GLOBCORINE

Validation Report

Milestone	Final meeting 1 – KO+9
Authors	Pierre Defourny, Sophie Bontemps, Eric van Bogaert, Jean-Louis Weber, Tomas Soukup
Distribution	ESA : Olivier Arino, Vasileios Kalogirou

Table of contents

Acronyms

CLC	: Corine Land Cover
EEA	: European Environmental Agency
ENVISAT	: European Space Agency Environmental Satellite
ESA	: European Space Agency
FRS	: Full Resolution Full Swath
FSG	: Full Swath Geo-corrected
GLOBCOVER	: ESA DUE (http://dup.esrin.esa.it/invitations.asp)
LCCS	: Land Cover Classification System
MERIS	: Medium Resolution Imaging Spectrometer Instrument (http://envisat.esa.int)
NDVI	: Normalized Difference Vegetation Index
SDR	: Surface Directional Reflectance
SRTM	: Shuttle Radar Topography Mission
SWBD	: SRTM Water Body Dataset
SWIR	: Short-Wave InfraRed
VGT	: SPOT-VEGETATION

Applicable documents and references

Applicable documents

[AP-1] GLOBCORINE, Technical Specification, 2.2.

[AP-2] SRTM Water Body Data Product Specific Guidance, v 2.0, March 12 2003 (http://edc.usgs.gov/products/elevation/swbd.html)

[AP-3] GLOBCOVER, GlobCover Validation Report I2.1, December 2008

Reference

[RD-1] European Environment Agency. 2006. Land accounts for Europe 1990-2000, EEA Report No 11/2006 prepared by Haines-Young, R. and Weber, J.-L. (http://reports.eea.europa.eu/eea_report_2006_11/en)

[RD-2] Vancutsem, C., Bicheron, P., Cayrol, P. and Defourny, P. 2007. Performance assessment of three compositing strategies to process global ENVISAT MERIS time series, *Canadian Journal of Remote Sensing*, 33, 492-502.

[RD-3] Strahler, A.H., Boschetti, L., Foody, G.M., Friedl, M.A., Hansen, M.A., Mayaux, P., Morisette, J.T., Stehman, S.V. and Woodcock, C.E. 2006. Global Land Cover Validation: recommendations for evaluation and accuracy assessment of global land cover maps, Office for Official Publications of the European Communities, Luxembourg.

1. General introduction

The GlobCorine project, which was initiated by the European Space Agency (ESA), focused on the production of a land cover map dedicated to the pan-European continent and driven by the European Environmental Agency (EEA) recommendations and needs. The GlobCorine project aims to address this issue by making the full use of the potential of the ENVISAT's Medium Resolution Imaging Spectrometer Instrument (MERIS) Full Resolution Full Swath (FRS) time series and by further developing the GlobCover classification approach.

The GlobCover classification module has to be adjusted to produce from the 300-m MERIS dataset a land cover product as compatible as possible with the Corine Land Cover (CLC) aggregated typology which is more land use oriented than the GlobCover legend.

This document describes the GlobCorine product, as it is released to the users. It is structured as follows:

- A presentation of the GlobCorine land cover product is provided in section 2.
- A quantitative validation of the GlobCorine land cover product is proposed in section 3.
- The GlobCorine land cover product is discussed and general recommendations are reported in section 4.

2. Description of the land cover product

2.1. Data coverage

The GlobCorine land cover map is based on the MERIS data acquired and pre-processed in the framework of the GlobCover project.

The GlobCover project was based on global MERIS FRS time series available from December 2004 until June 2006 processed at level 1B, i.e. calibrated top of atmosphere gridded radiances. Each MERIS FRS mosaic is organized on a 5° by 5° tiling without any overlap (Figure 1). The pan-European continent was therefore covered by 204 tiles (17 horizontal tiles x 12 vertical tiles from 15°N to 75°N and from 20°W to 65°E).

Figure 1. Scheme of the Globcover tiles location

The very first challenge of Globcover was the global acquisition of a MERIS time series while the instrument was not initially designed to do so. As shown in Figure 2, the data coverage was uneven due to programmatic constraints. As expected, the number of valid observations after the pre-processing steps, in particular that of cloud screening, was also rather variable (Figure 3).

Nevertheless, the dataset available over most of Europe is significantly better than in many places of the world and the GlobCorine project should make use of the spatially consistent time series.

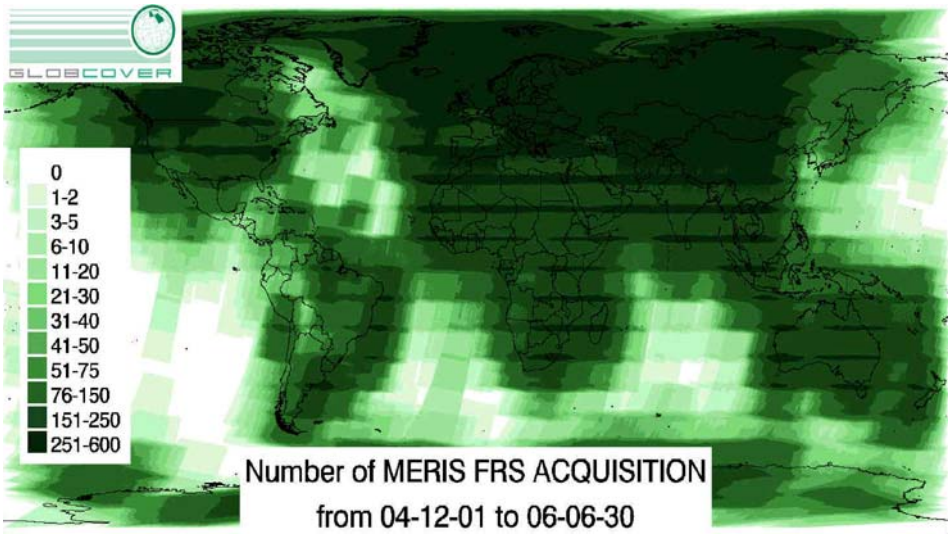

Figure 2. Density of MERIS FRS acquisitions from 1st December 2004 to 30th June 2006.

.

