GPHN
IN-43-CR
179494
P- 27

Final Report

to

National Aeronautics and Space Administration
Grant NAGW-1431, NRA 87-106

Satellite Inventory of Minnesota Forest Resources

by

Marvin E. Bauer, Thomas E. Burk, Alan R. Ek,
Pol R. Coppin, Stephen D. Lime, Terese A. Walsh, and David K.Walters
Department of Forest Resources
University of Minnesota
St. Paul, MN 55108

William Befort and David F. Heinzen
Division of Forestry
Minnesota Department of Natural Resources
Grand Rapids, MN 55744

July 1993

(NASA-CR-193565) SATELLITE INVENTORY OF MINNESOTA FOREST RESOURCES Final Report (Minnesota Univ.) ~~27 p~~ N94-11521

Unclas

455647 29p G3/43 0179494

Table of Contents

Satellite Inventory of Minnesota Forest Resources

Abstract

The methods and results of using Landsat Thematic Mapper (TM) data to classify and estimate the acreage of forest covertypes in northeastern Minnesota are described. Portions of six TM scenes covering five counties with a total area of 14,679 square miles were classified into six forest and five nonforest classes. The approach involved the integration of cluster sampling, image processing, and estimation. Using cluster sampling, 343 plots, each 88 acres in size, were photo interpreted and field mapped as a source of reference data for classifier training and calibration of the TM data classifications.

Classification accuracies of up to 75% were achieved; most misclassification was between similar or related classes. An inverse method of calibration, based on the error rates obtained from the classifications of the cluster plots, was used to adjust the classification class proportions for classification errors. The resulting area estimates for total forestland in the five-county area were within 3% of the estimate made independently by the USDA Forest Service. Area estimates for conifer and hardwood forest types were within 0.8 and 6.0%, respectively, of the Forest Service estimates. A trial of a second method of estimating the same classes as the Forest Service resulted in standard errors of 0.002 to 0.015. A study of the use of multidate TM data for change detection showed that forest canopy depletion, canopy increment, and no change could be identified with greater than 90% accuracy. The project results have been the basis for the Minnesota Department of Natural Resources and the Forest Service to define and begin to implement an annual system of forest inventory which utilizes Landsat TM data to detect changes in forest cover.

Introduction

Forests covering 16.7 million acres or nearly a third of the land area of Minnesota are a significant component of the state's natural resource base and a significant contributor to its economy. In spite of growing demands for information about the state's forest resources, statewide inventories are conducted only at about 15-year intervals. Although forest stand growth models have become increasingly important for updating inventory information and projecting future forest conditions (Ek, 1983), such updates are limited because of their concentration on existing forested plots. Total forest area and area by cover type change because of cropland abandonment, harvesting, and urban development. Such changes are extremely difficult to model; that is a major reason for wanting to use satellite data to determine forest areas. Since the mean characteristics of forest strata change relatively little, the major inventory problem is to estimate the amount and location of the strata.

For many years foresters have effectively utilized aerial photography as a tool to help monitor and manage forest resources and aerial photographs are an integral part of most forest inventory procedures. The launch of Landsat-1 in 1972 added an entirely new dimension to the capability to acquire Earth resources information and there has been much interest in the potential of satellite data and computer-aided analysis techniques to identify and map forest resources. Although it has generally not been possible with Landsat MSS data to achieve satisfactory classification accuracy for any but the most general classifications in the Great Lakes States, a number of studies have shown that the information content of Landsat Thematic Mapper (TM) data is considerably higher than that of MSS data (Price, 1984; DeGloria, 1984) and that the additional spectral bands and finer spatial and radiometric resolution of TM data result in significant improvements in classification accuracy for more specific information classes describing forest species (Horler and Ahearn, 1986; Moore and Bauer, 1990) and forest stand characteristics (Peterson et al., 1986; Williams and Nelson, 1986). The results of Moore and Bauer (1990), which provided much of the impetus for this research, showed a 15-20% increase in classification accuracy of TM data over MSS data, with TM accuracies of greater than 80% for seven classes.

Objectives

The overall objective of the research was to develop and test procedures for using multispectral satellite data to inventory forest resources in the state of Minnesota. Specific objectives were to:

1. Develop a methodology to use digital satellite data and computer-aided pattern recognition to classify forest covertypes which will be compatible with and complementary to the other surveys conducted by the Minnesota Department of Natural Resources and the U.S. Forest Service.

2. Estimate forest areas and produce digital maps of the state's forest resources by species group at county, region and state levels, and determine the accuracy and precision of the forest area estimates and maps derived from satellite data, compared to traditional forest inventory estimates.

3. Investigate alternative innovative approaches to satellite data classification, sampling and estimation and determine to what degree satellite data classifications can be used to obtain additional information such as stand density, size class, and disturbance.

4. Integrate the satellite data classification, sampling and estimation designs and procedures into the Forest Resource Assessment and Analysis Program of the Minnesota Department of Natural Resources (DNR) and Forest Inventory and Analysis (FIA) of the U.S. Forest Service.

The goal of the project was to develop and implement a methodology to provide cover type area estimates of $\pm5\%$ at the 95% confidence level at the state level and $\pm10\%$ with 90% confidence at the county level. Other performance goals of the final inventory procedure include: cost $.01-.02 per forested acre, one year to acquire and analyze the data, procedures that can be implemented by the DNR with reasonable personnel and capital costs, and enough flexibility to meet changing conditions or requirements.

Two important underlying premises of the objectives and approach tested in the investigation are: (1) the synoptic view of Landsat provides the opportunity to obtain forest inventory information over large areas (i.e. state) and (2) by using computer data analysis methods to classify pixels distributed over counties and unique sampling designs, it is also possible to make accurate and precise estimates for local areas (i.e., counties). This approach offers a means to improve upon the sampling methods now used for making area estimates from ground-based, two-phase surveys. At the same time the ground data (which is also used to estimate other parameters such as forest stand characteristics and therefore its collection cannot be abandoned) will be used to remove the bias from (i.e., correct) the satellite-based estimates. An important consideration in the proposed approach is that improved covertype estimation techniques would lead to more efficient and timely forest descriptions for a variety of purposes.

Multispectral Landsat TM data, together with ground-based forest inventory data were used to produce an inventory of Minnesota forest lands, along with a methodology for frequent updates. The following sections describe the sample design and survey model, satellite data classification, calibration and evaluation of Landsat area estimates, change detection using multidate Landsat data, and, lastly, the use of satellite remote sensing for operational forest inventories in Minnesota. The emphasis is on describing new approaches to forest inventory using satellite data.

Background

Study Area

The study area for the project consisted of five forested counties in northeastern Minnesota (Figure 1), totaling some 8.8 million acres. The region stretches nearly 230 miles east/west and 170 miles north/south. The study area corresponds to the USDA Forest Service's Forest Inventory and Analysis (FIA) survey unit one, the Aspen/Birch unit (Miles and Chen, 1992). The study area is comprised of a variety of forest types, with primary types being aspen-birch, spruce-fir and pine.

The geology of this region is largely the result of glaciation (Wright, 1972). The northeastern portion of the area, Lake and Cook counties, is marked with an abundance of Canadian Shield lakes, and displays relatively large topographic variations. The western half of the region, Koochiching county, is a vast lowland, with intermittent moraines. The central portion of the region is characterized by a variety of geologic formations, and is heavily forested. The central portion of St. Louis county is dominated by granite highland termed the "Iron Range," and is home to numerous mining operations. In southern St. Louis and Carlton counties, agriculture and other non-forest landuses, are more prevalent. Even so, these areas are still predominately forested.

Landsat Data

The image data for the project consisted of portions of six Landsat TM scenes (Figure 2). Data were collected between 5/29/88 and 6/14/88, with scenes along the same path being collected on the same day. All images were virtually cloud free, with any clouds occurring outside of the study area. Haze was observed over water bodies within the path 26, row 26 scene covering the northeast portion of the study area.

All scenes were rectified to the UTM (zone 15) projection and coordinate system using a nearest neighbor resampling scheme to preserve the original DNs. Rectification allowed for relatively easy overlay of reference data sets for training and accuracy assessment. Because of the inherent problems of working across scenes of differing acquisition dates (changes in atmosphere, sun angle, and phenology), processing was limited to within scene (or path) processing. Consequently, three separate image datasets were processed, and merged only as classified (GIS) files.

