The IGBP-DIS Global 1-Km Land-Cover Data Set DISCover: A Project Overview

Alan S. Belward, John E. Estes, and Karen D. Kline

Abstract

Since 1992 the International Geosphere Biosphere Programme's (IGBP) Data and Information System (DIS) has been working towards the completion of a validated global land-cover data set, DISCover. This 1-km resolution data set consists of 17 cover classes identified on the basis of the science requirements of the IGBP's core projects. DISCover has been created from over 4.4 Terabytes of data from the Advanced Very High Resolution Radiometer collected from 23 receiving stations. These data were processed and assembled into a coherent set of monthly Normalized Difference Vegetation Index composites (April 1992 to April 1993) and classified using unsupervised techniques with post-classification refinement. The first global land-cover classification was completed in July 1997. The IGBP-DIS Land Cover Working Group, in turn, convened a Validation Working Group to provide and implement a validation method to provide statistical statements concerning the accuracy of the global land-cover product and to allow the estimation of the error variance in areal totals of classes globally and within regions. The validation workshop was completed in September 1998 and the analysis by March 1999. This paper describes the history of the DISCover version 1.0 implementation.

Introduction

There is a growing demand for global data sets describing land cover. Scientific communities, such as the International Geosphere Biosphere Programme (IGBP) use these data in the study of global change. Policy makers need such data when assessing the consequences of environmental treaties. Many non-governmental organizations are concerned with the global environment, and development programs and resource managers increasingly need to consider the global perspective.

While there is no lack of issues for which science quality thematic global data sets are required, few such data sets exist (Estes and Mooneyhan, 1994; Estes *et al.*, 1994). We assume that the maps we require exist, contain the information we seek, and are accurate and up-to-date. Unfortunately, this is not always the case. Some mapped information is more perishable than others—continental outlines are clearly more durable than maps of forest clear cutting—yet, as soon as data about the Earth are collected or a map made, those data and map products are dated.

Land cover is a case in point. Many global assessments have used models to delimit potential land-cover types. These have typically used the correlation between major biome patterns and climate for the description of land cover. Such models tend to be based on a simplified set of climatic parameters to delineate potential land-cover patterns. Although these climate-based land-cover models have been important in the analysis of changing land-cover patterns, there are still major limitations to their use in any predictive sense. Land cover representing baseline conditions is an important input to global dynamic models, and the quality and validity of such data, in essence, defines the reliability of simulated future scenarios.

This is clearly articulated by the IGBP core projects BAHC (Biospheric Aspects of the Hydrological Cycle), GCTE (Global Change and Terrestrial Ecosystems), IGAC (International Global Atmospheric Chemistry Project), and LUCC (Land Use Cover Change) which have all identified the need for improved baseline land-cover data sets for studies of global change (IGBP, 1994). This is because the existing global land-cover databases (e.g., Matthews, 1983; Olson *et al.*, 1985), although valuable, have too coarse a spatial resolution for many purposes, unknown accuracy, inappropriate classes for some core project needs, and variable nomenclatures. Equally important, these data sets are derived from disparate primary data sources, which raises significant issues of internal consistency and thus overall accuracy. An additional concern is that all existing global land-cover maps are dated. As the late James Anderson of the U.S. Geological Survey (USGS) pointed out, few things commonly depicted on maps change as fast as land cover (Anderson, 1977). An additional problem is that most of these thematic data sets have a climate element as part of the classification scheme and, therefore, often have an undesirable mixture of potential versus actual land cover (Townshend *et al.*, 1994).

Although the IGBP core projects recognized the shortcomings of existing databases, the wide range of the projects' requirements made specification of any new product a far from straightforward process. The land-cover categories identified by the IGBP are related to the needs of gas exchange studies, vegetation attributes for modeling Net Primary Production (NPP), burn emissions and gas exchange, wetlands cover and wetland water regimes, changes in vegetation/land cover over time, biological attributes, physical attributes, and landscape characteristics. Finding a set of land-cover characteristics that satisfy as many of these requirements as possible was a key objective. Optimum spatial resolution for such a global map was identified as being at the 1-km level, and ideally the map should be derived from a single data source obtained within a fixed and sequential time period. In this light, the IGBP's Data and Information System (IGBP-DIS) began the 1-km global land-cover project (Townshend, 1992) to provide appropriate land-cover information to the core projects.

A.S. Belward is with the European Commission Joint Research Centre, Space Applications Institution, 21020 Ispra, Varese, Italy.

J.E. Estes and K.D. Kline are with the Remote Sensing Research Unit, Department of Geography, University of California, Santa Barbara, CA 93106-4060 (estes@geog.ucsb.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 65, No. 9, September 1999, pp. 1013–1020.

0099-1112/99/6509–1013$3.00/0

The IGBP-DIS Land Cover Working Group (LCWG) coordinated development of the global 1-km land-cover product. The LCWG recognized that no single land-cover product would meet the needs of all the core projects, but that a land-cover product could be optimized to meet many of the requirements. Initial work focused on the creation and validation of a science quality 1-km resolution land-cover product, called DISCover. "Science quality" here means that the errors inherent in the overall production of a map have been documented (Estes and Mooneyhan, 1994), and that the DISCover product would consist of the classification scheme, the classification methodology, the land-cover data-base, the validation methodology, the validation database, and the validation data layer (Belward, 1996; Loveland and Belward, 1997).

