International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: repository@iiasa.ac.at
Web: www.iiasa.ac.at

Working Paper **WP-17-021**

Evaluation of ESA CCI prototype land cover map at 20m

Myroslava Lesiv, lesiv@iiasa.ac.at

Steffen Fritz, IIASA, fritz@iiasa.ac.at

Ian McCallum, IIASA, mccallum@iiasa.ac.at

Nandika Tsendbazar, Wageningen University and Research, the Nertherlands

Martin Herold, Wageningen University and Research, the Nertherlands

Jean-Francois Pekel, Joint Research Center, Ispra, Italy

Marcel Buchhorn, VITO

Bruno Smets, VITO

Ruben Van De Kerchove, VITO

Approved by

Michael Obersteiner
Program Director, Ecosystems Services and Management (ESM)

27 November 2017

Contents

Abstract

In September 2017, the ESA CCI Land Cover Team released a prototype land cover (LC) map at 20 m resolution over Africa for the year 2016. This is the first LC map produced at such a high resolution covering an entire continent for the year 2016. To help improve the quality of this product, we have assessed its overall accuracy and identified regions where the map should be improved. We have compared the product against two independent datasets developed within the Copernicus Global Land Services (CGLS): a reference land cover dataset at a 10 m resolution, which has been used as training data to produce the LC map at 100 m over Africa for the year 2015 (http://land.copernicus.eu/global/products/lc); and an independent validation dataset at a 10 m resolution, which has been developed by CGLS for independent assessment of land cover maps at resolutions finer than 100 m. According to our estimates, overall accuracy of the African CCI LC at 20 m is approximately 65%. We have highlighted regions where the spatial distribution of such classes as shrubs, crops and trees should be improved before the map at 20 m could be used as input for research questions, e.g. conservation of biodiversity, crop monitoring and climate modelling.

Acknowledgments

This work has been supported by the Copernicus Global Land Services (CGLS) and the ERC project CrowdLand (617754).

About the Authors

Myroslava Lesiv, PhD, is a Research Scholar in the Earth Observation Group (EOS) in the Ecosystems Services and Management Program (ESM), IIASA. Steffen Fritz, PhD, is the Deputy Program Director of ESM and group leader of EOS. Ian McCallum, PhD, is a Research Scholar in the EOS group.

This report has been prepared in collaboration with:

- Nandika Tsendbazar, PhD, Research Scientist at Wageningen University and Research, the Netherlands
- Jean-Francois Pekel, PhD, Research Scientist, European Commission, Joint Research Centre (JRC), Italy
- Martin Herold, PhD, Professor of Geoinformation Science and Remote Sensing at Wageningen University and Research, the Netherlands
- Marcel Buchhorn, PhD, Senior R&D Professional with VITO and is within the biodiversity team responsible for algorithm development / implementation for global land cover products and other biodiversity variables for multiple RS platforms and sensors.
- Bruno Smets, MSc, R&D Project Manager at VITO, Belgium and is leading the consortium that implements the Copernicus Global Land Service.
- Ruben Van De Kerchove, PhD, Senior R&D professional at VITO, Belgium.

1. Introduction

In recent decades, increasing amounts of satellite imagery have become freely available for land surface monitoring. Furthermore, the quality and spatial resolution of this imagery is constantly improving.

In 2008, the US Geological Survey (USGS) made freely available the entire Landsat archive at 30 m resolution, and, in 2014, ESA launched the Sentinel missions at a resolution of 10 m, with all data being freely available. With such an impressive evolution in earth observation, the products derived from remote sensing, in particular, land cover (LC) maps, have also become of finer resolution, moving from the defacto standard of 1 km resolution e.g. GLC2000 (Fritz et al., 2003) and 300 m ESA GlobCover products (http://due.esrin.esa.int/files/20160624100912.pdf), down to 100m e.g. Copernicus Global Land (http://land.copernicus.eu/global/products/lc) and 30 m products e.g. Globeland (Jun et al., 2014). While the resolution of LC products becomes finer over time, the reported accuracies do not always satisfy user requirements (Tsendbazar et al., 2017b).

This evaluation report is a response to the ESA Climate Change initiative (CCI) request to provide feedback and comments regarding the quality of the first-ever prototype LC map at 20 m resolution over Africa (http://2016africalandcover20m.esrin.esa.int/). This report includes an independent accuracy assessment (based on two validation datasets) of the LC map at 20 m resolution, as well as feedback and comments on possible improvements.

Evaluation of the CCI LC map at 20m

To evaluate the LC map at 20 m, we have used two independent datasets: (1) reference dataset that has been developed as training data to produce the Copernicus LC map at 100 m (Lesiv et al., 2017); and (2) an independent validation land cover dataset for independent assessment of land cover maps at resolutions finer than 100 m. Both datasets were developed within the Copernicus Land Cover project, which is on-going. The datasets will be publicly available at the end of the project. In this report, we refer to both datasets as the CGLS (Copernicus Global Land Services) datasets.

1.1. Evaluation based on the CGLS reference dataset at 10m resolution

This reference dataset (Lesiv et al., 2017) at a 10 m resolution has been collected by experts at the International Institute for Applied Systems Analysis (IIASA). It contains approximately 24 K sample sites, from which we have used 19,548 sample sites with a high level of confidence. The data has been collected through the Geo-Wiki web-application (https://geo-wiki.org/). The experts were asked to visually interpret high resolution imagery (Google and Bing) and to analyze NDVI time-series and historical imagery in Google Earth at each sample site.

The sample design of reference data has been systematic (with the same distance between sample sites – 35 km) in order to represent well the African landscapes. Some parts of homogenous landscapes such as deserts and rainforest were excluded. To ensure more training points for rare classes and areas with low accuracies, additional sample sites were added accordingly (see Figure 1).

Figure 1: Spatial distribution of the reference sample sites in Africa

Each sample location is a point that corresponds to a centre of a PROBA-V 100 m pixel. The Geo-Wiki tools generate a PROBA-V 100 m pixel extent for each sample location and split the pixel area into 100 equal subpixels (polygons) that are being validated one by one. Figure 2 illustrates a sample location (red point) and a generated PROBA-V 100 m pixel (yellow box) that is validated in the Geo-Wiki.

Figure 2: Sample location that corresponds to a PROBA-V 100 m pixel which is split into 100 equal subpixels (polygons)

To evaluate the CCI LC map at 20 m resolution, we have aggregated it to a 100 m resolution by applying a majority rule. We assumed that land cover class “lichens/mosses” corresponds to bare land in the training dataset. We have also aggregated the training dataset to the 100 m resolution by applying a majority rule.

We have calculated confusion matrices, based on Olofsson et al (2014), see Tables 1 and 2. Table 1 summarizes the confusion errors as a proportion of the total reference sites. In this case the overall accuracy of the CCI LC 20 m product was 58%. This error matrix was corrected for the area of the land cover types in Africa (Table 2). Overall area weighted accuracy was 65 % +/- 1 %.

Table 1: Confusion matrix without area bias correction

Mapped Classes	Reference class								Total	User accuracies
	Trees	Shrubs	Grass-land	Crops	Wet-lands	Bare lands	Urban/ built-up	Water		
Trees	3035	436	1339	82	21	12	0	8	4933	62%
Shrubs	207	422	2287	248	7	91	2	5	3269	13%
Grassland	330	313	2680	191	51	410	10	7	3992	67%
Crops	122	90	1202	1334	6	121	22	6	2903	46%
Wetlands	4	0	5	4	3	5	0	4	25	12%
Bare lands	1	12	189	19	1	2787	64	3	3076	91%
Urban/ built-up	19	1	164	6	1	80	573	0	844	68%
Water	0	1	14	7	3	12	3	466	506	92%
Total	3718	1275	7880	1891	93	3518	674	499	19548	
Producer accuracies	82%	33%	34%	71%	3%	79%	85%	93%		**58%**

Table 2: Confusion matrix with area bias correction

Mapped Classes	Reference class								Total proportions	User accuracies	Confidence intervals
	Trees	Shrubs	Grass-land	Crops	Wet-lands	Bare lands	Urban/ built-up	Water			
Trees	13.96	2.01	6.16	0.38	0.10	0.06	0.00	0.04	22.69	62%	1%
Shrubs	0.75	1.54	8.32	0.90	0.03	0.33	0.01	0.02	11.89	13%	1%
Grassland	1.48	1.40	12.02	0.86	0.23	1.84	0.04	0.03	17.90	67%	2%
Crops	0.53	0.39	5.21	5.79	0.03	0.52	0.10	0.03	12.59	46%	2%
Wetlands	0.02	0.00	0.02	0.02	0.01	0.02	0.00	0.02	0.12	12%	13%
Bare lands	0.01	0.13	2.05	0.21	0.01	30.22	0.69	0.03	33.35	91%	1%
Urban/ built-up	0.00	0.00	0.04	0.00	0.00	0.02	0.13	0.00	0.19	68%	3%
Water	0.00	0.00	0.04	0.02	0.01	0.03	0.01	1.17	1.27	92%	2%
Total proportions	16.75	5.47	33.85	8.17	0.41	33.04	0.98	1.33	100		
Producer accuracies	83%	28%	36%	71%	4%	92%	13%	88%		**65%**	**1%**
Confidence intervals +/-	1%	2%	1%	2%	4%	1%	2%	4%			

1.2. Evaluation based on the CGLS independent validation dataset at 10m resolution

A second independent land cover map validation dataset was collected in order to validate the CGLS-LC100 map of Africa (Tsendbazar et al., 2017a). The validation dataset is based on

stratified sampling following the global stratification suggested by Olofsson et al. (2012). Figure 3 shows spatial distribution of validation sample sites. It has land cover information at a 10 m resolution over approximately 100 m x 100 m areas for 3716 unique locations. The data was collected by regional experts from Africa through the Geo-Wiki web-application. Similar to the previous reference data, the experts were asked to visually interpret high resolution imageries (using Google and Bing imagery), historical imageries in Google Earth and NDVI time-series profiles. These data were collected independently from the reference dataset stated in Section 2.1., on a different branch of the Geo-Wiki and by different experts. In addition to the general stratifications, additional sample sites were also collected for rare classes based on the CGLS-LC100 product. The sample unit areas corresponded to the pixels of the Proba-V 100 m data and within this area, land cover information was recorded for 10x10 subpixels (each covering ~10 m x 10 m areas.) similar to Figure 3.

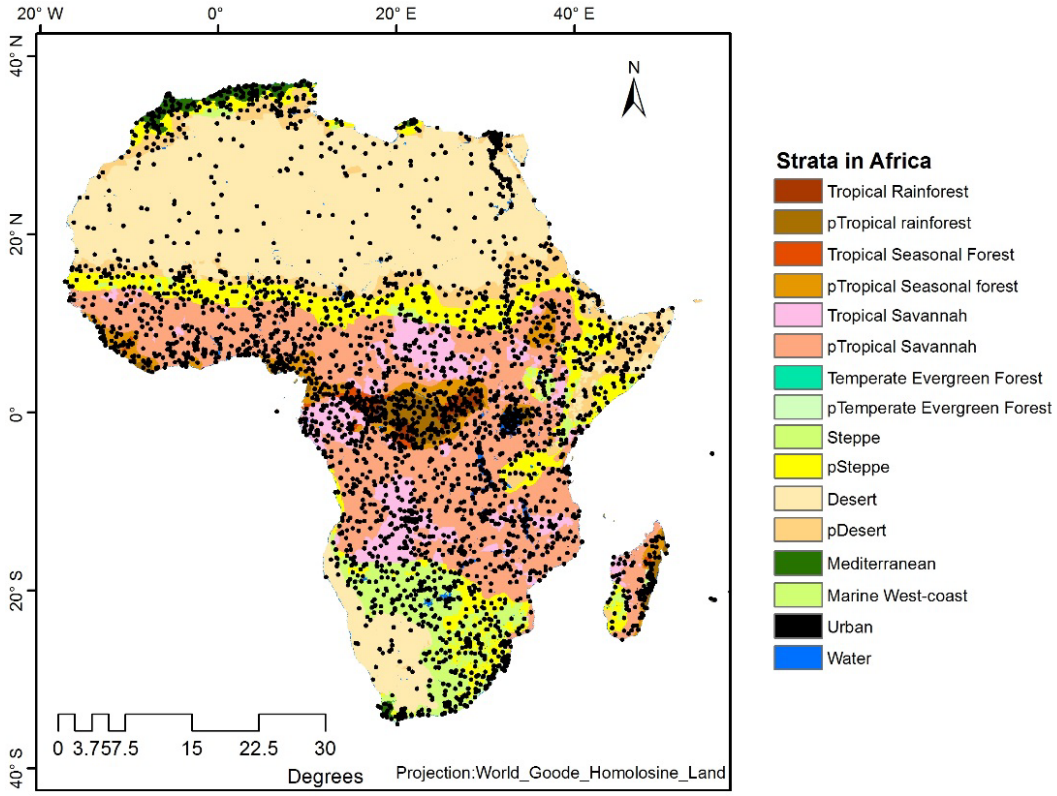


Figure 3: Spatial distribution of the validation sample sites (Tsendbazar et al., 2017a)

To evaluate the CCI map at 20 m resolution, we selected sample pixels from the CCI LC map at 20 m that contain at least 4 center points of the subpixels of the validation data to represent approximately 20 m x 20 m areas. To reduce the impact of possible geo-location shift between the 4 subpixels of the validation data and the pixels of the CCI LC 20 m map, we selected 4 subpixels with homogeneous land cover types, with the assumption that impact of the shift can be less in homogeneous land cover areas. In total 41, 059 sample pixels were used for the validation.

We merged litchen/mosses and bare classes of the CCI LC 20 m map to the bare class as the validation data does not separate litchen/mosses and bare land. Furthermore, snow and/or ice classes were not assessed and there can be a difference in the wetland definition since the CGLS land cover validation data defines wetland vegetation as wetland herbaceous vegetation. Therefore, the land cover types used for the assessments are trees, shrubs, grassland, cropland wetland, bare/sparse vegetation, built up areas and open water.

We calculated the confusion matrix before and after area bias correction following the method of Olofsson et al (2014), see Tables 3 and 4.

