ACCURACY ASSESSMENT OF THE 20 M LAND COVER MAP OF AFRICA

Deliverable 4 (WP4)

LACO-Wiki

Table of Contents

1 Introduction

This is deliverable D4 of the CrowdVal project, which presents the overall and spatial accuracy assessment of the ESA 20 m prototype land cover map for Africa for four countries: Kenya, Gabon, Ivory Coast and South Africa. The data were collected using the online LACO-Wiki land cover validation tool. The next section describes the methodology followed by the accuracy assessment. The report concludes with some suggestions for improvements to future high-resolution land cover mapping exercises.

2 Methodology

2.1 Sample design

The sample design for the validation consists of a systematic sample. The sample units were placed at a spacing of approximately 12 km for Kenya, Gabon and Ivory Coast; Table 1 indicates the total number of samples units for each of these countries.

Table 1: Sample size in each of the four countries

Country	Number of points per country
Kenya	4364
Gabon	1949
Ivory Coast	2428
South Africa	920

For South Africa, it was agreed during a progress meeting held at the ESA Living Planet Symposium (May 2019) that we would add around 1,000 extra points. In the end 920 extra points were added but the online validation process was undertaken in a different manner, as outlined in section 2.2.

2.2 Visual interpretation of the samples using LACO-Wiki

The number of points listed in Table 1 is not the number of points that were visually interpreted; these values are listed in Table 2. In Kenya, 4369 points were points validated twice by the participants of the workshop. This was for quality assurance purposes. The validations were then compared at each location, where 1898 indicated disagreement between the workshop participants. These points were then validated by an expert (Dr Myroslava Lesiv) to produce a final data set with high quality validations.

Table 2: Number of points visually interpreted using LACO-Wiki

Campaigns	Number of points visually interpreted
Kenya	10636
Gabon	1949
Ivory Coast	6128
South Africa	92000 (resulting in 23000 20m x 20m pixels validated)

In Gabon, there was only enough time during the workshop to do a single validation by the workshop participants since most of the time was devoted to in situ collection with the mobile app. However, as the results show, Gabon is one of the countries with the highest accuracy for the land cover map since the country is mostly forest.

Similar to Kenya, in Ivory Coast the points were validated twice by the workshop participants. In total 2428 points were validated twice, with an additional 1272 disagreeing areas validated by 2 experts (Dr Brice Mora and Dr Myroslava Lesiv). This quality assurance process was particularly important for Ivory Coast as this country has a highly heterogeneous land cover, as shown in the results below.

Finally, for South Africa, a different approach was adopted. Instead of validating a single 20 m pixel (as done for Kenya, Gabon and Ivory Coast), a grid of 100 m by 100 m was placed on the location and 100 validations of 10 m pixels were undertaken, where the 10 m pixels are consistent with the grid for Sentinel-2. Hence in total, there were 92000 points validated. This results in 23000 20 m x 20 m pixels.

2.3 Accuracy assessment

The accuracy assessment involved the following calculations:

- A confusion matrix;
- Overall accuracy and producer's/user's accuracy by land cover class including 95% confidence intervals;
- Percentage area mapped and the adjusted area estimates taking user's and producer's accuracy into account; and
- Spatial accuracy maps, overall and user's accuracies by class.

As a comparison, the spatial accuracy map calculated previously by Lesiv et al. (2017) is provided in Figure 1, where a previous validation exercise resulted in an overall accuracy of 65%. This provides a reference for accuracy comparisons of the individual countries.

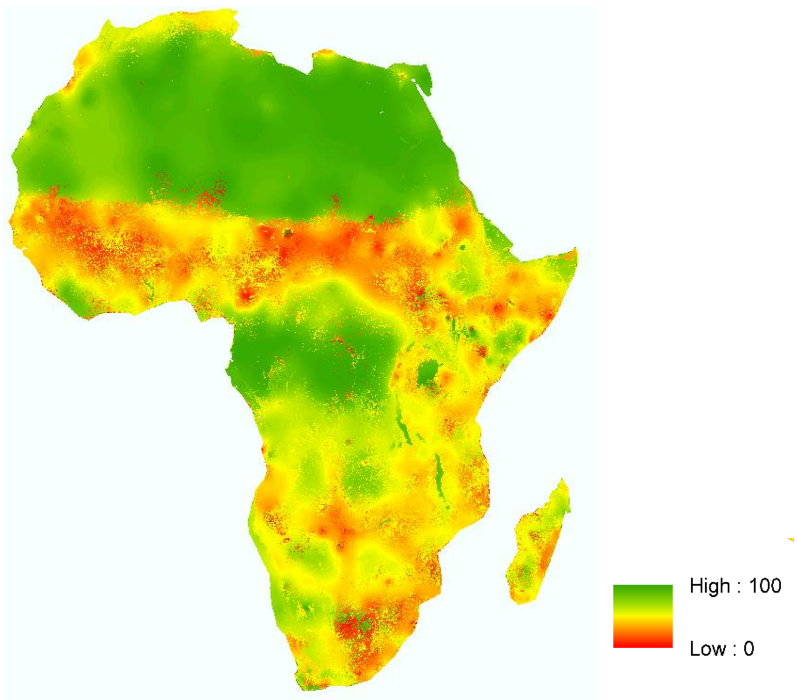


Figure 1: Map of spatial accuracy of the ESA 20 m land cover map (Lesiv et al. 2017)

The idea behind the accuracy assessment of the CrowdVal project was to produce a much denser validation sample in order to examine the spatial accuracies associated with the ESA African land cover map for Kenya, Gabon, Ivory Coast and South Africa.

3 Accuracy Assessment of Kenya

The systematic sample for Kenya is shown in Figure 2 and can be visualized and downloaded from the CrowdVal branch of Geo-Wiki (https://www.geo-wiki.org). The country is largely covered by shrubs and grassland although there are also areas of cropland in the center and to the west of the country.

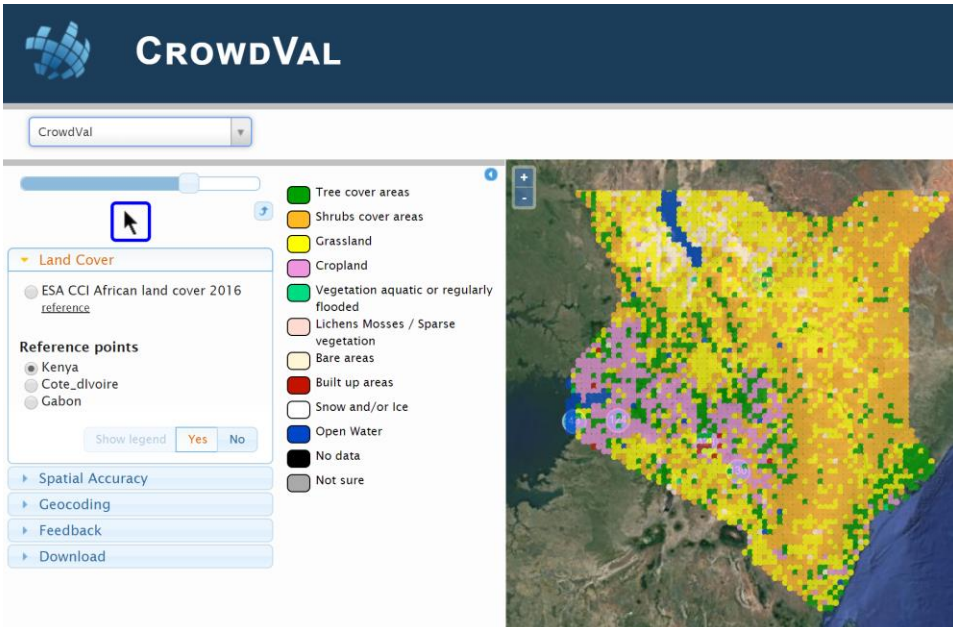


Figure 2: Systematic sample of locations visually interpreted for Kenya

3.1 Overall accuracy

Based on the confusion matrix shown in Table 3, the overall accuracy for the ESA land cover map for Kenya is 56%. If, however, the confusion between grassland and shrubs is not considered to be important, one can weight the matrix (see e.g., Fritz and See, 2008) and the overall accuracy would increase to 79%.

Table 3: Confusion matrix for Kenya

Mapped Classes	Reference class								Total	UA	CI (0.95)
	Trees	Shrubs	Grass-land	Crops	Flooded	Bare	Built-up	Water			
Trees	**217**	58	63	12	0	2	0	1	353	61%	5%
Shrubs	78	**749**	391	31	0	22	1	3	1275	59%	3%
Grassland	119	624	**819**	15	3	45	0	2	1627	50%	2%
Crops	108	139	134	**410**	1	6	3	1	802	51%	4%
Flooded	0	0	1	0	**1**	0	0	1	3	33%	65%
Bare	0	16	47	0	0	**141**	0	0		69%	6%
Built-up	2	0	0	1	0	0	**8**	0	11	73%	28%
Water	0	0	0	0	0	0	0	**89**	89	100%	0%
Total	524	1586	1455	469	5		12	97	**4364**		
PA	41%	47%	56%	87%	20%	65%	67%	92%		**56%**	2%
CI (0.95)	3%	2%	2%	3%	35%	23%	23%	5%			

CI = Confidence Intervals; PA = Producer's Accuracy; UA = User's Accuracy

The user's and producer's accuracies are plotted in Figure 3 while the percentage areas mapped and then adjusted by the user's and producer's accuracies are given in Figure 4. The main issues identified in Kenya are that shrub cover is underestimated and both grassland and cropland are overestimated. This is clearly reflected in Figure 4 but also in the confusion matrix, which shows the confusion between these three classes. There is also confusion between grassland and bare areas as well as between trees and shrubs, grassland and cropland but the overall area of trees is much smaller than these other classes.

