The DISCover Validation Image Interpretation Process

Joseph Scepan, Gunter Menz, and Matthew C. Hansen

Abstract

Thematic validation of the International Geosphere Biosphere Data and Information System (IGBP-DIS) Global 1-Kilometer Land-Cover Data Set (DISCover) was performed utilizing a "state-of-the-practice" technique by a team of Expert Image Interpreters (EII) examining subscenes extracted from 379 digital Landsat TM (Thematic Mapper) and SPOT images.

The 15 validated IGBP land-cover classes (Snow/Ice and Water were not assessed) were not equally interpretable on the TM and SPOT imagery. Interpreter confidence was highest for Evergreen Broadleaf Forests and Urban/Built-up DISCover classes while Grasslands and Permanent Wetlands were interpreted with relatively less confidence. Analysis of image interpretation in each of the 13 validation regions indicates that confidence in interpretations for North America/Canada (Region 1) and Central Asia/Japan (Region 11) are lower than average. Confidence in interpretations is significantly higher than average for North America/US (Region 2), Northern and Southern South America (Regions 4 and 5), and Southest Asia and China (Region 12). In this study, variations in interpretation confidence are also noted between regions or based upon the geographic location of samples.

This exercise demonstrates that Landsat TM and SPOT imagery can be efficiently used to validate high-resolution global land-cover products. The results suggest that the utility of and confidence that may be placed in this technique depends upon the land-cover classification scheme used and the quality of digital and ancillary data available to aid interpreters. Another important factor is the relative confidence of the interpreters to verify the land cover within their respective areas of the globe.

Introduction

State-of-the-practice statistical analysis of the thematic accuracy of geospatial products calls for the use of manual interpretation of higher resolution imagery in the validation process (Fitzpatrick-Lins, 1980; Rosenfield *et al.*, 1981; Belward, 1996). The assumption is that the data product to be validated is accurate to a certain level and that the constituent land-cover classes are unambiguously interpretable from the higher resolution images. An assumption is also made that the validation interpreters are experts with experience in interpreting the classes represented within the thematically classified product for their specific area of the globe.

With these assumptions in mind, the International Geosphere Biosphere Programme Data and Information System (IGBP-DIS) Land-Cover Working Group (LCWG) and the Validation Working Group (VWG) identified a cadre of Expert Image Interpreters (EII) for verification of samples drawn from IGBP-DIS global land-cover data set, called DISCover 1.0. These interpreters provided expert analyses of the thematic classes represented in DISCover 1.0 from Landsat Thematic Mapper (TM) and SPOT (System Probatoire de la Observation de la Terre) imagery. In order to identify individuals capable of serving as EIIs, the IGBP LCWG and VWG first identified a group of Regional Validation Advisors to provide guidance on EII selection.

The material that follows provides a background on the DISCover Validation Image Interpretation effort. The methodology employed in this research and a discussion of the interpretation procedure follows. Results are discussed, as are comments from the EIIs during the Validation Workshop and, specifically, questions posed in writing by VWG personnel. Findings are presented and conclusions and recommendations made.

Background

The thematic validation of the IGBP DISCover data set was performed by a team of EIIs examining subscenes extracted from 432 digital Landsat TM and SPOT images. This approach has been demonstrated to be accurate and reproducible (Jensen, 1996; Vogelmann *et al.*, 1998). The protocol for validation of DISCover was developed by the IGBP VWG and reviewed and approved by the LCWG (Belward,1996). The validation procedure was tested prior to the IGBP Global Validation Workshop (GVW) in a preliminary evaluation session held at the University of California, Santa Barbara (UCSB) during February 1998. The GVWwas held during 07–18 September 1998 in the Donald Moore Training Facility at the U.S. Geological Survey EROS Data Center (USGS/EDC) in Sioux Falls, South Dakota.

Expert Image Interpreters

For validation purposes, the Earth's land surface was segmented into 13 regions. This regionalization was based loosely on the IGBP START (Global Change System for Analysis, Research and Training) regions. Validation metadata were compiled for each sample. Three-hundred seventy-nine core samples in 15 of the 17 IGBP Land-Cover Classes were verified.

Researchers and other personnel from institutes around the globe participated in this effort as Regional Validation Advisors, Cooperating Laboratories, and Expert Image Interpreters (EII). Regional Advisors assisted in the selection of the EIIs. The Regional Advisors were selected to participate in the

J. Scepan is with the Remote Sensing Research Unit, Department of Geography, University of California, Santa Barbara, CA 93106-4060 (scepan@goeg.ucsb.edu)

G. Menz is with the Remote Sensing Research Group, Department of Geography, University of Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany.

M.C. Hansen is with the Laboratory for Global Remote Sensing Studies, Department of Geography, University of Maryland at College Park, College Park, MD 20742.

