Copernicus Global Land Operations "Vegetation and Energy"

"CGLOPS-1"

Framework Service Contract N° 199494 (JRC)

PRODUCT USER MANUAL

MODERATE DYNAMIC LAND COVER CHANGE

COLLECTION 100M

AFRICA

VERSION 2.1

(RELEASE CANDIDATE 1)

Issue 3.10

Organization name of lead contractor for this deliverable: VITO

Book Captain: Mr. Bruno Smets (VITO)

Contributing Authors: Dr. Marcel Buchhorn (VITO)

Luc Bertels (VITO)

Myroslava Lesiv (IIASA)

Nandin-Erdene Tsendbazar (WUR)

Linlin Li (WUR)

Dainius Masiliunas (WUR)

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Document Release Sheet

Book captain:	Bruno Smets	Sign	Date 10/11/2019
Approval:	Roselyne Lacaze	Sign	Date 27.11.2019
Endorsement:	Michael Cherlet	Sign	Date
Distribution:	Public		

Change Record

Issue/Rev	Date	Page(s)	Description of Change	Release
Draft	24.06.2017	All	Initial Version	D1.00
Draft	25.07.2017	All	Finished all paragraphs	D1.10
	10.10.2017	All	Revision after internal review	I1.20
Draft	29.04.2019	All	Revision for algorithm version 2	D2.00
	10.05.2019	All	Revision after internal review	I2.10
Draft	12.09.2019	All	Major revision for demonstration Africa change maps, derived from global reference map	D3.00
	10.11.2019	All	Revision after internal review	I3.10

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

Acronym	Meaning
AD	Applicable Documents
ARD	Analysis Ready Data
ATBD	Algorithm Technical Basis document
BFAST	Break for Additive Season and Trend
CCI	Change Consistency Indicator
CF V1.6	Climate & Forecast conventions compliant with version 1.6
CGLS	Copernicus Global Land service
COG	Cloud-Optimized GeoTIFF
CRS	Coordinate Reference System
DDI	Data Density Indicator
DOI	Digital Object Identifier
EO	Earth Observation
EPSG	European Petroleum Survey Group
EU	European Union
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization of the united nation
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
GeoTIFF	Geospatial Tagged Image File Format
GSD	Ground Sampling Distance
IIASA	International Institute for Applied Systems Analysis
JRC	Joint Research Center
LAI	Leaf Area Index
LC	Land Cover
LC100	Land Cover map at 100 m resolution
LCCS	Land Cover Classification System
MAE	Mean Absolute Error
MODIS	Moderate Resolution Imaging Spectroradiometer
NDMI	Normalized Difference Moister Index
NDVI	Normalized Difference Vegetation Index
netCDF	Network Common Data Form
NIR	Near Infra-Red reflectance
NIRv	Near-Infrared reflectance of vegetation
OECD	Organisation for Economic Cooperation and Development
PROBA-V	Vegetation instrument on board of PROBA satellite
PSD	Product Specifications Document
PUM	Product User Manual
RF	Random Forest classifier
RGB	Red Green Blue

RMSE	Root Mean Square Error
SDG	Sustainable Development Goal
SEEA	System of Environmental and Economic Accounting
SIPI	Structure Intensive Pigment Index
SM	Status Mask
SSD	Service Specifications Document
SVP	Service Validation Plan
SWIR	Short Wave Infra-Red reflectance
UN	United Nations
UNCCD	United Nations Convention to Combat Desertification
UNFCCC	United Nations Framework Convention on Climate Change
URD	Users Requirements Document
UTM	Universal Transverse Mercator
VI’s	Vegetation Indices
VITO	Vlaamse Instelling voor Technologisch Onderzoek (Flemish Institute for Technological Research), Belgium
VR	Validation Report
WGS84	World Geodetic System 1984
WUR	Wageningen University and Research

EXECUTIVE SUMMARY

The Copernicus Global Land Service (CGLS) is earmarked as a component of the Land service to operate "a multi-purpose service component" that provides a series of bio-geophysical products on the status and evolution of land surface at global scale. Production and delivery of the parameters take place in a timely manner and are complemented by the constitution of long-term time series.

Since January 2013, the Copernicus Global Land Service is continuously providing Essential Variables like the Leaf Area Index (LAI), the Fraction of Absorbed Photosynthetically Active Radiation absorbed by the vegetation (FAPAR), the surface albedo, the Land Surface Temperature, the soil moisture, the burnt areas, the areas of water bodies, and additional vegetation indices, which are generated every hour, every day or every 10 days on a reliable and automatic basis from Earth Observation satellite data.

The Dynamic Land Cover map at 100 m resolution (CGLS-LC100) delivers a yearly global land cover map at 100 m spatial resolution. Land cover plays a major role in the climate and biogeochemistry of the Earth system. The CGLS Land Cover product provides a primary land cover scheme at three levels, 12 classes at level 1 up to 23 classes at level 3, classes according to the Land Cover Classification System (LCCS). Next to these discrete classes, the product also provides proportional estimates for vegetation/ground cover for all basic land cover classes. This continuous classification scheme may depict areas of heterogeneous land cover better than the standard classification scheme and, as such, can be tailored for application use (e.g. forest monitoring, crop monitoring, biodiversity and conservation, monitoring environment and security in Africa, climate modelling, etc.)

The first Land Cover map (V1.0), reaching an overall accuracy of 74% at Level1, was provided for the 2015 reference year over the African continent, derived from the PROBA-V 100 m time-series, a database of high quality land cover training sites and several ancillary datasets.

This second Land Cover map (V2.0), reaching an overall accuracy of 80 % at Level1, is provided for the 2015 reference year over the entire Globe, derived from the PROBA-V 100 m time-series, a database of high quality land cover training sites and several ancillary datasets.

This second Land Cover map is extended (V2.1) to detect yearly land cover changes, demonstrated over the Africa continent covering the period 2015 to 2018.

A third Land Cover map collection (V3.0) is in production to provide new improved yearly maps from 2015-2019 over the entire Globe and further improves the change detection. From 2020, the yearly updates will be continued through the use of Sentinel missions data time-series.

1 BACKGROUND OF THE DOCUMENT

1.1 SCOPE AND OBJECTIVES

This Product User Manual (PUM) is the primary document that users have to read before handling the yearly land cover change products over Africa, covering the years 2015, 2016, 2017 and 2018.

It gives an overview of the product characteristics, in terms of algorithm, technical characteristics, and main validation results.

1.2 CONTENT OF THE DOCUMENT

This document is structured as follows:

- Chapter 2 summarizes the retrieval methodology,
- Chapter 3 describes the technical properties of the product,
- Chapter 4 summarizes the results of the quality assessment,
- Chapter 5 lists all references to cited literature

The users' requirements are recalled in the Annex.

1.3 RELATED DOCUMENTS

1.3.1 Applicable documents

AD1: Annex I – Technical Specifications JRC/IPR/2015/H.5/0026/OC to Contract Notice 2015/S 151-277962 of 7th August 2015

AD2: Appendix 1 – Copernicus Global land Component Product and Service Detailed Technical requirements to Technical Annex to Contract Notice 2015/S 151-277962 of 7th August 2015

AD3: GIO Copernicus Global Land – Technical User Group – Service Specification and Product Requirements Proposal – SPB-GIO-3017-TUG-SS-004 – Issue I1.0 – 26 May 2015.

1.3.2 Input

Document ID	Descriptor
CGLOPS1_SSD	Service Specifications of the Global Component of the Copernicus Land Service.
CGLOPS1_SVP	Service Validation Plan of the Global Component of the Copernicus Land Service
CGLOPS1_URD_LC100m	User Requirements Document of the moderate

	dynamic land cover product
CGLOPS1_PSD_LC100m	Product Specifications Documents of the moderate dynamic land cover product
CGLOPS1_PUM_LC100_V2.0	Product User Manual document of the 100m moderate dynamic land cover product for Version 2.0.
CGLOPS1_ATBD_LC100_V2.0	Algorithm Theoretical Basis Document of the 100m moderate dynamic land cover product for Version 2.0
CGLOPS1_VR_LC100m-V2.0	Validation report for global Moderate Dynamic Land Cover V2.0 product
CGLOPS1_VR_LCC100_V2.1	Report describing the results of the scientific quality assessment of the change demonstration maps over Africa of the 100 m moderate dynamic land cover product for Version 2.1

1.3.3 External documents

PROBA-V	http://proba-v.vgt.vito.be/
PROBA-V User Manual	User Guide of the PROBA-V data, available on http://www.vito-eodata.be/PDF/image/PROBAV-Products_User_Manual.pdf

2 ALGORITHM

This chapter presents the concept of the land cover change detection applied to create the yearly Dynamic Land Cover change map v2.1 product at 100m resolution (section 2.1). The land cover change detection algorithm is based upon the land cover mapping approach used to generate the dynamic land cover map v2.0 over the globe, which is described in the ATBD [CGLOPS1_ATBD_LC100_V2.0]. The updates of the v2.1 as well as the specificities of change detection method are detailed in Section 0.

2.1 OVERVIEW

The CGLS Dynamic Land Cover change detection approach relies on several proven individual methodologies:

1. Data pre-processing including atmospheric & geometric correction, data cleaning and temporal outlier detection techniques,
2. Applying data fusion techniques at multiple levels,
3. Supervised classification, and
4. established third party datasets via expert rules.

The workflow is built on top of the regular CGLS Dynamic Land Cover change (2015 base reference year), shown in Figure 1.

1. From satellite data to metrics [CGLOPS1_PUM_LC100_V2.0]. The metrics are calculated per epoch (3-years period):
   a. PROBA-V UTM Analysis Ready Data (ARD) generation,
   b. data cleaning & compositing,
   c. data fusion as well as quality indicator generation,
   d. metrics extraction,
2. training data generation,
3. ancillary datasets preparation,
4. machine learning (classification / regression),
5. break detection
6. land cover change map and cover fraction layers generation plus final quality layer assembling.

