TAKING THE " " OUT OF "GROUND TRUTH": OBJECTIVE ACCURACY ASSESSMENT

Timothy B. Hill GIS/Remote Sensing Analyst
Geographic Resource Solutions, USA

ABSTRACT

Interest in remotely sensed data as a source for generating thematic maps has increased greatly with the expanded use of Geographic Information Systems (GIS). Landsat Thematic Mapper (TM) data and the California Department of Fish and Game's Wildlife Habitat Relationships vegetation classes were used as a means to classify the project area. Characteristics estimated include canopy closure, species type, average tree size, and canopy structure. The accuracy assessment incorporated a stratified random sampling scheme with a minimum of 30 samples in each of the 28 strata. Sample stands were visited on the ground to collect quantitative vegetation data used to compare with the habitat unit characteristics estimated using satellite image processing techniques. Conventional methods of accuracy assessment utilize photo interpretation and/or qualitative "ground truth" techniques. Studies have shown photo interpretation to be inaccurate when compared to quantitative data. The same can be said of qualitative estimates of vegetation (i.e. ocular estimates). Preliminary analyses have shown conventional methods of match determination between map and reference data were significantly lower when comparing data described by continuous estimates. New methodologies were developed to compare ground truth estimates with the mapped polygon estimates. Data estimates were continuous and didn't necessarily coincide with class boundaries used to generalize the habitat characteristics. The major concern is the determination of what constitutes a match of the characteristics being evaluated. Four different types of matches are defined. Error matrices were developed to demonstrate correspondence of characteristics as well as errors of omission and commission. Sampling methodology and the different error matrices representing the types of matches are presented, compared and discussed.

INTRODUCTION

The remote sensing community has been called upon by both public and private agencies to provide data layers for their respective GIS'. These data are being used to help find answers for complex, present and future, resource issues. Decisions based in part on information from these thematic maps can affect wildlife, ecosystems and people. Thousands of jobs can be eliminated, ecosystems needlessly destroyed, and species driven to the brink of extinction if these maps are inaccurate. Classification of remotely sensed data have used the error matrix for evaluating map accuracy (Congalton, Oderwald, and Mead, 1983; Rosenfield and Fitzpatrick-Lins, 1986). In the past, accuracy assessment has focused on pixel level accuracy of classified remotely sensed data compared to assumed correct reference data (Congalton, 1988). A pixel map resulting from classified imagery is not the end product of many remote sensing projects in the 1990's. Final maps of current remote sensing projects are polygon GIS databases in a vector format. The processes applied to a pixel map such as filtering, smoothing, and resampling can significantly alter the information between the original pixel map and the desired GIS database (Stumpf, 1993). Accuracy assessment at the polygon level has relied heavily on airphoto interpretation for reference data ("ground truth") because the cost of acquiring actual ground visits is expensive and time consuming (Congalton and Green, 1993). Results from polygon level accuracy assessment methods also use the error matrix as a means of expressing map accuracy. Congalton (1988) writes "The overriding assumption then in the entire accuracy assessment procedure is that the error matrix must be indicative or representative of the entire area mapped from the remotely sensed data." Taking this, and the subjective (biased) and inaccurate nature of reference data derived from airphoto interpretation (Biging, Congalton and Murphy 1991) into consideration, the need to base map accuracy on sound sampling techniques and ground truth instead of "ground truth" is evident.

Geographic Resource Solutions (GRS) was contracted by the California Department of Forestry and Fire Protection (CDF) Forest and Rangeland Resource Assessment Program (FRRAP) for a pilot mapping project of a 6 million acre portion of the Klamath Province in Northwestern California in May of 1990. The pilot's main goal was to evaluate the use of satellite data for mapping wildlife habitat over a large area. For this reason, image classification did not include any masking, modeling or ancillary data. The project also incorporated a rigorous accuracy assessment. Sampled stands were visited on the ground where quantitative data were collected. This information was summarized and used as reference data in the error matrices. The accuracy assessment was used as an iterative step in the mapping methodology, and providing feed back for the second phase of the project. Phase II of the project entails mapping the entire Klamath Province (18 million acres). GRS completed maps of wildlife habitat characteristics for the study area in May of 1992. In the resulting polygon GIS database, each polygon representing the vegetation had a distribution of percent cover by quadratic mean diameter (qmd) and species. Field data collection, used as reference data in the accuracy assessment, took place during the summer of 1992. Existing methodologies were used and new methodologies were developed to compare reference data and map data, both of which are described by continuous and discrete estimates. This paper presents the methodology of the accuracy assessment and compares the results.

