Validation of the Global Land Cover 2000 Map

Philippe Mayaux, Hugh Eva, Javier Gallego, Alan H. Strahler, *Member, IEEE*, Martin Herold, *Student Member, IEEE*, Shefali Agrawal, Sergey Naumov, Evaristo Eduardo De Miranda, Carlos M. Di Bella, Callan Ordoyne, Yuri Kopin, and P. S. Roy

***Abstract*—The Joint Research Centre of the European Commission (JRC), in partnership with 30 institutions, has produced a global land cover map for the year 2000, the GLC 2000 map. The validation of the GLC2000 product has now been completed. The accuracy assessment relied on two methods: a confidence-building method (quality control based on a comparison with ancillary data) and a quantitative accuracy assessment based on a stratified random sampling of reference data. The sample site stratification used an underlying grid of Landsat data and was based on the proportion of priority land cover classes and on the landscape complexity. A total of 1265 sample sites have been interpreted. The first results indicate an overall accuracy of 68.6%. The GLC2000 validation exercise has provided important experiences. The design-based inference conforms to the CEOS Cal-Val recommendations and has proven to be successful. Both the GLC2000 legend development and reference data interpretations used the FAO Land Cover Classification System (LCCS). Problems in the validation process were identified for areas with heterogeneous land cover. This issue appears in both in the GLC2000 (neighborhood pixel variations) and in the reference data (cartographic and thematic mixed units). Another interesting outcome of the GLC2000 validation is the accuracy reporting. Error statistics are provided from both the producer and user perspective and incorporates measures of thematic similarity between land cover classes derived from LCCS.**

***Index Terms*—Quality control, statistics, vegetation mapping.**

I. Introduction

SINCE the early 1990s, the scientific community started to produce consistent global land-cover information from remotely-sensed data. The International Geosphere-Biosphere Program (IGBP) Data and Information System [1] published the first global map at 1.1-km spatial resolution (DISCover) derived from a single data source (the AVHRR Local Area Coverage), and made over a fixed time period (April 1992 to end of 1993). Recently, new sensors, MODIS on board the Terra and Aqua platforms [2] and VEGETATION on board SPOT-4 and SPOT-5 [3], allowed for a spatial and thematic refinement of the previous global maps (respectively MODIS Land Cover and Global Land Cover 2000 or GLC 2000) due to the greater stability of the platforms and spectral characteristics of the sensors. Future global land-cover maps are now planned from medium resolution sensors, such as MERIS.

These maps are extensively used in Global Change research for model parameterization or regional stratification, by the biodiversity community for identifying areas suitable for conservation management and to support the work of other groups such as Non-Governmental Organizations (NGO) and development assistance programs. As this wide-ranging user community has gained experience with global land cover datasets, the map producing community is receiving requests for new products; products offering increased spatial and thematic detail and products bringing the global land-cover data base more up-to-date.

The multiplicity of existing products and of potential users clearly poses the question of the adequacy of a particular map for a specific use. Many people use land-cover data in their applications without concern for their accuracy, even though this check could improve the quality of the final results. The choice of a map should be dictated for a particular application according to the focus in the legend, to the differences in the regional accuracy or to the spatial pattern. For example, the IGBP DISCover and the MODIS Land Cover products were primarily designed for carbon cycle studies and climate modeling at the global scale, while the global map GLC2000 is derived from regional products adapted to the local context and has for its main customer the Millennium Ecosystem Assessment.

This paper aims at presenting the strategy developed for validating the GLC2000 regional products and the global synthesis. Final results of this validation will be detailed in a further article.

II. Global Land Cover 2000 Product

The general objective of the European Commission's "Global Land Cover 2000" is to provide for the year 2000 a harmonized land cover database over the whole globe. To achieve this objective GLC 2000 makes use of the VEGA 2000 dataset: a dataset of 14 months of preprocessed daily global data acquired by the VEGETATION instrument on board SPOT 4, made available through a sponsorship from members of the VEGETATION program.

The GLC2000 land cover database has been chosen as a core dataset for the Millennium Ecosystems Assessment. This means in particular that the GLC2000 dataset is a main input dataset to define the boundaries between ecosystems such as forest, wetlands, and cultivated systems, which were defined by the MA secretariat as priority classes (http://www-gvm.jrc.it/glc2000/defaultGLC2000.htm).

Manuscript received October 30, 2004; revised September 30, 2005.

P. Mayaux, H. Eva, and J. Gallego are with the Institute for Environment and Sustainability of the European Commission's Joint Research Centre, I-21020 Ispra, Italy (e-mail: philippe.mayaux@ jrc.it).

A. H. Strahler and C. Ordoyne are with the Department of Geography, Boston University, Boston, MA 02215-1406 USA.

M. Herold is with the ESA GOFC GOLD Land Cover project office at the Friedrich Schiller University, Jena 07743, Germany.

S. Agrawal and P. S. Roy are with the Indian Institute of Remore Sensing, Dehradun 248001, India.

S. Naumov and Y. Kopin are with Data+, Moscow 123242, Russia.

E. E. De Miranda is with Embrapa-CNPM, Campinas 13088-300, Brazil.

C. M. Di Bella is with the Institute of Climate and Water, Buenos Aires 1712, Argentina.

Digital Object Identifier 10.1109/TGRS.2006.864370

The project was based on a partnership of some 30 institutions from around the World. Teams of regional experts mapped each continent independently. Each regional team participating in the project had experience of mapping their area through the use of data from Earth Observing satellites. This ensures that optimum image classification methods were used, that the land cover legend was regionally appropriate and that access could be gained to reference material. This bottom-up approach is novel for mapping land-cover at a global scale, compared to the previous IGBP DISCover and MODIS Land Cover which are based on a top-down approach: one method applied to the same dataset over the globe.

The GLC2000 philosophy dictates that these regionally detailed classes also be aggregated into a thematically simpler global legend, so that African or Eurasian classes may be put into the full global context and to provide traceability to earlier map legends, especially that of [1]. To achieve this, the regional classes have been described through the Land Cover Classification System (LCCS). LCCS was developed by the FAO to analyze and cross-reference regional differences in land cover descriptions [4]. LCCS describes land cover according to a hierarchical series of classifiers in a dichotomous phase (vegetated or nonvegetated surfaces, terrestrial or aquatic/flooded, cultivated/managed or natural/semi-natural) and by four main attributes (life-form, fractional cover, leaf type and phenology).

The dual nature of the GLC2000 products, i.e., the regional maps (5, 6, 7, 8) assembled in a global synthesis, dictated a validation scheme based on two methods: a confidence-building method (also called quality control and based on a comparison with ancillary data) for the regional maps and a quantitative accuracy assessment based on a sampling of high-resolution sites for the global synthesis.

III. Quality Control

A. Objectives

Systematic quality control is imperative because recent global land-cover products, although of good overall quality, exhibit in some areas major errors that could be avoided by a careful review of the draft products. Such errors reduce the user's overall confidence in the products, even if the quantitative accuracy is high. Errors affecting accuracy of thematic maps can be caused by confusion between the land-cover classes (wrong label, missing classes) or can be spatial errors (wrong position of the boundary between classes, disappearance of small patches).

Systematic quality control is intended to meet two main objectives: the elimination of macroscopic errors and an increase in the overall acceptance of the land cover product by users. This quality control should be integrated into the classification procedure, with the results of the analysis employed for removing errors and improving the map.

Accuracy indices derived from the error matrix provide information on the quality of the map as a whole but cannot be used to characterize distinct areas of the map. Even when global land-cover maps are produced applying the same global algorithm to a homogenous dataset, the quality of the final product is not uniform in all the regions, but instead depends on the quality of observation conditions (cloud coverage, haze, etc.) and ancillary data used to parameterize the classification. In many cases, the land cover map is obtained using a complex classification procedure involving different steps where different algorithms are applied. As a consequence, it is not possible to derive a per-pixel confidence value and it is necessary to evaluate the accuracy of the results using reference data. The systematic quality control is a way of describing the spatial distribution of the macroscopic errors of a land cover classification.

