Region One
Vegetation Classification, Mapping, Inventory and Analysis Report

Numbered Report 12-35 **April, 2012**

The Region 1 Existing Vegetation Mapping Program (VMap) Flathead National Forest Methodology; Version 12

Robert S. Ahl
Stephen R. Brown Jr.
USDA Forest Service, Region 1
Engineering Staff
200 E. Broadway, Missoula, MT 59802

1. DATA

Image Data

The two cost effective data sets available for interpretation and mapping are Landsat Thematic Mapper (TM) and National Agriculture Imagery Program (NAIP) imagery. Both of these datasets have properties that add value to mapping work, and are available at no additional cost to the public.

The TM is considered a moderate resolution sensor, with 30 meter pixels, but it has high radiometric resolution allowing for increased discrimination between vegetation types. The footprint of a Landsat scene covers nearly 100 square miles, providing for an effective for landscape-level mapping tool because it provides consistent radiometric values over large areas. The Landsat constellation of satellites has been in orbit since 1972 and provides a long history of use in vegetation mapping and monitoring through that time period.

Imagery provided by NAIP affords a higher spatial resolution, using 1 meter pixels, yet lacks some of the radiometric characteristics provided by TM data. The NAIP imagery consists of just 4 bands of data, spanning the visible spectrum into the near-infrared (NIR). The main drawback of this imagery is that each image tile covers approximately 30 square miles and therefore radiometric readings can be less consistent across large landscapes. The polygon-based map units delineated from these data are very accurate as compared to what can be accomplished from Landsat alone, and the secondary statistics derived from NAIP are useful for detailed delineation of various cover types.

In short, both Landsat TM, and NAIP imagery have useful properties for interpretation and mapping purposes. Landsat provides consistent and refined spectral values over large areas, while NAIP provides high spatial resolution which is useful for delineation, texture analysis, and visual interpretation. When used in combination these two image products complement each other very well, and provided the foundation for the development of the Flathead VMap V12 Database. Below is a brief description of the image products used in this project.

- Landsat Thematic Mapper imagery: A mid-summer image collected on July 21, 2009 was selected to capture “peak greenness” vegetation prior to senescence. Landsat TM images are distributed with 30 meter pixel resolution and seven bands of spectral information. We used bands 1, 2, 3, 4, 5, and 7 in this project. All TM images were orthorectified to the color infrared NAIP imagery, and radiance/reflectance corrected. The 30 meter pixel product was ultimately resampled to 10 meters and used in combination with NAIP data for quantitative analysis.

- National Agriculture Imagery Program data: NAIP imagery used in this project was also collected in July of 2009, and is provided with four spectral bands including the blue, green, red and NIR components. The original digital images were delivered with a 1 meter ground sample distance (GSD) and rectified to National Mapping Standards at the 1:24,000 scale. This imagery was used in two distinct ways. In one application, the original 1 meter resolution data were used for visual inspection and interpretation in the mapping process. Secondly, the high resolution data were ultimately resampled to a 10 meter pixel product and used in combination with TM data for quantitative analysis.

Image Derivatives and Ancillary Data

In addition to the values provided by the raw imagery, a variety of image derivatives and vegetation indices were computed from both datasets. The combination of the two image sources provides abundant spectral and texture-based information that is very useful for landcover mapping.

Image derivatives computed from the TM data include: a tasseled cap (TC) transformation, the first three principal components (PCA) of the TM data , and the first three principal components calculated on the TC transformation.

Derivatives of the NAIP imagery include: calculation of a normalized difference vegetation index (NDVI), quantification of intensity-hue-saturation (IHS), and the extraction of the first three principal components of the four band data. In addition to these spectral interpretations, two measures of image texture, which are based on the standard deviation of the first principal component, were computed for the four band NAIP image, within a 5x5 pixel window. The first measure of texture accounts for the mean standard deviation within the analysis window, while the second measure records the minimum standard deviation within the analysis window. The mean texture is useful for delineating edges of patches and the minimum texture is useful for discriminating differences within patches. Texture derivatives in general are useful for interpretations of roughness related to vegetation types, canopy cover, and tree size estimates.

Ancillary datasets used to describe biophysical setting are also important variables that are incorporated to better model the influence that topographic factors have on the type, structure, distribution, and abundance of vegetation across the landscape. A 10 meter resolution digital elevation model (DEM), obtained from the National Elevation Dataset (NED) was used to characterize and quantify topography, and produce a variety of topographic derivatives that provide biophysical interpretations.

All of the direct and derived classification variables used in the production of Flathead VMap V12 Database, are listed in Table 1. The various image, image derivatives, and topographically based products are used throughout the VMap production process.

Table 1. *A description of image and topographic variables used in the production of the Flathead VMap, V12 Database*

Image Input	Image Description
MEANIHSC1	NAIP CIR intensity
MEANIHSC2	NAIP CIR hue
MEANIHSC3	NAIP CIR saturation
MEANIHSR1	NAIP RGB intensity
MEANIHSR2	NAIP RGB hue
MEANIHSR3	NIAP RGB saturation
MEANCNDVI	NAIP CIR normalized difference vegetation index
MEANCPCA1	NAIP CIR 1st principal component
MEANCPCA2	NAIP CIR 2nd principal component
MEANCPCA3	NAIP CIR 3rd principal component
MEANNAIP1	NAIP band 1: red
MEANNAIP2	NAIP band 2: green
MEANNAIP3	NAIP band 3: blue
MEANNAIP4	NAIP band 4: near infrared
MEANTM1	LANDSAT TM band 1: blue
MEANTM2	LANDSAT TM band 2: green
MEANTM3	LANDSAT TM band 3: red
MEANTM4	LANDSAT TM band 4: near infrared
MEANTM5	LANDSAT TM band 5: mid infrared
MEANTM7	LANDSAT TM band 7: mid infrared
MEANTPCA1	LANDSAT TM 1st principal component
MEANTPCA2	LANDSAT TM 2nd principal component
MEANTPCA3	LANDSAT TM 3rd principal component
MEANTC1	LANDSAT TM tassled cap transformation: brightness
MEANTC2	LANDSAT TM tassled cap transformation: greenness
MEANTC3	LANDSAT TM tassled cap transformation: wetness
MEANTCPCA1	LANDSAT TM 1st principal component of the tassled cap transformation
MEANTCPCA2	LANDSAT TM 2nd principal component of the tassled cap transformation
MEANTCPCA3	LANDSAT TM 3rd principal component of the tassled cap transformation
MEANTRS1	DEM derived fully illuminated hillshade: band 1
MEANTRS2	DEM derived fully illuminated hillshade: band 2
MEANTRS3	DEM derived fully illuminated hillshade: band 3
MEANTXTME	NAIP mean texture within a 5x5 5m window
MEANTXTMI	NAIP minimum texture within a 5x5 5m window
MEANASR	DEM derived annual solar radiation estimate
MEANELE	DEM derived elevation
MEANTRI	DEM derived topographic ruggedness index based on 5x5 10m window
MEANTSP	DEM derived slope and aspect transformation
MEANDIS	DEM derived topographic dissection index based on 5x5 10m window