METHODS

Classification Scheme

The California Department of Fish and Game's Wildlife Habitat Relationships (WHR) rules were adapted and used for the pilot. Table 1 shows the 28 vegetation strata.

WHR Tree Types:

- Subalpine Conifer (**SCN**)
- Ponderosa Pine (**PPN**)
- Lodgepole Pine (**LPN**)
- Redwood (**RDW**)
- Douglas-fir (**DFR**)
- Montane Hardwood-Conifer (**MHC**)
- Red Fir (**RFR**)
- Closed Cone Pine-Cypress (**CPC**)
- Mixed Conifer (**KMC**)
- White Fir (**WFR**)
- Juniper (**JUN**)
- Montane Hardwood (**MHW**)

Other WHR Cover Types:

- Herbaceous (**HRB**)
- Barren (**BAR**)
- Water (**WAT**)
- Shrub (**SHR**)
- Urban (**URB**)

WHR Canopy Closure Classes:

Class		Canopy Closure
Sparse	**S**	10 - 24%
Open	**P**	25 - 39%
Moderate	**M**	40 - 59%
Dense	**D**	>= 60%

WHR Size Classes:

Class		Average Tree Size
Sapling	**1**	0.0 - 5.5" qmd
Pole	**2**	5.6 - 10.5" qmd
Small Tree	**3**	10.6 - 23.5" qmd
Large Tree	**4**	23.6 - 35.5" qmd
Giant Tree	**5**	>= 35.6" qmd

WHR Canopy Structure Classes:

Class	Structure
E	Even
U	Multi-layered

Table 1: WHR Map Categories

Landsat Thematic Mapper (TM) data were specified by FRRAP as the data source to be used in the project. Portions of 4 TM scenes were required for the pilot ranging in dates from May of 1990 to July of 1990. Since GRS relies heavily on ground data in their methodology, field personnel quantitatively sampled vegetation (during 1991 and early 1992) throughout the project area for use as training data. GRS used a combination of supervised and unsupervised image classification techniques (Brown and Fox, 1992.) Pixel grids resulting from classification were aggregated into stands (polygons) of a minimum mapping unit (mmu) of 40 acres for forest types and 10 acres for non-forest types as required by the contract. The aggregation process uses an ecologically based set of rules that consider cover type, size and density of pixels and the similarity to surrounding group of pixels in the formation of polygons (Stumpf, 1993). The resulting raster polygon maps were vectorized and data were loaded into the relational database.

Sampling Scheme

Academia has produced varying opinions on sampling schemes used for accuracy assessment. Population research by Stehman (1992) found systematic sample designs more precise than simple random sampling. He also infers that systematic sampling uses sampling resources more efficiently than simple random sampling. Congalton (1988) concluded that simple random sampling compared to systematic sampling techniques, performed the best when comparing sample means to population means. He also reported "stratified random sampling also performed well and should be used especially when it is necessary to make sure that small, but important, areas are represented in the sample."

GRS used the stratified random sampling approach. This sampling scheme provides information on all the map categories and increases the overall accuracy of population estimates (Cochran, 1977 p 89). The first step was to develop a GIS theme of sample points. Points were randomly placed by a computer program throughout the study area. The resulting 7000 sample points were assigned a unique sequential number representing their order of placement. By overlaying the sample point theme with the polygon habitat maps in the GIS, WHR characteristics were determined for the sample points. The goal of the accuracy assessment (as specified by the contract) was to have a minimum of 30 samples in each of the 28 strata (Table 1) and 75 samples in the more heterogeneous types (MHC, KMC, PPN etc...). Sample points were selected sequentially on the basis of their sample number by the WHR characteristics that they were associated with from the polygon theme. Because sampling with replacement was used some polygons were tested for only one WHR characteristic while others were tested for all four (species type, canopy closure, size and canopy structure). Some polygons were sampled more than once for one or a combination of reasons: large polygon size, frequency of the characteristic, and/or random chance. Once a polygon was selected for sampling the sample point was assigned a "SELECTED" sample status. GRS developed plot packets for all "SELECTED" samples.

