

SWReGAP Land Cover Mapping Methods Documentation

Functional Unit or Mapping Zone: Arizona 3	1
Organization: U.S. Geological Survey, Southwest Biological Science Center	1
1) Predictor layer preparation:	1
a) <i>Image standardization</i> :	1
b) <i>Image dates and mosaicking</i> :	2
d) <i>DEM derived datasets</i> :	3
2) Samples:	3
a) <i>Sample collection methods</i> :	3
b) <i>Summary of samples</i> :	4
c) <i>Preparing Sample Data for Classification Tree (CT) Modeling</i> :	5
3) Cover types:	6
a) <i>Classification Tree modeled cover types</i> :	6
b) <i>Non CT modeled cover types</i> :	7
4) Summary of predictor layers used:	9
5) Modeling Methods:	10
a) <i>See5 Classification Tree modeling</i> :	10
b) <i>Post-classification, recoding and other modeling steps</i> :	11
c) <i>Generalizing to MMU and map completion</i> :	12
6) Validation:	13
7) Citations:	17

Functional Unit or Mapping Zone: Arizona 3

Organization: U.S. Geological Survey, Southwest Biological Science Center

Person Preparing Document: Kathryn Thomas, Cynthia Wallace, Jess Kirby

Date of Preparation: October 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), with the addition of a dark object subtraction step to remove atmospheric haze, as suggested by Chavez (1988, 1996). The equation used was as follows:

$$P_{\text{BandN}} = \frac{\pi((L_{\text{BandN}} * \text{Gain}_{\text{BandN}} + \text{Bias}_{\text{BandN}}) - (H_{\text{BandN}} * \text{Gain}_{\text{BandN}} + \text{Bias}_{\text{BandN}})) * D^2}{E_{\text{BandN}} * (\text{COS}((90 - \theta\pi/180)))}$$

Where,

P_{BandN} = Reflectance for Band N

L_{BandN} = digital number for Band N

H_{BandN} = digital number representing dark object for Band N

D = Normalized earth-sun distance

H_{BandN} = solar irradiance for Band N

b) Image dates and mosaicking:

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 date (yr Julian date)
35/34	2000-096	2000-256	1999-285
35/35	2000-096	2000-256	1999-317
36/34	2000-103	2000-167	1999-292
36/35	2000-103	2000-167	1999-292

The ERDAS Imagine mosaic files showing overlap arrangement can be found at:

AZ\ARCHIVE\AZ3\MOSAIC\AZ3_fa.mos
AZ\ARCHIVE\AZ3\MOSAIC\AZ3_sp.mos
AZ\ARCHIVE\AZ3\MOSAIC\AZ3_su.mos

In addition, a description of the overlap arrangement is presented in:

AZ\ARCHIVE\AZ3\MOSAIC\mosaic\AZ3_readme.txt

c) Image derived datasets:

Normalized Difference Vegetation Index (NDVI): Used a modified version of the NDVI model provided by ERDAS Imagine 8.6. This model performs the band ratio (band4-band3)/(band4+band3) then scales the output by 200 to create a continuous unsigned 8-bit image ranging from 0-200. For an example of the *.gmd file go to AZ\ARCHIVE\AZ3\IMG_FILES.

Tasseled cap: Brightness, Greenness & Wetness band transformations were created using coefficients derived for the Landsat 7 ETM+ sensor, by Huang et al (2001b). An example of the *.gmd file can be found at AZ\ARCHIVE\AZ3\IMG_FILES.

All imagery derived predictor layers can be found at:

AZ\ARCHIVE\AZ3\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. DEM's were converted from floating point grids to integer grids and mosaicked for the region, then clipped to the mapping area.

Aspect: A nine-class grid was created. 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 was created from a topographic relative moisture index grid (values ranging from 0-28) (Manis et al 2001).

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

AZ\ARCHIVE\AZ3\IMG_FILES\

2) Samples:

a) Sample collection methods:

USGS SBSC field crews collected the majority of samples on the ground, others came from existing data sources. Samples were assigned a label corresponding to either an Ecological System (Comer et al 2003) or a cover type uniquely defined for the SWReGAP project.

The CSOURCE field in the \AZ\ARCHIVE\AZ3\TRAIN_DATA\az3_polys_all identifies the source for each sample except where noted. The PLOTNUM field indicates the unique identifier for that record. In the case of samples taken from existing data, the unique identifier was assigned either by the project generating the data or by the SWReGAP field team for the purposes of this project. Following is an explanation of each data source.

