure maintained the strength of the relationship but increased the steepness of the slope of the regression line between measured and modeled values from 0.66 to 0.84. Thus, low values became lower and high values became higher. The final relationship between measured and modeled tree size values is given in Figure 8 – *Scatter plot of final residual adjusted and stretched tree size model*..

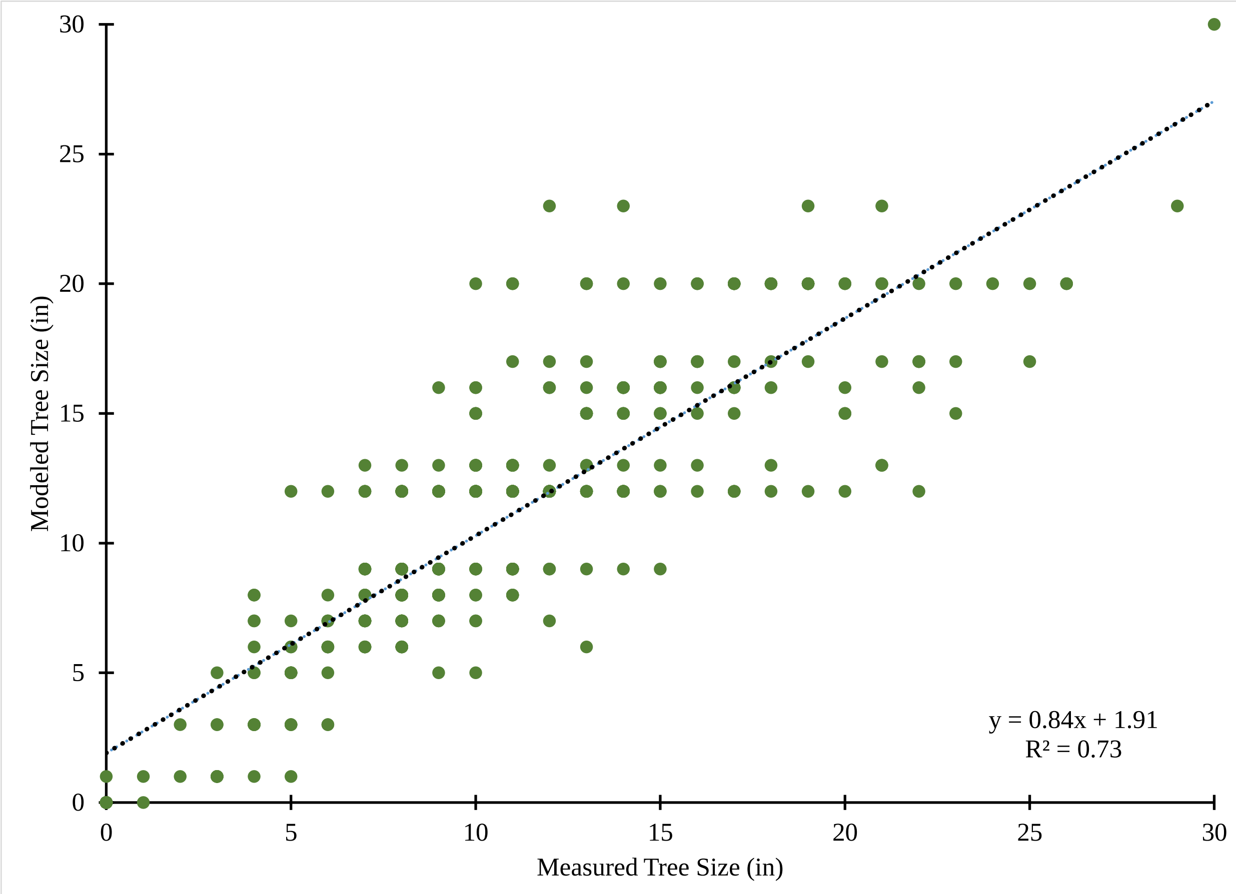


Figure 8 – *Scatter plot of final residual adjusted and stretched tree size model.*

This final continuous tree size model has a mean absolute error of 2.2 inches when compared to measured values. The raster containing these final values was summarized to the image segmentation and then grouped into tree size classes based on the specifications of the Region 1 Existing Vegetation Classification System, as described above. Within the designated Regional canopy cover classes, the mean class absolute error is:

Table 3 – *Mean tree size class absolute error.*

Regional Tree Size Classes	Percent
Mean Absolute Error Overall	2.2
Mean Absolute Error 0-4.9”	1.6
Mean Absolute Error 5-9.9”	1.7
Mean Absolute Error 10-14.9%	2.4
Mean Absolute Error 15-19.9”	2.4
Mean Absolute Error >= 20”	4.7

A polygon-based accuracy assessment is given in the Accuracy Assessment of this document.

3.4 Species Distribution Modeling

A statistical model known as "Maximum Entropy" (Maxent) (Phillips et al., 2006) was used to develop an envelope of likely areas of occurrence for each individual species across the B-D project. The species distribution models (SDMs) were used in conjunction with other predictor layers to determine species' abundance across the landscape as detailed in section 3.5.

In our application of Maxent we used 10 meter pixel resolution raster datasets that included a digital elevation model, minimum and maximum temperature (Holden et al., 2015), Daymet precipitation (Thornton et al.,1997), time integrated NDVI, heatload, and the first 3 principle components of Sentinel-2 satellite imagery for the B-D. Heatload was calculated using the 2010 10m DEM as an input into an ERDAS model that implements heatload equations from McCune et al., (2002). The resulting SDMs represent the affinity of each species to certain biophysical settings as well as relations to spectral information contained in the Sentinel-2 data.

3.5 Tree Dominance Type Modeling

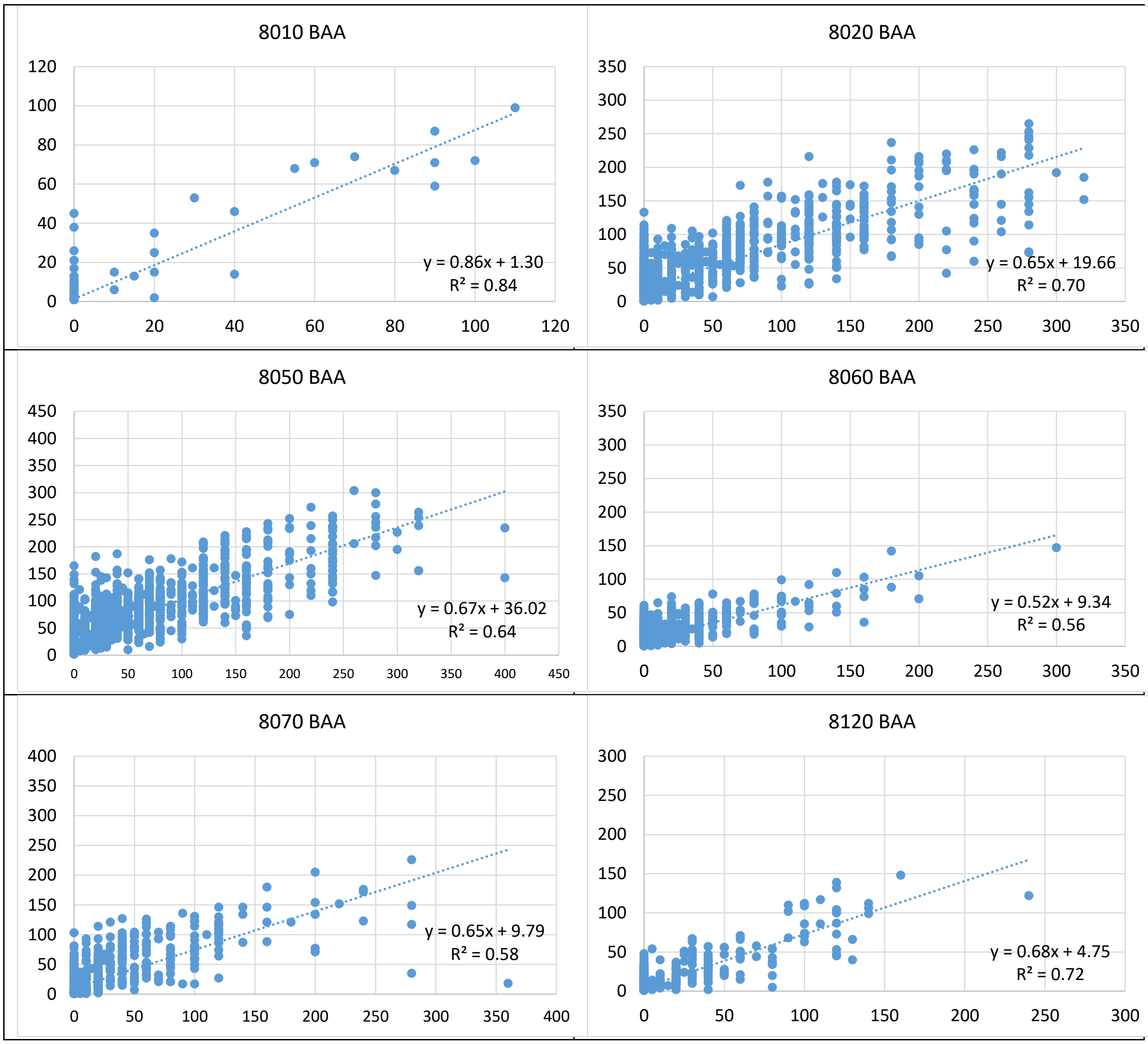
Tree dominance type was modeled using a Random Forest regression (Breiman 2001, Liaw and Wiener 2002, Liaw 2015) based on individual tree species abundance information collected at the field plot level as the response variable and independent variables of Sentinel-2 image derivatives, a biophysical derivative, a vegetation index derivative, and lastly a Maxent SDM that was converted to a distribution mask. In the SDM masks, values of 1 indicated potential species presence and values of 0 indicated the probable lack of presence. Species abundance was described in terms of square feet of basal area per acre (BAA), for each species encountered at each reference site. Using all of this information, a separate raster surface was built for each species, where the continuous range of BAA values was estimated for any given pixel.