While the IGBP core projects represent the key DISCover constituency, the LCWG was aware of programs outside IGBP with similar or related data requirements such as the Global Terrestrial Observing System (GTOS) and Global Climate Observing System (GCOS) and the United Nations Food and Agriculture Organization's (FAO) land-cover harmonization effort. Liaisons with these and similar programs were maintained throughout the life of the DISCover program.

Data Set Development

The requirements for consistency of data, timeliness, and resolution all pointed to the use of Earth observation data as a basis for creating a new global land-cover map. An appropriate Earth observation data set had to be found to satisfy the user requirements made by the IGBP. The most appropriate operational system at the time was the Advanced Very High Resolution Radiometer (AVHRR) on the National Oceanic and Atmospheric Administration's (NOAA) polar-orbiting satellites. These data are only available when the satellite is in direct line of sight of a ground receiving station, or by the limited capacity of on-board data recorders. The first task, therefore, was to establish a globally consistent daily AVHRR data set.

In April 1992 the U.S. Geological Survey's (USGS) EROS Data Center (EDC) and the European Space Agency (ESA), working with the National Oceanic and Atmospheric Administration (NOAA), the National Aeronautic and Space Administration (NASA), and the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) began to coordinate data collection from 23 receiving stations from around the world (Figure 1). More than 4.4 Terabytes of 1-km resolution AVHRR data have been archived. The data collection has continued to date (with the exception of the period September 1994 to February 1995 when there was no operational afternoon satellite). There are currently over 126,230 satellite scenes available. Most of these data have been assembled into a coherent global archive (Eidenshink and Faundeen, 1994). Data and information are available via the World Wide Web, using the following URL, **http://edcwww.cr.usgs.gov/landdaac/1KM/1kmhomepage.html**. The establishment of such a data set is in itself a major achievement, and these data are now the foundation for a suite of global terrestrial products (see, for example, Hansen *et al.* (1999) and Stroppiana *et al.* (1999)).

The DISCover Classification Scheme

To meet the diverse requirements of the IGBP core projects, the DISCover classification scheme was conceived and reviewed at workshops and by IGBP scientists intending to use the new database. The categories embrace the climate-independence and canopy component philosophy presented by Running *et al.* (1994) but modified to be compatible with classification systems currently used for environmental modeling; to provide, where possible, land-use implications; and to represent landscape mosaics. The classification scheme also embraced modifications proposed as the validation strategy developed.

DISCover contains 17 land-cover classes. The scheme is based on definitions of three canopy components: above ground biomass, leaf longevity, and leaf type. The requirements placed on the classification scheme were that every part of the Earth's surface be assigned to a class, and that each class must be exclusive, i.e., no class overlaps with another. The classification had to discriminate between perennial and annual above-ground biomass, as this is important for seasonal climate and carbon-balance modeling. It is also a major vegetation determinant of the surface roughness length parameter required by climate models for energy and momentum transfer equations. The classification had to discriminate between evergreen and deciduous canopy types because this is a critical variable in carbon cycle dynamics of vegetation, and affects seasonal albedo and energy transfer characteristics of the land surface. Leaf longevity was included as this indicates whether a plant annually must completely regrow its canopy, or a portion of it, with inferred consequences for carbon partitioning, leaf litter fall dynamics, and soil carbon. The classification separated needleleaf, broadleaf, and grasses because leaf type affects gas exchange characteristics. Finally, so that validation of the classification was possible, it was structured so that the categories could be consistently and equally interpreted from either the 1-km data or from higher resolution imagery (Belward, 1996).

The DISCover Classification Methodology

Approaches to classification of digital imagery are many and varied. The LCWG decided to base the DISCover mapping exercise on methods already proven for classification of large areas. The project used an adaptation of the method developed by the USGS EROS Data Center and University of Nebraska-Lincoln (Loveland *et al.*, 1991; Brown *et al.*, 1993).

The approach taken in this project involves mapping land cover for each continent in turn. The first mapping step involves unsupervised classification of Normalized Difference Vegetation Index (NDVI) composites to identify greenness classes using a clustering algorithm based on the K-Means algorithm (Kelly and White, 1993) run on monthly NDVI maximum value composites. Each cluster was assigned a preliminary cover-type label on the basis of the spatial patterns and spectral or multi-temporal statistics of each class, and on comparison with ancillary data. Ancillary data included descriptive land-cover information, NDVI and elevation statistics, and class relationships to other land-cover legends. Related "single-category" classes were then grouped under DISCover labels using a convergence-of-evidence approach.

Identification of DISCover classes water, snow and ice, and barren from greenness classes is inappropriate, and these were identified from AVHRR 0.72- to 1.10-μm data also retained by the IGBP-DIS processing chain, and a 12-month maximum value NDVI composite. The urban class was added from the "urban" layer information from the Digital Chart of the World (Danko, 1992). Loveland *et al.*, (1999, in this issue) provides full details of the classification methodology.

The DISCover Validation Strategy

Accuracy assessments of digital land-cover classifications are typically based on contingency tables, or confusion matrices, where accuracy is expressed in terms of errors of omission and commission, or in terms of agreement analysis using the Kappa test statistic (Rosenfeld and Fitzpatrick-Lins, 1986; Congalton, 1991; Stehman, 1996a). The contingency table is created by comparing on a class-by-class basis the land-cover classification with an independent data source—field observations, existing maps, higher resolution imagery—collected using a statistically valid sampling strategy (Robinson *et al.*, 1983; Rosenfeld, 1986; Stehman, 1996b).