Az2_pts: These data were collected in the portion of functional unit AZ2 that overlapped with AZ3. The sources of these data are various and can be identified by the PLOTNUM structure. See descriptions under field_pts for identification of PLOTNUM and description of these data sources.

Esdig: These data are from interpretation of Digital Orthophoto Quads obtained from the Arizona Land Resource Information System (ALRIS, <http://www.land.state.az.us/alris/>). Interpretation was done by the U.S. Geological Survey SWReGAP team. An example plot code is: AZ091703DQ014.

Field_CPRS: Data collected by the U.S. Geological Survey SWReGAP field crew from 2001 to 2003. These data were collected from 2001 to 2003. An example plot code is: AZ092803MA009.

Field_pts: These data come from various sources. The source can be determined from the structure of the PLOTNUM.

AZ3DQT<idnum>: These data are from interpretation of aerial photography available on the TerraServer web site, <http://terraserver.microsoft.com/>. The U.S. Geological Survey SWReGAP team did the interpretation. An example plot code is: AZ3DQT006.

AZ<month/day><year><teamcode><idnum><sequence letter, if applicable>: Samples collected by the U.S. Geological Survey SWReGAP field crew. These data were collected from 2001 to 2003. An example plot code is: AZ093002KM040C.

<idnum>: NoAzGAP data collected by the U.S. Geological Survey Colorado Plateau Research Station in 1997 for accuracy assessment of the first generation Arizona Gap Analysis land cover map. An example plot code is 2832.

Terra_CPRS: These plots are from interpretation of aerial photography available on the TerraServer web site, <http://terraserver.microsoft.com/>. The U.S. Geological Survey SWReGAP team did the interpretation. The plot identification number was accidentally stripped off in the CSOURCE column. An example plot code is: AZ2DQT119.

Utah

Utdoq_79: Data with an identifier with this sort of structure is interpretation of black and white Digital Orthophoto Quads available from the state of Utah from AGRC. DOQ interpretation was done by the RSGIS lab personnel.

UT100202MD20: Data with an identifier with this sort of structure was collected by the Utah State University RSGIS lab SWReGAP field team.

b) Summary of samples:

1325 samples were available to model this mapping area. A polygon coverage containing all samples is located at: AZ\ARCHIVE\AZ3\TRAIN_DATA.

Code	Ecological System Name	#
D04	INVASIVE SW RIPARIAN WOODLAND AND SHRUBLAND	42
D09	INVASIVE ANNUAL FORBLAND	3
N31	BARREN	33
S010	COLORADO PLATEAU MIXED BEDROCK CANYON AND	353

	TABLELAND	
S011	INTER-MOUNTAIN BASINS SHALE BADLANDS	35
S012	INTER-MOUNTAIN BASINS ACTIVE AND STABILIZED DUNES	42
S032	ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND	3
S036	ROCKY MOUNTAINS PONDEROSA PINE WOODLAND	2
S039	COLORADO PLATEAU PINYON-JUNIPER WOODLAND	93
S046	ROCKY MOUNTAINS GAMBEL OAK - MIXED MONTANE SHRUBLAND	3
S054	INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND	41
S059	COLORADO PLATEAU BLACKBRUSH-MORMON TEA SHRUBLAND	111
S065	INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB	111
S075	INTER-MOUNTAIN BASINS JUNIPER SAVANNA	5
S079	INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE	156
S090	INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND	112
S093	ROCKY MOUNTAINS LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND COMPLEX	13
S096	INTER-MOUNTAIN BASINS GREASEWOOD FLAT COMPLEX	76
S136	SOUTHERN COLORADO PLATEAU SAND SHRUBLAND	59
	**TOTAL	1325

c) Preparing Sample Data for Classification Tree (CT) Modeling:

The point sample data were prepared for decision tree modeling input by the following steps:

1. Create UTM point coverage of all point data with attributes (ArcInfo commands generate, dbaseinfo, additem, tables, joinitem. Note that additional fields are idnum, to accomplish the join, and N4RPTS, which is the number of random points that will be extracted from each sample polygon and is initially set to 20)
2. Project UTM points to Albers
3. Flag points that are within 30 meters of a road and increase their buffer value if they are smaller than 90m (this step is to ensure enough sample points once the roads are erased):
 - a. Identity az_pt1 az_tran_b30m az_pt2
 - b. Tables
 - c. Sel az_pt2.pat
 - d. Resel inside = 100
 - e. Resel buffer < 90 (smaller)
 - f. Resel buffer > 15 (not riparian)
 - g. Calc buffer = 90 (to get enough to sample once roads are erased)
 - h. Calc num4rpts = 40 (define more samples near roads as some will be erased to eliminate pixels on the roads)

4. Buffer the field points according to the field notes as modified by step 3. Use ArcView extension BufferThemeBuilder, choose attribute 'buffer' as the distance field.
5. Find overlapping polygons and eliminate (either re-label, if appropriate, or delete)
6. Merge all neighboring data (adjacent mapzones and states).
7. Merge with data points identified from imagery. For example, class S010 was digitized as polygons but included in the decision tree modeling by: buffering polygons by 100m (to avoid mixed pixels at edges) and assigning the number of random points to be extracted for these polygons to 5 (N4RPTS). These polygons were added after they were 'erased' by the buffered polygons of field data, so that field data took precedence.
8. Recode the character field containing the Ecosystem Code (ES_code) to a number (ES2NUM) using ArcView extension Recode.
9. Create 80% training and 20% validation polygon sets from the sample polygons. Use ArcView with featureselect extension, select 20 percent for validation set, recorded in field 'ESWH20'.
10. Create random points within each sample polygon. Use ArcView extension Random Point Generator v. 1.27, specify N4RPTS as the field with the number of points desired, convert shapefile to coverage az_rpts_all in ArcInfo.
11. Delete random points within 45m of roads. (In ArcInfo, identify the random points with the 45m road buffer coverage; in ArcEdit, delete points with INSIDE = 100).
12. Attached sample polygon attributes to the random points using Identity in ArcInfo and the azx_polys_all coverage.
13. In ArcInfo, copy the sample polygon set (azx_polys_all) twice to create azx_polys_trn and azx_polys_val. In ArcEdit, delete the polygons with ESWH20 = 1 in azx_polys_trn and delete the polygons with ESWH20 = 0 in azx_polys_val.
14. In ArcInfo, copy the sample point set (azx_rpts_all) twice to create azx_rpts_trn and azx_rpts_val. In ArcEdit, delete the polygons with ESWH20 = 1 in azx_rpts_trn and delete the polygons with ESWH20 = 0 in azx_rpts_val.
15. Convert the three random point coverages (azx_rpts_all, azx_rpts_trn and azx_rpts_val) to grids using the ES2NUM field as the grid value. Use ArcView with Spatial analysis tool, specify extent and grid cells to match the ETM+ imagery.
16. Import the grids azx_rpts_all and azx_rpts_trn into ERDAS Imagine as *.img files for Cart modeling.

3) Cover types:

a) Classification Tree modeled cover types:

The following cover types were modeled using the See5 Classification Tree:

Code	Ecological System Name
D04	INVASIVE SW RIPARIAN WOODLAND AND SHRUBLAND
D09	INVASIVE ANNUAL FORBLAND

N31	BARREN
S010	COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND
S011	INTER-MOUNTAIN BASINS SHALE BADLANDS
S012	INTER-MOUNTAIN BASINS ACTIVE AND STABILIZED DUNES
S032	ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND
S036	ROCKY MOUNTAINS PONDEROSA PINE WOODLAND
S039	COLORADO PLATEAU PINYON-JUNIPER WOODLAND
S046	ROCKY MOUNTAINS GAMBEL OAK - MIXED MONTANE SHRUBLAND
S054	INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND
S059	COLORADO PLATEAU BLACKBRUSH-MORMON TEA SHRUBLAND
S065	INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB
S075	INTER-MOUNTAIN BASINS JUNIPER SAVANNA
S079	INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE
S090	INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND
S093	ROCKY MOUNTAINS LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND COMPLEX
S096	INTER-MOUNTAIN BASINS GREASEWOOD FLAT COMPLEX
S136	SOUTHERN COLORADO PLATEAU SAND SHRUBLAND

b) Non CT modeled cover types:

The following types were included as non-modeled cover types:

Code	Ecological System Name
S014	INTER-MOUNTAIN BASINS GREASEWOOD WASH
N11	OPEN WATER
N21	DEVELOPED, LOW INTENSITY
N22	DEVELOPED, MED-HIGH INTENSITY
N80	AGRICULTURE
N82	AGRICULTURE, CULTIVATED CROPS
D03	RECENTLY MINED OR QUARRIED

S014: INTER-MOUNTAIN BASINS GREASEWOOD WASH

For AZ3, the nested unsupervised classification accomplished for the open water class was used to extract linear features of moderately abundant vegetation associated with drainage patterns. These features were recoded to produce a binary image. The image was converted to a grid and the grid was converted to a polygon coverage.

The polygons extracted from the imagery were then given to a field mapper, who labeled or deleted polygons as Greasewood Wash based on field knowledge and ancillary data, including digital orthophoto quarter quad data.

N11: OPEN WATER

The majority of the open water class was extracted from the ETM+ imagery using a nested unsupervised classification procedure. The initial unsupervised classification used an input image stack consisting of the following 13 images:

- 12 pca bands resulting from a PCA analysis of the 12 band stack of the fall image and the summer image. This pca was calculated to enhance the seasonal differences between the two dates.
- Slope image, derived from the 30 meter DEM.

The 13 image stack was classified into 40 clusters using ERDAS Imagine's unsupervised classification module. Water was identified by visually inspecting the images for open water visible in either season. Several of the original 40 clusters were isolated and reclassified into 20 clusters. The classified water image was recoded to a binary image, clumped (allowing 8 neighbors), and eliminated (removing clusters less than 5 pixels in size), to produce a water image that retains only water bodies of 1 acre MMU. The image was converted to a grid and the grid was converted to a polygon coverage.

The polygons extracted from the imagery were then given to a field mapper, who labeled or deleted polygons by comparing them to information in the DeLorme Atlas (1993) and ancillary data sources. Areas where the image classification incorrectly identified pixels as water (e.g., areas of shadow) were captured via screen digitizing as a polygon coverage that was used to erase the water coverage. Additional areas of open water visible in either the fall or summer image (but missed by the nested unsupervised classification) were screen digitized over ETM+ imagery at a scale of 1:24,000 or larger using ArcMap.

N21 DEVELOPED, LOW INTENSITY-OPEN SPACE

N22 DEVELOPED, MED-HIGH INTENSITY

The urban areas were identified on ETM+ imagery and screen digitized at a scale of 1:50,000 or larger (1:24,000) using ArcMAP. Urban areas were recognized by their distinctive rectilinear patterns of intersecting road networks and variable patterns of spectral brightness, including extremely bright visible reflectance associated with roofing materials and concrete. Fall, spring and summer images were variously used. When the image pattern was questionable, urban areas were verified by referring to ancillary data, including the DeLorme Atlas (1993) and the ALRIS urban area coverage (based on census tract data). The shapefile created using the fall image was then overlain on the spring and summer images and additional urban areas visible on these images were captured. Two classes of urban areas were recognized: N21 (open space - low intensity developed) and N22 (medium - high intensity developed). These were differentiated on the basis of the density of 'green space' included, i.e. yards, parks, golf courses, etc.

- N80 AGRICULTURE
- N82 AGRICULTURE, CULTIVATED CROPS

The agricultural areas were identified on ETM+ imagery and screen digitized at a scale of 1:50,000 or larger (1:24,000) using ArcMAP. Typically, agricultural fields were easily recognized by their rectangular geometries and distinctive patterns of uniform spectral brightness, including extremely bright near-infrared reflectances found in at least one of the seasonal images. Fall, spring and summer images were variously used. When the image pattern was questionable, agricultural areas were verified by referring to ancillary data, including the DeLorme Atlas (1993). Where the distinction between cultivated crops and pastureland could not be made, N80 was used as a label.

D03 RECENTLY MINED OR QUARRIED

Recently mined areas were identified on the fall ETM+ imagery and screen digitized over ETM+ imagery at a scale of 1:24,000 or larger. The image was overlaid with locations of surface mining activities extracted from the USGS Minerals point coverage found at: <http://mrdata.usgs.gov/>. Areas of apparent disturbance (anomalous reflectance on the image) that were collocated with known surface mining activity were identified and digitized as land cover type D01. A few larger mined areas were digitized with the Urban Areas and extracted from that shapefile to include in this coverage.

