ACCURACY ASSESSMENT OF GLOBAL LAND COVER MAPS - LESSONS LEARNT FROM GLOBCOVER AND GLOBCORINE EXPERIENCES

Defourny P.\(^{(1)}\), Bontemps S.\(^{(1)}\), Obsomer V.\(^{(1)}\), Schouten L.\(^{(2)}\), Bartalev S.\(^{(3)}\), Cacetta P.\(^{(4)}\), de Wit A.\(^{(5)}\), di Bella C.\(^{(6)}\), Gérard B.\(^{(7)}\), Giri C.\(^{(8)}\), Gond V.\(^{(9)}\), Hazeu G.\(^{(5)}\), Heiniman A.\(^{(10)}\), Herold M.\(^{(11)}\), Jaffrain C.\(^{(12)}\), Latifovic R.\(^{(13)}\), Lin H.\(^{(14)}\), Mayaux P.\(^{(15)}\), Mücke S.\(^{(5)}\), Nanguierma A.\(^{(16)}\), Stibig H.J.\(^{(15)}\), Van Bogaert E.\(^{(1)}\), Vancutsem C.\(^{(15)}\), Bicheron P.\(^{(17)}\), Leroy M.\(^{(18)}\), Arino O.\(^{(19)}\)

\(^{(1)}\)Université catholique de Louvain, Earth and Life Institute, Croix du Sud 2-16, B-1348 Louvain-la-Neuve, Belgium, Email : (Sophie.Bontemps, Pierre.Defourny, Eric.Vanbogaert)@uclouvain.be
\(^{(2)}\)INFRAM B.V., Postbus 16, 8316 ZG Marknesse, The Netherlands, Email : leon.schouten@infram.nl
\(^{(3)}\)Space Research Institute, Russian Academy of Science, 117997, 84/32 Profsoyuznaya str., Moscow, Russia, Email : bartalev@d902.iki.rssi.ru
\(^{(4)}\)Commonwealth Scientific and Industrial Research Organization, Floreat, Western Australia Leeuwinn Centre, 65 Brockway Road, Floreat WA 6014, Australia, Email : Peter.Caccetta@csiro.au
\(^{(5)}\)Alterra, Team Earth Observation, Postbus 47, 6700AA, Wageningen, The Netherlands, Email : (Allard.Dewit, Gerard.Hazeu, Sander.Macher)@uw.nl
\(^{(6)}\)Instituto Nacional de Tecnologia Agropecuaria and Consejo Nacional de Investigaciones Cientificas y Tecnicas, Los Reseros y Las Cabañas S/N (CD 1686), Hurlingham, Buenos Aires, Argentina, Email : edibella@cnia.inta.gov.ar
\(^{(7)}\)International Livestock Research Institute, PO Box 5689, Addis Ababa, Ethiopia, Email : B.Gerard@CGIAR.ORG
\(^{(8)}\)United States Geological Survey, EROS Data Center, Sioux Falls, SD, 57198, United States, Email : cgiri@usgs.gov
\(^{(9)}\)CIRAD - ES, UPR 105 Biens et services des écosystèmes forestiers tropicaux, Campus de Baillarguet TA C-105/D-214, 34398 Montpellier Cedex 5, France, Email : valery.gond@cirad.fr
\(^{(10)}\)National Centre of Competence in Research North-South – Centre for Development and Environment (CDE), Hallerstrasse 10, 3012 Bern, Switzerland, Email : Andreas.Heinimann@cde.unibe.ch
\(^{(11)}\)GOFC-GOLD, University Jena, Loebbergraben 32, 07737 Jena, Germany, Email : m.h@uni-jena.de
\(^{(12)}\)IGN France International, Maison Mozart, 8 rue du Sentier, 75002 Paris, France, Email : gjaffrain@ignfi.fr
\(^{(13)}\)Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Canada, Email : latifovic@nrcan.gc.ca
\(^{(14)}\)Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A, Datun Road, Chaoyang District, Beijing, 100101 China, Email : liujy@igsnrr.ac.cn
\(^{(15)}\)EU- Joint Research Center, Via Enrico Fermi, 1, 21020 Ispra, Italy, Email : { philippe.mayaux, hans-juergen.stibig, christelle.vancutsem}@jrc.it
\(^{(16)}\)Economic Commission for Africa, Menelik II Avenue, P. O. Box 3001, Addis Ababa, Ethiopia, Email : andre.nonguierma@yahoo.com
\(^{(17)}\)Infoterra, 31 rue des Cosmonaute, 31402 Toulouse Cedex 4, France, Email : Patrice.Bicheron@infoterra.fr
\(^{(18)}\)AMESD Program Coordination Team, African Union Commission, P. O. Box 200288 Addis Ababa, Ethiopia, Email : marc.leroy@amesd.org
\(^{(19)}\)European Space Agency, European Space Research Institute, Via Galileo Galilei, 00044 Frascati, Italy, Email: Olivier.arino@esa.int