The quality control can be considered as the last step of the map production or as the first level of accuracy assessment. For that reason, the quality control is conducted at the regional level, which is the same geographical scale as the production.

B. Procedure

Qualitative validation is based on a systematic descriptive protocol, in which each cell of the map is visually compared with reference material and its accuracy documented in terms of type of error, landscape pattern and land-cover composition. The grid size is adapted to the characteristics of the landscape, the map, and the reference material. For example, in the heart of the Sahara, the grid cells can be much larger than in the complex landscapes of Western Europe. A cell size of 200 to 400 km is proposed as a target for providing a good idea of the overall quality of a global product, keeping in mind that the goal of this exercise is a quick survey.

Each cell examined during the quality control procedure is characterized in detail by a few parameters: the composition and the spatial pattern of the cell, its comparison with other existing global land cover products, the overall quality of the cell, and the nature of any problems.

The cell composition is a key factor affecting the precision of a map because some land-cover classes (e.g., evergreen forests, deserts, water bodies) are easier to discriminate than others (e.g., deciduous forests or woodlands, grasslands, extensive agriculture). Information on the composition of the cell contributes to a better understanding of the errors and can help to stratify the population to improve the sampling design for the quantitative accuracy assessment. On the other hand, some users focus on specific land-cover classes and will be interested in a spatial representation of the errors for cells dominated by their class of interest.

It is widely recognized that the spatial pattern of the landscape influences the appearance or disappearance of land cover classes at varying resolution [9] and the area estimates derived from coarse resolution maps [10], [11]. Our quality control procedure allows some explanatory analysis, such as the spatial pattern of a given land-cover class and its associated accuracy.

C. Analysis of the Errors

The analysis of the data systematically recorded in the database allows for a definition of a typology of the errors.

- The delineation of a land-cover feature is accurate, but the label is wrong. In this case, the type of confusion must be specified in order to derive a thematic "distance" between the right and the wrong labels. It is more problematic to classify tropical forests as grasslands than to classify woodlands as savannas.

- The labels present in the cell are correct, but the delineation of the various features is wrong. If this case is the most frequent, it means that the spatial resolution (and eventually the preprocessing steps) precludes any accurate delineation of land-cover features. The first global land-cover products derived from AVHRR suffered from limitations, such as geolocation. The extreme case of this category occurs when no clear structures appear on the map. The land-cover map then corresponds more to a climatic stratification.
- One important land-cover feature is missing in the map or a feature is mapped while it is not present in the field. This is a particular case of combining a wrong label and an inaccurate delineation of the land-cover features. For example, it happens when specific features are derived from erroneous ancillary data, like planned infrastructures never actually built (dams).

D. Results

Asia, Africa, Europe, and Northern Eurasia were systematically examined with this procedure, while Oceania, North, and South America were not due to a lack of partners. Since results can vary from one region to another, we present here Northern Eurasia as a methodological example. A detailed presentation of the four continents would be too long for the objective of this paper.

The regional map of Northern Eurasia [5] contains 25 land-cover classes. This map was overlaid by 385 validation cells. We can compute the number of validation cells in which a class is considered as well identified and the number of cells in which it is poorly identified. Because a validation cell may be covered by up to five different classes, the total number of occurrences (1554) is higher than the number of validation cells. Table I identifies the classes with low accuracy in this region: needleleaf forest (18 to 27% of error) and coniferous shrubs (43%). Note that a box is covered by up to five different classes, which explains the high number of occurrences.

The label error is the most common (62), with a very limited number of errors in the definition of the limits (4), and missing classes (2). The absence of errors in the limit delineation illustrates the remarkable geometrical properties of the VEGETATION sensor, which allow for recognition of the landscape pattern even in the composite images.

Table II shows the interactions between the spatial pattern and the errors. As expected, most errors are found in heterogeneous landscapes. Fig. 1 illustrates the utility of the quality control for locating the errors on the map. The examination of the spatial distribution of errors in an exhaustive way provides to the user a reliable assessment of the strengths and the caveats of the product.

TABLE I
LAND-COVER CELLS CORRECTLY AND BADLY CLASSIFIED ACCORDING TO THE QUALITY CONTROL IN NORTHERN EURASIA

Land-cover classes	Correct	Wrong	Error
Light Everg. Needle. Forest	127	23	18.1%
Dark Everg. Needle. Forest	92	19	20.7%
Decid. Needle Forest	154	42	27.3%
Deciduous Broadleaf Forest	127	8	6.3%
Mixed Forest (50Nee-50Bro)	67	1	1.5%
Mixed Forest (60-80%Bro)	85		-
Mixed Forest (60-80% Nee)	65	3	4.6%
Regrowth forest	61	5	8.2%
Recent Burns	17		-
Coniferous Shrubs	21	9	42.9%
Deciduous Scrabs	103	2	1.9%
Tundra-Lichen Dominated	106	9	8.5%
Tundra-Moss Dominated	25	4	16.0%
Tundra-Shrub/Lichen Dominated	36	3	8.3%
Dry Grass (steppe)	148	3	2.0%
Cropland	65	1	1.5%
Riparian Vegetation	18	1	5.6%
Wetland/Shrublands	62	6	9.7%
Bogs	113	8	7.1%
Water	22		-
Bare soil and rocks	40	3	7.5%
Total	1554	150	9.7%

TABLE II
RELATIONSHIP BETWEEN THE SPATIAL PATTERN AND THE ERRORS

Spatial Pattern	Good	Medium	Bad	Total
Very homogeneous	48	5		53
Homogeneous	19	9		28
Heterogeneous	219	44	1	264
Very heterogenous	29	11		40
Total	315	69	1	385

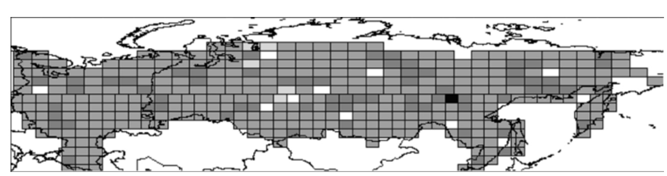


Fig. 1. Errors estimated in Northern Eurasia by the quality control procedure (four increasing levels of error from white to black).

IV. ACCURACY ASSESSMENT

A. Methods

The accuracy assessment of the GLC2000 map has a number of initial requirements.

1) The assessment should test in priority the main classes of interest of the GLC2000 map, i.e., forests, croplands and wetlands.
2) To provide a global accuracy of the product, sampling units should have a worldwide distribution since different teams have produced the regional products with different accuracies.
3) The landscape complexity has a major impact on the map accuracy (reference) and should be taken into account for the sampling and in the accuracy reporting.
4) For cost/efficiency reasons, the validation is derived from one single sensor dataset, Landsat, and the sampling design is adapted to it.
5) The sampling design should be based on an equal-area projection, since the sampling probability of a pixel should not be biased by its latitude.

6) To gain additional sampling units at low cost, we should include clustered sampling in the design.
7) As much as possible, regional experts should interpret the reference data. Most are independent from the "production" teams.
8) The interpretation key should be flexible and consistent with the rules used during the map production (Di Gregorio and Jansen, 2000).
9) The absolute location error of SPOT VEGETATION is about 300 m, which means that a validation protocol based on the analysis of single pixels can include up to 30% of inaccuracy. A single pixel-based is then too subject to error and we decided on an analysis of pixel blocks of 3×3 km.

1) Sampling Strategy: With all these constraints, we established a two-stage stratified clustered sampling. The stratification was based on the proportion of priority classes and on the landscape complexity. The two-stage clustering was selected for clear advantages of cost [12] and applied on the Landsat World Reference 2 System (WRS-2). The sampling strategy involves the following steps.