Ground Data Collection

The quantitative method for determining the vegetative characteristics of a stand uses a line transect method of sampling. Transects (plots), located within polygon boundaries, were 18 chains (about 1/4 mile) in length with points 12 feet apart. A total of 100 points were sampled along the transect. At each point a vertical sighting was taken using the GRS canopy densitometer (vertical sighting device). If the cross hairs of the densitometer intercepted a tree crown field personnel recorded: species, diameter at breast height (DBH) (to the nearest inch), crown diameter (to the nearest foot) and whether the tree was a spectral contributor or not for that point. If there was no tree cover at that point the crews recorded the appropriate cover codes for shrub species, grass, bare soil, rock, or duff. Only the top canopy layer for each point was recorded. Upon completing the plot, the crews filled out the back side of the plot cards and estimated the WHR characteristics. These qualitative estimates were later compared to the quantitative estimates. In addition to the stand characteristics the crew noted: slope, aspect, landform, soils, stand history, snags, and other features of interest. This method was used for both training and accuracy assessment ground data collection efforts. At no time during the accuracy assessment were the field personnel aware of what WHR characteristics they were sampling. Any prior knowledge can introduce bias into the samples.

The most important part of field data collection was the accurate location of the initial sample point. This was used as the start of the transect. The field data coordinator constructed plot packets containing the following: a USGS 1:24,000 quad map; orthophoto quad; and a 1:24,000 map showing polygon boundaries (without labels), 1000 meter ticks, section lines, sample point and transect locations. Compass and pacing were the basis for locating the sample points in the woods. Reference points (RP's) used to locate sample points were established using benchmarks, section corners, road intersections, points on a road (sharp curves), 'K' tags along section lines, ridges or creeks, or other prominent landmarks. (GRS is currently using GPS for training data collection and plans to use it for future accuracy assessment data collection efforts.)

The only way to actually know the **true** vegetation composition of a "SELECTED" polygon was to sample every thing in it. Due to the impracticality of a 100% sampling method it was desirable to sample as much of the polygon as possible. As mentioned above transects were 1/4 mile in length. The cardinal direction of the transects was north. If the transects crossed a polygon boundary the initial transect azimuth was rotated clock-wise at increments of 45 degrees until the entire transect fit inside the polygon. After initial rotation if the transect did not fit the polygon the first half of the transect (9 chains) was rotated using the above rotation rules. Once the first half fell inside the polygon boundary, the second half of the transect was rotated in the same manner, clockwise at 45 degrees until it fit. If it still didn't fit the first part was rotated to the next 45 degree increment and second half rotation was performed. This systematic transect rotation approach was taken to avoid bias in sampling polygons.

Data Analysis

Transect data were loaded into a relational database. The data were then processed with *GRS_polysum*. This program processes the vegetation data and outputs cover matrices similar to those in Table 2.

		SAMPLED POLYGON SIZE ESTIMATES									
		0 non-tree	1 0-5"	2 6-11"	3 12-23"	4 24-35"	5 36"+	TOTAL	PERCENT CORRECT	ACREAGE	CORRECT ACRES
	0							0			
MAPPED	1	7	16	4	3			30	53.3%	3,689	1,967
POLYGON	2	4	4	11	14	1	1	35	31.4%	87,303	27,438
CLASSES	3	6	2	17	78	11	4	118	66.1%	3,939,502	2,604,078
	4		1	5	33	13	3	55	23.6%	1,084,329	256,296
	5			4	23	14	17	58	29.3%	566,969	166,181
	TOTAL	17	23	41	151	39	25	296		5,681,792	3,055,960
PERCENT			69.6%	26.8%	51.7%	33.3%	68.0%		45.6%		

