SWReGAP Land Cover Mapping Methods Documentation

Functional Unit or Mapping Zone: Arizona 5

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

Person Preparing Document: Kathryn Thomas, Cynthia Wallace

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:

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

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 cutline" 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)
36/37	2000-103	2000-167	1999-292
36/38	2000-103	2000-167	1999-292
37/36	2000-110	2000-158	1999-283
37/37	2000-078	2000-142	1999-283
37/38	2000-078	2000-110	1999-283
38/35	2000-117	2000-165	1999-290
38/36	2000-101	2000-165	1999-290
38/37	2000-085	2000-117	1999-274
38/38	2000-085	2000-117	1999-274
39/35	2000-092	2000-156	1999-297

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

AZ\ARCHIVE\AZ5\MOSAIC\AZ5_fa.mos
AZ\ARCHIVE\AZ5\MOSAIC\AZ5_sp.mos
AZ\ARCHIVE\AZ5\MOSAIC\AZ5_su.mos

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

AZ\ARCHIVE\AZ5\MOSAIC\mosaic\AZ5_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\AZ5\IMG_FILES.

Tasseled cap: Brightness, Greenness & Wetness band transformations wee 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\AZ5\IMG_FILES.

All imagery derived predictor layers can be found at:

AZ\ARCHIVE\AZ5\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\AZ5\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 SOURCE field in the \AZ\ARCHIVE\AZ45TRAIN_DATA\az5_polys_all identifies the source for each sample. 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.

AzGF: Data were collected by the Arizona Department of Game and Fish along riparian corridors from 1993 to 2000. An example plot code is: AZSM004.

BuenosAiresNWR: Data collected by the Bureau of Reclamation for vegetation mapping in 2002 and 2003. An example plot code is: BAR.11

Dig: 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: Kat3.

Field_CPRS: Data collected by the U.S. Geological Survey SWReGAP field crew from 2001 to 2003. Plot code structure is:
AZ<month/day><year><teamcode><idnum><sequence letter, if appl>. An example plot code is: AZ110902KM043.

Malusca: Data collected by Jim Malusca of the U.S. Geological Survey Southwest Biological Science Center at Organ Pipe National Park in 2002. An example plot code is: AZ01040202011.

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: 15421

NvGap: Samples collected by the Nevada SWReGAP field crew. They were collected in 2002. An example plot code is: NV053002TS21

PSI: Data collected by the Pacific Biodiversity Institute at the Sonoran Desert National Monument in 2002 and 2003. An example plot code is: 177.

SoAzGAP: Identifies data from an existing database collected for accuracy assessment of the first generation Arizona Gap Analysis land cover map. Data were collected by the U.S. Geological Survey Sonoran Desert Research Station in 1997. An example plot code is: 1109.

TOH: 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: AZ040203DQ014/

YumaProvingGrounds: Data collected at the Department of Defense Yuma Proving Grounds in 1998. An example plot code is: AZ061091YP172.

b) Summary of samples:

3207 samples were available to model this mapping area. A polygon coverage containing all samples is located at: AZ\ARCHIVE\AZ5\TRAIN_DATA.

Code	Ecological System Name	#

D04	INVASIVE SW RIPARIAN WOODLAND AND SHRUBLAND	41
N31	BARREN	94
S020	NORTH AMERICAN WARM DESERT WASH	13
S036	ROCKY MOUNTAINS PONDEROSA PINE WOODLAND	16
S040	GREAT BASIN PINYON-JUNIPER WOODLAND	77
S057	MOGOLLON CHAPARRAL	101
S058	MESQUITE UPLAND SCRUB	91
S060	MOJAVE MID-ELEVATION MIXED DESERT SCRUB	175
S063	SONORAN PALO VERDE-MIXED CACTI DESERT SCRUB	1270
S069	SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB	1029
S070	SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB	81
S097	NORTH AMERICAN WARM DESERT RIPARIAN WOODLAND AND SHRUBLAND	37
S098	NORTH AMERICAN WARM DESERT RIPARIAN MESQUITE BOSQUE	36
S129	SONORAN MID-ELEVATION DESERT SCRUB	85
	**TOTAL	3207