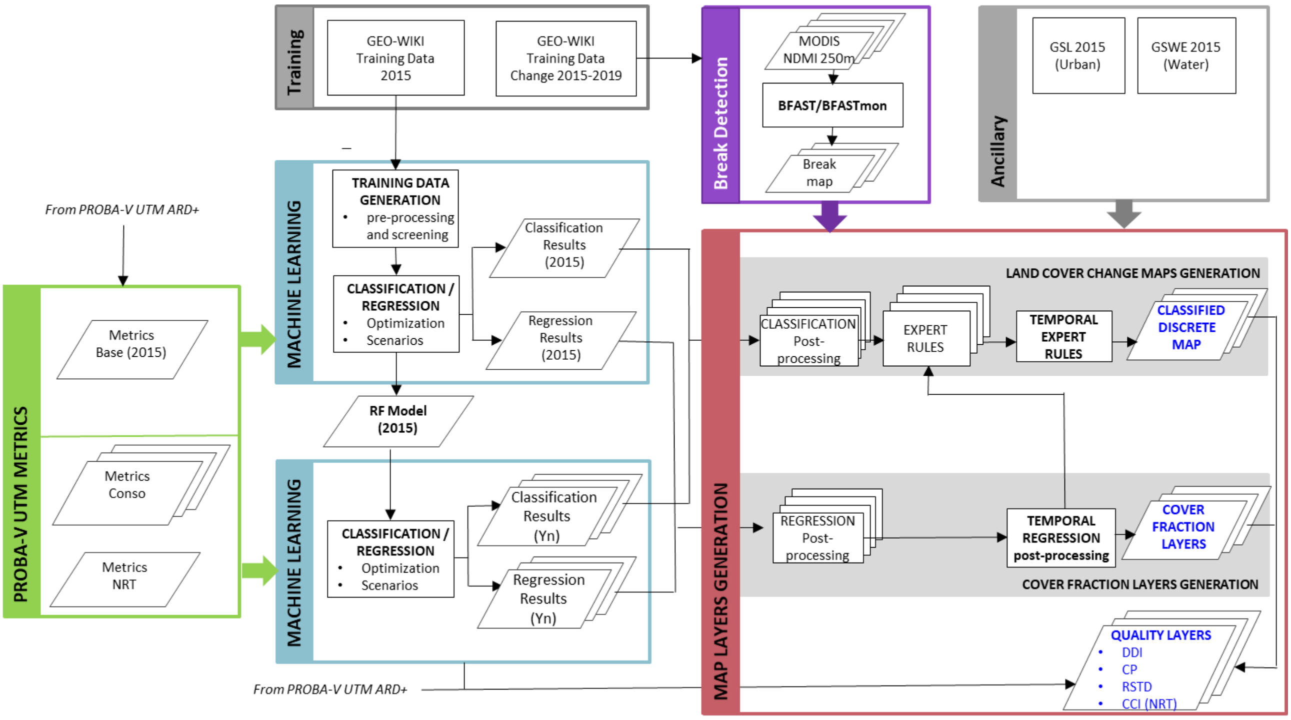

Figure 1: Workflow diagram for the CGLS Dynamic Land Cover 100 m (CGLS-LC100) change products (in blue layers included in the final product)

To generate the Metrics, several pre-processing steps have to be performed:

- To reduce distortion in the High North, to make our land cover products better usable with other data and to allow continuity of the service, the PROBA-V archive, used as current main input data source, was reprocessed with a new geometric correction and an improved atmospheric correction. The complete PROBA-V archive was, in this way, translated into the PROBA-V UTM ARD which is produced in Universal Transverse Mercator (UTM) coordinate system and fully aligned with the Sentinel-2 tiling grid in tile naming as well as tile dimensions.
- The PROBA-V UTM ARD includes the 5-daily PROBA-V multi-spectral image data with a Ground Sampling Distance (GSD) of ~0.001 degree (~100 m) which is used as primary earth observation (EO) data, and the PROBA-V UTM daily multi-spectral image data with a GSD of ~0.003 degree (~300 m) which is used secondarily. Next to a Status Mask (SM) cleaning using the internal quality flags of the PROBA-V EO data, a temporal cloud and outlier filter built on a Fourier transformation is applied. From this cleaned and outlier screened dataset, a Data Density Indicator (DDI) is calculated which is used as an input quality indicator in the supervised learning process.
- Next, the 5-daily PROBA-V 100 m and daily PROBA-V 300 m datasets are fused using a Kalman filtering approach. The Kalman-filled 100 m data set is then automatically checked for consistency before extracting several metrics. Therefore, a harmonic model is fitted through each of the additional derived vegetation indices for each time series step.

- Next to the parameters of the harmonic model, which are used as metrics for the overall level and seasonality of the time series, descriptive statistics and textural metrics are generated. Overall, 154 metrics are extracted from the PROBA-V EO data.

The training data is collected through manual classification using Google Maps and Bing images at 10 m spatial resolution using the Geo-Wiki Engagement Platform (http://www.geo-wiki.org/). Therefore, the training data not only includes the land cover type, but also the cover fractions of the main land cover classes in PROBA-V UTM 100 m resolution. In the classification preparation, the metrics of the training points are analysed for intra- and inter- specific outliers, as well as screened for the best metrics combinations. The optimized training data together with the input data quality indicators (DDI dataset) are then used in a supervised classification/regression using Random Forest (RF) machine learning techniques, initially performed on the base 2015 layer. The 2015 RF model is stored and re-used in the subsequent supervised learning for the change maps (2016 – 2018). The input Metrics for this supervised learning are either calculated in 'Consolidation' (Conso) mode meaning using the date of the year before and after, or in 'Near Real Time' (NRT) mode meaning using a three years period ending 3 months after the reference year.

To support the change detection algorithm and assess whether a given pixel is temporally stable or not, an algorithm (BFAST - Verbesselt et al., 2010) is used to detect all breaks in a given time series and provide an estimation of the breakpoint timing together with an uncertainty measure about the time of the breaks. To have a sufficient length of the time series to detect breaks, a MODIS 250m derived Normative Difference Moisture Index (NDMI) from 2009-2018 is used. The algorithm is optimized using change training data collected through the same Geo-Wiki Engagement platform. The break maps are resampled/warped to PROBA-V UTM 100m spatial resolution and included in the land cover map generation step.

Finally, during the map generation step, the temporal dimension of the single land cover results (classes and fractions) are used together with the break maps and ancillary datasets to generate the final yearly dynamic land cover change maps. The produced maps, both discrete and continuous cover fraction, use a hierarchical legend based on the United Nations (UN) Land Cover Classification System (LCCS).

2.2 MODES

The yearly land cover maps are generated using an 'epoch', where an epoch consists of 3 years of input data (one year before and one year after). The classification/regressor model is generated on the **Base** reference year, in this case 2015. The machine learning model is re-used in the subsequent generation of the change maps. As shown in Figure 2, the actual land cover can change over time and, depending on when the change happens, it can be detected. The yearly land cover map updates are delivered at the end of the 1st quarter of the year-end, and hence changes during the latter part of the year can be detected but the new class cannot yet be identified (depicted in red crosses). This mode is called **Near-Real-Time**.

Therefore the yearly map is re-generated the year after, known as **Consolidated**, and enough data is available to identify all new classes including the land covers which changed at late stage.

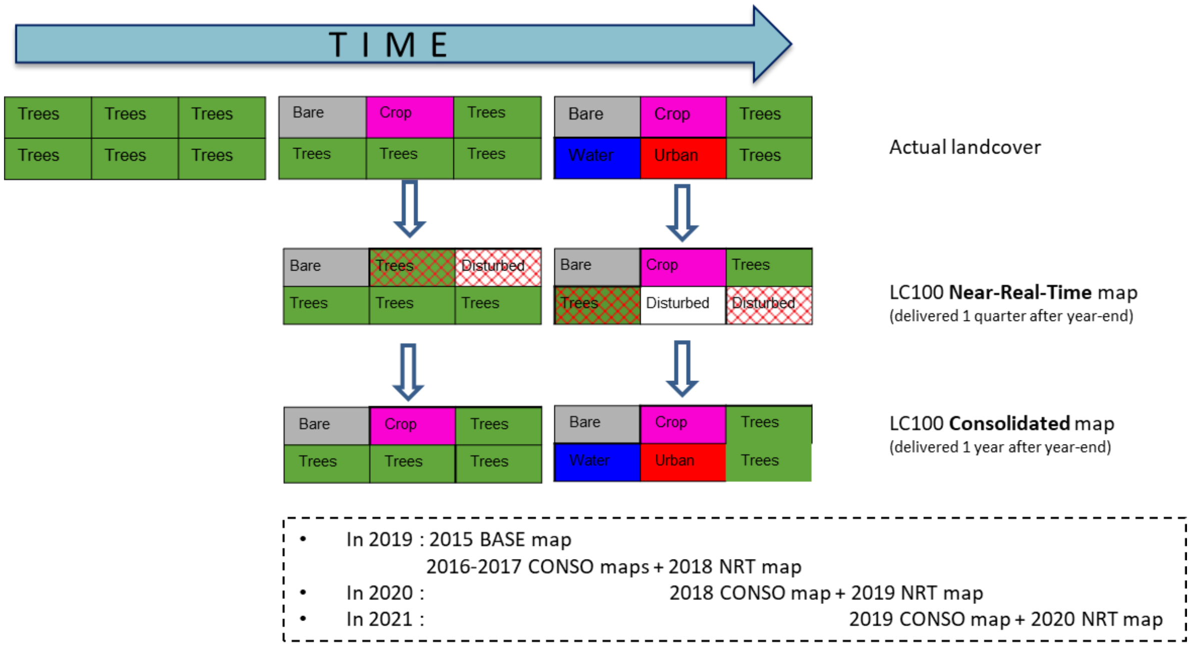

Figure 2 : Processing modes of land cover change maps

The result of this 'mode' processing implies that during the beginning of a new year (Yr), two maps are provided:

1. Near-Real-Time (NRT) map of previous year (Yr-1)
2. Consolidated (CONSO) map of the year before (Yr-2)

The latter map consolidates the detected changes of the NRT map and hence substitutes the former map.

2.3 THE RETRIEVAL METHODOLOGY

The details of the land cover v2.0 algorithm can be found in the ATBD [CGLOPS1_ATBD_LC100_V2.0]. The elements, updated in v2.1, are presented hereafter.

2.3.1 Metrics

The number of descriptive statistics of the PROBA-V time series is limited in v2.1 compared to v2.0. They are the mean, standard deviation, sum, median, 10^{th} percentile, 90^{th} percentile, 10^{th} – 90^{th} percentile range. These overall 7 descriptive metrics are extracted for each of the vegetation indices time series sets. An additional descriptive metric is calculated using a 3x3 moving window calculating the standard deviation of the box for all calculated median statistics (10 VI's). This metric can be interpreted as textural metric representing the uniformity of the pixel in its box (low values show a homogeneous area, high values a more heterogeneous land cover).

Overall, 154 metrics (8 descriptive metrics and 7 harmonic metrics for the 10 VI's time series sets plus 4 topographic metrics) are generated from the PROBA-V UTM ARD++ archive at 100 m and are input in the classification/regression step of the automated processing chain.

2.3.2 Training Data

The training data was collected with reference imagery of the year 2015 which matches our base reference map of 2015. On top of this, the Geo-Wiki engagement platform has been extended to cope more easily with temporary imagery, i.e. integrating the Sentinel-2 time-series. This new feature was used to start collecting reference data for training the change map generation. At this stage the amount of points collected has been limited to about 200 points for 2016, as there was a clear need to have a good change stratification map to optimize the effort of collecting change reference information. The reference data is being extended with more points covering the period 2016-2019, initially over the African continent and thereafter scaling up to the globe. This change reference training data will be used in the production of the yearly global change maps.

For the machine learning, the approximately 36,000 sample sites (Figure 3) located on the African continent, were re-used from the 2015 global training dataset.