Photogrammetric Engineering & Remote Sensing
Vol. 65, No. 9, September 1999, pp. 1075–1081.

0099-1112/99/6509–1075$3.00/0

GVW based upon their demonstrated knowledge of remote sensing. Of particular importance was experience working with personnel and facilities with expertise in land-cover analysis for a given region of the globe. Regional Advisors worked with the VWG to develop a list of potential EIIs for each GVW Region. The objective was to recruit for each region at least one EII who had native or first-hand knowledge. Three EIIs were identified for each region (in two cases, individual EIIs served in two regions). Table 1 contains the names of the EIIs for each GVW Region.

In order to accommodate a number of technical and logistical considerations, the Global Validation Workshop was conducted in two sessions. Session 1 was held during 07–11 September and included Validation Regions 3 (Central America), 4 (Northern South America), 7 (North Africa), 8 (Central Africa), 9 (Southern Africa), 11 (Central Asia/Japan), and 12 (Southeast Asia/China). Session 2 included Validation Regions 1 (North America/Canada), 2 (North America/US), 5 (Southern South America), 6 (Europe), 10 (Russia), and 13 (Australia/New Zealand). This group met during 14–18 September.

Landsat TM and SPOT spectral imagery covering each of the globally distributed core sample sites were used as the validation interpretation source data. Both TM and SPOT data have been used extensively to classify and map a broad suite of land-cover types, including those contained in the DISCover data set (Ali, 1989; Franklin, 1991; Cherrill *et al.*, 1994). Landsat TM data were provided by the USGS EROS Data Center. The European Commission Joint Research Center (EC/JRC), the Centre Nationale Etudes Spatial (CNES), and the IGBP Secretariat acquired the SPOT data for use in the GVW.

Methodology

High-resolution image interpretation was accomplished using digital imagery displayed on a video monitor. Each core sample was presented to the interpreter centered in a 40- by 40-square-kilometer image subscene. The 40- by 40-km size was selected in order to provide the interpreter with sufficient regional land-cover area to provide context for each sample location. Each TM subscene was provided to the EIIs as complete six-band data, while SPOT imagery was provided in three bands. TM imagery has the spectral bands shown in Table 2. For the GVW, the false color band combination of 4-5-3 (RGB) was used as the display default for TM data. After evaluation and review during the initial validation test, this band combination was chosen for initial display over the more common 4-3-2 (RGB) TM band combination, principally due to perceived enhanced image visual characteristics, particularly for analysis of subtle vegetation variations in shrub land, savanna, grassland, and wetland areas. However, during the validation, EIIs made frequent use of alternate band combinations felt to be most appropriate for each region or land-cover regime, and also made selective use of band stretches (e.g., Gaussian) for image enhancement.

TABLE 1. GVW EXPERT IMAGE INTERPRETERS

Region	EII Name
1	G.K. Eulert, Alisa Gallant, Jim Vogelmann
2	Matthew Hansen, Gabriela Gamez Rodriguez, Jim Vogelmann, John Estes
3	John Foster, Doug Muchoney, James C. Smoot
4	Doug Muchoney, Jim Verdin, Eric Waller
5	David Cunningham, Juan Carlos Salazar Lea Plaza, Carlos Scoppa
6	Allard de Wit, Gunter Menz, Fabio Vescovi
7	David Cunningham, Fabio Vescovi, Bruce Wylie
8	Matt Hansen, Massaer Mbaye, Gunter Menz
9	Sharon Gomez, Dominick Kwesha, Mark Thompson
10	Koji Kajiwara, Wayne Miller, Yuzo Suga
11	Shi Hua, Koji Kajiwara, Charles R. Larson, Mikiyasu Nakayama
12	Chandra Prasad Giri, Laili Nordin, Hans-Jurgen Stibig
13	Susan Maxwell, Peter Newsome, Philip Tickle

TABLE 2. LANDSAT THEMATIC MAPPER AND SPOT HRV VERIFICATION IMAGERY

Instrument Spectral Band	Spectral Region (Micrometers)	GVW Interpretation Subscene
	Landsat Thematic Mapper	
1	0.45-0.52	
2	0.52-0.60	
3	0.63-0.69	Displayed in blue
4	0.76-0.90	Displayed in red
5	1.55-1.75	Displayed in green
7	2.08-2.35	
6	10.40-12.50	
	SPOT HRV	
1	0.50-0.59	Displayed in blue
2	0.61-0.68	Displayed in green
3	0.79-0.89	Displayed in red

The spectral characteristics of SPOT XS image data are also shown in Table 2. SPOT data were provided to EIIs in 3-2-1 (RGB) default band combination although, as with the Landsat TM data, EIIs were free to employ alternate band combinations as appropriate.