1) The WRS-2 grid provides a convenient sampling frame for a sample of Landsat scenes, as it is the case in the current validation. However, at high latitudes, a WRS-2 based sampling becomes very complicated due to the overlap between adjacent scenes that can represent 60% at 60° of latitude. To take into account this issue, Voronoï polygons are computed from the WRS-2 centroids in order to assign each GLC2000 pixel to one and only one scene. The Voronoï polygons are used for the sampling procedure.
2) The proportion of the priority classes (forests, wetlands and croplands) is calculated from the GLC 2000 map in each polygon. The polygon is flagged as "Priority" as soon as one of the three thresholds is satisfied: $> 30\%$ forests, $> 10\%$ croplands, $> 10\%$ wetlands.
3) The complexity of each polygon is estimated by the Shannon index **[13]**, which is a measure of diversity

$$H = -\sum_{k=1}^{m} (Pk) \cdot \log (Pk)$$

where Pk is the proportion of the landscape in cover type k, and m is the number of land cover types observed. The larger the value of H, the more diverse the landscape. The Shannon index of the GLC2000 cells follows a normal distribution centered on 0.5. The population is then split in two strata around this average value (> 0.5 and < 0.5).
4) Four strata are defined from the two criteria: homogenous landscapes in priority land-cover classes ($\mathrm{N} = 2267$), heterogeneous landscapes in priority land-cover classes ($\mathrm{N} = 2936$), homogenous landscapes in nonpriority land-cover classes ($\mathrm{N} = 2174$) and heterogeneous landscapes in nonpriority land-cover classes ($\mathrm{N} = 649$), where N is the total number of polygons in each stratum.
5) A sample grid of blocks (cells of 1800 $\times$ by 1200 km) is overlaid on the GLC2000 map reprojected in an equal-area projection. In each block, six fixed points are selected at a distance of 600 km in the two directions (Fig. 2). This distance was determined by the target sample size, which is defined by the budget available for the validation. This grid is crossed with the stratification and a number of replicates retained in each block is defined by stratum, with the highest sampling rate in the complex landscapes covered by priority classes (all the replicates are selected) and a minimum in the homogenous landscapes covered by nonpriority classes (2/6 replicates are selected). Through this procedure, also called systematic sampling on an irregular stratification with different sampling rates for each stratum **[14]**, we extract 253 Primary Sampling Units (PSU), with the following distribution by stratum: homogenous landscapes in priority land-cover classes ($\mathrm{n} = 65$), heterogeneous landscapes in priority land-cover classes ($\mathrm{n} = 120$), homogenous landscapes in nonpriority land-cover classes ($\mathrm{n} = 40$) and heterogeneous landscapes in nonpriority land-cover classes ($\mathrm{n} = 28$). Although the sampling was based on the Voronoï polygons, we use for efficiency reasons the full Landsat scenes during the validation and not only the Voronoï polygons (Fig. 3). That means that pixels present in the overlap area are sampled with higher probabilities than pixels in nonoverlap areas because they could possibly be selected from several different Landsat scenes. However, no significant differences have been found between overlap and nonoverlap areas, and, therefore, this local perturbation in the geographic distribution

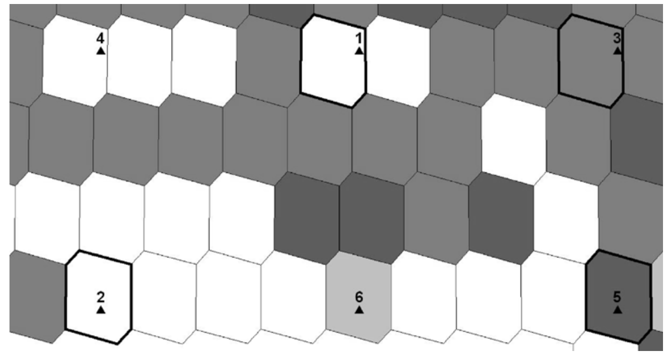


Fig. 2. Four strata are defined with high to low sampling probabilities. The sampling rate is as follows: 6/6 (gray), 4/6 (medium gray), 3/6 (light gray), and 2/6 (white). The grid of the Voronoï polygons is generated for the globe with equal numbers of each stratum randomly assigned to the polygons. To achieve the sampling rates, a network of points 600 km apart strung over the grid, with each point being randomly assigned a number from 1 to 6. A polygon is selected if the number assigned to it is 1 or 2 for the low strata; 1, 2, or 3 for the medium low; 1, 2, 3, or 4 for the medium high strata; and 1, 2, 3, 4, 5, or 6 for the high strata. Hence, in Fig. 3, the point assigned number 4 that falls on the white strata is not selected. Whereas the point 5 that falls on the dark gray strata is selected.

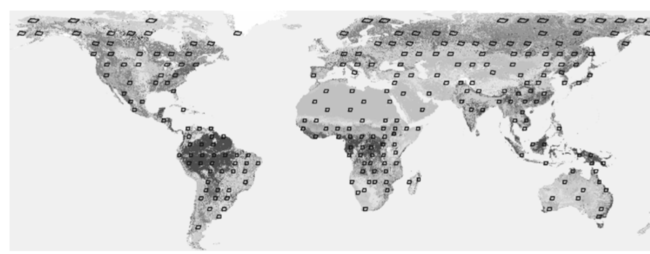


Fig. 3. Distribution of the Primary Sampling Units.

of sampling probability will have a minimal effect on the ultimate results. If we consider the orbital paths as a systematic sample with a random starting point, independent of the land cover, the final sampling probability is nearly constant within each stratum

$$\pi_h = \frac{a \times m \times n_h}{A_h}$$

where A_h and n_h are the area and sample size (first stage) in stratum h, $a = 9\ \text{km}^2$ is the area of each SSU and $m = 5$ is the number of SSU sampled in each SSU. π_h has a slight perturbation in the boundaries between strata.

6) For each Landsat scene, we extract five boxes of 3×3 km, the Secondary Sampling Units (SSU), at the centre of the Landsat scene and at each corner of a rectangle of 100×100 km centered on this first box. The procedure avoids sample units (boxes) too close to each other, reducing the impact of spatial autocorrelation and improving the precision of the accuracy estimates. Boxes were chosen for the interpretation in order to reduce the misregistration impact. In total, 1,265 SSU are interpreted.

2) Reference Data: The GLC2000 validation profited from the Landsat dataset for the year 2000 sponsored by NASA [15]. These Landsat scenes were orthorectified with a nominal accuracy of 50 meters. Data were downloaded from the Global Land Cover Facility (http://glcf.umiacs.umd.edu/index.shtml) or provided by the U.S. Geological Survey. Landsat channels 3, 4, 5, and 7 are used during the interpretation process. For each scene, a quick-look is created at 142.5-m spatial resolution and the SSU are extracted at full resolution. Regional interpreters used ancillary data like aerial photographs, thematic maps and NDVI profiles at coarse resolution in support to the Landsat interpretation.

3) Interpretation of the Reference Material: The analysis of each SSU is done by one partner with ecological knowledge of the local situation and expertise in fine resolution data interpretation. Within one continent, only a few teams were involved in the procedure for keeping the consistency of the interpretations.

A key challenge of the interpretation protocol is to respect the logic used during the classification scheme, i.e., a scheme based on objective classifiers that could be aggregated at different levels.

Each 3×3 km box is interpreted according to a series of classifiers describing the basic parameters of the landscape (vegetated/nonvegetated, natural/artificial, dominant layer), the water conditions (regime, seasonality, quality), detailing the tree, shrub and grass layers (cover, height, leaf type, and phenology). Some indications are also given on the reference material used for supporting the interpretation. When the box is covered by many spatially distinct land-cover classes, the two largest classes are described with the fraction of the box covered by each type. A simple interface was developed for storing the interpretations in a database. Table III details the different fields used for the characterization of the blocks.