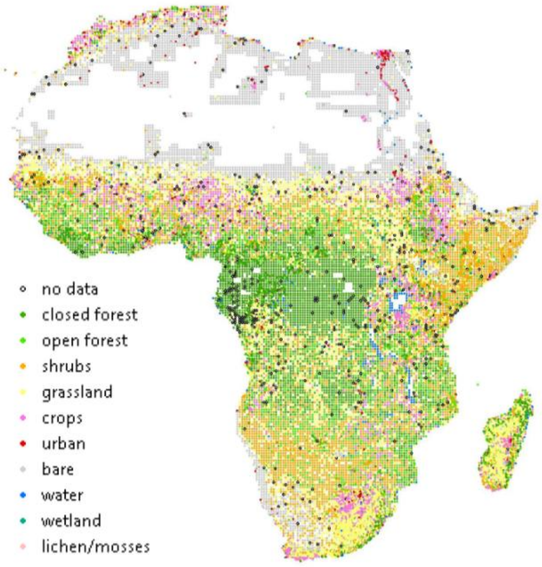

Figure 3: Training dataset used in the CGLS-LC100 workflow showing the ~36K training locations over African continent.

2.3.3 Break dataset preparation

For producing yearly map updates, a change detection algorithm is needed to assess whether a given pixel is temporally stable or not. If it is stable, then it is desirable to keep the previous land cover classification, in order to avoid spurious changes that come from minor changes in the input data for the classifier.

There are several algorithms available for land cover change detection. Based on an evaluation, the BFAST (Verbesselt et al., 2010) and BFAST Monitor algorithm have been selected. For consolidated map production, we chose BFAST, since the algorithm is able to detect all breaks in a given time series, is more precise with the estimation of breakpoint timing, and gives an uncertainty measure about the time of the breaks. This allows potential optimisation of the classifier, since knowing the time of the breaks allows avoiding training the classifier with metrics produced from unstable periods of change. In addition, knowing all the breaks in the time series, rather than just the most significant one, allows for more efficient consolidation, as newly incoming data is reprocessed and thus contributes to a more precise estimation of breaks in all years simultaneously. This allows for updating (consolidating) previously-released land cover maps as part of the map updating process for a new year.

In contrast, BFAST Monitor was designed for detecting change at the end of a time series, and thus reports only whether there was a break or not and the timing of the first break since the start of the monitoring period. This break is not necessarily the largest break, and only a single break is reported. Similarly, a t-test run over the history and compared with a monitoring period is a simple way to test whether there has been a break in the time series, however, it only indicates whether the mean of the monitoring period is different from the mean of the history period, with no indication about the time of the break.

The BFAST algorithm works by first decomposing the input (cloud-free) time series into a seasonal (harmonic: sine and cosine functions, with a defined frequency order of a year or half-year), a trend and a remainder component. Next, the time series is segmented into segments via piecewise linear regression, with a minimal segment size that is controlled by the parameter h. The number of breaks in the time series is optimised by iteratively minimising the Bayesian Information Criterion (BIC). This procedure is only run in case a structural change test indicates with $p < 0.05$ that there is, at least, one break in the time series. The algorithm has several configurable parameters: vegetation index that it is run on, minimal segment size h, components on which the breaks are being detected, and the harmonic order of the seasonal component(s).

We set the parameter h to equal the maximum number of timesteps within a single year of the time series. This means that at most one break can be detected in any given year. Also, this implies that the first break that can be detected is at least one year away from the start or end of the time series. As such, the algorithm is not suited for detecting changes at the end of the time series, i.e. for NRT updating; in that case, BFAST Monitor is used instead.

We detected breaks on both seasonal and trend components. The harmonic order of the seasonal component was set to 3, i.e. the mean, annual and semi-annual frequencies were included in the model. This was determined by testing the algorithm output over Sahel when using order 3 vs order 2 harmonics and comparing the change detection accuracy.

The algorithm was run on three vegetation indices (NDMI, EVI and NIRv) time series derived from MODIS 250 m Terra+Aqua reflectance product from 2009 until 2018. The NDMI has been found to provide the best results for the African continent and hence was used during the production of the land cover change maps.

The algorithm parameters were tuned using limited change training reference data, as described in above. Hence, it is expected to further optimize these parameters in the next version of the Copernicus Global Land Cover maps (version 3).

Some examples of detected processes can be seen in the figures below: change in crop types (see Figure 4), desertification (see Figure 5), wetland dry-up (see Figure 6) and deforestation (see Figure 7).

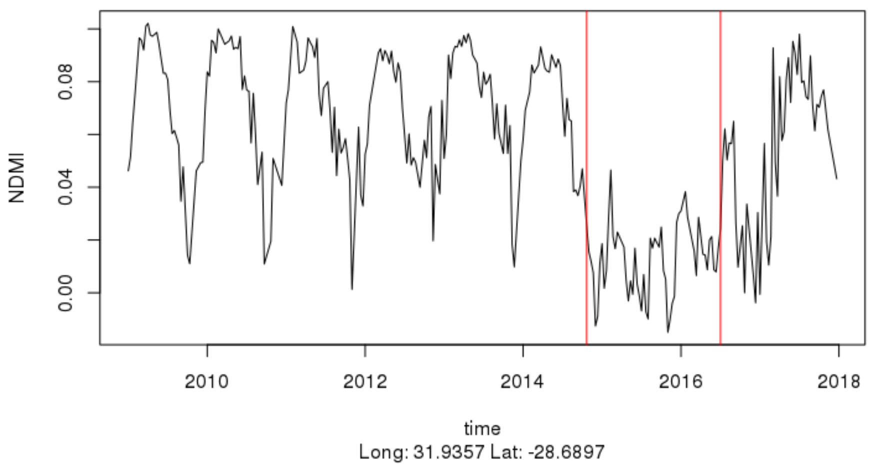

Figure 4 : Re-cultivation change process

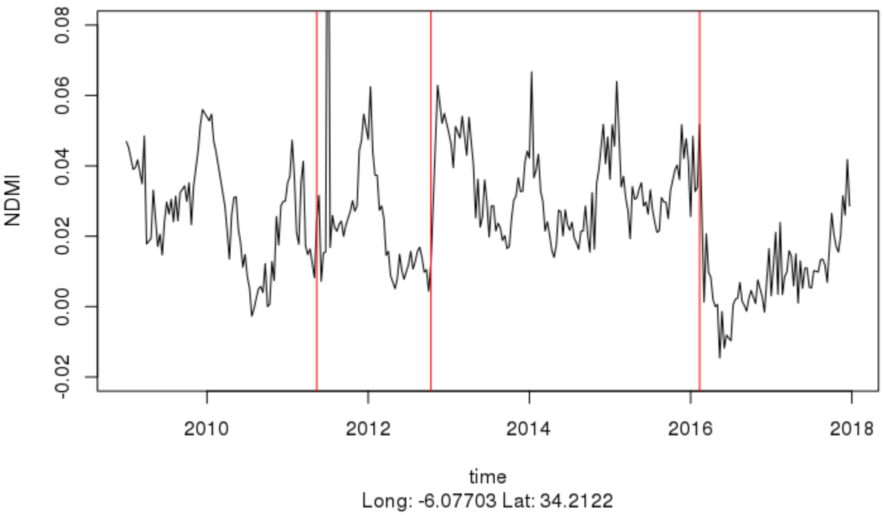

Figure 5 : Desertification change process

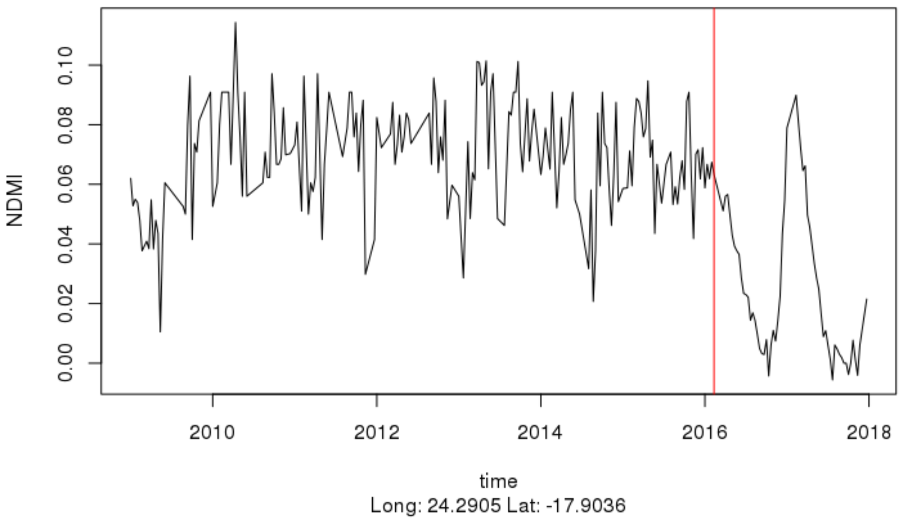

Figure 6 : Wetland dry-up change process

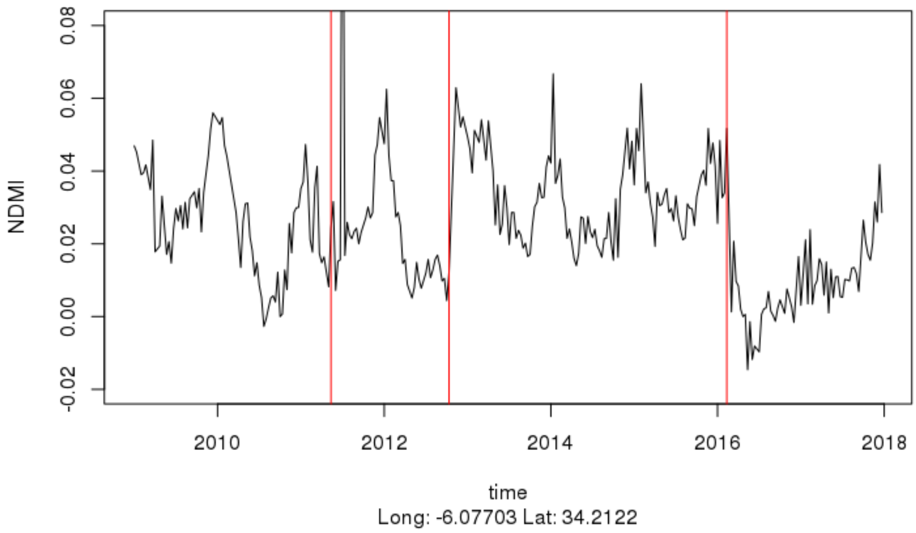

Figure 7 : Deforestation change process

Once the breaks are determined, the break maps are resampled/rewarped to 100m UTM projection and are then used in further processing, in combination with the classifier output and expert rules based on extra stability metrics to determine which pixels are likely to have changed.

2.3.4 Classification and Regression

During the base run (reference year 2015), the entire RF model (all trees in the forest) is saved, as shown in Figure 1. This model is thereafter used during each classification/regression subsequent yearly run (reference years 2016 to 2018).