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, identity 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:

Code	Ecological System Name
D04	INVASIVE SW RIPARIAN WOODLAND AND SHRUBLAND
N31	BARREN
S020	NORTH AMERICAN WARM DESERT WASH
S036	ROCKY MOUNTAINS PONDEROSA PINE WOODLAND
S040	GREAT BASIN PINYON-JUNIPER WOODLAND
S057	MOGOLLON CHAPARRAL
S058	MESQUITE UPLAND SCRUB

S060	MOJAVE MID-ELEVATION MIXED DESERT SCRUB
S063	SONORAN PALO VERDE-MIXED CACTI DESERT SCRUB
S069	SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB
S070	SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB
S097	NORTH AMERICAN WARM DESERT RIPARIAN WOODLAND AND SHRUBLAND
S098	NORTH AMERICAN WARM DESERT RIPARIAN MESQUITE BOSQUE
S129	SONORAN MID-ELEVATION DESERT SCRUB

b) Non CT modeled cover types:

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

Code	Ecological System Name
S018	NORTH AMERICAN WARM DESERT ACTIVE AND STABILIZED DUNES
S019	NORTH AMERICAN WARM DESERT VOLCANIC ROCKLAND
S020	NORTH AMERICAN WARM DESERT WASH
S022	NORTH AMERICAN WARM DESERT PLAYA
N11	OPEN WATER
N21	DEVELOPED, LOW INTENSITY
N22	DEVELOPED, MED-HIGH INTENSITY
N81	AGRICULTURE, PASTURE/HAY
N82	AGRICULTURE, CULTIVATED CROPS
D02	RECENTLY BURNED
D03	RECENTLY MINED OR QUARRIED
D14	DISTURBED- OIL WELL

S018: NORTH AMERICAN WARM DESERT ACTIVE AND STABILIZED DUNES

The active and stabilized dunes class was extracted from the ETM+ imagery using an unsupervised classification procedure. The initial unsupervised classification used an input image stack consisting of the raw bands of the fall image. The AOI was selected for identifying dunes based on field knowledge and the DeLorme Atlas (1993). The 6 band image stack was classified into 20 clusters using Erdas Imagine's unsupervised classification module. Dunes were identified by visually inspecting the images for bright areas known to contain dunes. The classified dunes image was recoded to a binary image and then filtered, applying a 3x3 majority filter three times, to reduce speckle and produce an image of contiguous dune fields. For AOIs, see \AOI\dune_aoi folder in the ARCHIVE.

S019: NORTH AMERICAN WARM DESERT VOLCANIC ROCKLAND

The volcanic rock and cinder land class was extracted from the ETM+ imagery using an unsupervised classification procedure. The unsupervised classification used an input image stack consisting of the 6 non-thermal raw bands of the summer image. Several

distinct AOIs were selected for identifying volcanics based on field knowledge and inspection of the imagery. Volcanics were identified by visually inspecting the images for dark areas known to contain volcanics, commonly with distinctive crater or flow textures. Areas thought to contain volcanics based on inspection of the imagery were confirmed by referring to the Roadside Geology of Arizona (Chronic 1983) and/or the Delorme Atlas (1993). The 6 band image stack was classified into 20 or 30 clusters using Erdas Imagine's unsupervised classification module. The classified volcanics image was recoded to a binary image, clumped (allowing 4 neighbors), and eliminated (removing clusters less than 5 pixels in size), to produce a volcanics image that retains only features of 1 acre MMU. For AOIs, see \AOI\lava_aoi folder in the ARCHIVE.

S020: NORTH AMERICAN WARM DESERT WASH

This class was extracted from the ETM+ imagery using a nested unsupervised classification procedure. The 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 Desert Wash based on field knowledge and ancillary data, including digital orthophoto quarter quad data.