Table 3 summarizes the confusion errors as a proportion of the total validation sites. Overall accuracy here was 67.2 %. This error matrix was corrected for the area of the land cover types in Africa (Table 4). Overall area weighted accuracy was 64.3 % +/- 0.5 %. Lower class specific accuracies of large-area classes (e.g., grassland and cropland) influenced to the reduction of the overall area weighted accuracy compared with overall accuracy before area bias correction. Furthermore, the producers accuracy for several rare classes such as water and urban dropped after area bias correction due to confusion of some sample pixels with large-area classes namely grassland, cropland and bare/sparse vegetation.

Table 1: Confusion matrix for the CCI20m map as proportions of total validation sites

Mapped classes	Reference class								**Correct proportion**	Total proportion	**User's accuracy**
	Trees	Shrubs	Grass - land	Crops	Wet-lands	Bare lands	Urban/ built-up	Water			
Trees	**10144**	249	1290	190	276	8	3	41	**10144**	12201	**83.1**
Shrubs	215	**565**	1643	321	331	80	3	10	**565**	3168	**17.8**
Grassland	618	273	**2901**	511	1753	510	21	159	**2901**	6746	**43.0**
Crops	450	89	1828	**2965**	396	100	21	29	**2965**	5878	**50.4**
Wetlands	34	0	43	10	**493**	5	0	152	**493**	737	**66.9**
Bare lands	20	15	267	122	92	**3834**	72	168	**3834**	4590	**83.5**
Urban/ built-up	93	4	133	29	0	85	**736**	14	**736**	1094	**67.3**
Water	26	6	6	19	592	42	0	**5954**	**5954**	6645	**89.6**
Correct proportion	**10144**	**565**	**2901**	**2965**	**493**	**3834**	**736**	**5954**	**27592**		
Total proportion	11600	1201	8111	4167	3933	4664	856	6527		**41059**	
Producer's accuracy	87.4	**47.0**	**35.8**	**71.2**	**12.5**	**82.2**	**86.0**	**91.2**			**67.2**

Table4: Confusion matrix for the CCI20m map as proportions of African continent area

Mapped classes	Reference class								**Correct proportion**	Total proportion	**User's accuracy**	Confidence interval +/-
	Trees	Shrubs	Grass-land	Crops	Wet-lands	Bare lands	Urban/ built-up	Water				
Trees	19.09	0.47	2.43	0.36	0.52	0.02	0.01	0.08	**19.09**	22.96	**83.1**	0.7
Shrubs	0.79	2.07	6.02	1.18	1.21	0.29	0.01	0.04	**2.07**	11.61	**17.8**	1.3
Grassland	1.64	0.73	7.72	1.36	4.67	1.36	0.06	0.42	**7.72**	17.95	**43.0**	1.2
Crops	0.98	0.19	3.97	6.45	0.86	0.22	0.05	0.06	**6.45**	12.78	**50.4**	1.3
Wetlands	0.01	0	0.01	0.002	0.10	0.001	0	0.03	**0.10**	0.14	**66.9**	3.4
Bare lands	0.15	0.11	1.940	0.89	0.669	27.86	0.523	1.221	**27.86**	33.36	**83.5**	1.1
Urban/ built-up	0.02	0.001	0.03	0.01	0	0.02	0.14	0.003	**0.14**	0.21	**67.3**	2.8
Water	0.004	0.001	0.001	0.003	0.09	0.01	0	0.89	**0.89**	1	**89.6**	0.7
Correct proportion	**19.09**	**2.07**	**7.72**	**6.45**	**0.10**	**27.86**	**0.14**	**0.89**	**64.3**			
Total proportion	22.67	3.57	22.12	10.24	8.11	29.77	0.78	2.74		**100**		
Producer's accuracy	**84.2**	**58.0**	**34.9**	**63.0**	**1.2**	**93.6**	**17.9**	**32.5**			**64.3**	**0.5**
Confidence interval +/-	0.7	2.7	0.8	1.6	0.1	0.4	2.9	2.3				

1.3. Summary of feedback

The overall accuracies of the CCI LC 20 m product estimated by using two independent datasets and two different approaches come to matching results around 65%. User and producer accuracies in Tables 1 and 2 are much lower, because the reference dataset contains sample sites that were added in regions that are difficult to map. So, Tables 1 and 2 have a bias towards areas mapped with lowest accuracy. We consider Tables 1 and 2 as the worst case accuracy estimates. Tables 3 and 4 present the objective and valid accuracy estimates, based on the validation dataset, which has been designed for independent validation of land cover maps at a resolution finer than 100 m.

Tables 1 - 4 show the same patterns. Based on these tables, the following improvement possibilities for the CCI LC map at 20 m have been identified:

- Massive overestimation of shrub lands. Shrub lands are mapped with the lowest users and producers accuracies. The highest confusion is found to be between shrub lands and grasslands.
- Massive overestimation of croplands. Croplands are mapped with low users and producers accuracies. The highest confusion is between croplands and grasslands. Croplands are mainly overestimated in dry areas. Visually, the distribution of croplands at 20 m is very similar to the croplands on ESA CCI land cover maps at 300 m resolution (https://www.esa-landcover-cci.org/).
- Overestimation of tree cover. We have observed that the regions with highly fragmented landscapes due to shifting cultivations are mapped as 100 % forests. Users would expect that the CCI LC map at 20 m resolution should capture individual fields.
- Underestimation of grasslands. Grasslands are highly confused with shrubs, trees and crop lands.
- Underestimation of wetlands. This is a very difficult class to map from a remote sensing point of view. The CCI LC team may consider using an ancillary layer instead.
- Water objects are mapped with the highest accuracy and the urban layer is of reasonable quality, BUT both of these layers were implemented from auxiliary layers:
  (1) Global Water Surface (https://global-surface-water.appspot.com/) and
  (2) Global Urban Footprint
  (http://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-9628/16557_read-40454/).
  Moreover, the wrong Global Water Surface layer was implemented showing the maximum water extent of the last 30 years instead of the water seasonality for the year 2016.

  If two classes would be removed from the confusion matrices in order to show classification results based on only using Sentinel-2 data, the overall accuracies would drop to 57% and 64% in Tables 1 and 3, respectively.
- Bare lands are mapped well.

To conclude, the overall and per-class accuracies do not yet meet user requirements. The map therefore, should be improved.

2. Spatial assessment of overall accuracies

3.1. Methodology

To produce spatial overall accuracies, we have applied geographically weighted logistic regression (GWR) (Comber et al., 2012; Lesiv et al., 2016). As input reference data we used the reference dataset presented in Section 2.1.

Here, we refer to overall accuracy as a probability that a LC map is correct within a moving kernel window. GWR estimates the model parameters at each geographical location using a kernel. In addition, the observations are weighted by distance, so those closer to the studied location will have more influence on the parameter estimates (Brunsdon et al., 1998).

All calculations have been done in R (packages “raster”, “sp,” and “spgwr”).

3.2. Results

Figure 4 shows the spatially explicit overall accuracy (confidence interval 95%) of the CCI LC map at 20 m resolution.

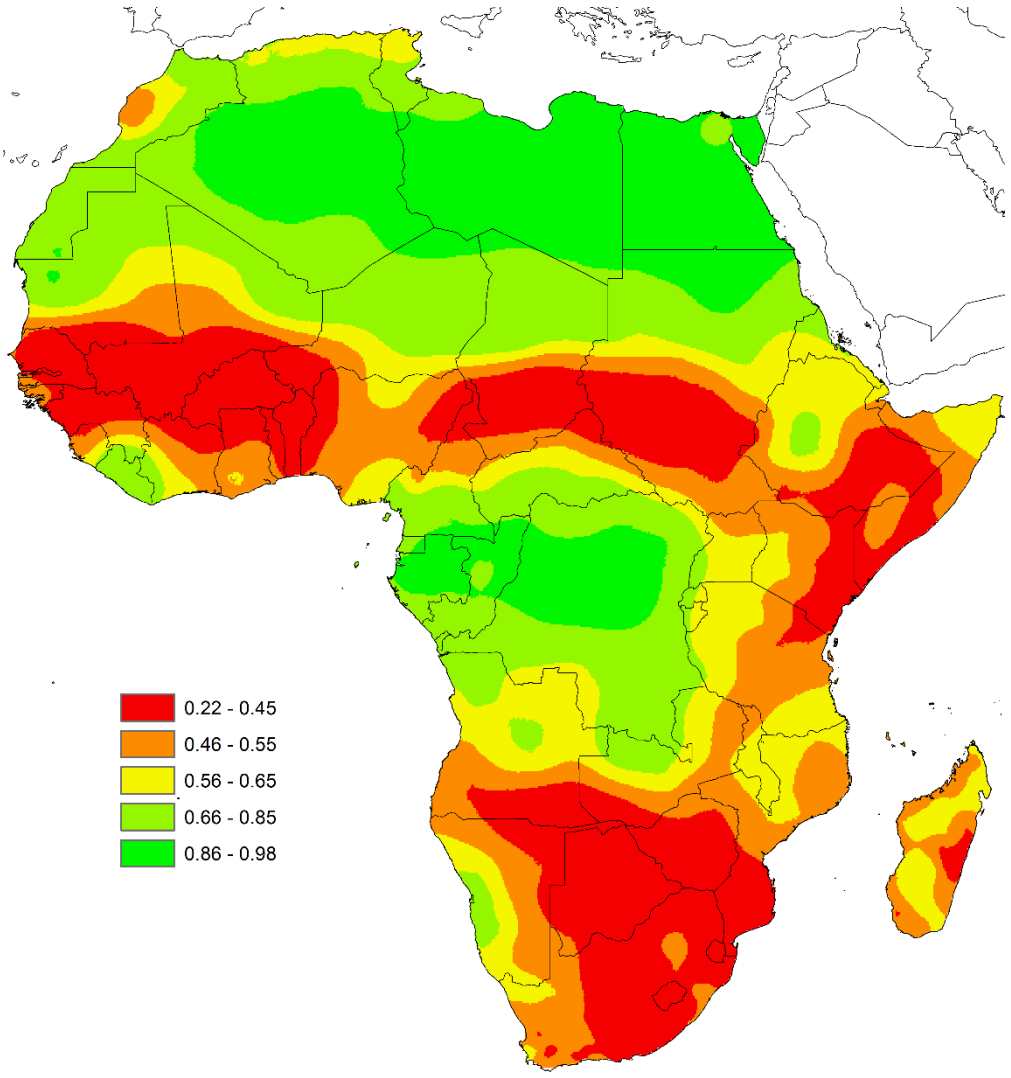


Figure 4: Spatially explicit overall accuracy of the CCI LC map at 20 m

The regions with the lowest overall accuracies are highlighted in red color. Those are mainly areas of high confusion between grasslands and shrubs, grasslands and croplands. The red spot in Madagascar shows where tree cover is overestimated, and it is a very fragmented landscape due to shifting cultivations. The orange spot in Morocco is where huge cropland areas are missing. The red spot at the border between Ethiopia and Somalia contains the areas where shrubs are mapped as grasslands, croplands are also overestimated along the sea coast of Somalia. Croplands are massively overestimated in Chad and Sudan. There is a high confusion between crops, shrubs and grasslands in Senegal, Mali, Burkina Faso and Niger. In Botswana and South Africa, shrubs are highly confused with grasslands and bare lands, and sparse trees are not mapped (it is savannas). In overall, the CCI LC map at 20 m does not show accuracy improvements in comparison with previous mapping efforts (Tsendbazar et al., 2015).

3.3. Summary of feedback

Africa is a continent with complex heterogeneous landscapes, which are challenging to map from a remote sensing point of view. By analyzing spatially explicit accuracies, we have identified both regions and classes that should be improved:

(1) Western Sahara region: wrong croplands in the middle of the desert, high confusion between shrub-grass-cropland;

(2) Chad and Sudan: massive overestimation of croplands

(3) Ethiopia and Somalia: high confusion between shrub lands and grassland.

(4) Transition areas from rainforest to savannas to shrub lands (e.g. Botswana and South Africa): shrubs are overestimated;

(5) Regions with shifting cultivations (e.g. Madagascar): overestimated tree cover;

(6) Morocco and Algeria: missing croplands, natural vegetation (trees, shrubs, grass) is not well mapped, shrubs are overestimated.

3. Overall comments based on visual inspection

We would like to highlight three major visual observations:

1. Significant spatial inconsistencies likely related to production tiling, see Figure 5.

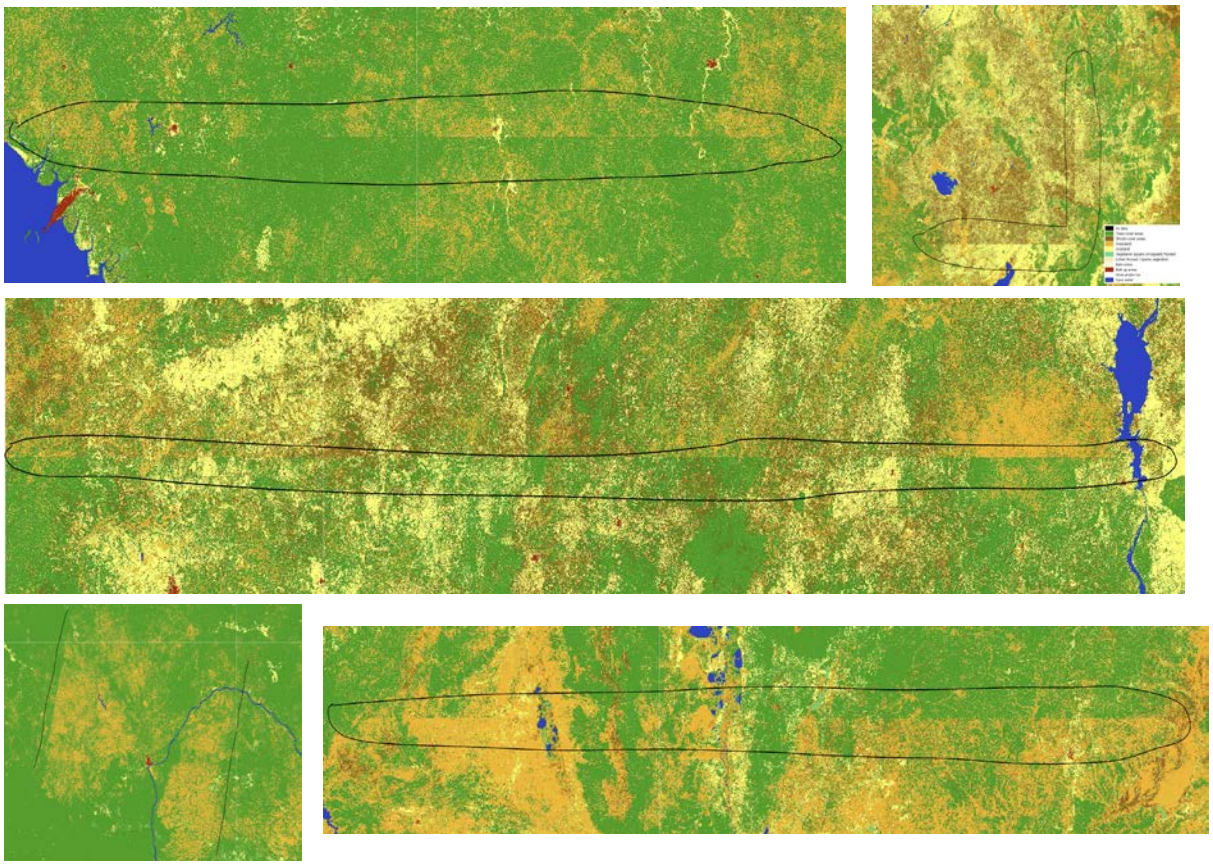
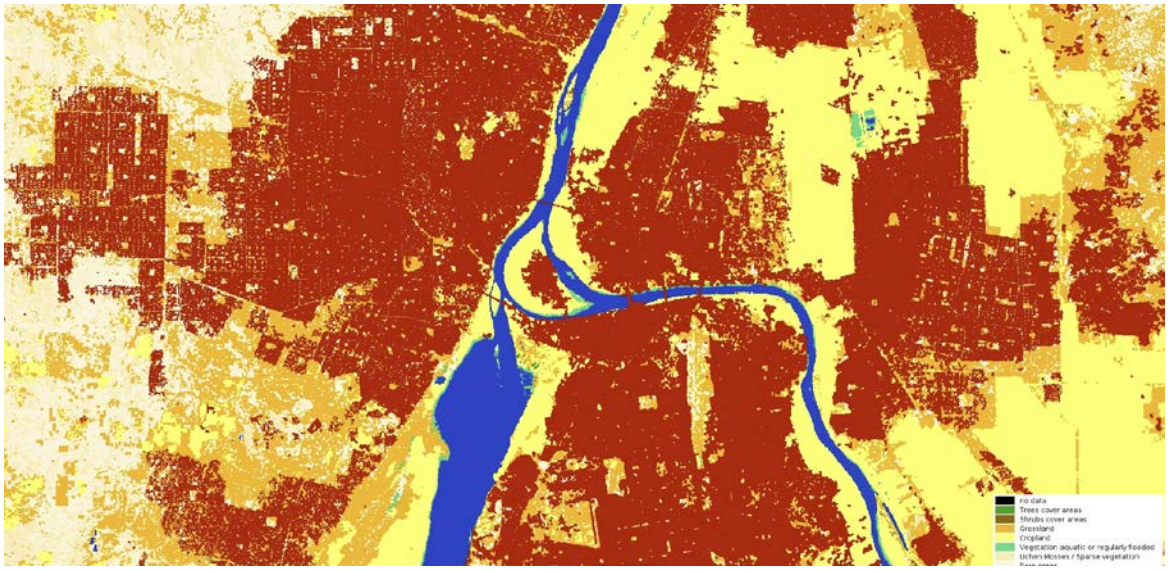


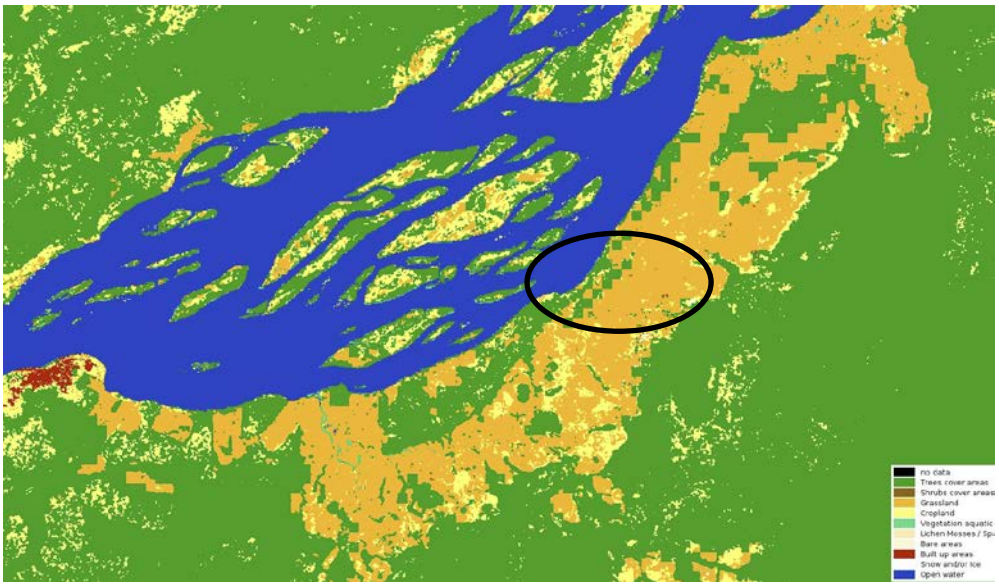
Figure 5: Spatial inconsistencies likely related to a production tiling

2. Inappropriate representativeness of the product derived from 20 m spatial resolution images. The first impression of users is that the CCI LC map does not actually correspond to a 20 m resolution, see Figure 6. The built up areas are largely identified thanks to the Global Human Settlement and the Global Urban Footprint. The open water is largely identified thanks to the Global Surface Water explorer. The high-resolution nature of both classes is obvious (moreover they match each other, Figure 6A). But for the other classes, large pixels are sometimes visible, which gives the feeling that the product is a mix of medium and high spatial resolution, see Figure 6B and 6C.

A



B



C

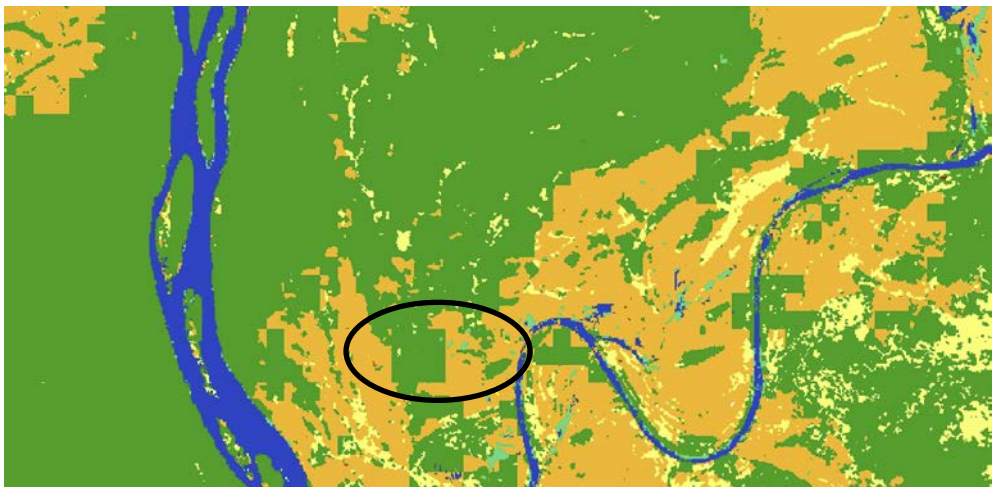
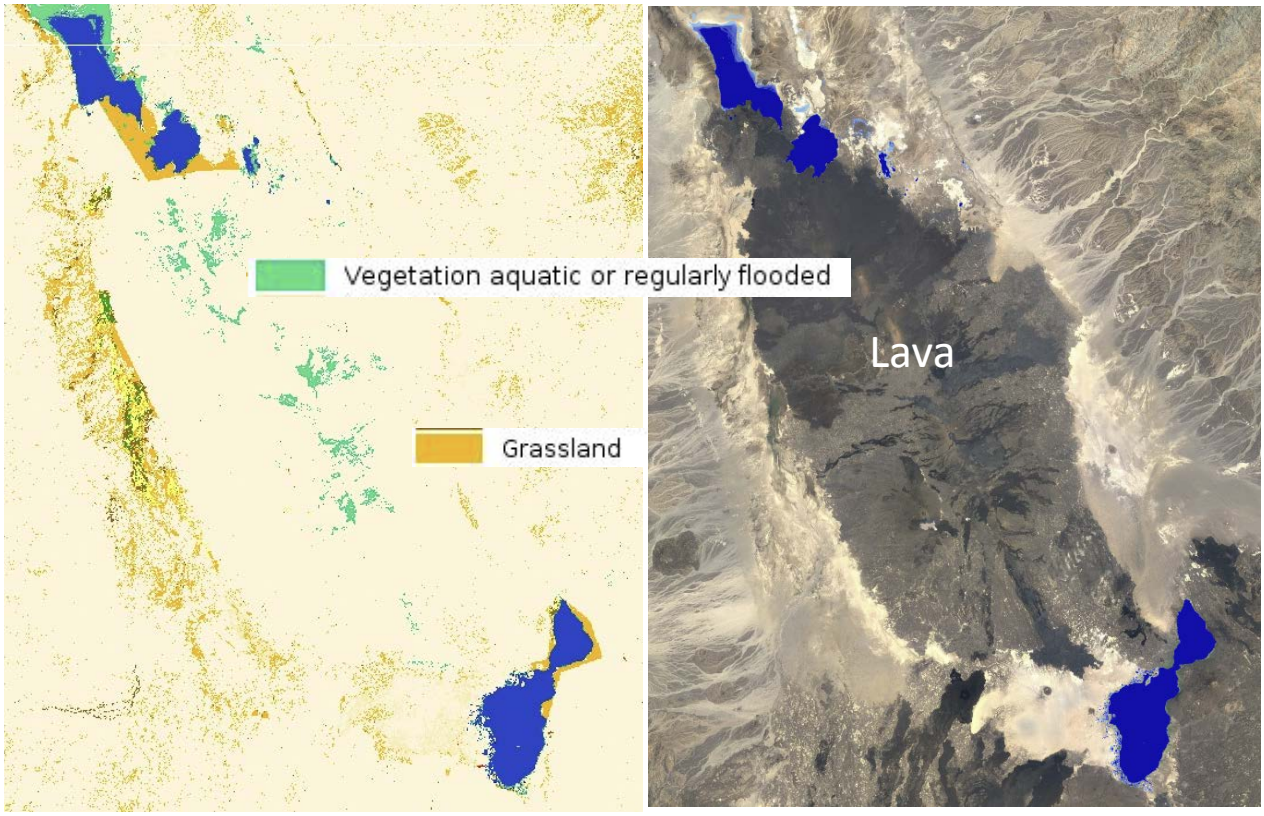


Figure 6: Examples of inappropriate representativeness of the product derived from 20m spatial resolution imagery

3. Significant classification errors from a thematic point of view error. Figure 7A shows that grassland and “Vegetation aquatic or regularly flooded” are partially confused with lava flow; Figure 7B shows cropland partially confused with “Trees cover” areas. More examples could be found in the Annex.

A



B

Figure 7: Examples of classifications errors

Final remark

We think that the feedback provided in this report could be considered by the CCI LC team to improve the African CCI LC 20 m map from the current estimated overall accuracy of 64% derived from an independent validation according well defined protocols. Furthermore, the CCI prototype LC map at 20 m could potentially be improved methodologically to remove a number of visual artefacts. In addition, the CCI LC team may consider investing into high quality training data at 20 m resolution.

In producing this prototype 20 m African LC product the CCI team has processed 180,000 Sentinel-2A images representing 90 terrabytes of data. This demonstrates that recent technological developments now allow for the processing of large amounts of remote sensing data at continental and global levels at high spatial resolution. However, the challenge still remains in satisfying user needs and producing highly accurate maps, with accuracies per LC class bigger than 85%.

References

Brunsdon, C., Fotheringham, S., and Charlton, M. (1998). Geographically Weighted Regression-Modelling Spatial Non-Stationarity. J. R. Stat. Soc. Ser. Stat. *47*, 431–443.

Comber, A., Fisher, P., Brunsdon, C., and Khmag, A. (2012). Spatial analysis of remote sensing image classification accuracy. Remote Sens. Environ. *127*, 237–246.

Fritz, S., Bartholomé, E., Belward, A., Hartley, A., Stibig, H.-J., Eva, H., and Mayaux, P. (2003). Harmonisation, mosaicing and production of the Global Land Cover 2000 database (Beta Version) (Luxembourg: Office for Official Publications of the European Communities).

Jun, C., Ban, Y., and Li, S. (2014). China: Open access to Earth land-cover map. Nature *514*, 434–434.

Lesiv, M., Moltchanova, E., Schepaschenko, D., See, L., Shvidenko, A., Comber, A., and Fritz, S. (2016). Comparison of Data Fusion Methods Using Crowdsourced Data in Creating a Hybrid Forest Cover Map. Remote Sens. *8*.

Lesiv, M., Fritz, S., and See, L. (2017). A New Open Reference Global Dataset for Land Cover Mapping at a 100m Resolution (Frascati, Italy).

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., and Wulder, M.A. (2014). Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. *148*, 42–57.

Tsendbazar, N., Herold, M., Lesiv, M., and Fritz, S. (2017a). Validation Report of Moderate Dynamic Land Cover ,Collection 100M,Version 1 (WUR).

Tsendbazar, N.-E., de Bruin, S., Fritz, S., and Herold, M. (2015). Spatial Accuracy Assessment and Integration of Global Land Cover Datasets. Remote Sens. *7*.

Tsendbazar, N.-E., de Bruin, S., and Herold, M. (2017b). Integrating global land cover datasets for deriving user-specific maps. Int. J. Digit. Earth *10*, 219–237.

Annex: Visual validation of the ESA CCI Land Cover map at 20 m in Geo-Wiki

We have uploaded the CCI LC map at 20 m into the Geo-Wiki Land Cover page (geo-wiki.org). Users can have a look at the map that's overlaid on top of very high resolution Google and Bing imagery.

We have performed visual validation of selected plots across Africa to support the summary of feedbacks (section 3.3). The annex lists the screenshots taken in different locations.

Figure 8 below shows the spatial distribution of the taken screenshots.