Finally, each box is translated to the GLC2000 legend to measure the accuracy of this specific product. This translation was made easy by the fact that the classifiers were defined using the GLC2000 classification scheme.

TABLE III
DETAILED DESCRIPTION OF THE CRITERIA USED IN THE DATABASE FOR THE INTERPRETATION OF THE BOXES

Vegetated	Artificial	Dominant layer	
Vegetated	Artificial / managed	Bare soil	
Unvegetated	Natural	Herbaceous	
Mixed	Mixed	Lichens/mosses	
		Shrubs	
		Trees	
		Unknown	
Water regime	**Water seasonality**	**Water quality**	
Terrestrial	<4 months a year	Brackish	
Aquatic/regularly flooded	> 4 months a year	Fresh	
Mixed	Daily variations	Saline	
	Permanent	Unknown	
	Waterlogged		
	Unknown		
Tree cover	**Tree height**	**Leaf type Tree**	**Phenology Tree**
< 5%	5-15m	Broadleaved	Deciduous
5-10%	15-30m	Needleleaved	Evergreen
10-40%	>30m	Mixed	Mixed
40-70%	Unknown	Aphyllous	Unknown
70-100%		Spiny	
Unknown		Unknown	
Shrub cover	**Shrub height**	**Leaf type Shrub**	**Phenology shrub**
< 5%	< 1m	Broadleaved	Deciduous
5-10%	1-3 m	Needleleaved	Evergreen
10-40%	3-5m	Mixed	Mixed
40-70%	Unknown	Aphyllous	Unknown
70-100%		Spiny	
Undetermined		Unknown	
Grass cover	**Grass height**	**Grass phenology**	**Bare soil**
< 1%	0.03-0.3m	Deciduous	Ice/snow
1-5%	0.3-1m	Evergreen	Rocks
5-10%	1-3m	Mixed	Salt hardpans
10-40%		Unknown	Sands
40-70%			Stony
70-100%			Urban
Unknown			Unknown
Agriculture cycle	**Agriculture intensity**	**Agriculture pattern**	**Agriculture water**
Annual	Permanent intensity high	Contiguous fields	Rainfed
Perennial	Permanent intensity low	Scattered fields	Postfloodings
2 crops / year	Shifting cultivation	Sparse fields	Irrigated
	Fallows	Grouped fields	Unknown
	Unknown		

B. Preliminary Results

The analysis of the interpretations is undertaken using a series of tests, designed both to evaluate the validity of the GLC 2000 map and to understand better possible causes and the magnitude of disagreement. The use of pixel boxes as Secondary Sampling Units, although useful for mitigating for the misregistration effects, made the statistical analysis of the interpretations difficult. Indeed, each box can be covered in the GLC2000 map and in the Landsat interpretation by many land-cover classes. Therefore, we first present the heterogeneity of the blocks. Then, we examine the classical confusion matrix and present an adaptation of the confusion matrix to take into account class similarities both from the producer and from the user point of view.

1) Analysis of Heterogeneity: Due to the different spatial resolutions of the two data sets, we can expect to find many

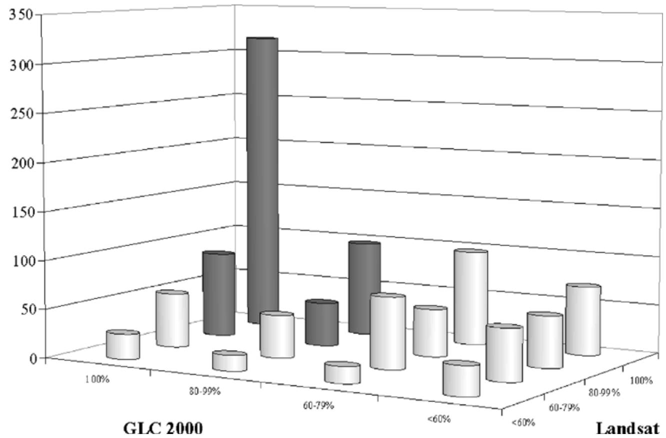


Fig. 4. Distribution of blocks according to the proportion of the dominant land-cover in the GLC 2000 map and in the Landsat interpretation. In dark gray, the blocks present a high homogeneity in GLC 2000 and in the Landsat interpretations.

cases in which the reference data have more than one class compared to the map data. We can reasonably expect in the case of a correct classification that a composite class in the map data (e.g., mosaic of forest and agriculture) would correspond to two classes in the reference data. We, therefore, examined the number of reference sites that contain single and multiple classes and compare these to our map data so as to give an idea as to the magnitude of this problem and what its effect is on the overall classification accuracy.

From the proportion of the box covered by the dominant land-cover, we can define four situations (Fig. 4).

1) The main land-cover class of each sample (GLC 2000 map and Landsat interpretation) covers more than 80% of the box. The cell is then considered as homogenous and is processed as a single point. It represents 544 boxes on a total of 1178, with 300 boxes purely covered by one class in the two populations.
2) The box is covered at 80% by one land-cover class in the GLC 2000 map, but by two classes in the Landsat interpretation. It means that the map overestimates the proportion of this land-cover class, even in case of correct classification (commission error). 301 boxes are in this case, with 196 covered at 100% by one class.
3) The box is covered by more than one land-cover class in the GLC 2000 map, but by one class in the Landsat interpretation. It means that the map underestimates the proportion of this land-cover class, even in the case of correct classification (omission error). 146 boxes are in this case, with 83 covered at 100% by one class.
4) The box is covered by more than one land-cover class in the GLC 2000 map and in the Landsat interpretation. This more complex situation is present in 196 blocks. The only way to measure the map accuracy in this case is to work with fuzzy logic [16].

As a first conclusion, we can say that dataset is very homogeneous. About 50% of the population is dominated by one land-cover class in both datasets. At the opposite, the situation mixing different land-cover types in both datasets represents less than 20% of the population.

2) Classical Accuracy Metrics: We produced a confusion matrix [17] for the 21 land cover classes so as to obtain a first measure of overall, the user's and the producer's accuracies. At this stage, we just give a flavor of the accuracy, taking into account only the 544 blocks dominated by one land-cover class at least 80%. Table VI details the classical confusion matrix and the user's and producer's accuracy.

We have used a two-stage systematic sampling plan slightly modified to take into account the geometry of Landsat TM frames. Systematic sampling is generally more efficient than random sampling, but there are no unbiased estimators for the variance under systematic sampling. We have used classical formulae for two-stage random sampling [18, Ch. 10], that over-estimate the variance. Therefore, we give pessimistic values for the standard error, although estimators with a smaller bias are possible [19].

For the global accuracy

$$v(\overline{p}_h) = \frac{1-f_1}{n_h(n_h-1)}\sum_i^{nh}(p_i-\overline{p}_h)^2 + \frac{f_1(1-f_2)}{n_h^2(\overline{m}-1)}\sum_i^{nh} p_i q_i$$

$$v(\overline{p}) = \frac{\sum_h v(p_h) a_h^2}{a^2}$$

where $\overline{p}_h$ is the average proportion of pixels correctly classified in stratum h, $p_i = 1 - q_i$ is the proportion of pixels correctly classified in PSU i,, n_h is the number of PSU sampled in the stratum, $\overline{m} \leq 5$ is the average number of boxes sampled in each PSU, and f_1 and f_2 are the sampling fractions in the first and second sampling stages, and a_h and a are the areas for each stratum and the total area. We obtain a standard error for the accuracy of 2.6%. This gives a (conservative) 95% confidence interval of $68.6 \pm 5\%$ for the overall accuracy.

The overall GLC2000 (21 classes) accuracy is similar to the IGBP DISCover (17 classes) [20]. We must recognize that the results expected for the other blocks (with no dominant land cover) should be less favorable. Their analysis is still on-going through methods derived from the fuzzy logic.