TOTAL PERCENT CORRECT ACRES 53.8%

KHAT 0.2563 Var(Khat) = 0.001406

Table 3: WHR Size Error Matrix by Absolute Match Type

Stand Cover Density Summary:
Stand: 876

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover	Non-Tree Cover	Total Cover
Species								
Douglas-fir	2.0%	1.0%	2.0%	12.1%	17.1%	34.1%		34.1%
ponderosa pine		1.0%	1.0%	5.1%	2.0%	9.0%		9.0%
Jeffrey pine		1.0%	2.0%		5.1%	8.0%		8.0%
sugar pine		1.0%	2.0%	8.1%		11.1%		11.1%
white fir			1.0%			1.0%		1.0%
hardwoodC		2.0%				2.0%		2.0%
hardwood			5.0%			5.0%		5.0%
misc shrub							1.0%	1.0%
rock							4.0%	4.0%
BAR exp soil							2.0%	2.0%
duff/debris							22.8%	22.8%
Total Cover	2.0%	6.0%	13.0%	25.3%	24.2%	70.2%	29.8%	100.0%
Total Tree Cover						70.2%		

Stand Tree Density Summary:
Stand: 876

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	All Sizes
Species						
Douglas-fir	2.8%	1.4%	2.8%	17.2%	24.4%	48.6%
ponderosa pine		1.4%	1.4%	7.3%	2.8%	12.8%
Jeffrey pine		1.4%	2.8%		7.3%	11.4%
sugar pine		1.4%	2.8%	11.5%		15.8%
white fir			1.4%			1.4%
hardwoodC		2.8%				2.8%
hardwood			7.1%			7.1%
Total Tree Cover	2.8%	8.5%	18.5%	36.0%	34.5%	100.0%

Stand Quadratic Mean DBH Summary:
Stand: 876

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	All Sizes
Species						
Douglas-fir	2.9"	7.4"	15.4"	27.4"	44.9"	35.9"
	2.0pts	1.0pts	2.0pts	12.0pts	17.0pts	34.0pts
ponderosa pine		7.4"	12.8"	27.8"	36.0"	27.3"
		1.0pts	1.0pts	5.0pts	2.0pts	9.0pts
Jeffrey pine		9.6"	21.0"		37.2"	31.4"
		1.0pts	2.0pts		5.0pts	8.0pts
sugar pine		8.6"	18.9"	31.8"		28.5"
		1.0pts	2.0pts	8.0pts		11.0pts
white fir			17.1"			17.1"
			1.0pts			1.0pts
hardwoodC		9.1"				9.1"
		2.0pts				2.0pts
hardwood			12.4"			12.4"
			5.0pts			5.0pts
Quad Mean DBH	2.9"	8.6"	15.9"	29.0"	42.7"	31.3"
	2.0pts	6.0pts	13.0pts	25.0pts	24.0pts	70.0pts
Quad Mean DBH - Con	2.9"	8.3"	17.8"	29.0"	42.7"	32.8"
	2.0pts	4.0pts	8.0pts	25.0pts	24.0pts	63.0pts
Quad Mean DBH - Hwd		9.1"	12.4"			11.6"
		2.0pts	5.0pts			7.0pts

Table 2. Example Cover Matrix

GRS_polysum also summarizes the WHR classes expressed by the cover matrices. Error matrices were developed for each WHR category (canopy closure, size, type and structure) and KAPPA (Khat) coefficients were calculated (Congalton, Oderwald and Mead 1983; Rosenfield and Fitzpatrick-Lins 1986). An Absolute match occurred when the WHR characteristics from the polygon and sample (reference) data were the same. The error matrix for WHR Size Classes can be seen in Table 3. The Absolute match correspondence was poor. Field personnel also estimated WHR characteristics of the polygon being sampled after completing the transect. A Qualitative Match was assigned when the ocular estimates matched the polygon estimates. The qualitative information was incorporated into error matrices by WHR category. During the accuracy assessment questions arose as to whether the transects adequately sampled the polygons and whether the transects were picking up heterogeneity in the polygons. To help answer these questions the Half Transect match type was developed. In a Half Transect Match the 100 transect points were split in half. WHR characteristics were calculated for both the first 50 points and second 50 points of the transect. A Half Transect match occurred when the polygon and at least one of the transect halves had the same WHR characteristic.

