This is to certify that the
thesis entitled

COMPARATIVE ANALYSIS OF FOREST CLASSIFICATION IN
FOREST MANAGEMENT INFORMATION DATABASES IN
MICHIGAN

presented by

Nirmal Subedi

has been accepted towards fulfillment
of the requirements for the

Master of Science degree in Forestry

Larry A. Leefers
Major Professor's Signature

August 18, 2005
Date

COMPARATIVE ANALYSIS OF FOREST CLASSIFICATION IN FOREST MANAGEMENT INFORMATION DATABASES IN MICHIGAN

By

Nirmal Subedi

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Forestry

2005

ABSTRACT

COMPARATIVE ANALYSIS OF FOREST CLASSIFICATION IN FOREST MANAGEMENT INFORMATION DATABASES IN MICHIGAN

By

Nirmal Subedi

In Michigan, there are four primary sources of forest management information for public forest lands, namely a raster land-cover map (IFMAP), Forest Inventory and Analysis (FIA) plot-level information, Natural Resource Information System Field Sampled Vegetation (NRIS-FSVeg) for national forest lands, and Operations Inventory (OI) for state-owned forest lands. The objective of this study is to compare forest classifications between and among the forest management databases with FIA data as the reference location for comparison. Difference matrices were created between and among forest classifications and descriptive accuracy assessments for overall accuracy, producer's accuracy and user's accuracy were computed. The overall accuracy of IFMAP with FIA as reference was 63.6% for state forest lands and 64.8% for national forest lands. Overall accuracy of IFMAP with OI as a reference was 60.3% and IFMAP with NRIS-FSVeg as a reference was 68.3%. Overall accuracy of OI with FIA was 84.5% and NRIS-FSVeg with FIA was 82.2%. Overall accuracy of three-way forest classification was 54.8% and 58.5% for state and national forests lands, respectively. Kappa statistic, calculated from three approaches, ranged from 0.568 to 0.628 for state forest lands and 0.555 to 0.612 for national forest lands. This finding is consistent with a previous study of IFMAP.

ACKNOWLEDGEMENTS

This study is a collaborative effort of the Department of Forestry, MSU, the Michigan Department of Natural Resources, the national forests of Michigan, and FIA unit, NCRS. Many people working in different organizations contributed significantly for the successful completion of this thesis. Committee members, Dr. David MacFarlane and Dr. Rique Campa provided valuable contributions in improving the quality of work. Dr. Larry Pedersen, another committee member, was very instrumental for creating amicable environment for inter-agency collaboration and providing empirical understanding of the problem. In particular, my advisor, Dr. Larry A. Leefers, guided me in the MS program, provided generous support for better understanding the research problem, resolving the issues during data analysis and improving the quality of research work. Similarly, Mr. Geoff Holden, NCRS, played a very crucial and cooperative role to accomplish this study on time. Including my committee members, Dr. Donald Dickmann, Dr. Mark Hansen, Mr. Jason Stephens and Mr. Joseph Gates contributed significantly to improve the crosswalk tables for forest classifications. MSU Faculty Dr. Steven Friedman, Dr. Kelly F. Millenbah and Dr. Murari Suvedi were generous in providing valuable inputs for my research as well. The fellow colleagues at MSU particularly Eric White, Anita Morzillo, Qing Xiang and Murari Regmi were very helpful to me while learning new skills required for this research.

I would like to thank Fulbright Scholarship for providing an excellent opportunity to study MS degree at Michigan State University. Special thanks are due to my spouse Manjana, my daughter, Swastika for their patience and understanding and support. Without their cooperation, it would have been impossible to accomplish my goal.

TABLES OF CONTENTS

LIST OF TABLES.....vii
LIST OF FIGURES.....v

CHAPTER 1
INTRODUCTION.....1
Michigan Statewide Raster Map or IFMAP.....3
Operational Inventory.....5
Field Sampled Vegetation of Natural Resource Information System (NRISFSVeg).....6
Forest Inventory Analysis (FIA).....10
Problem Statement.....12
Study Purpose.....13
Organization of Thesis.....14

CHAPTER 2
LITERATURE REVIEW.....16
Land Use Classification.....16
Michigan Resource Inventory System.....21
National Land Cover Database (NLCD) for United States.....22
IFMAP Classification.....23
Forest Classification of FIA/OI/NRIS-FSVeg.....25
FIA Sampling..... 26
Digital Imagery Classification.....29
Remote Sensing Based Cover Type Mapping Studies.....35
Accuracy Measurement Methods Used for Remote Sensing Imagery Classification.....41

CHAPTER 3
DATA SOURCES, STUDY AREA, CLASSIFICATION APPROACHES AND ANALYSIS METHODS 51
Data Sources.....51
Study Area.....53
Classification Approaches.....54
Analysis Methods.....55
Chapter Summary 63

CHAPTER 4
RESULTS AND DISCUSSION.....64
Using FIA Forest Types as Reference Classification.....68
Accuracy Assessment for State Forest Lands.....72
Accuracy Assessment for National Forest Lands.....77
Three-way accuracy of FIA/OI/IFMAP Assessment in the State Forest Lands 85

Three-way accuracy of FIA/NRIS-FSVeg/IFMAP Assessment for the National Forest Lands....88
Comparison of Classification Agreement Between IFMAP and FIA for State Forest Lands.... 91
Comparison of Classification Agreement Between IFMAP and FIA for National Forest Lands....98
Explanation of Differences in Classification of IFMAP, OI with FIA....104

CHAPTER 5
SUMMARY AND CONCLUSION....110
Conclusions....112
Limitations.... 113
Policy Implications....115
Additional Research....116
REFERENCES CITED....117

APPENDIX A
A1. NRIS-FSVeg FIA Crosswalk....125
A2. IFMAP-FIA Crosswalk....126
A3. OI-FIA Crosswalk....127
A4. IFMAP-OI Crosswalk....128
A5. IFMAP-NRIS-Veg Crosswalk....129

LIST OF TABLES

Table 1. Forest Classification of Michigan Statewide Raster Map (IFMAP)..................5

Table 2. Forest Type Classification in Operational Inventory.................................7

Table 3. Forest Types Used in Michigan National Forest......................................9

Table 4. FIA Forest Type Classification...12

Table 5. Anderson et al. (1976) Land Use Classification....................................17

Table 6. Comparison of Forest Management Information Database in Michigan.........65

Table 7. Difference Matrix of IFMAP and FIA for State Forest Lands.....................73

Table 8. Difference Matrix of OI and FIA for State Forest Lands...........................74

Table 9. Difference Matrix of IFMAP and OI for State Forest Lands.......................75

Table 10. Difference Matrix of IFMAP and FIA for National Forest Lands...............78

Table 11. Difference Matrix of NRIS-FSVeg and FIA for National Forest Lands.........80

Table 12. Difference Matrix of IFMAP and NRIS-FSVeg for National Forest Lands....82

Table 13. Three-way accuracy comparison of FIA, OI and IFMAP for State Forest Land 86

Table 14. Three-way accuracy comparison of FIA, NRIS-FSVeg and IFMAP for National Forest Land..89

Table 15. Error Matrix of IFMAP-FIA for State Forest Land (UA constant)...............93

Table 16. Error Matrix of IFMAP-FIA for State Forest Land (PA constant)...............94

Table 17. Error Matrix of IFMAP-FIA for State Forest Land (Aggregated)................95

Table 18. Individual Error Matrix Kappa Analysis Results for State Forest Land.........96

Table 19. Kappa Analysis Results for Pairwise comparison of Error Matrices...........98

Table 20. Error Matrix of IFMAP-FIA for National Forest Land (UA constant).........99

Table 21. Error Matrix of IFMAP-FIA for National Forest Land (PA constant)..........100

Table 22. Error Matrix of IFMAP-FIA for National Forest Land (Aggregated)..........101

Table 23. Individual Error Matrix Kappa Analysis Results for National Forest Land...102

Table 24. Kappa Analysis Results for Pairwise comparison of Error Matrices...........102

LIST OF FIGURES

FIGURE 1. National Forest and Digitized State-Owned Forest in Michigan 52

FIGURE 2. Illustration of an Error Matrix or a Confusion Matrix............................57

CHAPTER 1

INTRODUCTION

Forest resources are increasingly important in providing ecosystem stability, human recreation and other goods and services. Information about forest resources is needed to meet the demand for goods and services from forests, whether in Michigan or elsewhere. "Michigan's forests are making important contributions to the quality of life by providing a wide array of benefits including wildlife habitat, biological diversity, outdoor recreation, and improved air and water quality. Economic contributions are also significant—an estimated $12 billion of value added and 200,000 jobs annually are supported by forest based industries/ tourism/ recreation" (Schmidt et al. 1997). Michigan forestland area totals 19.3 million acres, and timberland totals 18.7 million acres (MDNR, 2002, Leatherberry et al. 2005). The state timberland acreage is the sixth largest in United States exceeded only by the states of Oregon, Georgia, Alabama, Montana and North Carolina (Smith et al. 2004).

Management of forest resources requires information about the resources, such as their location, forest type and other characteristics. In addition, demographic, social and economic information are important to assess demand. Nowadays, publics are demanding a variety of goods and services from forestlands, and multiple use management is the main principle for providing such demands. Economic multiple-use forestry planning is a relatively challenging task and it requires information about resources and their demand.

In this way, multiple use forestry planning has enhanced the importance of spatial information of forest stand characteristics, as it compares the economic benefits from joint production and specialized production. This thesis focuses on spatial information regarding forest type as defined by various information databases.

Forest type information is one of the most valuable pieces of information for forest managers and planners. It can reveal the types of vegetation that are growing which relates to information about the potential production of goods and services at a particular location. In Michigan, four primary sources provide information for forest types available on public lands. However, they sometimes provide conflicting information. The first is a raster map based on remote sensing satellite imagery information, locally named Integrated Forest Monitoring Assessment and Prescription (IFMAP); it provides information statewide (PMR, 2001a). The second is Operations Inventory (OI) developed by the Forest Minerals and Fire Management Division of the Michigan Department of Natural Resources (MDNR) which provides information for state-owned forestlands. The third is the Field Sampled Vegetation of the Natural Resource Information System (NRIS-FSVeg), developed by the United States Department of Agriculture-Forest Service (USDA-FS) which has information for national forests. The other information source is the sampled plot information inventoried by the USDA-FS Forest Inventory and Analysis (FIA) staff. For every FIA plot on state and national forests, there is a corresponding IFMAP pixel (raster) and either an OI (MDNR) or NRIS-FSVeg (USDA-FS) classification.

Michigan Statewide Raster Map or IFMAP

IFMAP is a statewide map of land cover over all ownerships within the state of Michigan. It also provides a means for the MDNR to develop management prescriptions for their lands. Only the former is used in this research. Basically, this map is derived from Landsat Thematic Mapper 5 and 7 Enhanced Thematic Mapper plus (ETM +) with some ground verification. There are two separate maps, one for the Upper Peninsula and the other for the Lower Peninsula. This is a coarse resolution map for producing a land cover map and dataset, which can serve multiple functions (MDNR 2003b, MDNR 2003c). It is anticipated that this map will provide a resource for ecosystem-scale management and a statewide planning tool for wildlife habitat (PMR, 2001a). In addition, the pixel-level information of land cover can be integrated to assess the resources on a unit or county-wide scale. Each pixel has spatial resolution of 30 meters X 30 meters (PMR, 2001a).

The inventory module of IFMAP records information on present land cover, and the activity tracking module records information on treatments and disturbances (PMR, 2001b). These two types of information are kept in separate GIS map layers. In IFMAP, a proposed treatment polygon can be a subset for a land cover stand or even from several adjacent stands.

The percentage of canopy occupied by a given species was used for rules to derive land cover and association categories (PMR, 2001a). The percentage of canopy cover is determined by a sensing mechanism, either satellite-borne or airborne. Because it is overhead, there is an extremely limited capability for canopy penetration; below-canopy data is not used in this classification. The classification scheme has been examined by representatives of all regions of the state, and resulting changes/additions made help account for the variety of species, species associations, and land cover types found in Michigan.

Land use in IFMAP is classified into eight land-use classes: Urban, Agriculture, Upland Open land, Upland Forest, Lowland Forest, Non-forested Wetland and Bare/Sparsely Vegetated. In the Upland classifications lands not periodically flooded or not on hydric soils are included, and Lowlands refer to lands that are periodically flooded and/or on hydric soils. Land is classified as forest when the canopy cover exceeds the 25% of the ground. The algorithm for the forest type classification can be assessed via the Michigan Geographical Data Library website (MDNR 2003a).

IFMAP has classified forestland into 12 different forest types in the Lower Peninsula and into 11 different forest types (excludes other upland deciduous) in the Upper Peninsula (Table 1). There are 20 other non-forestry land use classifications to cover the entire State. Thus, IFMAP facilitates statewide identification of the land-use within 32 land-use categories (for details see, MDNR, 2003e).

Table 1. Forest Classification of Michigan Statewide Raster Map (IFMAP) [a]

Forest Land Use Class	Forest Type	Grid Value
Upland Deciduous Forest	Northern Hardwood Association	14
	Oak Association	15
	Aspen Association	16
	Other Upland Deciduous	17
	Mixed Upland Deciduous	18
Upland Conifer Forest	Pines	19
	Other Upland Conifers	20
	Mixed Upland Conifers	21
	Upland Mixed Forest	22
Lowland Forest	Lowland Deciduous Forest	24
	Lowland Coniferous Forest	25
	Lowland Mixed Forest	26

Source: Michigan Geographic Data Library, 2003 (http://www.mcgi.state.mi.us/mgdl/)
[a] Non-forested land with lower canopy class (less than 25%) include Herbaceous openland, Upland and lowland shrub / Low density trees, Mixed non-forested wetland and Other bare/ sparsely vegetated.

Operations Inventory

Operations Inventory (OI) is an inventory system developed by the Forest, Minerals and Fire Management Division (FMFMD) of the MDNR. OI locates and identifies physical, biological, economic, and social information on each unit of land (MDNR, n. d.). OI is expected to provide operational level information related to resource management issues regarding timber, wildlife, forest recreation, water quality, reforestation and land use. It provides the descriptive information at the stand level (the smallest record keeping unit) and a plan of operation for the stand after a multidisciplinary review of a preliminary prescription.

Within OI, stands are grouped into compartments (group of stands with average area of 1500 to 3000 acres) based on proximity, common access, landform and soil properties and uniformity on distribution of major cover type acreages (MDNR, n. d.). The OI definition of forest is very similar to FIA, it identifies a stand as a forest when it has at least 16.7% stocked and land capable of producing 20 cubic feet per year (Pedersen, pers. comm.). Compartments are grouped by year-of-entry (YOE). The OI analysis proceeds by compartment within a given YOE. Each YOE contains approximately 10 percent of the compartments in a forest area (MDNR, n. d.). By the end of Fiscal Year 2004, the inventory of compartments with YOE 2003, 2004, 2005 and 2006 was completed, and the boundaries of compartments and stands were digitized using a geographic information system (GIS). The inventory of whole forest area owned by MDNR will be completed in six more years. This research is limited to the compartments within these four YOE. The OI has classified the forest type (referred to cover types) into 16 categories (Table 2).

Field Sampled Vegetation of Natural Resource Information System (NRIS-FSVeg)

The USDA-FS's Field Sampled Vegetation of Natural Resource Information System (NRIS-FSVeg) combines a standard corporate database and computer applications designed to support field-level users. NRIS databases contain basic natural resources data in standard formats built to run within the Forest Service computing environment (NRIS FSVeg, 2005), and it provides agency personnel with the information needed to respond to public concerns and to address complex issues. Basically, it provides a diverse range of basic and calculated information in standard formats that can be shared across administrative and ecological boundaries.

Table 2. Forest Type Classification in Operations Inventory [a]

Cover Type	Description	Short Description
A	ASPEN (UPLAND)	Aspen
B	PAPER BIRCH	Paper Birch
C	CEDAR	Cedar
E	SWAMP HARDWOODS	Swamp Hrdwds
F	SPRUCE-FIR (UPLANDS-INCLUDING UPLAND BLACK SPRUCE)	Spruce Fir
H	HEMLOCK	Hemlock
I	LOCAL USE	Local Name
J	JACK PINE	Jack Pine
M	NORTHERN HARDWOOD	Upland Hdwds
O	OAK	Oak
P	BALSAM POPLAR & SWAMP ASPEN and SWAMP WHITE BIRCH	Lowlnd Poplr
Q	MIXED SWAMP CONIFER	Mx Swmp Cnfr
R	RED PINE	Red Pine
S	BLACK SPRUCE-SWAMP	Black Spruce
T	TAMARACK	Tamarack
W	WHITE PINE	White Pine

Source: Operations Inventory Manual, MDNR (Unpublished)
[a] Non-forest classes include Tree bog, Grass, Rock, Lowland brush, Marsh, Upland brush, Bog or Muskeg, Other non-stocked or non-forest or non-productive land, Sand dunes, and Water.

Field Sampled Vegetation (FSVeg), one of the resource modules of NRIS, provides guidelines for National Forest management planning. This component includes point and plot vegetation data from field surveys. Data on trees, surface cover, understory vegetation and down woody material are included in this component. The USDA-FS defines Forest land as "[l]and at least 10 percent stocked by forest trees of any size, including land that formerly had such trees cover and that will be naturally or artificially regenerated. Forest land includes transition zones, such as areas between heavily forested and nonforested lands that area at least 10 percent stocked with forest trees and forest areas adjacent to urban and built-up lands. The minimum area of classification for forest land is 1 acre. Roadside, streamside and shelterbelt strips of trees must have a crown width of at least 120 ft to qualify as forest land. Unimproved roads and trails, streams, and clearings in forest areas are classified as forest if less than 120 ft wide" (Smith et al. 2004).

NRIS conducts ongoing strategic inventories, tactical inventories and stand examinations. The strategic inventory scale is national, the tactical inventory scale is for the Region or Forest, and the stand examination scale is for a project area, stand or vegetation condition. Basically, the stand examinations are one "type" of inventory conducted by the USDA-FS. Its purpose is to obtain the site and setting characteristics required to identify stand conditions and capabilities. The information may be collected by simple observations, or by formal, intensive examinations. The appropriate method to be chosen depends on factors such as stand complexity, the decisions to be made, and the purpose of the exam. Each examination method has varying data and accuracy requirements. However, stand examination data provide information for a wide variety of uses ranging from determining silvicultural treatments to evaluating wildlife habitat and modeling water yields.

Integration of the stand level information, based on inventory design, produces national forest planning information. Without knowing the compatibility of scope, scale, and objective, there is a risk of introducing bias into a combined data set (NRIS FSVeg 2004). However, the design considerations and data acquisition guidelines prepared by NRIS are conceptual for national forest management planning rather than directional.

The Ottawa, Hiawatha and Huron-Manistee national forests are located in Michigan. The forest cover classification used in the NRIS-FSVeg database of the national forest includes 60 forest cover types and has separate codes for each class (Table 3).

Table 3. Forest Types Used by Michigan National Forests

Code	Forest type	Code	Forest type
1	Jack Pine	55	Northern Red Oak
2	Red Pine	57	Scarlet Oak
3	White Pine	59	Mixed Oak
4	White Pine-Hemlock	60	Oak-Hardwoods
5	Hemlock	63	Northern pin Oak
6	Scotch Pine	70	Sugar Maple-Black cheery
7	Norway Spruce	71	BI Ash-Elm-R Maple
8	White Spruce	74	White Ash
9	Conifers	76	Red Maple(Wet)
10	Spruce	79	Mixed Lowland Hdwd
11	Balsam Fir-Asp-PB	80	Sugar-Maple-northern red oak
12	Black Spruce	81	S Maple-Beech -YB
14	Northern Wh Cedar	82	S Maple-Basswood
15	Tamarack	83	BI CH-W Ash-Y Pop
16	Wh-Sp-BF	84	Red Maple(Dry)
17	Upland BI Spruce	85	Sugar Maple
18	Mix Swamp conifer	86	Beech
19	Cedar-Aspen-PB	87	Sugar maple-beech/yellow birch/red spruce
20	Northern Hdwd-Heml		
21	Mixed Northern Hdw	88	Black Locust
22	N. White Cedar- UP	89	Mixed Upland Hdwd
23	W. Spruce-BF-Aspen	90	Sugar maple-beech/ basswood
24	Balsam Fir	91	Quaking Aspen
41	Wh Pine NRO-W Ash	92	Paper Birch
42	E. Red cedar Hardwood	93	Bigtooth Aspen
43	Oak-Eastern white pine	94	Balsam Poplar
47	Oak-Aspen	95	Asp-W Spruce BF
48	Jack Pine-Oak	97	Lowland Brush
49	Red Pine-Oak	98	Upland Brush
53	Black Oak	99	Open
54	White Oak		

Source: NRIS FSVeg Data Dictionary Version 1.7 January 2005, Appendix (pp E-6 and E-8)

Forest Inventory Analysis (FIA)

The primary objective of FIA is to determine the extent, condition, volume growth, and depletions of timber on the Nation's forestland (Miles et al. 2001). The FIA's continuing endeavor is mandated by Congress in the McSweeney-McNary Forest Research Act of 1928 and the Forest and Rangeland Renewable Resources Planning Act of 1974. Further, the 1998 Farm Bill required FIA to collect data on 20 percent of plots annually within each State (Miles et al. 2001).

FIA defines forest areas as "[l]and at least 16.7 percent stocked by forest trees of any size, or formerly having had such tree cover, and not currently developed for nonforest use. The minimum area for classification of forest land is one acre. Roadside, streamside, and shelterbelt strips of timber must have a crown width of at least 120 feet to qualify as forest land. Unimproved roads and trails, streams or other bodies of water or clearings in forest areas shall be classified as forest if less than 120 feet wide" (Hahn and Hansen 1985).

"The FIA inventory is based on aerial photo and/or remote sensing activity used to characterize the acreage of forest and non-forest land in the US. These classes are based on land-use. For forested land, more detailed classes are sometimes defined based on criteria such as forest type, volume per acre, stand size, stand density, ownership and/or stand age" (Miles et al. 2001). Then, ground plots are measured to adjust the remote sensing sample for changes since its data acquisition and to correct any misclassification. "The remote sensing classification of these ground plots, together with the area estimates from the remote sensing sample is used to assign area expansion factors to all ground plots" (Miles et al. 2001). These area expansion factors are used to weigh plot-level estimates when computing estimates for selected strata of the population.

FIA plots are designed to cover a 1-acre sample area; however, not all trees on the area are measured. Recent inventories use a national standard, fixed-radius plot layout for sample tree selection. Ground plots may be new plots that have never been measured during a previous inventory. For all plots several observations are recorded for each sample tree, including its diameter, species and other measurements that enable the prediction of the tree's volume, growth rate, quality, and forest health data. These tree measurements form the basis of the data on the tree records in the FIA database. According to the sixth FIA inventory online information, there were 45 forest types reported as classified by field crew for the public forest land in Michigan (Table 4).

A variety of tools like maps, aerial photographs/imagery, and Global Positioning System (GPS) units are utilized to properly install the ground plots (Burkman 2005). Once a ground plot location has been selected on an aerial photograph, it is established and measured in the field. On all forested field plots, quantitative and qualitative measurements are made for conditions such as tree diameter, length, damage, amount of rotten or missing wood and tree quality, tree regeneration, site quality information, stocking, and general land use. And the general stand characteristics are gathered for forest type, stand age and disturbance, change in land use, general stand characteristics and estimates of growth, mortalities, and removals are gathered (Burkman 2005).