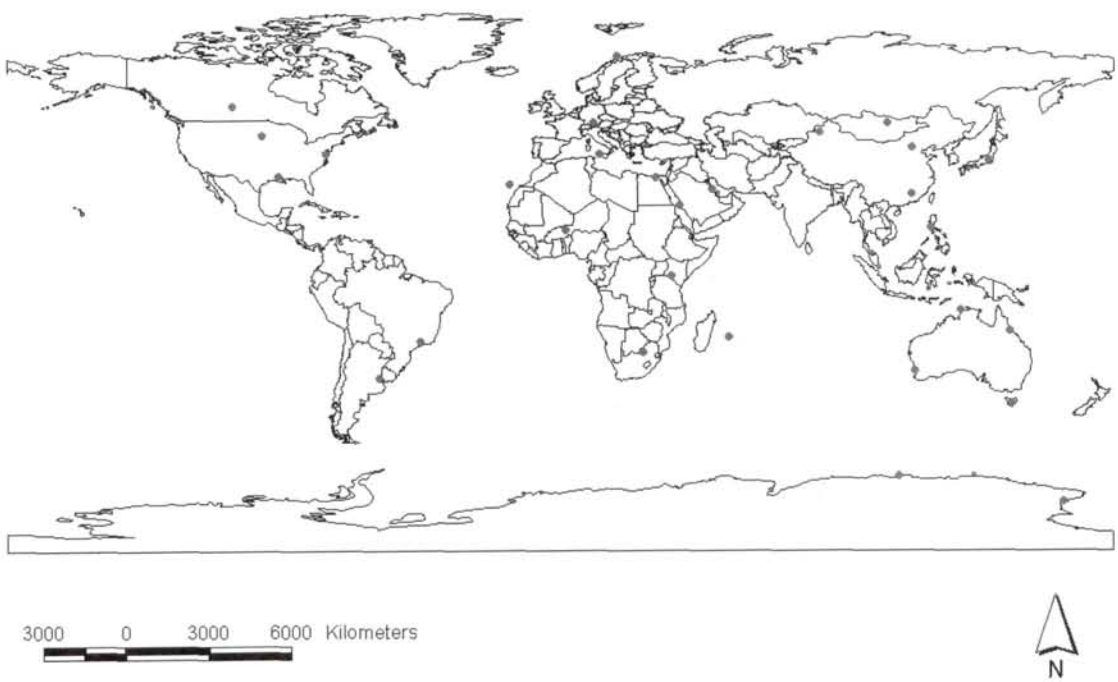

Figure 1. The locations of the 23 HRPT Receiving Stations that participated in the IGBP-DIS data collection effort.

While such methods are well established, they have generally only been applied to local scale classifications, occasionally to regional scale work, and only in isolated instances on scales greater than these (Hay, 1988; Manshard, 1993; Estes and Mooneyhan, 1994; Jeanjean *et al.*, 1996). The lack of peer reviewed documentation describing validation of global land-cover products led the LCWG to convene a specific sub group, the Validation Working Group (VWG), to examine how previous approaches to regional work could be adapted to the global scale of DISCover.

The VWG's mandate was "to specify and coordinate the implementation of a practical and achievable methodology for the validation of global land-cover products", where validation was defined as "the process that verifies the accuracy of the land-cover information provided by the Land Cover Working Group in as independent a means as practical" (Belward, 1996). The key here is the word "practical". Although pragmatic, the approach adopted by VWG included documenting all processes and stating assumptions and potential errors at each step.

Consultation with IGBP core projects suggested that an acceptable level of accuracy for the final global land-cover map was 85 percent correct classification for each class. It was the goal of the image classification exercise to reach these targets and the goal of the validation exercise to determine if they were reached.

Developing a validation strategy for the DISCover classification involved two assumptions. First, that at 1 km, the thematic and spatial accuracy of the classified data sets cannot be separated, and second, that higher resolution satellite images from sensors on Landsat or SPOT will provide accurate independent reference data describing the true land-cover class. The use of high-resolution data as a surrogate for reference data at least has precedents. Visual interpretation of higher resolution satellite imagery, either from hard copy prints or on screen using image processors, has been used for operational crop yield forecasting on a pan-European basis and has proved to be cost effective, capable of delivering consistent results over large areas, and easy to apply in an operational setting (De Boissezon *et al.*, 1993). This approach was thus to form the basis of the VWG validation strategy.

Using high-resolution data as a reference requires that, for each class, the class definitions are matched by a description of the AVHRR data characteristics associated with this class and a description of the higher resolution image characteristics associated with the class. The latter two descriptions are needed because the higher resolution data will not have the same temporal dimensions as the AVHRR. The higher resolution characteristics will also vary within the classes depending on the time of year that the data were obtained. If visual interpretation is to be consistent, precise decision rules must be provided to the interpreters. Ambiguity in what constitutes class definition must be avoided. The question of what percent of a given class must occupy a given sample site for that sample to be designated, that class must also be clear. Finally, independence of interpretations was maintained by using a team of three interpreters.

With these assumptions in mind, a two-tier strategy was adopted. The first was a core sampling strategy to provide simple statistical statements of accuracy. Second, a parallel activity based on data rich "confidence sites" where researchers would address more complex questions of accuracy.

