ELSEVIER

Remote Sensing of Environment

# Using a Mensuration Approach with FIA Vegetation Plot Data to Assess the Accuracy of Tree Size and Crown Closure Classes in a Vegetation Map of Northeastern California

*Samantha J. Gill,* Jeff Milliken,† Debby Beardsley,‡ and Ralph Warbington§*

*The accuracy of the Lassen-Modoc vegetation map was assessed using data from a grid of permanent Forest Inventory and Analysis (FIA) plots in conjunction with a fuzzy logic approach. This paper presents the results for the Modoc National Forest. The main foci of this paper are the methods and results for assessing the accuracy of tree size and crown closure classes using a mensurational approach based on calculations made from the plot data. This approach proved very useful in conjunction with fuzzy accuracy methods. Higher error rates reported for crown closure classes may be due to a variety of factors including error associated with crown closure modeling methods, error associated with crown closure calculations for a plot, and the variability of crown closure within a plot. Using the FIA data has proven to represent a cost-effective means of generating accuracy assessment information based on a probability sample.* ©*Elsevier Science Inc., 2000*

* Natural Resources Management Department, California Polytechnic and State University, San Luis Obispo, CA

† U.S. Department of Interior, Bureau of Reclamation, Sacramento, CA

‡ Pacific Northwest Research Station, U.S. Department of Agriculture, Forest Service, Portland, OR

§ U.S. Department of Agriculture, Forest Service, Region 5 Remote Sensing Lab, Sacramento, CA

Address correspondence to Samantha J. Gill, Natural Resources Management Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA

*Received 24 September 1999; revised 10 February 2000.*

## INTRODUCTION

The Lassen-Modoc vegetation-mapping project is a USDA Forest Service Region 5 and California Department of Forestry and Fire Protection cooperative vegetation-mapping program covering nine million acres of the northeastern portion of California. Vegetation maps were produced using remotely sensed processing and GIS modeling techniques (Miller et al., 1994). For each polygon (minimum mapping unit of 1 ha), a lifeform type and Classification and Assessment with Landsat of Visible Ecological Groupings (CALVEG) (U.S. Forest Service Regional Ecology Group, 1981) type are mapped. Crown closure and tree size are mapped for hardwood and conifer types. The purpose of this paper is to report the accuracy assessment methods and results for size and crown closure classes for the Modoc National Forest portion of the Lassen-Modoc vegetation map and address aspects of the methods used in the accuracy assessment.

Methods and results of lifeform and CALVEG for this project have been published elsewhere (refer to Milliken et al., 1998). Generally, the overall accuracy for lifeform and CALVEG were high. Using the MAX operator, overall accuracy of lifeform classes for the Modoc National Forest is 86.9% and increases to 95.5% when using the RIGHT operator. Overall accuracy for conifer CALVEG types is 59.5% using the MAX operator and 89.8% using the RIGHT operator. In the shrub classes, overall accuracy significantly increases between the MAX (44.8%) and RIGHT (90.8%) operators. For more detailed results of lifeform and CALVEG refer to Milliken et. al (1998).

REMOTE SENS. ENVIRON. 73:298–306 (2000)

655 Avenue of the Americas, New York, NY 10010

0034-4257/00/$–see front matter
PII S0034-4257(00)00105-X

The primary motivation for doing an accuracy assessment of a map is to provide map users with information that will aid in effectively utilizing the map and will provide map producers with information that can be used for improving current maps and refining procedures used in generating new maps. The four types of information about errors that are of interest to users and/or producers of a map are the nature, frequency, magnitude, and source of the errors (Gopal and Woodcock, 1994). Traditional accuracy assessment procedures compare the label assigned to a polygon in the map (map label) to the label assigned to the same polygon using ground data (often referred to as the expert evaluation or ground truth). The results are often displayed in an error or confusion matrix in which map labels are displayed on one axis (rows or columns) and labels from ground truth on the other axis (Story and Congalton, 1986). This approach assumes that each polygon in a map can be assigned one accurate label from the ground truth data. Additional methods for analyzing an error matrix are reviewed by Congalton (1991) or Janssen and van der Wel (1994). In addition, traditional methods of assessing the accuracy of a map often include additional collection of ground reference data obtained from fieldwork or photointerpretation (Congalton, 1991; Muller et al., 1998). One common approach to collecting ground truth information for the purpose of assessing the accuracy of a map is to visit a site in the field corresponding to a polygon on the map and assign to it a label based on the classification categories used in the mapping project. This method may be time-consuming since there may be considerable distance between points, and polygons may be far from roads and thus travel between points will be expensive. However, if the observer simply visits the sites, writes down what he/she believes is the correct classification, and then compares it to the map the method is rather inexpensive. In some cases sampling designs are developed and employed, but in other cases a sample design is not followed. Other approaches to collecting ground data include the use of simple random and stratified sampling (Congalton, 1991).

