

## SWReGAP Land Cover Mapping Methods Documentation

Functional Unit or Mapping Zone: CO-3 (MMZ8) .....	1
Organization: Colorado Division of Wildlife, Denver, CO .....	1
1) Predictor layer preparation: .....	2
a) <i>Image standardization</i> : .....	2
b) <i>Image dates and mosaicking</i> : .....	2
c) <i>Image derived datasets</i> : .....	3
d) <i>DEM derived datasets</i> : .....	4
2) Samples: .....	4
a) <i>Sample collection methods</i> : .....	4
b) <i>Summary of samples</i> : .....	5
3) Cover types: .....	7
a) <i>Classification Tree (CT) modeled cover types</i> : .....	7
b) <i>Non CT modeled cover types</i> : .....	8
4) Summary of predictor layers used: .....	8
a) <i>Multi band predictors</i> : .....	8
b) <i>Single band predictors</i> : .....	8
5) Modeling Methods: .....	8
a) <i>See5 Classification Tree modeling</i> : .....	8
b) <i>Post-classification, recoding and other modeling steps</i> : .....	10
c) <i>Generalizing to MMU and map completion</i> : .....	11
6) Validation: .....	12
a) <i>CT model validation</i> : .....	12
b) <i>Discussion of mapped cover types</i> : .....	13
7) Citations: .....	18

### **Functional Unit or Mapping Zone: CO-3 (MMZ8)**

(Colorado's portion of a southern Rockies "Mega-mapzone." Occasionally referred to as MMZ8 or MMZ8CO (the latter to reflect that it is only the portion of the original mega-mapzone 8 within Colorado)

### **Organization: Colorado Division of Wildlife, Denver, CO**

**Person Preparing Document:** Eric Waller

**Date of Preparation:** 23 July, 2004

**1) Predictor layer preparation:**

**a) Image standardization:**

Standardization from DN values to at-sensor reflectance was performed on Landsat 7 ETM+ imagery using methods presented by Huang et. al (2001a). The resulting reflectance values were multiplied by 400, to make use of the range of byte data. The equation used for reflectance was as follows:

$$\rho_{BandN} = (PI((DN_{BandN} * Gain_{BandN} + Bias_{BandN}) * D^2) / (E_{BandN}(SIN\theta)))$$

Where,

$\rho_{BandN}$  = Reflectance for Band N

$DN_{bandN}$  = Digital Number for Band N

D = Normalized Earth-Sun Distance

$E_{bandN}$  = Solar Irradiance for Band N

$\theta$  = Solar Elevation

*[Colorado specifically avoided performing any dark object subtraction. This is consistent with the U.S.G.S. EROS Data Center (EDC), which does not perform the dark object subtraction on their imagery. It is thought that this technique may not offer much improvement, accounting for path radiance but not accounting for atmospheric transmission. It can also be difficult to ascertain appropriate dark object values, especially in images with variable atmospheric conditions or regions with variable atmospheric depth (such as mountainous areas).]*

**b) Image dates and mosaicking:**

Good summer mosaics were especially hard to create for the core Rockies. Snow is acceptable and not surprising in the spring and fall mosaics, but it is critical to get summer imagery that is as snow free as possible. This is compounded by a serious problem of cloud cover in the summer months. In several cases (33/33, 33/34, 34/32, and 34/33), image alternates from outside of our standard temporal window (1999-2001) had to be acquired to patch up problem areas of our imagery (clouds for 33/33 and 33/34, snow for 34/32 and 34/33). In two cases we had to resort to Landsat 5 alternates. While the replacement of entire images might have been easier, replacement was restricted to the problem areas, in deference to the temporal window, through the use of AOI's. The AOI's (per scene) can be found at:

**/CO/archive/CO3\_MMZ8/AOI/PRE/.**

Images were mosaicked using ERDAS Imagine 8.6 Mosaic Tool with “no outline” for type, and the “Overlay” option for *overlap function*.

Image dates and scenes were as follows:

ETM+ Scene (path/row)	Spring (yr- Julian date)	Summer (yr- Julian date)	Fall (yr- Julian date)
33/33	2000-113	2000-226 2003-162 (L5 alt.)	1999-303

33/34	2000-130	1999-191 1996-175 (L5 alt.)	1999-287
33/35	2000-130	2000-258	1999-287
34/32	2001-107	2000-153 2002-174 (alt.)	1999-310
34/33	2001-107	1999-182 2002-174 (alt.)	1999-310
34/34	2001-107	2000-201	1999-310
34/35	2000-105	1999-182	1999-310
35/32	2001-162	2000-256	1999-317
35/33	2000-144	2000-256	1999-285
35/34	2000-144	2000-256	1999-285

Images showing overlap arrangement (precedence given to lower numbers) of the scenes and their dates for each of the mosaics are available at:

**/CO/ARCHIVE/CO3\_MMZ8/MOSAICS/mz8\_sceneseq\_sprcl.img**  
**/CO/ARCHIVE/CO3\_MMZ8/MOSAICS/mz8coseq\_sum\_newimag.img**  
**/CO/ARCHIVE/CO3\_MMZ8/MOSAICS/mz8sseq\_fall.img**

These images are highly useful in modeling. Without them, the model would have no information on which particular image the spectral information is derived from. This information is critical, since image standardization does not account for all of the differences between images (e.g. atmospheric and seasonality differences).