3) Adjusted Accuracy Matrices—The Producers Perspective: The heterogeneity analysis leads us to present a set of adjusted accuracy matrices, where we take into account class similarities both from the producer's point of view and from the user's point of view. These new matrices aim to quantify the magnitude of the error from different perspectives. In the classical confusion matrix a misclassification of a desert area as an evergreen forest has the same impact as classifying a semi-deciduous forest as an evergreen forest. From both a producer and a user's point of view, we need to present a matrix where misclassifications between similar classes are weighted lower than misclassifications between dissimilar classes. On what basis do we measure "similarity" then?

From the producer's point of view, two related parameters influence the separability between classes: the spectral separability in the dataset and the confusion in the legend definition. As many different regional classification techniques were used during the map production [3], it is not possible to measure globally the spectral separability. For estimating thematic similarity from the legend point of view, we measured the proximity

TABLE IV

(a) Matrix of Thematic Distance (in %) Between GLC 2000 Classes Based on the LCCS Classifiers With the Same Weight Given to the Dichotomous Phase and to the Attribute-Based Phase. (b) Matrix of Thematic Distance Between GLC 2000 Classes Based on the Four LCCS Classifiers With the Highest Impact on Separability (Aquatic Versus Terrestrial, Life Form, Phenology and Leaf Type)

	Tree Cover, broadleaved, evergreen (1)	Tree Cover, broadleaved, deciduous, dense (2)	Tree Cover, broadleaved, deciduous, open (3)	Tree Cover, needle-leaved, evergreen (4)	Tree Cover, needle-leaved, deciduous, open (5)	Tree Cover, mixed leaf type (6)	Tree Cover, regularly flooded, fresh (7)	Tree Cover, regularly flooded, saline (8)	Mosaic: Tree cover / Other natural vegetation (9)	Shrub Cover, closed-open, evergreen (10)	Shrub Cover, closed-open, deciduous (11)	Herbaceous Cover, closed-open (12)	Sparse Herbaceous or sparse shrub cover (13)	Regularly flooded shrub and/or herbaceous cover (14)	Cultivated and managed areas (15)	Mosaic: Cropland / Tree Cover / Other vegetation (16)	Mosaic: Cropland / Shrub and/or grass cover (17)	Bare Areas (18)	Water Bodies (natural & artificial) (19)	Snow and Ice (natural & artificial) (20)	Artificial surfaces and associated areas (21)
1	0																				
2	14	0																			
3	21	06	0																		
4	17	28	29	0																	
5	29	15	16	13	0																
6	17	15	16	13	13	0															
7	27	32	33	29	35	23	0														
8	21	32	33	29	42	29	06	0													
9	20	18	19	16	16	03	26	32	0												
10	14	24	26	16	28	16	26	26	16	0											
11	26	12	13	28	16	16	32	39	16	13	0										
12	25	17	18	33	27	21	37	37	18	18	11	0									
13	34	20	14	43	30	30	47	47	27	27	14	09	0								
14	37	41	42	39	45	32	09	16	29	23	29	28	37	0							
15	48	46	47	53	53	42	43	42	39	39	39	31	41	33	0						
16	22	26	28	24	30	18	28	34	31	31	38	43	52	31	29	0					
17	35	26	28	36	30	24	41	47	38	38	31	30	39	31	23	13	0				
18	65	63	57	60	60	60	77	77	60	60	54	51	42	68	75	68	66	0			
19	88	88	88	88	88	88	71	71	88	88	88	88	88	71	88	88	88	33	0		
20	83	83	83	83	83	83	83	83	83	83	83	83	83	83	92	92	92	29	33	0	
21	100	100	100	100	100	100	100	100	100	100	100	100	99	100	100	91	92	33	38	38	0

(a)

	Tree Cover, broadleaved, evergreen (1)	Tree Cover, broadleaved, deciduous, dense (2)	Tree Cover, broadleaved, deciduous, open (3)	Tree Cover, needle-leaved, evergreen (4)	Tree Cover, needle-leaved, deciduous, open (5)	Tree Cover, mixed leaf type (6)	Tree Cover, regularly flooded, fresh (7)	Tree Cover, regularly flooded, saline (8)	Mosaic: Tree cover / Other natural vegetation (9)	Shrub Cover, closed-open, evergreen (10)	Shrub Cover, closed-open, deciduous (11)	Herbaceous Cover, closed-open (12)	Sparse Herbaceous or sparse shrub cover (13)	Regularly flooded shrub and/or herbaceous cover (14)	Cultivated and managed areas (15)	Mosaic: Cropland / Tree Cover / Other vegetation (16)	Mosaic: Cropland / Shrub and/or grass cover (17)	Bare Areas (18)	Water Bodies (natural & artificial) (19)	Snow and Ice (natural & artificial) (20)	Artificial surfaces and associated areas (21)
1	0																				
2	25	0																			
3	25	0	0																		
4	25	50	50	0																	
5	50	25	25	25	0																
6	25	25	25	25	25	0															
7	38	50	50	50	63	38	0														
8	25	50	50	50	75	50	13	0													
9	31	31	31	31	31	06	44	56	0												
10	19	44	44	31	56	31	44	44	31	0											
11	44	19	19	56	31	31	56	69	31	25	0										
12	41	29	29	66	54	41	66	66	35	35	23	0									
13	49	24	24	74	49	49	74	74	43	43	18	08	0								
14	56	69	69	69	81	56	19	31	50	38	50	48	55	0							
15	49	49	49	69	69	46	48	46	40	41	41	25	33	28	0						
16	19	31	31	31	44	19	31	44	38	38	50	60	68	38	38	0					
17	44	31	31	56	44	31	56	69	50	50	38	35	43	38	25	25	0				
18	75	75	75	75	75	75	100	100	75	75	63	56	50	83	83	74	70	0			
19	100	100	100	100	100	100	75	75	100	100	100	100	100	75	88	100	100	50	0		
20	100	100	100	100	100	100	100	100	100	100	100	100	100	100	88	100	100	50	25	0	
21	100	100	100	100	100	100	100	100	100	100	100	100	98	100	100	99	100	25	50	50	0

(b)

between classes in the different LCCS classifiers. LCCS scheme combines two classifying phases, one dichotomous at three levels (i.e., a choice between two options—vegetated OR nonvegetated, artificial OR natural, aquatic OR terrestrial) and the other four based on attributes: life form (trees, shrubs, grasses, bare soil), phenology (evergreen, deciduous, mixed), leaf type (broadleaf, needleleaf, mixed), and vegetation cover (dense, open, sparse, bare). For each of the seven classifiers we developed a similarity matrix showing the correspondence between class pairs from 0 to 1. For example, the distance of "Tree cover broadleaved evergreen" in terms of phenology is 1 with "Tree cover broadleaved deciduous" and 0.5 with "Tree cover Mixed." The overall similarity is the combination of the seven parameters, ideally with the same weight given to the dichotomous phase and to the attribute-based phase for strictly respecting the LCCS approach. However, some classifiers can be highly correlated, like vegetal versus nonvegetal and the life form. Many land-cover classes are characterized by the same dichotomous classifiers and distinct by only one attribute, their overall similarity is then very high and it artificially augments the resulting map accuracy. For example, 11 classes covering 61% of the land belong to one single category defined during the dichotomous phase (vegetal, natural, terrestrial), that means that the numerical distance between these classes will reflect only the differences in the attributes (four classifiers on seven). For accounting for this artifact, we also calculated a similarity matrix with the four factors giving the best separability, i.e., aquatic/terrestrial, life form, phenology and leaf type. Table IV shows the similarity matrices with the two methods. The resultant matrix is now applied to the original confusion matrix by removing the errors due to thematic proximity in the legend definition, and by adding the similarities to the diagonal of the matrix.

