

Owyhee Upland Mapping Zone Documentation: Mapping Methods and Accuracy

1) Predictor Layer Preparation:

a) *Image Standardization:*

Standardization from DN values to at-sensor reflectance was performed by EROS Data Center as part of the MRLC Preprocessing Procedure. This procedure uses the method presented by Huang et. al (2001a) to transform Landsat 7 ETM+ imagery (Note, Landsat 5 used for some dates). The equation used for reflectance was as follows:

$$r_{BandN} = (\Pi((DN_{Band} * Gain_{Band} + Bias_{BandN}) * D2) / (E_{BandN} (E_{BandN} (SINq))))$$

r_{BandN} = Reflectance for Band N

DN = Digital Number for Band N

D = Normalized Earth-Sun Distance

E_{BandN} = Solar Irradiance for Band N

q = Solar Elevation

Gain $_{BandN}$ = Provided within header file, and band specific

Bias $_{BandN}$ = Provided within header file, and band specific

b) *Image Dates and Mosaic:*

Images were mosaicked using ERDAS Imagine 8.7 Mosaic Tool with "no outline" for *type*, and the "Overlay" option for *overlap function*.

Image dates and scenes were as follows:

Path/Row	Spring Sat	Spring Julian Date	Spring Calender	Spring Overlay Order	Summer Sat	Summer Julian Date	Summer Calender	Summer Overlay Order	Fall Sat	Fall Julian Date	Fall Calender	Fall Overlay Order
41030	7	01124	5/4/2001	3	7	41030	7/10/2002	3	7	41030	10/8/2000	3
41031	7	00154	6/2/2000	4	7	41031	7/10/2002	4	7	41031	9/4/1999	4
42031	7	01131	5/11/2001	7	7	42031	7/27/2000	7	7	42030	9/11/1999	7
42030	7	02134	5/14/2002	6	7	42030	7/1/2002	6	7	42031	9/27/1999	6
39031	7	00092	4/1/2000	5	7	39031	7/22/2000	5	7	39031	10/8/1999	5
40030	7	01149	5/29/2001	2	7	40030	8/14/2000	2	7	40030	9/13/1999	2
40031	7	00099	4/8/2000	1	7	40031	7/11/1999	1	7	40031	10/17/2000	1

c) *Image Derived Datasets:*

Normalized Difference Vegetation Index (NDVI): This dataset was provided by the EROS Data Center as part of the MRLC processing output.

Tasseled cap: Brightness, Greenness & Wetness band transformations were provided by the EROS Data Center as part of the MRLC processing output. This transformation followed the methods of Huang et. al 2001b.

Shrub Cover: Overall percent of shrub cover was described following Jennings et. al 2004. Overlap in the top three occurring shrubs strata where addressed by the following:

$$C_i = \left(1 - \prod_{j=1}^n (1 - \% \text{cov } j / 100) \right) * 100$$

where C_i is the percent cover of stratum i for species or growth form j in stratum i .

Fractional Vegetation: The percent of ground covered by photosynthetic vegetation was estimated by the equation of Carlson and Ripley (1997). Reference values used in the equation were identified by examination of NDVI histograms and locating known sites of bare soil and irrigated agricultural fields. The equation is $N^* = (\text{SQRT}((\text{NDVI} - \text{NDVI}_{\text{min}})/(\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}}))) * 100$, where NDVI_{min} is the NDVI value for bare soil pixels in the scene, and NDVI_{max} is the NDVI value for fully vegetated pixels in the scene. Fractional vegetation is related to calculations of impervious surface (Imp. Surface = 100 - fractional vegetation). The output is an integer layer with values between 0 and 100.

$\text{ou_summer_fv} = (\text{SQRT}((\text{sawt_ndvi_summer} - 0.07) / (0.9 - 0.07))) * 100$

Image Texture: The texture of the image is

d) DEM Derived Datasets:

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

Slope: A slope layer was created using the ARC/INFO SLOPE command. Values represent slope in degrees.

Topographic Relative Moisture Index: A TRMI grid (values ranging from 0-28) was created using an Arc/INFO AML obtained from the Southwest Regional GAP Project created by G. Manis (Manis et. al 2001). The TRMI model is based on the methods defined by Haplin, P. N. 1999, and Parker, A. J. 1982.

Landform: A 10 class landform grid was created using an Arc/INFO AML obtained from the Southwest Regional GAP Project created by G. Manis (Manis et. al 2001). This product was derived from the topographic relative moisture index.