The prohibitive cost associated with collecting accuracy assessment data often results in no accuracy assessment being conducted. This study's approach utilized existing Forest Inventory and Analysis (FIA) Permanent Plot grid inventory data to assign ground truth labels, significantly reduced the costs associated with collecting the accuracy assessment data. Considerations in using FIA data for accuracy assessment will be discussed later in this paper.

A modified fuzzy logic accuracy assessment approach based on Gopal and Woodcock (1994) was used in this project. Muller et al. (1998) found that using a fuzzy logic approach allowed for more feasible field observation methods. The concept of a fuzzy set was introduced by Zadeh (1963, 1965) to describe imprecision that is characteristic of much of human reasoning. With fuzzy sets, there are different grades of membership within a class. In the case of a vegetation map, one label may be absolutely correct, but other labels may be considered good or acceptable. For example, for a given site (in this case an inventory plot within a map polygon) a map label of 30% crown closure may be considered absolutely correct, but a map label of 40% crown closure may still be within an acceptable range. Using the traditional error matrix, only one possible answer (considered to be the best answer by an "expert" in the field) is compared to the map label. Fuzzy set theory allows the user and producer to look at ranges of acceptable answers and understand more about the magnitude of errors in the map.

## METHODS

### Data Collection and Assignment of Ground Truth Labels

The data used in this study came from two sources, USFS Region 5 FIA Permanent Plot grid data for areas within a National Forest and Pacific Northwest Research Station (PNW) FIA permanent plot grid data for areas outside of National Forest boundaries. Only the data for the Modoc National forest will be presented in this paper. FIA data are used to provide current estimates of forestland area, timber volume, net annual growth and mortality, and harvest. Plots are located on a 3.4-mile grid across California. Plot installation on the National Forest is administered by the USFS Region 5 (R5) inventory staff and plots outside the National Forests are administered by the PNW Research Station. These independent data sets were not used in any way as part of the vegetation-mapping project.

Between 1993 and 1994, 307 five-point cluster plots were installed on the Modoc National Forest. Five-point cluster plots were installed at each grid location and GPS control was provided. Each cluster (all five points) covered approximately 1 ha. This 1-ha plot area was equivalent to the minimum mapping unit of the vegetation map being assessed. Map polygons were formed using image segmentation algorithms (Woodcock et al., 1992). FIA inventory plot locations were checked by the USFS to ensure they fell within a single map polygon as part of the integrated inventory process. At each point of the cluster plot, a variable radius plot was installed using a Basal Area Factor of 20 or 40. A 1/4-acre circular fixed plot was used at each point for data on very large live trees, large snags, vegetation/ground cover, special features, and woody debris. Plot measurements included all live and dead trees, shrubs, grasses and forbs, growth and mortality, duff and fuel bed depth, stand structure and history, CALVEG type, terrain data, and other attributes [see FIA User's Guide (U.S. Forest Service—Region 5, 1995) for detailed information].

*Table 1.* Size Classes Used in Analysis

| *Size Class Name* | *Range DBHs in cm (inches)* | *Approximate Class Width in cm (inches)* |
|---|---|---|
| 0 | <2.5 (1.0) | 2.5 (1) |
| 1 | 2.5–12.4 (1.0–4.9) | 10.2 (4) |
| 2 | 12.4–30.2 (5.0–11.9) | 17.8 (7) |
| 3 | 30.2–60.7 (12.0–23.9) | 30.5 (12) |
| 4 | 60.7–101.3 (24.0–39.9) | 40.6 (16) |
| 5 | >101.3 (40.0) | 50.8 (20) |

The grid was densified where necessary to capture undersampled forest types. The densification increases the sample size for the assessment of individual classes. However, it could possibly influence the overall accuracy because the sampling weights required for estimation would be different for different classes. Also, by assuming the samples are of equal probability when in fact densification has occurred could lead to biased results. For purposes of this study, it was assumed that results within the densified regions are the same as those outside the densification and thus all sample points are treated as if the sample design is equal probability. In doing this, we are assuming that the accuracy of the area represented by the densified sample is similar to accuracy found elsewhere on the map.

At each point within a plot, the inventory crew filled out an accuracy assessment form indicating best and second-best lifeform (general growth forms: conifer, hardwood, mixed conifer and hardwood, shrub, herbaceous, nonvegetated). Within the best and second-best lifeform type, best and second-best CALVEG types were assigned. If the lifeform (best or second best) was a conifer or hardwood type, best and second-best size and crown closure classes were also assigned. Even though best and second-best size and crown closure classes were recorded in the field, this information was not used in this study. Instead, size and closure were calculated from the trees measured at each point. Refer to the tables in the results section for the number of sites for each size and crown closure category, and refer to Milliken et al. (1998) for the number of sites in lifeform and CALVEG.