The following cover type was modeled with a post-classification model (see section 5b for details):

Code	Ecological System Name
S034	ROCKY MOUNTAINS MONTANE MESIC MIXED CONIFER FOREST AND WOODLAND

4) Summary of predictor layers used:

a) Multi band predictors:

Multi band predictors:

- AZ3_fa.img (ETM bands 1-5 & 7 for fall)
- AZ3_sp.img (ETM bands 1-5 & 7 for spring)
- AZ3_su.img (ETM bands 1-5 & 7 for summer)

All multi-band predictors can be found at: *AZ\ARCHIVE\AZ3\IMG_FILES*

b) Single band predictors:

- aspect.img Categorical 5 class aspect
- elev.img Continuous (integer) elevation

fabrt.img	Fall brightness band
fagrn.img	Fall greenness band
spbrt.img	Spring brightness band
spgrn.img	Spring greenness band
subrt.img	Summer brightness band
sugrn.img	Summer greenness band
landf.img	Categorical 10 class landform (modeled from DEM)

All multi-band predictors can be found at: AZ\ARCHIVE\AZ3\IMG_FILES

5) Modeling Methods:

a) See5 Classification Tree modeling:

Samples: 20% of the all sample polygons were withheld for validation. With the remaining 80%, 20 sub-samples were randomly selected for each sample polygon. This was done by first randomly generating points within each polygon and then converting the points to a raster *.img file. Pixels in the *.img (each to be considered a separate observation for the See5 classifier) were ‘drilled’ through predictor layers using the Sampling tool from CART Module for Imagine (EarthSatellite Corp. 2003), producing two important files for See5: the *.names and *.data files.

See5 Classification Tree: See5 (Release 1.8) data mining software (Rulequest 2004) was used for generating classification trees. Boosting was employed using 15 trials.

The See5 files are located: AZ\ARCHIVE\AZ3\OUTPUT\. The following briefly describes these files (Rulequest 2004).

*.names file: Identifies the dependent variable *.img file and the predictor *.img files created from the CART Module Sampling tool. Required by See5 software.

*.data file: Contains the training cases from which See5 extracts rules. This is also produced from the CART Module Sampling tool, by ‘drilling’ the dependent variable pixels through the specified predictor images. Required by See5 Software.

*.test file: Produced from the CART Module Sampling tool, but not used by SWReGAP. This file, if populated, would contain a separate ‘test’ set of cases to evaluate the rules generated from See5. The SWReGAP mapping procedures did not populate this file, and it was not used.

*.names.hst file: Produced from the CART Module Sampling tool. Details the distribution of samples available within the dependent input, and those output to the *.data and *.test file. Not required by See5, but produced by CART Module Sampling tool.

*.set file: Produced from See5 software. This file contains the settings for the classification tree run. For example the third value ‘15’ indicates the number of boosts used for boosting.

*.tree file: Produced from the See5 software. This file contains the classification tree in 'tree' format. This along with the *.data and *.names file are required by the CART Module Classifier tool to spatially apply the tree.

*.out file: Output file generated by See5 and displayed when See5 classification tree model has run. This file provides a visual representation of the classification tree that is somewhat easier to interpret than the *.tree file.

As a result of spatially applying the classification tree using the CART Module's Classifier tool, two files are created: an *.img file, which is the spatial application of the tree's rules, and the *_error.img file which is spatial depiction of confidence in the rules generated by the tree and displayed pixel by pixel.

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

The post-classification model described below was created to assist in differentiating between ecological system cover types that possess similar characteristics. In these cases species composition, species presents and/or ecoregional location was not enough to distinguish between the two types especially when the types share similar physiognomic and hydrologic characteristics. The following model is a tool for discriminating between these ecological system cover types.

The .gmd file used for this post-classification model is:
AZ\ARCHIVE\AZ3\POST_MODEL\swgap_az3.gmd

Step 1: Discriminating Mesic Conifer from Non-Mesic Conifer Forests Types.

The Southwest Regional GAP-National Vegetation Classification Standard legend has divided Rocky Mountain montane and subalpine forest types into two distinct classes; mesic forest types and dry-mesic forest types (Comer et al 2003). It was determined early in the modeling process that these types are difficult to discriminate using standard CT modeling methods alone due to vegetation composition similarities. As a result, it was decided that these types would be post modeled based strictly on conceptual meaning rather than spectral and ancillary differences.