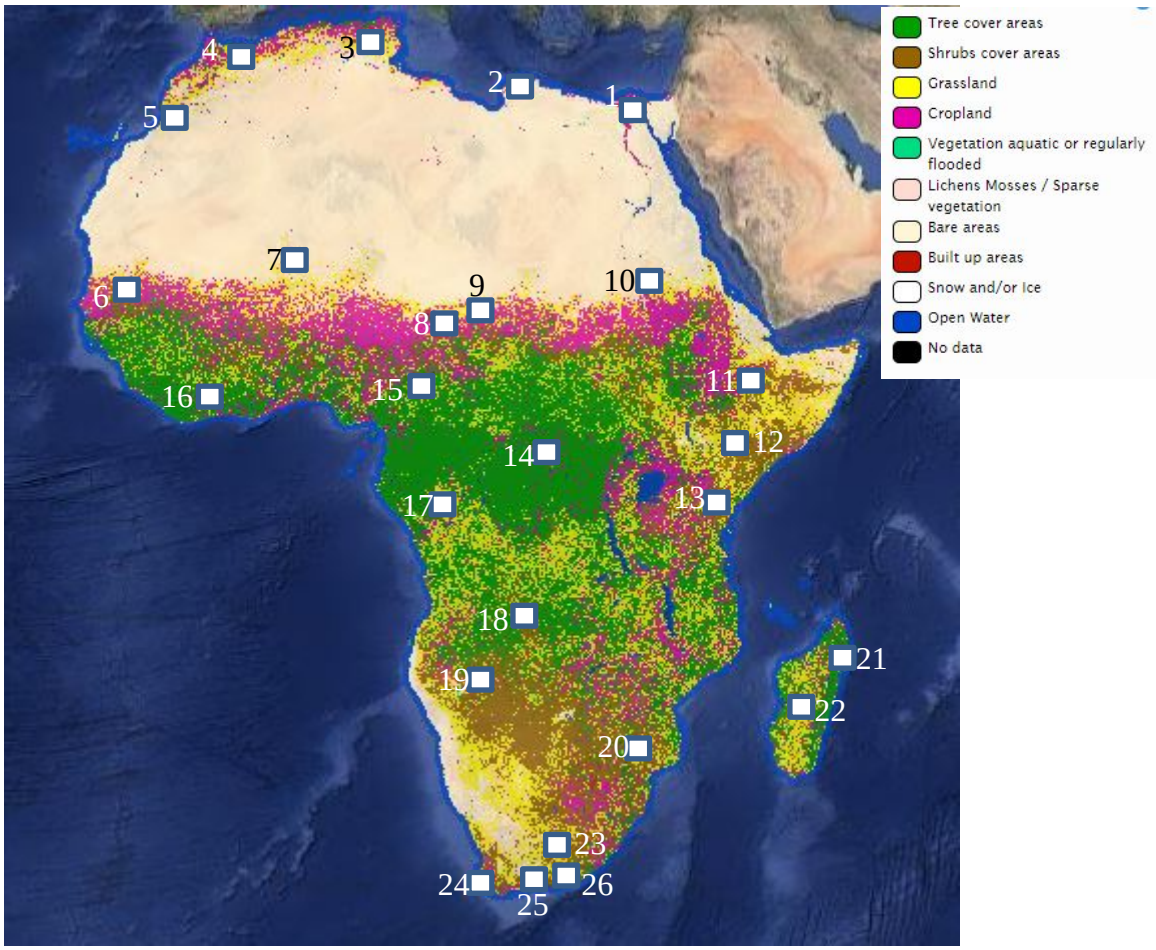


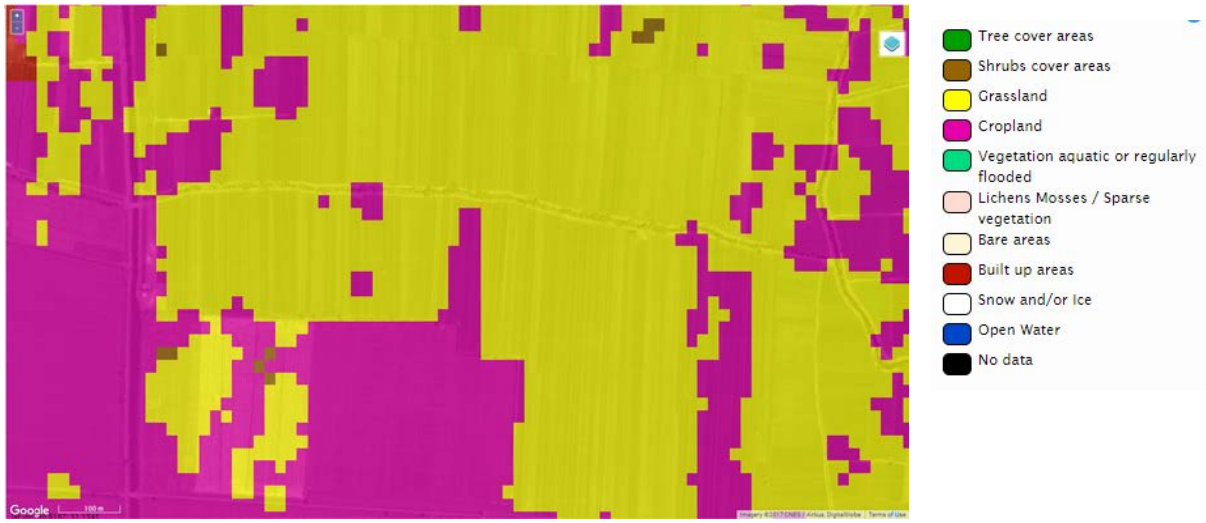
Figure 8: Spatial distribution of plots for visual validation of the CCI 20 m map

Example 1: Nile Delta

Coordinates in Lat/Lon: 30.8187, 31.1348

Croplands are confused with grassland. According to the images, available in Google Earth, these are long-shaped cropland fields. In Google Earth, for this area, there are images from 2015, 2016 and 2017, which confirms that it is all cropland.

The CCI LC map at 20 m:



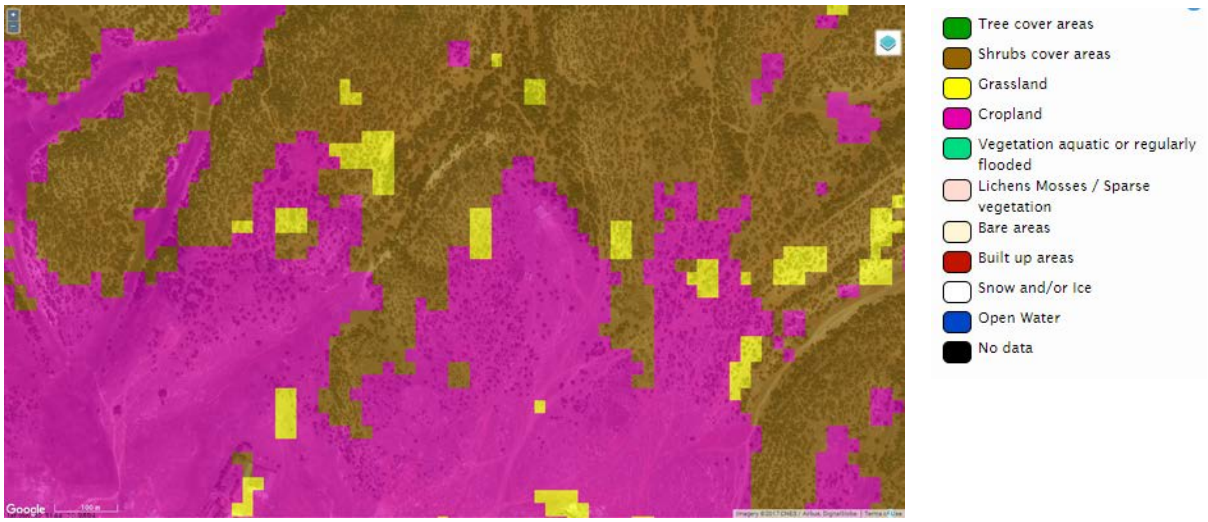
Corresponding Google image:

Example 2: Al Jabal Akhdar, Libya

Coordinates in Lat/Lon: 32.3144, 20.9894

Croplands are overestimated. Google Earth image is very recent here (2016), it confirms that these are grasslands. In addition, there are Panoramio pictures (http://www.panoramio.com/photo/96224522?source=wapi&referrer=kh.google.com) confirming presents of trees.

The CCI LC map at 20 m:



Corresponding Google image:

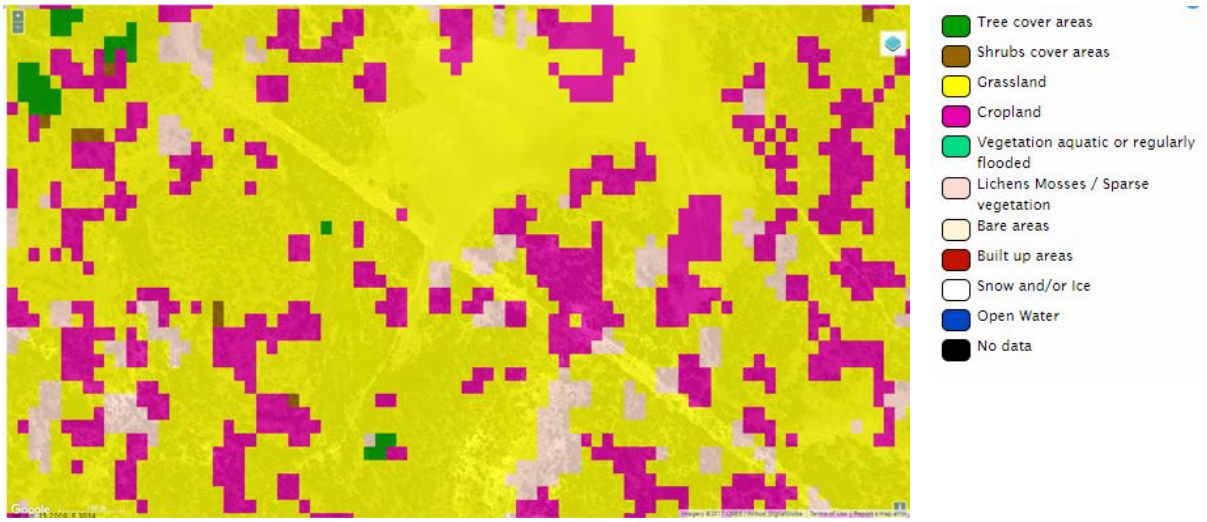
Example 3: North of Algeria

Coordinates in Lat/Lon: 35.2009, 8.3034

Grasslands are overestimated. Trees and shrubs are considerably underestimated. The Google Earth image is from 2016.

The dark green spots are trees, smaller spots are shrubs, but not grassland. Croplands “spots” are not croplands. There is a field, which could be cropland, but it is anyway classified as grassland.

The CCI LC map at 20 m:



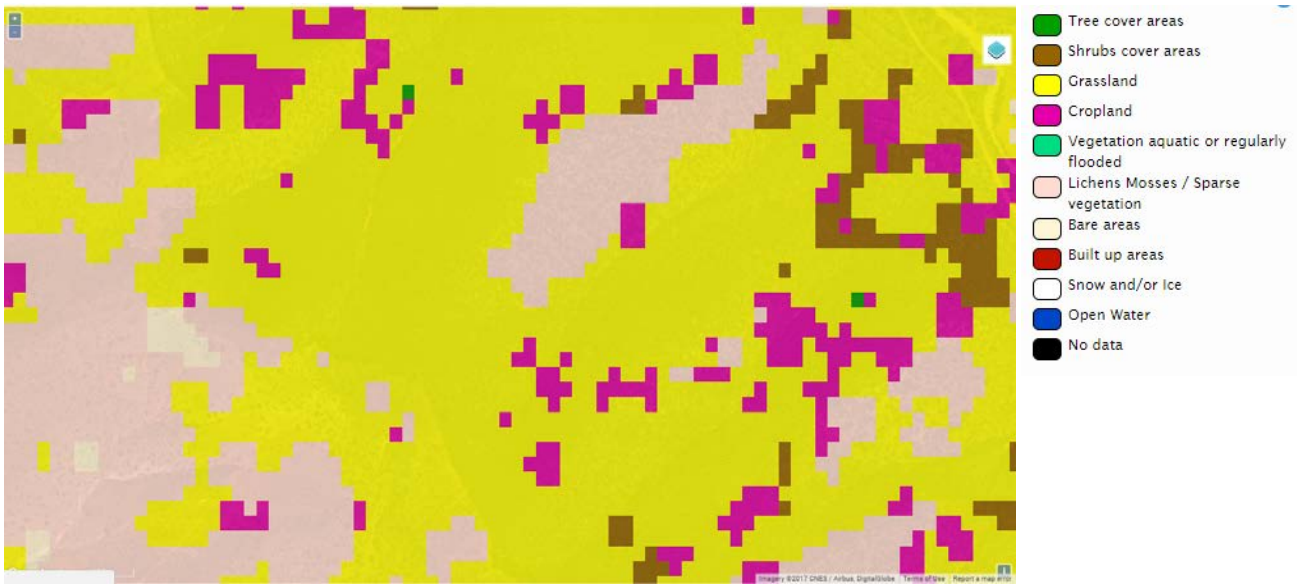
Corresponding Google image:

Example 4: North of Morocco

Coordinates in Lat/Lon: 35.0973, -4.2885

Grasslands are overestimated. The small green-brown spots are shrubs but not grassland. Google Earth image is from 2016.