## Assigning Fuzzy Ratings for Map Labels

To utilize a fuzzy logic approach, a fuzzy rating system indicating the degree of correctness must be developed. The fuzzy rating system used in this study was:

5: Absolutely right—there is no doubt about the match.
4: Good label—it would be above satisfactory to find this label given on the map.
3: Acceptable label—this maybe not the best possible label but it is acceptable.
2: Understandable but wrong—this is not an acceptable label. There is something about the site that makes the label understandable but there is clearly a better label.
1: Absolutely wrong—this label is absolutely unacceptable.

Fuzzy ratings were determined for lifeform, CALVEG type, tree size, and tree crown closure.

### Size

The quadratic mean diameter (QMD) of overstory trees (crown classes of predominants, dominants, and codominants) on each plot was calculated and used to determine to which size class a plot should be assigned. For this analysis, only overstory trees were used because the remote sensing techniques used for mapping tree size primarily represent overstory trees. For the analysis of size, six size classes were used (refer to Table 1). The mapped diameter class containing the calculated QMD was given a fuzzy rating of 5 and the other mapped classes were assigned fuzzy ratings based on percentages of the class width in which they fell. The percentage of class widths used to assign fuzzy ratings were: 10% of the class width for a fuzzy rating of 5; 30% of the class width for a fuzzy rating of 4; 60% of the class width for a fuzzy rating of 3; 120% of the class width for a fuzzy rating of 2; and greater than 120% of the class width for a fuzzy rating of 1. With these rules for assigning fuzzy ratings, a plot that is very close to the boundary of two classes (within 10%) would be considered absolutely right (a rating of 5) for either class. As an example of assigning fuzzy ratings using this scheme, consider a plot with a calculated QMD of 56.1 cm (22.1 inches). This value falls within class 3, which has a class width of 30.5 cm (12 inches). Plus or minus 10% of the 30.5 cm (12 inches) class width, or 3.0 cm (1.2 inches), gives the range of 53.1 cm to 58.9 cm (20.9–23.2 inches). Because the range of 53.0 cm to 58.9 cm (20.9–23.2 inches) is entirely within class 3, no other classes would receive a rating of 5. Plus or minus 30% of this class width gives a range of 47.0 cm to 65.3 cm (18.5–25.7 inches). Class 4 would be assigned a rating of 4 because the range of 30% of the class width overlaps class 4. Similarly, class 2 would be assigned a fuzzy rating of 2 because it is within 120% of the class width.

### Crown Closure

Crown closure was calculated from the inventory data using a mensurational approach. For each tree on a point,

*Table 2.* Results of the MAX and RIGHT Operators for Size

| | | *Expert Evaluation* | | | |
|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *Matches Using MAX* | *Matches Using RIGHT* | *Increase* | *Weight* |
| 0 | 0 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.0266 |
| 1 | 0 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.0154 |
| 2 | 86 | 29 (33.7%) | 67 (77.9%) | 38 (44.2%) | 0.5366 |
| 3 | 109 | 95 (87.2%) | 105 (96.3%) | 10 (9.2%) | 0.4003 |
| 4 | 11 | 4 (36.4%) | 8 (72.7%) | 4 (36.4%) | 0.0210 |
| 5 | 0 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.0000 |
| Total | 206 | 128 (62.1%) | 180 (87.4%) | 52 (25.2%) | 1.0000 |
| Weighted | | (53.8%) | (81.9%) | (28.2%) | |

maximum crown width was predicted using the species-specific linear or quadratic equations developed by Warbington and Levitan (1992). Because these equations estimate maximum crown widths, an estimate of overlapping crown closure was obtained. However, the mapped crown closure calculated from remotely sensed data was nonoverlapping crown closure. To estimate nonoverlapping crown closure from the ground data, the following equations based on the assumption of random placement of tree crowns from Warbington and Levitan (1992) were used:

$$\text{noc} = -0.0319 + \{1.1510 * [1.0 - \exp(-\text{oc}/43560.0)]\} \quad (1)$$

with the following restriction if (noc>1.0) noc=1.0, where noc=percentage of nonoverlapping crown closure and oc=overlapping crown closure in square feet. The oc value is obtained by summing the area obtained from the crown width equations for all trees on the plot.

Because a conversion from overlapping to nonoverlapping crown closure was applied, both overstory and understory trees on the plot were used to calculate crown closure. The average crown closure over all points in a cluster plot was calculated and used to calculate fuzzy ratings for a site. As an estimate of the standard error of calculated crown closure, the following formula was used: se(noc)=(noc/ba)·se(ba), where se(noc)= standard error of calculated crown closure for a plot; noc=mean crown closure for a plot; ba=mean basal area for a plot; and se(ba)=standard error of the basal area for all five points on a plot. This assumes that the crown closure standard error is proportional to the basal area standard error. This assumption was made based on the fact that crown closure is a function of basal area and is only intended to be a rough approximation of the standard error. In fact, the standard error of crown closure should be at least that of basal area because crown closure is estimated as a function of basal area. Some plots had very large variances for calculated crown closure, and it was assumed that for plots with large variances, an adequate assessment of the map label could not be accomplished. These large variances came about because of inadequate precision from the FIA data. In these cases, the use of the existing FIA data produced a sample that was inadequate for characterizing precisely the classification of some polygons. Because of this inadequate precision, when the se(crown closure) spanned more than 1.8 crown closure classes, the sites were dropped from the analysis. Twenty-four of the 206 sites were dropped from the analysis.