The resulting six band Landsat ETM+ mosaics for each season can be found at:

**/CO/ARCHIVE/CO3\_MMZ8/MOSAICS/mmz8co\_ref\_spr\_aea.img**  
**/CO/ARCHIVE/CO3\_MMZ8/MOSAICS/mmz8co\_ref\_sum4.img**  
**/CO/ARCHIVE/CO3\_MMZ8/MOSAICS/mmz8co\_ref\_fall\_aea1.img**

***c) Image derived datasets:***

Tasseled cap: Brightness, Greenness & Wetness band transformations were created using coefficients derived for the Landsat 7 ETM+ sensor, by Huang, et. al (2001b). The .gmd file can be found at:

**/CO/ARCHIVE/CO3\_MMZ8/IMG\_FILES/MODELS/eros\_tcap\_mod.gmd**

Normalized Difference Vegetation Index (NDVI): Initially calculated as float data (possible data ranging between -1 and 1), with a model (**CO/ARCHIVE/CO3\_MMZ8/IMG\_FILES/MODELS/my\_ndvi\_float.gmd**) that performed the normalized band ratio (band4 - band3)/ (band4 + band3). Later, when it was determined that byte data were required, another model (**CO/ARCHIVE/CO3\_MMZ8/IMG\_FILES/MODELS/nd2nd\_byte.gmd**) was used to

make all values positive with a possible range from 0 to 200 (by adding one and then multiplying by 100).

Non-Photosynthetic Vegetation Index (NPV): A newly developed index for this project, similar to NDVI in that it is a normalized ratio that ranges from -1 to +1, this ratio highlights ETM+ band 5 information that was not very evident in the new Tasseled Cap transforms. In earlier versions, the Tasseled Cap “Wetness” was somewhat inversely related to non-photosynthetic vegetation (and band 5). Non-photosynthetic vegetation (and band 5) information can help differentiate many cover types. The lack of NPV information in the new transform made this ratio useful. The ratio is  $(\text{band5} - \text{band4} - \text{band3}) / (\text{band5} + \text{band4} + \text{band3})$ . One is added and the result multiplied by 100 to scale the ratio between 0 and 200. The .gmd file can be found at:  
**/CO/ARCHIVE/CO3\_MMZ8/IMG\_FILES/MODELS/npv\_ratio\_c.gmd**

All imagery derived predictor layers can be found at:  
**/CO/ARCHIVE/CO3\_MMZ8/IMG\_FILES/.**

***d) DEM derived datasets:***

Thirty meter digital elevation models were obtained from the EROS Data Center, National Elevation Database (NED). The date for these data was October 1999. DEMs were converted from floating point grids to integer grids and mosaicked for the region, then clipped to the mapping area.

**Slope: Expressed in degrees, as calculated by ERDAS Imagine’s slope routine.**

**Aspect: A nine class aspect grid. Values 1=N, 2, NE, 3=E, 4=SE, 5=S, 6=SW, 7=W, 8=NW, 9 = FLAT.**

**Landform: A 10 class landform grid derived from a topographic relative moisture (values ranging from 0-28) index grid (Manis et. al 2001).**

For modeling purposes all arcinfo grids were converted to ERDAS Imagine .img files and can be found at:

**/CO/ARCHIVE/CO3\_MMZ8/IMG\_FILES/.**

**2) Samples:**

***a) Sample collection methods:***

Samples were collected in a variety of ways. Originally, it was thought that most, if not all, of the sampling would be derived from field collected information – polygons delineated over imagery in the field by field crews. Classification trees, however, require substantial amounts of training data so that additional information had to be acquired. Each type of data can be distinguished by the SITEID field in the polygon coverage:

SWReGAP data – polygons collected in the field by Colorado field crew personnel. The SITEID contains CO for Colorado, followed by the date (mmddy), the collector’s initials, and a two digit number to distinguish the site from others collected on the same day.

“Basinwide” data - points collected in the field by another recent Colorado vegetation mapping effort: the Colorado Vegetation Classification Project (<http://ndis.nrel.colostate.edu>). As points, they had to be buffered (45 meters) to provide polygons (approximately 3 pixels by 3 pixels) that could be used in our modeling. They usually contain a field person’s initials and a date etc., but they might be best distinguished by not following the naming convention of the other sources.

“Augmentation” data – these are polygons screen digitized by image analysts guided by field information, aerial photography (<http://terraserver-usa.com>), satellite imagery and/or context. These were often generated in areas that were persistently and obviously misclassified in modeling, and to replace imprecise field data, such as a Basinwide point located on a road. The SITEID begins with an “AD” for “Augmentation Data”, followed by the date (mmddy), the analyst’s initials, and a two digit number to distinguish the polygon from others collected on the same day.

Extensive use was made of the U.S. Forest Service’s “R2VEG” vegetation coverage (e.g. <http://www.fs.fed.us/r2/sanjuan/projects/gis/index.shtml>) to guide the augmentation of polygons. Some consideration was given to “burning in” these forest service polygons directly, but the guided augmentation approach permitted the better modeling of areas lacking R2VEG coverage, and provided a more independent mapping that was consistent with the rest of our effort; a separate assessment could be conducted to determine whether a direct “burn in” would have been more accurate.

***b) Summary of samples:***

11,144 samples were available for modeling. The polygon coverage containing all samples is located at: **CO/ARCHIVE/CO3\_MMZ8/TRAIN\_DATA/**. In this particular mapzone, there is only a post-“shrunk” coverage, relating to a step necessary for modeling described in section 5.a). A pre-“shrunk” coverage would have had to have been created from scratch, as successive sets of shrunken augmentation polygons were added to a shrunken version of base layer (SWReGAP and Basinwide) polygons. (This approach saved a tremendous amount of time computationally.)