In total, information from 1,405 reference sites were used to build regression models of BAA for each tree species being considered in this mapping project. The mapped species and number of associated reference sites for each species is given below in Table 4.

Table 4 - *Mapped species and number of associated reference sites for each species*.

Tree Species	Vmap Code	Sites with BAA > 0
Ponderosa pine	8010	19
Douglas fir	8020	544
Lodgepole pine	8050	897
Subalpine fir	8060	359
Engelmann spruce	8070	300
Whitebark pine	8120	183
Limber pine	8150	61
Juniper	8180	24
Total		2,387

The relationship between modeled and measured BAA values for all species is given graphically in Figure 9 – *Scatter plots representing relationships between measured and modeled basal area per acre plot description across the Beaverhead-Deerlodge National Forest, collected and processed in the 2017 field season.*, where models tended to explain more than 60% of the variation in measured BAA.



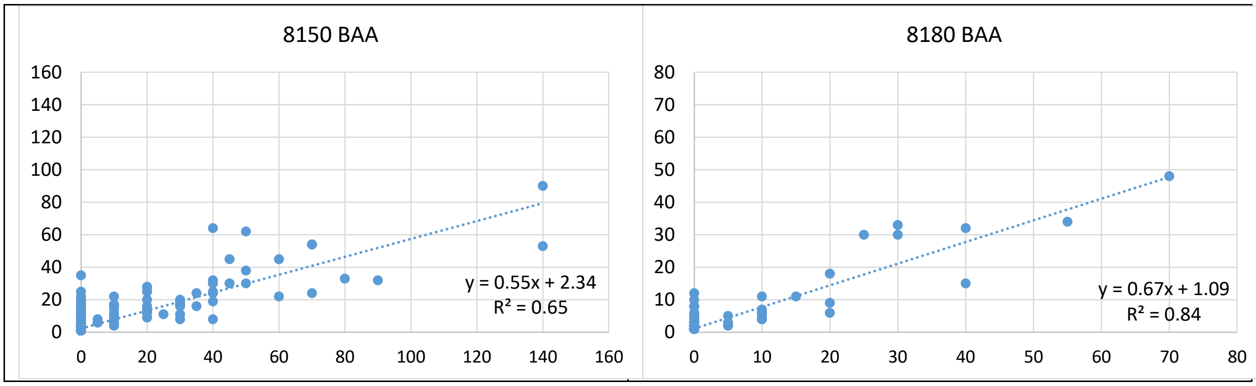


Figure 9 – *Scatter plots representing relationships between measured and modeled basal area per acre plot description across the Beaverhead-Deerlodge National Forest, collected and processed in the 2017 field season.*

The suite of species BAA raster data were then summarized to the VMap polygons to determine percent composition, and a dominance type label was assigned based on Region 1 Existing Vegetation Classification System tree dominance type rules.

As a final product, each polygon with a lifeform classification of Coniferous Tree, is assigned a tree species dominance type label. An ancillary table with all species associated with a given polygon is also produced, and can be used to infer cohort information. In the VMap database, fields describing the mean basal area per acre estimate for each species are given. They are labeled with the four letter species code followed by mean BAA. For example, the field for mean square feet of basal area per acre of Douglas-fir is: PSME_MEAN_BAA. A polygon-based accuracy assessment is given in the accuracy assessment section of this document.

3.6 Tree Mortality Modeling

The proportion of standing dead trees in every tree measurement plot was recorded in the field sampling during the 2017 season. Proportions were based on the number of dead standing trees versus live trees identified in variable radius plots.

In total 1,068 sites were used for model development, and 385 sites were used for model validation and assessment of residual errors. Figure 10 - *A comparison of tree mortality model development and model evaluation training data distributions.* illustrates the relative distribution of model development and model evaluation data.

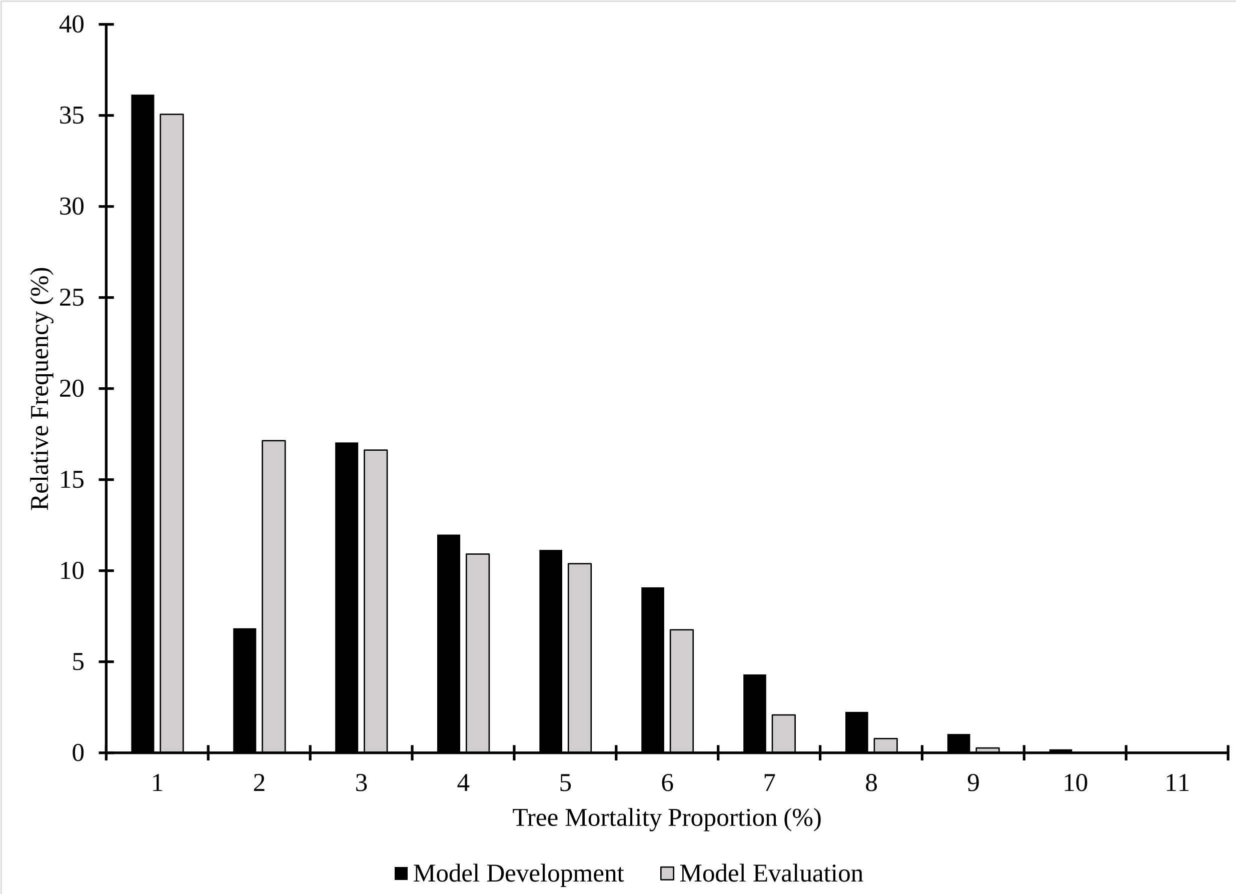


Figure 10 - *A comparison of tree mortality model development and model evaluation training data distributions.*

For input as independent variables, the same raster stack of spectral values from Sentinel-2 imagery, vegetation index, and biophysical index, and coordinate space used for tree dominance type modeling were used to model tree mortality.

