

SWReGAP Land Cover Mapping Methods Documentation

Functional Unit or Mapping Zone: UT4.....	1
Organization: RSGIS Laboratory, Utah State University.....	1
1) Predictor layer preparation:.....	1
a) <i>Image standardization</i> :.....	1
b) <i>Image dates and mosaicking</i> :.....	2
c) <i>Image derived datasets</i> :.....	2
d) <i>DEM derived datasets</i> :.....	2
2) Samples:.....	3
a) <i>Sample collection methods</i> :.....	3
b) <i>Summary of samples</i> :.....	3
3) Cover types:.....	4
a) <i>Classification Tree modeled cover types</i> :.....	4
b) <i>Non CT modeled cover types</i> :.....	5
4) Summary of predictor layers used:.....	5
a) <i>Multi band predictors</i> :.....	5
5) Modeling Methods:.....	6
a) <i>See5 Classification Tree modeling</i> :.....	6
b) <i>Post-classification, recoding and other modeling steps</i> :.....	7
c) <i>Generalizing to MMU and map completion</i> :.....	11
6) Validation:.....	11
a) <i>CT model validation</i> :.....	11
b) <i>Final map</i> :.....	11
c) <i>Discussion of mapped cover types</i> :.....	11
7) Citations:.....	16

Functional Unit or Mapping Zone: UT4

Organization: RSGIS Laboratory, Utah State University

Person Preparing Document: John Lowry/Jessica Kirby/Lisa Langs

Date of Preparation: 27 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:

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

Where,

• ρ_{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

E_{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 (yr-Julian date)	Spring Overlay Order	Summer Overlay Order	Fall Overlay Order
37/32	2001-128	2000-158	1999-283	2	1	1
37/33	2000-126	2000-158	1999-283	3	1	1
36/32	2000-103	2000-247	1999-292	1	3	4
36/33	2000-103	2000-167	1999-307	1	2	2
35/33	2000-144	2000-256	1999-285	4	4	3

Three coverages (one for each season) showing overlap arrangement and date and scene attributed in the SOURCE field can be found:

```
/ut/archive/UT4/mosaic/spscn.img  
/ut/archive/UT4/mosaic/suscscn.img  
/ut/archive/UT4/mosaic/fascscn.img
```

Six band ETM mosaics for each season can be found at:

```
/ut/archive/UT4/mosaics/
```

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 /ut/archive/UT4/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: /ut/archive/UT4/img_files/.