ABSTRACT

The GlobCover project supported by ESA has developed an operational service dedicated to the generation of global land cover maps through an automated classification of MERIS FRS time series. This paper reports the independent accuracy assessment of the global GlobCover product as the first global exercise implemented according to the CEOS Land Product Validation Subgroup recommendations. Based on a network of 16 international experts and on-line tools, a unique, globally distributed reference dataset was collected in a standardized manner and used to derive mapping accuracy figures. The overall accuracy, weighted by the area proportions of the various land cover classes, is 73% based on a set of 3167 samples. These results are discussed with regards to the previous experiences.

1. INTRODUCTION

The global GlobCover land cover product, which was derived from data acquired by the ENVISAT MERIS instrument, is a scientific and technical demonstration of the first automated mapping of land cover on a global scale. This product of 300-m spatial resolution is available to a broad-level stakeholder community on the ESA website [1]. This was achieved by the GlobCover consortium (mainly MEDIAS, UCL-Geomatics and Brockman Consult), sponsored by the European Space Agency (ESA) and supported by an international partnership including EU-JRC, FAO, EEA, UNEP,

Proc. ‘ESA Living Planet Symposium’, Bergen, Norway
28 June – 2 July 2010 (ESA SP-686, December 2010)
GOFC-GOLD and IGBP. It was also deemed essential to deliver detailed information on the GlobCover product accuracy which allows a potential user to determine the map’s “fitness for use” for a given application.

The accuracy of a land cover map has two components: the geometric accuracy and the thematic accuracy. In the GlobCover context, the geometric navigation of the MERIS Fine Resolution full Swath (FRS) images was the very first challenge because the technical accuracy specification of this ocean color instrument was 2 km. This would have been a major constraint for land cover mapping applications. The GlobCover preprocessing chain managed to reduce the absolute geo-location root mean square error (RMSE) in the end to 77 meters [1, 2]. This allowed the team to use the MERIS time series for the classification and it also enables end-users to work with the final land cover product. The thematic accuracy assessment of the 300-m spatial resolution global land cover map was another challenge. This was also the opportunity to implement the CEOS Land Product Validation Subgroup recommendations for the first time at global scale [3].

The quantitative validation of the thematic accuracy aimed at assessing the accuracy of the 22-classes land cover map using an independent reference dataset. In the case of the IGBP DISCover global land cover map [4] and of the GLC2000 map [5], the reference dataset was built by visually interpreting a 50-m spatial resolution orthorectified Landsat color composite. [6] assessed the overall accuracy of the GLC2000 land cover map (21 classes with a spatial resolution of 1 km) using 544 blocks dominated by one land cover class (>80% of the area) and selected by a two-stage sampling using the Landsat World Reference 2 System. They found an area-weighted overall accuracy of 68.6±5% for a 95% confidence interval which was very similar to the IGBP DISCover accuracy for 17 classes at 1.1 km.

Based on the CEOS recommendations [3], the validation process was designed to be scientifically sound, internationally acceptable and feasible from cost and time point of views. The validation plan was adopted from the very beginning of the GlobCover project, before starting land cover map production. An independent stakeholder – namely the private company Infram International B.V. – developed the data collection tool and completed the data analysis for the accuracy assessment.

2. METHODOLOGY

A dozen of GlobCover land cover maps, including both global and regional products, were produced using the MERIS FRS time series acquired from December 2004 to June 2006 based on the same automatic processing chain. The land cover typology, fully described according to the UN Land Cover Classification System (LCCS) [7] consists in 22 land cover classes for the global level and is extended to 51 land cover classes when including those consistently discriminated only at the regional level.

The validation process that is reported here concerns only the global GlobCover map. It included three different steps: collecting reference data, elaborating the sampling strategy and assessing the product’s accuracy.

2.1. Reference data collection

The reference data collection could only rely on already existing expertise distributed all over the world. The creation of an international expert network was the key element of the validation process. The experts were selected according to the following criteria: undisputed expertise on land cover over relative large areas, familiarity with interpreting remote sensing imagery, commitment, complementarities to the other experts and belonging to well-known international networks.

Sixteen experts from all over the world were invited for six different 5-day workshops hosted by UCL (Louvain-la-Neuve, Belgium). The experts truly committed themselves to build the GlobCover reference dataset. Some of them could not join the working session but were familiar enough with the LCCS system and the validation process to complete their job remotely using the same tools.