2. MODELING UNIT CONSTRUCTION

Model Areas

To make the 30 meter Landsat TM and the 1 meter NAIP data useable for image processing, both sets of data were resampled to 10 meters using a cubic convolution procedure. At 10m resolution, data sets are still quite large, and to accommodate the capabilities of current USFS computers, discrete mapping areas were created. The individual mapping areas are referred to as sub-models, or simply models. Another advantage of creating smaller modeling units is that different vegetation types could be modeled more effectively as all types do not occur in the same proportions in all models. The model delineations were based on the combination of sixth code watershed boundaries and the Flathead National Forest administrative boundary. Specifically, the overall mapping boundary was established by the intersection of watershed boundaries and the Flathead National Forest (FNF) administrative boundary, where all watershed areas that intersected the FNF boundary were selected for mapping. This provided full coverage of the area within a reasonable distance from National Forest System (NFS) lands. The individual model areas were defined in a similar way, focusing on the interaction between Ranger District and watershed boundaries. In all, seven sub- models were created to cover the entire FNF, ranging in size from 360,000 to 680,000 acres, with an average size of roughly 550,000 acres. The largest mapping area was in the Bob Marshall Wilderness Complex. Final model area boundaries are shown in Figure 1.

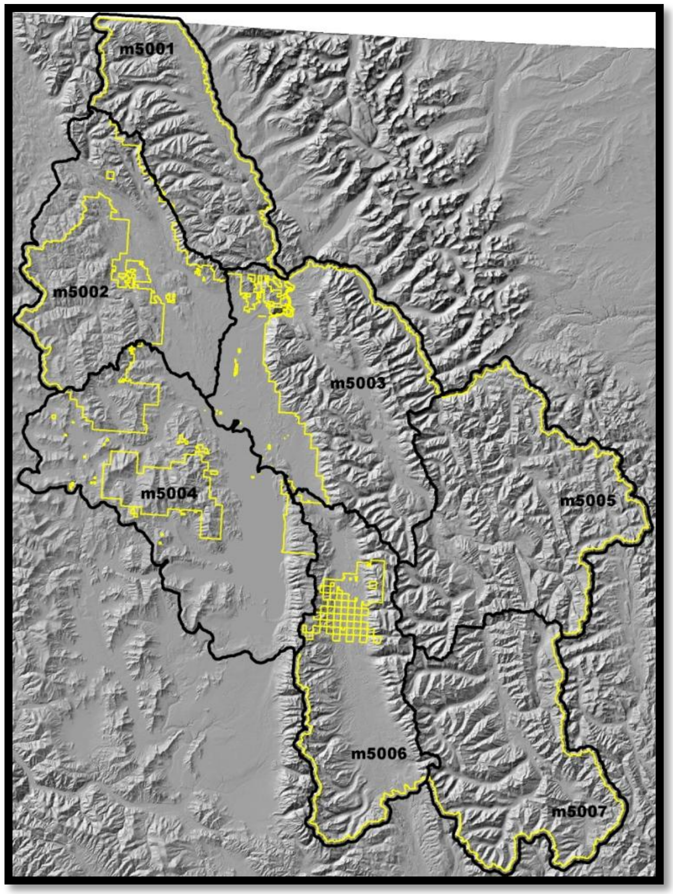


Figure 1. *Vegetation modeling units within the Flathead National Forest are labeled as m5001- m5007 and illustrated by heavy black lines. The thin yellow outline represents the Flathead NF administrative boundary*

Image Segmentation

Image segmentation is the process of combining pixels within digital images into spatially cohesive regions. These individual regions are called image objects and represent distinct areas within the image. The resulting image objects are inherently more data rich than individual pixels, and form the building blocks upon which image classifications are built (Haralick and Shapiro 1985, Ryerd and Woodcock 1996). Ultimately, the raster-based image objects are converted to vector-based polygons with associated image statistics as attributes. The segmentation process is performed using a proprietary software package, Definiens' eCognition, and is based on the local variance structure within imagery and User defined shape indices. These image objects effectively depict elements of vegetation and other patterns on the landscape (McDonald et al. 2002).

The initial segmentation, completed on a model area basis, is of a moderate resolution, based on defined scale parameter along with shape and spectral metrics. The segmentation is then classified into the basic lifeform classes of 1) sparsely-vegetated, 2) nonforest vegetation (a combination of Herbaceous and Shrub types), 3) forest, 4) and water, using membership functions and/or nearest neighbor algorithms within the eCognition software. A classification-based segmentation is subsequently applied to each of the mapped lifeform classes. By focusing on each defined class, variance appropriate delineations can be achieved. Specifically, multiple polygons that constitute a lake will be merged into a single polygon representing the lake. Likewise, many small polygons representing rocky ridges will be allowed to grow into bigger polygons because distinctions between rock types are generally not considered important to maintain. Polygons representing the nonforest vegetation will generally be re-segmented into smaller polygons to capture elements of detail that are important in rangeland communities. Conversely, polygons representing forest vegetation will be re-segmented to yield larger units to allow for some variation within forest stands. Results of the classification-based segmentation produce the base level polygons in the VMap database.