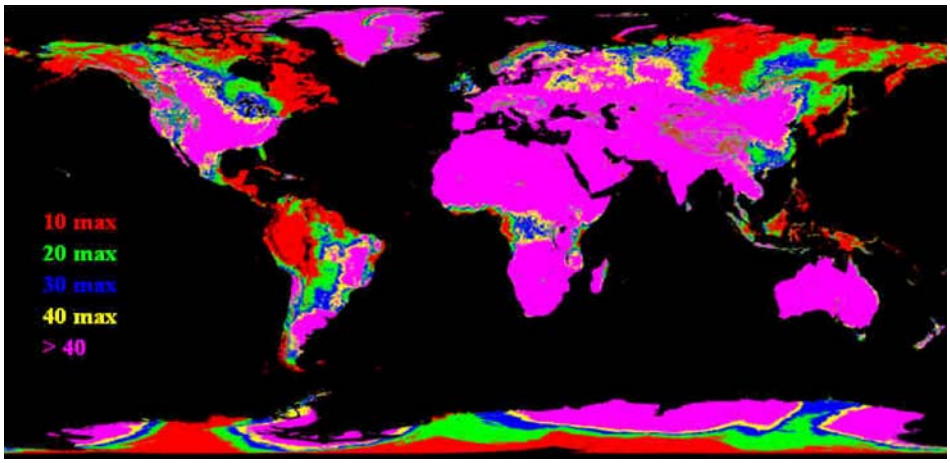

Figure 3. Number of valid observations obtained after 19 months of MERIS FRS acquisitions. Magenta areas are defined as well covered (>40 observations)

2.2. Description of the algorithm

The GlobCorine classification chain consists in transforming the MERIS multispectral mosaics produced by the GlobCover pre-processing modules into a meaningful pan-European land cover map. As explicitly requested by ESA and EEA, the challenge is twofold: (i) producing, in an automatic, repeatable and global way, a pan-European land cover map at 300-m resolution and (ii) defining and documenting a legend as compatible as possible with the CLC aggregated typology and the UN Land Cover Classification System (LCCS). The classification module has been designed by UCL-Geomatics to combine both the spectral and temporal range of the MERIS FR time series and to be globally consistent while regionally-tuned.

Table 1 presents a summary of the methodology used to produce the GlobCorine land cover map. The classification algorithm is exhaustively described in the GlobCorine TS document [AP-1].

Processing step	Algorithmic description
Pre-processing level Step 1: seasonal syntheses production	The daily images, as pre-processed in the framework of the GlobCover project ([AP-4]), are composited into MERIS FRS seasonal syntheses, through a Mean Compositing (MC) algorithm ([RD-2]). The status of each pixel is validated and daily surface reflectance values are averaged for a specified period, ranging from two months to one year. The compositing period is determined according to the seasonality and the availability of daily images. This step results in MERIS FRS seasonal syntheses, organised by 5° x 5° tiles.
Pre-processing level Step 2: pan-European mosaics production	MERIS FRS seasonal syntheses are merged to produce MERIS FRS seasonal mosaics. The GlobCorine area is covered by 176 tiles (16 horizontal tiles x 11 vertical tiles)
Classification level Step 0: stratification	The pan-European continent is stratified in 5 equal-reasoning areas from an ecological and a remote sensing point of view, with a twofold objective: (1) reducing the land surface reflectance variability in the dataset in order to improve the classification efficiency and (2) allowing a regional tuning of the classification parameters to take into account the regional characteristics (vegetation seasonality, cloud coverage, etc). The stratification mainly relies on natural discontinuities (oceans, seas, mountains area…) and on sharp interfaces clearly depicted from a remote sensing point of view (e.g. forest-grassland interfaces). Each delineated equal-reasoning area is then classified independently.
Classification level Step 1a: spectral supervised classification	The supervised classification aims to identify land cover classes that are not well represented in the strata (urban areas and wetlands). Region of Interest (ROI) are defined for each land cover class. First, an unsupervised spectral classification is performed in each ROI in order to delineate

	several clusters representative of the corresponding land cover classes. Second, in each MASK area including the previous ROI, a supervised spectral classification is achieved to identify the land cover classes.
Classification level Step 1b: spectral unsupervised classification	An unsupervised spectral classification is performed on the pixels that were not classified by the step 1a, with the aim of creating clusters of spectrally similar pixels.
Classification level Step 1c: merge of the spectral classifications	The step 1c merges the two spectral classifications obtained from the steps 1a and 1b.
Classification level Step 2: first reference-based labeling	The step 2 transforms the spectral classes identified in the step 1 into land cover classes. The labeling procedure is automated. The GlobCorine land cover labels are decided according to the correspondence between the spectral classes obtained from the step 1 and the classes described with the reference land cover map. Several decision rules have been defined to derive unique label for each spectral class. A code that expresses the ambiguity of the labeling is associated with each class. The reference land cover map consists of a pan-European reference land cover database compiled from CLC maps (from 2000 and 2006), Globcover 2005 map and other local reference land cover maps selected as the most accurate map available for a given region, with the highest spatial resolution and with a GlobCorine-compatible legend.
Classification level Step 3a: temporal characterization	The step 3 focuses on pixels classified as mosaic classes after the step 2. The step 3a applies an unsupervised classification on MERIS 10-day NDVI profiles to create clusters of pixels similar in the temporal space.
Classification level Step 3b: temporal classification	The mosaic classes obtained in step 2 are then sub-divided using the step 3a temporal information. That disaggregation aims to create new smaller and more homogeneous spectro-temporal classes.

Classification level Step 4: second reference-based labeling	A second labeling procedure is performed in order to transform the temporal classes into land cover classes. The step 4 is based on the same procedure than step 2 but on different labeling rules.
Classification level Step 5: merge of classifications	The step 5 merges the land cover classes derived from the spectral classification (step 1 and 2) and those derived from the spectro-temporal classification (step 3 and 4) in order to produce the GlobCorine land cover map.
Post-classification edition Step 6a: gap filling	In step 6a, gaps in the GlobCorine land cover map are filled with the reference land cover.
Post-classification edition Step 6b: water bodies	To deal with imprecision in the land/water mask used in the pre-processing ([AP-3]), the step 6b makes uses of an external dataset in order to improve the "water bodies" delineation in the GlobCover classification. This dataset is the Shuttle Radar Topography Mission (SRTM) Water Body Dataset (SWBD) ([AP-2]).

Table 1. Summary of the GlobCorine classification methodology

2.3. Description of the legend

It was one of the project's requirements that the GlobCorine legend had to be as compatible as possible with the CLC aggregated typology ([RD-1]) and with the LCCS system.

The GlobCorine land cover product has been designed to be a consistent continental land cover map. Therefore, the legend is determined by the level of information that is available and that makes sense at the scale of the pan-European continent. The GlobCorine legend counts 14 classes and meets this requirement (Table 2).

NB_LAB	LAB	Color
10	Urban and associated areas	
20	Rainfed cropland	
30	Irrigated cropland	
40	Forest	
50	Heathland and sclerophyllous vegetation	
60	Grassland	
70	Sparsely vegetated area	
80	Vegetated low-lying areas on regularly flooded soil	

90	Bare areas	
100	Complex cropland	
110	Mosaic cropland / natural vegetation	
120	Mosaic of natural (herbaceous, shrub, tree) vegetation	
200	Water bodies	
210	Permanent snow and ice	

Table 2. 14 classes of the GlobCorine legend

The LCCS definition of each GlobCorine class and their correspondence with the CLC aggregated typology are provided in Appendices A and B.

A 15th class (coded as "230") has been added to the final legend to account for no data pixel-values or for areas that are not covered by the project.

2.4. Basic review

The GlobCorine land cover product is the second 300-m land cover map – after the GlobCover land cover map – produced for the pan-European continent for the period December 2004 – June 2006. The map projection is a Plate-Carrée (WGS84 geoid).