The initial model representing the percentage of standing dead trees on a site was developed by creating a Random Forest regression of model development plot data against the independent variables raster stack. Using the withheld plot data, model error was assessed and modeled. The modeled residuals were then added back to the original output, and referred to as the residual adjusted model. Simple linear regressions of the initial mortality model and the residual adjusted models are given in Figure 11 - *Scatter plot of the initial tree mortality model.* and Figure 12 – *Scatter plot of the residual adjusted tree mortality model.*.

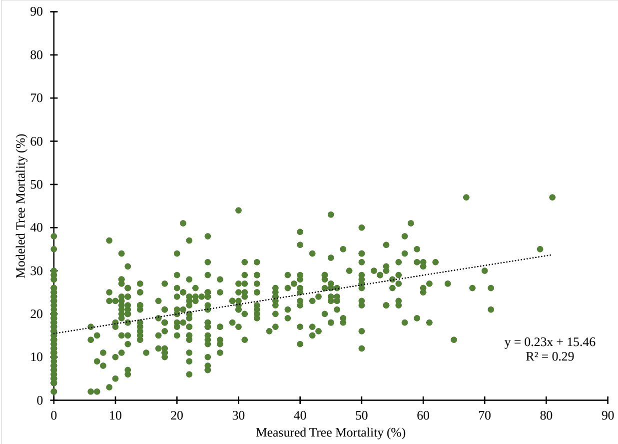


Figure 11 - *Scatter plot of the initial tree mortality model.*

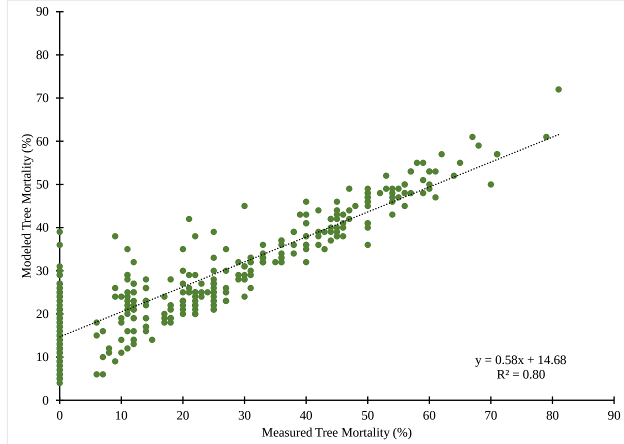


Figure 12 – *Scatter plot of the residual adjusted tree mortality model.*

The residual adjusted model represents a significant improvement in fit, where the proportion of explained variation increased from 0.29 to 0.80, and the slope of the relationship increased from 0.23 to 0.58. While model fit was greatly improved with the residual adjustment, low values were still being overestimated and high value were somewhat underestimated. To address this, the modeled output was further stretched. The stretched model explained the same amount of variability as the residual adjusted model, but the slope of the relationship between measured and modeled values increased from 0.58 to 0.73, and the intercept was lowered from 14.68% to 12.07%. The regression between the measured and stretched model output is shown graphically in Figure 13 – *Scatter plot of residual adjusted and stretched tree mortality model, given as a percentage of a measurement site on the Beaverhead-Deerlodge National Forest.*. The stretched raster output was then summarized to VMap polygons where the mean percent value is reported in an attribute named MORTALITYPERC.

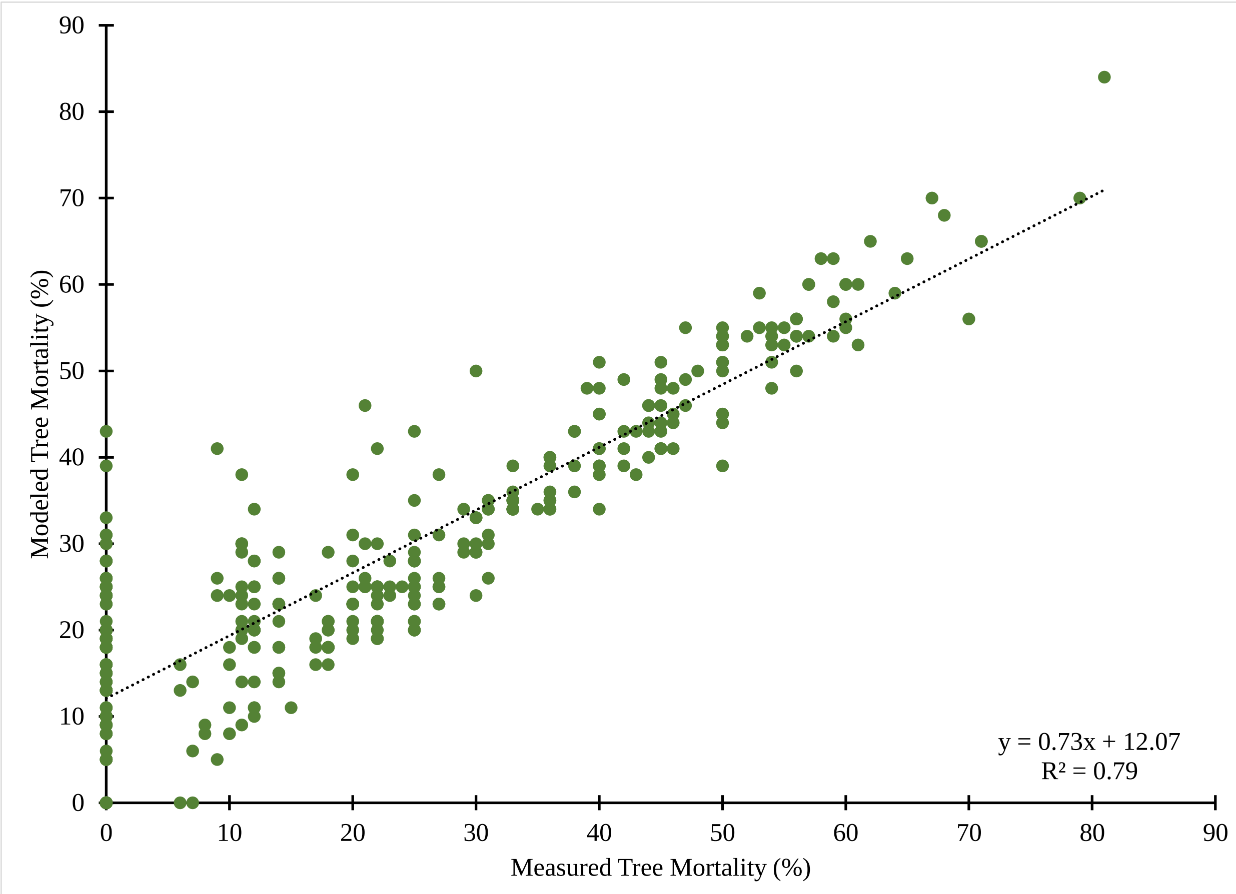


Figure 13 – *Scatter plot of residual adjusted and stretched tree mortality model, given as a percentage of a measurement site on the Beaverhead-Deerlodge National Forest.*

3.7 Rangeland Productivity Modeling

As part of this project we modeled rangeland productivity in terms of dry pounds of forage per acre across an area that encompasses the Beaverhead-Deerlodge National Forest. Productivity reference data were obtained from the Natural Resources Conservation Service, Soil Survey Geographic Database (SSURGO) soil dataset and used in combination with a time-integrated vegetation index, and the precipitation heatload and elevation adjusted topography biophysical index (PHEAT) to estimate forage production for low, average, and above average precipitation conditions.

These spatially explicit estimates of forage production across a range of precipitation conditions were produced as raster datasets with 10 m pixel resolution. Each estimate for low, average, and above average condition was summarized to the polygons in the VMap database, where the mean value for each polygon is represented for each condition is represented in the following attributes LOW_PRODUCTIVITY, AVERAGE_PRODUCTIVITY and HIGH_PRODUCTIVITY. The full Rangeland Productivity Report is available online at https://usfs.box.com/s/3mihxv8sozhij4615ukv317igx5rzko8.

3.8 Rangelands Characteristics Modeling

During the summer of 2017 the Montana Natural Heritage Program (MNHP) collected vegetation data on nonforested lands throughout the B-D project area. After field sampling was conducted, MNHP modeled cheatgrass presence, bunchgrass vs single-stem grass, dry grass canopy cover, shrub canopy cover and sagebrush canopy cover using collected plot data, biophysical characterization data, Sentinel-2 and NAIP imagery, and a number of derivatives from the imagery. The models produced by MNHP were used to create raster datasets of shrub canopy cover, as well as cheatgrass presence, and single-stem grass and bunch grass distributions across the Beaverhead-Deerlodge mapping area. These raster data were then summarized as a percent cover to the following four attributes within the VMap database: SHRUB_CCV, BRTE_CCV, BUNCHGRASS_CCV, and SINGLESTEM_CCV. Contents of the Montana Natural Heritage Program Rangeland Mapping Report (Fortier 2018) is given as Appendix I.