Reference Data

Reference data sets were generated using 35-mm color infrared aerial photography flown at a scale of approximately 1:10,000. Seven lines of photography were gathered at equal intervals across the study area (Figure 3). Interpretation units were defined as 88-acre clusters spaced roughly one mile apart over the entire length of the flightlines (Figure 4). A total of 343 clusters were interpreted into approximately 100 covertype, size and density classes (Table 1). Subsequent to photo interpretation, all clusters were visited on the ground, or viewed from the air if ground access was limited. Clusters were located on USGS 7.5' minute quadrangles, and then digitized using PC Arc/Info. Reference datasets were then rasterized and linked to attribute data (i.e. type, size, and density) and registered to the Landsat image data using ERDAS.

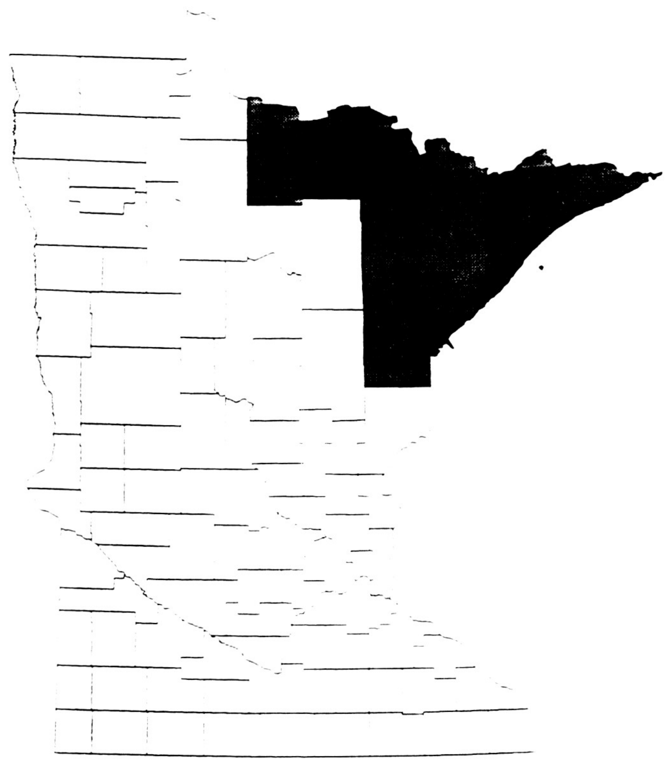


Figure 1. Five-county project study area in northeastern Minnesota (FIA aspen-birch unit).

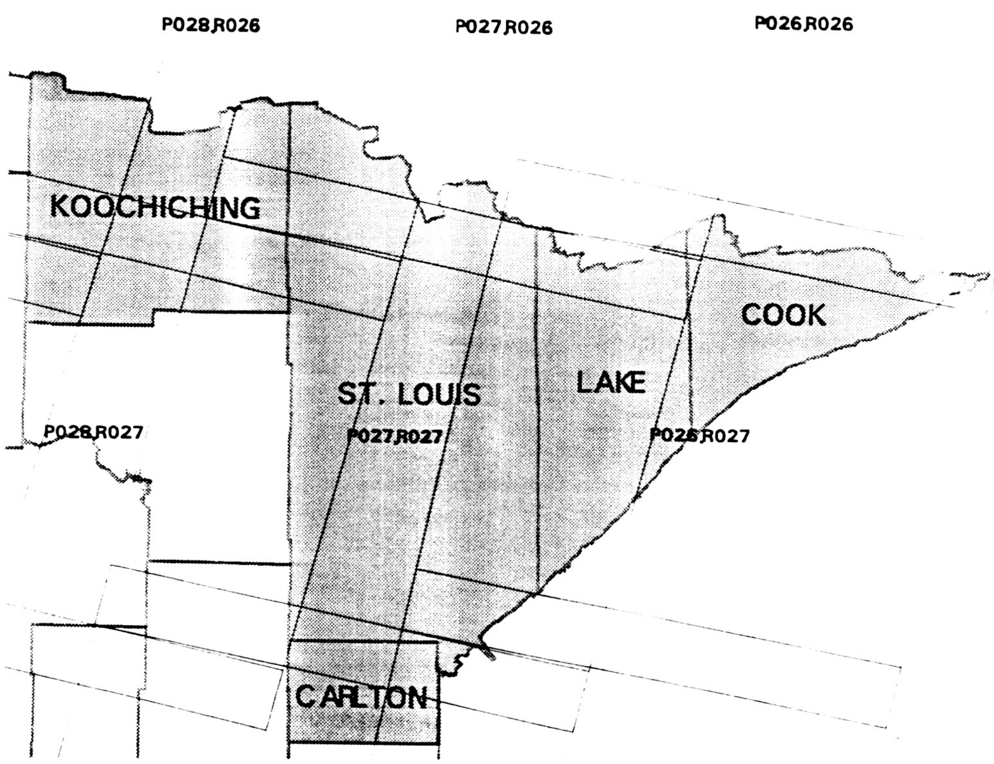


Figure 2. Project study area and Landsat TM scenes.

Hardware and Software Environment

Image processing was performed on SUN SparcStations and 386 microcomputers. Raster-based image processing and GIS procedures were completed using workstation ERDAS. Vector GIS and data development procedures were completed using PC Arc/Info 3.4D. Additional routines were developed in-house using C and AWK.

Survey Design

Given the emphasis on forest area estimation and satellite data, there is much potential for gains from refinements in survey design. Designs using large clusters have been shown to be an effective means of collecting forest inventory information (Scott et al., 1983). In particular, Benessalah (1985) has shown that such layouts have decided advantages for ground checking of remotely sensed data. Advantages include ease of field work, variance reduction, and the provision of area data as proportions rather than binary (0-1) counts. Proportion data facilitate the use of remote sensors because it is relatively insensitive to scale problems. The sampling design considered for covertype area statistics was a single phase design involving remote sensing-based classification of the entire area and ground checking of a sample of large clusters. This exploited the synoptic coverage of satellite-acquired digital remote sensing information and included all of the pixels in the population in addressing area estimation.

A sample of clusters on the "imagery" was observed on the ground to assess forest type. The clusters might be 10 to 100 or more acres in size, however the size used here was a consistent 88 acres (a size equivalent to 17 x 23 TM pixels). The ground sampling (field visit) established an accurate type map for the cluster, i.e., polygons in the cluster were identified and labeled as to covertype. This meant usage of boundaries and labels according to standards of interest to forest management. Clusters were distributed systematically (with a random start) across the survey unit. The ground sample clusters were then used to train the classifier and subsequently the classifications became the predictor variables for regression estimates of the proportion of the clusters in various covertypes.

The ground clusters were very inexpensively mapped and field checked using large scale aerial photography. Cluster visitation costs were approximately $100 each, including travel, preliminary covertype mapping of the cluster on large scale 1:10,000 color infrared aerial photography and field verification of covertype boundaries and lines. The photography itself cost approximately $23 per cluster and digitizing the corrected cluster covertype map cost approximately $15. The photography was employed here to help locate the clusters and assist with access (a nontrivial task) and to assist in developing accurate ground reference data.

Landsat TM Classification Methodology and Results

Although more than 100 landcover classes resulted from interpretation of the aerial photography, it was apparent that such a detailed Landsat-data classification was not possible. In many cases there was simply not enough training information in the cluster samples to permit further work. In other cases, initial tests showed no promise in separating certain classes (e.g., aspen vs. birch). Ground truth classes (both forest and non-forest) were then condensed into 11 (6 forest, 5 non-forest) covertypes based on spectral and management considerations as listed in Table 1.

Table 1. Hierarchy of information classes for photo interpretation and classification of Landsat TM data.