A post-classification model was used to discriminate *S032 ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND* and *S034 ROCKY MOUNTAINS MONTANE MESIC MIXED CONIFER FOREST AND WOODLAND*. This post-classification model was designed to extract *S034 ROCKY MOUNTAINS MONTANE MESIC MIXED CONIFER FOREST AND WOODLAND* from the standard CT modeled *S032 ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND* ecological system cover type. NatureServe NVCS concept (Comer et al 2003) 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.”

A standard CT model was run labeling all montane conifer forests as *S032 ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND*. A conditional statement was then created to extract *S034 ROCKY MOUNTAINS MONTANE MESIC MIXED CONIFER FOREST AND WOODLAND* using the above aspect and landform characteristics in order to identify this montane mesic mixed conifer system.

The conditional statement is:

```
EITHER 34 IF (($n3_landf==2 OR $n3_landf==5 OR $n3_landf==6 OR $n3_landf==9)
AND ($n4_asp_az3_9cls==1 OR $n4_asp_az3_9cls==2 OR $n4_asp_az3_9cls==0)
AND ($n2_az3_allwr_mod2_rec==32)) OR $n2_az3_allwr_mod2_rec OTHERWISE
```

Where:

34 *S034 ROCKY MOUNTAINS MONTANE MESIC MIXED CONIFER FOREST AND WOODLAND*

32 *S032 ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND*

\$n3_landf 10 class landform grid
2= toe slopes, bottoms, and swales
5= very moist steep slopes
6= moderately moist steep slopes
9= cool aspect scarps, cliffs, canyons

\$n4_aspaz2 Nine class aspect image
1=North facing slope
2=Northeast facing slope
0=slope less than 3 degrees

\$n2_az3_allwr_mod2_rec AZ3 standard CT output image. Recoded to standard integer code.

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.

After the Clump & Eliminate step, the non CT modeled classes were ‘burned in’ to the final map using an overlay function. The final map can be found at:
\\AZ\ARCHIVE\AZ3\FINAL_MAP\.

After CT model validation, a final map was generated using the methods described in section 5a this time using 100% of the sample data. Similarly to the steps involved with the 80% CT model, post-classification modeling, inclusion of the non-modeled cover types, and generalization to the MMU were completed to create **az3_final.img**, which can be found at \\AZ\ARCHIVE\AZ3\FINAL_MAP.

6) Validation:

a) CT model validation: 20% of sample polygons were randomly selected and withheld from CT modeling using the featuresselect.avx as reference data. The CT model was run as described in section 5a using the remaining 80% of the samples. The 20% withheld samples were used to assess the predictive capability of the CT modeled map using the 80% training data. The CT-modeled classes were assessed using kappa.avx, which works by intersecting the validation sample polygons through the CT modeled land cover map, and considers the site correctly mapped when the majority of pixels within the sample polygon agree 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. These files can be found at:
\\AZ\ARCHIVE\AZ3\VALIDATION\.

b) Discussion of mapped cover types:

The following narrative provides qualitative assessments by the mapping team for each cover type mapped in this mapping area. It is intended to elaborate on the quantitative results of the CT model validation from the perspective of those most familiar with the map and the mapping process and is hoped to be of value to potential map users. User's accuracy indicates how well the model predicts a withheld sample point; producer's accuracy indicates how well a withheld sample was predicted by the model. When quantitatively evaluated, the error tendency can be described as confusion between types indicated by the producer's accuracy and the tendency of the mapped types to actually represent some other type as indicated by the user's error. Accuracy of 65% or above indicates good model performance. It is important to note that this is not a complete accuracy assessment of map performance, but is rather a preliminary measure of map performance. Also, the total number of withheld samples may differ from the total number of samples reported in 2b above. Overlapping training sites create duplicate ID fields and replicates site information; these training sites were ignored in the validation totals.

D03: RECENTLY MINED OR QUARRIED. Not quantitatively assessed. Qualitative assessment indicates that most recently mined or quarried areas contemporary with the date of the imagery are included.

D04: INVASIVE SOUTHWEST RIPARIAN WOODLAND AND SHRUBLAND. Quantitatively assessed, validation 67% (users) and 44% (producers) based on 9 independent validation samples. This type was mostly confused with S093 ROCKY MOUNTAIN LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND (3 of 9). Confusion can be explained by the similar vegetation structure and habitat of the two ecological systems.