3.9 Whitebark Pine Characteristics Modeling

Whitebark pine is an ecologically and culturally valuable asset, and to help the Beaverhead-Deerlodge National Forest with management decisions associated with this species, potential range, restoration suitability, and relative percent canopy cover were estimated. Based on disturbance history and species preferences, estimates were made to describe sites where restoration of whitebark pine may be successfully focused. These estimates, were converted to a binary response, where the value 1 suggests suitability while 0 suggests a lower likelihood of successful restoration. In this version of the VMap database, the attribute is listed as PIAL_SUITABILITY. Further information about whitebark pine characteristics model can be found in Fortier 2018b, and online at
https://usfs.box.com/s/m9kh3583e1boocldi4fhv78gp9qk7wr4

4.0 Accuracy Assessment

An independent accuracy assessment of the VMap products was conducted across the entire Beaverhead-Deerlodge National Forest (B-D) mapping area to provide a validation of the issued data. An estimate of overall map accuracy and confidence of individual map classes was computed with a standard error matrix derived from the comparison of known reference sites to mapped data classed through the Region 1 Exiting Vegetation System. In general, the delivered B-D map products were validated with accuracies ranging from 63-98% depending on the class attribute. While the accuracy assessment was generally good for classified attributes, a comparison of independently observed versus modeled continuous outputs for tree canopy cover percent and average tree diameter was also evaluated with favorable results.

4.1 Error Matrices

Following the recommendations of Stehman and Czaplewski (1998), a stratified random sample design was used to select comparison sites across the B-D mapping area for the lifeform and tree canopy cover attributes and used to construct a standard accuracy assessment error matrix (Congalton, 1991). Sampling strata were constructed for the lifeform and tree canopy attributes and a minimum of 100 and 70 (respectively) spatially distributed samples per class were drawn from each strata. Assessments were conducted somewhat differently for the tree dominance type (DOM40) and tree size class attributes because it is difficult to accurately assess both of those attributes using straight image interpretation. Assessment of tree dominance type and tree size class attributes was therefore conducted by comparing classified values to a dataset of reference sites that comprised 10% of each assessment class and that was not used in the classification process.

For the lifeform attribute, evaluation sites were selected from 8 sampling strata: dry grass, wet grass, xeric shrub, mesic shrub, sparsely-vegetated, water, deciduous tree, and coniferous tree, with 5,088 sample sites selected and compared to the mapped VMap lifeform class in a standard error matrix, shown in Table 5.

Table 5 - *Beaverhead-Deerlodge National Forest VMap 2018 lifeform error matrix.*

	Reference Lifeform Class									
Mapped Lifeform Class	**Dry Grass**	**Wet Grass**	**Xeric Shrub**	**Mesic Shrub**	**Coniferous Tree**	**Water**	**Sparsely Vegetated**	**Deciduous Tree**	**Grand Total**	**Comission Error**
Dry Grass	612	9	23	1	2	0	3	0	650	94%
Wet Grass	37	591	2	12	4	0	0	0	646	91%
Xeric Shrub	28	6	597	4	7	0	2	2	646	92%
Mesic Shrub	4	32	7	566	11	0	1	5	626	90%
Coniferous Tree	8	2	2	2	627	1	4	2	648	97%
Water	1	3	1	0	1	629	7	2	644	98%
Sparsely Vegetated	26	0	6	1	8	2	598	0	641	93%
Deciduous Tree	4	9	2	19	17	1	3	532	587	91%
Grand Total	720	652	640	605	677	633	618	543	5,088	**Overall Accuracy**
Omission Error	85%	91%	93%	94%	93%	99%	97%	98%		93%

Tree dominance type was evaluated on 12 classes for DOM40: PIPO-IMIX, PSME-IMIX, , PICO-IMIX, ABLA-TMIX, PIEN-TMIX, PIAL-IMIX, PIFL-IMIX, JUNIP-IMIX, IMIX, and TMIX. A total of 1,362 samples, was compared to the resulting map to yield the error matrix shown in Table 6, below.

Table 6 - *Beaverhead-Deerlodge National Forest VMap 2018 DOM40 error matrix.*

	Reference DOM_MID_40 Class											
Mapped DOM_MID_40 Class	**PIPO-IMIX**	**PSME-IMIX**	**PICO-IMIX**	**ABLA-TMIX**	**PIEN-TMIX**	**PIAL-IMIX**	**PIFL-IMIX**	**JUNIP-IMIX**	**IMIX**	**TMIX**	**Grand Total**	**Comission Error**
PIPO-IMIX	6	0	0	0	0	0	0	0	0	0	6	100%
PSME-IMIX	6	342	73	3	10	0	2	2	0	0	438	78%
PICO-IMIX	0	59	493	28	21	16	9	0	4	3	633	78%
ABLA-TMIX	0	0	2	12	4	0	1	0	0	0	19	63%
PIEN-TMIX	0	4	13	3	65	0	0	0	0	2	87	75%
PIAL-IMIX	0	0	0	0	0	30	1	0	0	0	31	97%
PIFL-IMIX	0	2	0	0	0	0	6	0	0	0	8	75%
JUNIP-IMIX	0	0	0	0	0	0	0	3	0	0	3	100%
IMIX	0	7	26	12	11	8	1	0	2	1	68	3%
TMIX	0	8	19	15	23	2	1	0	0	1	69	1%
Grand Total	12	422	626	73	134	56	21	5	6	7	1,362	**Overall Accuracy**
Omission Error	50%	81%	79%	16%	49%	54%	29%	60%	33%	14%		70%

Tree canopy cover class evaluation sites were drawn from four sampling strata representing: low canopy cover tree (10-24.9%), moderate-low canopy cover tree (25-39.9%), moderate-high canopy cover tree (40-59.9%), and high canopy cover tree (60% +), with 100 initial sample sites selected from each strata. By selecting a minimum of 100 evaluation sites from each strata, a sufficient sample is still available if unsuitable sites are encountered due to excessive shadowing or site variability. In the case of tree canopy cover, 280 sites were evaluated. The results are displayed in Table 7 below.

Table 7 - *Beaverhead-Deerlodge National Forest VMap 2018 tree canopy cover error matrix.*

	Reference Canopy Cover Class					
Mapped Canopy Cover Class	**10-24.9%**	**25-39.9%**	**40-59.9%**	**60+%**	**Grand Total**	**Comission Error**
10-24.9%	**55**	**15**	**0**	**0**	**70**	**79%**
25-39.9%	**5**	**53**	**12**	**0**	**70**	**76%**
40-59.9%	**0**	**9**	**57**	**4**	**70**	**81%**
60+%	**0**	**3**	**13**	**54**	**70**	**77%**
Grand Total	**60**	**80**	**82**	**58**	**280**	**Overall Accuracy**
Omission Error	**92%**	**66%**	**70%**	**93%**		**78%**

For the tree size class assessment, evaluation sites were selected from four sampling strata, consisting of: seedling tree (0-4.9” DBH), small tree (5-9.9” DBH), medium tree (10-14.9” DBH), and large/very large tree (15”+ DBH), by a 10% withholding of the field sampled data within each class, for a total of 232 samples. These sites were then evaluated for classification into a corresponding VMap Tree Size class and compared with the existing Map. The results are displayed in Table 8 below.

Table 8 - *Beaverhead-Deerlodge National Forest VMap 2018 tree size class error matrix.*

	Reference Tree Size Class						
Mapped Tree Size Class	0-4.9" DBH	5-9.9" DBH	10-14.9" DBH	15-19.9" DBH	20" + DBH	Grand Total	Comission Error
0-4.9" DBH	10	0	0	0	0	10	100%
5-9.9" DBH	11	66	21	0	1	102	65%
10-14.9" DBH	0	17	56	16	4	94	60%
15-19.9" DBH	0	1	1	8	6	16	50%
20" + DBH	0	0	1	3	6	10	60%
Grand Total	22	84	79	30	17	232	Overall Accuracy
Omission Error	45%	79%	71%	27%	35%		63%

4.2 Discussion

There are tradeoffs to constructing a post-classification, stratified random sample-based accuracy assessment. The biggest advantage will be a guarantee of a sufficiently large sample size so that a full assessment of each represented class is possible. A disadvantage may be that the ability to estimate a true quantification of omission error is lost due to the biased nature of the sample selection. All things considered, however, the advantage of having the ability to assess within class accuracy outweighs this disadvantage.

Since not all of the map attributes lend themselves to confident visual interpretation, specifically tree size class and tree dominance type, it is necessary to withhold a certain amount of the field collected reference information in order to compute an independent estimate of the map class accuracy. The draw back to using withheld data is that there may not be enough data to withhold in some classes to provide a meaningful quantification of the error for such classes. This is the case in the B-D database for the classes of MX-PIPO, MX-JUNIP, IMIX, and TMIX where there are not enough withheld samples to provide a statistically valid estimate of the class accuracy. Fortunately these classes comprise a very small percentage of the overall landscape area mapped.