The calculated crown closure for a point was then assigned to 10% crown closure classes with class 1 ranging from 10% to 20% crown closure and class 9 ranging from 90% to 100% crown closure. Fuzzy ratings were then assigned to each of these categories using the following system. When the mapped class was within 7% of the calculated crown closure, the class was assigned a fuzzy rating of 5; when it was within 10% (one class width) of the calculated crown closure, the class was assigned a fuzzy rating of 4; when it was within 15%, the class was assigned a rating of 3; when it was within 18%, the class was assigned a rating of 2; and when it was greater than 18%, it received a rating of 1. Thus, a class

*Table 3.* Results of the DIFFERENCE Operator for Size

| | | *Mismatches* | | | | *Matches* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *−4* | *−3* | *−2* | *−1* | *0* | *1* | *2* | *3* | *4* | *Mean* |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |
| 2 | 86 | 1 | 18 | 22 | 16 | 9 | 15 | 5 | 0 | 0 | −1.08 |
| 3 | 109 | 0 | 4 | 7 | 3 | 23 | 38 | 34 | 0 | 0 | 0.71 |
| 4 | 11 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 0 | 0 | −1.18 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |
| Total | 206 | 2 | 24 | 31 | 21 | 34 | 54 | 40 | 0 | 0 | |

*Table 4.* Results of the CONFUSION (C) and the AMBIGUITY (A) Operators for Size

| | | 0 | | 1 | | 2 | | 3 | | 4 | | 5 | | No. of Mismatches | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* |
| 0 | 0 | X | X | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | X | X | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 86 | 0 | 1 | 0 | 2 | X | X | 57 | 9 | 19 | 16 | 1 | 4 | 77 | 32 |
| 3 | 109 | 0 | 4 | 4 | 2 | 9 | 16 | X | X | 5 | 7 | 0 | 3 | 18 | 32 |
| 4 | 11 | 0 | 2 | 1 | 2 | 3 | 2 | 7 | 2 | X | X | 0 | 1 | 11 | 9 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | X | X | 0 | 0 |
| Total | 206 | 0 | 7 | 5 | 6 | 12 | 18 | 64 | 11 | 24 | 23 | 1 | 8 | 106 | 73 |

was given a rating of at least acceptable (fuzzy rating of 3 or greater) when it was within 15% (one and one-half class widths) of the calculated crown closure. For example, if the calculated crown closure were 48%, a crown closure map label of 4 (40% to 50%) would be assigned a rating of 5, as would a map label of class 5 because 48%+7%=55%, which was within class 5. A map label class of 3 did not receive a rating of 5 because 48%−7%=41%, which was still within class 4. A map label of class 3 would be assigned a rating of 4 (within 10% of the calculated crown closure) and a map label of class 6 was assigned a rating of 3 (within 15% of the calculated crown closure). All other classes were assigned a rating of 1 because the range for a rating of 2 is 30 to 66, which falls into mapped classes already assigned ratings (above). Because many people use broader classes for crown closure, the same procedures were also used to evaluate crown closure using the "sparse" (10–20%), "poor" (light, 20–40%), "normal" (medium, 40–70%), and "good" (heavy, 70–100%; referred to as SPNG classes) of crown closure.

### Creating Fuzzy Accuracy Assessment Tables

Standard fuzzy accuracy assessment procedures were completed and tables displaying results were computed using procedures developed by the Boston University Center for Remote Sensing (Gopal and Woodcock, 1994). The accuracy assessment is done using a hierarchical approach with lifeform accuracy evaluated first. CALVEG accuracy is evaluated for polygons that are considered acceptable (rating=3) in the lifeform classification. Size and crown closure accuracy are evaluated for forested lifeforms.

Two fuzzy measures were used in this study, the MAX and RIGHT operators. The MAX operator evaluates whether the best label was assigned to a map polygon; a match was assigned if the map label was the same as the class receiving the highest rating. The RIGHT operator evaluates whether the map label for an accuracy site is acceptable; a match was assigned if the map label received a fuzzy rating of at least 3. For most purposes, it is reasonable to use the RIGHT operator as an indication of the accuracy of the map, but this is a function of both the tolerances used for the fuzzy ratings as well as the sensitivity of an analysis that is based on the map. For both the MAX and RIGHT operators, a standard error for the percent of matches can be computed using the following formula: $\sqrt{p(1-p)/n}$, where $p$ is the estimated proportion of matches and $n$ is the number of sites in that class. Because mapped categories occupy different amounts of area, accuracy using the MAX and RIGHT operators was computed using a weighted approach (i.e., weighting by the percentage of mapped area occupied by each category).