S022: NORTH AMERICAN WARM DESERT PLAYA

This class was captured in two ways: screen digitizing over the ETM+ imagery and unsupervised classification of the ETM+ imagery within local AOIs. Some playas were identified on the fall ETM+ imagery and screen digitized over ETM+ imagery at a scale of 1:24,000 or larger. Other playas were identified on the imagery as distinct, closed, highly reflective patches in topographic lows and verified in the DeLorme Atlas (1993). AOIs were created around these features and an unsupervised classification of the 6 raw bands of the fall image was performed. The classification specified 20 clusters, which were labeled as a binary playa image based on inspection of the image. For AOIs, see \AOI\playa_aoi folder in the ARCHIVE.

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 re-classified 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
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.

N81: AGRICULTURE, PASTURE/HAY
N82: AGRICULTURE, CULTIVATED CROPS

D02: RECENTLY BURNED

Recently burned areas were identified on ETM+ imagery as dark patches, typically angular, commonly bordered by roads. These were captured by: defining an AOI around each burn identified in each season, performing an unsupervised classification within the AOI specifying 20 clusters, and recoding the clusters corresponding to the burned area. For AOIs, see AZ\ARCHIVE\AZ5\AOI\FIRE_AOI folder in the ARCHIVE.

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.

D14: DISTURBED-OIL WELL

The following cover types were modeled with a post-classification model (see section 5c for details):

Code	Ecological System Name
S016	NORTH AMERICAN WARM DESERT BEDROCK CLIFF AND OUTCROP
S062	CHIHUAHUAN MIXED DESERT AND THORN SCRUB
S112	MADREAN PINYON-JUNIPER WOODLAND

4) Summary of predictor layers used:

a) Multi band predictors:

Multi band predictors:

AZ5_fa.img	(ETM bands 1-5 & 7 for fall)
AZ5_sp.img	(ETM bands 1-5 & 7 for spring)
AZ5_su.img	(ETM bands 1-5 & 7 for summer)

All multi-band predictors can be found at: AZ\ARCHIVE\AZ5\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\AZ5\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\AZ5\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 too. 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 creates: 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\AZ5\POST_MODEL\swgap_az5.gmd

Step 1: A common artifact of using a thirty-meter digital elevation model (DEM) in the CT modeling process is the creation of artificial banding in the output. This banding is especially evident when representing vegetation on low to moderate elevation gradients. While recognizing the overall cover type benefit of incorporating the DEM in the modeling process, it was decided to post process the lower gradient banding areas rather than removing the DEM altogether.

A standard CT model was created which harnessed the overall benefits of the DEM as an input dependent variable. Another CT model was run, this time withholding elevation. A post processing modeler was created to remove the lower elevation areas that exhibited banding relics and replace them with pixels that have no elevation banding artifacts. Only cover types that where greatly influenced by the presence of the DEM where replaced.

The conditional statement is:

EITHER $n5_az5all_m3_j3_noelev_rec IF (($n2_elev < 322) AND
($n1_az5all_m3_j3_rec==231 OR $n1_az5all_m3_j3_rec==304 OR
$n1_az5all_m3_j3_rec==70 OR $n1_az5all_m3_j3_rec==63)) OR
$n1_az5all_m3_j3_rec OTHERWISE

Where:

$n5_az5all_m3_j3_noelev_rec	AZ5 output image ignoring the thirty meter digital elevation model (DEM). Recoded to standard integer code.
$n2_elev	Thirty meter digital elevation model (DEM)

$n1_az5all_m3_j3_rec	AZ5 standard output image. Recoded to standard integer code.
231	N31 NON-SPECIFIC BARREN
304	D04 INVASIVE SW RIPARIAN WOODLAND AND SHRUBLAND
70	S070 SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB
63	S063 SONORAN PALO VERDE-MIXED CACTI DESERT SCRUB

Step 2: Restrict *S020 NORTH AMERICAN WARM DESERT WASH* to riparian corridors. A post-classification model was used to limit the mappable area of S020 by restricting it to wash corridors.

S020 NORTH AMERICAN WARM DESERT WASH was found to be mapping in undesirable locations that were uncharacteristic of wash environments. It was the intention of this conditional statement to restrict S020 to a more realistic landscape.