The GlobCorine classification counts 14 land cover classes that are well documented and comparable all over the pan-European continent. Figure 4 presents the pan-European GlobCorine land cover map.

Figure 5 presents the area covered by the 14 classes, computed after a projection of the land cover map into a Lambert equal area projection and expressed in percentage. Some classes are little represented as expected: class 10 "Urban and associated areas", class 30 "Irrigated cropland", class 50 "Heathland and sclerophyllous vegetation", class 60 "Grassland", class 80 "Vegetated low-lying areas on regularly flooded soil" and class 100 "Complex cropland".

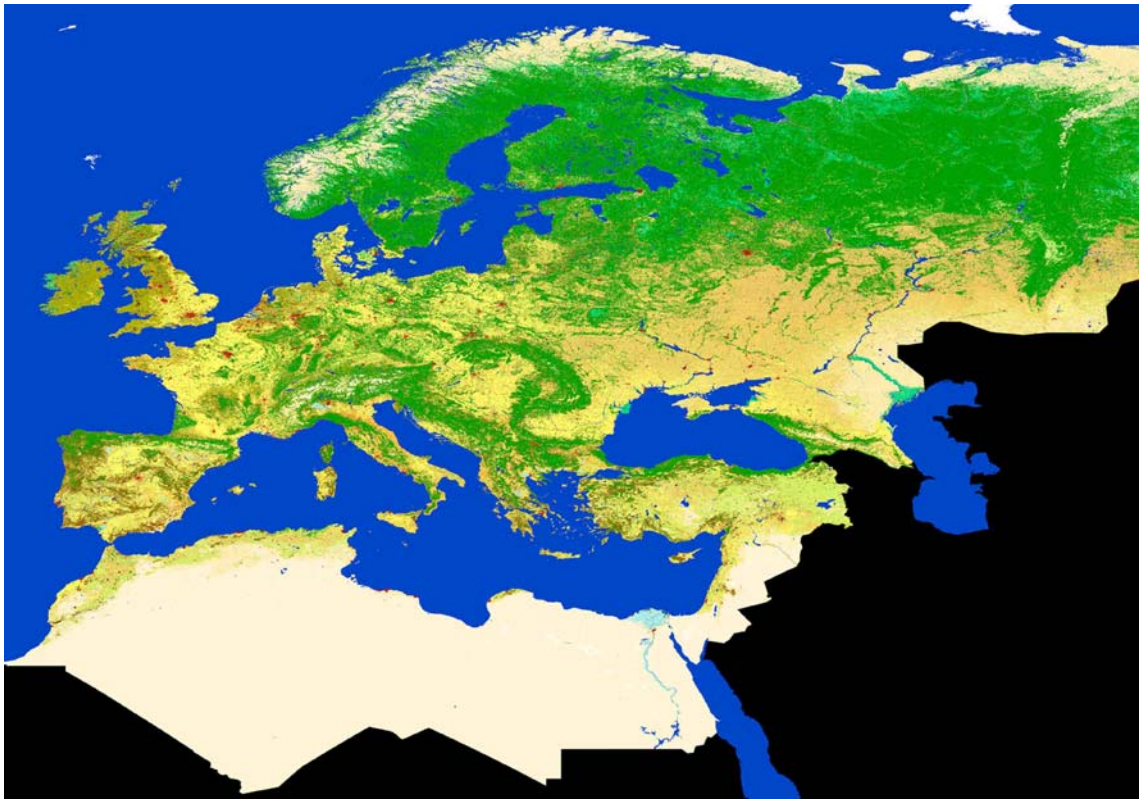

Figure 4. The GlobCorine product as the second 300-m pan-European land cover map for the period December 2004 - June 2006

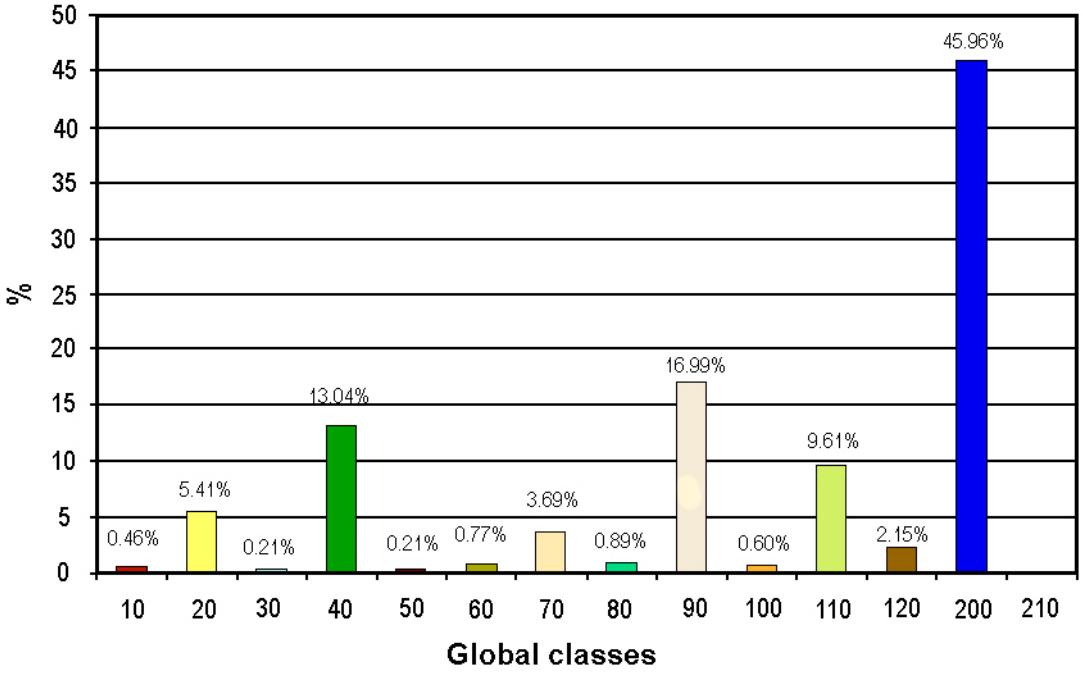

Figure 5. Area (%) covered by the 14 classes

The use of medium resolution brings a considerable improvement in comparison with other global land cover products at lower spatial resolution. Figure 6 provides a comparison

between GLC2000 (1km spatial resolution) and GlobCorine (300-m spatial resolution) in Europe. This improvement was already noticed in the GlobCover product ([AP-3]).

