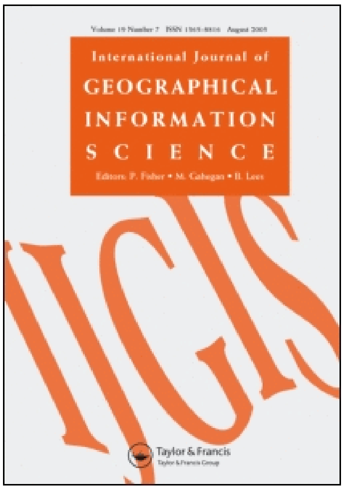


International Journal of Geographical Information Science
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713599799

Comparison of land cover maps using fuzzy agreement

S. Fritz [a]; L. See [b]
[a] Global Vegetation Monitoring Unit (GVM) Unit, Institute for Environment and Sustainability (IES), EC Joint Research Centre, I-21020 Ispra (VA), Italy [b] School of Geography, University of Leeds, Leeds, LS2 9JT, UK

Online Publication Date: 01 August 2005

To cite this Article Fritz, S. and See, L.(2005)'Comparison of land cover maps using fuzzy agreement',International Journal of Geographical Information Science,19:7,787 — 807

To link to this Article: DOI: 10.1080/13658810500072020

URL: http://dx.doi.org/10.1080/13658810500072020

International Journal of Geographical Information Science
Vol. 19, No. 7, August 2005, 787–807

Research Article

Comparison of land cover maps using fuzzy agreement

S. FRITZ*† and L. SEE‡
†Global Vegetation Monitoring Unit (GVM) Unit, Institute for Environment and Sustainability (IES), EC Joint Research Centre, TP 440, Via Enrico Fermi, 1, I-21020 Ispra (VA), Italy
‡School of Geography, University of Leeds, University Road, Leeds, LS2 9JT, UK

(*Received 30 June 2004; in final form 16 December 2004*)

A generic problem associated with different land cover maps that cover the same geographical area is the use of different legend categories. There may be disagreement in many areas when comparing different land cover products even though the legend shows the same or very similar land cover class. To capture the uncertainty associated with both differences in the legend and the difficulty in classification when comparing two land cover maps, expert knowledge and a fuzzy logic framework are used to map the fuzzy agreement. The methodology is illustrated by comparing the Global Land Cover 2000 data set and the MODIS global land cover product. Overall accuracy measures are calculated, and the spatial fuzzy agreement between the two land cover products is provided. This approach can be used to improve the overall confidence in a land cover product, since areas of severe disagreement can be highlighted, and areas can be identified that require further attention and possible re-mapping.

Keywords: Fuzzy logic; Global land cover mapping; Map comparison; Accuracy assessment

1. Introduction

National, regional and global land cover maps are largely produced from remotely sensed data, and provide information about the types of features found on the Earth's surface. Land cover information is a vital input to the planning, management and monitoring of a wide range of areas including forestry, agriculture, transportation and telecommunications. Issues such as sustainable management of forests and other land resources, forest conservation and restoration, and extension of surfaces dedicated to agriculture, desertification or watershed degradation will all substantially benefit from the availability of accurate baseline land cover information (United Nations 2002). Land cover also provides the boundary conditions for a number of climate and land surface process models. Accurate information on land cover is therefore essential.

Accuracy assessment is an important part of land cover mapping, but it is recognized as difficult to quantify and express (Foody 2002). One of the main approaches used is design-based inference, which involves sampling the map and assessing the accuracy against known values. As part of an accuracy assessment exercise, it is also possible to compare two different land cover products in order to

*Corresponding author. Email: steffen.fritz@jrc.it

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online © 2005 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/13658810500072020

highlight any differences and/or obvious errors. This is referred to as a confidence building approach (Strahler 2002) and is specifically addressed in this paper for the comparison of the Global Land Cover Map 2000 (GLC-2000) (Fritz *et al.* 2003) with the MODIS land cover product (Friedl *et al.* 2002). Giri *et al.* (2005) have already undertaken an initial comparison of these land cover products, but the approach presented here takes two additional considerations into account: (1) how to compare maps that have two different legend categories; and (2) how to capture classification uncertainty in order to create a map of spatial agreement/disagreement. Uncertainty can arise when classifications are created by different people using different methods (Lund 1993), which is true for both the GLC-2000 and MODIS classifications. The methodology outlined in this paper first reconciles the different map legends of the GLC-2000 and MODIS land cover maps and then outlines how classification uncertainty is incorporated into the map comparison using fuzzy logic. The resulting fuzzy agreement/disagreement map allows one to identify (1) where differences between the two land cover maps occur; and (2) the degree of agreement or disagreement. This methodology does not indicate which map is more accurate; rather, it can be used as a starting-point for further investigation and indicate where more effort is required to improve the quality of either the GLC-2000 or MODIS land cover products. Moreover, the methodology presented here can be generalized for use with any map comparison. Existing approaches to map comparison are first reviewed, followed by details of the GLC-2000 and MODIS land cover products. The methodology is then outlined followed by the results of the map comparison. Specific issues that were raised as part of the map comparison process are discussed in further detail, followed by a brief description of future work in this area.

2. Review of existing approaches

There are few existing approaches of map comparisons that take into account both the problem of legend reconciliation and classification uncertainty. Hagen (2003) has developed fuzzy evaluation measures for comparing raster maps of categorical data. The uncertainty associated with the categories as well as neighbourhood effects that incorporate distance decay have been considered in his approach. A fuzzy kappa measure is derived, which represents an overall measure of similarity, as well as a map showing the degree of similarity on a cell-by-cell basis. Hagen then demonstrates how his approach can be used to validate cellular automata model results that predict temporal land-use changes in Dublin, Ireland. However, in his example, the two maps have identical legend categories.

The issue of comparing maps where legend definitions differ has recently been tackled by Comber *et al.* (2004), who have used expert knowledge to gather information on how different land cover data sets (in this case, the Land Cover Map of Britain 1990 and 2000) can be compared when the land cover classes have also changed their meaning. In the end, it was decided that it was not possible to compare these two maps directly so instead they used Dempster–Shafer theory and expert knowledge to capture the likelihood of a change over time.

The uncertainty approach adopted in this paper has been modelled on the work done by Gopal and Woodcock (1994) and Woodcock and Gopal (2000) although the application of a map comparison is somewhat different. In their studies, they asked experts to use a linguistic scale to capture their perception of how well a land use class at a given reference site was described by a map category. Experts were asked to express this comparison as one of five linguistic values ranging from absolutely wrong (1) to absolutely right (5). These fuzzy values were then hardened using MAX and RIGHT Boolean functions to create a set of tables describing the nature, distribution, magnitude and frequency of the errors. Although their approach has the advantage of providing information on errors to the end user, the spatial distribution of the error is not one of the outputs. In our approach, we have also used a linguistic scale ranging from 1 to 5 to capture how difficult experts found it to distinguish between legend classes (see figure 1). These values have then been mapped onto a fuzzy set characterizing the degree of agreement between the two land cover maps. More details are provided in section 4.