2.3.4.1 Cleaning regression results

To cope with changes, a linear regression is performed across the yearly maps (2015-2018) using a set of rules on the stability of the pixel. This step is known as Temporal Regression Post-processing (TRP). First the median over the years per cover fraction is calculated. Each pixel is then labelled: (i) stable if the temporal values are within 6% range, (ii) reliable if within 16% range and (iii) unstable if temporal values are outside 16% range. The TRP process is limited to the six fractions resulted directly from the regressor (shrubland, forest, moss&lichen, bare/sparse vegetation, cropland, herbaceous vegetation).

Fractions are not altered if:

- SUM of the six cover fractions is less than 10%, since the pixel will mainly be defined by external layers (built-up, water, snow&ice);
- Cover labelled as 'stable'.

Fractions will be altered, through applying the linear regression function (slope and aspect) over the years if:

- Covers labelled as 'reliable', and the standard error of the cover is higher than the standard error of the mean of all covers.
- Covers labelled as 'unstable', if they comply to one of the following rules:
  - outlier rule (one single non-zero if not last year) : remove single non-zero value;
  - outlier inverse rule (one single zero if not last year) : remove single zero value;
  - shrub - grass confusion rule (only shrub and herbaceous vegetation layers are labelled unstable): alter all years of the cover fraction with highest standard error and slope < 5
  - jumpy tree rule (only tree cover is labelled unstable, and consecutive values above & below mean) : alter all years of the cover fraction

After the Temporal Regression Post-processing step (TPR), the regular regression post-processing per single year is performed using the same normalization approach as described earlier. These results are then used to re-generate the final cover fractions layers.

2.3.4.2 Cleaning classification results

Based on the cleaned post-processed regression results, the standard Expert Rules are applied to combine the existing knowledge represented by the ancillary datasets and the classification and regression results for each single map. Thereafter, a temporal cleaning (Temporal Expert Rules or TER) is performed across the time-series to generate the yearly land cover change maps.

In order to incorporate the vegetation cover fraction layers, a discrete map per year was generated by applying the training data rules on the 9 'temporary cleaned' cover fraction layers naming forest, shrubland, herbaceous vegetation, moss & lichen, bare/sparse vegetation, cropland, built-up, snow, and permanent water bodies.

In a second step, these yearly discrete maps are combined in a temporal stack and a set of Temporal Expert Rules (TER) are applied to make the discrete maps consistent over time. The rules make use of special classification, performed over the entire period (2014-01-01 to 2019-03-31), using the standard workflow. The result of this classification is known as 'base classes'.

The following rules are applied in TER :

- Ancillary rule (not applied in NRT-mode) :
  - pixels classified as snow, urban, permanent water or wetland are kept consistent over time
- Cover fraction stability rules (not applied in NRT-mode) : the following rules are applied to clean unreliable changes and/or inconsistencies with cover fractions:
  - Base class if propagated to yearly updates if the six cover fractions of the TRP are stable;
  - Base class is propagated to yearly updates if no break is detected;
  - Class of previous year (Y-n) is propagated till a break is detected;
- Jumpy case rules (not applied in NRT-mode), classes switching from one type to another type from year to year due to classifier noise : the long term series class is used to clean unreliable jumpy patterns if:
  - jumpy open <-> closed forest classes,
  - jumpy herbaceous vegetation <-> shrubland classes
  - jumpy shrubland <-> open forest classes
- Near Real Time rules (only applied in NRT-mode) : overwrite the Y_n class with class of Y_{n-1} if:
  - no satellite data available
  - no break is detected in Y_n
  - break is detected in second half of Y_n

2.3.4.3 Adding metadata and quality indicators

In a final step, metadata attributes compliant with version 1.6 of the Climate & Forecast conventions (CF V1.6) and the colour tables translating the discrete class code into the legend are injected. Moreover, the probability layer indicating the classifier certainty was produced out of the predicted class probabilities of the classification results. This classification quality layer was bundled together with the cover fraction quality layers and the Data Density Indicator layer, as quality indicator for the input data, to create the overall product quality layers.

Figure 8 shows an overview of the discrete map with 23 classes over the African continent for 2015 (base mode) and 2018 (NRT mode) while Figure 9 shows the legend in more detail.

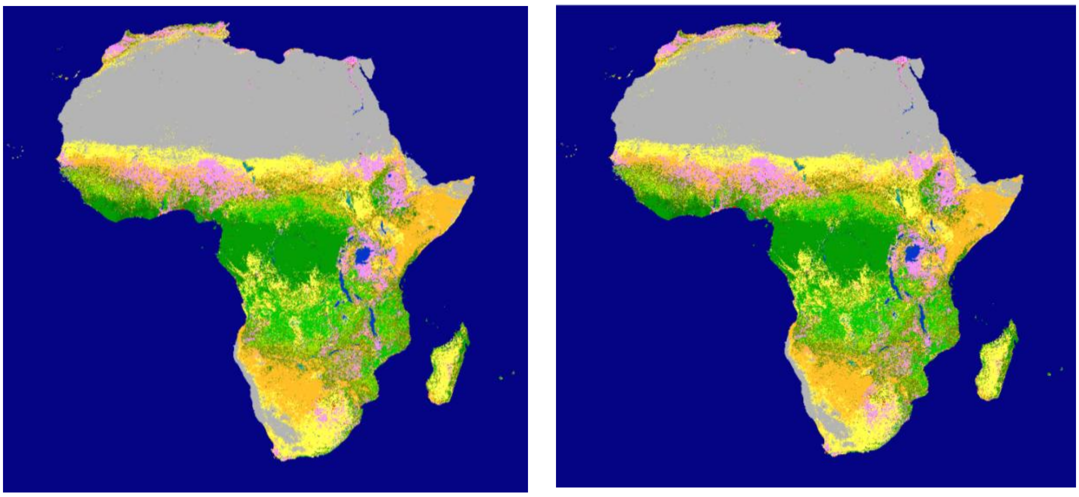

Figure 8: The CGLS Dynamic Land Cover Map at 100 m for epoch 2015 base (left) and epoch 2018 NRT (right) with 23 discrete classes (detailed legend in Figure 9)

Figure 9: Legend for the 23 discrete classes of the CGLS Dynamic Land Cover Map at 100 m

Figure 10 shows the ten cover fraction layers in a collage on global scale.

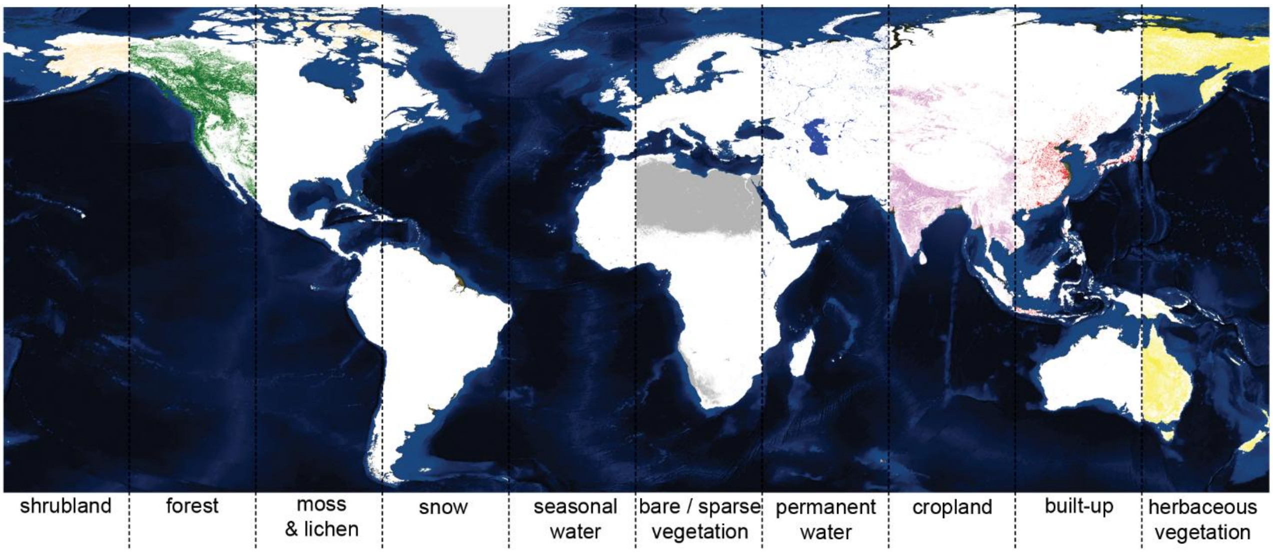

Figure 10: The cover fraction layers for the 9 base land cover classes and the seasonal inland water cover fraction of the CGLS Dynamic Land Cover product at 100 m for epoch 2015 (shown as a collage on global scale).

2.4 LIMITATIONS OF THE METHOD

Since the methodology for generating the yearly change maps is based on the methodology of the global 2015 version 2 base map, all limitations of the global map apply. The user is referred to [CGLOPS1_ATBD_LC100_V2.0] to know the base limitations.

The yearly change maps are currently in demonstration stage (https://land.copernicus.eu/global/products/development-stages), hence resulting in the following limitations:

- Areas of changes that are smaller than the 250m Minimum Mapping Unit (MMU) are likely not detected and reported as no change, due to the MODIS NDMI used to detect breaks.
- Some limited areas of changes can be overestimated, as the break detection algorithm currently overestimates the number of breaks in the time series. This overestimation is currently partly solved through constrain the change by combining it with the classifier output.
- An additional Change Consistency Indicator (CCI) of the reliability of the change is currently only provided for the 2018 (NRT) map, based on a simple agreement between the break map and the classifier maps.
- Built-up class is not updated in these demonstration change maps, hence all yearly maps have the same built-up class than the base 2015 map.
- Permanent Water Bodies class is not updated in these demonstration change maps, hence all yearly maps have the same permanent water bodies class than the base 2015 map.

- The cover fractions over the temporal domain should be used with care, as some unexpected variations can still be present. A number of rules were applied to clean the cover fractions over time, however these temporal regression rules will be further improved in the Version 3.
- The new (version 2.1) base map (2015) over the African continent slightly deviates from the global (version 2.0) base map (2015) since the 2015 is regenerated through removing some metrics compared to V2.0 (see §2.3.1).

2.5 DIFFERENCES WITH THE PREVIOUS VERSION

Table 1: Successive versions of Collection 100m Land cover products. The V2.1 described in this PUM is highlighted in bold.

Version	Coverage	Status	Main characteristics
1.0	Africa 2015	Demonstration	PROBA-V S1 time-series (plate carree) Random Forest for classification
2.0	Global 2015	Operational	PROBA-V L1C time-series (gridded to Sentinel-2 UTM) Random Forest for classification
2.1	**Africa 2015-base, 2016-conso, 2017-conso, 2018-NRT**	**Demonstration**	**PROBA-V L1C time-series (gridded to Sentinel-2 UTM)** **Random Forest for classification** **BFAST for break detection** **linear regression**
3.0	Global 2015-2019		PROBA-V L1C time-series (gridded to Sentinel-2 UTM) Random Forest for classification Optimized BFAST for break detection

2.6 ROADMAP

A version 3.0 of Collection 100 m of LC products will be generated from the PROBA-V 100 m and 300 m sensor data and will cover the years 2015-2019 with global coverage.