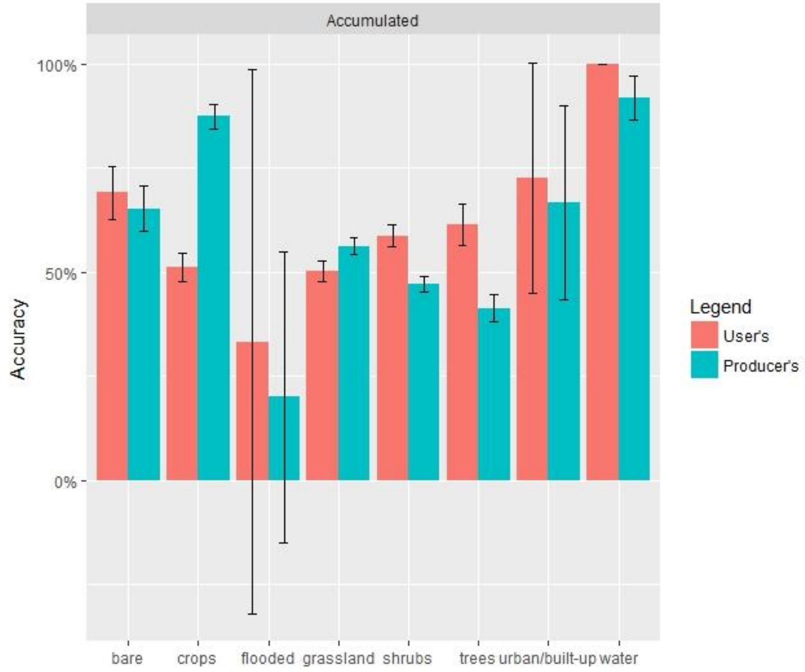


Figure 3: User and producer accuracies of the ESA African land cover product for land cover classes in Kenya

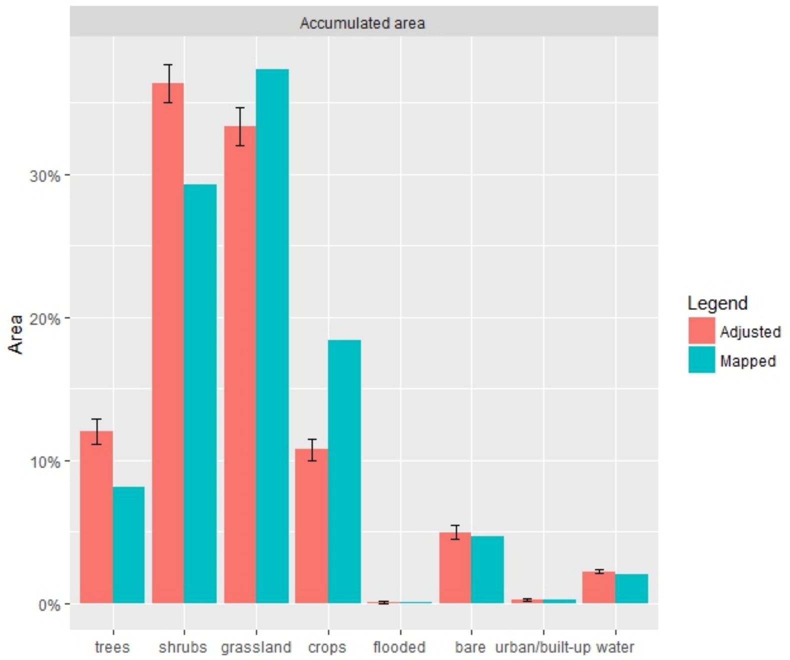


Figure 4: Percentage areas mapped by the ESA African land cover product for Kenya and the adjusted areas based on the user/producer accuracies

3.2 Spatial accuracy

Figure 5 shows the spatial accuracy of the ESA African land cover map for Kenya. The results show examples of low accuracy throughout the country. In particular, the north-eastern part of the country, which is due to the confusion between grassland and shrubs. Even though relatively dry, there is a substantial proportion of shrubs present in this northeastern area, which have been incorrectly classified in as grassland.

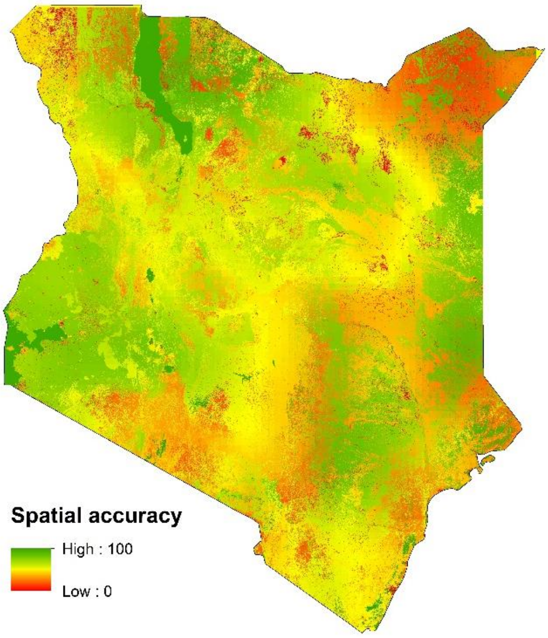
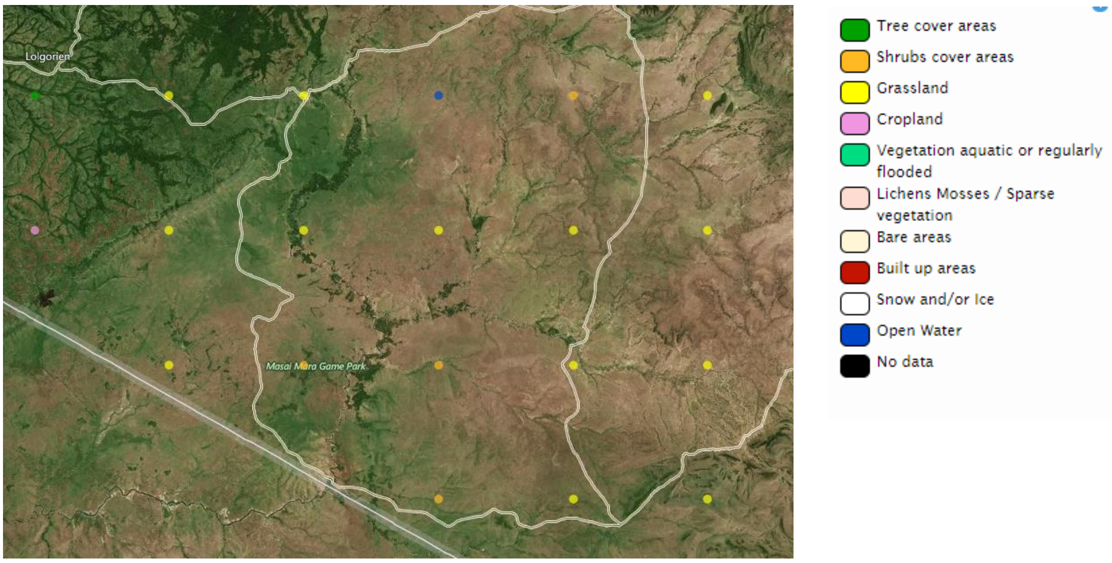


Figure 5: Map of the spatial accuracy of the ESA African land cover map for Kenya

Another area incorrectly classified as cropland occurs in the Maasai Mara Reserve, but it contains natural, mostly grassland, areas (see Figure 6 and 7).



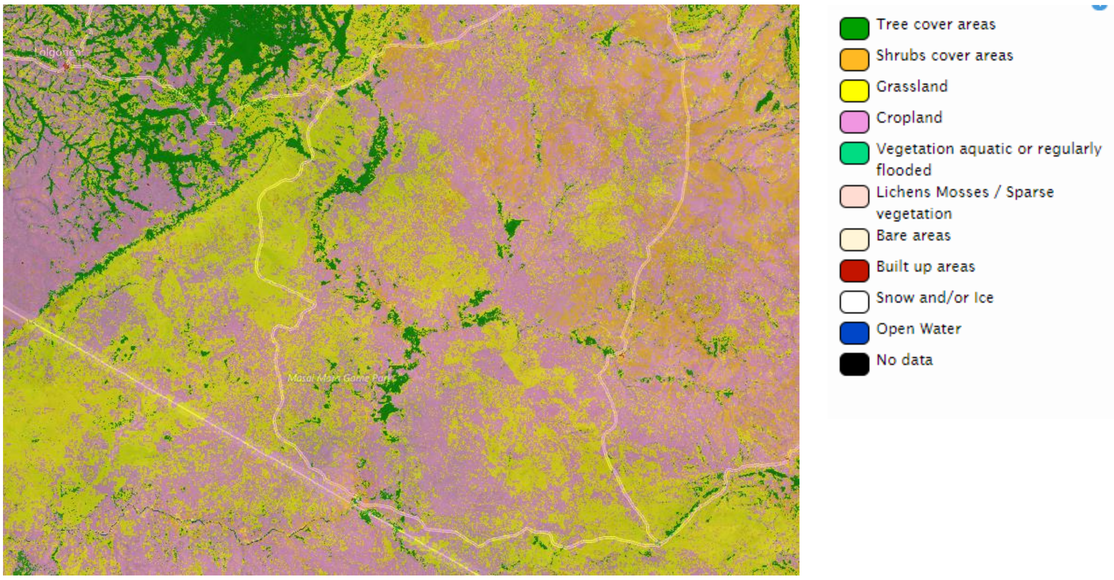


Figure 6: The Maasai Mara Reserve shown using Microsoft Bing very high-resolution satellite imagery with the data points collected using LACO-Wiki overlaid on the map (top image) and the ESA 20m African land cover map (bottom image)

Figure 7: Photographs from the Maasai Mara Reserve showing Grassland

Figures 8 to 10 show the spatial user's accuracy. Tree cover shows high user's accuracy (Figure 8a) compared to the shrub class (Figure 8b), which shows areas of lower user's accuracy in the western part of the country. Figure 9a shows the spatial user's accuracy for grassland where, in particular in the north east, low user's accuracies for that class can be identified. Overall the cropland class (Figure 9b) shows a high user's accuracy although one can see issues in the southern natural areas (e.g., the Maasai Mara). Both bare soils (Figure 10a) and urban areas (Figure 10b) show little confusion error with a high user's accuracy.

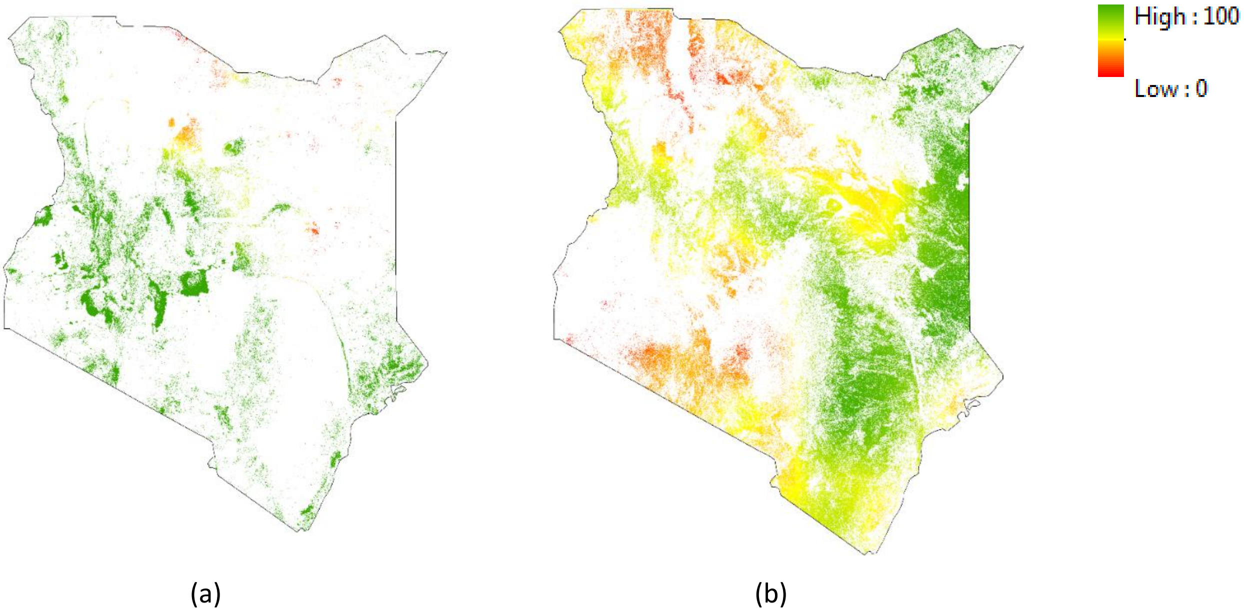


Figure 8: Spatial accuracy of the ESA African land cover map for Kenya for (a) the tree cover class and (b) the shrub class

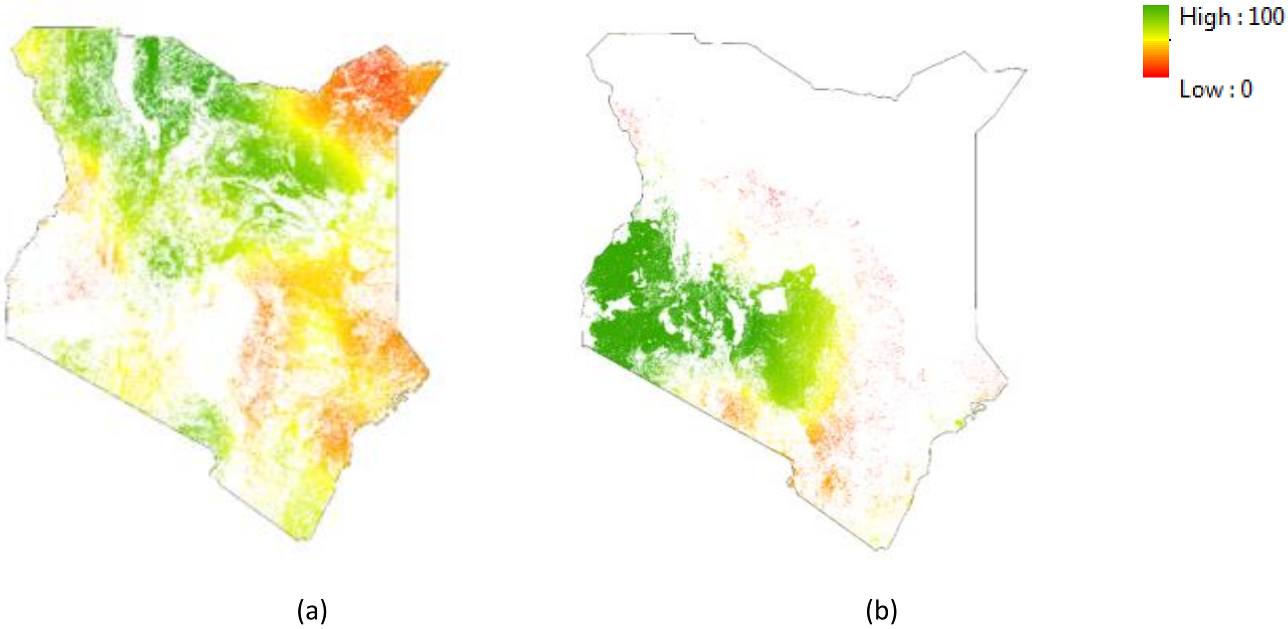
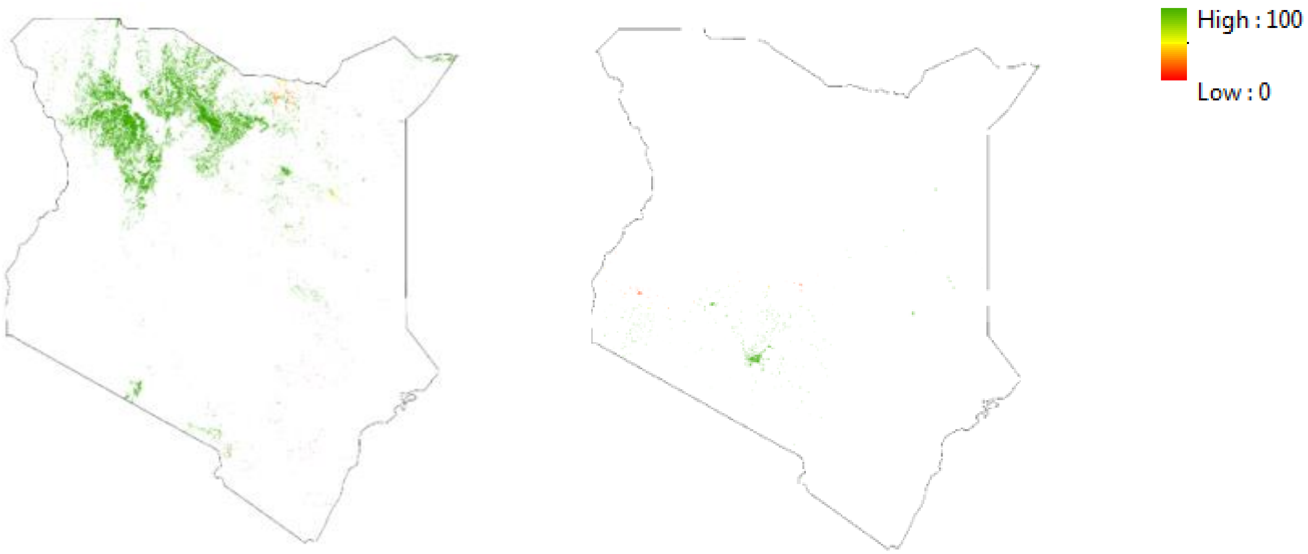


Figure 9: Spatial accuracy of the ESA African land cover map for (a) the grassland and (b) the cropland class



(a) (b)

Figure 10: Spatial accuracy of the ESA African land cover map for (a) the bare soils and sparse vegetation and (b) the built-up class

4 Accuracy Assessment of Gabon

In Figure 11, the systematic sample for Gabon is shown. As with Kenya, the data can be visualized and downloaded from the CrowdVal branch of Geo-Wiki. The land cover for Gabon shown in Figure 11 is entirely different to the other three countries because it is largely forest cover.

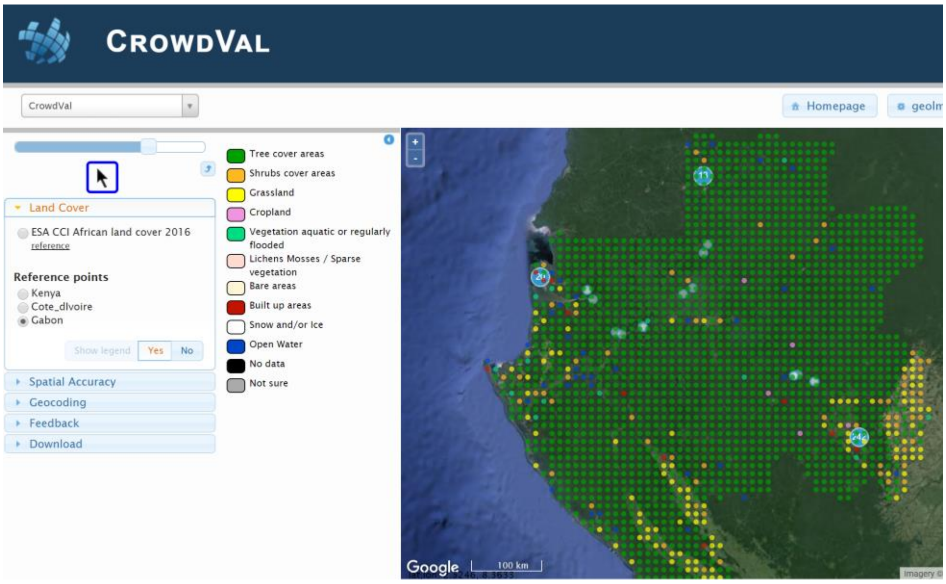


Figure 11: Systematic sample of locations visually interpreted for Gabon

4.1 Overall accuracy

The overall accuracy for Gabon is 91% based on the confusion matrix shown in Table 4. This is unsurprising because of the high amount of forest cover, which is a relatively easy class to map using remote sensing.

Table 4: Confusion matrix for Gabon

Mapped Classes	Reference class							Total	UA	CI (0.95)
	Trees	Shrubs	Grass-land	Cropland	Flooded	Built-up	Water			
Trees	**1691**	26	26	4	8	4	17	1776	95%	1%
Shrubs	1	**0**	0	0	0	0	0	1	0%	-
Grassland	7	24	**57**	0	1	4	0	93	61%	10%
Crops	14	7	19	**1**	0	1	1	43	2%	5%
Flooded	0	2	5	0	**1**	1	2	11	9%	20%
Built-up	0	0	0	0	0	**2**	0	2	100%	0%
Water	1	0	1	0	3	0	**18**	23	78%	17%
Total	1714	59	108	5	13	12	38	**1949**		
PA	99%	0%	53%	20%	8%	17%	47%		**91%**	1%
CI (0.95)	0%	-	7%	35%	14%	9%	12%			

CI = Confidence Intervals; PA = Producer's Accuracy; UA = User's Accuracy

The user's and producer's accuracies are plotted in Figure 12 while the percentage areas mapped and adjusted by the user's and producer's accuracies are given in Figure 13. The UA and PA for forest cover are very high with small confidence intervals, further confirming the overall high accuracy of the map. There is some confusion between grassland, shrubs and cropland but the areas are very small.

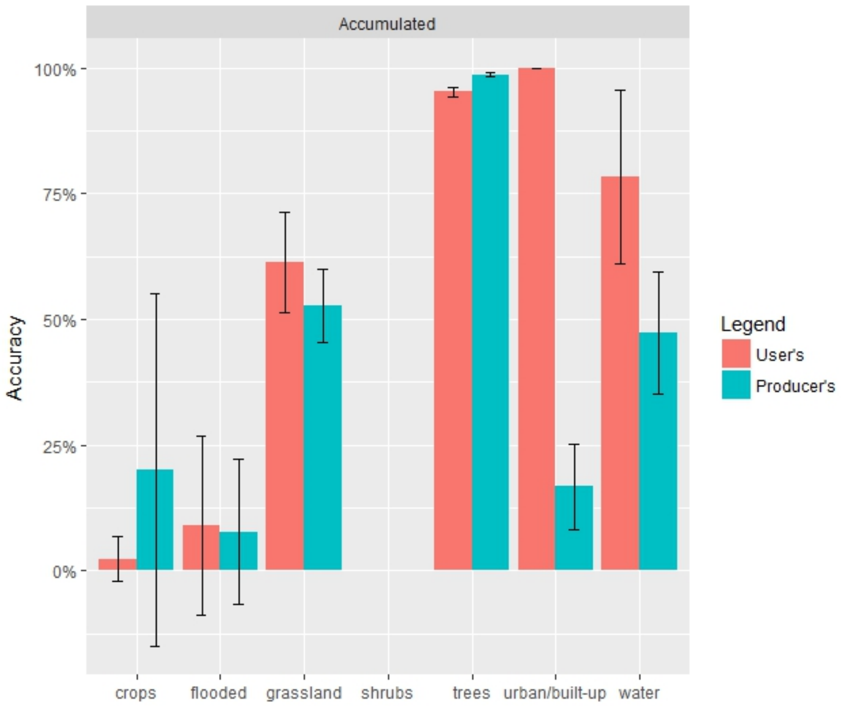


Figure 12: User and producer accuracies of the ESA African land cover product for land cover classes in Gabon

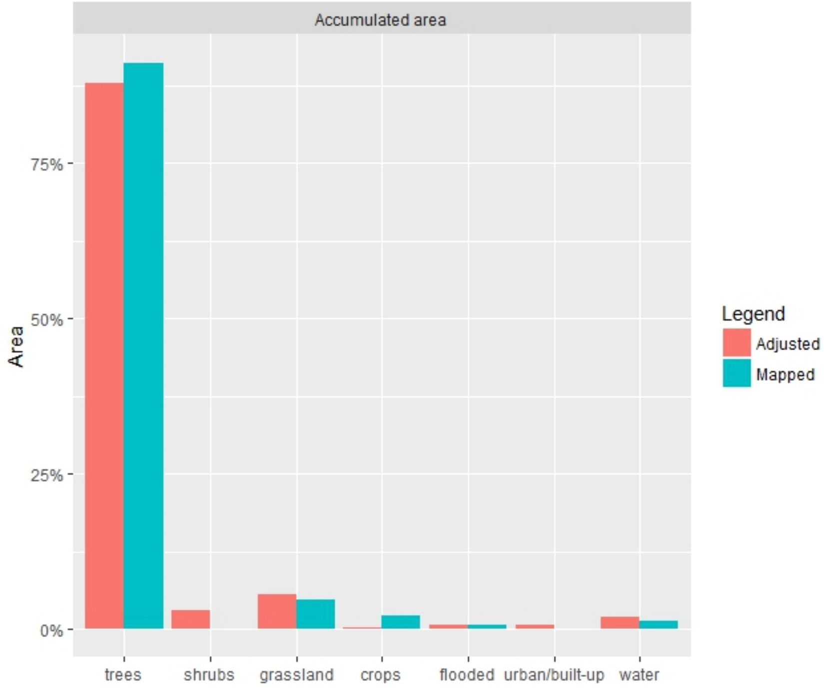
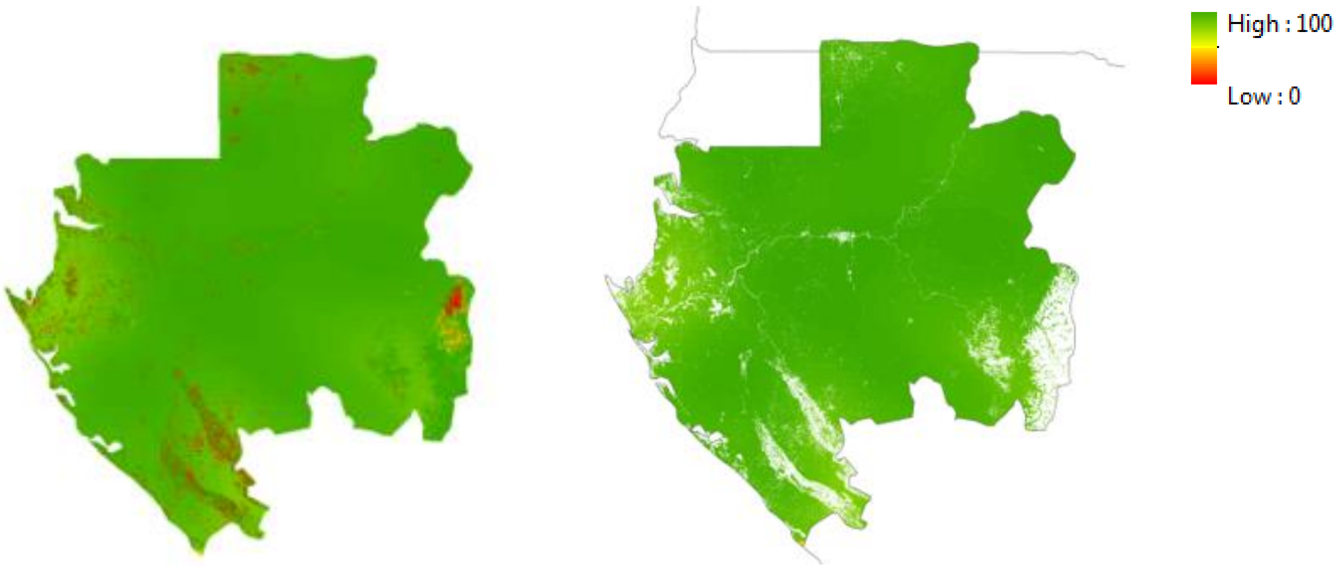


Figure 13: Percentage areas mapped by the ESA African land cover product for Gabon and the adjusted areas based on the user/producer accuracies

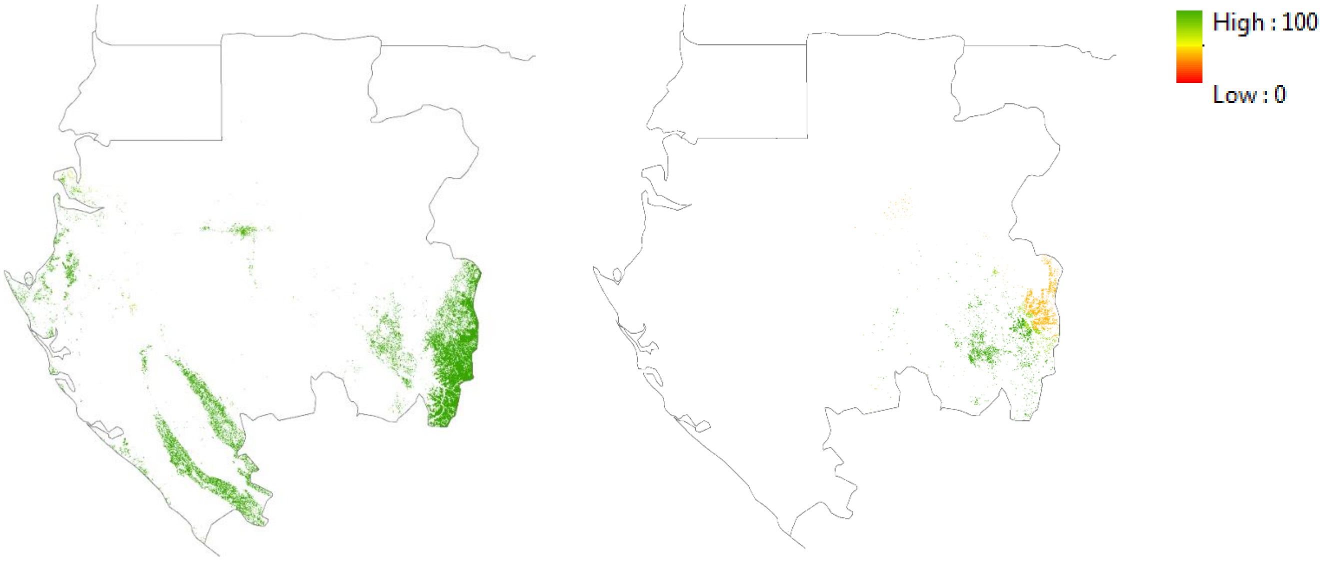
4.2 Spatial accuracy

Figure 14a shows the spatial accuracy of the ESA African land cover map for Gabon while spatial user accuracies by land cover class are shown in Figure 14b and Figures 15 and 16. There are some issues related to cropland in the east of the country (Figure 15b) as cropland is overestimated. Other classes appear to be mapped relatively well.



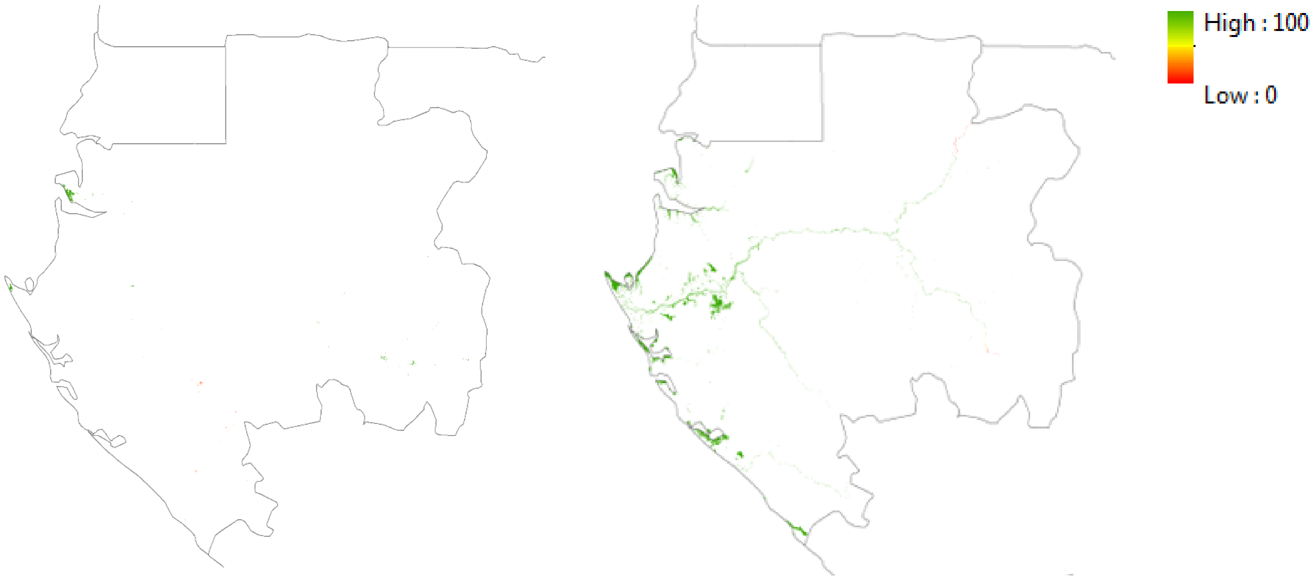
(a) (b)

Figure 14: (a) Map of the spatial accuracy of the ESA African land cover map for Gabon and for (b) the tree cover class



(a) (b)

Figure 15: Spatial accuracy of the ESA African land cover map for (a) the grassland and (b) cropland classes



(a) (b)

Figure 16: Spatial accuracy of the ESA African land cover map for (a) the built-up and (b) water classes

5 Accuracy Assessment of Ivory Coast

In Figure 17, the systematic sample for Ivory Coast is shown in Geo-Wiki, where it can be viewed and downloaded. This figure clearly shows how heterogenous the land cover is in Ivory Coast.

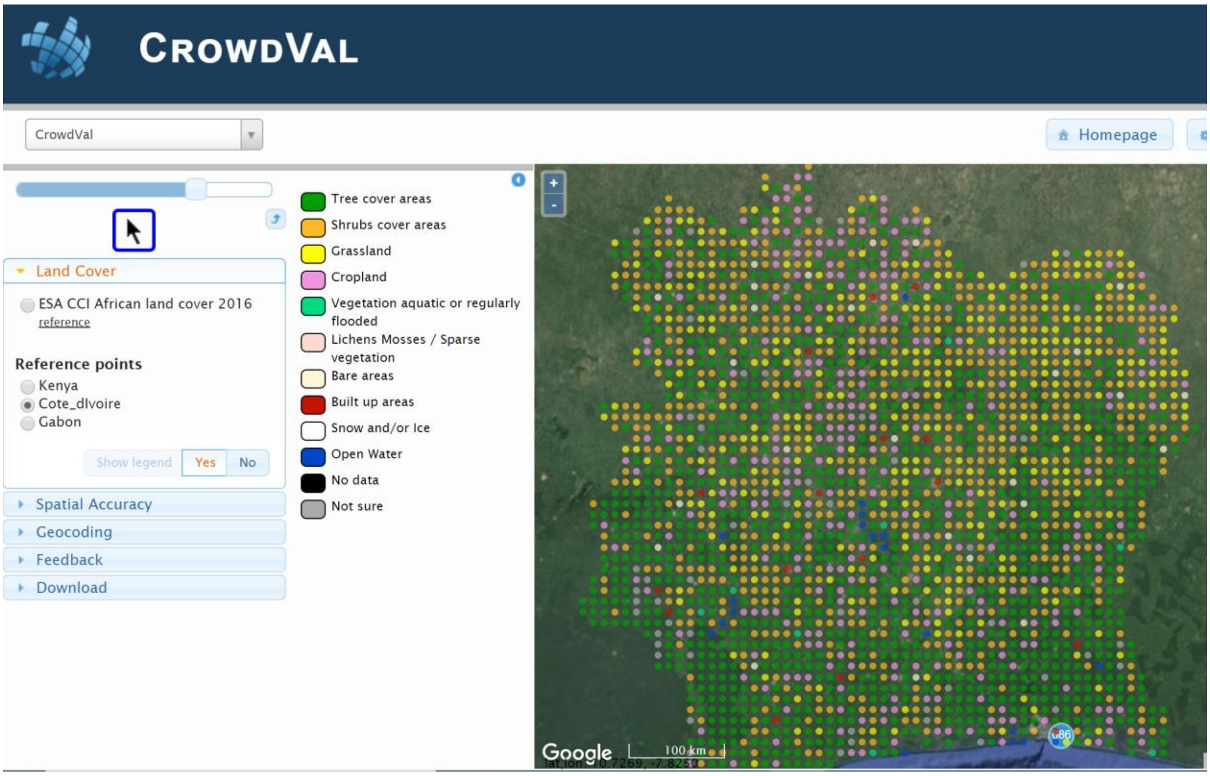


Figure 17: Systematic sample of locations visually interpreted for Ivory Coast

5.1 Overall accuracy

Given the heterogeneity of the land cover, it is unsurprising that the overall accuracy for Ivory Coast is 47% (derived from the confusion matrix shown in Table 5).

Table 5: Confusion matrix for Ivory Coast

Mapped Classes	Reference class								Total	UA	CI (0.95)
	Trees	Shrubs	Grass-land	Crops	Flooded	Bare	Built-up	Water			
Trees	748	382	113	108	1	8	0	0	1360	55%	3%
Shrubs	21	94	44	31	0	4	0	0	194	48%	7%
Grassland	95	176	132	45	3	10	1	0	462	29%	4%
Crops	96	63	38	102	3	4	6	0	312	33%	5%
Flooded	1	1	1	0	0	0	0	0	3	0%	0%
Bare	0	0	0	0	0	0	0	0		100%	0%
Built-up	0	2	2	0	0	0	9	0	13	69%	26%
Water	2	0	1	0	1	0	0	19	23	83%	16%
Total	963	718	331	286	8		16	19	2367		
PA	78%	13%	40%	36%	0%	0%	56%	100%		**47%**	2%
CI (0.95)	2%	2%	5%	5%	0%	0%	20%	0%			

CI = Confidence Intervals; PA = Producer's Accuracy; UA = User's Accuracy

The user's and producer's accuracies are plotted in Figure 18 while the percentage areas mapped and adjusted by the user's and producer's accuracies are given in Figure 19.

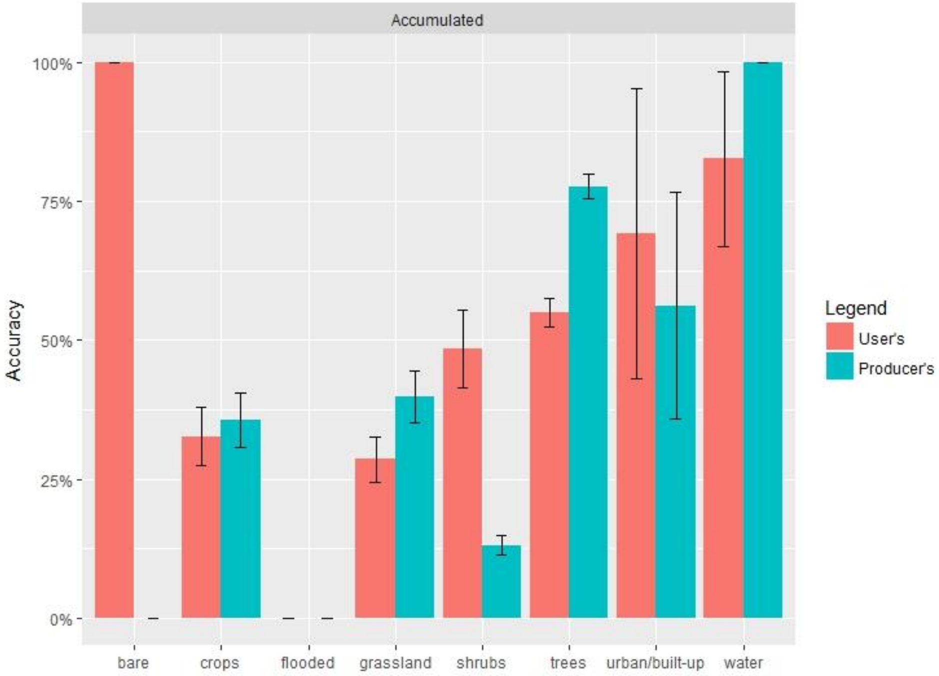


Figure 18: User and producer accuracies of the ESA African land cover product for land cover classes in Ivory Coast

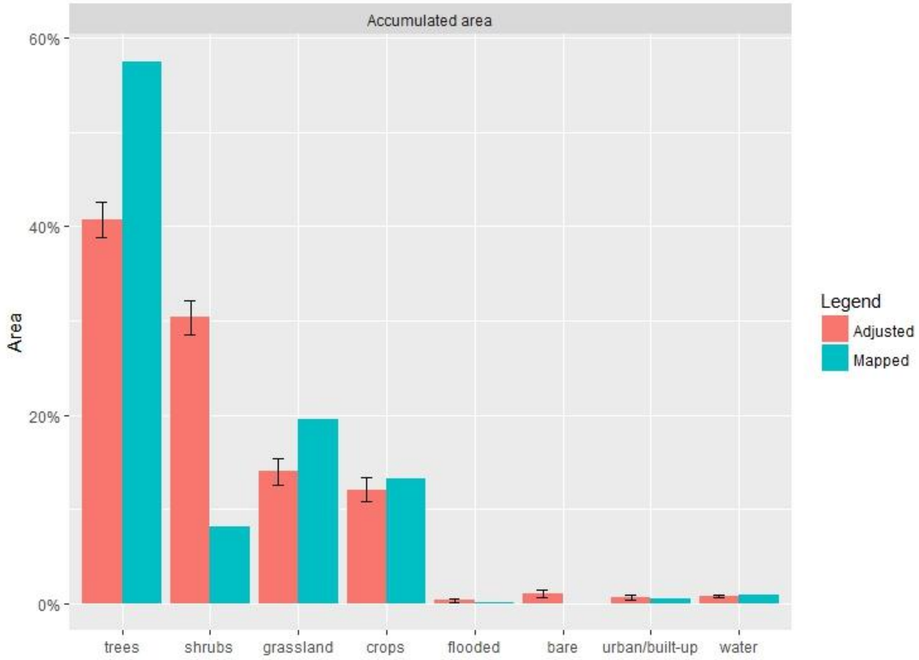


Figure 19: Percentage areas mapped by the ESA African land cover product for Ivory Coast and the adjusted areas based on the user/producer accuracies

Although wetland covered a small amount of the area, some wetlands in Ivory Coast are fully covered by tree species. These wetlands were observed during the fieldwork activity. Such environments may be misinterpreted as tree cover (and misclassified if not appropriately represented in the training data set). Here is an example where the field work produced some interesting insights into the local land cover that could only be determined through interaction with local experts.

5.2 Spatial accuracy

Figure 20 shows the spatial accuracy of the ESA African land cover map for Ivory Coast while spatial user accuracies by land cover class are shown in Figures 21 and 22. The overall spatial accuracy once again shows the heterogeneity of the country, with some areas mapped well and other less well. Tree cover is mapped well in some areas but not others (Figure 21a), with Figure 19 showing that tree cover is generally overestimated. There are also problems with the mapping of shrubs (which are underestimated), grassland (overestimated) and cropland as shown spatially but also in Figures 18 and 19.

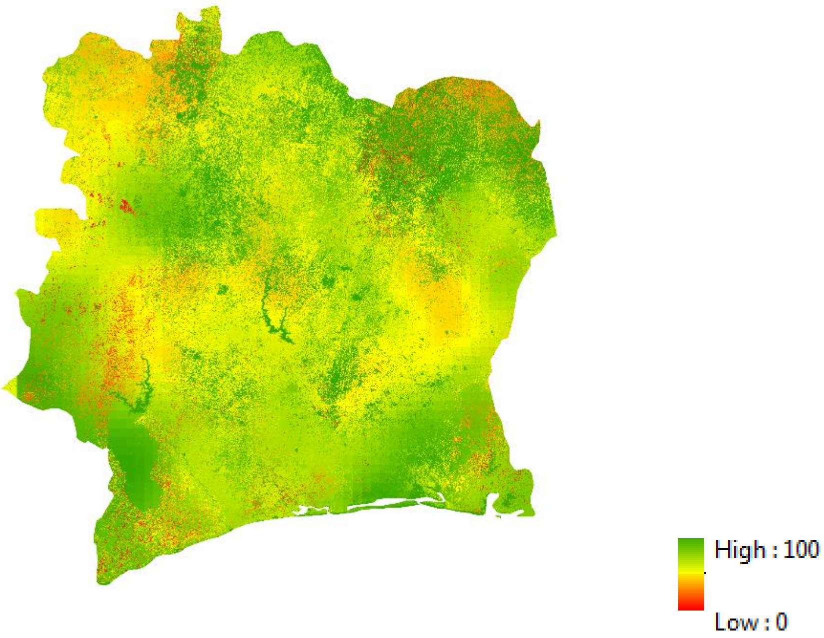
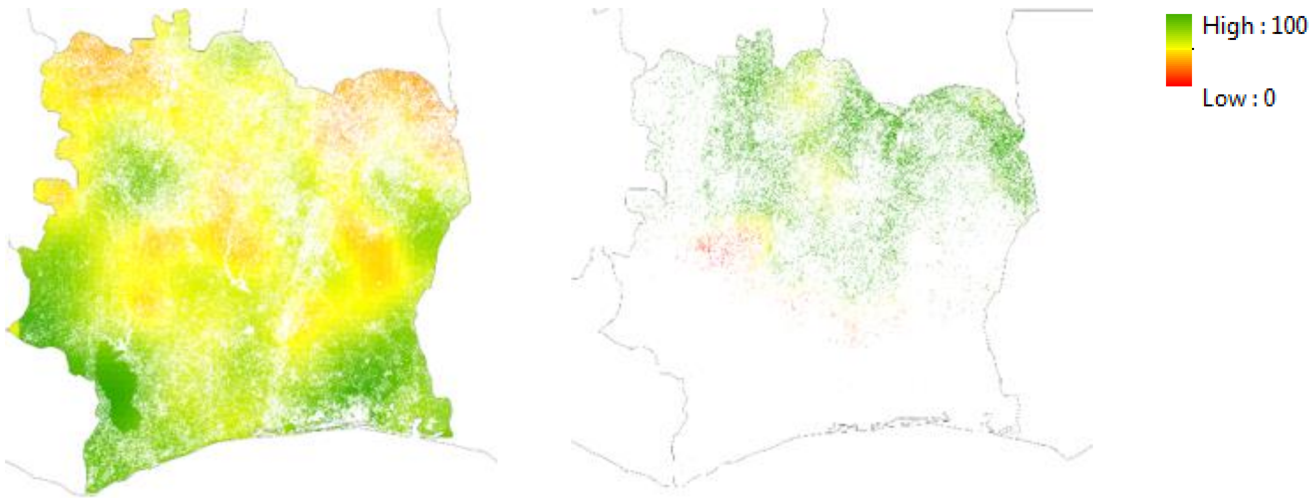
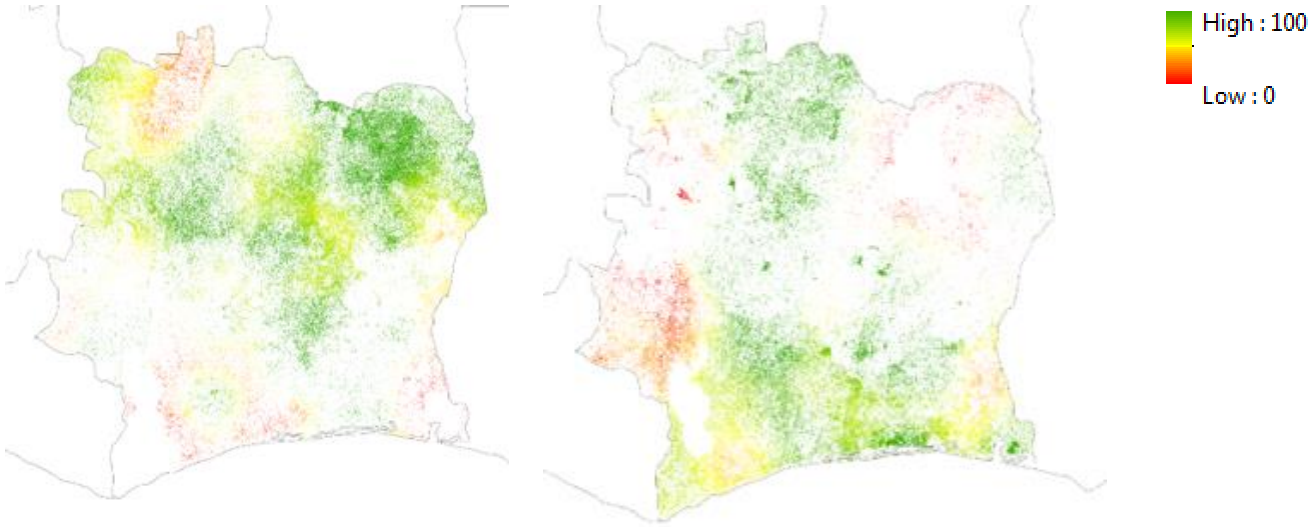


Figure 20: Map of the spatial accuracy of the ESA African land cover map for Ivory Coast



(a) (b)

Figure 21: Spatial accuracy of the ESA African land cover map for Ivory Coast for the (a) tree cover and (b) shrub classes



(a) (b)

Figure 22: Spatial accuracy of the ESA African land cover map for Ivory Coast for the (a) grassland and (b) cropland classes

6 Accuracy Assessment of South Africa

The final country assessed as part of the CrowdVal project is South Africa, which is shown in Figure 23 in Geo-Wiki. The gap in the data shown is filled using reference data from the C-GLOPS project for the purpose of the accuracy assessment but is not available for downloading. Only the data shown in Figure 23 are available for downloading. Hence the current accuracy numbers reported here CANNOT be reproduced by downloading this data set. In the future, this data set will be made available by the C-GLOPS project and then the numbers will be reproducible.

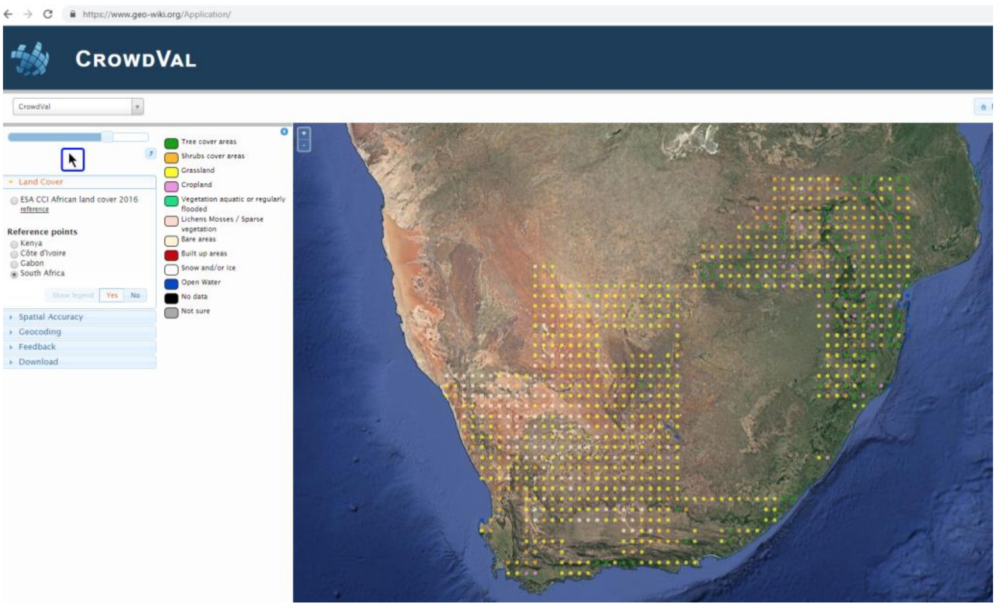


Figure 23: Systematic sample of locations visually interpreted for South Africa

As mentioned in section 2, each point shown in Figure 23 is actually 100 validation points in a 100 m grid, where each grid cell is 10 m. We did not do the validation at a 20 m pixel level. Instead, we used the dominant land cover class across the 100 m grid to address any geolocation errors and hence the original visually interpreted data set is provided in Geo-Wiki. Users could aggregate the data to 20 m pixels if desired.

5.1 Overall accuracy

Based on the confusion matrix shown in Table 6, the overall accuracy for South Africa is 44%.

Table 6: Confusion matrix for South Africa

Mapped Classes	Reference class								Total	UA	CI (0.95)
	Trees	Shrubs	Grass-land	Crops	Flooded	Bare	Built-up	Water			
Trees	**197**	78	100	10	4	0	0	0	389	51%	5%
Shrubs	82	**268**	1199	122	1	20	0	2	1694	16%	2%
Grassland	28	92	**728**	31	3	91	1	1	975	75%	3%
Crops	24	24	205	**406**	2	1	2	3	667	61%	4%
Flooded	0	0	0	0	**0**	1	0	0	1	0%	
Bare	0	4	103	0	0	**166**	2	0		60%	6%
Built-up	6	0	31	4	0	18	**56**	1	116	48%	9%
Water	0	0	2	0	0	1	0	**18**	21	86%	15%
Total	337	466	2368	573	10		61	25	**4138**		
PA	58%	58%	31%	71%	0%	56%	92%	72%		**44%**	1%
CI (0.95)	5%	4%	1%	3%		5%	7%	15%			

CI = Confidence Intervals; PA = Producer's Accuracy; UA = User's Accuracy

The user's and producer's accuracies are plotted in Figure 24 while the percentage areas mapped and adjusted by the user's and producer's accuracies are given in Figure 25.

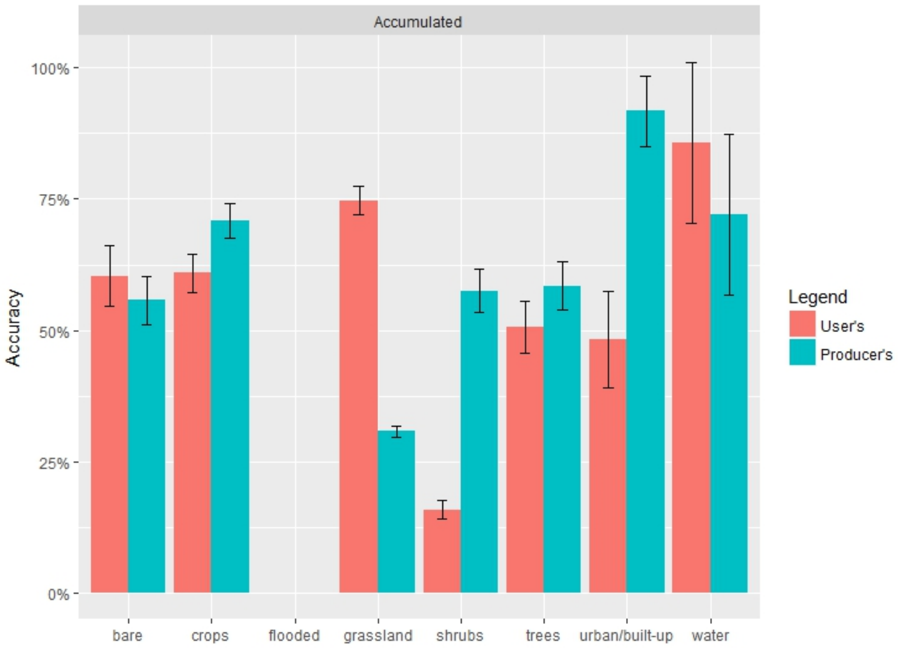


Figure 24: User and producer accuracies of the ESA African land cover product for land cover classes in South Africa

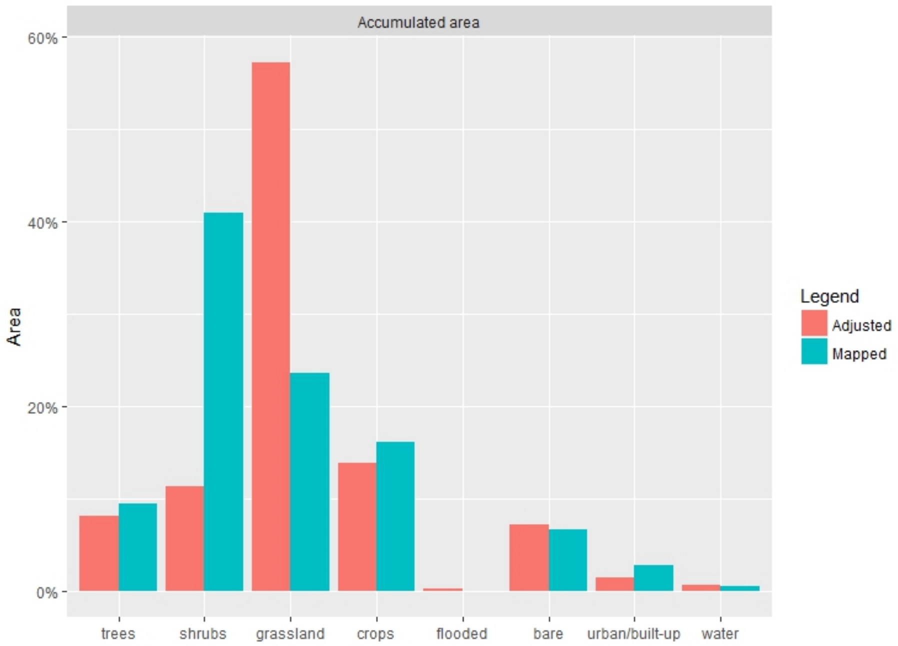


Figure 25: Percentage areas mapped by the ESA African land cover product for South Africa and the adjusted areas based on the user/producer accuracies

5.2 Spatial accuracy

Figure 26 shows the spatial accuracy of the ESA African land cover map for Ivory Coast, which shows large areas in which the accuracy is low, while spatial user accuracies by land cover class are shown in Figures 27 to 31. The main issue is overestimation of shrubland and underestimation of grassland as shown in Figure 25 as well as in the spatial accuracy maps (e.g., see Figure 28).

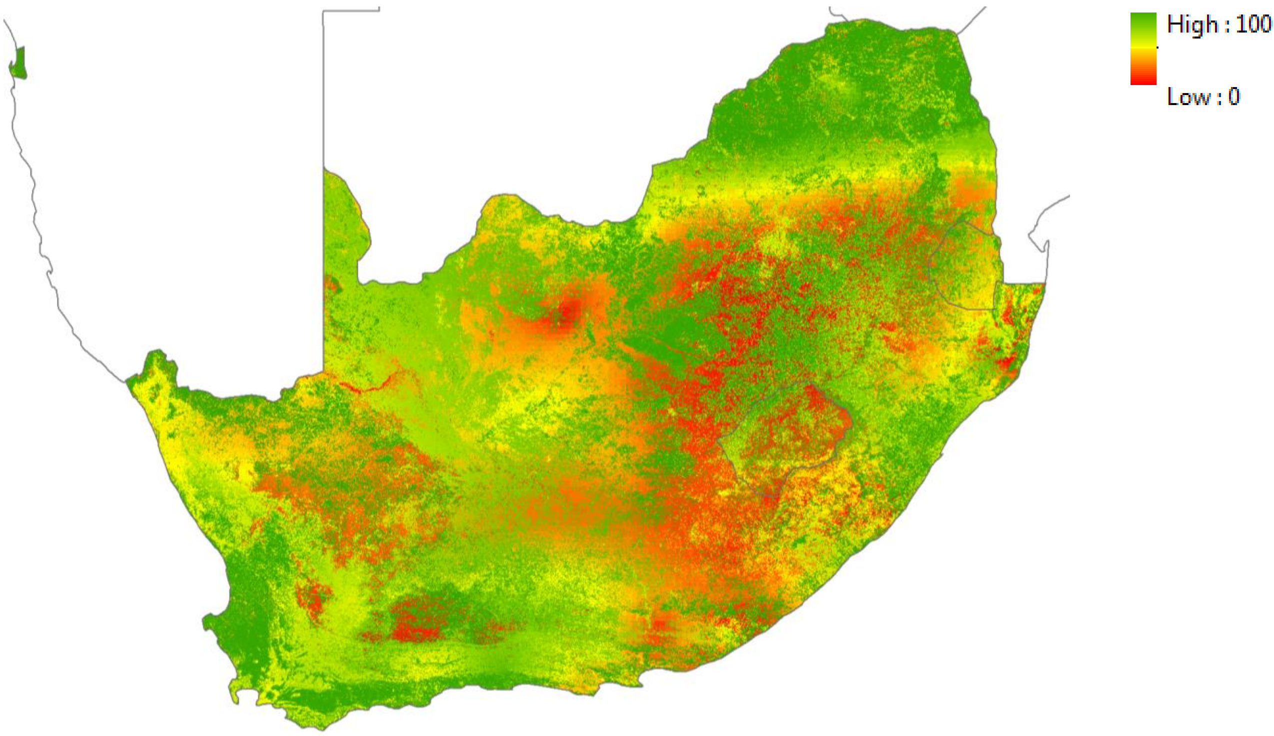


Figure 26: Map of the spatial accuracy of the ESA African land cover map for South Africa

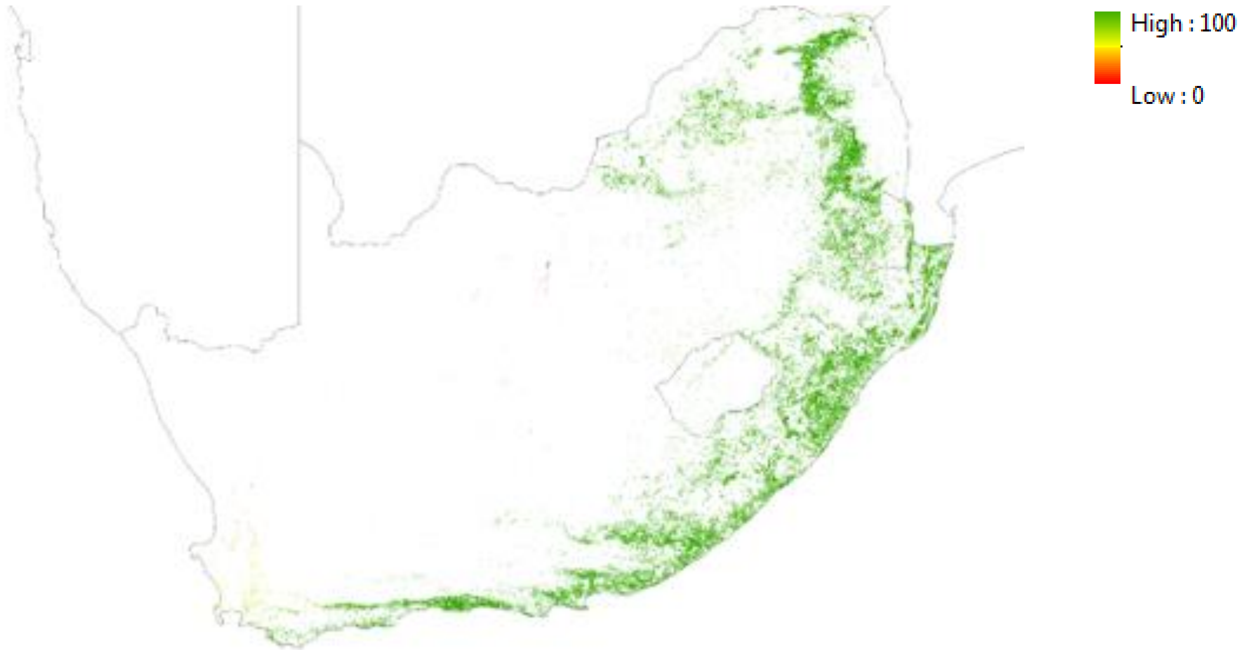


Figure 27: Spatial accuracy of the ESA African land cover map for South Africa for the tree cover class

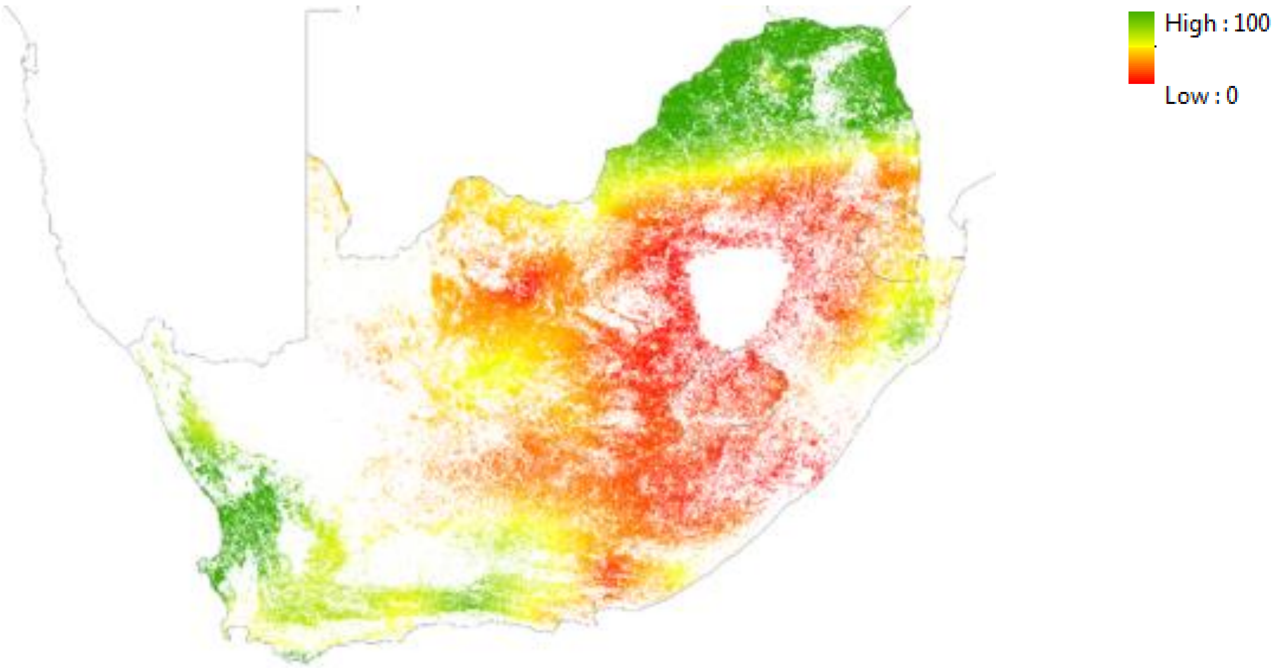


Figure 28: Spatial accuracy of the ESA African land cover map for South Africa for the shrub cover class

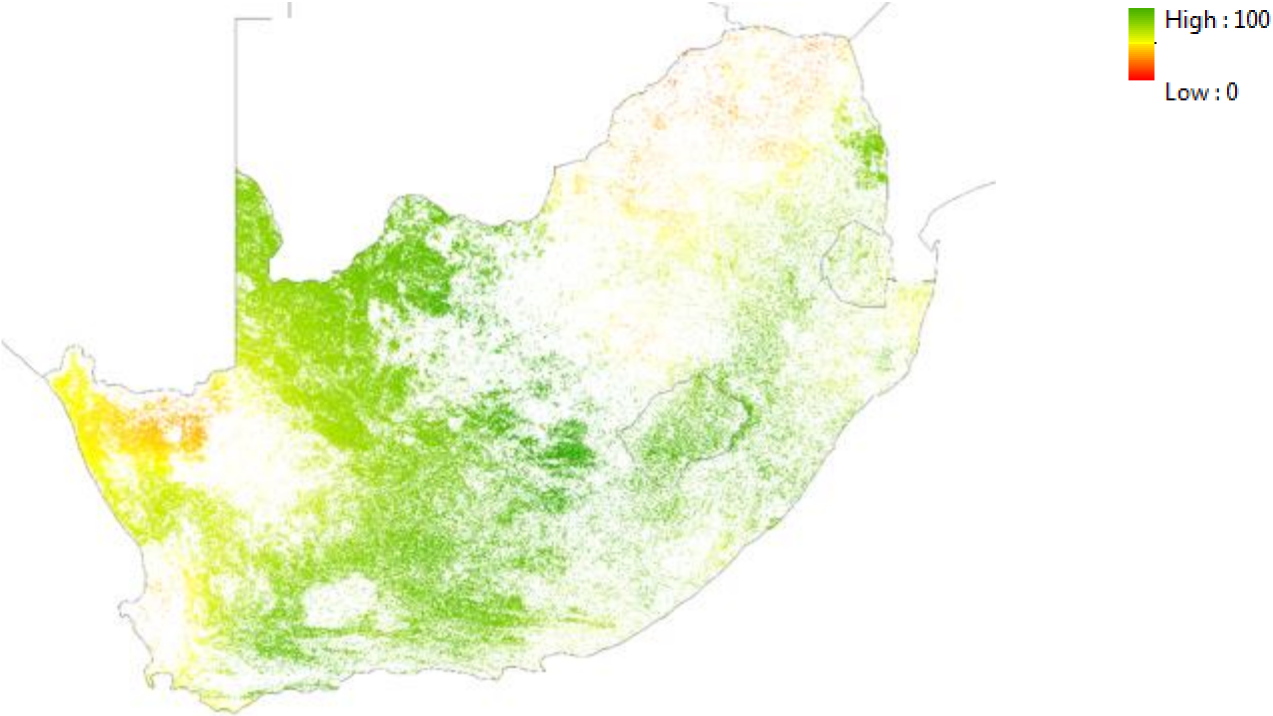


Figure 29: Spatial accuracy of the ESA African land cover map for South Africa for the grassland class