D09: INVASIVE ANNUAL FORBLAND. Not quantitatively assessed due to low number of validation samples. In general this type did not model well as its occurrence is often due more to disturbance than edaphic factors and it is spectrally similar to native or only partially infested native grasslands.

N11: OPEN WATER. Not quantitatively assessed. Qualitative assessment indicates that most perennial open water mapped well. Ephemeral water bodies were not a focus of the effort and may be missing.

N21: DEVELOPED, LOW INTENSITY. Not quantitatively assessed. Some of this class may have been excluded due to the difficult nature of detecting low intensity development.

N22 DEVELOPED, MED-HIGH INTENSITY. Not quantitatively assessed. Qualitative assessment indicates that most medium and high density developed areas contemporary with the date of the imagery are included.

N31: BARREN. Quantitatively assessed, validation 50% (users) and 29% (producers) based on 7 independent validation samples. This type was confused with S059 COLORADO PLATEAU BLACKBRUSH-MORMON TEA SHRUBLAND, (1 of 7), S065 INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB (2 of 7), and S096 INTER-MOUNTAINS BASINS GREASEWOOD FLAT (1 of 1). Each of these types may have patches of low vegetation cover that could spectrally appear as N31.

N80 AGRICULTURE. Not quantitatively assessed. Qualitative assessment indicates that most agricultural areas contemporary with the date of the imagery are included.

N82 AGRICULTURE, CULTIVATED CROPS. Not quantitatively assessed. Qualitative assessment indicates that most agricultural areas contemporary with the date of the imagery are included; however, distinction between cultivated crops and pasturelands may not be consistent.

S010: COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND. Quantitatively assessed, validation 86% (users) and 86% (producers) based on 69 independent validation samples. In some cases this type was confused with S039 COLORADO PLATEAU PINYON-JUNIPER WOODLAND (4 of 69) or S079 INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE (2 of 69). Elements of S010 can be found within S039 and to a lesser extent S079.

S011: INTER-MOUNTAIN BASINS SHALE BADLAND. Quantitatively assessed, validation 57% (users) and 57% (producers) based on 7 independent validation samples. This type was confused with S065 INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB, S096 INTER-MOUNTAIN BASINS GREASEWOOD FLAT or S0136 SOUTHERN COLORADO PLATEAU SAND SHRUBLAND (1 of 7 in each case). Elements of S011 intergrade within each of these types.

S012: INTER-MOUNTAIN BASINS ACTIVE AND STABILIZED DUNES. Quantitatively assessed, validation 0% (users) and 33% (producers) based on 9 independent validation samples. This type was mostly confused with S059 COLORADO PLATEAU BLACKBRUSH-MORMON TEA SHRUBLAND (5 of 9). In this mapping zone, S012 intergrades with S059.

S014: INTER-MOUNTAIN BASINS GREASEWOOD WASH. Not quantitatively assessed.

S032: ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND. Quantitatively assessed, validation 0% (users) and 0% (producers) based on 1 independent validation samples. This type was confused with S010 COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND (1 of 1). There are too few independent validation samples to adequately assess the modeling of S032.

S034: ROCKY MOUNTAINS MONTANE MESIC MIXED CONIFER FOREST AND WOODLAND. Not quantitatively assessed. This cover type was mapped using a post processing modeler. Please reference methods section entitled *Post-classification, recoding and other modeling* section 5b for a detailed explanation

S036: ROCKY MOUNTAINS PONDEROSA PINE WOODLAND. Quantitatively assessed, validation 0% (users) and 0% (producers) based on 1 independent validation samples. This type was confused with S039 COLORADO PLATEAU PINYON-JUNIPER WOODLAND (1 of 1). There are too few independent validation samples to adequately assess the mapping of S032.

S039: COLORADO PLATEAU PINYON-JUNIPER WOODLAND. Quantitatively assessed, validation 54% (users) and 68% (producers) based on 19 independent validation samples. This type was confused with S010 COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND (3 of 19). These two types can intergrade in this functional unit.

S046: ROCKY MOUNTAINS GAMBEL OAK-MIXED MONTANE SHRUBLAND. Quantitatively assessed, validation 100% (users) and 100% (producers) based on 1 independent validation sample. Although this ecological system shows perfect mapping there are too few independent validation samples to adequately assess the mapping of S046.