A subsequent segmentation is them performed that allows for a “coarsening” of the dataset by increasing the allowable size of the base-level polygons, thereby forming the mid-level database structure. The base polygons are then completely nested within the larger mid-level polygons. This is achieved by forcing the mid-level segmentation to constrain its delineation on the boundaries of the base polygons. A comparison of base and mid polygon dimensions is given in Figures 2 and 3.

Figure 2 illustrates results of image segmentation on the FNF in Submodel 5001, displayed over 1m NAIP color infrared imagery. Distinctions between lifeform classes such as sparse vegetation, grass, water, and forest can easily be determined. Similarly, differences in forest canopy cover and reflectance are also clearly visible, and delineated by the segmentation process. Figure 3 is an illustration of mid-level polygons over the same area, showing greater generalization while maintaining a similar pattern.

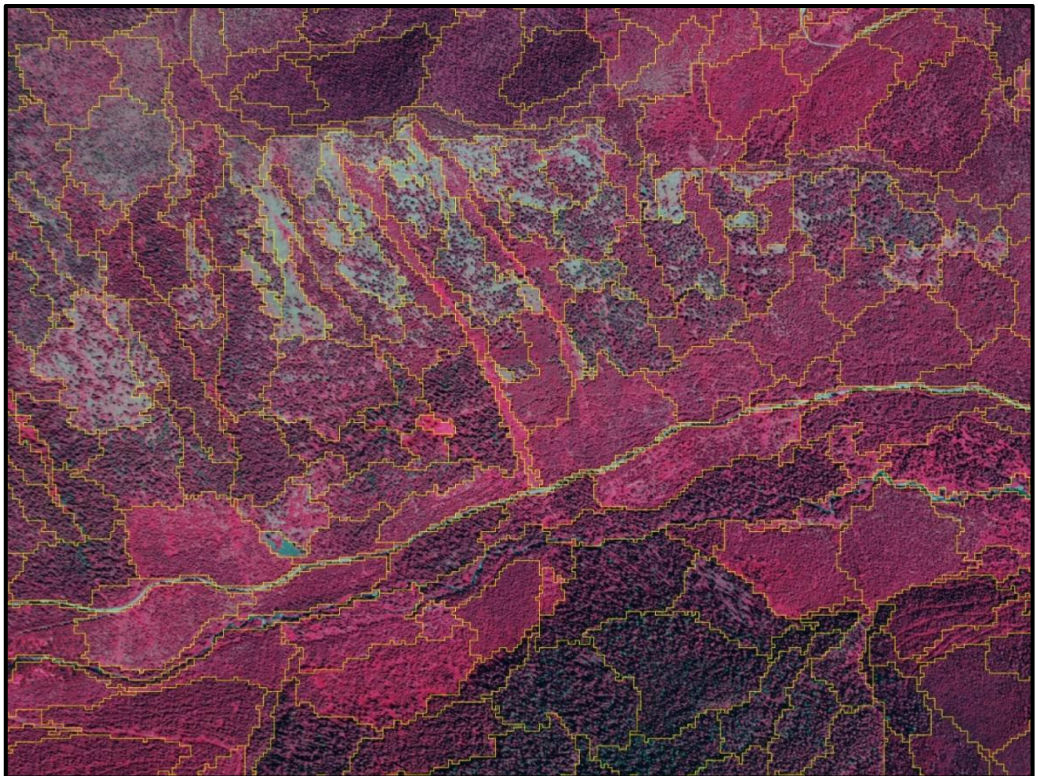


Figure 2. *Illustration of base-level VMap V12 polygons in the Flathead National Forest*

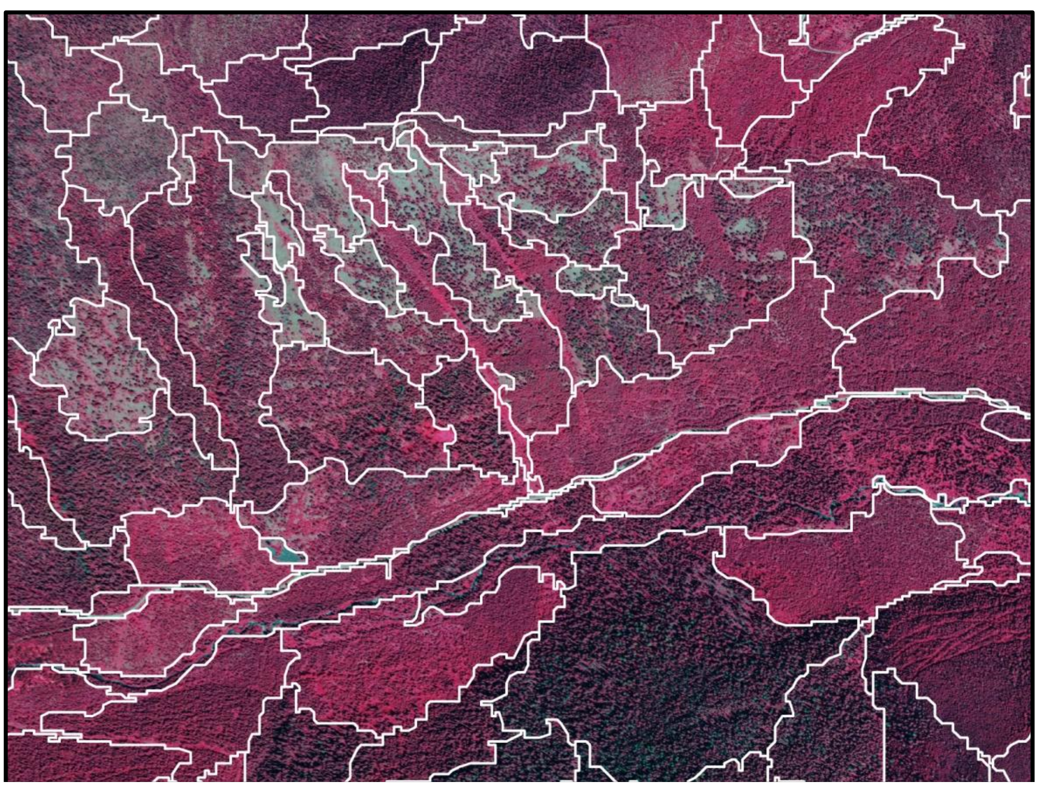


Figure 3. *Illustration of mid-level VMap V12 polygons in the Flathead National Forest*