Core Sampling

The principal objective of the core sampling strategy was to provide a set of simple statistical statements that (1) characterize the accuracy of the global land-cover product and (2) allow the estimation of the error variance in areal totals of land-cover types globally and within regions. Accepting that there is no current state-of-the-art for global land-cover classification validation, the most appropriate core sampling strategy must be based on state-of-the-practice procedures. Four such approaches were identified (Congalton, 1991).

Random sampling or random sampling stratified by continent (on the basis of percent total land area only) are the simplest approaches and would allow sample points to be reused for validation of future land cover classifications. However, without stratification and sub-sampling on the basis of class area, less frequent classes would be sampled less frequently, and thus the accuracy of their labeling would be less well known than for other, more frequently occurring classes (Rosenfield *et al.*, 1982). Accuracy statistics would only be valid on a global basis, and confidence intervals would be unequal and would vary widely as the number of sample points per class would vary. There would be strong bias towards the larger classes because these will tend to be sampled more frequently, and some classes may not be sampled at all.

Random sampling stratified by class would provide roughly equal confidence ranges *a priori*, and would produce unbiased estimates of cover areas and proportions at the global scale. The optimum sampling strategy would be random sampling stratified by class and processing region (where processing region refers to all areas classified in the same production run, using the same ancillary data, same clusters, etc.) because this would highlight any systematic variation in error structure between processing regions, and the errors of omission and commission would also be applicable on a regional scale. Though optimum, the cost is high because the number of samples per class is then multiplied by the number of processing regions. For this reason, the VWG recommended random sampling stratified by class.

The sampling strategy aims to estimate the probability p that a 1-km square, truly of type I, gets classified according to the AVHRR data as type I. In terms of conditional probabilities, this is expressed as

$$p = P(\text{classify as I}|\text{truly type I}) \quad (1)$$

where p is the probability of classification as I, given truly an I.

To achieve this, n pixels of type I will be sampled (known to be of type I on the basis of reference data from high-resolution imagery) and a check will be made of how many are allocated to type I by the classification scheme being validated. If r is the number correctly classified by the scheme, then r/n is a natural estimate of the accuracy probability above, and a 95 percent confidence interval around r/n tells us how reliable this estimate is. Ideally, the accuracy of each class would be calculated with the same confidence interval. For practical reasons, the sample size per class needs to be kept down to a minimum. Standard charts for Binomial confidence intervals (Pearson and Hartley, 1966) confirm that a sample of 25 would be adequate, but this gives wide confidence intervals ($\pm$14.3 percent) and the accuracy statements are thus somewhat limited. Basically, problem classes can be identified, but the wide confidence interval means it is impossible to say if a class has truly been classified to 85 percent correct. Bearing in mind the Validation Working Group's key word "practical," this limited validation was judged to be of far greater value than no validation, and the sample of 25 pixels was accepted.

The core sites were thus identified using a stratified random sample of 25 points per class. Two-hundred eighty-five TM scenes and 150 SPOT HRV scenes were acquired (the total allows a slight margin for undersampling in some classes) thanks to the efforts of the USGS EROS Data Center (EDC), the European Commission's Joint Research Centre, the IGBP-DIS office, SPOT Image, and the support of the Centre National d'Etudes Spatiales (CNES). The high-resolution reference data were then prepared by the University of California Santa Barbara and transported to EDC where they were analyzed at a dedicated workshop involving participation of expert interpreters from around the world. Husak *et al.*, (1999, in this issue) provide more details on the data preparation, Kelly *et al.*, (1999, in this issue) cover the image interpretation key development efforts, and Scepan (1999, in this issue) provides full details of the validation process.

Confidence Sites

Land cover is not static, and the LCWG recognizes that there will be a need for additional updates of the global land-cover product. With this understanding, the second element of the DISCover validation exercise was proposed around the concept of "confidence sites." Confidence sites comprise a set of locations at which fine-resolution land-cover information for a substantial region exists or is readily available. The VWG envisioned that local land-cover information would be available at or provided by a sponsoring research group, institute, or other collaborating body with local knowledge of the confidence site. Care would be taken to ensure that confidence sites were chosen such that they represent areas disturbed by human activity, in addition to natural vegetation sites. These confidence sites should be employed in a variety of ways. Important among these are to

- Explore ways to improve the land-cover classification scheme and the methodology involved in the development of the land-cover database,
- Test ways to improve the core thematic sampling strategy,
- Assess the applicability of advanced validation strategies to improve both thematic and positional accuracy assessment, and
- Address ways to improve processing related methodologies.

Plans for confidence site selection and implementation are provided in Belward (1996). A description of the confidence site database is presented by Muchoney *et al.* (1999, in this issue). Collection and establishing of global validation sites is a long term objective of the project, and one that will continue to test the international scientific community's collective will for collaboration.

Conclusions

This IGBP-DIS lead effort is a significant attempt to break through the myths surrounding global thematic mapping, myths that perpetuate the notion that the maps we need exist, are readily accessible, are up-to-date, are accurate, are inexpensive to produce, and that there is little research to do in the mapping sciences—in essence, that mapping is easy (Estes and Mooneyhan, 1994).

This is the first time a global-scale geospatial land-cover product has been statistically validated. The DISCover project has served as a forerunner to many new global land-cover products (DeFries and Belward, 1999). We all need to be more aware that mapping is an important, complex, expensive, and time consuming task, a task that many in the geospatial science community believe is not being performed in an acceptable fashion (Estes *et al.*, 1994). We need to invest more in mapping and map-update research.