If we call B the confusion matrix, the global accuracy can be written

$$p = \frac{\sum_{ij} B_{ij}\left(1-\delta_{ij}\right)}{\sum_{ij} B_{ij}}$$

where $\delta_{ij} = \begin{cases} 0, & \text{if } i=j \\ 1, & \text{if } i \neq j \end{cases}$ is the "sharp" thematic distance.

If δ is substituted by the fuzzy thematic distance φ given in Table IV, we get the fuzzy global accuracy

$$p_{\varphi} = \frac{\sum_{ij} B_{ij}\left(1-\varphi_{ij}\right)}{\sum_{ij} B_{ij}}.$$

In Table VII, we show the example of the confusion matrix adjusted by the similarity matrix computed from the four factors giving the best separability. The overall accuracy is now 90.3% between the 21 classes, with a very drastic improvement for the accuracy of the mosaic classes, and 92.6% with the similarity matrix weighting in the same way the dichotomous phase and the attribute-based phase. A similar approach was developed for the IGBP DISCover dataset, calculating the similarity of land-cover classes according to their Leaf Area Index and Surface Roughness properties. In this case, the overall accuracy increased by 10% to 13% [21].

TABLE V
CORRESPONDENCE BETWEEN THE GLC2000 CLASSES AND THE TREES CLASSES (ACHARD *et al.*, 2002)

GLC 2000 classes	TREES classes
Tree Cover, broadleaved, evergreen	Closed forests
Tree Cover, broadleaved, deciduous, closed	
Tree Cover, needle-leaved, evergreen	
Tree Cover, needle-leaved, deciduous	
Tree Cover, mixed leaf type	
Tree Cover, regularly flooded, fresh	
Tree Cover, regularly flooded, saline	
Tree Cover, broadleaved, deciduous, open	Open forests
Cropland / Tree Cover / Other vegetation	Degraded forests
Tree cover / Other natural vegetation	
Shrub Cover, closed-open, evergreen	Woody non-forest vegetation
Shrub Cover, closed-open, deciduous	
Mosaic: Cropland / Shrub and/or grass cover	
Herbaceous Cover, closed-open	Non woody vegetation
Sparse Herbaceous or sparse shrub cover	
Regularly flooded shrub and/or herbaceous	
Cultivated and managed areas	
Bare Areas	Unvegetated
Water Bodies (natural & artificial)	
Snow and Ice (natural & artificial)	
Artificial surfaces and associated areas	

4) Adjusted Accuracy Matrices—The Users Perspective: A number of different user perspectives can be envisaged for a global land cover database. Requirements for climate studies, biodiversity, land-cover change, carbon accounting may all have different requirements. Here, we present one example representing the needs of the climate change community for measuring land cover change in the tropics, a major uncertainty in IPCC calculations [22]. The classification relates to deforestation in the tropics (TREES project) [23]. The user requirements are to be able to establish changes between closed forests, open forests, mosaics, and nonforest classes. These categories in turn enable a measure of deforestation, degradation, and regrowth. While it is obvious that the reduction in classes will radically improve any accuracy measure, it must be remembered that this is not an accounting trick but reflects the real value of the classification to a particular set of users.

V. DISCUSSION AND CONCLUSIONS

- The two approaches (quality control and statistical accuracy assessment) are totally complementary. They do not evaluate the same products (regional on the one hand and global on the other hand) and provide different information (contextual and qualitative versus statistical). The classical accuracy assessment based on a sample of reference data gives a quantitative figure of the map accuracy, while the wall-to-wall quality control provides more exhaustive information on the nature of errors, their location and their relationship with the spatial pattern.
- As the secondary sampling units are interpreted according to classifiers and not according to a predefined

TABLE VI
CLASSICAL CONFUSION MATRIX OF THE 21 GLC 2000 CLASSES

	GLC2000																						
Landsat	Tree Cover, broadleaved, evergreen (1)	Tree Cover, broadleaved, deciduous, dense (2)	Tree Cover, broadleaved, deciduous, open (3)	Tree Cover, needle-leaved, evergreen (4)	Tree Cover, needle-leaved, deciduous, open (5)	Tree Cover, mixed leaf type (6)	Tree Cover, regularly flooded, fresh (7)	Tree Cover, regularly flooded, saline (8)	Mosaic: Tree cover / Other natural vegetation (9)	Shrub Cover, closed-open, evergreen (10)	Shrub Cover, closed-open, deciduous (11)	Herbaceous Cover, closed-open (12)	Sparse Herbaceous or sparse shrub cover (13)	Regularly flooded shrub and/or herbaceous cover (14)	Cultivated and managed areas (15)	Mosaic: Cropland / Tree Cover / Other vegetation (16)	Mosaic: Cropland / Shrub and/or grass cover (17)	Bare Areas (18)	Water Bodies (natural & artificial) (19)	Snow and Ice (natural & artificial) (20)	Artificial surfaces and associated areas (21)	Producer's accuracy (%)	Total
1	**62.0**		0.8	1.1					1.1							1.1						94.0	66.0
2		**11.6**										0.4										96.3	12.0
3	2.1	6.2	**2.1**						2.3		1.1											15.4	13.8
4		1.5		**19.3**																		92.7	20.9
5				1.1	**2.1**																	66.7	3.2
6	0.8	5.5		12.3	1.1	**12.5**				0.4	1.1											37.1	33.7
7	3.0			0.8			**1.5**															28.6	5.3
8	1.1							**1.1**														50.0	2.1
9									**1.5**						1.1							58.9	2.6
10	1.1								1.1		3.9	1.5	5.0		1.8								14.4
11		2.6	5.8	0.8					3.0		**12.1**	9.1	10.6	1.1	2.1			0.4				25.5	47.5
12		1.1	0.8						2.3		4.9	**12.4**			2.1		3.4					46.1	26.9
13											1.5	6.9	**20.1**		1.5			2.4				61.9	32.4
14	1.1	2.1		1.1								0.4		**3.7**	0.8		1.1					35.9	10.2
15				1.2								5.6	0.8		**49.1**	1.1	2.3	3.0	1.1			76.6	64.0
16	2.1	1.1	1.1	2.6		0.8									4.3	**9.7**	1.8	1.5				39.0	25.0
17		1.1									1.1				3.4								5.5
18												0.8	3.3		1.1		0.4	**111**				95.2	116.6
19																			**39.2**			100.	39.2
20																				**2.0**			2.0
21				0.8																			0.8
User's	84.6	35.3	20.3	47.2	66.7	94.2	100	100	13.6	-	47.3	33.4	50.5	77.4	73.0	82.0	0	93.8	97.4	100	-	**68.6**	
Total	73.3	32.8	10.5	41.0	3.2	13.3	1.5	1.1	11.2	0.4	25.6	37.1	39.7	4.7	67.2	11.9	9.0	118.	40.3	2.0	0.0	**373**	544

TABLE VII
CLASSICAL CONFUSION MATRIX OF THE 21 GLC 2000 CLASSES AFTER WEIGHTING BY THE THEMATIC DISTANCE CALCULATED FROM .THE FOUR MOST DISCRIMINANT LCCS CLASSIFIERS

