

A TIMER INVENTORY BASED UPON MANUAL AND AUTOMATED ANALYSIS OF ERTS-1 AND SUPPORTING AIRCRAFT DATA USING MULTISTAGE PROBABILITY SAMPLING

James D. Nichols, Mike Gialdini and Sipi Jaakkola, *Center for Remote Sensing Research, 260 Space Sciences Laboratory, University of California, Berkeley, California*

ABSTRACT

A quasi-operational study demonstrating that a timber inventory based on manual and automated analysis of ERTS-1, supporting aircraft data and ground data was made using multistage sampling techniques.

The inventory proved to be a timely, cost-effective alternative to conventional timber inventory techniques. The timber volume on the Quincy Ranger District of the Plumas National Forest was estimated to be 2.44 billion board feet with a sampling error of 8.2 percent. Costs per acre for the inventory procedure at 1.1 cent/acre compared favorably with the costs of a conventional inventory at 25 cents/acre. A point-by-point comparison of CALSCAN-classified ERTS data with human-interpreted low altitude photo plots indicated no significant differences in the overall classification accuracies.

SECTION I: INTRODUCTION AND PROCEDURES

In order to test the usefulness of ERTS imagery for wildland resource inventories, a timber inventory was performed in which the ERTS imagery acted as the first stage of a multistage sampling design. The objective of the inventory was to estimate the standing volume of merchantable timber within the Quincy Ranger District (215,000 acres) of the Plumas National Forest in California. Secondary objectives of the inventory were: (1) to test the operational efficiency of the sampling procedures of the multistage sampling design; (2) to test the effectiveness of the CALSCAN classifier on the ERTS data; (3) to determine the value of ERTS data and the aircraft data in reducing the sample error; and (4) to compare the costs of this timber inventory with other inventories that utilize conventional procedures.

A three stage sampling design was tested in which "timber volume" was the variable estimated. At each stage timber volume estimates were made from sampling units whose probability of selection in the sample was proportional to the predicted volume. Timber volume estimates were made from three stages: (1) the first stage involved automatic classification of the timberland on the ERTS data tapes into four timber volume classes. Within the classified area subsamples were selected (called primary sampling units; PSU) from which a more refined estimate

FN 74 30717

of timber volume could be made in the second stage; (2) the second stage involved the acquisition of low altitude photograph of selected primary sampling units to choose photo plots based on a second timber estimate made by comparison with photo-volume tables; (3) the third stage involved selecting individual trees within selected sample photo plots by photo measurement of all merchantable trees. The selected trees were then precisely measured for volume on the ground and these volume measurements in turn were expanded through the various stages of the sample design to estimate total timber volume over the national forest land within the Quincy Ranger District.

The statistical procedures for expanding the timber volume estimates through the various stages of the timber inventory are discussed in Section II of this study. The procedures for selecting sampling sites and estimating timber volumes from ERTS and aircraft imagery at each of the stages of the sample design are described in subsequent sections in order to demonstrate how the ERTS and supporting aircraft data were necessary components in performing the timber inventory.

Stage I. CALSCAN Classification of ERTS Data and Primary Sample Unit Selection

ERTS-1 data tapes of the Quincy Ranger District, Plumas National Forest, were classified on the CRSR interactive human-computer system using a CALSCAN point-by-point classification routine. The coordinates of the Ranger District boundary and those of non-national forest land within the District were identified on the tapes so that only those picture elements associated with national forest land were classified and incorporated into the inventory. This procedure considerably reduces the costs of classification.

The classification was based on four timber volume classes, namely (1) non forest; (2) forest sites containing less than 10,000 Bd. ft./ac.; (3) forest sites containing 10,000-20,000 Bd. ft./ac.; and (4) forest sites containing more than 20,000 Bd. ft./ac. The classifier was trained to recognize each of the four timber volume classes based upon photo-interpretation selection of 33 training cells whose range of timber volumes was based on crown closure and average crown diameter. The training cells were selected from interpretation of high altitude color-infrared photography (scale 1/120,000). Each of the training cells was located on the ERTS imagery and digitized. Point-by-point classification of all ERTS data points within the Quincy Ranger District proceeded by matching each data point (picture element) with the training cells. The results were grouped into the four timber volume classes. The accuracy achieved by automatic classification of ERTS tapes (Tables 1 and 2) demonstrates the efficiency which can be attained in the timber inventory through analysis of ERTS data tapes in the initial stages of the sample design.