The infrastructure needed to support the development of DISCover was put in place through international cooperation and coordination. This process must continue. More than 100 individuals from 30 nations participated. These participants came from seven international organizations; 26 research institutes and academic organizations; 21 national agencies, centers, and/or laboratories; and seven industrial groups. At the very least, 15 separate "formal" international meetings were held in connection with this effort. Appendix A lists many of those individuals who participated directly in these meetings and this effort as we could assemble for this article. A major role was played by those dedicated individuals at the HRPT receiving stations, the USGS EROS Data Center, and at CNES and SPOT Image who worked so hard and so diligently to provide the image data employed in this effort. Special acknowledgment of and thanks for these efforts is required.

No unique permanent infrastructure was put in place to accomplish the development and validation of DISCover. Instead, through the auspices of the IGBP-DIS Working Groups, a wide variety of participating organizations were involved on the basis of their already established activities. What is remarkable about this effort was that participation was mostly voluntary and self funded. The dollars allocated specifically to fund DISCover related development and validation activities totaled no more than $500,000. Many organizations and individuals involved in this effort used existing operational funding and/or funding from projects they considered related to ensure their participation in the project.

What have we proven from the DISCover project? We have, we believe, shown that, through cooperation and collaboration, the international community can pull together to achieve a scientifically significant goal: the creation of a validated global land-cover data set. In common with other global data sets, the DISCover product is already dated. Yet the need for global terrestrial information continues to grow. The importance of the global perspective now has new significance as environmental information input to the policy making, policy development, and policy implementation processes grows. The Kyoto Protocol to the UN framework convention on climate change, for example, recognizes the need to monitor measures promoting the protection and enhancement of greenhouse gas sinks and reservoirs (Kyoto Protocol Article 2); measure changes in carbon stocks resulting from human-induced land-use change and forestry activities (Article 3) and monitor such activities under the clean development mechanism (Article 12); monitor transfer of emission reduction units resulting from projects which enhance anthropogenic removals by sinks (Articles 6 and 17); provide systematic observation and the development of data archives to reduce uncertainties related to the climate system (Article 10); and be internationally recognized and accepted (Articles 5 and 17). The legal and economic implications of the Kyoto Protocol and other treaties place new demands and constraints on the reliability and accuracy of global land-cover information. Future research and infrastructure development efforts are needed to continue to produce science quality thematic data sets, such as DISCover. This will most realistically be achieved through international collaboration.

The Committee on Earth Observation Satellites (CEOS), formed in 1984 under the aegis of the Economic Summit of Industrialized Nations' Working Group on Growth Technology and Employment, is probably the major forum for international cooperation in space. CEOS membership includes all the world's civil agencies responsible for Earth observation satellite programs along with international user organizations. The goals of CEOS are to promote cooperation so as to maximize the benefits of space-borne Earth observations, to aid members and users by acting as a focal point for international coordination of space related Earth observation activities, and to exchange policy and technical information. CEOS established the Working Group on Calibration and Validation (WGCV) to address sensor-specific calibration and validation, and geophysical parameters and derived product validation. Dedicated forums such as the IGBP-DIS Land Cover Working Group proved a key asset for the creation of this first validated global land-cover product; building on examples such as the DISCover validation effort the WGCV is increasingly addressing validation of derived products (Belward, 1997). The importance and complexity of international collaboration are more fully explored in Estes *et al.* (1999).

Finally, we cannot close this article without extending our many thanks to all the participating organizations and individuals (see Appendices A and B). The logistics of this effort were complex. The collaboration needed to accomplish the production and validation of DISCover was significant, and we firmly believe that a detailed acknowledgment of project participation is more than justified. So many individuals and organizations donated resources (e.g., funding, time, materials, advice, etc.) to this effort that not to acknowledge them all would do a disservice to them and to the reader.

References

Anderson, J.R., 1977. Land Use and Land Cover Changes—A Framework for Monitoring, *Journal of Research*, U.S. Geological Survey, 5:143–153.

Belward, A.S. (editor), 1996. *The IGBP-DIS Global 1 Km Land Cover Data Set "DISCover" Proposal and Implementation Plans*, Report of the Land Cover Working Group of IGBP-DIS, IGBP-DIS Working Paper #13, IGBP Data And Information System Office, Toulouse, France, 64 p.

Belward, A.S., 1997. The CEOS Working Group on Calibration and Validation: Plans and Perspectives, *Canadian Journal of Remote Sensing*, 2.3 (4):414–418.

Belward, A.S., and T.R. Loveland, 1995. The IGBP-DIS' 1 Km Land Cover Project, *Remote Sensing in Action, Proceedings of The Remote Sensing Society Annual Conference*, Southampton, 11–14 September, Remote Sensing Society, Nottingham, pp. 1099–1106.

Brown, J.F., T.R. Loveland, J.W. Merchant, B.C. Reed, and D.O. Ohlen, 1993. Using Multisource Data in Global Land Characterization: Concepts, Requirements, and Methods. *Photogrammetric Engineering & Remote Sensing*, 59(6):977–987.

Congalton, R.G., 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data, *Remote Sensing of Environment*, 37:35–46.

Danko, D.M., 1992. The Digital Chart of the World, *Geoinfosystems*, 2:29–36.

De Boissezon, H., G. Gonzales, B. Pous, and M. Sharman, 1993. Rapid Estimates of Crop Acreage and Production at a European Scale Using High Resolution Imagery–Operational Review, *Proceedings of the International Symposium on Operationalization of Remote Sensing*, Enschede, The Netherlands, 19–23 April, International Institute for Aerospace Survey and Earth Sciences, Enschede, 2:94–105.