	GLC2000																						
Landsat	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Producer's accuracy (%)	Total
1	**67.5**		0.2	0.3					0.3							0.2						98.6	68.5
2		**19.8**										0.1										99.4	19.9
3	0.5		**10.5**						0.7		0.2											88.0	12.0
4		0.8		**26.7**																		97.2	27.5
5				0.3	**2.9**																	91.7	3.2
6	0.2	1.4		3.1	0.3	**20.7**				0.1	0.3											79.3	26.1
7	1.1			0.4			**2.7**															63.6	4.2
8	0.3							**1.5**														84.6	1.7
9									**5.1**						0.4							92.3	5.6
10	0.2								0.3	**4.9**	1.0	0.5	2.1		0.8							49.8	9.8
11		0.5	1.1	0.4					0.9		**30.7**	2.0	1.8	0.5	0.9			0.3				78.3	39.2
12		0.3	0.2						0.8		1.1	**27.3**			0.5		1.2					86.9	31.4
13											0.3	0.5	**31.9**		0.5			1.2				92.8	34.3
14	0.6	1.5		0.7								0.2		**5.4**	0.2		0.4					59.8	9.0
15				0.8								1.4	0.2		**58.9**	0.4	0.6	2.5	0.9			89.6	65.7
16	0.4	0.3	0.3	0.8		0.1									1.6	**15.5**	0.5	1.1				75.0	20.7
17		0.3									0.4				0.8		**5.0**					76.0	6.6
18												0.4	1.7		0.9		0.3	**113**				97.2	116.6
19																			**39.3**			100	39.3
20																				**2.0**		100	2.0
21				0.8																		0.0	0.8
User's	95.3	79.6	85.2	78.0	91.7	99.3	100	100	62.2	97.2	90.4	83.9	84.4	91.0	89.9	96.3	63.1	95.7	97.7	100	-	**90.3**	
Total	70.8	24.8	12.4	34.3	3.2	20.8	2.7	1.5	8.3	5.0	33.9	32.6	37.8	5.9	65.5	16.1	7.9	118	40.2	2.0		491	544

classification, the validation dataset can be used for validating other products using a translation scheme adapted to this map. As long as the sampling probability of each sampling unit is known, we can recycle the sampling design for this specific map.

- The validation protocol is supposed to measure the accuracy of a map, but this protocol it self suffers from different types of errors. First, the interpretation of the validation samples although provided by regional experts can be wrong, especially in the case of fragmented and seasonal landscapes, where one image may not be sufficient to catch the correct land-cover type. A confidence flag given to each box could be an improvement. Secondly, the interpreter describes the box according the LCCS classifiers. These classifiers must be then translated to the GLC2000 legend, but conflicts between contradictory classifiers appear during this translation. Finally, when a box is composed by many land-cover classes, a decision is taken by the analyst to affect the box to one single value from the different values. Here again, some cases can provoke conflicts.
- The interpretation phase of the validation procedure can be definitely improved. In particular, the confusion between the thematic mixing and cartographic mixing within the box is not explained during the procedure and could be taken into account by a description of the spatial pattern in the Landsat box.
- Following IGBP DISCover and TREES, this is the third time that a global land-cover product has been validated according to a statistically designed method. The main innovation consists of the systematic quality control developed for evaluating the regional products. This approach was announced by [24]. Another innovation is the systematic use of user-oriented accuracy metrics, as suggested by [25]. A complete presentation of the results will be soon available.

References

[1] T. R. Loveland, Z. Zhu, D. O. Ohlen, J. F. Brown, B. C. Reed, and L. Yang, "An analysis of the IGBP global land-cover characterization process," *Photogramm. Eng. Remote Sens.*, vol. 65, pp. 1021–1032, 1999.

[2] M. A. Friedl, D. K. McIver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. Schaaf, "Global land cover mapping from MODIS: Algorithms and early results," *Remote Sens. Environ.*, vol. 83, pp. 287–302, 2002.

[3] E. Bartholomé and A. S. Belward, "GLC2000: A new approach to global land cover mapping from earth observation data," *Int. J. Remote Sens.*, vol. 26, pp. 1959–1977, 2005.

[4] A. Di Gregorio and L. Jansen, *Land Cover Classification System, Classification Concepts and User Manual*. Rome, Italy: Food and Agriculture Organization of the United Nations, 2000.

[5] S. Bartalev, A. S. Belward, D. Erchov, and A. S. Isaev, "A new SPOT4-VEGETATION derived land cover map of Northern Eurasia," *Int. J. Remote Sens.*, vol. 24, pp. 1977–1982, 2003.

[6] H. D. Eva, A. S. Belward, E. E. De Miranda, C. M. Di Bella, V. Gond, O. Huber, S. Jones, M. Sgrenzaroli, and S. Fritz, "A land cover map of South America," *Global Change Biol.*, vol. 10, pp. 1–14, 2004.

[7] P. Mayaux, E. Bartholomé, S. Fritz, and A. S. Belward, "A new land-cover map of Africa for the year 2000," *J. Biogeography*, vol. 31, pp. 861–877, 2004.

[8] H.-J. Stibig, F. Achard, and S. Fritz, "A new forest cover map of continental Southeast Asia derived from satellite imagery of coarse spatial resolution," *Appl. Veg. Sci.*, vol. 7, pp. 153–162, 2004.

[9] C. E. Woodcock and A. H. Strahler, "The factor of scale in remote sensing," *Remote Sens. Environ.*, vol. 21, pp. 311–325, 1987.

[10] A. Moody and C. E. Woodcock, "Scale-dependent errors in the estimation of land-cover proportions—Implications for global land-cover datasets," *Photogramm. Eng. Remote Sens.*, vol. 60, pp. 585–594, 1994.

[11] P. Mayaux and E. F. Lambin, "Estimation of tropical forest area from coarse spatial resolution data: A two-step correction function for proportional errors due to spatial aggregation," *Remote Sens. Environ.*, vol. 53, pp. 1–16, 1995.

[12] S. V. Stehman, "Practical implications of design-based sampling inference for thematic map accuracy assessment," *Remote Sens. Environ.*, vol. 72, pp. 35–45, 2000.

[13] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana, IL: Univ. Illinois Press, 1962.

[14] T. Richards, J. Gallego, and F. Achard, "Sampling for forest cover change assessment at the pan-tropical scale," *Int. J. Remote Sens.*, vol. 21, pp. 1473–1490, 2000.

[15] C. J. Tucker, D. M. Grant, and J. D. Dykstra, "NASA's global orthorectified landsat data set," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 3, pp. 313–322, 2004.

[16] C. E. Woodcock and S. Gopal, "Fuzzy set theory and thematic maps: Accuracy assessment and area estimation," *Int. J. Geograph. Inf. Sci.*, vol. 14, no. 2, pp. 153–172, 2000.

[17] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data; Principles and Practices*. New York: Lewis, 1999.

[18] W. Cochran, *Sampling Techniques*. New York: Wiley, 1977.

[19] K. M. Wolter, "An investigation of some estimators of variance for systematic sampling," *J. Amer. Statist. Assoc.*, vol. 79, no. 388, pp. 781–790, 1984.

[20] J. Scepan, "Thematic validation of high-resolution global land-cover data sets," *Photogramm. Eng. Remote Sens.*, vol. 65, no. 9, pp. 1051–1060, 1999.

[21] R. DeFries and S. Los, "Implications of land cover misclassifications for parameter estimates in global land surface models: An example from the simple biosphere model (SiB2)," *Photogramm. Eng. Remote Sens.*, vol. 65, no. 9, pp. 1083–1088, 1999.

[22] *IPCC, Climate Change 1995- The Science of Climate Change. The Second Assessment Report of the Inter-Governmental Panel on Climate Change*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[23] F. Achard, H. Eva, H. J. Stibig, P. Mayaux, J. Gallego, T. Richards, and J. P. Malingreau, "Determination of deforestation rates of the world's humid tropical forests," *Science*, vol. 297, pp. 999–1002, 2002.

[24] D. Muchoney, A. Strahler, J. Hodges, and J. LoCastro, "The IGBP DISCover confidence sites and the system for terrestrial ecosystem parametrization: Tools for validating global land-cover data," *Photogramm. Eng. Remote Sens.*, vol. 65, pp. 1061–1067, 1999.

[25] M. Herold, C. Woodcock, A. Di Gregorio, P. Mayaux, A. S. Belward, J. Latham, and C. C. Schmullius, "A joint initiative for harmonization and validation of land cover datasets," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1719–1727, Jul. 2006.