3. TRAINING DATA

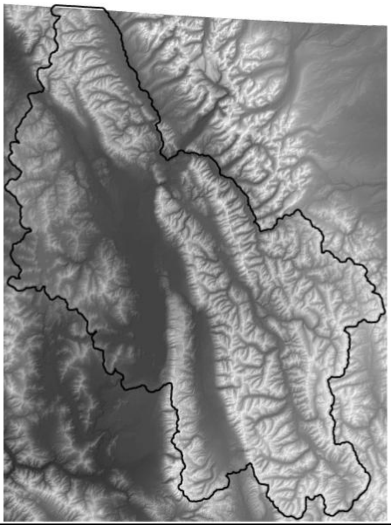
A remote sensing product is only as good as the ground truth data associated with it. Ground or other reference data is used to build the relationships between the observed phenomena and the spectral and biophysical information derived from remotely sensed and ancillary data. Collectively, ground and 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 unsampled areas within a study area. Thus, they “train” the algorithm to distinguish between and label the unknown areas within a modeling area.

In the VMap process, image object-based polygons are the units within which training data are collected. Collection of training data is primarily ground-based sampling, and is supplemented with image interpretation when/where appropriate. For instance, data such as lifeform, dominance type, and tree canopy cover could be interpreted from the 1m NAIP if personnel are familiar with the area.

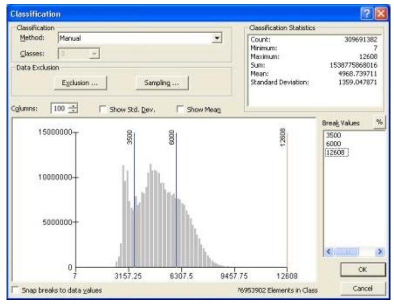
Landscape Stratification

One of the primary goals of field data collection is to capture the variability of the vegetation types that occur across the landscape. Based on previous experience on the Beaverhead-Deerlodge National Forest VMap Database production (Brown and Ahl, 2011) it was found that a landscape stratification based sample design that accounts for variation in climatic, geologic, vegetative, and topographic characteristics can be accomplished by modeling the interaction between basic lifeform and elevation classes across a study area. Since many of the layers used to describe biophysical properties of the landscape are modeled from elevation values, the modeling process was simplified by focusing directly on elevation values as a primary component of the stratification.

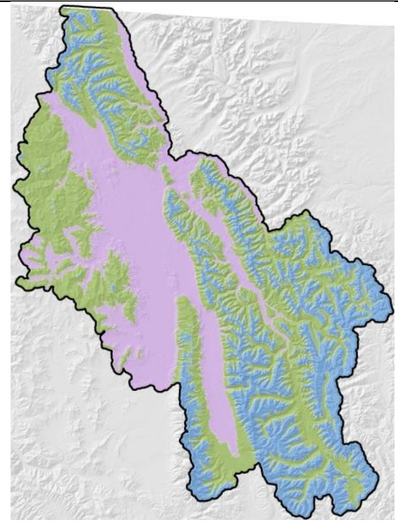
To begin, data from the National Elevation Dataset (NED) originally provided continuous elevation estimates rounded to the nearest foot, but this level of detail was difficult to work with. Therefore the dataset was reclassified into three classes, essentially representing low, medium, and high elevation landscape units. The Natural Breaks classification algorithm was used to parse the elevation histogram into the three specified classes, which is shown in the example below (Figure 2).



A) Continuous Elevation in ft.



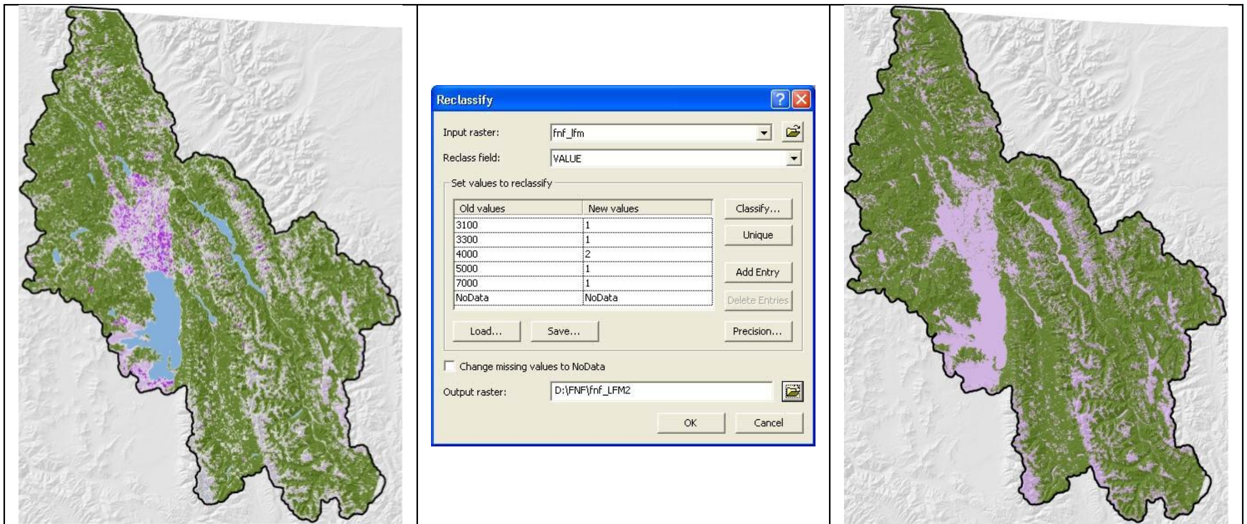
B) Classification Algorithm



C) Low, Medium, High Elevation

Figure 2. C*lassification of continuous elevation data using the natural breaks algorithm to produce three classes ranging from 1) 0 – 3,500, 2) 3,501 - 6,000) and 3) greater than 6,001 ft, shown in pink, green, and blue, respectively*