The CCI LC map at 20 m:



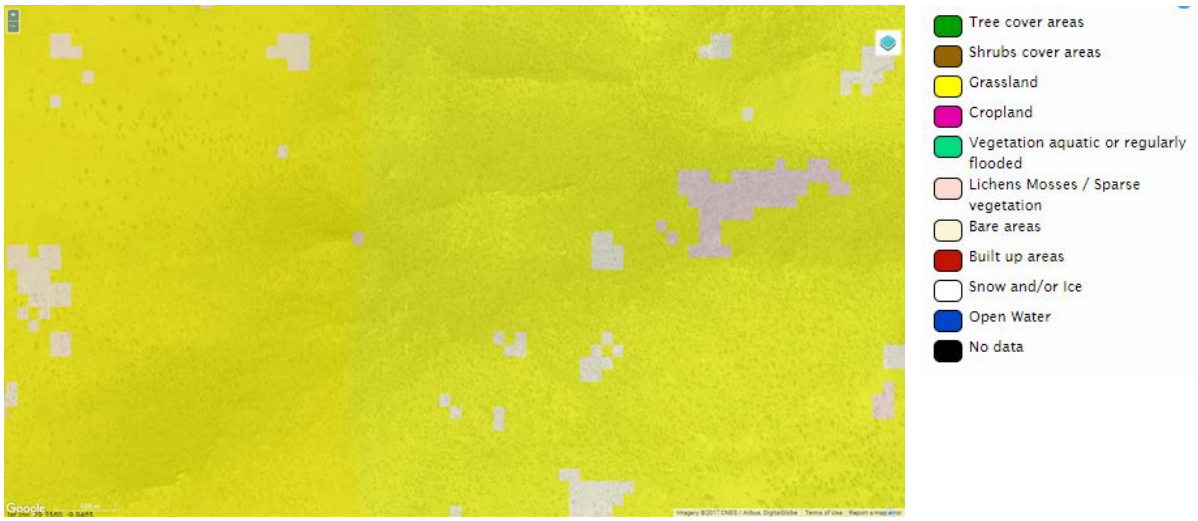
Corresponding Google image:

Example 5: North of Morocco

Coordinates in Lat/Lon: 29.3560,-9.9465

In the Google Earth image below, those are very small shrubs. The image is from 2016.

The CCI LC map at 20 m:



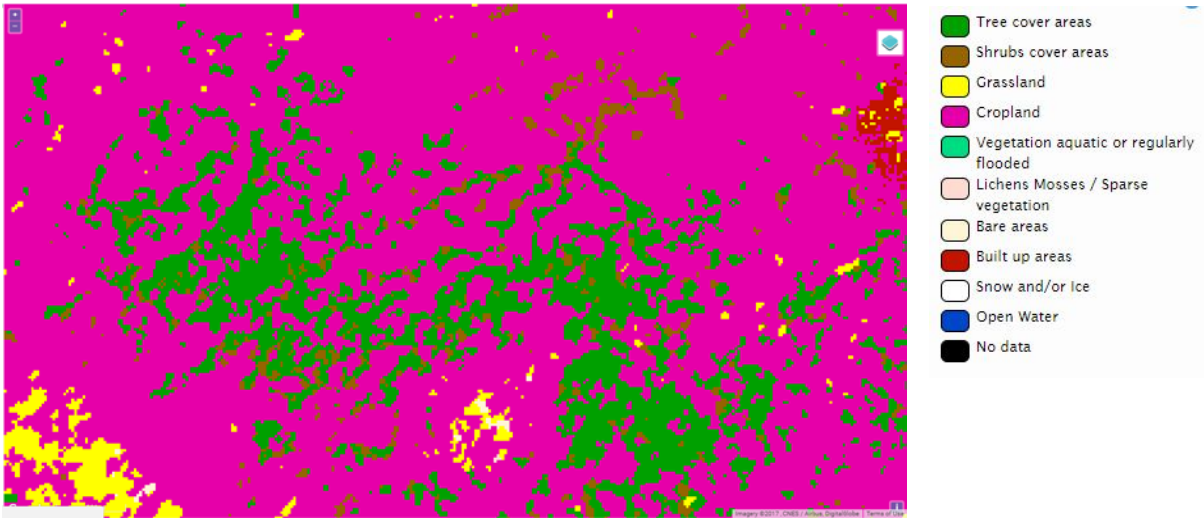
Corresponding Google image:

Example 6: Senegal

Coordinates in Lat/Lon: 15.3815,-13.1844

As it has been detected earlier on Figure 4, croplands are overestimated in this region. The Google Earth high resolution imagery confirms it is a mix of small shrubs and grassland.

The CCI LC map at 20 m:



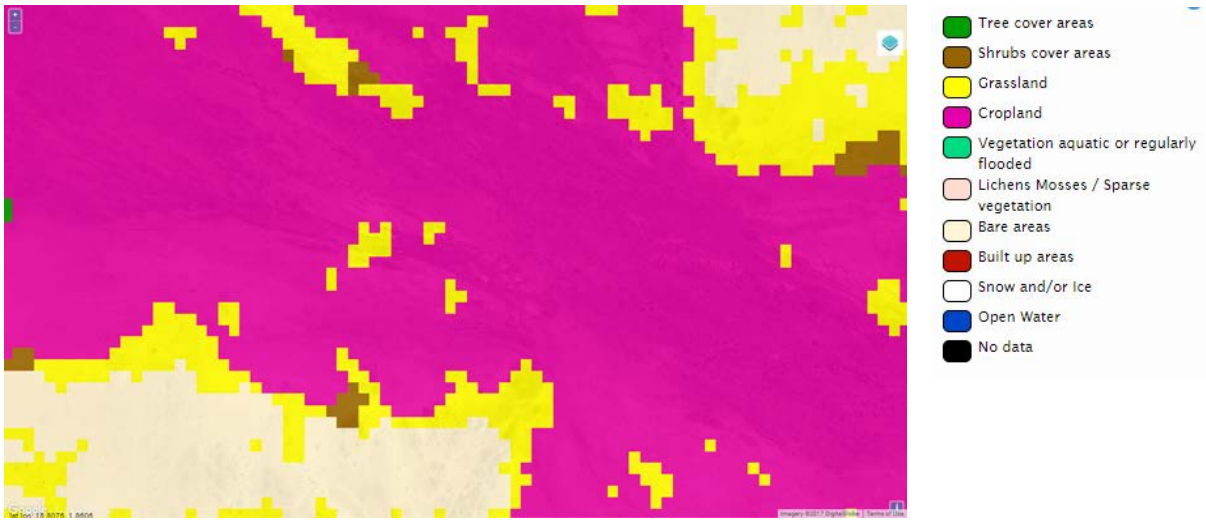
Corresponding Google image:

Example 7: Mali

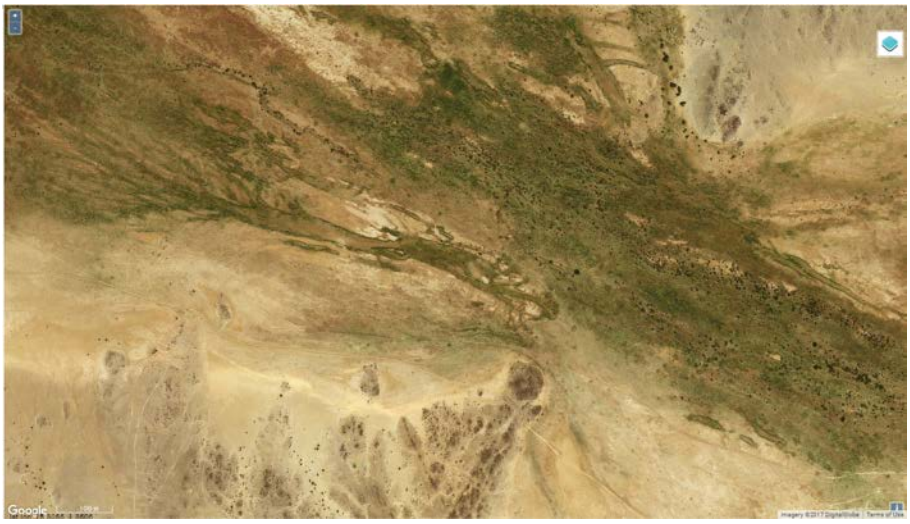
Coordinates in Lat/Lon: 18.8076, 1.8606

That's a very dry region. Croplands are overestimated due to the confusion with natural vegetation, e.g. shrubs. This region is also highlighted on Figure3.

The CCI LC map at 20 m:



Corresponding Bing image:

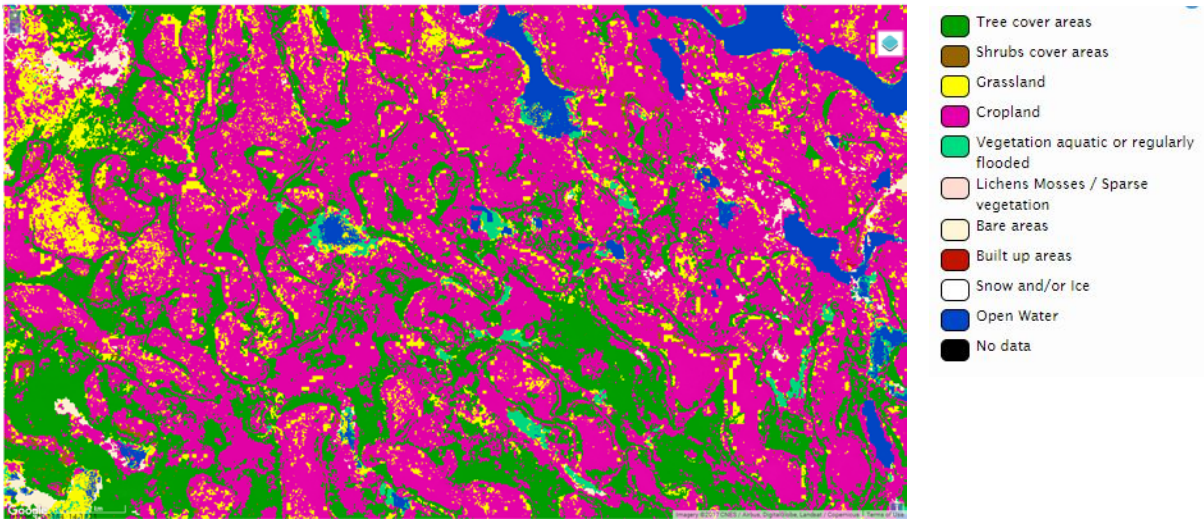


Example 8: Lake Chad

Coordinates in Lat/Lon: 13.3034, 14.3006

Trees and croplands are overestimated due to confusion with temporary flooded objects. The Bing image below shows that in these areas there are a lot of herbaceous wetlands, grasslands and some crops. The CCI LC map at 20 m identify these areas as tree cover and croplands.

The CCI LC map at 20 m:



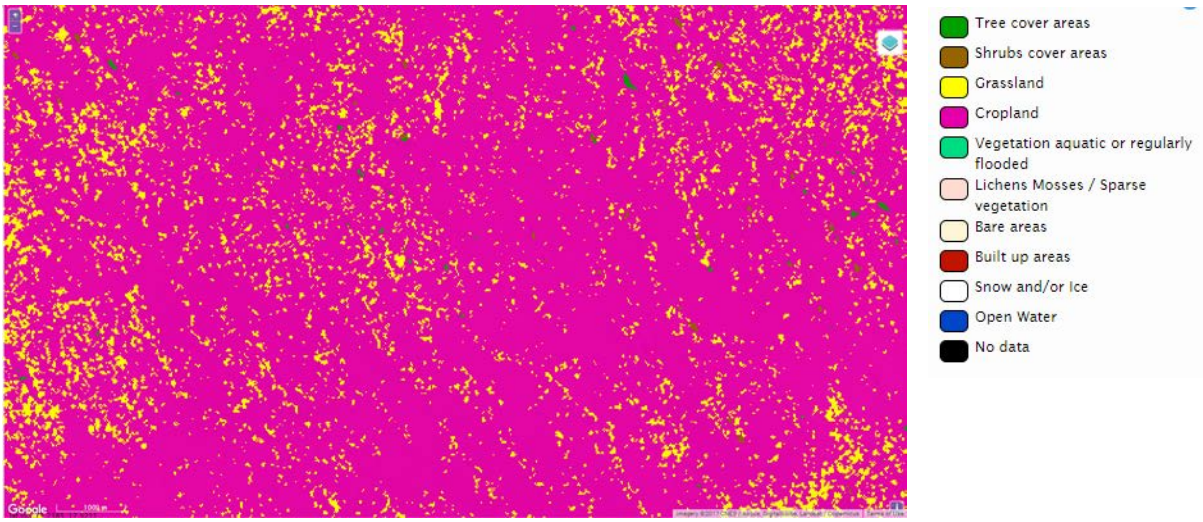
Corresponding Bing image:

Example 9: Chad

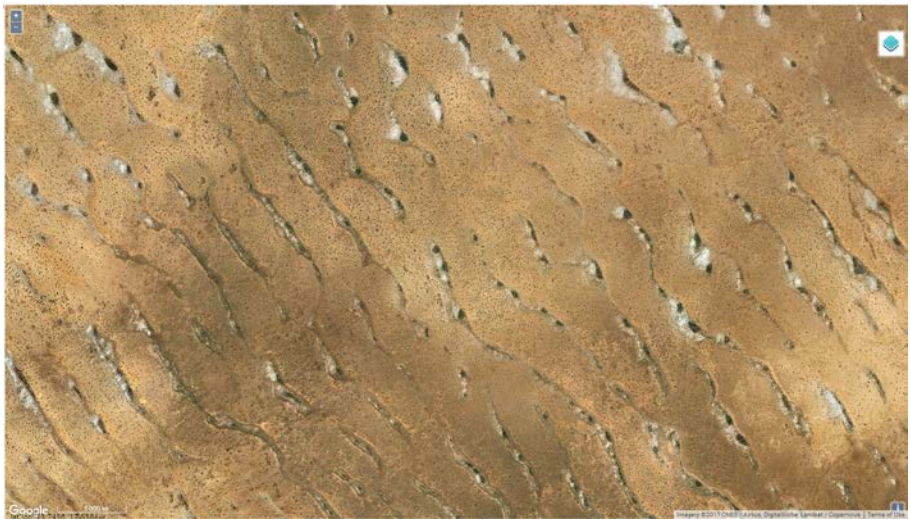
Coordinates in Lat/Lon: 13.7383, 17.5215

This is a very dry area, and as it has been shown on Figure 4 croplands are overestimated. The Google Earth high resolution imagery confirms this.