The classified ERTS data (CALSCAN classification) of the Quincy Ranger District was divided into rectangular sampling units (called primary sampling units). Each unit measured 1325 ft. wide by one and a half miles long. The size of these sampling units was based upon a practical area which could be photographed in a single flight line by a light aircraft using a 35 mm camera system, the ability of the ground crew to complete the ground work for a flight line in one day, and the variation between sampling units. For each primary sampling unit, the following information was computed.

- (1) The number of points in each volume class (within the unit).
- (2) The weighted total volume for each volume class.*
- (3) The sum of the weighted totals for all classes.
- (4) A cumulative sum of the weighted totals.
- (5) The mean volume for all sampling units.
- (6) The variance of the sampling units.

Based upon the information either estimated or computed for each primary sampling unit (PSU), four units were selected for further sampling in the timber inventory. The four units were selected with probability of selection proportional to their estimated volumes. The location of the four selected PSU's was transferred from the ERTS classified images to the color-infrared high altitude aerial photography (scale 1/120,000) to facilitate locating them accurately from the air when they are photographed from a lower altitude as part of the second stage of the timber inventory.

Stage II. Volume Estimation on Low Altitude Photography

Two 35 mm cameras were used to obtain low altitude photography of the selected primary sampling units at two different scales. A 24 mm focal length wide-angle lens was used to acquire complete coverage of each sampling unit at an approximate scale of 1/7500, and a 200 mm focal length lens obtained large scale stereo triplicates, scale approximately 1/1000, from which to make precise photo estimates of timber volume. The camera with the telephoto lens was equipped with a motorized film drive which enabled each stereo triplicate to be taken within one second at five second intervals, while the camera with the wide-angle lens was operated manually to obtain single frames at five second intervals. The photo coverage for each PSU consisted of 10 stereo triplicates and 10 wide-angle photographs.

The wide-angle photos of each primary sampling unit were formed into a mosaic to show its full area. The center of the middle photo for each stereo triplicate was used as a plot center and was located and marked on the mosaic. Each plot center was also located on a topographic map and its elevation was determined.

*The weighted total volume was determined by multiplying the number of points in each volume by the assigned weight for that class:
non-timber = 0; less than 10,000 Bd. ft./ac. = 1; 10,000-20,000 Bd. ft./ac. = 2; 20,000 Bd. ft./ac. and above = 3.

Discriminant Analysis

		NF	T1	T2	T3
Photo Interpretation on 1:1000 color photography	NF	90.0	5	0	5
	T1	0	93	7	0
	T2	7	19	67	7
	T3	0	0	25	75
	% change in class	+5%	+20	-7%	-8

NF = Non forest

T1 = Less than 10,000
board feet per acre

T2 = 10,000 to 20,000
board feet per acre

T3 = Greater than 20,000
board feet per acre

Table 1

The accuracy obtained in a comparison of classified ERTS data with photo interpreter classification from a photo scale of 1/1000. The diagonal of the matrix shows the percentage correct classification when comparing the discriminate analysis of ERTS data with the large scale photographs. The values not along the diagonal represent the percentage of points which were classified into adjacent timber volume classes. The values at the bottom of each column indicate the percentage of points classified by discriminate analysis of ERTS data which were greater than (+) or less than (-) the points classified by interpreters on large scale aerial photographs.

Discriminant Analysis of ERTS-1 data

	Coni	Soil/R	BR	H.W.	Wat.	Conv.	Grass
Coni	95.7	.9	0	3.8	0	0	.9
Soil/R	0	71.8	0	14.1	0	7.0	7.0
BR	14.3	0	48	13.0	0	1.3	23.4
H.W.	13.2	.9	1.9	74.5	0	3.8	5.7
Wat.	0	0	0	0	100	0	0
Conv.	12.5	12.5	0	43.8	0	31.25	0
Grass	15.6	9.4	3.1	6.35	0	6.3	59.4
% change in class	+56	-16.9	-48.1	+13.2	0	+6.25	+594

Photo Interpretation on 1:120,000
CIR and ground data

CONI = Commercial conifer consisting of pure fir and mixed conifer stands.

SOIL/R = Soil and Rock.

BR = Brush lands made up several species in mixed and pure stands.