A standard output CT model was run ignoring cover class S020 (see AZ\ARCHIVE\AZ2\POST_MODEL\POST_MODEL_IMAGES\az5all_m3_j3_nos020.img). This output was used as replacement vegetation pixel values for areas that were identified as unsuitable terrain for S020. Suitable terrain was identified by a 250 meter buffered stream corridor mask image.

The conditional statement is:

EITHER $n10_az5all_m3_j3_nos020_rec IF ($n11_az5bf250 <99 and $n3_memory==20) OR $n3_memory OTHERWISE

Where:

20	*S020 NORTH AMERICAN WARM DESERT WASH*
$n10_az5all_m3_j3_nos020_rec	AZ5 output image ignoring ecological system class S020. Recoded to standard integer code.
$$n11_az5bf250	250m buffered stream corridor 1-99= all other terrain 1000= suitable wash terrain
$n3_memory	AZ5 standard output image. Recoded to standard integer code. Note that this output is a temporary

memory that has been modified by previous steps in the post classification model.

Step 3: A post-classification model was used to discriminate S040.*S040 GREAT BASIN PINYON-JUNIPER WOODLAND* from *S112 MADREAN PINYON-JUNIPER WOODLAND*. A post-classification model was used to map S040.

The *S040 GREAT BASIN PINYON-JUNIPER WOODLAND* ecological system requires the presence of *Pinus monophylla*, among other characteristic, to be eligible for this ecological system nomenclaure. *S040 GREAT BASIN PINYON-JUNIPER WOODLAND*, typically found in the Great Basin ecoregion, was discovered to inhabit a small portion of the landscape in the northeast section of AZ5 functional mapping unit. This system differs from its widespread Madrean counterpart *S112 MADREAN PINYON-JUNIPER WOODLAND* by the absence of dominating *Pinus cembroides, Pinus discolor*, or other Madrean trees. Differenctiating between these two covertypes can be proplematic using standard CT modeling methods. Spectral and composition similarities utimately leads to misclassification by allowing this Great Basin system to be mapped far from its biogeographic range. In order to restrict S040 to the northeast section of AZ5 it was necessary to post-classify this type.

A standard CT model was run mapping all Pinyon-Juniper types as *S040 S040 GREAT BASIN PINYON-JUNIPER WOODLAND*. An area of interest (aoi) was created to identify S040 known area and the aoi was then converted to a mask image. This mask image (see AZ\ARCHIVE\AZ2\POST_MODEL\POST_MODEL_IMAGES\s112tos040.img) allows the model to restrict *S040 GREAT BASIN PINYON-JUNIPER WOODLAND* to its known range mapping the remaining Pinyon-Juniper as S112. Information on the known range for S040 was gathered from on site Arizona field crew observations.

The conditional statement is:

EITHER 112 IF ($n12_memory== 40 and $n78_s112tos040_rec==0) OR $n12_memory OTHERWISE

Where:

40	*S040 GREAT BASIN PINYON-JUNIPER WOODLAND*
112	*S112 MADREAN PINYON-JUNIPER WOODLAND*
$n78_s112tos040_rec Binary	AOI image indicating S040 range 1 = area identified by AOI to be know range 0 = all other area
$n12_memory	AZ5 standard CT output image. Recoded to standard integer code. Note that this output is a temporary memory

that has been modified by previous steps in the post classification model.

Step 4: A post-classification model was used to limit the mappable area *S060 MOJAVE MID-ELEVATION MIXED DESERT SCRUB* to the Mojave Desert and correctly mapping *S129 SONORAN MID-ELEVATION DESERT SCRUB* in its place.

S060 MOJAVE MID-ELEVATION MIXED DESERT SCRUB was found to be mapping out of its ecoregional habitat of the Mojave Desert. It was the intention of the following conditional statement to restrict S060 to known range. A standard output CT model was run allowing S060 to be mapped throughout AZ5. The TNC Ecoregion Boundary image (see AZ\ARCHIVE\AZ2\POST_MODEL\POST_MODEL_IMAGES\ecoreg_az5.img) was used to identify the Mojave Desert Ecoregion. This image allows the model to restrict *S060 MOJAVE MID-ELEVATION MIXED DESERT SCRUB* to its known Mojave range and recodes the remaining misclassified pixels as the Sonoran scrub counterpart S129.