Table 4. FIA Forest Type Classification

Code	Forest Type	Code	Forest Type
101	Jack pine	515	Chestnut oak / black oak / scarlet oak
102	Red pine	519	Red maple / oak
103	Eastern white pine	520	Mixed upland hardwoods
104	White pine / hemlock	700	Elm / Ash / Cottonwood Group
105	Eastern hemlock	701	Black ash / American elm / red maple
121	Balsam fir	703	Cottonwood
122	White spruce	704	Willow
125	Black spruce	706	Sugarberry / hackberry / elm / green ash
126	Tamarack	708	Red maple / lowland
127	Northern white-cedar	709	Cottonwood / willow
380	Exotic Softwoods Group	800	Maple / Beech / Birch Group
381	Scotch pine	801	Sugar maple / beech / yellow birch
383	Other exotic softwoods	802	Black cherry
400	Oak Pine Group	803	Cherry / ash / yellow-poplar
401	White pine / Red oak / white ash	805	Hard maple / basswood
409	Other pine / hardwood	807	Elm / ash / locust
500	Oak Hickory Group	809	Red maple / upland
503	White oak / red oak / hickory	900	Aspen / Birch Group
504	White oak	901	Aspen
505	Northern red oak	902	Paper birch
507	Sassafras / persimmon	904	Balsam poplar
509	Bur oak	999	Non-Stocked
513	Black locust		

Problem Statement

When there are multiple sources of forest resource information many users of these information sources try to make comparisons to meet their policy, planning and management needs. However, the adoption of different definitions for forest land and forest types makes it difficult to infer similar conclusions in some cases. For example, FIA and OI define forest area as land at least 16.7 percent stocked by forest trees of any size or formerly having had such tree cover, and not currently developed for non forest use. IFMAP defines forest as land with the proportion of crown cover exceeding 25% of the land area (MDNR 2003d). The difference has resulted in classification of a number of non-forest types that are not delineated in FIA (Pedersen 2002). The OI and IFMAP have non-forest types like Treed bog, Grass, Rock, Lowland/Upland Brush and Marsh. The other distinction between the IFMAP and NRIS-FSVeg/OI is that the former will identify a recent clear cut as the non-forest while the latter will consider it as forest. This is problematic when only a single date of imagery from all medium- to coarse-resolution sensors in a landscape where agriculture and forestry are interwoven (Wynne et al. 2000).

IFMAP has classified forest land into the least number of forest types followed by OI which has 16 categories. NRIS-FSVeg is classified into 60 types, and FIA has 45 forest types in public forest land in Michigan. FIA plot data can be viewed as the most precise classification scheme given rigorous data collection procedures at the plot level. Other schemes are used, but their relationship to FIA classifications is largely unknown. There are no studies which compare the IFMAP, OI and NRIS-FSVeg classifications with FIA classifications in Michigan. Taking FIA plot location as the point of reference for comparisons, I will provide information on the similarity and dissimilarity between forest classification approaches. The comparison will be helpful in explaining discrepancies in total public land acreage of different forest types computed from these data bases. Results from a comparative forest classification study may be useful developing recommendations to harmonize the available databases.

Study purpose

The purpose of this research is to compare classifications of forest types between forest resource databases for public lands in Michigan. The general research questions are:

"What are the classifications made by the four data sets? And how consistent are the four classification systems?"

This study addresses the following objectives:

- To compare and contrast the four forest resources inventory/database designs in forest type classification (FIA, IFMAP, OI and NRIS-FSVeg),
- To examine consistency among FIA, NRIS-FSVeg, and IFMAP classifications on national forest lands,
- To examine consistency among FIA, OI, and IFMAP classifications on state forest lands,
- To compare the agreement of IFMAP and FIA classifications on national and state forest lands, and
- To explain differences in classifications using plot and stand level characteristics for the state forest lands.

Organization of Thesis

The thesis has five chapters. A synopsis of each chapter is presented in this section. Chapter 2, Literature Review, examines previous studies on forest classification compared against ground truth plots. Previous studies comparing the accuracy of classification of satellite imagery and forest classifications are reviewed. Studies particularly relevant to Michigan land use are also reviewed.

Chapter 3, Classification Approaches, Data Sources, Study Area and Analysis Methods, presents background information on forest classifications for FIA, NRIS-FSVeg, OI, and IFMAP. Details on data sources regarding spatial locations and data attributes are summarized. Methods used to assess accuracy of the various information sources with reference to FIA sample plot information are described. Specifically, construction of error matrices, errors of omission, errors of commission, and Kappa statistics are the major tools used to classify the accuracy of the databases.

Chapter 4, Results and Discussion, presents analysis of the error matrices and IFMAP, OI and NRIS-FSVeg accuracy in classifying forest types on the publicly-owned forestland. Major findings, strengths and weaknesses of the comparisons and data limitations are discussed.

In Chapter 5, Summary and Conclusions, policy, planning and management implications and several additional issues on forest classification are discussed. Among them are the definition of the forest area and inclusion of the non-forested land in classification systems. Also, policy recommendations to reconcile the forest classification in public forest land in Michigan are discussed. Additional research ideas are presented.

CHAPTER 2

LITERATURE REVIEW

This study compares the forest type classification among the forest management planning databases of national forest, state-owned forest and the coarse-resolution state map taking the exact FIA plot location as the reference for comparison. The previous chapter introduced major public forest management databases in Michigan. This chapter reviews the land use classification systems broadly, the Michigan Resource Inventory System (MIRIS), the National Land Cover Database (NLCD) for the United States, forest classification of FIA/OI/ NRIS-FSVeg, FIA sampling, forest classification of IFMAP, digital imagery classification, IFMAP imagery classification, Remote Sensing (RS) based forest cover type studies, and studies related to accuracy measurement of RS imagery classification.

Land Use Classification

Existing use of land is the main characteristic for defining land use. Anderson et al. (1976) pioneered land use classification from remote sensing information and outlined criteria to effectively utilize the remote sensing information for land use and land cover classification. The land use classification developed by Anderson et al. (1976) has level I and level II classifications (Table 5). Information at levels I and II are a basis for national level or statewide aggregation. In addition, more detailed land use and land cover data, categorized at level III and IV, will be used more frequently by those who need and generate location information at the intrastate, regional, county, or municipality level (Anderson et al. 1976). Further, land use and land cover classification level V can be added if a finer level of classification is desired.

Table 5. Anderson et al. (1976) Land Use Classification

Level I	Level II
1 Urban or Built-up Land	11 Residential.
	12 Commercial and Services.
	13 Industrial.
	14 Transportation, Communications, and Utilities.
	15 Industrial and Commercial Complexes.
	16 Mixed Urban or Built-up Land.
	17 Other Urban or Built-up Land.
2 Agricultural Land	21 Cropland and Pasture.
	22 Orchards, Groves, Vineyards, Nurseries, and Ornamental Horticultural Areas.
	23 Confined Feeding Operations.
	24 Other Agricultural Land.
3 Rangeland	31 Herbaceous Rangeland.
	32 Shrub and Brush Rangeland.
	33 Mixed Rangeland.
4 Forest Land	41 Deciduous Forest Land.
	42 Evergreen Forest Land.
	43 Mixed Forest Land.
5 Water	51 Streams and Canals.
	52 Lakes.
	53 Reservoirs.
	54 Bays and Estuaries.
6 Wetland	61 Forested Wetland.
	62 Nonforested Wetland.
7 Barren Land	71 Dry Salt Flats.
	72 Beaches.
	73 Sandy Areas other than Beaches.
	74 Bare Exposed Rock.
	75 Strip Mines, Quarries, and Grave Pits.
	76 Transitional Areas.
	77 Mixed Barren Land.
8 Tundra	81 Shrub and Brush Tundra.
	82 Herbaceous Tundra.
	83 Bare Ground Tundra.
	84 Wet Tundra.
	85 Mixed Tundra.
9 Perennial Snow or Ice	91 Perennial Snowfields.
	92 Glaciers.

Source: Anderson et al. (1976), Public Database Warehouse, USGS (http://www.wsdot.wa.gov/environment/envinfo/docs/RSPrj_USGS_lulcclass.pdf)

However "[t]he Level II is the fulcrum of the classification system as Level II can be created by aggregating the similar Level III categories classification system" (Anderson et al. 1976).

The U. S. Geological Survey (USGS) classification system provides flexibility in developing categorization at the more detailed levels and provides freedom to the users to develop categories that meet their particular needs. To retain the compatibility of the information, whatever categories are used at the various classification levels, special attention should be given to providing potential users of the data with sufficient information so that they may either compile the data into more generalized levels or aggregate more detailed data into the existing classes. Basically, the system satisfied the three major attributes process: (1) it gives names to categories by using accepted terminology, (2) it enables information to be transmitted, and (3) it allows inductive generalizations to be made (Anderson et al. 1976). This classification system is capable of further refinement on the basis of more extended and varied use.

Regarding the forest land classification, Anderson et al. (1976) stated that "Forest lands have a tree-crown areal density (crown closure percentage) of 10 percent or more, are stocked with trees capable of producing timber or other wood products, and exert an influence on the climate or water regime". Further, the authors noted when the "trees reach the marketable size"... they may be harvested and replanted and ... "there will be large areas that have little or no visible forest growth. The pattern can sometimes be identified by the presence of cutting operations in the midst of a large expanse of forest". Such areas should be included in the forest land category, "[u]nless there is evidence of other use". And, "[l]ands that meet the requirements for forest land and also for an urban or built-up category should be placed in the latter category. The only exceptions in classifying forest land are those areas which would otherwise be classified as wetland if not for the forest cover. Since the wet condition is of much interest to land managers and planning groups and is so important as an environmental surrogate and control, such lands are classified as forested wetlands" (Anderson et al. 1976).

At Level II, forest land is divided into three categories: Deciduous, Evergreen, and Mixed. According to the Anderson et al. (1976) classification, Deciduous Forest Land includes all forested areas having a predominance of trees that lose their leaves at the end of the frost-free season or at the beginning of a dry season. In most parts of the United States, these would be hardwoods such as oak (*Quercus*), maple (*Acer*), or hickory (*Carya)* and "soft" hardwoods, such as aspen *(Populus tremuloides).* Deciduous forest types characteristic of wetland, such as tupelo (*Nyssa*) or cottonwood (*Populus*), also are not included in this category. Evergreen Forest Land included all afforested areas in which the trees are predominantly those which remain green throughout the year. Both coniferous and broadleaf trees are included in this category. The coniferous evergreens are commonly referred to or classified as softwoods. They include eastern species such as balsam fir (*Abies balsamea*), various spruces (*Picea*), white pine (*Pinus strobus*), red pine (*Pinus resinosa*), jack pine (*Pinus banksiana*), and hemlock (*Tsuga canadensis*) (Anderson et al. 1976). Evergreen species commonly associated with wetlands, such as tamarack (*Larix laricina*) and black spruce (*Picea mariana*), are not included in this category.

Similarly, Anderson et al. (1976) identified Mixed Forest Land as all forested areas where both evergreen and deciduous trees are growing and neither predominates. When more than one-third intermixture of either evergreen or deciduous species occurs in a specific area, it is classified as mixed forest land. Where the intermixed land use or uses total less than one-third of the specified area, the category appropriate to the dominant type of forest land is applied, whether deciduous or evergreen. Further to this, Forested Wetlands are wetlands dominated by woody vegetation. Forested wetland includes seasonally flooded bottomland hardwoods, mangrove swamps, shrub swamps, and wooded swamps including those around bogs (Anderson et al. 1976). As the forested wetlands can be detected and mapped by the use of seasonal (winter/summer) imagery, and delineation of Forested Wetlands is needed for many environmental planning activities. For these reasons, they are separated from other categories of forest land.

Though several classification approaches are available, the land-use and land-cover classification system devised for the USGS program, developed by Anderson et al. (1976), has become one of the most widely used classification systems for land use maps prepared by interpretation of remotely sensed images (Campbell 2002).

Michigan Resource Inventory System

The Michigan Resource Inventory System (MIRIS) has two types of inventories about the State's land and water resources: (1) a current use inventory to illustrate land cover and land use and (2) a land resource inventory which includes resources, unique areas, and areas hazardous to development (Goodwin et al. 2002). The first current use inventory was compiled from photo interpretation of color infrared aerial photography (1:24,000 scale to 1 inch to 2,000 ft) obtained in 1978/79. Aerial photography obtained in 1985 was used for the inventory of Detroit and seven highly urbanized countries in Southeast Michigan. The second Michigan Land Cover/ Use Classification System (Division of Land Resource Programs 1981) is similar to the national system developed by the U.S. Geological Survey (Anderson et al. 1976). It is multi-level, hierarchical system that classified Michigan's land cover/use into approximately 500 categories (Goodwin et al. 2002). The current use inventory is a subset (approximately 60 level I and II categories) of the Michigan land cover/use classification system (MIRIS, 1981). However, some of the category numbers and category definitions were changed.

The MIRIS was upgraded in 2000 by using the 1998/1999 National Aerial Photography Program (NAPP) imagery, flown by the USGS for the entire state of Michigan (RS and GIS, n. d.). These changes in MIRIS version II were made in order to correct problem areas that existed with the design and application of the earlier version. Changes in II version were based on several criteria. These criteria included:

(1) "describing the major components of each category group within the confines of a three-level hierarchy, (2) assigning map codes to categories in the lowest classification levels within each group, (3) creating separate categories for those items that may need to be cross-referenced under various user defined aggregation schemes and (4) maintaining clear distinction between "upland" and "wetland" natural cover types" (MIRIS 2000). In addition, the current use inventory of 22 sampled counties was supplemented with a detailed land component (level IV and V data which included species designation, stand size and stocking classification).

National Land Cover Database (NLCD) for the United States

In the last decade, a major provider of land cover information within the federal government has been the Multi-Resolution Land Characteristics Consortium (MRLC) (Homer et al. 2004). The MRLC was originally formed in 1993, to meet the needs of several federal agencies: United States Geological Survey (USGS), Environmental Protection Agency (EPA), National Oceanic and Atmospheric Administration (NOAA), and U.S. Forest Service (USFS) for Landsat 5 imagery and land cover information (Loveland and Shaw 1996). For NLCD 2001 land cover classification, a method that optimally classifies many database layers in a single step, with the ability to document this relationship in a rule base was highly desirable, and the decision tree classification method was chosen (Homer et al. 2004). The authors used the commercial decision tree program C5©. They claimed that this decision tree classification provided an efficient, robust method for classifying large quantities of information in documentable form, and moreover, it allowed them to export mutually exclusive rules generated by the classification into generic textual rule sets allowing users access to classification parameters. NLCD 2001 defined land cover into 29 land cover classes. And forest cover is defined as "areas dominated by trees generally greater than 5 meters tall, and greater than 20 percent of total vegetation (Homer et al. 2004). This database is developed using the Mapping Zone approach, with 66 Zones in the continental United States and 23 Zones in Alaska.

IFMAP Classification

IFMAP has a hierarchical scheme of classification of land use. The hierarchical scheme contains various levels from very broad categories to more detailed ones. Pacific Meridian Resources (PMR 2001a) stated that the objectives of the classification scheme of IFMAP are: (1) to provide useful land cover labels for forest and wildlife management, (2) to provide suitable strata to support stratified inventory data, and (3) to generate land cover information for use by managers and researchers. In addition, the IFMAP classification kept "the number of classes of land use to a minimum, and ensure that classes agree with the definitions of the other ecoregions" (PMR 2001a). The rule established for this classification system was that each class should have a local management objective. The vegetation classification rules were based on percentage of the ground covered by specific vegetation covers. Vegetation was broken into woody and non-woody with the definition of woody being a plant that contains a secondary xylem (dicotyledons and conifers), then into shrub and tree. The cut off point between forest and non-forest, and vegetated and non-vegetated was set as 25% of ground covered by canopy.

The IFMAP classification is very close to the Gap Analysis natural terrestrial cover classification system. In fact, it is a modified UNESCO natural terrestrial cover classification that has six hierarchical levels namely: forests, woodlands, shrublands, dwarf-shrublands, grasslands, and barren (Jennings 1993). The Gap Analysis classification defined forest as areas dominated with a total canopy cover of 61% or more, trees crown usually interlocking; woodlands as areas dominated by trees with a total canopy cover of 26 to 60%, most of the trees not touching each other; shrublands, dwarf–shrublands and grassland as areas with less than 26% total canopy of tree; and barren land as areas with vegetation cover less than 5% (Jennings 1993). At level II, the Gap land use classification combined the land with morphologically similar main vegetation into a class. For the classes of forests, woodlands, shrublands and dwarf-shrublands, the similarities were based on evergreen, deciduous and xeromorphic characteristics (Jennings 1993). The IFMAP level II classification has adopted the 60% rule of canopy cover between coniferous and deciduous so that only those stands that were neither dominated by coniferous nor deciduous fell into mixed stands (PMR 2001a). The Level III classification of IFMAP was developed based on the majority of a particular tree species in a forest stand. And the classification scheme has been viewed as a series of sequential if-then statements. The detail of the IFMAP classification can be accessed in the Michigan Geographical Data Library website (http://www.mcgi.state.mi.us/mgdl/)

IFMAP uses imagery obtained from the TM 5 and ETM 7+. To understand the IFMAP imagery classification it is important to review the remote sensing techniques of imagery classification, after first describing on-the-ground management and research databases like FIA, OI, and NRIS-FSVeg.

Forest Classification of FIA/OI/NRIS-FSVeg

FIA and the forest management planning operational databases, OI and NRIS-FSVeg, provide more detailed forest classification (See Chapter 1). These databases have Level IV/V classification. IFMAP has information up to Level III land use classification (SI 2004). According to the FIA National Core Field Guide Version 2.0 (2004), the forest types of the Continental U.S. and Alaska have been classified into 28 forest groups, and there are 140 types that best describe the plurality of stocking for all live trees. Similarly, based on online FIA condition table information, the public forest of Michigan has 14 forest groups and 45 distinct forest types (Table 4). FSVeg has classified the national forests of Michigan into 60 forest types, including lowland brush, upland brush and open as a type (Table 3). OI has classified state forests into 16 forest types (Table 2). OI has defined treed bog, grass, rock, lowland brush, upland brush, marsh, bog, musk, sand dunes, water and other non-stocked or non-forest or non-productive land as non-forest classes.

There are differences in the minimum percentage of stocking required to define an area as forest among various databases. The minimum level of forest tree stocking requirement is 10% for FSVeg (Smith et al. 2004), 16.7% for FIA (Hahn and Hansen 1985) and for OI (Pedersen, per. comm..), 25% for IFMAP (Michigan Geographical Data Library, SI, 2004). Thus there are big differences in inclusion and non-inclusion of areas with lower levels of forest tree stocking. For example, the NRIS-FSVeg has forest type as lowland brush and upland brush and OI and IFMAP has defined the lowland brush, upland brush, treed bog, as non-forest areas.

FIA Sampling

For this study, FIA information is serving as the ground truth for assessing the accuracy of forest type classifications of the databases. Therefore, it is important to understand the nature of FIA. With passage of the 1998 Farm Bill, formerly known as the Agricultural Research, Extension, and Education Research Act of 1998 (PL 105-185), Congress required that the Forest Service conduct annual forest inventories in all states (McRoberts 1999). This Bill made some changes in FIA sampling procedures and intensity. For example, the Farm Bill established requirements that (1) each year, 20 percent of total plots are to be measured in each eastern state and 10 percent of plots are to be measured in each western state, (2) the annual data are to be made available each year, and (3) statewide resource reports are to be published every five years with integration of FIA and the Forest Health Monitoring (FHM) program (McRoberts 1999, Brand 2005). FHM is a national program that uses data from ground plots, aerial surveys, and other sources to produce annual estimates of status changes and trends in indicators of health.

FIA inventories are commonly designed to meet the specified sampling errors at the State level at the 67 percent confidence limit (Miles et al. 2001). FIA precision standards require sampling intensity of one plot for approximately every 6,000 acres in the North Central region (McRoberts 1999). To satisfy this requirement, the geographical hexagons established for the FHM programs were divided into 27 smaller FIA hexagons, each of which contains approximately 5,900 acres. A grid of field plots was established by selecting or establishing a plot in each smaller hexagon: (1) if an FHM plot fell within a hexagon, it was selected as the grid plot; (2) if no FHM plot fell within a hexagon, the plot from existing network of permanent FIA plots that was nearest the hexagon center was selected as the grid plot; and (3) if neither the FHM nor the existing FIA plot fell within the hexagon a new permanent FIA plot was established at the hexagon center and selected as the grid plot (McRoberts 1999, Brand 2005). The grid of plots is called the federal base sample and is considered as an equal probability sample. In this way, FIA uses grid sampling (Schreuder et al. 2003) that covers a 1-acre sample area (Miles et al. 2001). Recent inventories use a common sampling design consisting of four 24.0-foot radius subplots (approximately (1/24th acre) for trees at least 5 inches in diameter and four 6.8-foot radius microplots (1/300th acre) for smaller trees (Miles et al. 2001). Another characteristic of the new design is the mapping of differing forest conditions. If two or more conditions occur within a plot, the boundary between them is mapped and the proportion of the plot in each condition is recorded.

Forested plots are installed and measured regardless of intended use or any restrictive management policy. After the adoption of the national plot design in the mid 1990's, all FIA units have implemented a common sampling design consisting of four 24.0-foot radius subplots (approximately 1/24th acre) for trees at least 5 inches in diameter and four 6.8-foot radius microplots (approximately 1/300th acre) for smaller trees (Miles et al. 2001). In this way, tree expansion factors are approximately 6 for trees at least 5 inches in diameter and approximately 75 for smaller trees. Subplot 1 is the center of the cluster with the other three subplots located 120 ft away at azimuths of 360^0, 120^0, and 240^0, respectively (Miles et al. 2001). In addition, the temporal regularity was incorporated by systematically assigning each hexagon to one of 5 inter-penetrating panels. Plots located in panels 1 to 5 hexagons were to be measured in first to fifth years, respectively. Once the five- or ten-year cycle is complete, the sequence starts again. In fact, FIA inventories are extensive inventories that provide reliable estimates for large sampling areas. As data are subdivided into smaller and smaller areas, such as geographic unit or a county, the sampling errors increase and the reliability of the estimates goes down. There are nine tables in the FIADB (Forest Inventory and Analysis Database) Version 1.0. In this study Condition table and Plot table information is used.

Digital Imagery Classification

Digital image classification is the process of assigning pixels to classes (Campbell 2002). The aim of the classification is to assign each pixel in an image to a distinct cover class or theme, broadly called an "information class" (Foody 2003). By comparing pixels to one another, and to pixels of known identity, it is possible to assemble groups of similar pixels into classes that are associated with the informational categories of interest to users of remotely sensed data (Campbell 2002). The informational classes are the categories of interest to the users of the data. Informational classes are, for example, different kinds of forests or land uses that convey information to the remotely sensed data users (e.g. policy planners, resource managers, the scientific community and publics).

Unfortunately, the information classes are not recorded directly on remote sensing images; they can only be derived indirectly by using evidence contained in the brightnesses recorded by each image. A group of pixels that are uniform with respect to the brightnesses in their several spectral channels is called a spectral class (Campbell 2002). Thus, remote sensing classification proceeds by matching spectral categories to information categories. In general, each pixel is assigned into the class that it most nearly resembles or to which it is closest, according to some measure of distance in a broad sense. In this way, information classes are typically composed of numerous spectral subclasses.

Unsupervised classification is defined as the identification of natural groups, or structures, within spectral data (Campbell 2002). In unsupervised classification these natural groups are defined, identified, labeled and mapped. On the other hand, Campbell (2002) informally defines supervised classification as the process of using samples of known identity (i.e., pixels already assigned to informational classes) to classify pixels of unknown identity (i.e., to assign unclassified pixels to one of several informational classes). Samples of known identity are obtained from those pixels located within training areas, or training fields. Pixels located within these areas from the training samples are used to guide the classification algorithm to assign specific spectral values to appropriate informational classes (Campbell 2002). However, there is no unique criterion to use as the basis for classification; different criteria yield different classifications resulting in a trade-off between accuracy and efforts (Foody 2003).