All imagery derived predictor layers can be found at:

```
/ut/archive/UT4/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.

Aspect: A nine class aspect 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 (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:

/ut/archive/UT4/img_files/

2) Samples:

a) Sample collection methods:

The majority of samples were collected on the ground as polygons delineated over imagery in the field by USU field crews, but others come from other 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 method and/or source for each sample can be distinguished by the SITEID field in the polygon coverage:

UT070501LL02: Identifies a sample taken by USU RSGIS lab SWReGAP field crew. Indicates it was the second site taken on May 7, 2002 by Lisa Langs.

DEAD-02-0018 (e.g. <quadname>-<date>-<site>): Indicates sample from Utah Department of Natural Resources, Division of Wildlife. East Manti Vegetation Coverage Project(2003).

Doq-ut4_1, doq_ut4add_1 and Ut4doq040704_1: Indicates sites obtained from interpretation of black and white Digital Orthophoto Quads available for the state of Utah from AGRC (date of DOQs is either 1993 or 1997). DOQ interpretation by RSGIS lab personnel.

UT4_colo: Obtained from Colorado Basin Wide database.

b) Summary of samples:

2,379 samples were available to model this mapping area of which 2,358 samples were used in the final model. Qualifiers for a discarded sample site was sample quality, Ecological System (ES) mapability and sample abundance. Sample quality can be defined as the likelihood that the sample has a correct geospatial location and that the site sample characteristics correspond with site species data. ES mapability can be defined as the sites likelihood to map accurately given the map model conditions. Landuse types, in comparison to landcover types, typically are seen as sites that were discarded for ES mapability problems (example: Recently Burned, Recently Mined or Quarried and Developed areas). Sample abundance can be defined as all Ecological System cover types with roughly less then 20 samples sites collected in the field. Polygon coverage containing all samples is located at: ut/archive/UT4/train data/.

VALUE	CODE	ECOLOGICAL SYSTEMS DESCRIPTION	NUMBER OF SAMPLES
304	D04	INVASIVE SW RIPARIAN WOODLAND AND SHRUBLAND	20
308	D08	INVASIVE ANNUAL GRASSLAND	46
0	S000	IGNORED	20
6	S006	ROCKY MOUNTAIN CLIFF AND CANYON COMPLEX	45
10	S010	COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND	90

11	S011	INTER-MOUNTAIN BASINS SHALE BADLANDS	28
23	S023	ROCKY MOUNTAIN ASPEN FOREST AND WOODLAND	173
28	S028	ROCKY MOUNTAINS SUBALPINE DRY-MESIC SPRUCE-FIR FOREST AND WOODLAND	31
32	S032	ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND	132
39	S039	COLORADO PLATEAU PINYON-JUNIPER WOODLAND	244
42	S042	INTER-MOUNTAIN BASINS ASPEN-MIXED CONIFER FOREST AND WOODLAND	32
45	S045	INTER-MOUNTAIN BASINS MAT SALTBUUSH SHRUBLAND	22
46	S046	ROCKY MOUNTAINS GAMBEL OAK - MIXED MONTANE SHRUBLAND	186
52	S052	COLORADO PLATEAU PINYON-JUNIPER SHRUBLAND	203
54	S054	INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND	344
56	S056	COLORADO PLATEAU MIXED LOW SAGEBRUSH SHRUBLAND	88
65	S065	INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB	128
71	S071	INTER-MOUNTAIN BASINS MONTANE SAGEBRUSH STEPPE	212
79	S079	INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE	62
85	S085	SOUTHERN ROCKY MOUNTAINS MONTANE GRASSLAND	36
90	S090	INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND	27
93	S093	ROCKY MOUNTAINS LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND COMPLEX	66
96	S096	INTER-MOUNTAIN BASINS GREASEWOOD FLAT COMPLEX	101
102	S102	ROCKY MOUNTAINS ALPINE/MONTANE WET MEADOW	43
Total number of samples			2379

3) Cover types:

a) Classification Tree modeled cover types:

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

VALUE	CODE	ECOLOGICAL SYSTEMS DESCRIPTION	NUMBER OF SAMPLES
304	D04	INVASIVE SW RIPARIAN WOODLAND AND SHRUBLAND	20
308	D08	INVASIVE ANNUAL GRASSLAND	46
6	S006	ROCKY MOUNTAIN CLIFF AND CANYON COMPLEX	45
10	S010	COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND	90