3. GLC-2000 and MODIS land cover maps

The GLC-2000 dataset was created by the Global Vegetation Monitoring Unit of the Joint Research Centre (JRC) of the European Union in collaboration with partners around the world (Fritz *et al.* 2003). The objective of the project was to provide a global, harmonized land cover database for the year 2000, a reference year for environmental assessment and in particular the United Nation's Ecosystem-related International Conventions. The GLC-2000 makes use of a dataset of 14 months of pre-processed daily global data acquired by the VEGETATION instrument on board the SPOT 4 satellite (Bartholomé and Belward 2005). In contrast to other global mapping initiatives, the GLC-2000 project employed a bottom-up approach to global mapping in which more than 30 research teams contributed to 19 regional windows. Each defined region was mapped by local experts using local knowledge. Partners were given free control to use the most appropriate mapping techniques for their region in both the production of the seasonal mosaics and the classification techniques employed. Nevertheless, almost all the partners chose to use an unsupervised classification approach, followed by labelling of clusters based on ancillary datasets. The consistency between the regional classifications was ensured through the use of the LCCS (Land Cover Classification System) produced by FAO and UNEP (Di Gregorio and Jansen

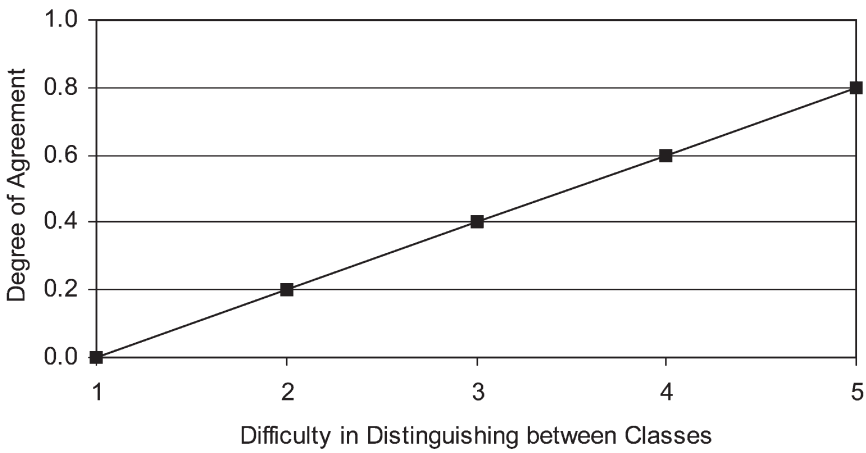


Figure 1. Fuzzy set characteriszing the degree of agreement between the two land cover maps. Linguistic values of 1 (easy to distinguish between two classes) to 5 (hard to distinguish) are mapped onto a linear fuzzy set with values between 0 and 0.8. The fuzzy membership value of 1 is reserved for 100% agreement between land cover classes.

2000). This hierarchical classification system allowed each partner to choose the most appropriate land cover classification for their region, while also providing the possibility to translate regional classes to a more generalized global legend. A preliminary accuracy assessment has been carried out with Landsat TM images within Africa, South America, and Canada, and a comparison has been made with national forest statistics in Russia (Bartalev *et al.* 2003), showing that these regional products correspond quite well with forest cover (Bartalev *et al.* 2003, Eva *et al.* 2004, Mayaux *et al.* 2004, Stibig *et al.* 2004). A global validation exercise is now under way for the GLC-2000 product using a stratified random sampling approach. The validation exercise and the work presented here fit in well with the current activities of the GOFC-GOLD (Global Observation of Forest and Land Cover) implementation team and the GLN (Global Land Cover Network), a FAO/UNEP initiative. One of their aims is to increase the availability of reliable and standardized information on land cover validation and comparative validation for existing and future datasets, and the harmonization of land cover definitions (Anonymous 2004).

The Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Terra Satellite was launched in December 1999. It is the primary sensor for monitoring the terrestrial ecosystem for the NASA Earth Observing System at a resolution of 250–1000 m (Liang *et al.* 2002, Morisette *et al.* 2002). MODIS, which includes seven bands (out of 36) that are particularly designed for land applications, has enhanced spatial, spectral, radiometric and geometric qualities when compared with the AVHRR satellite (Friedl *et al.* 2002). A number of higher-level products relevant to global change research are produced including radiation budget and ecosystem variables, and land cover characteristics (Morisette *et al.* 2002). Two main parameters are provided by the MODIS land cover product suite: a land cover parameter at 1 km spatial resolution updated quarterly and a land cover dynamics parameter (Friedl *et al.* 2002). In order to produce the land cover parameter, a supervised classification methodology has been used, based primarily on the C4.5 decision tree methodology, which also includes robust and mature procedures for missing data (Friedl *et al.* 2002). Training sites from the System for Terrestrial Ecosystem Parameterization (STEP) database were used and consist mainly of interpreted Landsat TM scenes, acquired after the 1990s. The 1370 sites were interpreted by Boston University analysts (Friedl *et al.* 2002), and additional sites were then selected from the Olsen land cover classes to ensure a good global representation of the different ecosystems. The MODIS version 3 land cover product has been validated, and the global level-1 product is estimated to have an accuracy of 75–80% globally (Hodges 2002). However, the sites for validation were not chosen randomly as in the case of the GLC-2000 validation exercise.

4. Methodology

4.1 *Reconciling the GLC-2000 and MODIS map legends*

The GLC-2000 dataset has 22 classes derived from the Land Cover Classification System provided by the FAO, while the MODIS land cover map has only 17 categories in the IGBP (International Global Biosphere Project) legend (Loveland *et al.* 2000). Tables 1 and 2 compare the GLC-2000 and the IGBP legends. Table 1 shows how the IGBP legend compares to the GLC-2000, and table 2 provides the opposite case. These tables clearly show that some classes are identical, some

Table 1. GLC-2000 land cover class definitions and the corresponding IGBP land cover classes.

GLC Class	GLC-2000 land cover class definitions	Contains/overlaps IGBP land cover class (see table 2)
1	Tree Cover, broadleaved, evergreen, LCCS >15% tree cover, tree height >3 m closed >40% tree cover; open 15–40% tree cover	2, 8, 9
2	Tree Cover, broadleaved, deciduous, closed LCCS >15% tree cover, tree height >3 m Examples of sub-classes at regional level: closed >40%	4, 8
3	Tree Cover, broadleaved, deciduous, open tree cove; open 15–40% tree cover	8, 9
4	Tree Cover, needle-leaved, evergreen LCCS >15% tree cover, tree height >3 m closed >40% tree cove; open 15–40% tree cover	1, 8, 9
5	Tree Cover, needle-leaved, deciduous LCCS >15% tree cover, tree height >3 m closed >40% tree cove; open 15–40% tree cover	3, 8, 9
6	Tree Cover, mixed leaf type	5, 8, 9
7	Tree Cover, regularly flooded, fresh water (& brackish)	1, 2, 3, 4, 5, 8, 9, 11
8	Tree Cover, regularly flooded, saline water, (daily variation of water level)	1, 2, 3, 4, 5, 8, 9, 11
9	Mosaic: Tree cover/Other natural vegetation	1, 2, 3, 4, 5, 8, 9
10	Tree Cover, burnt	1, 2, 3, 4, 5, 8, 9
11	Shrub Cover, closed-open, evergreen More than 10–20% vegetation, cover, closed 60% to 70%, open 10–20% vegetation, refers to the time of its fullest development	6, 7
12	Shrub Cover, closed-open, deciduous More than 10–20% vegetation, cover, closed 60% to 70%, open 10–20% vegetation, refers to the time of its fullest development	6, 7
13	Herbaceous Cover, closed-open Examples of sub-classes at regional level: (i) natural, (ii) pasture, (iii) sparse trees or shrubs	9, 10
14	Sparse Herbaceous or sparse Shrub Cover Vegetation cover between 1% and 10–20% vegetation cover	7, 10, 16
15	Regularly flooded Shrub and/or Herbaceous Cover	6, 7, 8, 9, 10, 11
16	Cultivated and managed areas Examples of sub-classes at reg. level: (i) terrestrial; (ii) aquatic (=flooded during cultivation), and under terrestrial: (iii) tree crop & shrubs (perennial), (iv) herbaceous crops (annual), non-irrigated, (v) herbaceous crops (annual), irrigated	12
17	Mosaic: Cropland/Tree Cover/Other natural vegetation	14
18	Mosaic: Cropland/Shrub or Grass Cover	14
19	Bare Areas	16
20	Water Bodies (natural & artificial)	17
21	Snow and Ice (natural & artificial)	15
22	Artificial surfaces and associated areas	13

overlap partially, and a number of classes are not present in either the IGBP or the GLC-2000 legend. This means that a direct comparison of the GLC-2000 map and the MODIS map (with the IGBP legend) is not straightforward. A number of steps are required in order to make such a comparison. First, the definition of the classes that do not directly correspond must be taken into account. Because of differences in the vegetation percentage cover, only a partial overlap exists. This means that, for

Table 2. IGBP land cover class definitions and the corresponding GLC-2000 land cover classes.

IGBP class	IGBP legend definitions	Contains/overlaps GLC class (see table 1)
1	Evergreen needleleaf forests are dominated by trees with a percent canopy cover of greater than 60% and height exceeding 2 m. Almost all of its trees remain green all year. Its canopy is never without green foliage.	4, 7, 8, 9, 10
2	Evergreen broadleaf forests are dominated by trees with a percent canopy cover of greater than 60% and height exceeding 2 m. Almost all of its trees remain green year all year. Its canopy is never without green foliage.	1, 7, 8, 9, 10
3	Deciduous needleleaf forests are dominated by trees with a percent canopy cover of greater than 60% and height exceeding 2 m. It consists of seasonal needleleaf tree communities with an annual cycle of leaf-on and leaf-off periods.	5, 7, 8, 9, 10
4	Deciduous broadleaf forests are dominated by trees with a percent canopy cover of greater than 60% and height exceeding 2 m. It consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.	2, 7, 8, 9, 10
5	Lands dominated by trees with a percentage cover >60% and height exceeding 2 m. Consists of tree communities with interspersed mixtures or mosaics of the other four forest types.	6, 7, 8, 9, 10
6	Closed shrublands are lands with woody vegetation less than 2 m tall and with shrub canopy cover greater than 60%. The shrub foliage can be either evergreen or deciduous.	11, 12
7	Open shrublands are lands with woody vegetation less than 2 m tall and with shrub canopy cover is between 10–60%. The shrub foliage can be either evergreen or deciduous.	11, 12
8	Woody Savannahs are lands with herbaceous and other understory systems, and with forest canopy cover between 30–60%. The forest cover height exceeds 2 m.	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
9	Non-Woody Savannahs are lands with herbaceous and other understory systems, and with forest canopy cover between 10–30%. The forest cover height exceeds 2 m.	1, 3, 4, 5, 6, 7, 8, 9, 10
10	Grasslands are lands with herbaceous types of cover. Tree and shrub cover is less than 10%.	13, 15
11	Permanent wetlands are lands with a permanent mixture of water and herbaceous or woody vegetation that cover extensive areas. The vegetation can be present in salt, brackish, or fresh water.	7, 8, 15
12	Croplands are lands covered with temporary crops followed by harvest and a bare soil period (e.g., single and multiple cropping systems). Note that perennial woody crops will be classified as the appropriate forest or shrub land cover type.	16
13	Urban and built-up areas are covered by buildings and other man-made structures. Note that this class will not be mapped from the AVHRR imagery but will be developed from the populated places layer that is part of the Digital Chart of the World (Defense Mapping Agency, 1992).	22

Table 2. *(continued)*

IGBP class	IGBP legend definitions	Contains/overlaps GLC class (see table 1)
14	Cropland/natural vegetation mosaics are lands with a mosaic of croplands, forests, shrublands, and grasslands in which no one component comprises more than 60% of the landscape.	17, 18
15	Snow and ice covered areas are lands under snow and/or ice cover throughout the year.	21
16	Barren and sparsely vegetated areas are lands of exposed soil, sand, rocks, or snow and never has more than 10% vegetated cover during any time of the year.	14, 19
17	Water bodies are oceans, seas, lakes, reservoirs, and rivers. They can be either fresh or salt water bodies.	20

example, the class forest cover with more than 15% tree cover (GLC-2000) can be either woody savannah or savannah according to the definition of the IGBP. Moreover, one IGBP class can also correspond to two or more classes of the GLC-2000 legend. This is the case for the single 'Barren or sparsely vegetated' class of the IGBP, which corresponds to both the 'Bare areas' and the class 'Sparse Herbaceous or sparse Shrub Cover' of the GLC-2000 legend. Second, if a similar class between the GLC-2000 and IGBP does not exist, a number of possibilities must be considered. One class in the GLC-2000 could, therefore, potentially correspond to more than one IGBP class and vice versa. The different possibilities that must be considered when the GLC-2000 land cover map is compared with the IGBP legends of the MODIS land cover map are provided in a matrix in table 3. A cross (X) indicates that the two land cover types of the GLC-2000 and MODIS are comparable and are therefore not in disagreement. Hence, a comparison between the land cover maps is only possible if such a matrix or lookup table of legend combinations is taken into consideration.

4.2 *Uncertainty in classification from expert knowledge*

Classification of the 19 regions that make up the GLC-2000 was carried out by different teams of people. Although the basis of the classification was the FAO LCCS to provide a standardized system of defining classes, the difficulty involved in classifying different areas will vary between classes and between different experts. To capture this uncertainty in classification, expert knowledge was gathered from the GLC-2000 partners. A questionnaire was administered which took the form of an empty matrix, similar in form to a confusion matrix containing the 22 GLC-2000 classes. The experts were then asked to express how easy or difficult it was for them to differentiate between different land cover classes using five linguistic categories: very easy (1); easy (2); medium (3); difficult (4); very difficult (5). For example, it is very easy to classify water as opposed to herbaceous cover, but the degree of difficulty may be much higher when considering two similar types such as herbaceous and shrub cover. The latter is particularly difficult, since the under story layer of the shrub might be herbaceous, and the shrub cover might be relatively sparse. An expert involved in the development of the MODIS classification was also

Table 3. Correspondence between GLC2000 and MODIS land cover classes in matrix form.

	Evergreen Needleaved Forest	Evergreen Broadleaf Forest	ous Needleleaf Forest	ous Broadleaf Forest	Mixed Forests	Closed Shrublands	Open Shrublands	Woody Savannahs	Savannahs	Grasslands	Permanent Wetlands	Croplands	ban and Built-Up	Cropland/Natural Veg	Snow and Ice	Barren or Sparsely Vegetated	Water Bodies
Tree Cover, broadleaved, evergreen		X						X	X								
Tree Cover, broadleaved, dec. (closed)				X				X									
Tree Cover, broadleaved, dec. (open)								X	X								
Tree Cover, needle-leaved, evergreen	X							X	X								
Tree Cover, needle-leaved, deciduous			X					X	X								
Tree Cover, mixed leaf type					X			X	X								
Tree Cover, regularly flooded, fresh water (& brackish)	X	X	X	X	X			X	X		X						
Tree Cover, reg. flooded, saline water	X	X	X	X	X			X	X		X						
Mosaic: Tree cover/Other natural veg.	X	X	X	X	X			X	X								
Tree Cover, burnt	X	X	X	X	X			X	X								
Shrub Cover, closed-open, evergreen						X	X										
Shrub Cover, closed-open, deciduous						X	X										
Herbaceous Cover, closed-open									X	X							
Sparse Herbaceous or sparse Shrub Cover																X	
Regularly flooded Shrub and/or Herbaceous Cover						X	X	X	X	X	X						
Cultivated and managed areas												X					
Mosaic: Cropland/Tree Cover/Other natural vegetation														X			
Mosaic: Cropland/Shrub or Grass Cover														X			
Bare Areas																X	
Water Bodies (natural & artificial)																	X
Snow and Ice (natural & artificial)															X		
Artificial													X				

given a questionnaire, but the GLC-2000 categories were replaced with the 17 classes used in the MODIS data set. An example of the questionnaire given to the MODIS expert with responses is provided in table 4.

4.3 *Mapping expert responses onto a fuzzy agreement membership function*

This section explains how the expert responses, which were collected in terms of how difficult it is to distinguish between two classes, map onto a fuzzy membership function that represents the degree of agreement between two land cover products. If, at a given pixel, the two land cover types match (as given by the correspondence table in figure 2), the cell is automatically given a value of 1, i.e. there is 100% agreement or no disagreement between the two land cover types. For all other situations in which the two land cover maps do not agree, the cell is assigned a value between 0 and 0.8, denoting increasing degrees of agreement between the maps, corresponding to values of 1–5. If an expert found it very easy to distinguish between the two classes, but they were incorrect when comparing the two land cover products, the severity of disagreement is high or assigned a value of 0 (i.e. no

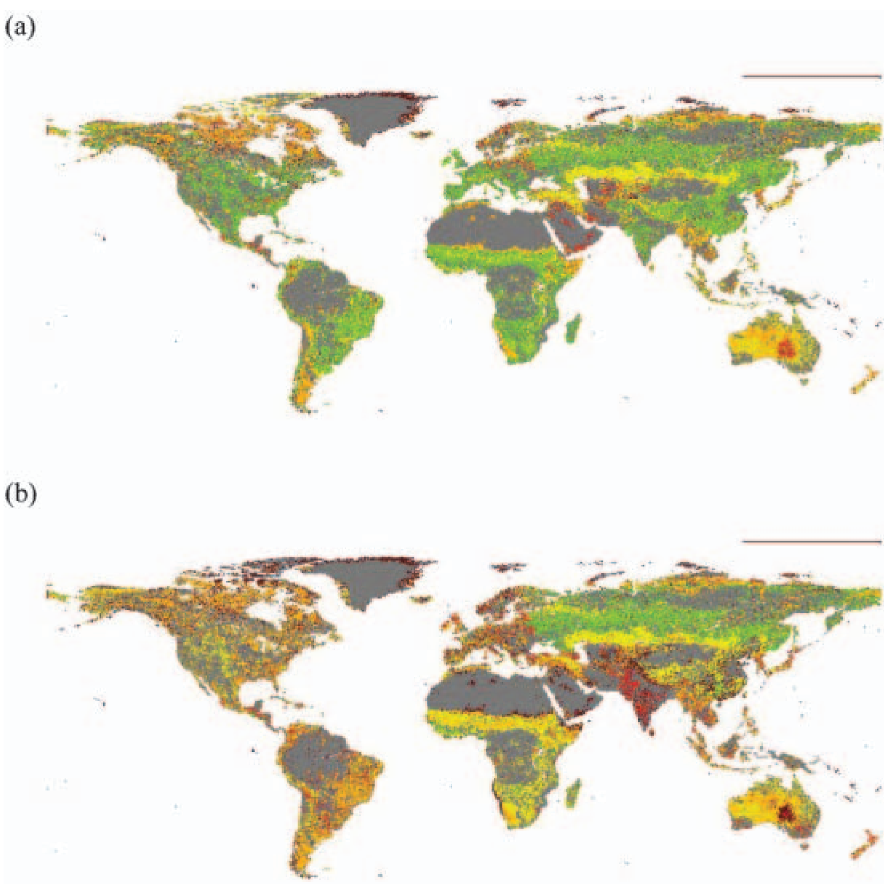


Figure 2. Fuzzy map of spatial disagreement between GLC-2000 and MODIS using (*a*) a maximum and (*b*) a minimum operator. Note the color sceme from dark red (very severe disagreement) to red (severe disagreement) to orange (moderate severity of disageement) to yellow (low severity of disageement) to green (very low severity of disageement).

Table 4. Format of the questionnaire provided to the experts (categories 1–17 correspond to the 17 IGBP land cover categories as they appear in table 3)[a].

MODIS Classes	Evergreen Need-leaved Forest	Deciduous Need-leleaf Forest	Deciduous Broadleaf Forest	Mixed Forests	Closed Shrublands	Open Shrublands	Woody Savannans	Savannans	Grasslands	Permanent Wetlands	Croplands	Urban and Built-Up	Cropland/ Natural Veg Mosiac	Snow and Ice	Barren or Sparsely Vegetated	Unclassified	Water Bodies
Evergreen Needleleaf	—	3	3	2	4	3	2	4	3	2	5	2	2	4	1	1	3
Deciduous Needleleleaf		—	2	2	4	3	2	4	3	2	5	2	2	4	1	1	3
Deciduous Broadleaf			—	3	4	3	2	4	4	2	5	2	2	4	1	1	1
Mixed Forests				—	4	3	2	4	4	2	5	2	2	4	1	1	1
Closed Shrublands					—	3	2	5	4	2	5	2	2	4	1	1	1
Open Shrublands						—	4	5	5	4	5	3	3	5	1	1	1
Woody Savannas							—	3	5	5	2	5	4	5	3	3	1
Savannas								—	5	4	3	4	3	5	1	3	1
Grasslands									—	5	2	4	4	5	2	3	1
Permanent Wetlands										—	2	5	4	5	2	4	1
Croplands											—	3	2	3	2	3	5
Urban and Built-Up												—	5	5	2	4	2
Cropland Veg.													—	3	3	5	3
Snow and Ice														—	2	3	2
Barren/Sparse Veg.															—	4	4
Unclassified																—	2
Water Bodies																	—

[a]Please indicate above the diagonal the degree of difficulty in distinguishing between the two land cover classes using the following numbers:
1=very easy;
2=easy;
3=medium;
4=difficult;
5=very difficult; 0=land cover class not used.

Table 5. Results of applying the minimum and maximum operators to combine the GLC-2000 expert responses (cells contain the minimum and maximum values separated by a comma).

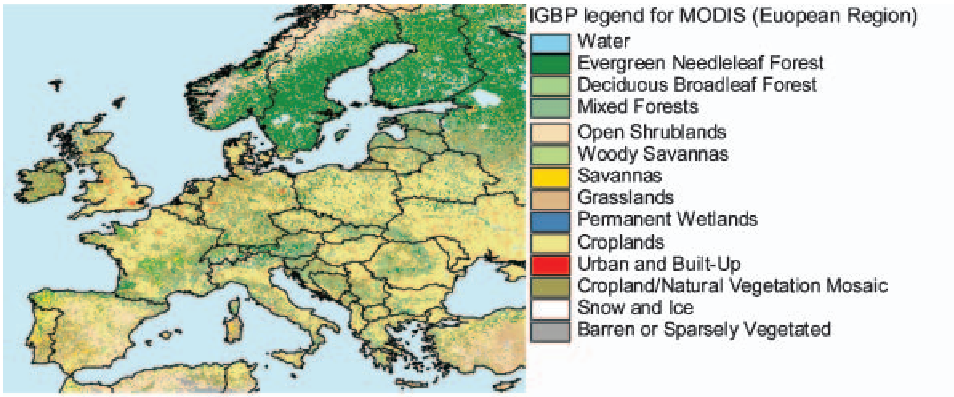
GLC-2000 Classes	Tree Cover, broadleaved, evergreen	Tree Cover, broadleaved, deciduous, closed	Tree Cover, broadleaved, deciduous, open	Tree Cover, needle-leaved, evergreen	Tree Cover, needle-leaved, deciduous	Tree Cover, mixed leaf type	Tree Cover, regularly flooded, fresh water (& brackish)	Tree Cover, regularly flooded, saline water	Mosaic: Tree Cover/Other Natural Vegetation	Tree Cover, burnt	Shrub Cover, closed-open, evergreen
Tree Cover, broadleaved, evergreen	–	0.2, 0.6	0, 0.2	**0, 0.8**	**0, 0.6**	0.6, 0.6	0.6, 0.6	0, 0.6	**0, 0.6**	0.2, 0.2	0, 0.6
Tree Cover, broadleaved, deciduous, closed		–	0.6, 0.6	**0, 0.6**	**0, 0.8**	0.6, 0.6	0.6, 0.6	0, 0.6	**0, 0.6**	0, 0.2	0, 0.6
Tree Cover, broadleaved, deciduous, open			–	0, 0.2	0, 0.6	0.6, 0.6	0.2, 0.6	0, 0.6	**0, 0.6**	0.2, 0.2	**0, 0.6**
Tree Cover, needle-leaved, evergreen				–	0.2, 0.6	0.6, 0.6	0.2, 0.2	0, 0.2	**0, 0.6**	0, 0.6	**0, 0.6**
Tree Cover, needle-leaved, deciduous					–	0.6, 0.6	0.2, 0.2	0, 0.2	**0, 0.6**	0, 0.2	**0, 0.6**
Tree Cover, mixed leaf type						–	0.2, 0.2	0.2, 0.2	0.2, 0.6	0, 0.2	0.2, 0.6
Tree Cover, regularly flooded, fresh water (& brackish)							–	0.6, 0.6	0.2, 0.2	0.2, 0.2	0.2, 0.2
Tree Cover, regularly flooded, saline water								–	0, 0.2	0.2, 0.2	0, 0.2
Mosaic: Tree Cover/Other Natural Vegetation									–	0, 0.2	0.2, 0.6
Tree Cover, burnt										–	0.2, 0.6
Shrub Cover, closed-open, evergreen											–

Table 5. *(continued)*

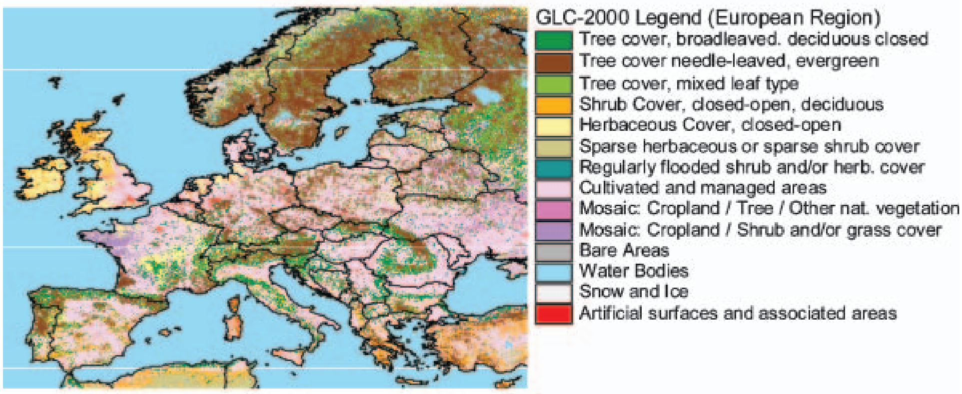
GLC-2000 Classes	Shrub Cover, closed-open, deciduous	Herbaceous Cover, closed-open	Sparse Herbaceous or Sparse Shrub Cover	Regularly Flooded Shrub and/or Herbaceous Cover	Cultivated and Managed Areas	Artificial Surfaces and Associated Areas	Mosaic: Cropland/Tree Cover/Other Natural Veg	Mosaic: Cropland/Shrub or Grass Cover	Bare Areas	Water Bodies (natural & artificial)	Snow and Ice (natural & artificial)
Tree Cover, broadleaved, evergreen	0, 0.2	0, 0.2	0, 0.2	0, 0	0, 0.2	0, 0.6	0, 0.2	0, 0	0, 0	0, 0	0, 0.2
Tree Cover, broadleaved, deciduous, closed	**0, 0.6**	0, 0.6	0, 0.2	0, 0.2	0, 0.6	0, 0.6	0, 0.6	0, 0	0, 0	0, 0	0, 0.2
Tree Cover, broadleaved, deciduous, open	**0, 0.6**	0, 0.2	0, 0.2	0, 0.2	0, 0.6	**0, 0.6**	0, 0.6	0, 0	0, 0	0, 0	0, 0
Tree Cover, needle-leaved, evergreen	0, 0.2	**0, 0.6**	0, 0.2	0, 0.6	0, 0.6	0, 0.2	0, 0.2	0, 0	0, 0.6	0, 0	0, 0.2
Tree Cover, needle-leaved, deciduous	0, 0.6	**0, 0.6**	0, 0.6	0, 0.2	0, 0.6	0, 0.6	0, 0.6	0, 0.2	0, 0	0, 0	0, 0.2
Tree Cover, mixed leaf type	0.2, 0.6	0.2, 0.6	0.2, 0.2	0.2, 0.2	0, 0.6	0.2, 0.6	0, 0.6	0, 0	0, 0.2	0, 0	0, 0.2
Tree Cover, regularly flooded, fresh water (& brackish)	0.2, 0.2	0.2, 0.2	0, 0.2	0.2, 0.6	0, 0.2	0, 0.2	0, 0.2	0, 0	0, 0	0, 0	0, 0
Tree Cover, regularly flooded, saline water	0, 0.2	0, 0.2	0, 0.2	0.2, 0.6	0, 0.2	0, 0.2	0, 0.2	0, 0	0, 0	0, 0	0, 0
Mosaic: Tree Cover/Other Natural Vegetation	0.2, 0.6	0.2, 0.6	**0, 0.8**	**0, 0.8**	**0, 0.6**	**0.2, 0.8**	0.2, 0.6	0, 0.2	0, 0	0, 0	0, 0.2
Tree Cover, burnt	0.2, 0.6	0.2, 0.6	0.2, 0.6	0.2, 0.2	0.2, 0.2	0.2, 0.6	0.2, 0.6	0.2, 0.6	0, 0.2	0, 0	**0, 0.6**
Shrub Cover, closed-open, evergreen	0, 0.6	0.2, 0.6	**0, 0.6**	**0, 0.6**	**0, 0.6**	0.2, 0.6	0.2, 0.6	0, 0.2	0, 0	0, 0	0, 0.6
Shrub Cover, closed-open, deciduous	—	**0.2, 0.8**	0.2, 0.6	0, 0.6	**0, 0.6**	**0.2, 0.8**	0.2, 0.6	0, 0.2	0, 0	0, 0	0, 0
Herbaceous Cover, closed-open		—	0.6, 0.8	**0, 0.6**	0.2, 0.6	**0, 0.8**	0.6, 0.6	**0, 0.6**	0, 0.2	0, 0	0, 0.2
Sparse Herbaceous or Sparse Shrub Cover			—	0.2, 0.6	0.2, 0.6	0.2, 0.6	0.2, 0.6	0.2, 0.6	0, 0.2	0, 0	**0, 0.6**
Regularly Flooded Shrub and/or Herbaceous Cover				—	0, 0.2	0, 0.6	0, 0.6	0, 0.2	0, 0.6	0, 0	0, 0.6

Table 5. *(continued)*

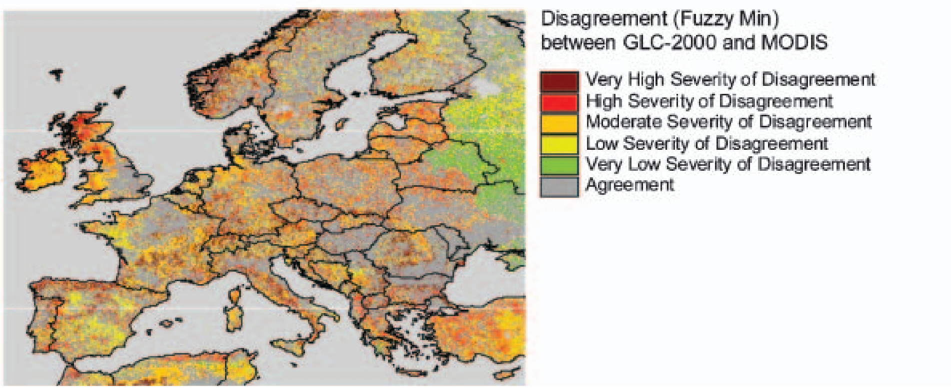
GLC-2000 Classes	Shrub Cover, closed-open, deciduous	Herbaceous Cover, closed-open	Sparse Herbaceous or Sparse Shrub Cover	Regularly Flooded Shrub and/ or Herbaceous Cover	Cultivated and Managed Areas	Artificial Surfaces and Associated Areas	Mosaic: Cropland/Tree Cover/Other Natural Veg	Mosaic: Cropland/Shrub or Grass Cover	Bare Areas	Water Bodies (natural & artificial)	Snow and Ice (natural & artificial)
Cultivated and Managed Areas					—	0.6, 0.6	0.2, 0.6	**0, 0.6**	0, 0.2	0, 0	**0, 0.8**
Artificial Surfaces and Associated Areas						—	0.2, 0.8	0, 0.2	0, 0	0, 0	0, 0
Mosaic: Cropland/Tree Cover/ Other Natural Veg							—	0, 0.2	0, 0	0, 0	**0, 0.6**
Mosaic: Cropland/Shrub or Grass Cover								—	0, 0	0, 0	**0, 0.8**
Bare Areas									—	0, 0.2	**0, 0.8**
Water Bodies (natural & artificial)										—	0, 0.2
Snow and Ice (natural & artificial)											—



(a)



(b)



(c)

Figure 3. GLC-2000, MODIS and fuzzy agreement map for Europe. (*a*) MODIS land cover (v3) for European region. (*b*) GLC 2000 (version 1.1) for European region. (*c*) Disagreement between MODIS and GLC 2000.

agreement). An example would be distinguishing between the GLC-2000 classes 'Tree Cover, broadleaved, evergreen' and 'Bare areas', which all experts found to be very easy, or a value of 1. If the GLC-2000 had 'Tree Cover, broadleaved, evergreen' in a given pixel, and the MODIS map had the equivalent of 'Bare areas', this would result in a severe disagreement at that point. Likewise, if an expert found it very difficult to distinguish between two classes and responded with a value of 5, this would map onto a fuzzy value of 0.8. For example, the experts found it much harder to distinguish between the categories ‘Tree Cover, broadleaved, evergreen’ and ‘Tree Cover, needle-leaved, evergreen’, especially one expert who responded with a value of 5. Thus, if the land cover products did not agree at a given point, but the two categories were difficult to distinguish, then the severity of disagreement is correspondingly much lower and given a value of 0.8. Values in between therefore represent different degrees of severity of disagreement.

The range of results provided by the GLC-2000 experts is given in table 5. Those categories with the largest range in response serve to illustrate the classes that are the most uncertain and indicate that one expert had little difficulty in separating two given classes, whereas the other expert may have found this task much harder. For example, if one compares the category ‘Tree Cover, broadleaved, evergreen’ with ‘Bare areas’, the minimum and maximum values are both 0. This result indicates that these categories are easy to distinguish from one another based on expert knowledge, with no variation in opinion between the different experts. However, if we look at the category ‘Tree Cover, broadleaved, evergreen’ and compare that with ‘Tree Cover, needle-leaved, evergreen’, the minimum and maximum values range from 0.0 to 0.8. The reason for this range of uncertainty is because the experts have used different methods of classification, and some have also used multiple sensors, which have allowed them to distinguish a better separation between the classes. Another reason might be regional differences of vegetation composition and the occurrence of complex landscapes in some regions. Other situations in which these wide ranges of uncertainty exist are highlighted in bold in table 5.

4.4 *Application of a maximum operator to allow different legend categories to be compared*

Once each of the expert responses was mapped onto fuzzy values of agreement, there were k matrices G of dimension n by n, where k is the number of GLC-2000 experts, and n is 22 GLC-2000 classes. Likewise, there was one matrix M of dimension m by m, where m is 17 MODIS classes representing the responses from the one MODIS expert. The next step is to create an n by m matrix for each expert, which contains the fuzzy membership values of agreement between the GLC-2000 and MODIS classes. This matrix expresses the degree of agreement that is assigned spatially if there is disagreement when comparing a particular GLC-2000 and MODIS class. If $A[n,m,k]$ is the desired matrix for each expert k and $L[n,m]$ is the correspondence matrix in table 3, where the Xs are replaced by 1s and all other cells are 0, then the algorithm is as follows:

1. For $k=1$, loop through the entire matrix $A[n,m,k]$ one cell at a time.
2. For each cell in $A[n,m,k]$, determine the corresponding GLC-2000 categories that are equivalent to that MODIS class using the information in $L[n,m]$ and place these in vector $x[n]$. If we take the operation in the first cell, the calculation would be as follows: $x[n]=G[1,n,1]\ L[n,1]^{T}$.
3. Apply a maximum operator to vector x and place the result in $A[n,m,k]$. A maximum operator is used so that the most uncertain situation is captured. For example, we may be comparing a GLC-2000 category against a MODIS class for which there are three corresponding GLC-2000 categories. Suppose

that when you extract vector $x[n]$, you yield three positive values of 0.2, 0.4 and 0.6, indicating a range of fuzzy agreements. Using the maximum operator, a value of 0.6 would be assigned to $A[n,m,k]$.

4. Repeat steps 1–3 for the remaining experts.

This algorithm is repeated for the response from the MODIS expert creating an additional matrix $B[n,m]$. However, in step 2, the reverse would take place, i.e. for each cell, the corresponding MODIS classes that are equivalent to the GLC-2000 class are extracted, and a maximum operator is applied. At this stage, we now have fuzzy matrices $A[n,m,k]$ and $B[n,m]$ for experts $k+1$.

4.5 *Combining experts using a maximum and minimum operator*

The expert responses were combined using the maximum operator to denote the most conservative approach and the minimum operator to indicate the most optimistic expert view. For each region for which an expert response was available, a fuzzy minimum and maximum operator were applied to create a new matrix D as follows:

$$D_{\max}[n, m] = \max(A[n, m, k], B[n, m])$$

$$D_{\min}[n, m] = \min(A[n, m, k], B[n, m])$$

where k is the regional expert. For regions where no expert response was available, a fuzzy operator was used to combine the responses from all the GLC-2000 experts with that of the MODIS expert:

$$D_{\max}[n, m] = \max(A[n, m, 1], A[n, m, 2], \ldots, A[n, m, k], B[n, m])$$

$$D_{\min}[n, m] = \min(A[n, m, 1], A[n, m, 2], \ldots, A[n, m, k], B[n, m])$$

For tropical Asian areas, two GLC-2000 experts were combined with the responses of the MODIS expert.

5. Comparison of the GLC-2000 and MODIS land cover products

The GLC-2000 (version 1.1) and MODIS (v.3) land cover maps were compared on a pixel-by-pixel basis using $D_{\max}$ and $D_{\min}$ to produce maps of spatial agreement. For the Boolean approach, whenever the classes of the two maps are not comparable and hence do not map onto each other as given in the legend category lookup table (table 3), they are in complete disagreement or assigned a value of 0. This is in contrast to the fuzzy approach, which maps degrees of agreement ranging from 0 (complete disagreement) to 1 (complete agreement). For example, if a pixel-by-pixel comparison shows a bare area class on one map and a forest class on the other, the degree of agreement will be low, as these classes are generally easy to distinguish. In contrast, if a needle-leaved evergreen forest class corresponds to a mixed forest class, then there will be a degree of agreement, but it will be higher (or conversely, the amount of disagreement will be less severe).

Table 6 provides global agreement values for the Boolean and fuzzy approaches. In addition, the number of pixels that fall in each category of fuzzy agreement are listed. The percentage of fuzzy agreement was calculated by multiplying the degree of fuzzy agreement by the number of pixels, and then calculating the sum as a percentage of the total pixels. The overall agreement for the Boolean map is 52.7%,

Table 6. Number of pixels showing full, partial and no agreement[a].

Operator	Fuzzy agreement						Fuzzy percentage agreement
	1.0	0.8	0.6	0.4	0.2	0.0	
Boolean	101 507 493	NA	NA	NA	NA	89 841 995	52.7
Fuzzy Max	101 507 493	35 530 703	21 146 087	23 054 104	4 834 583	5 276 518	79.9
Fuzzy Min	101 507 493	11 747 237	26 084 496	31 785 836	7 158 105	1 306 6321	73.5

[a]1 = full agreement; 0 = no agreement; values in between = partial degrees of agreement. NA = not applicable.

but this value increases significantly when taking a fuzzy agreement approach, i.e. 79.9% for the map generated using a maximum operator and 73.5% for a minimum operator.

The fuzzy agreement between the GLC-2000 (version 1.1) and MODIS (v.3) land cover maps is shown in figure 2. Figure 2(*a*) results from expert knowledge combined through a maximum operator or a conservative approach, while figure 2(*b*) shows the same map created using a minimum operator. The colour scheme used in figure 2 ranges from little agreement or severe disagreement (dark red, red) towards medium disagreement (orange) to higher degrees of agreement (yellow, green). Severe disagreement indicates that in those areas, either one (when using the minimum operator) or both the GLC-2000 and MODIS experts (when using the maximum operator) agree that the two classes at a particular point on the GLC-2000 and MODIS map were very easy to classify and would therefore be penalized more heavily. In contrast, a very low severity of disagreement indicates that the two classes of GLC-2000 and MODIS are considered to be very difficult to classify.

Since the use of the maximum operator represents the most conservative approach, it therefore considers the largest uncertainty in classification difficulty. In most cases, the MODIS expert, who took a global classification approach, found differentiation between classes generally more difficult than local experts, so the MODIS expert's opinion weighed more heavily in the comparison of the land cover maps. Looking at figure 2(*a*), it is apparent that large areas of disagreement (dark red, red and orange) appear in Australia, Middle East, Australia, northern Canada, Scandinavia and southern South America. Much of the global pattern of disagreement occurs where the GLC-2000 land cover type is 'Bare Areas' or 'Sparse Herbaceous or Sparse Shrub Cover', while the MODIS map classifies most of these same areas as 'Open Shrubland'. This discrepancy in the two maps is clearly an area for further investigation.

The minimum operator (figure 2(*b*)), in contrast, generally favoured the local expert, so the errors are more severe but still considerably less than indicated using a Boolean approach. When looking at figure 2(*b*), additional patterns of severe to medium disagreement appear, including areas in western India, Europe (Scotland, Italy and northern Spain in particular) and eastern Brazil. Figure 3c shows a more detailed view of Europe. If we look at areas of very severe/severe disagreement in Scotland, there are very high percentages of Cropland/Natural Vegetation in MODIS (see figure 3(*a*)), but this appears as Shrubland in GLC-2000 (see figure 3(*b*)). Moreover, Permanent Wetlands appear on the MODIS map which are classified as Shrubland on the GLC-2000. Many areas of a very high or high severity of disagreement in Italy and Northern Spain (see figure 3(*c*)) are related to the fact that areas which appear as forest on the GLC-2000 (see figure 3(*b*)) are classified as Cropland, Cropland/Natural Vegetation or Grassland in the MODIS product (see figure 3(*a*)). There are many other continental, regional and small-scale patterns of disagreement revealed by comparing the two maps in this way. However, it is beyond the scope of this paper to describe all these patterns here. The identification of areas of disagreement is the subject of ongoing work and will feed into the accuracy assessment and validation process of the GLC-2000.

6. Discussion

The methodology in this paper has raised an important issue regarding the need to specifically consider the differences in map legend definitions when comparing products that have been created using different sources and approaches. It is extremely important to determine how the legend types from the two different products map onto one another. The approach in this paper has been to create a lookup table between legend categories. Generally, methods that compare maps try to force one legend completely into another as a one-to-one mapping, which can have catastrophic effects. This approach has mistakenly been used by Giri *et al.* (2005), in their recent comparison of the GLC-2000 and MODIS land cover data sets. Even though the definitions of the GLC-2000 and MODIS are fundamentally different, the GLC-2000 legend was mapped with a one-to-one relationship into the MODIS legend. For example, the GLC-2000 Shrub cover class, which has an LCCS definition of percentage cover between 15 and 100%, was set equal to the MODIS Open Shrubland class with a definition of 10–60% cover. The problem with this approach is that it is now no longer possible to understand if a disagreement is real or if it results from the incompatibility of the legend definitions.

In contrast, by using the approach described in this paper, we can capture the real disagreement that results, which is independent of the legend definitions. In the situation where legends overlap partly, we simply record 100% agreement, which is a very conservative approach. Some disagreement may therefore be lost, especially in situations where there is a great deal of overlap between legends. It is not possible to use a fuzzy approach here because we do not know the degree to which these legend classes overlap, which is a basic requirement for use of fuzzy memberships. For example, the GLC-2000 class Tree cover broadleaved closed (defined as 40%—100% forest cover) overlaps with the woody Savannah class (60%—30% forest cover). In the situation where a large area with a broadleaved forest cover of 50% exists, this area is mapped correctly in the GLC-2000 as Tree cover broadleaved closed and on MODIS as woody Savannah, and hence a 100% agreement exists.

Another important issue that has emerged from this research is the need for developing these types of methodologies: (1) for identifying incompatibility problems of current definitions of global land cover legends and (2) as a precursor to the creation of a global hybrid map. The methodology presented here can be used to highlight the most severe disagreements between two different land cover products in order to direct a ground truthing exercise on these highlighted areas. This will also avoid the need to investigate or re-map those areas where there is already a high level of confidence in the classification. The second process, i.e. the creation of a hybrid GLC-2000/MODIS land cover map, would result in a higher-quality combined global land cover product. Such an approach does not necessitate the remapping of large areas and therefore requires very few additional resources. In creating such a map, the areas of severest disagreement could be validated with freely available TM scenes and Quicklooks, regional and national maps, and expert knowledge. It is then possible to determine which map is more correct in a given area or whether the disagreement simply results from differences in land cover definition. Moreover, if one class from one map is correct at the specific point of validation, the probability is quite high that it is mapped correctly in the surrounding neighbourhood.

Finally, this type of methodology lends itself well to highlighting disagreement for specific users of the two land cover products, an issue flagged by both Foody (2002) and Hagen (2003). For example, the World Conservation Monitoring Centre (WCMC) is using both land cover products in assessing the spatial distribution of global tree cover. The WCMC places emphasis on the presence of tree cover, rather than on the type of trees present such as needleleaved or broadleaved. It would be possible to provide the WCMC with the same type of questionnaire that was given to the GLC-2000 and MODIS experts. The results from this questionnaire would then be used to generate a map highlighting areas of disagreement specific to their needs. In contrast, if the land cover products were used in the field of ecology, different forest types have a strong impact, so the resulting map of spatial disagreement would look very different to that produced for the WCMC. What is a high level of disagreement for one application may be irrelevant for another, and this can be displayed spatially. Other applications that would benefit from this type of approach would be global biomass assessment and global climate modelling.

7. Conclusions

Assessing the accuracy of land cover maps is an important but difficult task. The confusion matrix and other global accuracy measures are commonly used in accuracy assessment, but they have fundamental flaws (Foody 2002) and provide no information on spatial errors. The methodology outlined in this paper allows different land cover maps to be compared in order to highlight areas of spatial disagreement. Although this does not indicate which map is correct, it can provide input to further validation and accuracy assessment exercises. The method specifically took into account the differences in legend categories and used expert knowledge about classification uncertainty to compare the two most recent global land cover maps for the year 2000. Identifying areas of disagreement can also be the first step in producing a hybrid map that combines the most accurate features of the two individual maps. Approaches of this type lead to better use of existing global land cover products without the need for a large additional investment. The next step in creating a hybrid map would be to validate those areas of highest disagreement followed by combining those parts of the two validated maps that are closest in agreement to the validation site. This is currently the subject of ongoing work.

Acknowledgements

The authors would like to thank the British Council/Ministero dell'Instruzione, dell'Università e della Ricerca in Italy for the funding provided through the British Council–MIUR/CRUI Agreement 2002–2003. The authors would also like to thank Alan Belward and Etienne Bartholomé of the Joint Research Centre for coordinating the GLC2000 project and all the experts who filled out the questionnaire.

References

ANONYMOUS, 2004, The ESA GOFC GOLD Land Cover Project Office newsletter land cover and change, November 2004. Available online at: http://www.gofc-gold.unijena.de/documents/newsletter/GOFC-newsletter_no4.pdf (accessed 12 December 2004).

BARTALEV, S.A., BELWARD, A.S., ERCHOV, D.V. and ISAEV, A.S., 2003, A new SPOT4-VEGETATION derived land cover map of Northern Eurasia. *International Journal of Remote Sensing*, **24**, pp. 1977–1982.

BARTHOLOMÉ, E. and BELWARD, A.S., 2005, GLC2000: a new approach to global land cover mapping from Earth Observation data. *International Journal of Remote Sensing*, **26**, pp. 1959–1977.

COMBER, A., FISHER, P. and WADSWORTH, R., 2004, Integrating land cover data with different ontologies: identifying change from inconsistency. *International Journal of Geographic Information Science*, **18**, pp. 691–708.

DEFENSE MAPPING AGENCY, 1992, Development of the Digital Chart of the World (Washington DC: US Government Printing Office).

DI GREGORIO, A. and JANSEN, L., 2000, *Land Cover Classification System: Classification Concepts and User Manual* (Rome: Food and Agriculture Organization of the United Nations).

EVA, H.D., BELWARD, A.S., DE MIRANDA, E.E., DI BELLA, C.M., GOND, V., HUBER, O., JONES, S., SGRENZAROLI, M. and FRITZ, S., 2004, A land cover map of South America. *Global Change Biology*, **10**, pp. 731–744.

FOODY, G.M., 2002, Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, **80**, pp. 185–201.

FRIEDL, M.A., MCIVER, D.K., HODGES, J.C.F., ZHANG, X.Y., MUCHONEY, D., STRAHLER, A.H., WOODCOCK, C.E., GOPAL, S., SCHNEIDER, A., COOPER, A., BACCINI, A., GAO, F. and SCHAAF, C., 2002, Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, **83**, pp. 287–302.

FRITZ, S., BARTHOLOMÉ, E., BELWARD, A., HARTLEY, A., STIBIG, H.J., EVA, H., MAYAUX, P., BARTALEV, S., LATIFOVIC, R., KOLMERT, S., ROY, P., AGRAWAL, S., BINGFANG, W., WENTING, X., LEDWITH, M., PEKEL, F.J., GIRI, C., MÜCHER, S., DE BADTS, E., TATEISHI, R., CHAMPEAUX, J-L. and DEFOURNY, P., 2003, *Harmonisation, Mosaicing and Production of the Global Land Cover 2000 Database (Beta Version)* (Luxembourg: Office for Official Publications of the European Communities).

GIRI, C., ZHU, Z. and REED, B., 2005, A comparative analysis of the global land cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, **94**, pp. 123–132.

GOPAL, S. and WOODCOCK, C., 1994, Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*, **60**, pp. 181–188.

HAGEN, A., 2003, Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science*, **17**, pp. 235–249.

HODGES, J., 2002, Validation of the consistent-year V003 MODIS land cover Product. Available online at: http://geography.bu.edu/landcover/userguidelc/consistent.htm (accessed 13 April 2004).

LIANG, S., FANG, H., CHEN, M., SHUEY, C.J., WALTHALL, C., DAUGHTRY, C., MORISETTE, J., SCHAAF, C. and STRAHLER, A., 2002, Validating MODIS land surface reflectance and albedo products: methods and preliminary results. *Remote Sensing of Environment*, **83**, pp. 149–162.

LOVELAND, T.R., REED, B.C., BROWN, J.F., OHLEN, D.O., ZHU, Z., YANG, L. and MERCHANT, J., 2000, Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. *International Journal of Remote Sensing*, **21**, pp. 1303–1330.

LUND, H.G., 1993, Politically correct global mapping and monitoring. In *Mapping Tomorrow's Resources, Natural Resources and Environmental Issues Volume II,* 23–24 April 1992, Logan, UT, by A. Falkoner (Eds), pp. 47–54 (Logan, UT: Utah State University, College of Natural Resources).

MAYAUX, P.M., BARTHOLOMÉ, E., FRITZ, S. and BELWARD, E., 2004, A new land-cover map of Africa for the year 2000. *Journal of Biogeography*, **31**, pp. 861–877.

MORISETTE, J.T., PRIVETTE, J.L., CHRISTOPHER, O. and JUSTICE, C.O., 2002, A framework for the validation of MODIS land cover products. *Remote Sensing of Environment*, **83**, pp. 77–96.

STIBIG, H-J., ACHARD, F. and FRITZ, S., 2004, A new forest cover map of continental Southeast Asia derived from satellite imagery of coarse spatial resolution, *Applied Vegetation Science*, **7**, pp. 153–162.

STRAHLER, A., 2002, Global Land Cover: Approaches to Validation. GLC2000 Meeting, JRC, Ispra, Mar 2002. Available online at: http://www.gvm.jrc.it/glc2000/Workshops/March2002/Presentations/strahler_Validation.pps (accessed 13 April 2004).

UNITED NATIONS, 2002, *Report of the World Summit on Sustainable Development,* Johannesburg, South Africa, 26 August to 4 September 2002.

WOODCOCK, C.E. and GOPAL, S., 2000, Fuzzy set theory and thematic maps: accuracy assessment and area estimation. *International Journal of Geographical Information Systems*, **14**, pp. 153–172.