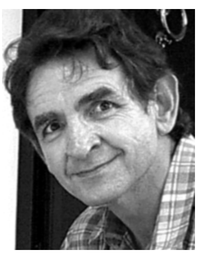
Philippe Mayaux was born in 1963 in Belgium. He received the Engineer in Forestry degree in 1986, the M.S. degree in land planning in 1988, and the Ph.D. degree in 1999 from the University of Louvain, Louvain, Belgium.

In the early 1990s, he worked for three years in field projects in Africa. He has worked for over 15 years in land cover mapping and tropical ecosystem monitoring from satellite imagery. He coordinated the African TREES project's activities, produced the Global Land Cover 2000 Map of Africa and Australia and is involved in many projects dealing with biodiversity and land-cover changes, in particular, in Central Africa. His research focused also on the scaling issues in remote sensing and on validation of regional to global scale land-cover products.

Dr. Mayaux is member of the Scientific and Technical Board of the GOFC-GOLD project.

Hugh Eva was born in 1958 in the U.K. He received the B.Sc. degree from Manchester University, U.K., the M.Sc. degree from Cranfield University, U.K., and the Ph.D. degree from the University of Louvain, Louvain, Belgium.

He specialized in the use of remotely sensed data for mapping fires and forests in tropical ecosystems and has published two land cover maps of South America derived from satellite imagery. He has been the Latin America coordinator of the TREES project, which was set up to monitor and measure changes in the tropical forest belt, using remote sensing.



Javier Gallego was born in 1953 in Spain. He graduated with a degree in mathematics in 1975 and the Ph.D. degree in statistics in 1979 from the University of Valladolid, Valladolid, Spain, and received the Doctorat de troisième cycle from the University of Paris VI, Paris, France, in 1982.

In 1983, he became Head of the Department of Statistics in the University of Valladolid, then Full professor in 1986. In 1987, he joined the European Commission, first in DG Agriculture, and then, in 1988, he joined the European Commission's Joint Research Centre, Ispra, Italy, where he has worked on the use of remote sensing for land cover and agricultural statistics (MARS Project), geographic sampling, and downscaling techniques.

Alan H. Strahler (M'86) received the B.A. and Ph.D. degrees in geography from The Johns Hopkins University, Baltimore, MD, in 1964 and 1969, respectively.

He is currently a Professor of geography and a Researcher in the Center for Remote Sensing, Boston University, Boston, MA. He has held prior academic positions at Hunter College, City University of New York; University of California, Santa Barbara; and the University of Virginia, Charlottesville. Originally trained as a Biogeographer, he has been actively involved in remote-sensing research since 1978. He has been a Principal Investigator on numerous NASA contracts and grants. His primary research interests are directed toward modeling the bidirectional reflectance distribution function (BRDF) of discontinuous vegetation covers and retrieving physical parameters describing ground scenes through inversion of BRDF models using directional radiance measurements. He is also interested in the problem of land cover classification using multitemporal, multispectral, multidirectional, and spatial information as acquired in reflective and emissive imagery of the earth's surface.

Dr. Strahler was awarded the AAG/RSSG Medal for Outstanding Contributions to Remote Sensing in 1993 and the honorary degree Doctorem Scientarum Honoris Causa from the Université Catholique du Louvain, Belgium, in 2000. He was also honored as a Fellow of the American Association for the Advancement of Science in 2003.

Martin Herold (S'98) was born in Leipzig, Germany, in 1975. He received the the graduate degree (Diplom in Geography) from the Friedrich Schiller University of Jena, Jena, Germany, and the Bauhaus University of Weimar, Weimar, Germany, in 2000, and the Ph.D. degree from the Department of Geography, University of California, Santa Barbara, in 2004.

He is currently coordinating the ESA GOFC GOLD Land Cover project office at the Friedrich Schiller University Jena. Earlier in his career, his interests were in multifrequency, polarimetric and interferometric SAR-data analysis for land surface parameter derivation, and modeling. He joined the Remote Sensing Research Unit, University of California Santa Barbara, in 2000, where his research has focused on the remote sensing of urban areas and the analysis and modeling of urban growth and land use change processes. His most recent interests are in international coordination and cooperation toward operational terrestrial observations with specific emphasis on the harmonization and validation of land cover datasets.

Dr. Herold is a member the German Society of Photogrammetry and Remote Sensing (DGPF) and the Thuringian Geographical Union (TGG).

Shefali Agrawal was born in India. She received the degree in physics, the M.S. degree in physics, and the PG Diploma in computer management and received an Advance Diploma in Geoinformatics from ITC, The Netherlands.

She joined the Indian Institute of Remore Sensing, Dehradun, in 1992, and, since then, she has been working in the field of image processing and GIS-based modeling related to land use/land cover characterization. She contributed to the GLC map of South Central Asia and its validation, as well as the TREES-II project (for Indian region), sponsored by the European Commission.

Sergey Naumov was born in 1977 in Russia. He graduated as a Research Engineer in GIS in 1999 and received the Ph.D. degree from the Moscow State University of Geodesy and Cartography, Moscow, Russia, in 2002.

After graduation, he worked for a Russian satellite imagery provider as a specialist for imagery processing and three-dimensional modeling of urban areas. For the past three years, he has been working with the Remote Sensing Data Processing Department, DATA+, Moscow, as a GIS and Remote Sensing Specialist.

Evaristo Eduardo de Miranda was born in 1952 in São Paulo, Brazil. He graduated from the "Institute Superieur d'Agriculture Rhône Alpes" in Lyon, France, and received the M.Sc. and Ph.D. degrees in ecology from the University of Montpellier, France, in 1980.

In his professional career, he has been working with rural development issues, linking environmental protection to agricultural production and satellite monitoring. He was a UN Consultant for the World Environmental Summit, Rio de Janeiro, Brazil, in 1992, and he is a Consultant for organizations such as FAO, UNESCO, FAPESP, World Bank, AEO, and other national and international research organizations. He has published hundreds of technical and scientific papers in Brazil, as well as abroad. He is a teacher and advisor for the University of São Paulo, São Paulo, Brazil, and at the State University of Campinas, Campinas, Brazil.

Carlos M. Di Bella was born in Argentina. He graduated as an Agronomist from the University of Buenos Aires, Buenos Aires, Argentina, in 1994, and received the Ph.D. degree from the Institut National Agronomique Paris-Grignon, France, in 2002.

Since 1998, he has been a Staff Researcher at the Institute of Climate and Water, Buenos Aires. His research focuses on the application and development of remote-sensing and GIS techniques to natural resources management and monitoring.

Callan Ordoyne received the B.A. degree in biological sciences from Mount Holyoke Colleg, South Hadley, MA, in 2003. She is currently pursuing the M.S. degree, for which she is studying the remote sensing of wetlands, tracking the annual hydrological pattern of the Florida Everglades, Boston University, Boston, MA.

Her undergraduate field was ecology, with a research focus on wetland ecology and arachnology. At Boston University, she worked on the MODIS Global Land Cover project, with the primary responsibility for developing and updating the System for Terrestrial Ecosystem Parameterization database of training sites.

Yuri Kopin was born in 1981 in Russia. He received the B.S. degree in cartography from the Lomonosov Moscow State University, Mosow, Russia, in 2002.

His main activities include land cover classification, remote sensing in cadastre, and radar mapping.

P. S. Roy received the M.S. degree in botany in 1973 and the Ph.D. degree in ecology in 1980.

He is the Deputy Director (RS and GIS–AA) National Remote Sensing Agency, Department of Space, Government of India, Hyderabad, India. He specializes in Geoinformatics for Ecosystem Analysis. Presently, he is leading a national forum on Biodiversity Characterization at Landscape Level and Land Use Land Cover in different scales.