11	S011	INTER-MOUNTAIN BASINS SHALE BADLANDS	28
23	S023	ROCKY MOUNTAIN ASPEN FOREST AND WOODLAND	173
28	S028	ROCKY MOUNTAINS SUBALPINE DRY-MESIC SPRUCE-FIR FOREST AND WOODLAND	31
32	S032	ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND	132
39	S039	COLORADO PLATEAU PINYON-JUNIPER WOODLAND	244
42	S042	INTER-MOUNTAIN BASINS ASPEN-MIXED CONIFER FOREST AND WOODLAND	32
45	S045	INTER-MOUNTAIN BASINS MAT SALTBUUSH SHRUBLAND	22
46	S046	ROCKY MOUNTAINS GAMBEL OAK - MIXED MONTANE SHRUBLAND	186
52	S052	COLORADO PLATEAU PINYON-JUNIPER SHRUBLAND	203

54	S054	INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND	344
56	S056	COLORADO PLATEAU MIXED LOW SAGEBRUSH SHRUBLAND	88
65	S065	INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB	128
71	S071	INTER-MOUNTAIN BASINS MONTANE SAGEBRUSH STEPPE	212
79	S079	INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE	62
85	S085	SOUTHERN ROCKY MOUNTAINS MONTANE GRASSLAND	36
90	S090	INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND	27
93	S093	ROCKY MOUNTAINS LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND COMPLEX	66
96	S096	INTER-MOUNTAIN BASINS GREASEWOOD FLAT COMPLEX	101
102	S102	ROCKY MOUNTAINS ALPINE/MONTANE WET MEADOW	43
Total number of samples			2358

b) Non CT modeled cover types:

Screen digitized over ETM imagery at a scale between 1:24,000 and 1:100,000: Updated existing **Land Use | Water Related Use** GIS coverage (screen digitizing over ETM imagery). Source of GIS coverage: Utah Department of Natural Resources, Division of Water Resources, collection period: 1985-2001. FGDC metadata for **Land Use | Water Related Use** coverage is found at: /ut/archive/UT4/non_model/

VALUE	CODE	Description Ecological System (Name)
211	N11	Open Water
221	N21	Developed, Open Space - Low Intensity
222	N22	Developed, Medium - High Intensity
281	N81	Agriculture, Pasture/Hay
282	N82	Agriculture, Cultivated Crops
302	D02	Recently Burned
303	D03	Recently Mined or Quarried
311	D11	Recently Chained Pinyon-Juniper Areas
314	D14	Disturbed, Oil well

Modeled with a post-classification model (see section 5c for details):

VALUE	CODE	Description Ecological System (Name)
30	S030	Rocky Mountain Subalpine Mesic Spruce-Fir Forest and Woodland
34	S034	Rocky Mountain Montane Mesic Mixed Conifer Forest and Woodland
47	S047	Rocky Mountain Lower Montane-Foothill Shrubland

4) Summary of predictor layers used:

a) Multi band predictors:

UT4_fall_2000.img (ETM bands 1-5 & 7 for fall 2000)
 UT4_spr_2000.img (ETM bands 1-5 & 7 for spring 2000)
 UT4_sum_2000.img (ETM bands 1-5 & 7 for summer 2000)

All multi-band predictors can be found at:
 /ut/archive/UT4/mosaics/

b) Single band predictors:

asp_9cls.img Categorical 9 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 (from DEM)

All single-band predictors can be found at:
/ut/archive/UT4/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. 20 random points were placed within each polygons using an Arcview Avenue script. The points were converted to pixels while ensuring that the resulting pixels (the new grid) aligned with the raster predictor layers. The resulting sub-sampled pixels would often be less than 20 per sample polygon, if random points fell within the same pixel. This new grid was converted to an Imagine IMG file (e.g.m4all20ppp.img) and is available at: UT/ARCHIVE/UT2/TRAIN_DATA/

Withholding Validation Sites: 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: ut/archive/UT4/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 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:

Introduction: This post-classification model was designed to be used as a tool to assist in differentiating between ecological systems that have similar characteristics. Similar characteristics can be defined as one or all of the following: Shared physiognomic structure, shared species composition and/or shared ecological process. This model has also been designed to correct for poor TM image quality. Clouds, fire smoke or dramatic scene lines are example of such quality issues.

The .gmd file used for this post-classification model is:

```
/archive/UT4/post_model/swgap_UT4.gmd
```

Step 1: Adjusting scene line output error

Modeling Methods: Correcting sceneline error. A conditional statement was written for the purpose of correcting misclassified landcover. The presence of scene lines can be attributed seasonal differences in Landsat imagery and is a relic of the mosaic processing. Errors were identified by visually inspecting the mosaiced Landsat imagery and CT output map model image. The details of the conditional statement follow:

```
1A EITHER 46 IF ($n2_m6all1_nosum_rec ==46 AND $n1_m6all1_rec==71  
AND$n4_scenemask==1) OR $n1_m6all1_rec OTHERWISE
```

Where:

46 *Rocky Mountain Gambel Oak-Mixed Montane Shrubland*

71 *Intermountain Basins Montane Sagebrush Steppe*

scenemask Mask image indicating where sceneline error occurred
1= area to be adjusted
0= all other area

m6all1_rec UT4 standard CT output image. Recoded to standard
integer code.

1B EITHER 23 IF (\$n3_memory==71 AND \$n2_m6all1_nosum_rec==23 AND \$n4_scenemask==1) OR \$n3_memory OTHERWISE

Where:

23 *S023 Rocky Mountain Aspen Forest and Woodland*

71 *S071 Intermountain Basins Montane Sagebrush Steppe*

Scenemask Mask image indicating where sceneline error occurred
1= area to be adjusted
0= all other area

Memory UT4 standard CT output image. Recoded to standard integer code. Note: This output is a temporary memory file that has been modified by previous steps in the post classification model.

Step 2: Discriminating Mesic Conifer from Non-Mesic Conifer Forests Types

The National Vegetation Classification Standard (NVCS) legend used for the Southwest Regional Gap Analysis Project divided Rocky Mountain conifer forest types into two distinct moisture classes; Mesic Forest types and Dry-Mesic Forest types (Comer et. al 2003). Early indications revealed that these types would be difficult to accurately discriminate using standard CT modeling methods due to site composition and spectral similarities. As a result, these types have been post modeled using their Ecological System conceptual description as a basis for modeling.

Step 2A: Discriminating Between Rocky Mountain Montane Dry-Mesic and Rocky Mountain Montane Mesic Conifer Types

Model methods: The original CT output modeled *S032 Rocky Mountain Montane Dry-Mesic Mixed Conifer Forest and Woodland* as the only montane conifer type. This post model was designed to discriminating between *S032* and, its moister ecologic counterpart, *S034 Rocky Mountain Montane Mesic Mixed Conifer Forest and Woodland*. The logic and parameters for the model were as follows:

This post-classification model was designed to extract *S034 Rocky Mountain Montane Mesic Mixed Conifer Forest and Woodland* from the standard CT modeled conifer system, *S032 Rocky Mountain Montane Dry-Mesic Mixed Conifer Forest and Woodland*.

NatureServe's NVCS concept describes *S032* montane mesic 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." (Comer et. al 2003)

Using the output of step 1, a conditional statement was created to extract *S034 Rocky Mountain Montane Mesic Mixed Conifer Forest and Woodland*. The conditional statement was created by using the aspect and landform characteristics described above. Affected pixels were then recoded to represent montane mesic mixed conifer system *S034*.

The details of the conditional statement follow:

```
EITHER 34 IF (($n14_landf == 2 OR $n14_landf == 5 OR
$n14_landf==6 OR $n14_landf == 9) AND ($n15_asp9cls == 1 OR
$n15_asp9cls == 2 OR $n15_asp9cls == 0) AND ($n15_memory ==
32)) OR $n5_memory OTHERWISE
```

Where:

34 *S034 Rocky Mountain Montane Mesic Mixed
Conifer Forest and Woodland.*

32 *S032 Rocky Mountain Montane Dry-Mesic
Mixed Conifer Forest and Woodland*

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

asp9cls Nine class aspect image
 1=North facing slope
 2=Northeast facing slope
 0=slope less than 3 degrees

memory UT4 standard CT output image. Recoded to
standard integer code. Note: This output is a temporary
memory file that has been modified by previous steps in the
post classification model.

Step 2B: Discriminating Between Rocky Mountain Subalpine Dry-Mesic
Spruce-Fir and Rocky Mountain Subalpine Mesic Spruce-Fir Types.

Using the output of step 2b, *S030 Rocky Mountain Subalpine Mesic
Spruce-Fir Forest and Woodland* was extracted from its Dry-Mesic
counterpart *S028* following similar parameters as step 2A.

NatureServe NVCS concept (Comer et. al 2003) describes the
subalpine mesic system 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 conditional statement was created to extract *S030
Mountain Subalpine Mesic Spruce-Fir Forest and Woodland* using the
above aspect and landform characteristics in order to identify
this subalpine mesic conifer system. The details of the
conditional statement follow:

```
EITHER 30 IF (($n14_landf== 2 OR $n14_landf == 5 OR
$n14_landf==6 OR $n14_landf== 9) AND ($n15_asp9cls== 1 OR
$n15_asp9cls == 2 OR $n15_asp9cls == 0) AND ($n15_memory ==
28)) OR $n16_memory OTHERWISE
```

Where:

30 *S030 Mountain Subalpine Mesic Spruce-Fir Forest and Woodland*

28 *S028 Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland*

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

aspect_9cls Nine class aspect image
 1=North facing slope
 2=Northeast facing slope
 0=slope less than 3 degrees

memory UT4 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 3: Mapping Foothill Shrubland in Buffer

EITHER 47 IF ((\$n17_memory == 71 AND \$n33_landf == 6 AND \$n41_s047 == 1)
 OR (\$n17_memory == 46 AND \$n33_landf == 6 AND \$n41_s047 == 1)) OR
 \$n17_memory OTHERWISE

Where:

47 *S047 Rocky Mountain Lower Montane-Foothill Shrubland*

71 *S071 Inter-Mountain Basins Montane Sagebrush Steppe*

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

S047 Mask image indicating where S047 will be mapped
 1= area to be adjusted
 0= all other area

Memory UT4 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: Removing Cloud Error

A conditional statement was written for the purpose of correcting misclassified landcover. The presence of seasonal clouds, captured on the Landsat 7 ETM+ imagery, can be attributed for the error. Error was identified by visually inspecting Landsat imagery and CT model output images. The details of the conditional statement follow:

EITHER \$n43_m6all1_nosum_rec IF (\$n35_memory > 0 AND
 \$n46_cld_mask == 1) OR \$n35_memory OTHERWISE

Where:
m6all1_nosum_rec CT output that contains no summer imagery
or related summer variables.

memory UT4 standard CT output image. Recoded to
standard integer code. Note: This output
is a temporary memory file that has been
modified by previous steps in the post
classification model.

c) Generalizing to MMU and map completion:

Once the CT model and the post-classification steps were employed, the map was generalized using the Clump tool in ERDAS Imagine 8.6. The parameter of 4 connected neighbors (or "rook's move") was used in the clumping process. Isolated pixels that fell under the specified 1-acre minimum mapping unit (MMU) were removed using the Eliminate tool.

Following the Clump & Eliminate step, the non-CT modeled classes were then "burned in" to the final map using the Overlay function. The final map can be found at: /ut/archive/UT4/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: /ut/archive/UT4/validation/.

b) 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 **the final map product**, which can be found at /ut/archive/UT4/final_map/.

c) 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.

N11 OPEN WATER:

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

N21 DEVELOPED, OPEN SPACE-LOW INTENSITY:

Not quantitatively assessed. Qualitative assessment indicates that most low density developed areas contemporary with the date

of the imagery are included. Some confusion may occur with the N22 DEVELOPED, MEDIUM-HIGH INTENSITY class.

N22 DEVELOPED, MEDIUM-HIGH INTENSITY:

Not quantitatively assessed. Qualitative assessment indicates that most low density developed areas contemporary with the date of the imagery are included. Some confusion may occur with the N21 DEVELOPED, OPEN SPACE-LOW INTENSITY class.

N81 PASTURE/HAY:

Not quantitatively assessed. Qualitative assessment indicates that most pasture/hay agricultural areas contemporary with the date of the imagery are included. This class included alfalfa and other non-irrigated agricultural types, which may be confused with N82 CULTIVATED CROPS.

N82 CULTIVATED CROPS:

Not quantitatively assessed. Qualitative assessment indicates that most cultivated agricultural areas contemporary with the date of the imagery are included. This class may be confused with N81 PASTURE/HAY, where crops were harvested prior to the date of the imagery or where crops were out of rotation for the season producing spectral signatures similar to that of N81.

D02 RECENTLY BURNED:

Not quantitatively assessed. Qualitative assessment indicates that most recently burned areas contemporary with the date of the imagery are included. This class may be confused with areas that have had fires in the past, but where the site has become somewhat revegetated since the time of the burn.

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. Some confusion may occur with the N21 DEVELOPED, OPEN SPACE-LOW INTENSITY class.

D04 INVASIVE SOUTHWEST RIPARIAN WOODLAND AND SHRUBLAND:

Quantitatively assessed, validation 0% (producers) and 0% (users) based on 4 independent validation samples. Confusion is evenly divided with S093 ROCKY MOUNTAINS LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND COMPLEX (2 of 4 samples) and S096 INTER-MOUNTAIN BASINS GREASEWOOD FLAT COMPLEX (2 of 4 samples). Confusion can be explained by similarities between these two cover types and by low training data sampling numbers. Where training data was available D04 mapped very well.

D08 INVASIVE ANNUAL GRASSLAND:

Quantitatively assessed, validation 20% (producers) and 25% (users) based on 10 independent validation samples. Confusion is primarily with S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND (4 of 10 samples). Confusion can be partially explained by the common presence of Big Basin and Wyoming sagebrush intermixing with cheatgrass. Where training data was available D08 mapped very well.

D11 RECENTLY CHAINED PINYON-JUNIPER AREAS:

Not quantitatively assessed. Qualitative assessment indicates that most recently chained pinyon-juniper areas contemporary with the date of the imagery are included. This class may be confused with areas that have been chained in the past, but where the site has become somewhat revegetated since the time of the chaining.

D14 DISTURBED - OIL WELL:

Not quantitatively assessed. Qualitative assessment indicates that most disturbed areas due to oil well development contemporary with the date of the imagery are included. There are areas known to have oil well development but were not mapped due to minimum mapping unit constraints. Some confusion may occur with the N21 DEVELOPED, OPEN SPACE-LOW INTENSITY class.

S006 ROCKY MOUNTAIN CLIFF AND CANYON COMPLEX:

Quantitatively assessed, validation 67% (producers) and 67% (users) based on 9 independent validation samples. S006 mapped very well.

S010 COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND:

Quantitatively assessed, validation 78% (producers) and 58% (users) based on 18 independent validation samples. Confusion is primarily with S006 ROCKY MOUNTAIN CLIFF AND CANYON COMPLEX (2 of 18 samples) and S052 COLORADO PLATEAU PINYON-JUNIPER SHRUBLAND (2 of 18 samples). Confusion can be explained by the spectral and substrate similarities shared by the systems. S010 is mapping very well as expected.

S011 INTER-MOUNTAIN BASINS SHALE BADLANDS: Quantitatively assessed, validation 33% (producers) and 67% (users) based on 6 independent validation samples. Confusion is primarily with S010 COLORADO PLATEAU MIXED BEDROCK CANYON AND TABLELAND (2 of 6 samples) and S065 INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB (2 of 6 samples). Confusion can be explained partially by similarities in species composition and substrate.

S023 ROCKY MOUNTAIN ASPEN FOREST AND WOODLAND: Quantitatively assessed, validation 100% (producers) and 90% (users) based on 35 independent validation samples. Users error is related to spectral similarity among morphologically similar cover types. S023 mapped very well.

S028 ROCKY MOUNTAIN SUBALPINE DRY-MESIC SPRUCE-FIR FOREST AND WOODLAND: Quantitatively assessed, validation 57% (producers) and 57% (users) based on 7 independent validation samples. Confusion is primarily with S032 ROCKY MOUNTAINS MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND (3 of 7 samples). Confusion can be explained by morphological similarities among these two cover types.

S030 ROCKY MOUNTAIN SUBALPINE MESIC SPRUCE-FIR FOREST AND WOODLAND: Not quantitatively assessed. This cover type was attained using a post processing modeler. Please reference methods section entitled *Post-classification and recoding, Step 2a*, for a detailed explanation. Post-classification mapped S030 very well.