Chuvieco and Congalton (1988) developed a hybrid method which combines the training statistics generated from both supervised and unsupervised classification approaches using two multivariate statistical techniques: cluster and discriminant analysis. This method was used to classify the pixels for IFMAP. In this method, the cluster analysis is used to group together training statistics generated from both classification approaches. The strength of these groupings was then tested using discriminant analysis. "The cluster analysis is not simply a reduction process for the unsupervised approach, but rather a way of combining similar groupings from both the supervised and unsupervised approach and the groupings that contain both supervised and unsupervised statistics provide a powerful match between the informational categories and the spectral classes" (Chuveico and Congalton 1988).

The cluster analysis proposed by the authors uses the hierarchical method in which the classes are merged progressively, two in each step, until all classes belong to the same cluster or group. This grouping can result either in a merging of two single classes, or in a class being merged with an already formed grouping, or in merging two groupings. At the end of the process, the user selects the step at which the clustering stops, usually when two classes very remote in the distance matrix are merged. The squared Euclidian distance is selected to calculate the distance between the classes. Once the cluster analysis of the training statistics is completed, discriminant analysis is used to evaluate the strength of the grouping.

Discriminant analysis is a multivariate technique that attempts to find a new set of functions which maximize the ratio of the variance between and within groups (Chuvieco and Congalton 1988). The analysis involves a linear transformation of the original variables, which are orthogonal, in such a way that the new functions maximize the separation between the already formed groups. After these functions have been found, it is possible to regroup each one of the original training classes to test its membership in the correct grouping. After running a discriminant analysis on the groups formed by the clustering of the training statistics, each one of the final groups was defined by the average value of its members. Then these average statistics were input into the classification algorithm to be used in the assignment phase of the classification process. In this way, by the clustering and discriminant analysis, "an improvement in the classification results is expected because of the improved grouping of training statistics" (Chuvieco and Congalton 1988).

In IFMAP, classification schemes were first developed for the Southern Lower Peninsula in the spring of 2000, and then updated to reflect the differing ecoregions of the Northern Lower Peninsula and the Upper Peninsula in the winter and spring of 2001 (PMR 2001a). Imagery were acquired for different seasons namely, spring (leaf-off imagery), summer (growing season) and senescence imagery (fall images). The images were used to enable separation of species based on their phenological differences, that is, different cropping cycles and moisture regimes that allowed differentiation of these categories (e.g., aspen from northern hardwoods and from oaks). A listing of the classification schemes and the rules developed to derive the classes can be accessed via the Michigan Spatial Data Library website (http://www.dnr.state.mi.us/spatialdatalibrary/sdl2/land_use_cover/2001/IFMAP_lp_landcover.htm).

The acquisition of imagery from Landsat 5 and Landsat 7 sensors for the rest of the state of Michigan took place in Fall 2001. In 2002, imageries from both sensors were collected in leaf-off, mid-season, and senescence for the Northern Lower Peninsula and the Upper Peninsula. A stratified approach was adopted to obtaining the training data. The training data were collected comprising reflectance samples, taken from the images to be classified of the cover types to be classified. These reflectances may differ from region to region, because of phenologic differences, soil variability, hydrologic differences, and elevation changes, though the latter was not a big factor in the State of Michigan (PMR 2001a). To obtain an adequate stratification of training samples, a coverage of ecosystems was intersected with a coverage of the TM scene boundaries, and the resulting polygons were named eco-scenes (PMR 2001a). Eco-scenes of interest were identified and specific areas were chosen for field visits. Since the IFMAP project is specifically concerned with gaining information on forested lands, efforts were made to identify areas of high natural land cover diversity.

Training fields are areas of known identity delineated on the digital image, usually specifying corner points of a square or rectangular area using line or column numbers within the coordinate system of the digital image. Specific training areas need to be identified for each informational class following the standard guideline of image classification (Campbell 2002). The key characteristics of training areas are number of pixels, size, location, number, placement, and uniformity (Campbell 2002). According to PMR (2001a), while performing the training of IFMAP imagery, the raw information, such as the canopy closure of the three size classes, various shrub, wetland, and herbaceous species, were recorded from training sites.

Field data were digitized onto the imagery using the air photographs and the manual delineations (on the mylar overlays) as guides. The supervised training sites were extracted from this image using a region-growing method. The region growing method ensured a spectrally pure signature, which was appropriate for the supervised classification. It involved specifying a spectral threshold for inclusion in the area of interest (AOI), which eliminated spectrally outlying portions of the sites. After digitizing, the training sites were used to recode a single band of a TM mosaic, with unique identifiers for each training sample. This layer was used to extract training statistics from any of the imagery mosaics and/or derivatives without re-drawing the training sites. As the atmosphere can substantially impact the signatures derived from imagery, scenes were selected that had a minimum of haze and clouds and no atmospheric correction was made. Images of Southern Michigan were divided into areas with similar date scenes, usually taken on the same date. They were mosaiced, and areas that represented different conditions on the ground were classified separately (PMR 2001a).

The primary classifier was a cluster analysis method developed by Chuveico and Congalton (1988), which matched clusters from an unsupervised classification with the training site data collected by the field crews. In the first iteration, the summer image mosaic is classified using the ISODATA unsupervised method (PMR 2001a). The resulting classes were examined with the supervised classes using an agglomerative hierarchical procedure by the application of Mathsoft 2000 (PMR 2001a). This procedure begins by considering each signature, the pixel reflectance value, as a separate group; then it combines and divides groups based on spectral similarity until all signatures are in a single group, displayed in a hierarchical structure according to the order in which the groups were merged or divided. The resulting clustering tree is then examined for matches of supervised and unsupervised signatures, and in the case of a close one-to-one relationship, the unsupervised signature in question can be labeled with the land cover label of the supervised signature. After the tree is fully examined for clusters of this type, the labeled classes are subset from the imagery, and the procedure is run again. This procedure continues until it is no longer advantageous to do so.

The cluster analysis method was used initially on the entire image to achieve an Anderson Level I classification, and subsequently on the individual level I cover types to classify the land cover to the desired level. Once the imagery was subset to level I, the signatures derived from the individual subsets will contain less overall spectral diversity, and as a result the subtle differences were more evident and therefore easier to separate. Level II classifications were based on analysis with leaf off imagery and Level III classification were derived using different masks and use of texture bands.

Remote Sensing-Based Cover Type Mapping Studies

Ross Nelson and his colleagues (1987) developed a method to assess the continental forest land cover using Landsat Multispectral Scanner (MSS) data. In their study, the authors assessed the Anderson et al. (1976) level II classification for conifer and hardwood forest area in the continental United States. The authors used a stratified random sampling (SRS) approach to allocate Landsat MSS scenes country-wide so that the aerial extent of the conifer/hardwood resources of the United States could be evaluated (Nelson et al. 1987). The authors generalized *a priori* information in the form of a major forest cover type map (USGS 1980) and used it to stratify the country into six forest strata, namely northern conifer, northern hardwood, central hardwood, southern conifer, western conifer, pinyon-juniper and one non-forest strata.

The MSS scenes were allocated to each forest stratum with the constraint that each contain at least two MSS scenes (Nelson et al. 1987). Within each MSS scene, four 200 X 200 pixel sample blocks were located systematically, and each block's forest cover was identified. The authors' selection of fewer large blocks was derived from their preliminary study. The sample blocks were first located in the Landsat MSS data and on 1:250,000 USGS topographic map sheets and to facilitate the acquisition of National High Altitude Photography (NHAP). NHAP are color infrared photos flown at a scale of 1:58,000. These photos were used to help identify spectral classes in MSS data and to refine the digital land cover classifications. The authors used most of the photos which were acquired leaves off, facilitating conifer/hardwood identification. A multicluster blocks procedure was used to classify the MSS data. In this method, for a given scene, small blocks of MSS digital data ranging from 40 X 40 to 60 X 60 pixels were clustered into spectral classes using an unsupervised classifier using IDIMS (Interactive Digital Image Manipulation System, ESL 1978). Each spectral class was identified using ancillary data (NHAP, B&W photo quad, false CIR composites) (Nelson et al. 1987). Forest was defined as any spectral class which included forested area >30% canopy closure. Conifer or hardwood forest was defined as a forested area where > 50% of the canopy was coniferous or hardwood, respectively.

The Landsat MSS data products and the statistical estimates of the conifer, hardwood, and water resources of the continental United States were evaluated to determine their reliability. The authors used the two different assessments to characterize the accuracy of this MSS classification. The first assessment compared the MSS classification and airphotos on a point by point (pixels by pixel) basis. The second assessment compared aerial estimates of the four cover types derived from the MSS data and from the corresponding areas on the airphotos. A third assessment was carried out to determine the accuracy of the MSS-based national estimates of conifer, hardwood, and water compared with national estimates generated by other U.S. government agencies. The findings of this study revealed that the national estimates of conifer and hardwood derived using this sampling method is within 3% of the total USDA Forest Service acreage. Comparison of the MSS classification products and airphotos showed that the conifer cover class was correctly identified 74% of the time and hardwood 80% of the time (Nelson et al. 1987). The average classification accuracy countrywide for the four types considered (conifer, hardwood, water and "other") is 74%, the overall accuracy is 85%.

Pax-Lenney et al. (2001) developed a generalized classifier method to monitor temperate conifer forests, ultimately at the global scale with Landsat TM and ETM+ data. Within this context, the generalization refers to a concept in which a classifier is trained with data from one domain, but applied to data from different domains (e.g., different geographic location, time, and /or imaging sensors). The authors mentioned that analytical methods based on image-by-image interpretation are too timeconsuming and labor intensive for studies of large areas to be undertaken with any degree of frequency. The authors found generalization is well suited for multitemporal classifications of one Landsat scene using simple dark-object-subtraction (DOS) atmospheric corrections to produce classifications with comparable accuracies as classifications from the more complex radiative transfer corrections, based on over 200 classifications. However, the high degree of variability in the classification accuracies underscores the importance of extensive, in-depth analysis of remote sensing techniques and applications, and highlights the potential problem for misleading results based on just a few tests (Pax-Lenney et al. 2001).

Miguel-Ayanz and Biging (1997) compared the performance of TM and SPOT data for cover type mapping on the Central Sierra of Spain. The authors used three single-stage and one multistage iterative classification in their study. The three-single stage classifications were: (1) supervised classification with band selection by spectral analysis, (2) supervised classification with band selection with spectral separability indices, and (3) supervised classification with prior probabilities and band selection by spectral separability indices. The multi-stage supervised classification was performed using an iterative classification method. In this method the class that attains the highest average accuracy (Congalton 1991) in each iteration is masked and classified as part of the image. In this way, GIS analysis was used to obtain the prior probabilities for classes being discriminated and these probabilities were used in the band selection and classification processes. However, prior probabilities did not significantly improve the band selection process. In most cases the best band combinations were selected both by the weighted (including the prior probabilities) and nonweighted separability indices. The best overall accuracy, 66%, was obtained for TM data with the iterative classification approach. Accuracy of 61% was the best overall accuracy for SPOT data, which was obtained with the iterative classification methods. For TM imagery, the five most abundant classes, which account for over 72% of the study area, where classified with 90% overall accuracy.

Similarly, Wang et al. (2003) compared the dry season ETM+ and 1-m panchromatic sharpened IKONOS imagery classified as tree canopies and open area taking the latter imagery as valid ground truth to assess the tropical deforestation in the in the Amazonian state of Mato Grosso, Brazil. The authors found the squared correlation coefficients (R^2) between the canopy cover values derived from ETM+ and IKONOS were 0.92 and 0.96 at the 30-m and 90-m scales respectively. Thus, the authors argued ETM+ imagery can be used to estimate canopy cover across large areas of tropical forest.

Silbernagel et al. (1997) compared the distribution of the landscape measure among landtype association groups in historic (1840's) and present (1990-1992) landscapes in the Eastern Upper Peninsula of Michigan comprising six counties: Alger, Chippewa, Delta, Luce, Mackinac and Schoolcraft. In addition, the authors compiled quantitative information on landscape metrics, to supplement existing qualitative descriptions of landtype associations (LTA) in the study area. Cover type boundaries between each section line were interpolated using elevation lines, surface geology maps, and other early vegetation maps. The authors used the cover classes based on expanded MIRIS land cover codes. Prevalence or dominance of cover classes was based on class area (CA) and landscape similarity index (LSIM), or class area weighted by total landscape area. "In the four physiographically based land type association groups studied: bedrock-controlled, lowland sand lake plain, morainal origin, and outwash-northern hardwoods and mixed conifer were most prevalent cover types of the 10 studied, historically and currently" (Silbernagel et al 1997). Northern hardwoods were especially prevalent in the moraine groups, while the mixed conifer type was more prevalent in the bedrock group. Wetlands and mixed pines, in addition to northern hardwoods, were also prominent in the lowland group. In the outwash group, these types were also present, but mixed and white pine were more prevalent. Largest single patches (LPI) were found redundant to the LSIM, and therefore were not assessed to the same extent as other indices. The highest LPI values were found in the northern hardwoods in the moraine groups.

Skole et al. (2002) developed a model "Forecast Michigan" to forecast the land use or urban sprawl in the context of spatial decision system support. The authors claimed, "the *Forecast Michigan* models are process models using most of the state's standard GIS data layers, as well as inputs from economic and demographic models." They also utilized network analysis algorithms to include transportation routing and traffic demand with enough spatial resolution and sensitivity to provide transportation planners a way to evaluate different corridor alignments and access points in the context of secondary and cumulative impacts on land use change in a project region.

There are a large number of studies about detailed forest classification, see for example, Saatchi and Rignot (1996), Mayaux and Lambin (1997), Martin et al.(1998). However, they are based on high spectral resolution remote sensing data like Synthetic Apertures Radar (SAR), NOAA's Advanced Very High Resolution Radiometer (AVHRR) and Airbrone Visible/Infrared Imaging Spectrometer (AVIRIS), respectively.

Accuracy Measurement Methods Used for Remote Sensing Imagery Classification

Stehman (1999) reviewed several basic probability sampling designs useful for accuracy assessment. According to the author, the first step in choosing the appropriate sampling design is to "define the population for which the accuracy assessment is needed, and to determine if this population will be partitioned into pixels, polygons, or some other aerial unit. Then the probability sampling design forms the statistical foundation of the assessment." Basically, "[c]hoosing a design from among a basic probability sampling design options should be guided by the project objectives and the relative importance of other remaining design criteria" (Stehman 1999). In addition, the criteria to consider when planning the sampling design are that "the sample should: (1) satisfy the probability sampling protocol, (2) be simple to implement and analyze, (3) result in low variance for the key estimates of the assessment, (4) permit adequate variance estimation, (5) be spatially well distributed, and (6) be cost effective" (Stehman 1999).

Congalton (1991) mentioned "researchers and users of remotely sensed data have a strong knowledge of both the factors needed to be considered as well as techniques used in performing any accuracy assessment". The accuracy assessment task can be defined as one of comparing two maps, one based upon analysis of remotely sensed data (the map to be evaluated), and another based upon a different source of information (Campbell 2002). Basically, the second map is designated the reference map, assumed to be accurate, that forms the standard for comparison. The reference data are of obvious significance; if they are in error, the attempt to measure accuracy will be in error (Campbell 2002).

The simplest method of evaluation is to compare the two maps with respect to the areas assigned to each category and the result of such comparison is to report the areal proportions of categories. These values report the extent of the agreement between the two maps with respect to total areas in each category, but do not take into account compensating errors in misclassification that cause this kind of accuracy measure to be itself inaccurate (Campbell 2002, Congalton and Green 1999). In addition, this form of error assessment is sometimes called non-site specific accuracy, because it does not consider agreement between the maps at specific locations, but only the overall figures for the two maps. The second form of accuracy, site-specific accuracy or classification error, is based upon the detailed assessment of agreement between maps at specific locations (Campbell 2002). This computation is performed by comparing a sample of locations on the map with the same locations on the reference data and keeping track of the number of times there is agreement (Congalton and Green 1999). In the majority of analyses, the units of comparison are simply pixels derived from the remote sensing data, although if necessary, a pair of matching maps can be compared using any network of uniform cells (Campbell 2002). After the maps are evaluated on over all accuracy, the need to evaluate individual categories within the classification scheme is recognized, and so began the use of the error matrix to represent map accuracy (Congaltan and Green 1999).

An error matrix is a square array of numbers set out in rows and columns which express the number of sample units (i.e. pixels, clusters of pixels, polygons) assigned to a particular category in one classification relative to the number of sample units assigned to particular category in another classification (Congalton 1991, Congalton and Green 1999). As noted, one of the classifications is considered to be correct (i.e. the reference data) and may be generated from aerial photography, airborne video, ground observation or ground measurement. The columns usually represent this reference data, while the rows indicate the classification generated from the remotely sensed data. In this way, the error matrix is a very effective way to represent map accuracy in that the individual accuracies of each category are plainly described along with both the errors of inclusion (commission errors) and errors of exclusion (omission errors) present in classification (Congalton and Green 1999, Congalton 1991). Sometimes "the error matrix is referred to as a confusion matrix because it identifies not only overall errors for each categories but also misclassifications (due to confusion between categories) by a category" (Campbell 2002).

In addition to clearly showing errors of omission and commission, the error matrix can be used to compute other accuracy measures, such as overall accuracy, producer's accuracy and user's accuracy (Story and Congalton 1986). Overall accuracy is simply the sum of the major diagonal (i.e., the correctly classified sample units) divided by the total number of sample units in the entire matrix and this is the most commonly reported accuracy assessment statistic and is probably most familiar to the readers (Congalton and Green 1999). Producer's and user's accuracy are ways of representing individual category accuracies instead of just the overall classification accuracy.

The inspection of the error matrix only reveals the overall nature of errors present; there is often a need for more objective assessment of classification (Campbell 2002). For example, if we are interested to know "are the two maps in agreement?" –this is a question very difficult to answer without the help of just an error matrix. The notion of agreement is difficult to define and implement. The error matrix is an example of a more general class of matrices, known as contingency tables. Some of the procedures that have been developed for analyzing contingency tables can be applied to examination of the error matrix (Campbell 2002).

Congalton (1981), Congalton et al. (1983) and Congalton and Green (1999) proposed application of techniques described by Bishop et al. (1975) and Cohen (1960), a discrete multivariate technique, as a measure of improving interpretation of an error matrix. A shortcoming of the usual error matrix is that even chance assignments of pixels to classes can result in surprisingly good results, as measured by percentage correct (Campbell 2002). Hord and Brooner (1976) and others have noted that the use of error matrix accuracy measures is highly dependent upon the samples, and therefore upon the sampling strategy used to derive the observations used in analysis.

Kappa is a discrete multivariate technique used in accuracy assessment for statistically determining if one error matrix is significantly different than another (Bishop et al. 1975, Congalton and Green 1999). The result of performing a Kappa analysis is a KHAT statistic

$$\hat{k} = \frac{\text{Observed} - \text{expected}}{1 - \text{expected}}$$

There are numerous studies on the assessment of the accuracy of remotely sensed data using the error matrix that calculate overall accuracy, producer's accuracy, user's accuracy and the Kappa statistic. Several studies that are relevant to this study were found.

Appropriate sample size requirement for the Error Matrix is one of the very important aspect, however "[a] balance between what is statistically sound and what is practically attainable must be found" (Congalton and Green, 1999). Further to this, the authors mention, a general guideline or good "rule of thumb" is to collect a minimum of 50 samples for each vegetation or land cover category in the error matrix. However, for the larger area (i.e. more than a million acres) or the classification with a large number of vegetation or land cover categories (i.e., more than 12 categories), the minimum number of samples needs to be increased to 75 or 100 samples per category (Congalton and Green, 1999).

Berlanga-Robels and Ruiz-Luna (2002) studied land use change mapping and change detection in the costal zone of Northwest Mexico using remote sensing. The authors used a multitemporal post-classification study with data from Landsat Multispectral Scanner (MSS) and TM to detect landscape changes. The authors compared four thematic maps (1973, 1986, 1990 and 1997) and classified the land-use into six classes as direct indicators of landscape condition. The accuracy of the classification (only in 1997 scene) was calculated from an error matrix, using overall accuracy assessment and the Kappa coefficients (Berlanga-Robels and Ruiz-Luna 2002).

Lawson (n. d.) used the National Resource Inventory (NRI) data, a compilation of natural resource information on non-Federal land in the United States, with Landsat TM scenes in Iowa. The author demonstrated the use of NRI point data for image classification and assessment of accuracy of a broad cover/use digital layer (TM scence) obtained from the Iowa Department of Natural Resources participating in the GAP program. The author used accuracy assessment tools including an error matrix, overall accuracy, user's accuracy and producer's accuracy.

Schreuder et al. (2003) made a number of recommendations for accuracy assessment of percent canopy cover, cover type and size class. The authors recommended to "well define vegetation types, stand size and canopy cover percentage to... [e]xplore the use of ambiguous classes, compute the contingency table and Kappa statistics for each of.... mapped categories, producers and users accuracy".

Kurvonen and Hallikainen (1999) studied the accuracy assessment of multitemporal ERS-1 and JERS-1 synthetic aperture radar (SAR) images in Finland test sites. The authors used the confusion matrix for land-cover type and forest type classification accuracy assessment. The authors mentioned that the use of textual parameters significantly improved the classification of land-cover and forest type classification.

Liu et al. (2003) compared the neural networks and statistical methods in classification of ecological habitats using FIA data. The authors used two artificial neural networks (ANN) and three traditional statistical classification methods to classify FIA plots into six ecological habitats in the U.S. Northeast and found four variables (overstory, understory species composition, hardwood basal area percentage, and current FIA forest type) as the most important discriminating variables for habitat classification. In this study also the authors used classification accuracy, Kappa statistics, and a classification success index to compare the classification of ecological habitats.

Wicham et al. (2004) assessed the thematic accuracy of the 1992 National Land Cover Data (NLCD) for the six western mapping regions of United States. The authors collected the reference data in each region for a probability sample of pixels stratified by map land-cover class. The authors assessed the thematic accuracy using overall accuracy percentage and an error matrix for Anderson Level II classification.

Wessels et al. (2004) compared the classification of the Moderate Resolution Imaging Spectroradiometer (MODIS) with the existing Landsat TM land cover maps as reference data for two major conservation areas (Greater Yellowstone Ecosystem-GYE, USA and the Parà State, Brazil). In this study, the Landsat TM land cover was processed to their fractional composition at the MODIS resolution (250 m and 500 m). The authors used the error matrix, overall accuracy, producer's accuracy and user's accuracy to assess the accuracy of MODIS thematic maps. The findings of this study suggested, in GYE, the MODIS land cover was very successful at mapping extensive cover types (e.g. coniferous forest and grasslands and far less successful at mapping smaller habitats (e.g. wetlands, deciduous tree cover) that typically occur in patches that were smaller than the MODIS pixels. For Parà State it was successful at producing a regional forest/non-forest product (Wessels et al. 2004). However, a single 500 m MODIS forest/non-forest product cannot be expected to reflect all the complex human impacts on biodiversity such as secondary regrowth, local land-use matrix dynamics and low intensity logging.

Similarly, Powell et al. (2004) studied the sources of error in an accuracy assessment of thematic land-cover maps in the Brazilian Amazon. The authors tried to quantify the subjectivity in reference data labeling and compared reference data produced by five trained interpreters. In addition, the authors identified the impact of other error sources, including geolocational errors between the map and reference data, land-cover changes between the dates of data collection, heterogeneous reference samples, and edge pixels. By the findings of this study the author suggested " (1) labels of continuous land-cover types are more subjective and variable than the commonly assumed, especially for the transitional classes[1]; (2) validation data sets that include only non-mixed, non-edges samples are likely to result in overly optimistic accuracy estimates, not representative of the map as a whole" (Powell et al. 2004).