The conditional statement is:

EITHER 129 IF (($n20_memory == 60) and ($n32_ecoreg_az5==0 OR $n32_ecoreg_az5==1 OR $n32_ecoreg_az5==2 OR $n32_ecoreg_az5==4)) OR $n20_memory OTHERWISE

Where:

60	*S060 MOJAVE MID-ELEVATION MIXED DESERT SCRUB*
129	*S129 SONORAN MID-ELEVATION DESERT SCRUB*
$n32_ecoreg_az5	AZ5 TNC Ecoregion Boundaries 0 = null 1 = Apache Highlands 2 = Colorado Plateau 3 = Mojave Desert 4 = Sonoran Desert
$n20_memory	AZ5 standard output image. Recoded to standard integer code. Note that this output is a temporary memory that has been modified by previous steps in the post classification model.

Step5: Discriminate between S069 *SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB* (<700 meters) and *S062 CHIHUAHUAN MIXED DESERT AND THORN* SCRUB (>700 meters) while restricting S069 from mapping in the Mojave Desert. *S060 MOJAVE MID-ELEVATION MIXED DESERT SCRUB* was mapped in place of S069 <700 meters.

Mapping regionally specific desert scrub types can be problematic when left to standard modeling methods. The majority of regional ecological system names change due to changes in plant species composition associated with geographical locality. These changes do not necessarily affect the cover type density or spectral reflectance. It is for this reason that the following conditional statement was created. Elevation and ecoregional boundaries play a strong role in defining these regionally specific types. Both were used to drive this post model.

The conditional statement is:

5a:

EITHER 62 IF (($n69_ecoreg_az5==0 OR$n69_ecoreg_az5==1 OR $n69_ecoreg_az5==2 OR $n69_ecoreg_az5==4) AND ($n68_memory==69) AND ($n70_elev > 700)) OR $n68_memory OTHERWISE

Where:

62	S062 CHIHUAHUAN MIXED DESERT AND THORN SCRUB
69	SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB
$n69_ecoreg_az5	AZ5 TNC Ecoregion Boundaries 0 = null 1 = Apache Highlands 2 = Colorado Plateau 3 = Mojave Desert 4 = Sonoran Desert
$n70_elev	Thirty meter digital elevation model
$n68_memory	AZ5 standard output image. Recoded to standard integer code. Note that this output is a temporary memory that has been modified by previous steps in the post classification model

5b:

EITHER 60 IF ($n70_elev< 700 AND $n71_memory==69 AND $n69_ecoreg_az5==3) OR $n71_memory OTHERWISE

Where:

60	MOJAVE MID-ELEVATION MIXED DESERT SCRUB
69	SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB
$n70_elev	Thirty meter digital elevation model
$n69_ecoreg_az5	AZ5 TNC Ecoregion Boundaries 0 = null 1 = Apache Highlands 2 = Colorado Plateau 3 = Mojave Desert 4 = Sonoran Desert
$n71_memory	AZ5 standard output image. Recoded to standard integer code. Note that this output is a temporary memory that has been modified by previous steps in the post classification model

Step6: Mapping *S016 NORTH AMERICAN WARM DESERT BEDROCK CLIFF AND OUTCROP* on the Copper Mountains. A conditional statement was created to map S016.

The conditional statement is:

EITHER 16 IF ($n37_memory==63 AND $n48_elev >500 AND $n57_s063tos016_rec==1) OR $n37_memory OTHERWISE

Where:

16	*S016 NORTH AMERICAN WARM DESERT BEDROCK CLIFF AND OUTCROP*
63	S063 *SONORAN PALO VERDE-MIXED CACTI DESERT SCRUB*
$n48_elev	Thirty meter digital elevation model.
$n57_s063tos016_rec	Binary mask image. Identifies area where S063 SONORAN PALO VERDE-MIXED CACTI DESERT SCRUB was recoded to S016 NORTH AMERICAN WARM DESERT BEDROCK CLIFF AND OUTCROP. 1= Copper Mountain area 0= All other area

$n37_memory	AZ5 standard output image. Recoded to standard integer code. Note that this output is a temporary memory that has been modified by previous steps in the post classification model

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\AZ5\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 **az5_final.img**, which can be found at \AZ\ARCHIVE\AZ5\FINAL_MAP.\

6) Validation:

a) CT model validation: 20% of sample polygons were randomly selected and withheld from CT modeling using the featureselect.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\AZ5\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% of 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.