Attributes from the pilot project maps were derived from quantitative training data. Final map polygon attributes not only included estimates of the WHR classes, but also had continuous estimates of diameter (qmd), crown closure (density), percent conifer, and percent hardwood. The Absolute, Qualitative and Half Transect match types compare categorical data to categorical data. This traditional approach used in generating error matrices ignores the fact that class boundaries used in most mapping projects are artificial and rarely occur in nature (Congalton, 1991). Table 4 illustrates the problem.

	stand ID	WHR TYPE	Dominant Species	Closure Class	Closure	Size Class	Size	Canopy Structure
sample	271	RFR	red fir	3	59%	3	23.1"	E
polygon	12254	RFR	red fir	4	61%	4	24.6"	E

Table 4. Comparison of Sample 271 and Stand 12254

The Class Width match was developed to handle the situation shown in Table 4. If the sample data and the polygon data were not in the same class, the polygon's class span is added to the sample's class span. The resulting class span is divided by 2 to establish a sliding class width. If the continuous estimates from the polygon and sample data were within the newly established class width they were considered a match. Class spans are shown in Table 1. A span of 12" for size class 5 (36"+) and a span of 20% for closure class D (>= 60%.) were used. This is straight forward for estimates of density and size. For example, if the polygon estimate of canopy closure was 61% (dense) and the transect was 59% (moderate) by categorical types of matches, these estimates do not agree. By applying the class width match method to this example the polygon closure class dense 20% span was added to the sample closure class moderate (40-59%) 19% span and divided by 2 to yield a sliding class width of 19.5%. Since both closure estimates fall within the established sliding class width a match is awarded for the sample and map estimates. The same methodology may be applied to estimates of size.

The WHR Classification System rules used for defining species types vary (Mayer and Laudenslayer 1988.) Class Width match methods were applied to forest types. WHR forest types are those stands with at least 10% tree cover. Pure species types (DFR, RFR, WFR etc...) used a "50% rule" for their respective designations. The "50% rule" utilizes a 50% threshold of the tree cover for single species types. If a stand had >=50% of the tree cover as Douglas-fir, then the stand was assigned a DFR species type. WHR species types are shown in Table 1. The Montane-Hardwood Conifer type (MHC) is a mixed type in which there is > 33% hardwood **and** > 33% conifer. The Montane Hardwood type (MHW) has > 66% tree cover of hardwood species. The Closed-Cone Pine-Cypress and Subalpine Conifer type are

Polygon Data
Stand Tree Density Summary:
Stand: 249195

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	All Sizes
Species						
Douglas-fir	2.9%	2.3%	0.9%	1.6%		7.7%
ponderosa pine	30.7%	9.1%	10.3%	3.8%		53.9%
sugar pine	4.1%	0.1%	0.9%			5.2%
white fir	4.7%	5.1%	3.8%	0.9%		14.5%
cedar	2.2%	1.8%	5.8%	5.7%		15.5%
hardwoodC	2.4%	0.4%	0.4%			3.2%
Total Tree Cover	47.0%	18.8%	22.1%	12.0%		100.0%

****

Sample Data
Stand Tree Density Summary:
Sample Number: 993

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	All Sizes
Species						
Douglas-fir	7.5%	1.9%	9.4%			18.8%
ponderosa pine	5.7%	7.8%	22.6%	9.4%		45.5%
cedar	7.5%	1.9%	13.2%	11.3%		33.9%
hardwood		1.8%				1.8%
Total Tree Cover	20.7%	13.4%	45.2%	20.7%		100.0%