In general, the accuracies exhibited in the VMap 2018 database are very good. Classes with higher error rates, such as IMIX (shade intolerant species mix) and TMIX (shade tolerant species mix) may be over represented across the landscape, and are generally difficult to detect and describe because of their variable species composition. Therefore, it is possible that a mislabeled polygon could still be considered “OK” in most analysis situations.

The same can be said of the tree canopy cover and tree size class attributes, where most of the error occurs between adjacent classes and can easily be attributed to either interpretation error or just the inherent fact that when a continuous world is parceled into discrete classes not everything will always fit as neatly as expected. For example, if a given polygon is estimated to have 61% tree canopy cover, but the analyst estimates that it has 59%, the true difference is only 2%, but 59.9% is the cutoff between two classes so that the polygon would then be assessed as incorrect.

The take home message is that even the accuracy assessment, which is judged as “truth”, needs to be taken with a grain of salt. While the accuracy assessment attempts to quantify the error structure in the B-D map products, this is no substitute for a qualitative map evaluation prior to its use in any analysis.

5.0 References

Barber, J., D. Berglund, R. Bush. 2009. Region 1 Existing Vegetation Classification System and its Relationship to Inventory Data and the Region 1 Existing Vegetation Map Products. USDA Forest Service, Northern Region, CMIA Numbered Report 09-03 5.0. Available online at: http://fsweb.r1.fs.fed.us/forest/inv/classify/r1_ex_veg_cmi_4_09.pdf. Last accessed September 8, 2017.

Barber, J., S.R. Jr. Brown, R. Ahl. 2012. Mid-level and Base-level Databases of the R1 Existing Vegetation (VMap) Products. USDA Forest Service. Region One Vegetation Classification, Mapping, Inventory and Analysis Report 12-38Available online at: https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5364531.pdf; last accessed January 25, 2017.

Bechtold, W. A., P. Patterson. Editors. 2005. The Enhanced Forest Inventory and Analysis Program – National Sampling Design and Estimation Procedures. USDA Forest Service General Technical Report SRS-80. 85 p.

Berglund, D., R. Bush, J. Barber, and M. Manning. 2009. R1 Multi-level Classification, Mapping, Inventory, and Analysis System. USDA Forest Service, Northern Region, CMIA Numbered Report 09-01 v2.0. Available for download at: http://fsweb.r1.fs.fed.us/forest/inv/classify/cmia_r1.pdf. Last accessed September 8, 2017.

Breiman, L. 2001. Random forests. Machine Learning 45(1):5-32. doi:10.1023/A:1010933404324

Brohman, R. and L. Bryant. 2005. Existing vegetation classification and mapping technical guide. USDA Forest Service, Washington Office, Ecosystem Management Coordination Staff. Available for download at: https://www.fs.fed.us/emc/rig/documents/integrated_inventory/FS_ExistingVEG_classif_mapping_TG_05.pdf. Last accessed September 8, 2017.

Brown, S., J. Barber. 2012. The Region 1 Existing Vegetation Mapping Program (VMap) Flathead National Forest Overview; Version 12. Numbered Report 12-34. Available online at: https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5366381.pdf. Last accessed September 8, 2017.

Congalton, R. G. 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote sensing of environment. 37: 35-46.

European Space Agency (ESA). 2015. Sentinel-2 User Handbook. Available online at: https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook, last accessed November 14, 2018.

Fortier, J. 2018. Beaverhead-Deerlodge National Forest: 2017 VMap Rangeland Mapping Report. *Unpublished Report to United States Forest Service Northern Region District Office.* Montana Natural Heritage Program, Helena, Montana. 21 pp. iii.

Fortier, J. 2018b. Modeling Whitebark Pine in the Beaverhead-Deerlodge National Forest: potential range, restoration suitability, and relative percent of canopy cover. Unpublished report to the USDA Forest Service. Montana Natural Heritage Program, Missoula, MT. 13 pp.

Gorelick, N. M. Hancher, M. Dixon, S. Iiyushchenko, D. Thau, R. Moore. 2017. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone.

Haralick, R.M., L.G. Shapiro. 1985. Image segmentation techniques. Computer Vision, Graphics and Image Processing. 29(1): 100-132.

Holden, Z. A., Swanson, A., Klene, A. E., Abatzoglou, J. T., Dobrowski, S. Z., Cushman, S. A., and Oyler, J. W. (2015). Development of high-resolution (250 m) historical daily gridded air temperature data using reanalysis and distributed sensor networks for the US Northern Rocky Mountains. International Journal of Climatology.

Jolliffe, I.T. 2002. Principal Component Analysis and Factor Analysis. P. 150-160 in Principal Component Analysis. Springer Series in Statistics. Springer, New York. ISBN: 978-0-387-95442-4

Kellner, F., S. Brown. 2018. Beaverhead-Deerlodge Rangeland Productivity Modeling. USDA Forest Service, Northern Region Geospatial Group, Project Report ???. Available online at: https://usfs.box.com/s/3mihxv8sozhij4615ukv317igx5rzko8

Liaw, A. 2015. Documentation for R package randomForest. Available online at: https://cran.r-project.org/web/packages/randomForest/randomForest.pdf; Last accessed January 25, 2017.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News 2(3):18-22. Available online at: http://CRAN.R-project.org/doc/Rnews/; Last accessed January 25, 2017.

Lillesand, T., Kiefer R. W., Chipman J. 2015. Remote Sensing and Image Interpretation, 7th Edition. John Wiley and Sons, Inc. New York. ISBN: 978-1-118-34328-9.

McCune, B., & Keon, D. (2002). Equations for potential annual direct incident radiation and heat load. Journal of vegetation science, *13*(4), 603-606.

Phillips, S. J., R.P. Anderson, and R.E. Schapire. Maximum entropy modeling of species geographic distributions, Ecological modelling 190.3-4 (2006): 231-259

Reed, B. C., J. F. Brown, D. VanderZee, T. R. Loveland, J. W. Merchant, and D. O. Ohlen, 1994: Measuring phenological variability from satellite imagery. J. Veg. Sci., 5, 703–714.

Rouse, J.W, R.H. Haas, J.A. Scheel, and D.W. Deering. 1974. Monitoring Vegetation Systems in the Great Plains with ERTS. Vol 1. P. 48-62 in Proceedings, 3rd Earth Resource Technology Satellite (ERTS) Symposium. Available online at: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022592.pdf. Last accessed September 8, 2017.

Stehman, S. V., and Czaplewski, R. L. 1998. Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. Remote Sensing of Environment. 64: 331-344.

Thornton, P.E., Running, S.W., White, M.A. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology 190: 214 - 251.

USDA Farm Service Agency. 2016. Four Band Digital Imagery Information Sheet. Available online at: https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/APFO/support-documents/pdfs/fourband_infosheet_2016.pdf ; Last accessed February 14, 2017.

USDA Farm Service Agency. 2015. National Agriculture Imagery Program (NAIP) Information Sheet. Available online at: https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/APFO/naip_info_sheet_2015.pdf; Last accessed January 25, 2017.

USDA Forest Service. 2017. Geospatial Data: Northern Region Existing Vegetation Mapping Program. Available online at: https://www.fs.usda.gov/detailfull/r1/landmanagement/gis/?cid=stelprdb5331054&width=full; Last accessed September 8, 2017.

Zaitoun, N.M., M.J. Aqel. 2015. Survey on Image Segmentation Techniques. Procedia Computer Science. 65: 797-806.

6.0 Appendix I

Beaverhead-Deerlodge National Forest:
2017 VMap Rangeland Mapping Report

Prepared for:
US Forest Service Northern Region Geospatial Group; Engineering
Prepared by:
Joseph Fortier

P.O. Box 201800 ● 1515 East Sixth Avenue ● Helena, MT 59620-1800 ● 406-444-5354ii

For full report refer to:

Fortier, J. 2018. Beaverhead-Deerlodge National Forest: 2017 VMap Rangeland Mapping Report. *Unpublished Report to United States Forest Service Northern Region District Office.* Montana Natural Heritage Program, Helena, Montana. 21 pp. iii

SUMMARY

Vegetation Map (VMap) is a geospatial database of existing vegetation that relates to spatially unique polygons across the Northern Region of the USDA Forest Service. This landscape scale mapping product utilizes consistent vegetation classification that is updated every 5 to 10 years in order to remain current with large scale landscape disturbances.

Vegetation management is one of the Forest Service's primary responsibilities. The VMap database provides existing vegetation data that facilitates broad scale analysis and enables managers to accurately address resource planning and monitoring objectives.

During the field season of 2017 the Northern Region Geospatial Group (NRGG) began the process of revising the VMap database for the Beaverhead-Deerlodge National Forest (B-D NF) and surrounding landscape with support from the Montana Natural Heritage Program (MTNHP). MTNHP was tasked with gathering rangeland vegetation data to inform the VMap process in non-forested lands and to explore what degree of rangeland vegetation types, species, or relationships could be accurately portrayed spatially by applying similar data and methods used in VMap forest mapping.