S032 ROCKY MOUNTAIN MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND: Quantitatively assessed, validation 70% (producers) and 68% (users) based on 27 independent validation samples. Confusion is primarily with S039 COLORADO PLATEAU PINYON-JUNIPER WOODLAND (4 of 27 samples) and S028 ROCKY MOUNTAINS SUBALPINE DRY-MESIC SPRUCE-FIR FOREST AND WOODLAND (2 of 27 samples). Confusion can be explained by similarities in ecological process among these three cover types.

S034 ROCKY MOUNTAIN MONTANE MESIC MIXED CONIFER FOREST AND WOODLAND: Not quantitatively assessed. This cover type was attained using a post processing modeler. Please reference methods section entitled *Post-classification and recoding, Step 2b*, for a detailed explanation. Post-classification mapped S034 very well.

S039 COLORADO PLATEAU PINYON-JUNIPER WOODLAND: Quantitatively assessed, validation 80% (producers) and 70% (users) based on 49 independent validation samples. This widespread cover type is most commonly confused with S052 COLORADO PLATEAU PINYON-JUNIPER SHRUBLAND (6 of 49 samples). Confusion can be explained by similarities in species composition among these cover types. Overall, this upper pinyon-juniper belt woodland is mapping very well.

S042 INTER-MOUNTAIN WEST ASPEN-MIXED CONIFER FOREST AND WOODLAND COMPLEX: Quantitatively assessed, validation 57% (producers) and 100% (users) based on 7 independent validation samples. Confusion is distributed among S023 ROCKY MOUNTAIN ASPEN FOREST AND WOODLAND (1 of 7 samples), S032 ROCKY MOUNTAIN MONTANE DRY-MESIC MIXED CONIFER FOREST AND WOODLAND (1 of 7 samples), and S028 ROCKY MOUNTAINS SUBALPINE DRY-MESIC SPRUCE-FIR FOREST AND WOODLAND (1 of 7 samples). Confusion can be explained by similarities in species composition among these three cover types. Where training data was available S042 mapped very well.

S045 INTER-MOUNTAIN BASINS MAT SALTBUUSH SHRUBLAND: Quantitatively assessed, validation 60% (producers) and 100% (users) based on 5 independent validation samples. Confusion is distributed among S011 INTER-MOUNTAIN BASINS SHALE BADLANDS (1 of 5 samples) and S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND (1 of 5 samples). Overall, S045 is mapping well.

S046 ROCKY MOUNTAINS GAMBEL OAK - MIXED MONTANE SHRUBLAND: Quantitatively assessed, validation 76% (producers) and 81% (users) based on 38 independent validation samples. Confusion is distributed among several systems. Confusion can be partially explained by the presence of *Quercus gambelii* in the understory of a variety of systems and the patchwork growing habits which allows for S046 to intermingle with other systems.

S047 ROCKY MOUNTAINS LOWER MONTANE-FOOTHILL SHRUBLAND: Not quantitatively assessed. This cover type was attained using a post processing modeler. Please reference methods section entitled *Post-classification and recoding, Step 3*, for a detailed explanation. Post-classification mapped S047 very well.

S052 COLORADO PLATEAU PINYON-JUNIPER SHRUBLAND: Quantitatively assessed, validation 73% (producers) and 58% (users) based on 41 independent validation samples. Confusion is primarily with S039 COLORADO PLATEAU PINYON-JUNIPER WOODLAND (6 of 41 samples) Confusion with S039 can be explained by specie composition similarities between the cover types. Overall, this lower pinyon-juniper belt shrubland is mapping well.

S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND: Quantitatively assessed, validation 65% (producers) and 56% (users) based on 69 independent validation samples. Confusion is primarily with S052 COLORADO PLATEAU PINYON-JUNIPER SHRUBLAND (7 of 69 samples). Confusion can be partially explained by the physiognomic similarities among the two systems. Otherwise, S054 is mapping well.

S056 COLORADO PLATEAU MIXED LOW SAGEBRUSH SHRUBLAND: Quantitatively assessed, validation 44% (producers) and 53% (users) based on 18 independent validation samples. Confusion is primarily with S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND (6 of 18 samples). These two systems are commonly found adjacent