To assess the magnitude of the errors, the DIFFERENCE operator is used. The DIFFERENCE operator is the difference between the rating for the map label class and the highest rating given to any possible map class (Gopal and Woodcock, 1994). For the ideal case, where the mapped class is perfectly right (score=5) and all other classes are absolutely wrong (score=1), the DIFFERENCE operator yields a value of 4. All sites

*Table 5.* Results of the MAX and RIGHT Operators for Crown Closure by SPNG Classes

| | | Expert Evaluation | | | |
|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *Matches Using MAX* | *Matches Using RIGHT* | *Increase* | *Weight* |
| S | 15 | 8 (53.3%) | 13 (86.7%) | 5 (33.3%) | 0.1170 |
| P | 49 | 34 (69.4%) | 46 (93.9%) | 12 (24.5%) | 0.3425 |
| N | 77 | 49 (63.6%) | 64 (83.1%) | 15 (19.5%) | 0.3984 |
| G | 41 | 26 (63.4%) | 30 (73.2%) | 4 (9.8%) | 0.1422 |
| Total | 182 | 117 (64.3%) | 153 (84.1%) | 36 (19.8%) | 1.0000 |
| Weighted | | (64.4%) | (85.8%) | (21.4%) | |

*Table 6.* Results of the DIFFERENCE Operator for Crown Closure by SPNG Classes

| | | *Mismatches* | | | | *Matches* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *−4* | *−3* | *−2* | *−1* | *0* | *1* | *2* | *3* | *4* | *Mean* |
| S | 15 | 2 | 0 | 4 | 1 | 5 | 1 | 2 | 0 | 0 | −0.93 |
| P | 49 | 2 | 1 | 7 | 5 | 26 | 8 | 0 | 0 | 0 | −0.45 |
| N | 77 | 7 | 6 | 9 | 6 | 34 | 7 | 8 | 0 | 0 | −0.61 |
| G | 41 | 10 | 1 | 2 | 2 | 19 | 4 | 1 | 0 | 2 | −0.83 |
| Total | 182 | 21 | 8 | 22 | 14 | 84 | 20 | 11 | 0 | 2 | |

that are matches using the MAX operator have DIFFERENCE values greater than or equal to 0 and all mismatches are negative. A DIFFERENCE operator score of −1 indicates a case where the map label received a rating one less than the highest possible rating [this magnitude of difference (−1) is not as troublesome as those where a difference of −4 is found]. The mean of the DIFFERENCE operator is reported in the tables.

Another important kind of information is the categorical nature of the errors, or between which classes confusion occurs. CONFUSION and AMBIGUITY matrices were created for each of the classes evaluated. A CONFUSION matrix indicates which categories are being confused with each other. The CONFUSION matrix for fuzzy accuracy displays classes with a rating higher than that of the map label class and is identical to a traditional confusion matrix except that more than one class can have a rating higher than the mapped class at a single site. An AMBIGUITY matrix lists classes with the same rating as that of the map label.

## RESULTS

Results are reported here for size and closure for the Modoc National Forest component of the map. For each category and area of the map that was assessed, three accuracy assessment tables (MAX and RIGHT operators, DIFFERENCE operator, and a CONFUSION and AMBIGUITY matrix) were created.

### SIZE Classes within Forested Stands

*Overall Accuracy*

Results for the accuracy assessment of size can be found in Tables 2, 3, and 4. Using the MAX operator, the accuracy is 62.1% and increases to 87.4% for the RIGHT operator (Table 2). Using the weighted operators, the accuracy decreases because approximately 53% of the map is size class 2 and the accuracy for this class is 33.7% and 77.9% for the MAX and RIGHT operators, respectively. The samples for size class do not follow the same distribution as the map due to densification of the grid inventory for undersampled forested types.

*Magnitude and Categorical Nature of the Error*

The DIFFERENCE operator (Table 3) indicates that size class 2 has the highest magnitude of error. Of the 86 sites, 18 with a map label of size class 2 have a magnitude of error of −3. This error generally corresponds with 19 sites in the CONFUSION operator (Table 4) that have a map label of size class 2 for which size class 4 received a higher rating. However, the majority of confusion is not of this magnitude and is between size class 2 and size class 3. Fifty-seven sites with a map label of size class 2 received higher ratings for size class 3 (Table 4). Size class 2 has many more errors of commission (77 instances where sites with a map label of size class 2 received a higher rating for one or more other classes) than there are of omission (12 instances). The confusion matrix is not symmetrical, and in this case it indicates that errors in the map more often are the result of a smaller size class being mapped rather than that calculated from the accuracy assessment data. Most of the confusion is between size classes 2 and 3. Only 11 sites corresponded with map labels of size class 4, so this information should be used with caution. However, Table 4 indicates there are 24 instances where size class 4 received a higher rating than that of the map label, indicating that size class 4 may be somewhat "undermapped"