This report describes the mapping and modeling efforts for rangeland shrub and grass in the B-D NF and surrounding lands. Five ecosystem components were modeled based on the 2017 field work, including total shrub and total grass canopy cover, as well as subcategories sage cover, bunchgrass and single-stem grass cover. A sixth component, *Bromus tectorum* (Cheatgrass) abundance, was modeled using a combination of VMap field data and other presence location information provided by the B-D NF. All ecosystem components were modeled using a Random Forest classification and regression algorithm and produced as continuous cover raster products across the rangeland in southwest Montana. 5

1 Study Area

The Beaverhead-Deerlodge (B-D) National Forest and surrounding landscape considered in the 2017 VMap update includes over 8.25 million acres in Southwest Montana. Within the study area approximately 60% of the land is publicly owned, including over 3 million acres of National Forest land, 1 million+ acres under Bureau of Land Management administration, and the remainder a mix of state and local ownership. This report focuses on the rangeland areas within the larger B-D study area, where dominant vegetation is a mix of shrubs and perennial grass/forbs (Figure 1). The area includes large proportion of sagebrush habitat, dominated by Big Sagebrush (*A. tridentata*) and Threetip Sagebrush (*A. tripartita*), as well as associated species like Sagewort and Rabbitbrush, and common grass species Idaho Fescue (*Festuca idahoensis*), Bluebunch Wheatgrass (*Elymus spicatus*), and Junegrass (*Koeleria macrantha*).

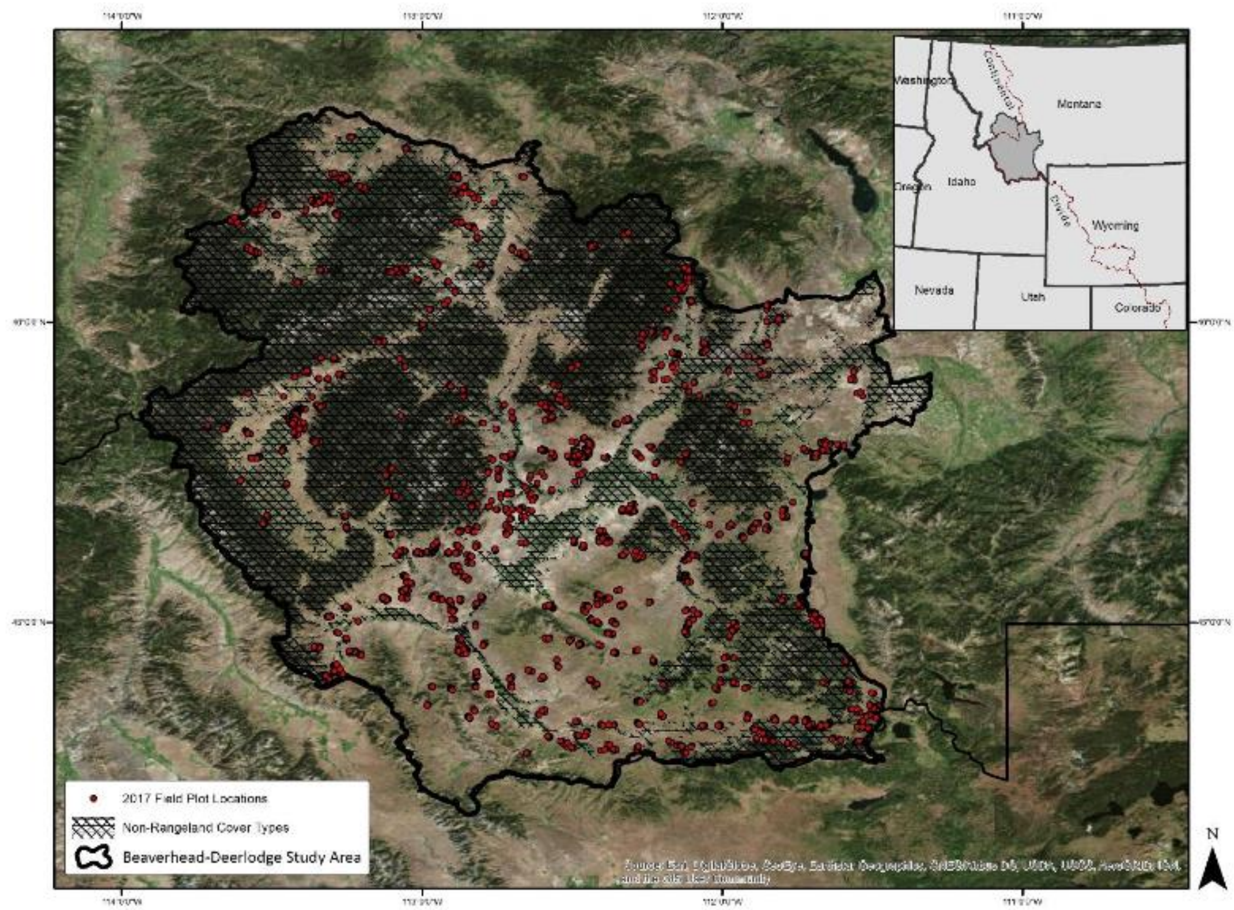


Figure 1. Study area of the Beaverhead-Deerlodge National Forest VMap mapping effort. Location of rangeland field collection sites are identified in red.

2 Data

The initial mapping process began by testing established VMap data and modeling procedures (Ahl and Brown 2017), specifically the use of 10-meter NAIP aerial photography, to model non-forest ecosystem components in the B-D, producing only limited success. Subsequent efforts aimed to tease out the complex signatures of each component by testing a variety of additional data sources from spectral transformations to biophysical parameters to muti-seasonal and multi-platform imagery. Table 1 lists all the variables tested as independent predictors over the course of the model testing process, though not all were subsequently used in the final products.

Dataset	Native Resolution	Source
Image Data		
2013 NAIP	1m	US Farm Service Agency
Spectral Bands 1-4 (Visible & NIR)		
Principal Components		
NDVI		
Green NDVI		
Variance Texture		
2015 NAIP	1m	US Farm Service Agency
Spectral Bands 1-4 (Visible & NIR)		
Variance Texture		
Sentinel-2	10 - 15m	US Geological Survey / GoogleEarth engine
Spectral Bands 2-8, 11 &12 (Visible, Red edge, NIR, SWIR)		
NDVI		
NDWI		
Landsat (average 2013-2018)	30m	US Geological Survey / GoogleEarth engine
mean April NDVI		
mean May NDVI		
mean June NDVI		
mean July NDVI		
April-May NDVI Change		
May-June NDVI Change		
June-July NDVI Change		
April-July NDVI Range		
Biophysical Variables		
Digital Elevation Model	10m	US Geological Survey
Elevation		
Slope		
Slope Curvature		
Folded Aspect (Southwestness)		
Topographic Position Index		
Hillshade		
Precipitation Heat & Elevation Adjusted Topography (PHEAT)	10m	USFS Region 1 Geospatial Group
Combined PHEAT-NDVI	30m	USFS Region 1 Geospatial Group
Human Distrubance Index	30m	Montana Spatial Data Infastructure

Table 1. List of predictor variables assembled for random forest modeling of rangeland vegetation.

2.1 Image Data

Three sources of remotely sensed imagery were acquired for this analysis: NAIP aerial photography collected at 1m spatial resolution, Sentinel-2 multi-spectral satellite imagery collected at 10 to 15m spatial resolution, and Landsat satellite imagery collected at 30m resolution. All data were resampled to 10m spatial resolution to match the minimum mapping unit/pixel size of this analysis and to match the sampling size of the field data, which were collected in 50-foot diameter plots.

National Agriculture Imagery Program (NAIP) 4-band aerial photography is the preferred spectral dataset used for the VMap project because of its high spatial resolution and cloud-free image coverage statewide, collected on a 3-to-5 year acquisition cycle. For this study 2013 NAIP was selected as the primary NAIP imagery, being the most current image collection over the study area where the data were captured during the summer months (June & July). 2015 NAIP was also tested in this analysis to provide a multi-seasonal spectral view of the landscape since the aerial collection of that year' data occurred from late-September through November. Unfortunately, due to the late season collection select flight lines crossing the project boundary included heavy snow cover leaving large data gaps in the 2015 product. Both years were processed to generate a stack of Principal Components as well as a spectral variance texture product and degraded from 1m to 10m resolution following the VMap methods outlined by Ahl and Brown (2017). Additionally, both a Normalized Difference Vegetation Index (NDVI =NIR-Red/NIR+Red) and Green NDVI (gNDVI =NIR-Green/NIR+Green) were calculated from the 10m spectral data. These spectral transformations and the original spectral bands (blue, green, red, NIR) were all provided as possible predictor variables to the mapping algorithm.

