Space Climate Observatory, Climate Change in the Chad LAc Region (ECLAT)

Final Report

CLS-ENV-RP-22-0142

Issue: 2.1

Date: 04/07/2022

Limited distribution/Diffusion limitée/Distribución limitada

CLS Lille: 61 rue de la Cimaise-Green Park-Building C - 59650 Villeneuve d'Ascq Tel +33 (0)3 20 72 53 64
CLS Head Office :11 rue Hermès, Parc Technologique du Canal - 31520 Ramonville-Saint-Agne, France
Tel +33 (0)5 61 39 47 00 Fax +33(0)5 61 75 10 14

SUMMARY

LIST OF TABLES :

LIST OF FIGURES :

LIST OF ABBREVIATIONS IN FRENCH

BDB Database
CESBIOCentre for Space Studies of the Biosphere
CNESCentre national d'études spatiales
CSECentre de suivi écologique au Sénégal
ECLATlimate change in the Lake Chad region
IOTA-2 Infrastructure for automatic land use processing OCSOland use
UN United Nations
FIFO Sentinel Product Exploitation Platform
RGB Red Green Blue
S2 Sentinel-2
TCPT Population Growth Rate
TCT Land Consumption Rate
TCTTCP TCT*/TCP*
EUEuropean Union

LIST OF ABBREVIATIONS IN ENGLISH

AOIArea of Interest
ICC Climate Change Initiative
DHI Dynamic Habitat Index
ESAEuropean Space Agency
FAOFood and Agriculture Organisation
G3WBMGlobal 3 arc-second Water Body Map
GFSAD30AFCEGlobal Food Security-support Analysis Data Africa Cropland Extent 30 m
GHSLGlobal Human Settlement Layer
GLCGlobal Land Cover
GLCNMOGlobal Land Cover by National Mapping Organisations
GSFGlobal Surface Water
GUF Global Human Footprint
ISCGM International Steering Committee for Global Mapping
ISRICInternational Soil Reference and Information Centre LCLand Cover
LCCS Land Cover Classification System
MODIS Moderate-Resolution Imaging Spectroradiometer
NDVINormalized Difference Vegetation Index
NDWI Normalized Difference Water Index
NOAA National Oceanic and Atmospheric Administration
PSUPrimary Sampling Unit
SCO Space Climate Observatory
SDG Sustainable Development Goal
SRTMWB Shuttle Radar Topography Mission Water Body
SSUSecondary Sampling unit
UNUnited Nations
WSF World Settlement Footprint

1 Introduction

Following the inventory of databases carried out in Batch 1, a quantitative evaluation was carried out in this Batch 2 in order to select the databases used for training the Iota- 2 chain and producing land use maps. This quantitative analysis complements the qualitative analysis carried out in the first part and allowed a final score to be given to each database by combining the two parts. Some databases, depending on their results, can also serve as a reference for a given variable.

2 General methodology

In order to carry out this qualitative study, an analysis of the "sensitivity" and accuracy of each database was carried out. Sensitivity represents the number of items that are present in the reference database as well as in the database under study. Accuracy is defined as the number of items in the study database that are well classified in the same class as the reference database. In order to meet our sampling objective, we therefore favoured a selection from a database with a higher rate of precision than of sensitivity. Indeed, it is preferable to select samples that are certain to belong to the right class than to select a larger number of samples with a higher proportion of "misclassified".

The methodology first requires the construction of a reference set that will be used for the analysis of the databases. The construction method consists of carrying out a sampling plan of the area. Subsequently, a breakdown of the study area for each database was carried out and finally a comparison was made with the reference database.

2.1 Study areas

The Lake Chad Basin Region is bordered by 4 countries: Chad, Cameroon, Niger and Nigeria and covers an area of approximately 33 km². Following interest from other stakeholders such as the Centre de Suivi Écologique (CSE) in Senegal and the African Union, two other study sites were also analysed: the Tocc Tocc Park in Senegal with an area of about 12 km² and the W Park straddling Benin, Niger and Burkina Faso, with an area of about 42 km² (see Table 1 & Figure 1)

Table 1: Study areas and corresponding Sentinel-2 tiles

ID	Nom	Tuiles S2 traitées	Zones d'étude élargie pour les tests
AOI01	Lac Tchad	33PWP 33PUP 33PTP	33PTQ 33PTR 33PUQ 33PUR 33PVP 33PVQ 33PVR 33PWQ 33PWR
AOI02	Parc Tocc Tocc Sénégal	28QCD	28QDD
AOI03	Parc W Niger	-	31PCN 31PCP 31PDN 31PDP

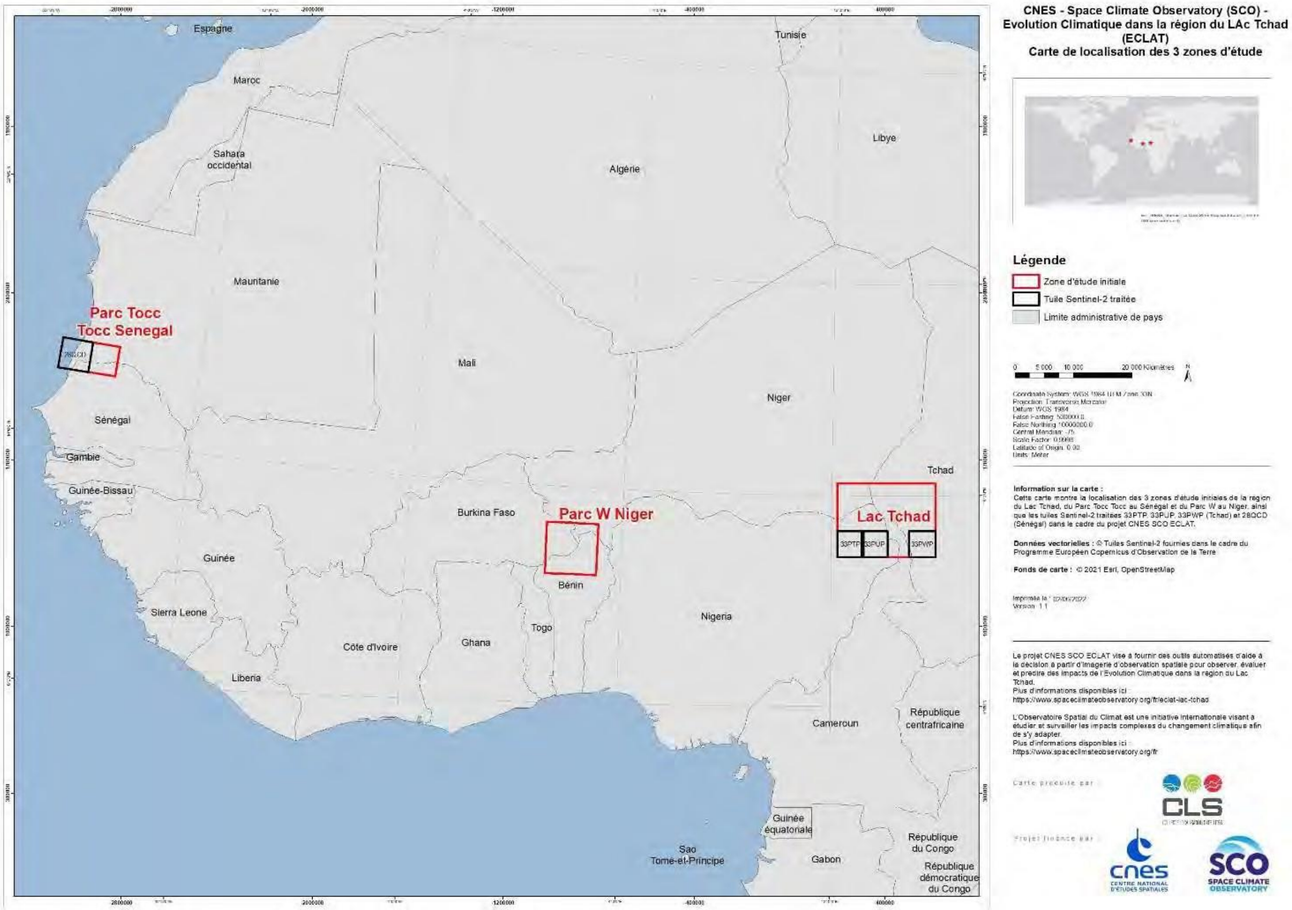


Figure 1: Location map of the study areas

2.2 Construction of the reference database by means of a sampling plan

A baseline database was constructed over the Lake Chad study area using a sampling method. A random selection followed by visual interpretation was carried out. In a first step, the study area was divided into a grid of 325 cells of 20 km x 20 km. Within each cell, a 2 km x 2 km *Primary Sampling Unit* (PSU) was randomly selected. This random sampling design ensures that the entire study area is covered and accounted for. In addition, to ensure that the baseline observations are spatially contiguous and independent, a random selection of 10 points was made within each PSU. This selection takes the form of a single pixel the size of a Sentinel-2 pixel at 10m. These pixels are called Secondary *Sampling* Units (SSUs). The 3250 SSUs were interpreted visually using the Collect Earth tool from Openforis. In addition, all available imagery was used to support the photo-interpretation, such as Google Earth and Bing Map backgrounds, high-resolution time series, in particular Sentinel-2, as well as Google Earth Engine for the calculation of the NDVI of the pixel over one year from Sentinel-2, Landsat-8 and Modis images. A time series visualisation was also used in particular to discern permanent water. Almost all visual interpretation (2950 SSU) was carried out by a photo-interpreter independent of the SCO ECLAT project production team, making the baseline independent of biases that may have come from the team already working on the databases in Batch 1. This method ensures a solid baseline for the rest of the project. However, errors of interpretation cannot be excluded. In

Indeed, the transition zones, especially concerning vegetation, are sometimes complex to interpret and can be a source of erroneous classification. In addition, the interpretation of agricultural areas can be difficult because some plots may be fallow or not in use at the time of acquisition. The lack of information on agricultural practices in the area was unfortunately a hindrance to the interpretation of some samples, which may have been misclassified as "arable land".

Nevertheless, the following definitions were used by the photo-interpreter to differentiate the classes:

- Artificial area: area covered by man-made or transformed by man, including urban sprawl, as well as all man-made structures
- Permanent water area: includes all surfaces covered by water at the time of image acquisition. These surfaces must be visible, in a liquid state and on the surface for > 95% of the year.
- Cultivable land: Cultivable land is an area that has the appearance of an agricultural plot and where the soil is temporarily covered by a crop and then by bare soil. We decide to classify points in this category if the NDVI shows one or two cycles over the year with NDVI values > 0.5 and < 0.3 on the bare soil phase (depending on the harvest phase).
- Wetland: an area that is not permanently wet but is covered by water for a significant part of the year
- Bare soil: area with exposed soil, sand or rocks and with NDVI never > 0.25
- Vegetation: Area with persistent plants and a constant NDVI over the year > 0.4
- Forest: Area with tree presence > 66% of the pixel

Table 1 below shows the number of SSUs analysed for each of the land use class variables.

Table 2: Number of SSU samples per class

Classes	Number of SSUs
Artificial area	12
Water surface	93
Cultivable land	441
Wetland	349
Bare ground	735
Vegetation	1418
Forest	17

No stratification was applied prior to this sampling phase due to time and budget constraints.

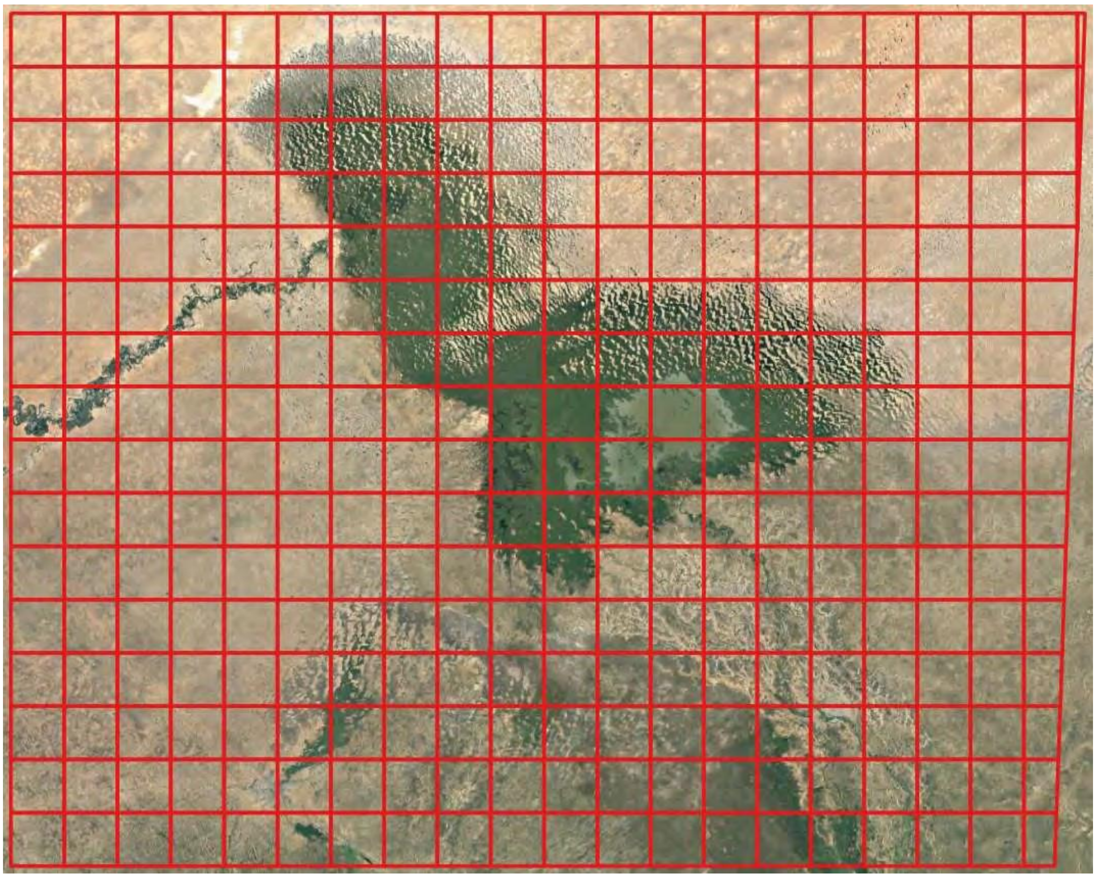


Figure 2: Sampling grid

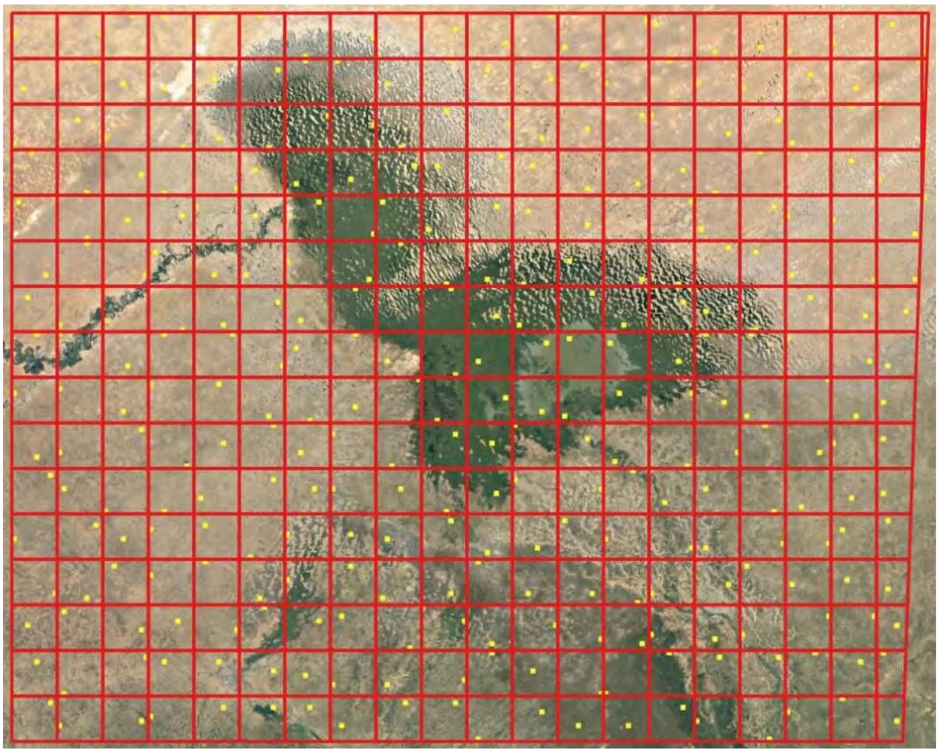


Figure 3: Map of randomly selected PSUs

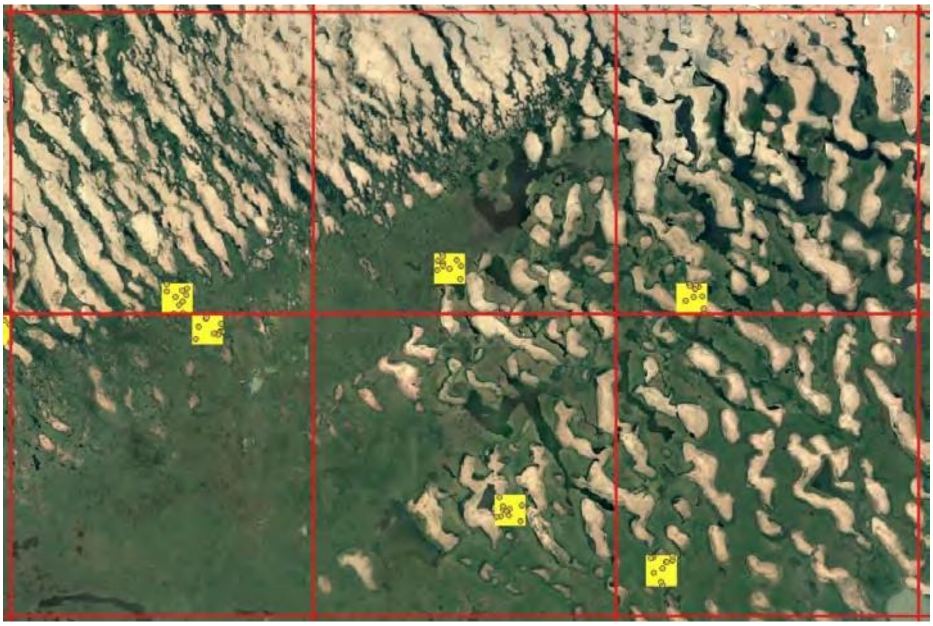


Figure 4: Example of SSU

2.3 Harmonisation of databases

The following Table 2 shows the groupings of classes that have been made on the databases to coincide with the nomenclature of the SCO ECLAT project.

Table 3: Harmonisation of databases with the SCO ECLAT nomenclature

	ESA CCI LC		Copernicus Land Cover		World Settlement Footprint	
SCO ECLAT	Code	Nomenclature	Code	Nomenclature	Code	Nomenclature
Artificial area	190	Urban areas	50	Urban / built up	255	Urban
Water surface	210	Water bodies	80	Permanent water bodies		
	10	Cropland rainfed				
	20	Cropland irrigated				
Cultivable land	30	Mosaic cropland	40	Cultivated and managed vegetation/agriculture		
	160	Tree cover flooded, fresh water				
	170	Tree cover flooded, saline water				
Wetlands	180	Shrub, herbaceous flooded, fresh / saline / brackish water	90	Herbaceous wetland		
Bare ground	200	Bare areas				
	40	Mosaic natural vegetation				
	110	Mosaic herbaceous > 50% / tree and shrub < 50				
	120	Shrubland	20	Shrubs		
	130	Grassland	30	Herbaceous vegetation		
	140	Lichens and mosses	100	Moss and lichen		
Vegetation	150	Sparse vegetation (tree, shrub, herbaceous cover)	60	Bare / sparse vegetation		
			111	Closed forest, evergreen needle leaf		
			113	Closed forest, deciduous needle leaf		
			112	Closed forest, evergreen, broad leaf		
			114	Closed forest, deciduous broad leaf		
			115	Closed forest, mixed		
			116	Closed forest, unknown		
	50	Tree cover, broadleaved, evergreen, closed to open	121	Open forest, evergreen needle leaf		
	60	Tree cover, broadleaved, decidious, closed to open	123	Open forest, deciduous needle leaf		
	70	Tree cover, needleaved, evergreen, closed to open	122	Open forest, evergreen broad leaf		
	80	Tree cover, needleaved, decidious, closed to open	124	Open forest, deciduous broad leaf		
	90	Tree cover, mixed leaf type	125	Open forest, mixed		
Forest	100	Mosaic tree and shrub > 50 % / herbaceous cover < 50	126	Open forest, unknown		

	ISRIC World Soil Information		FROM - GLC		bal Land Cover National Mapping Organization (GLCN)	
SCO ECLAT	Code	Nomenclature	Code	Nomenclature	Code	Nomenclature
Artificial area			80	Impervious surface	18	Urban
Water surface			60	Water	20	Water bodies
Cultivable land			10	Cropland	 11 12 13	Cropland Paddy field Cropland / other vegetation mosaic
Wetlands			50	Wetland	14 15	Mangrove Wetland
Bare ground	1 2 3 4 5 6 7 8 9 10 11 12	Clay Silty Clay Sandy Clay Clay Loam Silty clay loam Sandy clay loam Loam Silty Loam Sandy Loam Silt Loamy sand Sand	90	Bareland	16 17	Bare area consolidated Bare area not consolidated
Vegetation			30 40 70	Grassland Shrubland Tundra	7 8 9 10	Shrub Herbaceous Herbaceous with sparse shrub Sparse vegetation
Forest			20	Forest	1 2 3 4 5 6	Broadleaf Evergreen Forest Broadleaf Deciduous Forest Needleleaf Evergreen Forest Needleleaf Deciduous Forest Mixed Forest Tree Open

	MODIS Land Cover		Copernicus Water Bodies		Natural Earth water bodies	
SCO ECLAT	Code	Nomenclature	Code	Nomenclature	Code	Nomenclature
Artificial area						
Water surface			70 71 72 73 74 75 76	Water Very low occurrence Low occurrence Average occurrence High occurrence Very high occurrence Permanent Occurrence	2	Water body
Cultivable land						
Wetlands						
Bare ground						
Vegetation						
Forest	0-100	Percent tree cover				

	Earth Stat - Pasture & Cropland		Global Food Security support analysis Data 30m Africa Cropland Extent (GFSAD30AFCE)			
SCO ECLAT	Code	Nomenclature	Code		Nomenclature	
Artificial area						
Water surface						
Cultivable land	0-100%	Cropland > 15%.		2	Cropland	
Wetlands						
Bare ground						
Vegetation	0-100%	Pasture				
Forest						

	Natural Earth Urban		Africapolis		Global Tree Cover	
SCO ECLAT	Code	Nomenclature	Code	Nomenclature	Code	Nomenclature
Artificial area	1	Urban	1	Urban		
Water surface						
Cultivable land						
Wetlands						
Bare ground						
Vegetation						
Forest					0-100	Tree cover

	Global Surface Water (Pekel)		Global Lake and wetland database		Global 1 sec Water body Map (G1WBM) Yamazaki	
SCO ECLAT	Code	Nomenclature	Code	Nomenclature	Code	Nomenclature
Artificial area						
Water surface	0-100%	Water occurrence			30 50 51 99	Salt Marsh Permanent Water Permanent Water (Added by SWBD) Ocean (Given by external land/sea mask)
Cultivable land						
Wetlands					20 40	Wet Soil / Wet Vegetation / Lava Temporal Flooded Area
Bare ground						
Vegetation						
Forest						

	SRTM Water bodies (SRTMWB		Global Urban Footprint		Global Human Settlement	
SCO ECLAT	Code	Nomenclature	Code	Nomenclature	Code	Nomenclature
Artificial area			255	Urban	0 - 100 %	Urban
Water surface	255	Water				
Cultivable land						
Wetlands						
Bare ground						
Vegetation						
Forest						

2.4 Methodology for comparing databases

The databases were compared using statistical criteria derived from the evaluation of each database with the reference constructed by the sampling plan. These statistical data are intended to assess two important criteria:

- *Producer's Accuracy*, also known as *Recall*, which measures the number of well-ranked items relative to the total number in the class database
- *User's Accuracy*, which measures the accuracy of classification of items against the same class in the reference database

Each database was first cut out of the area and its nomenclature was reclassified to harmonise it with the project nomenclature. The two criteria described above are statistically translated into a confusion matrix which will be used to rank the databases by taking the criteria mentioned above, to which the *F-Score* and *Overall* Accuracy are added.

To summarise the statistical indices used are as follows:

- Reminder $\frac{\text{True positive}}{(\text{True positive+False positive})}$
- User accuracy = $\frac{\text{Vrai positif}}{(\text{True positive+False positive})}$
- F-Score = $2 * \frac{(\text{Accuracy*Sensitivity})}{(\text{Accuracy+Sensitivity})}$ é)
- Overall accuracy = $\frac{(\text{True positive+True negative})}{(\text{True positive+False positive+True negative+False negative})}$

Two other indicators are also possible to calculate:

- Omission rate = 1 - Recall
- Commission rate = 1 - User accuracy

These two indicators are used more generally in GIS to quickly assess the quality of a classification.

It should be noted that the rates and indices calculated above give an assessment of the database in the study area and may not reflect the rates obtained through the validation processes of the data producers if they exist.

3 Comparison of databases

3.1 Lake Chad

The Lake Chad Basin Region is bordered by 4 countries: Chad, Cameroon, Niger and Nigeria and covers an area of about 33 km² (see Figure 5: Location map of the Lake Chad Basin study area).

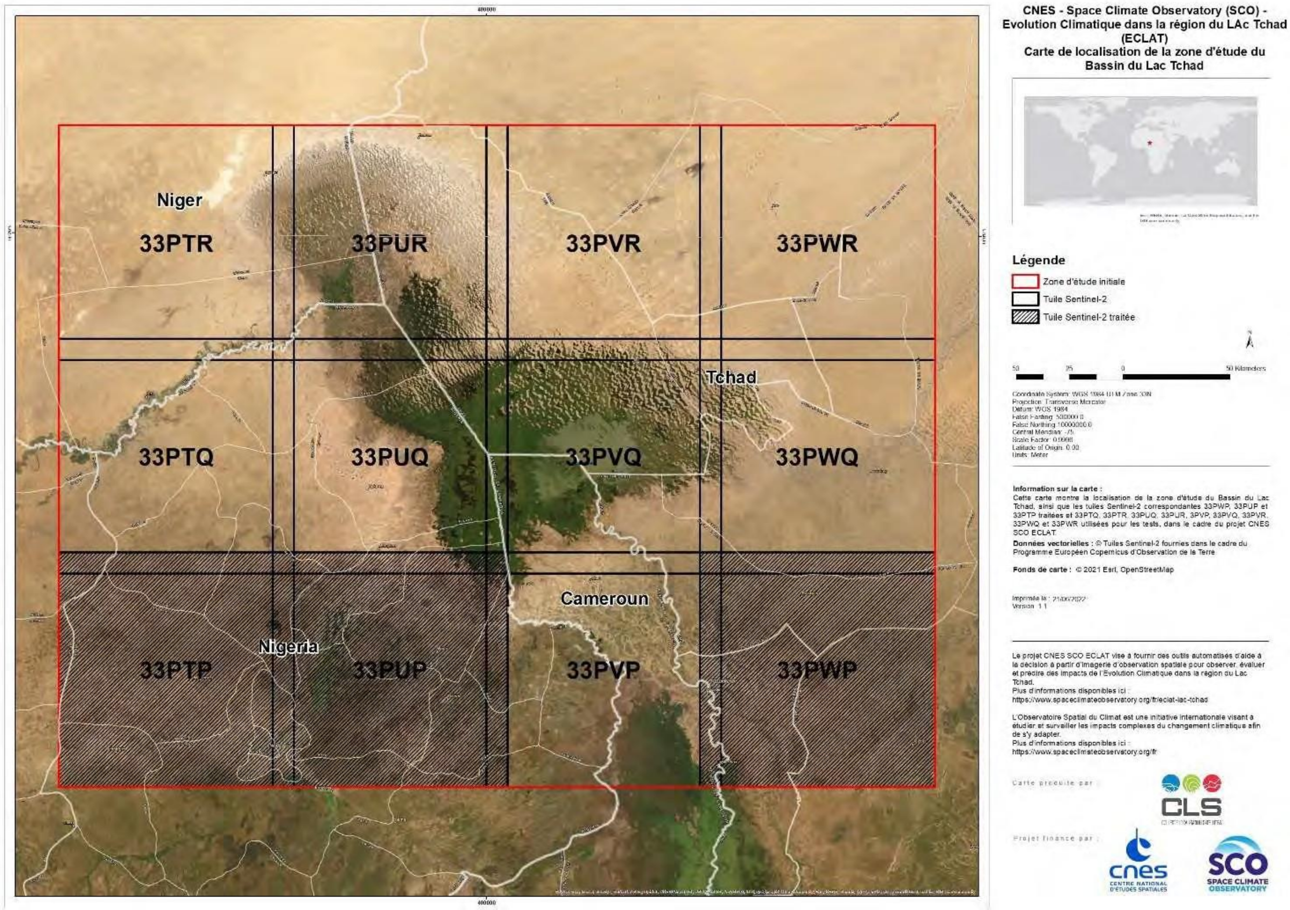


Figure 5: Location map of the Lake Chad Basin

3.1.1 Summary tables

In this section we present the results obtained for the statistical evaluation of the databases for all variables. Table 3 presents the results of the comparisons with the sample design. It groups the 3 statistical values described above for each variable available in each database. The colour code represents the result between 0 and 1, (e.g., 0 for the worst result and 1 for the best) ranging from red for 0 to green for 1.