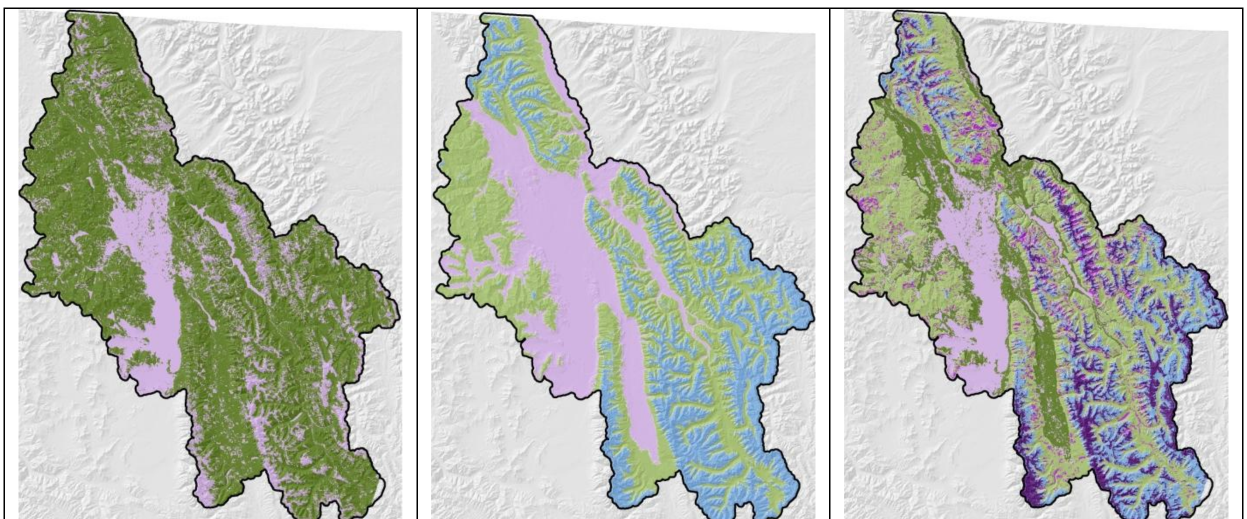
Further division of the landscape focused on the distribution of vegetation. While more complex datasets were considered (i.e., mapped distributions of geomorphic land types and their various associations (R1 LTA), regional geology, and Level 4 Ecoregion data layers) it was found that a basic classification of forest versus non-forest lifeforms provided the most meaningful and straight-forward interpretation. The four basic classes of lifeform established during the segmentation process were reduced into two categories describing the basic forest and non-forest lifeforms across the FNF (Figure 3).



A) Lifeform with all classes	**B**) Classification Algorithm	**C**) Reclassified Lifeform

Figure 3. *Reclassification of the initial herbaceous, shrub, deciduous tree, coniferous tree, and rock types into basic forest versus non-forest lifeform classes*

The final land unit stratification was completed by combining both the vertical and horizontal elements of the landscape. The vertical elements represented the low, moderate and high elevation classes, and the horizontal elements were composed of forest and non-forest vegetation types.



A) Lifeform with 2 classes	**B**) Elevation with 3 classes	**C**) ELE-LFM with 6 classes

Figure 4. *Development of the final landscape stratification dataset based on forest and non-forest lifeforms and elevation zones. Two classes of A) lifeform were combined with three classes of B) elevation to create 6 unique combinations (strata) of vertical and horizontal landscape features.*

Sampling vegetation within the unique combinations of forest and non-forest types over a range of elevation classes ensures that the range of expected environmental conditions in the FNF landscape is fully covered.

Table 2. *Spatial characteristics of Flathead NF vegetation modeling units and associated Strata*

		Strata Proportion by Model Area							
STRATA CODE	**STRATA DESCRIPTION**	**m5001**	**m5002**	**m5003**	**m5004**	**m5005**	**m5006**	**m5007**	**Average Proportion**
11	Low Elevation Forest (~2,500 - 3,500 ft)	14	16	35	12	3	37	0	17
12	Low Elevation Nonforest (~2,500 - 3,500 ft)	3	1	4	2	0	4	0	2
21	Mid Elevation Forest (3,501 - 6,000 ft)	61	76	49	77	55	47	47	59
22	Mid Elevation Nonforest (3,501 - 6,000 ft)	12	5	7	7	5	6	3	6
31	High Elevation Forest (6,001 > 9,000 ft)	8	2	4	2	26	5	37	12
32	High Elevation Nonforest (6,001 > 9,000 ft)	2	0	1	0	11	1	13	4