Sentinel-2 data was provided by the USFS Northern Region Geospatial Group as a 10-band multispectral product (visible through short-wave infrared wavelengths) already pre-processed into 10m resolution. To provide seamless, cloud-free Sentinel coverage of the study area the USFS acquired the imagery through Google Earth Engine, extracting mean spectral values across multiple dates and image paths. From the Sentinel data both NDVI and a Normalized Difference Wetness Index (NDWI =NIR-SWIR/NIR+SWIR) were calculated and provided alongside the multispectral bands as model variables.

Google Earth Engine was also used to calculate, and extract 5-year mean NDVI values by month from April through July. These months were selected to capture the primary growing season of non-irrigated grass in Southwest Montana, which typically experiences final snow melt and early green-up in April and nears complete senescence by late July. Monthly NDVI change (April to May, May to June, June to July) and seasonal range products were also generated as inputs to the modelling process.

2.2 Biophysical Data

A 10 meter DEM was used to produce a slope surface, aspect, slope curvature and a topographic position index (Jeness et al. 2013) for use in the modelling process under the assumption that topographic location correlates highly with access to moisture and nutrients important for vegetation growth. Two products developed by the USFS Northern Region Geospatial Group were included as ancillary variables to the modeling algorithm: a Precipitation Heat & Elevation Adjusted Topography (PHEAT), which attempts to synthesize precipitation, heat load, and water holding capacity, as well as a combination of PHEAT with a 30-year mean historical NDVI calculated from Landsat (Ahl and Brown 2017). Finally, a contiguous map of human disturbance index developed by the Montana Natural Heritage Program (Newlon 2015) was included in the modelling of potential Cheatgrass invasion.

All data were resampled to 10m resolution, projected into Universal Transverse Mercator zone 12N, and clipped to the study area.

2.3 Field Data

A total of 1,916 field plots from the 2017 VMap/MTNHP field season were used for this analysis. Plot data were collected on public lands using a stratified sampling method, where strata were defined along unique moisture and vegetation gradients and the allocated density of plots in each were weighted proportional to the area of the respective strata within the overall study region. At each plot 25 foot transects were laid out in four cardinal directions from the plot center and species were recorded at 2 foot intervals along each line. Species data and cover estimates for shrub, grass, and forb were calculated from these 50 sample points; ocular estimates of cover were used to validate or modify the final data recorded for each plot. Percent of ground cover occupied by rock, litter, or bare ground were also recorded, as well as number and species of trees present. Following field collection, the data points were individually reviewed on computer screen for spatial inaccuracy and data anomalies in order to flag and remove from analysis any questionable, redundant or incomplete data. Field plot locations within the study area are displayed in Figure 1 and a full breakdown of the collection methodology, as well as data gathered is covered in a companion report by Dressing and Fortier (2018).

An additional 581 locations of Cheatgrass presence/abundance were provided by the B-D National Forest. While the MTNHP plot data identifies *Bromus tectorum* on a continuous scale of 0-100% canopy cover the additional point dataset identified Cheatgrass abundance as falling in one of 5 categories: 0-1%, 1-5%, 5-25%, 25-50% and >50%. As a result, *Bromus tectorum* abundance was modeled as a categorical output following the aforementioned breakdown with exception of the >50% category, which was dropped due to lack of examples (only 5 out of 2,651 records).

Field data for each ecosystem component were randomly partitioned such that 70% of the point locations would be used to calibrate the statistical model and the remaining 30% set aside to provide independent validation of the model outputs. This resulted in a 1,212/704 split of calibration to validation points for the five canopy cover products and a 1,856/795 split for *Bromus tectorum*.

3 Modeling Process

Modeling was performed using a Random Forest (RF) classification and regression algorithm (Breiman 2001, Liaw and Wiener 2002). Each RF was set to build forests out to 2,001 trees, with each tree selecting 66% of the calibration data available to it for growing the model and the remaining 33% used to estimate Out Of Bag (OOB) error.

For each ecosystem component multiple RF model runs were performed to test the array of independent variables with increasing or decreasing complexity, ex. Model 1 – NAIP spectral only, Model 2 – NAIP and biophysical data, Model 3 – NAIP PCA, etc. The purpose of this process was not only to highlight the combination of variables which produce the highest statistical accuracy output but also to show how each variable affected the statistical output and provide insight into their efficacy as predictors of rangeland vegetation patterns. After an average of 10-15 model tests per component, including one that provided all possible predictors, the top 8-12 variables were selected according to their impact on RF accuracy (i.e., their inclusion resulted in a noticeable reduction in Root Mean Square Error). The top predictors selected as important for modeling each ecosystem component were then included in a final RF model used to produce a spatial prediction across the entire rangeland area (Table 2).

Random Forest Model Variable Selection

	Grass CC	Bunchgrass CC	Single-stem Grass CC	Shrub CC	Sagebrush CC	Cheatgrass
1	Texture	Total Grass CC	Total Grass CC	Sentinel B11 (SWIR)	Total Shrub CC	MT Human-Distubance Index
2	April-July NDVI Range	April NDVI	DEM	Sentinel B5 (Red edge)	DEM	April-May NDVI Change
3	Sentinel B11 (SWIR)	Sentinel B12 (SWIR)	April NDVI	Texture	Texture	Slope
4	Sentinel B8 (NIR)	May NDVI	NAIP B1 (Red)	NAIP B4 (NIR)	Sentinel NDWI	May-June NDVI Change
5	NAIP B4 (NIR)	DEM	Sentinel B8 (NIR)	NAIP B1 (Red)	Sentinel B8 (NIR)	May NDVI
6	NAIP B1 (Red)	Sentinel B8 (NIR)	May NDVI	Sentinel B8 (NIR)	Slope	April NDVI
7	NAIP PCA comp1	Slope	Sentinel B11 (SWIR)	NAIP B2 (Green)	Sentinel B3 (Green)	June-July NDVI Change
8	Slope	Sentinel B4 (Red)	Sentinel NDWI	NAIP B3 (Blue)	NAIP B4 (NIR)	April-July NDVI Range
9	PHEAT	April-July NDVI range	April-July NDVI Range		NAIP B3 (Blue)	July NDVI
10		Texture	NAIP B4 (NIR)		Sentinel B4 (Red)	Texture
11						Sentinel B11 (SWIR)
12						Sentinel B3 (Green)

Table 2. Predictor variables used in the final classification or regression model for each ecosystem component. Variables are listed in descending order of importance; variables at the top of each list provided the highest reduction in RMSE according to the OOB error calculated by the RF, variables not selected were either superfluous or redundant as their inclusion provided little additional benefit to the statistical accuracy of the model.

The ecosystem components bunchgrass, single-stem grass, and sagebrush are each subcategories of either the total grass canopy cover or total shrub cover components. In those three models their overarching components (grass, shrub) were included in their respective RF models as predictor variables. Despites these predictors obvious correlations with the subcomponents being modeled this step was deemed necessary to act as a type of top down constraint, reducing errors in which the subcomponent might otherwise predict a higher canopy cover than the super-component (i.e., modeled sage cc > total shrub cc). Doing so not only reduced statistical errors it also provides an ease of understanding for any end user who might interact with the full collection of modeled results.

4 Results

The five percent canopy cover components and Cheatgrass abundance were produced as continuous maps across the entire study region. These maps are displayed at the end of this paper in APPENDIX I, cropped to areas of rangeland vegetation as identified by the VMap lifeform classification process.

4.1 Regression Analyses

Using the field data set aside for independent validation Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and r2 regression statistics were calculated for each of the continuous, canopy cover models (Table 3). Rangeland grass cover proved the most difficult to model at this spatial resolution, RMSE of 14.17 and r2 of 0.215, due in part to its spectral signature being near-indistinguishable from surrounding litter and shrub once senesced. The two key predictor variables selected by the RF were texture, which showed an inverse correlation with grass, and seasonal range in Landsat measured NDVI. The bunchgrass and single-stem grass models struggled to find a spectral signature for the same reasons, although their predictive accuracy measures were improved from total grass due, in part, to their unique spatial arrangements. Single-stem grasses were more prevalent in the lowland field plots with decreasing abundance as sample locations moved away from the private and heavily managed valley bottoms, and vice versa.

Modeled Component	RMSE	MAE	r^2
% Grass Canopy Cover	14.17	10.93	0.215
% Bunchgrass Canopy Cover	11.90	8.68	0.525
% Single-stem Grass Canopy Cover	7.61	3.30	0.616
% Shrub Canopy Cover	10.06	5.17	0.443
% Sagebrush Canopy Cover	6.90	4.55	0.795

Table 3. Statistical accuracy of the regression models calculated using independent validation plots.

Total shrub cover fared slightly better than grass (RMSE 10.06) bolstered by the texture variable, which showed the highest direct correlation to shrub cover as any other variable/component pair tested. Sagebush was perhaps the least complicated variable for the RF to model as 90% of sage points were dominated by only a single species (Big Sagebrush), meaning that there were not a lot of intra-component variations in the spectral signature. Sage cover also followed certain geographical trends that would make it easier to model, specifically sage density increased as field plots moved away from the valley floors, and higher concentrations of sage exist in the southern half of the study area compared to the northern half.

The r2 values for all five regression models indicate that no prediction was a perfect correlation to the field estimates. Despite the abundance of predictor variables available to the classifier their combination were still not enough to properly capture the unique signals both within and between cover types across the landscape. As a result, the RF generated what amounts to a statistical "best-fit" models that tends to under classify both the low (0-10%) and high ends of the cover predictions.

One promising result of this analysis is that the statistical accuracies are similar to, or within range of other research published on rangeland vegetation mapping. Table 4 lists reported RMSE statistics from three other studies that modeled similar ground cover components in sagebrush habitats of the western United States. All three studies listed worked with Landsat data at the 30m resolution to predict biomass, which shows that this mapping approach can provide an improvement on the spatial resolution of the final product without a significant sacrifice to accuracy.

	Grass/Herb Cover	Shrub Cover	Sage Cover
Jones et al., 2018	14.9	9.9	*NA*
Xian et al., 2015	7.52	8.93	4.25
Homer et al., 2009	8.55	6.53	6.75
This Study	14.17	10.06	6.9

Table 4. Comparison of published Root Mean Square Error statistics from similar studies modeling studies that focus on rangeland cover types in western USA.

4.2 Categorical Analysis

The RF classification model was able predict Cheatgrass abundance with 95% overall accuracy; per class accuracies ranged from 78-98% for the four broad abundance categories mapped. The model seemed to perform the best when predicting the lowest (no Cheatgrass) and highest (25-50% Cheatgrass) abundance categories where the presence or absence signals are likely the clearest (Figure 2). Initial tests modeling *Bromus tectorum* as a continuous (0-100%) estimate of abundance proved challenging due to the relative scarcity of occurrence in the MTNHP field plots as well as the sparse cover exhibited where it was recorded. The inclusion of the additional occurrence data and partitioning of abundance into cover categories dramatically improved the model results, although remnants of those early predictive challenges are still evident in the class confusion common between the low to moderate Cheatgrass abundance categories.

Cheatgrass *(Bromus tectorum)* Cover Percentage

		Field Data						
		Category 1 0-1%	Category 2 1-5%	Category 3 5-25%	Category 4 25-50%		Error of Ommission	Error of Comission
Classifier Result	0-1%	603	7	12	3	625	1.6%	3.5%
	1-5%	7	43	5	0	55	20.4%	21.8%
	5-25%	2	4	62	0	68	21.5%	8.8%
	25-50%	1	0	0	46	47	6.1%	2.1%
		613	54	79	49	795		

Overall Accuracy =	94.8%
Kappa =	0.863

Figure 2. Confusion matrix and accuracy statistics for the categorical Cheatgrass (Bromus tectorum) abundance model.

5 Discussion

Iterative variable testing showed that, at the 10-meter resolution, image texture was consistently one of the most often selected and important variable for decreasing RMSE in all cover components modeled. Shrub height relative to the surrounding vegetation was shown to increase textural variance as the taller vegetation cast shadows visible in the 1m spectral data. Larger shrub types, like Big Sagebrush (*A. tridentata*), therefore produced the most consistent correlation between abundance and image texture and the ubiquity of that species across the study area (present in 62% of all shrub plots / 90% of sage plots) helps explain the increased accuracy in modeling shrub and sage, when compared to the grass cover models. Similar relationships of shrub presence to spectral texture were not as obvious when a texture product was generated on Sentinel or Landsat spectral data, making it clear that the utility of this variable is H-resolution dependent (Strahler et al., 1986) and almost certainly a product of shadow visibility. To exploit this, multiple tests were performed where texture metrics extracted from 2015 NAIP were included to the model in addition to 2013, which due to their collection at different times of the year offer changing shadow lengths and positions. This provided an almost multi-temporal texture dimension to the RF algorithm resulting in significantly improved results in both shrub and grass models, however the large swaths of the snow covered ground in 2015 NAIP imagery made it impractical to apply those results to the entire study area.

The inclusion of spectral data proved important to each of the model's predictive power, but the type of spectral data did not seem to matter significantly; NAIP, PCA, or Sentinel data improved the results relatively equally. While the top variables selected for the final models seemed to identify some key bandwidths of importance, NIR and SWIR were common, testing showed that swapping out the NAIP near-infrared band for the Sentinel equivalent had little to no significant change on the model's accuracy. Their properties and the information the RF was extracting from them were highly correlated, even despite the Sentinel data providing additional spectra. Similar results were found with the inclusion of biophysical variables; the inclusion of each had some positive effect on the RMSE but which variable(s) included to the model within each of those categories were, in most cases, trivial.

The pseudo-phenology from the monthly Landsat NDVI did prove useful, particularly in mapping Cheatgrass abundance, which fits with the findings other published studies that have leveraged Landsat time-series for modeling rangeland biomass. In this case the relative discrepancy between the field plot size (182 m2) and the 900 m2 area of a Landsat pixel used to develop the NDVI data probably complicated any signal the RF was able to identify. Future work generating more robust phenology metrics with either Sentinel or a fusion of Sentinel and Landsat NDVI at a more appropriate scale would likely improve 10m resolution modeling going forward.

The final observation from iterative testing was in the relationship between the training data and RF results: While not perfect matches, comparing the RF reported OOB RMSE against the independent validation RMSE showed remarkably similar values in every test. This provided confidence in the OOB error metrics, showing that model comparison could be done without the use of independent validation and allowing for multiple tests on the number of calibration points provided to the algorithm, including using 100% of the field data to train the RF. Accuracies consistently improved with the increase in training examples, demonstrating the importance of a robust set of training data and making an argument for attaining more whenever possible. Similarly, since the RF algorithm generates a "best-fit" to model the statistical trends in the data given, it is important that the field data accurately capture the variance of the landscape being modeled. By using a stratified field data collection approach the VMap methodology provided confidence that the training set used here adequately covered the B-D rangeland, however without equal access to collect examples on private land it is possible that these models lose predictive power in those locations.

Overall the results of this work are promising and informative. NAIP imagery alone offers little clarity for rangeland components, with its limited spectral and temporal ranges, but it does provide the high resolution, contiguous coverage necessary to exploit some important textural variances in rangeland vegetation. Landsat scale analyses, on the other hand, have shown promise in capturing biomass signatures from the phenology of rangeland vegetation but are limited in their ability to identify clear spectral signatures due to the scale of the sensor and the generally sparse, mixed vegetation cover in rangelands. This analysis attempted to work on both levels by letting the modeling algorithm choose which variables provide the clearest picture of each rangeland component and from that predict those rangeland features across the entire study area.

REFERENCES

Ahl, R. and S. Brown. 2017. The Bitterroot and Lolo National Forests Region 1 Existing Vegetation Database (VMap) Revision of 2016. *USDA Forest Service. Northern Region Geospatial Group Project Report.* Available online at: https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fseprd573432.pdf

Breiman, L. 2001. Random Forests. *Machine Learning*. 45(1): 5-32.

Dressing, E. and J. Fortier. 2018. Beaverhead-Deerlodge National Forest: 2017 VMap Rangeland Report. *Unpublished Report to United States Forest Service Northern Region District Office.* Montana Natural Heritage Program, Helena, Montana. 20 pp.

Homer, C.G., C.L. Aldridge, D.K. Meyer, M.J. Coan, and Z.H. Bowen. 2009. Multiscale sagebrush rangeland habitat modeling in Southwest Wyoming. *U.S. Geological Survey Open-File Report.* 2008-1027, 14p.

Jenness, J., B. Brost and P. Beier. 2013. Land Facet Corridor Designer: Extension for ArcGIS. Jenness Enterprises. Available at: http://www.jennessent.com/arcgis/land_facets.htm

Jones, M.O., B.W. Allred, D.E. Naugel, J.D. Maetas, P. Donelly, L.J. Metz, J. Karl, R. Smith, B. Bestelmeyer, C. Boyd, J.D. Kerby, and J.D. McIver. 2018. Innovation in rangeland monitoring: annual, 30 m, plant functional type percent cover maps for U.S. rangelands 1984-2017. *Ecosphere.* 9(9)

Liaw, A. and M. Wiener. 2002. Classification and Regression by randomForest. *R News* 2(3): 18-22.

Newlon, K. R. 2015. Montana Human Disturbance Index. Available from https://mslservices.mt.gov/Geographic_Information/Data/DataList/datalist_Details.aspx?did=%7B639e7c86-8224-11e4-b116-123b93f75cba%7D

Strahler, A.H., C.E. Woodcock, J.A. Smith. 1986. On the nature of models in remote sensing. *Remote Sensing of Environment.* 20, 121-139

Xian, G., C. Homer, M. Rigge, H. Shi, D. Meyer. 2015. Characterization of shrubland ecosystem components as continuous fields in the northwest United States. *Remote Sensing of Environment.* 168: 286-300 17

APPENDIX I.