S054: INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND. Quantitatively assessed, validation 60% (users) and 33% (producers) based on 9 independent validation samples. This type was confused with S039 COLORADO PLATEAU PINYON-JUNIPER WOODLAND (4 of 9). These two types can intergrade in this functional unit.

S059: COLORADO PLATEAU BLACKBRUSH-MORMON TEA SHRUBLAND. Quantitatively assessed, validation 55% (users) and 52% (producers) based on 23 independent validation samples. This type was confused with S012 INTER-MOUNTAIN BASINS ACTIVE AND STABILIZED DUNES (2 of 23) and S090 INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND (2 of 23). S039 and S012 can intergrade in this functional unit.

S065: INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB. Quantitatively assessed, validation 50% (users) and 39% (producers) based on 23 independent validation samples. This type was confused with S079 INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE (4 of 23), S090 INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND (3 of 23), and S096 INTER-MOUNTAIN BASINS GREASEWOOD FLAT (4 of 23). All of these are shrublands that are expected to have similar spectral signal as S065 and have some floristic overlap.

S075: INTER-MOUNTAIN BASINS JUNIPER SAVANNA. Quantitatively assessed, validation 33% (users) and 100% (producers) based on 1 independent validation sample. This type was sometimes mapped as S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND or S079 INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE. Both S054 and S079 can intergrade with S075. There are too few independent validation samples to adequately assess the modeling of S075.

S079: INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE. Quantitatively assessed, validation 45% (users) and 41% (producers) based on 32 independent validation samples. This type was mainly confused with S090 INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND (8 of 32). These two ecological systems occur in a mosaic in this functional unit.

S090: INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND. Quantitatively assessed, validation 32% (users) and 43% (producers) based on 23 independent validation samples. This type was mainly confused with S079 INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE (6 of 23). These two ecological systems occur in a

mosaic in this functional unit and in the field it is sometimes difficult to determine exact lines of distinction between them.

S093: ROCKY MOUNTAIN LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND. Quantitatively assessed, validation 0% (users) and 0% (producers) based on 3 independent validation samples. This type was confused with D04 INVASIVE SOUTHWEST RIPARIAN WOODLAND AND SHRUBLAND (2 of 3) or S032 ROCKY MOUNTAIN MONTANE DRY-MESIC MIXED CONIFER FOREST (1 of 3). Confusion between S093 and D04 is understandable since the types have similar structure and occur in the same habitat. S093 and S032 also have similar structure but S032 occurs upslope of S093. However, are too few independent validation samples to adequately assess the mapping of S093.

S096: INTER-MOUNTAIN BASINS GREASEWOOD FLAT. Quantitatively assessed, validation 27% (users) and 40% (producers) based on 15 independent validation samples. This type was confused with S059 COLORADO PLATEAU BLACKBRUSH-MORMON TEA SHRUBLAND (2 of 15), S065 INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB (2 of 15), or S136 SOUTHERN COLORADO PLATEAU SAND SHRUBLAND (3 of 15). All of these types are shrublands with expected similar spectral properties. S096 can intergrade or occur as a mosaic with each of the other types.

S136: SOUTHERN COLORADO PLATEAU SAND SHRUBLAND. Quantitatively assessed, validation 33% (users) and 42% (producers) based on 12 independent validation samples. This type was confused with S079 INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE (2 of 12), S090 INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND (3 of 12), or S096 INTER-MOUNTAIN BASINS GREASEWOOD FLAT (2 of 12). All of these types are shrublands with expected similar spectral properties. S136 can intergrade or occur as a mosaic with each of the other types.

7) Citations:

Chavez, P. S. Jr. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data.

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.

DeLorme Mapping. 1993. Arizona Atlas and Gazetteer. Freeport, Maine.

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. 2001a. t-sensor reflectance: A first order normalization of Landsat 7 ETM+ Images. (<http://landcover.usgs.gov/pdf/huang2.pdf>)

Huang, C., B. Wylie, L. Yang, C. Homer and G. Zylstra. 2001b. Derivation of tasseled cap transformation based on Landsat 7 at-satellite reflectance. *International Journal of Remote Sensing*, 8: 1741-1748.