Table 5: Class Width Match Example for KMC

comprised of the tree species which grow endemic to their respective ecosystems. The last forest type is the Mixed Conifer type; this type is assigned when at least 3 conifer species are present with at least 10% tree cover and no one species is over 50% of the tree cover. The Class Width match methodology for WHR species type utilizes the Tree Density Summary portion of the cover matrix (Table 2.) The Tree Density Summary of the sample data were compared to the polygon Tree Density Summary. For comparing mixed conifer types (KMC,CPC,SCN) to single species conifer types (DFR,RDW,WFR,RFR,PPN,LPN,JUN), a Class Width of 10% tree species cover was applied to the 50% threshold for specific species type calls. Table 5 illustrates the application of the agreement testing. If the predominant conifer species of KMC was the same as the pure species type and the % tree cover was above 40% (a 10% Class Width) the sample was considered a match. The reverse case also resulted in a match if at least 3 conifer species, each over 10% tree cover, were present and the predominant conifer species is < 60% tree cover. In checking MHC and pure species types, a 30% Class Width to the % hardwood tree cover was used. If the MHC predominant conifer species was the same as the pure species type and the % hardwood tree cover estimates were within the 30% Class Width the sample was considered a match. This 30% Class Width around the % hardwood tree cover was also applied to the comparison of MHC to MHW.

Polygon Data
Stand: 247512

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover	Non-Tree Cover	Total Cover
Species								
Douglas-fir	5.3%	10.4%	13.4%	4.4%	6.5%	40.1%		40.0%
redwood	2.0%	4.3%	10.6%	3.6%	10.4%	30.9%		30.0%
white fir	0.1%	1.0%	2.1%	0.4%		3.7%		3.0%
hardwoodC	2.0%	5.2%	4.2%	0.6%		12.0%		12.0%
cham/art						1.8%	1.0%	
duff/debris						1.8%	1.0%	
noncontributor						5.3%	5.0%	
Total Cover	9.5%	21.9%	31.5%	9.2%	17.1%	89.2%	10.8%	100.0%
Total Tree Cover						89.2%		

****

Sample Data
Sample Number: 416

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover	Non-Tree Cover	Total Cover
Species								
Douglas-fir			17.0%	5.0%	4.0%	26.0%		26.0%
redwood			17.0%	5.0%	10.0%	37.0%		37.0%
white fir	1.0%	2.0%	9.0%	5.0%		17.0%		17.0%
hardwoodC	1.0%	2.0%	2.0%			5.0%		5.0%
misc hardwood		1.0%				1.0%		1.0%
cham/art							1.0%	1.0%
misc shrub							6.0%	6.0%
prairie							2.0%	2.0%
forb/herbac							5.0%	5.0%
Total Cover	2.0%	5.0%	45.0%	15.0%	14.0%	86.0%	9.0%	100.0%
Total Tree Cover						86.0%		

Table 6: Class Width Match for Canopy Structure

Table 6 exemplifies the application of the Class Width match approach to canopy structure. The Cover Density Summary portion of the cover matrix was used. An uneven structure designation was given to those conifer types that exhibited a skip, <10% conifer cover, in a size class and have >10% conifer cover above the skipped class and >25% conifer cover below the skip. A sliding class width of 10% conifer cover was applied to the skipped class for the structure test. If the data showed one of the estimates having a skip (or <10% conifer cover) in a size class and the other with no skip (>10% conifer cover) in the same size class by categorical matching techniques these estimates do not match. As in Table 6, the polygon data had a skipped class (9% conifer cover in the 24-35" size class) and the conifer cover in the 24-35" size class from the sample data was 15%. A Class Width match was awarded for the sample and polygon data for canopy structure because the tree cover in the skipped size class fell within the 10% sliding class width and the other distributions for uneven structured designation were present.