For modeling purposes all arcinfo grids were converted to ERDAS Imagine .img files.

2) Samples:

a) Sample Collection Methods:

Samples were collected in a variety of ways. Originally, it was thought that most, if not all, of the sampling would be derived from field collected information polygons

delineated over imagery in the field by field crews. Classification trees, however, require substantial amounts of training data so that additional information had to be acquired. All samples were assigned a label corresponding to an Ecological System (Comer et. al 2003). On the ground data was collected as polygons delineated over imagery in the field by ORNHIC and USGS Snake River Field Station, and SW-ReGap field crews . A listing of the number of ground points by source is provided below. On screen digitizing was done using ETM imagery as a backdrop.

Source Sites

Map Region	Shrub Sample (2003)	Nevada ReGap	Bureau of Land Mangement			US Forest Service								Screen Digitized
	Training		Burns	Lakeview	Vale	Freemont NF	Gifford Pinchot NF	Malheur NF	Mt Hood NF	Ochoco NF	Umatilla NF	Wallowa NF	Winema NF	
Owyhee Uplands	204	847												184

3) Cover Types:

a) Classification Tree Modeled Cover Types:

Thirty-nine cover types were modeled in this zone. All forest and barren types were modeled and later reclassified into single classes in the overall Tree modeling.

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

Ecological System	Total	Code
Agriculture-Irrigated and Non-irrigated	32	280
Breat Basin Semi-Desert Chaparral	10	53
Breat Basin Xeric Mixed Sagebrush Shrubland	77	55
Colorado Plateau Pinyon-Juniper Woodland	1	39
Columbia Plateau Low Sagebrush Steppe	6	154
Columbia Plateau Western Juniper Woodland	5	41
Columbia (<i>Great</i>) Basin Foothill and Lower Montane Riparian Woodland and Shrubland	62	137
Great Basin Pinyon-Juniper Woodlands	1	40
Rocky Mountain (<i>Inter-Mountain Basin</i>) Aspen - Mixed Conifer Forest and Woodland	3	23
Inter-Mountain Basin Big Sagebrush Shrublands	137	89
Inter-Mountain Basin Big Sagebrush Steppe	362	78
Inter-Mountain Basin Cliff and Canyon	43	9
Inter-Mountain Basin Greasewood Flat	9	96
Inter-Mountain Basin Greasewood Wash	1	14
Inter-Mountain Basin Mixed Salt Desert Scrub	12	65
Inter-Mountain Basin Montane Sagebrush Steppe	133	71
Inter-Mountain Basin Mountain Mahogany Woodland and Shrubland	34	50
Inter-Mountain Basin Playa	7	15
Inter-Mountain Basin Semi-Desert Grassland	82	90
Inter-Mountain Basin Semi-Desert Shrub Steppe	7	79
Inter-Mountain Basin Subalpine Limber-Bristlecone Pine Woodland	2	242
Invasive Annual Grassland	14	308
Invasive Annual and Biennial Forbland	4	307

Invasive Perennial Grassland	2	306
Invasive Southwest Riparian Woodland and Shrubland	1	304
North American Arid West Emergent Marsh	5	100
Open Water	9	211
Pasture/Hay, Irrigated Agriculture	15	281
Recently Burned	1	302
Rocky Mountain Alpine Bedrock and Scree	5	2
Rocky Mountain Alpine-Montane Wet Meadow	1	102
Rocky Mountain Aspen Forest and Woodland	63	23
Rocky Mountain Dry Tundra	2	81
Rocky Mountain Subalpine - Montane Riparian Shrubland	1	151
Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	53	242
Rocky Mountain Subalpine Mesic Meadow	11	102
Rocky Mountain Subalpine Mesic Spruce-Fir Forest and Woodland	21	242
Rocky Mountain Subalpine-Montane Limber-Bristlecone Pine Woodland	1	242
Total	1235	

b) Non Tree modeled Cover Types:

Screen digitized over ETM:

Recently Burned (302) areas were developed from Fires in Western North America by visually comparing the shapefile to the ETM imagery and selecting those fire scars visible on the imagery and those fires that occurred after the image date. Source of shapefile: USGS Snake River Field Station, 2004, <http://sagemap.wr.usgs.gov>

Interstates and highways were buffered by 30m to develop a transportation mask. Sources: Idaho Department of Lands Interstates and Highways and Oregon BLM Highways.