Photo Interpretation Classes*	Satellite Classes	Code
Ash	Lowland Hardwoods	LH
Elm		
Aspen/Birch	Aspen/Birch	A/B
Aspen		
Birch		
Northern Hardwoods	Northern Hardwoods	NH
Upland Conifers	Upland Conifers	UC
Red Pine		
White Pine		
Jack Pine		
Balsam Fir/White Spruce	Balsam Fir/White Spruce	BF/WS
Balsam Fir		
Spruce		
Lowland Conifers	Lowland Conifers	LC
Lowland Black Spruce		
Tamarack		
Northern White Cedar		
Stagnant Spruce		
Stagnant Cedar		
Cutover Area	Shrub/Cutover/Grass	S/C/G
Lowland Grass		
Upland Grass		
Lowland Brush		
Agriculture	Agriculture	Ag
Urban and Industrial	Developed	Dev
Recreational Development		
Water	Water	W
Marsh	Marsh	M/M
Muskeg		

* Forest cover types were further subdivided into three crown closure and three size classes.

In addition to the six reflective TM bands, a set of six vegetation indices (VIs) were added to form twelve feature sets used for processing. The VIs were chosen to be a representative sample of possible such indices. The first set of indices consisted of the first three Tasseled Cap transformations, greenness, brightness and wetness, with coefficients given by Crist and Cicone (1984). While greenness is the most indicative of vegetative cover, brightness is also related to vegetative cover. Wetness is related to moisture content and may be useful for wetland delineation and separation of upland and lowland forest types. The second group of VIs were ratios that have been found to intensify forest canopy characteristics. This group consists of TM4/TM3, TM4/TM2 and TM5/TM4 ratios. Jensen (1983), among others, states that TM4/TM3 provides information with respect to vegetation and canopy condition and that TM4/TM2 may be a promising feature for wetland studies. The TM5/TM4 ratio has been used in studies related to conifer canopy structure.

A number of classification processing approaches were evaluated for their utility in large area classification. Supervised, unsupervised and combination approaches were examined. Supervised techniques were judged inappropriate for a number of reasons: extreme forest complexity, low ability to automate processing, and narrow covertype spectral separability. Several types of clustering methodologies, from standard ISODATA clustering (ERDAS, 1991) to hierarchical strategies defining more than 1,000 classes, were tested. In all cases, the ability to name the resulting classes was limited by the dominance and variability of a few well represented classes. In most cases, less than 50% of unsupervised classes could be named, and in no cases could classes be developed for all target classes.

An alternative to the supervised and unsupervised approaches is what we have called "guided" clustering. The procedure makes use of both supervised and unsupervised techniques, but avoids many of problems associated with each individually. The method uses of analyst-defined training data, in our case digitized covertype polygons from the interpreted and field-checked aerial photo plots, and uses this data to isolate image pixels that are clustered into spectrally homogenous sub-classes. The processing stream is as follows:

1. Isolate image pixels for target class A.
2. Cluster resulting pixels into sub-classes A1,...An.
3. Repeat 1 and 2 for all 11 target classes.
4. Perform maximum likelihood classification using all sub-classes.
5. Collapse subclasses back to original 11 target classes.
6. Perform any post-processing procedures.

Guided clustering provided consistently superior results to any of the other methods tested. The process is highly automated, so was ideal for large area application. The approach combines the training and classification approach with statistical information from the cluster sampling units.

The sheer size of the area to be classified created many obstacles that had to be overcome. It was clear from visual assessment of the imagery that landcover gradients existed within the scenes. In addition, atmospheric and phenological differences existed from north to south, and to a lesser extent from east to west, through the study area. To compensate for such differences, physiographic regions delineated by Wright (1972) were used to segment the study area into eight sub-regions (Figure 3). TM images for paths 26 and 28 were not segmented into sub-regions because there was not sufficient training data for all of physiographic regions falling within each image. The resulting covertype distributions for the reference data are shown in Table 2. The distribution of covertypes across physiographic regions is not constant, and is actually quite varied, a good indication that the segmentation scheme accomplished what

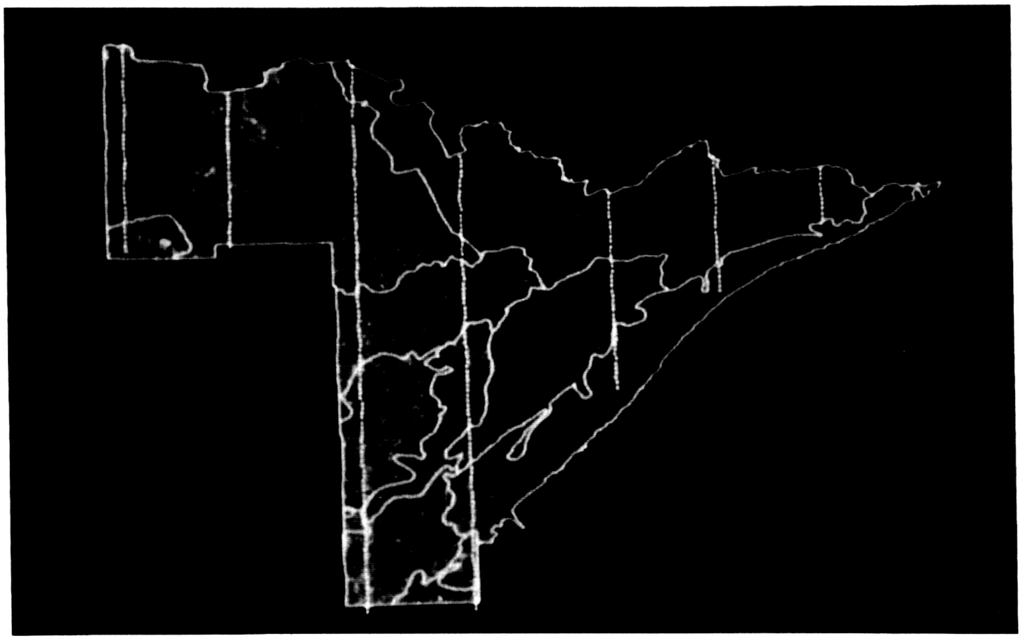


Figure 3. Overlays on Landsat TM image mosaic of aerial photography flightlines and boundaries of physiographic strata.

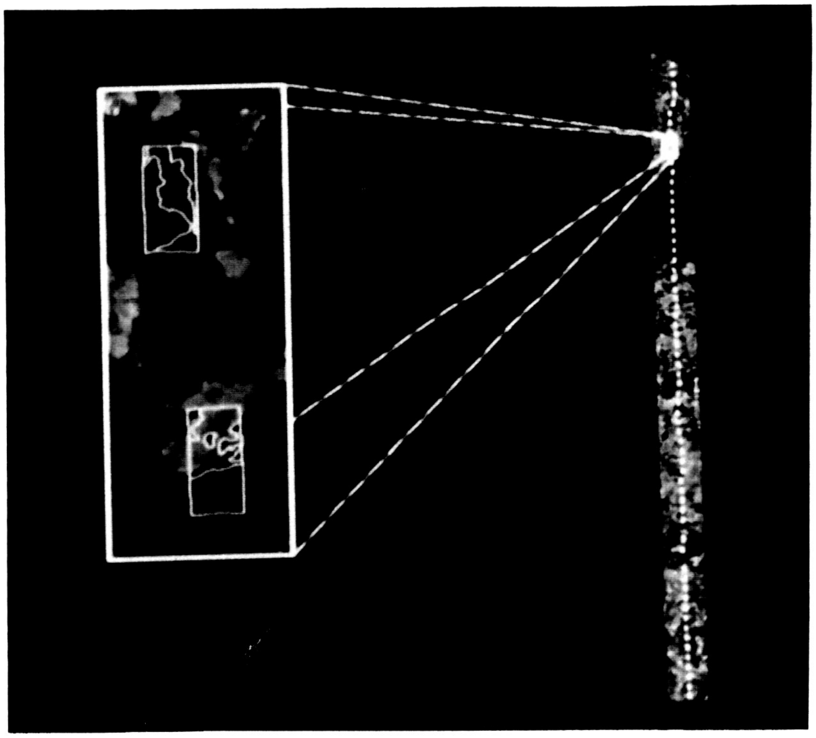


Figure 4. Overlays for two cluster samples of forest type polygons on Landsat TM imagery (left). Locations of aerial photography flightlines and sample clusters are shown on right image.

Table 2. Distribution (%) of covertypes for each physiographic region from reference data.