to one another, and share codominating species. Where training data was available S056 mapped very well.

S065 INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB:

Quantitatively assessed, validation 54% (producers) and 56% (users) based on 26 independent validation samples. Confusion is distributed among several desert scrub communities. Confusion can be partially attributed to the diversity of habitats in which *Atriplex confertifolia* can reside. Where training data was available S065 mapped very well.

S071 INTER-MOUNTAIN BASINS MONTANE SAGEBRUSH STEPPE:

Quantitatively assessed, validation 84% (producers) and 78% (users) based on 43 independent validation samples. Confusion is predominantly with S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND (7 of 43 samples). Confusion can be partially explained by similarities in species and/or could likely be attributed to sub-specie misidentification.

S079 INTER-MOUNTAIN BASINS SEMI-DESERT SHRUB STEPPE:

Quantitatively assessed, validation 31% (producers) and 57% (users) based on 13 independent validation samples. Confusion is primarily with S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND (4 of 13 samples), and S065 INTER-MOUNTAIN BASINS MIXED SALT DESERT SCRUB (2 of 13 samples). Confusion can be partially explained by the variability of species composition within S079. Where training data was available S079 mapped very well.

S085 SOUTHERN ROCKY MOUNTAINS MONTANE GRASSLAND: Quantitatively assessed, validation 0% (producers) and 0% (users) based on 8 independent validation samples. Confusion is primarily with S071 INTER-MOUNTAIN BASINS MONTANE SAGEBRUSH STEPPE (3 of 8 samples) and S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND (4 of 8 samples). It is unclear why these confusions occurred. Where training data was available S085 mapped very well.

S090 INTER-MOUNTAIN BASINS SEMI-DESERT GRASSLAND: Quantitatively assessed, validation 33% (producers) and 100% (users) based on 6 independent validation samples. Confusion is with a verity of systems. It is unclear why this system is mapping poorly. Where training data was available S090 mapped very well.

S093 ROCKY MOUNTAINS LOWER MONTANE RIPARIAN WOODLAND AND SHRUBLAND COMPLEX:

Quantitatively assessed, validation 64% (producers) and 100% (users) based on 14 independent validation samples. Confusion is primarily with S023 ROCKY MOUNTAIN ASPEN FOREST AND WOODLAND (2 of 14 samples). Confusion can be partially explained by morphologic similarities between the two cover types. Where training data was available S093 mapped very well.

S096 INTER-MOUNTAIN BASINS GREASEWOOD FLAT COMPLEX:

Quantitatively assessed, validation 38% (producers) and 57% (users) based on 21 independent validation samples. Confusion is primarily with S052 COLORADO PLATEAU PINYON-JUNIPER SHRUBLAND (3 of 27 samples) and S054 INTER-MOUNTAIN BASINS BIG SAGEBRUSH SHRUBLAND (3 of 27 samples). It is unclear why there is confusion with S052. It is possible that the physiognomic characteristics are to blame. The confusion with S054 can be partially explained by the sometimes dense presence of *Atrémisia tridentata* in greasewood stands. Where training data was available S096 mapped very well.

S102 ROCKY MOUNTAINS ALPINE/MONTANE WET MEADOW:

Quantitatively assessed, validation 44% (producers) and 80% (users) based on 9 independent validation samples. Confusion is primarily with S046

ROCKY MOUNTAINS GAMBEL OAK - MIXED MONTANE SHRUBLAND (4 of 9 samples). Morphologic similarities can explain this confusion. Where training data was available S126 mapped very well.

7) Citations:

Chaves, 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. Menard, M. Pyne, M. Reid, 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, 2001a, t-sensor reflectance: A first order normalization of Landsat 7 ETM+ Images.

Huang, C., B. Wylie, L. Yang, C. Homer and G. Zylstra, 2001b, Derivation of tasseled cap transformation based on Landsat 7 at-satellite reflectance.

Rulequest, 2004. See5: An informal tutorial. WWW URL <http://www.rulequest.com/see5-win.html> (23 August, 2004).