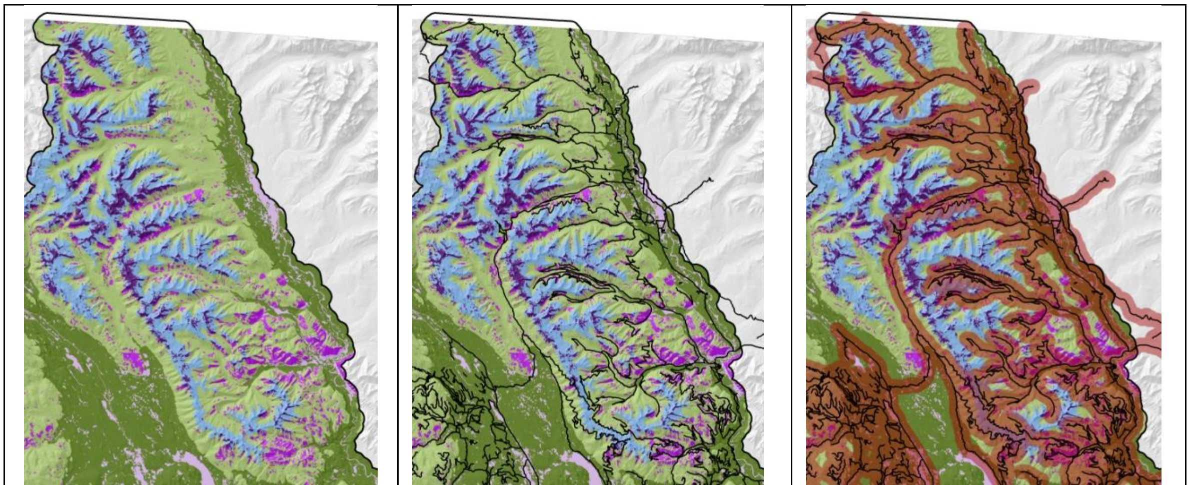
Sampling within Strata

Upon development of the biophysical strata composing the FNF model areas, the next stage of the VMAP sampling strategy is to identify potential sites for field review. There are three essential considerations in the development of a proposed sample network. First, an appropriately proportioned distribution across the landscape. Second, it is desirable to collect as many high quality samples as possible. In keeping with these principles, the time and effort needed to access suggested sample sites must be balanced against the need to acquire a certain number of samples. In short, spending excessive effort to visit a few remote sample sites is not as efficient as collecting more, but easier to obtain samples. Thus establishing the premise that a purposive sampling design meets the needs of a mapping project.

To set up a spatially proportionate sample design, a systematic grid of points with 500 meter spacing across the entire study area was created, where each point represents a potential field review site. Each point was attributed with a vegetation model identification number, and relevant Strata code. The basic assumption is that if all potential sites are reviewed, a proportionate sample of landscape features and associated vegetation characteristics will be sampled. Given that it will not be possible to visit all sites, further stratification is necessary to derive a realistic proposed sample network.

Sample Reduction

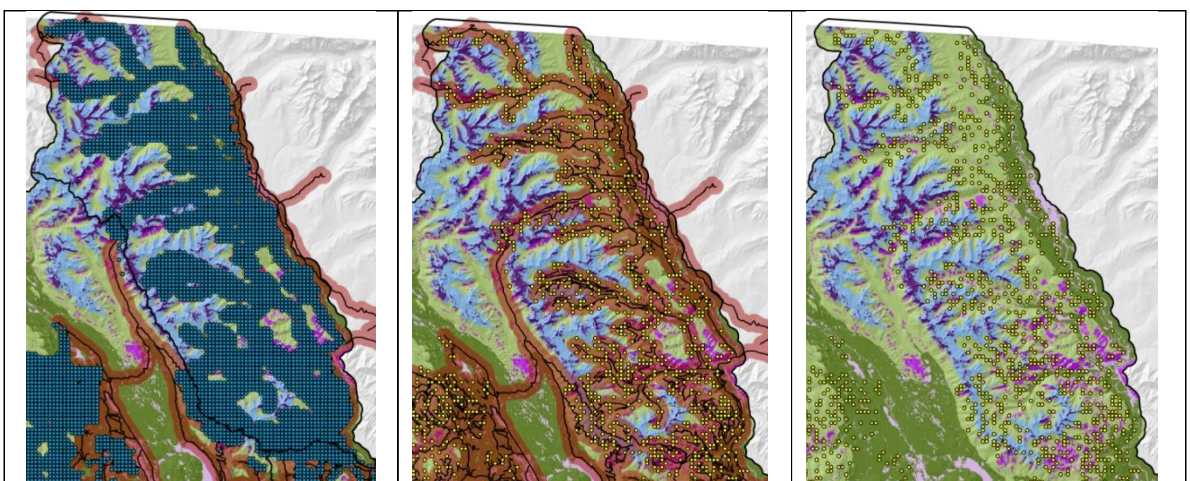
As a first step towards reducing the potential sample points down to a reasonable number it was assumed that the existing roads & trails network will determine the primary access to proposed sites. Realizing the amount of time required to record sample data is limited, we applied a 1 km buffer (about 0.5 mile) buffer around the road network. The zone identified by the buffered network then represents potential areas within which vegetation modeling units, constrained by Forest Service ownership, may be visited by a sample collection crew with a reasonable amount of effort. An example of this buffer network is given below for the North Fork Flathead vegetation sub-model (m5001) in Figure 5.



A) Model 5001 STRATA | **B**) STRATA with Roads | **C**) STRATA with Buffered Road

Figure 5. *Sample reduction, phase 1*

While collection of sample data from sites outside of the road buffer zone may be valuable, the amount of time and effort to reach them can be prohibitively excessive. Therefore field time was focused on the proposed sample sites within the buffered network. This reduced the number of suggested sample points by roughly 50%. Despite the buffer-based reduction, 28,652 point still represents approximately 4,000 sample sites for each vegetation modeling unit, which is more than time and budget constrained sampling efforts can accomplish in a field season. With this in mind, a random selection of 25% of the buffered points within each strata was performed, within each vegetation modeling unit of the FNF, to provide a more reasonable goal. This resulted in 7,163 suggested sample points across the area, with a min. of 487, max. of 1,704, and mean of 1,023 locations in each of the 7 vegetation modeling units. The process of selecting points within the buffer zone is illustrated schematically below in Figure 6, using the North Fork Flathead vegetation sub-model (m5001) as an example.