D02: RECENTLY BURNED. Not quantitatively assessed. Qualitative assessment indicates that many recently burned areas contemporary with the date of the imagery are included. Recently burned areas were identified on ETM+ imagery as dark patches, typically angular, commonly bordered by roads. Burned areas, however, can be confused with recently logged and regenerating areas or shadows. Since the date of the imagery some substantial fires have occurred, such as the Aspen fire. The perimeters of these recent fires may be obtained through Geomac (geomac.usgs.gov) at ftp://ftp.geomac.gov/outgoing.

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 56% (users) and 63% (producers) based on 8 independent validation samples. This type was confused with S063 SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB (1 of 8), S070 SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB (1 of 8), or S097 NORTH AMERICAN WARM DESERT RIPARIAN WOODLAND AND SHRUBLAND (1 of 8). D04 was confused with other riparian systems and with adjacent upland systems.

D14: DISTURBED OIL WELL. Not quantitatively assessed. This system probably reflects some other type of mining than oil exploration.

N11: 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 70% (users) and 37% (producers) based on 19 independent validation samples. This type was confused with D04 INVASIVE SOUTHWESTERN RIPARIAN WOODLAND AND SHRUBLAND (1 of 19), S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (7 of 19), and S070 SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB (4 of 19). N31 may occur in a mosaic with any of the systems with which it was confused.

N81: AGRICULTURE, PASTURE-HAY. Not quantitatively assessed.

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.

S016: NORTH AMERICAN WARM DESERT BEDROCK CLIFF AND OUTCROP. Not quantitatively assessed.

S018: NORTH AMERICAN WARM DESERT ACTIVE AND STABLIZED DUNES. Not quantitatively assessed. Qualitative assessment indicates that dunes visible on the satellite imagery are generally captured in the mapping.

S019: NORTH AMERICAN WARM DESERT VOLCANIC ROCKLAND. Not quantitatively assessed.

S020: NORTH AMERICAN WARM DESERT WASH. Quantitatively assessed, validation 0% (users) and 0% (producers) based on 3 independent validation samples. This type was confused with D04 INVASIVE SOUTHWESTERN RIPARIAN WOODLAND AND SHRUBLAND (2 of 3) and S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (1 of 3). S020 was confused with other riparian systems and with adjacent upland systems.

S021: NORTH AMERICAN WARM DESERT PAVEMENT. Quantitatively assessed, validation 50% (users) and 13% (producers) based on 16 independent validation samples. This type was confused with S057 MOGOLLON CHAPARRAL (7 of 16), MOJAVE MID-ELEVATION MIXED DESERT SCRUB (3 of 16), S063 SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB (2 of 16), S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (1 of 16), and S0129 SONORAN MID-ELEVATION DESERT SCRUB (1 of 16). S021 may occur in a mosaic with any of the systems with which it was confused.

S022: NORTH AMERCIAN WARM DESERT PLAYA. Not quantitatively assessed. Qualitative assessment indicates that playas visible on the satellite imagery are generally captured in the mapping.

SO36: ROCKY MOUNTAIN PONDEROSA PINE WOODLAND. Quantitatively assessed, validation 33% (users) and 25% (producers) based on 4 independent validation samples. This type was confused with S057 MOGOLLON CHAPARRAL (3 of 4) which can intergrade with S036 in this functional unit.

S057: MOGOLLON CHAPARRAL. Quantitatively assessed, validation 60% (users) and 86% (producers) based on 21 independent validation samples. This type was confused with S036 ROCKY MOUNTAIN PONDEROSA PINE WOODLAND (2 of 21) and S0129 SONORAN MID-ELEVATION DESERT SCRUB (1 of 21). S057 is generally well mapped.