All verification image interpretation activities were performed using commercially available hardware and image processing and geographic information system software packages. The USGS/EDC Donald Moore Training Facility includes IBM compatible personal computers supporting the Microsoft Windows-NT operating system. Raster/vector image processing activities for the GVW were performed using ERDAS/Imagine (Version 8.3.1). Direct support for the GVW was received from ERDAS, Incorporated, of Atlanta, Georgia in the form of 20 temporary Imagine licenses for use on the EDC computers during the exercise.

Interpreter Procedure

The DISCover data set includes 17 land-cover classes (Loveland *et al.*, 1999, in this issue). Fifteen of these classes were validated. Samples for Classes 15 (Snow and Ice) and 17 (Water) were not verified. The IGBP validation protocol specified that a minimum of 25 primary samples points be verified for each of the 15 classes. Due principally to the sparse satellite image coverage, only 11 core sample points were acquired for Deciduous Needleleaf Forests (Class 3) and 17 for Permanent Wetlands (Class 11). Each primary sample point had two associated secondary sample points included in a 40- by 40-km^2 Validation Subscene, oriented north/south and east/west from the primary point and separated by 20 km. Each sample point corresponded to a 1-km DISCover data set pixel. The validation procedure required that each sample point be verified by three EIIs, each working independently. A majority rule was used for determining the correctness of each sample.

Within each region, all primary and secondary samples were displayed and verified by each EII. A majority rule was employed to determine the accuracy of each sample: at least two of the three EIIs were required to independently verify that a given sample was classed correctly in order for it to be counted as correct. A standard hardcopy sample verification form was used by EIIs to record relevant information, including EII name, date, interpreted class, and interpretation confidence for each sample set.

Sixty hardcopy keys produced from TM imagery, representing 15 land-cover classes in ten regions, were available for interpreter use (Kelly *et al.*, 1999, in this issue). These keys demonstrated only one phenology for each class. Classes in the remaining three regions were represented on SPOT imagery for which, due to logistical reasons, hardcopy keys were not available.

The styles of interpretation and levels of familiarity with their prospective regions varied tremendously among the interpreters. Where some interpreters preferred to remain unbiased and would not make use of ancillary references, others prefer to collect as much reference material as possible before making a decision. A number of maps and atlases were provided as references by the EDC library for interpreter use during the GVW.

Results

The raw thematic accuracy of DISCover 1.0 was compiled for each of the 15 individual classes verified. Overall area-weighted accuracy was also computed for the data set. Raw by-class accuracy varied from a low of 0.294 for permanent wetlands (Class 11) to 1.00 for Barren (Class 16). The lowest by-class accuracy for any class with more than 25 samples verified was deciduous broadleaf forest (Class 4) at 0.40. For statistical analyses that are based on a stratified random sample, the area-weighted statistic is the appropriate measure of overall data set accuracy. The area-weighted accuracy for the DISCover 1.0 product is 66.9 percent. For several classes that cover large areas (Class 16 Barren and Class 7 Open Shrublands), accuracy is relatively high. A more detailed discussion of the thematic accuracy of the DISCover 1.0 data set is found in Scepan (1999, in this issue)

In order to consider inconsistencies in individual interpretations for the accuracy analysis, an additional analysis was performed. This procedure applied the majority interpretation rule to the core samples that were verified as in error as well as those verified as correct. If a sample was verified as incorrect, and the three EIIs did not agree on the correct class, the sample was not included in the accuracy assessment. Although this approach reduces the number of samples that are included in the analysis, it removes an element of verification confusion by excluding those samples which could be considered "uninterpretable" in the sense that there was no EII agreement on their proper class; 73 of the total 154 incorrect are thus removed from the statistical analysis. Using this method, individual class accuracies of DISCover 1.0 range between 38.5 percent and 100 percent. Calculating the average of all the class accuracies, the raw accuracy of the data set is 73.5 percent (225/306). Aggregating the class accuracies using an area-weighted calculation results in an overall DISCover 1.0 data set accuracy of 77.6 percent.

The overall quality of the DISCover validation effort was the direct result of the interpretation of the high-resolution verification data. The increase in accuracy (both for individual classes and for the overall data set) which results from employing the majority rule for errors as well as correct samples is a direct result of interpreter agreement and consistency.

In order to compare the variation in regional accuracy, a compilation of EII interpretations was also performed for each of the 13 regions as summarized in Table 3 and shown in Plate 1. North Africa, Central Africa, and Central America scored the highest accuracies (over 70 percent), while North America/Canada and Southern South America had the lowest accuracies (below 30 percent).

The results of the core sample verification were evaluated by conducting a quantitative analysis of the interpretation confidence metrics compiled by the EIIs during the GVW. Qualitative responses of the EII individual reviews were also analyzed.