Region/Class	LH	A/B	NH	UC	BF/WS	LC	S/G/C	Ag	D	W	M/M
Agassiz	2.0	44.0	0.0	2.3	3.0	13.3	6.0	3.6	1.0	20.3	4.5
Border Lakes	0.0	22.9	0.0	29	2.0	12.6	1.1	0.0	0.2	31.1	1.1
Laurentian	0.0	40.4	0.7	19.1	4.3	26.5	5.3	0.0	0.0	1.5	2.2
Mille Lacs	4.1	23.9	7.0	0	0.4	13.3	13.4	25.2	1.9	3.0	7.8
Iron Range	1.0	41.2	1.9	7.4	2.6	9.8	8.6	2.8	18.8	5.2	0.7
Tamarack	4.1	22.2	4.0	0.0	3.7	34.5	18.2	3.9	4.9	0.1	4.4
LS Highlands	1.1	39.9	11.8	3.0	3.6	10.2	9.4	7.1	5.1	6.4	2.4
Tamarack II	2.1	39.9	1.3	3.3	1.3	19.8	12.4	2.9	1.9	7.5	7.6

we had hoped. Each physiographic region image was classified using the guided clustering approach described above. The physiographic stratification prior to training and classification increased classification accuracy an average of 15% over full-scene classifications. Once classified, the subregion GIS files were re-assembled into the county and unit files.

After classification the images were filtered to remove "salt-and-pepper" artifacts in an attempt to re-create the forest stand (polygon) structure inherent in the reference data. Tests against reference data indicated an optimal window size between 4 and 5 pixels square. A 5 x 5 window was chosen for ease of application and to minimize bias. Summary statistics for use in area estimation procedures, discussed in later sections of this paper, were created using AWK scripts.

Final classifications were completed for the 11 target classes for the entire study area (Figures 5 and 6). Overall classification accuracies ranged from 64 to 80%, with average class accuracies from 63 to 76% (Table 3). Kappa statistics (Congalton and Mead, 1986) ranged from 0.56 to 0.76. Example error matrices are given in Tables 4 and 5. Overall accuracy of classification of forest, nonforest, and water for these two examples was 86% for Koochiching County and 85% for Lake Superior Highlands. The overall accuracy of classifying conifer vs. hardwood forest, along with nonforest land and water, was 79% in Koochiching County and 77% for the Lake Superior Highlands.

The majority of classification errors for forest classes occurred in adjacent classes such as lowland hardwood and aspen/birch, lowland conifer and balsam fir/white spruce, and northern hardwood and aspen/birch. The classes of agriculture (cropland/pasture), developed/other, water, and marsh were classified relatively accurately, generally 75% or higher. The class, shrub/grassland/cutover, was the most prone to misclassification; it was confused with all classes, especially other types of forest and nonforest vegetation. In general, similar-related classes were more likely to be confused with each other than with different classes.

Figure 5. Final classification of Landsat TM data of St. Louis County.

Figure 6. Final classification of Landsat TM data of entire five-county study area (FIA unit 1).

Table 3. Number of clusters used, Kappa statistic, overall and average class accuracy for counties (or image subsets) based on training data.

County or physiographic region	N	Kappa	Overall Accuracy (%)	Average Class Accuracy (%)
Koochiching (path 28)	67	.61	73.0	63.0
Lake and Cook Counties (path 26)	68	.60	68.2	68.8
Agassiz (path 27)	43	.62	69.2	74.6
Border Lakes (path 27)	38	.64	72.9	72.0
Laurentian (path 27)	22	.67	75.8	74.9
Mille Lacs (path 27)	19	.76	80.2	75.9
Iron Range (path 27)	31	.58	67.6	65.6
Tamarack (path 27)	26	.60	67.5	68.5
Lake Superior Highlands (path 27)	42	.64	70.3	73.3
Tamarack II (path 27)	21	.56	64.0	73.7

Part of the misclassification results from the considerable difficulty in assigning a unique label to each polygon in the reference data or pixel in the Landsat data. The traditional concept of a forest stand, which we used in developing the reference data, often does not relate to or capture the variability that is present in the multispectral satellite imagery. Forest stands tend to be defined by management considerations, whereas the multispectral radiances measured by the TM data are determined by the biophysical properties of the covertye(s) within each pixel.

Two additional observations can be made about classification accuracy. First, while satellite imagery is typically referred to as having low resolution (at least in comparison to aerial photography), it actually has much higher resolution than forest covertype maps. The minimum mapping unit for our reference data was 2.5 acres, compared to the 1/4-acre pixel size of the TM data. Many inclusions of spatially small covertypes were not mapped in the reference data, and even if correctly classified in the TM data would show up as classification errors. Second, at the pixel level the accuracy of the reference data is probably no better than 75-80%. Therefore, it is impossible to achieve measured classification accuracies of more than that. It also can be noted that large area classification is a more difficult problem, likely to result in lower classification accuracy, than found in previous studies which were typically classifications of small areas.

Table 4. Classification accuracy assessment for Koochiching County.

Ref. Class	Landsat Class								
	LH	A/B	BF/WS	LC	S/C/G	Ag	Dev	Water	M/M
LH	798	727	80	138	244	76	7	21	4
A/B	147	2364	111	220	332	59	18	51	0
BF/WS	86	287	801	307	78	1	3	4	0
LC	16	626	69	11471	522	10	4	18	52
S/C/G	23	309	81	344	838	43	3	8	9
Ag	28	137	59	55	169	700	26	0	0
Dev	6	64	4	0	8	3	188	5	0
Water	30	128	25	44	28	1	2	403	0
M/M	8	85	25	542	26	1	0	2	530
% Corr.	69.9	50.0	63.8	87.4	37.3	78.3	78.3	78.7	89.1

Overall Accuracy = 73.1%, Average Class Accuracy = 69.4%, KHAT = .61

Table 5. Classification accuracy assessment for Lake Superior Highland physiographic region.

Ref. Class	Landsat Class										
	LH	A/B	NH	UC	BF/WS	LC	S/C/G	Ag	Dev	Water	M/M
LH	172	17	5	0	0	0	22	0	0	0	0
A/B	3	4329	174	83	70	126	331	25	43	132	54
NH	11	370	1667	20	0	23	103	39	29	0	0
UC	0	117	2	348	5	15	27	2	7	8	2
BF/WS	0	238	2	3	404	27	61	3	17	2	6
LC	0	574	86	0	79	1449	76	0	29	39	11
S/C/G	3	471	3	20	18	6	727	38	50	5	7
Ag	0	304	11	18	15	24	139	1050	119	1	0
Dev	0	122	31	4	7	24	68	34	507	6	0
Water	0	76	0	5	0	0	0	0	53	844	19
M/M	0	65	0	8	4	20	18	0	10	36	289
% Corr.	91.0	64.7	84.1	68.4	67.1	84.5	46.2	88.2	58.7	78.7	74.5

Overall Accuracy = 70.3%, Average Class Accuracy = 73.3%, KHAT = .64

Calibration and Evaluation of Landsat Area Estimates

Calibration of Landsat Estimates

The satellite imagery provided estimates of type acreages for the area as a whole. However, it is probable that those estimates have a bias associated with them. The ground survey clusters provided observations on how the satellite data classifications compared to ground classifications, with the ground classifications assumed to be truth. Observations from the clusters can thus be used to adjust the satellite-data-based estimates, a procedure we hope will result in a reduction or elimination of bias. This adjustment is commonly referred to as calibration.

As part of the project, Walsh and Burk (1993) compared the classical and inverse methods of calibrating satellite classifications where the ground survey units were pixels. They found, through extensive simulations, that the inverse method was superior. We based our choice of the inverse method of calibration for cluster-based ground units on that previous result and more theoretical reasons. With the inverse method, the satellite classifications are used as "independent" (without error) variables in the regression-based calibrator, while the ground classifications act as "dependent" (with error) variables. These are reasonable roles for these variables as the satellite classifications are surely fixed once the classifier is trained (results are conditional on that training) and interest lies directly in predicting ground classifications (truth).