DeFries, R.S., and A.S. Belward, 1999. Characterizing Global and Regional Land Cover from Satellite Data: An Introduction to the Special Issue, *International Journal of Remote Sensing*, (in press).

Eidenshink, J.E., and J.L. Faundeen, 1994. The 1 Km AVHRR Global Land Data Set: First Stages in Implementation, *International Journal of Remote Sensing*, 15(17):3443–3462.

Estes, J.E., and D.W. Mooneyhan, 1994. Of Maps and Myths, *Photogrammetric Engineering & Remote Sensing*, 60:517–524.

Estes, J.E., J. Lawless, and D.W. Mooneyhan, 1994. *Proceedings of the International Symposium on Core Data Needs for Environmental Assessment and Sustainable Development Strategies*, Bangkok, Thailand, 15–18 November, United Nations Development Programme and United Nations Environment Programme.

Estes, J.E., A.S. Belward, J.O. Justice, T.R. Loveland, J. Scepan, J. Townshend, and A. Strahler, 1999. The Way Forward, *Photogrammetric Engineering & Remote Sensing*, 65(9):1089–1093.

Hansen, M.C., R.S. DeFries, J.R.G. Townshend, and R. Sohlberg, 1999. Global Land Cover Classification at 1 km Spatial Resolution Using a Classification Tree Approach, *International Journal of Remote Sensing*, (in press).

Hay, A.M., 1988. The Derivation of Global Estimates from a Confusion Matrix, *International Journal of Remote Sensing*, 9:1395–1398.

Husak, G.J., B.C. Hadley, and K.C. McGwire, 1999. Registration Accuracy of High-Resolution Satellite Data Used in the IGBP Validation, *Photogrammetric Engineering & Remote Sensing*, 65(9):1033–1039.

International Geosphere Biosphere Programme, 1994. *IGBP in Action: Work Plan 1994–1998*, IGBP, Stockholm, 151 p.

Jeanjean, H., F. Achard, and J.-P. Malingreau, 1996. Large Scale Tropical Forest Change Monitoring Using Multiple Resolution Satellite Data: From Hot Spot Detection to Global Deforestation Assessment? *Proceedings of the 2nd International Symposium on the*

Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Fort Collins, Colorado.

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):1041–1049.

Kelly, P.M., and J.M. White, 1993. Preprocessing Remotely-Sensed Data for Efficient Analysis and Classification, *Proceedings, SPIE Applications of Artificial Intelligence 1993: Knowledge-Based Systems in Aerospace and Industry,* pp. 24–30.

Loveland, T.R., J. W. Merchant, D.O. Ohlen, and J.F. Brown, 1991. Development of a Land-Cover Characteristics Database for the Conterminous U.S., *Photogrammetric Engineering & Remote Sensing,* 57(11):1453–1463.

Loveland, T.R., and A.S. Belward, 1997. The IGBP-DIS Global 1 km Land Cover Data Set, DISCover First Results, *International Journal of Remote Sensing,* 18(15):3289–3295.

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

Manshard, W., 1993. Geography and the International Global Change Programmes, *European Review,* 1:309–317.

Matthews, E., 1983. Global Vegetation and Land Use: New High Resolution Data Bases for Climate Studies, *Journal of Climate and Applied Meteorology,* 22: 474–487.

Muchoney, D., A. Strahler, J. Hodges, and J. Locastro, 1999. The IGBP DISCover Confidence Sites and the System for Terrestrial Ecosystem Parameterization: Tools for Describing and Validating Global Land Cover Data, *Photogrammetric Engineering & Remote Sensing,* 65(9):1061–1067.

Olson, J.S., J.A. Watts, and L.J. Allison, 1985. *Major World Ecosystems Ranked by Carbon in Live Vegetation: A Database,* NDP-017, Carbon Dioxide Information Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 397 p.

Pearson, E.S., and H.O. Hartley, (editors), 1966. *Biometrika Tables for Statisticians,* Cambridge University Press, Cambridge, Volume 1, Third Edition, 84 p.

Robinson, J.W., F.J. Gunther, and W.J. Campbell, 1983. *Ground Truth Sampling and Landsat Accuracy Assessment,* Document N83-26161, National Technical Information Services.

Rosenfield, G.H., 1986. Analysis of Thematic Map Classification Error Matrices, *Photogrammetric Engineering & Remote Sensing,* 52:681–686.

Rosenfield, G.H., and K. Fitzpatrick-Lins, 1986. A Coefficient of Agreement as a Measure of Thematic Classification Accuracy, *Photogrammetric Engineering & Remote Sensing,* 52:223–227.

Rosenfield, G.H., K. Fitzpatrick-Lins, and H.S. Ling, 1982. Sampling for Thematic Map Accuracy Testing, *Photogrammetric Engineering & Remote Sensing,* 48:131–137.

Running, S.W., T.R. Loveland, and L.L. Pierce, 1994. A Vegetation Classification Logic Based on Remote Sensing for Use in Global Biogeochemical Models, *Ambio,* 23(1):77–81.

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

Stehman, S.V., 1996a. Estimating the Kappa Coefficient and its Variance under Stratified Random Sampling, *Photogrammetric Engineering & Remote Sensing,* 62:401–407.