A) Full Network of Systematic Points within the 1km Buffer | **B**) Random Selection of 25% of Buffered Points in Each Strata | **C**) Final Illustration of Suggested Sample Sites in Model 5001

Figure 6. *Sample reduction, phase 2*.

Comparison of relative proportions of land area occupied by the various Strata to the percentage of sample points within the Flathead NF suggests a close agreement (Figure 7, and Table). This indicates that proportionate sampling of the landscape is possible using the procedure outlined herein.

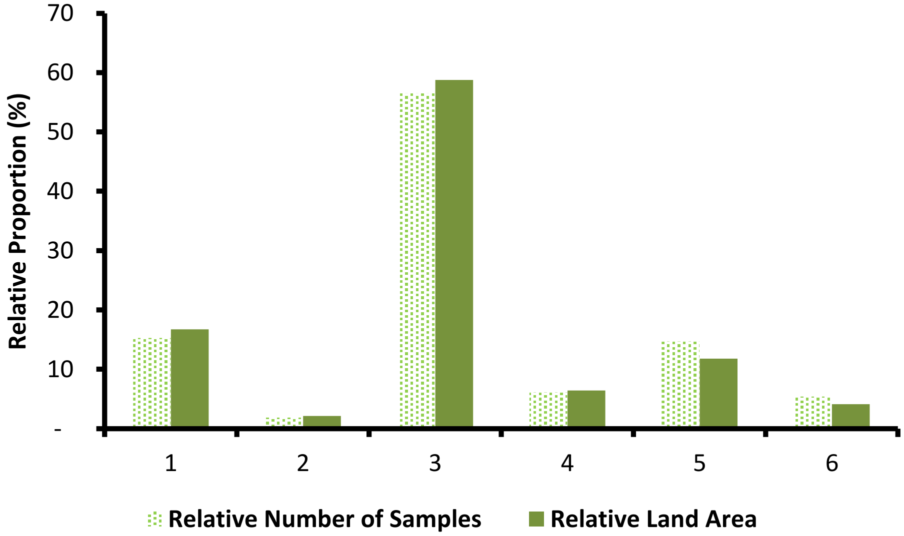


Figure 7. *Proof of concept for proportionate sampling in the Flathead NF, using the average land area percentage in each strata versus the relative number of suggested samples in associated strata*

Table 3.*Two Sample t-test comparing the relative distributions of suggested training data sites and associated land area in each of the defined strata of the Flathead National Forest*

	Variable 1	*Variable 2*
Mean	16.667	16.667
Variance	410.354	453.360
Observations	6	6
Pearson Correlation	0.997	
Hypothesized Mean Difference	0	
df	5	
t Stat	1.93996E-16	
P(T<=t) one-tail	0.5	
t Critical one-tail	2.015048373	
P(T<=t) two-tail	1	
t Critical two-tail	2.570581836	

4. IMAGE CLASSIFICATION

Labeling Algorithms

The Federal Geographic Data Committee (FGDC) Vegetation Classification Standards (FGDC 1997) establishes a hierarchy of existing vegetation classification with nine levels. The top seven levels are primarily based on physiognomy. The two lowest levels, alliance and association, are based on floristic attributes. The USDA Forest Service has set the national direction for classification and mapping of existing vegetation to implement the FGDC standards, and to provide direction for classifying and mapping structural characteristics (Brohman and Bryant 2005). This direction applies to a variety of geographic extents and thematic resolutions characterized as map scale levels. The Northern Region Vegetation Mapping Program (VMap), and resulting existing vegetation database, is specifically designed to meet this national program direction at the mid-level.

Attribute labeling of the VMap products is accomplished using a multi-step process. The image classification process begins with the segmentation procedure. Image-objects created during the segmentation routine are first labeled according to lifeform classes using algorithms within the eCognition software (version 4.6). eCognition operates using a hierarchical classification scheme, and for features that are fairly easy to discern from image statistics, such as 1) tree, 2) non-tree, 3) water, and 4) sparse vegetation, membership functions were used to properly label these cover types. Figure 8 provides an example of one of these functions.

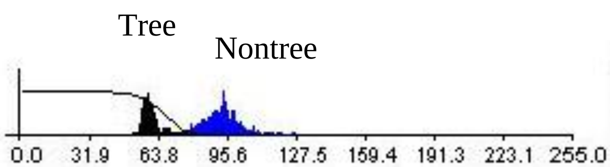
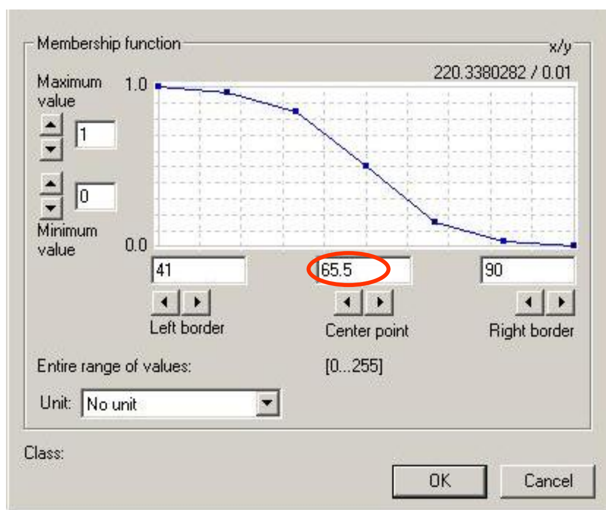


Figure 8. *Illustration of an eCognition membership function, where 'tree' samples are in the blue histogram, and 'nontree' sample data are represented by the black histogram for one of the image inputs. The histogram is used to create a membership function that excludes 'tree' at 65.5 for this input. A series of functions can be created for all image inputs that show separation and combined to create classified outputs.*



Following segmentation and initial lifeform classification, a polygon layer with associated image and biophysical statistics from each model area is exported. This data is then associated with the field collected training data and brought into the data mining software Random Forests (Breiman and Cutler, 2008) using a custom built user interface, to derive tree dominance type, tree canopy cover, and tree size classifications within the Tree lifeform. Using a similar approach, grass and shrub types were defined from within the initially determined non-forest lifeform.