S058: CHIHUAHUAN MESQUITE UPLAND SCRUB. Quantitatively assessed, validation 69% (users) and 47% (producers) based on 19 independent validation samples. This type was confused with S063 SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB (5 of 19), S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (4 of 19), and S070 SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB (1 of 19). S058 was confused with other upland shrublands.

S060: MOJAVE MID-ELEVATION MIXED DESERT SCRUB. Quantitatively assessed, validation 56% (users) and 57% (producers) based on 35 independent validation samples. This type was confused with S063 SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB (9 of 35), S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (2 of 35), and S0129 SONORAN MID-ELEVATION DESERT SCRUB (3 of 35). S060 was confused with other shrublands in the functional unit.

S062: CHICHUAHUAN MIXED DESERT AND THORN SCRUB. Not quantitatively assessed.

S063: SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB. Quantitatively assessed, validation 75% (users) and 85% (producers) based on 255 independent validation samples. This type was confused with SO57 MOGOLLON CHAPARRAL (1 of 255), S058 CHIHUAHUAN MESQUITE UPLAND SCRUB (2 of 255), S060 MOJAVE MID-ELEVATION MIXED DESERT SCRUB (4 of 255), S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (28 of 255), and S0129 SONORAN MID-ELEVATION DESERT SCRUB (2 of 255). S063 is generally well mapped.

S069: SONORA-MOJAVE CREOSTOEBUSH-WHITE BURSAGE DESERT SCRUB. Quantitatively assessed, validation 73% (users) and 72% (producers) based on 206 independent validation samples. This type was confused with N31 BARREN (2 of 206), SO21 NORTH AMERICAN WARM DESERT PAVEMENT (1 of 206), S058

CHIHUAHUAN MESQUITE UPLAND SCRUB (1 of 206), S060 MOJAVE MID-ELEVATION MIXED DESERT SCRUB (6 of 206), S063 SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB (47 of 206), and S0129 SONORAN MID-ELEVATION DESERT SCRUB (1 of 206). Although S069 can be confused with other shrublands or with N31 and S021 which can intergrade with it, the system is generally well mapped.

S070: SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB. Quantitatively assessed, validation 36% (users) and 24% (producers) based on 17 independent validation samples. This type was confused with N31 BARREN (1 of 17), S063 SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB (1 of 17), and S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (11 of 17). S070 was confused with other shrublands in the functional unit.

S097: NORTH AMERICAN WARM DESERT RIPARIAN WOODLAND AND SHRUBLAND. Quantitatively assessed, validation 50% (users) and 25% (producers) based on 8 independent validation samples. This type was confused with D04 INVASIVE SOUTHWEST RIPARIAN WOODLAND AND SHRUBLAND (1 of 8), S020 NORTH AMERICAN WARM DESERT WASH (1 of 8), S058 CHIHUAHUAN MESQUITE UPLAND SCRUB (1 of 8), S069 SONORA-MOJAVE CREOSOTEBUSH-WHITE BURSAGE DESERT SCRUB (1 of 8), S070 SONORA-MOJAVE DESERT MIXED SALT DESERT SCRUB (1 of 8), and S098 NORTH AMERICAN WARM DESERT RIPARIAN MESQUITE BOSQUE (1 of 8). S097 was sometimes confused with other riparian systems or with adjacent upland systems.

S098: NORTH AMERICAN WARM DESERT RIPARIAN MESQUITE BOSQUE. Quantitatively assessed, validation 83% (users) and 71% (producers) based on 7 independent validation samples. This type was confused with S063 SONORAN PALOVERDE-MIXED CACTI DESERT SCRUB (1 of 7) and S097 NORTH AMERICAN WARM DESERT RIPARIAN WOODLAND AND SHRUBLAND (1 of 7). S098 is generally well mapped.

S112: MADREAN PINYON-JUNIPER WOODLAND. Not quantitatively assessed.

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.

Chronic, H. 1983. Roadside Geology of Arizona. Mountain Press Publishing Company. 322 p.

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. Vogelman 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.