HW = Hardwoods included Oak and riparian.

WAT = Water.

CONV = Brush lands converted to conifer plantations.

GRASS = Grass lands included dry and wet meadows.

TABLE 2. The results obtained in comparing classified ERTS data with photo-interpretation-classification of land vegetation categories on small scale aerial photographs (1/120,000). The diagonal of the matrix shows the percentage correct classifications when comparing the discriminant analysis of ERTS data with the training cell classes delineated on the small scale photograph. The values not along the diagonal represent the percentage of points which were classified into adjacent categories. The values at the bottom of the columns indicates the percentage of points classified by discriminant analysis which are greater than (+) or less than (-) the points associated with the training cells.

The scale of each photo plot was determined, and a .4 acre circular plot was drawn about the photo plot center. Timber volume in each .4 acre photo plot was estimated by referring to photo-volume tables based upon interpretation of % crown closure and measurement of average stand height using a parallax bar (Chapman, 1965). Within each primary sampling unit, two out of the ten possible photo plots were chosen with probability of selection proportional to the estimated volume.

Stage III. Selection of Trees for Precise Ground Measurement of Timber Volume

In the third stage, all trees of merchantable size within each selected photo plot were pin-pricked and numbered. For each of these trees, the average crown diameter was determined based on the longest and shortest dimension of their crowns. After adjustments for scale, the average crown diameter value was cubed (raised to the third power) to be used as a relative measure of individual tree volume for the third stage volume estimation. Four trees were selected from the population of merchantable trees found within each photo plot to be measured by a dendrometer on the ground. (Selection again was based upon probability proportional to the estimated volume of each tree.)

A two-man crew went into the Quincy Ranger District with a Barr and Stroud optical dendrometer to measure the selected trees. The large scale (low altitude) photographs were used to locate the photo plot centers as well as the trees within the plots to be measured. In addition to the dendrometer measurements, an easily recognizable feature on the ground near the plot center was measured in order to get a more accurate estimation of the photo scale of the plot. The dendrometer measurements were brought back from the Quincy Ranger District and entered into a computer program that calculated volumes for the individual trees. The tree volumes were then expanded through each stage of the sample design to estimate total volume on the District, consistent with the statistical methods for variable probability sampling (see Section II).

SECTION II. STATISTICAL METHODS FOR TIMBER VOLUME ESTIMATION

Timber volume predictions were made from three stages of the timber inventory of the Quincy Ranger District for the purpose of selecting sample plots whose probability of selection was proportional to the volume predictions. Thus, variable probability sampling methods were used to estimate the total volume in this timber inventory. Three variables proportional to timber volume were used in generating the selection probabilities: (1) "Volume" estimate of the ERTS picture element based on the spectral signatures on four bands and subsequent training and classification; (2) volume estimate of a plot on a 1:1000 scale color print based on the photo-volume tables (Chapman, 1965); and (3) volume estimate on a large scale photo based on a rough volume estimate - crown diameter cubed.

Using a scheme where probability of selection is proportional to the estimated volume, the effort is focused on the areas of higher timber volume and adds to the overall cost-efficiency. The ability to list the populations at each stage prompted the selection of list sampling as the variable probability sampling scheme.

Method of Estimation

The method of estimation was based on "unequal expansion" as implied by the probability scheme discussed above. At each of the three stages, the probability proportional to estimated size (p_i) was obtained by listing the volume estimates of the sampling units (x_i), and dividing them by the total of volume estimates ($\sum_{i=1}^n x_i$):

$$P_i = \frac{x_i}{\sum_{i=1}^n x_i}$$

A sample of a chosen size was then drawn by applying random integers from 1 to $\sum_{i=1}^n x_i$ and observing the probability interval and the corresponding $\sum_{i=1}^n$ sampling unit which contains the randomly selected integers.

In the first and second stages the timber volumes of the selected sampling units were estimated by subsequent sampling, whereas in the third stage the volume was carefully measured by a precision dendrometer. The entire three stage estimation procedure was as follows:

Stage I: A sample of n_h out of the N_h PSU's was drawn from stratum h with probability proportional to estimated size (ppes). The estimate of the total volume then becomes:

$$\hat{V} = \sum_{h=1}^L \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{Y_{hi}}{P_{hi}}$$

where: L = total number of strata

P_{hi} = selection of probability of the i th PSU in the h th stratum

\hat{y}_{hi} = total volume of the ith PSU in the hth stratum (remains to be estimated by subsequent stages).