Table 4: Comparison of DBs with the reference DB created by sampling

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
Africapolis	Artificial area	0,97	1	0,98	0,97
Copernicus Land cover	Total				0,58
Copernicus Land cover	Artificial area	0,9	0,35	0,51	

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
Copernicus Land cover	Water surface	1	0,43	0,6	
Copernicus Land cover	Cultivable land	0,45	0,42	0,44	
Copernicus Land cover	Wetland	0,68	0,79	0,73	
Copernicus Land cover	Bare ground	0,87	0,11	0,2	
Copernicus Land cover	Vegetation	0,56	0,85	0,68	
Copernicus Land cover	Forest	0,11	0,31	0,16	
Copernicus Water bodies	Water surface	0,68	1	0,8	0,68
Cropland2000	Cultivable land	0,1	1	0,16	0,1
ESA_CCI_LC	Total				0,4
ESA_CCI_LC	Artificial area	1	0,01	0,04	
ESA_CCI_LC	Water surface	0,77	0,47	0,59	
ESA_CCI_LC	Cultivable land	0,25	0,75	0,38	
ESA_CCI_LC	Wetlands	0,63	0,6	0,61	
ESA_CCI_LC	Bare ground	0,97	0,05	0,1	
ESA_CCI_LC	Vegetation	0,49	0,54	0,51	
ESA_CCI_LC	Forest	0	0	0	
From GLC	Total				0,47
From GLC	Artificial area	0,75	0,05	0,1	
From GLC	Water surface	0,96	0,58	0,73	
From GLC	Cultivable land	0,29	0,24	0,26	
From GLC	Wetland	0	0	0	
From GLC	Bare ground	0,96	0,1	0,1	
From GLC	Vegetation	0,47	0,89	0,61	
From GLC	Forest	0	0	0	
G3WBM_Yamazaki	Total				0,84
G3WBM_Yamazaki	Water surface	1	0,93	0,97	
G3WBM_Yamazaki	Wetland	0,73	1	0,85	
GFSAD30AFCE_2015	Total				0,39
GFSAD30AFCE_2015	Cultivable land	0,39	1	0,56	
GLCNMO	Total				0,45
GLCNMO	Artificial area	1	0	0,01	
GLCNMO	Water surface	0,86	0,4	0,55	
GLCNMO	Cultivable land	0,24	0,33	0,28	
GLCNMO	Wetland	0,57	0,62	0,59	
GLCNMO	Bare ground	0,36	0	0,01	
GLCNMO	Vegetation	0,49	0,74	0,59	

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
GLCNMO	Forest	0	0	0	
Global Human Settlement	Total				0,92
Global Human Settlement	Artificial area	0,92	1	0,96	
Pasture2000	Total				0,39
Pasture2001	Vegetation	0,39	1	0,56	
SRTMWB	Total				0,77
SRTMWB	Water surface	0,77	1	0,86	
World settlement footprint	Total				0,95
World settlement footprint	Artificial area	0,95	1	0,97	
WWF Global lake wetland	Total				0,07
WWF Global lake wetland	Water surface	0,07	0,94	0,14	
WWF Global lake wetland	Wetland	0,01	0,04	0,05	
Natural Earth lake	Water surface	0,9	1	0,95	0,9
Natural Earth urban areas	Artificial area	TBD	TBD	TBD	TBD
Modis Tree cover	Forest	0	0	0	0
NOAA Nightlights	Artificial area	0,66	1		0,8
MODIS Water mask	Water surface	0,81	1	0,89	0,81
Pekel_occurence (> 60%)	Water surface	0,96	1	0,98	0,96
Pekel_occurence (> 95%)	Water surface	1	1	1	1
Modis_land_cover_maryland	Total				0,43
Modis_land_cover_maryland	Artificial area	1	0,02	0,04	
Modis_land_cover_maryland	Water surface	1	0,32	0,48	
Modis_land_cover_maryland	Cultivable land	0,12	0,24	0,17	
Modis_land_cover_maryland	Wetland	0,28	0,02	0,04	
Modis_land_cover_maryland	Bare ground	0,86	0,23	0,36	
Modis_land_cover_maryland	Vegetation	0,49	0,77	0,6	
Modis_land_cover_maryland	Forest	0	0	0	
GLC_Share_FAO	Total				0,35
GLC_Share_FAO	Artificial area	1	0,004	0,008	
GLC_Share_FAO	Water surface	0,57	0,47	0,52	

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
GLC_Share_FAO	Cultivable land	0,18	0,74	0,29	
GLC_Share_FAO	Wetland	0,43	0,03	0,06	
GLC_Share_FAO	Bare ground	0,8	0,19	0,3	
GLC_Share_FAO	Vegetation	0,5	0,49	0,5	
GLC_Share_FAO	Forest	0	0	0	
World_pop	Artificial area	0,84	1	0,92	0,84
MODIS_bare_soil	Bare ground	0,66	0,99	0,8	0,67
MODIS_vegetation	Vegetation	0,49	0,99	0,65	0,5
Global Tree Cover	Forest	0	0	0	0
Radiant ML Hub		Interpretation impossible			

The "Radiant ML Hub" database was not evaluated because there is not enough data in the study area.

The following tables summarise the 3 parameters User Accuracy, Recall and F1-Score by class.

Table 5: Summary of DB results for the "Artificial area" variable

BD	Variable	User accuracy	Reminder	F1-Score
Africapolis	Artificial area	0,97	1	0,98
World settlement Footprint	Artificial area	0,95	1	0,97
Global Human Settlement	Artificial area	0,92	1	0,96
World_pop	Artificial area	0,84	1	0,92
NOAA Nightlights	Artificial area	0,66	1	0,8
GLC_Share_FAO	Artificial area	0,57	0,47	0,52
Copernicus Land cover	Artificial area	0,9	0,35	0,51
From GLC	Artificial area	0,75	0,05	0,1
ESA_CCI_LC	Artificial area	1	0,01	0,04
Modis_land_cover_maryland	Artificial area	1	0,02	0,04
GLCNMO	Artificial area	1	0	0,01

Table 6: Summary of DB results for the variable "Water surface".

BD	Variable	User accuracy	Reminder	F1-Score
Pekel_occurence (> 60%)	Water surface	0,96	1	0,98
G3WBM_Yamazaki	Water surface	1	0,93	0,97

BD	Variable	User accuracy	Reminder	F1-Score
Natural Earth lake	Water surface	0,9	1	0,95
MODIS Water mask	Water surface	0,81	1	0,89
SRTMWB	Water surface	0,77	1	0,86
Copernicus Water bodies	Water surface	0,68	1	0,8
From GLC	Water surface	0,96	0,58	0,73
Copernicus Land cover	Water surface	1	0,43	0,6
ESA_CCI_LC	Water surface	0,77	0,47	0,59
GLCNMO	Water surface	0,86	0,4	0,55
GLC_Share_FAO	Water surface	0,57	0,47	0,52
Modis_land_cover_maryland	Water surface	1	0,32	0,48
WWF Global lake wetland	Water surface	0,07	0,94	0,14

Table 7: Summary of DB results for the variable "Arable land".

BD	Variable	User accuracy	Reminder	F1-Score
GFSAD30AFCE_2015	Cultivable land	0,39	1	0,56
Copernicus Land cover	Cultivable land	0,45	0,42	0,44
ESA_CCI_LC	Cultivable land	0,25	0,75	0,38
GLC_Share_FAO	Cultivable land	0,18	0,74	0,29
GLCNMO	Cultivable land	0,24	0,33	0,28
From GLC	Cultivable land	0,29	0,24	0,26
Modis_land_cover_maryland	Cultivable land	0,12	0,24	0,17
Cropland2000	Cultivable land	0,1	1	0,16

Table 8: Summary of DB results for the "Wetland" variable

BD	Variable	User accuracy	Reminder	F1-Score
G3WBM_Yamazaki	Wetland	0,73	1	0,85
Copernicus Land cover	Wetland	0,68	0,79	0,73
ESA_CCI_LC	Wetland	0,63	0,6	0,61
GLCNMO	Wetland	0,57	0,62	0,59
GLC_Share_FAO	Wetland	0,43	0,03	0,06
WWF Global lake wetland	Wetland	0,01	0,04	0,05
Modis_land_cover_maryland	Wetland	0,28	0,02	0,04
From GLC	Wetland	0	0	0

Table 9: Summary of DB results for the variable "Bare soil".

BD	Variable	User accuracy	Reminder	F1-Score
MODIS_bare_soil	Bare ground	0,66	0,99	0,8
Modis_land_cover_maryland	Bare ground	0,86	0,23	0,36
GLC_Share_FAO	Bare ground	0,8	0,19	0,3
Copernicus Land cover	Bare ground	0,87	0,11	0,2
ESA_CCI_LC	Bare ground	0,97	0,05	0,1
From GLC	Bare ground	0,96	0,1	0,1
GLCNMO	Bare ground	0,36	0	0,01

Table 10: Summary of DB results for the variable "Vegetation

BD	Variable	User accuracy	Reminder	F1-Score
Copernicus Land cover	Vegetation	0,56	0,85	0,68
MODIS_vegetation	Vegetation	0,49	0,99	0,65
From GLC	Vegetation	0,47	0,89	0,61
Modis_land_cover_maryland	Vegetation	0,49	0,77	0,6
GLCNMO	Vegetation	0,49	0,74	0,59
Pasture2001	Vegetation	0,39	1	0,56
ESA_CCI_LC	Vegetation	0,49	0,54	0,51
GLC_Share_FAO	Vegetation	0,5	0,49	0,5

Table 11: Summary of DB results for the "Forest" variable

BD	Variable	User accuracy	Reminder	F1-Score
Copernicus Land cover	Forest	0,11	0,31	0,16
ESA_CCI_LC	Forest	0	0	0
From GLC	Forest	0	0	0
GLCNMO	Forest	0	0	0
Modis Tree cover	Forest	0	0	0
Modis_land_cover_maryland	Forest	0	0	0
GLC_Share_FAO	Forest	0	0	0
Global Tree Cover	Forest	0	0	0

3.1.2 Analysis of statistical results by variable

3.1.2.1 Zone artificial

The results obtained for the artificial areas are in line with those expected before the tests. The databases obtaining the best results are databases that have been developed with the objective of extracting precisely the urbanised areas on the surface of the globe or in the case of "Africapolis" for the African continent. Thus for "Africapolis", the "World Settlement Footprint" and "Global

Human Settlement", the user accuracy and sensitivity are very high and will probably be used in their final ranking.

A future approach is to "fix" the artificial areas that intersect the most relevant databases so that these areas do not have to be recalculated in the SCO ECLAT products. This also guarantees a higher classification accuracy by knowing a priori the class of a group of pixels. This can be achieved in particular for large agglomerations. For small villages or towns, a follow-up will always be carried out afterwards.

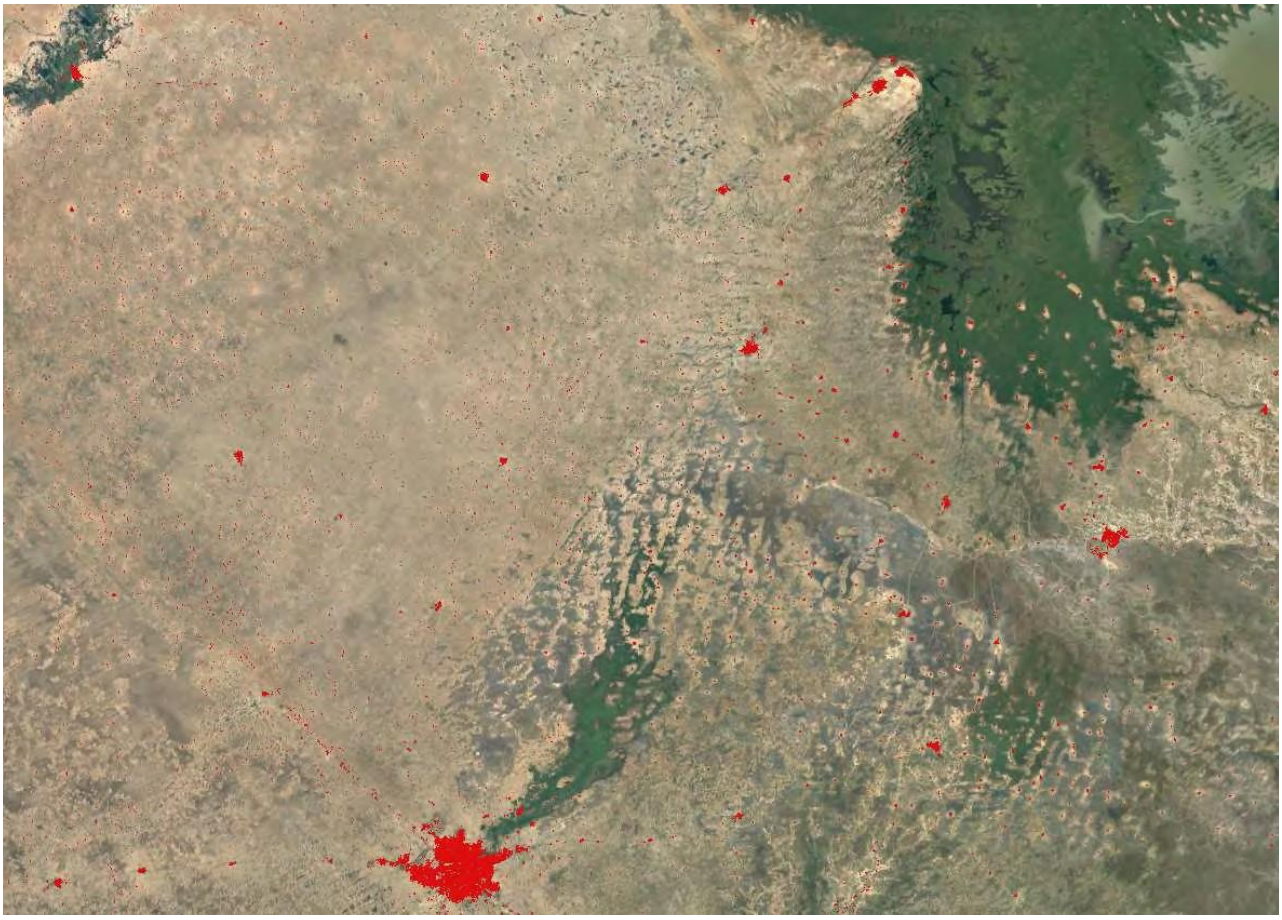


Figure 6: Illustration of the World Settlement Footprint product (red) on a Google satellite background

3.1.2.2 Water surface & area

The number of relevant databases on the water surface variable are numerous, as the topic has been the subject of a very large number of studies. The best databases are "Global Surface Water - Peke" and "G3WBM - Yamazaki". For the GSW - Pekel database, the comparison was made for an occurrence greater than or equal to 95%. Indeed, this is considered the only occurrence for permanent water1. The delineation of permanent open water is mainly located in the central part of Lake Chad where water is clearly visible. Almost all the random sampling of points was carried out on this stretch of water, which also explains the good results of low spatial resolution databases such as "Natural Earth" or "SRTMWB". A study of long data series over smaller areas will probably be necessary to delimit the other permanent water areas. The case of "Copernicus Water Bodies" is more

---

[1] Pekel, J., Vancutsem, C., Bastin, L., Clerici, M., Vanbogaert, E., B. E., & Defourny, P. (2014). A near real-time water surface detection method based on HSV transformation of MODIS multi-spectral time series data. Remote Sensing of Environment, 704-716

questionable. Some pixels in the database were incorrectly assigned to the class "Wetland" (commission error).

The Yamazaki database contains a "Wetland" class in its nomenclature. Its relevance is good on this variable. Other databases such as "GSW-Pekel" and "Copernicus Water Bodies" have water occurrence layers that can indicate the location of wetlands. Their results are inferior to Yamazaki's, however they could be used as a basis for cross-checking wetlands.

The detectability of wetlands can also be done with floating aquatic vegetation. Indeed, on Lake Chad, a large part of the wetlands are covered by vegetation and the a priori knowledge of a wetland (for the sampling plan) was done by visual interpretation of a series of images. However, automatic image classification will not be able to classify some areas completely covered by vegetation over long periods of time if water is not detectable on some dates. This may explain why some databases have not classified pixels as wetlands because some water bodies are never free. However, vegetation can be an indirect indicator of hydrodynamics. Its difficult characterisation can nevertheless be an obstacle in the detection of these zones.

For the wetland variable, a better assessment will be carried out during the production of the land use maps with Iota-2 where the extent of wetlands will be better characterised for these annual studies.

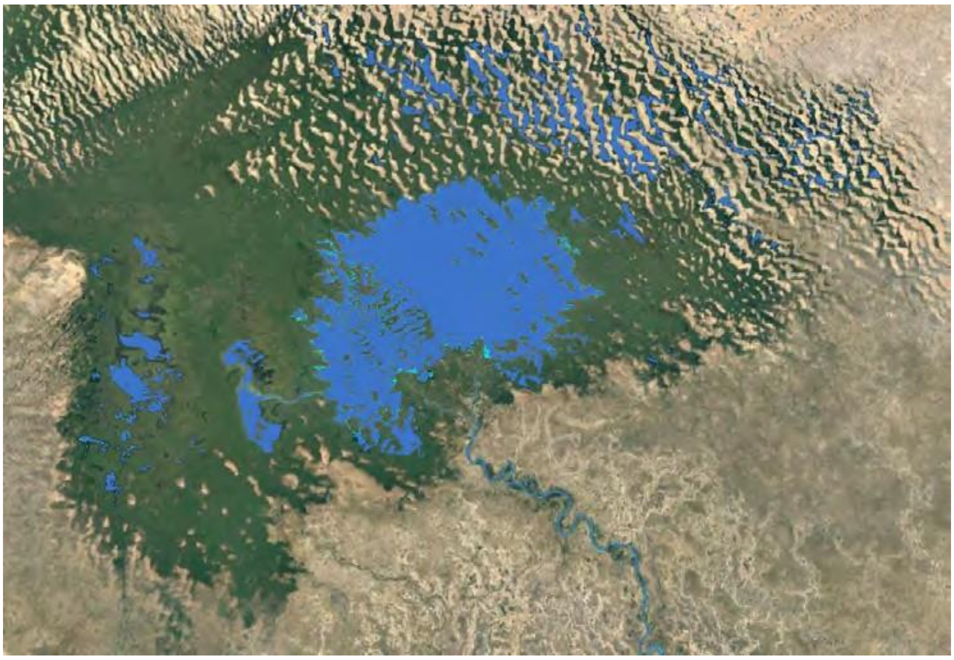


Figure 7: Illustration of the G3WBM Yamazaki database (blue) on a Google satellite background

3.1.2.3 Land cultivable

The databases for arable land have relatively low statistical results in terms of accuracy. Indeed, the values do not exceed the 45% threshold which will be a potential problem for future sample selection steps. The sensitivity values are better, with one database reaching 100% (GFSAD30AFCE_2015) and others evolving in values around 75%. These low values in terms of accuracy can a priori be explained by a dichotomy between the date of production of the database and the date of the images on which the sampling plan was carried out. In the hypothesis that some agricultural plots may be abandoned or others appear between these different dates, this would lower the accuracy rate drastically in some cases. During the sample design and characterisation activities, questions were raised about this problem.

Confusion in the area between arable land, wetlands (in case of presence of water) and bare soil (in case of dry weather and/or drought) can also explain potential discrepancies between the databases and the sampling plan.

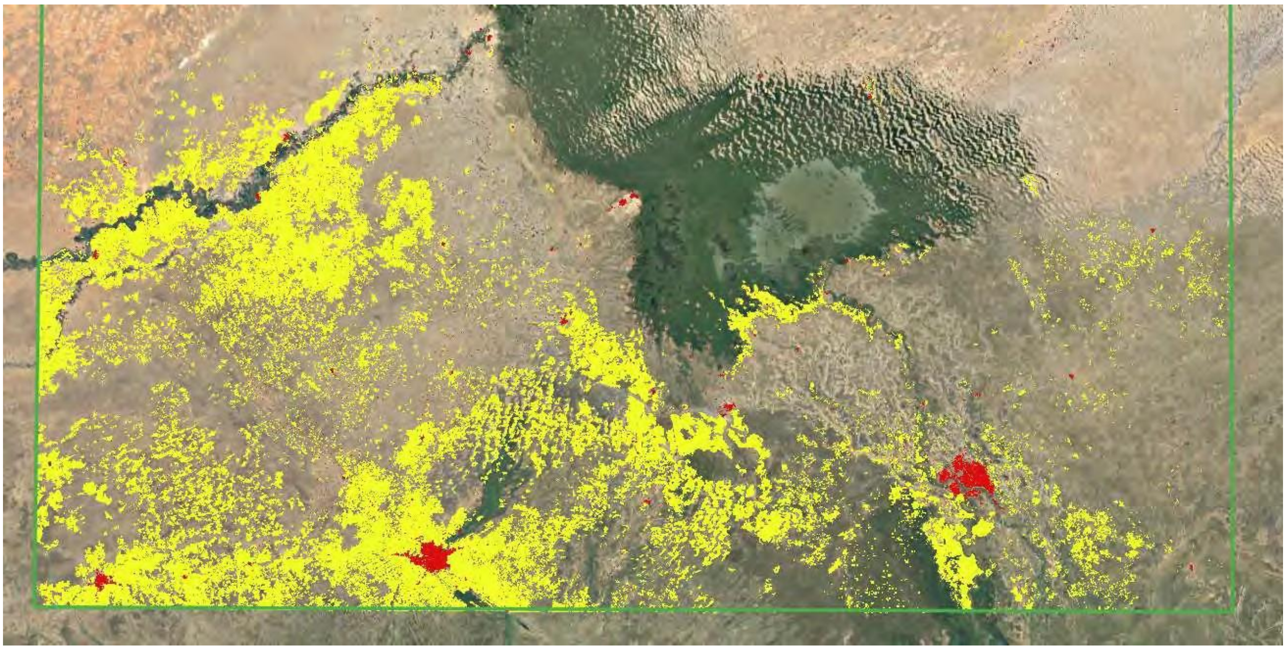


Figure 8: Illustration of arable land (GFSAD30AFCE 2015) in yellow and urban area (red) on Google background

3.1.2.4 Soil bare

The results of the databases for the bare soil variable are good in accuracy but poor in sensitivity. Only the MODIS Bare soil database gives better results than the others. Paradoxically, its accuracy is lower than the other databases and its sensitivity is much higher. From a sample selection point of view, the reverse is desirable. Nevertheless, the two leading databases are from MODIS, and could potentially be cross-referenced in order to draw on the results of both. The other databases have extremely low sensitivity rates despite good statistical accuracy.

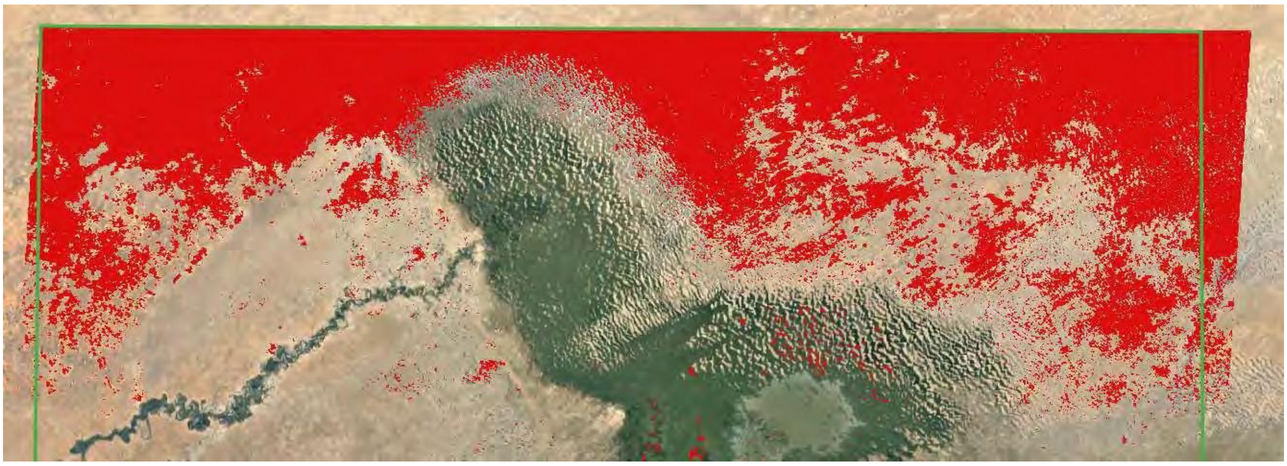


Figure 9: Bare soil illustration (MODIS Bare soil, in red) on Google Satellite background

3.1.2.5 Vegetation

For the vegetation variable, the results are average and will not allow the use of a certain number of databases, notably because the precision values are not very good. Indeed, it is preferable to have a database with average sensitivity and very good precision for the purpose of drawing samples. Since the commissions are low, it is highly likely that samples will be drawn from the target class.

It should also be noted that the vegetation here does not include floating vegetation, which is part of the wetlands. In fact, the term vegetation in this case includes low vegetation, high vegetation, pastures and erratic vegetation. As the wetlands of the lake are almost entirely covered by vegetation, we have decided to classify these areas as wetlands because the presence of water favours the development of vegetation which disappears during periods when water is no longer present.