EII Confidence Analysis

As a part of the GVW interpretation process, each EII placed a confidence value on the interpretation of each core sample. EIIs rated their confidence in each interpretation as Low, Medium, or High. Quantifying the level of overall interpretation confidence as well as the mean interpretation confidence levels for each land-cover class and each region provide useful measures of the utility of the validation methodology. For analytic purposes, the descriptive confidence metrics have been transformed to numerical values: 1 (Low Confidence), 2 (Medium Confidence), and 3 (High Confidence).

Confidence metrics were compiled by class in order to identify the IGBP classes which the EIIs believed were relatively more or relatively less interpretable (Table 4 and Figure 1). The overall interpretation confidence value of 2.23 demonstrates a group interpretation confidence level between medium and high. Only two classes (Class 10 - Grasslands, and Class 11 - Permanent Wetlands) have average EII interpretation confidence values below 2.0, indicating relatively low EII confidence. All other class values fall between 2.0 and 3.0, with confidence levels for Class 2 (Evergreen Broadleaf Forests; 2.71), Class 12 (Croplands; 2.47), Class 13 (Urban and Built-up; 2.66), and Class 16 (Barren; 2.49) being highest.

To identify any regional differences in EII interpretation confidence, ratings were also compiled by IGBP Validation Region (Table 5). Average EII confidence values are below 2.0 (medium) confidence for only Regions 1 (North America/Canada) and 11 (Central Asia/Japan). This indicates that EIIs felt that improvements were needed for the validation methodology and/or the data available for use in these areas. Confidence levels for Region 2 (North America/United States; 2.51), Region 4 (Northern South America; 2.58), Region 5 (Southern South America; 2.52), and Region 12 (Southeast Asia/China; 2.55) were highest

It is important to emphasize that these confidence metrics are subjective. They are, however, useful in gaining perspective on the relative utility of the validation procedure and data sets among IGBP Classes or GVW Regions.

Following the analysis of EII interpretation confidence by region and DISCover class, the data were analyzed in order to determine whether there was a correlation between the verified individual DISCover class accuracies and the EII confidence. The data were plotted and a correlation coefficient was determined. This analysis shows a positive relationship between DISCover 1.0 class accuracy and EII interpretation confidence for the DISCover classes. Discover classes which were interpreted as most accurate also tended to be verified by EIIs with higher confidence ($r = 0.460$; as shown in Figure 1). We thus conclude that there is a general relationship between the efficacy of the DISCover classification procedure and the ability of the analyst to confidently interpret and verify samples from the data set.

EII Response and Comments

Discussions were held among IGBP participants and EIIs to incorporate input from regional experts into the global validation process. Several meetings regarding the validation procedures were held between individual EIIs and GVW participants concurrent with Core and Confidence validation activities, and summary discussions were held among all EIIs and GVW participants at the conclusion of each Validation session. Finally, a short written questionnaire was provided to each EII to allow them the opportunity to provide written comments (anonymously, if desired) about the Validation Workshop and Core sample verification procedures.

The responses and comments provided by the EIIs agreed generally that the effort of the global validation workshop was a valid and positive experiment. As a first attempt, the overall effort seemed to be very promising and, at the same time, very challenging. It was the general consensus that team of EIIs would gladly contribute their expertise again in subsequent activities related to global land-cover validation.

Additional questions were asked regarding the Confidence Site Mapping exercise, and general comments were also solic-