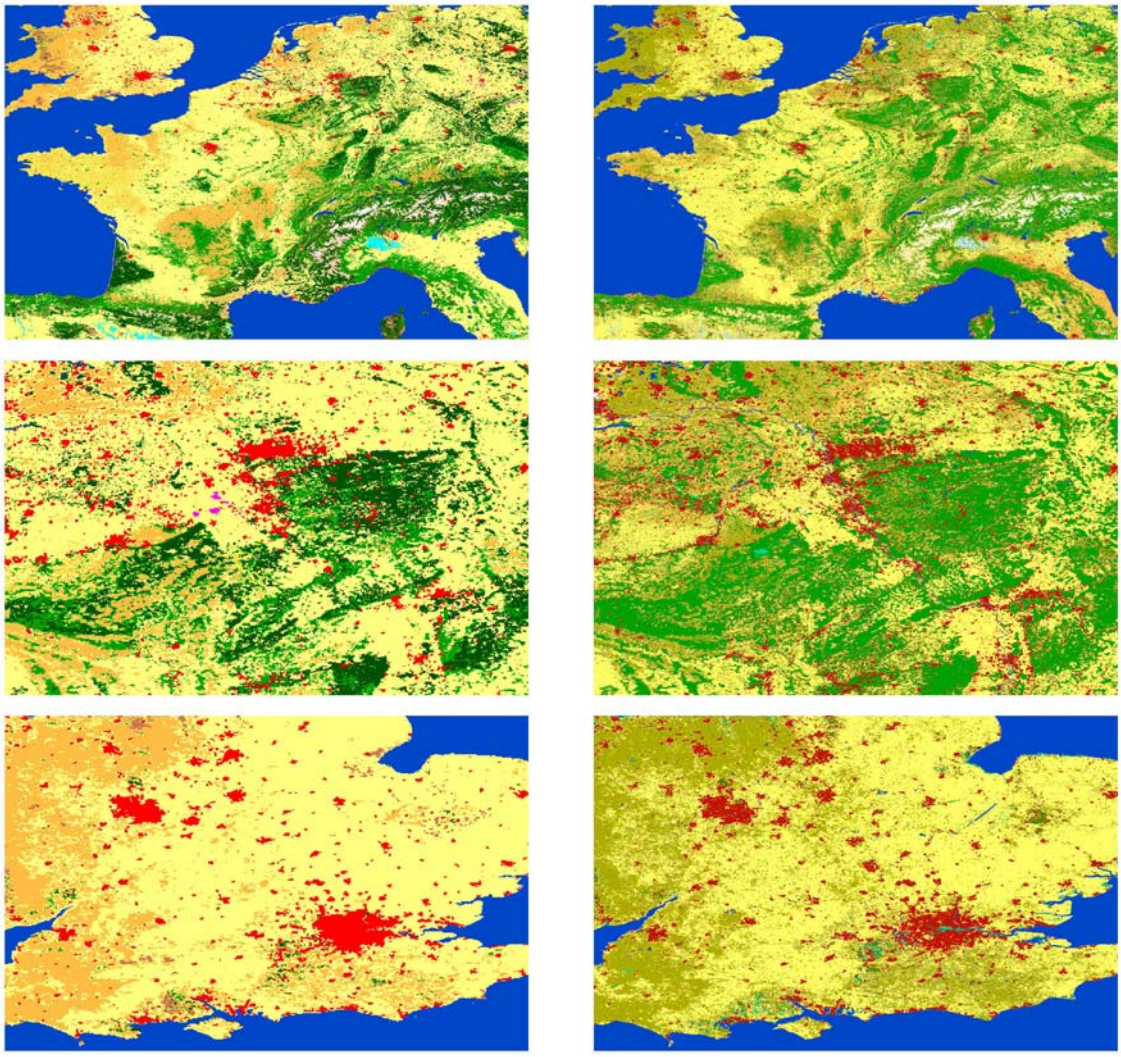

Figure 6. Improvement of the spatial detail due to the 300-m spatial resolution, when compared with the 1-km spatial resolution GLC2000 map. Illustration in Europe (top), with specific examples in Germany/Belgium (centre) and in England (bottom)

The quality of the GlobCorine product is highly dependent on the reference land cover database used for the labeling process and on the number of valid observations available as input. When the reference dataset is of higher spatial resolution with a high thematic detail, the GlobCorine product also shows a high accuracy. Figure 7 shows a comparison between areas in Europe, where the reference dataset is the CLC map, and areas in North Africa, where the reference dataset is either GlobCover or Africover.

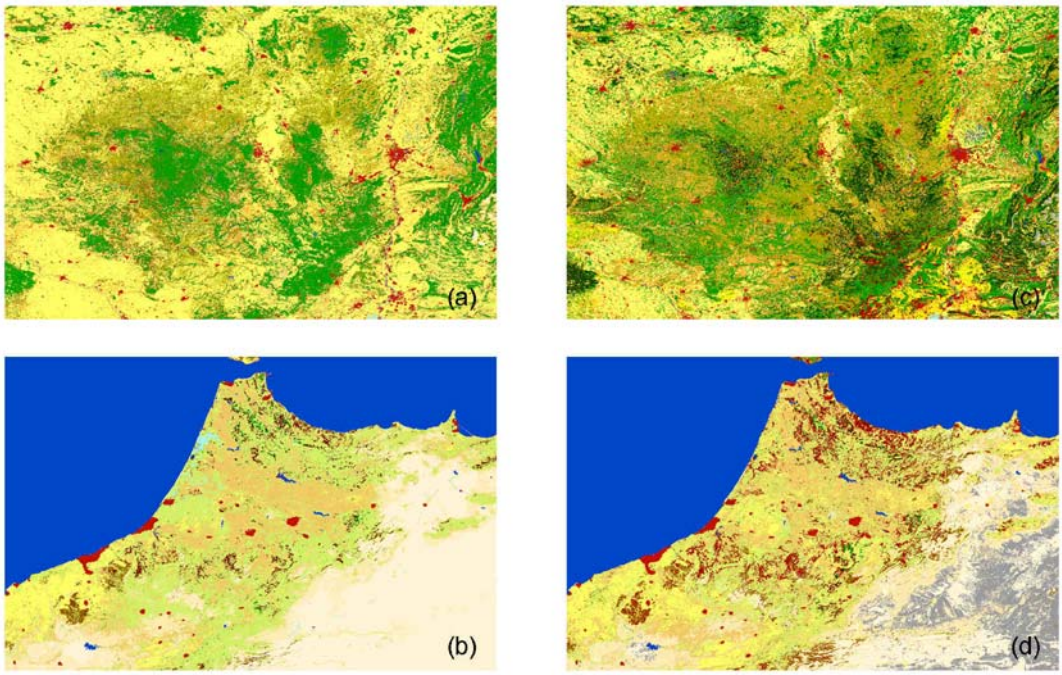

Figure 7. GlobCorine product over France (a) and Morocco (b), compared with the reference. The European reference is based on CLC2000 and 2006 (c) while the African reference uses the GlobCover map (d). Spatial patterns and details of the reference are also present in GlobCorine, which illustrates the dependency of the GlobCorine quality on the reference dataset.

The adjustment of the GlobCover classification algorithm also brings a considerable improvement in comparison with the GlobCover land cover product. First, the algorithm was running on a smaller – and thus more homogeneous – area. Second, the classification parameters were tuned to improve discrimination of the European spatial patterns (typical land cover and/or landscapes, etc.). The three most significant improvements concern the urban areas, the mosaic classes and the sparsely vegetated areas.