c) Cover Types Modified with a Post-Classification model:

(see section 5c for details)

Ecological System	Action
Breat Basin Xeric Mixed Sagebrush Shrubland	Shrub Cover applied to split for steppe vs shrubland
Colorado Plateau Pinyon-Juniper Woodland	Limited to Cliff and Canyon terrain region
Columbia Plateau Western Juniper Woodland	Split along ecoregion boundaries
Great Basin Pinyon-Juniper Woodlands	Split along ecoregion boundaries
Inter-Mountain Basin Big Sagebrush Shrublands	Shrub Cover applied to split for steppe vs shrubland
Inter-Mountain Basin Big Sagebrush Steppe	Shrub Cover applied to split for steppe vs shrubland
Inter-Mountain Basin Cliff and Canyon	Terrain model to identify
Inter-Mountain Basin Montane Sagebrush Steppe	Shrub Cover applied to split for steppe vs shrubland
Inter-Mountain Basin Semi-Desert Grassland	Shrub Cover applied to convert category 1 shrub cover to grassland
Inter-Mountain Basin Semi-Desert Shrub Steppe	Shrub Cover applied to split for steppe vs shrubland
Conifer Forest	All classified forest type converted to Conifer Forest (242)
Recently Burned	Updated from LandFire boundaries

4) Summary of Predictor Layers Used:

a) Multi band predictors:

ETM bands 1-5 & 7 for fall

ETM bands 1-5 & 7 for spring

ETM bands 1-5 & 7 for summer

b) Single Band Predictors:

Fall Tasseled Cap brightness band
Fall Tasseled Cap greenness band
Fall Tasseled Cap wetness band
Summer Tasseled Cap brightness band
Summer Tasseled Cap greenness band
Summer Tasseled Cap wetness band
Spring Tasseled Cap brightness band
Spring Tasseled Cap greenness band
Spring Tasseled Cap wetness band

Tree model Shrub Cover Percentage (integer) **

Continuous (integer) slope
Continuous (integer) elevation
Categorical 10 class landform (from DEM)
Topographic Relative Moisture Index (from DEM)
** Post Classification

5) Modeling Methods:

a) See5 Classification Tree Modeling:

Sub-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. 30 random points were placed within each polygon using the Arcview tool Hawth's Tool. 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 30 per sample polygon, if random points fell within the same pixel.

Training and Validation Sites: Twenty percent of the all sample polygons were withheld for validation using the Feature Select extension in ArcView.

With the remaining 80%, 30 sub-samples were randomly generated within each sample polygon. This was done by first randomly generating points within each polygon using the Random Points extension in Hawth's Tool for ArcVGIS, later converted to *.img files. Individual points were converted to *.img files (each to be considered a separate observation for the See5 classifier) and 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 2.01) data mining software (Rulequest 2004) was used for generating classification trees. Boosting was employed using 20 trials.

The following briefly describes the files used by the See5 Program (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.

*.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 completed the final 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 an *.img file, which is the spatial application of the tree's rules was created.

b) Shrub Cover: Sampling protocols followed by project teams (section 2a) required the collection of a visual estimate of percent coverage of individual shrub strata. Following similar methodology used in trial regions of SW-ReGap (Huang et. al 2003, Jennings et. al 2004) a overall percent shrub cover was estimated for each training site (80%/20% training/validation). The total percent coverage is represented as a continuous surface at each site, and was reclassified to five categorical types following guidelines suggested by LandFire.

Category	Range %
Very High	> 45%
High	36-45%
Moderate	26-35%
Steppe	11-25%
Grassland	<10%

Tree model Validation: Twenty percent of the sample polygons were randomly selected and withheld from Tree modeling. The preliminary Tree models were run as described in section 5a using the remaining 80% of the training site data. The 20% withheld samples were used to assess the predictive capability of the Tree modeled map via the kappa_stats.avx extension for ArcView by intersecting the reference polygons with the Tree modeled land cover map. This extension considers the site correctly mapped when the majority of pixels within the reference polygon agree with the reference label. Output from kappa_stat.avx includes the kappa statistic and an error matrix indicating errors of omission and commission.

c) *Post Classification, Recoding and Other Modeling Steps*: Post-classification modeling was done to map classes where there were not enough training data to map using the full Tree model, to differentiate between ecological systems that have similar characteristics, use ancillary data sets, or correct problems found during qualitative review. The processes are listed in the order in which they were implemented.