The Copernicus Global Land service will continue the 100 m production, from 2020, through the combination of Sentinel-1 and Sentinel-2 mission. The adaptation of the retrieval methodology to the Sentinels missions data is currently under test.

3 PRODUCT DESCRIPTION

The Africa 100m Land Cover version 2.1 products are provided in **ZIP** files per year, per 20 x 20 degree tile (Figure 11), containing up to 21 **GeoTIFF** files corresponding to the following **layers**:

- One discrete land cover map
- Fractional cover map for 10 classes (bare/sparse vegetation, cropland, herbaceous grassland, moss&lichen, shrubland, tree, permanent snow&ice, built-up, permanent water, seasonal water);
- One Forest type map
- One discrete classification probability map (quality indicator of the classifier) (only for 2015 base map)
- Standard deviation maps for the fractional cover of 6 classes (bare/sparse vegetation, cropland, herbaceous grassland, moss&lichen, shrubland, tree) (quality indicator of the regressor) (only for 2015 base map)
- Data Density Indicator for PROBA-V UTM 100m input data
- Change Consistency Indicator (quality indicator) of the discrete land cover map (only for NRT map)

3.1 FILE NAMING

The ZIP files, and GeoTIFF files they contain, follow this naming convention:

<TILE>_<SENSOR>_LC100_epoch<YEAR>-<MODE>_Africa_<VERSION>_<LAYER>_<CRS>.tif

where

- <TILE> is the designation of the 20 x 20 degree tile, composed of the 3-digit longitude and 2-digit latitude of the top-left corner (see Figure 11)
  Example: W020N20 for the tile covering the area from 20W to 00W and 20N to 00N.
- <SENSOR>: the EO sensor used. Here is "ProbaV".
- "LC100" indicates this is a 100 m resolution Land Cover product
- epoch< YEAR> indicates the epoch year in four digits.
- <MODE> indicates the run mode, as is "base" for the 2015 reference map, "conso" for the 2016 and 2017 consolidation map and "nrt" for the 2018 Near Real Time map.
- "Africa" indicates that the tile is part of a Land Cover product that covers the African land surface.

- <VERSION> shows the processing line version used to generate this product. The version denoted as vM.m.r (e.g. v2.1.1), with 'M' representing the major version (e.g. v2), 'm' the minor version (starting from 0) and 'r' the production run number (starting from 1) (Table 2).
- <LAYER> gives the name of the data layer (see Table 3)
- <CRS> is the Coordinate Reference System used. The current tiles are provided in EPSG:4326, geographic latitude/longitude CRS.

Figure 11: Scheme of the 20 x 20 degree tiles for the African demonstration change maps

Table 2: Version numbering and recommendations for handling version updates

Versions	Differences	Recommendations
Major	Significant change to the algorithm.	Do not mix various major versions in the same applications, unless it is otherwise stated.
Minor	Minor changes in the algorithm	Can be mixed in the same applications, but require attention or modest modifications
Run	Fixes to bugs and minor issues, updates in input data.	Later run is a drop-in replacement of all former runs.

Table 3: Land Cover layer names in the filename

Layer in filename	Description
discrete-classification	Main discrete land cover class according to FAO LCCS scheme
discrete-classification-proba	Classification probability, a quality indicator for the discrete land

Layer in filename	Description
	cover map
forest-type-layer	Forest type for all pixels where tree cover fraction is bigger than 1 %
bare-coverfraction-layer	Fractional cover (%) for the bare and sparse vegetation class
crops-coverfraction-layer	Fractional cover (%) for the cropland class
grass-coverfraction-layer	Fractional cover (%) for the herbaceous vegetation class
moss-coverfraction-layer	Fractional cover (%) for the moss & lichen class
shrub-coverfraction-layer	Fractional cover (%) for the shrubland class
snow-coverfraction-layer	Fractional cover (%) for the snow & ice class
tree-coverfraction-layer	Fractional cover (%) for the forest class
urban-coverfraction-layer	Fractional cover (%) for the built-up class
water-permanent-coverfraction-layer	Fractional cover (%) for the permanent inland water bodies class
water-seasonal-coverfraction-layer	Fractional cover (%) for the seasonal inland water bodies class
bare-coverfraction-StdDev crops-coverfraction-StdDev grass-coverfraction-StdDev moss-coverfraction-StdDev shrub-coverfraction-StdDev tree-coverfraction-StdDev	Quality indicator (standard deviation) of the percentage vegetation cover regression, for the respective class.
DataDensityIndicator	Data density indicator showing quality of the EO input data between 0 – 100 (0 = bad, 100 = perfect data)
ChangeConsistencyIndicator	A quality indicator of the NRT discrete land cover change map

3.2 File Format

The 21 Land Cover layers are provided as single-band GeoTIFF files that are internally compressed with standard metadata attributes, and include overview pyramids on levels 2, 4, 8 and 16 for faster loading in GIS.

Note that this format may not be fully compliant with Cloud-Optimized GeoTIFF (COG) requirements. The GeoTIFF format will be further improved towards COG in upcoming Land Cover products.

3.3 Product Content

All land cover layers are stored as single bytes per pixel, without scaling or offset.

3.3.1 Discrete classification

The discrete classification map provides 23 classes (Table 4) and is defined using the Land Cover Classification System (LCCS) developed by the United Nations (UN) Food and Agriculture

Organization (FAO). The UN-LCCS system was designed as a hierarchical classification, which allows adjusting the thematic detail of the legend to the amount of information available:

- The “level 1” legend contains classes with codes that are multiples of ten (10, 20, 30, etc.).
- The “level 2”, also known as regional legend, has class codes of two digits that is not a multiple of ten (i.e. 11, 12 are sub-classes of 10, and so on).
- The “level 3” classes have three digits (i.e. 111 – 116 and 121 – 126) and are used to further distinguish the forest types (sub-classes of 11 – open forest and 12 – closed forest).

The discrete map is coded with special values 200 for sea pixels and 0 signifying missing input data (i.e. not observed by PROBA-V sensor).

Table 4: Discrete classification coding

Map code	UN LCCS level	Land Cover Class	Definition according UN LCCS	Color code (RGB)
0	-	No input data available	-	40, 40, 40
111	A12A3A10B2D2E1	Closed forest, evergreen needle leaf	tree canopy >70 %, almost all needle leaf trees remain green all year. Canopy is never without green foliage.	88, 72, 31
113	A12A3A10B2D2E2	Closed forest, deciduous needle leaf	tree canopy >70 %, consists of seasonal needle leaf tree communities with an annual cycle of leaf-on and leaf-off periods	112, 102, 62
112	A12A3A10B2D1E1	Closed forest, evergreen, broad leaf	tree canopy >70 %, almost all broadleaf trees remain green year round. Canopy is never without green foliage.	0, 153, 0
114	A12A3A10B2D1E2	Closed forest, deciduous broad leaf	tree canopy >70 %, consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.	0, 204, 0
115	A12A3A10	Closed forest, mixed	Closed forest, mix of types	78, 117, 31
116	A12A3A10	Closed forest, unknown	Closed forest, not matching any of the other definitions	0, 120, 0
121	A12A3A11B2D2E1	Open forest, evergreen needle leaf	top layer- trees 15-70 % and second layer- mixed of shrubs and grassland, almost all needle leaf trees remain green all year. Canopy is never without green foliage.	102, 96, 0
123	A12A3A11B2D2E2	Open forest, deciduous needle leaf	top layer- trees 15-70 % and second layer- mixed of shrubs and grassland, consists of seasonal needle leaf tree communities with an annual cycle of leaf-on and leaf-off periods	141, 116, 0
122	A12A3A11B2D1E1	Open forest, evergreen broad leaf	top layer- trees 15-70 % and second layer- mixed of shrubs and grassland, almost all broadleaf trees remain green year round. Canopy is never without green foliage.	141, 180, 0
124	A12A3A11B2D1E2	Open forest, deciduous broad leaf	top layer- trees 15-70 % and second layer- mixed of shrubs and grassland, consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.	160, 220, 0
125	A12A3A12	Open forest, mixed	Open forest, mix of types	146, 153, 0
126	A12A3A12	Open forest, unknown	Open forest, not matching any of the other	100, 140, 0

Map code	UN LCCS level	Land Cover Class	Definition according UN LCCS	Color code (RGB)
			definitions	
20	A12A4A20B3(B9)	Shrubs	These are woody perennial plants with persistent and woody stems and without any defined main stem being less than 5 m tall. The shrub foliage can be either evergreen or deciduous.	255, 187, 34
30	A12A2(A6)A20B4	Herbaceous vegetation	Plants without persistent stem or shoots above ground and lacking definite firm structure. Tree and shrub cover is less than 10 %.	255, 255, 76
90	A24A2A20	Herbaceous wetland	Lands with a permanent mixture of water and herbaceous or woody vegetation. The vegetation can be present in either salt, brackish, or fresh water.	0, 150, 160
100	A12A7	Moss and lichen	Moss and lichen	250, 230, 160
60	B16A1(A2)	Bare / sparse vegetation	Lands with exposed soil, sand, or rocks and never has more than 10 % vegetated cover during any time of the year	180, 180, 180
40	A11A3	Cultivated and managed vegetation/agriculture (cropland)	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.	240, 150, 255
50	B15A1	Urban / built up	Land covered by buildings and other man-made structures	250, 0, 0
70	B28A2(A3)	Snow and Ice	Lands under snow or ice cover throughout the year.	240, 240, 240
80	B28A1B1	Permanent water bodies	lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies.	0, 50, 200
200	B28A1B1[1]	Open sea	Oceans, seas. Can be either fresh or salt-water bodies.	0, 0, 128

3.3.2 Fractional Cover layers

The Fractional Cover layers, also referred to as cover fractions, give the percentage of a 100 m pixel that is filled with a specific land cover class (Table 3). As such it provides more detailed information than the dominant class that is shown in the discrete classification.

The Fractional Cover layers are coded as a number between 0 and 100, in steps of 1 %. The sum of all fractional cover layers for a given pixel is 100. Missing values are set to 255.

3.3.3 Forest type layer

The Forest Type layer provides discrete values per type of forest (see Table 5), for all pixels where the tree (forest) cover fraction exceeds 1 %. Value 255 is used for missing values.

[1] Note: a distinction is made between Open sea (oceans) = 200 and other permanent water bodies = 80, despite they're mapped to the same UN LCCS layer legend.

Table 5: Forest type coding

Value	Short name	Description
0	Unknown	Doesn't match any of the other types
1	ENF	Evergreen needle leaf forest
2	EBF	Evergreen broad leaf forest
3	DNF	Deciduous needle leaf
4	DBF	Deciduous broad leaf
5	Mixed	Mix of forest types

This layer is provided for all maps.