Finally, the high spatial consistency of the GlobCorine product has to be pointed out. Areas not covered by CLC2000 (i.e. by the European reference dataset) are coherently classified, as illustrated in Figure 8.

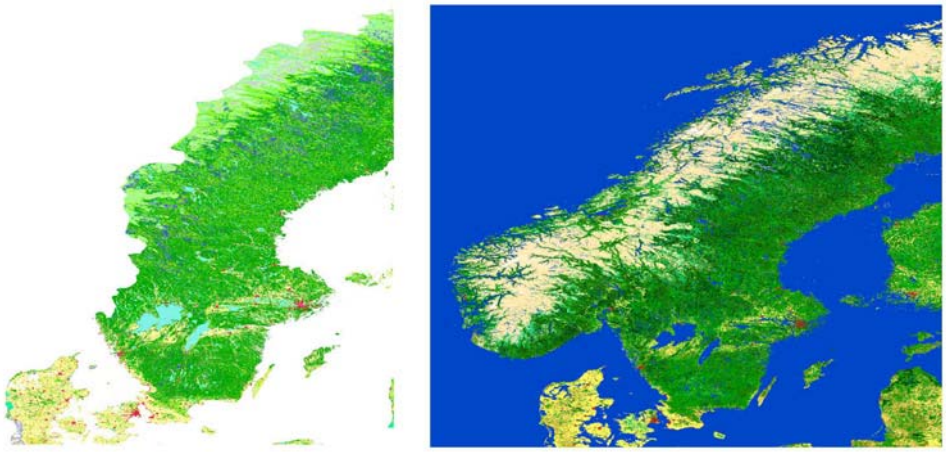

Figure 8. The classification of Norway (bottom), which is not covered by the reference dataset (top), is spatially consistent with surrounding areas

3. Land cover validation

3.1. Land cover validation based on the GlobCover validation dataset

The quantitative validation of the GlobCorine land cover product aims to assess the accuracy of the 14 classes of the land cover map from an independent reference dataset. The validation results allow a potential user determining the map's "fitness for use" for his or her application. The validation process is designed to be scientifically sound, internationally acceptable and feasible from a cost and a time point of view. This is based on the document of the CEOS Land Product Validation subgroup: "*Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment and of Global Land Cover Maps*" ([RD-3]).

The validation process of the GlobCorine land cover map makes uses of the validation dataset that has been constituted in the GlobCover project ([AP-3]).

First, the validation samples related to the GlobCorine area (Europe, Northern Africa and European Russia) were extracted. Second, the set of classifiers and attributes that was collected in order to characterize the land cover of a particular site were transformed to the legend of the GlobCorine product to make them comparable. Finally, the GlobCorine product's accuracy was assessed.

3.1.1. GlobCorine validation dataset

Over the 4258 points contained in the GlobCover validation dataset, 788 points overlaid the GlobCorine area. The distribution of these 788 points is shown in Figure 9. The validation dataset was further reduced to 430 points by removing all the points where the experts were not explicitly certain that the information they provided was correct.

Figure 9. Distribution of the 788 points in the GlobCorine validation dataset.

3.1.2. GlobCover validation dataset interpretation

For observational units where the experts only report one land cover type, this is a relative straightforward process. The set of selected classifier values is transformed into a single GlobCorine class, according to the GlobCorine-LCCS correspondence given in Appendix A.

In case the experts described two or three land cover types to describe the area covered by an observational unit, the translation process becomes less obvious. Indeed, the fact that several land cover types have been identified for one observational unit gives cause to consider mosaic classes as well. The GlobCorine legend includes 3 mosaic classes (NB_LAB 100, 110 and 120), which correspond to the combinations of land cover types indicated by the experts.

For illustration purpose, Table 3 reports the values of LCCS classifiers selected by the expert to describe a given observational unit. In Europe, the experts assigned maximum two land cover types to the observational units. These two sets of classifiers can then be translated into two different “pure” GlobCorine classes and into one “mosaic” GlobCorine class, as shown in the Table 4. These different possible translations of the classifiers set provided by the expert must be taken into account to analyze the confusion matrix comparing the GlobCorine product with the validation dataset.

Furthermore, it is worth mentioning that many combinations of land cover types cannot be transformed into a GlobCorine mosaic class. Indeed, a legend that would cover for all these potential combinations is not desirable because the mosaic classes are often considered less informative and therefore less useful from an end-user point of view.

Land cover 1 (LC1)	**Land cover 2 (LC2)**
Natural & Semi-natural terrestrial vegetation	Cultivated & managed lands
Shrubs	Herbaceous
Open (70-60 - 20-10%)	Rainfed
5-0.3 m	
Broadleaved	

Table 3. Two sets of LCCS classifiers that describe the land cover for an observational unit out of the validation dataset

Pure GlobCorine class describing LC 1	**Pure GlobCorine class describing LC 2**	**GlobCorine mosaic class describing the combination of LC1 and LC2**
Heathland and sclerophyllous vegetation	Rainfed herbaceous cropland	Mosaic cropland / natural vegetation

Table 4. GlobCorine classes to which the land cover types from Table 3 have been assigned

3.1.3. Validation results

Over the 430 “certain” points, 403 could be translated into a GlobCorine class. The reference dataset of these 403 “certain and valid” points is then matched to the GlobCorine map codes extracted for all the validation points in order to build a confusion matrix (Table 5). It is important to mention that the dominance between land cover types identified by the expert for a given sample has not been taken in account in the validation process.

The accuracy level is found to be 79.9% Next to the overall accuracy, the user’s and producer’s accuracies have also been computed.

		Reference															
		10	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**	**110**	**120**	**200**	**210**		
GlobCorine	**10**	5														5	100%
	20	1	40	6	1		1				3	12				64	85.9%
	30		1	7							1					9	88.9%
	40		2		61	1	2			1	1	14	9			91	92.3%
	50															0	-
	60		8		2		1					1	3			15	33.3%
	70	2	1		2	2	7	1		5	1		2	1	2	26	3.9%
	80												1			1	0%
	90		1			1				63						65	96.9%
	100		10													10	100%
	110	3	40	20	4		4			2	7	10	1			78	85.7%
	120		4		4	2	1					8				7	36.8%
	200													7		7	33.3%
	210															0	100%
		11	107	33	71	6	16	1	0	71	13	37	12	8	2	403	
		45.5%	84.1%	81.8%	93.2%	33.3%	37.5%	100%	-	88.7%	30.8%	82.2%	75%	87.5%	0%		**79.9%**

Table 5. Adjusted contingency matrix that considers the product and the validation dataset. Green cells are cells that show direct agreement between classification and validation. Yellow cells show cells that have been considered to show agreement, accounting for the definition of mosaic classes. The value in red shows the overall accuracy

The previous accuracy results are derived with equal weighting for each of the stratified randomly sampled reference points. Classes that cover only small surfaces are overrepresented in the sample set and conversely, classes that cover large surfaces may have been underrepresented in the set. As recommended by the CEOS, the overall accuracy value has to be weighted by the area proportions of the various land cover classes, resulting in an accuracy of **89.25%**. The weighting factor corresponding to the area proportion of the given class has been derived from the GlobCorine product that has been projected in an equal area projection.