The table below shows the classes that were mapped in this map zone, and the number of sample polygons for each class. These classes make up the hybrid/alliance classification. Simple recoding of the alliances to systems for the systems classification is described in section 5.b).

*(Technically, a few more cover types not shown were actually modeled and then aggregated for the final map; this recoding is described in section 5.b and in*

**CO/ARCHIVE/CO3\_MMZ8/POST\_MODEL/recodes.doc**; *the numbers below reflect the aggregation into the final classes.*)

The absence of any samples for *S030* and *S034* relates to the fact that these mesic versions of *S028* and *S032*, respectively, were difficult to identify in the field (and from other data sources) and were therefore only mapped with post classification tree modeling as described in section 5.b).

<b>CODE</b>	<b># SAMPLES</b>	<b>DESCRIPTION</b>
S001	2	NORTH AMERICAN ALPINE ICE FIELD
S002	239	ROCKY MOUNTAIN ALPINE BEDROCK AND SCREE
S004	100	ROCKY MOUNTAINS ALPINE FELL-FIELD
S006	153	ROCKY MOUNTAIN CLIFF AND CANYON COMPLEX
S008	2	WESTERN GREAT PLAINS CLIFF AND OUTCROP
	20	INTER-MOUNTAIN BASINS ACTIVE AND STABILIZED
S012		DUNES
S015	9	INTER-MOUNTAIN BASINS PLAYA
S023	782	ROCKY MOUNTAIN ASPEN FOREST AND WOODLAND
	87	ROCKY MOUNTAIN SUBALPINE-MONTANE LIMBER-
S025		BRISTLECONE PINE WOODLAND
	1197	ROCKY MOUNTAINS SUBALPINE DRY-MESIC SPRUCE-
S028		FIR FOREST AND WOODLAND
	0	ROCKY MOUNTAINS SUBALPINE MESIC SPRUCE-FIR
S030		FOREST AND WOODLAND
S031	696	ROCKY MOUNTAINS LODGEPOLE PINE FOREST
	767	ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED
S032		CONIFER FOREST AND WOODLAND
	0	ROCKY MOUNTAINS MONTANE MESIC MIXED
S034		CONIFER FOREST AND WOODLAND
S036	644	ROCKY MOUNTAINS PONDEROSA PINE WOODLAND
S038	293	ROCKY MOUNTAIN PINYON-JUNIPER WOODLAND
S039	146	COLORADO PLATEAU PINYON-JUNIPER WOODLAND
	232	INTER-MOUNTAIN WEST ASPEN-MIXED CONIFER
S042		FOREST AND WOODLAND COMPLEX
	420	ROCKY MOUNTAINS GAMBEL OAK - MIXED MONTANE
S046		SHRUBLAND
	185	ROCKY MOUNTAINS LOWER MONTANE-FOOTHILL
S047		SHRUBLAND
	331	INTER-MOUNTAIN BASINS BIG SAGEBRUSH
S054		SHRUBLAND
	10	INTER-MOUNTAIN BASINS MIXED SALT DESERT
S065		SCRUB
	532	INTER-MOUNTAIN BASINS MONTANE SAGEBRUSH
S071		STEPPE

	443	INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB
S079		STEPPE
S081	213	ROCKY MOUNTAINS DRY TUNDRA
S083	238	ROCKY MOUNTAIN SUBALPINE MESIC MEADOW
	739	SOUTHERN ROCKY MOUNTAINS MONTANE
S085		GRASSLAND
S086	68	ROCKY MOUNTAINS FOOTHILL GRASSLAND
S088	10	WESTERN GREAT PLAINS SHORTGRASS PRAIRIE
S090	85	INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND
	478	ROCKY MOUNTAINS SUBALPINE/MONTANE RIPARIAN
S091		SHRUBLAND
	164	ROCKY MOUNTAINS SUBALPINE/MONTANE RIPARIAN
S092		WOODLAND
	112	ROCKY MOUNTAINS LOWER MONTANE RIPARIAN
S093		WOODLAND AND SHRUBLAND COMPLEX
	9	WESTERN GREAT PLAINS RIPARIAN WOODLAND AND
S095		SHRUBLAND
	147	INTER-MOUNTAIN BASINS GREASEWOOD FLAT
S096		COMPLEX
S100	25	NORTH AMERICAN ARID WEST EMERGENT MARSH
S102	344	ROCKY MOUNTAINS ALPINE/MONTANE WET MEADOW
N11	116	OPEN WATER
N21	31	DEVELOPED, OPEN SPACE-LOW INTENSITY
N22	46	DEVELOPED, MEDIUM - HIGH INTENSITY
N31	5	BARREN LANDS
N81	448	PASTURE/HAY
N82	135	CULTIVATED CROPS AND IRRIGATED AGRICULTURE
D02	12	RECENTLY BURNED
D03	34	RECENTLY MINED OR QUARRIED
D06	126	INVASIVE PERENNIAL GRASSLAND
D07	8	INVASIVE PERENNIAL FORBLAND
D08	24	INVASIVE ANNUAL GRASSLAND
D09	58	INVASIVE ANNUAL AND BIENNIAL FORBLAND
D10	154	RECENTLY LOGGED AREAS
D11	3	RECENTLY CHAINED PINYON-JUNIPER AREAS
A788	2	ARCTOSTAPHYLOS PATULA SHRUBLAND ALLIANCE
	5	PURSHIA TRIDENTATA SHRUB HERBACEOUS
A1523		ALLIANCE

### 3) Cover types:

#### a) Classification Tree (CT) modeled cover types:

See the above table.