3.3.4 Quality layers

3.3.4.1 Probability of the discrete classification

The probability of the discrete classification is provided as a number between 0 and 100, in steps of 1 %. Closer to 100 is the probability, higher is the quality of the discrete classification. Value 255 is used for missing values.

This layer is only provided with the base map (2015).

3.3.4.2 Standard Deviation for the Fractional Cover

The Standard Deviation of the percentage cover regression is provided as a number between 0 and 100, in steps of 1 %. Closer to 100 is the standard deviation, higher is the quality of the associated fractional cover layer. Value 255 is used for missing values.

These layers are only provided with the base map (2015).

3.3.4.3 Data Density Indicator (DDI)

The Data Density Indicator indicates the availability of input data from the 5-daily PROBA-V UTM ARD+ composites for 100 m and 300 m resolutions. It is a score between 0 = no input data available and 100 = best data availability. Missing DDI values are coded as 255.

This layer is provided for all maps.

3.3.4.4 Change Consistency Indicator (CCI)

The Change Consistency Indicator indicates the reliability of the change in the discrete class (and hence implicitly also in the fractional covers). It is provided as a number between 0 and 2013, as indicated in Table 6. The value of 0 corresponds to no value.

Table 6: Change Consistency Indicator coding

Value	Short description	Description
0	Unknown	No information from classifier and/or break map.
210	Potential change, low confidence break	Discrete class did no change from previous year. Classifier results detected no change, but break was detected
211	Potential change, low confidence classifier	Discrete class did not change from previous year. Classifier results detected change, but no break was detected
212	**Change**, high confidence	Discrete class changed from previous year. Classifier detected change and break was detected in 1st half of reference year
213	Potential change, high confidence	Discrete class changed from previous year, but new class is not yet known. Classifier detected change, but break was detected in 2nd half of reference year
255	No change, high confidence	Discrete class did not change from previous year. Classifier results were consistent and no break was detected

This layer is only provided with the NRT map (2018).

3.3.5 Metadata attributes

The GEOTIFF files provide the metadata attributes as key value pairs, according to the Climate and Forecast Convention (CF, version 1.6):

- on the file-level (Table 7);
- on the band-level, with an example values given for the discrete classification layer (Table 8).

Table 7: Description of GEOTIFF file attributes

Attribute name	Description	Example 2015 map
archive_facility	Specifies the name of the institution that archives the product	VITO NV
copyright	Text to be used when referring to this product in publications (copyright notice)	Copernicus Service information 2019
creator	Principal investigator of the algorithm	Dr. Marcel Buchhorn (VITO)
delivered_product_crs	Land Cover product is delivered in this Coordinate Reference System	WGS84 (EPSG:4326)
delivered_product_grid	Land Cover product is delivered in this tile grid	20x20 deg tiling grid
doi	Data Object Identifier	10.5281/zenodo.3518056
file_creation	File creation timestamp	Fri Apr 19 11:46:08 2019
history	A global attribute for an audit trail. One line, including date in ISO-8601 format, for each invocation of a program that has modified the dataset.	2019-04-26 Processing line LC100
Info	Additional comment on the processing history.	MasterTile W160N20 for product discrete-classification of CGLOPS LC100 layers for epoch 2015.
institution	The name of the institution that produced the product	VITO NV
long_name	Extended product name	Land Cover
orbit_type	Orbit type of the orbiting platform(s)	LEO
platform	Name(s) of the orbiting platform(s)	Proba-V
processing_level	Product processing level	L3
processing_mode	Processing mode used when generating the product (Near-Real Time, Consolidated or Reprocessing)	Offline
production_crs	Coordinate Reference System used for the pre-processed input data and during the different production steps	UTM
production_grid	Grid used for the pre-processed input data and during the different production steps	MGRS (Sentinel-2 tiling grid)
product_version	Version of the product	V2.1.1
references	Published or web references with more product information.	https://land.copernicus.eu/global/products/lc
region_name	Name of the geographic area covered, e.g. name of the 20x20 degree tile	W160N20
sensor	Name(s) of the sensor(s) used	VEGETATION
source	The method of production of the original data	Derived from 100m EO satellite imagery
time_coverage_end	End date and time of the temporal coverage of the input data.	2016-12-31T23:59:59Z
time_coverage_start	Start date and time of the temporal coverage of the input data.	2014-01-01T00:00:00Z
title	A description of the contents of the file	Dynamic Land Cover Map 100m 2015

Table 8: Description of GEOTIFF band attributes.

Attribute	Description	Examples for LCCS layer
CLASS	Dataset type	DATA
band_crs	Coordinate Reference System used for this GeoTIFF band.	WGS84 (EPSG:4326)
flag_meanings	Description for each flag value	unknown, ENF_closed, EBF_closed, DNF_closed, DBF_closed, mixed _closed, unknown_closed, ENF_open, EBF_open, DNF_open, DBF_open, mixed _open, unknown_open, shrubland, herbaceous_vegetation, cropland, built-up, bare_sparse_vegetation, snow_ice, permanent_inland_water, herbaceous_wetland, moss_lichen, sea
flag_values	Provides a list of the specific values used.	0, 111, 112, 113, 114, 115, 116, 121, 122, 123, 124, 125, 126, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200
long_name	A non-standardized, descriptive name that indicates a variable's content.	Land Cover Classification
missing_value	Single value, outside of valid_range, used to represent missing or undefined data, for applications following older versions of the standards.	255
short_name	A shortened, non-standardized name.	discrete-classification
unit	Physical unit. None or omitted when the data is dimensionless.	None
valid_range	Smallest and largest valid values.	0, 254

3.4 Product Characteristics

3.4.1 Projection and Grid Information

The Land Cover v2.1 products are delivered in a regular latitude/longitude grid (EPSG:4326) with the ellipsoid WGS 1984 (Terrestrial radius=6378 km). The resolution of the grid is 1°/1008 or approximately 100 m at equator.

3.4.2 Spatial Information

The Land Cover v2.1 products cover the geographic area from longitude 60°E to 20°W and latitude 40°N to 60°S, i.e. the African continent. They are provided in 20 x 20 degree tiles (see Figure 11).

The position of a pixel is by standard, through GeoTIFF format, given by the "upper left corner", as marked wih the GDALMD_ARE_OR_POINT[2] geotransform metadata entry.

3.4.3 Temporal Information

The Land Cover v2.1 product is provided with yearly updates, each representing the land cover for the epoch or reference calendar Year (from 01 January to 31 December). The data 1 year prior and past of the reference year is used in its processing. As such, the temporal coverage provides a start date of 01 January Year-1 to 31 December Year+1.

The temporal coverage for the Near Real Time (NRT) Land Cover v2.1 product provides a start date of 01 April Year-1 to 31 March Year+1.

3.5 DATA POLICIES

EU law[3] grants free and open access to Copernicus Sentinel Data and Service Information, which includes Global Land Service products, for the purpose of the following use in so far as it is lawful:

a) reproduction;
b) distribution;
c) communication to the public;
d) adaptation, modification and combination with other data and information;
e) any combination of points (a) to (d).

EU law allows for specific limitations of access and use in the rare cases of security concerns, protection of third party rights or risk of service disruption.

By using (Sentinel Data or) Service Information the user acknowledges that these conditions are applicable to him/her and that the user renounces to any claims for damages against the European Union and the providers of the said Data and Information. The scope of this waiver encompasses any dispute, including contracts and torts claims that might be filed in court, in arbitration or in any other form of dispute settlement.

Where the user communicates to the public on or distributes the **original** Land Cover products, he/she is obliged to refer to the data source with (at least) the following statement (included as the copyright metadata item):

Copernicus Service information [Year]

With [Year]: year of publication

Where the user has adapted or modified the products, the statement should be:

---

[2] GDAL identifies GDALMD_AOP_AREA as "geotransform is the position of the upper left corner of the area spanned by the pixel, which center is the upper left point in the dataset if interpreted as DTM".
[3] European Commission, Regulation (EU) No 377/2014 and Commission Delegated Regulation (EU) No 1159/2013.

Contains modified Copernicus Service information [Year]

For complete acknowledgement and credits, the following general statement can be used:

"The product was generated by the Global component of the Land Service of Copernicus, the Earth Observation programme of the European Commission. The research leading to the current version of the product has received funding from various European Commission Research and Technical Development programs. The product is based on PROBA-V data provided by Belgian Science Policy Office (BELSPO) and distributed by VITO."

For citation in scientific publications, the following statement can be used*:*

Buchhorn, M., Smets, B., Bertels, L., Lesiv, M., Tsendbazar, N.-E., Masiliunas, D., Herold, M., Fritz, S. (2019). Copernicus Global Land Service: Land Cover 100m, Collection 2, epoch <year> Africa Demo (Version V2.1.1) [Data set]. Zenodo. DOI: <doi>;

with <year>, the actual year of dataset used, i.e. 2015, 2016, 2017 or 2018; and <doi> the identifier as shown in Table 9.

The user accepts to inform Copernicus about the outcome of the use of the above-mentioned products and to send a copy of any publications that use these products to the scientific & technical support (help desk) contact specified in the next section.

3.6 Access and Contacts

The Land Cover products are available through the Global Land Cover viewer, available at https://land.copernicus.eu/global/lcviewer (see Figure 12). It displays the various land cover layers (discrete map, cover fractions, false-colour combinations of cover fractions) on a map, allows downloading the data in 20x20 degree tiles and reports on land cover statistics per administrative area.

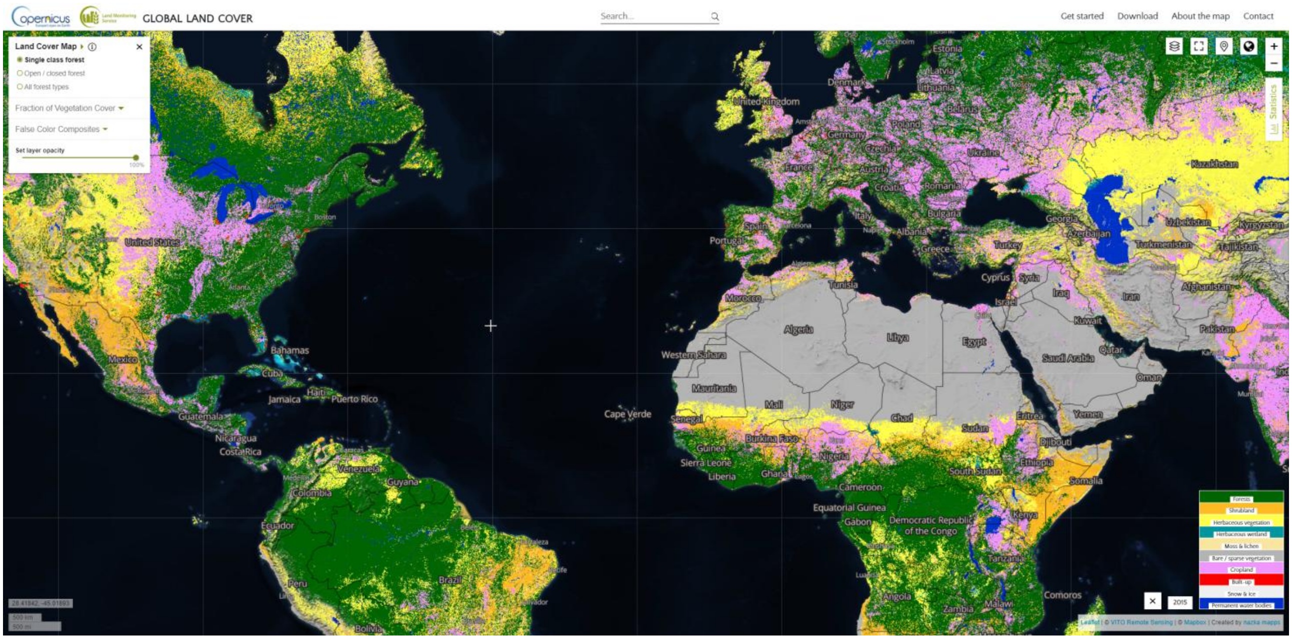

Figure 12: Screenshot of the Global Land Cover viewer

The yearly Land Cover change v2.1 maps will become available soon in this viewer.