These final accuracy results document the quality of the GlobCorine product. These figures appear quite acceptable for a pan-continental product. Nevertheless, these figures have to be used cautiously, i.e. they need to be moderated by the users and producers accuracies and by the sampling validation strategy.

First, Table 5 reveals a clear contribution of the mosaic classes in the high global accuracy figure. Indeed, their agreement with several (both pure and mosaic) classes increases the global accuracy. However, these classes are not easily interpretable and they should thus be avoided as much as possible. According to the GlobCorine classification chain, the mosaic classes can be created from two different ways. First, they can be created because either the input data or the classification algorithm did not allow a proper discrimination between different land cover types. Second, they can be created because they are present in the reference dataset that is used in the labeling step. This second explanation is principally valid over the European Russia where the reference dataset is the GlobCover 2005 land cover map. In this region, the mosaic classes are, in a way, “inherited” from the reference dataset and have nothing to do with the performance of the classification algorithm or with the suitability of the MERIS data. In other regions, the surface of mosaic classes is smaller than in GlobCover. Even if the discrimination issue of mosaic landscape is still not totally solved, there are significant improvements, which demonstrate the pertinence of the GlobCorine classification algorithm.

Second, Table 5 reveals that the classes that perform best have been validated with the largest numbers of points, which confirms the high accuracy of these classes.

Table 5 also indicates that the number of validation points highly varies between classes. The stratified sampling that generated the validation dataset was indeed achieved on a global scale, based on the GlobCover product. The stratification, which ensures that each class is representatively sampled, is thus not necessarily valid at the GlobCorine pan-European scale. This validation exercise has to deal with this slight bias, which mainly affects the classes of “Heathland and sclerophyllous vegetation”, of “Sparsely vegetated areas”, of “Vegetated low-lying areas on regularly flooded soil” and of “Permanent snow and ice”. This bias has also a clear influence on the overall accuracy value weighted by the class area, which is then artificially increased. Accordingly, it is strongly recommended to also account for user’s and producer’s accuracies to have a clear idea of the GlobCorine map accuracy.

3.2. Land cover validation based on the CLC2006 dataset

A quantitative evaluation of the GlobCorine product has also been achieved on the pan-European level by the ETCLUSI (GISAT), using the CLC2006 data as reference dataset. It has to be stated that large countries like Finland, Germany, Greece, Italy, Norway, Spain, Sweden and United Kingdom were missing in this reference dataset.

The CLC2006 dataset has been re-mapped into the GlobCorine legend and GlobCorine reference grid. The correspondence between the GlobCorine and the CLC classes can be found in Appendix B. No sampling was defined, thus considering all pixels with valid GlobCorine product.

Table 6 shows the agreement table between the GlobCorine product and the re-mapped CLC2006 dataset and Table 7 presents the corresponding accuracy statistics.

		GlobCorine													
		10	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**	**110**	**120**	**200**	**210**
Re-mapped CLC2006 dataset	**10**	72	11.1	0.1	3.4	0.5	2.3	0.1	0.3	0.1	5.1	2.8	0.9	1.2	0
	20	4.7	66.7	0	4.7	0.5	9	0.4	0.3	0.1	8.2	3.9	1.1	0.3	0
	30	0.7	5.5	75.3	4.9	0.4	1	0.3	0.5	0.1	2.9	1	0.7	0.2	0
	40	0.9	3.5	0	71.6	0.9	7.3	0.1	0.3	0.1	3	5.4	6.4	0.4	0
	50	0.3	4.1	0.2	3.8	72.2	3.1	2	2.8	1.4	1.5	3.4	4.4	0.4	0
	60	1.5	8.4	0	10	0.7	61.4	0.3	2.4	0.1	7.3	5.2	2.4	0.2	0
	70	7.7	5.8	0.1	1.2	5.3	5.9	55.1	1.6	11.7	0.7	0.4	1.4	1	0.1
	80	0.9	7.5	0	5.8	2.1	3.9	0.4	67.4	0.7	0.4	1.2	7.2	2.1	0
	90	1.1	0.2	0	0.2	1.4	1	4.1	0.3	89.1	0	0	0	0.4	1
	100	6.7	33.3	0.6	9.1	1.4	15.9	0.3	0.3	0.1	23.1	6.4	2.3	0.3	0
	110	5.6	33.6	0.4	13.1	2.5	16.8	1.2	0.7	0.5	11.7	8.5	4.7	0.7	0
	120	2.3	13.4	0.1	27.7	2.9	19.4	0.8	0.9	0.3	10.9	11.5	9.4	0.3	0
	200	1.3	0.9	0	1.7	0.5	1.5	0.2	7.8	1.2	0.4	0.7	0.4	83.3	0
	210	0	0	0	0	0	0	0.2	0	4.2	0	0	0	0	95.4

Table 6. Matrix showing the agreement between the GlobCorine product and the re-mapped CLC2006 dataset. Green cells are cells that show direct agreement between GlobCorine and CLC2006.

Average accuracy	60.75%
Overall accuracy	52.39%
Kappa coefficient	0.51304

Table 7. Accuracy statistics showing the agreement between the GlobCorine product and the re-mapped CLC2006 dataset

Table 6 shows that the overall accuracy is affected by mosaic classes present in the GlobCorine legend (classes 100, 110 and 120). That is mainly due to the difference in mosaic/non-mosaic classes share as well as to the mosaic class structure, which are quite distinct between the GlobCorine product and the re-mapped CLC2006 dataset. Figure 10 confirms that observation.