———, 1996b. Cost-Effective, Practical Sampling Strategies for Accuracy Assessment of Large-Area Thematic Maps, *Proceedings of the 2nd International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences,* Fort Collins, Colorado.

Stroppiana, D., S. Pinnock, and J.M. Grégoire 1999. The Global Fire Product: Daily Fire Occurrence from April 1992 to December 1993, Derived from NOAA-AVHRR Data, *International Journal of Remote Sensing,* (in press).

Townshend, J.R.G., (editor), 1992. *Improved Global Data for Land Applications, IGBP* Report Number 20, IGBP Secretariat, Royal Swedish Academy of Sciences, Stockholm, Sweden, 75 p.

Townshend, J.R.G., C.O. Justice, D. Skole, J.-P. Malingreau, J. Cihlar, J. Teillet, F. Sadowski, and S. Ruttenberg, 1994. The 1-km AVHRR Global Data Set: Needs of the International Geosphere Biosphere Program, *International Journal of Remote Sensing,* 15:3319–3332.

Appendix A.

List of Participants in IGBP DISCover Production and Validation Meetings and the Number of Meetings (see Appendix B) They Attended.

Name	Affiliation	Country	Number of Meetings
Frederic Achard	EC Joint Research Centre	Italy	4
Roman Alvarez	UNAM	Mexico	1
Clive Anderson	University of Sheffield	UK	1
Olivier Arino	European Space Agency	Italy	1
Etienne Bartholomé	EC Joint Research Centre	Italy	4
Peter Beedlow	Environmental Protection Agency	USA	1
Alan Belward	EC Joint Research Centre	Italy	13
Jordan Borak	Boston University	USA	1
Jesslyn F. Brown	EROS Data Center	USA	2
Alessandra Buongiorno	European Space Agency	Italy	2
Michael Cairns	Environmental Protection Agency	USA	1
Jean-Louis Champeaux	Meteo-France	France	1
Josef Cihlar	Canada Centre for Remote Sensing	Canada	3
Mark Collins	World Conservation Monitoring Centre	UK	1
David Cunningham	Earth Satellite Corporation	USA	1
Gérard Dedieu,	CESBIO	France	1
Pierre Defourny	Universite Catholique de Louvain	Belgium	3
Ruth DeFries	University of Maryland	USA	1
Jim Eastwood	Institute of Terrestrial Ecology	UK	1
Daniele Ehrlich	EC Joint Research Centre	Italy	3
Jeff Eidenshink	USGS/EROS Data Center	USA	5
Christopher Elvidge	Desert Research Institute	USA	1
John E. Estes	University of California, Santa Barbara	USA	9
	US Geological Survey	USA	
Gene Eulert	EROS Data Center	USA	1
David Evans	US Forest Service	USA	1
John Foster	Stennis Space Center	USA	1
Robin Fuller	Institute of Terrestrial Ecology	UK	1
Alisa Gallant	USGS/EROS Data Center	USA	1
Kevin Gallo	NOAA/National Climate Data Center	USA	2
Sig Gerstl	Los Alamos National Laboratory	USA	1
Chandra Prasad Giri	UNEP/GRID-Bangkok	Thailand	1
Sharon N. Gomez	SIRDC	Zimbabwe	1
Gabriela Gomez-Rodriguez	UNAM	Mexico	1
Sam Goward	University of Maryland	USA	1
Brian Hadley	University of California, Santa Barbara	USA	2
Matt Hansen	University of Maryland	USA	2
Jeff Hemphill	University of California, Santa Barbara	USA	1
Greg Husak	University of California, Santa Barbara	USA	2
John Ingram	University of Oxford	UK	1
Mary James	NASA/Goddard Space Flight Center	USA	1
Louisa Jansen	UN Food and Agriculture Organization	Italy	1
Hervé Joannes	SPOT-IMAGE	France	1
Chris Justice	University of Virginia	USA	7
	NASA/Goddard Space Flight Center	USA	