The CCI LC map at 20 m:



Corresponding Google image:

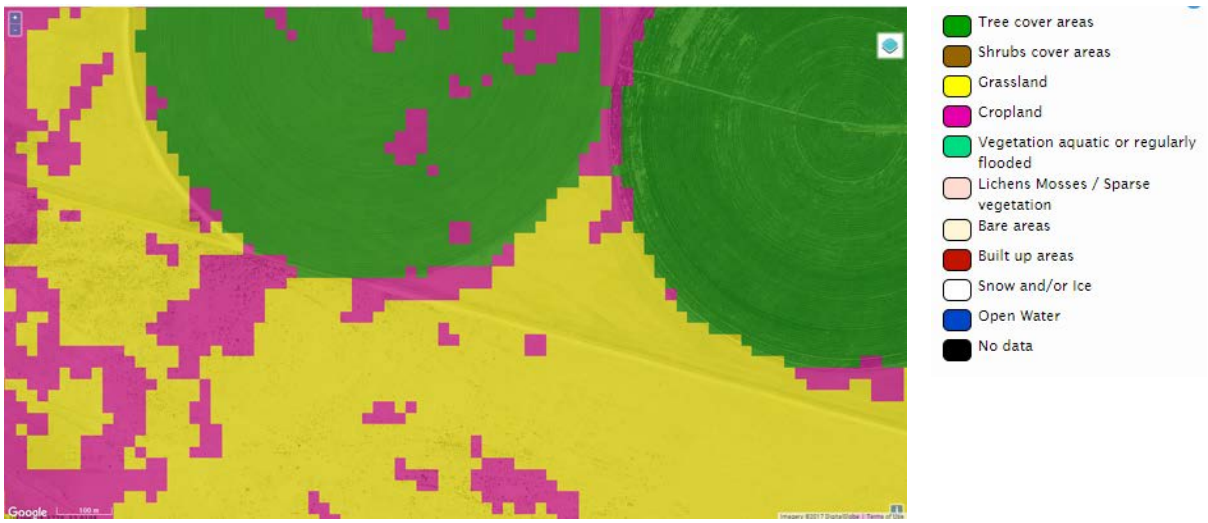


Example 10: Sudan

Coordinates in Lat/Lon: 15.3778, 32.9195

One of the widely occurring errors that happen during image classification (not only on the CCI LC map at 20 m) is misclassified irrigated croplands due to the confusion with forest. It could be corrected by adding more training data.

The CCI LC map at 20 m:



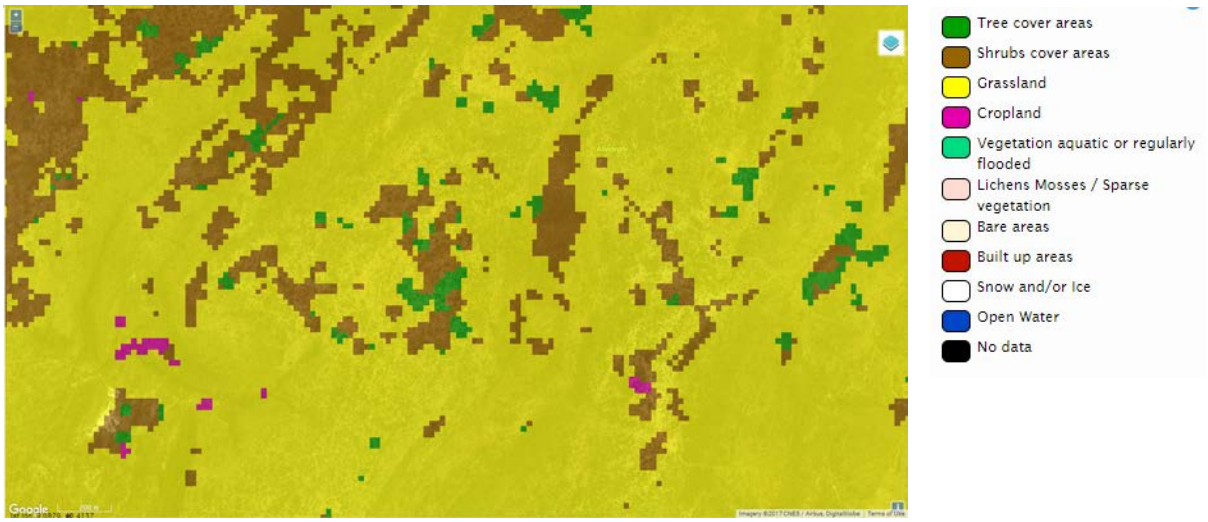
Corresponding Google image:

Example 11: Ethiopia

Coordinates in Lat/Lon: 9.0879, 40.4137

In many places in Ethiopia, shrubs are confused with grasslands. In the example below, there is an area where shrubs have wide crowns and low height. See Panoramio pictures (http://www.panoramio.com/photo/49164200?source=wapi&referrer=kh.google.com#). The image in Google Earth is from 2017.

The CCI LC map at 20 m:



Corresponding Google image:

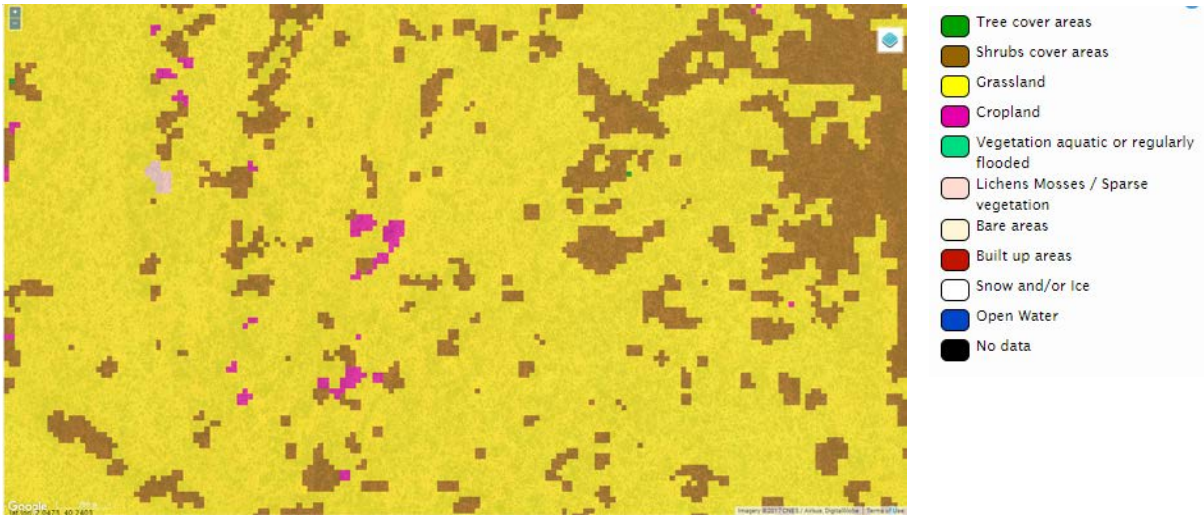
Example 12: Somalia

Coordinates in Lat/Lon: 2.0473, 40.2403

In Somalia, shrub lands are very often confused with grassland. See also explanation to Example 13.

The image in Google Earth is from 2015

The CCI LC map at 20 m:



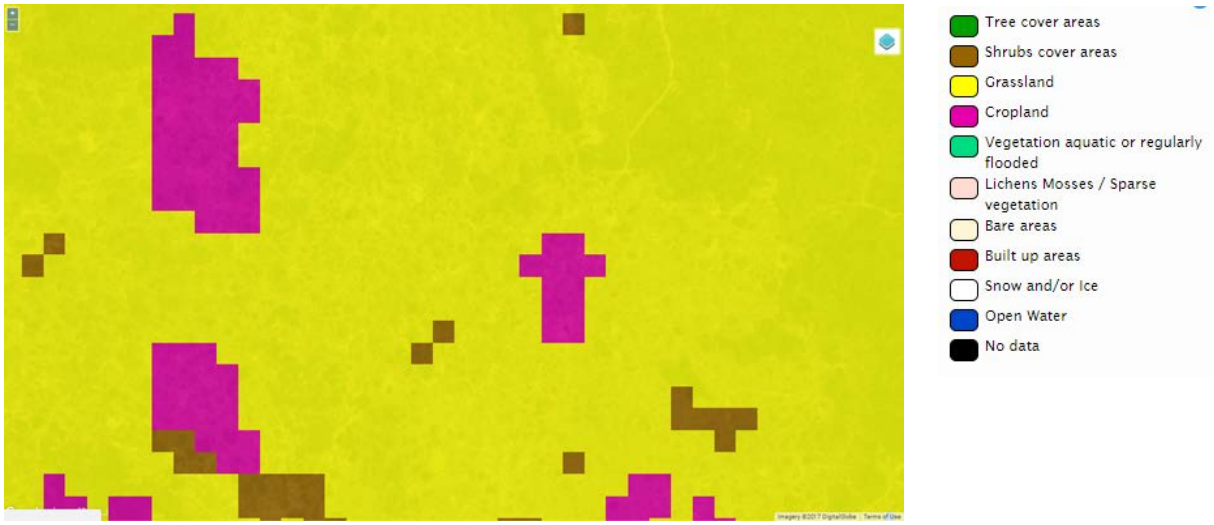
Corresponding Google image:

Example 13: Kenya

Coordinates in Lat/Lon: -3.3819,39.6481

This is an example from Kenya, where shrubs and trees are classified as grassland. Croplands are wrong too, as it can be seen from the Google Earth image. However, the newest image in Google Earth is from 2012.

The CCI LC map at 20 m:



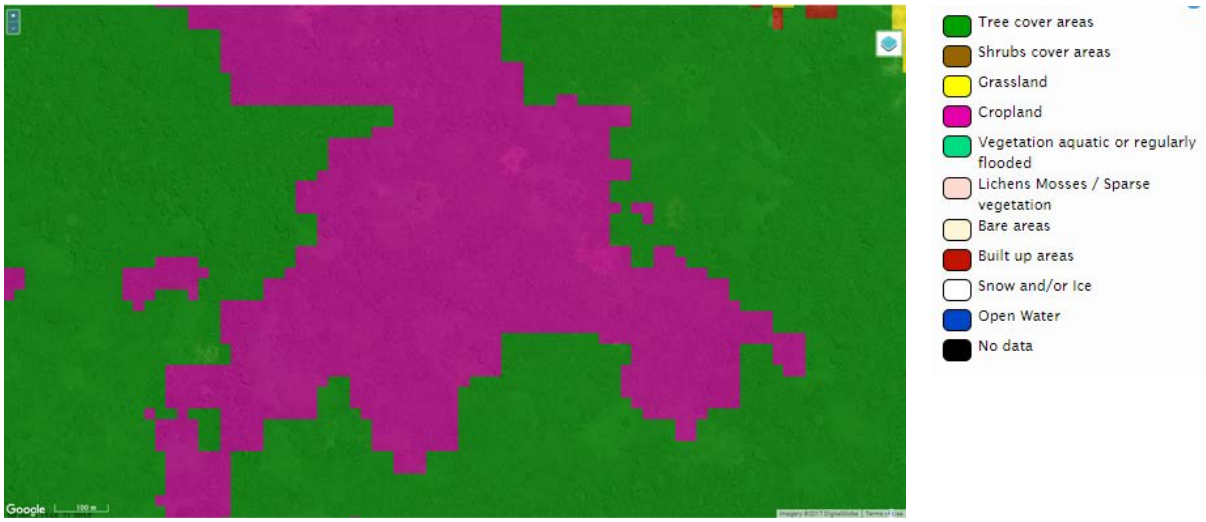
Corresponding Google image:

Example 14: Democratic Republic of Congo (DR of Congo)

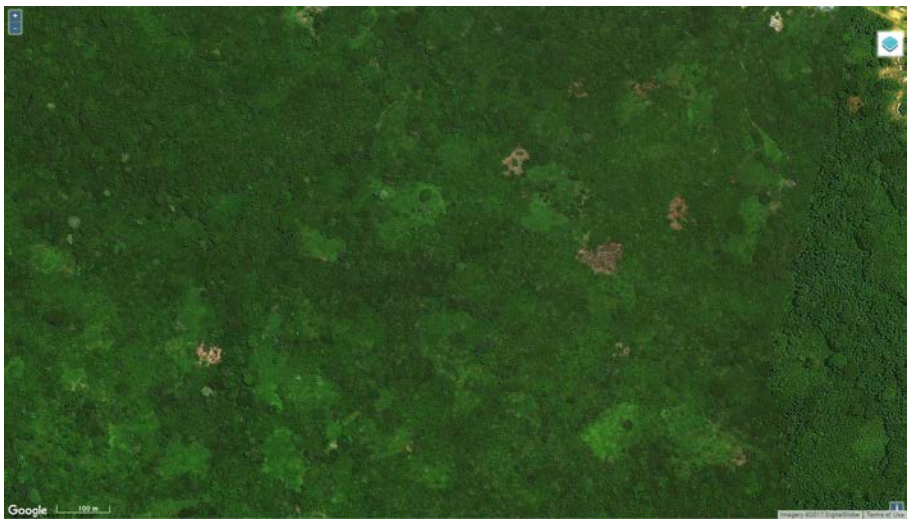
Coordinates in Lat/Lon: 0.5598, 21.0058

During the last years, cropland expansion has followed deforestation activities in the DR of Congo. In the example below, those are cropland fields, which are also difficult to recognize visually. The CCI LC map captures the cropland area but at the same time overestimates potential croplands. Users would expect that LC map at 20 m should better delineate individual fields.

The CCI LC map at 20 m:



Corresponding Google image:

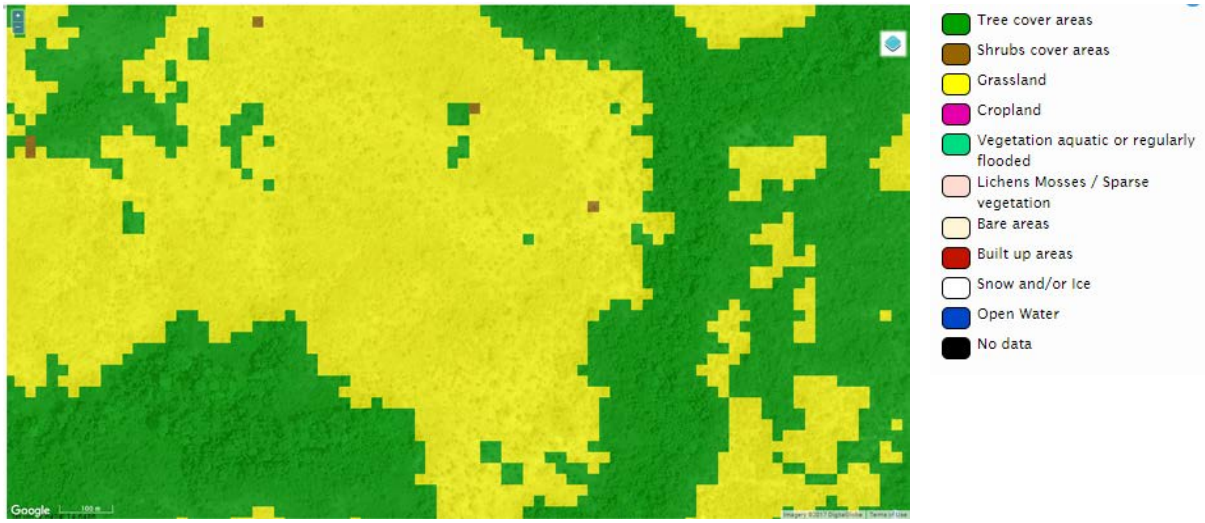


Example 15: Cameroon

Coordinates in Lat/Lon: 4.6424, 14.6486

The CCI LC map at 20m captures very well riparian vegetation, which is evergreen. However, in dryer areas it does not map very well deciduous shrubs as shown in this example.