Post-Classification Tree modeling

Shrub Coverage: Unique combinations of all shrub dominated ecological systems and the independent shrub cover were modified to represent a Shrub -> Steppe -> Grassland type based upon the ecoregion of occurrence.

Ecological System	Shrub Cover				
	Very High	High	Moderate	Steppe	Grassland
Breat Basin Xeric Mixed Sagebrush Shrubland	Shrubland	Shrubland	Shurbland	Steppe	Steppe
Inter-Mountain Basin Big Sagebrush Shrublands	Shrubland	Shrubland	Shurbland	Steppe	Steppe
Inter-Mountain Basin Big Sagebrush Steppe	Shrubland	Shrubland	Shurbland	Steppe	Grassland
Inter-Mountain Basin Montane Sagebrush Steppe	Shrubland	Shrubland	Shurbland	Steppe	Grassland
Inter-Mountain Basin Semi-Desert Grassland	Shrubland	Shrubland	Steppe	Steppe	Grassland
Inter-Mountain Basin Semi-Desert Shrub Steppe	Shrubland	Shrubland	Shurbland	Steppe	Grassland

Overmapping Corrections

A visual examination of the 100% Tree model was done using ETM imagery. Columbia Plateau Scabland Shrubland and Open Water were over-mapped throughout the mapping area.

Juniper Dominated Ecological System:

There are three distinct juniper dominated ecological systems with the map zone, which tree models are not sufficient in breaking the types apart. Unique types were assigned based upon the direction of Idaho Natural Heritage ecologists to specific regions of occurrence and all pixels representing a juniper system within that region were confirmed/reclassified. Using the cliff terrain model as a mask (value = 1) all juniper in the southern portion of the map zone were assigned to Colorado Plateau Pinyon-Juniper Woodland (39) if they occurred in a canyon landscape. The ecoregional boundary for the Columbia Plateau Ecoregion was used to delineate the additional two Juniper types, Columbia Plateau Western Juniper Woodland (41) is limited to the target ecoregion and Great Basin Pinyon-Juniper Woodlands (40) occur in the remaining portions of the mapzone.

Forest

Conifer Forest (242) is a composite type composed of all individual ecological systems present within the map zone.

Water

The Open Water cover type was developed using a Feature Extraction (Feature Analyst™) on the summer Landsat imagery. Each cluster was assigned a water or non-water class through visual inspection. The water classes were then combined using

conditional statements where featWater = 1. Through a visual examination of the output compared to ETM imagery any necessary corrections to the classes were done using a spatial recode. The result was applied to the final Tree model as a mask.

Recent Fires

The recent fires layer was applied as a mask over the landcover map superseding all underlying ecological systems with the Recently Burned (302) land cover type

Transportation

Following the lead of the SWGAP project, Interstates and Highways were applied to the final product. The transportation layer, recoded as Developed Medium (223), was applied as a mask over the landcover map using a conditional statement. See 3b for development

c) Generalizing to MMU and Map Completion:

Once the Tree model and the post-classification steps were employed, the map was generalized using the Clump tool in ERDAS Imagine 8.7. The parameter of 4 connected neighbors was used in the clumping process. Isolated pixels that fell under the specified 1-acre (5 pixels) minimum mapping unit (MMU) were removed using the Eliminate tool. Following the Clump & Eliminate step, the non-Tree modeled classes were then "burned in" to the final map using the Overlay function.

6) Accuracy Assessment

Accuracy statistics were calculated using the 20% withheld samples on the preliminary Tree model. These statistics include an error matrix, kappa statistics and breakdown of user's/producer's accuracy and error.

