1996

Estimation of forest stand structure attributes from aerial photographs| An accuracy assessment

Charlie C. Patton
The University of Montana

Maureen and Mike
MANSFIELD LIBRARY

The University of **MONTANA**

Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

*** Please check "Yes" or "No" and provide signature ***

Yes, I grant permission X

No, I do not grant permission ____

Author's Signature ____

Date 16 MAY 1996

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.

ESTIMATION OF FOREST STAND STRUCTURE ATTRIBUTES FROM AERIAL PHOTOGRAPHS: AN ACCURACY ASSESSMENT

by

Charlie C. Patton

B.A., Lewis and Clark College, 1985

Presented in partial fulfillment of

the requirements for the degree of

Master of Science in Resource Conservation

University of Montana

1996

Approved by:

Robert D. Pfister

Chairperson

[signature]

Dean, Graduate School

5-16-96

Date

Patton, Charlie C., M.S. May 1996 Resource Conservation

Estimation of Forest Stand Structure Attributes from Aerial Photographs: An Accuracy Assessment. (102 pp.)

Director: Robert D. Pfister

Abstract

Attributes of forest stand structure were estimated by two different techniques and accuracy assessments conducted by comparing the estimates to objective plot data. Over 500 stands were photo-interpreted (1:15,840 nominal scale) for the following structure attributes: DBH, crown diameter, height, canopy cover, canopy layers, cover type, and spatial aggregation (a measure of the horizontal distribution of forest vegetation). Estimates for the attributes were also obtained using the field relevé plot method. Systematic, unbiased data for each stand were obtained by averaging data from 5 plots within each stand; these average values were considered 'ground truth'. The estimates (photo-interpreted and relevé plot) were compared to ground truth data for 50 sample stands using error matrix tables and the chi-square test of a hypothesized variance (Freese, 1960). Percentage accuracy adjusted for chance agreement (Tau coefficient, Ma and Redmond, 1995) ranged from 0.63 to 0.29. The chi-square coefficients ranged from 0.77 to 0.52. The relevé estimates were generally more accurate than the photo estimates for most variables. No obvious sources of error were detected in the photo-interpretation methods. Acceptable accuracy levels will vary depending on project scale and objectives. Land managers will continue to look to remote sensing technology as a cost effective way to obtain data as they assess forest resources and processes across larger landscapes. Awareness of the limitations of this technology and of potential inaccuracies in these data are critical factors to consider when decisions are to be made based on a landscape scale analysis

Abstract ii

Acknowledgments v

List of Tables vi

List of Figures vii

Chapter I: Introduction/Statement of problem 1

A. Literature review 2

Chapter II: Methods/Pilot Study 8

A. Study area 8

B. Pilot study 10

1. Objectives 10

2. Methods 10

2. Analysis 14

4. Discussion 23

5. Conclusions 24

C. Landscape study 26

1. Photo interpretation 26

2. Sampling design 30

3. Field methods 31

4. Data compilation and analysis methods 35

Chapter III: Results 40

A. Photo-interpretation error matrices 40

B. Error analysis 53

C. Summary of accuracy measures 53

Chapter IV: Discussion 59

Chapter V: Summary/Conclusions **71**

Literature cited **73**

Appendices **78**

Appendix A: Pilot study plot error matrices 78
Appendix B: Estimation templates 87
Appendix C: Statistical methods 88
Appendix D: Relevé plot error matrices 91
Appendix E: Scatterplots of estimated vs ground truth data 97
Appendix F: Coefficient of variation analysis 102

Acknowledgments

I extend my thanks to my committee members, Dr. Richard Hutto, Dr. Kevin O'Hara, and Dr. Lloyd Queen, for their valuable input during this project. A special note of thanks goes out to Dr. Robert Pfister, my major advisor. I was very fortunate in being able to work closely with Bob during the design and implementation of this project; the knowledge that I gained from this association extends well beyond the pages of this thesis.

This research was supported by a Federal McIntire-Stennis Grant through the Montana Forest and Conservation Experiment Station, and the Confederated Salish and Kootenai Tribes of the Flathead Reservation who provided access to the study area, associated data, and numerous consultations. The personnel at the USDA Forest Service R-1 Insect and Disease Lab generously provided access to their zoom stereroscope.

List of Tables

Table		Page
1	Initial pilot study variables	11
2	Photo-interpreted pilot study variables	12
3	Number of trees used in the DBH regression equations	15
4	Adjusted R^2 values for the regression equations	16
5	Regression equations used to predict DBH	17
6	Example error matrix and calculations	18
7	Allowable error for chi-square analysis technique	20
8	Accuracy coefficients for the pilot study	22
9	Classes for error matrix tables	36
10	Hypothetical error matrix table	37
11	Photo-interpreted DBH (modeled) error matrix	43
12	Photo-interpreted DBH (ocular estimate) error matrix	46
13	Photo-interpreted tree height error matrix	47
14	Photo-interpreted crown diameter error matrix	48
15	Photo-interpreted canopy cover error matrix	49
16	Photo-interpreted canopy layers error matrix	50
17	Photo-interpreted cover type error matrix	51
18	Photo-interpreted spatial aggregation error matrix	52
19	Accuracy measures summary table	58

List of Figures

Fig.		Page
1	Map locating the Finley Creek study area	9
2	Spatial aggregation stand types	28
3	Stand map with field inventory plots...	32
4	Relationship between the datasets...	36
5	Scatterplot error analysis..	39
6	Map of sample stands	41
7	Accuracy measure comparison scatterplots...............................	56
8	Histogram of the ground truth canopy cover data	63

Chapter I: Introduction/Statement of problem

Defining relatively homogeneous patches (stands) across forested landscapes is critical in assessing wildlife habitat, fire risk, forest health, and resource outputs for project planning. The issue of how these patches are defined will take on greater importance as landscape pattern and process models are refined and applied. Ground-based survey methods for defining forest stand structure are prohibitive in time and cost at the landscape scale. Efficient and accurate patch characterization from remotely sensed images will be critical to landscape level studies.

The history of aerial photograph interpretation of forest stands dates back to 1887 (Spurr, 1960), and it remains an important part of resource planning today (Lillesand and Keifer, 1994). The primary focus of interpretation efforts has been to delineate and characterize stands using criteria important for timber production.

Ecologically based stand characterization criteria will gain importance with the current emphasis on ecosystem management. This emphasis involves making resource management decisions considering smaller spatial scales than in the past; remote sensing will play a critical role in assessing these landscapes. It will be important for managers and researchers to know the accuracy of data collected from remotely sensed images. Accuracy assessments have frequently been conducted for maps based on satellite imagery, but not for most photo-interpreted datasets.

The objective of this study was to determine the accuracy and methodology of obtaining forest stand structure variables from 1:15,840 nominal scale standard color aerial photographs. Additionally, the relevé concept of vegetation sampling (Mueller and Dombois-Ellenberg, 1974) was tested. This concept involves selecting a single plot that represents the average conditions observed in the vegetation community (stand), and measuring this single plot to obtain a relatively efficient and accurate view of the stand.

A pilot study was conducted for the investigator to gain familiarity with the study area, to explore the use of a forest inventory database in defining forest stand structure, and to define photo-interpretation and field methodologies. Techniques developed in the pilot study were used in a landscape assessment of the entire study area.

A. Literature review

Aerial photograph interpretation techniques have been and continue to be an important tool in forest resource management. Common applications include estimation of timber volume, cover type mapping, and insect and disease damage estimation (Lillesand and Keifer, 1994; Hall et al, 1992; Balice, 1979; Spurr, 1960; Nyyssonen, 1957; Moessner and Jensen, 1951). Reliable timber volume estimates from aerial photographs require accurate estimations of stand height, crown diameter, and stocking (usually canopy cover); these are often the independent variables in timber volume models (Lillesand and Keifer, 1994). The accuracy of estimating or measuring these variables from aerial photographs of various scales has been reported (e.g. Paine, 1981; Spurr, 1960; Worley and Meyer, 1955).

Some authors state that photo measurements of structure attributes cannot be directly related to ground measurements. Worley and Meyer (1955) report that photo measurements of crown diameter and cover in hardwood stands cannot be reliably checked by ground measurements because of the complexities of multiple stem and interlocking crowns, and shadow variations. Paine (1981) also determined that it is difficult to make a direct comparison of field and photo measurements of crown diameter and crown cover because in the field it is not possible to note which portion of the crowns or openings in the stand are visible to the interpreter on the photos. Thin and narrow branches in the tree crowns may be visible from the ground but not visible on the photos, leading to an underestimation of crown diameter and canopy cover from the photos. Because of these discrepancies, only the relative accuracy (accuracy among and within observers) can be investigated (Worley and Meyer, 1955). The accuracy of crown diameter estimates cannot be investigated by comparison to ground-measured values, rather the "consistency in differences" between the two can be explored (Paine, 1981). The discrepancies pointed out by these authors undoubtedly do exist, but they do not invalidate an accuracy test between photo and ground measurements of these variables. The value of an accuracy assessment lies not only in reporting the ability of the technique to estimate field-determined values, but also in reporting the nature of the errors associated with the technique. The differences described above (Paine, 1981; Worley and Meyer, 1955) may result in the photo-interpretation technique being biased; knowledge of this type of error can be extremely valuable. The problems associated with interlocking crowns and multiple stems described by Worley and Meyer (1955) may not be as critical in coniferous stands where individual crowns are more distinct and multiple stems the exception rather than the norm. Spurr (1960) reports that because of the differences described above, aerial photo estimates of crown diameter and canopy cover will tend to be less than ground measurements.

Worley and Meyer (1955) found that tree crowns can be measured to within 3 - 4 feet on 1:12,000 scale photos and 5 - 10 feet on 1:20,000 scale photos using a dot transparency. Spurr (1960) reports that tree crowns can be consistently classified into 0.9 m (3-ft) diameter classes from 1:15,840 scale photos.

Shadows, resolution and scale can lead to an overestimation of crown cover from photos, but this may be compensated by the tendency to underestimate crown cover from the ground (Spurr, 1960). Several tests have shown standard errors of crown cover estimates by independent observers of around 10% (Spurr, 1960).

Unlike crown diameter and cover, tree height can be objectively measured on the ground and from aerial photos, so accuracy relative to ground measurements can be investigated (Worley and Meyer, 1955). Worley and Landis (1954) report standard errors of

8 - 10 feet in tree height measurements using a parallax bar on 1:12,000 scale photos; the standard error varied among trees of different height classes. Similar standard errors (5 feet) are reported by Spurr (1960) for parallax measurements from 1:15,840 photos. Factors affecting tree height measurement from photos include image sharpness, scale, forest structure, shape of tree crowns, topography, and the skill of the observer. Tree species with narrow crowns may be underestimated because the tip of the crown may not be resolved on the photo (Spurr, 1960; Paine, 1981). Errors in tree height estimation are not consistent among trees of different heights; shorter trees tended to be underestimated and taller trees overestimated (Maclean and Pope, 1961).

Linear regression has been used to predict tree diameter at breast height (DBH) using photo estimated tree heights and crown measurements as dependent variables; this approach involves the testing of a model as well as the accuracy of the photo estimated variables (Hall, et al 1992; Hall, et al 1989; Hagan and Smith, 1986).

A number of studies have related photo-derived estimates of forest site characteristics to ground estimates of these same characteristics.

Larson and others (1971) measured average height and crown cover from aerial photographs and used tables developed by Moessner (1963, 1964) to estimate volume and basal area. These photo-derived estimates were highly correlated with ground estimates of basal area and volume.

In their study of inventory methods using quad-centered 1:24,000 scale photography, Martin and Gerlach (1981) made some direct photo observations of stand structure (overstory height, canopy cover, and crown size), some indirect observations (pattern, texture) and some physical site observations (slope angle, aspect, elevation). These photo-interpreted categorical variables were then related to timber inventory attributes determined from the ground (volume, site index, growth and yield, habitat type). Equations were developed using multiple regression techniques to predict the timber inventory attributes from the photo-interpreted attributes. Correlation coefficients for these equations ranged from 0.55 to 0.82.

Teuber (1983) used methods similar to Martin and Gerlach (1981) except that larger-scale color aerial photos were used for interpretation. Stands were delineated based on texture, tone, pattern, and topographic characteristics. Descriptive attributes (tree height, crown diameter, forest type) were recorded for each stand and then used in developing predictive models. These models were then tested on 9 previously unsampled stands. The equations predicted the observed conditions fairly well for most of the variables.

Deegan and Befort (1990) field sampled 216 stands that had previously been photo interpreted for cover type from 1:15,840 scale black and white infrared photos. The percentage agreement adjusted for chance agreement (Kappa) was 70% for forest cover type.

Recently, photo-interpreted data have been used in wildlife habitat models (Short, 1988) old growth surveys (Rutledge and Hejl, 1990), and large scale landscape assessments (Kalkhan, et al, 1995; Allen, 1994; Gonzales, 1994; Lehmkuhl, et al, 1994; Green, et al, 1993; Lehmkuhl, et al, 1992; Deegan and Befort, 1990). In these types of studies it is necessary to be aware of the potential inaccuracy of photo interpreted data and the effect this inaccuracy may have upon the results. Deegan and Befort (1990) showed that inaccurate photo-interpretation can have a substantial effect on forest cover type acreage estimations. The large discrepancy between the theoretical (no photo misclassifications) sampling error and the photo misclassification sampling error indicates the importance of recognizing the accuracy of photo-interpreted data. Some researchers have used aerial photo data as "ground truth" reference data for maps constructed from satellite digital images (Green et al, 1993). Error in reference data will be attributed to the digital images unless otherwise accounted for (Congalton, 1991). Others have used double sampling techniques where a combination of ground and aerial photo data are assembled into a matrix which is then used to assess the accuracy of classified satellite imagery (Kalkhan, et al, 1995).

Obtaining accurate estimates of forest stand structure variables for each stand across the landscape is important for the modeling of landscape level processes of vegetation change. J. Chew (1995) developed a model to simulate vegetation change considering processes such as stand development, fire, and insect and disease outbreaks. Data on the density (canopy cover), cover type, size class (DBH), and canopy layering for each stand across the landscape are necessary to run the model; the landscape pattern created by stands varying in structure affects how the model simulates vegetation change. Frequently, these stand level attributes are obtained from aerial photographs. Models such as Chew's are becoming important tools in implementing ecosystem management. Knowledge of the accuracy of the data input into such models can provide the researcher/land manager with a measure of confidence in the output of these models.

Previous investigations into the accuracy of estimating stand structure variables from aerial photographs have been conducted at the plot or single tree level (e.g., Spurr, 1960; Worley and Meyer, 1955; Worley and Landis, 1954); few studies have examined this issue at the landscape scale (see Deegan and Befort, 1990). This study will explore this accuracy at the landscape scale using structure attributes averaged at the stand level.

Data obtained from aerial photographs continue to be important in natural resource management because of widespread accessibility, cost effectiveness, and relative ease in developing interpretation skills. Like all remotely sensed data, however, the accuracy of the data should be assessed and incorporated appropriately into the analysis.

Chapter II: Methods/Pilot Study

All data collection and analysis on this project was conducted by a single investigator to eliminate bias between observers. All aspects of the pilot study will be described in this chapter, followed by the methods for the full landscape assessment.

A. Study area

The Finley Creek management area is located on the Flathead Indian Reservation, north of Missoula, Montana (fig. 1). The area ranges in elevation from 3800 ft to 6000 ft (1158 m to 1829 m) and is generally west-facing in aspect. A wide range of habitats are present in the 11,200 acre (4532 ha) management area; the drier, lower elevation slopes contain ponderosa pine (Pinus ponderosa)/Douglas fir (Pseudotsuga menziesii) forest types while the highest portions of the area are dominated by whitebark pine (Pinus albicaulis) and subalpine fir (Abies lasiocarpa).

The lower elevation western part of the Finley Creek area has been extensively managed silviculturally using both even and uneven-aged methods, while the eastern portion of the area is currently in an unmanaged state with the possible exception of fire suppression activities. The eastern part of the management area is characterized by steep, rocky slopes, high alpine lakes and meadows.

The Finley Creek area was selected for this study because inventory at this time fits in with Confederated Salish/Kootenai Tribal Forestry plans for management activities in this area in the future. Additionally, the study area was conveniently located near Missoula, MT; this resulted in substantial savings in logistical costs.

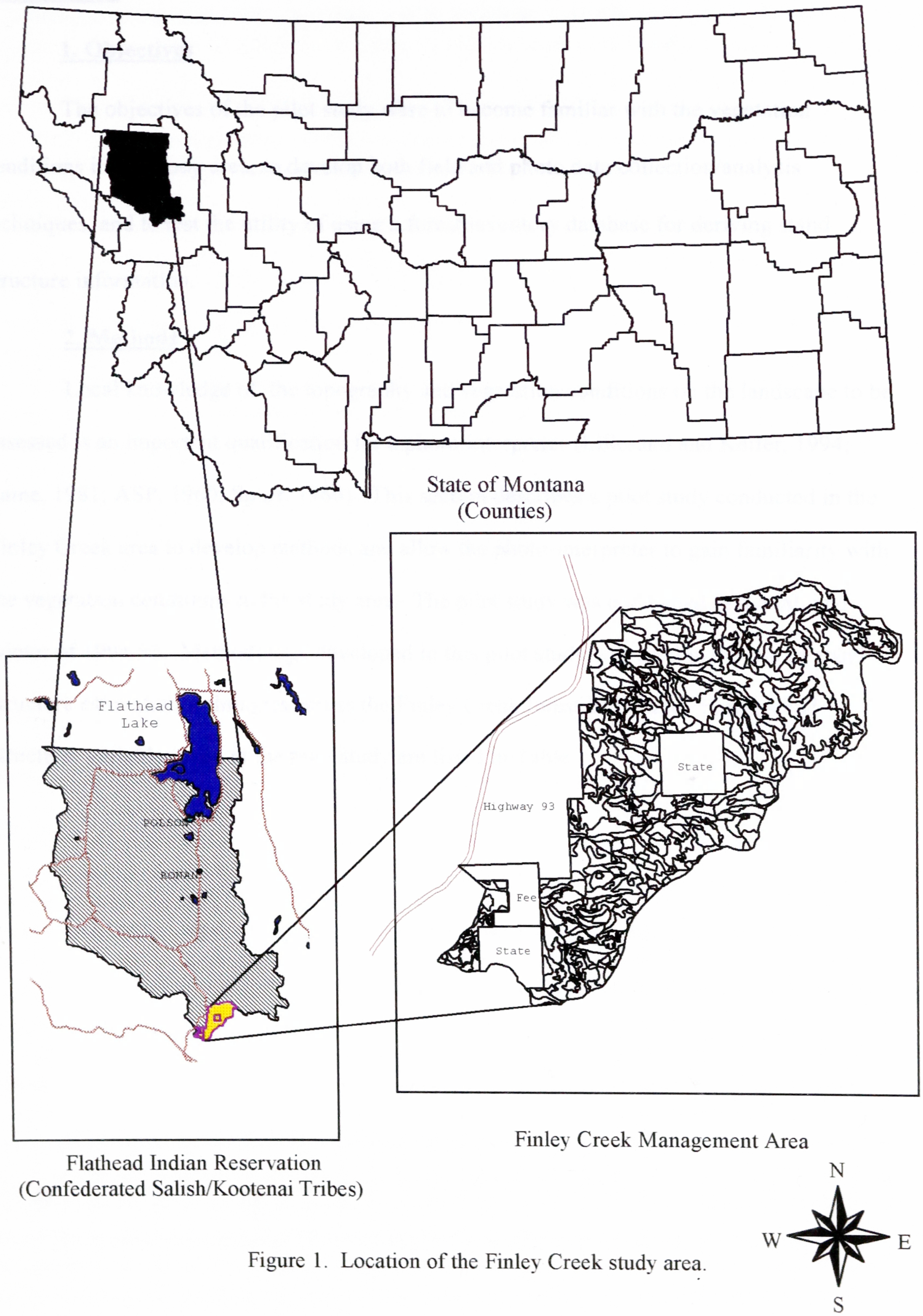


Figure 1. Location of the Finley Creek study area.