Katila et al. (2000) introduced a statistical calibration method aimed at reducing the effects of map errors on multisource forest resource estimates. The authors developed a correction method based on the confusion matrix between land use classes of the field sample plots and the corresponding map information with the empirical example from the ninth National Forest Inventory of Finland.

There is a key concern that "the land cover maps derived are often judged to be of insufficient quality for operational applications" (Foody, 2003). For this reason, there is a need to compare the forest management databases available in Michigan. However, comparing the different forest management databases/maps, generated for specific management purposes, is a challenging task because of differences in objectives, spatial resolution and in the definitions of forest type classification. Different government entities like the MDNR, national forests and Forest Service

[1] However, using multiple interpreters to produce the reference data classification increases reference data accuracy.

research are using their own forest type classifications based on their management objectives. In this study, IFMAP is a thematic map and the other databases like FIA, OI and NRIS-FSVeg are developed based on field observation. In other words, the FIA, OI and NRIS-FSVeg are the databases developed from on-the-ground observation and IFMAP is generated from space observation. In addition, the spatial resolution of the each database is different. For example the OI spatial resolution is 900 m^2 and FIA is 4 X 1/24th acre. The NRIS-FSVeg and OI spatial resolution are based on timber stand sizes, the smallest forest management unit. Objectives of the NRIS-FSVeg and OI database are more or less the same; however, they are not spatially overlapping because of different ownerships. In addition, the IFMAP is a coarse-resolution map, derived from Landsat TM with cluster analysis (Chuvieco and Congalton 1988) using FIA data as the ground truth.

A number of authors (see for example, Katila et al. 2000, Wessels et al. 2004, Wicham et al. 2004, Kurvonen and Hallikainen 1999, Lawson, n. d.) have compared the accuracy of thematic maps with the other thematic maps, topographic maps, or ground truth data collected for different objectives. The review of previous studies on the accuracy of thematic maps suggests that, regardless of differences in various databases, the comparison of the accuracy using the standard accuracy assessment statistics like overall accuracy, producer's accuracy, user's accuracy and the Kappa coefficient of agreement can be used for this study. The rigorous field data collection of FIA and its exact plot location provide a sound basis for comparative analyses for a large number of forest types representing the majority of the forest area of Michigan.

CHAPTER 3

DATA SOURCES, STUDY AREA, CLASSIFICATION APPROACHES AND ANALYSIS METHODS

This research was a collaborative project among the Department of Forestry at Michigan State University, the Michigan Department of Natural Resources, national forests of Michigan and USDA-FS Forest Inventory Analysis unit at the North Central Research Station. Details of the data sources, study area, analysis methods are presented in the following sections.

Data Sources

Forest resource management and planning databases were obtained from the USDA Forest Service and the Michigan Department of Natural Resources. The NRIS-FSVeg information was obtained from the Huron-Manistee National Forest, Hiawatha National Forest and Ottawa National Forest. OI data for state-owned forests, which included the compartments with YOE 2003 to 2006 available in digitized GIS environment, were obtained from the Forest, Minerals and Fire Management Division of the MDNR. The national forest data were in vector GIS and were re-projected into Michigan Georef State Plane Coordinate System 1983 using the Projection function of ArcToolbox of ArcGIS Desktop (8.2), (Environmental Systems Research Institute, Redlands, California, USA). The OI data were already in Michigan Georef State Plane Coordinate System 1983.

Similarly. IFMAP, land cover 2001 which was in image format, was downloaded from the Michigan Spatial Data Library (SDL). The image file of IFMAP was transferred to the grid file using the Import-Export function of Imagine 8.7 (Leica Geosystems GIS & Mapping, LLC, St. Gallen, Switzerland ERDAS). IFMAP is in Michigan Georef, so the raster data was not re-projected. IFMAP land cover 2001 satellite imageries were taken during 1997 to 2001.

The data fields of the NRIS-FSVeg, national forest database, and OI, state-owned forest were carefully reviewed. Initially, a limited number of the stand parameters were chosen which may be useful in explaining differences/similarities of the forest type among the databases. Four key attributes of the stand were selected for each database. From NRIS-FSVeg the selected stand attributes were size density, stand dbh, stand age and survey year. From OI, the stand attributes of size density, stand age class, total basal area and understory type were selected. All the selected attributes were aggregated (generalized) to help maintain confidentiality of FIADB. By law, the USDA-FS must protect the confidentiality of FIA plot locations. Spatial location of forest for both NRIS-FSVeg and OI were aggregated to broader categories.

Similarly from FIADB, two plot-level information variables were selected, namely: (1) measurement year, and (2) kind code, and three attributes of stands were selected from the condition table, namely: (1) code for forest type of the condition (assigned by field crew), (2) aggregated stand age-the average total age, and (3) growing stock stocking code. The Ecological subsection code, forest types derived from an FIA algorithm and attributes information about present level of stocking similar to stand dbh from NRIS-FSVeg and total basal area from OI were initially selected. These attributes were dropped later from the analysis due to confidentiality concerns. The revision of the attributes useful in explaining differences in classification results was based on the potential problems in maintaining the confidentiality of FIADB. In the end, only the equivalent of forest type, from three databases, was selected.

Study Area

Out of the 19.28 million acres of forestland in Michigan, 7.14 million acres of forestland is owned publicly (Smith et al. 2004). Public ownership of forestland is distributed among the Federal government (National Forest, Bureau of Land Management and other), State government, counties and municipalities. In Michigan, there are three national forests, namely the Hiawatha, Huron-Manistee and Ottawa, covering about 2.68 million acres (Smith et al. 2004) (Figure 1). Forestland under State ownership is about 3.95 million acres (Smith et al. 2004). This study covers all national forests and the state forest land owned by the MDNR for which OI information has been digitized.

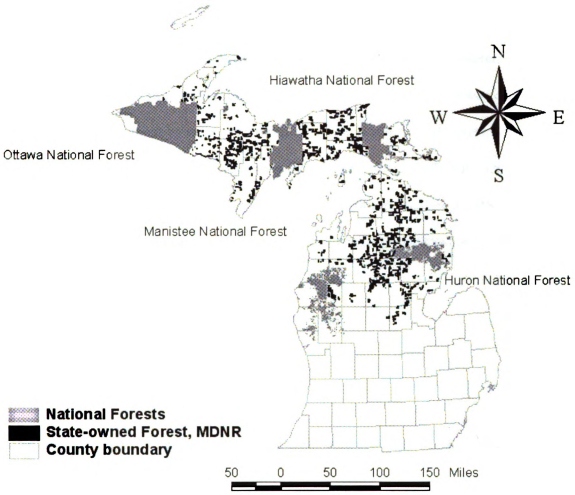

Figure1. National Forests and Digitized State-Owned Forests in Michigan

Classification Approaches

When every information source or map or database has its own classification system, making comparisons between and among them is not straightforward. The classification of the reference database or map needs to be regrouped by making it more compatible with other databases. In order to make forest classifications more compatible to each other, all the available forest types in each database were enumerated and a separate list of forest classification of each database was made (See Tables 1 to 4, Chapter I). Then pair-wise crosswalk tables were made by the author, and they were reviewed and revised based of expert judgment by experienced professionals working at Michigan State University (MSU, Drs. Donald Dickmann and Larry Leefers), the MDNR (Dr. Larry Pedersen and Mr. Jason Stephens) , and USDA-FS (Dr. Mark Hansen, NCRS and Mr. Joseph Gates, HMNF). Feedback and comments were incorporated to create the final crosswalk table (Appendix A).

Analysis Methods

After compiling agency data, the spatial and stand attributes were aggregated at the Forest Social Science and Economics Lab at Michigan State University (MSU). Methods to pick up data field from the grid and vector data were developed in the ESRI ArcView 3.2 environment. There were two types of data in this research. The NRIS-FSVeg and OI were in vector data shape files, and IFMAP data was in image or raster data. To compile the forest classification of these two types of data set for a point location (e.g an FIA plot) two approaches were proposed. For vector data the Mila Grid extension was downloaded from the ESRI website, which can be used to pick up the grid value from underlying raster data; this picks up the forest type values from the raster layer (IFMAP). Similarly for vector data sets, a script was downloaded from the ESRI website, which can be used to pick up the data field from the vector data. In this way, two methods were proposed to append the stand level data field like forest type, and the other variables selected to explain the difference in the forest types at the sample points. By overlaying the FIA plot vector data with the plot level and stand level information requested for this study as additional data fields to the FIA plot location, all the studied databases' forest type and the other selected stand level attributes could be added to the FIA plot level record.

In this study, Mr. Geoff Holden, North Central Research Station (NCRS), carried out this data compilation. Data extraction at the FIA plot locations was performed using two approaches depending on the data format. For vector data (OI and NRIS-FSVeg), a point in polygon overlay was performed using the "Intersect" tool in ArcGIS 9.0 to pick up the OI and NRIS-FSVeg forest classification. For raster data (IFMAP), the "Extract Values to Points" tool (from the Spatial Analyst extension in ArcGIS 9.0) was used to pickup the IFMAP classification. In this way, Mr. Holden carried out the overlay operation and combined the NRIS-FSVeg, OI and IFMAP information at the exact location of the FIA permanent plots. The resulting database table had a plot number (sequential, 1 to n), the FIA field crew defined (FLDTYPCD) classification, the OI or FS-Veg classification, and the IFMAP classification. Hence, each plot/record had data from three sources.

In this study, the approach to study remote sensing thematic mapping classifications was adopted to examine the site-specific accuracy of classification of both the satellite imagery and the OI and NRIS-FSVeg forest classification databases.

An error matrix compares the reference condition to the classified condition. Assume that n samples are distributed into k^2 cells where each sample is assigned to one of k categories in map or database and independently (usually in the rows), to one of the same k categories in the reference data set (usually the columns) (Figure2).

		j = columns (reference)				row total
		1	2	3	k	n_{i+}
i = rows	1	n_{11}	n_{12}	n_{13}	n_{1k}	n_{1+}
(classification)	2	n_{21}	n_{22}	n_{23}	n_{2k}	n_{2+}
	3	n_{31}	n_{32}	n_{33}	n_{3k}	n_{3+}
	k	n_{k1}	n_{k2}	n_{k3}	n_{kk}	n_{k+}
Column total n_{+j}		n_{+1}	n_{+2}	n_{+3}	n_{+k}	n

Note: Adopted from Congalton and Green (1999).

Figure 2. Illustration of an Error Matrix or a Confusion Matrix

Following Congalton and Green (1999), let n_{ij} denotes the number of samples classified into categories i (i = 1, 2, 3,..., k) in the remotely sensed classification or any classification under comparison (for e.g. OI, NRIS-FSVeg and IFMAP) and category j (j =1,2,3, ...k) in the reference data set (FIA data).

Let $$n_{i+} = \sum_{j=1}^{k} n_{ij}$$ (column sum) be the number of samples classified into category i in the remotely sensed

classification, and $n_{+j} = \sum_{i=1}^{k} n_{ij}$ (row sum)

be the number of samples classified into category j in the reference data set. The overall accuracy between the defined classification (IFMAP, NRIS-NRIS-FSVeg and OI) and the reference data (FIA) was computed as follows:

$$\text{Overall accuracy} = \frac{\sum_{i=1}^{k} n_{ii}}{n} \qquad (1)$$

Producer's accuracy can be computed by

$$\text{Producer's accuracy}_{j} = \frac{n_{jj}}{n_{+j}} \qquad (2)$$

And user's accuracy can be computed by

$$\text{User's accuracy}_{i} = \frac{n_{ii}}{n_{i+}} \qquad (3)$$

The user's accuracy (UA) and producer's accuracy (PA) can be used to calculate the error of commission and error of omission. Errors of omission (EO) refers to the samples of a certain class of the reference data that were not classified as such and errors of commission refers to the samples of a certain class of the classified data that were wrongly classified (Janssen and Van der Wel 1994).

Campbell (2002) applied the reference classification column sum as the denominator to calculate the Errors of Commissions however, Janssen and Van der Wel (1994) published the following relationships

User's accuracy % = 100% - Errors of Commission (%), and (4)

Producer's accuracy (%) = 100% - Error of Omission (%). (5)

A commission error is simply defined as including an area into a category when it does not belong to that category and an omission error is excluding that area from the category in which it truly does belong (Congalton and Green 1999). In this way, "every error is an omission from the correct category and a commission to a wrong category" (Campbell 2002).

Cohen (1960) developed a coefficient of agreement (called Kappa statistic) for nominal scales which measure the relationship of two classifications beyond chance agreement to expected disagreement. This measure of agreement uses all cells in the matrix, not just diagonal elements.

Again, let p_{ij} denote the proportion of samples in the i,jth cell, then corresponding to n_{ij}. In other word

$$p_{ij} = n_{ij} / n.$$

Then let p_{i+} and p_{+j} be defined by

$$p_{i+} = \sum_{j=1}^{k} p_{ij} \text{ and}$$

$$p_{+j} = \sum_{i=1}^{k} p_{ij}$$

The estimate of Kappa (K) is the proportion of agreement after chance agreement is removed from consideration, that is,

$$\hat{K} = \frac{(p_0 - p_c)}{(1 - p_c)} \tag{6}$$

in which

$p_o =$ proportion of units which agree

$p_c =$ proportions of units for expected chance agreement, and

$$p_o = \sum p_{ii}, \quad p_c = \sum (p_{i+} p_{+i}), \quad p_{ij} = X_{ij} / N$$

+ represents summation over the index (Rosenfield and Fitzpatrick-Lins 1986, Congalton and Green 1999) where N = total number of counts in matrix and X_{ij} = number of counts in *ij*th cell.

For computation purposes

$$\hat{K} = \frac{n \sum_{i=1}^{k} n_{ii} - \sum n_{i+} n_{+i}}{n^2 - \sum n_{i+} n_{+i}} \tag{7}$$

with n_{ii}, n_{i+} and n_{+i} as previously defined (Congalton and Green 1999).

The KHAT values are a measure of agreement or accuracy (Congalton and Green 1999).

Cohen mentions $\hat{K} = 0$ when obtained agreement equals chance agreement. Positive values of Kappa occur from greater than chance agreement; negative vales of Kappa are from less than chance agreement (Rosenfield and Fitzpatrick-Lins 1986). Congalton and Green note the KHAT values can range from +1 to -1. However, since there should be positive correlation between the forest management databases (NRIS-FSVeg, OI and IFAMP) and the FIA, positive KHAT values are expected.

The approximate large sample of Kappa is computed using the Delta methods as follows.

$$Var(K) = \frac{1}{n}\left\{\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2-\theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2\left(\theta_4-4\theta_2^2\right)}{(1-\theta_2)^4}\right\} \tag{8}$$

where $\theta_1 = \frac{1}{n}\sum_{1=1}^{k} n_{ii}$

$$\theta_1 = \frac{1}{n}\sum_{1=1}^{k} n_{ii}$$

$$\theta_2 = \frac{1}{n^2}\sum_{1=1}^{k} n_{i+}n_{+i}$$

$$\theta_3 = \frac{1}{n^2}\sum_{1=1}^{k} n_{ii}(n_{i+}+n_{+i})$$

$$\theta_4 = \frac{1}{n^3}\sum_{i=1}^{k}\sum_{j=1}^{k} n_{ij}(n_{j+}+n_{+i})^2$$

The variance of KHAT was calculated by using LabView 7.1 (National Instruments) Mr. Murari Regmi, a physics graduate student of MSU, wrote the program as per the author's instruction.

The test statistic for testing the significance of a single error matrix is expressed by

$$Z = \frac{\widehat{K}_1}{\sqrt{Var(\widehat{K}_1)}}. \qquad (9)$$

Z is standardized and normally distributed. Given the null hypothesis $H_o : K_1 = 0$, and the alternative $H_1 : K_1 \neq 0$, Ho is rejected if $Z \geq Z_{\alpha/2}$, where $\alpha/2$ is the confidence level if the two-tailed Z test and the degrees of freedom are assumed to be ∞ (infinity).

The KHAT values are a measure of agreement or accuracy. A KHAT value is computed for each error matrix, and it is a measure of how well the classification of a forest management database (NRIS-FSVeg, OI and IFMAP) agrees with the FIA forest classification. This provides a means for testing the significance of the KHAT statistics for a two independent KHAT values, and therefore two error matrices that are significantly different. With this test, it is possible to compare each forest management database with the FIA forest classification and the classification with higher accuracy was identified.

The test statistics for testing if two independent error matrices are significantly different is expressed by

$$Z = \frac{\left|\widehat{K}_1 - \widehat{K}_2\right|}{\sqrt{Var(\widehat{K}_1) + Var(K_2)}}. \qquad (10)$$

Z is standardized and normally distributed (Congalton and Green 1999). Given the null hypothesis $H_o : (K_1 - K_2) = 0$, and the alternative $H_1 : (K_1 - K_2) \neq 0$, Ho is rejected if Z

$\geq Z_{\alpha/2}$, where $\alpha/2$ is the confidence level if the two-tailed Z test and the degrees of freedom are assumed to be ∞ (infinity).

Chapter Summary

In summary, this chapter outlined the procedures to assess the accuracy of the thematic maps or databases with any reference data sets, which may be ground truth field data or another thematic map or any other source of reliable information. In this study, data were obtained from the MDNR, the Spatial Data Library of Michigan and national forests of Michigan. These data were compared for accuracy using the FIA field crew defined forest type classification (FLDTYPCD) as the reference for comparison. The research team requested the Forest Service North Central Research Station (NCRS) to compile each forest type classifications (FIA, FS Veg and IFMAP in national forest and FIA, OI and IFMAP in state forest) taking FIA exact location as the point of reference for comparison. The objectives of consistency assessments of the NRIS-FSVeg in the national forest lands and OI in the state forest lands and IFMAP in the public forest land of Michigan were carried out by computing the error matrix and computation of the overall accuracy, producer's accuracy and user's accuracy and Kappa statistic. Each pair of forest classification NRIS-FSVeg, OI and IFMAP was compared for the agreement in classification by comparing the KHAT statistic. The Z test was carried out to test two independent KHAT values, and therefore two error matrices, for significant differences. By performing this test, the forest classification of higher accuracy was identified both for national forest land and state forest land.

CHAPTER 4

RESULTS AND DISCUSSION

There are four primary sources of forest management information for public forest land in Michigan, namely IFMAP, OI (for state forest lands) and NRIS-FSVeg (for national forest lands) and FIA, each with different characteristics (Table 6). IFMAP is a statewide raster map at the resolution of 30 meters by 30 meters and has information on land cover classification. OI is a forest resource management database developed by the MDNR with the objective of supporting its day-to-day operational activity relating to resource management issues and land use (MDNR n. d.). OI has detailed information in databases covering information ranging from trees resources to special wildlife practices. Similarly, NRIS-FSVeg which is one of several natural resource information systems (NRIS) developed by USDA-FS, has stand-level information for the federally managed national forest land. NRIS-FSVeg has stand examinations, inventories and regeneration survey information to support management of national forests. It has wide range of information ranging from the trees to animal habitats. Besides these, the FIA sample plots have extensive information at the individual plot level.

Selected aspects of these four databases are summarized in Chapter 1. The overall comparison of these forest management information databases in terms of defining forest, coverage, objective and other details is summarized in Table 6.

Table 6. Comparison of Forest Management Information Databases in Michigan
(a)

Details	FIA	NRIS-FSVeg	OI	IFMAP-land cover
Definition of Forest	At least 16.7 % stocked by forest trees of any size, or formerly having had such tree cover, and not currently developed for nonforest use (Hahn and Hanson 1985).	At least 10 % stocked by forest trees of any size, including land that formerly had such trees cover and that will be naturally or artificially regenerated (Smith et al. 2004).	At least 16.7% stocked and land capable of producing 20 cu ft per year and preponderance of cover type	Proportion of trees canopy exceeds at least 25% of the ground cover (PMR 2001a)
Minimum area of forest	One acre	Generally one acre	Generally one acre	One pixel (0.2224 acre)
Minimum width to qualify as forest	Strips of timber with a crown width of at least 120 ft		No minimum width	
Objective of database	To assess information on the status and trends of U.S. forests	To meet forest resource management needs of a national forest	To meet the need of forest resource management of state-owned forest	To support the stratified forest inventory of MDNR including ecosystem management
Coverage of Database	Entire continental U.S.	National forests managed by USDA Forest Service	State-owned forest managed by FMFM, currently ~40% digitized	State of Michigan including wildlife game areas and state forest lands
Nature of inventory	Detail information and derived from extensive sample	Management oriented	Management oriented	Management oriented
Type of forest classification	Detailed and same across the continental U.S.	Detailed and more forest management friendly (e.g. dominant and successional tree listing)	Broader and aggregated from management perspectives	More coarse level aggregation, but can be integrated with stand detail

Note: FMFM is Forest Minerals and Fire Management.

Table 6. (Cont'd.)

(b)

Details	FIA	NRIS-FSVeg	OI	IFMAP land cover
Non timbered forestland classification	No Only Non stocked	Non-stocked lands, openings, shrubs, water, etc.	Bogs, rock, marsh, lowland brush, upland brush, bog, muskeg, sand dunes, water as non stocked	Herbaceous openland, shrub / Low-density trees, other bare / sparsely vegetated
Application	Large number of users of this database	Generally used within the USDA-FS work environment	Generally used within the MDNR FMFM and WLDF work environment	Enterprise GIS with broader applications (e.g. ecosystem management and resource assessments)
Anticipated level of accuracy	Detail and most accurate at plot level	Reasonably accurate for entire stand to meet ecosystem based forest management objective	Fairly accurate to meet natural resources management objective	Acceptable level of accuracy among the remotely sensed data at Level III classification

Note: FMFM is Forest Minerals and Fire Management and WLDF is Wildlife Division.

In defining the forest area, FIA, NRIS-FSVeg and OI are based on the stocking of the forest trees. OI has the requirement of 20 cu ft of timber production per annum per acre. In contrast, IFMAP defines a forest based on the percentage of tree canopy cover. As discussed in Chapter 1, IFMAP and OI have forest types for the lower density cover or stocking of trees. Regarding the objective of each database, OI and NRIS-FSVeg are very similar to each other. On the other hand, the FIA objectives are very different, and IFMAP's objectives are very broad. Similarly, regarding the coverage, the IFMAP is a "wall-to-wall" coverage of the entire state of Michigan, and OI and NRIS-FSVeg covers the state forest land and national forest land, respectively. Although, the FIA has only sample plots, it covers the continental U.S.

From an inventory point of view, OI, NRIS-FSVeg and IFMAP are very close to each other in that they yield a map. In contrast, FIA is extensive in terms of coverage and intensive at the sample unit level. In terms of forest type classification, FIA and NRIS-FSVeg are more detailed and OI forest classification is more specialized to suit the needs of forest management in Michigan. IFMAP has Level II and III land cover classification.

The users of these databases are different; OI and NRIS-FSVeg are used within the state and federal forest management agencies. IFMAP and FIA data can be used for a variety of reasons by large numbers of users. The anticipated levels of accuracy of these databases are very close to each other. It will be fair to say that the accuracy of these databases are more related to amount of resources spent to create them. In summary, there are multiple sources of forest management information in Michigan, developed for specific purposes. Prospective users should use either one or multiple sets of databases based on their coverage and objective of use. Only FIA and IFMAP information have complete statewide forest coverage in Michigan. To the extent that these databases are compatible to each other, they will provide more robust information for forest resource and wildlife habitat management and for sustainable use of ecosystem services.