Stage II: To estimate the total volume (y_{hi}) of the ith PSU, a sample of n_{hi} out of the N_{hi} secondary sampling units (.4 acre plots) is drawn with probability proportional to estimated volume. This would give:

$$y_{hi} = \frac{1}{n_{hi}} \sum_{j=1}^{n_{hi}} \frac{y_{hij}}{p_{hij}}$$

However, in order to include area expansion from circular sample plots to the full PSU, plus stratify the second stage plots into four volume strata, the estimator becomes:

$$\hat{y}_{hir} = \sum_{r=1}^R \frac{1}{p_{hir}} \frac{A_{hir}}{a_{hir}} \frac{1}{n_{hir}} \sum_{j=1}^{n_{hir}} \frac{\hat{y}_{hirj}}{p_{hirj}}$$

where: $r = 1, 2, \dots, R$ refers to the CALSCAN volume strata

p_{hir} = selection probability of the rth volume stratum of the ith PSU in the hth stratum

A = area (indices as above)

a = sample area (indices as above)

n = sample size (indices as above)

p_{hirj} = selection probability of the jth plot of the rth volume stratum, of the ith PSU in the hth stratum

\hat{y}_{hirj} = plot volume (to be estimated by Stage III)

Stage III: To estimate the total volume of the jth plot, a sample of n_{hirj} out of the N_{hirj} tertiary sampling units (trees) is drawn with pps. Then ...

$$\hat{y}_{hirj} = \frac{1}{n_{hirj}} \sum_{k=1}^{n_{hirj}} \frac{Y_{hirjk}}{P_{hirjk}}$$

where: P_{hirjk} = the selection probability of the k th sample tree of the j th plot of the r th volume stratum of the i th PSU of the h th stratum.

Y_{hirjk} = the dendrometer measured volume of the k th sample tree of the j th plot of the r th volume stratum of the i th PSU of the h th stratum.

Combining the various stages above, the entire estimator becomes:

$$v = \sum_{h=1}^L \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{r=1}^R \frac{1}{P_{hir}} \frac{A_{hir}}{a_{hir}} \frac{1}{n_{hir}} \sum_{j=1}^{n_{hir}} \frac{1}{P_{hirj}} \frac{1}{n_{hirj}} \sum_{k=1}^{n_{hirj}} \frac{y_{hirjk}}{P_{hirjk}}$$

Variance of the Estimator

In multistage sampling, when the number of first stage units is large, most of the variability in the population is due to the first stage. Therefore, it suffices to consider only the first stage values (here \hat{y}_{hi}) to estimate the population variance and, consequently, the variance of the estimator (Durbin, 1953, p. 262; Kendall and Stuart, 1967, vol. 3, p. 200; Langley, 1971, p. 131).

Thus for the first stage our stratified sampling estimator becomes (Cochran, 1963, p. 260):

$$\hat{v} = \sum_{h=1}^L \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{y_{hi}}{P_{hi}}$$

Its variance is:

$$\text{Var}(\hat{v}) = \sum_{h=1}^L \frac{1}{n_h} \sum_{i=1}^{N_h} P_{hi} \left(\frac{y_{hi}}{P_{hi}} - v_h \right)^2$$

which has an unbiased estimator

$$\text{Var } (\hat{V}) = \sum_{h=1}^L \frac{1}{n_h (n_h - 1)} \sum_{i=1}^{n_h} \left(\frac{y_{hi}}{P_{hi}} - \hat{V}_h \right)^2$$

For proportional allocation, $n_h = n (N_h/N)$ and

$$\text{Var } (\hat{V}) = \sum_{h=1}^L \frac{N}{n N_h} \sum_{i=1}^{N_h} P_{hi} \left(\frac{y_{hi}}{P_{hi}} - V \right)^2$$

$$\text{Var } (\hat{V}) = \sum_{h=1}^L \frac{N^2}{n N_h (n N_h - 1)} \sum_{i=1}^{n_h} \left(\frac{y_{hi}}{P_{hi}} - \hat{V}_h \right)^2$$

The last equation is an unbiased estimator of $\text{Var } (\hat{V})$ and can be used for the estimation of the sampling error of the inventory.