***b) Non CT modeled cover types:***

NONE. Colorado relied on classification tree (CT) modeling, with post-classification modeling and recoding as necessary, to map all cover types. Some ancillary coverages were used in this process, but they were never directly “burned in”.

**4) Summary of predictor layers used:**

***a) Multi band predictors:***

NONE. In particular, raw ETM+ bands were not used. The Tasseled Cap Transforms and the ratios captured most of the spectral variability in the imagery in a way that might be more efficiently utilized by the classification tree algorithm, and in a way that might be more meaningfully interpreted. (As a test, we conducted a Kappa evaluation of a map created with raw bands and a map created with the transforms plus ratios; the results were essentially identical.)

***b) Single band predictors:***

In addition to the scene sequence images, the model used the spring, summer, and fall Tasseled Cap products (Brightness, Greenness, and Wetness), the spring, summer, and fall NDVI and NPV ratio, and elevation. Slope, aspect, and landform were not used, as it was determined that these often had deleterious effects on the modeling.

All single-band predictors can be found at:  
[/CO/ARCHIVE/CO3\\_MMZ8/IMG\\_FILES/](#)

**5) Modeling Methods:**

***a) See5 Classification Tree modeling:***

Sampling:

Pseudo-replication within each sample polygon was conducted in order to increase the number of samples used by the classification algorithm. While this use of non-independent data is not ideal for classification tree modeling, it has been found to improve classification accuracies, particularly when there are limited amounts of training data.

The pseudo-replication process involved several steps to generate random pixels within each sample polygon. Each sample polygon was first shrunk by 13 meters. This somewhat arbitrary distance was selected so that when random points placed within the shrunken polygon were converted to pixels in a raster grid, a good part of any of the resulting pixels would not lie outside the original sample polygon. Next, 20 random points were placed within each of the shrunken polygons using an **Arcview Avenue**

**script.** Finally, the points were converted to pixels while ensuring that the resulting pixels (the new grid) aligned with the raster predictor layers (failure to ensure this would result in a future resampling that would be more likely to place pixels outside the original sample polygon). The resulting sub-sampled pixels would often be less than 20 per sample polygon, if random points fell within the same pixel. This was far more likely on smaller polygons, such as the buffered Basinwide points, for which there would often only be five or six sub-samples per polygon. This new grid was converted to an Imagine IMG file and is available at:

**CO/ARCHIVE/CO3\_MMZ8/TRAIN\_DATA/rp200623\_100.img**

These sub-sampled pixels were then individually drilled through predictor layers using the sampling tool from the CART Module for Imagine (EarthSatellite Corp. 2003). 10% of the *pixels* were set aside as “test data” (although these were not considered to be at all independent, and this is not the same as the validation *polygon* data set described in Section 6).

See5 Classification Tree:

See5 (Release 1.8) data mining software ([www.rulequest.com](http://www.rulequest.com)) was used for generating classification trees. Boosting was employed using 10 trials with no pruning. Boosting involves successive modeling runs on a portion of the data (90% with 10% set aside for evaluation each time, given 10 trials). After the first run, the model tries to focus on the types it erred on in the previous run(s). After 10 runs, voting occurs, whereby the most common classification for a given pixel is assigned to that pixel.

The See5 related files (\*.names, \*.names.hst, \*.data, \*.test, \*.out, \*.tree, and .set) are located at: **CO/ARCHIVE/CO3\_MMZ8/OUTPUT/**.

The \*.names file describes the dependent variable and the independent variables. The \*.data file contains all of the raw training data for the dataset. The \*.test file contains the raw test data. The \*.out file describes all the rules – at the very bottom there is information on how the model performed on each of the 10 runs and on the final boost, both against the 90% training data and the 10% “test data”. The \*.tree is an application file used for applying the final rules.

Spatial application of rules:

The CART Module for Imagine (EarthSatellite Corp. 2003) was used to spatially apply the rules from the .tree file in See5 to create a classified image for the entire mapzone. In the process, an associated “\*\_error.img” was also created. This latter image is more of a “confidence” map, as higher values indicate greater classification confidence (i.e. less class confusion for similar pixels in the training data). Confidence should not be equated with “accuracy”, as the absence of class confusion may simply result from a lack of similar training data, but this map does provide a potentially useful spatial evaluation.

The classified image and the associated “error.img” can be found at:  
**/CO/ARCHIVE/CO3\_MMZ8/OUTPUT/**.

***b) Post-classification, recoding and other modeling steps:***

The output of the C5 classification was evaluated on a class by class basis by isolating individual classes on the map (this can be done in ERDAS IMAGINE by keeping the opacity of the evaluated class at 1 while changing the opacity of the other classes to 0). Anomalous distributions that were not thought to be naturally disjunct outliers were often removed through the generation of “areas of interest” (AOI’s) in Imagine, and recoding the problematic type within those areas to a more likely class value. AOI’s were named according to the type\_to\_type conversion and can be found at:

**CO/ARCHIVE/CO3\_MMZ8/AOI/**

Some of the recoding was more automated, such as the recoding of certain types within known urban areas. A Colorado road network coverage draped over Landsat ETM+ imagery was used to guide the digitizing of an urban layer; an AOI of these urban areas can be found at: **CO/ARCHIVE/CO1\_CO2\_MMZ7/AOI/mmz8\_all\_urb2ed.aoi**. This AOI was used to recode anything that mapped as a naturally barren type within an urban area to N22 (Developed, Medium – High Intensity). Similarly, any non-forested cover type that fell within a logged AOI

**(CO/ARCHIVE/CO1\_CO2\_MMZ7/AOI/mmz7\_gmug\_rg\_rt\_sj.aoi)** was recoded to D10 (Recently Logged). This AOI was derived from the U.S. Forest Service R2VEG coverage.