Using FIA Forest Types as the Reference Classification

FIA has sample plot information, which are called "conditions" that were derived by "the discrete combination of landscape attributes that define the condition" (Miles et al. 2001) of the particular sampling unit. Forest type is one of the important information items included in the condition attributes. In addition, FIA has tree information that describes each tree over 1 inch in diameter found on subplots. In this study, the FLDTYPCD (a forest type assigned by FIA field crew) information from the condition attributes of FIA plots was used as the ground truth forest classification for assessing the accuracy of IFMAP, NRIS-FSVeg and OI forest type information.

The FIA plots were established using a systematic grid approach (USDA-FS 2004). The sample units for this study were chosen exactly at the FIA plot center location, so the sample units selected for this study could be taken as a sub-population of systematic sampling. The sets of data were filtered by a complex process of FIA data filtering before release. After completing the data filtering process of FIA, NCRS released the classification of 239 FIA sample plots (out of 397 possible) for state forest land and 523 sample plots (out of 753 possible) for national forest land (National FIA Database System). The 239 datasets for state forest land was due to the partial digitization (only 40%) of the state forest land database in OI. Accuracy assessments of this study were based on these sample units.

As discussed in Chapter 1, there were differences between the FIA, OI/NRIS-FSVeg and IFMAP classification schemes. While developing the crosswalk table between classifications, one-to-one relationships between the reference classifications to the "map" classification were expanded to include cases where categories were "acceptable" or "probably right". In a previous study, Congalton and Green (1999) used a similar approach in the California hardwood rangeland monitoring project. In this way, one-to-many relationships between the reference classification and "map" classification were identified for some reference classes. The most likely category is referred to as the "primary classification", and "acceptable" or "probably right" category is referred as the "secondary classification", of classified map hereafter.

In this study, for example, Jack pine in FIA may be reasonably comparable to Jack pine and Oak-Jack pine (secondary) in NRIS-FSVeg. The author had only the FIA field crew defined forest classification, the IFMAP classification and OI or NRIS-FSVeg classification. Due to FIA confidentiality concerns, further details from any data sources would have significantly reduced the number of sample plots available.

Congalton and Green (1993) identified eight factors that can affect error matrix or difference matrix results, and these factors provide useful framework for interpretation of the results of this study. The factors were: (1) land-cover change between the time of reference data acquisition and satellite data acquisition, (2) reference sample location error, (3) reference label data entry error, (4) reference label photo interpretation error, (5) inconsistent labeling of reference data due to land-cover heterogeneity surrounding the sample location, (6) difference in map and reference data registration, (7) map delineation error, and (8) map classification error (Wickham et al. 2004 cited as Congalton and Green 1993). Wickham and his colleagues (2004) also realized the usefulness of these factors for the interpretation of the results of error matrix. In this study, mainly factors (4) and (5) were important. To deal with these factors, measures were taken while developing the forest classification crosswalks between the reference classification and the classified map or database classification. This measure permitted matches for a number of reference forest types to multiple "map" forest classification types as a potential agreement between the two classifications. In other word, this assumption had made one reference forest classification to many classified forest classifications as "acceptable" classifications. For example, the Black Spruce of FIA was matched with the Lowland Coniferous Forest, Lowland Shrub and Mixed Non-Forested Wetland. These types of relaxation in matching of two classifications may overstate the accuracy of a map or database.

In addition, FIA represents "the four 1/24th acre" subplots, IFMAP represents 900 square meters of area and the OI and NRIS-FSVeg represents an entire stand with a particular forest type. In other words, the relative size of the area represented by these classifications was different. A recent study by Smith et al. (2002) revealed that "accuracy decreases as land cover heterogeneity increases and as patch size decreases". Further the author emphasized these landscape variables remain significant factors in explaining classification accuracy. As in this study, both the information on classified map/database and the reference map/database were secondary source information, so the author had to rely on a crosswalk between classifications to make comparisons.

In the forest classification crosswalk tables (Appendix A), there were multiple matches of FIA forest classification with a classified classification of IFMAP, OI and NRIS-FSVeg and OI/NRIS-FSVeg as reference classification with IFMAP as the classified classification. This type of one-to-many relationship between reference and classified map limited direct formulation of the Error matrix and consequently the assessment of the chance agreement and calculation of the KHAT statistic.

The accuracy assessments in this study were made from the difference classification matrix. A user's accuracy of 98.0% (Table 7) for Lowland Conifer Forest means that 50 of 51 of the pixels classified as Lowland Conifer Forest were lowland conifer forest types according to FIA. User's accuracy is sometimes called "reliability" in the comparison scheme (Janssen and Van der Wel 1994). For example, 18 of 37 Jack pine plots (FIA, 101) are classified as Pines or acceptable Lowland Coniferous Forest (Table 7). Similarly, dividing the number of correctly classified samples by the total column total yields the producer's accuracy: it indicates the percentage of samples of certain (reference) class that were correctly classified in the comparison scheme (Janssen and Van der Wel 1994).

Accuracy Assessment for State Forest Lands (MDNR)

To assess the accuracy of the IFMAP classification, FIA and OI classifications were considered as reference classifications. And to assess the agreement between OI and FIA forest classification, the FIA forest classification was used as reference classification. In other words, the accuracy of IFMAP and OI classification was assessed based on FIA classification (Tables 7-9). IFMAP overall accuracy was found to be 63.6% with FIA as the reference classification and 60.3% with OI as the reference classification. The overall accuracy of OI was 84.5% compared to FIA as the reference classification. There were differences between the FIA, OI and IFMAP classification schemes that limited the formulation of the error matrix and simultaneously calculating the chance agreement and calculation of KHAT statistic.

The match between the reference classification and classified classification were not along the diagonal as usual. The matched cells were shaded to identify them (Tables 7-12). Producer's and user's accuracy are in the last row and column of the difference classification matrix. The user's accuracy for the Lowland Conifer was more than 94% in comparison with FIA and OI (Tables 7-8). Thus, IFMAP did well in identifying the Lowland Conifer Forest type. The user's accuracy for Oak Association was as low as23.5% in comparison to the FIA classification (Table 7) and 29.4% in comparison to OI

Table 7. Difference Matrix of IFMAP and FIA for State Forest Lands

Classified Map — Reference Map (FIA)

Code	IFMAP	101	102	121	125	126	127	503	520	801	805	809	901	Total	Match	UA%
10	Herbaceous Openland	10	2	1									2	15	0	0.0
12	Upland Shrub / Low-density tree	4											6	10	0	0.0
14	Northern Hardwood Association			1			1			28	3	2	5	40	33	82.5
15	Oak Association							4				2	11	17	4	23.5
16	Aspen Association		1				1	2	1	7	2		21	35	21	60.0
19	Pines	17	10		3		1			1			2	34	27	79.4
21	Mixed Upland Conifers									1				1	0	0.0
22	Upland Mixed Forest	4	2				1	1		1			5	14	7	50.0
24	Lowland Deciduous Forest						1			1			2	4	0	0.0
25	Lowland Coniferous Forest	1	1	1	15	1	32							51	50	98.0
28	Lowland Shrub	1			6		1			1			4	13	7	53.8
30	Mixed Non-forest Wetland		1		2		1						1	5	3	60.0
	Total	37	17	3	26	1	39	7	1	40	5	4	59	239	152	
	Match	18	10	1	23	1	34	5	0	29	3	2	26	Overall accuracy		
	PA %	48.6	58.8	33.3	88.5	100	87.2	71.4	0.0	72.5	60.0	50.0	44.1	63.6%		

Note: PA, Producer's accuracy; UA, User's accuracy. FIA codes: 101, Jack pine; 102, Red pine;121 Balsam fir; 125 Black Spruce 126, Tamarack; 127, NW cedar; 503, White/Red oak/hickory; 520, Mixed upland hardwood; 801, Sugar maple/beech/yellow birch; 809, Red maple (upland); 901, Aspen.

Table 8. Difference Matrix of OI and FIA for State Forest Lands

Classified Database

Reference Map (FIA)

Code	OI Classification	101	102	121	125	126	127	503	520	801	805	809	901	Total	Match	UA %
J	Jack pine	27	3		1									31	27	87.1
R	Red Pine	3	13										2	18	13	72.2
S	Black Spruce				7									7	7	100.0
T	Tamarack					1								1	1	100.0
C	Cedar				3		25							28	25	89.3
O	Oak	4						7					1	12	7	58.3
M	Northern Hardwood									34	5	2	3	44	41	93.2
A	Aspen	3	1	2	1				1	5		2	53	68	53	77.9
D	Treed Bog				3									3	3	100.0
E	Swamp hrdwds						1							1	0	0.0
L	Lowland brush				5									5	5	100.0
Q	Mixed Swamp Conifer			1	6		13			1				21	20	95.2
	Total	37	17	3	26	1	39	7	1	40	5	4	59	239	202	
	Match	27	13	1	21	1	38	7	0	34	5	2	53	Overall accuracy		
	PA %	73.0	76.5	33.3	80.8	100	97.4	100	0.0	85.0	100	50.0	89.8	84.5%		

Note: PA, Producer's accuracy; UA, User's accuracy. FIA codes: 101, Jack pine; 102, Red pine; 121 Balsam fir; 125 Black Spruce; 126, Tamarack; 127, NW cedar; 503, White/red oak/hickory; 520, Mixed upland hardwood; 801, Sugar maple/beech/yellow birch; 809, Red maple (upland); 901, Aspen.

Table 9. Difference Matrix of IFMAP and OI for State Forest Lands

	Classified Map	Reference Database (OI)														
Code	IFMAP	A	C	D	Q	E	J	R	M	O	L	S	T	Total	Match	UA %
10	Herbaceous Openland	3					7	3		2				15	0	0.0
12	Upland Shrub / Low-density trees	6					4							10	0	0.0
14	Northern Hardwood Association	10	1						29					40	29	72.5
15	Oak Association	12								5				17	5	29.4
16	Aspen Association	21			1	1		1	9	2				35	21	60.0
19	Pines	2		1	2		14	12	1	2				34	26	76.5
21	Mixed Upland Conifers								1					1	0	0.0
22	Upland Mixed Forest	5			1		4	1	2	1				14	7	50.0
24	Lowland Deciduous Forest	2	1						1				1	5	0	0.0
25	Lowland Coniferous Forest	2	25	2	14			1				6		50	47	94.0
28	Lowland Shrub	4			2		1		1		4	1		13	7	53.8
30	Mixed Non-Forest Wetland	1	1		1		1				1			5	2	40.0
	Total	68	28	3	21	1	31	18	44	12	5	7	1	239	144	
	Total match	26	25	2	17	0	14	12	31	5	5	7	0	Overall accuracy		
	PA %	38.2	89.3	66.7	81.0	0.0	45.2	66.7	70.5	41.7	100	100	0.0	60.3%		

Note: PA, Producer's accuracy; and UA, User's accuracy. OI Codes: A, Aspen; C, Cedar; D, Treed Bog; Q; Mixed Swamp Conifer; E, Swamp hardwood; J, Jack pine; R, Red pine; M, Northern Hardwood; O, Oak; L, Lowland brush; S, Black Spruce; T, Tamarack.

classification (Table 9). Thus, IFMAP did a poor job in identifying the Oak Association. It was confused with Aspen in both comparisons. Regarding the producer's accuracy, IFMAP did well in identifying the Northern White Cedar (87.1%) and Black Spruce (88.5%) of the FIA classification and Cedar (89.3%) and Mixed Swamp Conifer (81%) of the OI classification. However, for Northern White Cedar and Black Spruce of FIA and Cedar and Mixed Swamp Conifer of OI, the crosswalk allowed multiple IFMAP classifications as "acceptable". This is one of the reasons for the small error of omission in these categories of classification. Due to lack of further information, the author could not verify any existence of map classification error in these forest classification categories.

IFMAP was inaccurate in identifying Aspen (38.2 % PA for OI and 44.1% PA for FIA) and Jack pine (48.6 % PA) for FIA and Oak (41.7% PA) for OI. In classifying Aspen, IFMAP had confusion with the Northern Hardwood and the Oak Association (Table 9) and the Oak Association and low density cover forest types (Table 7). However, due to the small number of samples in the Oak classification, the results of this study were not conclusive and require further verification with adequate number of samples to infer about Oak confusion with other categories of IFMAP.

From the difference matrix of OI and FIA, for most of OI classification there were similarities with the FIA classification. From an OI data user's perspective, Oak had the least user's accuracy. The Oak classification of OI had confusion with the Jack Pine. In nature these two species are often found together. This type of Oak trees in Jack pine forest type of FIA was reported in Michigan Forest Statistics (Table 47), 1993 (Leatherberry and Spencer 1996). However, the small number of sample units and lack of other stand information from the OI database relative to FIA plots, made it difficult to derive further conclusions. There was also evidence of confusion in OI in classifying Red pine and Jack pine. Similarly from the producer's perspective, OI had confusion in identifying Jack Pine with Oak, Red Pine and Aspen, and Red Pine had confusion with Jack Pine.

Accuracy Assessment for the National Forest Lands

To assess the accuracy of IFMAP classification, the FIA and NRIS-FSVeg classifications were regarded as reference classification. And to assess the agreement between NRIS-FSVeg and FIA forest classification, the FIA forest classification was used as the reference classification (Table 10-12).

The overall accuracy of IFMAP, without reducing chance agreement, was 64.8% taking FIA as the reference. The agreement of forest classification between the NRIS-FSVeg and FIA without reducing chance agreement was 82.2%.

By comparing user's accuracy, IFMAP was very accurate for the Lowland Coniferous Forest (89.6%), the Northern Hardwood Association (86.0%) and the Pines (82.7 %) relative to FIA (Table 10). And it did well for Lowland Deciduous Forest (87.5 %), and

Table 10. Difference Matrix of IFMAP and FIA for National Forest Lands

(a)

Classified Map — Reference Classification (FIA)

Code	IFMAP	101	102	103	105	121	122	125	127	400	401	503
10	Herbaceous Openland	3	2									
12	Upland Shrub / Low-density trees	1										
14	Northern Hardwood Association				1							
15	Oak Association		3									10
16	Aspen Association		1						2		2	3
18	Mixed Upland Deciduous	1										
19	Pines	38	71				1	6	2	1		
20	Other Upland Conifers	3	5									
21	Mixed Upland Conifers		5									
22	Upland Mixed Forest	11	10						1			
24	Lowland Deciduous Forest		1									
25	Lowland Coniferous Forest					1		13	29			
28	Lowland Shrub	1						7	2			
30	Mixed Non-Forest Wetland	4		1				5	4			
	Total	62	98	1	1	1	1	31	40	1	2	13
	Match	38	71		1	1	0	25	35	1	0	10
	PA%	61.3	72.4	0.0	100	100	0.0	80.6	87.5	100	0.0	76.9

Note: PA, Producer's accuracy; UA, User's accuracy. FIA codes:101, Jack pine; 102, Red pine; 103 Eastern white pine; 105, Eastern hemlock;121, Balsam fir;122, White spruce; 125, Black spruce; 127, Northern white-cedar; 400, Oak Pine Group; 401, White pine / Red oak / white ash; 503, White oak / red oak / hickory.

Table 10. (Cont'd).

(b)

Code	IFMAP	504	505	701	800	801	805	809	901	Total	Match	UA %
10	Herbaceous Openland	3				1				9		0
12	Upland Shrub / Low-density trees								1	2		0
14	Northern Hardwood Association			4	4	94	11	1	14	129	111	86.0
15	Oak Association	2	4			1			15	35	16	45.7
16	Aspen Association		1		1	30	1	1	9	51	9	17.6
18	Mixed Upland Deciduous									1		0.0
19	Pines	1		1		1		3	8	133	110	82.7
20	Other Upland Conifers					1			1	10		0.0
21	Mixed Upland Conifers					1				6		0.0
22	Upland Mixed Forest	2		4	1	15		1	13	58	30	51.7
24	Lowland Deciduous Forest			2					5	8	2	25.0
25	Lowland Coniferous Forest							2	3	48	43	89.6
28	Lowland Shrub					2			5	17	9	52.9
30	Mixed Non-Forest Wetland								2	16	9	56.3
	Total	8	5	11	6	146	12	8	76	523	339	
	Match	2	4	2	5	109	11	2	22	Overall accuracy		
	PA %	25.0	80.0	18.2	83.3	74.7	91.7	25.0	28.9	64.8%		

Note: PA, Producer's accuracy; UA, User's accuracy. FIA codes: 504, White oak; 505, Northern red oak; 701, Black ash/American elm/red maple; 800, Maple / Beech / Birch Group; 801, Sugar maple / beech / yellow birch; 805, Hard maple / basswood; 809, Red maple / upland; 901, Aspen.

Table 11. Difference Matrix of NRIS-FSVeg and FIA in National Forest Lands

(a)

Classified Map / Reference Map (FIA)

Code	NRIS-FSVeg	101	102	103	105	121	122	125	127	400	401	503	504
1	Jack Pine	45	5					1					
2	Red Pine	3	80					1					
3	White Pine		2										
5	Hemlock				1								
11	Balsam Fir-Asp-PB							1	2				
12	Black Spruce							13					
14	Northern Wh Cedar								19				
18	Mix Swamp conifer							8	12				
19	Cedar-Aspen-PB								2				
21	Mixed Northern Hdw												
48	Jack Pine-Oak	9	1										
49	Red Pine-Oak		4										
53	Black Oak	1	1							1			2
55	Northern Red Oak											4	
59	Mixed Oak	1	2								2	9	6
71	BI Ash-Elm-R Maple												
81	S Maple-Beech -YB												
82	S Maple-Basswood												
84	Red Maple(Dry)												
85	Sugar Maple												
87	SM-beech-YB/red spruce												
89	Mixed Upland Hdwd												
91	Quaking Aspen	1				1	1	1	1				
93	Bigtooth Aspen		2										
95	Asp-W Spruce BF								1				
97	Lowland Brush	1						5					
99	Open	1	1	1				1	3				
	Total	62	98	1	1	1	1	31	40	1	2	13	8
	Match	54	84	0	1	0	0	26	33	0	0	13	6
	PA %	87.1	85.7	0.0	100	0.0	0.0	83.9	82.5	0.0	0.0	100	75.0

Note: PA, Producer's Accuracy; and UA, User's Accuracy. FIA codes:101, Jack pine; 102, Red pine; 103, Eastern white pine; 105, Eastern hemlock;121, Balsam fir;122, White spruce; 125, Black spruce; 127, Northern white-cedar; 400, Oak Pine Group; 401, White pine / Red oak / white ash; 503, White oak / red oak / hickory

Table 11. (Cont'd).
(b)

Code	NRIS-FSVeg	504	505	701	800	801	805	809	901	Total	Match	UA %
1	Jack Pine									51	45	88.2
2	Red Pine							1	1	86	80	93.0
3	White Pine									2	0	0.0
5	Hemlock					3				4	1	25.0
11	Balsam Fir-Asp-PB				1	2			3	9	0	0.0
12	Black Spruce									13	13	100.0
14	Northern Wh Cedar									19	19	100.0
18	Mix Swamp conifer					2			1	23	20	87.0
19	Cedar-Aspen-PB									2	2	100.0
21	Mixed Northern Hdw								1	1	0	0.0
48	Jack Pine-Oak									10	9	90.0
49	Red Pine-Oak									4	4	100.0
53	Black Oak	2								5	0	0.0
55	Northern Red Oak									4	4	100.0
59	Mixed Oak	6	5							25	20	80.0
71	BI Ash-Elm-R Maple			3		1				4	3	75.0
81	S Maple-Beech -YB			5	2	51	4	2	3	67	53	79.1
82	S Maple-Basswood					6			1	7	6	85.7
84	Red Maple(Dry)					3				3	0	0.0
85	Sugar Maple				2	25	7			34	34	100.0
87	Sugar maple-beech-yellow birch/red spruce					6				6	6	100.0
89	Mixed Upland Hdwd			1		42	1	2	1	47	45	95.7
91	Quaking Aspen			2	1	4		3	42	57	42	73.7
93	Bigtooth Aspen								12	14	12	85.7
95	Asp-W Spruce BF					1			7	9	7	77.8
97	Lowland Brush								3	9	5	55.6
99	Open								1	8	0	0.0
	Total	8	5	11	6	146	12	8	76	523	430	
	Match	6	5	3	4	130	8	2	61	Overall accuracy		
	PA %	75.0	100	27.3	66.7	89.0	66.7	25.0	80.3	82.2%		

Note: PA, Producer's Accuracy; and UA, User's Accuracy. FIA codes: 504, White oak; 505, Northern red oak; 701, Black ash / American elm / red maple; 800, Maple / Beech / Birch Group; 801, Sugar maple / beech / yellow birch; 805, Hard maple / basswood; 809, Red maple. / upland; 901, Aspen

Table 12. Difference Matrix of IFMAP and NRIS-FSVeg for National Forest Lands

(a)

Classified Map — Reference Map (FSVeg)

Code	IFMAP	1	2	3	5	11	12	14	18	19	21	48	49	53	55	59
10	Herbaceous Openland	3	1													3
12	Upland Shrub / Low-density trees	1														
14	Northern Hardwood Association				2	2			1							
15	Oak Association		2								1				3	15
16	Aspen Association					2			1						1	5
18	Mixed Upland Deciduous	1														
19	Pines	31	64	2		3	2		2			6	4	4		1
20	Other Upland Conifers	3	5													
21	Mixed Upland Conifers		5													
22	Upland Mixed Forest	8	8		2				2			4		1		1
24	Lowland Deciduous Forest		1													
25	Lowland Coniferous Forest					2	5	18	12	2						
28	Lowland Shrub	1					3	1	3							
30	Mixed Non-Forest Wetland	3					3		2							
	Match	31	64	2	4	4	11	19	17	2	0	10	4	1	3	16
	Total	51	86	2	4	9	13	19	23	2	1	10	4	5	4	25
	Producer's accuracy	60.8	74.4	100	100	44.4	84.6	100	73.9	100	0.0	100	100	20.0	75.0	64.0

Note: PA, Producer's accuracy; UA, User's accuracy. NRIS-FSVeg codes: 1, Jack Pine; 2, Red Pine ; 3, White Pine; 5, White Pine-Hemlock; 11, Balsam Fir-Asp-PB; 12, Black Spruce; 14, Northern Wh Cedar; 18, Mix Swamp conifer; 19, Cedar- Aspen-PB; 21, Mixed Northern Hdw; 48, Jack Pine-Oak; 49, Red Pine-Oak; 53, Black Oak; 55, Northern Red Oak; 59, Mixed Oak;

Table 12. (Cont'd).
(b)

Code	IFMAP	71	81	82	84	85	87	89	91	93	95	97	99	Total	Match	UA %
10	Herbaceous Openland							1					1	9	0	0.0
12	Upland Shrub / Low-density trees								1					2	0	0.0
14	Northern Hardwood Association	2	45	7	2	24	2	26	13	1	1		1	129	108	83.7
15	Oak Association		1						6	7				35	18	51.4
16	Aspen Association		9		1	6	2	11	8	4	1			51	15	29.4
18	Mixed Upland Deciduous													1	0	0.0
19	Pines		3			1			7	1	1		1	133	107	80.5
20	Other Upland Conifers		1						1					10	0	0.0
21	Mixed Upland Conifers		1											6	0	0.0
22	Upland Mixed Forest		7			2	2	7	8	1	5			58	40	69.0
24	Lowland Deciduous Forest	2							5					8	7	87.5
25	Lowland Coniferous Forest							1	4		1	2	1	48	39	81.3
28	Lowland Shrub					1		1	3			4		17	11	64.7
30	Mixed Non-Forest Wetland								1			3	4	16	12	75.0
	Match	2	52	7	2	26	4	33	21	5	6	7	4	523	357	
	Total	4	67	7	3	34	6	47	57	14	9	9	8	Overall accuracy		
	Producer's accuracy	50.0	77.6	100	66.7	76.5	66.7	70.2	36.8	35.7	66.7	77.8	50.0	68.3%		

Note: PA, Producer's accuracy; UA, User's accuracy. NRIS-FSVeg codes: 71, BI Ash-Elm-R Maple; 81, S Maple-Beech – YB; 82, S Maple- Basswood;84, Red Maple(Dry);85, Sugar Maple ; 87, Sugar maple-beech-yellow birch/red spruce; 89, Mixed Upland Hdwd; 91, Quaking Aspen; 93, Bigtooth Aspen; 95, Asp-W Spruce BF; 97, Lowland Brush; 99 Open.

the Northern Hardwood Association (83.7 %), and the Lowland Coniferous Forest (81.3%) in comparison to NRIS-FSVeg. IFMAP did a very poor job for the Aspen Association (17.6 %), the Lowland Deciduous forest (25.0%) and Oak Association (45.7 %) in comparison to FIA classification.

