Forest Inventory Stratification and Mapping based on the Classification of Landsat Thematic Mapper Imagery

Kenneth A. Stumpf
Geographic Resource Solutions
1125 Sixteenth Street, Suite 213
Arcata, CA 95521
(707) 822-8005

Abstract

The classification of forestland vegetation using Landsat Thematic Mapper (TM) data and quantitative field data collection techniques can produce maps that are consistently correlated to the various forestland characteristics that might represent a stratified forest. Differences in stand characteristics such as species composition, density, and size can be recognized and mapped over large areas using these objective and consistent methodologies. Stand volumes can also be estimated and used to guide the aggregation process. The resulting maps can be used as the basis of stratification (first stage sampling) for large scale forest inventory applications. Stand characteristics may be grouped into fairly narrow classes or generalized to form broader classes and reduce the overall number of strata. Second stage sampling involving conventional forest inventory plot measurements can be designed on the basis of this stratification. Data collection efforts can be estimated and prioritized based on initial stratum volume estimates and variances, desired statistical measures of reliability, costs of data acquisition, and accessibility. Stand boundaries may be developed and attributes estimated based on field data collection efforts and/or image classification results.

Introduction

Forest inventories conducted over large ownerships often consist of multi-stage samples. Remote sensing techniques are commonly used to stratify forestlands and provide a basis for the design and implementation of field data collection techniques. Aerial phototyping has traditionally been one method applied to provide the basis for the stratification. Areas or types of forest vegetation are identified, delineated, and grouped (typed) in an effort to reduce the variance of the inventory estimates and decrease the sample intensity and cost of such inventories. Large scale inventories that encompass tens or hundreds of thousand hectares necessitate the use of large quantities of current aerial photography. Phototyping must be applied consistently across the subject area to maintain the benefits of stratified sampling. The uniform mapping and analysis of the typed information for use as a first stage sample is a difficult and formidable task with or without the assistance of a geographic information system (GIS). Recent developments in the field of image classification (Brown and Fox, 1992) and the post-classification process of rule-based aggregation of pixels into groups or stands of similar forest types (Stumpf and Koltun, 1992) show the potential to replace airphoto interpretation to produce maps of forestland strata. The proper application of these techniques can produce stratified forest type maps that are consistently representative over large areas and less subjective relative to photo interpreted data. The number of strata can be defined based on the variability of forest stand characteristics as they are related to timber inventory volumes. Strata can be defined using volume classes as well as vegetation types. Second stage inventory sampling techniques can be designed, prioritized, and applied based on the preliminary strata variances, sample sizes, and area distributions.

Image Classification

Image classification processes are used to associate different spectral data as represented in the satellite imagery with different vegetation characteristics as measured on the ground. Multi-spectral data, such as Landsat Thematic Mapper (TM) data with a resolution of 25 meters, provides a means of estimating and mapping different types of vegetation features provided these features can be differentiated through image classification processes.

GRS utilizes a combination of supervised and unsupervised classification processes to produce pixel maps of the forestland cover to be stratified. The supervised process is primarily used to classify vegetation types that characterize the areas of forestland types. The unsupervised techniques are primarily used to classify the non-forestland areas and identify any unclassified forestlands requiring training data. The key to the success of the supervised classification process is the collection of quantitative data at sites selected as training areas. Quantitative measurements and estimates are the basis of the training data forestland estimates. This reduces confusion resulting from subjective visual assessments of forestland characteristics (Biging and Congalton,1992). Visual estimates are not used without accompanying quantitative data and only then to recognize situations where the quantitative data sample may not have provided an accurate description of a particular training area. Selected training sites are areas of relatively low spectral homogeneity that are also fairly homogeneous in terms of their forest type characteristics and timber inventory estimates. The training data is collected based on the pin-point sampling method and the GRS crown canopy-densitometer. Groups of three to five transects with no less than thirty-three points per transect are established and measured in each training area. Sample points are equally spaced between twelve and twenty feet depending on the stand's average crown size. The vegetation characteristics specific to each sample point are collected along the transects. This data includes tree species, diameter breast height(dbh), total height, percent defect, crown area, and canopy level. Overstory or top level canopy vegetation is also identified to distinguish it from other levels of the canopy that may be present in the stand but may not contribute spectrally to the imagery. The training data is processed to estimate the stand characteristics that will be used to accomplish the stratification of these forest lands. Each group of transects yields a cover and frequency distribution by species and size. A sample stand distribution is shown in Table 1.

Table 1: Sample training area cover and frequency distribution.

Training Area Cover Density Summary:
Training Area: 110
Number of Transects: 3

Cover by Covertype:

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover	Non-Tree Cover	Total Cover
Species								
Doug-fir	0.0%	0.0%	0.0%	1.0%	3.0%	4.0%		4.0%
redwood	0.0%	0.0%	7.0%	9.0%	48.9%	64.9%		64.9%
W. hemlock	0.0%	0.0%	0.0%	0.0%	4.0%	4.0%		4.0%
tanoak	0.0%	2.0%	0.0%	0.0%	0.0%	2.0%		2.0%
							25.1%	25.1%
Total Cover	0.0%	2.0%	7.0%	10.0%	55.9%	74.9%	25.1%	100.0%
Total Tree Cover						74.9%		

Cover by Tree Type:

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover
Species						
Doug-fir	0.0%	0.0%	0.0%	1.3%	4.0%	5.4%
redwood	0.0%	0.0%	9.3%	12.1%	65.3%	86.6%
W. hemlock	0.0%	0.0%	0.0%	0.0%	5.3%	5.3%
tanoak	0.0%	2.7%	0.0%	0.0%	0.0%	2.7%
Total Tree Cover	0.0%	2.7%	9.3%	13.4%	74.6%	100.0%

Quadratic Mean DBH Summary:

Size Class:	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover
Species						
Doug-fir	0.0"	0.0"	0.0"	34.0"	64.0"	45.9"
	0.0tpa	0.0tpa	0.0tpa	1.4tpa	0.7tpa	2.1tpa
redwood	0.0"	0.0"	20.2"	30.6"	70.2"	53.2"
	0.0tpa	0.0tpa	8.1tpa	7.1tpa	15.9tpa	31.1tpa
W. hemlock	0.0"	0.0"	0.0"	0.0"	40.0"	40.0"
	0.0tpa	0.0tpa	0.0tpa	0.0tpa	1.7tpa	1.7tpa
tanoak	0.0"	7.5"	0.0"	0.0"	0.0"	7.5"
	0.0tpa	2.5tpa	0.0tpa	0.0tpa	0.0tpa	2.5tpa
QMDBH	0.0"	7.5"	20.2"	31.2"	67.7"	50.5"
	0.0tpa	2.5tpa	8.1tpa	8.5tpa	18.2tpa	37.4tpa
QMDBH - Con	0.0"	0.0"	20.2"	31.2"	67.7"	52.3"
	0.0tpa	0.0tpa	8.1tpa	8.5tpa	18.2tpa	34.9tpa
QMDBH - Hwd	0.0"	7.5"	0.0"	0.0"	0.0"	7.5"
	0.0tpa	2.5tpa	0.0tpa	0.0tpa	0.0tpa	2.5tpa

Volume Summary:

Volume/Acre			2.0mbf	10.3mbf	158.2mbf	170.5mbf
Volume/Acre - Con			2.0mbf	10.3mbf	158.2mbf	170.5mbf
Volume/Acre - Hwd						0.0mbf

This distribution is processed to estimate stand characteristics that include species composition and type, average size, crown canopy closure, and stand density, in terms of stems per unit area. Timber inventory plots of the same type sampling methodology that will be used during the field inventory process are also measured in the training areas. These data will yield estimates of volume per unit area as well as variances for the inventory estimates. These stand characteristics and inventory estimates are treated as characteristics of the training area. The result of the classification processes is that individual pixels are assigned the value of the training class whose spectral data is most similar to the spectral values of the pixels. The classified pixels comprise a classification grid representative of the forestland types of the classified area and each classified forestland pixel is associated with its training area stand characteristics and inventory estimates.

Pixel Data Aggregation to Form Stand Boundaries

The pixel data produced during the classification processes is often too heterogeneous to immediately form stand boundaries suitable for stratification and sampling. The classified grid contains hundreds of thousands of individual pixels and small groups of pixels representative of different forestland characteristics and inventory estimates that are too small to map as individual stands. These pixels must be aggregated to form stands of sufficient size as defined by the inventory project objectives. The definition of the minimum stand size is relative to the characteristics being mapped and the sampling methods employed. Variable minimum size limits may be defined for different stand and inventory characteristics. For example, high volume per unit area stands may have smaller minimum size limits than low volume stands; stands comprised of commercially valuable species types may have smaller minimum size limits than stands comprised of poor quality or non-commercial species types; and stands comprised of large trees may have smaller limits than stands comprised of small trees.

GRS has developed a rule-based approach to aggregating pixels to form stand types of similar groups of pixels. This technique is based on evaluating the similarity of adjacent stands and merging sub-minimum size stands into the adjacent stand that is the most similar stand. Similarity may be expressed in terms of a number of forest inventory characteristics such as average tree size, stems per unit area, species composition, and volume per unit area. By aggregating sub-minimum size stands with the most similar neighboring type the variance of the forestland types is maintained at the lowest possible level while still enabling the inclusion of these sample stands in the inventory.

If the criteria for stratification is the estimated timber volume per unit area and the species type and size are not used for stratification then the similarity of adjacent types can be judged quite simply. The volume per unit area estimates are compared and the adjacent stands that are in the same stratum and have the smallest volume difference are aggregated (this approach will tend to force types towards the midpoint of each stratum). In this case the forest inventory stratification would be indifferent to species and size differences. However, if the objectives of the inventory include the development of reliable estimates by species and/or size, then these stand characteristics should also be included in the evaluation of stand similarity.

Other characteristics such as the average piece size may also be used to assist in the evaluation of similarity. However, as multiple characteristics are included in the estimate of similarity the evaluation becomes more complex. The different characteristics may not contribute equally to the estimate of similarity. The relative level of contribution of each characteristic can be estimated and incorporated in the rules that guide the aggregation process. The user may therefore control the manner in which aggregation decisions are made to emphasize forest inventory goals and objectives.

Aggregated Stand Characteristics

Following the completion of the aggregation process stand summaries are produced for all stands. These summaries include all the pixels that are contained within the boundaries of each stand. These summaries are basically cover and frequency distributions by species and size and are similar to the cover and frequency distributions developed for each training area(see Table 1). The average stand characteristics are developed from these summaries. These characteristics provide the basis for the stratification of the mapped forestlands. Additional information derived from these distributions includes an estimate of the homogeneity of each mapped stand. Stands comprised of similar types are rated high in homogeneity whereas stands comprised of dissimilar types are rated low in homogeneity. This estimate of homogeneity is a stand characteristic that can also be included in the strata definitions. Stands may be assigned to the various strata based their stand characteristics, including their different levels of homogeneity.

Frequency and Area Distributions by Strata

Summarizing aggregated stands by stratum produces area and frequency distributions that characterize the forestlands being inventoried. Each area's distribution can be used to identify the relative occurrence and magnitude of stands of the different strata as well as estimations of the average size stand (the acreage by stratum distributions can also be used to provide initial estimates of the total inventory of the area under consideration). This area distribution is also significant as a check of the aggregation process that was used to form the stand type boundaries. If the aggregation process is operating in an unbiased manner then the stand (polygon) area distributions by stratum should approximate the area distributions by stratum of the initial pixel grids. However, an acreage comparison may be difficult to evaluate as the area of non-forest type pixels and forest type pixels can change radically as they are aggregated together. Small groups of non-forest pixels typically become parts of forest type stands and cause a shift in non-forest to forest type area. A better measure of correspondence of the two datasets is a comparison the estimated total inventory of the inventory project areas as represented by the two datasets. This is accomplished by summing the volume by stratum for each of the area distributions and comparing the two estimates of total volume. The results of such a comparison are shown in Table 2.

Table 2: Summary of Volume by Stratum

	------ Pixel Map -------		------- Polygon Map ------	
Volume Stratum	Area (Acres)	Volume by Stratum	Area (Acres)	Volume by Stratum
1	239,289.0	1,196,445	81,049.1	405,245
2	180,118.5	3,152,074	217,046.1	3,798,306
3	139,150.6	4,522,394	260,151.8	8,454,933
4	198,140.4	9,907,020	232,482.8	11,624,140
5	277,090.9	24,938,181	243,059.6	21,875,364
Totals	1,033,789.4	43,716,114	1,033,789.4	46,157,988

Volume Class	Volume per Acre	Midpoint
1	0 - 9.9 mbf	5.0 mbf
2	10 - 24.9 mbf	17.5 mbf
3	25 - 39.9 mbf	32.5 mbf
4	40 - 59.9 mbf	50.0 mbf

Additional Field Data Collection Requirements

Inventory estimates are not enough to design the field sampling stage of an inventory of this magnitude. Variances of timber inventory values must be developed for each stratum using the same measurement techniques used during the timber inventory data collection. Sample stand data collected during the training data collection phase of the project may be used to estimate the variances so long as the training site was classified and aggregated into one stand as indicated by the training data being collected within the final stand boundaries. Strata that represent mixes of forest types, such as those stands indicating high within stand diversity, may require additional field sampling as they were not sampled as training areas due to their high within stand variability. A small number of additional sample stands may be selected from each stratum requiring additional data to estimate within stratum variances of these strata. Inventory plots are then measured in the selected stands and the data is processed to produce estimates of inventory volumes and variances. This information in combination with desired statistical levels of reliability, the acreage distributions, budget constraints, and project objectives enables the determination of the minimum sample size by and the allocation of inventory plots by stratum. Scheduling of inventory activities is also facilitated by applying efforts to those strata determined to be of high priority for sampling relative to areas considered low priority.

Summary

Photo interpretative methods have been applied to generate stratified maps for forest inventory purposes. These efforts may be difficult due to the extent of the inventory project and the subjectivity of the classification effort. Image classification processes and rule-based pixel aggregation techniques can be applied to satellite imagery to produce stratified maps for timber inventory purposes. These techniques can be applied consistently across large areas. Data used to train the image classification and pixel aggregation processes can also be used to estimate inventory volumes and variances. The distributions of area by stratum combined with sample estimates of volumes and variances provide useful information for the design and implementation of field sampling efforts.

Literature Cited

Biging, G. and R. Congalton, 1992. A pilot study evaluating ground reference data collection efforts for use in forest inventory. Photogrammetric Engineering & Remote Sensing Vol. 53, No. 12, pp. 1669-1671.

Brown, G. and L. Fox, 1992. Digital classification of Landsat Thematic Mapper imagery for recognition of wildlife habitat characteristics. Proc:1992 ASPRS/ACSM Convention, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, Vol. 4, pp. 251-260.

Stumpf, K. and J. Koltun, 1992. Rule based aggregation of raster image classifications into vector GIS databases with five- and forty-acre minimum size mapping units. Proc:1992 ASPRS/ACSM Convention, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, Vol. 4, pp. 261-278.