An exact description of the conversions within these AOI’s is given in:

**CO/ARCHIVE/CO3\_MMZ8/AOI/MMZ8\_AOI\_explanations.doc**

One other recoding step involved converting the classes that were used in the modeling to those that were used in the final maps. (There are different reasons for different classes going into modeling. Some old class labels were grandfathered in to the modeling. In other cases, there was some uncertainty as to which alliance/system a cover type should go into, and class confusion in the final map was a good way of helping that decision.)

This recoding is described in:

**CO/ARCHIVE/CO3\_MMZ8/POST\_MODEL/recodes.doc**

Finally, a model was used to separate the Mesic from the Dry-Mesic versions of systems S028 and S032 to create systems S030 and S034, respectively. The .gmd file used for this post-classification model can be found at:

**CO/ARCHIVE/CO3\_MMZ8/POST\_MODEL/mesic\_conifer\_classes.gmd**

This model is actually a combination of the two models described below:

*Discriminating S028-Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland and S030-Rocky Mountain Subalpine Mesic Spruce-Fir Forest and Woodland:*

A post-classification model was used to discriminate S028 from S030. The logic and parameters for the model were as follows:

This post-classification model was designed to extract *S030-Rocky Mountain Subalpine Mesic Spruce-Fir Forest and Woodland* from the standard modeled *S028-Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland*. Nature Serve describes S030 to be “*typically found in locations with cold air drainage or ponding, or where snowpacks linger late into the summer, such as north-facing slopes and high elevation ravines. They can extend down in elevation below the subalpine zone in places*”

*where cold air ponding occurs; northerly and easterly aspects predominate. These forests are found on gentle to very steep mountain slopes, high elevation ridgetops and upper slopes, plateaulike surfaces, basins, alluvial terraces, well-drained benches, and inactive stream terraces.”*

Model methods: A standard model was run withholding *S030*, consequently allowing *S028* to be mapped in its place. A condition statement was created to extract *S030* using the above aspect and landform characteristics in order to identify this subalpine mesic conifer system. The details of the conditional statement follow:

EITHER *S030* IF ((Landform == 2 OR Landform == 5 OR Landform == 6 OR Landform == 9) AND (Aspect == 1 OR Aspect == 2 OR Aspect == 0) AND (Standard Modeled Vegetation Output == *S028*)) OR Standard Modeled Vegetation Output OTHERWISE

*Discriminating S032-Rocky Mountain Montane Dry-Mesic Mixed Conifer Forest and Woodland and S034-Rocky Mountain Montane Mesic Mixed Conifer Forest and Woodland*: A post-classification model was used to discriminate *S032* from *S034*. The logic and parameters for the model were as follows:

This post-classification model was designed to extract *S032-Rocky Mountain Montane Dry-Mesic Mixed Conifer Forest and Woodland* from the standard modeled *S034-Rocky Mountain Montane Mesic Mixed Conifer Forest and Woodland* ecological system cover type. Nature Serve describes this system as “*occurring predominantly in cool ravines and on north-facing slopes. Elevations range from 1200 to 3300 m. Occurrences of this system are found on cooler and more mesic sites than the Rocky Mountain Montane Dry-Mesic Mixed Conifer Forest and Woodland. Such sites include lower and middle slopes of ravines, along stream terraces, moist, concave topographic positions and north and east-facing slopes which burn somewhat infrequently.*”

Model methods: A standard model was run withholding *S034*, consequently allowing *S032* to be mapped in its place. A condition statement was created to extract *S034* using the above aspect and landform characteristics in order to identify this montane mesic mixed conifer system. The details of the conditional statement follow:

EITHER *S034* IF ((Landform == 2 OR Landform == 5 OR Landform == 6 OR Landform == 9) AND (Aspect == 1 OR Aspect == 2 OR Aspect == 0) AND (Standard Modeled Vegetation Output == *S032*)) OR Standard Modeled Vegetation Output OTHERWISE

***c) Generalizing to MMU and map completion:***

After the spatial application of the CT model to create an \*.img file, and post classification steps were taken, the map was generalized using ERDAS Imagine 8.6 GIS Analysis, Clump tool using *4 connected neighbors* (rooks move), and then using Eliminate with a setting of a *minimum of 1 acre*.

Finally, alliances were recoded to systems to create a “Systems only” final map. A description of this simple recoding is found at:

**/CO/ARCHIVE/CO3\_MMZ8/POST\_MODEL/alliance\_recodes.doc**

Both maps can be found at: **/CO/ARCHIVE/CO3\_MMZ8/FINAL\_MAP/**.

## **6) Validation:**

### ***a) CT model validation:***

Classification assessment was an internal validation in that it used a subset of the reference data collected for modeling purposes. 20% of the reference sample polygons were randomly selected (using the `featureselect.avx`) and set aside for validation. A CT model was then generated with the remaining 80% of sample polygons, using the same procedures described in section 5.a). This model was recoded to final “hybrid” classes as described in section 5.b) and

**CO/ARCHIVE/CO3\_MMZ8/POST\_MODEL/recodes.doc**, and from there to final “system” classes as described in section 5.c) and

**/CO/ARCHIVE/CO3\_MMZ8/POST\_MODEL/alliance\_recodes.doc**. Without any post-classification modeling or clump/eliminate, these recoded land cover maps were then assessed with the appropriate recoded version of the 20% reference sample. (The lack of post-processing increased the fidelity between the map and the training data; for example, there were no mesic conifer classes in the training data. This was thought to provide a better indication of model performance.)