The CCI LC map at 20 m:



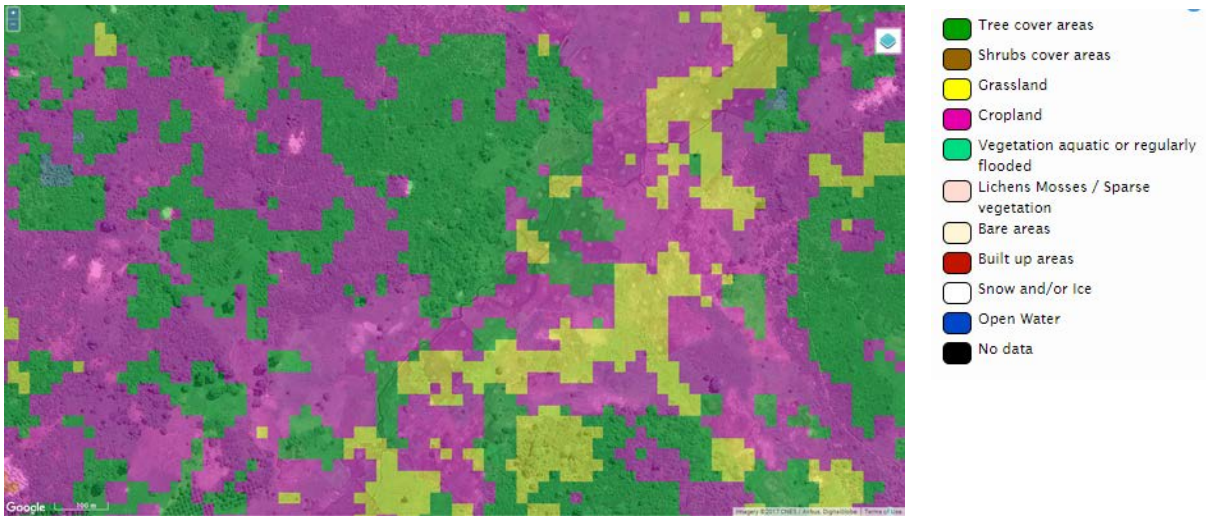
Corresponding Google image:

Example 16: Cote d'Ivoire

Coordinates in Lat/Lon: 6.5166,-5.8197

The landscapes are very fragmented in this country. There are no clear error patterns: sometimes croplands are classified as grassland or trees, trees are classified as cropland or grassland. One of the reasons is that there is not enough spectral data due to very frequent clouds.

The CCI LC map at 20 m:



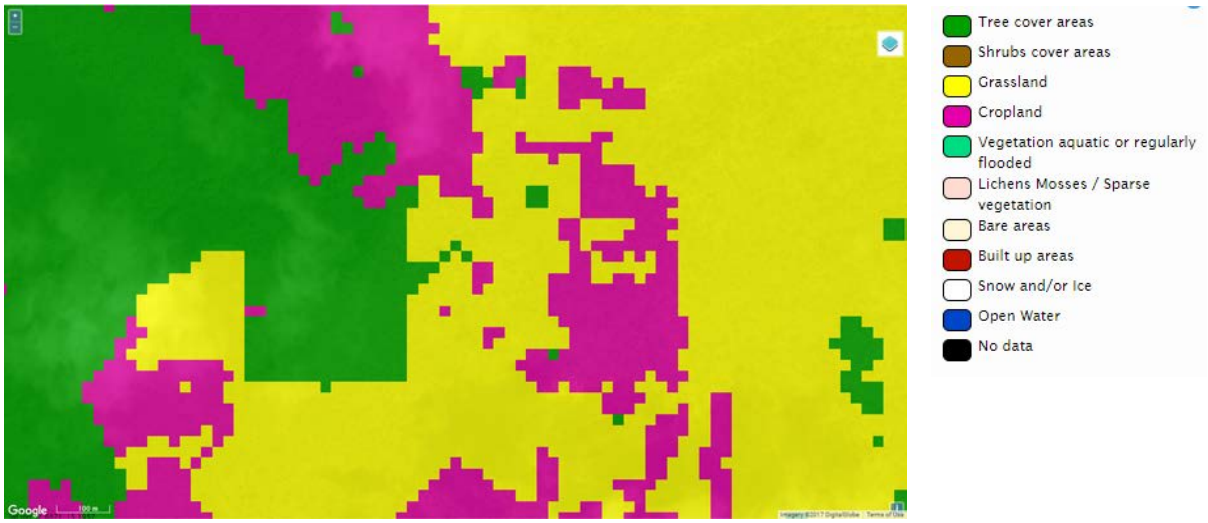
Corresponding Google image:

Example 17: Congo

Coordinates in Lat/Lon: -2.8871, 15.1957

This is another example with very strange forest boundary and wrongly classified shrubs. The image in Google Earth is from 2015, but forest could not regrow in one year. It is possible that this error is related to cloud masking.

The CCI LC map at 20 m:



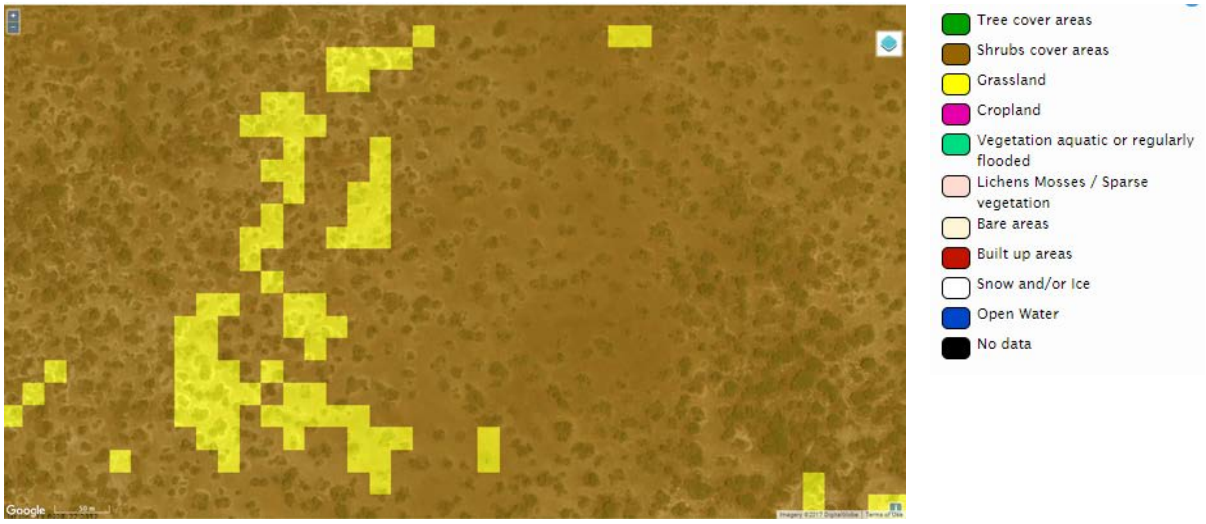
Corresponding Google image:

Example 18: Zambia

Coordinates in Lat/Lon: -13.6928, 22.2332

In this region, there are a lot of shrub lands, in particular, sparse shrub lands. In the example, in coarser resolutions, it is shrub lands. However, users would expect better delineation of small grasslands at 20 m resolution.

The CCI LC map at 20 m:



Corresponding Google image:

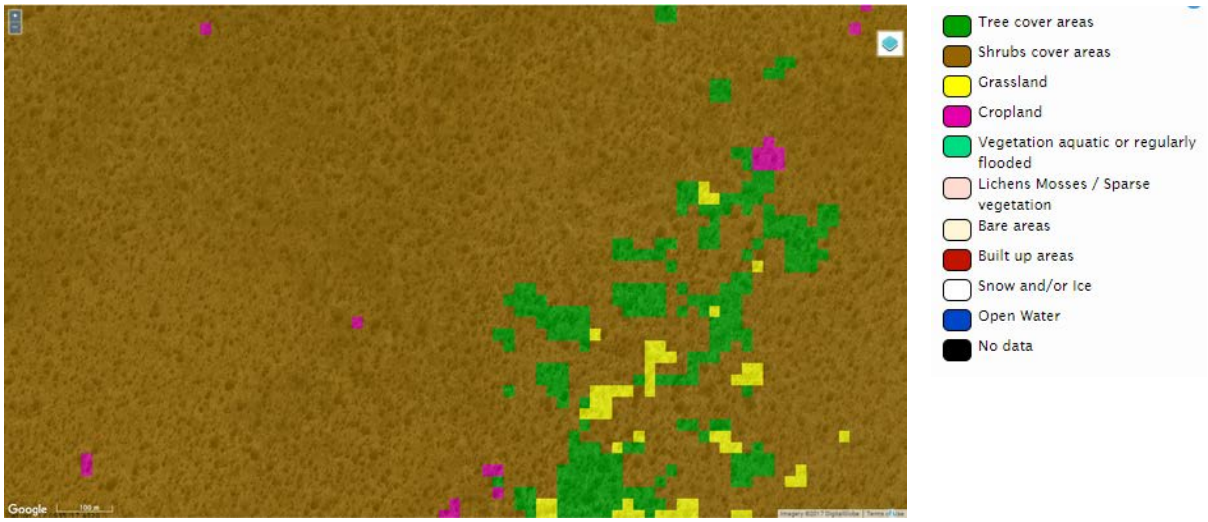
Example 19: Namibia

Coordinates in Lat/Lon: -17.0535,17.4320

This is savanna. In the image from Google Earth, dark spots with wider crowns are actually trees, below the trees are shrubs. Some areas with more dense trees should have been mapped as tree cover.

It is a question for discussion if savannas should not be mapped as a separate land cover class, or as open forest but then the resolution should be coarser than 20 m.

The CCI LC map at 20 m:



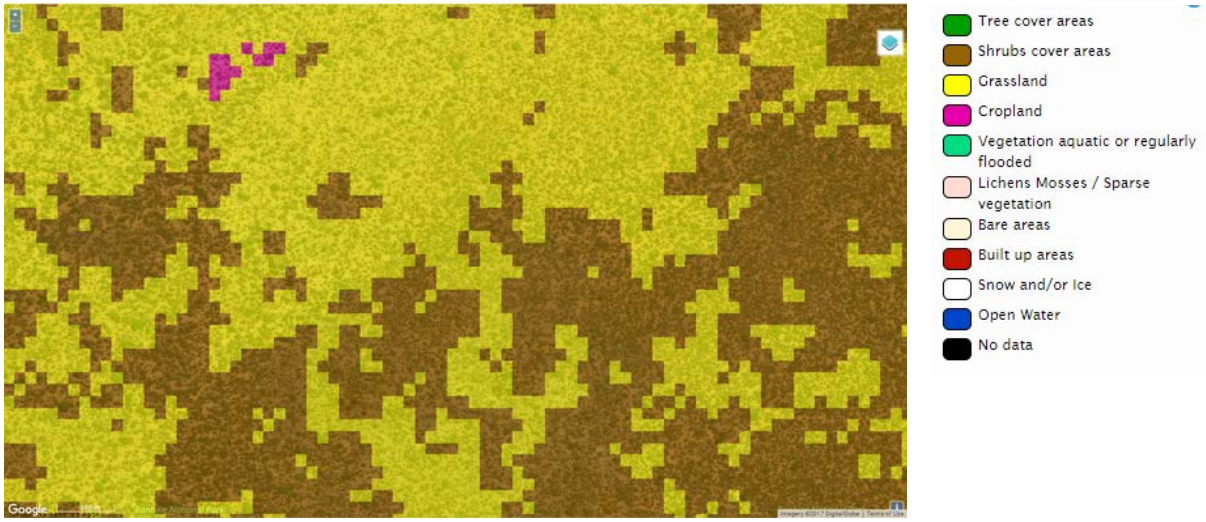
Corresponding Google image:

Example 20: Zimbabwe

Coordinates in Lat/Lon: -22.8883, 32.6396

The Google Earth imagery shows this area is covered by shrubs (small brown and dark green spots). However, on the CCI LC map at 20m, shrubs are highly confused with grasslands.