*Table 7.* Results of the CONFUSION (C) and the AMBIGUITY (A) Operators for Crown Closure by SPNG Classes

| | | *S* | | *P* | | *N* | | *G* | | *No. of Mismatches* | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* |
| S | 15 | X | X | 7 | 5 | 6 | 0 | 1 | 1 | 14 | 6 |
| P | 49 | 5 | 14 | X | X | 10 | 14 | 2 | 0 | 17 | 28 |
| N | 77 | 14 | 1 | 17 | 21 | X | X | 11 | 20 | 42 | 42 |
| G | 41 | 3 | 7 | 11 | 0 | 13 | 21 | X | X | 27 | 28 |
| Total | 182 | 22 | 22 | 35 | 26 | 35 | 35 | 14 | 21 | 199 | 194 |

*Table 8.* Results of the MAX and RIGHT Operators for Crown Closure by 10% Classes

| | | *Expert Evaluation* | | | |
|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *Matches Using MAX* | *Matches Using RIGHT* | *Increase* | *Weight* |
| 1 | 15 | 8 (53.3%) | 13 (86.7%) | 5 (33.3%) | 0.1170 |
| 2 | 23 | 9 (39.1%) | 16 (69.6%) | 7 (30.4%) | 0.1676 |
| 3 | 26 | 16 (61.5%) | 23 (88.5%) | 7 (26.9%) | 0.1748 |
| 4 | 22 | 6 (27.3%) | 12 (54.6%) | 6 (27.3%) | 0.1619 |
| 5 | 32 | 14 (43.7%) | 21 (65.6%) | 7 (21.3%) | 0.1588 |
| 6 | 23 | 9 (39.1%) | 18 (78.3%) | 9 (39.1%) | 0.0777 |
| 7 | 28 | 16 (57.1%) | 19 (67.9%) | 3 (10.8%) | 0.0842 |
| 8 | 7 | 1 (14.3%) | 3 (42.9%) | 2 (28.6%) | 0.0429 |
| 9 | 6 | 1 (16.7%) | 1 (16.7%) | 0 (0.0%) | 0.0150 |
| Total | 182 | 80 (44.0%) | 126 (69.2%) | 46 (25.3%) | 1.0000 |
| Weighted | | (43.6%) | (70.4%) | (26.8%) | |

(error of omission). However, it is important to note that there is not a direct correlation between the number of errors of omission in the CONFUSION operator and the number of sites mismapped. For example, in the case of errors associated with a map label of size class 2, both size class 3 and size class 4 may have been given a better rating for the same accuracy assessment site, which would be counted as a mismatch in both size classes. Therefore, size class 4 may have been a better answer than size class 2 but still not as good an answer as size class 3.

The AMBIGUITY operator (Table 4) indicates that size class 2 and size class 4 had the same rating at 16 accuracy assessment sites. These sites most likely had a calculated size at the midpoint of size class 3, resulting in an acceptable (or equal) rating for the adjacent classes.

## Crown Closure within Forested Stands—SPNG Classes

*Overall Accuracy*

Results for crown closure using SPNG classes are given in Tables 5, 6, and 7. Using SPNG classes, accuracy is 62.5% and 84.1% for the MAX and RIGHT operators, respectively, with similar results for both weighted and unweighted methods. The S (sparse) class has the most error in the MAX classification. There is a substantial increase in accuracy for all classes when using the RIGHT operator. The CONFUSION operator substantiates that this is due to the majority of error being with "adjacent" classes.

*Magnitude and Categorical Nature of Error*

Based on frequencies in the CONFUSION operator (Table 7), 63% (63 of 100 sites—refer to the lower half of Table 7) of mismatches indicates that the map overpredicts crown closure, and 37% of mismatches indicates that the map underpredicts crown closure. For the P crown closure class, there are many more errors of omission (35) than of commission (17), and most of these errors of omission are with the N class. Most of the ambiguity occurs with adjacent classes. The DIFFERENCE operator (Table 6) indicates that the majority of confusion is in the lower −1 and −2 magnitudes.

## Crown Closure within Forested Stands—10% Classes

*Overall Accuracy*

Results for 10% closure classes can be found in Tables 8, 9, and 10. Not surprisingly, the breakdown of crown

*Table 9.* Results of the DIFFERENCE Operator for Crown Closure by 10% Classes

| | | *Mismatches* | | | | *Matches* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Map Label* | *No. of Sites* | *−4* | *−3* | *−2* | *−1* | *0* | *1* | *2* | *3* | *4* | *Mean* |
| 1 | 15 | 2 | 0 | 4 | 1 | 5 | 1 | 2 | 0 | 0 | −0.80 |
| 2 | 23 | 5 | 2 | 5 | 2 | 9 | 0 | 0 | 0 | 0 | −1.65 |
| 3 | 26 | 2 | 1 | 4 | 3 | 16 | 0 | 0 | 0 | 0 | −0.85 |
| 4 | 22 | 7 | 3 | 5 | 1 | 6 | 0 | 0 | 0 | 0 | −2.18 |
| 5 | 31 | 8 | 3 | 5 | 2 | 14 | 0 | 0 | 0 | 0 | −1.71 |
| 6 | 23 | 3 | 2 | 6 | 3 | 9 | 0 | 0 | 0 | 0 | −1.57 |
| 7 | 28 | 9 | 0 | 1 | 2 | 16 | 0 | 0 | 0 | 0 | −1.43 |
| 8 | 7 | 2 | 2 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | −2.29 |
| 9 | 6 | 5 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | −3.00 |
| Total | 182 | 43 | 13 | 32 | 14 | 77 | 1 | 2 | 0 | 0 | |