The CT-modeled classes were assessed using `kappa.avx`, which intersects the validation sample polygons through the CT modeled land cover map, and considers the site correctly mapped when the most common pixel value within the sample polygon agrees with the sample label. Output from `kappa.avx` consists of a \*.txt, \*.dbf, and \*.shp file. The \*.txt file presents the kappa statistic, the \*.dbf file is an error matrix indicating errors of commission and omission and the \*.shp file indicates for each reference sample site whether the sample location was considered correctly mapped, or incorrectly mapped and what it was mapped as. The \*.dbf files are a bit difficult to read with the default output; they have been modified into cleaner looking matrices and saved as \*.xls files. All of these files can be found at: **/CO/ARCHIVE/CO3\_MMZ8/VALIDATION/**.

Several (internal) validations were performed. The first Kappa run (and the corresponding .txt file, .xls file, and .shp file) is an evaluation of the hybrid product (generated from the 80% sample) against 20% of all the training data (including augmentation polygons). The second Kappa run is a similar evaluation of the systems product (simply the hybrid with the few alliances recoded to systems). The third run is an evaluation of the systems map against only the portion of the 20% withheld data that was collected in the field (both Basinwide and SWReGAP). The fourth run is an evaluation of the systems map against only the withheld data that was collected by SWReGAP field crews. Finally, the fifth run is an evaluation of the final map (generated using 100% of the training data) against all 100% of the reference data. This last product

should not be considered a validation and could be very misleading – but it may demonstrate which classes are simply very difficult for the model to separate.

***b) Discussion of mapped cover types:***

The following is a somewhat qualitative evaluation of the first Kappa run, reviewing the model's performance on each of the classes of the hybrid scheme against a 20% sampling of all the training data (including augmented polygons). This evaluation is not much different than it would be for the systems scheme, as there is not much of a difference between the hybrid product and the systems product. Neither is there a tremendous difference between evaluating with all of the training data or with SWReGAP field-collected polygons only.

The producer's and user's accuracies described below relate to omission and commission errors, respectively, for the different classes of the hybrid (alliance/systems) level map. [Actual tables are found under **ARCHIVE/CO3/VALIDATION** at **errmtx1.dbf** and **matrix1\_adj.xls**.] In some cases a given class was not mapped in any of the training sites; while there were no errors of commission, the class was not mapped correctly anywhere either, and so the resulting 0/0 has no user's accuracy.

**A1523 (Bitterbrush):** 100% producer's and 50% user's may be reasonable, but not enough sites to adequately evaluate. [1 site]

**A788 (Manzanita):** 100% producer's and 100% user's appears good, but not enough sites to adequately evaluate. [1 site]

**D02 (Recent Burn):** 67% producer's and 100 user's suggest this is mapping well, but more sample sites would be preferable. [3 sites]

**D03 (Mining):** 14% producer's and 20% user's suggest that mining is not being mapped very well beyond the training data; confusion is with a variety of classes. [7 sites]

**D06 (Invasive Perennial Grassland):** 12% producer's and 50% user's suggest that it is generally being omitted, primarily to pasture, montane sagebrush, and montane grassland, but that at least it's not overmapping extensively, except for a fair amount of commission of montane grassland. [26 sites]

**D07 (Invasive Perennial Forbland):** 0% producer's (0/0 user's) suggests that it's being omitted but no commission suggests that at least it's not overmapping. It would need more sites to be better evaluated. [2 sites]

**D08 (Invasive Annual Grassland):** 0% producer's and 0% user's (3 errors of commission) suggest that it's both being omitted and overmapped a little bit. [5 sites]

**D09 (Invasive Annual and Biennial Forbland):** 17% producer's and 50% user's suggest that it's being heavily omitted – mostly to semi-desert shrub steppe, but that it's not overmapping extensively. [12 sites]

**D10 (Recently Logged):** 29% producer's and 90% user's indicates that it is has a strong tendency to be omitted, primarily in favor of spruce-fir and alpine/montane wet meadow, but that it is not overmapping very much at all. [31 sites]

**D11 (Chained Pinyon-Juniper):** 100% producer's and 100% user's are encouraging, but not enough sites to adequately evaluate. [1 site]

**N11 (Water):** 88% producer's and 100% user's indicate that it is mapped very well, except for some slight omission to alpine bedrock, which may be related to water looking like steep shaded bare slopes. [24 sites]

**N21 (Low Intensity Development):** 29% producer's and 50% user's indicate that it is often being omitted, especially in favor of pasture, but not overmapping extensively. [7 sites]

**N22 (High Intensity Development):** 50% producer's and 42% user's indicate that it is fairly regularly omitted, often to Rocky Mountain Cliff and Canyon, but also overmapped, particularly in place of Rocky Mountain Cliff and Canyon. The former problem was handled with urban area post-processing. [10 sites]

**N30 (Barren):** 100% producer's and 100% user's are encouraging, but not enough sites to adequately evaluate. [1 site]

**N81 (Pasture):** 71% producer's and 54% user's indicate that there is some omission of pasture to such classes as big sagebrush and Gambel oak, but, more commonly, it is being mapped in place of alpine/montane wet meadow, agriculture, and, to a lesser extent, Gambel oak and invasive perennial grass. [90 sites]