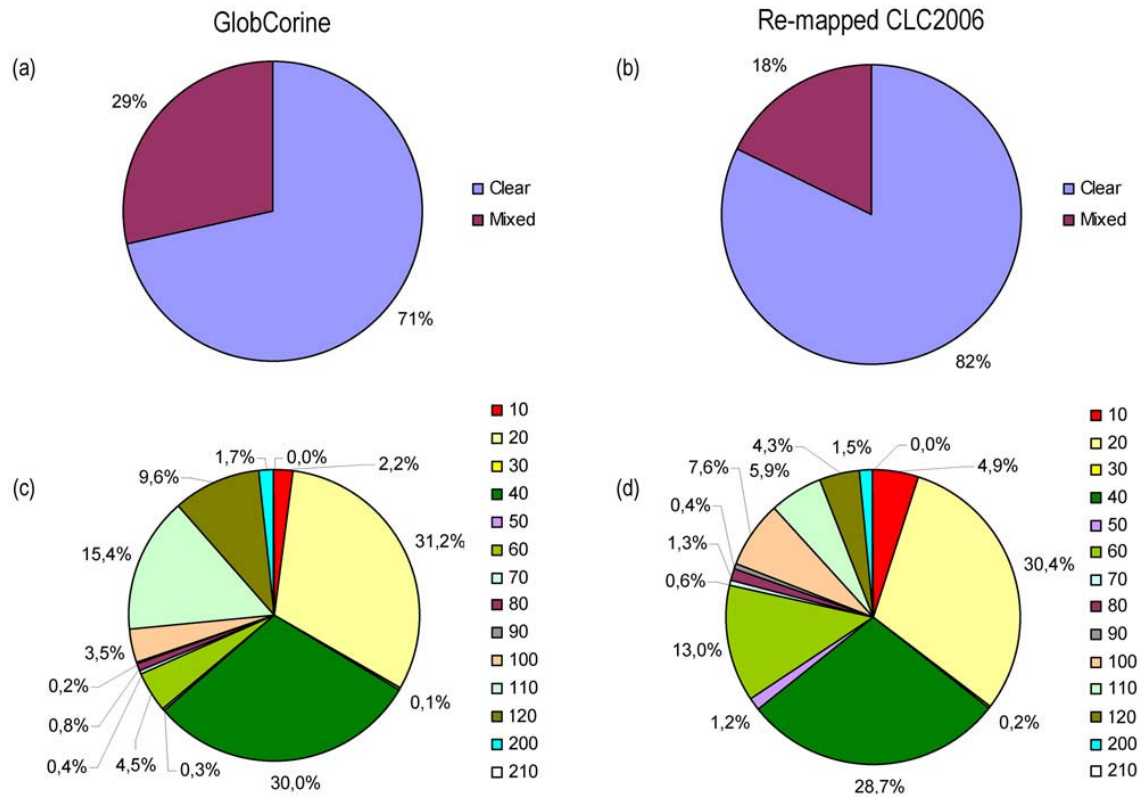

Figure 10. Proportions of pure and mosaic classes (top (a) and (b) figures) and of each class of the legend (bottom (c) and (d) figures) in the GlobCorine product (left column) and in the re-mapped CLC2006 dataset (right column)

Table 8, Table 9 and Figure 11 demonstrate the increase of the global agreement between the GlobCorine product and the re-mapped CLC2006 dataset when the mosaic classes are not considered. Overall class share is similar, except the grassland and urban areas which are clearly underestimated in the GlobCorine product. Water and wetland areas are represented quite similarly.

		GlobCorine										
		10	20	30	40	50	60	70	80	90	200	210
Re-mapped CLC2006 dataset	10	72	11.1	0.1	3.4	0.5	2.3	0.1	0.3	0.1	1.2	0
	20	4.7	66.7	0	4.7	0.5	9	0.4	0.3	0.1	0.3	0
	30	0.7	5.5	75.3	4.9	0.4	1	0.3	0.5	0.1	0.2	0
	40	0.9	3.5	0	71.6	0.9	7.3	0.1	0.3	0.1	0.4	0
	50	0.3	4.1	0.2	3.8	72.2	3.1	2	2.8	1.4	0.4	0
	60	1.5	8.4	0	10	0.7	61.4	0.3	2.4	0.1	0.2	0
	70	7.7	5.8	0.1	1.2	5.3	5.9	55.1	1.6	11.7	1	0.1
	80	0.9	7.5	0	5.8	2.1	3.9	0.4	67.4	0.7	2.1	0
	90	1.1	0.2	0	0.2	1.4	1	4.1	0.3	89.1	0.4	1
	200	1.3	0.9	0	1.7	0.5	1.5	0.2	7.8	1.2	83.3	0
	210	0	0	0	0	0	0	0.2	0	4.2	0	95.4

Table 8. Matrix showing the agreement between the GlobCorine product and the re-mapped CLC2006 dataset without considering the mosaic classes 100, 110 and 120. Green cells are cells that show direct agreement between GlobCorine and CLC2006.

Average accuracy	79.17%
Overall accuracy	79.73%
Kappa coefficient	0.70919

Table 9. Accuracy statistics showing the agreement between the GlobCorine product and the re-mapped CLC2006 dataset without considering the mosaic classes 100, 110 and 120

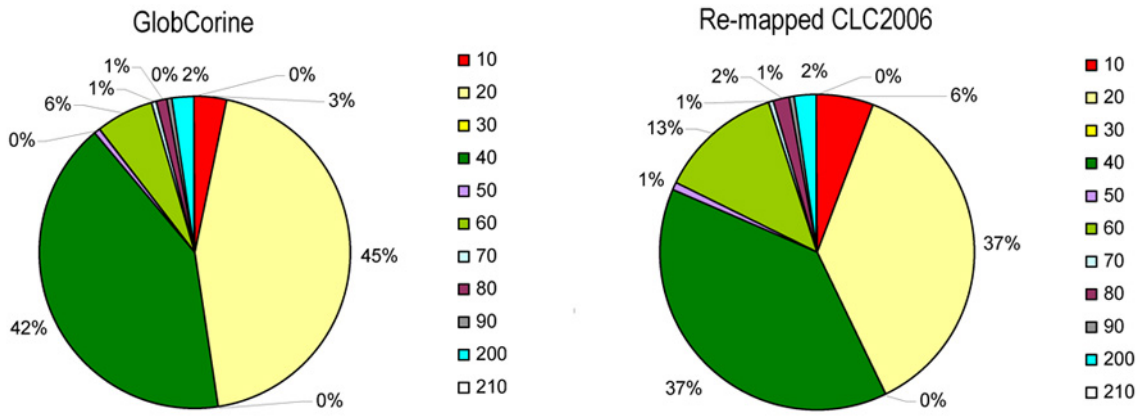

Figure 11. Proportions of each pure class in the GlobCorine product (left column) and in the re-mapped CLC2006 dataset (right column)

4. Recommendations – discussion - conclusion

This document reports on the accuracy of the 300-m pan-European land cover product delivered for the period December 2004 - June 2006 in the course of this GlobCorine project.

There were at the beginning of the project two main challenges which have been met: adjusting the GlobCover classification chain to produce from the 300-m MERIS dataset a land cover product dedicated to the pan-European continent and as compatible as possible with the CLC aggregated typology which is more land use oriented.

The 14-classes GlobCorine legend is coherent and links both the LCCS and the CLC map.

Furthermore, two comprehensive validation exercises have been completed providing quantitative figures of the product accuracy. According to the exercise based on the GlobCover validation dataset, the overall accuracy reaches 79.9% (and up to 89.25% when the overall accuracy value is weighted by the class area) using 403 points globally distributed and including homogeneous and heterogeneous landscapes. This accuracy is slightly higher than that of GlobCover which yet brought improvement with regard other existing global land cover products. According to the exercise involving a direct comparison with the CLC2006 map, the overall agreement reaches 52.39% considering all classes and 79.73% without considering the mosaic classes.

These very positive figures must be balanced by the fact that the GlobCorine map quality varies according to the thematic class and to the region of interest. Looking at the contingency matrices (Table 5, Table 6 and Table 8) gives indication about the accuracy of each class. Land cover classes such as the “Rainfed cropland”, the “Forest” and the “Bare areas” were found quite accurately mapped. On the other hand, classes such as the “Heathland and sclerophyllous vegetation”, the “Grassland” and the “Complex cropland” can be affected by errors. The number of points used for the validation exercise based on the GlobCover dataset has also to be taken into consideration, since this dataset does not ensure that each class is representatively sampled. Typically, the classes of “Sparsely vegetated areas”, of “Vegetated low-lying areas on regularly flooded soil” and of “Permanent snow and ice” is under-sampled.