*Table 10.* Results of the CONFUSION (C) and the AMBIGUITY (A) Operators for Crown Closure by 10% Classes

| *Map Label* | *No. of Sites* | *1* | | *2* | | *3* | | *4* | | *5* | | *6* | | *7* | | *8* | | *9* | | *No. of Mismatches* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* | *C* | *A* |
| 1 | 15 | X | X | 6 | 6 | 7 | 3 | 6 | 0 | 2 | 0 | 2 | 0 | 1 | 1 | 0 | 2 | 0 | 2 | 24 | 14 |
| 2 | 23 | 3 | 11 | X | X | 10 | 4 | 11 | 1 | 9 | 0 | 7 | 0 | 2 | 3 | 1 | 4 | 0 | 5 | 28 | 28 |
| 3 | 26 | 6 | 3 | 6 | 9 | X | X | 3 | 10 | 4 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 23 | 28 |
| 4 | 22 | 9 | 3 | 10 | 6 | 10 | 8 | X | X | 5 | 6 | 6 | 5 | 6 | 4 | 4 | 4 | 2 | 5 | 52 | 41 |
| 5 | 32 | 8 | 2 | 13 | 2 | 14 | 6 | 13 | 12 | X | X | 3 | 12 | 4 | 8 | 3 | 6 | 3 | 6 | 61 | 54 |
| 6 | 23 | 2 | 1 | 4 | 0 | 5 | 0 | 5 | 2 | 5 | 4 | X | X | 7 | 11 | 7 | 5 | 5 | 3 | 43 | 26 |
| 7 | 28 | 3 | 6 | 4 | 5 | 8 | 1 | 7 | 3 | 9 | 7 | 8 | 16 | X | X | 1 | 12 | 1 | 8 | 41 | 58 |
| 8 | 7 | 0 | 2 | 0 | 2 | 2 | 0 | 4 | 0 | 6 | 0 | 6 | 0 | 6 | 1 | X | X | 0 | 2 | 24 | 7 |
| 9 | 6 | 0 | 5 | 0 | 5 | 2 | 3 | 3 | 2 | 5 | 0 | 5 | 0 | 3 | 2 | 3 | 3 | X | X | 21 | 20 |
| Total | 182 | 31 | 33 | 43 | 35 | 59 | 25 | 53 | 30 | 46 | 18 | 39 | 34 | 30 | 31 | 20 | 37 | 11 | 33 | 332 | 276 |

closure into finer classes results in a loss in overall accuracy. The accuracy for 10% closure classes is 44.0% for the MAX operator and 69.2% for the RIGHT operator (Table 8). As with the SPNG classes, overall accuracy is not appreciably different after weighting by number of acres. Classes 8 and 9 are undersampled, and additional sites for class 1 are also recommended.

*Magnitude and Categorical Nature of Error*

Based on frequencies in the CONFUSION operator (Table 10), 60% of mismatches indicates that the map overpredicts crown closure, and 40% of mismatches indicates that the map underpredicts crown closure (a similar trend to the SPNG classes). This represents a tendency for the map to overpredict crown closure. All classes have high magnitudes of error. The highest magnitudes of error based on the DIFFERENCE (Table 9) and CONFUSION operator (Table 10) occur within class 2 and class 7. With respect to sites mapped as class 2, there are a significant number of higher ratings given for classes 3, 4, 5, and 6. This indicates areas mapped as class 2 are underestimating crown closure. However, areas mapped as class 7 received a significant number of higher ratings for classes 6, 5, 4, and 3, indicating that the map is overestimating crown closure in this class. Areas mapped as class 5 also appear to be overestimating crown closure.

## DISCUSSION AND CONCLUSIONS

In general, the accuracy of crown closure classes was poorer than that of other mapped categories. Biging et al. (1991) also found that accuracy of crown closure was lower than other features. Interestingly, accuracy assessment figures indicate the canopy model used for mapping crown closure may be overpredicting crown closure in some cases. This trend warrants further investigation since such a trend (if it is substantiated) may represent a potential to "correct" a significant amount of error by recalibrating the canopy model. In this analysis, crown closure was evaluated by comparing the map label of crown closure to the crown closure calculated from the inventory tree list. This type of mensurational approach is not often done in an accuracy assessment. The lower accuracies for crown closure classes are most likely due to problems with the canopy model, but also may be a function of error associated with equations used to calculate crown closure from the inventory data. In addition, for some stands with greater variability, five subplots within a plot may be too few to calculate a crown closure estimate with an acceptable standard deviation.