**N82 (Agriculture):** 37% producer's and 67% user's indicate that it is being fairly heavily omitted in favor of pasture, but that it's not being overmapped extensively. [27 sites]

**S001 (Alpine Ice Field):** 0% producer's (0/0 user's) suggests that it's being omitted but no commission suggests that at least it's not overmapping. It would need more sites to be better evaluated. [1 site]

**S002 (Alpine Bedrock and Scree):** 83% producer's and 80% user's suggest that it is being mapped fairly well – just a little bit of omission to Alpine Fell-Field and Rocky Mountain Cliff and Canyon, as well as some commission of Alpine Fell-Field. [48 sites]

**S004 (Alpine Fell-Field):** 45% producer's and 60% user's indicate that it is fairly frequently omitted to Dry Tundra and Alpine Bedrock, but also commits those two classes to some extent. [20 sites]

**S006 (Rocky Mountain Cliff and Canyon):** 58% producer's and 55% user's suggest that there is some omission to Rocky Mountain P-J Woodland and High Intensity Development, as well as some commission of several classes, primarily Alpine Bedrock and Scree and Montane Grassland. [31 sites]

**S008 (Great Plains Cliff and Outcrops):** 0% producer's (0/0 user's) suggests that it's being omitted but no commission suggests that at least it's not overmapping. It would need more sites to be better evaluated. [1 site]

**S012 (Active and Stabilized Dunes):** 50% producer's and 50% user's indicate that this class is being mapped reasonably well, with no clear pattern of confusion with other classes, but more sites would be needed to thoroughly evaluate its accuracy. [4 sites]

**S015 (Playa):** 50% producer's and 100% user's are somewhat encouraging, but there are too few sites to be conclusive. [2 sites]

**S023 (Aspen):** 86% producer's and 73% user's indicate that it's only rarely being omitted to other classes such as Subalpine/Montane Riparian Shrubland and Mixed Conifer, but that it more frequently is mapped instead of Subalpine/Montane Riparian Shrubland and Aspen-Mixed Conifer, and, to a lesser extent, Alpine/Montane Wet Meadow and Subalpine/Montane Riparian Woodland. [157 sites]

**S025 (Subalpine Limber-Bristlecone Pine):** 6% producer's and 25% user's suggest that it is being heavily omitted in favor of Spruce-Fir, Lodgepole Pine, and Ponderosa Pine, while also sometimes mapping in the place of similar high elevation classes. [18 sites]

**S028 (Spruce-Fir):** 83% producer's and 70% user's suggest that it is only occasionally mapped as Lodgepole Pine or Mixed Conifer, but that, more often, it maps in place of these classes, in addition to the Aspen-Mixed Conifer and Recently Logged areas. [240 sites]

**S031 (Lodgepole Pine):** 60% producer's and 60% user's suggest that it is somewhat frequently mapped as Spruce-Fir, and occasionally as Mixed Conifer, and that it also maps instead of these classes fairly often. [140 sites]

**S032 (Mixed Conifer):** 60% producer's and 54% user's suggest that it is fairly often being replaced by Ponderosa Pine, Lodgepole Pine, and Spruce-Fir, but that it is also mapping in place of these classes, in addition to Aspen-Mixed Conifer and Subalpine / Montane Riparian Woodland. [154 sites]

**S036 (Ponderosa Pine):** 74% producer's and 65% user's indicate that this class is only replaced by Mixed Conifer, but that it also replaces Mixed Conifer, in addition to Rocky Mountain Pinyon-Juniper, Lodgepole Pine, and Subalpine Limber-Bristlecone Pine. [129 sites]

**S038 (Rocky Mountain Pinyon-Juniper):** 70% producer's and 77% user's demonstrate that is fairly often mapped as Ponderosa Pine, and occasionally as Montane Grassland, while it also maps instead of Ponderosa occasionally. [59 sites]

**S039 (Colorado Plateau Pinyon-Juniper):** 73% producer's and 76% user's indicates that this class is being mapped reasonably well, with no strong confusion with any other class. [30 sites]

**S042 (Aspen-Mixed Conifer):** 28% producer's and 65% user's indicate that this class is strongly replaced by spruce-fir, aspen, and mixed conifer, but that it does not overmap extensively, only occasionally replacing aspen and Subalpine/Montane Riparian Woodland. [47 sites]

**S046 (Gambel Oak):** 76% producer's and 63% user's suggest that it is occasionally mapped as montane sagebrush and pasture, but that it is more frequently mapped instead of other classes, such as Lower Montane Riparian Woodland, Subalpine/Montane Riparian Shrubland, and Pasture. [84 sites]

**S047 (Foothill Shrubland):** 38% producer's and 58% user's indicate that it is often replaced by such classes as montane grassland, montane sagebrush, and pasture, and that it occasionally replaces a variety of classes. [37 sites]

**S054 (Big Sagebrush):** 76% producer's and 64% user's indicate that this class is only occasionally replaced by montane sagebrush and semi-desert shrub steppe, but that it more frequently replaces such classes as semi-desert shrub steppe and pasture. [67 sites]

**S065 (Mixed Salt Desert Scrub):** 0% producer's and 0% user's (1 error of commission) indicate that this class is getting replaced by Pasture, but that it also maps in place of Big Sagebrush in one instance. More sites are required to better evaluate this class. [2 sites]