IFMAP had confusion with Aspen in classifying the Lowland Deciduous Forest and the Oak Association in comparison to the FIA classification. Similarly, it was inaccurate for the Aspen Association (29.4%) and the Oak Association (51.4%) in comparison to NRIS-FSVeg classification. In classifying Aspen it had confusion with Sugar Maple-Beech-Yellow Birch and Oak, and in classifying the Oak Association it had confusion with Quaking Aspen and Bigtooth Aspen classification of NRIS-FSVeg.In this way, IFMAP did a more accurate job in classifying the Northern Hardwood Association and the Lowland Coniferous Forest in comparison with both classifications (FIA and NRIS-FSVeg). Similarly, it was inaccurate for the Aspen Association and the Oak Association in comparison with both classifications.

The difference matrix of NRIS-FSVeg and FIA revealed that there was excellent correspondence between most of the forest types between two classifications except the Hemlock forest type defined by NRIS-FSVeg (Table 11). The Hemlock forest and Lowland brush of NRIS-FSVeg had indication of confusion with the Aspen forest type of FIA. However, due to small numbers of sample units in these forest types, these pieces of evidence are not enough to support the hypothesis of confusion. Similarly, by comparing the producer's accuracy, it can be inferred that Red maple (upland), Black ash/American elm/Red maple of FIA classification had lower match with the corresponding NRIS-FSVeg corresponding classification. Red maple (upland) and Black ash/American elm/red maple had confusion with Quaking Aspen and Sugar maple/beech/YB. However, due to small numbers of sampling units in these classifications these hypotheses were not conclusive. Similar to OI and FIA comparison, the NRIS-FSVeg Red pine, classification has confusion with Jack pine, and Jack pine classification has confusion with Red pine. In a case of NRIS-FSVeg classification, there were Jack pine-Oak and Red pine-Oak classifications which gave more opportunity for matches for Red pine or Jack pine classified in FIA classification. However, in the OI classification, there were no such transitional forest classifications so there was marked confusion between both Red pine and Oak (Table 8).

Three-way Accuracy of FIA/OI/IFMAP Assessment for State Forest Lands

The three-way accuracy of the 239 sample units in the state forest land was assessed by computing the producer's accuracy of FIA classification and user's accuracy of OI and IFMAP. The overall accuracy in these three classifications was found to be 54.8%, that is, on average 54.8% of the plots would have comparable classifications in all the three schemes (Table 13).

The producer's accuracy of FIA gave information about how well each FIA classification matched concurrently with OI and IFMAP. The percentage match of the Northern White Cedar forest type of FIA was highest followed by White/Red oak hickory and Black spruce. Among the poorly matched forest types of FIA were Jack pine, Balsam fir, and Aspen. The user's accuracy of OI and IFMAP might be useful to infer about the errors of commission in these classified forest categories. The Aspen classification of OI had the least user's accuracy followed by Jack pine and Oak. The best performing classifications of OI were Black Spruce, Cedar and Lowland Brush. There was similar evidence from

Table 13. Three way accuracy assessment of FIA, OI and IFMAP for State Forest Lands

(a) Producer's accuracy

Code	Description	Total Match	Column Total	Overall accuracy %
101	Jack Pine	12	37	32.4
102	Red Pine	9	17	52.9
121	Balsam Fir*	1	3	33.3
125	Black Spruce	17	26	65.4
126	Tamarack*	1	1	100.0
127	NW cedar	34	39	87.2
503	W/R oak/hickory*	5	7	71.4
520	Mixed upland hardwood*	0	1	0.0
801	Sugar maple/beech/yellow birch	23	40	57.5
805	Hard maple/ basswood*	3	5	60.0
809	Red maple(upland)*	2	4	50.0
901	Aspen	24	59	40.7
	Overall accuracy %	131	239	54.8

(b) User's Accuracy OI

Code	Description	Total Match	Column Total	User's accuracy %
A	Aspen	24	68	35.3
C	Cedar	23	28	82.1
D	Treed Bog*	2	3	66.7
E	Swamp Hardwds*	0	1	0.0
J	Jack pine	12	31	38.7
L	Lowland Brush*	4	5	80.0
M	Northern Hardwood	28	44	63.6
O	Oak	5	12	41.7
Q	Mx Swamp Conifer	16	21	76.2
R	Red Pine	9	18	50.0
S	Black spruce*	7	7	100.0
T	Tamarack*	1	1	100.0
	Overall accuracy %	131	239	54.8

*refers to small number of sample units

Table 13. (Cont'd).

(c) User's accuracy IFMAP

Code	Description	Total Match	Column Total	User's accuracy %
10	Herbaceous Openland	0	15	0.0
12	Upland Shrub / Low-density trees	0	10	0.0
14	Northern Hardwood Association	28	40	70.0
15	Oak Association	4	17	23.5
16	Aspen Association	20	35	57.1
19	Pines	21	34	61.8
21	Mixed Upland Conifers*	0	1	0.0
22	Upland Mixed Forest	5	14	35.7
24	Lowland Deciduous Forest*	0	4	0.0
25	Lowland Coniferous Forest	46	51	90.2
28	Lowland Shrub	6	13	46.2
30	Mixed Non-Forest Wetland*	1	5	20.0
	Overall accuracy	131	239	54.8

* refers to small number of sample units

the OI-FIA difference matrix that, Aspen had the third lowest user's accuracy (Table 8). However, the three-way accuracy assessment was dependent on all three sets of However, the three-way accuracy assessment was dependent on all three sets of classifications.

From IFMAP user's accuracy assessment, Oak had the least accuracy of 23.5%. From this evidence it could be inferred that when IFMAP classified the Oak Association, then there would be less than 25% likelihood that the two other classifications also classify the given sample unit as Oak Association. IFMAP had the highest user's accuracy for Lowland Conifer followed by the Northern Hardwood Association. These conclusions from the three-way comparison were very similar to the previous IFMAP-FIA comparison (Table 7).

Three-way Accuracy of FIA/NRIS-FSVeg/IFMAP Assessment for National Forest Lands

The three-way accuracy assessment of 523 sample units in the national forest land was carried out by computing the producer's accuracy of FIA classification and user's accuracy of IFMAP and NRIS-FSVeg classifications. The overall agreement among these three classifications was found to be 58.5%, which was slightly higher than the similar comparison in state forest land (Table 14).

Producer's accuracy of FIA in this classification was useful to provide information about how well the other classification matched with FIA plots and the user's accuracy of two classified maps or databases were useful to provide information about how well these classifications correctly classified the corresponding FIA information. The producer's accuracy of FIA classification was excellent for Hard maple/basswood (91.7%) and good for Northern red oak (80%). The producer's accuracy was low for Black ash/American elm/Red maple (18.2%), Red maple/upland (25%), and Aspen (26.3%). In these categories the NRIS-FSVeg and IFMAP were unable to correctly classify these FIA forest types in their corresponding classified map category.

The user's accuracy of IFMAP was excellent for Lowland Conifer Forest (81.3%) and the Northern Hardwood Association (79.8%). However, the user's accuracy of IFMAP was very low for Other Upland Conifers (0.0%), the Aspen Association (15.7%), and Lowland Deciduous Forest (25%). The IFMAP classifications were also not so promising

Table 14. Three-way accuracy comparison of FIA, NRIS-FSVeg and IFMAP in National Forest Lands

(a) Producer's accuracy of FIA classification

Code	Description	Total Match	Column Total	Producer's Accuracy %
101	Jack pine	32	62	51.6
102	Red pine	63	98	64.3
103	Eastern white pine	0	1	0.0
105	Eastern hemlock	1	1	100.0
121	Balsam fir		1	0.0
122	White spruce		1	0.0
125	Black spruce	24	31	77.4
127	Northern white-cedar	30	40	75.0
400	Oak Pine Group	0	1	0.0
401	White pine / Red oak / white ash		2	0.0
503	White oak / red oak / hickory	10	13	76.9
504	White oak	3	8	37.5
505	Northern red oak	4	5	80.0
701	Black ash / American elm / red maple	2	11	18.2
800	Maple / Beech / Birch Group	3	6	50.0
801	Sugar maple / beech / yellow birch	101	146	69.2
805	Hard maple / basswood	11	12	91.7
809	Red maple / upland	2	8	25.0
901	Aspen	20	76	26.3
	Overall accuracy	306	523	58.5%

(b) User's accuracy IFMAP

code	Description	Total Match	Row Total	User's Accuracy %
10	Herbaceous Openland	0	9	0.0
12	Upland Shrub / Low-density trees	0	2	0.0
14	Northern Hardwood Association	103	129	79.8
15	Oak Association	16	35	45.7
16	Aspen Association	8	51	15.7
18	Mixed Upland Deciduous	0	1	0.0
19	Pines	95	133	71.4
20	Other Upland Conifers	0	10	0.0
21	Mixed Upland Conifers	0	6	0.0
22	Upland Mixed Forest	28	58	48.3
24	Lowland Deciduous Forest	2	8	25.0
25	Lowland Coniferous Forest	39	48	81.3
28	Lowland Shrub	9	17	52.9
30	Mixed Non-Forest Wetland	6	16	37.5
	Overall accuracy	306	523	58.5%

Table 14 (Cont'd)

(c) User's accuracy of NRIS-FSVeg

Code	Description	Total Match	Column Total	User's Accuracy %
1	Jack Pine	27	51	52.9
2	Red Pine	59	86	68.6
3	White Pine		2	0.0
5	Hemlock	4	4	100.0
11	Balsam Fir-Asp-PB		9	0.0
12	Black Spruce	11	13	84.6
14	Northern Wh Cedar	19	19	100.0
18	Mix Swamp conifer	17	23	73.9
19	Cedar-Aspen-PB	2	2	100.0
21	Mixed Northern Hdw	0	1	0.0
48	Jack Pine-Oak	5	10	50.0
49	Red-Pine-Oak	4	4	100.0
53	Black Oak	1	5	20.0
55	Northern Red Oak	3	4	75.0
59	Mixed Oak	13	25	52.0
71	BI Ash-Elm-R Maple	2	4	50.0
81	S Maple-Beech -YB	47	67	70.1
82	S Maple-Basswood	6	7	85.7
84	Red Maple(Dry)	0	3	0.0
85	Sugar Maple	26	34	76.5
87	Sugar maple-beech-yellow birch/red spruce	4	6	66.7
89	Mixed Upland Hdwd	31	47	66.0
91	Quaking Aspen	12	57	21.1
93	Bigtooth Aspen	3	14	21.4
95	Asp-W Spruce BF	5	9	55.6
97	Lowland Brush	5	9	55.6
99	Open		8	0.0
	Overall accuracy	306	523	58.5%

for Mixed Non-forest Wetland (37.5 %), Oak Association (45.7%) and Upland Mixed Forest (48.3%). These conclusions were very consistent with the results of IFMAP-FIA comparison in the national forest land (Table 10).

The user's accuracy of NRIS-FSVeg was found excellent for Northern White Cedar, Cedar-Aspen-Paper birch and Hemlock. The user's accuracy of NRIS-FSVeg forest types was poor ranging from 20 to 22% for Black oak, Quaking aspen, and Bigtooth aspen.

Similarly, the NRIS-FSVeg performance for Jack pine, Jack pine-Oak, and Mixed Oak was not so promising. However, these results were not similar to the NRIS-FSVeg-FIA comparison except for the Black oak (Table 11). The lower overall accuracy and lower user's accuracy in the above mentioned classifications of NRIS-FSVeg were mostly due to mismatch with the similar forest types of IFMAP classification.

Comparison of Classification Agreement Between IFMAP and FIA for State Forest Land

The well accepted method for comparing two classifications is to formulate an error matrix or confusion matrix and to calculate of a Kappa statistic, a coefficient of agreement between two classifications in comparison to random classification. As already noted, there were marked differences between the classification schemes of these four forest classification systems. When there were multiple categories of classified classifications which were acceptable or probably right to one category of reference classification, then formulation of a square error matrix with a matching category diagonal became difficult. The choices for formulating the error matrix were examined: (1) keeping the user's accuracy for each classified map constant, (2) keeping the producer's accuracy constant for each reference map, and (3) aggregating classified forest type categories with multiple matches in the reference classifications. In this study, these three approaches were adopted to assess effects of accuracy of IFMAP in comparison with FIA forest classification for both state forest land (in this section) and national forest land (in the next section). Overall, this provides a partial assessment of the utility of

IFMAP information. After formulation of error matrices using the three approaches, the first two approaches yielded the same overall accuracy as the difference matrices. For the last approach, aggregating the classified map categories, the overall accuracy will increase, but details of forest classification relationships will be lost. For many users, a loss of detailed information may be unacceptable.

In the UA constant approach, the "acceptable" or "probably right" classified map categories were moved left or right in the error matrix to the appropriate diagonal cell (Table 16). As a result the user's accuracy remained unchanged. However, the producer's accuracies (column) may be modified. Similarly, in the PA constant approach, in the "acceptable" or "probably right" map categories were moved up in the error matrix to the appropriate diagonal cell (Table 17). The user's accuracy of the "acceptable" matches or primary classification of IFMAP may be modified, but the producer's accuracy does not change. For example, in the UA constant approach, the PA for Upland Mixed Forest, Lowland Shrub, and Mixed Non-forest Wetland was 100 % (Table 15), which was not true in the original difference matrix (Table 7). And for the PA constant approach, the UA for the Pines, Lowland conifer, Northern Hardwood increased considerably (Tables 7 and 16); this was not the case in the original difference matrix. From the user's perspective, the approach of UA constant was found relatively fair for the "acceptable" match or primary classification types; however, this approach overestimated the producer's accuracy of the "probably right" or secondary classification types. In the UA constant approach, the classified map label was not changed as movements in the error

Table 15. Error Matrix of FIA-IFMAP for State Forest Lands (UA constant)

	Classified Map	Reference Map (FIA)													
Code	IFMAP	10	12	14	15	16	19	21	22	24	25	28	30	Total	UA %
10	Herbaceous Openland			0		2	12				1			15	0.0
12	Upland Shrub / Low-density tree			0		6	4				0			10	0.0
14	Northern Hardwood Association			33		5	0				2			40	82.5
15	Oak Association			2	4	11	0				0			17	23.5
16	Aspen Association			10	2	21	1				1			35	60.0
19	Pines			1		2	27				4			34	79.4
21	Mixed Upland Conifers			1		0	0				0			1	0.0
22	Upland Mixed Forest			0	0	0	6		7		1			14	50.0
24	Lowland Deciduous Forest			1		2	0				1			4	0.0
25	Lowland Coniferous Forest			0		0	1				50			51	98.0
28	Lowland Shrub			1		4	1				0	7		13	53.8
30	Mixed Non-forest Wetland			0		1	1				0		3	5	60.0
	Total			49	6	54	53		7		60	7	3	239	0.0
	Producer's accuracy %			67.3	66.7	38.9	50.9		100		83.3	100	100		
	KHAT				0.5683	Overall accuracy							63.6%		

Note: PA refers to Producer's Accuracy and UA refers to User's Accuracy.

Table 16. Error Matrix of IFMAP-FIA for State Forest Lands (PA constant)

	Classified Map	Reference Map													
Code	IFMAP	10	12	14	15	16	19	21	22	24	25	28	30	Total	UA %
10	Herbaceous Openland					2	12				1			15	0.0
12	Upland Shrub / Low-density tree					6	4				0			10	0.0
14	Northern Hardwood Association			34		5					2			41	82.9
15	Oak Association			2	5	11								18	27.8
16	Aspen Association			9	2	26	1		1		1			40	65
19	Pines			1		2	28				4			35	80
21	Mixed Upland Conifers			1			0							1	0.0
22	Upland Mixed Forest			0			6				1			7	0.0
24	Lowland Deciduous Forest			1		2					1			4	0.0
25	Lowland Coniferous Forest						1				59			60	98.3
28	Lowland Shrub			1		4	1				0			6	0.0
30	Mixed Non-forest Wetland					1	1				0			2	0.0
	Total			49	7	59	54		1		69			239	
	PA %			69.4	71.4	44.1	51.9		0.0		85.5				
	KHAT	0.5695				Overall accuracy				63.6%					

Note: PA refers to Producer's Accuracy and UA refers to User's Accuracy.

matrix were made only along the row and consequently the user's accuracy of the primary classification were remained the same. For the user's accuracy, the Northern Hardwood (82.5) remained the same and the producer's accuracy of Upland Mixed, Lowland Coniferous and Mixed non-forest Wetland forests were 100% (Tables 7 and 15).

In the aggregated approach the 12X12 matrix was reduced to 6X6, merging the six IFMAP categories to two new forest types, namely Pines, Conifers and Lowland Low-density Forest and Upland hardwood. The first new types were formed merging Pines, Lowland Conifers and Lowland Shrub and Mixed Non-forested Wetland groups and second by merging Aspen, Oak, and Northern Hardwood associations and Mixed Upland Forest (Table 17).

Table 17. Error Matrix of IFMAP-FIA in State Forest Lands (aggregated)

	Classified Map	Reference map (FIA)							
Code	IFMAP	A	B	21	24	10	12	Total	UA %
A	Pines, Conifers, and Low density Lowlands	94	9					103	91.3
B	Upland Hardwoods	11	95					106	89.6
21	Mixed Upland Conifers	0	1					1	
24	Lowland Deciduous Forest	1	3					4	
10	Herbaceous Openland	13	2					15	
12	Upland Shrub / Low-density tree	4	6					10	
	Total	123	116					239	
	PA %	76.4	81.9						
	KHAT	0.628	Over all accuracy		79.1 %				

Note: PA, Producer's accuracy; UA, User's accuracy; and new forest types A= Pines(19)+ Lowland coniferous forest (25)+ Lowland Shrub(28) + Mixed Non-forest Wetland (30) and B = Northern Hardwood (14)+ Oak (15)+ Aspen (16) associations and + Upland Mixed Forest (22).

By merging IFMAP classification the information on these classifications was lost. A number of sample units which were misclassified by IFMAP were now in as matched due to new classification. There were no sample units in four categories in so the cells in these categories are empty. The overall accuracy was 79.1% for the aggregated classification for the state forest land. The user's accuracy of the poorly classified forest types went up and well classified forest types of IFMAP went lower. For example the user's accuracy of Oak Association, a component of Upland Hardwood Forest, went up from 23.5% to 89.6% and Lowland Conifer Forest, a component of Pines, Conifers and Lowlands Low-density Forest, user's accuracy was lowered from 98.0% to 91.3%.

When interpreting the KHAT, the error matrix can be compared to how much better the classification was than the random allocation. The value of KHAT can range from -1 to +1 (Congalton and Green, 1999). To quantify the strength of agreement for the comparison of categorical data, Landis and Koch (1977) described the ranges for KHAT values into several groupings: a value greater than 0.80 (i.e., 80%) almost perfect; a value between 0.61 and 0.80 substantial; a value between 0.41 and 0.60 moderate; and less than 0.40 represents fair and slight agreement. Based on this classification of strength of agreement, there was a moderate level agreement between FIA and IFMAP classification for both UA constant and PA constant approaches (Table 18). However, for the aggregated forest type approach there was substantial agreement between FIA and IFMAP classification.

Table 18. Individual Error Matrix Kappa Analysis Results for State Forest Land

Classification Comparison	N	KHAT	Var (Khat)	Z statistic
IFMAP-FIA UA Constant	239	0.5683	0.001250	16.05882
IFMAP-FIA PA Constant	239	0.5695	0.001251	15.65486
IFMAP-FIA Aggregated Forest Types	239	0.6284	0.001772	14.92579

Note: UA refers to User's accuracy and PA refers to Producer's accuracy.

The comparison of the error matrix using the Z-test confirmed that all the three approaches of classification were significantly different than the random classification. Similarly, the aggregated forest types approach had highest KHAT value; it inferred that this approach of classification was 63% similar to FIA classification. The UA fixed and PA fixed approaches had almost the same KHAT value. The third approach raised the KHAT value due to aggregation of the forest categories; this approach made some of the misclassified sample units in the difference matrix as matched after redefining new categories. The limitation of this approach was that neither the user nor the producer had much information on the components of the merged forest types. The Z statistic values for all classifications are greater than 14 and it confirms that these classifications are better than the random classification.

Kappa analysis can compare two error matrices at a time to determine if they are significantly different. This test was based on the standard normal deviate and the fact that although remotely sensed data are discrete, the KHAT statistic is asymptotically normally distributed (Congalton and Green 1999). At the 95% confidence level, the critical value would be 1.96. Therefore, the results of the pairwise test of three matrices revealed that the comparisons were not significantly different, as the KHAT values of the UA constant and PA constant approaches were very close to the pairwise comparison Z statistics was found small for these two pairs (Table 19). Hence, aggregation improved overall accuracy, but not in a statistically significant sense. However, Z statistic for the comparison between aggregated and PA fixed were large but not enough to make them significantly different at 95% confidence interval. The Z values results were consistent

Table 19. Kappa Analysis Results for the Pairwise Comparison of the Error Matrices

	Z values		
Approach	UA constant	PA constant	Aggregated
UA constant		0.280893	1.102931
PA constant	0.280893		1.358207
Aggregated	1.102931	1.358207	

Note: UA refers to User's accuracy and PA refers to Producer's accuracy.

with pairwise Kappa test (Table 19). The strongest agreement was observed between the PA constant and UA constant approach.

Comparison of Classification Agreement between IFMAP and FIA for National Forest Land

A process identical to one used for state forest lands was employed to analyze national forest lands. Changes in UA and PA reflect cases for which there were multiple IFMAP matches to the FIA forest type (Table 10). For example, the user's accuracy of the Oak Association, Pines and Lowland Deciduous Forest remained the same (47.5%, 82.7%, and 25%, respectively) in both approaches (Tables 20 and 21) and producer's accuracy for the Pines and Lowland Deciduous Forest remained the same (68.3% and 18.2%, respectively). Similar to the state forest land, the aggregation approach increased the overall accuracy of IFMAP in national forest land from 64.8% to 73.6%. By aggregating the IFMAP classification the user's accuracy of Hardwood Forest (new aggregated classification) becomes of 82.8 %, lower than its Northern Hardwood Association component and higher than its Aspen and Upland Mixed Forest components. In this way, the user's accuracy of well matched classified map category may be reduced and the poorly defined classified map category may increase.