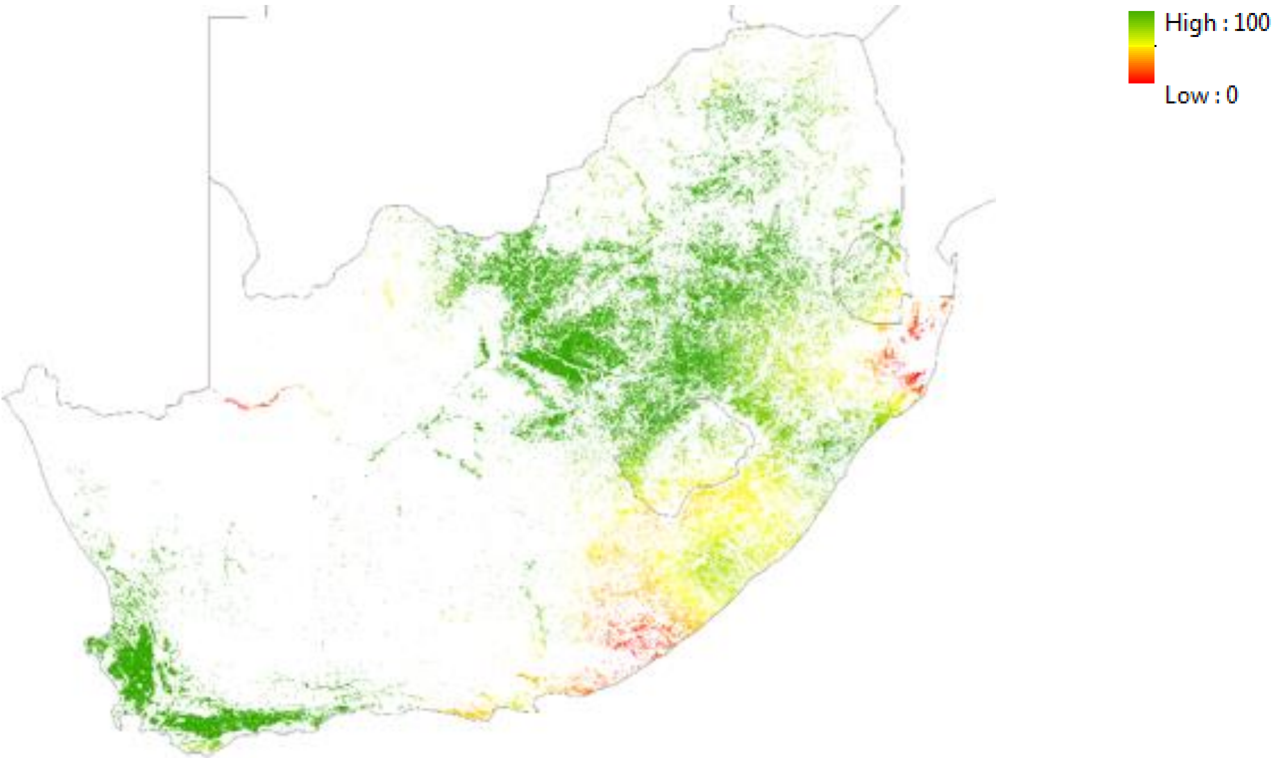


Figure 30: Spatial accuracy of the ESA African land cover map for South Africa for the cropland class

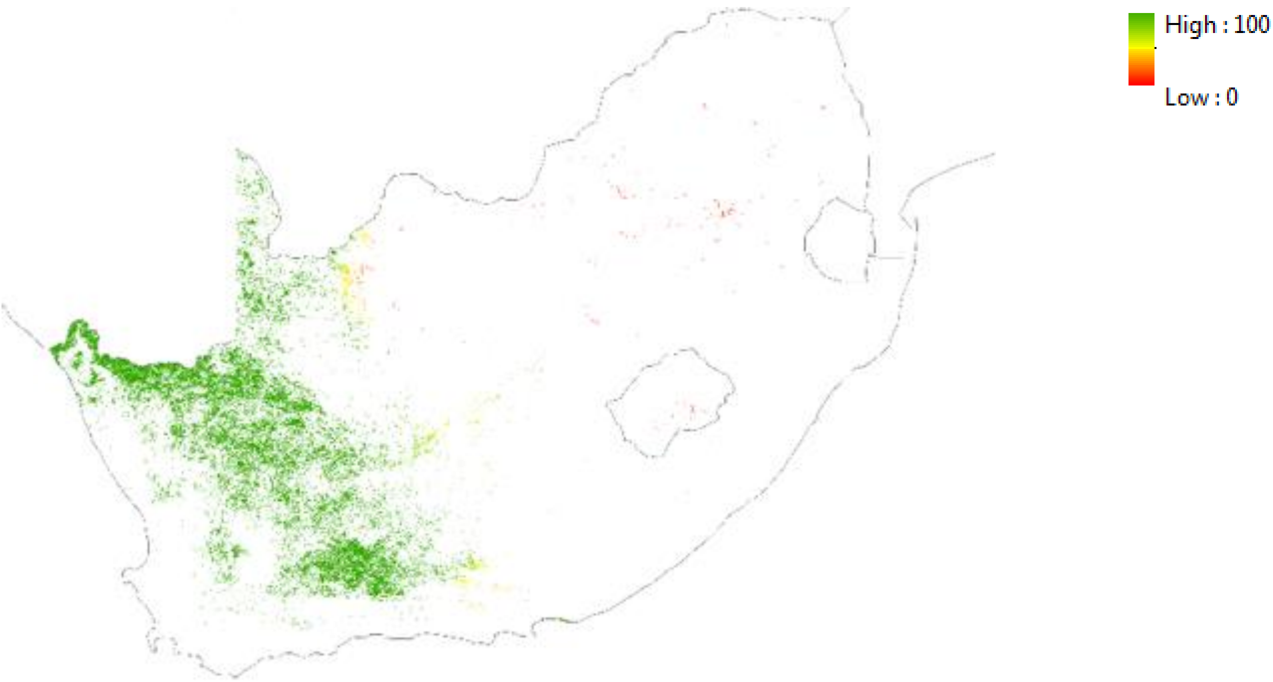


Figure 31: Spatial accuracy of the ESA African land cover map for South Africa for the bare area and sparse vegetation classes combine

7 Summary and Lessons Learned

This report has provided an accuracy assessment of the ESA 20 m land cover map of Africa for four African countries (Kenya, Gabon, Ivory Coast and South Africa). The results varied from 44% (for South Africa) to 91% (for Gabon). In the case of Kenya (56% overall accuracy) and South Africa, these values are largely caused by the confusion between grassland and shrubland. This may be due to the training data used by the classifier and should be carefully checked. The training data for the ESA African land cover map were partly taken from existing maps and may also go some way to explaining the classification errors. However, we have demonstrated that if a weighted confusion matrix is used, which diminishes the importance of the confusion between grassland and shrubs, the overall accuracy for Kenya increases to 79%.

The overall accuracy for Ivory Coast can be explained using different reasons. Ivory Coast has a highly fragmented land cover, which makes it a difficult country to map with remote sensing. Moreover, there will most likely be a low density of usable optical images that are cloud free, which may be compounding problems with the classification. The exception was Gabon with a high overall accuracy of 91% but can be explained by the high amount of tree cover across the country, which is a relatively easy class to map.

One might argue that doing a validation of a 20 m resolution map using 20 m resolution pixels is not the right approach due to geo-registration errors. However, we would argue that the issues are not related to resolution but rather misclassification of large areas. The South Africa example clearly demonstrates this since a different approach was used, i.e., the dominant land cover over a 100 m squared area was used in the validation yet this map had the lowest accuracy of all 4 countries. Hence aggregating to a larger area for validation does not improve the accuracy figures because the areas that are misclassified are very large. An example where aggregation might improve accuracy is actually Gabon. Although most is forest, there are occasional validation pixels at a 20 m resolution with a different land cover class such as cropland that could be considered noise. Aggregating to a larger area would remove these cases. However, as the overall accuracy for Gabon was already very high, this would be unlikely to make a huge difference.

Below is a list of suggestions for how high-resolution land cover mapping might be improved in the future:

- Improve the training data, particularly if they have been derived from coarser resolution maps rather than visual interpretation or in situ data collection. In areas where there are problems, LACO-Wiki could be used to gather a large training data set to improve those classes where there is currently large confusion, provided the resources exist to collect field data. The algorithm for creating a sample along a road network may help in more efficiently collecting the training data.
- Make use of additional training data that can be collected by using additional sources of data besides very high-resolution imagery (e.g., bioclimatic layers, field size maps, geo-tagged photographs (e.g., from Flickr and Mapillary)). There are numerous automatic object recognition algorithms that could classify photos into land cover types. This may be an additional source of training data to supplement the data that was used in creating the ESA African land cover map.
- Interact closely with local experts, e.g., to provide local insights into land cover types that are specific to an area of a country. The wetland example in Ivory Coast provides good evidence of the need to involve local experts. However, local experts will only provide additional value if they work very closely with the person who is involved as the global expert for training data collection. We have also experienced that validation points from local experts will need to be checked and possibly corrected since their personal view and interpretation of land cover classes many times does not match the general definition of the class applied. Hence just relying on local experts alone might result in unsatisfactory classifications. This issue becomes evident when classifications from different local experts are compared and large disagreements can occur. Hence there is a need to balance this interaction.
- Use additional sources of remote sensing imagery in the classification, e.g., Sentinel 1 in addition to Sentinel 2 or other imagery (e.g., Landsat) in a sensor fusion approach. This may help to filter out some of the noise, e.g., the occurrence areas of cropland in forest areas in Gabon.

References

Fritz, S. and See, L.M. 2008. Quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology*, 14, 1-23.

Lesiv, Myroslava, Steffen Fritz, Ian McCallum, N.E. Tsendbazar, Martin Herold, Jean-François Pekel, M. Buchhorn, B. Smets, and et al. 2017. 'Evaluation of ESA CCI Prototype Land Cover Map at 20m. IIASA Working Paper. IIASA, Laxenburg, Austria: WP-17-021'. 2017.