RESULTS

Percent Correct	Percent Correct Acres	Khat		M	H	Q	C
45.6%	53.8%	0.256	M	0.00 NS			
55.7%	64.7%	0.392	H	2.56 S	0.00 NS		
54.4%	62.7%	0.375	Q	2.23 S	-0.31 NS	0.00 NS	
69.3%	75.3%	0.582	C	6.31 S	3.68 S	3.98 S	0.00 NS

M - Absolute
H - Half Transect
Q - Qualitative
C - Class Width

S - Significantly Different
NS - Not Significantly Different

Table 7: Comparison of Match Types at a 95% Probability Level for Size

Percent Correct	Percent Correct Acres	Khat		M	H	Q	C
47.4%	38.3%	0.437	M	0.00 NS			
53.0%	47.4%	0.497	H	2.08 S	0.00 NS		
54.6%	48.6%	0.515	Q	2.68 S	0.60 NS	0.00 NS	
67.2%	72.3%	0.648	C	7.54 S	5.41 S	4.80 S	0.00 NS

M - Absolute
H - Half Transect
Q - Qualitative
C - Class Width

S - Significantly Different
NS - Not Significantly Different

Table 8: Comparison of Match Types at a 95% Probability Level for Species Type

Error matrices were developed for the above match types (Absolute, Qualitative, Half Transect, and Class Width) for the four WHR categories (canopy closure, size, cover type, and canopy structure.) Table 3 is an example of the size error matrix by Absolute match. The error matrices developed have individual class accuracies (errors of omission and errors of comission) and three measures of overall map accuracy. The first overall accuracy measurement was the total percent correct; the diagonal sum divided by the total number of samples. The second measure of overall map accuracy was the percent correct acres. This measure involves using the "user's accuracy" (Story and Congalton 1986). The class percent correct was multiplied by the acres in that class estimating the

Percent Correct	Percent Correct Acres	Khat		M	H	Q	C
75.5%	93.9%	0.003	M	0.00 NS			
76.9%	94.0%	0.090	H	1.00 NS	0.00 NS		
78.5%	94.9%	0.146	Q	1.54 NS	0.52 NS	0.00 NS	
88.1%	95.3%	0.617	C	6.50 S	4.83 S	4.14 S	0.00 NS

M - Absolute
H - Half Transect
Q - Qualitative
C - Class Width

S - Significantly Different
NS - Not Significantly Different

Table 9: Comparison of Match Types at a 95% Probability Level for Canopy Structure

Canopy Closure

Percent Correct	Percent Correct Acres	Khat		M	H	Q	C
54.5%	61.3%	0.372	M	0.00 NS			
63.8%	68.1%	0.500	H	2.33 S	0.00 NS		
61.8%	67.1%	0.473	Q	1.84 NS	-0.48 NS	0.00 NS	
75.6%	81.0%	0.667	C	5.65 S	3.02 S	3.67 S	0.00 NS

M - Absolute
H - Half Transect
Q - Qualitative
C - Class Width

S - Significantly Different
NS - Not Significantly Different

Table 10: Comparison of Match Types at a 95% Probability Level for Canopy Closure

number of correctly mapped acres for the class. The correct acres were summed and divided by the total number of acres. The result was the percentage of correctly type acres. Khat (Bishop, Fienberg and Holland 1975, p. 396) was the third measure of map accuracy. This statistic incorporates the off diagonal elements in the overall correspondence of sample and map data. Khat, along with its variance, was also used to test for differences between the different match type error matrices. There are too many tables to list in this paper. Tables 7 - 10 summarize the 16 error matrices produced and the results of the pairwise significance tests between match types. Significance test were performed using method described by Congalton and Mead (1983). Significant differences are those with Z values > 1.96.

DISCUSSION

The Half Transect match type exhibited higher map accuracies for Canopy Closure and Size than the Absolute and Qualitative match types. Agreement between the separate transect halves was around 65% for WHR characteristics, except for structure, which was 93%. This indicates the 40 acres mmu may be too large resulting in a high within polygon variation. Another possibility is the ground sampling technique does not adequately sample the vegetation. The Qualitative match yielded higher map accuracies than the Absolute Match for species type and size. This may be because field crews made the qualitative WHR estimates after completing the transects. Future research is needed to compare qualitative to quantitative estimates used as reference data in accuracy assessment. The Qualitative match was important when sampling low density tree types where the transect data indicated a non-tree type. The Qualitative and Half Transect match types resulted in similar measures of map accuracies with respect to each other. In all WHR map categories the Class Width Match produced significantly higher accuracies than the other match types. The other match types compare class estimates where as the Class Width methodology compares continuous estimates. These results reflect the artificial nature of the class boundaries used in the classification. The results are also a product of the rules used in determining the Class Width match.