The unit of observation for the calibration problem was the ground cluster. For each cluster we have a vector of satellite data classifications and a vector of ground classifications, the length of each vector being the total number of types being classified (each vector may have several zero elements). The elements of the vectors are percentages of the total cluster area classified as belonging to a particular type. The smallest area of interest in our application is the county, with each county having a number of ground clusters on which classification results were recorded. The total area of interest, an FIA survey unit, is a set of counties (five in the present case) that are contiguous and of similar forest composition.

Eleven types (classes) were used in classification. If there are n_I ground clusters present in county I, we can specify an n_I x 11 matrix **X** for the county that contains the satellite-data-based classification results for the clusters located in the county. This provides a system of (11) equations relating the true type classifications to **X**:

$$\begin{aligned} Y_1 &= X\beta_1 + \varepsilon_1 \\ Y_2 &= X\beta_2 + \varepsilon_1 \\ &\vdots \\ Y_{11} &= X\beta_{11} + \varepsilon_{11} \end{aligned} \qquad [1]$$

where: Y_i = percent of cluster area ground classified into type i for each of n_I clusters
X = matrix of satellite-data classifications
β_i = regression coefficients relating ground classifications for type i to satellite data classifications for a county
ε_i = error in predicting ground classification from the satellite-data-based classifications.

This system of equations was fit separately to each of the five counties in the survey unit. Ordinary least squares (OLS) was used. The results showed that 85% of the variation, on average, in the ground classifications was explained by **X**. With few exceptions, each regression was dominated by the satellite-data classification corresponding to the type being predicted. However, all elements of **X** were retained in each equation to insure additivity of the predicted percentages. Individual residual plots gave no indication of heterogeneous variance within any particular equation.

The system of equations [1] is most properly considered as a seemingly unrelated regressions (SUR) problem. While OLS provides unbiased estimates of β for SUR problems, accounting for cross-equation correlations (the ground type classifications would likely seem correlated) can result in a more efficient estimate of β. However, Ericksson (1989, p. 45) showed that when **X** is identical for each equation in the system, SUR solved by generalized least squares is equivalent to OLS on each equation separately (assuming homogeneous errors within an equation). That is the case in the present application.

For notational convenience we stack the system of equations [1] for a particular county and write them as:

$$Y_I = X_I\beta_I + \varepsilon_I \qquad [2]$$

where the subscript I denotes the county. Here Y_I and ε_I are vectors of length $11n_I$, $\mathbf{X}_I$ is a matrix with blocks of **X** along the diagonal, and β_I is a vector of length 121 (the coefficients vary by county).

While, overall, the results of fitting [2] were satisfactory, individual equations where there were few non-zero observations of ground type in a county had high standard errors of prediction, often exceeding 100 percent of the estimated proportion. As an alternative to calibrating counties separately, data across counties (within the survey unit) can be combined to estimate a pooled regression equation. The assumption needed to justify such an approach is that the coefficients relating ground types to satellite-data-based classifications are similar across counties within a survey unit (not that the proportions of the various types are similar). Under this assumption the calibration equation becomes:

$$Y_I = X_I\theta + v_I \qquad [3]$$

where the θ are pooled or survey unit wide coefficients. Each type has a separate set of calibration coefficients, but the coefficients for a type are constant across counties. Equation [3] was also fit to available cluster data.

If x_I is the vector of satellite-data-based classifications for county I (classifying the entire land area in the county), we have two calibrated estimates of percent land area by type:

$$y_I = x_I\beta_I$$
$$y_I^* = x_I\theta$$

The two calibrations y_I and y_I^* can be combined to produce an estimate that should have lower error than either of the two individually (Burk and Ek, 1982):

$$\tilde{y} = wy_I + (1 - w)y_I \qquad [4]$$

Both y_I and y_I^* are additive; the predicted type percentages add to 100. For the combined estimate to be additive requires that w be a constant across types. The optimal w for a type is a function of the ratio of the prediction variances of y_I and y_I^*. Computation of estimates of those variances indicated that their ratios were relatively constant across type. An average value (0.65) was used to obtain final estimates: this gives weights 0.606 and 0.394 for y_I and y_I^*, respectively. The final calibrated percentages of each covertype are given in Table 6.

Table 6. Calibrated Landsat estimates of area (000 acres) of six forest and five non-forest classes.

	County					
Class	Carlton	Cook	Koochiching	Lake	St. Louis	Total
Lowland Hardwood	3.1	0.0	131.6	0.0	34.3	169.0
Aspen/Birch	209.3	474.9	548.3	547.5	1,667.4	3,447.4
Northern Hardwood	50.7	68.7	0.0	78.8	56.4	254.6
Upland Conifer	9.1	135.7	3.8	269.1	418.6	836.3
Balsam Fir/White Spruce	2.7	83.9	144.5	44.2	112.9	388.2
Lowland Conifer	74.5	104.1	789.6	306.4	771.7	2,046.2
Shrub/Grass/Cutover	83.4	17.8	223.2	42.6	415.1	781.1
Cropland/Pasture	81.3	0.0	81.1	0.0	124.7	287.2
Developed/Other	13.1	3.6	9.3	13.4	202.8	242.1
Water	16.3	115.3	65.4	149.3	453.5	668.2
Marsh/Muskeg	18.0	30.1	21.6	15.3	58.6	143.6
Total	560.4	1,033.1	2,018.5	1,466.5	4,315.9	9,394.4

Comparison of FIA and TM Estimates

By aggregating the FIA statistics (Miles and Chen, 1992) we are able to make a rough comparison of estimates from the calibrated Landsat classifications and the FIA statistics for conifer, hardwood and total forestland at the county and region levels. Because of differences in definition of classes in the two inventories, it is difficult to make comparisons of more specific covertypes. The problem is that, on the one hand, the FIA covertype area statistics are reported only for commercial forests or "timberland" (defined as forest land capable of producing 20 cubic feet per acre of industrial wood crops...) and therefore does not include detailed breakdowns of the cover types for non-commercial, as well as reserved, forest land (reserved land includes state parks and the Boundary Water Canoe Area Wilderness). On the other hand, the Landsat classifications are for all forest lands.

In comparing the results of the two surveys we have assumed the same proportions of covertypes for unproductive and reserved lands as for the timberland. However, it is well understood that much of the unproductive land will be in lowland conifer types such as black spruce. We have therefore restricted the comparison to conifer, hardwood and total forest. The comparison of area estimates for these classes from the TM classifications and the FIA is shown in Table 7. At the region level, the differences in the two estimates (with FIA as the standard or base) are +0.8, -6.0, and -3.0% for conifer, hardwood, and total forest, respectively. Differences in total forest area estimates at the county level range from -5.0% to +3.9%. The sampling errors for timberland (not total forestland) range from 0.85 to 2.39% at the county level and is 0.57% at the region (survey unit) level.

Table 7. Comparison of FIA and Landsat TM estimates (in 000 of acres) of conifers, hardwoods, and total forest land.

Class	Estimate	County					Total
		Cook	Carlton	Koochiching	Lake	St. Louis	
Conifer	FIA	79.6	376.9	975.1	559.5	1,253.1	3,244.2
	TM	86.3	323.6	937.9	619.7	1,303.2	3,270.7
Hardwood	FIA	273.1	478.0	757.7	639.5	1,970.6	4,118.9
	TM	263.1	543.6	679.9	626.3	1,758.1	3,871.0
Total Forest	FIA	352.7	854.9	1,732.8	1,199.0	3,223.7	7,363.1
	TM	349.4	867.2	1,617.8	1,246.0	3,061.3	7,141.7

There are several sources or causes for differences in the two estimates. The first is that in three cases, the available reference data did not include samples of all FIA classes; in other words *n*, as reported in Table 3, was too small. A larger, more fundamental problem is that the forest lands in northern Minnesota are very complex ecosystems with a wide range of variability in composition and uniformity. This heterogeneity limits the accuracy of the photo interpretation and field checking, and in turn the reference data. This in turn affects the accuracy of class labeling. In many respects we are attempting to classify continuous data into discreet classes; doing so results in errors in the areas of transition from one type to another. The Forest Service recognizes the complexity of the forests and the presence of mixtures in its definitions of classes. For example, the description of red pine is, "forests in which red pine comprises a plurality of the stocking; common associates include eastern white pine, jack pine, aspen, birch, and maple.

Cluster-Based Estimation of FIA Covertype Areas

While satellite data classifications were limited to six forest types, forest management needs are often more specific. However, the Y_i need not be constrained to the same classes for the image classification; instead they can be defined as any ground truth variable. For example, the cluster samples used in this study were originally typed in 13 forest types. To estimate FIA covertype acreage directly, we define Y_i, $i = 1,...,13$, as the proportion of a cluster in FIA type i. The estimation model is then the same as [1]. Again, this system of equations is additive.

The approach is described below for St. Louis County. Figure 7 shows the standard errors of the mean proportions for 13 standard FIA covertypes, plus an other/nonforest category, as estimated from the 88-acre clusters, in comparison with theoretical standard errors. The theoretical standard errors were developed in two ways: (1) assuming each cluster location was instead the location of a standard FIA plot (covering approximately one-acre) and (2) a lower bound assuming the clusters were broken up and distributed as n x 88 one-acre FIA plots. The fact that the standard errors for the cluster sample lie approximately midway between the two theoretical estimates suggests that the clusters are far more effective at error reduction than a sample of n FIA plots and substantially less effective than the much larger number of plots that might have been distributed at random had the clusters been broken up and checked as one-acre components.

Ultimately, the choice of survey design must be considered on a cost basis. While the clusters are clearly less efficient than an equivalent acreage of random plots, the travel costs are dramatically reduced (fewer clusters than plots) and the cost of a cluster need not be much if any more expensive than the current FIA plots. The latter typically cost $150-300 each and involve 1-2 people and approximately one day of time including travel. Of that day, much of the effort goes into establishing and measuring a ten point cluster of small plots on an acre. We propose instead that those small plots be spread across the cluster and be used to verify the covertype of the polygons on the cluster. Use of large clusters is not unlike what has been done in Scandinavia (Kuusela 1978; Svenson 1980) and what was found as an optimal "supercluster" by Scott, et al. (1984).

A spreadsheet analysis of alternative forest inventory designs is currently being developed including this cluster design, the current multiphase FIA procedures and other designs. That effort will also consider optimal cluster size. However, the optimal cluster size will also depend on practical concerns for being able to locate it and potential data analysis as described below. For analysis, precision of this approach will be developed empirically from these results and additional cluster sizes to be tested. It is probable that a planning model useful to inventory design and analysis will express sampling error (or variance) as a function of the covertype proportion and the area or size of a cluster for any given classifier and costs.

Additional important aspects of the cluster based design are its statistical simplicity and its potential utility for monitoring landscape patterns. The simplicity comes from observation of operational sized land units and the fact that it requires only single phase estimation and ordinary least squares procedures for area estimates. The power of the approach is that it uses as covariates all of the pixels in the county or area of interest yet it doesn't necessarily require high accuracy in the satellite classification. Additionally, one could improve estimates by the methods described in the earlier calibration section. Statistically one may view the cluster as simply a large number of large image classification plots. The precision of clusters overall will be less than the same acreage small plots distributed randomly over the survey unit, but the cost of assessing them and particularly the realism it adds to classifier training, plus their modest ground assessment cost provide significant gains for the survey design.

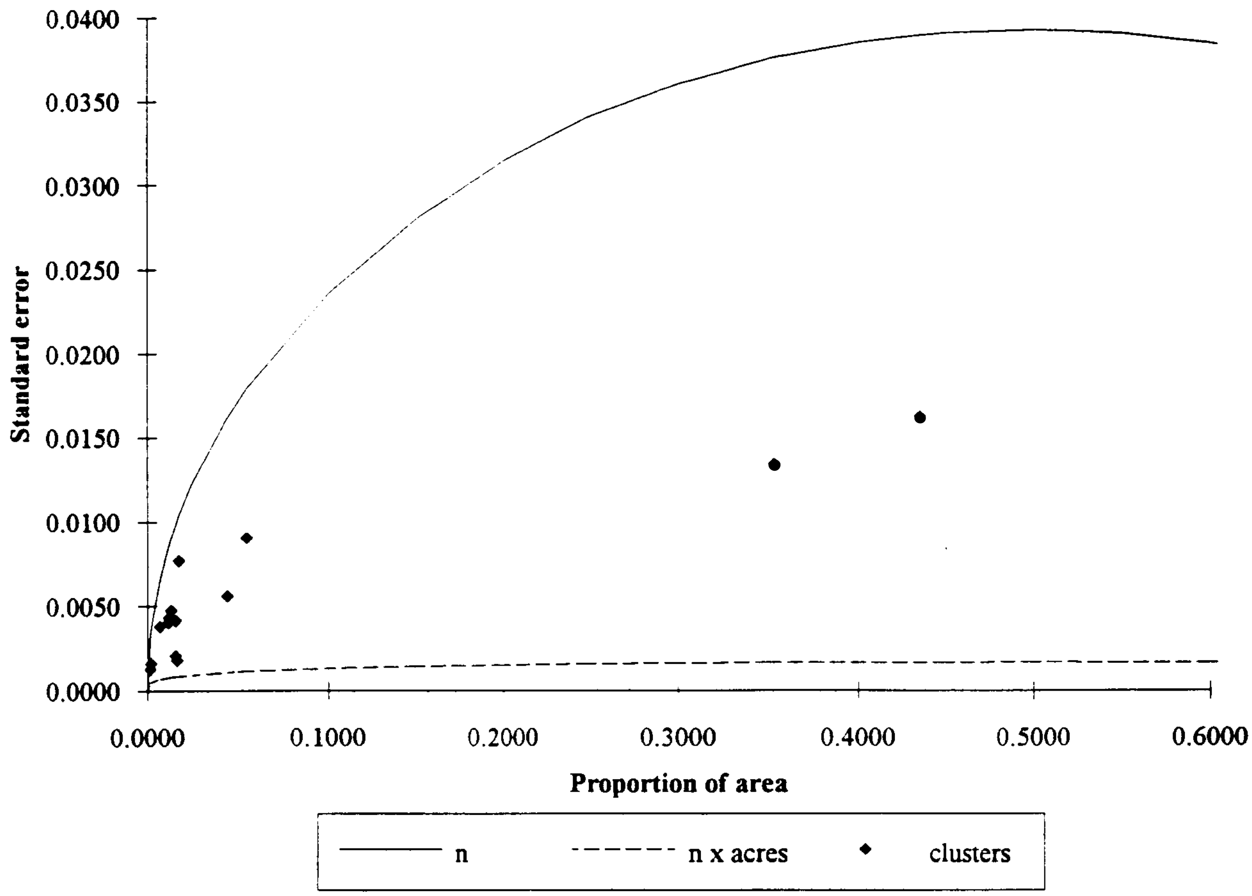


Figure 7. Standard error comparisons for estimations of proportions of 14 FIA covertypes and other category for 163 plots. n = 163 plots, n x acreage of cluster, and actual clusters as estimated for St. Louis County.

Change Detection using Multidate Landsat Data

Renewable natural resources such as forests are continually changing. Some forest cover modifications are human-induced, such as harvest, while others have natural causes, such as insect or disease damage. The rate of change may be abrupt (e.g. logging) or subtle/gradual (e.g. growth). The potential of using satellite data to detect and characterize changes in forest cover depends on the ability to quantify temporal effects using multitemporal data sets. As a part of the research under the EOCAP project we investigated the potential of multitemporal Landsat TM for forest cover change detection (Coppin, 1991).

TM data, along with detailed ground reference data, for three different years (1984, 1986, and 1990) covering a 400 km^2 (5 townships) test site in Beltrami County were acquired. To minimize sensor calibration effects and standardize data acquisition effects, the TM data were calibrated to exoatmospheric reflectance. After geometric rectification and registration, an atmospheric correction routine was applied, combining two major components: atmospheric normalization and transformation to ground reflectance. The normalization consisted of a statistical regression over time, based on five spatially well defined landscape features with unchanging spectral-radiometric characteristics. Linear correlation coefficients for all bitemporal band pairs ranged from 0.9884 to 0.9998 (Figure 8). The transformation made use of a simplified dark object subtraction technique incorporating published values of water reflectance.

For each time interval (two, four and six) years, 14 change features were determined using seven spectral bands and vegetation indices and two change detection algorithms (standardized differencing and pairwise principal components) were evaluated. The best four features for classification were selected based on J-M distance calculations of the best minimum separability between change signatures. A maximum likelihood classifier was used for the final classification. Classification accuracy and areal correspondence were evaluated from contingency matrices and KHAT values.

The results (Figure 9) demonstrated that disturbances and other changes can be detected very accurately if categorized in classes that relate to their effect on the forest canopy, and if their size exceeded one hectare. Forest stands as the classical management units were ascertained to be too spectrally heterogeneous to have the change phenomena differentiated at that level. However, for the three classes, canopy depletion, canopy increment, and no change, the methodology correctly identified 714 out of 759 (94%) stands reported as disturbed over the six-year interval (Table 8), indicating that the change event was portrayed in a majority of the stand's pixels. The Kappa statistic, which removes the contribution of correct classification due to chance was 0.82. The results show that the preprocessing sequence summarized above is critical to the forest cover monitoring; similar preprocessing and calibration procedures are being used in the large scale application of this technology by the Minnesota DNR described below.

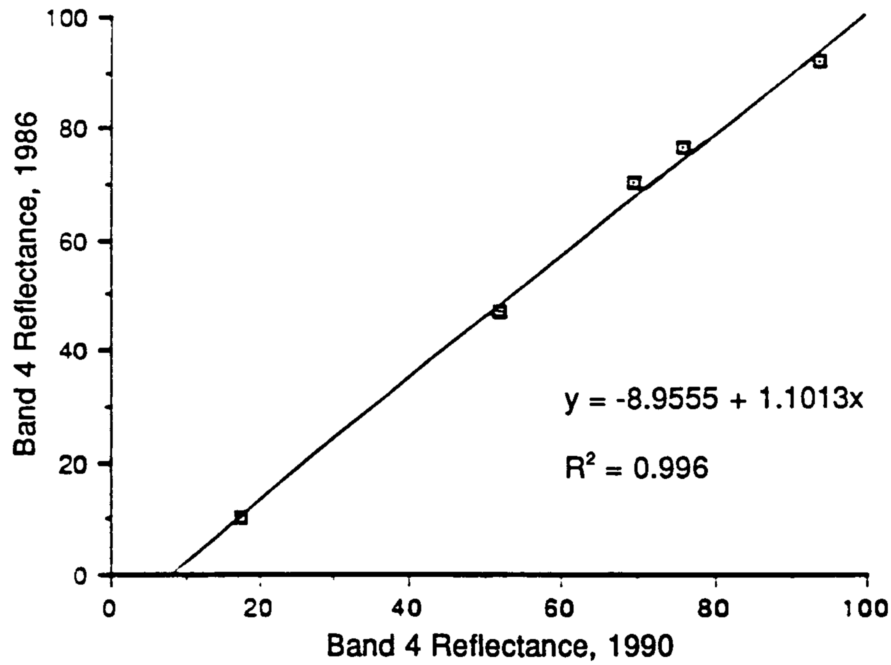


Figure 8. Example of the relationship between band-specific scaled reflectances of calibration sites over time.

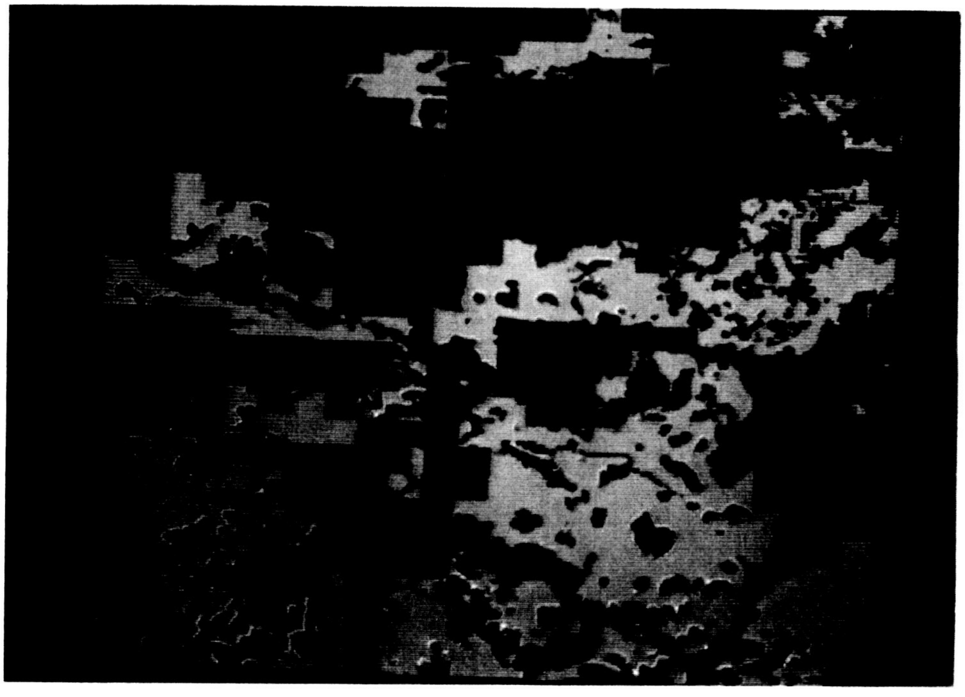
Table 8. Summary of statistics evaluating change detection classification accuracy.

Time Interval	Thematic Accuracy		
	Overall Accuracy (%)	Average Class Accuracy (%)	Kappa Statistic
2 years	97	79	76
4 years	96	90	83
6 years	94	91	82

Time Interval	Cartographic Accuracy*		
	Canopy Depletion	Canopy Increment	No Change
2 years	.64	.66	.97
4 years	.78	.74	.92
6 years	.79	.78	.89

* Pixel-based areal correspondence coefficients

(a) Reference data.



(b) Classification results.

Figure 9. Comparison of change events in the reference data and the Landsat TM classification of Jones Township, 1984-90. Red = canopy depletion, blue = canopy increase, green = no change.

Operational Use of Landsat Data for Forest Inventory in Minnesota

As a result of the research described above, the Minnesota DNR and the U.S. Forest Service have jointly undertaken to develop an Annual Forest Inventory System (AFIS) based on annual sampling of existing Forest Inventory and Analysis (FIA) ground plots in Minnesota (Befort and Heinzen, 1992). Satellite remote sensing plays an important role in the AFIS plan, and initial Landsat data analysis is well under way.

The objective of AFIS is to create and maintain a current and continuously updated FIA database. Under the proposed system, a relatively small proportion of FIA plots will be chosen each year for field remeasurement; information on other plots will be updated by use of forest growth models. Selection of plots for measurement is to be based on 1) likelihood of plot disturbance since the last field measurement, and 2) requirements of a 20-year sampling rotation in which all plots are ultimately field-visited. Satellite remote sensing has two roles: it will be used to stratify a statewide array of some 45,000 established FIA plot locations as a means to reduce variance, and second, to estimate the likelihood of change or disturbance on each plot in order to prioritize plots for field measurements. In the past, aerial photography has been used for both these purposes, but the cost of obtaining, handling and interpreting aerial photography is prohibitively expensive and time consuming. The use of satellite imagery is expected to reduce costs and allow an increase in the frequency of inventory updates.

The general remote sensing approach is to move through the state on a four-year rotation, covering one of the four major forest inventory regions each year. For each region, geo-referenced Landsat data of the most recent late summer date is obtained. (Work on the Aspen-Birch (unit 1) region is currently underway using Landsat data acquired in 1988 and 1992). After preprocessing, including rectification, radiometric calibration, and atmospheric normalization, a general land cover classification (stratification) is performed across multi-county survey units. Because of the difficulties in achieving accurate Landsat classifications of forest cover types, only broad categories of water, agriculture, other nonforest (e.g. developed/urban, clouds, etc.), conifer forest and hardwood forest will be mapped. The new imagery is then registered to that of the previous iteration and analyzed for change. Based on changes in the vegetation index (e.g. greenness) a change ranking or probability is generated for the pixels in the two forest classes. Change classes might be: no change between time 1 and time 2; minor change, but with the same stratum class, or major change from one stratum to another. Digitized locations of the FIA plots are then queried for stratum identity and change ranking. These two data elements, together with area expansion factors for all cover classes, are entered into the plot database for use in an algorithm that selects plots for field measurements in the following season or for projection forward by a forest growth model. Annual field measurements will serve as a check on image processing accuracy.

Because of its annual schedule, the AFIS plan requires current, inexpensive large-area imagery together with low-cost, but robust, interpretation methods for both stratification and disturbance detection. The research by Coppin (1991) has indicated that computer analysis of multidate Landsat data (summarized above) offers a cost-effective alternative to reliance on aerial photography. We believe that Minnesota is the first state to incorporate satellite remote sensing into its forest inventory system. If successful, the techniques could be easily modified for implementation in other states.

Summary and Conclusions

The objective of this research was to develop and test the use multispectral satellite data together with improved classification and sampling designs to inventory the forest resources of northeastern Minnesota. Two design alternatives were considered: one based on cluster sampling concepts and a second that considered disturbance classification as the basis for stratified, two-phase sampling.

Classification accuracies of up to 75% for six forest classes and five nonforest classes were achieved. Misclassification tended to be between similar-related classes. A major contributing factor to the difficulty in classification is the fact that the majority of forest stands are complex mixtures of two or more species which may also differ in size, density, crown closure, and age.

An inverse method of calibration was used to adjust the classifications for classification bias. At the survey unit level, the resulting estimates of forest land area were 3% less than comparable Forest Service estimates. Agreement between the two surveys at the county level ranged from -5.0 to +3.9%. The difference in estimates is attributed to differences in definitions and approaches used in the two surveys, as well as the complexity and variability of the forest landscape. A trial of estimating the acreage of 14 traditional forest covertypes as determined from sample 88-acre clusters as a function of the 11 satellite classes using a system of additive linear equations was also conducted. Errors were found to be comparable in magnitude to standard forest inventory estimates. This type of cluster-based estimate is potentially very cost-effective and provides data of increasing interest to the assessment of covertype and land use patterns over landscapes.

For the disturbance classification approach (also known as change detection) working with multidate imagery, classification accuracies of greater than 90% were obtained for the following four classes for time intervals of two, four and six years: active canopy depletion, active canopy growth, no change and storm damage (two years only). The success rate for the detection of stand-based canopy change events over the six year interval was 94%, while the pixel-based thematic accuracy resulted in an average class accuracy of 91% with a KHAT value of 0.82. The key to obtaining these results was a rigorous approach to reflectance calibration and normalization for atmospheric effects.

The results have provided the basis for the Minnesota Department of Natural Resources and the USDA Forest Service to define and begin to implement a statewide inventory system which utilizes multidate Landsat TM data to detect changes in forest cover. Landsat TM imagery acquired at four year intervals will be used to detect major changes in forest inventory plot characteristics. The likelihood of change as determined from the satellite data will be used to determine which plots should be revisited for field measurement. The general approach will be to classify one of the four major forest inventory regions of the state each year. Forest growth models will be used to project the growth of plots which are not measured in a given year. Satellite-acquired data are an integral part of the system along with model predictions, sampling and date base techniques. We believe that Minnesota is the first state to incorporate satellite remote sensing into its forest inventory system.

Acknowledgements

The support of the National Aeronautics and Space Administration, NRA-87-106 (Grant NAGW-1431, NRA 87-106), is gratefully acknowledged, along with that of the University of Minnesota Agricultural Experiment Station and Department of Forest Resources, and the Minnesota Department of

Natural Resources, Division of Forestry. The contribution of the Beltrami County Land Department to development of reference data sets was indispensable for the change detection work. Mark Carter provided the photo interpretation and field checking was conducted by Dean Halvorson and Michael Smith. Forest cover type maps used in the initial development of training and classification procedures were provided by the Minnesota Department of Natural Resources, USDA Forest Service, and St. Louis County Land Department.

References

Befort, W. and D.F. Heinzen. 1992. Use of satellite remote sensing for forest inventory by the Minnesota Department of Natural Resources. Technical Report, Minnesota DNR, Division of Forestry, Resource Assessment Unit. Grand Rapids, Minnesota. 28 pp.

Bennesslah, D. 1985. Forest area estimation using cluster sampling in single and double phase sampling designs. Ph.D. Dissertation, University of Minnesota, St. Paul (University Microfilm #DA85-26459). 171 pp.

Burk, T.E. and A.R. Ek. 1982. Application of empirical Bayes/James-Stein procedures to simultaneous estimation problems in forest inventory. Forest Science 28:753-771.

Congalton, R.G. and R.A. Mead. 1986. A review of three discrete multivariate analysis techniques used in assessing the accuracy of remotely sensed data from error matrices. IEEE Transactions on Geoscience and Remote Sensing 24:169-174.

Coppin, P.R. 1991. The change component in multitemporal Landsat TM images: Its potential for forest inventory and management. Ph. D. thesis, University of Minnesota, St. Paul. 173 pp.

Crist, E.P. and R.C. Cicone. 1984. Application of the tasseled cap concept to simulated Thematic Mapper data. Photogrammetric Engineering and Remote Sensing 50:343-352.

DeGloria, S.D. 1984. Spectral variability of Landsat-4 thematic mapper and multispectral scanner data for selected crop and forest cover types. IEEE Geoscience and Remote Sensing 22:303-311.

Ek, A.R. 1983. Use of forest growth models in management inventories. Proceedings, IUFRO Subject Group S4.02, meeting in Finland, Sept. 5-9. Published by Univ. Helsinki, Dept. Forest Mensuration and Management, Research Notes No. 17, pp. 129-136.

ERDAS Staff. 1991. ERDAS Field Guide. ERDAS Inc., Atlanta, Georgia.

Ericksson, M. 1989. Integrating forest growth and dendrochronological methodologies. Ph.D. thesis, University of Minnesota, St. Paul. 303 pp.

Horler, D.N.H. and F.J. Ahern. 1986. Forestry information content of thematic mapper data. Intl. J. of Remote Sensing 7:405-428.

Jensen, J.R. 1983. Biophysical remote sensing. Annals of Assoc. of American Geographers 73:111-132.

Kuusela, K. 1978. The National Forest Inventory in Finland. In: National Forest Inventory. Proc. IUFRO GROUPS S4.02a and S4.04. Institut de Cercetari Si Amenajara Silvice. Bucharest, Romania. pp 368-377.

Miles P.D. and C.M. Chen. 1992. Minnesota Forest Statistics, 1990. Resource Bulletin NC-141. U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, Minnesota.

Moore, M.M. and M.E. Bauer. 1990. Classification of forest vegetation in north-central Minnesota using Landsat Multispectral Scanner and Thematic Mapper data. Forest Science 36:330-342.

Peterson, D.L., W.E. Westman, N.L. Stephenson, V.G. Ambrosia, J.A. Brass, and M.A. Spanner. 1986. Analysis of forest structure using thematic mapper data. IEEE Geoscience and Remote Sensing 24:113-121.

Price, J.C. 1984. Comparison of the information content of Landsat-4 thematic mapper and the multispectral scanner. IEEE Geoscience and Remote Sensing 22:272-280.

Scott, C.T., A.R. Ek, T.R. Zeisler and D. Benesalah. 1984. Cluster sampling: From theory to practice. Proceedings, Symposium on Inventorying Forest and Other Vegetation of the High Latitude and High Altitude Regions. (V. LaBau and C. Kerr, eds.) Society of American Foresters. pp. 78-81.

Svendson, S.A. 1980. The Swedish National Forest Survey, 1973-77: State of forests, growth and annual cut. Swedish University of Agricultural Sciences, Department of Forest Survey Report 30, Umea, Sweden. 167 pp.

Walsh, T.A. and T.E. Burk. 1993. Calibration of satellite classifications of land area. Remote Sensing of Environment. In Press.

Williams, D.L. and R.F. Nelson. 1986. Use of remotely sensed data for assessing forest stand conditions in the Eastern United States. IEEE Geoscience and Remote Sensing 24:130-138.

Wright, H.E. Jr. 1972. Physiography of Minnesota. *in* Geology of Minnesota. Minnesota Geological Survey. St. Paul, Minnesota. pp. 561-578.