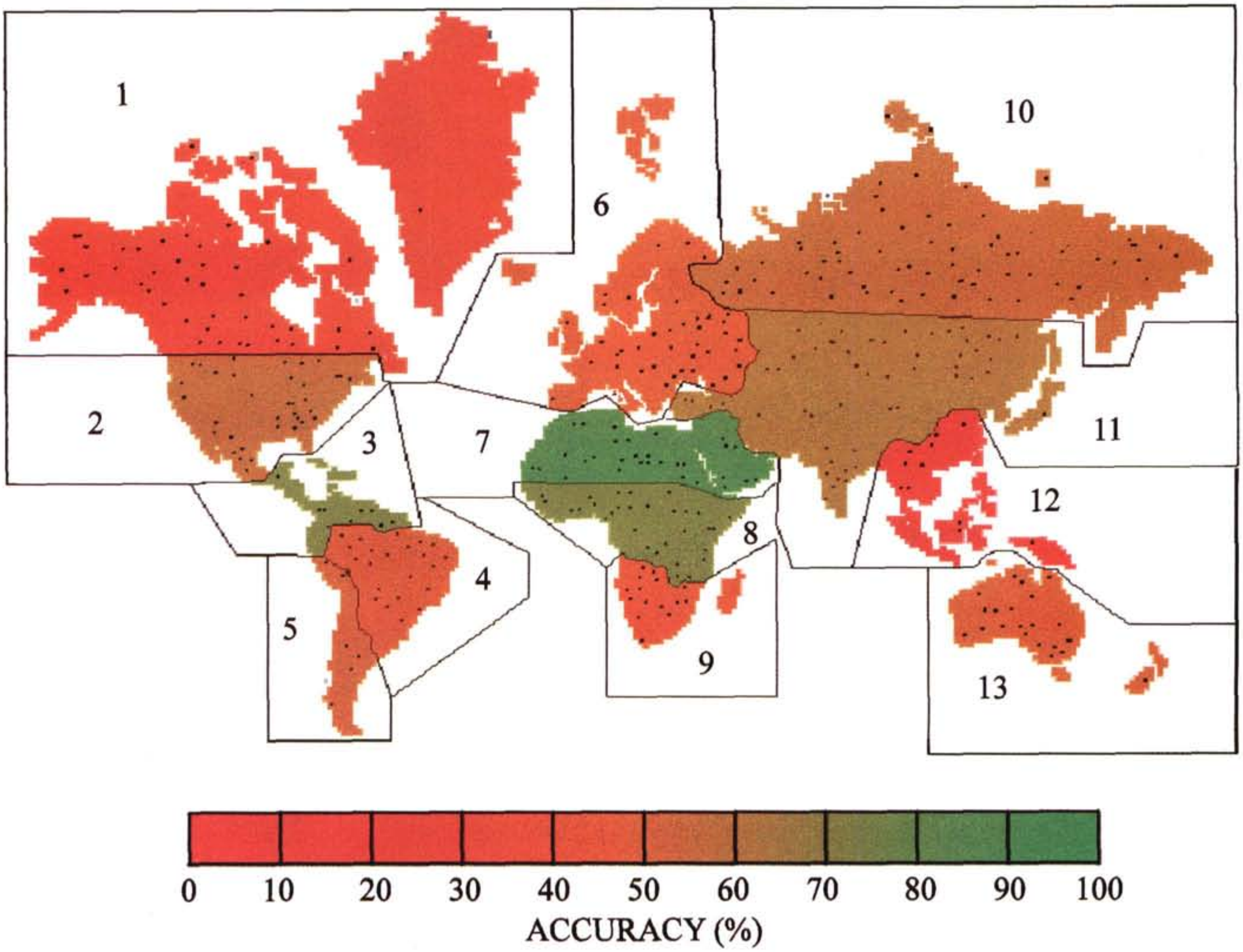

Plate 1. Regional accuracy of the core validation sample.

Table 3. IGBP DISCover 1.0 Global Validation Summary by Region

Region	Verified Samples	Samples Verified Correct	Region Accuracy (%)
1-North America/Canada	41	12	29.3
2-North America/US	69	38	55.1
3-Central America	14	10	71.4
4-Northern South America	34	16	47.1
5-Southern South America	6	3	50.0
6-Europe	45	21	46.7
7-North Africa	28	24	85.7
8-Central Africa	51	36	70.6
9-Southern Africa	14	6	42.9
10-Russia	21	12	57.1
11-Central Asia/Japan	58	35	60.3
12-Southeast Asia/China	13	2	15.4
13-Australia New Zealand	15	8	53.3
Total	409	225	54.5

Table 4. Discover EII By-class Interpretation Confidence

Class	EII1	EII2	EII3	Class Average
1-Evergreen Needleleaf Forests	2.30	2.04	2.15	2.16
2-Evergreen Broadleaf Forests	2.59	2.88	2.67	2.71
3-Deciduous Needleleaf Forests	2.09	2.00	2.36	2.15
4-Deciduous Broadleaf Forests	2.37	2.38	2.21	2.32
5-Mixed Forests	2.32	1.89	2.04	2.08
6-Closed Shrublands	2.11	2.30	2.11	2.17
7-Open Shrublands	1.89	2.19	1.96	2.01
8-Woody Savannas	2.45	2.47	2.13	2.35
9-Savannas	1.93	2.46	2.11	2.17
10-Grasslands	1.41	1.78	1.86	1.68
11-Permanent Wetlands	2.15	1.58	1.35	1.69
12-Croplands	2.48	2.52	2.41	2.47
13-Urban and Built-up	2.62	2.76	2.59	2.66
14-Cropland/Natural Vegetation Mosaics	2.31	2.42	2.15	2.29
16-Barren	2.72	2.29	2.45	2.49
Overall DISCover EII Interpretation Confidence				2.23

TABLE 5. DISCOVER EII REGIONAL INTERPRETATION CONFIDENCE

Region	EII1	EII2	EII3	Average EII
1-North America/Canada	2.22	1.58	1.67	1.82
2-North America/US	2.59	2.44	2.48	2.51
3-Central America	1.93	2.71	2.14	2.26
4-Northern South America	2.56	2.65	2.53	2.58
5-Southern South America	2.45	2.54	2.55	2.52
6-Europe	2.38	2.57	2.15	2.37
7-North Africa	2.55	2.45	2.41	2.47
8-Central Africa	2.27	2.57	2.27	2.37
9-Southern Africa	1.93	2.47	2.00	2.13
10-Russia	1.81	2.57	1.90	2.10
11-Central Asia/Japan	2.36	1.70	1.80	1.95
12-Southeast Asia/China	2.56	2.33	2.75	2.55
13-Australia/New Zealand	1.58	1.82	2.75	2.05