Of the measures of overall map accuracy, Khat was the only measure used to incorporate the off diagonal elements and chance agreement. This statistic is widely used by the remote sensing community as a measure of thematic map accuracy. While it does account for both errors of omission and comission, Khat does not take into account the area covered by the individual map classes. The percent correct acres is a measure of how well the map represents what is actually on the ground.

CONCLUSION

Quantitative data used as ground truth in accuracy assessment presents new questions as to how matches are defined. The Class Width match is an appropriate method for this mapping project because of the continuous estimates in the maps and the pixel aggregation process (Stumpf, 1993). Pixels were aggregated into polygons using a methodology that reduced within polygon variation, as opposed to aggregating towards the class midpoints. The incorporation of conditional KAPPA ,as described by Rosenfield and Fitzpatrick-Lins (1986), into an area based accuracy statistic is needed as an overall measure of map reliability. Results from the pilot project accuracy assessment are providing valuable quantitative information on the types of errors and their spatial distribution. This information is crucial in the further developments in mapping methodology used for mapping the entire Klamath Province. Future maps will also include measures of within polygon variance. This variance can be used in future agreement testing of the continuous estimates included in the new maps. CDF's foresight in using accuracy assessment as an iterative step in the mapping process shows their commitment to responsibly using satellite data for such an ambitious mapping effort.

LITERATURE CITED

Biging, G., R. Congalton and E. Murphy, 1991. A Comparison of Photointerpretation and Ground Measurements of Forest Structure. In: Proc. of the 57th Annual Meeting of American Society of Photogrammetry and Remote Sensing, Baltimore, MD, (3):6-15.

Bishop, Y., S. Feinberg, and P. Holland, 1975. Discrete Multivariate Analysis - Theory and Practice. MIT Press Cambridge, MA

Brown, G. and L. Fox, 1992. Digital Classification of Thematic Mapper Imagery for Recognition of Wildlife Habitat Characteristics. In: Proc. 1992 ASPRS/ACSM Convention, American Society of Photogrammetry and Remote Sensing, Bethesda, MD, (4):251-260

Cochran, W., 1977. Sampling Techniques. John Wiley and Sons. New York.

Congalton. R., 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sens. Environ. (37):35-46

Congalton, R., R. Oderwald, and R. Mead, 1983. Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques. Photogramm. Eng. Remote Sens. 49(12):1671-1678

Congalton, R., 1988. A Comparison of Sampling Schemes Used in Generating Error Matrices For Assessing the Accuracy of Maps Generated from Remotely Sensed Data. Photogramm. Eng. Remote Sens. 54(5):593-600.

Congalton, R. and K. Green, 1993. A Practical Look at the Sources of Confusion in Error Matrix Generation, Photogramm. Eng. Remote Sens. 59(5):641-644

Mayer, K. and W. Laudenslayer Jr. 1988. A Guide to Wildlife Habitats of California. California Department of Forestry and Fire Protection.

Rosenfield, G. and K. Fitzpatrick-Lins, 1986. A Coefficient of Agreement as a Measure of Thematic Map Accuracy, Photogramm. Eng. Remote Sens. 52(2):223-227.

Stehman, S., 1992. Comparison of Systematic and Random Sampling for Estimating the Accuracy of Maps Generated from Remotely Sensed Data, Photogramm. Eng. Remote Sens. 58(9):1343-1350.

Story, M. and R. Congalton, 1986. Accuracy Assessment: a User's Perspective. Photogramm. Eng. Remote Sens. 52(3):397-399.

Stumpf, K., 1993. From Pixel to Polygons: The Rule-Based Aggregation of Satellite Image Classification Data Using Ecological Principles. In: Proc. Seventh Annual Symposium on GIS in Forestry, Environment and Natural Resources. Vancouver B.C. Canada. (2):939-945