Table 20. Error Matrix of FIA-IFMAP in National Forest Lands (UA constant)

Code	Classified Map	Reference Classification (FIA)														
Code	IFMAP	10	12	14	15	16	18	19	20	21	22	24	25	28	30	Total	UA %
10	Herbaceous Openland			1	3			5				0	0	0	0	9	0.0
12	Upland Shrub / Low-density trees					1		1				0	0	0	0	2	0.0
14	Northern Hardwood Association			111		14		0				4	0	0	0	129	86.0
15	Oak Association			1	16	15		3				0	0	0	0	35	45.7
16	Aspen Association			33	4	9		1		2		0	2	0	0	51	17.6
18	Mixed Upland Deciduous							1				0	0	0	0	1	0.0
19	Pines			4	1	8		110	1			1	8	0	0	133	82.7
20	Other Upland Conifers			1		1		8				0	0	0	0	10	0.0
21	Mixed Upland Conifers			1				5				0	0	0	0	6	0.0
22	Upland Mixed Forest				2			21			30	4	1	0	0	58	51.7
24	Lowland Deciduous Forest					5		1				2	0	0	0	8	25.0
25	Lowland Coniferous Forest			2		3		0				0	43	0	0	48	89.6
28	Lowland Shrub			2		5		1				0	0	9	0	17	52.9
30	Mixed Non-Forest Wetland					2		4	1			0	0	0	9	16	56.3
	Total			156	26	63		161	2	2	30	11	54	9	9	523	64.8
	PA %			71.2	61.5	14.3		68.3	0.0	0.0	100	18.2	79.6	100	100		
	KHAT	0.5687				Overall accuracy					64.8%						

Note: PA, Producer's accuracy; UA, User's accuracy.

Table 21. Error Matrix of FIA-IFMAP for National Forest Lands (PA constant)

	Classified Map	Reference Classification (FIA)															
Code	IFMAP	10	12	14	15	16	18	19	20	21	22	24	25	28	30	Total	UA %
10	Herbaceous Openland			1	3			5		0			0			9	0
12	Upland Shrub / Low-density trees			0	0	1		1		0			0			2	0
14	Northern Hardwood Association			128	0	14		0		0		4	0			146	87.67
15	Oak Association			1	16	15		3		0			0			35	45.7
16	Aspen Association			33	6	22		1		0			2			64	34.4
18	Mixed Upland Deciduous			0	0			1		0			0			1	0.0
19	Pines			4	1	8		110		1		1	8			133	82.7
20	Other Upland Conifers			1	0	1		8		0			0			10	0.0
21	Mixed Upland Conifers			1	0			5		0			0			6	0.0
22	Upland Mixed Forest			0	2	0		21		0		4	1			28	0.0
24	Lowland Deciduous Forest			0	0	5		1		0		2	0			8	25.0
25	Lowland Coniferous Forest			2	0	3		0		0			61			66	92.4
28	Lowland Shrub			2	0	5		1		0			0			8	0.0
30	Mixed Non-Forest Wetland			0	0	2		4		1			0			7	0.0
	Total	0	0	173	28	76		161	0	2	0	11	72	0	0	523	0
	PA %			74.0	57.1	28.9		68.3		0.0		18.2	84.7				
	KHAT	0.5548			Overall accuracy					64.8%							

Note: PA Producer's accuracy; and UA User's accuracy.

Table 22. IFMAP-FIA Error matrix for National Forest Lands (Aggregated)

Classified Map — Reference classification (FIA)

Code	IFMAP	10	12	C	15	18	19	20	21	24	D	Total	UA %
10	Herbaceous Openland			1	3		5					9	0.0
12	Upland Shrub / Low-density trees			1			1					2	0.0
C	Hardwood Forest			197	6		22		2	8	3	238	82.8
15	Oak Association			16	16		3					35	45.7
18	Mixed Upland Deciduous				0		1					1	0.0
19	Pines			12	1		110	1		1	8	133	82.7
20	Other Upland Conifers			2			8					10	0.0
21	Mixed Upland Conifers			1			5					6	0.0
24	Lowland Deciduous Forest			5			1			2		8	25.0
D	Lowland Coniferous and Low density Lowlands			14			5	2			60	81	74.1
	Total			249	26		161	3	2	11	71		
	PA %			79.1	61.5		68.3			18.2	84.5		
	KHAT			0.6121	Overall accuracy						73.6%		

Note: PA, Producer's accuracy; UA User's accuracy and New Forest type Hardwood forest = Northern Hardwood Association (14) + Aspen Association (15) + Upland Mixed Forest (22) and Lowland Coniferous and Low Density Lowlands Forest (D) = Lowland Coniferous Forest (25)+ Lowland Shrub (28) + Mixed Non-Forest Wetland (30).

The KHAT value was 0.569 for the UA constant approach and 0.554 for the PA constant approach (Table 23). For both the PA constant and UA constant approaches, the KHAT value was at a moderate level. The KHAT value of the error matrix for aggregated categories was 0.612, a substantial level. The aggregated IFMAP classification was 61.2% better than the random classification.

Table 23. Individual Error Matrix Kappa Analysis Results for National Forest Land

	N	KHAT	Var (Khat)	Z statistics
IFMAP-FIA UA Constant	523	0.568698	0.000602	23.17888
IFMAP-FIA PA Constant	523	0.554815	0.000607	22.51480
IFMAP-FIA Aggregated	523	0.612099	0.000736	22.56409

The comparison of the error matrix with Z-test confirmed that all the three approach of classification were significantly different than the random classification (Table 24).

Table 24. Kappa Analysis Results for the Pairwise Comparison of the Error Matrices

	Z value		
Approach	UA constant	PA Constant	Aggregated
UA constant		0.399238	1.186575
PA Constant	0.399238		1.563060
Aggregated	1.186575	1.563060	

The results of the pairwise test of three matrices for significance at 95% confidence interval revealed that the comparisons were not significantly different. The KHAT value of the UA constant with PA constant approach was close to the pairwise comparison Z statistics. However, the Z statistic for the comparison between aggregated and PA constant were higher, but not enough to make them significantly different at the 95% confidence level.

A previous study carried out by Space Imaging (2004) found the overall accuracy of IFMAP as 67.9% and KHAT value 0.60 for state forest land (N=789). This study found the overall accuracy of IFMAP as 63.6% in State forest land and 64.8% in national forest land. Similarly, the KHAT values were almost 0.56 for both state forest and national forest lands. In both studies the FIA sample plots were used as the reference classification. The previous study also had additional information of "canopy call", a measure of crown cover percentage use. But in this study, robust crosswalk tables were used with some flexibility for including to the low density crown cover forest types of IFMAP.

Due to lack of additional information on the crosswalk table, it is difficult to critically compare the results across studies. For example the user's accuracy of the Oak Association and the Aspen Association in the previous study was 35.6% and 44.9%, and from this study Error Matrix keeping the UA constant user's accuracy for these forest types were 23.5% and 60% in state forest land and 45.7% and 17.6% for national forest land respectively. Both studies conclude that IFMAP had difficulty in classifying the Oak and Aspen associations. Similarly, the overall accuracy of Oak and Aspen associations from the three-way accuracy assessment user's accuracy of IFMAP were 23.5% and 57.1% for state forest land and 45.7% and 15.7% for national forest land, respectively. In this way, IFMAP had difficulty in classifying the Oak Association and Aspen Association. The unavailability of the crosswalk table of previous study between the FIA classification and IFMAP classification precludes more detailed comparisons. Findings from the previous study and this study confirm that the overall KHAT values of IFMAP are in the range of 55% to 60%.

Explanation of Differences in Classifications of IFMAP, OI with FIA

The final study objective focuses on explaining the differences in classification for the state forest lands. Analyses are limited due to the unavailability of plot and stand level characteristics. However, efforts were made to explain the difference in classification based the interpretation of the difference matrix and the personal communications with the FIA unit at NCRS. For a number of forest type categories, the reliability of the IFMAP was very poor (e.g. Oak Association, Aspen Association, Upland Mixed Forest, Mixed Non-forest Wetland, and Lowland Shrub). The findings of this study and the previous study were similar for these forest types. Thus, the satellite imagery classification techniques used for IFMAP classification needs to be further improved to enhance the producer's accuracy of these forest classifications.

The other reason for lower overall accuracy may be due to incompatibility between FIA and IFMAP classifications. To explain the differences in classification between IFMAP and FIA in state forest land, a number of questions were asked to FIA unit at NCRS. The questions clarified the relationships between the FIA plots and IFMAP classes.

The difference matrix of IFMAP and FIA (Table 7) was used to explain the differences in IFMAP with FIA classifications. There was noticeable misclassification error for Jack pine, Black spruce and Aspen of FIA classification into appropriate IFMAP classifications. Aspen Associations of IFMAP wrongly classified inappropriate FIA forest types into this classification.

In the 10 sample locations, Jack pine (FIA) was classified into Herbaceous Openland and in four locations into Upland Shrub/Low-density tree (IFMAP) (Table 7). The observation of FIA plot data for additional details about the age or stocking of trees indicated "all but one of 14 plots is in the small diameter size class. Ages are generally young (5-15 years), but a few were 30 plus. Algorithm typed those stands a variety of types including Northern red oak, other pine/hardwood, post oak/black jack oak, aspen, black cherry" (Holden pers. comm.). Thus, IFMAP and FIA classification approaches were non-compatible to one another for forest classifications of the non-forest types, less than 25% trees canopy cover, defined of IFMAP. The information of 13 plots in small diameter indicates that most of the trees in this location were smaller in size and younger in age and heterogeneous in vegetation. This explanation apparently seems to corroborate the IFMAP classification but not sufficiently confirm that in these locations the tree canopy covers were less than 25%. Thus the differences in defining forest areas and type classifications may have lowered the accuracy for low canopy cover classes of IFMAP.

Four misclassified plots of Jack pine (FIA) were called Upland Mixed Forest (IFMAP). Close examination of FIA plot data and FIA algorithm forest type classes indicated "only one of four plots was typed as 'other pine/hardwood' (409) by the algorithm" (Holden pers. comm.). So, IFMAP clearly misclassified these plots in three of four cases.

In other instances of misclassification, Black Spruce (FIA) was classified into Lowland Shrub or Mixed Non-forest Wetland (IFMAP) in eight sample units. The FIA plot-level information gave an impression that these stands "are between 25 and 70 years old, four of which are over 50 years. Half of stands are in the small diameter size class. The other half is distributed between large and medium. Most stands are medium or poorly stocked" (Holden pers. comm.). This additional information supported that IFMAP was right in 50% of these sample units.

In another example of misclassification of Aspen (FIA) into Oak and Northern Hardwood associations (IFMAP) in 16 sample units, out of total 59 Aspen forest types classified by FIA in this study (Table 7), the FIA plot-level information indicated that "[t]he algorithm typed almost all of these stands as aspen. Two were typed as Oak forest types and one as other pine/hardwood" (Holden pers. comm.). This information only supported IFMAP classification of Oak at two stands of out of total 11 stands; therefore IFMAP misclassified 9 sample stands.

The other misclassification of Aspen forest (FIA) into Herbaceous openland or Upland shrub/low-density (IFMAP) in eight sample units, the FIA plot-level detail information pointed out "four of eight stands are less than 10 years old. A couple of the older stands are poorly stocked. One stand was typed as 'white oak/red oak/hickory by algorithm, also a pretty young stand" (Holden, pers. comm.). In this case, IFMAP was right in half of the stands, however the FIA does not have compatible forest classifications with lower canopy cover or sparsely vegetated forest types. This non-compatibility between the reference and classified map in lower canopy cover of sparsely vegetated categories may have lowered the overall accuracy of IFMAP.

In case of Aspen Associations (IFMAP) misclassified as various oaks or Northern Hardwoods (FIA), the further query of FIA plot-level information implied that "[m]ost of these stands are 50 plus years old and all fall into the large to medium diameter size class. Two of 12 stands are typed as Aspen by the algorithm" (Holden pers. comm.). In this way, out of 12 stands, IFMAP was wrong in 10 stands.

Similar efforts were made to explain the misclassification between the OI and IFMAP (Table 8). In case of misclassification of Jack pine (FIA) into Red pine (OI) in three sample units (Table 8), the FIA plot level information hinted "[o]ne of three plots was half red pine" (Holden pers. comm.). In this way, OI was only right in one of three misclassifications of Jack pine. Similarly in case of misclassification of Jack pine (FIA) into Oak (OI) in four sample stands, the FIA plot-level information supported "Two of four plots had alternative forest type conditions from algorithm. One was other pine/hardwood and the other was oak/ black jack oak" (Holden pers. comm.). In this way OI was completely right in one of four misclassifications and partially right on one additional sample stands. Out of four misclassified Jack pine stands by OI, OI was wrong for two stands.

In addition to misclassification of Jack pine (FIA) into Aspen (OI) in three stands, the FIA plot-level information provided additional details "[o]ne of three plots typed as other pine/hardwood by algorithm. Some aspen on each plot (Holden pers. comm.)." In this way, the FIA algorithm typed another forest type supported the classification of OI in one plot, and the occurrence of aspen partially supported OI classification.

Misclassification of Red pine (FIA) into Jack pine (OI) occurred. The FIA plot-level in three sample units data indicated "[o]ne of three plots typed as jack pine by algorithm and most of plots are composed of jack pine" (Holden pers. comm.). In this case, the OI was completely right in one sample units. This example provided the errors of misclassification of reference forest types by FIA field crew due to the mixed vegetation of Red pine and Jack pine.

In case of misclassification of Black spruce (FIA) into Treed bog and Lowland brush (OI) in eight sample stands, the FIA plot-level data hinted "[s]tand ages on eight plots range from 25-80 and stand size classes (from algorithm) are large to medium and small diameter. Field crew tended to call plots small to medium. Growing stock codes were poor to medium stocking. One was considered overstocked" (Holden, pers. comm.). In this way, these sample stands which were poorly stocked or small diameter, the OI classification may be right. However, the non-compatibility of OI and FIA in forest classification and definition of forest limited the accuracy assessment.

In five sample units, the Sugar maple/beech/yellow birch (FIA) was classified as Aspen by OI. The request for further information from FIA in this misclassification indicated "[s]tand ages on five plots range from about 45-70. Stand size class mostly large diameter. Few aspen on any of the plots" (Holden, pers. comm.). In this way, this explanation revealed that in these plots there were mixes of both sugar maple and aspen, and the plot-level heterogeneity has made it difficult to assign the reference forest classification. As discussed earlier, the relative size of FIA plots and OI stands are different. The relative size of patch size may have biased the two classifications.

The non-compatibility of reference and classified map classification in lower canopy or stocking class of OI and IFMAP, the relative patch size, and the heterogeneity in the sample stands creates difficulty in correctly classifying the reference classification. In addition, the IFMAP is relatively more inaccurate in classifying the Aspen, Oak associations, Upland Mixed Forest and Lowland shrub; this may have aggravated the differences in classification of IFMAP and FIA. These issues may have consequently lowered the consistency of three-way forest classification.

CHAPTER 5

SUMMARY AND CONCLUSION

In Michigan, there are multiple sources of forest management information, developed for specific purposes. There are differences in defining forest, forest types, objectives of their development and details in defining the non-timbered forestland classes, application, and anticipated level of accuracy among these databases and map (Table 6). OI and NRIS-FSVeg are more similar to each other than the other two in terms of their objective, nature of inventory and application. IFMAP is the only database which can provide the "wall-to- wall" information in Michigan. FIA has sample plots throughout Michigan, which are used to statistically assess the status and trends of forests in Michigan. Application of FIA data for sub-state areas produces results with lower levels of accuracy. However, they provide very accurate point-level data for the reference data for this study. IFMAP uses a number of low density cover forest types based on a Level II classification scheme, and aggregates forest classes relative to the finer detail in FIA, OI, and NRIS-FSVeg.

The results of the consistency among FIA, NRIS-FSVeg and IFMAP were calculated by computing the three-way overall accuracy. Based on the sampling information of 523 sample points on national forest lands at the exact FIA plot locations, the overall accuracy was found to be 58.5%. Among the forest classification with a large number of samples, the Lowland Coniferous Forest and Northern Hardwood Association of IFMAP did well in user's accuracy or reliability and the Aspen Association, Oak Association and

Upland Forest did poorly in terms of user's accuracy. These findings were similar to those found in a study conducted by Space Imaging in 2004.

The findings of the consistency among FIA, OI and IFMAP on state forest land revealed that the overall accuracy of three-way classification was 54.8% (N=239). The three-way matching of Lowland Conifer and Northern Hardwood were promising on state forest land. The study indicates the performance in the Oak Association and Aspen Association were not very accurate. However, the strength of this conclusion is limited due to small number of sample plots in these categories.

The objective of comparison of the IFMAP and FIA classifications for national and state forests was limited due to the incompatibility of the two classifications. A crosswalk table was prepared, and based on that the overall accuracy were calculated. The overall accuracy of IFMAP in comparison to FIA on national forest land was 64.8% and on state forest was 63.6%. To compute the KHAT values and assessment, "acceptable" and "probably right" matches had to be re-assigned to other classes. First, the user's accuracy of IFMAP classification was kept constant while developing the error matrix. In this process all of the secondary FIA forest type matches were shifted to the diagonal of the error matrix. Second, the producer's accuracy was kept constant; in this approach, the secondary matches were added to the FIA primary classification to keep the PA constant. Third, all of the IFMAP classifications with matches in multiple categories were merged to new categories. After formulating the error matrix the KHAT values were computed, the KHAT values from the first two approaches were very close, 0.55 and 0.56. For the third approach, using aggregated IFMAP categories, the KHAT value was 0.612 for national forest land and 0.628 for state forest land. The third approach resulted in slightly higher KHAT values due to aggregation of categories. The problem with the third approach is that it gave larger overall accuracy and larger KHAT values at a cost. Specifically, the users will not be aware of the user's accuracy in the category of IFMAP for which information was aggregated.

The major reason for the difference in the classification was due to poor performance of IFMAP in classification of two important forest types, the Oak Association and the Aspen Association. The other reason for the lower overall accuracy and KHAT value was due to the inconsistency in labeling the reference data due to land cover heterogeneity surrounding the sample location. IFMAP has Level II classification for the low canopy cover land use classification, and FIA has a finer detailed forest classification for land with lower canopy densities. For example, regenerating jack pine is Jack pine in FIA plots, but non-forest in the IFMAP classification.

Conclusions

With the creation of IFMAP land cover information, many planners, resources managers and researchers will likely use it for landscape-level studies because it is the "best information available". This study highlights some of the shortcomings of IFMAP, OI and NRIS-FSVeg. These shortcomings should be caveats on the landscape-level studies. The standard of accuracy set forth by the USGS for the generalized first and second levels thematic map is 85% to 90%, IFMAP needs to improve its producer's accuracy to meet this requirement. For some forest types, it meets this standard.

Limitations

The findings of this study are based on the number of small sample plots from the NCRS FIA unit. As the rule of thumb, the number of sample size required for a map covering large area or the classification with large number of forest cover types (i.e. more than 12 categories), the minimum number of samples should be increased to 75 or 100 samples per category (Congalton and Green, 1999). The producer's accuracy, user's accuracy derived for the map category with lower number of samples may or may not be representative of the public forest lands of Michigan.

The findings of this study were based on the FIA field crew defined forest types. The heterogeneity in the sample area may have affected the reliability of the reference map classifications. In this study there was some evidence that supports the heterogeneity in the sample units and inconsistency in defining the reference forest classification (FIA field crew and forest typed by algorithm). FIA reporting of forest resources 1993 as well as 2003 clearly mentions the heterogeneity in FIA sample plots. For example the growing stock of Jack pine forest types only contains about 61% Jack pine, 11% Red pine, 10% other softwoods and other species growing stocks (Table 47 in Leatherberry and Spencer 1996, Table 6 in Leatherberry et al.2005).

In this study the patch size, a smallest unit for which the forest classifications were compared, are different. For FIA patch size is the four subplots; for IFMAP it is a 30m X 30 m pixel; and for the OI and NRIS-FSVeg it is the stand. Previous research found that as the patch size decreases, the accuracy decreases (Smith et al. 2002). Small patches may have affected the accuracy results in this study.

This study was only confined within the public forest lands of Michigan. The accuracy of the IFMAP may be similar for public and private forest lands, but this was not examined due to the public lands focus of this project.

This study was based on the current vegetation types reported on the forest management databases (FIA, OI, NRIS-FSVeg) and IFMAP land-cover 2001. Vegetation classes may change over time, and data from different times may lead to misclassifications. Data in this study were ca. 2000, but from different years.

As there are differences in defining forest and forest types among databases, this may have generated bias on accuracy estimation for the forest classifications with lower canopy cover and between the databases with differences in definition of forest.

In this study, the crosswalk between the reference and IFMAP classifications allowed multiple matches to the "acceptable" and "probably right" classification of IFMAP. This assumption may have over-estimated the accuracy for these forest types which were allowed for multiple matches by the crosswalk tables.

Policy Implications

Results of this study have several policy implications. Some of the important policy implications are noted below.

For state forest lands, OI provides more accurate information than IFMAP land cover, and OI information should be used by policy makers and planners when possible. Similarly, for national forest lands, the NRIS-FSVeg information is more accurate than IFMAP, so priority should be given to using this information. However, for all forest lands, public and private, the only available information is either IFMAP or FIA information.

There are undoubtedly many detailed implications of the reliability of forest management data. A few examples provides ideas regarding some concerns that may arise. User's accuracy-Jack pine and Red pine account for 80% user's accuracy for IFMAP's Pine category (Table 7). By using the IFMAP land cover map for Pine category, users will have non-pine types 20% of the time. If these point data were expanded to the landscape, say 500,000 acres, the actual Pine forest area may be 100,000 acres less. Even this high reliability of Pine could yield significant challenges in the field ranging from issues of short-term timber/habitat availability to long-term sustainability.

Producer's accuracy for Jack Pine is fairly low, 48.6% (Table 7). Over 25% of the Jack Pine plots are classified as Herbaceous Openland by IFMAP. In most cases, these are regenerating stands, but this is not evident from the simple comparison of forest cover; hence, producers must communicate potential problems of this type with users.

Misclassification or confusion among Upland hardwoods is common. For example, 21 Aspen stands are classified as Aspen Associations, but seven Sugar maple/beech/yellow birch stands are also classified as part of the Aspen Association (Table 7). Again these misclassifications can have significant habitat management implications, and users must be aware of these classification issues.

Additional Research

Based on the experiences of this study, the following areas of future study will be useful to expand the understanding of accuracy of the IFMAP land cover raster map.

(1) Study on impact of patch size on accuracy of IFMAP

(2) Study on impact of landscape heterogeneity on accuracy of IFMAP

(3) Study on crosswalk formulation between forest classifications

(4) Compare the acreages classified by the four databases, using FIA expansion factors

(5) Examine experiences from other states in improving raster land cover map accuracy.

For several studies (1), (2) and (4), the research will be greatly enhanced if researchers have direct access to the exact FIA plot locations and plot-level attributes.

REFERENCES CITED

Anderson, J. R., Hardy, E.E., Roach, J.T and Witmer, R.E. 1976. A land use and land cover classification system for use with remote sensor data, (a revision of the land use classification system as presented in U.S. Geological Survey Circular 671), USGS Publications Warehouse *Geological Survey Professional Paper* 964 In: *http://www.wsdot.wa.gov/environment/envinfo/docs/RSPrj_USGS_lulcclass.pdf* (downloaded 06-15-05).

Berlanga-Robels, C. and Ruiz-Luna, A. 2002. Land use mapping and change detection in the coastal zone of northwest Mexico using remote sensing techniques. *Journal of Coastal Research* 8:517-522.

Bishop, Y. M. M., Feinberg, S. E. and Holland, P. W. 1975. Discrete Multivariate Analysis-Theory and Practice: Cambridge, Mass, The MIT Press.

Brand, G. 2005. Forest inventory and analysis sampling hexagons. FIA fact sheet series, USFS. In: http://www.fia.fs.fed.us/library/fact-sheets/data-collections/Sampling_hexagons.pdf (downloaded 06-15-05).