More information and documentation about the product is available from the Copernicus Global Land Service web site at https://land.copernicus.eu/global/products/lc

Accountable Contact: European Commission, Directorate-General Joint Research Centre, Italy Email address: copernicuslandproducts@jrc.ec.europa.eu

Scientific & Technical support contact e-mail address:
https://land.copernicus.eu/global/contactpage.

Direct access to the data can also be achieved through the Digital Object Identifier (DOI), following the information as shown in Table 9 below.

Table 9 : Digital Object Identifier information

Attribute	Value
Dataset	Collection 2
Provided epochs	2015, 2016, 2017, 2018
Dataset version	2.1.1
Concept DOI	not applicable
Version DOI	10.5281/zenodo.3518056 (2015 base map) 10.5281/zenodo.3518061 (2016 conso map) 10.5281/zenodo.3518083 (2017 conso map) 10.5281/zenodo.3518087 (2018 NRT map)
Citation	Buchhorn, M., Smets, B., Bertels, L., Lesiv, M., Tsendbazar, N.-E., Masiliunas, D., Herold, M., Fritz, S. (2019). Copernicus Global Land Service: Land Cover 100m, Collection 2, epoch <year> Africa Demo (Version V2.1.1) [Data set]. Zenodo. DOI: <doi>
Direct Access	http://doi.org/<version DOI>
Data layers per epoch	13
File Size	7 GBytes

4 VALIDATION RESULTS

The CGLS_LCC100m V2.1 product for Africa (including both the discrete maps and fractional land cover layers for 2015-2018) is assessed using an independent validation dataset, which includes 3616 sample sites with labels collected for the year 2015. Detailed information about the collection of validation dataset can be found in [CGLOPS1_VR_LC100m-V2.0]CGLOPS1_VR_LC100m-V2.0.

This validation focus on the assessment of the temporal consistency between the yearly maps rather than the statistical map accuracy as the validation data is only covering the 2015 year. Indeed, for the statistical map accuracy, updated sample sites for changed areas are required.

In addition to the consistency assessments, we report on the change area within the period based on the maps and as well as visual evaluation of detected changes.

The details of the analysis and results are available in the validation report [CGLOPS1_VR_LCC100_V2.1]. Below, a summary of major outcomes is given.

4.1 TEMPORAL CONSISTENCY

For the discrete maps, a confusion or error matrix is calculated for each year based on the reference land cover dataset. For the fractional land cover layers, namely tree cover, shrub cover, grass cover, crops, urban, bare/sparse vegetation, lichen/moss and permanent water, the mean absolute error (MAE) and root mean square error (RMSE) were calculated.

Overall map accuracies are 80.4%, 80.1%, 79.9% and 79.5% for 2015, 2016, 2017 and 2018, respectively, showing consistency in high overall map accuracy over time. There is a slight decrease in the map accuracy over years. This could be explained by the fact that the validation data is collected for 2015 and the temporal gaps (between map and validation data) increase for the latter years. It is possible that changes in land cover might have occurred at validation sample locations after 2015 and that information is not included in the validation data.

Among the fractional land cover layers, rare classes such as lichen/moss, permanent water and urban area fraction maps show the lowest errors. Grass fraction maps have the highest errors for all years (MAE≥15.7%). Similar to the overall accuracies of the discrete maps, the errors tent to increase over time. For example there is 1.3% increase in MAE for grass and crops between 2015 and 2018. This could be partially due to the temporal gaps between the map and the validation data which may have larger impact on more dynamic land cover class such as crops, and partially due to the NRT mode of 2018 map as indicated above.

4.2 LAND COVER CHANGE ANALYSIS

4.2.1 Statistical analysis

The area of forest marginally increases from 28.80% of the total area of Africa in 2015 to 29.07% in 2018. Cropland area also marginally increases from 7.99% in 2015 to 8.23% in 2018. The area of wetland has a slight increase between 2015 and 2018. On the other hand, shrubs, herbaceous vegetation and bare area decrease from 2015 to 2018. For example, shrubs decrease from 13.30% in 2015 to 12.84% in 2018. Urban area and water are quite stable over time.

The change area for Africa between 2015 and 2018 was also assessed without considering the changes between “shrubs” and “herbaceous vegetation”, as these changes are not indicated as important by the users (Figure 16). Thus “shrubs” and “herbaceous vegetation” were combined as “another vegetation” class (Figure 16) in the estimation. Table 10 shows the change area in Africa between the years and for the whole period. The land cover changes observed in the V2.1 Africa maps are maximum of 1% yearly change. For the period between 2015 and 2018, 2.29% of the area is changed.

Table 10: Land cover area change in Africa between 2015 and 2018 (the changes between “shrubs” and “herbaceous vegetation” are not included in the calculation)

	2015-2016	2016-2017	2017-2018	2015-2018
Changed area ($\times 10^4$ km^2)	30.86	30.53	25.85	70.33
Total area ($\times 10^4$ km^2)	3070.50	3070.50	3070.50	3070.50
Change percentage	1.00%	0.99%	0.84%	2.29%

4.2.2 Visual assessment

To investigate if the land cover change information is well captured by the CGLS_LCC100m V2.1 product, we have selected several example locations (detailed information on how these locations are selected can be found in [CGLOPS1_VR_LCC100_V2.1]), and performed visual validation by comparing the change maps with available satellite imagery from Google Earth, Sentinel imagery and Bing maps. Typical examples are reported here below.

Southern Ghana; Coordinates: 6.458°N, 2.029°W

Description: The CGLS_LCC100m V2.1 maps have detected loss of forest and expansion of flooded areas along the Ghana river over 2015-2016 (Figure 13 a, b). Google Earth imagery confirms the deforestation and increase of regularly flooded pond areas (Figure 13 e, f). Figure 13 g and h clearly show the creation of new inundated ponds during 2015. Actually, the appearance of these man-made ponds is due to large mining activities in Ghana. The CGLS_LCC100m V2.1 maps capture well the process of deforestation and expansion of flooded and inundated areas.

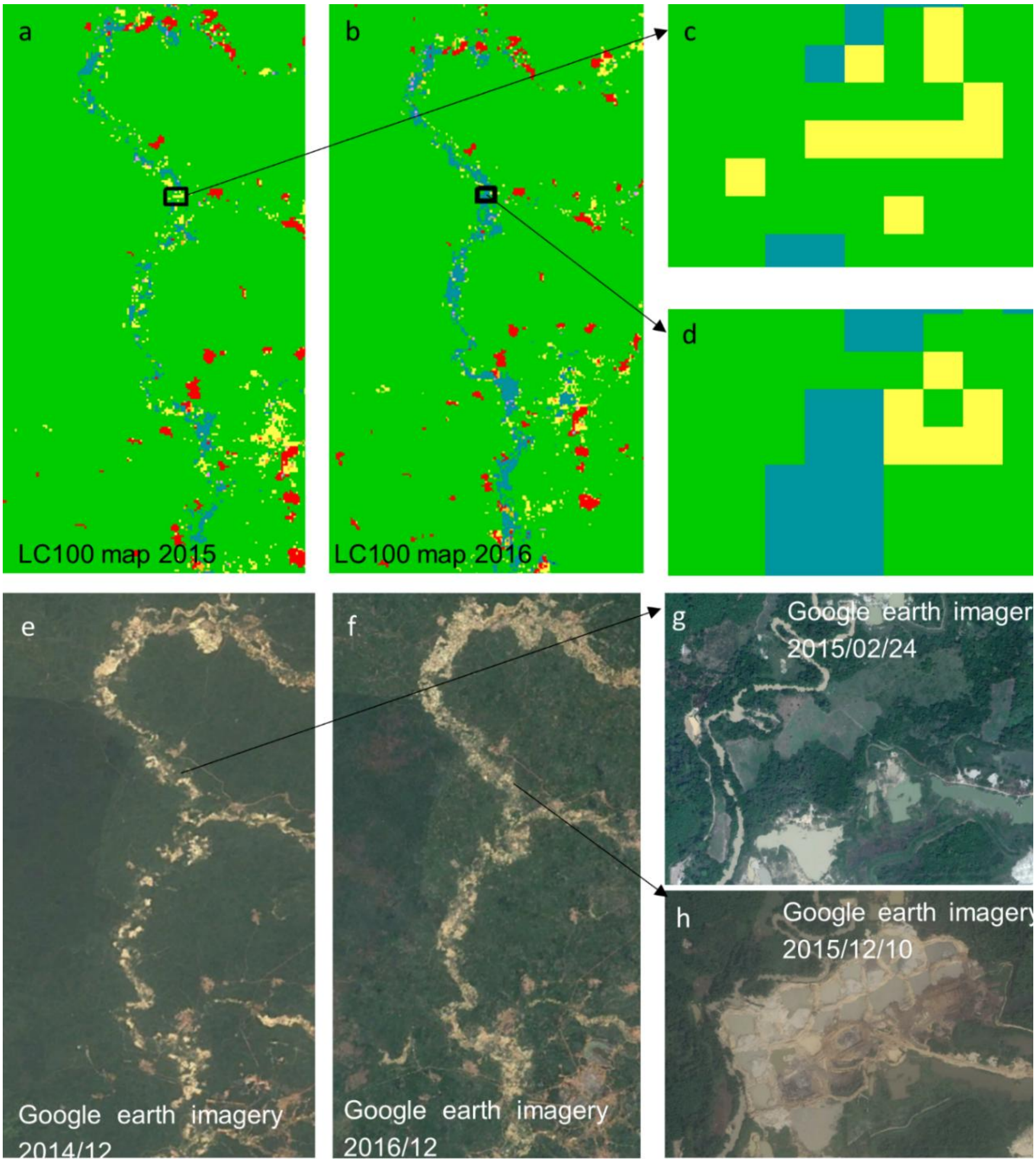

Figure 13: An example showing the CGLS_LCC100m V2.1 maps successfully detecting loss of forest and expansion of inundated area in Ghana. Color legend is shown in Table 4.