All experts used a dedicated working environment for on-screen collection of “ground truth” data. The validation samples were automatically overlaid either in Virtual Earth or Google Earth allowing a rapid access to recent remote sensing images with zooming capabilities. In addition, NDVI profiles (8 annual profiles from 2000 to 2007 and the corresponding average profile) were made available for each validation sample, thus complementing the interpretation of the high resolution imagery by its seasonal dynamics. The NDVI profiles were extracted from the 1-km spatial resolution 10-day SPOT VEGETATION time series and composited by UCL [8]. Finally, the experts could also support their work using any additional sources of information such as detailed maps.

2.2. Sampling design

In order to ensure that each pixel had an equal chance of being sampled, the GlobCover product was projected to the Lambert azimuthal equal area projection. As there is no equal area projection that does justice to the entire world, the world was divided into five regions (Africa,
Australia & Pacific, Eurasia, North America and South America) for which it was possible to apply an equal area projection. The samples were then selected using a stratified random sampling.

For a given sample, the expert saw not only the sample point but also a box that coincided with the so-called observational unit corresponding to 5x5 MERIS pixels (225 hectares). The effective observational unit was not necessarily a square or a circle around the point. Some land cover classes, notably lakes and wetlands, can indeed be rather elongated but cannot be discarded because of the shape of the observational unit. The main purpose of the box was to give an idea of the extent of an area of 225 hectares.

The experts labeled a single dominant land cover type when more than 75% of the observational unit belongs to the same type. If two or three land cover types covered each between 25 and 75% of the observational unit, they should be described as well. In addition, the experts were invited to give their level of confidence for the labelling according to three levels.

### 2.3. Accuracy assessment

To enhance the potential use of the GlobCover validation dataset beyond the scope of the actual project, it was strongly recommended to gather information on the LCCS classifiers in order to characterise the land cover of each validation sample irrespective of the current GlobCover typology. This is why this validation dataset was not specifically related to the current GlobCover legend and needed to be translated into the 22 land cover classes of the global GlobCover product afterwards.

For validation points or observational units where the international experts only reported one land cover type, this translation was a relative straightforward process. The set of selected classifier values was transformed by Infram B.V. into a single GlobCover class. When two or three land cover types described the area covered by a sample, this translation process became less obvious. In addition to the single translation into the respective classes, it was also necessary to consider assigning the sample to a mosaic class. As the combinations of three land cover types allowed various interpretations, the Infram team assigned up to 2 different mosaics to some of the samples.

For illustration purpose, Tab. 1 reports the values of LCCS classifiers selected by the expert to describe a given observational unit covered by three different land cover types. These three sets of classifiers can then be translated into three different GlobCover classes, as shown in the Tab. 2.

The fact that 3 land cover types were identified for one observational unit gives caused to consider mosaic classes as well. The expert described the most dominant land cover type first, followed by the land cover type that was second most dominant and, in some cases, a third land cover type was also described.

**Table 1. Three sets of LCCS classifiers that describe the land cover for an observational unit out of the validation dataset**

<table>
<thead>
<tr>
<th>Land Cover 1</th>
<th>Land Cover 2</th>
<th>Land Cover 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural &amp; Semi-natural terrestrial vegetation</td>
<td>Cultivated lands &amp; managed lands</td>
<td>Natural &amp; Semi-natural terrestrial vegetation</td>
</tr>
<tr>
<td>Shrub</td>
<td>Herbaceous</td>
<td>Trees</td>
</tr>
<tr>
<td>Open (70-60% - 20-10%)</td>
<td>Rainfed</td>
<td>Open to very open (40-20% - 10%)</td>
</tr>
<tr>
<td>5-0.3 m</td>
<td>&gt;3-30 m (for Trees)</td>
<td>Broadleaved evergreen</td>
</tr>
</tbody>
</table>

**Table 2. GlobCover classes to which the land cover types from Tab. 1 have been assigned**

<table>
<thead>
<tr>
<th>GlobCover class describing Land Cover 1</th>
<th>GlobCover class describing Land Cover 2</th>
<th>GlobCover class describing Land Cover 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed to open (&gt;15%) (broadleaved or needle-leaved, evergreen or deciduous) shrubland (&lt;5m)</td>
<td>Rainfed croplands</td>
<td>Closed to open (&gt;15%) broadleaved evergreen or semi-deciduous forest (&gt;5m)</td>
</tr>
</tbody>
</table>

In addition, there were two mosaic classes in the GlobCover legend which could also describe the land cover within the concerned observational unit:

- Mosaic vegetation (grassland / shrubland / forest) (50-70%) / cropland (20-50%)
- Mosaic cropland (50-70%) / vegetation (grassland / shrubland / forest) (20-50%)

These different possible translations of the classifier set provided by the expert to describe a given validation sample had to be taken into account to analyse the confusion matrix comparing the GlobCover product with the validation dataset.