**S071 (Montane Sagebrush Steppe):** 75% producer's and 63% user's indicate that it is occasionally replaced by montane grassland and subalpine meadow, but that it just as often replaces these classes, in addition to many other classes such as Gambel Oak and Big Sagebrush. [107 sites]

**S075 (Juniper Savanna):** 0% producer's (0/0 user's) indicates that this class is being heavily replaced by other classes such as montane sagebrush, but that it's not overmapping. [3 sites]

**S079 (Semi-Desert Shrub Steppe):** 63% producer's and 62% user's indicate that there is a fair amount of replacement by Greasewood, with some replacement by Montane Grassland and Big Sagebrush, while this class replaces Desert Grassland, Greasewood, and Big Sagebrush. [89 sites]

**S081 (Rocky Mountain Dry Tundra):** 86% producer's and 79% user's indicate that this class is being mapped very well except for occasionally replacing Alpine Fell-Field. [43 sites]

**S083 (Subalpine Meadow):** 50% producer's and 50% user's indicate that this class is replaced by a variety of classes, including Montane Grassland, Subalpine/Montane Riparian Shrubland, Montane Sagebrush, and Alpine/Montane Wet Meadow, while it just as frequently replaces these same classes. [48 sites]

**S085 (Montane Grassland):** 72% producer's and 62% user's indicate that it is sometimes replaced by montane sagebrush and subalpine meadow, but more frequently it replaces other classes such as Semi-Desert Shrub Steppe, Montane Sagebrush, Subalpine Meadow, and Foothill Shrubland. [148 sites]

**S086 (Foothill Grassland):** 64% producer's and 50% user's indicate that it is occasionally mapped as other classes such as Semi-Desert Shrub Steppe, but that more frequently it replaces other classes such as Foothill Shrubland and Invasive Perennial Grassland. [14 sites]

**S088 (Shortgrass Prairie):** 0% producer's (0/0 user's) indicates that it may get replaced by other grasslands away from the training sites, but that at least it's not overmapping extensively. More sites would be needed to evaluate it more thoroughly. [2 sites]

**S090 (Semi-Desert Grassland):** 12% producer's and 40% user's indicate that it's generally getting replaced by Semi-Desert Shrub Steppe and Greasewood, while it also occasionally replaces Semi-Desert Shrub Steppe. [17 sites]

**S091 (Subalpine/Montane Riparian Shrubland):** 53% producer's and 60% user's indicate that this class is often replaced by Aspen, sometimes by Gambel Oak, occasionally by Montane Grassland and Montane Sagebrush, and often replaces Alpine/Montane Wet Meadow, Subalpine Meadow, and Aspen. [96 sites]

**S092 (Subalpine/Montane Riparian Woodland):** 0% producer's (0/0 user's) is amazingly low for a class with 33 samples – but at least it's not overmapping extensively(!) – there's no commission error. This class is being entirely replaced by such classes as Mixed Conifer, Aspen, and Lodgepole Pine. [33 sites]

**S093 (Lower Montane Riparian Woodland and Shrubland):** 30% producer's and 70% user's indicate that this class is often being replaced by other classes such as Gambel Oak, but that it less frequently maps instead of other classes such as Arid West Emergent Marsh. [23 sites]

**S095 (Western Great Plains Woodland and Shrubland):** 50% producer's and 50% user's are probably inconclusive due to the lack of sample sites. [2 sites]

**S096 (Greasewood Flat):** 73% producer's and 46% user's indicate that it is occasionally replaced by Semi-Desert Shrub Steppe, but more frequently, it replaces Semi-Desert Shrub Steppe, and occasionally replaces Semi-Desert Grassland. [**30 sites**]

**S100 (Arid West Emergent Marsh):** 0% producer's (0/0 user's) indicates that this is either called Pasture or Lower Montane Riparian Woodland and Shrubland, but that it's not overmapping extensively. [**5 sites**]

**S102 (Alpine/Montane Wet Meadow):** 26% producer's and 41% user's suggest that it's being heavily replaced by such classes as Pasture, Subalpine Meadow, Montane Grassland, Aspen, and Subalpine/Montane Riparian Shrubland, but that it's also overmapping a lot of these same classes, in addition to Recently Logged Areas. [**69 sites**]

## 7) Citations:

Chavez, P. S. Jr., 1988, An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.*, 14, pp.459-479.

Chavez, P. S. Jr., 1996, Image-based atmospheric corrections—revisited and revised. *Photogrammetric Engineering and Remote Sensing* 62(9): 1025-1036.

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Mernard, M. Pyne, M. Ried, K. Schulz, K. Snow, and J. Teague. 2003. Ecological systems of the United States: A working classification of US terrestrial systems. NatureServe, Arlington, Virginia.

EarthSatellite. 2003. CART Software User's Guide, prepared by EarthSatellite Corporation for the US Geological Survey in support of the National Land Cover Database (NLCD) 2000. EarthSatellite, January 2003.

Manis, G., J. Lowry and R. D. Ramsey, 2001, Preclassification: An ecologically predictive landform model. GAP Analysis Bulletin No. 10. USGS.

Huang, C. L. Yang, C. Homer, B. Wylie, J. Vogelmann and T. DeFelice, 2001, At-sensor reflectance: A first order normalization of Landsat 7 ETM+ Images. (<http://landcover.usgs.gov/pdf/huang2.pdf>)

Huang, C., B. Wylie, C. Homer, L. Yang, and G. Zylstra, 2002, Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. *International Journal of Remote Sensing*, 8: 1741-1748.