The two validation exercises emphasize the impact of the mosaic classes on the overall accuracy. Pure GlobCorine classes (e.g. cropland, forest, wetland, etc.) reach good correspondence with the CLC map. However, this is not the case for the GlobCorine mosaic classes, which poorly match to the mosaic classes in the CLC map.

The surface of mosaic classes has been significantly reduced in the GlobCorine product with regard to the GlobCover product, which considerably helps data interpretation. Other notable improvements compared to GlobCover concern the better identification and delineation of the urban areas and of the wetlands.

The limitation of the GlobCorine product can mainly be explained by several strategic choices. Only MERIS data could be used, always missing the critical SWIR band for the forest discrimination. The automation of the interpretation chain requires us to rely on the class labeling on the already existing but sometimes coarser land cover products (typically in the Northern Africa and in the European Russia). From the end users point of view, the mosaic classes remain problematic despite the notable improvement with regard to the GlobCover product since they limit the thematic sharpness of the GlobCorine product.

Finally, it has to be stated that the GlobCorine classification project clearly demonstrates the operational service provided the automated processing chain previously developed by the GlobCover consortium.

It is important to recall that the GlobCorine land cover map cannot be used for any change detection application. Indeed, such land cover map accuracy surely prevents any consistent comparison with older maps to depict the change area. As a result, the change rate will always be much lower than the classification errors, thus hampering any relevant use for change mapping.

Appendix A. Correspondence between the GlobCorine legend and the LCCS

VALUE	LABEL	LCCS MAIN TYPE	LCCS CODES
10	Urban and associated areas	B15	XX
20	Rainfed cropland	A11	A1D1// A2D1// A3D1
30	Irrigated cropland	A11	A1D3//A2D3//A3D2//A3D3
40	Forest	A12	A3A11XXXXXXXX//A3A13XXXXXXXX//A3A20XXXXXXXX//A3A21XXXXXXX
50	Heathland and sclerophyllous vegetation	A12	A1A11XXXXXXXX//A1A13XXXXXXXX//A1A20XXXXXXXX//A1A21XXXXXXX//A4A11XXXXXXXX//A4A13XXXXXXXX//A4A20XXXXXXXX//A4A21XXXXXXXX
60	Grassland	A12	A2A11XXXXXXXX//A2A13XXXXXXXX//A2A20XXXXXXXX//A2A21XXXXXXX
70	Sparsely vegetated area	A12	A1A14XXXXXXXX//A2A14XXXXXXXX//A3A14XXXXXXXX//A4A14XXXXXXX//A7A14XXXXXXXX
80	Vegetated low-lying areas on regularly flooded soil	A24	A1A21XXXXXX//A2A21XXXXXX//A1A20XXXXXX//A2A20XXXXXX
90	Bare areas	B16	XX
100	Complex cropland	A11/A11	A1D1//A1D3//A2D1//A2D3//A3D1//A3D2//A3D3/A1D1//A1D3//A2D1//A2D3//A3D1//A3D2//A3D3

110	Mosaic cropland / natural vegetation	A11/A12	A1D1//A1D3//A2D1//A2D3//A3D1//A3D2//A3D3/A2A11XXXXXXXX//A2A13XXXXXXXX//A2A20XXXXXXXX//A2A21XXXXXXXX//A1A11B3XXXXXX//A1A13B3XXXXXX//A1A20B3XXXXXX//A1A21B3XXXXXX//A4A11B3XXXXXX//A4A13B3XXXXXX//A4A20B3XXXXXX//A4A21B3XXXXXX//A3A11B2XXXXXX//A3A13B2XXXXXX//A3A20B2XXXXXX//A3A21B2XXXXXX
120	Mosaic of natural (herbaceous, shrub, tree) vegetation	A12/A12	A2A11XXXXXXXX//A2A13XXXXXXXX//A2A20XXXXXXXX//A2A21XXXXXXX/A1A11B3XXXXXX//A1A13B3XXXXXX//A1A20B3XXXXXX//A1A21B3XXXXXX//A4A11B3XXXXXX//A4A13B3XXXXXX//A4A20B3XXXXXX//A4A21B3XXXXXX//A3A11B2XXXXXX//A3A13B2XXXXXX//A3A20B2XXXXXX//A3A21B2XXXXXX/A2A11XXXXXXXX//A2A13XXXXXXXX//A2A20XXXXXXXX//A2A21XXXXXXXX/A1A11B3XXXXXX//A1A13B3XXXXXX//A1A20B3XXXXXX//A1A21B3XXXXXX//A4A11B3XXXXXX//A4A13B3XXXXXX//A4A20B3XXXXXX//A4A21B3XXXXXX//A3A11B2XXXXXX//A3A13B2XXXXXX//A3A20B2XXXXXX//A3A21B2XXXXXX
200	Water bodies	B27/B28	A1
210	Permanent snow and ice	B27/B28	A2//A3

Appendix B. Correspondence between the GlobCorine and CLC nomenclatures

VALUE	LABEL	CLC CODES (LEVEL 3)	CLC LABELS (LEVEL 3)
10	Urban and associated areas	111	Continuous urban fabric
		112	Discontinuous urban fabric
		121	Industrial or commercial units
		122	Road and rail networks and associated land
		123	Port areas
		124	Airports
		131	Mineral extraction sites
		132	Dump sites
		133	Construction sites
		141	Green urban areas
		142	Sport and leisure facilities
20	Rainfed cropland	211	Non-irrigated arable land
		221	Vineyards
		222	Fruit trees and berry plantations
		223	Olive groves
30	Irrigated cropland	212	Permanently irrigated land
		213	Rice fields

40	Forest	311	Broad-leaved forest
		312	Coniferous forest
		313	Mixed forest
50	Heathland and sclerophyllous vegetation	322	Moors and heathland
		323	Sclerophyllous vegetation
60	Grassland	231	Pastures
		321	Natural grasslands
70	Sparsely vegetated area	333	Sparsely vegetated areas
80	Vegetated low-lying areas on regularly flooded soil	411	Inland marshes
		412	Peat bogs
		421	Salt marshes
		423	Intertidal flats
90	Bare areas	332	Bare rocks
		331	Beaches, dunes, sands
100	Complex cropland	241	Annual crops associated with permanent crops
		242	Complex cultivation patterns
110	Mosaic cropland / natural vegetation	243	Land principally occupied by agriculture, with significant areas of nat. veg.
		244	Agro-forestry areas

120	Mosaic of natural (herbaceous, shrub, tree) vegetation	324	Transitional woodland-shrub
200	Water bodies	511	Water courses
		512	Water bodies
		521	Coastal lagoons
		522	Estuaries
		523	Sea and ocean
210	Permanent snow and ice	335	Glaciers and perpetual snow