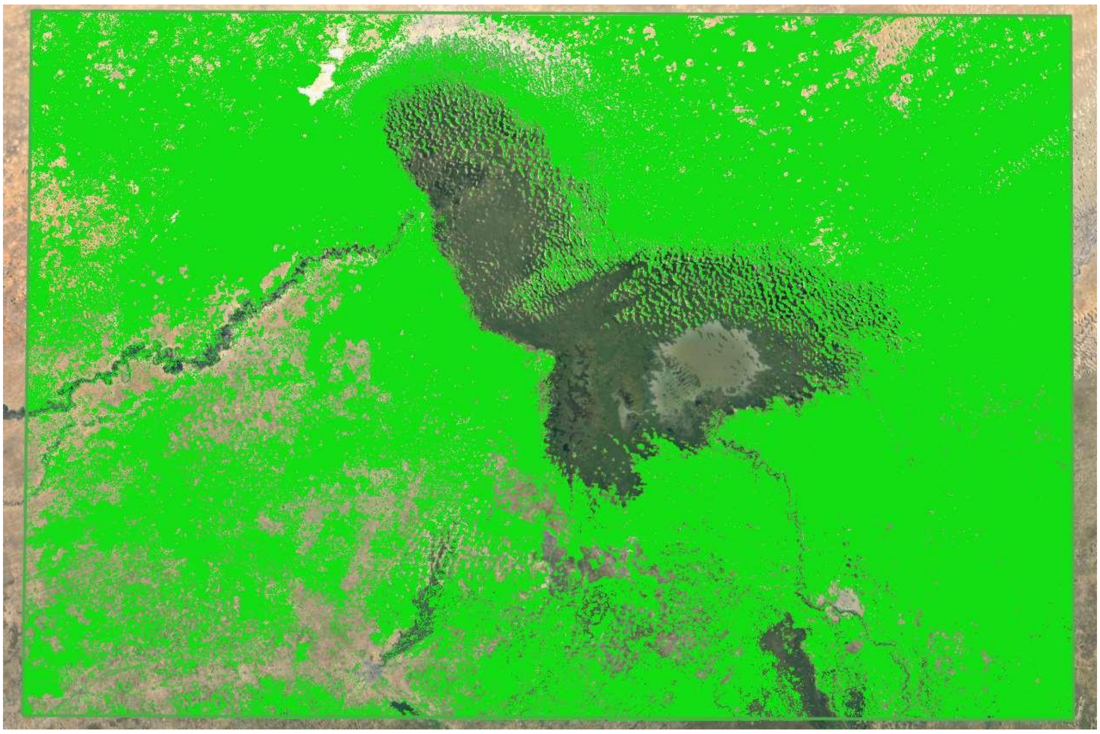


Figure 10: Illustration of vegetation seen by Copernicus Land Cover (green) on Google Satellite background

3.1.2.6 Forest

The forest variable has no database that can be used to select samples in this class. As the presence of forest is extremely low or non-existent, the relevance of keeping this variable in this study area is questionable. The number of forest points in the sample design is low and most databases do not have any pixels classified as forest in this area, while others have some but which do not intersect those in the sample design. These reasons explain the zero and near zero statistics (for one database) for this variable.

3.1.3 Final ranking for the Lake Chad area

The rankings achieved in Batch 1 have been included in the following tables. These have been supplemented by an additional F-Score criterion:

Table 12: F-Score

Criterion 5: Statistical Performance - F-Score	Points
F-Score > 20	Points ∈ [3:1] $\frac{1}{40} * (i - 40) + 1.5$
F-Score < 20	1

In order to make the statistical results more meaningful, it was decided to assign an F-Score between 1 and 3 to the databases according to an affine function. This makes it possible to differentiate the databases that have obtained a better result in terms of accuracy and sensitivity. All databases below 20% in F-Score automatically receive 1 point.

In addition, each criterion was given a different weighting according to the importance of each criterion compared to the others. For example, since criterion 5 is more important than all the others, it gets the maximum number of points and is therefore ranked first with the highest weight. We therefore have a distribution of weights as follows:

Table 13: Weight distribution

	Criterion 1	Criterion 2	Criterion 3	Criterion 4	Criterion 5
	1 > 2	2 > 3	3 >4	-	5 > 1
	1 > 3	2 > 4			5 > 2
	1 > 4				5 > 3
					5 > 4
Weight + 1	4	3	2	1	5

The following tables show the final ranking according to the different variables:

Table 14: Urban variable

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-
Global Human Settlement	Global / 10m	3	2	3	3	2,9	41,5
World Settlement footprint	Global / 10m	3	3	2	1	2,925	40,62
WorldPop	Global / 100m	2	2	3	3	2,8	37

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Global Urban Footprint	Global / 12m or 84m	3	3	1	1	2,5	36,5
Africapolis	Continental	3	2	1	1	2,95	35,75
FROM - GLC	Global / 30m	3	3	2	3	1	33
CCI Land Cover Africa 2016	Continental / 20m	3	3	3	1	0,5	30,5
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	2	1	30
NOAA Nightlights	Global / 1 km	1	2	1	1	2,5	25,5
MODIS Land Cover	Global / 500m	1	3	3	3	0,6	25
GLC - Share FAO	Global / 1km	1	2	1	1	1,8	22
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	0,525	17,62
Natural Earth	Overall / 1km	1	1	1	1	1	15

Table 15: Water surface variable

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-
GSW (Pekel)	Global / 30m	3	3	3	3	3	45
Yamazaki	Global / 100 & 30m	3	2	3	3	2,92	41,62
FROM-GLC	Global / 30m	3	3	2	3	2,32	39,62
SRTMWB	Global / 30m	3	3	2	1	2,65	39,25

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
MODIS Water surface	Global / 250m	1	3	3	3	2,72	35,62
Copernicus Global Land Service Water Bodies	Global / 30m	3	3	3	3	1	35
Natural Earth	Global / 10m	3	1	1	1	2,87	32,37
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1,97	31,87
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	1	31
CCI Land Cover Africa 2016	Continental / 20m	3	3	3	1	0,5	30,5
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1,87	24,37
GLC - Share FAO	Global / 1km	1	2	1	1	1,8	22
Global Lake and Wetlands database	Overall / 100m	2	1	1	3	0,85	20,25

Table 16: Variable Arable land

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
FROM-GLC	Global / 30m	3	3	2	3	1,15	33,75
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	1,45	33,25
GFSAD30AFCE	Global / 30m	3	3	2	1	1,2	32
MODIS Land Cover	Global / 500m	1	3	3	3	1,9	31,5
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	0,9	26,5
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1,22	21,12
EarthStat - Pasture and Cropland area	Overall / 10km	1	3	1	1	1	21
GLC - Share FAO	Global / 1km	1	2	1	1	1,22	19,12

Table 17: Wetland

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Yamazaki	Global / 100&. 30m	3	2	3	3	2,625	40,12
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	2,325	37,62
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	2,025	32,12
FROM-GLC	Global / 30m	3	3	2	3	0,5	30,5

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1,975	24,87
Global Lake and Wetlands database	Overall / 100m	2	1	1	3	1	21
GLC - Share FAO	Global / 1km	1	2	1	1	1	18

Table 18: Bare soil

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-
MODIS Bare soil layer	Global / 250m	1	3	3	3	2,5	34,5
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	1	31
MODIS Land Cover	Global / 500m	1	3	3	3	1,4	29
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27
ISRIC World Soil information	Global / 250m	1	3	3	2	0,75	24,75
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
GLC - Share FAO	Global / 1km	1	2	1	1	1	18
HARMONIZED WORLD SOIL DATABASE	Overall / 100m	2	2	1	1	0	17

Table 19: Vegetation

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
FROM-GLC	Global / 30m	3	3	2	3	2,02	38,12
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	2,2	37
MODIS Vegetation	Global / 250m	1	3	3	3	2,12	32,62
MODIS Land Cover	Global / 500m	1	3	3	3	2	32
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1,77	30,87
CCI Land Cover Africa 2016	Continental / 20m	3	3	3	1		28
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1,97	24,87
GLC - Share FAO	Global / 1km	1	2	1	1	1,75	21,75
EarthStat - Pasture and	Overall / 10km	1	3	1	1	1	21

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Cropland area							

Table 20: Forestry

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
FROM-GLC	Global / 30m	3	3	2	3	2,02	38,12
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	2,2	37
MODIS Vegetation	Global / 250m	1	3	3	3	2,12	32,62
MODIS Land Cover	Global / 500m	1	3	3	3	2	32
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1,77	30,87
CCI Land Cover Africa 2016	Continental / 20m	3	3	3	1		28
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1,97	24,87
GLC - Share FAO	Global / 1km	1	2	1	1	1,75	21,75
EarthStat - Pasture and Cropland area	Overall / 10km	1	3	1	1	1	21

We therefore use the following databases to draw the samples:

Table 21: DBs retained in the Lake Chad area

BD	Thematic
Global Human Settlement	Artificial area
GSW (Pekel)	Water surface
G3WBM Yamazaki	Water surface & Wetland
FROM-GLC	Arable land & vegetation
MODIS Bare soil layer	Bare ground

3.1.4 CESBIO products in the Lake Chad area

Tables 22 and 23 present the results of the classifications in the form of a confusion matrix and statistical indicators (i.e. global precision, user precision, recall and F-Score) over the Lake Chad Basin area. The classifications were produced with the Iota-2 chain developed by CESBIO and are based solely on Copernicus Land Cover (LC) data at 100m spatial resolution and on the Sentinel-2 annual time series for the reference year 2019.

Table 22: Confusion matrix for the Lake Chad area

Class	Artificial area	Water surface	Cultivable land	Wetland	Bare ground	Vegetation	Forest	Total classes	User accuracy
Artificial area	27	0	0	1	0	28	0	56	48,21
Water surface	0	45	0	44	0	4	0	93	48,38
Earth cultivable	0	0	169	2	2	254	0	427	39,57
Wetland	0	0	22	274	0	49	1	346	79,19
Bare ground	0	0	20	1	51	332	2	406	12,56
Vegetation	0	0	190	57	6	1104	3	1360	81,17
Forest	0	0	0	3	0	11	3	17	17,64
Total reference	27	45	401	382	59	1782	9	2705	
Producer accuracy	100	100	42,14	71,72	86,44	61,95	33,33		

Table 23: Summary of statistical indicators for the Lake Chad area

Class	n Truth	n Classified	Accuracy global	Accuracy user	Reminder	F Score
Artificial area	27	56	98,93	0,48	1	0,65
Water surface	45	93	98,23	0,48	1	0,65
Cultivable land	401	427	81,89	0,4	0,42	0,41
Wetland	382	346	93,35	0,79	0,72	0,75
Bare ground	59	406	86,58	0,13	0,86	0,22
Vegetation	1782	1360	65,47	0,81	0,62	0,7
Forest	9	17	99,26	0,18	0,33	0,23

Overall accuracy	61,85%

3.2 Park W and Park Tocc area Tocc

In this section, we present the studies on Park W and the Tocc Tocc area in Senegal, to which we applied the same database evaluation protocol as for the Lake Chad area. We therefore present the results of the studies directly.

3.2.1 Park W

Park W is a protected area straddling Benin, Niger and Burkina Faso, covering an area of about 42 km². Most of the Park is covered by vegetation, but this is tending to be reduced by agricultural activities (illegal or not).the Park area has been extended to cover the different land-use classes analysed (see Figure 11: Location map of Park W).

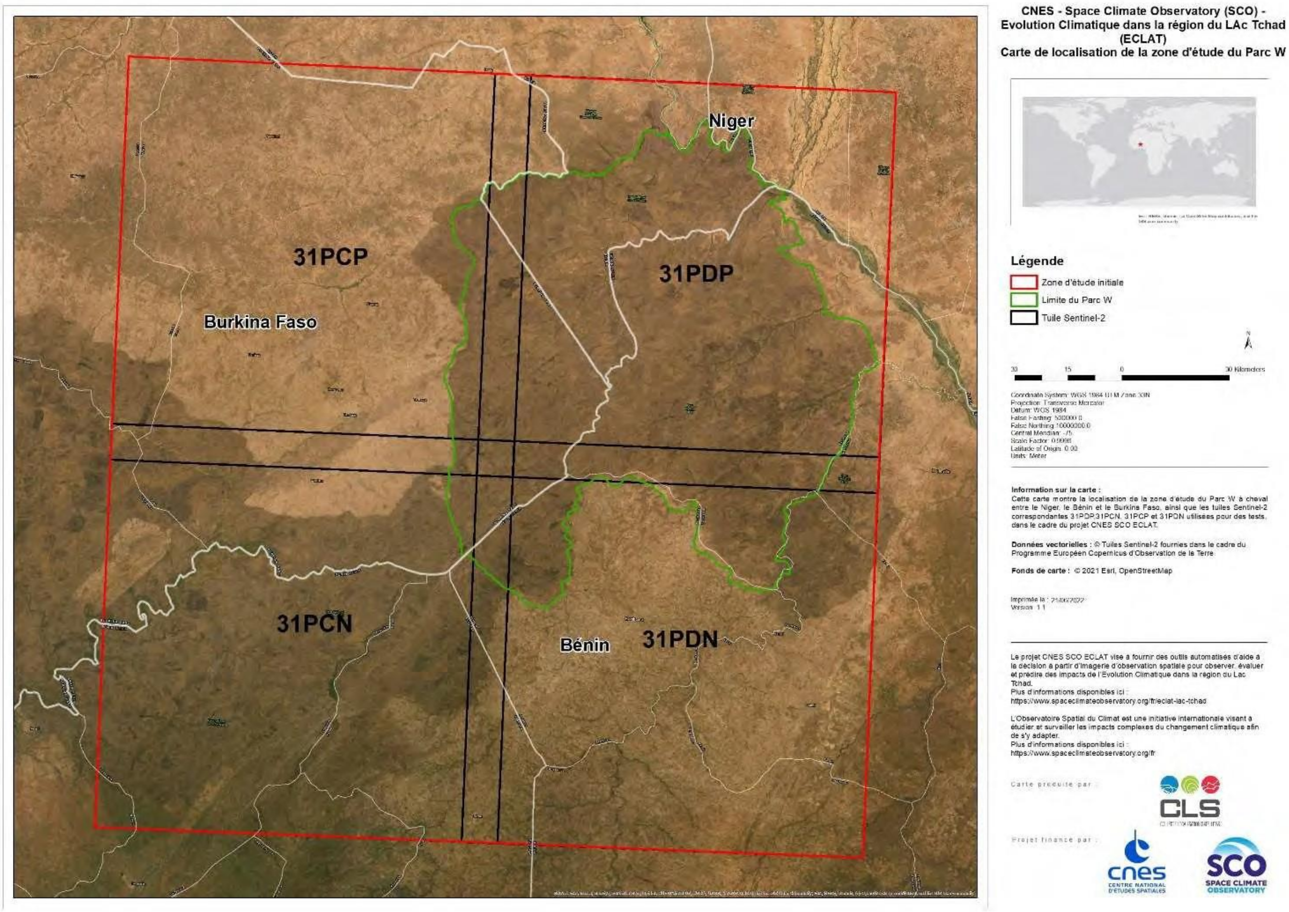


Figure 11: Location map of the Park W

3.2.1.1 Summary tables

Table 24: Database results for the Park area W

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
Africapolis	Artificial area				
Copernicus Land cover	Total				0,75
Copernicus Land cover	Artificial area	1	0,67	0,8	
Copernicus Land cover	Water surface	1	0,56	0,72	
Copernicus Land cover	Cultivable land	0,64	0,72	0,68	
Copernicus Land cover	Wetland	0,5	0,625	0,56	
Copernicus Land cover	Bare ground	0	0	0	
Copernicus Land cover	Vegetation	0,82	0,87	0,84	
Copernicus Land cover	Forest	0,22	0,3	0,26	
Copernicus Water bodies	Water surface	1	1	1	
Cropland2000	Cultivable land	0	0	0	
ESA_CCI_LC	Total				0,62
ESA_CCI_LC	Artificial area	1	0,1	0,17	
ESA_CCI_LC	Water surface	1	0,69	0,81	
ESA_CCI_LC	Cultivable land	0,31	1	0,48	
ESA_CCI_LC	Wetland	0	0	0	
ESA_CCI_LC	Bare ground	0	0	0	
ESA_CCI_LC	Vegetation	0,88	0,7	0,78	
ESA_CCI_LC	Forest	0,13	0,06	0,08	
From GLC	Total				0,64
From GLC	Artificial area	1	0,33	0,5	
From GLC	Water surface	1	0,75	0,86	
From GLC	Cultivable land	0,31	0,71	0,43	
From GLC	Wetland	0	0	0	
From GLC	Bare ground	0	0	0	
From GLC	Vegetation	0,78	0,78	0,79	
From GLC	Forest	0	0	0	
G3WBM_Yamazaki	Total				0,86
G3WBM_Yamazaki	Water surface	1	0,92	0,96	
G3WBM_Yamazaki	Wetland	0,33	1	0,5	
GFSAD30AFCE_2015	Total				0,68
GFSAD30AFCE_2015	Cultivable land	0,65	1	0,78	
GLC_Share_FAO	Total				
GLC_Share_FAO	Artificial area				
GLC_Share_FAO	Water surface				
GLC_Share_FAO	Cultivable land				
GLC_Share_FAO	Wetland				
GLC_Share_FAO	Bare ground				
GLC_Share_FAO	Vegetation				
GLC_Share_FAO	Forest				
GLCNMO	Total				0,32

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
GLCNMO	Artificial area	0	0	0	
GLCNMO	Water surface	0	0	0	
GLCNMO	Cultivable land	0,31	0,6	0,41	
GLCNMO	Wetland	0	0	0	
GLCNMO	Bare ground	0	0	0	
GLCNMO	Vegetation	0,64	0,3	0,41	
GLCNMO	Forest	0,1	0,66	0,17	
Global Human Settlement	Total				1
Global Human Settlement	Artificial area	1	1	1	
Global Tree Cover	Forest				
Modis Tree cover	Forest				
MODIS Water mask	Water surface				
MODIS_bare_soil	Bare ground	0,13	1	0,24	
Modis_land_cover_maryland	Total				0,68
Modis_land_cover_maryland	Artificial area	1	0,1	0,17	
Modis_land_cover_maryland	Water surface	0	0	0	
Modis_land_cover_maryland	Cultivable land	0,41	0,63	0,5	
Modis_land_cover_maryland	Wetland	0	0	0	
Modis_land_cover_maryland	Bare ground	0	0	0	
Modis_land_cover_maryland	Vegetation	0,76	0,91	0,82	
Modis_land_cover_maryland	Forest	0	0	0	
MODIS_vegetation	Vegetation	0,68	1	0,81	
Pekel_occurence	Water surface	1	0,86	0,92	
Pekel_occurence	Wetland	0,5	1	0,67	
SRTMWB	Total				
SRTMWB	Water surface				
World settlement footprint	Total				1
World settlement footprint	Artificial area	1	1	1	
World_pop	Artificial area	0,85	0,95	0,92	
WWF Global lake wetland	Total				0,35
WWF Global lake wetland	Water surface	0,4	0,86	0,55	
WWF Global lake wetland	Wetland	0	0	0	
ESA_CCI_LC_Africa_20m	Total				0,5
ESA_CCI_LC_Africa_20m	Artificial area	1	0,43	0,6	
ESA_CCI_LC_Africa_20m	Water surface	1	0,69	0,82	
ESA_CCI_LC_Africa_20m	Cultivable land	0,47	0,6	0,53	
ESA_CCI_LC_Africa_20m	Wetland	0	0	0	
ESA_CCI_LC_Africa_20m	Bare ground	0	0	0	
ESA_CCI_LC_Africa_20m	Vegetation	0,79	0,51	0,62	
ESA_CCI_LC_Africa_20m	Forest	0,11	0,66	0,19	

The following tables summarise the 3 parameters User Accuracy, Recall and F1-Score by class.

Table 25: Results for the artificial area variable in the Park area W

BD	Variable	User accuracy	Reminder	F1-Score
World settlement footprint	Artificial area	1	1	1
Global Human Settlement	Artificial area	1	1	1
World_pop	Artificial area	0,8	1	0,89
Copernicus Land cover	Artificial area	1	0,67	0,8
ESA_CCI_LC_Africa_20m	Artificial area	1	0,43	0,6
From GLC	Artificial area	1	0,33	0,5
ESA_CCI_LC	Artificial area	1	0,1	0,17
Modis_land_cover_maryland	Artificial area	1	0,1	0,17
GLCNMO	Artificial area	0	0	0
GLC_Share_FAO	Artificial area	0	0	0
Africapolis	Artificial area			

Table 26: Results for the variable water surface in the Park area W

BD	Variable	User accuracy	Reminder	F1-Score
Copernicus Water bodies	Water surface	1	1	1.0
MODIS Water mask	Water surface	1	1	1.0
G3WBM_Yamazaki	Water surface	1	0,92	0,96
Pekel	Water surface	1	0,86	0,92
From GLC	Water surface	1	0,75	0,86
ESA_CCI_LC_Africa_20m	Water surface	1	0,69	0,82
ESA_CCI_LC	Water surface	1	0,69	0,81
Copernicus Land cover	Water surface	1	0,56	0,72
WWF Global lake wetland	Water surface	0,4	0,86	0,55
GLC_Share_FAO	Water surface	0,04	0,81	0,1
Natural Earth lake	Water surface	0	0	0
GLCNMO	Water surface	0	0	0
Modis_land_cover_maryland	Water surface	0	0	0
SRTMWB	Water surface			

Table 27: Results for the variable cultivable land in the Park area W

BD	Variable	User accuracy	Reminder	F1-Score
GFSAD30AFCE_2015	Cultivable land	0,65	1	0,78
Copernicus Land cover	Cultivable land	0,64	0,72	0,68
ESA_CCI_LC_Africa_20m	Cultivable land	0,47	0,6	0,53
ESA_CCI_LC	Cultivable land	0,31	1	0,48
From GLC	Cultivable land	0,31	0,71	0,43
GLCNMO	Cultivable land	0,31	0,6	0,41

BD	Variable	User accuracy	Reminder	F1-Score
Modis_land_cover_maryland	Cultivable land	0,41	0,63	0,5
GLC_Share_FAO	Cultivable land	0,13	0,02	0,03
Cropland2000	Cultivable land	0	0	0

Table 28: Results for the wetland variable in the Park area W

BD	Variable	User accuracy	Reminder	F1-Score
Copernicus Land cover	Wetland	0,5	0,63	0,56
G3WBM_Yamazaki	Wetland	0,33	1	0,5
ESA_CCI_LC	Wetland	0	0	0
GLCNMO	Wetland	0	0	0
WWF Global lake wetland	Wetland	0	0	0
Modis_land_cover_maryland	Wetland	0	0	0
From GLC	Wetland	0	0	0
GLC_Share_FAO	Wetland	0	0	0
ESA_CCI_LC_Africa_20m	Wetland	0	0	0

Table 29: Results for the bare soil variable in the Park area W

BD	Variable	User accuracy	Reminder	F1-Score
MODIS_bare_soil	Bare ground	0,13	1	0,24
Modis_land_cover_maryland	Bare ground	0	0	0
GLC_Share_FAO	Bare ground	0	0	0
Copernicus Land cover	Bare ground	0	0	0
ESA_CCI_LC	Bare ground	0	0	0
From GLC	Bare ground	0	0	0
GLCNMO	Bare ground	0	0	0
ESA_CCI_LC_Africa_20m	Bare ground	0	0	0

Table 30: Results for the variable Vegetation in the Park area W

BD	Variable	User accuracy	Reminder	F1-Score
Copernicus Land cover	Vegetation	0,82	0,87	0,84
Modis_land_cover_maryland	Vegetation	0,76	0,91	0,82
MODIS_vegetation	Vegetation	0,68	1	0,81
From GLC	Vegetation	0,78	0,78	0,79
ESA_CCI_LC	Vegetation	0,88	0,7	0,78
ESA_CCI_LC_Africa_20m	Vegetation	0,79	0,51	0,62
GLCNMO	Vegetation	0,64	0,3	0,41
Pasture2001	Vegetation	0	0	0

BD	Variable	User accuracy	Reminder	F1-Score
GLC_Share_FAO	Vegetation	0	0	0

Table 31: Results for the variable Forest in the Park area W

BD	Variable	User accuracy	Reminder	F1-Score
Copernicus Land cover	Forest	0,22	0,3	0,26
ESA_CCI_LC_Africa_20m	Forest	0,11	0,66	0,19
GLCNMO	Forest	0,1	0,66	0,17
ESA_CCI_LC	Forest	0,13	0,06	0,08
From GLC	Forest	0	0	0
Modis Tree cover	Forest	0	0	0
Modis_land_cover_maryland	Forest	0	0	0
GLC_Share_FAO	Forest	0	0	0
Global Tree Cover	Forest	0	0	0

3.2.1.2 Analysis of results by variable

3.2.1.2.1 Artificial area

Artificial areas in this study area are less present but are still a variable where the database results are good. As in the case of the Lake Chad area, we can fix these areas from the World Settlement Footprint or the Global Human Settlement layer for example in order to use them as a mask in the classification so as not to recalculate these values.

3.2.1.2.2 Water surface & wetland

Water areas are also not very present in the area, apart from the Niger River in the north-eastern part of the area and one of its tributaries, the Mekrou, which crosses the park. The databases that obtain the best results in the zone are fewer in number than in the Lake Chad zone, but the main ones nevertheless remain at the top of the ranking. However, we can see that it is the higher resolution databases that stand out from the lower resolution ones. Indeed, in this area, unlike Lake Chad, the water areas are associated with a river and a stream, and therefore with narrower objects that are less easy to extract in terms of resolution.

Concerning wetlands, we note a regression of the Yamazaki database which reveals here a much lower relevance. Wetlands are more complicated to extract than on Lake Chad and less present. This may explain the disparity between the results obtained in the first and second study areas. It should be noted that the database with the best results for this area is the Copernicus Land Cover database.

3.2.1.2.3 Cultivable land

In this study area, the most relevant database is still "GFSAD30AFCE" (Global Food Security-support Analysis Data (GFSAD) Cropland Extent Africa 30m). It performs better than Lake Chad with an accuracy of 65% and a much higher F-Score of 0.78. Coperncius Land Cover also performs better. Both databases will be much more useful than on Lake Chad for extracting arable land. In terms of classification, better accuracy allows the selection of samples whose ground truth is indeed an agricultural plot.

We believe that some confusion still lies in the mixing of agricultural plots with bare soil for some and confusion with vegetation for others. The problem of abandoned plots such as on Lake Chad was also raised again during the design of the sampling plan and the confusion that could be caused in the classification of these plot areas.

3.2.1.2.4 Bare ground

This class shows poor results due to its near absence from the area. The subjective confusion between bare soil and low vegetation may also be at the origin of this confusion. Nevertheless, it is also noted that this class is not very present in the different databases studied. The very selection of samples in this area is questionable, as for the forest class on Lake Chad, in order not to introduce classification errors.

3.2.1.2.5 Forest

In this study area, as in Chad, the forest variable is almost non-existent and therefore produces poor results for all databases. As in the previous study area, the exclusion of this class should be considered when producing land cover maps.