Brand, G.J., Nelson, M.D., Wendt, D.G., and Nimerfro, K. K. 2000. The hexagon/panel system for selecting FIA plots under an annual inventory. In Proceedings of the First Annual Forest Inventory and Analysis Symposium (McRoberts, R.E., Reams, G.A., Van Deusen, P.C., 2000 Eds.) Gen. Tech. Rep. NC-213. St. Paul, MN: USDA Forest Service, North Central Research Station. pp. 8-13.

Burkman, B. 2005. FIA Inventory and Analysis Phase 2 and Phase 3: Ground Measurements. FIA fact sheet series. USFS. In: http://fia.fs.fed.us/library/fact-sheets/datacollections/Phase2_3.pdf (downloaded 06-15-05).

Campbell, J.B. 2002. Introduction to Remote Sensing (third edition). Guilford Publications Inc., New York.

Chuvieco, E. and Congalton R.G. 1988. Using cluster analysis to improve the selection of training statistics in classifying remotely sensed data. *Photogrammetric Engineering and Remote Sensing* 54:1275-1281.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20:37-46.

Congalton, R.G. 1981. The Use of Discrete Multivariate Analysis Techniques for the Assessment of Landsat Classification Accuracy. MS Thesis. Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA.

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

Congalton, R.G. and Green, K. 1993. A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering & Remote Sensing* 59:641-644.

Congalton, R.G. and Green, K.1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers, Boca Raton. 137 p.

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

Foody, G.M. 2003. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80:185-201.

Goodwin, R.F. and Hudson, W.D. 2002. Comparison of Air Photo and Satellite Image Sources for Updating Land Cover and Land Use Maps. Center for Remote Sensing and GIS, Michigan State University June 2002. In: http://www.rsgis.msu.edu/pdf/comparison.pdf, downloaded 06-15-05.In: http://www.globalchange.msu.edu/pdfs/Forecast_Michigan_Project.pdf (downloaded 06-15-05).

Hahn, J.T. and Hansen, M.H. 1985. Data Bases for Forest Inventory in the North Central Region. Gen Tech. Rep. NC- 101. St. Paul, MN: USDA, Forest Service, North Central Forest Experiment Station. 57 p.

Holden Geoff. Forest Inventory and Analysis, North Central Research Station. USDA-FS. St. Paul, MN.

Homer, C., Haung, C., Yang, L., Wylie, B., and Coan, M. 2004. Development of a 2001 national land-cover database for the United States, *Photogrammetric Engineering & Remote Sensing,* 70:829-840. In: http://www.mrlc.gov/pdfs/July_PERS.pdf (downloaded 06-15-05).

Hord, R. M. and Brooner, W. 1976. Land-Use map accuracy criteria. *Photogrammetric Engineering and Remote Sensing* 42:671-677.

Janssen, L.L.F. and Van der Wel, F.J.M. 1994. Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric Engineering & Remote Sensing, 60*: 419-426.

Jennings, M. 1993. Modified UNESCO Natural Terrestrial Cover Classification. Gap Analysis handbook. In: http://www.gap.uidaho.edu/handbook/LandCoverMapping/UNESCO/default.htm (downloaded 06-15-05).

Katila, M., Heikkinen, J. and Tomppo, E. 2000. Calibration of small-area estimates for map errors in multisource forest inventory. *Canadian Journal of Forestry*

Research 30:1329-1339.

Kurvonen, L. and Hallikainen, M.T. 1999. Textural information of multitemporal ERS-1 and JER-1 SAR images with applications to land and forestry type classification in boreal zone. *IIE Trans. Geosci. Remote Sensing* 37:680-689.

Landis, J.R. and Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.

Lawson, Glenn H. n. d. Use of NRI Point Data in Image Processing. USDA, NRCS/NRIA In: http://www.nriai.nrcs.usda.gov/technical/analysis/glenn/ (downloaded 06-16-2005. 2005).

Leatherberry, E.C., Haugen, D, and Brand, G. 2005. Michigan's Forest Resources in 2003. Resour. Bull. NC-245. USDA-FS, NCRS, St. Paul, MN. 39 p.

Leatherberry, E.C., Spencer, J.S, Jr. 1996. Michigan forest statistics 1993. Resour. Bull. NC-170. USDA-FS, NCRS, St. Paul, MN. 144 p.

Liu, C., Zhang, L., Davis, C.J., Solomon, D.S., Brann, T.B., and Caldwell, L.E. 2003. Comparison of neutral networks and statistical methods in classification of ecological habitats using FIA data. *Forest Science* 49:619-631.

Loveland, T.R., and Shaw, D.M. 1996. Multiresolution land characterization: building collaborative partnerships, Gap Analysis: A landscape Approach to Biodiversity Planning (Scoot, J.M., Teat, T., and Davis F. Eds.) *Proceedings of the ASPRS/GAP Symposium*, Charlotte, North Carolina, pp 83-89.

Martin, M. E., Newman, S. D., Aber, J. D., Congalton, and R. G. 1998. Determining forest species composition using high spectral resolution remote sensing data. *Remote Sensing of Environment* 65:249-254.

Mayaux, P., and Lambin, E.F. 1997. Tropical forest area measured form global land cover classifications: inverse calibration models based on spatial textures. *Remote Sensing of Environment* 59:29-43.

McRoberts, R.E. 1999. Joint annual forest inventory and monitoring system, the North Central perspective. *Journal of Forestry* 97:27-31.

MDNR. 2002. Michigan's Growing and Expanding Forest Resource. State of Michigan. Release date April 01, 2002. In http://www.michigan.gov/printerFriendly/0,1687,7-153--23833--,00.html (downloaded on 06-15-05).

MDNR. 2003a. Center for Geographic Information, Department of Information Technology, Michigan (http://www.mcgi.state.mi.us/mgdl/) (downloaded on 06-15-05).

MDNR. 2003b. IFMAP / GAP Lower Peninsula Land Cover, Michigan Spatial Data Library. In: http://www.dnr.state.mi.us/spatialdatalibrary/sdl2/land_use_cover/2001/IFMAP_lp_landcover.htm. (downloaded 06-15-05) 12 p.

MDNR. 2003c. IFMAP / GAP Upper Peninsula Land Cover, Michigan Spatial Data Library. In: http://www.dnr.state.mi.us/spatialdatalibrary/sdl2/land_use_cover/2001/IFMAP_up_landcover.htm. (downloaded 06-15-05) 12 p.

MDNR. 2003d. Decision Rules for classes, Classification system key for remote sensing and ground data gathering. Spatial Data library, MDNR. In: http://www.dnr.state.mi.us/spatialdatalibrary/sdl2/land_use_cover/2001/IFMAP%20Classification%20Rules.htm (downloaded on 06-15-05).

MDNR, 2003e. Description of the Classes Used in the Michigan Statewide Map, Michigan Department of Natural Resources. In: http://www.dnr.state.mi.us/spatialdatalibrary/sdl2/land_use_cover/2001/IFMAP%20TM%20Classification%20Description.htm (downloaded 06-15-05) 4 p.

MDNR. n. d. Operations Inventory Manuals, Michigan Department of Natural Resources, Lansing, MI. (Unpublished Technical Guideline).

Miguel-Ayanz, J.S. and Biging, G.S. 1997.Comparison of single-stage and multi-stage classification approaches for cover type mapping with TM and SPOT data. *Remote Sensing of Environment* 59:92-104.

Miles, P.D., Brand, G.J., Alerich, C.L., Bednar, L.F., Woudenberg, S.W., Glover, J.F., and Ezzell, E.N. 2001. The Forest Inventory And Analysis Database: Database Description and Users Manual Version 1.0. Gen Tech. Rep. NC-218. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 130 p.

MIRIS. 1981. Current Use Inventory Classification System Definitions (May 1981). Michigan Department of Natural Resources. In: http://www.rsgis.msu.edu/pdf/lclu/CUI.pdf (downloaded 06-15-05).

MRIS. 2000. Land Cover/Use Classification System, Version II. MIRIS, Michigan Resource Information System, Resource Mapping and Aerial Photography Section, Land and Mineral Services Division, Department of Natural Resources In: http://www.rsgis.msu.edu/pdf/lclu/MIRIS_2_definitions.pdf (downloaded 06-16-05).

Nelson, R., Case D., Horning, N., Anderson V., and Pillai, S. 1987. Continental land cover assessment using Landsat MSS data. *Remote Sensing of Environments* 21:61-81.

NRIS FSVeg. 2004. Common Stand Exam Users Guide Version 1.6 July 2004. Natural Resource Information System: Field Sampled Vegetation, USDA Forest Service.

NRIS FSVeg. 2005. FSVeg Data Dictionary Version 1.7, January 2005. Natural Resource Information System: Field Sampled Vegetation, USDA Forest Service.

Powell, R.L., Matzke, N., de Souza, C.J., Clark, M., Numata, I., Hess, L.L., and Roberts, D.A. 2004. Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon. *Remote Sensing of Environment* 90:221-234.

Pax-Lenny, M., Woodcock, C.E., Macomber, S.A., Gopal S., Song C. 2001. Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of Environment* 77:241-250.

Pedersen, Larry. 2000. Contrasts of State Forest FIA and OI Data: Can they be reconciled? Michigan Department of Natural Resources (Unpublished report). 9 p.

Pedersen, Larry. Forest Minerals and Fire Management, Michigan Department of Natural Resources. Lansing, MI.

PMR. 2001a. Integrated Forest Monitoring Assessment and Prescription IFMAP, Review of Remote Sensing Technologies for the IFMAP Project, 2nd Revision. (prepared for) State of Michigan, Michigan Department of Natural Resources, Pacific Meridian Resources (PMR), Ann Arbor, Michigan. 78 p. In: http://www.dnr.state.mi.us/spatialdatalibrary/sdl2/land_use_cover/2001/IFMAP%20Remote%20Sensing%20Report.doc (downloaded 06-15-05).

PMR. 2001b. Michigan DNR Integrated Forest Monitoring Assessment and Prescription, Current Field Inventory Design. (prepared for) State of Michigan, Michigan Department of Natural Resources, June, 2001. Pacific Meridian Resources (PMR), A Space Imaging Company, Ann Arbor, Michigan. 56 p.

Rosenfield, G.H. and Fitzpatric-Lins K. 1986. A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering & Remote Sensing* 52:223-227.

RS and GIS. n .d. RS&GIS History in Land Cover/Land Use Issues. Remote sensing and Outreach, Remote sensing and Geographic Information system, Michigan State University, East Lansing, MI, 48823. In: http://www.rsgis.msu.edu/pdf/lclu/lulc_history.pdf (downloaded 06-15-05).

Saatchi, S.S. and Rignot E. 1996. Classification of boreal forest cover types using ASR images. *Remote Sensing of Environment* 60:270-281.

Schmidt, T.L., Spencer, John S.J., and Bertsch, R. 1997. Michigan's forests, 1993: an analysis. Resour. Bull. NC-179. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experimental Station. 96 p.

Schreuder, H.T., Bain, S., Czaplewski, R.C. 2003. Accuracy Assessment of Percent Canopy Cover, Cover Type and Size Class. Gen. Tech. Rep. RMRS-GTR-108. Fort Collins, CO: USDA-FS, Rocky Mountain Research Station. 10 p.

SI. 2004. Integrated Forest Monitoring Assessment and Prescription (IFMAP), Review of Remote Sensing Technologies Used in the IFMAP Project, Final Report. (prepared for) State of Michigan, Michigan Department of Natural Resources (10/25/2004). Space Imaging (SI), Ann Arbor, Michigan.

Silbernagel J., Chen J., Gale, M.R., Pregitzer, K.S., and Probst, J. 1997. An Interpretation of Landscape Structure from Historic and Present Land Cover Data in the Eastern Upper Peninsula of Michigan. Gen. Tech. Rep. NC-192. St. Paul, MN: USDA-FS, North Central Forest Experiment Station. 30 p.

Skole, D.L., Batzli, S., Gage, S., Pijanowski, B., Chomentowski, W., Zhou, Y., Rustem, W., and Salas W. 2002. Land Use Issues and Policy: Forecasting Land Use Change in Michigan, August 2002. A Forecast Michigan Project, Center for Global Change and Earth Observations, Michigan State University.

Smith, J.H., Wickham, J.D., Stehman, S.V., Yang, Limin. 2002. Impacts of patch size and land-cover heterogeneity on thematic image classification accuracy. *Photogrammetric Engineering & Remote Sensing* 68:65-70.

Smith, W. B., Miles, P.D., Vissage, J.S., and Pugh, S.A. 2004. Forest Resources of the United States, 2002. Gen. Tech. Rep. NC-241. St. Paul, MN: USDA, Forest Service, NCRS. 137 p.

Stehman, S. V. 1999. Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing* 20:2423-2441.

Story, M. and Congalton, R.G. 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering & Remote Sensing* 52:397-399.

USDA-FS. 2004. Forest Inventory and Analysis National Core Field Guide Volume I: Field Data Collection Procedures For Phase 2 Plots Version 2.0. In: http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/core_ver_2-0_01_04.pdf (downloaded 06-05-05).

Wang, C., Qi, J., Skole, D.L., Cochrane M.A. and Matricardi, E.A.T . 2003. Tropical Forest Degradation and Recovery: Estimation and Validation of Canopy Cover with Landsat and IKONOS Imagery. CGCEO/RA01-03/w Michigan State University, East Lansing, Michigan.

Wessels K. J., De Fries, R.S., Dempewolf, J., Anderson, L.O., Hansen, A.J. Powell, S.L., Moron, E.F. 2004. Mapping regional land cover with MODIS data for biological conservation: examples from the Greater Yellowstone Ecosystem, USA and Pará State, Brazil. *Remote Sensing of Environment* 92:67-83

Wickham, J.D., Sethman, S. V., Smith J. H., and Yang, L. 2004. Thematic accuracy of the 1991 National Land-Cover Data for the western United States. *Remote Sensing of Environment* 91: 452-468.

Wynne, R.H., Oderwald, R.G., Reams, and Scrivani, J. A. 2000. Optical remote sensing for forest area estimation. *Journal of Forestry* 98:31-36

APPENDIX

APPENDIX A

Table A1. NRIS-FSVeg FIA Crosswalk

FSVeg Code	NRIS-FSVeg Classification	FIA CODE	FIA Classification
1	Jack Pine	101	Jack pine
2	Red Pine	102	Red pine
3	White Pine	103	Eastern white pine
5	Hemlock	105	Eastern hemlock
11	Balsam Fir_Asp-PB	121	Balsam Fir
12	Black Spruce	125	Black Spruce
14	Northern Wh Cedar	127	Northern white Cedar
18	Mix Swamp conifer	127	Northern white Cedar
18	Mix Swamp conifer	125	Black Spruce
19	Cedar-Aspen-PB	127	Northern white Cedar
21	Mixed Northeren Hdw	520	Mixed Upland Hdwd
48	Jack Pine-Oak	101	Jack pine
49	Red pine-Oak	102	Redpine
53	Black Oak	515	Chestnut oak /black oak/scarlet
55	Northern Red Oak	503	White oak / red oak / hickory
59	Mixed Oak	503	White oak / red oak / hickory
59	Mixed Oak	504	White oak
59	Mixed Oak	505	Northern red oak
71	Bl Ash-Elm-R Maple	701	Black ash/ American elm / red maple
81	S Maple-Beech -YB	800	Maple / Beech / Birch Group
81	S Maple-Beech -YB	801	Sugar maple / beech/ yellow birch
82	S Maple-Basswood	801	Sugar maple / beech/ yellow birch
82	S Maple-Basswood	805	Hard maple / basswood
84	Red Maple(Dry)	809	Red maple upland
85	Sugar Maple	800	Maple / Beech / Birch Group
85	Sugar Maple	801	Sugar maple / beech/ yellow birch
85	Sugar Maple	805	Hard maple / basswood
87	Sugar maple-beech-yellow birch-red spruce	801	Sugar maple/ beech/ yellow birch
89	Mixed Upland Hdwd	520	Mixed Upland Hdwd
89	Mixed Upland Hdwd	801	Sugar maple / beech/ yellow birch
89	Mixed Upland Hdwd	805	Hard maple / basswood
89	Mixed Upland Hdwd	809	Red maple / upland
90	Mixed Upland Hdwd	805	Hard maple / basswood
91	Quaking Aspen	901	Aspen
93	Bigtooth Aspen	901	Aspen
95	Asp-W Spruce BF	901	Aspen
97	Lowland Brush	125	Black spruce
99	Open	999	Nonstocked

APPENDIX A

Table A2. IFMAP-FIA Crosswalk

IFMAP code	IFMAP Classification	FIA Code	FIA Classification
10	Herbaceous Openland	999	Nonstocked
12	Upland Shrub /Lowdensity trees	999	Nonstocked
14	Northern Hardwood Association	105	Eastern hemlock
14	Northern Hardwood Association	800	Maple / Beech / Birch Group
14	Northern Hardwood Association	801	Sugar maple / beech / yellow birch
14	Northern Hardwood Association	805	Hard maple / basswood
14	Northern Hardwood Association	809	Red maple / upland
15	Oak Association	503	White oak / red oak / hickory
16	Aspen Association	901	Aspen
19	Pines	101	Jack pine
19	Pines	102	Red pine
21	Mixed Upland Conifer	101	Jack pine
21	Mixed Upland Conifer	102	Red pine
21	Mixed Upland Conifer	103	Eastern white pine
21	Mixed Upland Conifer	400	Oak Pine Group
22	Upland Mixed Forest	503	White oak / red oak / hickory
22	Upland Mixed Forest	801	Sugar maple / beech / yellow birch
22	Upland Mixed Forest	901	Aspen
22	Upland Mixed Forest	800	Maple / Beech / Birch Group
22	Upland Mixed Forest	809	Red maple / upland
21	Mixed Upland Conifer	401	White pine / Red oak / white ash
24	Lowland Deciduous Forest	701	Black ash / American elm / red maple
25	Lowland Coniferous Forest	121	Balsam fir
25	Lowland Coniferous Forest	125	Black spruce
25	Lowland Coniferous Forest	127	Northern white-cedar
28	Lowland Shrub	125	Black spruce
28	Lowland Shrub	127	Northern white-cedar
30	Mixed Non Forest Wetland	127	Northern white-cedar
30	Mixed Non Forest Wetland	125	Black spruce

APPENDIX A

Table A3. OI-FIA Crosswalk

OI Code	OI Classification	FIA Code	FIA Classification
A	Aspen	901	Aspen
C	Cedar	127	Northern White-cedar
D	Treed Bog	125	Black spruce
E	Swamp Hrdwds	700	Elm / Ash / Cottonwood Group
E	Swamp Hrdwds	701	Black ash / American elm / red maple
F	Spruce fir	121	Balsam fir
F	Spruce fir	125	Black spruce
J	Jack pine	101	Jack pine
L	Lowland Brush	125	Black spruce
M	Northern Hardwood	520	Mixed upland hardwoods
M	Nothern Hardwood	801	Sugar maple / beech / yellow birch
M	Nothern Hardwood	805	Hard maple / basswood
M	Nothern Hardwood	809	Red maple / upland
O	Oak	503	White oak / red oak / hickory
Q	Mx Swamp Cnfr	121	Balsam fir
Q	Mx Swamp Cnfr	125	Black spruce
Q	Mx Swamp Cnfr	126	Tamarack
Q	Mx Swamp Cnfr	127	Northern White-cedar
R	Red pine	102	Red pine
S	Black Spruce	125	Black spruce
T	Tamarack	126	Tamarack

APPENDIX A

Table A4. IFMAP-OI Crosswalk

IFMAP code	IFMAP Classification	OI Code	OI Classification
12	Upland Shrub/low density trees	U	Upland Brush
14	Northern Hardwood Association	M	Nothern Hardwood
15	Oak Association	O	Oak
16	Aspen Association	A	Aspen
19	Pines	J	Jack pine
19	Pines	R	Red pine
21	Mixed Upland Conifers	W	White pine
22	Upland Mixed Forest	A	Aspen
22	Upland Mixed Forest	M	Nothern Hardwood
25	Lowland Coniferous Forest	C	Cedar
25	Lowland Coniferous Forest	D	Treed Bog
25	Lowland Coniferous Forest	Q	Mixed Swamp Conifer
25	Lowland Coniferous Forest	S	Black Spruce
25	Lowland Coniferous Forest	T	Tamarack
28	Lowland Shrub	S	Black Spruce
28	Lowland Shrub	L	Lowland Brush
28	Lowland Shrub	Q	Mixed Swamp Conifer
30	Mixed Non-Forest Wetland	L	Lowland Brush
30	Mixed Non-Forest Wetland	Q	Mixed Swamp Conifer

Appendix A

Table A5. Crosswalk IFMAP-NRIS-FSVeg

(a)

IFMAP code	IFMAP Classification	FSVeg	NRIS-FS Veg Classificaiton
10	Herbaceous Openland	999	Nonstocked
12	Upland Shrub /Lowdensity trees	999	Nonstocked
14	Northern Hardwood Association	5	Hemlock
14	Northern Hardwood Association	81	S Maple-Beech -YB
14	Northern Hardwood Association	82	S Maple-Basswood
14	Northern Hardwood Association	84	Red maple(dry)
14	Northern Hardwood Association	85	Sugar maple
14	Northern Hardwood Association	85	Sugar maple / beech / yellow birch
14	Northern Hardwood Association	85	Sugar Maple
14	Northern Hardwood Association	90	Mixed upland hardwoods
15	Oak Association	59	Mixed Oak
16	Aspen Association	11	Balsam Fir_Asp-PB
16	Aspen Association	91	Aspen
16	Aspen Association	93	Bigtooth Aspen
16	Aspen Association	95	Asp-W Spruce BF
19	Pines	1	Jack pine
19	Pines	2	Red pine
19	Pines	3	White pine
19	Pines	48	Jack Pine-Oak
19	Pines	49	Red Pine-Oak
20	Other Upland Conifer	5	Hemlock
20	Other Upland Conifer	11	Balsam fir-Asp-PB
20	Other Upland Conifer	8	White spruce
20	Other Upland Conifer	19	Cedar-Aspen-PB
21	Mixed Upland Conifer	3	White pine
22	Upland Mixed Forest	5	Hemlock
22	Upland Mixed Forest	48	Jack Pine-Oak
22	Upland Mixed Forest	53	Black Oak
22	Upland Mixed Forest	59	Mixed Oak
22	Upland Mixed Forest	81	S Maple-Beech -YB
22	Upland Mixed Forest	85	Sugar Maple
22	Upland Mixed Forest	87	Sugar maple-beech-yellow birch/red spruce
22	Upland Mixed Forest	89	Mixed Upland Hdwd
22	Upland Mixed Forest	91	Quaking Aspen
22	Upland Mixed Forest	93	Bigtooth Aspen
22	Upland Mixed Forest	95	Asp-W Spruce BF
24	Lowland Deciduous Forest	71	Bl Ash-Elm-Red maple
24	Lowland Deciduous Forest	91	Quaking Aspen
25	Lowland Coniferous forest	11	Balsam fir-Asp-PB
25	Lowland Coniferous forest	12	Black spruce

Table 5. (Cont'd).
(b)

25	Lowland Coniferous forest	14	Northern Wh Cedar
25	Lowland Coniferous forest	18	Mix Swamp Conifer
25	Lowland Coniferous forest	19	Cedar-Aspen-PB
28	Lowland Shrub	12	Black Spruce
28	Lowland Shrub	14	Northern Wh Cedar
28	Lowland Shrub	18	Mix Swamp Conifer
28	Lowland Shrub	97	Lowland Brush
30	Mixed Non-Forest Wetland	12	Black spruce
30	Mixed Non-Forest Wetland	18	Mix Swamp Conifer
30	Mixed Non-Forest Wetland	97	Lowland Brush
30	Mixed Non-Forest Wetland	99	Open