B. Pilot study

1. Objectives

The objectives of the pilot study were to become familiar with the vegetation conditions in the study area, to develop both field and photo data collection/analysis techniques, and to test the utility of using a forest inventory database for deriving stand structure information.

2. Methods

Local knowledge of the topography and vegetation conditions on the landscape to be assessed is an important qualification for a photo-interpreter (Lillesand and Keifer, 1994; Paine, 1981; ASP, 1960; Spurr, 1960). This section describes a pilot study conducted in the Finley Creek area to develop methods and allow the photo-interpreter to gain familiarity with the vegetation conditions in the study area. The pilot study was conducted in the fall and winter of 1994-95. Methodology developed in this pilot study was used in the test of stand structure estimation techniques across the Finley Creek forest landscape. Forest stand structure variables used in the pilot study are listed in Table 1:

Table 1.--Initial pilot study variables. These variables were estimated in the initial CFI analysis and in field data collection for the pilot study.

1). Cover type (Single dominant overstory species)

2). Stem diameter 1.4 m above the ground (DBH) of dominant trees (Measured or estimated to the nearest inch [2.54 cm])

3). Height of dominant trees (Measured or estimated to the nearest foot [0.3 m])

4). Stand canopy layers (1, 2 or 3 distinguishable canopy layers)

5). Total tree canopy cover (Percentage)

6). Crown diameter of dominant trees (Measured or estimated to the nearest foot [0.3 m])

7). Shrub cover type (Dominant shrub species)

8). Snags (Number of snags > 10 inches DBH [25 cm])

9). Large woody debris (Number of pieces > 10 ft (3 m) long and > 10 inches (25 cm) in diameter)

These attributes were selected because they are conventional descriptors of stand conditions that are widely used in landscape assessments in this region (e.g. Lehmkuhl et al, 1994; Sweet and Wall, 1995).

Continuous Forest Inventory (CFI) plots are located in the southeast corner of those sections where the majority of the land is available for timber management. The CFI Plots were established by the Bureau of Indian Affairs and Confederated Salish/Kootenai Tribes to track timber inventory, evaluate land productivity, and describe stand and tree problems. These permanent plots are re-sampled every 10 years to gather data on growth and mortality.

Plot records for all of the 17 plots located in the Finley Creek area were analyzed and stand structure variables were derived from the raw CFI data. Forest cover type and DBH were determined by examining the individual tree records and averaging or weighting the values for the dominant trees for each plot. Dominance was determined by the relative number and size of the trees recorded for the plot. The number of stand layers was taken directly from the CFI database as a recorded variable. Height data for the trees was not always present in the database. Other structure attributes that could not be directly derived from the database include canopy cover and woody debris information.

Field visits to each of the 17 plots were conducted in the fall of 1994. The following stand attributes were recorded for three representative dominant or co-dominant trees using standard mensuration techniques: Tree species, size (DBH), visible crown diameter, and height. Averaging of these measurements gave values for the plot. Stand layers, canopy cover, cover type, shrub cover, snags, and large woody debris were recorded for the 1/8-acre (0.05 ha) radius around plot center. The exact plot location was pinpricked on the 1:15,840 scale normal color aerial photo while in the field.

The 17 plots were then photo-interpreted for the following stand attributes (Table 2):

Table 2.--Photo-interpreted pilot study variables. These variables were estimated from the aerial photos. Some variables were dropped from the initial list (table 1) because they were difficult to resolve on the photos (e.g. snags, large woody debris).

1). Cover type

2). DBH (ocular estimation to the nearest inch (2.54 cm)

3). Average height of dominant trees (nearest foot [30.5 cm])

4). Stand canopy layers

5). Total tree canopy cover (percent)

6). Visible crown diameter (nearest foot [30.5 cm])

7). Understory life form (e.g. grass, forb, shrub)

Snags, large woody debris, and shrub cover type could not be estimated from the aerial photographs due to low resolution and overstory canopy obscuring the forest floor. Instead of estimating shrub cover type, a simpler interpretation of the understory life form was made; plots were interpreted as having either a shrub, grass, forb, or rock understory.

The aerial photographs (1:15,840 nominal scale, 9"x9", normal color, date of exposure 8/90) were scaled and effective areas delineated. DBH, height, canopy cover, and crown diameter were measured with standard photo-interpretation tools such as estimation templates (transparencies), parallax bar, a Bausch and Lomb zoom stereoscope (6-10X power) with light table, and 10-power hand lens. Estimation templates were obtained from the Forest Service and are located in appendix B. Several methods of height measurement were explored. Shadow-length methods required extensive computation and the parallax wedge would involve setting up the photos for a non-mirrored stereoscope. The parallax bar could be used with the zoom mirror stereoscope and therefore no extra set up time was needed. Aerial photo stereo-pairs were aligned along flightlines and taped to a Plexiglas mounting board. The appropriate measurements between principal points and conjugate principal points were applied using parallax bar calibration methods described in Lillesand and Keifer (1994). The Plexiglas board could then be moved under the zoom stereoscope around the effective area of the stereo-pair without having to re-calibrate the parallax bar. Extensive practice tree height measurements with the parallax bar were conducted with trees of known (ground measured) height.

Cover type, stand canopy layers, and understory life form were estimated based on the texture, tone, and pattern visible on the image. Overstory cover type designations were aggregated into classes based on the primary species identified. For example, a plot classified as having a DF/WL cover type would be classified as DF for analysis. Four cover type classes were identified: DF (Douglas fir), WL (western larch [Larix occidentalis]), LP (lodgepole pine [Pinus contorta]), and PP (ponderosa pine).

DBH estimates from the photos were obtained in two ways: Simple ocular estimation, and using the regression equations described below in the analysis sub-section. Estimates by both methods were made to the nearest inch (2.54 cm).

2. Analysis

Multiple linear regression models were developed to predict ground DBH from the crown diameter and height measurements recorded from aerial photographs. The general form of the model is:

$$\text{DBH (estimated)} = a + b(\text{Crown diameter}) + c(\text{Height})$$

Data used to build the models were from the ground measurements of DBH, crown diameter, and height taken at each of the CFI plots and from Brown (1978). Data from both datasets (CFI plots and Brown, 1978) were combined to increase the sample size for the regressions; this compensated for low numbers of trees in either of the datasets. Regression models were constructed using only Brown's data, only the CFI data, and both datasets combined. For all of the species there was no appreciable difference in adjusted R^2 values between the Brown's data regressions and the regressions using both Brown's data and the CFI data. The CFI data were included in the equations because they provided some information on local variation of the tree stem diameter - height and crown width relationship and because they increased the number of sample trees in the dataset (table 3). Brown measured randomly selected trees on a range of different sites and stand density conditions throughout western Montana and northern Idaho. Because the trees he selected are representative of the natural variation seen in this region, using regressions based on his data is probably adequate without additional local sampling. However, Brown did not collect data on abnormal trees (i.e. wolf trees, trees with lopsided crowns, broken tops); if the ability to predict the stem diameter of these types of trees is desired, then addition of local tree data is recommended. Western larch tree measurements were absent from Brown's data, so the regression model was made using the CFI data only (table 3). Conversely, no grand fir (Abies grandis) trees were measured in the survey of the CFI plots, so only Brown's data were used to build the model for this species.

Table 3.--Number of trees used in the regressions in each of the datasets. Trees in the CFI data column were measured during field visits to the CFI plots.

Species	Brown (1978)	CFI data	Total
Ponderosa pine	35	11	46
Western larch	0	7	7
Subalpine fir	25	9	34
Douglas fir	28	23	51
Engelmann spruce (Picea engelmannii)	23	2	25
Lodgepole pine	19	8	27
Grand fir	30	0	30

Regressions were run using height as the single independent variable, crown diameter as the single independent variable, and both crown diameter and height as independent variables. For all of the species equations, an improvement in the adjusted R^2 was noted when both crown width and height were included in the equation (table 4).

The adjusted R^2 coefficient compensates for the addition of independent predictor variables to the multiple regression equation and is an appropriate coefficient to compare regression equations (Velleman, 1993). As new predictors are added to the regression equation, the R^2 value can only increase, even if the predictor is merely random numbers.

Table 4.--Adjusted R^2 values for the regression equations. Adjusted R^2 values are shown for equations using crown width only, height only, and both crown width and height as the independent variables.

Species	Crown width*	Height*	Crown width and height*
Ponderosa pine	0.87	0.87	0.93
Western larch	0.53	0.84	0.86
Subalpine fir	0.83	0.83	0.93
Douglas fir	0.77	0.88	0.93
Engelmann spruce	0.85	0.97	0.98
Lodgepole pine	0.74	0.92	0.94
Grand fir	0.85	0.97	0.99

* Independent variables included in the equation.

Adjusted R^2 accounts for this and generally does not increase if the added predictor is a nonsense one (Velleman, 1993). Because of the improvement in adjusted R^2 values resulting from inclusion of both crown width and height in the regression equation, these equations were used to predict DBH. A separate model was developed for each tree species (table 5). The models were then used to predict average DBH of the each of the plots using the average crown diameter and tree height as estimated from the photos.

The photo-interpreted cover type of the plot was used to select he model; if the cover type of the plot was western larch, then the western larch model was used.

Table 5.--Regression equations used to predict DBH. These equations were developed using CFI tree data and data from Brown (1978).

Species	Variable	Coefficient	S. E. of Coefficient	df*	R^2
Engelmann spruce	Constant	-1.25	0.383	22	0.98
	Height	0.142	0.010		
	Crown diameter	0.317	0.072		
	Variable	Coefficient	S. E. of Coefficient	df	R^2
Douglas fir	Constant	-2.69	0.609	48	0.93
	Height	0.157	0.0145		
	Crown diameter	0.465	0.0663		
	Variable	Coefficient	S. E. of Coefficient	df.	R^2
Lodgepole pine	Constant	-0.972	0.426	24	0.95
	Height	0.132	0.0141		
	Crown diameter	0.299	0.0860		
	Variable	Coefficient	S. E. of Coefficient	df.	R^2
Grand fir	Constant	-1.24	0.276	27	0.99
	Height	0.117	0.0068		
	Crown diameter	0.300	0.0456		
	Variable	Coefficient	S. E. of Coefficient	df	R^2
Western larch	Constant	-13.8	4.88	4	0.88
	Height	0.358	0.103		
	Crown diameter	0.375	0.177		
	Variable	Coefficient	S. E. of Coefficient	df	R^2
Subalpine fir	Constant	-1.80	0.391	31	0.96
	Height	0.0891	0.0109		
	Crown diameter	0.581	0.095		

* Degrees of freedom

A variety of analysis techniques were used to compare the values estimated for each attribute by the three methods (CFI database, photo-interpretation, field inventory). It was assumed for all attributes that the 'true' values are the field inventory values.

Ordinal and nominal variables (canopy layers, overstory cover type, understory life form) and those interval variables that were aggregated into classes (DBH, height) were analyzed using error matrix tables (Story and Congalton, 1986). Overall accuracy is calculated from the error matrix by dividing the sum of the stands in the major diagonal by the total number of stands (table 6).

Table 6.--Example error matrix and accuracy calculations. This hypothetical error matrix is used as an example of how to calculate the various types of accuracy (Overall, Producer's, and User's).

	Reference data					
	Class	A	B	C	D	Row total
Photo-interpreted data	A	3		3		6
	B		2			2
	C	4		3		7
	D	0		1	2	3
	Column total	7	2	7	2	18

Class	User's accuracy	Producer's accuracy
A	3/6 = 50%	3/7 = 43%
B	2/2 = 100%	2/2 = 100%
C	3/7 = 43%	3/7 = 43%
D	2/2 = 100%	2/3 = 67%

Sum of the major diagonal (3+2+3+2) divided by total sample (18) = 55%
55% = Overall accuracy (P_o)

Two different measures of accuracy can be calculated for each class in the error matrix. Errors of commission are determined by dividing the number of correctly classified stands in Class *A* by the total number of stands classified as Class *A* (row total). This is referred to as the user's accuracy and represents the probability that a stand classified as *A* on the image is actually *A* on the ground. Errors of omission are calculated by dividing the number of stands correctly classified in Class *A* and dividing by the total number of stands in Class *A* in the reference data (column total). This accuracy measure is called producer's accuracy and represents the probability that a stand identified as *A* on the ground has been correctly represented as *A* on the map.

Interval (continuous) variables (DBH, height, and visible crown diameter) were analyzed using the standard chi-square test of a hypothesized variance (Freese, 1960). This method involves calculating a probability [P(Z)] that the estimate is within the user-specified allowable error of the ground value. It also allows for the removal of two types of bias. This term [P(Z)] represents the probability that the estimate is within the allowable error specified for the variable. Allowable error terms were selected after considering the photo-interpretation and photogrammetry methods, and after consulting with Salish/Kootenai Tribal Forestry personnel (R. Becker, pers. comm., table 7). Details about this analysis technique are located in appendix C.

Table 7.--Allowable error for chi-square test (pilot study). These values were used in the chi-square test of a hypothesized variance (Freese, 1960) for the interval variables.

DBH	+/- 3" (7.62 cm)
Height	+/- 10' (3.05 m)
Crown diameter	+/- 4' (1.2 m)
Canopy closure	+/- 10%

The above allowable error values are also consistent with errors reported for images of this scale in the literature (Worley and Landis, 1954; Worley and Meyer, 1955; Spurr, 1960).

This method also allows for the removal of bias observed in the estimated data. The basic equations in Freese (1960) were algebraically rearranged so that the probability [P(Z)] the estimate (photo-interpreted or relevé) is within the allowable error of the ground value is easily determined. This method assumes a normal distribution of the data; normal probability plots were examined and this assumption was met. Details about this method are located in appendix C.

3. Results

Where possible, data were collected in interval (continuous) form and subjected to both the interval analysis technique (chi-square) and the categorical technique (error matrix). For the error matrix analysis, tree height data were collapsed into 5 classes with 10-ft (3.05-m) intervals and DBH data were collapsed into 5 classes with 5-inch (12.7-cm) intervals. Error matrices for each variable are located in appendix A.

Overall accuracies (P_o) for the photo-interpreted attributes ranged from 0.22 to 0.72 (table 8). The chi-square test of a hypothesized variance was only applicable to interval data (DBH, height, and crown diameter); the P(Z) values ranged from 0.60 - 0.80. CFI data on height, crown diameter, understory cover type, and canopy cover were not obtainable from the database. Theoretically, the CFI estimates of DBH and cover type should be 100% accurate if the same trees were chosen for the field inventory and the inventory of the CFI database. However, the criteria for selecting the trees in the field was different than the selection criteria used in the CFI database inventory (i.e. trees selected for CFI inventory were not always the representative dominant or co-dominants recognized in the field), so the accuracy coefficient values are somewhat less than 1.0. Also, the latest measurements in the CFI database were recorded in 1990, field measurements of DBH in 1995 may be significantly greater thus adding bias into the accuracy determination. Two estimates of DBH were obtained from the photos, an ocular estimate and an estimate using the regression models described above. The ocular estimates were nearly as accurate as the modeled estimates.

Overall, the P(Z) values are somewhat higher than the corresponding P_o values. This is probably due to the removal of bias using the P(Z) technique and the selection of classes for the error matrices. Some of the error between the ground measured variables and those derived from the CFI data can be attributed to the growth of the trees that has occurred since the latest CFI measurements were taken (1989 - 1990). This bias should have been removed in the chi-square test of hypothesized variance (Freese, 1960). For some variable pairs, very little bias was noted in this test.

Estimates of canopy layers from the photos were fairly accurate (72%). The small number of multi-layered plots (n=4) was partly responsible for the relatively high accuracy of this attribute. Four plots were classified as having multiple layers when they were single layered (table 20).

Table 8.--Accuracy coefficients for the pilot study. The accuracy coefficients for the photo and CFI data estimates for each of the variables are shown.

Variable	Photo-interpreted accuracy		CFI estimates accuracy
	(Estimate)	(Modeled)	
DBH	$P_o = 0.44$ $P(Z) = 0.60$	$P_o = 0.50$ $P(Z) = 0.70$	$P_o = 0.59$ $P(Z) = 0.85$
Height	$P_o = 0.22$ $P(Z) = 0.60$		N/A
Crown diameter	$P_o = 0.62$ $P(Z) = 0.80$		N/A
Canopy cover	$P_o = 0.38$		N/A
Canopy layers	$P_o = 0.72$		$P_o = 0.44$
Cover type	$P_o = 0.55$		$P_o = 0.94$
Understory cover type	$P_o = 0.42$		N/A

Note: The overall accuracy from the error matrix is designated by P_O. The error matrices are presented in appendix A. The probability corresponding to the standard normal deviate as calculated by the chi-square test of a hypothesized variance (Freese, 1960; appendix C) are also shown (P(Z)). This test was only applicable to the interval data.

Photo-interpreted cover type estimations were 55% accurate with the greatest confusion arising between ponderosa pine and Douglas fir stands (table 23).

Canopy cover estimations from the photos were surprisingly inaccurate (38%). It may be more realistic to aggregate canopy cover into 20% classes rather than 10% classes; this would result in an accuracy of 64%. There was no trend toward under-or-overestimation (table 30).

There was a trend toward underestimation of DBH from the aerial photos using the simple ocular estimation technique. Six plots were underestimated by one or more classes (table 25). The regression model estimates of DBH also were low; 8 plots were underestimated (table 26).

Heights were consistently underestimated; over half of the plots were underestimated by one or more classes (table 31).

The only plot measurements that could be reliably taken from the CFI data were DBH, cover type, and canopy layers. The low accuracy reported for deriving canopy layers (44%) from the CFI data was due to the different criteria for distinguishing layers between the two field methods (CFI methods and this pilot study). CFI data collection methods recognize canopy layers less than 6 ft (1.8 m) in height while the pilot study did not recognize vegetation less than 6 ft (1.8 m) in height as a canopy layer. CFI data collection methods did not require that tree heights be recorded for every plot; not enough tree heights were present in many of the sample plots to conduct an accuracy assessment.

4. Discussion

The slightly greater accuracy of the modeled estimates of DBH may indicate that with a little practice, ocular estimation may be nearly as accurate as the more time consuming process of estimation with the regression equations. DBH was predicted from the photo-estimates of crown width and height; errors in the measurement of these two attributes or in the interpretation of plot cover type (leading to the use of the wrong regression equation) would all contribute to inaccurate DBH estimation.

It is anticipated that overstory canopy layers will be hard to estimate accurately from the photos because of the difficulty in detecting layers beneath the overstory canopy. For this variable and cover type, the interpreters' knowledge of the vegetation conditions and how they are associated with the elevation, aspect and position of a site will be important.

In estimating cover type, crown texture patterns and color were distinctive where a single species dominated a plot (such as lodgepole pine), but were very difficult where several species co-dominated a plot. This made estimating the single dominant species for a mixed species plot difficult.

5. Conclusions

It was difficult to draw statistically significant conclusions with the small sample size used in this pilot study. Additionally, the possibility of observer bias precludes any detailed discussion of the photo interpretation results. Field visits to each of the plots were conducted prior to photo-interpretation; since there were only seventeen plots it was fairly easy to recall the stand conditions observed on the ground when conducting the photo-interpretation work. This bias was eliminated in the full landscape test by conducting the photo-interpretation work prior to field reconnaissance.

Estimates of DBH and cover type from the CFI data were more accurate than estimates from the photos; this indicates that when in-place stand inventory data are available these variables should be estimated using these data (table 8). The high accuracy for these variables is to be expected since they are directly measured in the CFI sampling design. The ability to derive canopy cover, height, and crown diameter data from the CFI database was limited; these variables were not directly recorded in the CFI inventory procedures. The CFI database was designed to track timber inventory using estimates of growth and yield from permanent plots that are re-measured periodically. The data from these plots can then be applied to similar stands across a landscape. Data derived from the CFI database may be adequate to characterize stands into structure types using DBH, cover type, and canopy layers, but the lack of data on height, crown diameter, and canopy cover precludes placement of stands into structure types based on these attributes. Stand-based inventory systems such as the US Forest Service stand exam program and EcoData inventory may yield more reliable stand structure information.

The field inventory of the CFI plots provided valuable knowledge of the vegetation conditions in the study area. Methods for data collection (photo-interpretation and field inventory) and analysis used in this pilot study were used in the full landscape study; however, some changes were made and are described below.

Some stand variables were not evaluated in the full test on the Finley Creek management area. Snags, large woody debris, and undergrowth life form could not be recognized or measured accurately on 1:15,840 nominal scale photos. Estimating these attributes from the photos would involve a lot of guesswork; therefore, they were dropped from the landscape study.

Stand attributes such as basal area and volume will not be tested because values for these attributes could be derived from tables using attribute values that will be tested (cover type, canopy closure, height).

A variable describing horizontal structure was defined and measured in the full study. This attribute, spatial aggregation, is a measure of clumpedness of vegetation distributed throughout the stand and is patterned after one used by Lehmkuhl, et al. (1994).

Because this is a stand-based variable, analysis in the plot-based pilot study was not appropriate.

C. Landscape study

1. Photo interpretation

A major assumption in this study was that the stands are more homogeneous internally and less homogeneous relative to their neighboring stands. Stand delineations for the study area have been provided by Confederated Salish/Kootenai Tribal Forestry personnel on a ARC/INFO GIS file and on hard copy maps. Each stand in the study area has a unique identifying number. Stand boundaries were transcribed on acetate sheets overlaying the aerial photos by hand using ortho-quad maps as a reference. Photo-interpretation was conducted in July, 1995 using techniques developed in the pilot study. Each stand was examined and the variables determined by visually averaging the conditions seen throughout the stand; no formal photo interpretation plots were measured within each stand. Knowledge of the relationships of the vegetation to site characteristics (elevation, aspect, topography) were used in conjunction with features seen on the photos to estimate stand attributes. Equipment and photo set-up were the same as in the pilot study. Photo-interpreted variables and methods are listed below:

1.) DBH of three representative co-dominant or dominant trees determined by ocular estimation. The average of these three trees was recorded as the DBH (ocular estimate). Crown class was estimated by observing the position of tree crowns in the stereomodel.

2). Height of the same three co-dominant or dominant trees measured with a parallax bar on the stereo model of a pair of photos.

3). Visible crown diameter of the same three co-dominant or dominant trees. A mylar overlay of circles of various diameters was used in conjunction with the scale of the location of the stand on the photo (see appendix B).

4). Stand overstory canopy layers - Each stand was placed into one of four categories based on the estimated presence of at least 20% canopy cover in recognizable layers. This variable was estimated based on the texture, tone, elevation, and aspect of the stand.

1 - Single canopy layer; this layer contains at least 85% of the total stand canopy cover.

2 - Two canopy layers; each layer contains at least 20% stand canopy cover.

3 - Three or more canopy layers; each layer contains at least 20% canopy cover.

4 - Continuous canopy layer spread among multiple height groups; no single height group contains 20% stand canopy cover. Plot does not fit categories 1-3 above.

5). Total tree canopy cover of all trees greater than 10 feet (3.05 m) in height, determined by ocular estimation and by referring to a mylar estimation template (appendix B) placed on the photo.

6). Cover type - Dominant single tree species present in the stand based upon the texture, color, tone and pattern seen on the photos and the physical (elevation, aspect) attributes of the stand. Society of American Forester's cover types were not used because of the difficulty in distinguishing the many species mixes on the photos.

7). Spatial aggregation (clumpiness) of the stand was determined by placing the stand into one of three categories (fig. 2):

U - Uniform tree cover throughout the majority of stand; openings in stand canopy 0.5 acre (0.2 ha) or less.

R - Regular pattern of openings in overstory canopy: openings range in size from 0.5 acre to 1 acre (0.2 ha to 0.4 ha). The mosaic of openings seems to be repeated throughout the stand.

I - Irregular pattern of openings in overstory canopy; openings variable in size but greater than 0.5 acre (0.2 ha).

Mylar templates were used to estimate the size of the openings (see appendix B).

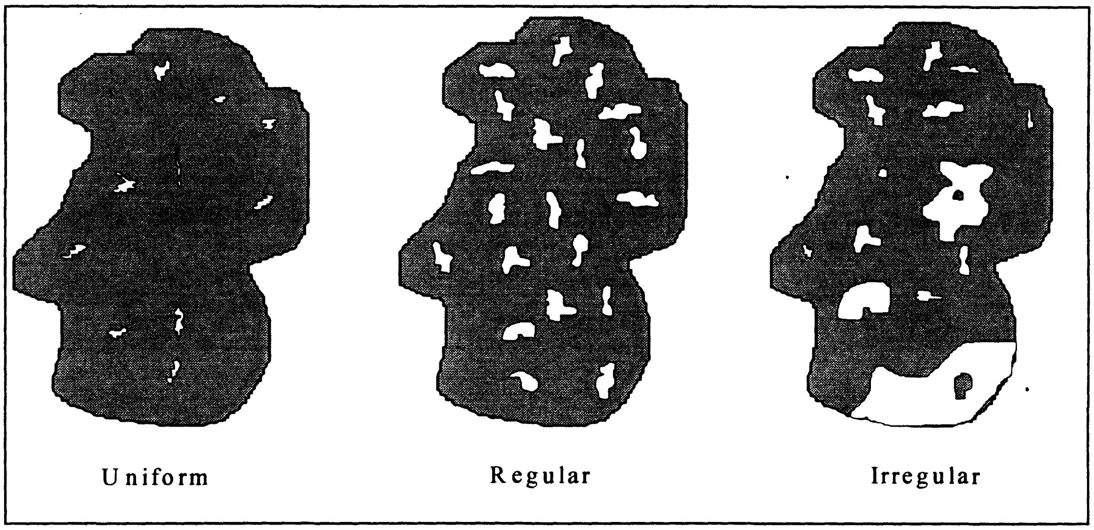


Figure 2.--Spatial aggregation stand types. Graphic examples of three hypothetical stands representing the photo-interpreted spatial aggregation stand types. Shaded areas represent continuous canopy cover, white areas openings.

Because of stereo overlap, many stands were covered on multiple photos. To maximize the accurate estimation of the canopy cover, crown diameter, and DBH attributes, each stand was interpreted on the photo pair where it was located nearest to the center of the photo (principal point). Thus measurement errors due to radial displacement and subsequent distortion on the edges of the photos were minimized. Some stands did not have adequate stereo coverage on the photos, either due to steep topography or lack of stereo coverage near the boundary of the study area. These stands were interpreted monoscopically with a 10-power hand lens.

A relevé plot was identified for the stand and pin-pricked on the photo. The requirements for the field selection of a relevé plot are 1). uniform habitat (vegetation conditions) within the plot, 2). homogeneous vegetation cover within the plot, and 3). the plot should be large enough to contain all species belonging to the plant community (Mueller-Dombois and Ellenberg, 1974). In order to meet these requirements, "a thorough reconnaissance" of the plant communities to be sampled is recommended (Mueller-Dombois and Ellenberg, 1974). In this study, the relevé concept of Mueller-Dombois and Ellenberg (1974) was modified somewhat, because the requirements (above) could not be met based on a photo reconnaissance of the stand. The relevé plot location selected best represented the average conditions seen throughout the stand from the photos, rather than a plot location selected to meet the requirements (above) based on a field inventory of the stand. Relevé plot size was 1/5 acre (0.08 ha); this size was thought to be large enough to capture most of the variation in overstory tree cover and yet still be practical for field sampling. Because the objective in selecting the relevé plot location was to represent the average conditions seen in the stand, in many cases the relevé plots were heterogeneous in vegetation cover to reflect heterogeneity seen in the stand. Photo-interpretation data specific to the relevé plots were not collected. Relevé plots were identified for all photo-interpreted stands in the study..

2. Sampling design

Following photo-interpretation, the stand database was sorted to exclude stands less than 10 acres (4.0 ha) in size. From this annotated database of 270 stands, 50 were randomly selected for field inventory. The random sample was drawn by writing a program using the RAN function in FOXPRO database management software based on the unique stand numbers assigned to each stand. The random draw was done after all stands had been photo-interpreted and the data entered into the database. Simple random sampling was selected as the sampling scheme because the Kappa analysis technique (see below) assumes a multi-nomial sampling model such as simple random sampling (Congalton, 1991; Congalton, 1988). Stratified random sampling was deemed inappropriate because of the potential low number of sample stands in any chosen class; i.e. if only 5 of the 50 sample stands were estimated to be the subalpine fir cover type then it would be difficult to draw any significant conclusions based on only 5 stands.

Five unbiased, systematic plots were selected in each of the 50 sample stands. First, a map was plotted of each sample stand from the ARC/INFO map file. A starting point within the stand that was easily located on the ground (such as a road intersection, or distinct curve in a road) was identified on the map. Next, the map distance of two chains was calculated using the scale of the map. From the starting point, a distance of at least two chains was measured in any cardinal direction (North, South, East or West) which remained within the stand boundary. This point was established as plot center for the first plot. This process was repeated until five plot locations were identified on the map (fig. 3). Care was taken to not place plot locations too close to stand boundaries because mapping and orienteering errors may result in plots being located outside the stand boundary on the ground. The relevé plot was not included as one of the five field inventory plots because it was intended to be an independent estimate of the ground truth data (averages of the 5 plots).

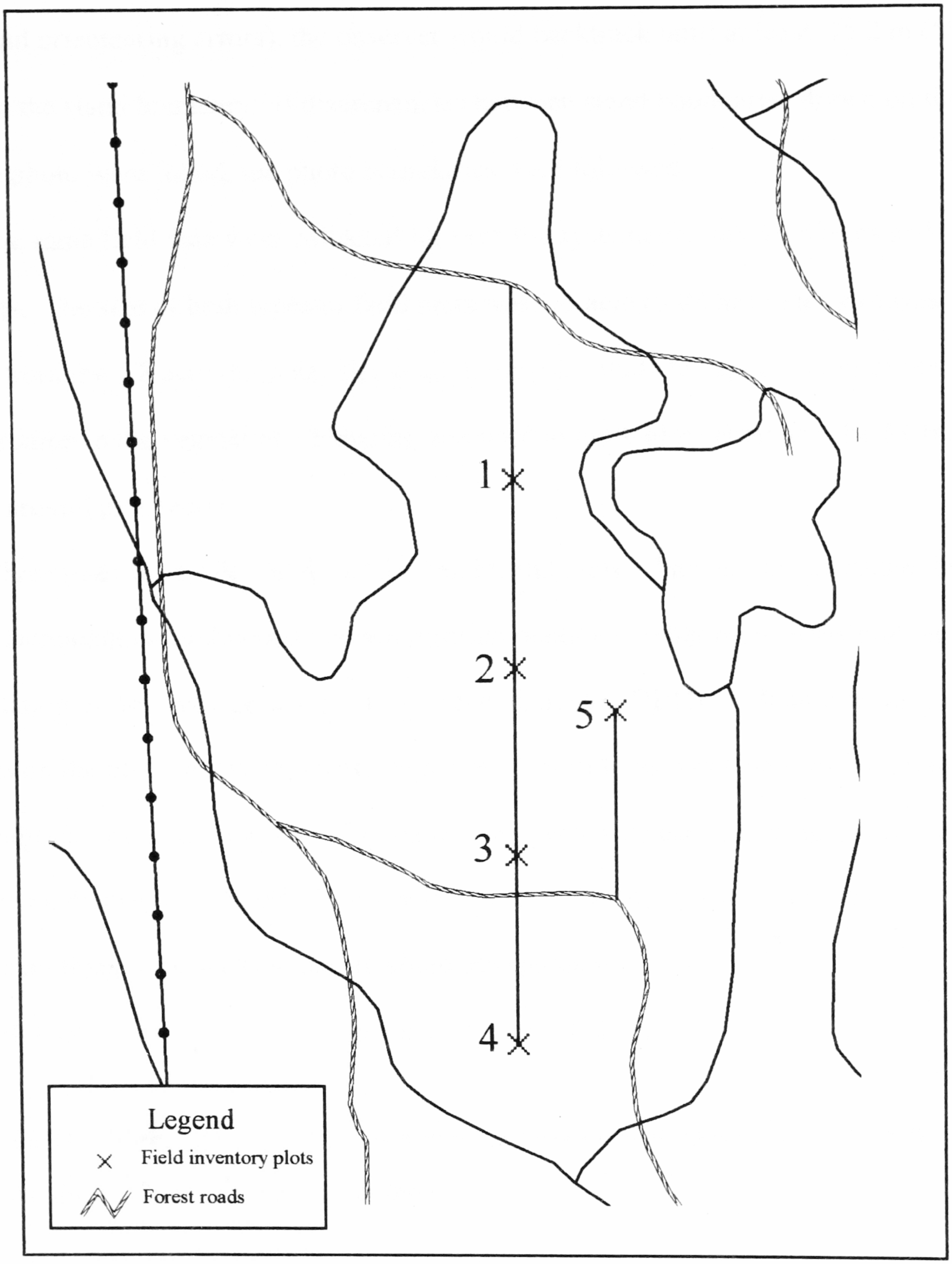
Maps were made in the office prior to going into the field to ensure that the method was objective and unbiased. A concerted effort was made to sample all parts of a stand; this was especially important in stands having irregular, convoluted boundaries. This effort sometimes necessitated establishing more than one starting point and having plots greater than two chains apart.

3. Field Methods

Field data collection methods were closely associated to the photo-interpretation methods to reduce differences due to methods between the datasets. For example, in the field, tree canopy cover was estimated to include only trees 10 feet (3.05 m) or taller because trees shorter than 10 feet are difficult to distinguish from brush on the photos.

Field data collection for the 51 sample stands was conducted in August and September, 1995. The intended sample size was 50; an extra stand was inadvertently sampled. Stands were accessed by truck, mountain bicycle, and foot.

The relevé plot location, identified previously during photo-interpretation, was visited first to reduce the possibility of observer bias between the two ground sampling methods. The aerial photo with the pinpricked relevé plot location was used to navigate to plot center. It was thought that if the five systematically placed plots were surveyed first, then the observer may be biased when recording data for the relevé plot which was to represent modal conditions seen throughout the stand. After collecting data at the relevé plot, the observer returned to the starting point in the stand and proceeded to the other five plots by pacing the appropriate number of chains at the appropriate azimuth.



(2 chains, 132 feet)

1:4965

Figure 3. Example stand map with field inventory plots. Plot locations were selected before going into the field.

If plot locations were obviously outside stand boundaries drawn on the photo (due to mapping and orienteering errors), the observer would backtrack until at least 15.2 m (50 feet) inside the stand boundary. If discrepancies between stand boundaries drawn on the map and on the photo were found, the photo boundaries were followed.

The same field data were collected for both the systematic, unbiased plots and the relevé plots. The size of both types of field plots was 1/5 acre (0.08 ha). Plot size was increased from the 1/8 acre (0.05 ha) size used in the pilot study to capture more of the spatial variation in the vegetation. Flagging was used to mark plot center and the 53-ft (16.2 m) radius around plot center.

Three trees were selected for DBH, height, and crown diameter measurement based upon: 1) Dominant or co-dominant crown position relative to other trees in the plot, and 2) Representation of the average size (Height, crown diameter, DBH) of all dominants and co-dominants on the plot, i.e. exceptionally large trees ("wolf trees") or small trees growing in dense conditions were not selected. The same three trees were used for measurement of height, crown diameter, and DBH. These selection criteria closely match the way measurement trees were selected during photo-interpretation.

Methods specific to each of the variables are listed below:

1). DBH of three representative co-dominant or dominant trees measured 4.5 ft (1.4 m) above groundline with a diameter tape.

2). Height of the same three trees measured with a clinometer at a distance of at least 66 ft (20 m) from each tree.

3). Visible crown diameter of the same three trees; two representative radii were measured on each tree from the center of the bole to the edge of the crown. The edge of the crown was determined by moving under the crown until a 90 degree vertical projection angle could be measured with a clinometer to the leading edge of the crown. The distance to the center of the bole was measured from this point. The sum of these two measurements was recorded as the diameter.

4). Overstory canopy layers - The plot was placed into one of four categories based on the estimated presence of at least 20% canopy cover in recognizable layers. Canopy less than 3.05 m (10 feet) in height was not considered a layer.

1 - Single canopy layer; this layer contains at least 85% of the total plot canopy cover.

2 - Two canopy layers; each layer contains at least 20% plot canopy cover.

3 - Three or more canopy layers; each layer contains at least 20% of the plot canopy cover.

4 - Continuous canopy layer spread among multiple height groups; no single height group contains 20% plot canopy cover. Plot does not fit categories 1-3 above.

5). Total tree canopy cover of all trees greater than 10 feet (3.05 m) determined by ocular estimation of percentage of ground that was obscured by canopy.

6). Cover type - Dominant single tree species (based on ocular estimation of the species with the greatest canopy cover) present in the plot.

7). Spatial aggregation of the stand was determined after all plots had been surveyed. After traveling to all plots in the stand, one of the following categories was assigned to the stand:

U - Uniform tree cover throughout the majority of stand; openings in stand canopy 0.5 acre (0.2 ha) or less.

R - Regular pattern of openings in overstory canopy: openings range in size from 0.5 acre to 1 acre (0.2 ha to 0.4 ha). The mosaic of openings seems to be repeated throughout the stand.

I - Irregular pattern of openings in overstory canopy; openings variable in size but greater than 0.5 acre (0.2 ha).

4. Data compilation and analysis methods

Stand values for each of the variables were determined from the plot data. The tree measurements (height, DBH, crown diameter) were averaged for each plot, and these five plot means were then averaged for each stand. Canopy cover estimates for each plot were averaged to get the stand canopy cover. The cover type for the stand was determined by choosing the majority cover type for the five plots. Cover type determination for the stands was conducted in the field so that if a majority cover type was not evident from the five plots, a tie-breaker determination could be made based upon the cover types observed while traveling between plots. Relevé plot values were not included in the above compilation.

The three datasets could now be compared for the 51 sample stands: A photo-interpreted dataset, a relevé dataset, and a dataset from the five unbiased, systematic plots. The five-plot dataset will hereafter be referred to as the ground dataset. In assessing the accuracy of each of the variables, the five-plot ground data averages were assumed to be the 'truth'. The relevé and photo data were then compared to the ground data (fig. 4):

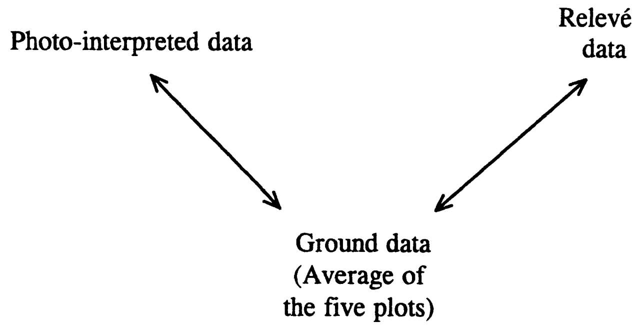


Figure 4.--Relationship of the three datasets. The estimates by the two different techniques (photo and relevé) were compared to the ground truth data. A comparison of the photo data to the releve data was not conducted.

Two different methods of accuracy assessment were used in the analysis of the data and are described below. Also, the data were analyzed to see if there were any consistent biases among the stands that were under-or-over-estimated (mis-estimated).

Error matrix tables (Story and Congalton, 1986; see pilot study above) - These tables were constructed for all variables. Interval variables (DBH, height, crown diameter, and canopy cover) were collapsed into the following classes (Table 9):

Table 9.--Classes for error matrix tables. The interval variables were collapsed into these classes for the error matrix analysis.

	DBH (inches)		Height (ft)		Canopy closure		Crown diameter (ft)
1.	< 5 (12.7 cm)	1.	< 20 (6.1 m)	1.	< 25%	1.	< 5 (1.3 dm)
2.	5 - 8.9 (22.8 cm)	2.	20 - 39 (11.9 m)	2.	26 - 59%	2.	5 - 8.9 (2.3 dm)
3.	9 - 14.9 (38.1 cm)	3.	40 - 59 (17.9 m)	3.	> 60%	3.	9 - 14.9 (3.8 dm)
4.	15 - 20.9 (53.3 cm)	4.	60 - 70 (21.3 m)			4.	15 - 20.9 (5.3 dm)
5.	> 21	5.	71 - 99 (30.2 m)			5.	> 21
		6.	> 99				

These classes were used because they seemed appropriate for the forest vegetation in this region and they were considered appropriate by Salish/Kootenai Tribal forestry personnel (R. Becker, pers. comm.).

In addition to the producer's, user's and overall accuracies (see pilot study above), the error matrix tables also display which classes were misinterpreted, and for the interval classes (table 9), which are misestimated. Errors for the interval variables can be determined by position of stands in the table relative to the major diagonal (table 10).

Table 10.--Hypothetical error matrix table. The data and classes are fictitious; this table is an example of how to interpret an error matrix.

		Reference data				
	Class	1	2	3	4	Row total
Photo interpreted data	1	80	64			144
	2		2	1		3
	3			1		1
	4	1	2	10	7	20
	Column total	81	68	12	7	168

Data in table 10 are fictitious. Stands that appear in cells to the right of the major diagonal (shaded) were underestimated and stands that appear to the left of the diagonal were overestimated. In table 10, 65 stands were underestimated for the variable and 13 were overestimated.

Three accuracy coefficients, Kappa (K) Tau (T_e) and overall accuracy (P_o), were calculated from the error matrices using methods described in Ma and Redmond (1995). The Kappa and Tau coefficients represent adjustments to overall accuracy to account for chance agreement. Tests for significant differences were conducted on the Tau coefficients using methods described in Ma and Redmond (1995).

Chi-square test of a hypothesized variance (Freese, 1960) - This test was applicable to the interval scale data only (DBH, height, crown diameter, and canopy cover) and is explained above in the pilot study methods and in detail in appendix C.

Error analysis - Scatterplots with regression lines were made for each of the interval variables in order to examine the sources of error. A line of equality was placed on each of the scatterplots; this line represents perfect prediction of the variable by the estimation technique. Stands that deviated from this line were examined to see if there were any consistent trends among the misestimated stands, such as position of the stand on the image or topography. Stands located near the edge of the images may have inaccurate estimates from distortion due to radial displacement and shadows. Fig. 5 illustrates the use of scatterplots in identifying misestimated stands. Stands plotted to the left of the line were underestimated; those to the right overestimated. A Least Squares Difference (LSD) regression line was also fitted to show how much of the variation in the ground values was captured by the estimate.

The coefficient of variation (CV) was calculated for each of the plot variables within each of the sample stands. This statistic provides an estimate of within stand variability for each of the attributes. The mean CV for the underestimated and overestimated stands was calculated and t-tests conducted to see if they were significantly different from the mean CV for those stands that were accurately interpreted. The mean CV for the stands in each of the three spatial aggregation categories were also compared to explore the relationship between measures of within stand heterogeneity (spatial aggregation) and within stand variance.

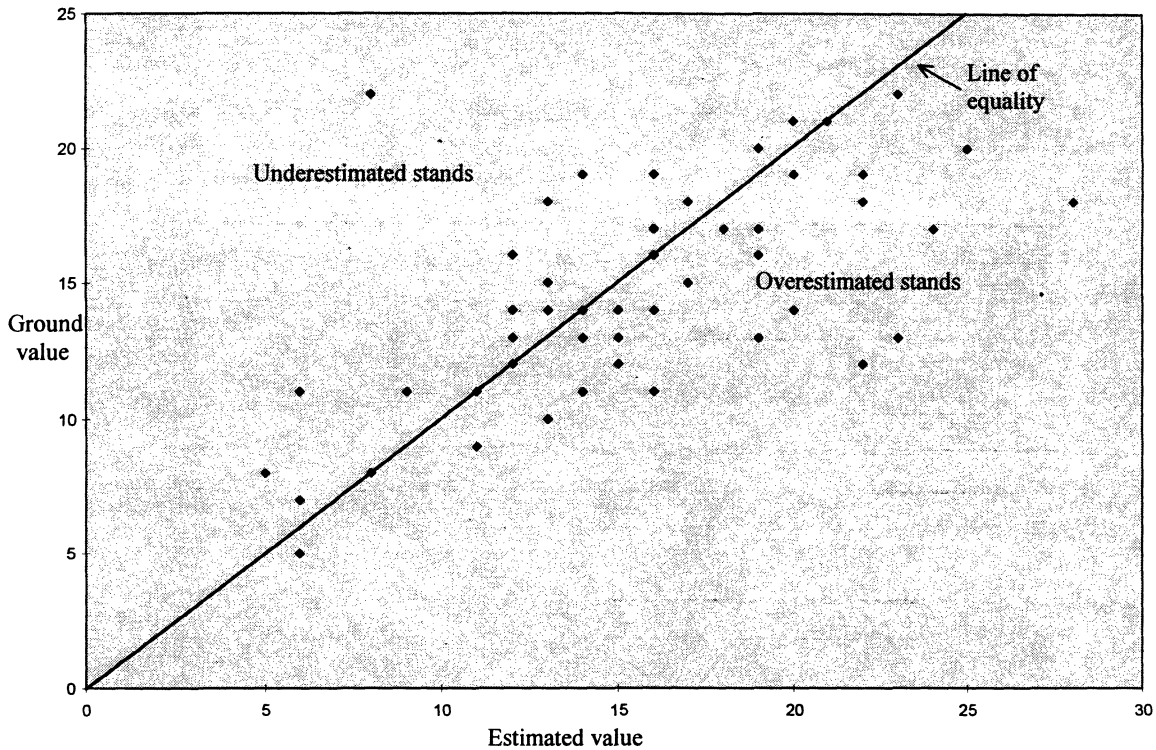


Figure 5.--Example scatterplot illustrating the error analysis technique. Each symbol represents one stand. Misestimated stands were identified based on their position relative to the line of equality.

All data were entered into FOXPRO database management software. Programs were written in FOXPRO programming language to perform many of the analyses. SPSS for Windows was used to construct the error matrix tables and a UNIX program provided by Zhenkui Ma of The Montana Cooperative Wildlife Research Unit was used to calculate the Kappa and Tau coefficients.

Chapter III: Results

The 51 randomly selected sample stands are displayed in fig. 6. Data from these 51 stands were used in the accuracy assessment analysis.

The results will be presented in three parts. First, the error matrices comparing the photo-interpreted estimates to the ground data are exhibited in section A. Error matrix tables for the relevé data are located in appendix D. Second, the results of the error analysis are presented in section B. Finally, the different accuracy measures will be summarized and the results from the relevé analysis compared to the photo-interpretation results (section C).

A. Photo-interpretation error matrices

Tables 11 through 18 are the error matrices for the photo-interpreted variables. A detailed explanation of the first table (table 11) provides familiarity with the format and interpretation of the remaining tables. The reference data are the objective field inventory data (averages of the five field inventory plots in each sample stand). Stands were placed in cells in the table based on the estimated value of the attribute (row), and the ground value of the attribute (column).

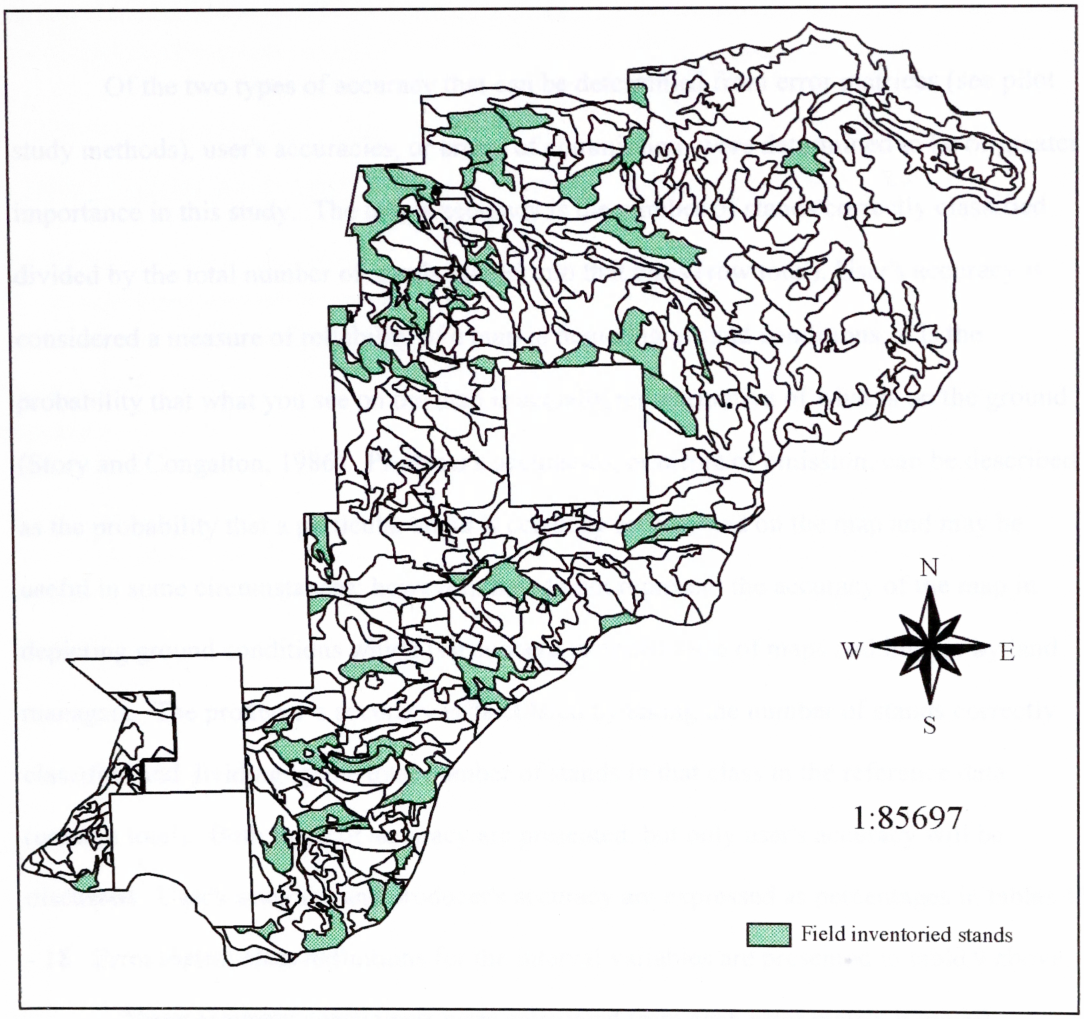


Figure 6. Map of Finley Creek Management Area showing field inventoried sample stands

Of the two types of accuracy that can be determined from error matrices (see pilot study methods), user's accuracies, or errors of commission, were determined to be of greater importance in this study. The user's accuracy is the number of stands correctly classified divided by the total number of stands placed into that class (row total). User's accuracy is considered a measure of reliability of a map in depicting ground conditions; it is the probability that what you see on the map is actually representative of what is on the ground (Story and Congalton, 1986). Producer's accuracies, or errors of omission, can be described as the probability that a particular stand is correctly represented on the map and may be useful in some circumstances; however, they do not represent the accuracy of the map in depicting ground conditions which is a concern in applied use of maps and images by land managers. The producer's accuracy is calculated by taking the number of stands correctly classified and dividing by the total number of stands in that class in the reference data (column total). Both types of accuracy are presented, but only user's accuracy will be discussed. User's accuracy and producer's accuracy are expressed as percentages in tables 11 - 18. Error matrix class definitions for the interval variables are presented in table 9 above.

Three accuracy coefficients were calculated from each error matrix [Overall accuracy (P_o), Tau (T_e), and Kappa (K)]; these coefficients are presented here and discussed in section D. In general, the Kappa (K) and Tau (T_e) coefficients represent adjustments to the overall accuracy (P_o) to compensate for chance placement of stands into the correct cells in the error matrix table.

Table 11.--Error matrix table of the aerial photo estimates of stand DBH using the regression models

Reference (ground) data

	Class	1	2	3	4	5	Row total
	1			4	1		5
Photo-interpreted data	2	1	3	4	1		9
	3			13	5	2	20
	4			4	10	2	16
	5				1		1
	Column total	1	3	25	18	4	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	33	100
3	65	52
4	63	56
5	-	-

Overall accuracy (P_o) = 0.51

Tau (T_e) = 0.39

Kappa (K) = 0.28

The major diagonal (shaded) are the cells that contain the correctly classified stands. The sum of the stands in the major diagonal divided by the total number of sample stands is the overall accuracy (In table 11, 26/51 = 0.51 = P_o). Stands in cells to the right of the major diagonal were placed in a class lower than the reference data; these stands were underestimated. For example in table 11, 4 stands were estimated to class 2, when actually they were in class 3. Likewise, stands in cells to the left of the major diagonal were overestimated; in table 11 a total of 6 stands were overestimated. Thus the stand DBH regression model underestimated ground DBH for 19 stands and overestimated DBH in 6 stands. The highest percentage of the incorrectly classified stands were located to the right of the major diagonal, so overall the photo-interpreted DBH regression models underestimated actual ground DBH. If a nearly equal number of stands were over and underestimated, then the estimating technique (photo-interpretation or relevé) would be considered balanced.

A more detailed analysis of producer's accuracy provides information on classes where stands were consistently omitted from the correct class; these classes can be identified by examining the reference data column for each class and noting where stands in that class were placed by the estimating technique. For example, 8 stands that were in DBH class 3 were underestimated into DBH classes 1 and 2 by the regression models. Similarly, 4 stands in DBH class 3 were overestimated and placed into class 4. DBH class 4 was also underestimated; 7 stands that were in class 4 were in placed in classes 1, 2 and 3.

User's accuracies can be explored in a similar way by examining the rows of the table and identifying those stands that were placed in the wrong class by the estimating technique. For example, looking at row 3 in table 11 reveals that 7 stands that were classified into class 3 by the estimating technique are actually in classes 4 and 5. Again the focus of the analysis is on the user's accuracies because this was deemed of greatest concern to land managers.

Obviously, this analysis of over-and-underestimation is applicable only to ordinal variables. For the nominal variables, this type of analysis shows classes that were misinterpreted; i.e. if Douglas fir stands are consistently misinterpreted as ponderosa pine stands then these stands would appear in the Douglas fir column and in the ponderosa pine row.

The DBH regression models were able to distinguish between DBH classes 3 and 4 with high accuracy (user's accuracies of 65% and 63%, respectively). The successful discrimination between these two classes was of great importance since the majority of the sample stands (43 out of 51), and presumably the majority of the stands in Finley Creek area, were in these classes. For the DBH, height, crown diameter, and canopy cover variables (tables 11, 12, 14, 15), user's accuracies were relatively low for the high and low classes (classes 1, 2, and 5). This may be due to few stands in these classes; one or two misclassified stands can result in low accuracies. But for the classes that contained the majority of the stands for these attributes the user's accuracies were fairly high, indicating that the photo-estimation techniques were more accurate than the coefficients [Overall accuracy (P_o), Tau (T_e), and Kappa (K)] indicate for the majority of the stands in the study area.

Table 12.--Error matrix table of the aerial photo estimates of stand DBH.

	Class	1	2	3	4	5	Row total
		Reference data					
	1			1			1
Photo-interpreted data	2	1	2	2	1		6
	3		1	14	5	1	21
	4			8	10	2	20
	5				2	1	3
	Column total	1	3	25	18	4	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	33	67
3	67	56
4	50	56
5	33	25

Overall accuracy (P_o) = 0.53

Tau (T_e) = 0.41

Kappa (K) = 0.27

Ocular estimates of DBH were nearly balanced between over and under estimation; 12 stands were underestimated while 11 were overestimated (table 12). Stands in higher classes were consistently underestimated into DBH classes 2 and 3. A significant number of stands in class 3 were overestimated into class 4 using the photo-interpretation technique. Like the estimates of DBH using the regression models (table 11), ocular estimates of DBH classes 3 and 4 were fairly accurate (user's accuracies of 67% and 50%, respectively); again this was encouraging since the majority of the stands in the study area are probably in these classes.

Table 13.--Error matrix table of the aerial photo estimates of stand height

Reference data

	Class	1	2	3	4	5	6	Row total
	1				1			1
	2			5	1			6
Photo-	3			9	6	2		17
interpreted	4			5	13	2		20
data	5				3	3		6
	6				1			1
	Column total	0	0	19	25	7	0	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	-	-
3	53	47
4	65	52
5	50	43
6	-	-

Overall accuracy (P_O) = 0.49

Tau (T_e) = 0.29

Kappa (K) = 0.14

Estimates of stand height from the photos were generally low, i.e. stands were frequently placed in lower classes than they actually were on the ground. Seventeen stands were underestimated, mostly in height classes 2 and 3, while 9 stands were overestimated in classes 4 and 5 (table 13). However, the user's accuracies for stand height classes 3 and 4 were greater than 50% and 86% of the stands were in these two classes.

Table 14.--Error matrix table of the aerial photo estimates of visible crown diameter

Reference data

Photo-interpreted data	Class	1	2	3	4	5	Row total
	1			1			1
	2	1	2	4	2		9
	3		1	15	2	2	20
	4			2	14	2	18
	5				3		3
	Column total	1	3	22	21	4	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	22	66
3	75	68
4	78	67
5	-	-

Overall accuracy (P_o) = 0.65

Tau (T_e) = 0.51

Kappa (K) = 0.42

Photo estimates of average crown diameter also tended to be low. Thirteen stands were underestimated while 6 were overestimated (table 14). There was considerable underestimation of crown diameter class 2; only 2 of nine stands were correctly classified while 6 of the remaining 9 were actually in classes 3 and 4. But, like the estimates of DBH and height, the user's accuracies were high in the classes that contained the majority of the stands (classes 3 and 4).

Table 15.--Error matrix table of the aerial photo estimates of canopy cover

		Reference data			
	Class	1	2	3	Row total
Photo interpreted data	1		4		4
	2		13	9	22
	3		11	14	25
	Column total	0	28	23	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	59	46
3	56	61

Overall accuracy (P_0) = 0.53

Tau (T_e) = 0.29

Kappa (K) = 0.13

Estimates of canopy cover were nearly balanced, with 12 stands underestimated and 11 overestimated (table 15). All of the stands were either in class 2 or 3. Most of the underestimated stands were in class 3 and all of the overestimated stands were in class 2 (11 stands misclassified as class 3); this indicates that it was difficult to distinguish the break between these classes. Accuracy is low considering that there were basically just two classes; random assignment to these classes would result in 50% overall accuracy.

Table 16.--Error matrix table of the aerial photo estimates of stand canopy layers

	Reference data					
	Class	1	2	3	4	Row total
Photo interpreted data	1	30	14			44
	2		2	1		3
	3			1		1
	4	1	2			3
	Column total	31	18	2	0	51

Class	User's accuracy	Producer's accuracy
1	68	96
2	67	11
3	100	50
4	-	-

Overall accuracy (P_0) = 0.63

Tau (T_e) = 0.50

Kappa (K) = 0.18

Most of the sample stands were single layered (class 1) and there was substantial misclassification between 1 and 2 layered stands (table 16). Fourteen stands were classified as 1 layered when they were actually 2 layered, indicating that it was difficult to detect the second layer of stand canopy. This also reflects the 'when in doubt, classify as a 1 layer stand' philosophy that was adopted because the majority of the stands seem to be 1 layered based on experience gained in the pilot study. The high user's accuracy for classes 2 and 3 is somewhat misleading because of the low number stands classified into class 2.

Table 17.--Error matrix table of the aerial photo estimates of stand cover type.

Reference data

	Class	A	D	G	L	P	S	W	Row total
	A	6	1		1				8
	D		20	2		3		2	27
Photo interpreted data	G								
	L	1	3		5			2	11
	P		2			1			3
	S						1		1
	W		1						·1
	Column total	7	27	2	6	4	1	4	51

Class	User's accuracy	Producer's accuracy
A	75	86
D	74	74
G	-	-
L	45	83
P	33	25
S	100	100
W	-	-

A = subalpine fir; D = Douglas fir; G = grand fir; L = lodgepole pine;
P = ponderosa pine; S = Engelmann spruce;
W = western larch.

Overall accuracy (P_o) = 0.64

Tau(T_e) = 0.59

Kappa (K) = 0.47

The majority of the sample stands were of the Douglas fir (D) cover type (27 out of 51); user's accuracy for Douglas fir was high (75%). User's accuracy for subalpine fir was also high, this may be because subalpine fir tend to have a distinctly pointed and tapered crown compared to the other cover type species and therefore were easier to distinguish on the photos. Accuracy percentages for the other cover types are deceiving because of low sample numbers in these types; one or two misclassified stands can have a significant impact on the percentages. Five stands were classified as Douglas fir when they actually were other cover types; conversely 6 stands erroneously classified as other cover types when actually they were Douglas fir.

Table 18.--Error matrix table comparing aerial photo estimates of stand spatial aggregation

	Reference data				
	Class	I	R	U	Row total
Photo-interpreted data	I	14	3	7	24
	R		3		3
	U	6		18	24
	Column total	20	6	25	51

Class	User's accuracy	Producer's accuracy
I	58	70
R	100	50
U	75	72

I = Irregular, R = Regular, U = Uniform

Overall accuracy (P_o) = 0.65

Tau (T_e) = 0.53

Kappa (K) = 0.46

For the spatial aggregation data (table 18), 6 stands were classified as being *uniform* while ground survey placed them in the *irregular* class. Conversely, 7 stands were classified as *irregular* but were actually *uniform* in the ground survey. The high user's accuracy for *regular* stands was deceptive since there were only 3 stands classified in this group.

B. Error analysis

The error analysis using the scatterplots (see fig. 5) was inconclusive; misestimated stands did not exhibit any tendency to be near the edge or the middle of the images, or in areas of steep topography. The scatterplots are presented in appendix E. Trends in underestimation and overestimation seen in the error matrix tables were also observed in each respective scatterplot.

Misestimated stands also did not have significantly different within stand variation for most variables from those stands that were estimated accurately (see appendix F). The exception was that stands underestimated for DBH (ocular estimation) had a significantly higher mean CV than those stands that were overestimated and accurately estimated.

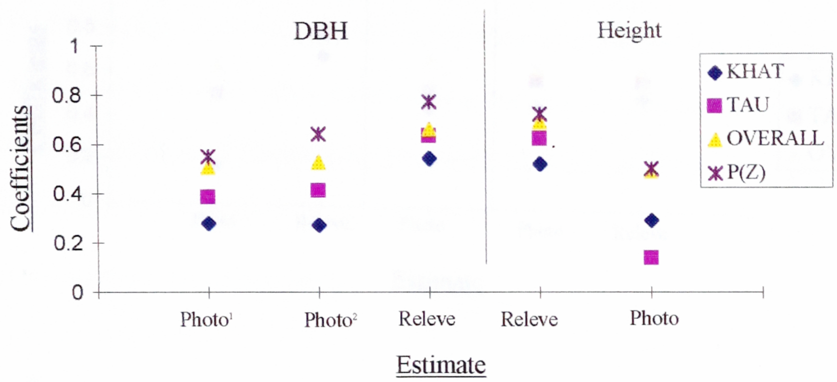
Mean CV values for stands in the three photo-interpreted spatial aggregation categories were not significantly different (see appendix F) for all variables.

C. Summary of accuracy measures

Two consistent trends can be seen in scatterplots of the accuracy coefficients (fig. 7). For many of the variables, relevé estimates were higher than the corresponding photo estimates, except for crown diameter and cover type. The P(Z) values were substantially higher than the corresponding error matrix accuracy measures.

To simplify analysis, it was necessary to eliminate some of the accuracy measures from consideration. For the interval scale variables, the chi-square probability {P(Z)} was calculated. For all variables, three coefficients were calculated from the error matrix tables: Kappa (K), Tau (T_e), and Overall accuracy (P_o). One accuracy measure from the error matrices was selected for analysis.

The Tau coefficient was selected because of its ability to compensate for chance agreement. P_o does not account for the chance placement of a stand into the correct cell in the matrix and therefore tends to overestimate accuracy (Ma and Redmond, 1995; Congalton and Mead, 1983). Foody (1992) demonstrated that the Kappa (K) coefficient overcompensates for chance agreement and thus under represents classification accuracy. When the three coefficients are calculated from the same matrix, K tends to be the highest value, P_o the lowest, and T_e falling somewhere in between (Ma and Redmond, 1995). The same pattern was noted in this study (fig 7). T_e is an improvement over K because it compensates for random chance agreement *and* for actual correct classification (Foody, 1992; Ma and Redmond, 1995).



[1]/ DBH estimate using the regression models
[2]/ DBH - ocular estimation

Fig. 7a

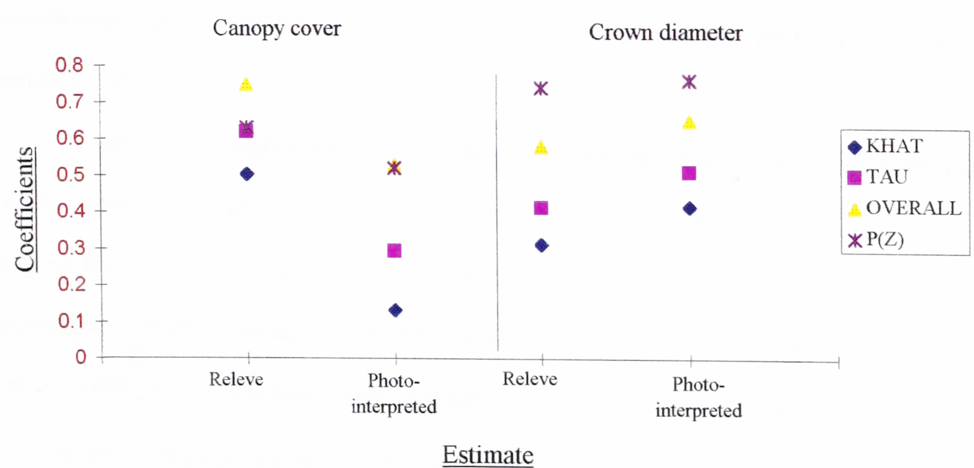


Fig. 7b

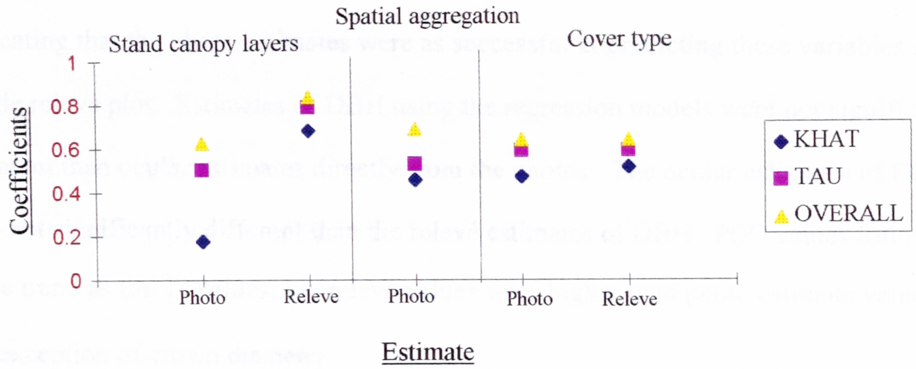


Fig. 7c

Figures 7a-c.--Scatterplots comparing the different accuracy measures. The chi-square test was not applicable to the nominal data (fig 7c).

The Tau coefficient (T_e) and the Z probability corresponding to the chi-square test of a hypothesized variance [P(Z)] were the accuracy measures used to evaluate the estimation techniques (table 19).

A critical assumption in the use of the chi-square technique is normal distribution of the data; normal probability plots of each variable in the dataset did not reveal any major deviations from the normal distribution. Removal of bias was conducted if it resulted in an improvement of 0.05 or greater for the P(Z) value; bias was removed for the photo-interpreted variables DBH (ocular estimation and modeled), height, and crown diameter.

Relevé estimates for most of the variables were more accurate than estimates from the photos; this trend was also seen in fig. 7. However, estimates of crown diameter, DBH (ocular estimate), and cover type were not significantly different at the 5% level (table 19), indicating that the photo estimates were as successful at predicting these variables as the single relevé plot. Estimates of DBH using the regression models were not significantly different than ocular estimates directly from the photos. The ocular estimates of DBH were also not significantly different than the relevé estimates of DBH. P(Z) values followed the same trend as the T_e values; i.e. relevé values were higher than photo estimate values, with the exception of crown diameter.

Table 19.--Accuracy coefficients for each of the stand structure variables derived by the two methods. This table summarizes the accuracy coefficients for each of the variables estimated from the relevé plots and from the photos.

Variable	Photo-interpreted accuracy	Relevé plot accuracy
DBH	(Estimate) $T_e = 0.41^{a,b}$ $P(Z) = 0.69^*$ (Modeled) $T_e = 0.39^a$ $P(Z) = 0.64^*$	$T_e = 0.63^b$ $P(Z) = 0.77$
Height	$T_e = 0.29^a$ $P(Z) = 0.55^*$	$T_e = 0.62^b$ $P(Z) = 0.72$
Crown diameter	$T_e = 0.51^a$ $P(Z) = 0.84^*$	$T_e = 0.41^a$ $P(Z) = 0.74$
Canopy cover	$T_e = 0.29^a$ $P(Z) = 0.52$	$T_e = 0.62^b$ $P(Z) = 0.68^*$
Canopy layers	$T_e = 0.50^a$	$T_e = 0.79^b$
Cover type	$T_e = 0.59^a$	$T_e = 0.59^a$
Spatial Aggregation	$T_e = 0.53$	N/A

Note: P(Z) values where mean sum of squares bias was removed (appendix C) are designated with an asterisk (*). The Kappa coefficients for equal prior probability (T_e) that are significantly different at the 5% level are designated by different superscript letters. P(Z) values are described in appendix C. Spatial aggregation was not determined from the relevé plots.

Chapter IV: Discussion

The transfer of the stand delineations from the maps provided by the Salish/Kootenai Tribes to the aerial photos proved to be somewhat problematic. Stand lines on the maps had been rectified, a process where variation in scale and image displacement from the photos is corrected. This process usually entails transfer of the linework from the original photos with stand delineations to orthophotos. It was necessary to transfer these lines from the maps back to the photos; basically to 'unrectifiy' the lines. The original photos (taken 1980) with the stand delineations provided some help, but management activities since 1980 changed many of the stand boundaries. Transfer of the lines by hand worked fairly well; most stand lines followed natural breaks in the topography or along cutting unit boundaries. But some lines were difficult to transfer because there was no obvious distinct boundary between stands on the photo and it was difficult to compensate for radial displacement, especially in steep terrain. This may have lead to some stands being more heterogeneous in structure; errors in line transfer could have resulted in stand inclusions different in structure than the rest of the stand. In many instances this problem was averted by moving stand boundaries or delineating new stands.

The error matrix classes used in this study were selected based upon their utility to Salish/Kootenai forestry personnel and because the class boundaries seemed logical for this region. Selection of class breaks and the number of classes can have a significant effect on error matrix accuracy coefficients. Tables 11 - 18 show that the majority of stands are in 1 or 2 of the classes with a few stands scattered among the remaining classes. In classes where there are few stands, just 1 or 2 misclassified stands can have a big affect on the user's or producer's accuracy. A larger number of sample stands would presumably add stands in these underrepresented classes and increase reliability of the producer's and user's accuracy figures. However, the user's accuracies for the classes where most of stands were placed by the reference data were fairly high for most of the variables, indicating successful placement of most of the stands across the landscape.

The relatively small sample size in this study may somewhat invalidate the error matrix results. Hay (1979) states that to adequately address errors using an error matrix, a sample size of 50 is needed in each class, based upon the specified confidence limits for the actual accuracy percentage. Many of the classes in this study contain less than 10 stands; the 95% confidence interval for a sample size of 10 would be so large that it was not possible with any measure of confidence to place the actual accuracy percentage. Error matrices are predominately used in assessing the accuracy of satellite digital imagery, where an area the size of Finley Creek MA would contain thousands of 30 meter pixels. In this situation, a stratified sample of 50 pixels per class is not unreasonable, but in this study the unit of resolution was forested polygons and the total population was 270; an error matrix with 5 classes would require 250 sample stands. The accuracy percentages obtained from the error matrices may not be conclusive statistically, but they do provide the interpreter with an idea of where errors are being committed and which classes are being confused. In landscape assessments where the unit of resolution is the stand or patch, the chi-square analysis technique may be more appropriate, unless the practicality of field inventory of a large number of stands is not a concern. This technique only applies to interval variables; attributes such as cover type would still need analysis with an error matrix table.

When using the linear regression model [DBH = a + b(crown diameter) + c(height)] to estimate stand DBH from photo-interpreted estimates of crown diameter, there are a number of potential sources of error. Estimates of crown diameter and height from the photos could be inaccurate and there is error associated with the models. Because the two dependent variables (crown diameter and height) are underestimated, it was no surprise that the model DBH estimates are also underestimated.

The ocular estimates of DBH have a nearly equal number of overestimated and underestimated stands (see table 12 and appendix E, fig. 12). It was interesting that while the chi-square analysis detected significant bias, both the error matrix and the scatterplot (appendix E) show no particular bias toward underestimation or overestimation. This was probably due to the use of the residual sum of squares bias term in the chi-square formula (see appendix C); this term removes bias that varies with the ground truth values rather than constant bias throughout the range of values. As an example, say there is an interpreter who consistently underestimates the canopy cover of dense stands and overestimates the canopy cover of open stands. Rather than a constant level of bias, where all trees are overestimated for example, this type of bias varies across the range of data. The chi-square technique removes this type of bias.

No significant difference at the 5% level between the ocular estimates of DBH and the modeled estimates indicates that taking the time to develop models and measure height and crown diameter from the photos did not result in an improvement in accuracy over ocular estimates of DBH. Modeling estimates of DBH may prove to be more accurate where the interpreter does not have familiarity with the study area or is working with different scale images.

Tree height estimation errors using a stereometer are reported to be inconsistent for trees of different sizes; trees smaller in height tend to be underestimated while taller trees tend to be overestimated (Maclean and Pope, 1961). Examination of the scatterplot (see appendix E, fig. 15) also shows this pattern; all the stands less than 12.2 m (40 feet) in height were underestimated while a large percentage of those over 21.3 m (70 feet) were overestimated.

The tendency to underestimate crown diameter seen in this study was also reported by Spurr (1960). He states that thin branches that are seen from the ground cannot be resolved on the photos, causing an underestimation of crown diameter. Removal of the bias toward underestimation in the chi-square technique results in a substantial improvement in the P(Z) value; the lack of substantial bias in the P(Z) calculation for the relevé estimate indicates that this measurement bias was not present in the relevé estimate technique.

Detection of significant bias in the chi-square calculations for photo estimates of DBH (ocular estimate and modeled), height, and crown diameter may be partially due to the growth of the trees since the time the photos were taken. The photos were taken in August of 1990 and field inventory was conducted in August - September of 1995; appreciable growth may have occurred in some of the stands.

Canopy cover showed no bias in the chi-square test and the scatterplot also shows a balance between underestimated and overestimated stands. The steep topography in parts of the study area resulted in shadows and changes in resolution that may have lead to an overestimation of canopy cover from the photos; this overestimation may be compensated by the tendency to underestimate canopy cover from the ground (Spurr, 1960). Estimates of canopy cover from aerial photographs may be closer to "truth" than estimations from the ground; the percentage of ground obscured by overstory canopy for the entire stand was easier to visualize from an aerial perspective than from a series of plots from the ground.

This discrepancy may be partially responsible for the low reported accuracy for this variable.

The error matrix table analysis for canopy cover provides an interesting example of the effect of the number of classes, class widths, and class boundaries on accuracy. Intuitively, we would expect that breaking down an interval variable like canopy cover into many different classes with small class widths would result in a lower overall accuracy than a smaller number of classes with large class widths. Changes in the reported accuracy that are a product of variation in the number of classes is one form of the "modifiable areal unit problem" (Openshaw, 1987). Different reported accuracies may not only be related to the number of classes; the class breaks and distribution of the data in the error matrix need to be considered. For example, 36 of the 51 sample stands had ground truth canopy coverages ranging from 50% - 70% (figure 8).

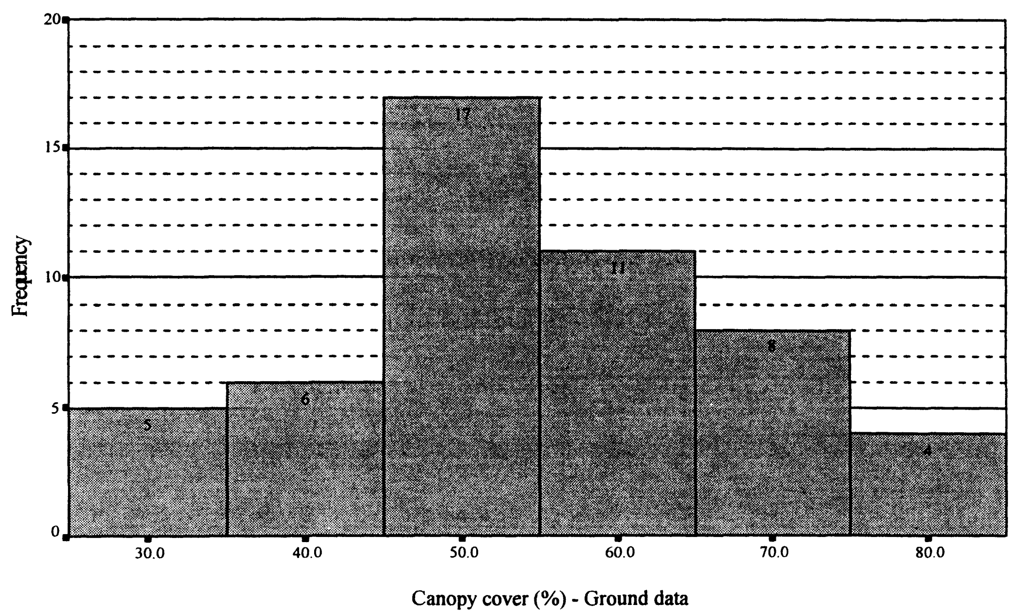


Figure 8.-- Frequency distribution (histogram) of the ground truth canopy cover data. The number of stands in each 10% canopy cover class are displayed in the histogram bars.

The class break between classes 2 and 3 was 60% (table 9); since the majority of the sample stands are clustered around the class break, relatively minor errors in canopy cover estimation (10%) may be responsible for the low reported accuracy. If the class breaks occurred at 25%, 40% and 70%, then the accuracy coefficients for the canopy cover data would probably have been higher because for the majority of the sample stands, small (10%) errors in estimation would not place a stand in the wrong class. In short, manipulating both the number and width of classes and the class breaks, can have a significant impact on reported accuracy from an error matrix. This is important to consider not only when collapsing interval data into classes, but also when defining classes for data collection.

Detecting multiple canopy layers on the photos was difficult and often depended more on the topography, elevation, and aspect of the stand rather than the texture, tone or patterns seen on the photograph. Local knowledge of the plant relationships to physical site characteristics also played an important role. The reported accuracy ($T_e = 0.50$) was fairly high considering the difficulties described above. Perhaps greater familiarity with the vegetation conditions and how they relate to the physical characteristics of the stands would have resulted in an increase in accuracy for this variable.

Over half of the sample stands were determined to be in the Douglas fir cover type using the ground truth data; the remaining 6 cover types had 6 or fewer stands in each cover type class. Single species cover types were chosen because of the large variety of species mixes that occur in the study area; the presumption was that it would be easier to identify the major species rather than try to define species mixes by canopy cover composition. Additionally, the variety of species mixes may have required a large number of cover types in the error matrix. Like stand layers, the location of the stand was often critical in making cover type calls from the photos. Accuracy ($T_e = 0.59$) is high when the difficulty of determining the dominant cover type species from the many mixed species stands is considered; the accuracy reported for cover type is not significantly different from that reported by Deegan and Befort (1990), who analyzed data from 1:15,840 scale black and white infrared photos and ground plots in northern Minnesota ($T_e = 0.54$). Martin and Gerlach (1983) report higher accuracy predicting cover types ($T_e = 0.71$) using multiple regression models from 1:24,000 scale photo interpreted variables.

Spatial aggregation was an attempt to measure the horizontal heterogeneity, or patchiness, of the stands. The categories were designed to represent the within-stand patterns created by disturbance (natural or man-caused) in the Finley Creek MA. Most of the patchiness observed in the stands seemed to be the result of timber harvest or disease-induced canopy openings 0.2 ha (0.5 acre) or greater in size; these types of stands were usually placed in the *irregular* class. Stands classified as *regular* were typically park-like stands of ponderosa pine and Douglas fir in the lower elevations of the study area, or stands that had been partially harvested. The high number of stands classified as *uniform* seemed to be because of invasion of openings by Douglas fir in stands that would have been placed in the *regular* class. Spatial aggregation is probably more accurately determined from the aerial photos; in this case the 'truth' dataset should probably be the photo-interpreted dataset. Estimation of spatial aggregation from the ground was done after all six plots were visited; if the survey route did not take the observer through a large opening seen on the photo, then the stand may be placed in the *uniform* or *regular* classes instead of the *irregular* class. This may be the source of error for 10 of the 16 misclassified stands in table 18. Spatial aggregation should correspond closely with the coefficient of variation (CV) for canopy cover; stands having high CV values for canopy cover would be classified in the *irregular* and *regular* classes. Results of this analysis were inconclusive; mean CV values in the three spatial aggregation classes were not significantly different (see appendix F). Not enough stands were classified as *regular* to discern any differences statistically.

There is a concern that forest structure variables measured on the ground cannot be related to data from the photos, due to the complexities of the forest canopy (Paine, 1981; Worley and Meyer, 1955). This may be true, however, quantitative data is only one piece of information that the interpreter uses to estimate ground conditions. The *art* of photo-interpretation takes measurements from the photos and combines them with the tone, color, texture, pattern on the image, and most importantly, the interpreter's local knowledge. The interpreter considers these factors when arriving at an estimate for a stand and it is impossible to sort out which one has the greatest influence on the interpreter's estimate. Another factor is the information gained from previous photo-interpretation efforts; for example if an interpreter knows that he or she has a tendency to misinterpret a specific cover type or overestimate tree height, then this information can be considered when making future estimates. In this research effort, the pilot study was an invaluable source of this information. At least 18 hours of interpretation training has been recommended (Getchell and Young, 1953), perhaps 20 hours were spent on interpretation techniques in the pilot study. A preliminary study such as this should be undertaken before any large landscape photo-interpretation effort.

The direct approach used in this study is in contrast to the methods used by Martin and Gerlach (1981) and Teuber (1983). Rather than making direct estimates of the ground attributes, these authors measured image features such as pattern, texture, height, and canopy cover; multiple regression models with these variables then predicted the ground attributes (volume, site index, yield, habitat type). In this indirect method, much of the *art* of photo-interpretation described above is replaced by the multiple regression models. An advantage to this method may be more consistent results between interpreters, since some of the subjectivity is removed; and it may be the only way to estimate site conditions like habitat type from aerial photos. However, no model can possibly account for all the information and combinations of factors better than the mind of the interpreter.

The observed tendency for the P(Z) coefficients to be higher than the corresponding Tau coefficient was probably related to the allowable error term in the chi-square equations. The P(Z) value is the probability that the estimate is within the allowable error of the ground estimate. Thus with a reported P(Z) value of 0.68, in 68 out of 100 stands we would expect the estimate to be within the allowable error of the ground value. Error matrix percentages (i.e., T_e) are the percent chance that a stand is in the correct class. They can also be interpreted as the percent improvement over a random placement of stands into cells in the error matrix table. It seems that the two measures should be closer than the results (table 19) indicate; the higher P(Z) values reported may be because this technique considers the difference between each variable pair (estimated - observed) whereas the error matrix technique lumps the interval variable pairs into categories. In directly comparing the predicted versus estimated values, the effects of gross estimation errors in any one stand may be smoothed over by other, more accurately interpreted stands.

The gap between the two measures is also a product of the class boundaries (error matrix tables) and the allowable error (chi-square test). Adjustment of these parameters, especially allowable error, can narrow the gap between the accuracy measures. For example, raising the allowable error for canopy cover from 10% to 20% increases the corresponding P(Z) values from 0.52 to 0.84. It is important to realize the impact that these parameters have on the accuracy measures and on the ability to compare results to other studies. The Tau coefficient overcomes a major limitation with the other error matrix coefficients in that it can be used to compare matrices having a different number of classes (Ma, pers. comm.).

The obvious difference between the two measures is that one measures interval scale data (chi-square) and the other measures ordinal scale data. Collecting interval scale data allows the analyst to collapse the data into many different error matrix schemes. Thus if a particular model requires different class breaks, interval data can be re-collapsed into the appropriate categories. This freedom is lost if the data was collected in ordinal form. Interval data can be collapsed into many classes, or just a few. The number of classes in an error matrix can affect the accuracy; the same data aggregated into a different number of classes may result in different accuracy coefficient values (modifiable areal unit problem [Openshaw, 1987]). It is interesting to speculate on whether collecting data in an ordinal form would have any effect on error matrix accuracy. Placing stands directly into categories seems to be popular (Lehmkuhl et al. 1994), probably because of the relative ease and speed with which stands can be interpreted. The decision on what measurement scale to collect the data is important and should be considered before any landscape assessment. Because of the problem of low sample sizes in the error matrices (see discussion above), greater confidence may be placed in the P(Z) accuracy coefficients than in the Tau coefficients. Although direct comparison of the two accuracy assessment techniques is not possible, the success of the estimating techniques in predicting the ground values of the interval variables may be higher than the Tau coefficients indicate.

The higher accuracy coefficients of the relevé estimates for most of the variables seem to indicate the increased accuracy associated with a ground sampling method versus a photo sampling method. A question that a land manager must address when faced with the need for landscape level data is, does the increased accuracy of a ground based inventory method justify the extra expense? In this study, the investigator was able to photo-interpret 50 stands per day, collect relevé plot data on 4 stands per day, and conduct full field inventories on 2 stands per day.

A reasonable approach would be to combine remotely sensed data, existing ground data (e.g., stand exams, EcoData plots), and field survey data into the landscape assessment, as in Morrison (1994). Ground data could be used to conduct an assessment of the photo-interpreted stands provided that the inventory methods were compatible, or to train photo-interpreters before data collection from the photos.

There was some difficulty in selecting relevé plot locations from the photos. Some stands were very heterogeneous in structure and species composition; this made it difficult to select a plot that best represented the average conditions seen throughout the stand. This was probably the major source of misclassification for the relevé plots. Additionally, it was sometimes hard to locate the relevé plots on the ground, especially in stands with uniform vegetation cover.

The finding of no significant trends in the scatterplot error analysis may be partly related to the relatively small sample size. With only 51 sample stands, often there were just 10 - 12 stands that were misestimated; it is difficult to draw any conclusions based on

10 - 12 stands. A sample size of 50 stands *per class* and a study design focused on the detection of error biases would perhaps find conclusive results (Hay, 1979; see discussion above).

Chapter V: Summary/Conclusions

As more is learned about landscape level processes, there will be a greater need for efficient methods of collecting data across landscapes. Satellite image technology is progressing, but accurate classification of some forest structure attributes has not been attained (Spies, 1994; Cohen, 1994). Aerial photograph interpretation is a relatively low cost and low technology method that is within the means of most land management agencies. A multi-stage approach is probably best; satellite images may be used for data collection in broad classes across large areas, and aerial photographs for more specific data on mid-scale landscapes. For detailed, site-specific data, field inventory will be necessary.

The photo-interpretation methods used in this study were intended to be similar to those used by most land management agencies. Other methods of collecting information from aerial photographs (i.e., Martin and Gerlach, 1981, and Teuber, 1983) are certainly valid, but the methods used in this study seem to be the most common and were used so that land managers and researchers may benefit from the accuracy assessment and error analysis techniques.

No attempts are made to determine whether the accuracy of these methods is acceptable or not -- that is left to the reader. It is important to recognize that these results apply to this study area, to these images, and most importantly, to this interpreter. Accuracy standards may be scale dependent and should be developed in conjunction with project objectives.

Accuracy assessments should be conducted on all projects where data from remotely sensed images are used. The accuracy assessment methods described in this study could easily be implemented on most datasets. The number of field plots or sample stands may be restricted by expense, but as few as 50 plots (stands) may provide insight into errors and misclassifications. Existing stand inventory data may be used in assessing the accuracy of remotely sensed or modeled data; this would minimize the amount of new field data needed. Knowledge of the accuracy of remotely sensed data will give increased confidence in decisions based upon the data, and also provide feedback to improve future interpretation and classification projects.

Forest structure attributes frequently are the defining characteristics for landscape elements such as the patch, matrix and corridor (Forman, 1995). To meet the challenge of implementing ecosystem management, models such as SIMPPLE (Chew, 1995), FIRE-BGC (Keane et al, 1996) and FRAGSTATS (McGarigal and Marks, 1995) have been developed to help us relate the pattern of these landscape elements to biological processes. These models frequently utilize remotely sensed data of these forest landscape elements. The accuracy of these input data and the effect of errors on model output are frequently overlooked (Hess, 1994); application of some models may be pointless if the input data are not accurate to a certain extent. Land cover weighting schemes, which are often used in wildlife habitat models, can be adjusted based on observed classification error (Prisley and Smith, 1987). Further research into the effect of errors in spatial data on landscape models, and methods to adjust models based on these errors, is needed.

Land managers will continue to look to remote sensing technology as a cost effective way to obtain data as they assess forest resources and processes across larger landscapes. Awareness of the limitations of this technology and of potential inaccuracies in these data are critical factors to consider when decisions are to be made based on a landscape scale analysis.

Literature cited

Allen, C.D. 1994. Ecological perspective: Linking ecology, GIS, and remote sensing to ecosystem management. *In*: Remote Sensing and GIS in Ecosystem Management. V. A. Sample, ed. Island Press, Washington, DC. 369 pp.

American Society of Photogrammetry, 1960. Manual of Photographic Interpretation. ASP, Washington, DC. 868 pp.

Balice, R. 1979. Mapping forest habitat types in the Unita Mts., Utah, utilizing natural color and color infrared photography. *In:* Color aerial photography in the plant sciences and related fields: Proceedings of the seventh biennial workshop on color aerial photography in the plant sciences. American Society of Photogrammetry, Falls Church, VA. 254 pp.

Brown, J. 1978. Weight and density of crowns of rocky mountain conifers. USDA Forest Service, Intermountain Research Station, INT-197.

Chew, J. D. 1995. Development of a system for simulating vegetative patterns and processes at landscape scales. Ph.D. dissertation, University of Montana, Missoula, MT. 181 pp.

Cohen, W. B. 1994. GIS applications perspective: Current research on remote sensing of forest structure. *In*: Remote Sensing and GIS in Ecosystem Management. V.A. Sample, ed. Island Press, Washington, DC. 369 pp.

Congalton, R. G., R. G. Oderwald, and R. A. Mead. 1983. Assessing LANDSAT classification accuracy using discrete multivariate analysis statistical techniques. Photogrammetric Engineering and Remote Sensing, 49(12): 1671-1678.

Congalton, R. G. and R. A. Mead. 1983. A quantitative method to test for consistency and correctness in photo-interpretation. Photogrammetric Engineering and Remote Sensing, 49(1): 69-74.

Congalton, R. G. 1988. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54(5): 593-600.

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

Congalton, R. G., K. Green, and J. Teply. 1993. Mapping old growth forests on national forest and park lands in the Pacific Northwest from remotely sensed data. Photogrammetric Engineering and Remote Sensing, 59(4):529-535.

Deegan, G., and W. A. Befort. 1990. Photointerpretation accuracy across two decades of forest inventory. *In*: State of the art methodology of forest inventory: A symposium. V. LauBau and T. Cunia, eds. USDA Forest Service, PNW Research Station. PNW-GTR-263.

Foody, G. M. 1992. On the compensation for chance agreement in image classification accuracy assessment. Photogrammetric Engineering and Remote Sensing, 58(10): 1459-1560.

Forman, R. T. 1995. Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press, New York. 632 pp.

Freese, F. 1960. Testing accuracy. Forest Science 6(2):139-145.

Getchell, W. A. and H. E. Young. 1953. Length of time necessary to attain proficiency with height finders on air photos. University of Maine, Forestry Dept. Tech. Notes No. 23. Orono, Maine.

Gonzales, J. 1994. GIS applications perspective: Using remote sensing and GIS for modeling old-growth forests. *In*: Remote Sensing and GIS in Ecosystem Management. V.A. Sample, ed. Island Press, Washington, DC. 369 pp.

Green, K., S. Bernath, L. Lackey, M. Brunengo, and S. Smith. 1993. Analyzing the cumulative effects of forest practices: Where do we start? Geo. Info. Systems 3(2): 37-29.

Hagan, G. F. and J. F. Smith. 1986. Predicting tree groundline diameter from crown measurements made on 35mm aerial photography. Photogrammetric Engineering and Remote Sensing 53(5): 687-690.

Hall, R. J., R. N. Nesby and R. T. Morton. 1989. A comparison of existing models for DBH estimation from large scale photos. The Forestry Chronicle, April 1989: 114-120.

Hall, R. J., S. J. Titus and W. J. A. Volney. 1992. Estimating top-kill volumes with large scale photos on trees defoliated by the jack pine budworm. Can. Jour. For. Res. 23:1337-1346.

Hay, A.M. 1979. Sampling designs to test land-use map accuracy. Photogrammetric Engineering and Remote Sensing, 45(4):529-533.

Hess, G. 1994. Pattern and error in landscape ecology: A commentary. Landscape Ecology 9(1):3-5.

Kalkhan, M. A., T. J. Stohlgren, and M. B. Coughenour. 1995. An investigation of biodiversity and landscape-scale gap patterns using double sampling: A GIS approach. *In:* Proceedings of the Ninth Annual Symposium on Geographic Information Systems, Vancouver, BC March 27-30, 1995.

Keane, R. E., P. Morgan, and S. Running. 1996. FIRE-BGC -- A mechanistic ecological process model for simulating fire succession on coniferous forest landscapes of the northern Rocky Mountains. USDA Forest Service INT-RP-484.

Larson, F. R., Moessner, K. E., and P. F. Ffolliott. 1971. A comparison of aerial photo and ground measurements of ponderosa pine stands. USDA Forest Service Research Note RM - 192.

Lehmkuhl, J. F., L. F. Ruggiero, and P.A. Hall. 1992. Landscape-scale patterns of forest fragmentation and wildlife richness and abundance in the southern Washington Cascade range. *In*: Wildlife and vegetation of unmanaged Douglas fir forests. L. Ruggiero, ed. USDA Forest Service. GTR - PNW - 285.

Lehmkuhl, J. F., P. F. Hessburg, R. L. Everett, M. H. Huff and R. D. Ottmar. 1994. Historical and current forest landscapes of eastern Oregon and Washington. Part I: Vegetation pattern and insect and disease hazards. USDA Forest Service, PNW-GTR-328. East side forest ecosystem health assessment, Volume III.

Lillesand, T. M. and R. W. Keifer. 1994. Remote Sensing and Image Interpretation. John Wiley and Sons, Inc. New York, NY. 723 pp.

Ma, Zhenkui and R. L. Redmond. 1995. Tau coefficients for accuracy assessment of classification of remote sensing data. Photogrammetric Engineering and Remote Sensing, 61(4): 435-439.

Maclean, C.D. and R. B. Pope. 1961. Bias in the estimation of stand height from aerial photographs. The Forestry Chronicle 37(2): 160-161.

MacGarigal, K. and B. J. Marks. 1995. FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. USDA Forest Service PNW-GTR-351.

Martin, F. C. and Gerlach, F. L. 1981. Forest stand mapping and evaluation on state and private lands in Montana. Bulletin 46, Montana Forest and Conservation Experiment Station, School of Forestry, University of Montana, Missoula, Montana 59812.

Moessner, 1949. A crown density scale for photo interpreters. Journal of Forestry, 47:569

Moessner, K. E. and C. E. Jensen. 1951. Timber cruising on aerial photographs. Central States Forest Experiment Station Technical Paper No. 126, USDA Forest Service.

Moessner, K. E. 1964. Two aerial photo basal area tables. USDA Forest Service Research Note INT-23.

Morrison, P. H. 1994. GIS applications perspective: Development and analysis of a chronosequence of late-successional forest ecosystem data layers. *In*: Remote Sensing and GIS in Ecosystem Management. V.A. Sample, ed. Island Press, Washington, DC.

Mueller-Dombois, D. and H. Ellenberg. 1974. Aims and Methods in Vegetation Ecology. John Wiley and Sons, New York, NY. 547 pp.

Nyyssonen, A. 1957. On the estimation of growing stock from aerial photographs. Metsantytkimuslaioksen Julkaisuja 46 pp.

Openshaw, S. 1987. Learning to live with errors in spatial databases. *In:* Accuracy of spatial databases, M. Goodchild and S. Gopal, eds. Taylor and Francis, New York. 290 pp.

Paine, D. P. 1981. Aerial Photography And Image Interpretation For Resource Management. John Wiley and Sons, New York. 571 pp.

Prisley, S. P. and J. L. Smith. 1987. Using classification error matrices to improve the accuracy of weighted land-cover models. Photogrammetric Engineering and Remote Sensing, 53(9): 1259-1263.

Rutledge, L. T. and Hejl, S. J. 1990. Inventorying old-growth by aerial photograph analysis and ground surveys: An initial assessment. The Northwest Environmental Journal. 6(2):397-398.

Sayn-Wittgenstein, L. 1961. Recognition of tree species on air photographs by crown characteristics. Photogrammetric Engineering 27(5): 792-809.

Short, H. L. 1988. A habitat structure model for natural resource management. Journal of Env. Management, (27) 289-305.

Spies, T. A. 1994. Ecological perspective: The nature of mature and old-growth forest ecosystems. *In*: Remote Sensing and GIS in Ecosystem Management. V.A. Sample, ed. Island Press, Washington, DC. 369 pp.

Spurr, S. H. 1960. Photogrammetry and Photo Interpretation. The Ronald Press, New York. 472 pp.

Story, M. and R. G. Congalton. 1986. Accuracy assessment: A user's perspective. Photogrammetric Engineering and Remote Sensing 52(3):397-399.

Sweet, M. D. and K. E. Wall. 1995. Spatial and temporal modeling of forest resources at the landscape-level on the Bitterroot National Forest. Final report for USDA Forest Service Agreement No. INT-94914-RJVA.

Teuber, K. B. 1983. Aerial photographic interpretation of forest stands on Lubrecht experimental forest, MT. Masters' Thesis, University of Montana, Missoula, MT. 157 pp.

Velleman, P. F. 1993. Learning Data Analysis With Data Desk. W. H. Freeman and Co., New York. 337 pp.

Worley, D. P., and Landis, G. H. 1954. The accuracy of height measurements with parallax instruments on 1:12,000 scale photographs. Photogrammetric Engineering 20(5): 823-829.

Worley, D. P., and Meyer, H. A. 1955. Measurement of crown diameter and crown cover and their accuracy on 1:12000 scale photographs. Photogrammetric Engineering, 21(3): 372-375.

Appendices

Appendix A: Pilot study plot error matrices

Tables 20 - 31 are error matrices comparing the photo and CFI estimates of the stand attributes to the ground values. Producer's and User's accuracies are expressed in percentages. Table 24 provides descriptions of the DBH and crown diameter classes.

Table 20.--Error matrix table of the photo estimates of plot canopy layers

	Reference data					
	Class*	1	2	3	4	Row total
Photo-data	1	12	2		2	16
	2		1	1		2
	3					0
	4					0
	Column total	12	3	1	2	18

Class*	User's accuracy	Producer's accuracy
1	75	100
2	50	66
3	-	-
4	-	-

(*) 1 = 1 layer, 2 = 2 layer, 3 = 3 layer,
4 = no distinct layers

Overall accuracy (P_O) = 0.72

Table 21.--Error matrix table of the CFI estimates of plot canopy layers

Reference data

	Class*	1	2	3	4	Row total
	1	5				5
CFI	2	4	3		1	8
data	3	3		1		4
	4				·	0
	Column total	12	3	1	1	17

Class*	User's accuracy	Producer's accuracy
1	100	42
2	38	100
3	-	-
4	-	-

(*) 1 = 1 layer, 2 = 2 layer, 3 = 3 layer,
4 = no distinct layers

Overall accuracy (P_O) = 0.44

Table 22.--Error matrix table of the CFI estimates of stand cover type.

Reference data

	Class	DF	LP	PP	WL	Row total
	DF	6		1		7
CFI	LP		2			2
data	PP			6		6
	WL				2	2
	Column total	6	2	7	2	17

Class	User's accuracy	Producer's accuracy
DF	85	100
LP	100	100
PP	100	85
WL	100	100

Overall accuracy (P_O) = 0.94

Table 23.--Error matrix table of the photo estimates of plot cover type

Reference data

	Class	DF	LP	PP	WL	Row total
Photo-data	DF	3		3		6
	LP		2			2
	PP	4		3		7
	WL			1	2	3
	Column total	7	2	7	2	18

Class	User's accuracy	Producer's accuracy
DF	50	43
LP	100	100
PP	43	43
WL	66	100

Overall accuracy (P_O) = 0.55

Table 24.--DBH and crown diameter classes for error matrix tables

	DBH (cm)		Crown diameter (m)
1.	< 12.7 (5")	1.	< 1.5 (5')
2.	12.8 - 22.8 (8.9")	2.	1.6 - 2.7 (8.9')
3.	22.9 - 38.1 (14.99")	3.	2.8 - 4.5 (14.9')
4.	38.2 - 53.3 (20.9")	4.	4.6 - 6.4 (20.9')
5.	53.4 - 63.5 (25")	5.	> 6.4
6.	> 63.5		

Table 25.--Error matrix table of the photo ocular estimates of plot DBH

Reference data

Photo data: Class	1	2	3	4	5	6	Row total
1							
2				1			1
3			1	2	1		4
4				6	2		8
5				2	1		3
6					1		1
Column total			1	11	5		17

Class	User's accuracy	Producer's accuracy
1	-	-
2	-	-
3	25	100
4	75	54
5	33	20
6	-	-

Overall accuracy (P_O) = 0.44

Table 26.--Error matrix table of the photo regression estimates of plot DBH

Reference data

Photo data: Class	1	2	3	4	5	6	Row total
1				1			1
2				1			1
3			1	2	1		4
4				7	3		10
5					1		1
6							
Column total			1	11	5		17

Class	User's accuracy	Producer's accuracy
1	-	-
2	-	-
3	25	100
4	70	64
5	100	20
6	-	-

Overall accuracy (P_O) = 0.50

Table 27.--Error matrix table of the CFI estimates of plot DBH

Reference data

	Class	1	2	3	4	5	6	Row total
CFI data	1							
	2							
	3			1	3			4
	4				8	4		12
	5					1		1
	6							
	Column total			1	11	5		17

Class	User's accuracy	Producer's accuracy
1	-	-
2	-	-
3	25	100
4	66	73
5	100	20
6	-	-

Overall accuracy (P_O) = 0.59

Table 28.--Error matrix table of the photo estimates of plot understory cover type

Reference data

	Class	G	S	T	Row total
Photo data	G			3	3
	S		2	4	6
	T			3	3
	Column total		2	10	12

Class	User's accuracy	Producer's accuracy
G	-	-
S	33	100
T	100	30

G = Grass, S = Shrub, T = Trees

Overall accuracy (P_O) = 0.42

Table 29.--Error matrix table of the photo estimates of plot crown diameter

Reference data

	Class	1	2	3	4	5	Row total
Photo data	1	2	1				3
	2						
	3			4	1		5
	4			3	4		7
	5				1	1	2
	Column total	2	1	7	6	1	17

Class	User's accuracy	Producer's accuracy
1	66	100
2	-	-
3	80	57
4	57	66
5	50	100

Overall accuracy (P_O) = 0.64

Table 30.--Error matrix table of the photo estimates of plot canopy cover

Reference data

	Class*	10	20	30	40	50	60	70	80	90	Row total
	10	1	1								2
	20										
Photo data	30										
	40						1	1			2
	50				2	1	1	2			6
	60					1	1				2
	70							2			2
	80							1	1	1	3
	90									1	1
	Column total	1	1		2	2	3	6	1	2	18

Class*	User's accuracy	Producer's accuracy
10	50	100
20	-	100
30	-	-
40	-	-
50	16	50
60	50	30
70	100	33
80	33	100
90	100	50

* 10% canopy cover classes

Overall accuracy (P_O) = 0.38

Table 31.--Error matrix table of the photo estimates of plot height

Reference data

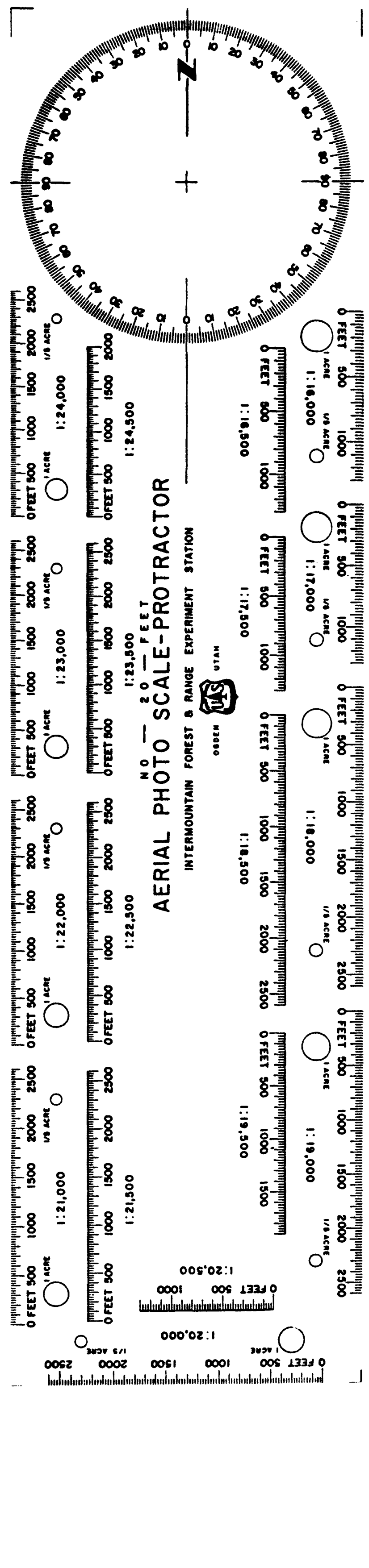
	Class*	10	20	30	40	50	60	70	80	90	Row total
	10										1
	20	1									
Photo	30										
data	40				1	1					2
	50					2					3
	60				1	1	1				3
	70				2	3	2				7
	80										
	90				1						1
	Column total	1		1	5	7	3				17

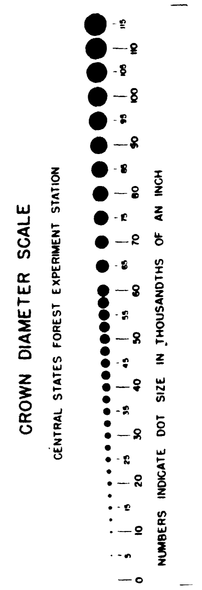
Class*	User's accuracy	Producer's accuracy
10	-	-
20	-	-
30	-	-
40	50	20
50	66	28
60	33	30
70	-	-
80	-	-
90	-	-

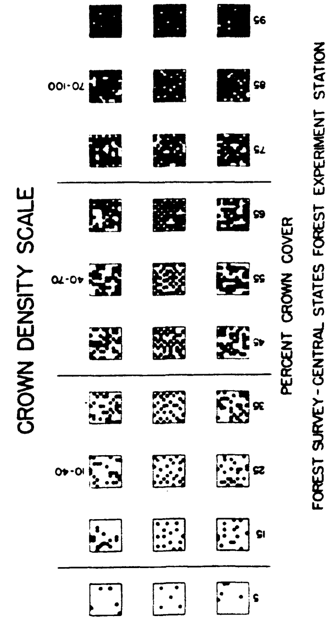
* 3.05 m (10 foot) height classes

Overall accuracy (P_O) = 0.22

Appendix B: Estimation templates







Appendix C: Statistical methods

Chi-square test of a hypothesized variance

The standard chi-square test of a hypothesized variance equation as presented by Freese (1960) is described below:

(1.0)
$$X^2_{n,\alpha} = \frac{\sum_{i=1}^{n} (x_i - \mu_i)^2}{\sigma^2}$$

Where μ_i = the value of the ith observation as determined by the estimate,

X_i = the "true" (ground truth) value of the ith observation, and

n = the number of units observed (sample size).

The accuracy as specified by the user is defined as:

(1.1)
$$\sigma^2 = \left[\frac{E^2}{\left(1/Z\right)^2} \right]$$

E is the allowable error specified by the user in the same measurement units as the estimating technique and Z is the standard normal value corresponding to the two tailed probability specified by the user.

Substituting eq. 1.1 into eq. 1.0:

$$X^2_{n,\alpha} = \frac{\sum_{i=1}^{n} (x_i - \mu_i)^2}{\left(\frac{E^2}{Z^2}\right)} \tag{1.2}$$

And rearranging algebraically, solving for Z:

$$Z = \sqrt{\frac{(E^2)(X^2_{n,\alpha})}{\sum_{i=1}^{n} (x_i - \mu_i)^2}} \tag{1.3}$$

Inclusion of the bias term nB^2, where $B = \frac{1}{n}\sum_{i=1}^{n} d_i = (\bar{x} - \bar{\mu})$, in the equation removes bias that is consistent with all values of μ_i:

$$Z = \sqrt{\frac{(E^2)(X^2_{n-1,\alpha})}{\sum_{i=1}^{n} (x_i - \mu_i)^2 - nB^2}} \tag{1.4}$$

Bias may also fluctuate with increasing or decreasing values of μ_i; removal of this type of bias involves replacing the denominator in eq. 1.4 with the residual sum of squares (SS_R):

$$Z = \sqrt{\frac{(E^2)(X^2_{n-2,\alpha})}{SS_R}} \tag{1.5}$$

The chi-square values $\left(X^2_{df}\right)$ at the 0.05% probability level were obtained from standard tables of the chi-square distribution (Koopmans, 1987).

Using values for μ_i and X_i from the appropriate datasets and the allowable error terms listed in Table 7, Equations 1.3 - 1.5 calculate the standard normal critical value (Z). This value corresponds to a two-tailed probability P(Z) which can be found in a table of standard normal probabilities (Koopmans, 1987). The P(Z) values reported can be interpreted as the probability at the 5% level that the estimate is within the allowable error of the ground values.

Appendix D: Relevé plot error matrices

Tables 32 - 37 are error matrices comparing the relevé estimates of the stand attributes to the ground values. Producer's and User's accuracies are expressed in percentages. Refer to Table 9 for descriptions of classes.

Table 32.--Error matrix table of the relevé estimates of stand canopy layers.

		Reference data				
	Class	1	2	3	4	Row total
Relevé data	1	27	2	1		30
	2	4	16	1		21
	3					0
	4					0
	Column total	31	18	2	0	51

Class	User's accuracy	Producer's accuracy
1	90	87
2	76	89
3	-	-
4	-	-

Overall accuracy (P_o) = 0.84
Tau (T_e) = 0.79
Kappa (K) = 0.68

Table 33.--Error matrix table of the relevé estimates of stand DBH

Reference data

Relevé data

Class	1	2	3	4	5	Row total
1						0
2	1	3	1			5
3			17	4		21
4			7	13	1	21
5				1	3	4
Column total	1	3	25	18	4	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	60	100
3	81	68
4	62	72
5	75	75

Overall accuracy (P_o) = 0.66

Tau (T_e) = 0.63

Kappa (K) = 0.54

Table 34.--Error matrix table of the relevé estimates of stand canopy cover

	Reference data				
	Class	1	2	3	Row total
	1				
Relevé data	2		17	2	19
	3		11	21	32
	Column total		28	23	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	89	61
3	66	91

Overall accuracy (P_o) = 0.75
Tau (T_e) = 0.62
Kappa (K) = 0.50

Table 35.--Error matrix table of the relevé estimates of crown diameter

Reference data

Relevé data

Class	1	2	3	4	5	Row total
1		1				1
2	1	2	1		1	5
3			11	4		15
4			8	12	1	21
5			2	5	2	9
Column total	1	3	22	21	4	51

Class	User's accuracy	Producer's accuracy
1		
2	40	67
3	73	50
4	57	57
5	22	50

Overall accuracy (P_o) = 0.58

Tau (T_e) = 0.41

Kappa (K) = 0.31

Table 36.--Error matrix table of the relevé estimates of stand height

Reference data

Relevé data: Class	1	2	3	4	5	6	Row total
1							0
2							0
3			16	4			20
4			1	15	2		18
5			2	6	4		12
6					1		1
Column total	0	0	19	25	7	0	51

Class	User's accuracy	Producer's accuracy
1	-	-
2	-	-
3	80	84
4	83	60
5	33	57
6	-	-

Overall accuracy (P_o) = 0.69

Tau (T_e) = 0.62

Kappa (K) = 0.52

Table 37.--Error matrix table of the relevé estimates of stand cover type

Reference data

Relevé data: Class	A	D	G	L	P	S	W	Row total
A	5	1				1		7
D		17	1	1	1		1	21
G		1	1				1	3
L	2	1		5				8
P		2			3			5
S		1						1
W		4					2	6
Column total	7	27	2	6	4	1	4	51

Class	User's accuracy	Producer's accuracy
A	71	71
D	81	63
G	33	50
L	63	83
P	60	75
S	-	-
W	33	50

Overall accuracy (P_O) = 0.64

Tau (T_e) = 0.59

Kappa (K) = 0.51

Appendix E: Scatterplots of estimated vs ground truth values for the stand attributes

Scatterplots of the estimates (photo and releve plot) vs. the ground values are presented below. The solid line in each scatterplot is the least squares difference fitted regression line, with the associated R^2 correlation coefficient. The broken line (— · — · -) is the line of equality.

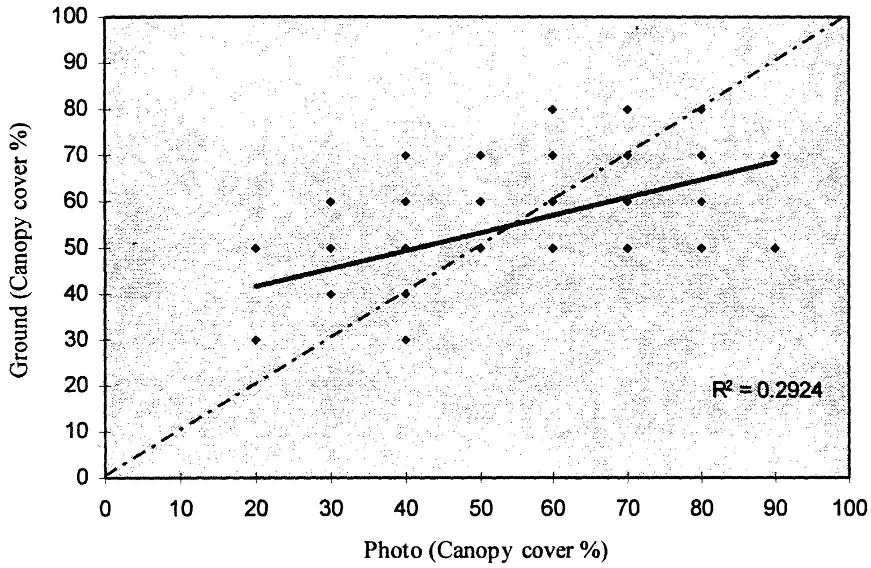


Figure 9.--Scatterplot of photo-estimates of canopy cover vs. ground values.

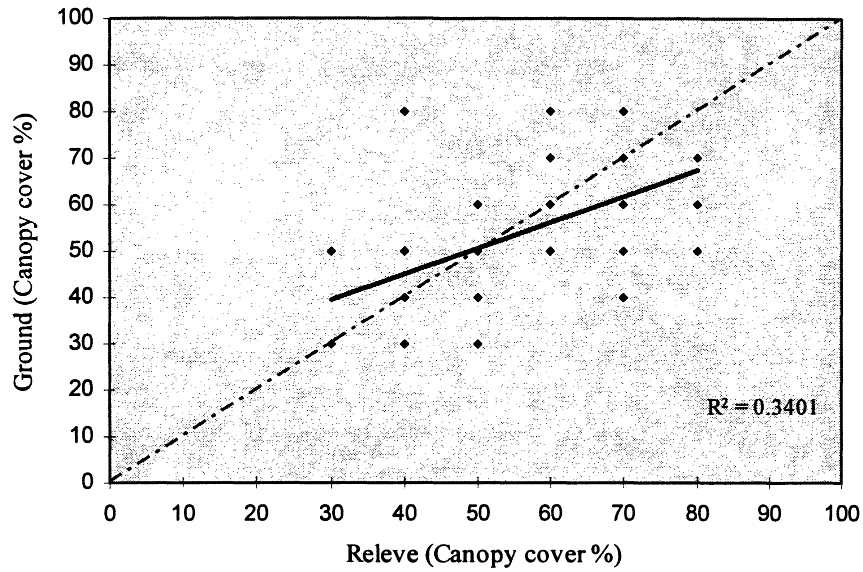


Figure 10.--Scatterplot of releve estimates of canopy cover vs. ground values.

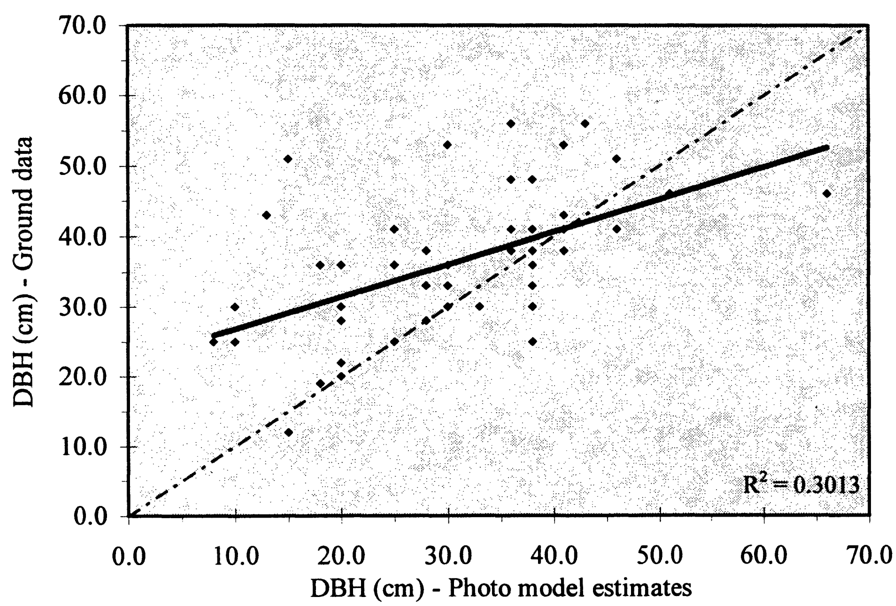


Figure 11.--Scatterplot of photo (model) estimates of DBH vs. ground values

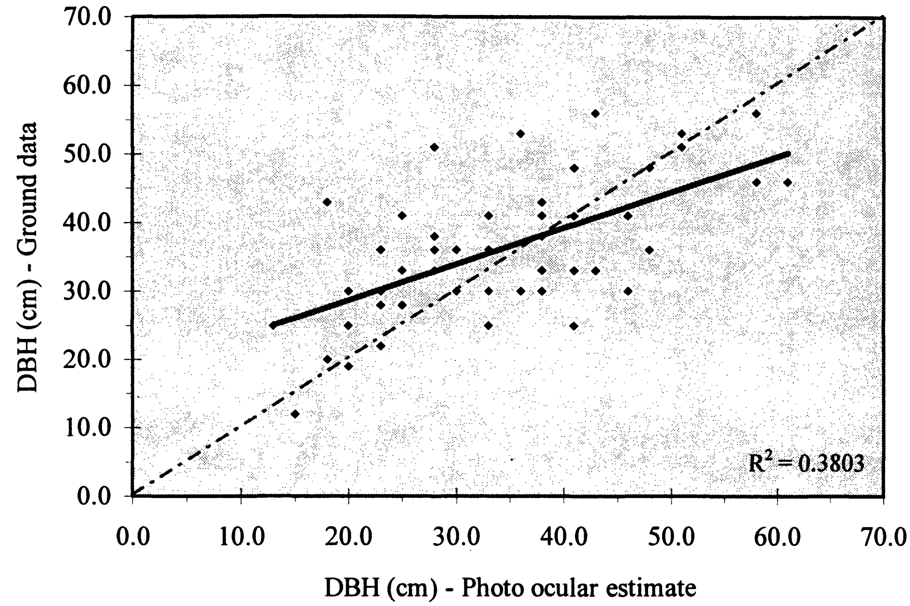


Figure 12.--Scatterplot of photo ocular estimates of DBH vs. ground values

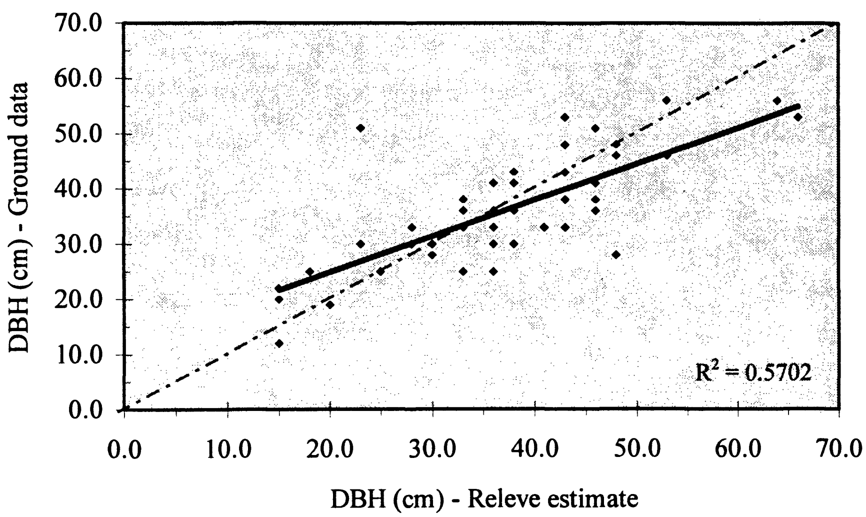


Figure 13.--Scatterplot of releve estimates of DBH vs. ground values

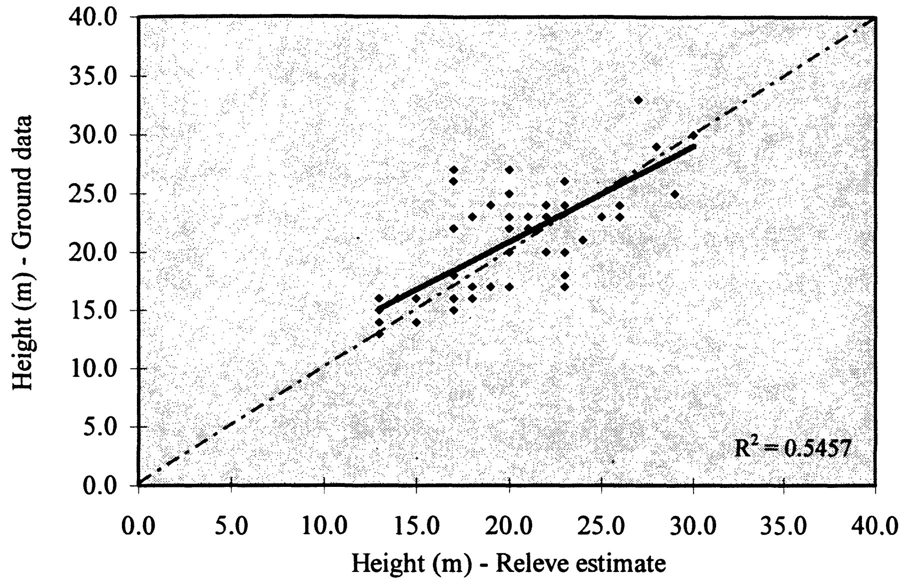


Figure 14.--Scatterplot of releve estimates of stand height vs ground values

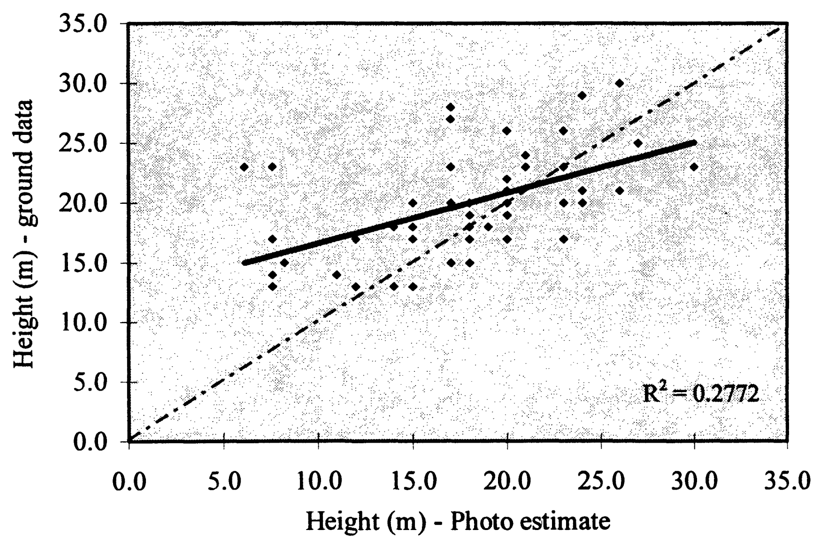


Figure 15.--Scatterplot of photo estimates of stand height vs. ground values

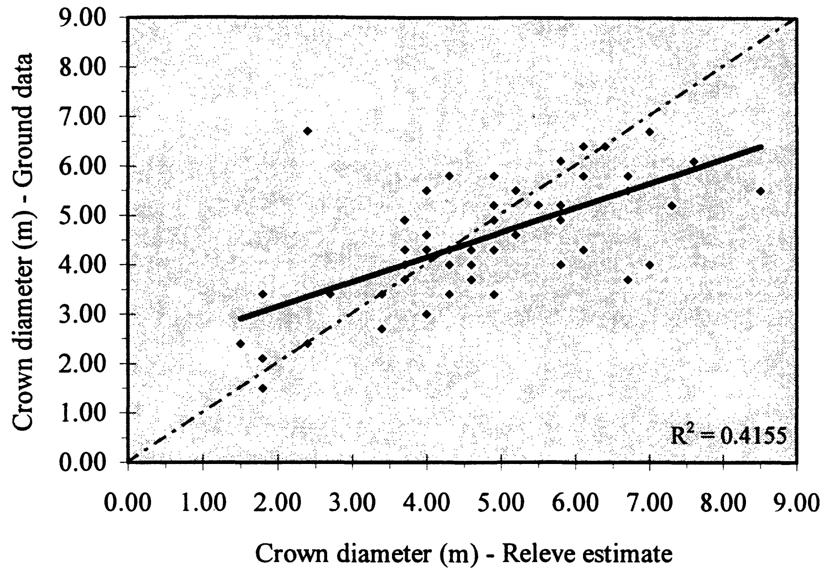


Figure 16.--Scatterplot of releve estimates of crown diameter vs. ground values

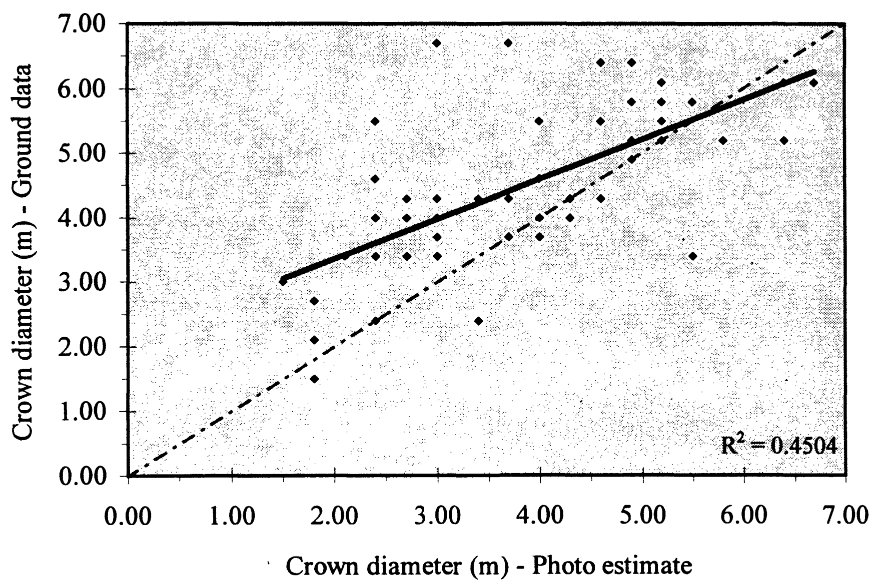


Figure 17.--Scatterplot of photo estimates of crown diameter vs. ground values

Appendix F: Coefficient of variation analysis

Table 38.--Mean coefficient of variation values for stands in the three spatial aggregation classes. Coefficient of variation (CV) values were calculated from the 5 plots within each stand. Mean CV values were then calculated for all the stands in each spatial aggregation category.

Attribute	Regular (n = 25)	Irregular (n = 3)	Uniform (n = 23)
Crown diameter	29.4	27.3	26.5
DBH	25.6	29.0	23.8
Height	18.6	18.3	19.2
Canopy cover	298	253	280

Table 39.--Mean coefficient of variation values for stands. Coefficient of variation (CV) values were calculated from the 5 plots within each stand; the mean CV for all the stands in each category was then calculated. Estimation categories were defined using scatterplots.

Attribute	Overestimated	Accurately estimated	Underestimated
Crown diameter	--*	24.8	32.6
DBH (ocular est.)	20.2[a]	23.5[a]	30.8[b]
DBH (model)	--*	23.8	27.1
Height	19.6	19.6	16.1
Canopy cover	--*	287	287

Note: All means were not significantly different at the 5% level, with the exception of the means for the DBH (ocular est.) attribute; in this attribute the underestimated stands had a significantly higher coefficient of variation than the overestimated or accurately estimated stands. (*) indicates no stands in this estimate category.