3.2.1.3 Final ranking on the Park area W

As for Lake Chad, we use the method with the additional statistical criterion of the F1- Score. The following tables summarise the different results of the analysed databases according to the variables:

Table 32: Results for the Urban Area variable in the Park area W

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-
Global Human Settlement	Global / 10m	3	2	3	3	3	42
World Settlement footprint	Global / 10m	3	3	2	1	3	41
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	2	2,2	36
WorldPop	Global / 100m	2	2	3	3	2,56	35,8
ESA_CCI_LC Africa_20 m	Continental e / 20m	3	3	3	1	1,4	35
FROM - GLC	Global / 30m	3	3	2	3	1	33
CCI Land Cover Africa 2016	Continental e / 20m	3	3	3	1	1	33
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ISCGM - GLOBAL MAP	Global / 500m	1	2	1	3	1	20

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
GLOBAL LAND COVER (GLCNMO)							
GLC - Share FAO	Global / 1km	1	2	1	1	1	18
Africapolis	Continentale	3	2	1	1	-1	16
Global Urban Footprint	Global / 12m or 84m	3	3	1	1	NA	NA
NOAA Nightlights	Global / 1 km	1	2	1	1	NA	NA
Natural Earth	Overall / 1km	1	1	1	1	NA	NA

Table 33: Results for the variable water surface in the Park area W

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-
GSW (Pekel)	Global / 30m	3	3	3	3	2,68	43,4
Yamazaki	Global / 100 & 30m	3	2	3	3	2,84	41,2
FROM-GLC	Global / 30m	3	3	2	3	2,44	40,2
CCI Land Cover Africa 2016	Continentale / 20m	3	3	3	1	2,28	39,4
MODIS Water surface	Global / 250m	1	3	3	3	3	37
Copernicus Global Land Land Cover Service	Overall / 100m	2	3	3	3	1,88	35,4
Copernicus Global Land Service Water Bodies	Global / 30m	3	3	3	3	1	35
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	2,24	33,2

Global Lake and Wetlands database	Overall / 100m	2	1	1	3	1,2	22
BD	**Scale and resolution**	**Evaluation criterion 1 - Resolution**	**Assessment criterion 2 - Reliability**	**Assessment criterion 3 - Setting up a day**	**Evaluation criterion 4 - Genealogy**	**Assessment criterion 5 - F-Score**	**Total**
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	-0,6	10
SRTMWB	Global / 30m	3	3	2	1	NA	NA
Natural Earth	Global / 10m	3	1	1	1	NA	NA

Table 34: Results for the variable cultivable land in the Park area W

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Evaluation criteria 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-
GFSAD30AFCE	Global / 30m	3	3	2	1	2,12	36,6
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	1,72	34,6
CCI Land Cover Africa 2016	Continentale / 20m	3	3	3	1	1,12	33,6
FROM-GLC	Global / 30m	3	3	2	3	0,72	31,6
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	0,92	26,6
EarthStat - Pasture and Cropland area	Overall / 10km	1	3	1	1	1	21

ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	0,64	18,2
GLC - Share FAO	Global / 1km	1	2	1	1	-0,88	8,6

Table 35: Results for the wetland variable in the Park area W

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
G1WBM	Global / 100 & 30m	3	2	3	3	1.88	36.4
FROM-GLC	Global / 30m	3	3	2	3	1	33
Copernicus Global Land Land Cover Service	Overall / 100m	2	3	3	3	1,24	32,2
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
Global Lake and Wetlands database	Overall / 100m	2	1	1	3	1	21
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1	18

Table 36: Results for the bare soil variable in the Park area W

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Evaluation criteria 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Weight		4	3	2	1	5	-
Copernicus Global Land Land Cover Service	Overall / 100m	2	3	3	3	1	31
MODIS Bare soil layer	Global / 250m	1	3	3	3	1	27
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27

ISRIC World Soil information	Global / 250m	1	3	3	2	1	26
BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
HARMONIZED WORLD SOIL DATABASE	Overall / 100m	2	2	1	1	1	22
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1	18

Table 37: Results for the variable Vegetation in the Park area W

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
FROM-GLC	Global / 30m	3	3	2	3	2,16	38,8
Copernicus Global Land Land Cover Service	Overall / 100m	2	3	3	3	2,36	37,8
CCI Land Cover Africa 2016	Continental / 20m	3	3	3	1	1,48	35,4
MODIS Land Cover	Global / 500m	1	3	3	3	2,28	33,4
MODIS Vegetation	Global / 250m	1	3	3	3	2,24	33,2
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	2,12	32,6
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	0,64	18,2
EarthStat - Pasture and Cropland area	Overall / 10km	1	3	1	1	-1	11
GLC - Share FAO	Global / 1km	1	2	1	1	-1	8

Table 38: Result for the variable Forest in the Park area W

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
FROM-GLC	Global / 30m	3	3	2	3	1	33
BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
CCI Land Cover Africa 2016	Continental / 20m	3	3	3	1	1	33
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	1	31
MODIS Tree Cover	Global / 250m	1	3	3	3	1	27
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27
EarthStat - Pasture and Cropland area	Overall / 10km	1	3	1	1	1	21
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1	18

The databases selected are the following:

Table 39: BD Retentions in the Park area W

BD	Thematic
Global Human Settlement	Artificial area
Global Surface Water	Water surface
G1WBM	Wetland
FROM-GLC	Vegetation & Forest
GFSAD30AFCE	Cultivable land
Copernicus Global Land Service Land Cover	Bare ground

Nevertheless, except for water surfaces and artificial areas, the Copernicus Land Cover database comes second for the other variables and with a maximum difference of 2 points in the classification.

We plan to use this database as the basis for drawing samples in order to facilitate the various selection stages, but also because its accuracy for the classes in which it ranks second in the classification is close to the first.

3.2.1.4 Use of the Copernicus Hot Spot product

The Hot Spot Land Cover Change Explorer is a web-based information system providing detailed information on land cover and land cover changes over specific areas of interest. This Copernicus service supports field projects and international policies developed by the EU.

The data and products are based on medium to high and very high resolution satellite imagery (approximately 1 to 30 m spatial resolution) with a change assessment frequency of 1 to 20 years.

The classification system follows the Land Cover Classification System (LCCS) developed by the Food and Agriculture Organization of the United Nations (FAO).

Independently produced and validated land cover and land cover change maps and statistics are available to users worldwide.

We validated this product using the points obtained from the selection of validation points in the area. The results are as follows:

Table 40: Confusion matrix for the Park area W

Class	Artificial area	Water surface	Cultivable land	Wetland	Bare ground	Vegetation	Forest	Total Classes	User accuracy
Artificial area	14	0	2	0	0	2	0	18	77,8
Water surface	0	12	0	4	0	0	0	16	75
Cultivable land	0	0	57	1	1	4	0	63	90,5
Wetland	0	1	2	3	0	2	0	8	37,5
Bare ground	0	0	7	0	0	14	0	21	0
Vegetation	0	0	13	2	0	303	0	318	95,3
Forest	0	0	0	3	0	27	0	30	0
Total Reference	14	13	81	13	1	352	0		
Producer accuracy	100	92,3	70,4	23,1	0	86,1	-		

Table 41: Summary of statistical indicators for the Park area W

Class	n Truth	n Classified	User accuracy	Reminder	F Score
Artificial area	14	18	1	0,78	0,88
Water surface	13	16	0,92	0,75	0,83
Cultivable land	81	63	0,7	0,9	0,79
Wetlands	13	8	0,23	0,38	0,29
Bare ground	1	21	0	0	0
Vegetation	352	318	0,86	0,95	0,9
Forest	0	30	0	0	0
Overall accuracy	82%				

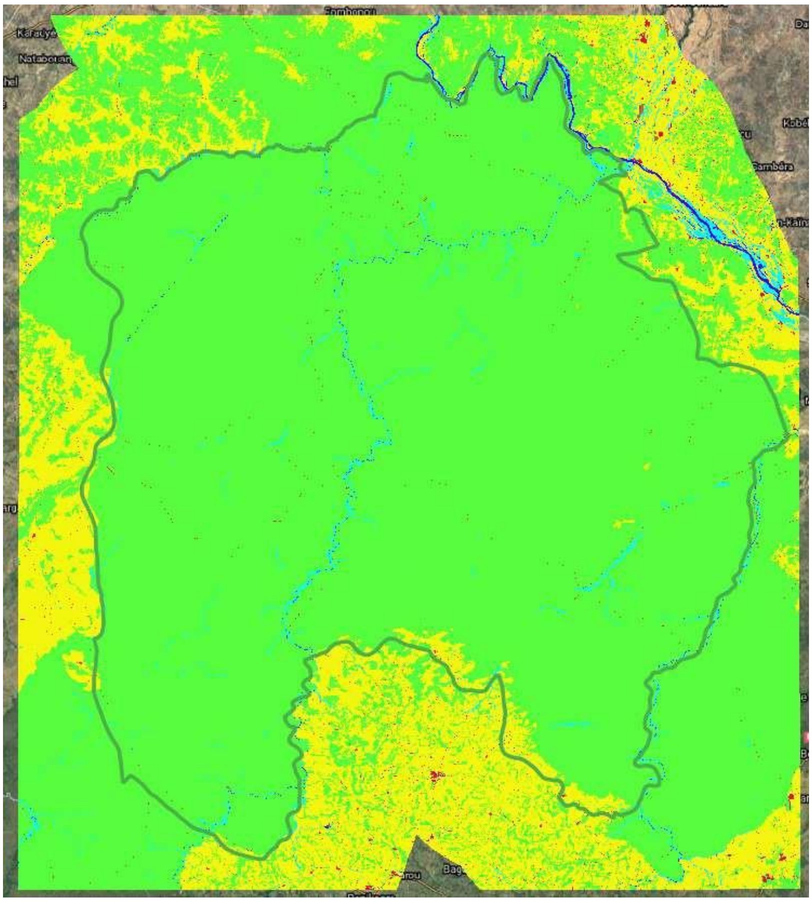


Figure 12: Illustration of the Hot Spot Land Cover Change product

3.2.2 Tocc Tocc Park at Senegal

Following the interest expressed by the Centre de Suivi Écologique (CSE) in Senegal, the Tocc Tocc Park in Senegal, with a surface area of approximately 12 km², was also analysed (see Figure 13: Map of the Tocc Tocc Park study area in Senegal). The Tocc Tocc Community Nature Reserve is classified as a UNESCO Biosphere Reserve and is a Ramsar site corresponding to a wetland of international importance. The Senegal River basin is a real reservoir of biodiversity, home to several vulnerable species of fish and water birds.

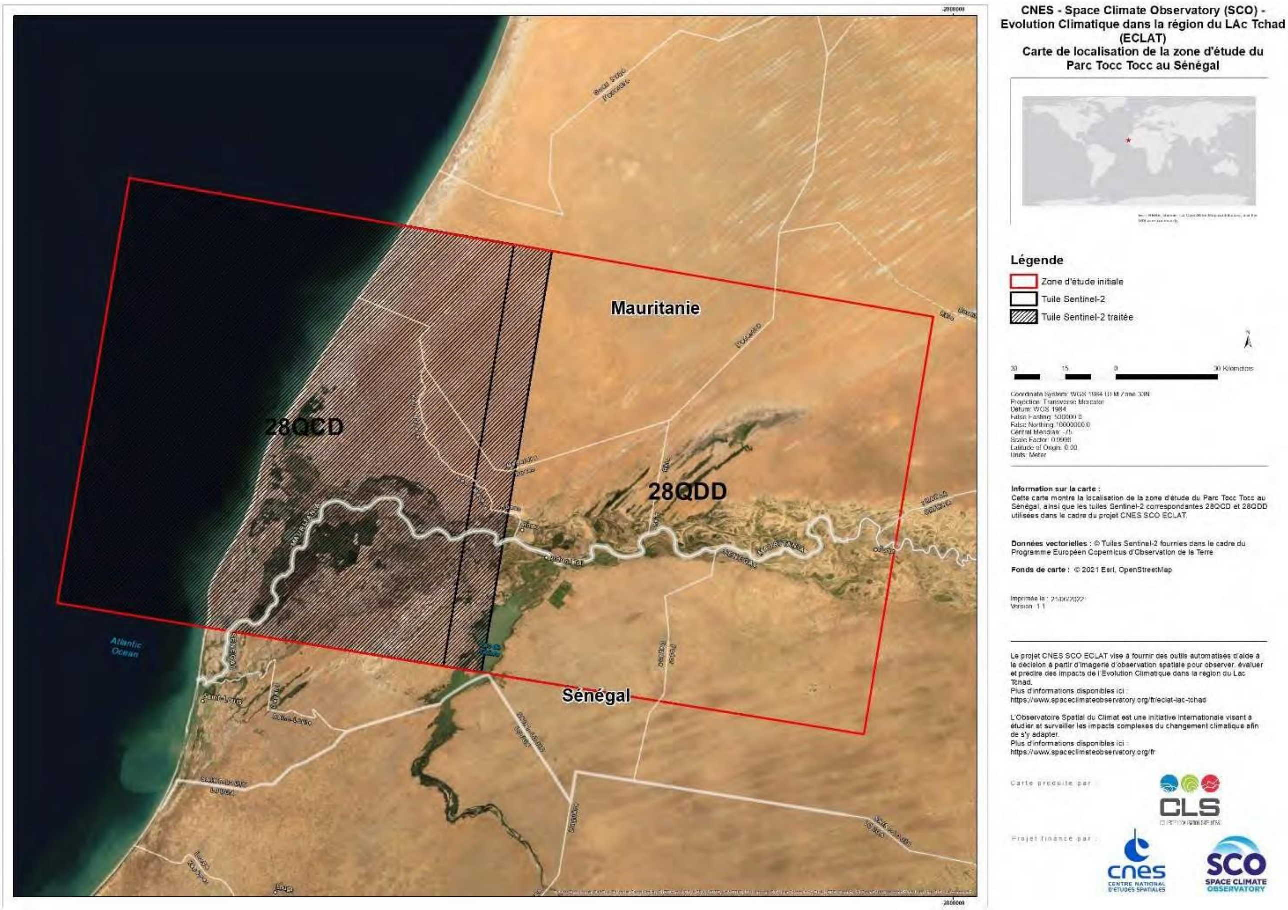


Figure 13: Location map of the Tocc Tocc Park in Senegal

3.2.2.1 Summary tables

Table 42: Database results for the Tocc Tocc Park area

BD	Variable	Accuracy user	Reminder	F1-Score	Accuracy global
Africapolis	Artificial area				
Copernicus Land cover	Total				0,42
Copernicus Land cover	Artificial area	1	0,87	0,93	
Copernicus Land cover	Water surface	0,93	0,84	0,89	
Copernicus Land cover	Cultivable land	0,56	0,74	0,64	
Copernicus Land cover	Wetland	0,85	0,61	0,71	
Copernicus Land cover	Bare ground	0,33	0,02	0,04	

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
Copernicus Land cover	Vegetation	0,21	0,73	0,32	
Copernicus Land cover	Forest	0	0	0	
Copernicus Water bodies	Water surface	0,48	1	0,65	
Cropland2000	Cultivable land	0	0	0	
ESA_CCI_LC_Africa	Total				0,34
ESA_CCI_LC_Africa	Artificial area	1	0,82	0,9	
ESA_CCI_LC_Africa	Area in water	0,63	0,81	0,71	
ESA_CCI_LC_Africa	Cultivable land	0,23	0,63	0,34	
ESA_CCI_LC_Africa	Wetland	0,71	0,19	0,3	
ESA_CCI_LC_Africa	Bare ground	0,31	0,04	0,1	
ESA_CCI_LC_Africa	Vegetation	0,22	0,57	0,32	
ESA_CCI_LC_Africa	Forest	0	0	0	
ESA_CCI_LC	Total				0,33
ESA_CCI_LC	Artificial area	1	0,35	0,52	
ESA_CCI_LC	Water surface	0,7	0,53	0,6	
ESA_CCI_LC	Cultivable land	0,26	0,38	0,31	
ESA_CCI_LC	Wetland	0,38	0,44	0,41	
ESA_CCI_LC	Bare ground	0,75	0,03	0,05	
ESA_CCI_LC	Vegetation	0,24	0,75	0,36	
ESA_CCI_LC	Forest	0	0	0	
From GLC	Total				0,34
From GLC	Artificial area	0,41	0,65	0,5	
From GLC	Water surface	0,68	0,94	0,79	
From GLC	Earth cultivable	0,37	0,49	0,42	
From GLC	Wetland	1	0,05	0,1	
From GLC	Bare ground	0,59	0,08	0,14	
From GLC	Vegetation	0,2	0,75	0,32	
From GLC	Forest	0	0	0	
G3WBM_Yamazaki	Total				0,81
G3WBM_Yamazaki	Water surface	1	0,81	0,89	
G3WBM_Yamazaki	Wetland	0,68	1	0,81	
GFSAD30AFCE_2015	Total				0,78
GFSAD30AFCE_2015	Cultivable land	0,88	1	0,94	
GLC_Share_FAO	Total				0,33
GLC_Share_FAO	Artificial area	0,89	0,73	0,8	
GLC_Share_FAO	Water surface	0,68	0,52	0,59	

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
GLC_Share_FAO	Cultivable land	0,35	0,42	0,39	
GLC_Share_FAO	Wetland	0,71	0,28	0,4	
GLC_Share_FAO	Bare ground	0,63	0,1	0,16	
GLC_Share_FAO	Vegetation	0,19	0,77	0,31	
GLC_Share_FAO	Forest	0	0	0	
GLCNMO	Total				0,28
GLCNMO	Artificial area	1	0,69	0,82	
GLCNMO	Water surface	0,89	0,31	0,46	
GLCNMO	Cultivable land	0,31	0,57	0,4	
GLCNMO	Wetland	1	0,1	0,1	
GLCNMO	Bare ground	0	0	0	
GLCNMO	Vegetation	0,2	0,86	0,33	
GLCNMO	Forest	0	0	0	
Global Human Settlement	Total				0,9
Global Human Settlement	Artificial area	0,91	1	0,95	
Global Tree Cover (> 30%)	Forest				
Modis Tree cover (> 30%)	Forest				
MODIS Water mask	Water surface	0,84	1	0,91	
MODIS_bare_soil	Bare ground	0,39	1	0,56	
Modis_land_cover_maryland	Total				0,28
Modis_land_cover_maryland	Artificial area	0,73	0,62	0,67	
Modis_land_cover_maryland	Water surface	0,75	0,43	0,55	
Modis_land_cover_maryland	Cultivable land	0	0	0	
Modis_land_cover_maryland	Wetland	0,8	0,11	0,19	
Modis_land_cover_maryland	Bare ground	1	0,1	0,13	
Modis_land_cover_maryland	Vegetation	0,19	1	0,32	
Modis_land_cover_maryland	Forest	0	0	0	
MODIS_vegetation	Vegetation	0,19	1	0,32	
Pekel_occurence	Total				0,91
Pekel_occurence	Water surface	1	0,83	0,91	
Pekel_occurence	Wetland	0,83	1	0,91	
SRTMWB	Total				0,71
SRTMWB	Water surface	0,71	1	0,83	
World settlement footprint	Total				0,93
World settlement footprint	Artificial area	0,93	1	0,96	
World_pop	Artificial area				
WWF Global lake wetland	Total				0,28

BD	Variable	User accuracy	Reminder	F1-Score	Overall accuracy
WWF Global lake wetland	Water surface	0,65	0,46	0,54	
WWF Global lake wetland	Wetland	0,2	0,94	0,33	

Table 43: Results for the artificial area variable in the Tocc Tocc Park area

BD	Variable	User accuracy	Reminder	F1-Score
World settlement footprint	Artificial area	0,93	1	0,96
Global Human Settlement	Artificial area	0,91	1	0,95
Copernicus Land cover	Artificial area	1	0,87	0,93
ESA_CCI_LC_Africa_20m	Artificial area	1	0,82	0,9
GLCNMO	Artificial area	1	0,69	0,82
GLC_Share_FAO	Artificial area	0,89	0,73	0,8
Modis_land_cover_maryland	Artificial area	0,73	0,62	0,67
ESA_CCI_LC	Artificial area	1	0,35	0,52
From GLC	Artificial area	0,41	0,65	0,5
World_pop	Artificial area			
Africapolis	Artificial area			

Table 44: Results for the variable water surface in the Tocc Tocc Park area

BD	Variable	User accuracy	Reminder	F1-Score
MODIS Water mask	Water surface	0,84	1	0,91
Pekel	Water surface	1	0,83	0,91
G3WBM_Yamazaki	Water surface	1	0,81	0,89
Copernicus Land cover	Water surface	0,93	0,84	0,89
SRTMWB	Water surface	0,71	1	0,83
From GLC	Water surface	0,68	0,94	0,79
ESA_CCI_LC_Africa_20m	Water surface	0,63	0,81	0,71
Copernicus Water bodies	Water surface	0,48	1	0,65
ESA_CCI_LC	Water surface	0,7	0,53	0,6
GLC_Share_FAO	Water surface	0,68	0,52	0,59
Modis_land_cover_maryland	Water surface	0,75	0,43	0,55
WWF Global lake wetland	Water surface	0,65	0,46	0,54
GLCNMO	Water surface	0,89	0,31	0,46

Table 45: Results for the variable Cultivable land in the Tocc Tocc Park area

BD	Variable	User accuracy	Reminder	F1-Score
GFSAD30AFCE_2015	Cultivable land	0,88	1	0,94

BD	Variable	User accuracy	Reminder	F1-Score
Copernicus Land cover	Cultivable land	0,56	0,74	0,64
From GLC	Cultivable land	0,37	0,49	0,42
GLCNMO	Cultivable land	0,31	0,57	0,4
GLC_Share_FAO	Cultivable land	0,35	0,42	0,39
ESA_CCI_LC_Africa_20m	Cultivable land	0,23	0,63	0,34
ESA_CCI_LC	Cultivable land	0,26	0,38	0,31
Modis_land_cover_maryland	Cultivable land	0	0	0
Cropland2000	Cultivable land	0	0	0

Table 46: Results for the wetland variable in the Tocc Tocc Park area

BD	Variable	User accuracy	Reminder	F1-Score
G3WBM_Yamazaki	Wetland	0,68	1	0,81
Copernicus Land cover	Wetland	0,85	0,61	0,71
ESA_CCI_LC	Wetland	0,38	0,44	0,41
GLC_Share_FAO	Wetland	0,71	0,28	0,4
WWF Global lake wetland	Wetland	0,2	0,94	0,33
ESA_CCI_LC_Africa_20m	Wetland	0,71	0,19	0,3
Modis_land_cover_maryland	Wetland	0,8	0,11	0,19
GLCNMO	Wetland	1	0,1	0,1
From GLC	Wetland	1	0,05	0,1

Table 47: Results for the bare soil variable in the Tocc Tocc Park area

BD	Variable	User accuracy	Reminder	F1-Score
MODIS_bare_soil	Bare ground	0,39	1	0,56
GLC_Share_FAO	Bare ground	0,63	0,1	0,16
From GLC	Bare ground	0,59	0,08	0,14
Modis_land_cover_maryland	Bare ground	1	0,1	0,13
ESA_CCI_LC	Bare ground	0,75	0,03	0,05
Copernicus Land cover	Bare ground	0,33	0,02	0,04
ESA_CCI_LC_Africa_20m	Bare ground	0,31	0,04	0,1
GLCNMO	Bare ground	0	0	0

Table 48: Results for the variable Vegetation in the Tocc Tocc Park area

BD	Variable	User accuracy	Reminder	F1-Score
ESA_CCI_LC	Vegetation	0,24	0,75	0,36
GLCNMO	Vegetation	0,2	0,86	0,33
Copernicus Land cover	Vegetation	0,21	0,73	0,32

BD	Variable	User accuracy	Reminder	F1-Score
Modis_land_cover_maryland	Vegetation	0,19	1	0,32
From GLC	Vegetation	0,2	0,75	0,32
ESA_CCI_LC_Africa_20m	Vegetation	0,22	0,57	0,32
MODIS_vegetation	Vegetation	0,19	1	0,32
GLC_Share_FAO	Vegetation	0,19	0,77	0,31

3.2.2.2 Analysis of results by variable

3.2.2.2.1 Water surface

The water surfaces are the parts of the Tocc Tocc Park site belonging to the lake. The validation points are therefore well clustered and it is not surprising to note that better results are obtained with lower spatial resolution databases (SRTMWB) at this site. Indeed, as the validation points are very clustered in a narrow area of the site, a pixel with a low spatial resolution could encompass them all and thus lead to a bias in the validation. It is more interesting to look at the lower spatial resolution databases where small areas of trees present in the middle of the water body of the area can also be identified. As the validation points are at the interface between the wetlands and the water body, it is not surprising that some databases, such as Pekel's, have somewhat lower results. We can imagine that if the water level is not constant, there could be confusion for some databases depending on the time of observation.

3.2.2.2.2 Wetland

Strangely, the best results on this variable are also obtained in the Tocc Tocc Park area with a low spatial resolution database. This can also be explained by the clustering of the points in a minimal area and unevenly distributed over the whole area. The Yamazaki database also still works well on this variable and could therefore remain as a reference database despite a lower accuracy than GLC_Share or GLCNMO. However, this can be explained by the difference in resolution between the databases. The ESA_CCI_LC database at 300m also performs well, with the same points as Yamazaki. The others are well below in terms of results, with accuracies and sensitivities falling sharply.

3.2.2.2.3 Bare ground

This variable is very poorly represented in all the validation points, and as in the Lake Chad or Park W areas, it is difficult to distinguish at times between bare soil and soil with very few trees. Some databases may have confused these two variables and therefore become less relevant in this area. The only two databases that do not have results at 0 are FROM_GLC and MODIS_Bare soil, although FROM_GLC is much more accurate.

3.2.2.2.4 Vegetation

The databases are a little more relevant with the vegetation variable despite the possible confusion with bare soil. However, the results remain very low with an F-Score of 0.42 for the Copernicus Land Cover database, which ranks first. The small size of the area and the possible confusion with bare soil may explain these low metrics on the vegetation variable. Moreover, as vegetation is less present than wetlands, this factor, added to the others, could also explain the poor results for lower resolution databases and the confusions for higher resolution ones.

3.2.2.3 Final ranking on the Tocc Park area Tocc

Table 49: Results for the artificial area variable in the Tocc Tocc Park area

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Global Human Settlement	Global / 10m	3	2	3	3	2,8	41
CCI Land Cover Africa 2016	Continentale / 20m	3	3	3	1	2,6	41
World Settlement footprint	Global / 10m	3	3	2	1	2,84	40,2
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	2	2,72	38,6
FROM - GLC	Global / 30m	3	3	2	3	1	33
MODIS Land Cover	Global / 500m	1	3	3	3	1,68	30,4
Global Urban Footprint	Global / 12m or 84m	3	3	1	1	1	29
WorldPop	Global / 100m	2	2	3	3	1	28
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	2,28	26,4
Africapolis	Continentale	3	2	1	1	1	26
GLC - Share FAO	Global / 1km	1	2	1	1	2,2	24
NOAA Nightlights	Global / 1 km	1	2	1	1	1	18
Natural Earth	Overall / 1km	1	1	1	1	1	15
Weight		4	3	2	1	5	-

Table 50: Results for the variable water surface in the Tocc Tocc Park area

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
GSW (Pekel)	Global / 30m	3	3	3	3	2,64	43,2
Yamazaki	Global / 100 & 30m	3	2	3	3	2,56	39,8

FROM-GLC	Global / 30m	3	3	2	3	2,16	38,8
BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	2,56	38,8
Copernicus Global Land Service Water Bodies	Global / 30m	3	3	3	3	1,6	38
SRTMWB	Global / 30m	3	3	2	1	2,32	37,6
CCI Land Cover Africa 2016	Continentale / 20m	3	3	3	1	1,84	37,2
MODIS Water surface	Global / 250m	1	3	3	3	2,64	35,2
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1,4	29
Natural Earth	Global / 10m	3	1	1	1	1	23
Global Lake and Wetlands database	Overall / 100m	2	1	1	3	1,16	21,8
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1,36	19,8
Weight		4	3	2	1	5	-

Table 51: Results for the variable Cultivable land in the Tocc Tocc Park area

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Evaluation criteria 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
GFSAD30AFCE	Global / 30m	3	3	2	1	2,76	39,8
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	1,56	33,8

FROM-GLC	Global / 30m	3	3	2	3	1	33
BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
CCI Land Cover Africa 2016	Continentale / 20m	3	3	3	1	1	33
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27
EarthStat - Pasture and Cropland area	Overall / 10km	1	3	1	1	1	21
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1	18
Weight		4	3	2	1	5	-

Table 52: Results for the wetland variable in the Tocc Tocc Park area

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
G1WBM	Global / 100 & 30m	3	2	3	3	2,24	38,2
Copernicus Global Land Land Cover Service	Overall / 100m	2	3	3	3	1,84	35,2
FROM-GLC	Global / 30m	3	3	2	3	1	33
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27
MODIS Land Cover	Global / 500m	1	3	3	3	1	27

Global Lake and Wetlands database	Overall / 100m	2	1	1	3	1	21
BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1	18

Table 53: Results for the variable "Bare soil" in the Tocc Tocc Park area

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Evaluation criteria 3 - Updating	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
Copernicus Global Land Service Land Cover	Overall / 100m	2	3	3	3	1	31
MODIS Bare soil layer	Global / 250m	1	3	3	3	1,24	28,2
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27
ISRIC World Soil information	Global / 250m	1	3	3	2	1	26
HARMONIZED WORLD SOIL DATABASE	Overall / 100m	2	2	1	1	1	22
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1	18
Weight		4	3	2	1	5	-

Table 54: Results for the variable Vegetation in the Tocc Tocc Park area

BD	Scale and resolution	Evaluation criterion 1 - Resolution	Assessment criterion 2 - Reliability	Assessment criterion 3 - Setting up a day	Evaluation criterion 4 - Genealogy	Assessment criterion 5 - F-Score	Total
FROM-GLC	Global / 30m	3	3	2	3	1	33
CCI Land Cover Africa 2016	Continentale / 20m	3	3	3	1	1	33
Copernicus Global Land Land Cover Service	Overall / 100m	2	3	3	3	1	31
MODIS Vegetation	Global / 250m	1	3	3	3	1	27
MODIS Land Cover	Global / 500m	1	3	3	3	1	27
ESA Climate Change Initiative - Land Cover led by UCLouvain (2017)	Global / 300m	1	3	3	3	1	27
EarthStat - Pasture and Cropland area	Overall / 10km	1	3	1	1	1	21
ISCGM - GLOBAL MAP GLOBAL LAND COVER (GLCNMO)	Global / 500m	1	2	1	3	1	20
GLC - Share FAO	Global / 1km	1	2	1	1	1	18

3.2.3 West Zone of Tocc Tocc Park in Senegal, Tile S2 28QCD

Following the involvement in the project of the Centre de Suivi Ecologique in Senegal, it was decided that the area to the west of the Tocc Tocc Park studied earlier would be of interest for a study. As this decision was taken at an advanced stage of the project, it was decided that the results obtained on the Tocc Tocc Park in Senegal would only be applied to the western part corresponding to the Sentinel-2 28QCD tile. We therefore use the previous results to produce the sample runs for the land cover production.

The databases selected are as follows:

Table 55: Selection of databases by theme in the Tocc Tocc Park area

BD	Thematic	
Global Human Settlement	Artificial area	
Global Surface Water	Water surface	
GFSAD30AFCE	Cultivable land	
G1WBM (Global 1sec Water Body Map)	Wetland	
Copernicus Global Land Service Land Cover	Bare ground	
FROM GLC & ESA CCI Land Cover Africa 20m	Vegetation	

4 Land use maps

4.1 Method

In order to carry out the land use on the different areas of the project and to take advantage of the previous database study, it is proposed to generate land use maps (OCS) from the Iota-2 chain.

The use of the databases will therefore take place in two ways:

- For a run of training samples to be provided as input to Iota-2
- For a correction of the chain output to optimise the overall accuracy of the OCS The

method of creating the land cover products follows 3 main steps:

- Data preparation :
    - Downloading the database of the area of interest
    - Split by number of processed Sentinel-2 tiles
    - Harmonisation of the base to the ECLAT nomenclature
    - Sample selection
- Iota-2 classification
- Validation

4.1.1 Preparation of the data

Harmonisation with the ECLAT nomenclature is done via a python code in order to perform the pixel value reclassification operations.

The selection of samples is done on the databases with the highest scores in the previous step. A python file was written to randomly select a number of samples from each database associated with each theme. The number of samples selected is the same for each theme. No stratification for each area was done beforehand in order to assess the number and proportion of each class in the different areas. The rebalancing of the number of samples is done later in Iota-2 during the "Sample Selection" step. This part will be explained later in this document.

The ground truth file is now ready for use in Iota-2.

4.1.2 Classification

The Iota-2 chain is launched to generate the land cover. In a second step in post classification a correction is applied to the land cover. In order to select the databases that will be used as a source of correction, it was decided to select databases only if they exceed a score of 2.5 on the "F-Score" scoring parameter, below which the database cannot be selected to perform corrections on the OCS. In addition, we also set an additional criterion for the correction of OCS. The databases have different production and update dates. Corrections should only be made when the production or update date is close enough to the production year of the OCS with Iota-2. The criterion therefore imposes a maximum of 2 years difference between the correction database and the OCS (example: production for the year 2020, correction with WSF of 2019 or GHSL of 2018 allowed, but forbidden with GUF of 2015).

Taking this rule into account, we can select the following databases for the SCO corrections for each region:

Table 56: DBs selected for the correction of OCS maps in the Lake Chad area

Database with F-Score > 2.5	Thematic
Global Human Settlement	Urban
Global Surface Water	Water surface
G1BMW	Wetland
MODIS bare soil	Bare ground

Table 57: DBs selected for the correction of the OCS maps in the Tocc Tocc Park area

Database with F-Score > 2.5	Thematic
Global Human Settlement	Urban
Global Surface Water	Water surface
GFSAD30AFCE	Cultivable land

Table 58: DBs selected for the correction of the OCS maps in the Park area W

Database with F-Score > 2.5	Thematic
Global Human Settlement	Urban
Global Surface Water	Water surface

4.1.3 Validation

After the Iota-2 classification and the application of the corrections, a validation is carried out using the validation points constituted during the sampling plan carried out upstream in the project. The results are presented in the following section and the detailed confusion matrices per production area are available in Annex I.

4.2 Land use maps

Land cover maps are generated using the Iota-2 tool, a processing chain developed at CESBIO for the production of land cover classifications from Sentinel-2 long time series.

The Sentinel-2 data used are L2A level and ordered on the CNES PEPS via a python file. For all the sites produced (3 Sentinel-2 tiles in the Lake Chad region and one tile in the Tocc Tocc reserve area), 3 years were downloaded: 2020; 2019; 2018. The L3A level that was tested as input to Iota-2 was produced using the WASP tool for 2020 over Chad.

The configuration of Iota-2 was made following several parameterisation tests taking into account the classification accuracy of each theme as well as the calculation time. Themes that cannot be merged with the databases were evaluated more specifically and favoured in configurations that gave them better accuracy rates.

The different configurations with output accuracies are summarised in Table 598 below.

Table 59: Iota-2 classification test parameters

Test number	Sentinel-2 level used	Classifier	*Sampling strategy*	Ratio	Additional parameters	Overall accuracy
1	L3A	Random Forest Parameters : Min: 25; Max: 5 nbtrees: 150	Smallest	0,5	No	49%
2	L3A	Random Forest Parameters : Min: 10; Max: 5 nbtrees: 150	Smallest	0,5	No	52%
3	L3A	Random Forest Parameters : Min: 10; Max: 20 nbtrees: 150	Smallest	0,5	No	52%
4	L3A	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Total	0,5	No	59%
5	L3A	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Percent (0.8)	0,5	No	59%
6	L3A	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Total	0,5	Smote: active	59%

Test number	Sentinel-2 level used	Classifier	*Sampling strategy*	Ratio	Additional parameters	Overall accuracy
7	L3A	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Smallest	0,5	Smote: active Features: NDVI, NDWI, brightness	56%
8	L2A (3 dates)	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Smallest	0,5	No	60%
9	L2A (3 dates)	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Total	0,5	No	62%
10	L2A (3 dates)	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Total	0,5	Features: NDVI, NDWI, brightness	63%
11	L2A (12 dates)	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Total	0,5	Features : NDVI, NDWI, brightness	62%
11	L2A (full series)	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Total	0,5	Features: NDVI, NDWI, brightness	64%
12	L2A (full series)	Random Forest Parameters : Min: 10; Max: 5; nbtrees: 150	Percent (0.8)	0,5	Features: NDVI, NDWI, brightness	65%
13	L2A (full series)	Random Forest Parameters : Min: 5; Max: 25; nbtrees: 150	Percent (0.8)	0,5	Features: NDVI, NDWI, brightness	63%

The configuration tests were conducted on Chad. The various parameters tested in Iota-2 were first carried out on one tile (33PWP) before being deployed on other tiles. The "autocontext" parameter could have been tested, but the computation times were too long and the difficulties in obtaining a classification at the output of Iota-2 with this parameter led to the tests not being carried out in depth. The parameters retained in the benchmarking are :

- The classifier: *shark random forest* or OTB *random forest*
- Sample selection strategy *Percent* (80%)
- The nature of the input Sentinel-2 images

- The addition of additional parameters such as *sample augmentation* or *features* additional

Regarding the classifiers, few differences were seen in the operation of the chain and in the output results. The decision was made to quickly focus on the *Random Forest* algorithm, a classifier that is widely used in classification operations and which has been proven to be robust and accurate.

The *sampling* strategy for the input samples had a real impact on the land cover outputs. Indeed, it was quickly seen that the smallest *smallest* class strategy, even if it was the most desirable to ensure a balance between classes, was biased by classes with too few samples. This was the case for the forest or artificial area class on some tiles. It was decided to remove the "forest" class, which was absent except for a few tiles, for future tests in order to determine the relevance of this strategy. Despite the removal of the forest class, the results of this strategy remain below the others available. The tests also focused on the *Total* or *Percent* strategies, which consist in taking the total of the samples or a certain percentage. The application of the *Total* strategy shows a significant gain in accuracy (5% to 6%), particularly for the transition from Sentinel-2 L3A to Sentinel-L2A. The application of the *Percent* strategy provides a further gain in accuracy (1% to 2%) and will be the strategy chosen for the further generation of land cover maps.

In this study, the tests of the input datasets were quickly conclusive. The Sentinel-2 images at the 3A pre-processing level are less robust than L2A for land cover generation over Chad. The gain in accuracy with the L2A time series data is substantial and the latter will therefore be applied for future production.

The addition of extra parameters was also tested but in a limited way compared to all the possibilities offered by Iota-2. As the aim of the project was not to test the chain, only two additional parameters were tested: the *sample augmentation smote* and the addition of *features*: NDVI, NDWI, *brightness*. No impact was observed with the application of *Smote* and this parameter was quickly discarded. Nevertheless, new investigations should undoubtedly be carried out on this parameter, which is very useful in several respects. The other parameter concerning the features clearly allowed a gain in precision on the output results and will therefore be retained for future productions.

With the configuration of Iota-2 ready, the launch of the land cover productions can be done. Initially, the entire Chad region was to be processed. Subsequently, the Senegal 28QDD area as well as the Park W area were also to be processed. However, during the reorganisation of the project's production areas, particularly following discussions with the SSC in Senegal, the 28QCD area in Senegal and the 33PWP, 33PUP and 33PTP tiles in Chad were finally endorsed as production areas for the land use products.

The results of the accuracies of the different land cover classes, raw and then corrected, are presented in Table 59 below.

Table 60: Precisions of OCS classifications before and after correction in the 3 study areas

Zone	Gross accuracy	Accuracy after correction
Chad 33PWP 2018	61%	68%
Chad 33PWP 2019	64%	67%
Chad 33PWP 2020	63%	69%
Chad 33PUP 2018	55%	60%
Chad 33PUP 2019	57%	66%
Chad 33PUP 2020	60%	63%
Zone	Gross accuracy	Accuracy after correction
Chad 33PTP 2018	65%	66%

Chad 33PTP 2019	63%	65%
Chad 33PTP 2020	65%	68%
Senegal 28QCD 2018		82%
Senegal 28QCD 2019		77%
Senegal 28QCD 2020		77%

The results of the land cover mapping on the 3 Sentinel-2 tiles in Chad were obtained with an overall gross accuracy of between 55 and 65% for the year 2020.

Indeed, many samples in the random draw of the artificial area class turn out to be heterogeneous pixels between artificial area and bare soil.

This is due to the environment of urban areas in Chad where buildings are separated by bare soil and little artificial soil exists. The figure below shows examples of the problem encountered:

(a) (b)

Figure 14: Illustration of bare soil (a) and buildings (b) in Chad on Google Satellite images

The other existing confusions are between the classes of bare soil / vegetation / cultivable land. Indeed, some crops seem to be abandoned and are close to bare soil, whereas the low shrubby vegetation which is very present in the region leads to confusion between the above classes.

To overcome the existing confusion between the built and bare soil classes, the tiles were classified without the artificial area class as input and then corrected with a correction from the World Settlement Footprint 2019. This approach allows the Chad tiles to obtain overall accuracies after correction between 60 and 69%.

The results for the 33PWP tile are shown in Figure 15 :

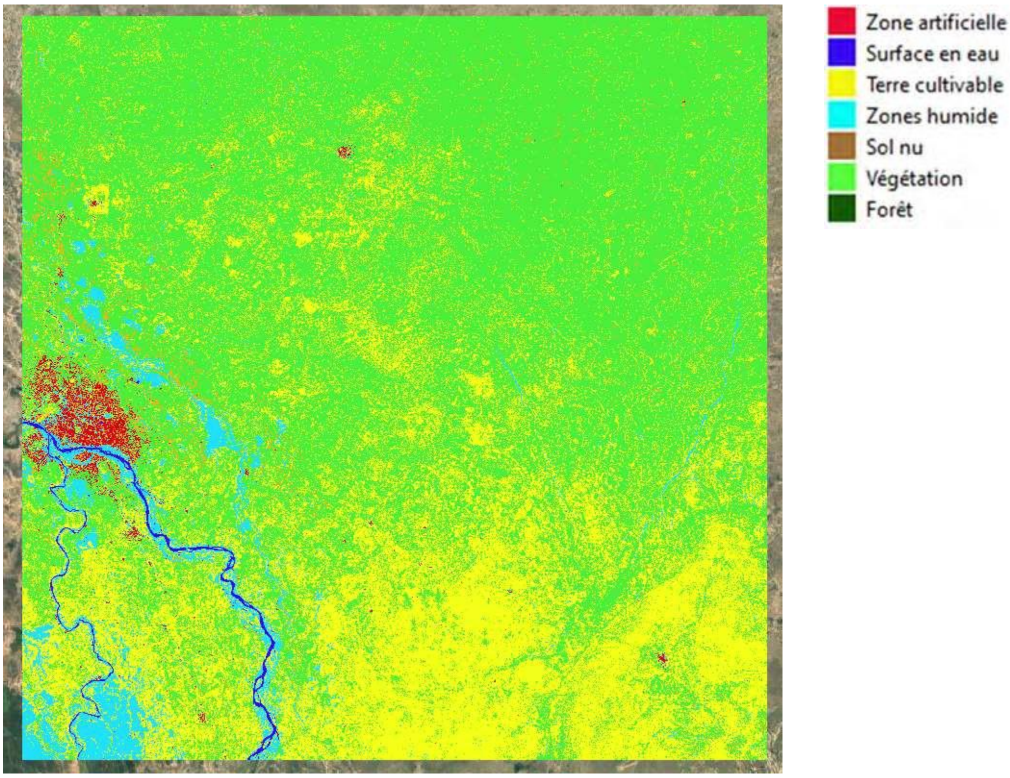


Figure 15: 33PWP - 2020

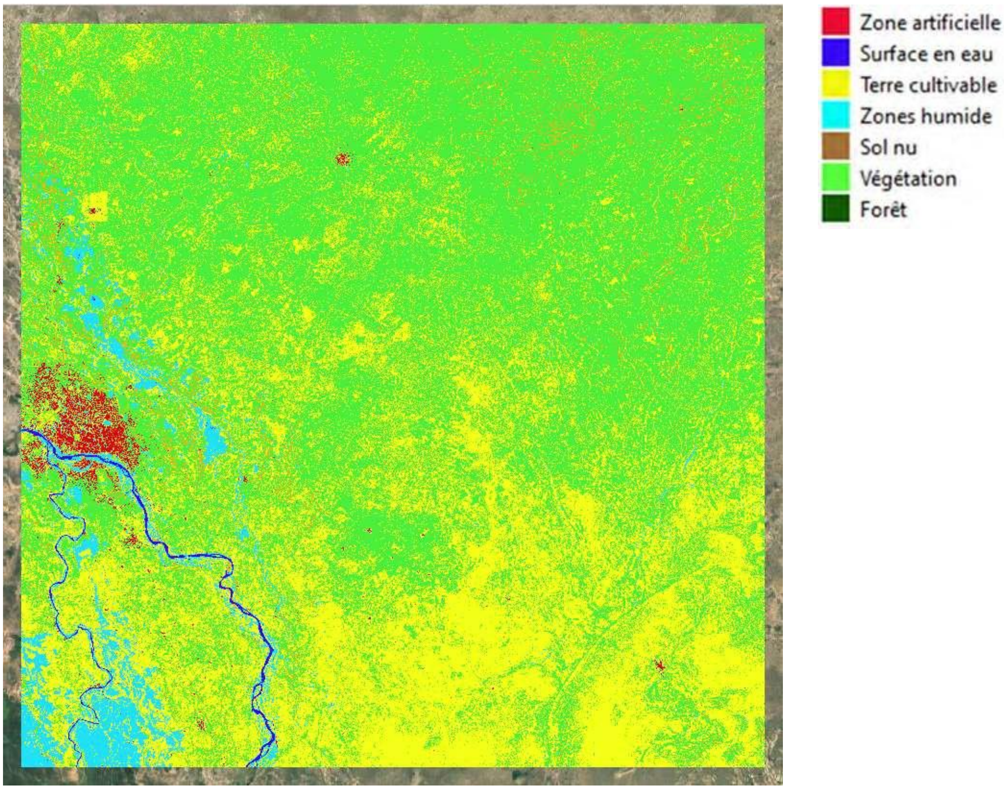


Figure 16: 33PWP - 2019

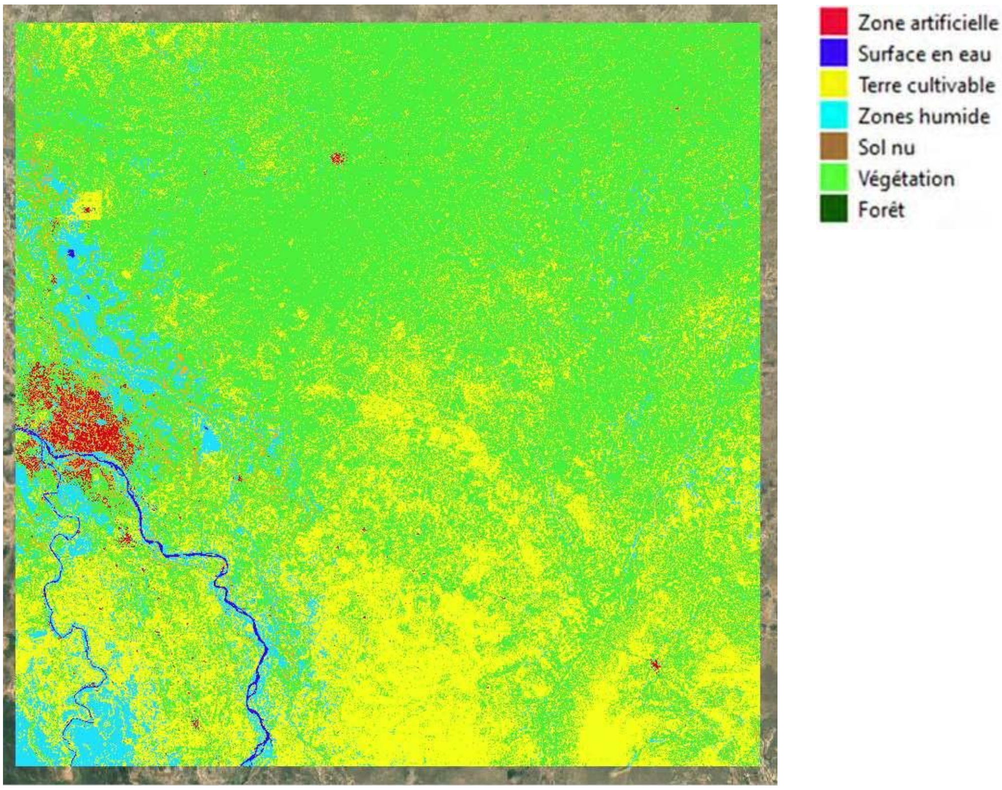


Figure 17: 33PWP - 2018

Confusions still exist, especially between the arable land and vegetation classes. The results for the 33PUP tile can be seen in Figure to 18 :

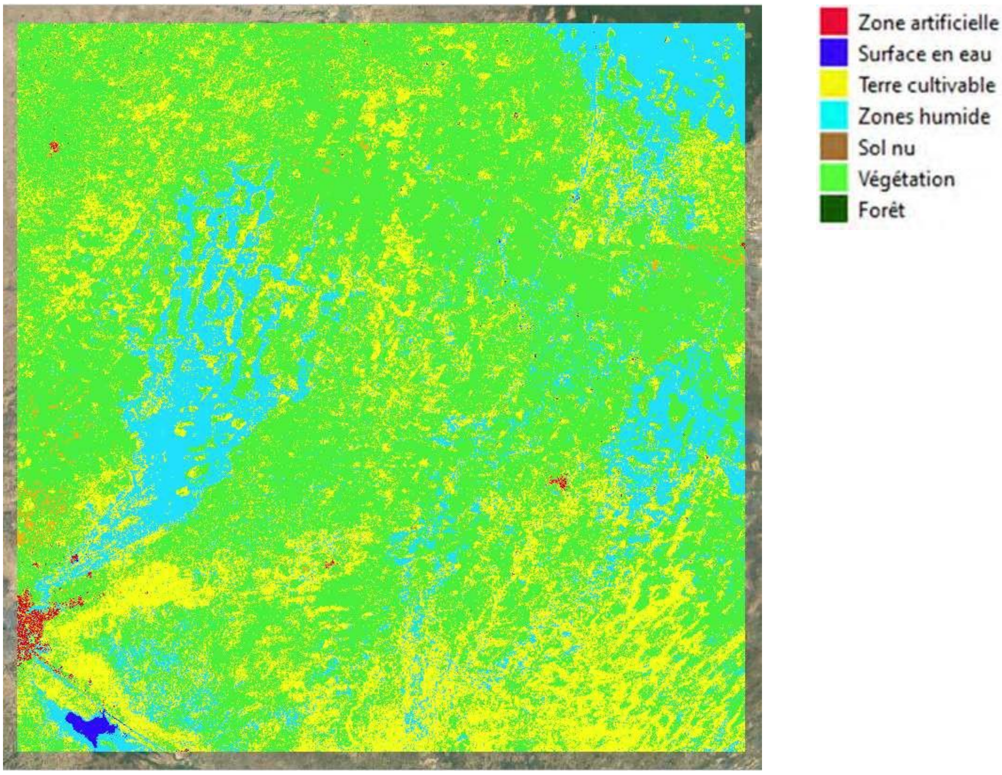


Figure 18: 33PUP - 2020

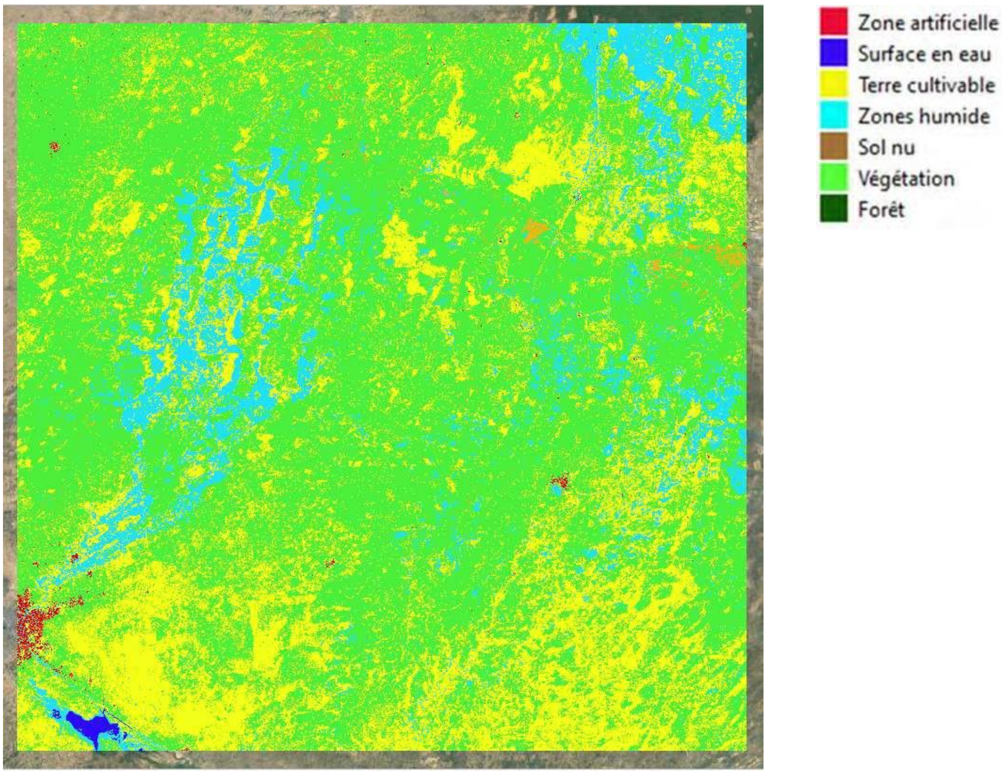


Figure 19: 33PUP - 2019

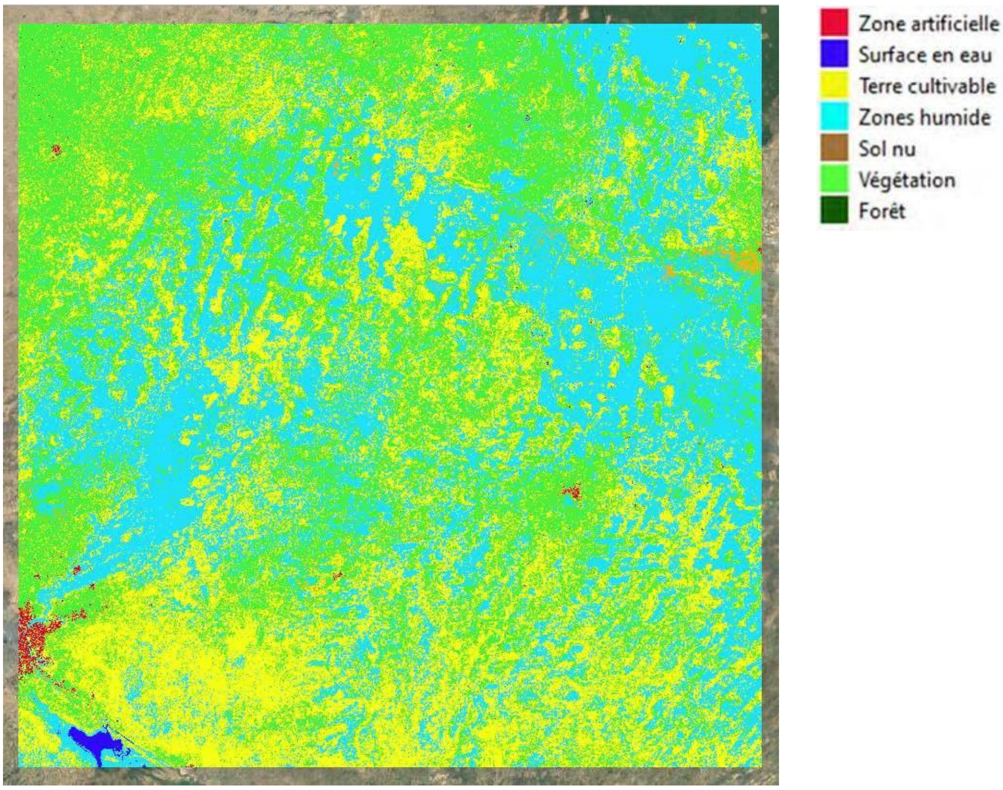


Figure 20: 33PUP - 2018

The results for the 33PTP tile are shown in Figure 21:

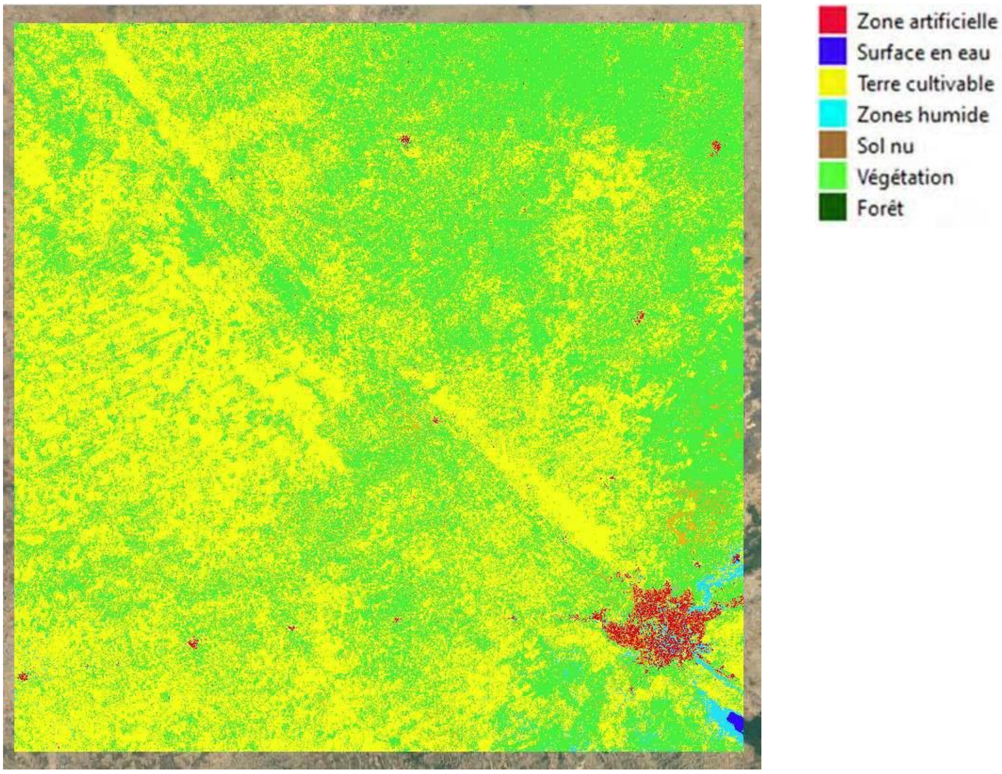


Figure 21: 33PTP - 2020

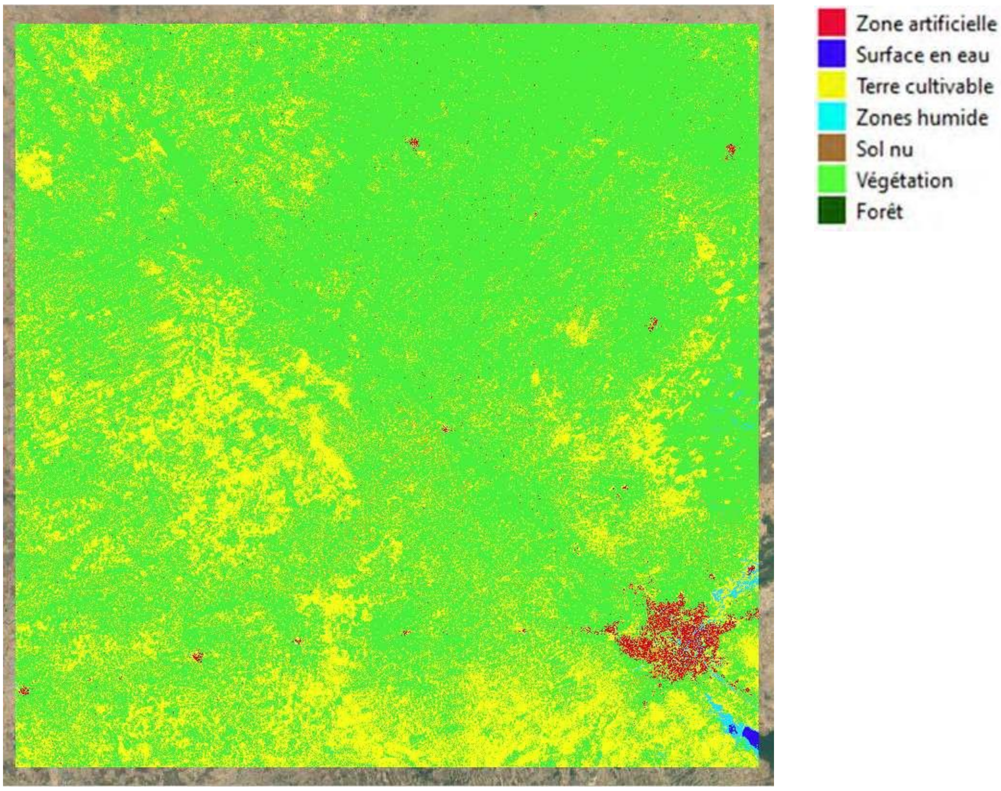


Figure 22: 33PTP - 2019

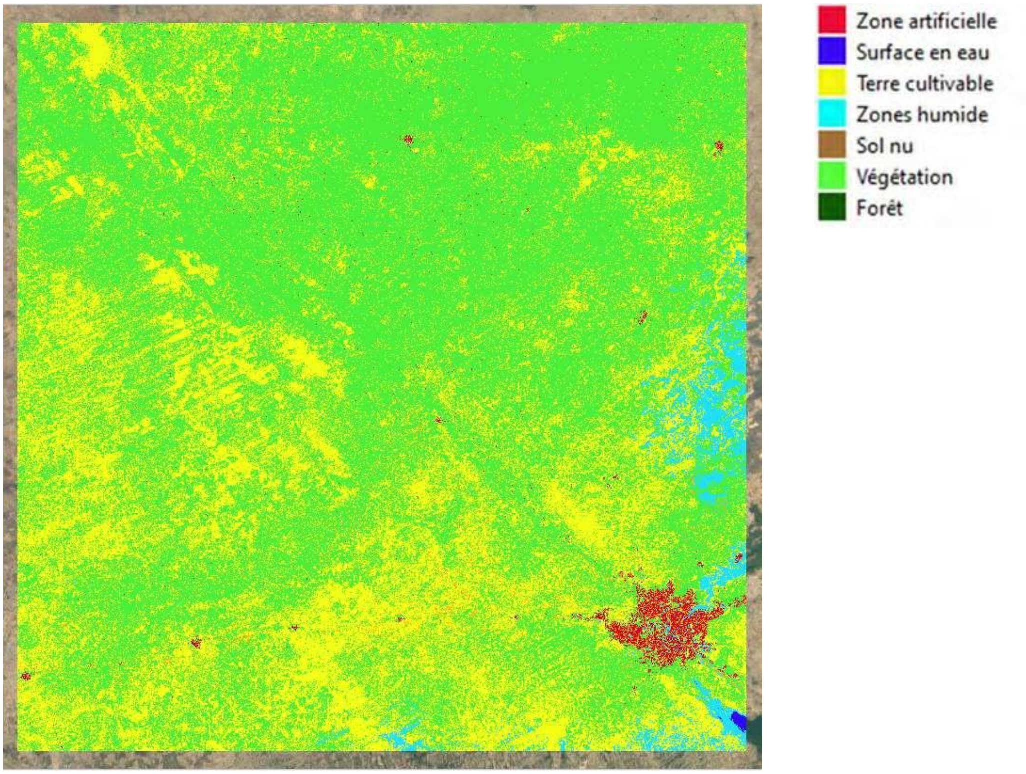


Figure 23: 33PTP - 2018

For the 33PTP tile, there were initially significant confusions in the artificial area class which was confused with bare soil but also with agricultural land. The application of corrections via the World Settlement Footprint 2019 corrects a significant part of these errors.

For all production areas, the class arable land or vegetation can be partly corrected with the indicator presented below, the DHI. This minor correction will not drastically increase the total classification accuracy.

The results for Senegal are presented in Figure 24.

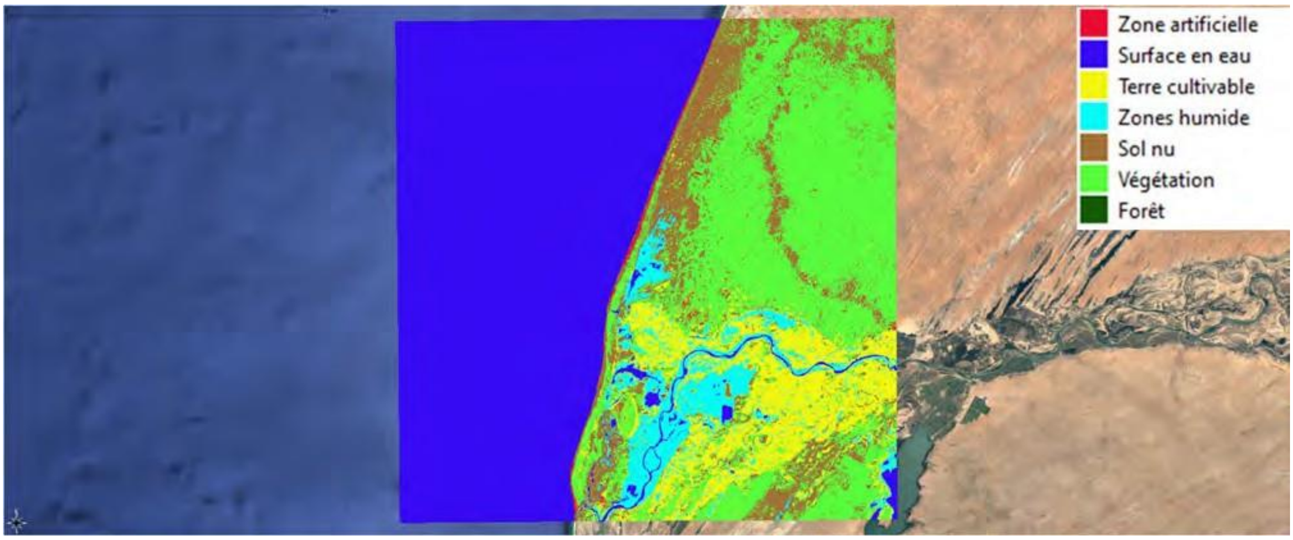


Figure 24: 28QCD - 2018

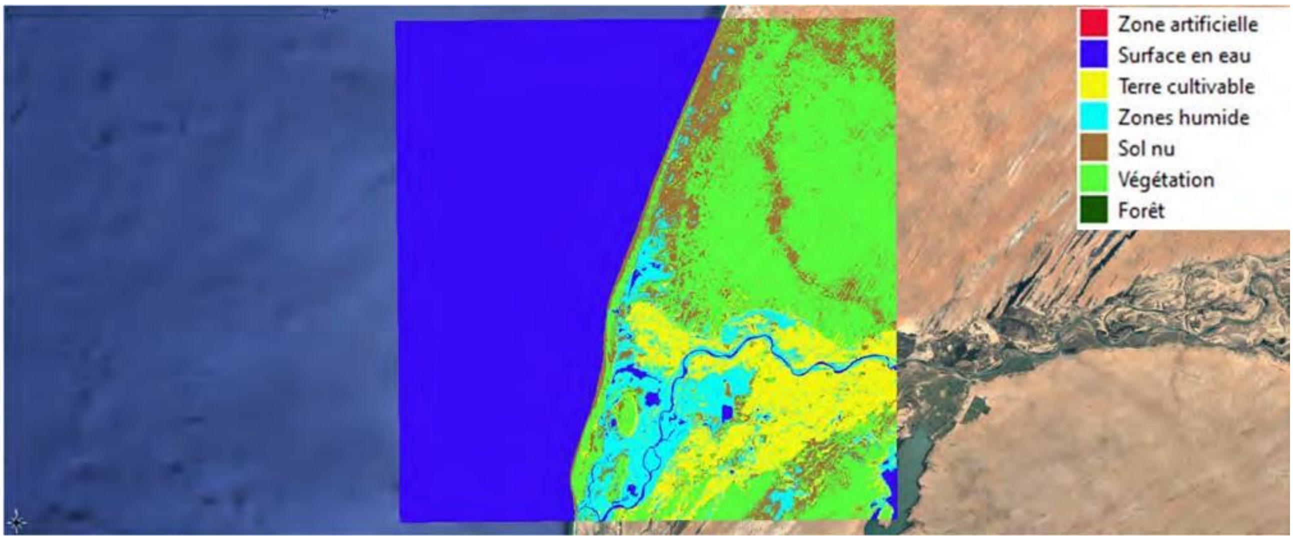


Figure 25: 28QCD - 2019

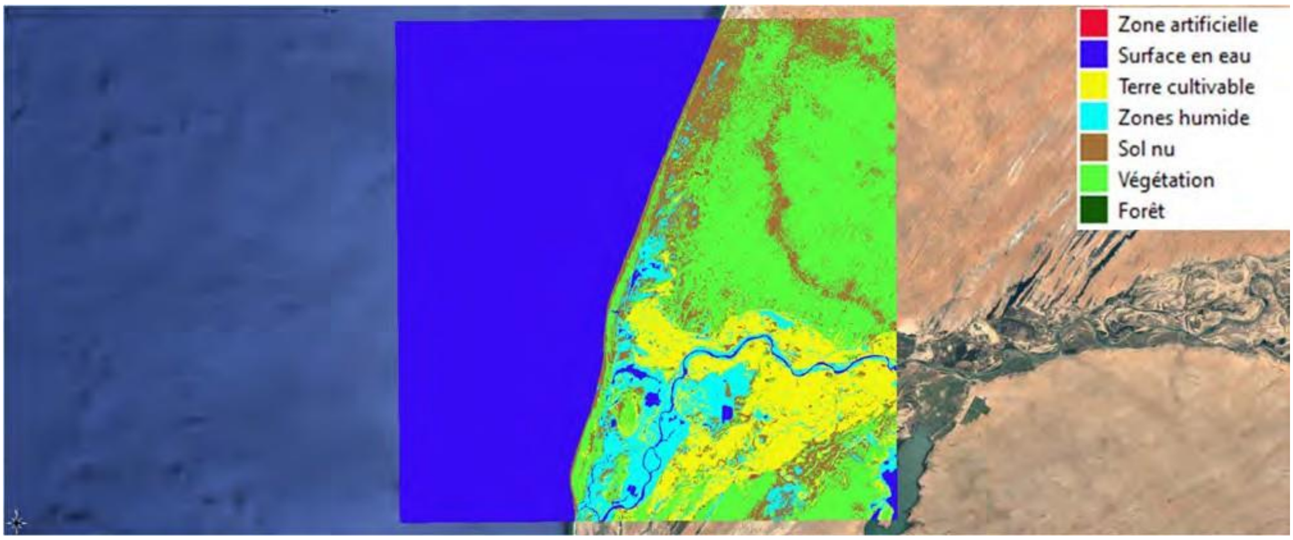


Figure 26: 28QCD - 2020

In general, the raw land cover results have an overall accuracy of about 60-65%. In spite of the corrections made, the gain in overall accuracy does not increase significantly and improves the different land cover classes by more or less a few percent. This is mainly due to the fact that the corrections of some are the errors of others. In other words, the correction of bare soil in artificial areas can induce other commission errors that will deteriorate the F-Score of the class. This will have an impact on the calculation of the overall accuracy, which in the best case will increase by 1 to 2% but will stagnate in most cases.

The initial desire to create land cover maps from data fusion would require further study, particularly when adding land cover from a classification chain such as Iota-2 to this fusion. This project only addresses this superficially, but it would be interesting to explore the possibility of merging with the free databases used to generate land use maps, in order to reduce confusion over certain classes such as built-up areas and bare soil by using more precise databases.

5 Calculation of indicators

5.1 Agricultural development indicator

The agricultural development indicators were strongly supported by the study of the *Dynamic* Habitat Index (DHI). This index corresponds to a statistical analysis of the Normalized *Difference Vegetation Index* (NDVI) over a given period. It contains the minimum observed, the cumulative and the standard variation of the values over the period. This index is most often represented in false colours with these 3 values respectively in the Red Green Blue (RGB) bands.

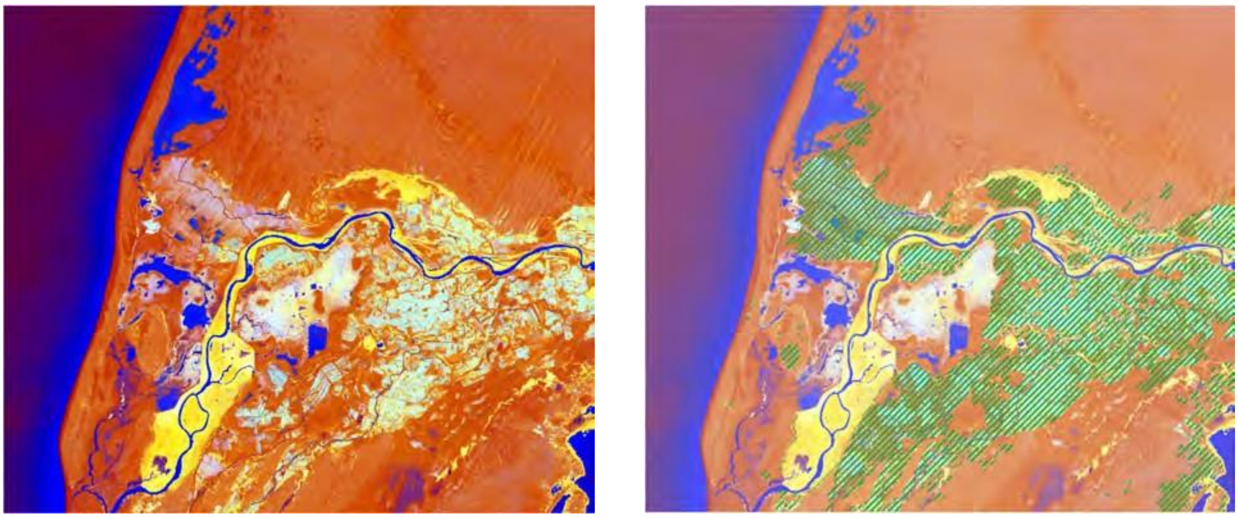
The NDVI calculated from satellite observations in the red and infrared channels makes it possible to discriminate the presence of vegetation in an area and its variations therefore inform us about its evolution. Thus, the DHI informs us about the presence of stable and persistent vegetation via a high accumulation and a low variation of NDVI over the year, but also about areas where the vegetation is episodic if we observe a large standard variation of NDVI. These zones correspond to the progressive and moderate natural growths that follow the rainy seasons but also to the strong and rapid appearance of vegetation on farms.

The difficulty in analysing this index lies, however, in separating natural and agricultural vegetation, a difficulty that is all the greater when the type of crops/climate/territory under study is not well known.

5.1.1 Characterisation of agriculture in Senegal

In this area of Senegal, the notable presence of agricultural areas has accentuated the demand for indicators to monitor this activity in particular. Whether it be for monitoring fallow periods, periods of land leaching, or the identification of possibly overexploited areas, the search for global or spatial, annual or seasonal statistical information was at the heart of the project.

With the use of IHD, it soon became apparent that the extraction of agricultural areas would be feasible with a good level of reliability, particularly due to the accurate databases of crop area boundaries and the clarity of activity on each plot.



(a) (b)

Figure 27: Illustration of IHD (a) and IHD with crop overlay (b)

It can be seen from the false colour representation (Figure a) of the IHD that the crops stand out immediately and that the agricultural base is reliable enough to serve as a pre-mask and facilitate their extraction (Figure b).

Different typical IHD values can then be identified.

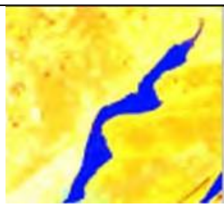
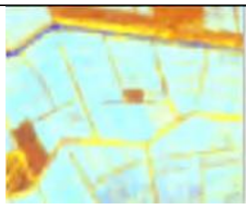
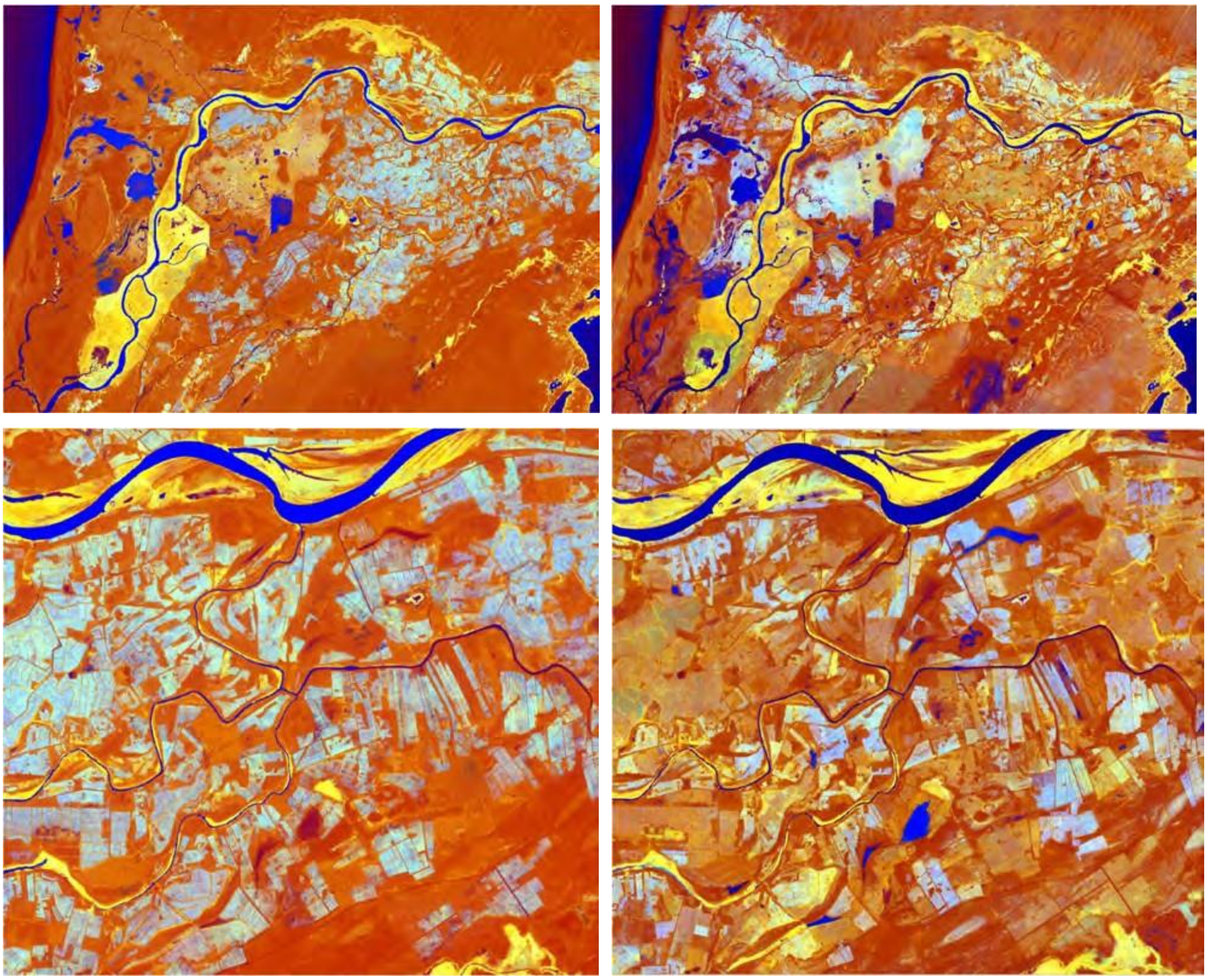
<table>
<tr><td>Areas of evergreen vegetation close to wet spots that maintain some greenery throughout the year have a rather high accumulation (band 2) and a low variation (band 3)</td><td>
Entité | Valeur
DHI2018 | 0
DHI2018
Bande 1 | 0.219971
Bande 2 | 100.12278
Bande 3 | 0.0354684</td></tr>
<tr><td>Areas with high accumulation (band 2) and variation (band 3) can be observed, which correspond to cultivated areas where greenery is present only part of the year.</td><td>
Entité | Valeur
DHI2018 | 0
DHI2018
Bande 1 | 0.0362091
Bande 2 | 132.37659
Bande 3 | 0.205857</td></tr>
<tr><td>There are also areas where agriculture is present but with a significantly and homogeneously lower accumulation (band 2).</td><td>
Entité | Valeur
DHI2018 | 0
DHI2018
Bande 1 | 0.110107
Bande 2 | 90.99314
Bande 3 | 0.18884</td></tr>
<tr><td>Finally, bare land where vegetation is never present has a low cumulative (band 2) and a (low) variation</td><td>
Entité | Valeur
DHI2018 | 0
DHI2018
Bande 1 | 0.0780029
Bande 2 | 29.266548
Bande 3 | 0.00903139</td></tr>
</table>

Figure 28: Illustration of different IHD types with corresponding NDVI values

The fact that 2 types of IHD values appeared for various agricultural areas made us look more closely at the evolution of NDVI at these locations. It was found that these differences were due to the use before and/or after the rainy season. It was then decided to calculate the IHD not for the whole year but for each of the two seasons to try to identify which plots were used in each season.

The result was very conclusive and we were indeed able to separate the 2 seasons as seen below:



(a) (b)

Figure 29: Illustration of IHD over 2 seasons in 2018: (a) January-July and (b) July-December

In this example of the year 2018 separated from January to July (left) and July to December (right), we see which plots are used and which are not, depending on the season.

It is then possible to extract by thresholding the areas concerned for each season and year and to establish operating indicators and statistics of interest to users.

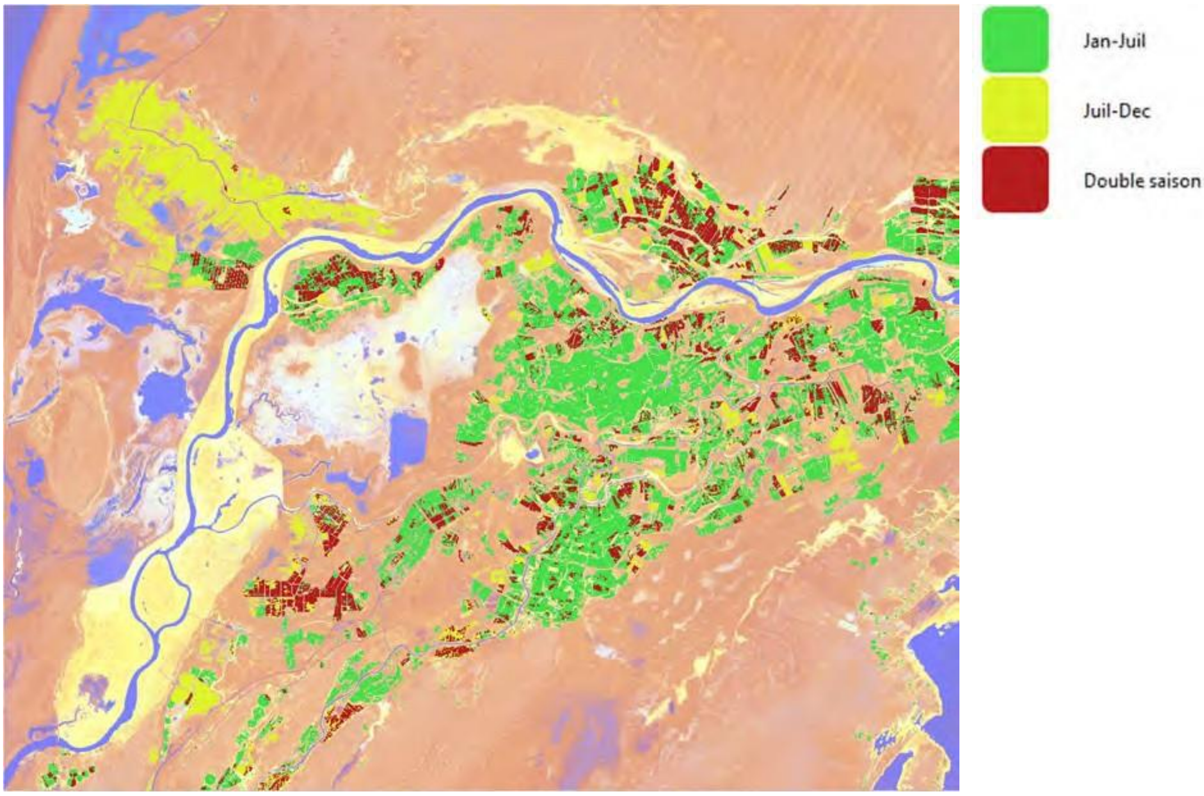


Figure 30: Extraction of agricultural areas from the 2018 IHD

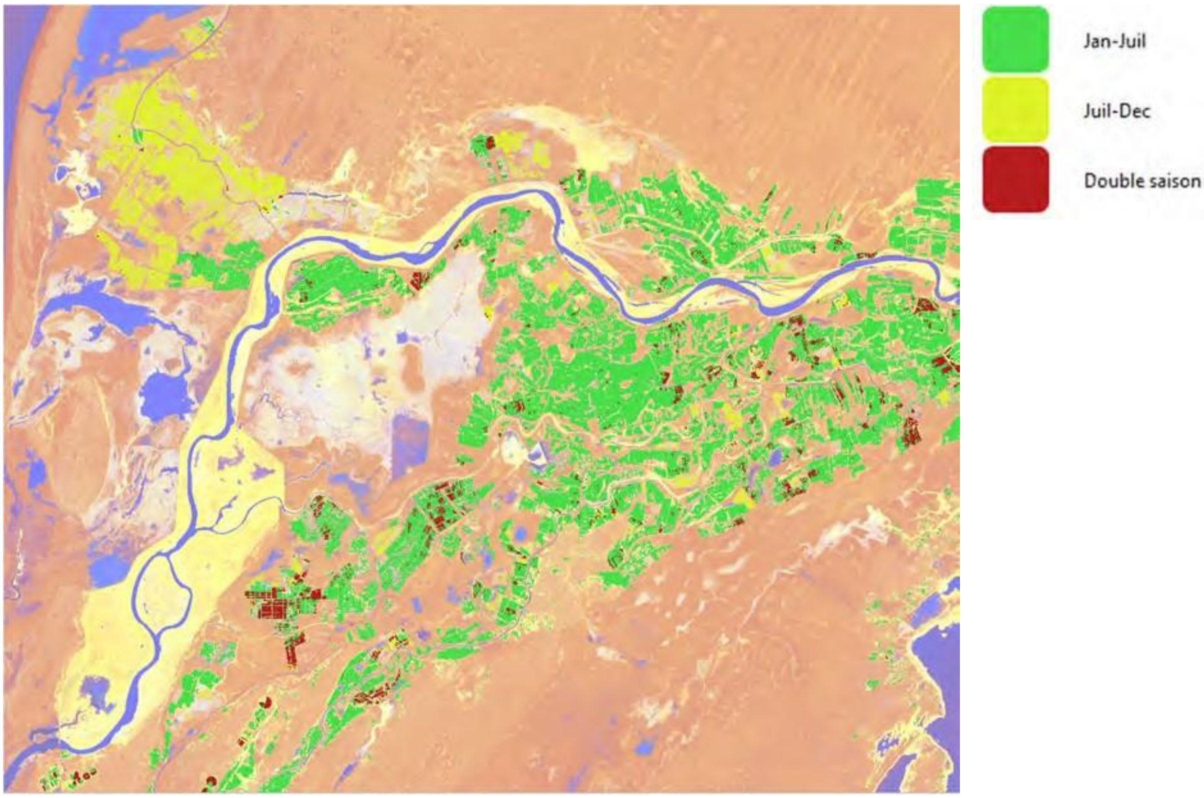


Figure 31: Extraction of agricultural areas from the 2019 IHD

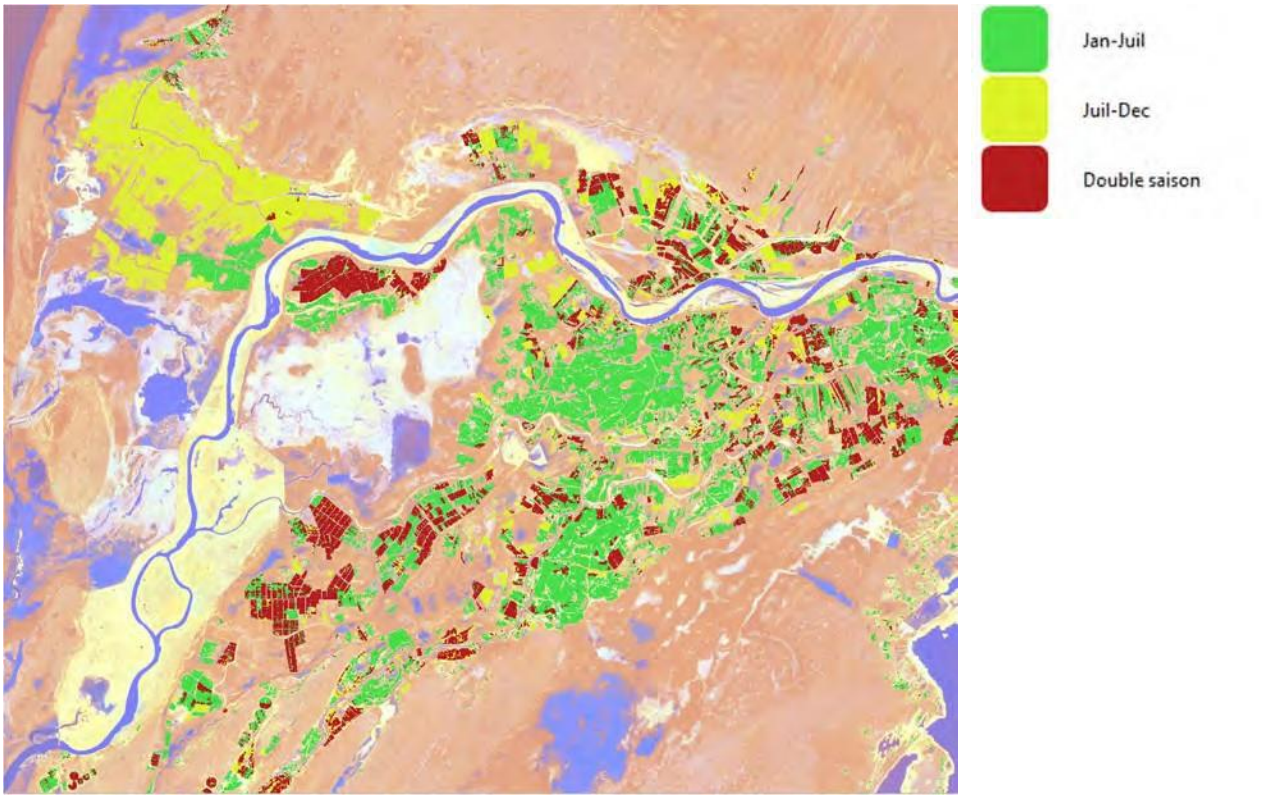


Figure 32: Extraction of agricultural areas from the 2020 IHD

These masks, although imperfect, are accurate enough to provide robust statistics. Visually, we can see a collapse of double cropping and July-December cropping in 2019 as a direct result of the drought that affected West Africa that year.

On the whole Sentinel-2 28QCD tile the following statistics can therefore be established:

Table 61: Sentinel-2 28QCD Tile Statistics

28QCD	2018	2019	2020
Season 1 total			
Area (km²)	393,18	389,63	385,65
Area (% agri base)	0,334	0,33	0,328
Productivity(sum(cumulative)/Area)	10,49	11,43	9,65
Total Season2			
Area (km²)	216,26	109,56	270,62
Area (% agri base)	0,18	0,093	0,23
Productivity(sum(cumulative)/Area)	12,46	12,98	14,43
Seasons 1 exclusive			
Area (km²)	292,3	359,44	259,67
Area (% agri base)	0,24	0,3	0,22
Productivity(sum(cumulative)/Area)	10,12	11,31	8,96
Season 2 exclusive			
Area (km²)	115,38	79,36	144,63
Area (% agri base)	0,098	0,067	0,123
Productivity(sum(cumulative)/Area)	12,67	13,39	14,83

Double seasons			
Area (km²)	100,87	30,19	125,98
Area (% agri base)	0,08	0,025	0,107
Productivity(sum(cumulative)/Area)	23,78	24,81	25,05
Free			
Area (km²)	665,36	704,93	643,64
Area (% agri base)	0,56	0,6	0,548

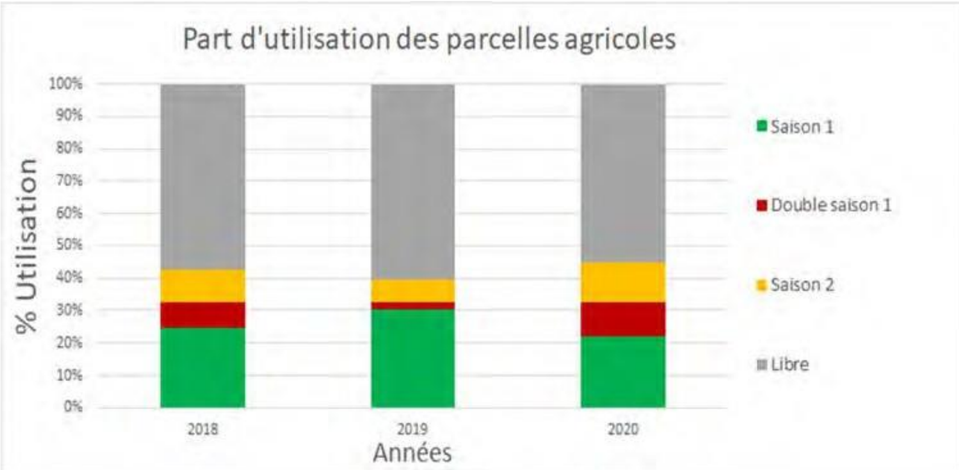


Figure 33: Share of agricultural land use per year

As visually predicted, the statistics confirm that 2019 was particularly difficult in the 2nd season from July to December. Many plots that were supposed to be in dual use were not and were switched to exclusive season 1. Season 2 areas show a 30% decrease compared to 2018.

Finally, it should be noted that the free area displayed is calculated on the basis of the agricultural data used for the extraction; its inclusive nature may lead to an overestimation of the free area actually available, but its integration into the graphs allows us to highlight the calculated statistics more explicitly.

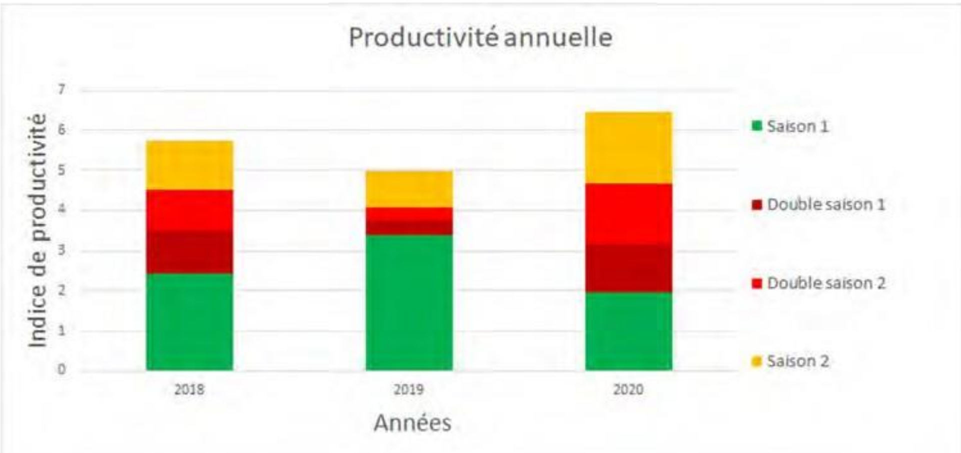
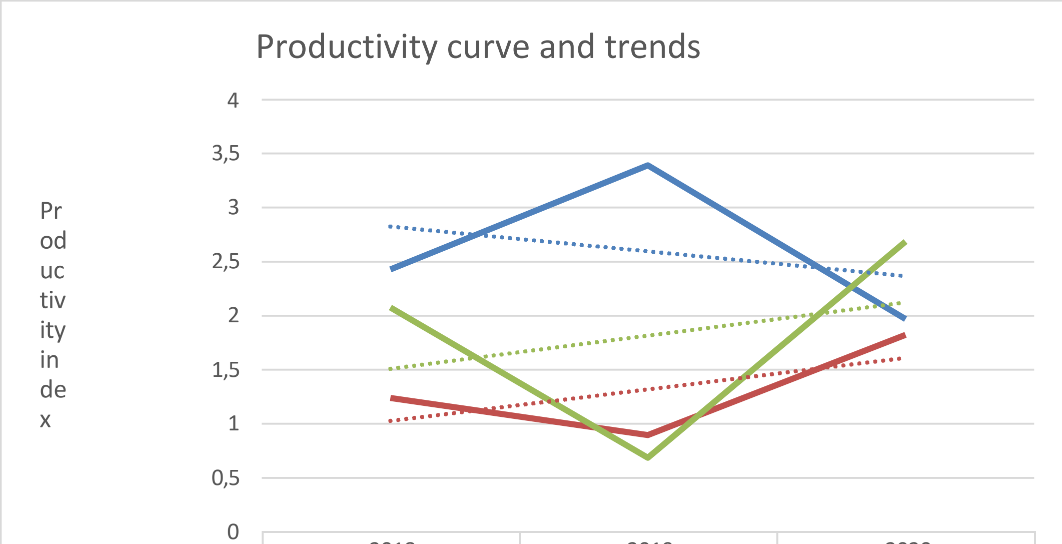


Figure 34: Productivity index by year

To estimate the health or even the productivity of the farms, we studied the cumulative IHD per year on the active plots, discriminating the contribution of each season, including for double use.

The cumulative NDVI that is thought to be related to productivity can also be seen in the curves below where trends for each season are plotted:



	2018	2019	2020
Season 1 total	2,4288	3,393	1,9712
Season 2	1,24166	0,89713	1,82409
Double season	2,076	0,68891	2,68881

Figure 35: Productivity curves and trends by year and season

In 2019, the drought from July to September has a strong impact on the use of plots, especially on those that had to produce in two consecutive seasons. Alternatively, land cover flow analysis can also be effectively visualised using Sankey graphs (see Figure a/b/c). The magnitude of individual fluxes can be highlighted, providing users with a sense of their importance or severity within a period or between several periods of change.

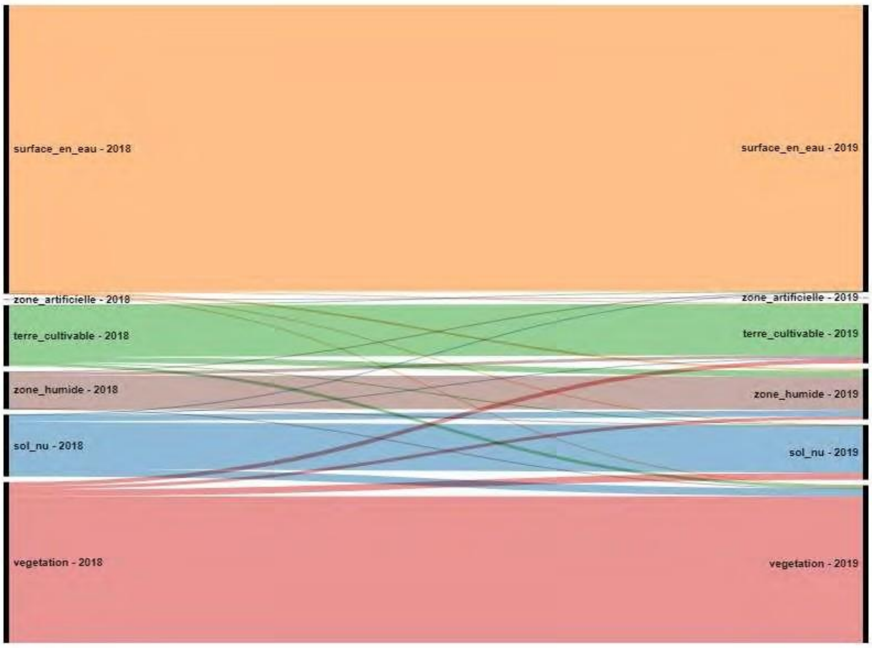


Figure 36: Flow chart between 2018 and 2019 on tile 28QCD

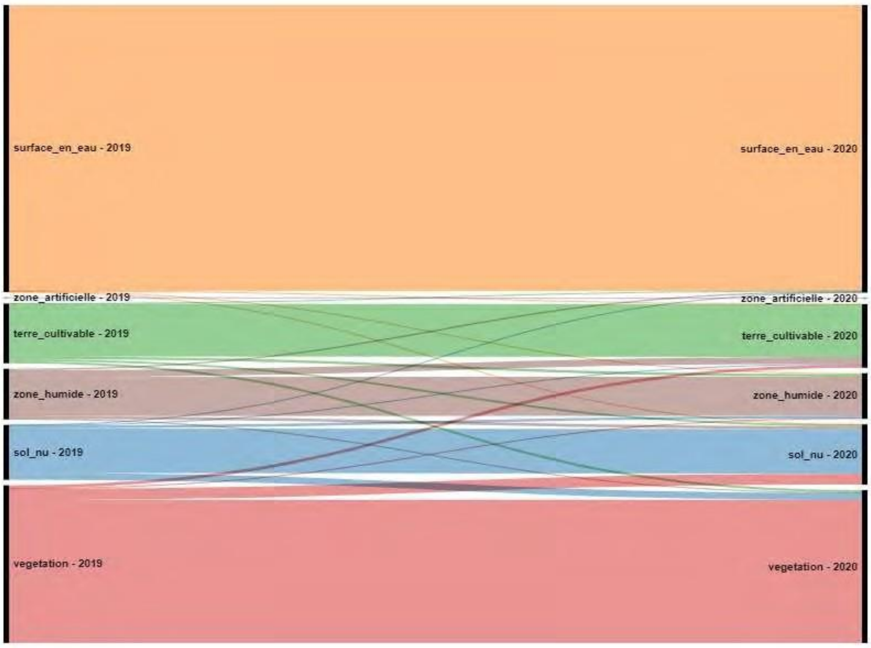


Figure 37: Flow chart between 2019 and 2020 on tile 28QCD

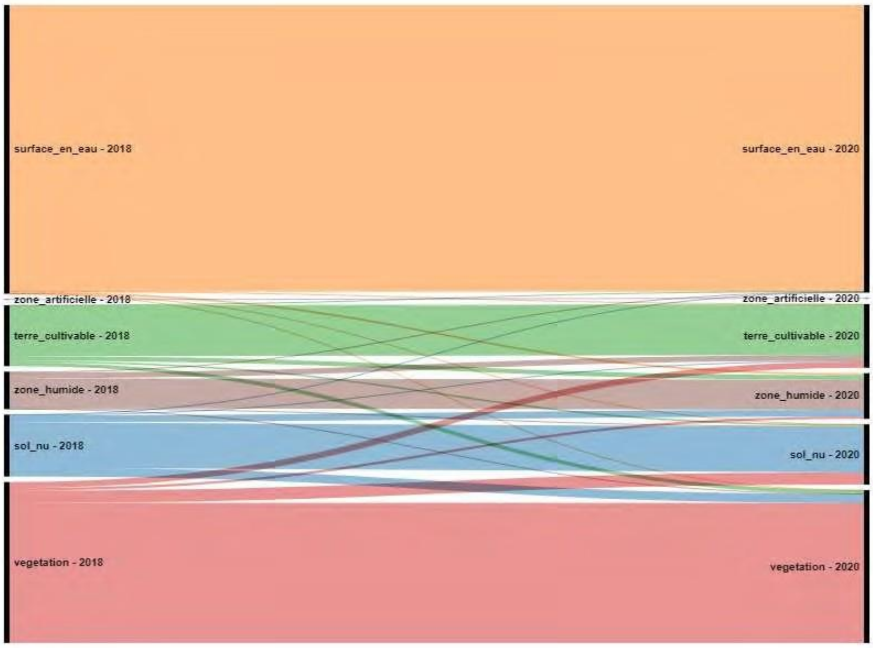


Figure 38: Flow chart between 2018 and 2020 on the 28QCD tile

These diagrams show, for example, that between 2018 and 2020 there is no significant net gain or loss of a land use type. Any flows from one item to another offset each other. For example, the loss of vegetation between 2018 and 2020 to mainly bare soil and arable land is compensated by the conversion of bare soil and arable land to vegetation between 2018 and 2020.

Thus, these cross transfers of surface areas suggest a spatial reorganisation of the territory without any significant variation in the total surface areas of the land use categories.

5.1.2 Characterisation of the vegetation in Chad

In Chad, despite the desire to analyse agricultural areas, we found that these areas were much less clear-cut and easier to identify than in Senegal. Moreover, we did not have the opportunity to discuss with partners on the ground who could help us to see more clearly the crops in the area. As a result, we reoriented our analysis towards vegetation in general rather than agriculture.

The analysis of Sentinel-2 annual time series showed that the local flora can be divided into at least 2 categories: persistent vs. ephemeral vegetation. This distinction between persistent and ephemeral vegetation was made by means of thresholding on the IHD and by visual interpretation of time series over several years to determine the optimal threshold.

- Evergreen vegetation, which is rather rare and therefore all the more important to monitor carefully as it can potentially serve as a refuge for certain species during dry periods

Figure 39: Illustration of persistent vegetation

- Ephemeral vegetation that appears from July to November, which includes part of the agricultural and natural environments, which should therefore also be carefully monitored for deterioration over time

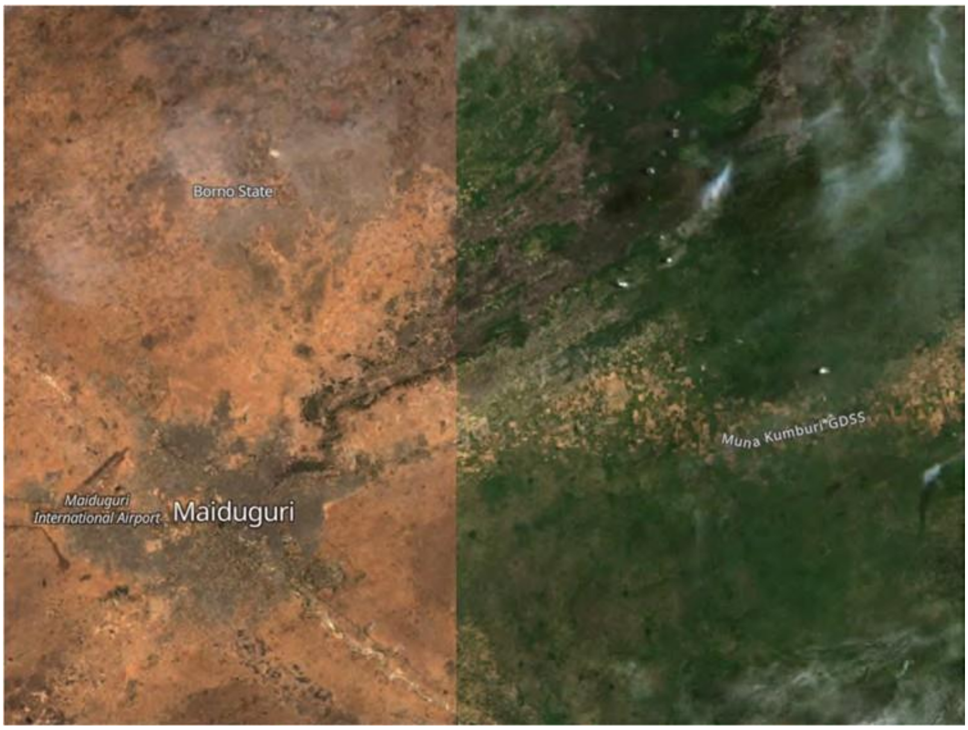


Figure 40: Illustration of ephemeral vegetation

In the same way as in Senegal, we found the different typical values of IHD to be identified in order to extract these two vegetation types.

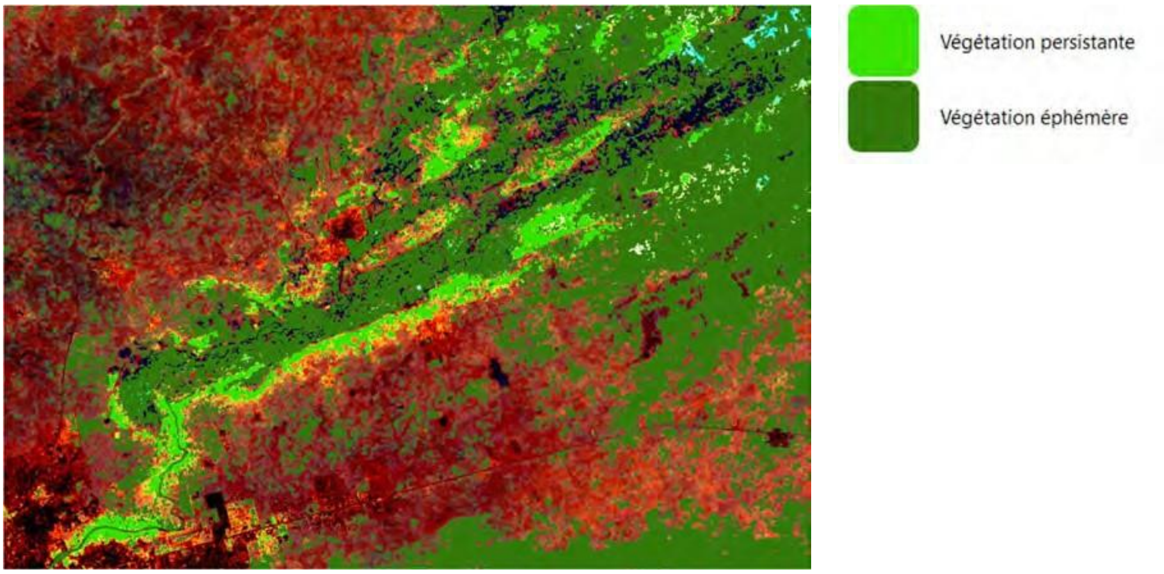


Figure 41: Illustration of classification results separating evergreen and ephemeral vegetation

We can therefore estimate the area covered by the two types of vegetation, study its evolution over time and observe two different trends below.