"Expert evaluation" for size classes was also calculated from the inventory data using a mensurational approach. The accuracy of size classes was generally high, ranging from 87% to 97% for the RIGHT operator. This mensurational approach of calculating size from inventory data was effective and represents a reduction in time and cost if the data can be collected for multiple purposes. In future work, other methods for assigning fuzzy ratings to crown closure and size, such as based on sampling error, could be considered.

As mentioned throughout this paper, a number of classes in the mapped categories are undersampled and figures for these should be used with caution or not used at all. The problems of not having enough sites in some categories when using simple random sampling, systematic, or any other equal probability sampling designs as reported by Congalton (1991) were sometimes apparent in this project. However, as a function of the FIA GRID inventory design, categories that represent the majority of acreage in the map tend to be adequately sampled and overall accuracy figures are quite useful to both the users and producers of the map. Some grid plot locations have been densified for rare forest types. In future studies, it would be possible to determine if this densification influences the reported overall map accuracy.

It is important to note that the total mapped area comprises nearly nine million acres and assessing the accuracy of an area this large is typically very expensive. A total of 1,171 five-point cluster plots were used in this

entire study (307 plots on the Modoc National Forest). With the exception of some of the PNW plots that were not ground visited, these are comprehensive inventory plots with an overall value of approximately $691,000. Using an independent data set such as this (installed for other purposes) affords a very cost effective means of generating good accuracy assessment data. In this project we utilized FIA data for this purpose, but other data sets that collect similar information on plot locations, tree sizes, and species and that are collected using a probability sample could be employed.

*Thanks to the California Department of Forestry and Fire Protection for joint funding of this effort. Thanks to Boston University for providing programming support. Thank Drs. Greg Biging and Steve Stehman for reviews of earlier versions of this work. Also, thanks to the anonymous reviewers for their considered reviews.*

## REFERENCES

Biging, G. S., Congalton, R. G., and Murphy, E. C. (1991), A comparison of photointerpretation and ground measurements of forest structure. In *Proceedings of the 57th Annual Meeting of the American Society of Photogrammetry and Remote Sensing*, Baltimore, MD, March 1991, Vol. 3 (Remote Sensing), pp. 6–15.

Congalton, R. G. (1991), A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 37:35–46.

Gopal, S., and Woodcock, C. E. (1994), Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogramm. Eng. Remote Sens.* 60(2):181–188.

Janssen, L. L. F., and van der Wel, F. J. M. (1994), Accuracy assessment of satellite derived land-cover data: A review. *Photogramm. Eng. Remote Sens.* 60(4):419–426.

Miller S., Eng, H., Byrne, M., Milliken, J., and Rosenberg, M. (1994), Northeastern California vegetation mapping: A joint agency effort. Remote Sensing and Ecosystem Management. In *Proceedings of the Fifth Forest Service Remote Sensing Applications Conference*, April, 1994, ASPRS, Bethesda, MD, pp. 115–125.

Milliken, J., Beardsley, D., and Gill, S. (1998), Accuracy assessment of a vegetation map of northeastern California using permanent plots and fuzzy sets. Natural Resources Management Using Remote Sensing. In *Proceedings of the Seventh Forest Service Remote Sensing Applications Conference*, April 1998, ASPRS, Bethesda, MD, pp. 218–229.

Muller, S. V., Walker, D. A., Nelson, F. E., Auerbach, N. A., Bockheim, J. G., Guyer, S., and Sherba, D. (1998), Accuracy assessment of a land-cover map of the Kuparuk River Basin, Alaska: Considerations for remote regions. *Photogramm. Eng. Remote Sens.* 64(6):619–628.

Story, M., and Congalton, R. G. (1986), Accuracy assessment: A user's perspective. *Photogramm. Eng. Remote Sens.* 52(3):397–399.

U.S. Forest Service Regional Ecology Group (1981), *CALVEG: A Classification of California Vegetation*, U.S.D.A. Forest Service, Region 5, San Francisco, CA.

U.S. Forest Service—Region 5 (1995), *Forest Inventory and Analysis User's Guide*, U.S.D.A. Forest Service, Region 5, San Francisco, CA.

Warbington, R. J., and Levitan, J. (1992), How to estimate canopy crown closure using maximum crown width/DBH relationships. In *Stand Inventory Technologies 1992: Proceedings of Conference* (G. Lund, E. Landis, T. Atterbury, Eds.), 1992 September 13–17, Portland, OR, The World Forestry Center, pp. 319–328.

Woodcock, C. E., and Harward, J. (1992), Nested-hierarchical scene model and image segmentation. *Int. J. Remote Sens.* 13(16):3167–3187.

Zadeh, L. (1963), Outline of a new approach to the analysis of complex or imprecise concepts. *IEEE Transactions: Systems, Man, and Cybernetics, SMC* 3:28–44.

Zadeh, L. (1965), Fuzzy sets. *Information and Control* 8: 328–353.