Furthermore, it is worth mentioning that many combinations of land cover types could not be transformed to a GlobCover 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 the end-user point of view.

3. RESULTS

The GlobCover validation dataset contained 4258 samples. This number included, in addition to the expert efforts, 341 samples corresponding to river basins ground truth which originate from International Water Management Institute (IWMI) to fill partially a gap in the Indian subcontinent.

In 3167 cases, the experts were explicitly certain of information they provided. In 797 cases, they were reasonably sure and in 294 cases, they had some reservations. The 3167 “certain” samples distribution is shown in Fig. 1.

Finally, to explore the effect of heterogeneous areas, the validation dataset was even further reduced to 2115 samples by removing all the points for which the experts needed to define more than one land cover type.

The reference dataset was then crossed with the GlobCover map and land cover codes were extracted for all the validation points in order to build a confusion matrix. As the notion of dominance between land cover types was not quantified for a given sample, it was not taken in account in the validation process.

As recommended by the CEOS recommendations, the overall accuracy values derived from the confusion matrix were weighted by the area proportions of the various land cover classes. The weighting factor corresponding to the area proportion of the given class was derived from the GlobCover product that was projected in an equal area projection. Tab. 3 reports the results.

Table 3. Accuracy of the global GlobCover map

<table>
<thead>
<tr>
<th>GlobCover validation dataset</th>
<th>Global accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>3167 “certain” points</td>
<td>73.14%</td>
</tr>
<tr>
<td>2115 “certain” &amp; “homogeneous” points</td>
<td>79.25%</td>
</tr>
</tbody>
</table>

These final accuracy results documented the quality of the GlobCover product. This accuracy was higher than that of GLC2000 with yet a spatial resolution improved by a factor 3.3, thus resulting in a product ten times better if the pixel area is considered.

This very positive figure must be balanced by the fact that the GlobCover map quality varied according to the region of interest. Looking at the number of valid observations available over a region (Fig. 2) gives a priori information about the input data quality and the expected classification reliability.

From a thematic point of view, land cover classes such as the evergreen and semi-deciduous forest, the irrigated croplands, the bare areas, the water bodies and the snow were found quite accurately mapped. On the other hand, classes such as the urban areas, the sparse vegetation and the herbaceous vegetation were more affected by errors.

As a very positive figure must be balanced by the fact that the GlobCover map quality varied according to the region of interest. Looking at the number of valid observations available over a region (Fig. 2) gives a priori information about the input data quality and the expected classification reliability.

Classification patterns of wetlands, grasslands and shrublands showed clear discrepancies with interpretations of experts. This might be due to the
absence of a mid-infrared channel that might affect the ability to identify these land cover types on the MERIS data.

4. DISCUSSION

Generally speaking, the classification methodology was still constrained by the quality and especially the amount of the reference database product. The global GlobCover product has 22 classes. It is obvious that the real world is far more heterogeneous than this model of the world. This aspect of a global land cover product needs to be emphasised and users of the product need to realise this!

There are still a number of known issues and artefacts which GlobCover users should consider. Some regions of the world (e.g. some areas in Amazonia) were poorly covered by MERIS FRS data. The limited number of valid MERIS FRS observations could have several effects on the land cover map. In areas of very low data coverage (about 2% of the terrestrial areas), the pixel values were derived from already existing dataset. When the data coverage was poor, there was still a tendency for GlobCover to overestimate forest areas.

It is important to recall that this land cover map accuracy surely prevents from any use of the map for land cover change detection or comparison with older maps to depict the change area. Indeed, the change rate would always be much lower than the land cover dynamics, thus hampering any relevant use for change mapping.

5. CONCLUSION

The concept of a global Land Cover service operational at global scale first requested by ESA has been developed and successfully validated. Implementing a globally consistent while regionally-tuned classification processing system allowed moving away from ad hoc interpretation strategies often used in the past. This system is an automatic and repeatable process enabling the production of a land cover map using imagery from other years. This significant step forward probably opens new avenues for the land cover community as well as for downstream applications.

Similarly, the validation dataset can be used to validate the forthcoming products assuming that most of the reference samples will not change.

The current GlobCover assessment used on-line dataset in a structural way for validation. This was probably the first time these new data sources were applied at global scale for this purpose. There is a huge potential of these dataset for this kind of purposes and we have merely started to tap this potential.

6. REFERENCES