Tree canopy cover is defined as "cover from above". This metric describes how much live canopy is present to intercept light/precipitation prior to its reaching the forest floor. This is measured as a percent and then divided into 4 classes: Low (10-25% cover), Moderate-Low (25-40% cover), Moderate-High (40-60% cover), and High (60%+ cover).

Tree size is mapped into four classes based on a canopy cover weighted average DBH. The classes are: Seedling/Sapling (0-5" DBH), Small (5-10" DBH), Medium (10-15" DBH), and Large/Very Large (15"+ DBH).

Tree dominance is mapped as two different, but related, classifications based on a basal area weighted plurality; Dominance of 40% (DOM40) and Dominance of 60% (DOM60). For more details about VMap dominance type, tree size, and tree canopy cover classes please refer to the *Region 1 Multi-level Classification, Mapping, Inventory, and Analysis System* (Berglund and others, 2009).

Implementation of this classification approach yields five primary attributes, consisting of lifeform, dominance type 40 and 60, tree canopy cover, and tree size class for the base level polygon feature class in each model area. Each attribute is exported to a raster of 10 meter pixel resolution, matching the original input imagery, and zonal statistics are computed for each polygon of the mid-level features. An aggregation algorithm is then implemented to attribute mid-level polygons with the majority features of the base level polygons for each attribute of interest.

For both the base and mid-level feature classes, individual models are appended to form a unified feature class that spans the entire Forest.

Map Product Review

As part of the review process, all models were visited in the field the summer of 2011 and revised based on data collected from that work. This review included only tree attributes since the expanded non-forest classes had not yet been mapped (non-forest data was collected during the review process however.) The field review process is critical for correction of errors associated with the classification and enables a refinement in the final output product that otherwise would not be possible. The resulting classification accuracy numbers (Brown, 2012) directly reflect the improvement that is seen when adequate field time is allowed.

Literature Cited

Bailey, R.G., P.E. Avers, T. King, and W.H. McNab, eds. 1994. Ecoregions and subregions of the United States (map). Washington, DC: U.S. Geological Survey. Scale 1:7,500,000; colored. Accompanied by a supplementary table of map unit descriptions compiled and edited by McNab, W. H. and Bailey, R. G. Prepared for the U.S. Department of Agriculture, Forest Service.

Berglund, D., Bush, R., Barber, J., and Manning, M. 2009. R1 Multi-level classification, mapping, inventory, and analysis system. Numbered Report 09-01 v2.0, Missoula, MT: U. S. Department of Agriculture, Forest Service, Region 1.

Breiman L, Cutler A (2008). "Random Forests – Classification Manual (website accessed in 1/2008)." http://www.math.usu.edu/~adele/forests/.

Breiman, L. (1996). Bagging predictors. *Machine Learning 26*(2), 123–140.

Brohman, R. and L. Bryant. 2005. Existing vegetation classification and mapping technical guide. U.S. Department of Agriculture, Forest Service, Washington Office, Ecosystem Management Coordination Staff. http://www.fs.fed.us/emc/rig/documents/integrated_inventory/FS_ExistingVEG_classif_mapping_TG_05.pdf

Brown, S.R. 2012. Flathead National Forest VMap Accuracy Assessment; Version 12. Numbered Report 12-06, Missoula, MT: U.S. Department of Agriculture, Forest Service, Region 1.

Brown, S.R. and Ahl, R.S. 2011. The Region 1 Existing Vegetation Mapping Program (VMap) Beaverhead-Deerlodge Methodology. Numbered Report 11-02, Missoula, MT: U.S. Department of Agriculture, Forest Service, Region 1.

Coburn C. A. and Roberts A. C. B. A multiscale texture analysis procedure for improved forest stand classification. *International Journal of Remote Sensing, October, 2004*: 25(20), 4287-4308

Federal Geographic Data Committee. 1977. Vegetation Subcommittee. Vegetation classification standard. FGDC-STD-005. Federal Geographic Data Committee, U.S. Geological Survey, Reston, Virginia, USA. [Available online: http://www.fgdc.gov/standards/documents/standards/vegetation/vegclass.pdf]

Federal Geographic Data Committee. 1996. FGDC Standards Reference Model. Federal Geographic Data Committee, U.S. Geological Survey, Reston, Virginia, USA. [Available online: http://www.fgdc.gov/standards/refmod97.pdf]

Gopal, S. and C. Woodcock. 1994. Theory and methods for accuracy assessments of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*, 60: 181-188.

Haralick, R.M. and L.G. Shapiro. 1985. Image segmentation techniques. Comput. *Vis. Image Understand*, 29: 100-132.

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Trans. on Pattern Analysis and Machine Intelligence, 20*(8), 832–844.

Kauth, R.J. and G.S. Thomas. 1976. The tasseled cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue, University, West Lafayette, IN, 4b: 41-51.

Moore I.D., Gessler P.E., Nielson G.A. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 1993;57:443-452

Press, W. H., S. A. Teukosy, W. T. Vetterling, and B. P. Flannery. 1992. Numerical recipes in fortran: the art of scientific computing, 2nd ed. London: Cambridge, 963 p.

Roberts, D.W., Cooper, S.V. 1989. Concepts and techniques of vegetation mapping. In: Ferguson, D.; Morgan, P.; Johnson, F.D., eds. Land classifications based on vegetation: applications for resource management. Gen. Tech. Rep. INT-257. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station: 90-96.

Ryherd, S. and C. Woodcock. 1996. Combining spectral and texture data in the segmentation of remotely sensed images. *Photogrammetric Engineering and Remote Sensing*, 62: 181-194.

Shandley, J., J. Franklin, and T. White. 1996. Testing the Woodcock-Harward image segmentation algorithm in an area of Southern California chaparral and woodland vegetation. *International Journal of Remote Sensing*, 17: 983-1004.