Name	Affiliation	Country	Number of Meetings
Satya Kalluri	University of Maryland	USA	1
Melissa Kelly	University of California, Santa Barbara	USA	2
Karen Kline	University of California, Santa Barbara	USA	2
Koji Kojiwara	Chiba University	Japan	1
Andrey Kushlin	World Bank	USA	1
Dominick Kwesha	Forestry Commission	Zimbabwe	1
Charles Larson	EROS Data Center	USA	1
John Latham	UN Food and Agriculture Organization	Italy	3
Michael Lawless	University of California, Santa Barbara	USA	1
Rik Leemans	Laboratory for Soil and Groundwater	Netherlands	1
Jan de Leeuw	ITC	Netherlands	1
David Llewellyn-Jones	Leicester University	UK	1
Massimiliano Lorenzini	UN Food and Agriculture Organization	Italy	1
Sietse Los	SSAI	USA	1
Tom Loveland	USGS/EROS Data Center	USA	9
Ross Lunetta	Environmental Protection Agency	USA	1
Martha Maiden	NASA/Goddard Space Flight Center	USA	1
Jean-Paul Malingreau	EC Joint Research Centre	Italy	4
Susan Maxwell	EROS Data Center	USA	1
Philippe Mayaux	EC Joint Research Centre	Italy	1
Massaer Mbaye	Centre de Suivi Ecologique	Senegal	1
Ken McGwire	Desert Research Institute	USA	5
Gunter Menz	University of Bonn	Germany	1
Martine Michou	IGBP-DIS Office	France	4
Wayne Miller	USGS/EROS Data Center	USA	1
Aaron Moody	Boston University	USA	1
Sander Mucher	DLO Winand Staring Centre	Netherlands	1
Doug Muchoney	Boston University	USA	3
Ranga Myneni	NASA/Goddard Space Flight Center	USA	1
Mikiyasu Nakayama	Utsunomiya University	Japan	1
Peter F. J. Newsome	Landcare Research New Zealand Ltd.	New Zealand	1
Ian Noble	Australian National University	Australia	1
Laili bin Nordin	Malaysian Centre for Remote Sensing	Malaysia	1
Dennis Ojima	Colorado State University	USA	1
Jose Palacio	UNAM	Mexico	1
Dieter Pelz	Albert-Ludwigs-Universitat	Germany	2
Bernard Pinty	EC Joint Research Centre	Italy	1
Giancarlo Pittella	European Space Agency	Italy	1
Brad Reed	EROS Data Center	USA	3
Ake Rosenqvist	EC Joint Research Centre SSC	Italy Sweden	1
Jean-Louis Roujean	Meteo-France	France	1
Steve Running	University of Montana	USA	1
Baldev Sahai	ISRO	India	1
Gilbert Saint	CNES	France	4
Bill Salas	University of New Hampshire	USA	1
Juan Carlos Salazar Lea Plaza	CIRN-INTA	Argentina	1
Nazmi el Saleous	University of Maryland	USA	1
Joseph Scepan	University of California, Santa Barbara	USA	5
Robert Scholes	CSIR	South Africa	2
Miriam Schomaker	UN Environment Programme	Kenya	1
Carlos Octavio Scoppa	CIRN-INTA	Argentina	1
Jeff Settle	University of Reading	UK	1
Hua Shi	University of Michigan	USA	1
Denis Sims	UN Food and Agriculture Organization	Italy	1
Ashbindu Singh	EN Environment Programme/GRID	USA	1
David Skole	University of New Hampshire	USA	1
James "Doc" Smoot	Stennis Space Center	USA	1
William Starmer	University of California, Santa Barbara	USA	2
Hans-Jürgen Stibig	EC Joint Research Centre	Italy	1
Tom Stone	Woods Hole Research Center	USA	1
Alan Strahler	Boston University	USA	8
Yuzo Suga	Hiroshima Institute of Technology	Japan	1
Gerard Szejwach	IGBP-DIS Office	France	1
Ryutaro Tateishi	Chiba University	Japan	1
Mark Thompson	CSIR	South Africa	1
Philip Tickle	Bureau of Resource Sciences	Australia	1
John Townshend	University of Maryland	USA	6
Rob van der Velde	RIVM	Netherlands	1
James Verdin	USGS/EROS Data Center	USA	1
Frank Veroustraete	VITO	Belgium	1
Michael Verstrate	EC Joint Research Centre	Italy	1
Fabio Vescovi	University of Bonn	Germany	1
Nicolas Viovy	LMCE CE SACLAY	France	1
James Vogelmann	USGS/EROS Data Center	USA	1
Eric Waller	University of California, Santa Barbara USGS/EROS Data Center	USA USA	2
Ben Waltenberger	University of California, Santa Barbara	USA	1
Allard de Wit	DLO Winand Staring Centre	Netherlands	1
Curtis Woodcock	Boston University	USA	1
Dorsey Worthy	Environmental Protection Agency	USA	1
Barry Wyatt	Institute of Terrestrial Ecology	UK	5
Bruce Wylie	EROS Data Center	USA	1
Yoshifumi Yasuoka	NIES	Japan	2
Zhiliang Zhu	EROS Data Center Hughes	USA USA	1

Appendix B.

List of Meetings Related to the Production and Validation of the IGBP DISCover Data Set.

Name	Location	Date
1. Identification of Fast Trace Project Production Issues, LCWG Meeting	Ispra, Italy	05 Jul 1993
2. Establishment of Land Cover Classification Methodologies and Validation Overview, LCWG Workshop	Las Vegas, Nevada	23–24 Feb 1994
3. Drafting of the Validation Strategy, VWG Meeting	Ispra, Italy	16–17 May 1994
4. Consolidation of the Validation Strategy, VWG Workshop	Monks Wood, UK	16–17 Sep 1994
5. Reconciliation of Land Cover Data Set Production Issues and Validation Implementation Issues, LCWG/VWG Workshop	Ispra, Italy	02–03 Feb 1995

Name	Location	Date
6. 1 km Data Set Evaluation, IGBP-DIS Workshop	Maryland	02–03 Mar 1995
7. Implementation and Project Review, LCWG	Sioux Falls, South Dakota	02 Jun 1995
8. Validation Sample Frame Specifications, VWG sub group meeting	Ispra, Italy	05–06 Jul 1995
9. Co-ordination and Planning, LCWG Chairs	Sioux Falls, South Dakota	28–30 Nov 1995
10. Core Validation and ad hoc Meeting, VWG	Ispra, Italy	29 Apr 1996
11. Implementation Meeting, LCWG	Ispra, Italy	16–18 Jul 1996
12. Validation Implementation Planning Meeting	Toulouse, France	09–10 Jul 1997
13. Validation Workshop dry run and Methodology Testing	Santa Barbara, California	01–04 Feb 1998
14. Validation Workshop	Sioux Falls, South Dakota	07–18 Sep 1998
15. LCWG Future Perspectives and Results Review	Ispra, Italy	10 Feb 1999