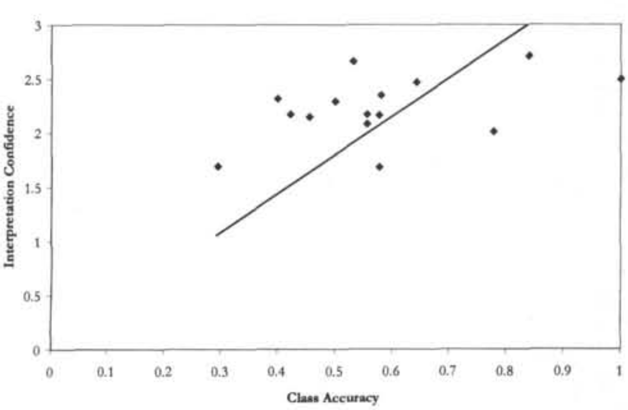

Figure 1. DISCover EII interpretation confidence and class accuracy.

ited from the EIIs. This questionnaire asked five questions regarding the Core sample verification; written responses were received for Regions 2(1 response), 4(1), 6(3), 8(2), 9(3), 10(2), 11(3), 12(3), and 13(3). The five questions and a summary of the information contained in the EII responses follows.

(1) In your opinion, how applicable is Landsat Thematic Mapper or SPOT data for analysis of the land-cover types present in your Region?

Generally, both TM and SPOT data are applicable for the major land-cover types except when images are acquired during winter when snow coverage, or in the dry season in arid lands. In some instances, depending on the complexity of the land cover being interpreted, it is more advantageous to use Landsat TM data because it has more band combination options than does the SPOT data. A number of comments outlined the need for multitemporal verification imagery. For example, the identification of classes such as deciduous forest often requires knowledge of both peak greenness and senescent conditions.

(2) Which (if any) DISCover land-cover classes were easily interpretable on the TM/SPOT data and which were difficult?

The only "easily interpretable" classes were those representing non-vegetated (barren) surfaces and those including a cultivation mosaic. Extensive undisturbed natural areas were also easily interpreted, especially if ancillary information were available. All the other open canopy vegetation categories, such as savanna and different types of shrublands, are difficult to identify on a purely spectral basis if there is no *a priori* knowledge of the specific site.

(3) Did you find any specific Landsat TM spectral band combinations particularly useful for interpretation in any particular area of your region or for any specific IGBP land-cover class?

The Thematic Mapper spectral combinations most useful for discriminating the DISCover land-cover categories are 4-5-3 and 4-3-2 and 3-2-1. The standard SPOT RGB 3-2-1 band combination was used exclusively for interpretation.

(4) Did you use the interpretation keys? Did you find them useful in any particular area of your region or for any specific IGBP land-cover class?

Generally, EIIs found the spectral interpretation keys useful, and some expressed a desire for additional interpretation keys. Principally due to logistic factors, coverage of the interpretation keys was not systematic in terms of area, thematic class, or temporal domain. It was also felt that more preparation time to collect ancillary data would have been helpful. The advance production of land-cover keys based on aerial photography (in addition to those generated from satellite data) for the vegetation classes would have been most useful on a region-specific basis. A few individual EIIs came to the GVWequipped with a range of representative aerial and ground photos of their regions land-cover classes; it would have been most useful for verification to provide this resource for all regions.

(5) What are your opinions of the core sample verification methodology in general? Please list any specific recommendations you have regarding the methodology used here and how it could be improved.

There is a large amount of regional land-cover variation present within the typical core sample subimage location. Due to the relatively small team of EIIs available, the possibility exists that in some regions land cover for large areas is not specifically familiar to the EIIs. For some sample points, there was an element of uncertainty in the verification process because EIIs unfamiliar with specific areas were required to provide verifications. Because the core sampling was stratified by land-cover class and not by continent, interpretations were not conducted using a statistically valid sample for each region. The verifications were also influenced by (indeed, dependant upon) the personal knowledge of the individual EIIs. These factors resulted in isolated cases where two out of the three EIIs came from a country within the region in which very few (or no) sample points were located. Although it would add logistical complexity, employing a sampling strategy that provided a sufficient number of sample points for all classes, validation regions, and EII locations (rather than just per class or per region) would have allowed more sites to be evaluated with higher interpreter confidence.