SECTION III. RESULTS AND CONCLUSIONS

The total volume of timber in the Quincy Ranger District of the Plumas National Forest was estimated to be 407 million cubic feet (approximately 2.44 billion board feet) based on eight selected photo plots located within four primary sampling units. The sampling error associated with this estimate was 8.2% which falls below the expected sampling error of 20% for the Quincy Ranger District. This indicates that the true volume of merchantable trees in the Quincy Ranger District will fall into the interval 2.11 - 2.77 billion board feet with 80% probability.

There were only 31 trees total measured by an optical dendrometer on the ground at the eight plots (32 trees should have been measured but one plot out of the eight contained only three merchantable trees which could be measured). The field work required one week's time from a two-man crew, and the total area of the ground plots measured was 3.2 acres, representing a sampling fraction of approximately 1/67,000.

Table 1 lists the costs of the timber inventory on 215,000 acres using the multistage sampling design as an operational system. The expected costs of an inventory using the above procedures, on the entire Plumas National Forest (1,161,554 acres) would be approximately \$15,000 and would take 5 months to complete. In comparison, the U. S. Forest Service in 1970 completed an inventory of the Plumas National Forest using the 10 point system, a conventional procedure, at a cost

of \$300,000 and took two years to complete. Some deficiencies of that inventory were: (1) that it did not provide in-place mapping of timber volumes and (2) it used interpolation to arrive at the volumes for each ranger district which tends to reduce the accuracy of the volume estimates for each district, as compared to the estimate for the entire forest. The CRSR system would inventory each ranger district as a separate unit, thereby achieving more precise estimates and could also provide in-place mapping of volume.

The Forest Service inventoried the entire El Dorado National Forest (858,496 acres) using WRIS (Wildland Resource Inventory System) at a cost of approximately \$31,300. While WRIS closely resembles the CRSR system in design, the results again lacked in-place mapping of volume as well as not treating each ranger district as a stratum. In terms of inventory costs per acre for each system, the CRSR system appears to be best (10-point system = 25¢/acre, WRIS = 3.6¢/acre, and CRSR = 1.1¢/acre) while still providing more information with little or no loss in precision. It should also be noted that the multistage design used here becomes more efficient as the total area considered increases. For example, if the entire State of California were inventoried, the per acre costs would be .175¢.

TABLE 3
 COST ESTIMATE-TIMBER INVENTORY
 BASED ON 215,000 ACRES - QUINCY RANGER DISTRICT

Computer Processing

CALSCAN Training and Photo Interpretation	\$ 120	
CALSCAN Classification	210	
Statistical Break-up	40	
CALSCAN Statistics	<u>12</u>	
		\$ 382

Aerial Photography

Aircraft and Crew	\$ 210	
Film and Processing	<u>90</u>	
		\$ 300

Supplies and Expenses

Travel	\$ 350	
Miscellaneous	<u>50</u>	
		\$ 400

Personnel

1 Project Scientist 2 mos. F.T.E. @ \$1100/mo.		\$2,200
1 Software Consultant .5 mo. F.T.E. @ \$1300/mo.		550
1 Statistician .5 mo. F.T.E. @ \$1100/mo.		550
2 Lab Assistants 1 mo. F.T.E. Each @ \$620/mo.		<u>1,240</u>
Salary Subtotal		\$4,640

Associated Overhead

\$968

\$5,568

TOTAL \$6,640

SUMMARY

The preliminary results of the timber inventory of the Quincy Ranger District indicate that the procedures employed in the multi-stage sampling design are valid and substantially reduce both the costs and the amount of time required to perform a timber inventory for a large area. This study demonstrates the value of ERTS data for accurately correlating picture elements with timber volume estimates as a fundamental first step in selecting primary sampling units in the first stage of the inventory. The inventory procedures utilized here will be applied to two additional districts of the Plumas National Forest in an effort to estimate the total timber volume on the forest to a desired sampling error of 10%.

Future activities will include the investigation of the ability of the CALSCAN classifier to correlate timber condition classes with picture elements to provide information for use in forest simulators. This would enable forest managers to project long-range results of management practices to determine their effectiveness. Also, similar inventory procedures will be developed for application to other vegetation types to provide timely, cost-effective, and useful information for resource management planning, such as fuel mapping for fire control and management and rangeland-condition mapping for carrying-capacity determinations.