Error Matrix

ERROR MATRIX: Reference Data in Columns/Classification Data in Rows	Columbia Plateau Low Sagebrush Steppe	Columbia Plateau Scabland Shrubland	Exotics	Inter-Mountain Basins Montaine Sagebrush Steppe	Northern Rocky Mountain Lower Montaine Mesic Deciduous Shrubland	Inter-Mountain Basins Mixed Salt Desert Scrub	Inter-Mountain Basins Semi-Desert Shrub-Steppe	Inter-Mountain Basins Big Sagebrush Steppe - ssp. wyoming	Inter-Mountain Basins Big Sagebrush Shrubland ssp. wyoming	Inter-Mountain Basins Juniper Savanna	Total	User's Accuracy
Columbia Plateau Low Sagebrush Steppe	0	0	0	0	0	0	0	0	0	0	0	0%
Columbia Plateau Scabland Shrubland	0	2	0	0	0	0	0	0	0	0	2	100%
Exotics	0	0	4	0	0	0	0	0	0	0	4	100%
Inter-Mountain Basins Montaine Sagebrush Steppe	1	0	0	0	0	0	0	0	0	0	1	0%
Northern Rocky Mountain Lower Montaine Mesic Deciduous Shrubland	0	0	0	0	1	0	0	0	0	0	1	100%
Inter-Mountain Basins Mixed Salt Desert Scrub	0	0	0	0	0	1	0	0	0	0	1	100%
Inter-Mountain Basins Semi-Desert Shrub-Steppe	0	0	0	0	0	0	2	0	0	0	2	100%
Inter-Mountain Basins Big Sagebrush Steppe - ssp. wyoming	0	0	0	0	0	0	0	5	8	0	13	38%
Inter-Mountain Basins Big Sagebrush Shrubland - ssp. wyoming	0	0	0	0	0	0	0	2	12	0	14	86%
Inter-Mountain Basins Juniper Savanna	0	0	0	0	0	0	0	0	0	3	3	100%
Total	1	2	4	0	1	1	2	7	20	3	41	
Producer's Accuracy	0%	100%	100%	0%	100%	100%	100%	71%	60%	100%		73%

User's/Producer's Accuracy and Omission/Commission Error

Ecological Systems	Producer Accuracy	User Accuracy	Specificity	Predictive Power	Omission Error	Commission Error
Columbia Plateau Low Sagebrush Steppe	0	Null	1	0.97561	1	0
Columbia Plateau Scabland Shrubland	1	1	1	1	0	0
Exotics	1	1	1	1	0	0
Inter-Mountain Basins Montaine Sagebrush Steppe	Null	0	0.97561	1	Null	0.02439
Northern Rocky Mountain Lower Montaine Mesic Deciduous Shrubland	1	1	1	1	0	0
Inter-Mountain Basins Mixed Salt Desert Scrub	1	1	1	1	0	0
Inter-Mountain Basins Semi-Desert Shrub-Steppe	1	1	1	1	0	0
Inter-Mountain Basins Big Sagebrush Steppe - ssp. wyoming	0.7142857	0.3846154	0.764706	0.928571	0.2857143	0.235294
Inter-Mountain Basins Big Sagebrush Shrubland - ssp. wyoming	0.6	0.8571429	0.904762	0.703704	0.4	0.095238
Inter-Mountain Basins Juniper Savanna	1	1	1	1	0	0
Columbian Basin Palouse Prairie	Null	Null	1	1	Null	0

Overall Statistics:

Overall Accuracy: $(30 / 41) = 0.731707317$ (57% prior to Shrub Cover Adjustment)

Overall Misclassification Rate: $(11 / 41) = 0.268292683$

Overall Sensitivity: 0.731707317

Overall Specificity: 0.973170732

Overall Omission Error: 0.268292683

Overall Commission Error: 0.026829268 *Kappa Statistics*

Kappa Statistics

KHAT = 0.646275

VARIANCE = 0.00936722

Z = 6.677

P = 0.00000000

7) Citations:

Carlson, T. N. and D. A. Ripley. 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*. 62:241-252.

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.

Haplin, P. N. 1999. "GIS Analysis for Conservation Site Design: A short-course developed for the Nature Conservancy" Nicholas School of the Environment--Landscape Ecology Lab, Duke University.

Huang, C.L., J. Vogelmann, B. Tolk, and Z. Zhu. 2003. Preliminary Landfire Vegetation Products in the Wasatch Range-Uinta Mountain Area of Utah. (<http://ams.confex.com/ams/pdfpapers/65875.pdf>)

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

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

Michael Jennings M., D. Faber-Langendoen, R. Peet, O. Loucks, D. Glenn-Lewin, A. Damman, M. Barbour, R. Pfister, D. Grossman, D. Roberts, D. Tart, M. Walker, S. Talbot, J. Walker, G. Hartshorn, G. Waggoner, M. Abrams, A. Hill, M. Rejmanek, 2004. Guidelines for describing associations and alliances of the U.S. National Vegetation Classification, V4. The Ecological Society of America, Vegetation Classification Panel.

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

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

Parker, A. J. 1982. The Topographic Relative Moisture Index: An Approach to Soil-Moisture Assessment in Mountain Terrain. *Physical Geography*, 1982, 3,2, pp.160-168.