Some pre-classification group standardization/calibration process might also have helped considerably to lower the variance in land-cover coding. For example, class assignments can vary greatly between analysts simply due to the manner in which they apply the classification scheme. In this exercise, the IGBP legend with class names and definitions was provided to the EIIs. An example where a variable application of class labeling may occur can be seen with the savanna class. Many definitions of savanna exist and most have in common the description of a formation which is a mixture of grasses with or without a woody overstory. While the areal extent of savannas across the globe differs among researchers, savannas do exist in all tropical regions (Collinson, 1988). In other words, savannas are very often tied to a definition which includes a tropical or sub-tropical climatic regime. However, the physignomic classification scheme employed by the IGBP defines savanna as a formation having a tree canopy cover of 10 to 30 percent with an herbaceous or other understory system (Belward, 1996). Using this definition, many formations from open boreal woodlands to areas of fragmented forest would be included in the savanna definition while other formations commonly called savanna, such as tropical grassland, would not. For this exercise, only three regions exist wholly outside of the tropics: Regions 1, 6, and 10. From the interpretation data, only 1 of 82 sites labeled by the EIIs as savanna occurred in these regions. This reveals a possible usage of a more biome-based definition of savanna rather than the structural one listed in the IGBP. An interpreter used to working in a boreal region might be less likely to employ a savanna label than would an interpreter from the tropics. A presentation on this issue was given to the EII team at the beginning of each week's session at the global validation workshop. However, additional instruction on how to use the definitions where the emphasis is placed on the structural component and not the nomenclature might have aided in standardizing the interpretation approach. Other classes such as woody savanna, agricultural mosaic, and permanent wetlands might likewise require a calibration exercise for the interpreters.

Conclusions

The effort carried out in the IGBP-DISCover validation project is the first attempt to compile a global vegetation data set through remotely sensed data in which the accuracy validation of the product was performed through an active contribution of different image interpreters from all over the world with native knowledge.

In the global context, the level of detail of information that may be extracted in a classification process depends on the spatial, spectral, and temporal characteristics of the source data; the definition of the land-cover classes in use; and the classification techniques and procedures employed. NOAA-AVHRR data with 1.1-km spatial resolution are typically employed in very large-area (i.e., continental or global scale) land-cover classification efforts. Currently, the validation of such data sets must, of necessity, depend upon data acquired by a number of satellite sensor systems, including Landsat TM and SPOT. Issues exist regarding the consistency of remotely sensed data derived from different systems. When utilizing imagery acquired by different sensor systems, the interpretation of cover types might be complicated due to the lack of spectral band correspondence or sample spatial correspondence.

Some specific conclusions may be reached regarding the role of high-resolution imagery for validating a global 1-km data set as well as to the general use and interpretation of satellite imagery within this context:

- Classification accuracy and high-resolution imagery interpretation confidence are relatively higher for those classes that cover large areas.
- Overall interpretation confidence values demonstrate a medium to high level of group interpretation confidence. Only two classes (Class 10—Grasslands, and Class 11-Permanent Wetlands) have relatively low EII interpretation confidence levels. Average EII confidence values are below a medium level for only Regions 1 (North America/Canada) and 11 (Central Asia/Japan).
- Those global regions containing the highest levels of homogeneity scored the highest classification accuracies and generally also were interpreted with greatest confidence. These regions included North Africa (comprised mainly of the Sahara desert), and Central Africa and Central America areas having large proportions of forest stands). This is partly a result of the reduced impact of geometric registration error in such homogeneous cover types.
- There is a general relationship between the efficacy of the DISCover classification procedure and the ability of the analyst to confidently interpret and verify samples from the data set.

Acknowledgments

The international effort to validate the DISCover 1.0 land-cover data set was broad based and would have been impossible without the participation and support of a number of individuals, groups, and organizations. Principal support for this research was provided by the National Aeronautics and Space Administration Earth Science Enterprise, Land-Cover and Land-Use Change Program Research Grant NAG5-6259. Additional support for this work was provided to the University of California Santa Barbara, Remote Sensing Research Unit by NASA Earth Science Enterprise under Research Grant NAG5-3620. Critical contributions of data, technical support, and logistic support were made to this study by the U.S. Geological Survey, EROS Data Center, Sioux Falls, South Dakota. Important support was also provided by the International Goesphere Biosphere Programme Data and Information System Office in Tolouse, France, as well as the European Commision Joint Research Center, Ispra, Italy.

A number of university researchers and faculty also provided direction and technical expertise to this effort. Validating the DISCover Data Set would have not been possible without the energy and leadership provided to the effort by Dr. John Estes of the University of California Santa Barbara. We also wish to thank Dr. Alan Strahler and Douglas Muchoney at Boston University; Dr. Kenneth McGwire at the Desert Research Institute, University of Nevada, Reno; and Drs. John Townshend and Ruth Defries at the University of Maryland.