The CCI LC map at 20 m:



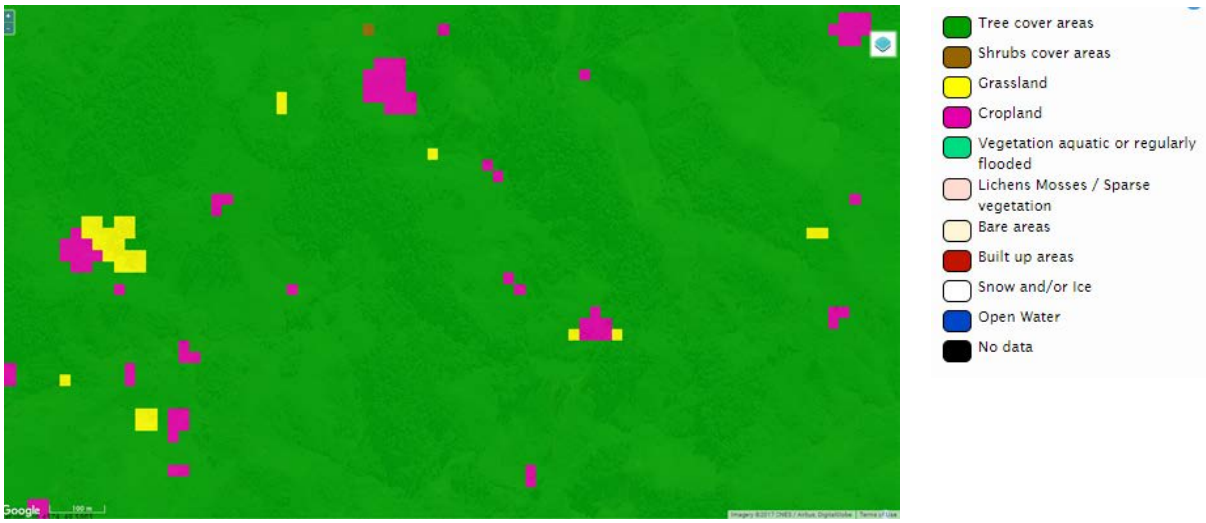
Corresponding Google image:

Example 21: Madagascar

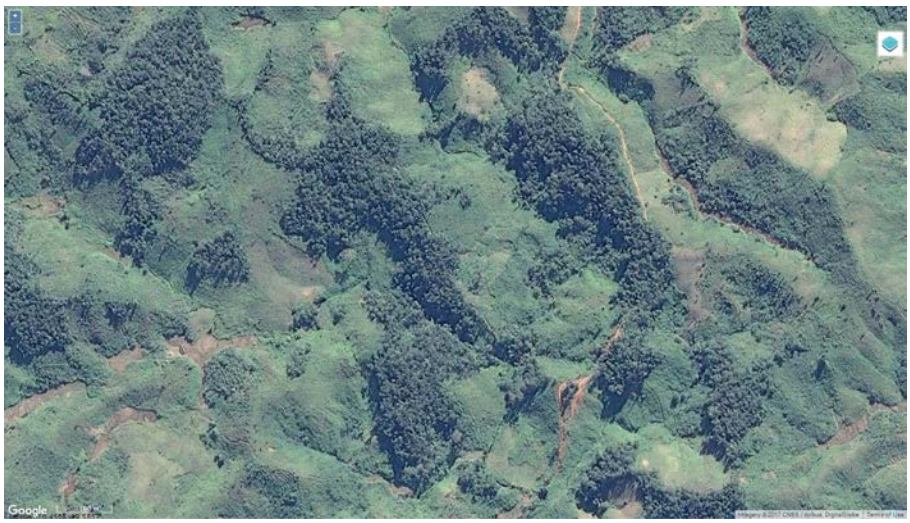
Coordinates in Lat/Lon: -16.4574, 49.5863

Figure 3 has shown that the lowest accuracies on Madagascar are observed along the Eastern coast. Those are areas of very intense shifting cultivations. The CCI LC map at 20 m identifies everything as forest cover, which is wrong. From the figures below, 20 m resolution is more than enough to capture individual fields.

The CCI LC map at 20 m:



Corresponding Google image:

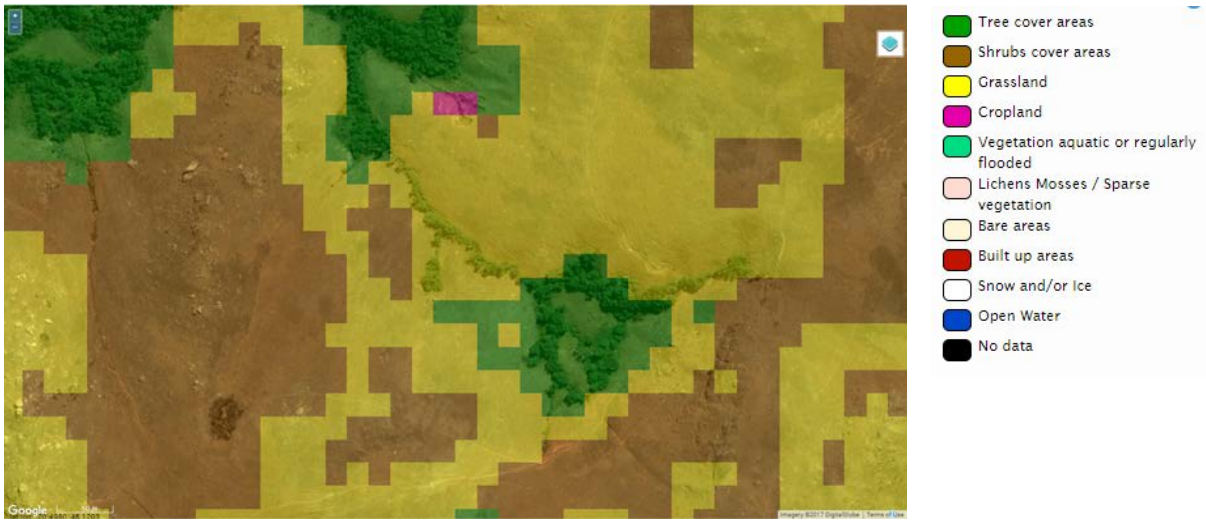


Example 22: Madagascar

Coordinates in Lat/Lon: -20.4980, 46.1203

In overall, there are a lot of pure grasslands on Madagascar. However, in many places they are confused with shrubs, as shown in the example below.

The CCI LC map at 20 m:



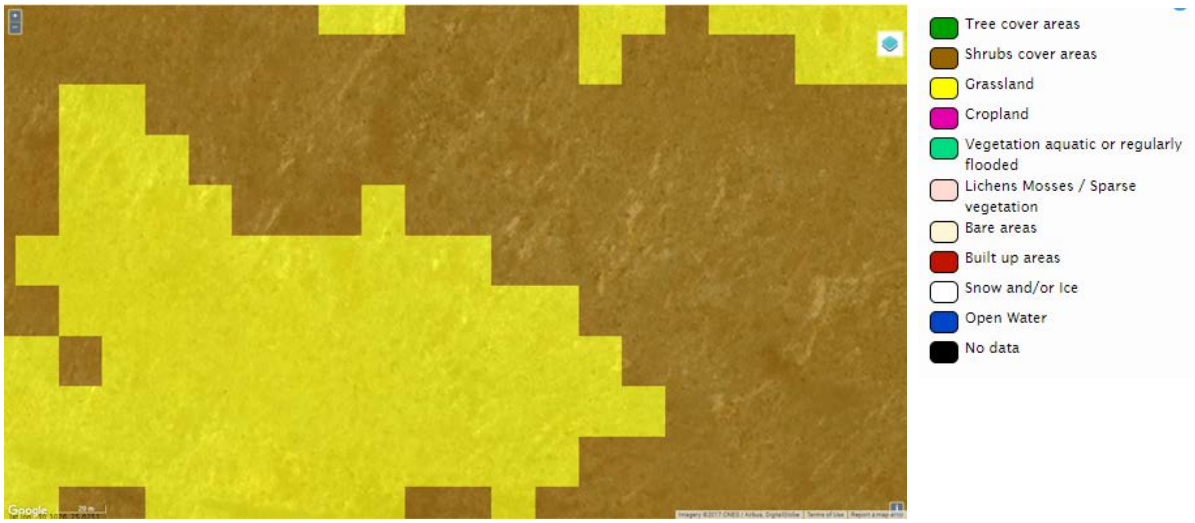
Corresponding Google image:

Example 23: South Africa

Coordinates in Lat/Lon: -30.3076, 25.6253

We have observed high confusion between shrubs and grasslands in the South Africa. In this example, this is mostly herbaceous land cover, while some parts of this region are mapped as shrub lands on the CCI LC map at 20 m.

The CCI LC map at 20 m:



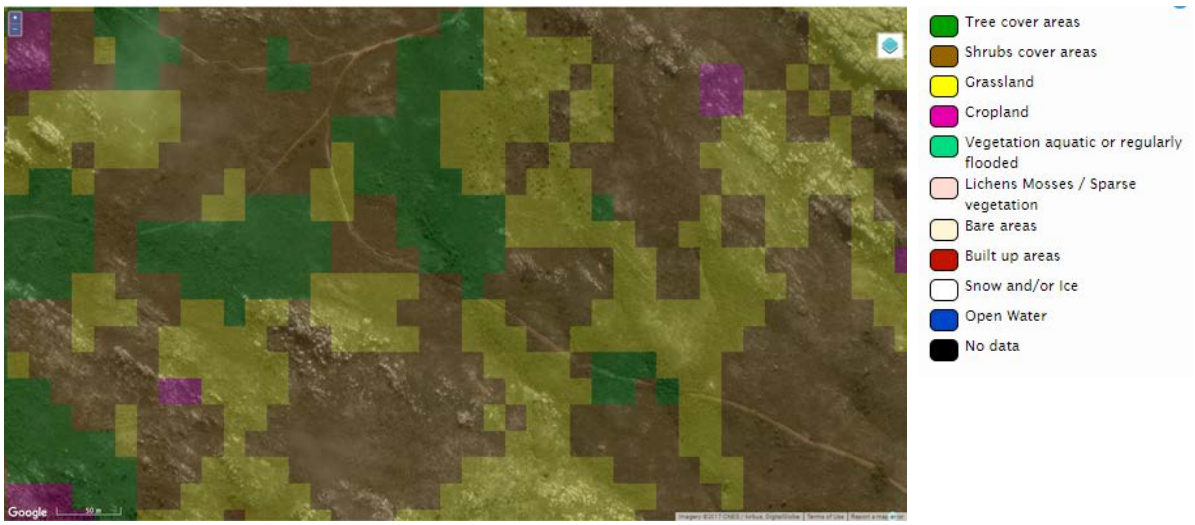
Corresponding Google image:

Example 24: South Africa

Coordinates in Lat/Lon: -34.0876, 18.4089

This area is strangely classified as trees or shrubs, or crops. Though, accordingly to the Google earth, it is grassland and some shrubs. In general, we observed that grassland class is highly confused with other classes in South Africa.

The CCI LC map at 20 m:



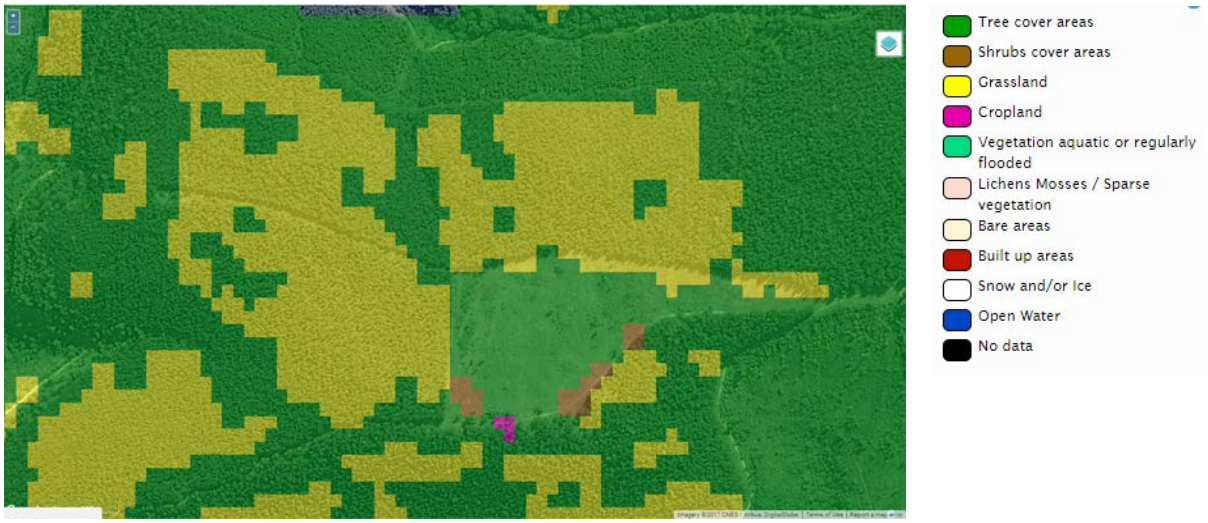
Corresponding Google image:

Example 25: South Africa

Coordinates in Lat/Lon: -33.9355, 23.2331

One more example, when closed forests are highly confused with grasslands in South Africa. The Google Earth image is from 2016 and it is clearly shows that areas classified as grassland are actually tree cover.

The CCI LC map at 20 m:



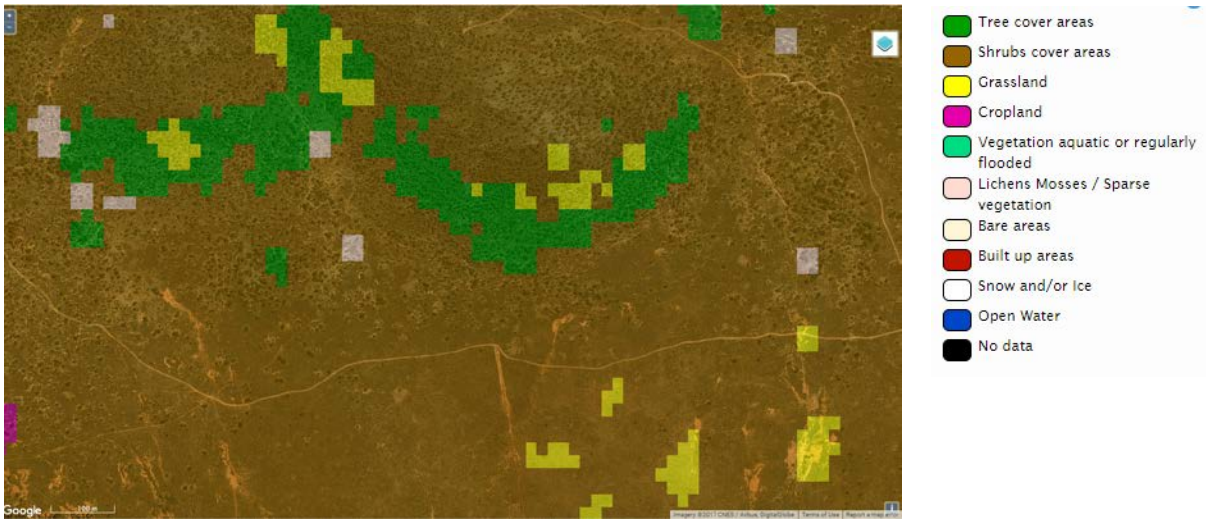
Corresponding Google image:

Example 26: South Africa

Coordinates in Lat/Lon: -33.1440, 25.9456

Here shrub lands are overestimated due the confusion with grassland. In the lower part of the screenshot from Google Earth, it can be seen that it is grassland.

The CCI LC map at 20 m:



Corresponding Google image: