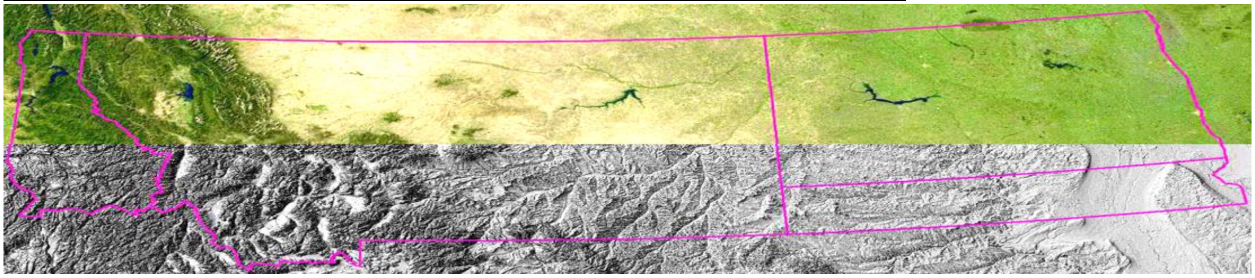
Project Report: NRGG_PR_IPKNF_VMAP2017_08_15_2017



The Idaho Panhandle and Kootenai National Forests Region 1 Existing Vegetation Database (VMap) Revision of 2017.

August 15, 2017

Northern Region Geospatial Group (NRGG)
USDA Forest Service
Authors: Steve Brown, Robert Ahl, PhD, and Fred Kellner

Table of Contents

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 2017) 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 2017, an updated VMap database was produced for the Idaho Panhandle and Kootenai National Forests (IPKNF). 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 (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 IPKNF is, however, located in a region that is persistently cloaked in clouds, and overcoming such challenging conditions required more processing than usual. In addition to frequent cloud cover, the mapping area of interest was also obscured by forest fire smoke in 2015. For these reasons, continuous high resolution NAIP imagery (USDA Farm Service Agency 2015, 2016) was not available for the full extent of the IPKNF update area. Thus, to obtain contemporary and full coverage imagery of the mapping area, RapidEye high resolution satellite data (Planet.com 2017) was sourced between July 18, 2016 and August 8, 2016 to capture existing vegetation patterns. The imagery was delivered with 5 meter pixel resolution, and five 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. To reveal the cloud obscured areas, cloud patches in the RapidEye data were masked, and coded as no data. Those areas of no data were then supplemented with cloud free Landsat 8 data. In 2016, no entirely cloud free Landsat data were available either. Nonetheless, a full area composite Landsat 8 scene (USDI Geological Survey 2017) was assembled with image data captured between June

4 and August 16, 2016. Areas obscured by clouds in this dataset were substituted with could free data acquired June 16, 2015.

Finally, a composite of cloud free RapidEye and Landsat 8 image data was used to segment the image and create a vector-based layer of polygons that represent a delineation of stand boundaries across the study area (Haralick and Shapiro 1985, Zaitoun and Aqel 2015). This set of polygons, or stands, are the elements attributed by the VMap process and issued in the final database. In the field, reference information was collected and used to make spatial predictions of the vegetation attributes contained in the database. Predicted raster surfaces of the attributes were summarized to the delineated polygons.

As draft map products were created, they were reviewed and appropriate changes were made in the labeling algorithms. 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 IPKNF, and ranged from 52-90%, 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 was 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 also 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 unknown areas. For the development of the VMap database, training data was specifically collected 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/fPiXL.

2.2 Image Data Collection and Pre-Processing

Three distinct types of spectral image data were used in the production of the VMap database, and include NAIP, Rapid Eye, and Landsat 8. 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 full coverage over the mapping area was not available for this project. While NAIP image data was not available for algorithm development, analysts did use it for visual inspection and to provide context when evaluating algorithm results. RapidEye image data was specifically obtained for the IPKNF project, and was delivered with 5 meter pixel resolution, and 5 bands of radiometric resolution (Planet.com 2017). For algorithm development in this project, Rapid Eye imagery had the finest spatial resolution but because it was just a single snapshot in time it was obscured by roughly 10% cloud cover. To compensate for the lack of data in the cloud obscured regions of the Rapid Eye image, Landsat 8 image data (USDI Geological Survey 2017) with 30 meter pixel resolution and 7 bands of radiometric resolution were also collected over the IPKNF mapping area.

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 IPKNF 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 IPKNF process were based on Rapid Eye and Landsat 8 data. As a first step in the derivative creation process, a principal component analysis (PCA) (Jolliffe 2002) of the five bands of RapidEye image data was conducted, and 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 their final application. In a similar fashion, Landsat 8 data were also transformed into a three band principal component raster. Focal and contrast derivatives were not created for the Landsat 8 data because of the course resolution of the raw imagery, but the PCA raster was resampled to 10 meter pixel resolution to facilitate integration with other datasets.

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). For our purposes we did not use the simple NDVI, but instead computed the TINDVI for a 30-year period record for the growing season months (July to September) from Landsat data and is referred to as the vegetation index derivative, TINDVI.

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 elements of vegetation and other patterns on the landscape, and all VMap attributes are associated with the polygons derived from the segmentation process.

3.0 Mapping Process

3.1 Lifeform Classification

The lifeform attribute is mapped with 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. Mapped lifeforms classes include tree, shrub, herbaceous, sparsely vegetated, and water with precedence order being tree, shrub, herbaceous in the lifeform key.

3.2 Tree Canopy Cover Classification

For polygons where a tree lifeform has been 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.

Canopy cover models were based on reference data obtained through analyst-based image interpretation, and a Random Forests regression algorithm (Breiman 2001, 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. 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 1,000 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 values greater than 60%, were generated and used as training data in the modeling process. The selection of reference sites were used in combination with the RapidEye image derivatives in a Random Forest regression model to estimate the full range of canopy cover values across the mapping area.

In the case of the IPKNF, a second round of modeling was accomplished in a similar way, using the same reference data to model continuous canopy cover values across the mapping area with Landsat 8 derivatives. The second round of modeling was implemented to fill in the holes created by the cloud mask. Thus, locations where a cloud mask was present, were filled with estimates based on Landsat 8 estimates. For all other locations, estimates are based on RapidEye image derivatives alone.

The resulting continuous canopy cover values were 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.

3.3 Tree Size Class

Tree size class is modeled from field collected data that quantifies basal area weighted average tree diameter at breast height (BAWDBH), as described in the Region 1 Exiting Vegetation Classification System. BAWDBH was computed from a variable radius plot to the nearest full inch. In a process similar to canopy cover modeling, data from reference sites were associated with image derivatives, a vegetation index derivative, and a biophysical derivative, and used in a Random Forest regression model (Breiman 2001, Liaw and Wiener 2002, Liaw 2015) to estimate continuous tree size values for every pixel.

For all polygons classified as the tree lifeform, individual BAWDBH pixel values were summarized and the mean BAWDBH was associated with those polygons. Due to the cloud mask, estimates were first generated with Rapid Eye image derivatives, and then repeated with Landsat 8 derivatives. No data values in the Rapid Eye estimates were replaced with values generated with Landsat 8 derivatives. The resulting mean values in all tree class polygons were grouped into the Region 1 Existing Vegetation Classification System tree size classes ranging from 0-4.9, to 5.0-9.9, 10.0-14.9, 15-19.9, 20.0 – 24.9, and greater than 25 inches dbh mean values for stands.

3.4 Tree Dominance Type

Similar to tree size, 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, RapidEye image derivatives, a biophysical derivative, and a vegetation index derivative. A separate raster surface was built for each species, where a continuous range of percent abundance values represent the potential abundance of a given species in any given pixel. This process was repeated using Landsat 8 image derivatives, the same biophysical derivative, and the same vegetation index derivative. In locations where cloud masks were present in the RapidEye based raster surfaces, no data values were filled in with Landsat 8 based estimates. Thus a combination raster surface was created for each species of interest. The suite of species abundance raster data were then summarized to the VMap polygons to determine percent composition and a dominance type label was then assigned based on R1 Existing Vegetation Classification System tree dominance type rules.

4.0 Accuracy Assessment

An independent accuracy assessment of the VMap products was conducted across the entire IPKNF 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 R1 Ex-Veg system. In general, the delivered IPKNF map products were confirmed with exceptional accuracies, ranging from 52-90+% depending on the class attribute. While the accuracy assessment was generally satisfactory 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 IPKNF mapping area for and used to construct a standard accuracy assessment matrix (Congalton, 1991). Sampling strata were constructed for both the lifeform and tree canopy cover attributes, and a minimum of 100 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 with 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 7 sampling strata: dry grass, wet grass, mesic shrub, sparsely-vegetated, water, deciduous tree, and coniferous tree, with 6,209 sample sites selected and compared to the mapped VMap lifeform class in a standard error matrix. The results are shown in Table 1 below.

Table 1. Idaho Panhandle & Kootenai NF VMap 2017 Lifeform Error Matrix

IPKNF VMap 2017 Lifeform Error Matrix									
Lifeform Class	Dry Grass	Wet Grass	Mesic Shrub	Coniferous Tree	Water	Sparsely Vegetated	Deciduous Tree	Grand Total	Comissin Error
Dry Grass	826	11	11	24	2	17	1	892	93%
Wet Grass	24	815	12	2	5	2	6	866	94%
Mesic Shrub	12	7	837	39	0	3	2	900	93%
Coniferous Tree	4	0	12	877	0	2	4	899	98%
Water	2	10	1	4	873	5	2	897	97%
Sparsely Vegetated	13	2	8	25	6	830	4	888	93%
Deciduous Tree	2	2	22	44	5	10	782	867	90%
Grand Total	883	847	903	1015	891	869	801	6209	Overall Accuracy
Omission Error	94%	96%	93%	86%	98%	96%	98%		94%

Tree dominance type was evaluated on 12 classes for DOM40: PIPO-IMIX, PSME-IMIX, ABGR-TMIX, LAOC-IMIX, PICO-IMIX, ABLA-TMIX, PIEN-TMIX, THPL-TMIX, TSME-TMIX, TSHE-TMIX, IMIX, and TMIX. A 10% withholding for each class, with a total of 2,583 samples, was compared to the resulting map to yield the error matrix shown in Table 2, below.

Table 2. Idaho Panhandle & Kootenai NF VMap 2017 DOM40 Error Matrix

IPKNF VMap 2017 DOM_MID_40 Error Matrix														
DOM_MID_40 Class	PIPO-IMIX	PSME-IMIX	ABGR-TMIX	LAOC-IMIX	PICO-IMIX	ABLA-TMIX	PIEN-TMIX	THPL-TMIX	TSHE-TMIX	TSME-TMIX	IMIX	TMIX	Grand Total	Comission Error
PIPO-IMIX	198	37	2	10	7	0	0	0	0	0	7	0	261	76%
PSME-IMIX	25	504	15	33	19	0	2	5	2	0	25	4	634	79%
ABGR-TMIX	1	12	119	4	0	0	0	0	6	1	4	7	154	77%
LAOC-IMIX	1	21	3	273	21	2	2	4	2	0	30	6	365	75%
PICO-IMIX	2	5	0	17	297	2	0	0	1	4	8	0	336	88%
ABLA-TMIX	0	0	0	2	6	190	9	0	0	0	0	2	209	91%
PIEN-TMIX	0	0	0	4	1	5	84	2	1	0	0	14	111	76%
THPL-TMIX	2	5	4	4	1	0	1	119	15	1	5	17	174	68%
TSHE-TMIX	0	1	1	1	2	0	0	1	73	0	4	5	88	83%
TSME-TMIX	0	2	1	0	2	6	0	0	1	101	6	6	125	81%
IMIX	8	11	1	15	7	0	0	3	1	1	15	1	63	24%
TMIX	0	6	4	1	4	9	10	3	6	1	7	12	63	19%
Grand Total	237	604	150	364	367	214	108	137	108	109	111	74	2583	Overall Accuracy
Omission Error	84%	83%	79%	75%	81%	89%	78%	87%	68%	93%	14%	16%		77%

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 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, 356 sites were evaluated. The results are displayed in Table 3 below.

Table 3. Idaho Panhandle & Kootenai NF VMap 2017 Tree Canopy Cover Error Matrix

IPKNF VMap V17 Tree Canopy Cover Class Error Matrix						
Canopy Cover Class	10-24.9%	25-39.9%	40-59.9%	60+%	Grand Total	Comission Error
10-24.9%	37	10	5		52	71%
25-39.9%	8	48	15	2	73	66%
40-59.9%	3	11	64	22	100	64%
60+%			13	118	131	90%
Grand Total	48	69	97	142	356	Overall Accuracy
Omission Error	77%	70%	66%	83%		75%

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 514 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 4 below.

Table 4. Idaho Panhandle & Kootenai NF VMap 2017 Tree Size Class Error Matrix

IPKNF VMap V2017 Tree Size Class Error Matrix							
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	24	19				43	56%
5-9.9" DBH		60	78			138	43%
10-14.9" DBH			116	120		236	49%
15-19.9" DBH				58	30	88	66%
20" + DBH					9	9	100%
Grand Total	24	79	194	178	39	514	Overall Accuracy
Omission Error	100%	76%	60%	33%	23%		52%

4.2 Regression Statistics

To better understand outcomes of the continuous variable modeling, from which tree canopy cover, and tree size classes were derived, some regression statistics were computed against the values of the withheld reference sites and the predicted surfaces.

Of the 400 sites assessed for percent canopy cover the average difference in observed – predicted was 0.43%, indicating a slight over prediction of canopy cover. The mean absolute error (MAE), which is the average of the absolute difference between observed and predicted values, and is similar to a margin of error (ME) figure, is +/- 9.5%. The multiple R, or correlation coefficient, which describes the degree of linearity between the two data sets, was 0.82 indicating a strong positive linear relationship between observed and predicted values. Finally, the p value was $<$ 0.05 (3.6684E -98) which indicates no statistical difference between the means of the observed and predicted populations.

For the tree size model, the average difference between observed and predicted was -1.1 inch, which indicates that the model generally under-predicted DBH. This is also confirmed in the error matrix for tree size class as all of the quantified error is due to the withheld samples being "undersized" by one class. This is further explained by the MAE value of +/- 3.2", which would put the precision of the model outside of the bounds of the 5" class breaks used by the R1 Existing Vegetation Classification System. As a frame of reference, we looked at the margin of error for both a 90% and 95% confidence interval from plot data for 1,884 stands containing 11,536 plots on the Idaho Panhandle National Forest. The margin of error for the 90% CI was +/-

2.7” and for the 95% CI it was +/- 3.2”. This suggests that the precision of the tree size model is within the bounds of what one could expect from average DBH estimates derived from summarized stand exams for the project area. Similar to percent canopy cover, the correlation coefficient was 0.74, illustrating a strong positive linear relationship between variables, and the *p* value of of < 0.05 (1.2076E -99) also suggests there is no statistical difference in the observed versus predicted population means.

4.3 Discussion

There are tradeoffs to constructing a post-classification, stratified random sample-based accuracy assessment. The biggest advantage may be a guarantee of a sufficiently large sample size so that a full assessment of each represented class is possible. A major 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 useful to withhold a certain amount of the field collected reference information and 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 was evident in the IPKNF database, where more classes were represented in the database (i.e., PIAL-IMIX) than there were enough samples of to provide a statistically valid estimate of the class accuracy.

In general, the accuracies exhibited in the VMap 2017 database are very good, and generally exceed national standards. Classes with higher error rates, such as IMIX (shade intolerant species mix) and TMIX (shade tolerant species mix) may be under 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 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 IPKNF 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.

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

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

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.

Planet.com. 2017. RapidEye Analytic Ortho Tiles. Description available at: https://www.planet.com/products/satellite-imagery/rapid-eye-analytic-ortho-tile/. Last accessed September 8, 2017.

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.

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.

USDI Geological Survey. 2017. Landsat 8 Imagery available for download at https://glovis.usgs.gov/. Last accessed: 8 Sep 2017.

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