We also express our deep appreciation to the following group of Expert Image Interpreters who participated in this research: David Cunningham, Earth Satellite Corporation, USA; Allard de Wit, DLO Winand Staring Centre, The Netherlands; John Estes, University of California, Santa Barbara, USA; G.K. Eulert, EROS Data Center, USA; John Foster, Stennis Space Center, USA; Alisa Gallant, EROS Data Center, USA; Gabriela Gamez Rodriguez, Instituto de Geografia, UNAM, Mexico; Chandra Prasad Giri, UNEP/GRID, Bangkok, Thailand; Sharon Gomez, Environment and Remote Sensing Institute, Harare, Zimbabwe; Matthew Hansen, University of Maryland, USA; Shi Hua, Michigan State University, USA; Koji Kajiwara, Chiba University, Japan; Dominick Kwesha, Forestry Commission, Harare, Zimbabwe; Charles R. Larson, EROS Data Center, USA; Susan Maxwell, EROS Data Center, USA; Massaer Mbaye, Centre de Suivi Ecologique, Senegal; Gunter Menz, University of Bonn, Germany; Wayne Miller, EROS Data Center, USA; Doug Muchoney, Boston University, USA; Mikiyasu Nakayama, Utsunomiya University, Japan; Peter Newsome, Landcare Research New Zealand Ltd., New Zealand; Laili Nordin, Malaysian Centre for Remote Sensing, Malaysia; Juan Carlos Salazar Lea Plaza, Instituto de Suelos, Argentina; Carlos Scoppa, Centre de Investigaciones Recursos Naturales, Argentina; Hans-Jurgen Stibig, Joint Research Centre, Italy; Yuzo Suga, Hiroshima Institute of Technology, Japan; James C. Smoot, Stennis Space Center, USA; Mark Thompson, CSIR Environmentek, Pretoria, South Africa; Philip Tickle, Bureau of Resource Sciences, Australia; Jim Verdin, EROS Data Center, USA; Fabio Vescovi, University of Bonn, Germany; Jim Vogelmann, EROS Data Center, USA; Eric Waller, EROS Data Center, USA; Bruce Wylie, EROS Data Center, USA.

The dedication, enthusiasm and expertise that these individuals brought to the Global Validation Workshop (often traveling long distances to do so) exemplified the best in international scientific cooperation and can serve as a model for future projects of this type. In a very real sense, they made this research possible.

References

Ali, A.E., 1989. Assessment of Visual Interpretation of Thematic Mapper (TM) Images for Land Use and Cover in Sudan, *Earth Surface Processes and Landforms*, 14(5):399–405.

Belward, A., 1996. *The IGBP-DIS Global 1 km Land Cover Data Set "DISCover" - Proposal and Implementation Plans,* IGBP Working Paper 13. International Geosphere-Biosphere Programme, Joint Research Center of the European Commission, Space Applications Institute, Ispra, Italy, 66 p.

Cherrill, A.J., A. Lane, and R.M. Fuller, 1994. The Use of Classified Landsat-5 Thematic Mapper Imagery in the Characterization of Landscape Composition: A Case Study in Northern England, *Journal of Environmental Management*, 40(4):357–378.

Collinson, A.S., 1988. *Introduction to World Vegetation*, Unwin Hyman Ltd., London, 325 p.

Fitzpatrick-Lins, Katherine, 1980. *The Accuracy of Selected Land Use and Land Cover Maps at Scales of 1:250,000 and 1:100,000*, Geological Survey Circular 829, U.S. Geological Survey, Washington, D.C., 24 p.

Franklin, J., 1991. Land Cover Stratification Using LANDSAT Thematic Mapper Data in Sahelian and Sudanian Woodland and Wooded Grassland, *Journal of Arid Environments*, 20(2):141–163.

Jensen, J.R., 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective, Second Edition*, Prentice Hall, New Jersey, 316 p.

Kelly, M., J.E. Estes, and K.A. Knight, 1999. Image Interpretation Keys for Validation of Global Land-Cover Data Sets, *Photogrammetric Engineering & Remote Sensing*, 65(9):1041–1079.

Loveland, T.R., Z. Zhu, D.O. Ohlen, J.F. Brown, B.C. Reed, and L. Yang, 1999. An Analysis of the IGBP Global Land-Cover Characterization Process, *Photogrammetric Engineering & Remote Sensing*, 65(9):1021–1032.

Rosenfield, George H., K. Fitzpatrick-Lins, and H.S. Ling, 1981. Sampling for Thematic Map Accuracy Testing, *Photogrammetric Engineering & Remote Sensing*, 48(1):131–137.

Scepan, J., 1999. Thematic Validation of High-Resolution Global Land-Cover Data Sets, *Photogrammetric Engineering & Remote Sensing*, 65(9):1051–1060.

Vogelmann, J.E., T.L. Sohl, P.V. Campbell, and D.M. Shaw, 1998. Regional Land Cover Characterization Using LANDSAT Thematic Mapper Data and Ancillary Data Sources, *Environmental Monitoring and Assessment*, 51(1–2):415–428.