Morocco; Coordinates: 30.589°N, 8.713°W

Description: The CGLS_LCC100m V2.1 maps have detected agriculture land expansion in an area in Morocco over 2015-2018, which is confirmed from Google Earth imagery (Figure 14).

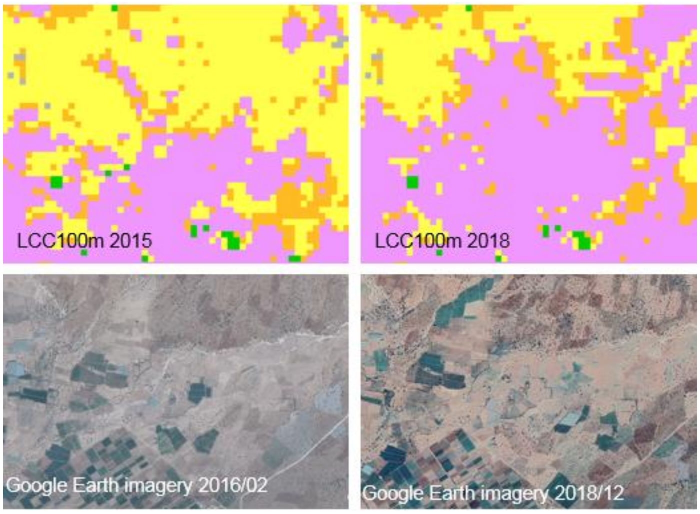

Figure 14: An example showing the CGLS_LCC100m V2.1 maps detecting the crop land expansion in Morocco. Color legend is shown in Table 4.

Although some of the processes related to land cover change have been correctly captured by the yearly maps, some spurious changes are also detected in the CGLS_LCC100m V2.1 maps. One example is shown below.

Southwestern Angola; Coordinates: 13.751°E, 16.512°S

Description: Shrubs were misclassified as forest in the 2017 and 2018 maps. The CGLS_LCC100m V2.1 maps have detected increased forest area (Figure 15). However, high resolution imagery confirms that the area is still covered by shrubs.

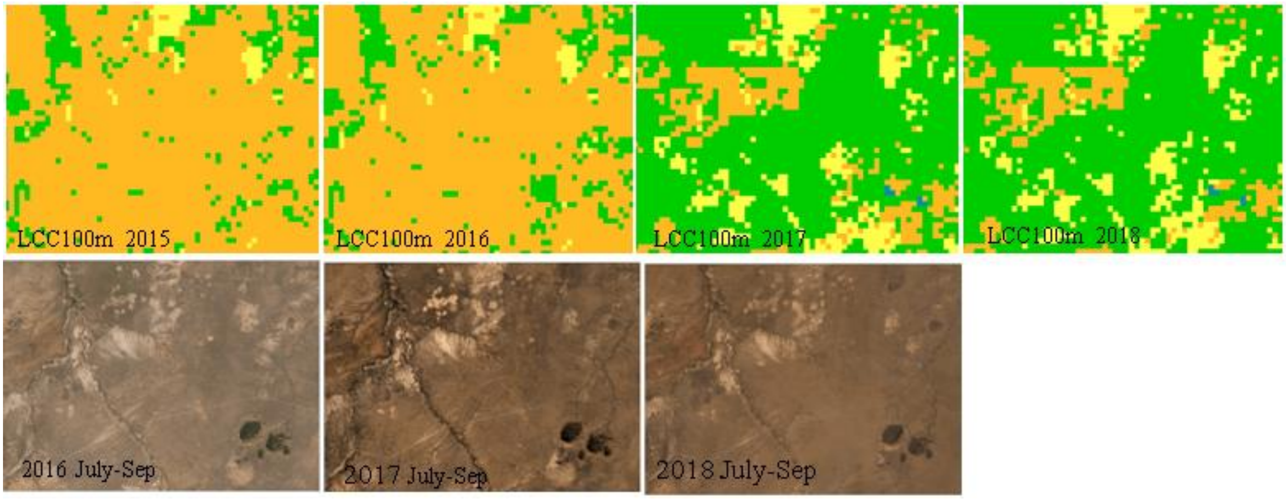

Figure 15: An example showing false detection of forest expansion by CGLS_LCC100m V2.1 maps. Color legend is shown in Table 4.

4.3 CONCLUSION

The CGLS_LCC100m V2.1 discrete maps for Africa 2015-2018 show high temporal consistency. The overall accuracies of the yearly maps are around 80% (80.4%, 80.1%, 79.9% and 79.5% for 2015, 2016, 2017 and 2018, respectively). The slightly decreasing trend in the overall accuracy could be a consequence of using a validation dataset representative of year 2015. Updated validation datasets corresponding to the mapping period are needed for the statistical accuracy assessment of 2016-2018 maps.

Among the fractional land cover layers, rare classes such as lichen/moss, permanent water and urban area fraction maps show the lowest errors. Fractional grass maps have the highest errors for all years (MAE≥15.7%). Similar to the overall accuracies of the discrete maps, the errors tent to increase over time. This could be also partially due to the temporal gaps between the map and the validation data.

Visual assessments of the land cover change areas revealed accurate detection of water expansion, deforestation, cropland expansion and wetland dynamics. Visual verification also identified false detection of change in forest and cropland areas that requires improvements.

Users are encouraged to provide feedback on disagreement areas through the GEO-WIKI platform feedback tool at https://application.geo-wiki.org/Application/index.php or through contacting the helpdesk (see §3.6).

5 References

Verbesselt, J., Hyndman R., Gewnham, G., Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. Remote Sensing of Environment 114 (2010) 106-115, doi: 10.1016/j.rse.2009.08.014.

Wagner, J. E., Stehman, S. V. (2015). Optimizing sample size allocation to strata for estimating area and map accuracy. Remote Sensing of Environment 168 (2015) 126-133, doi:10.1016/j.rse.2015.06.027

ANNEX: REVIEW OF USERS REQUIREMENTS

According to the applicable documents [AD2] and [AD3], the user's requirements relevant for Dynamic Moderate Land Cover are:

- **Definition**: Dynamic global land cover products at 300m and/or 100m resolution using UN Land Cover Classification System (LCCS)
- **Geometric properties**:
  - Pixel size of output data shall be defined on a per-product basis so as to facilitate the multi-parameter analysis and exploitation.
  - The baseline datasets pixel size shall be provided, depending on the final product, at resolutions of 100m and/or 300m and/or 1km.
  - The target baseline location accuracy shall be 1/3 of the at-nadir instantaneous field of view.
  - Pixel co-ordinates shall be given for centre of pixel.
- **Geographical coverage**:
  - geographic projection: lat long
  - geodetical datum: World Geodetic System 1984 (WGS84)
  - pixel size: 1/112° - accuracy: min 10 digits
  - global window coordinates:
    - Upper Left: 180°W-75°N
    - Bottom Right: 180°E, 56°S

- **Accuracy requirements**: Overall thematic accuracy of dynamic land cover mapping products shall be >80%. The overall accuracy assessment (including confidence limits) will be based on a stratified random sampling design and the minimum number of sampling points per land cover class relevant to the product shall be calculated as described in Wagner and Stehman, 2015.

Few workshops were held in 2016 to consult different stakeholders to understand users' needs for global land cover maps. A feasibility study was performed to define the guidelines to create the first LC100 map. More details can be found in [CGLOPS1_URD_LC100m]. Larger consultations in 2017 and 2018 allowed collecting the requirements of wide user communities which were translated in product specifications [CGLOPS1_PSD_LC100m]. Below land cover type and land change processes that were found to be important for different international actions and programmes are summarized in Table 11 and Figure 16.

Table 11: Usefulness of information on LC and LC change processes for different international actions and programmes.

	LC types	Related land change processes	UNFCCC	UNCCD	OECD	SEEA/FAO	SDGs
1	Urban/built-up areas	Urbanization	✓	✓	✓	✓	✓
2	Cropland	Crop expansion	✓	✓	✓	✓	✓
3	Cropland and other vegetation	Land abandonment	✓	✓	✓	✓	✓
4	Forest	Deforestation	✓	✓	✓	✓	✓
5	Forest	Reforestation	✓	✓	✓	✓	✓
6	Wetland	Wetland degradation	✓	✓	✓	✓	✓
7	Water body	Expansion of water surface			✓	✓	✓
8	Water body	Reduction of water surface			✓	✓	✓
9	Bare areas	Desertification			✓	✓	✓

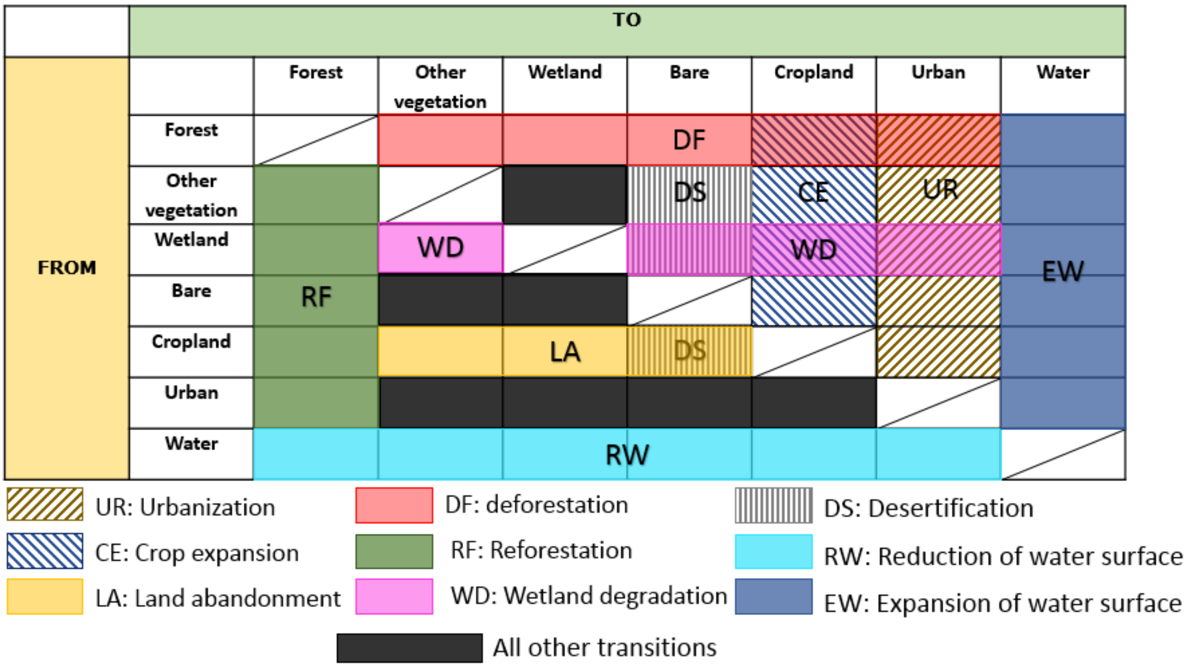

Figure 16: Land change processes and land cover transformations by user requirement