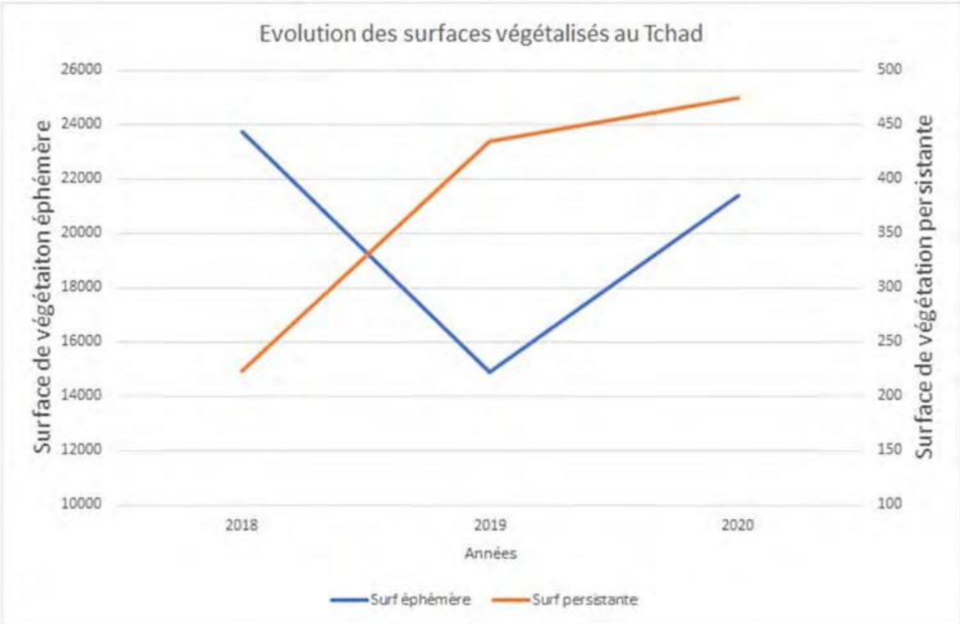


Figure 42: Evolution of vegetated areas in Chad

On the evolution of the evergreen vegetation, we can look at the area north-east of Maiduguri which is a reserve of evergreen vegetation and it appears that between 2018 (left) and 2020 (right), the vegetation bordering the river has clearly increased.

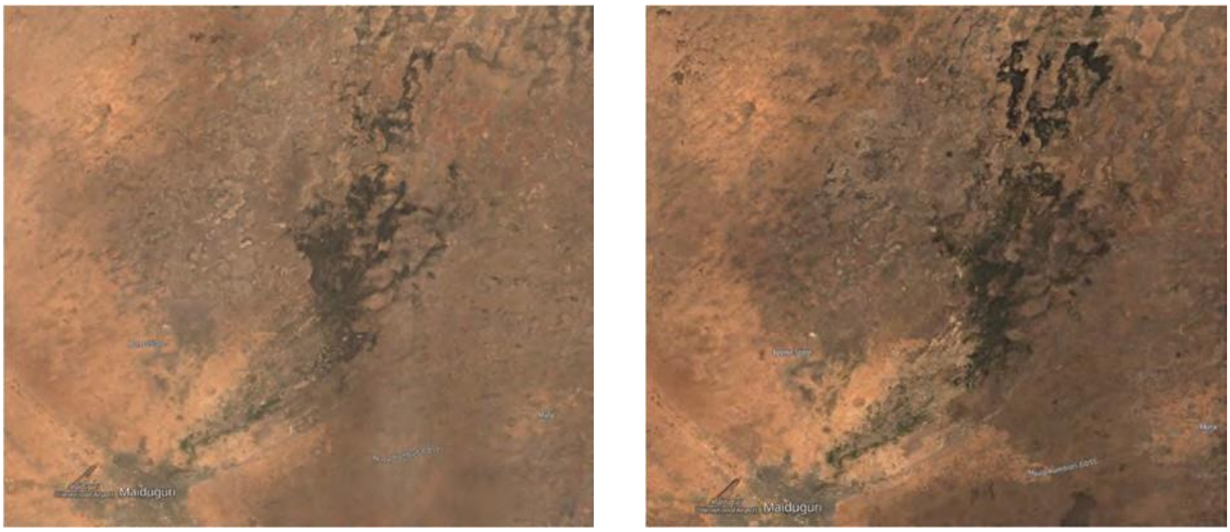


Figure 43: Illustration of the increase in the amount of trees along the river between 2018 (left) and 2020 (right)

5.1.3 Evolution agricultural

The construction of IHDs over several years has made it possible to highlight agricultural practices, particularly in Senegal, by discriminating between crops in different seasons. This also allows an estimation of agricultural production in the area and an estimation of the production of the different agricultural zones for the current season.

In addition, the monitoring of agricultural developments over several years makes it possible to address several specific issues in Senegal:

The first concerns leaching, which is necessary to remove salinisation from the soil. With the ECLAT project indicator, it would be relatively easy to determine the number of times each plot is used and thus optimise the rotation of plots to leach them at the best time.

The second is to meet production requirements. The monitoring of plots allows the anticipation of rotations and therefore the optimisation of yields. A gain/loss map can thus be produced between each year in order to quickly see if the season will be better than the previous one and to plan to optimise the next.

These monitoring maps were made with the previously produced IHDs. Thus, it is possible to see the different transitions from one season to another of the agricultural plots as shown in the following images:

Here is the colour legend to follow for each change:

The colours have been grouped for 3 changes (here the example is taken for 2018 and 2019 representing year 1 and year 2 respectively):

- Red = change from any activity in year 1 to season 1 in year 2
- Blue = change from any activity in year 1 to a double season in year 2
- Black = change from any activity to no activity in year 2
- Yellow = transition from any activity to season 2 in year 2

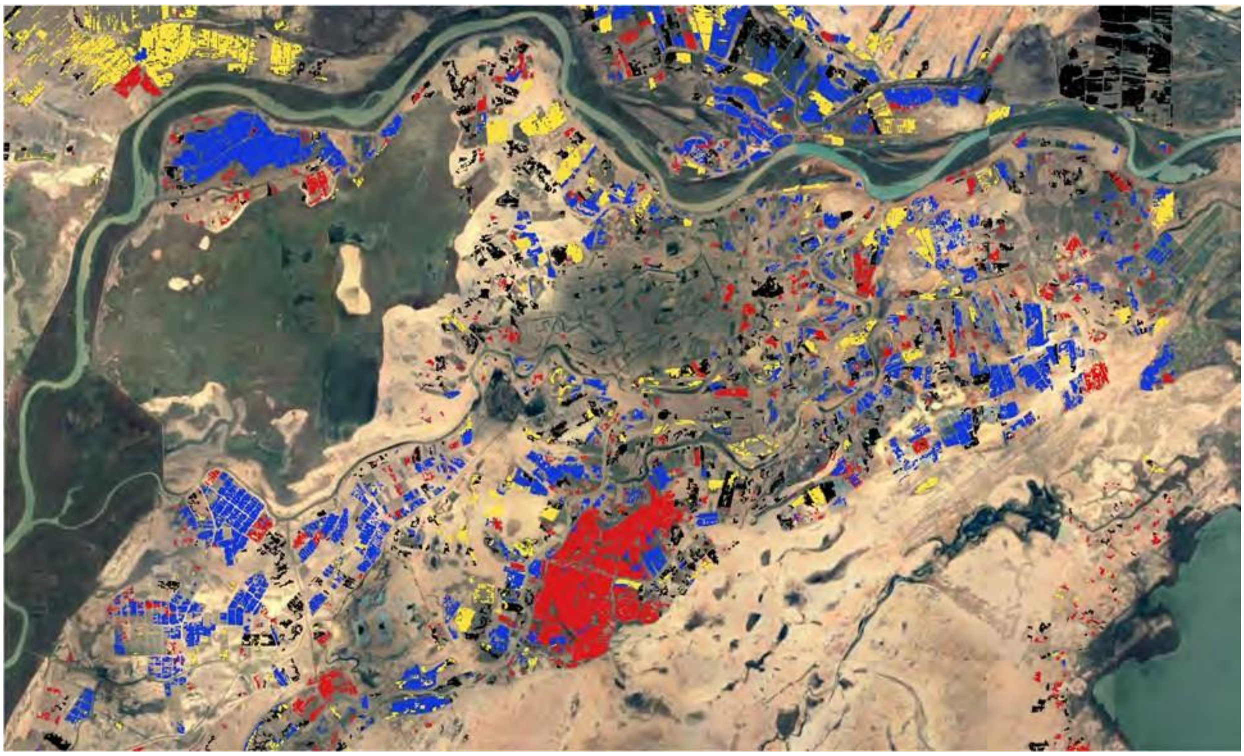


Figure 44: 2020 - 2019 change of practice in the seasons of some plots

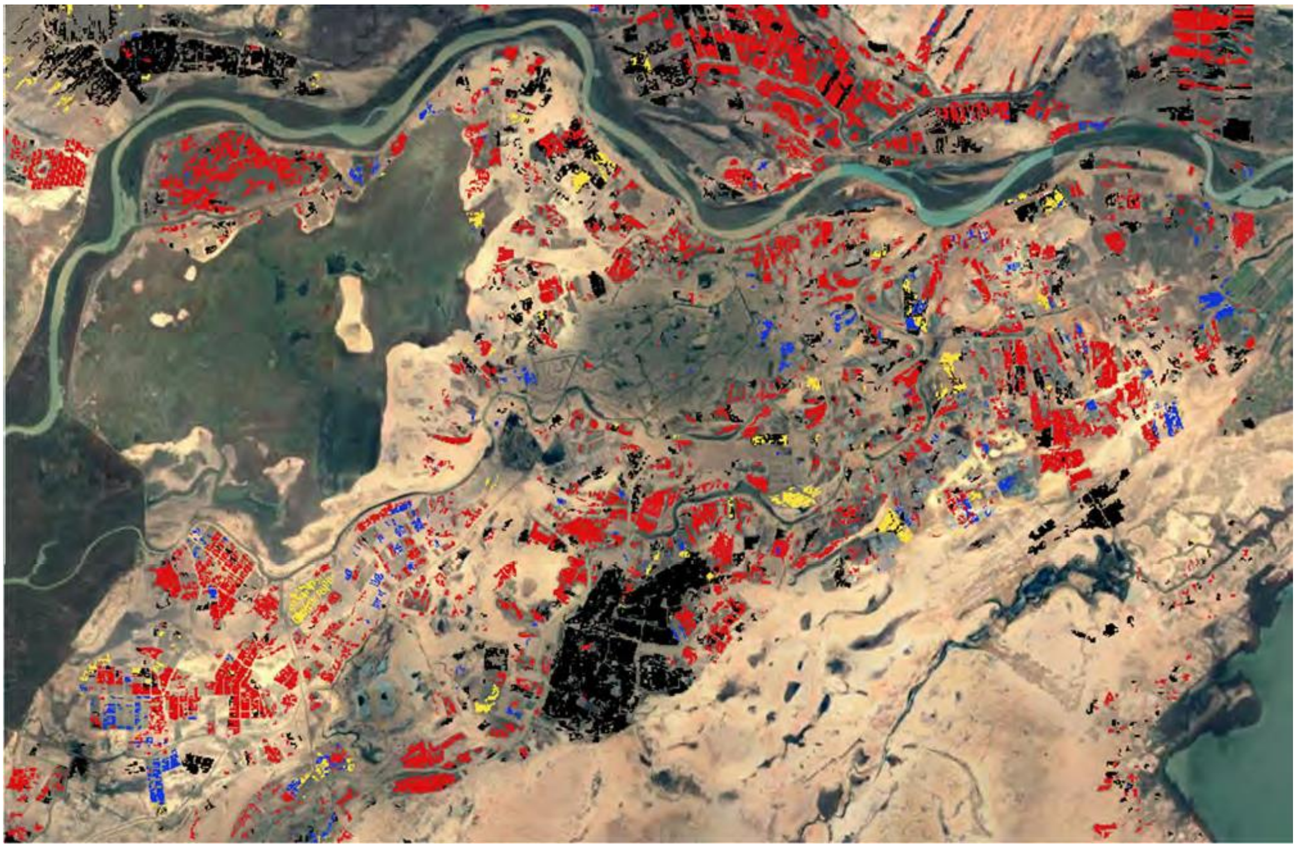


Figure 45: 2019 - 2018 change of practice in the seasons of some plots

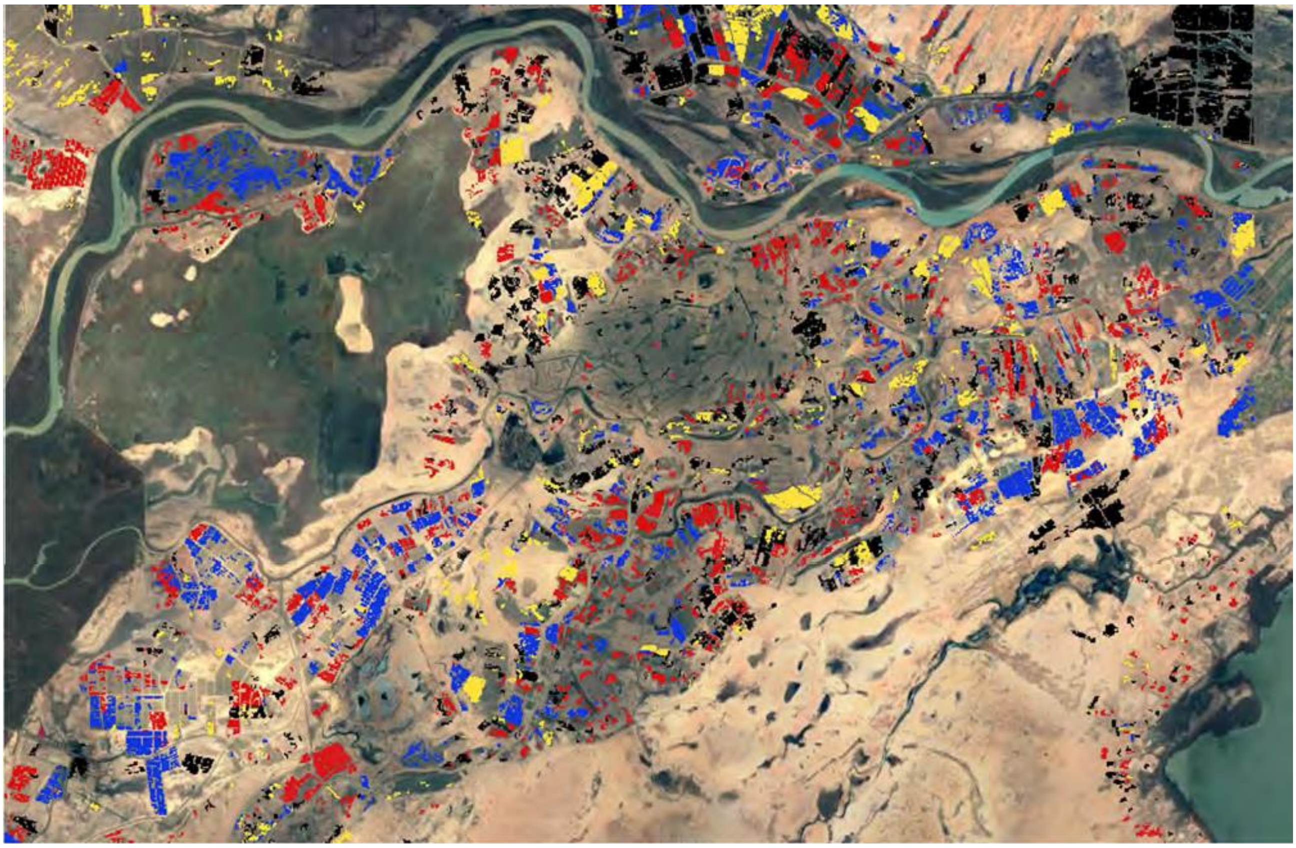


Figure 46: 2020 - 2018 change of practice in the seasons of some plots

To make interpretation easier, it was decided to group these changes into potential productivity gains and losses.

In red the losses:

- Double season in year 1 season 1 in year 2
- Double season in year 1 season 2 in year 2
- Season 1 in year 1 to season 2 in year 2: the explanation here lies in a description made by local actors who confirmed that season 2 (irrigated) is less productive than season 1 (rainfed)
- Activity in year 1 to no activity in year 2:

- No activity in year 1 to activity in year 2
- Season 2 in year 1 to season 1 in year 2
- Season 1 or 2 in year 1 to double season in year 2

The following images show the gain/loss maps for the 2018 to 2019; 2019 to 2020 and a 2018 to 2020 assessment respectively:

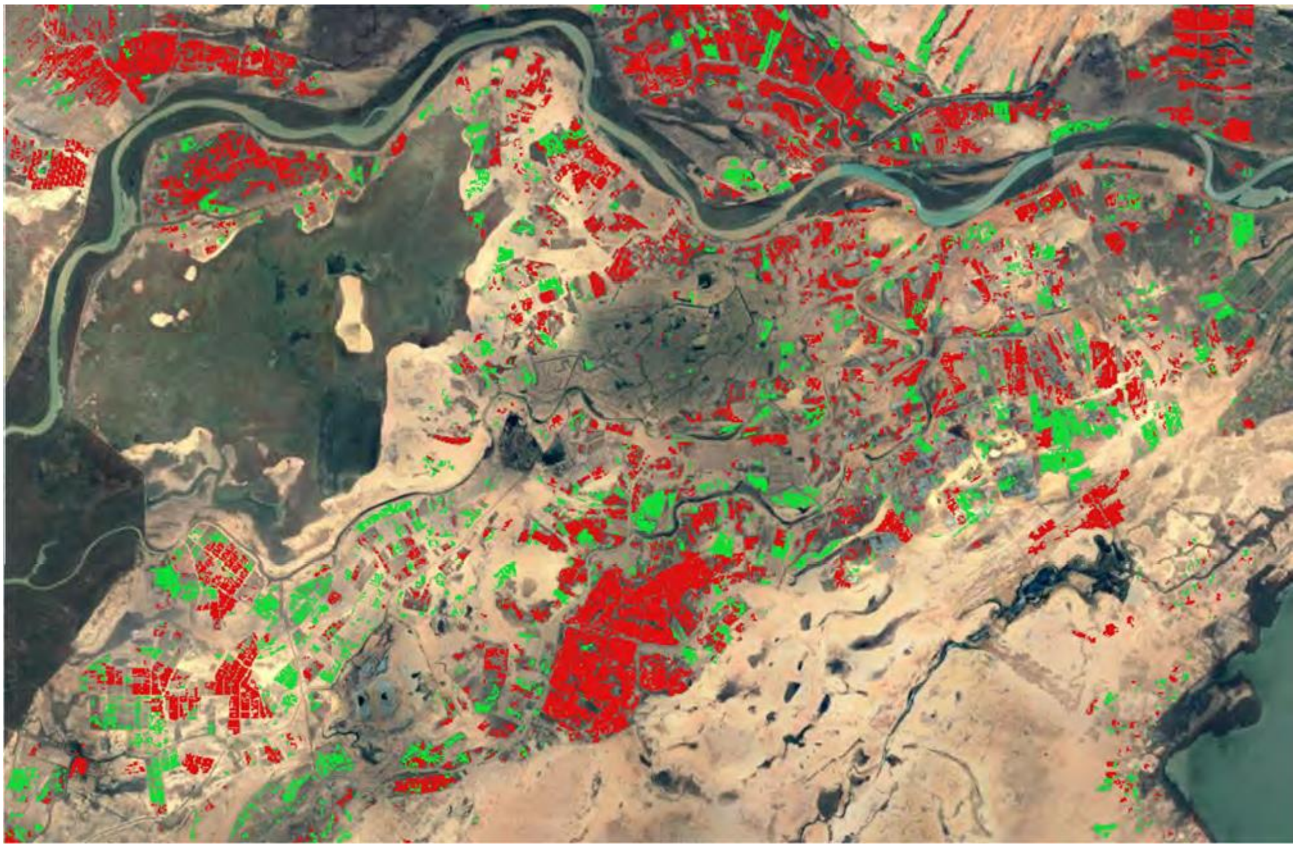


Figure 47: Gain/loss map 2018 - 2019

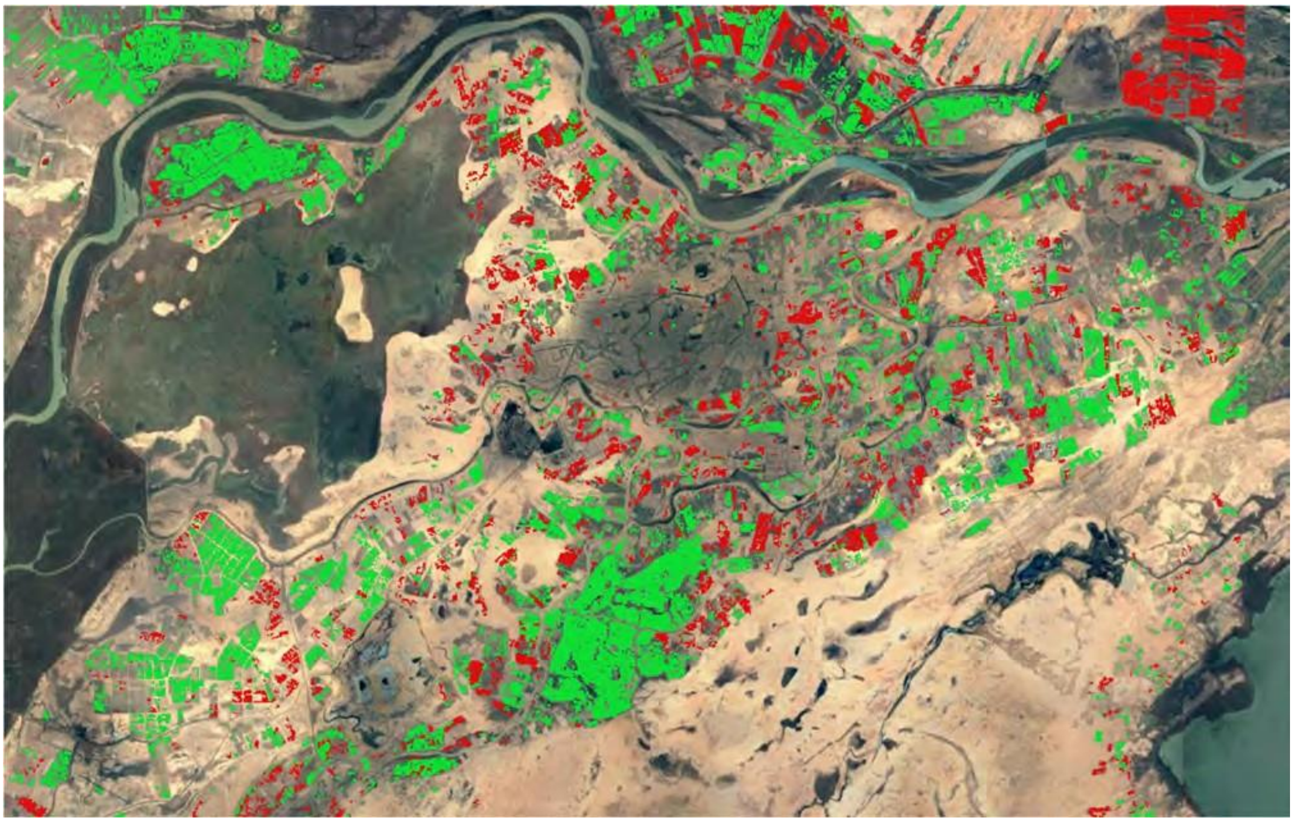


Figure 48: Gain/loss map 2019 - 2020

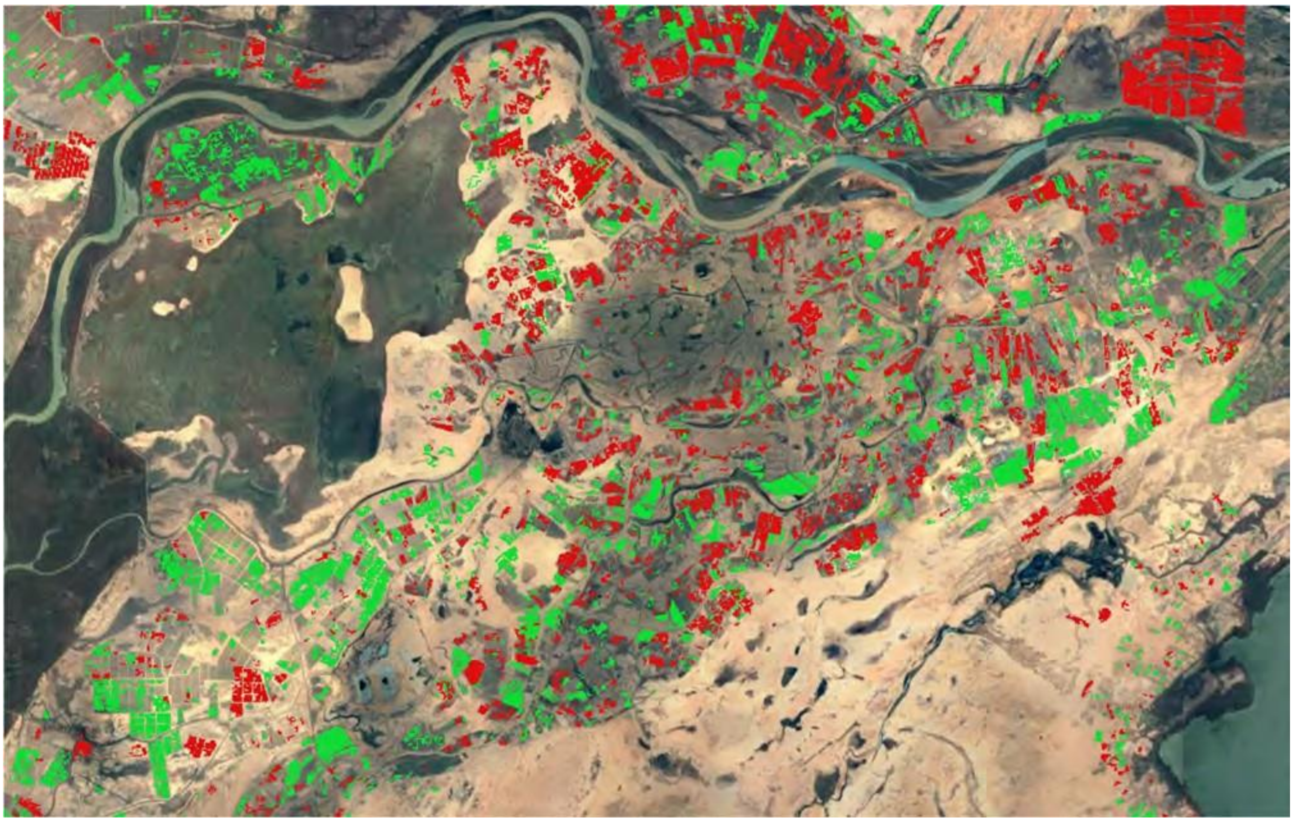


Figure 49: Gain/loss map 2018 - 2020

It is quite easy to see visually that the 2019 drought has had an impact on both losses and gains between 2018 and 2020, while the comparison between 2018 and 2020 shows changes that are less radical. Thus these maps can be used to assess the potential impact on productivity from one year to the next, but also to anticipate the year 2021, for example, in order to target plots that need to be changed, stopped (for leaching, for example) or reoriented. Data in hectare can be derived from these maps in order to have a more precise overview of the number of areas in gain or loss.

5.2 Development indicator on safeguarding water ecosystems

The establishment of indicators on the safeguarding of water systems aims to quantify the permanent and seasonal water surfaces and to safeguard the quality of this strategic resource by trying to minimise the potential pollution impacts in these water surfaces and particularly those coming from urban areas. The most significant pollution comes from dirty water (SDG 6.3) from homes, businesses and industries that is discharged and untreated and can cause serious disturbances in the quality of water ecosystems. In addition, water run-off from man-made areas and agricultural plots can also be a serious source of pollution. Therefore, this indicator in the glow project attempts to measure, from the data available in this project, the potential sources of pollution from man-made areas.

For this purpose, we use the data from the databases evaluated during the project and the land use maps generated by the Iota-2 chain to set up this indicator. This indicator represents only a sample and a globalised view of the proximity situation between water ecosystems and artificial areas. A known limitation of this indicator is the specification of the pollution (names of pollutants and origins) that can occur on water surfaces. The indicator does not aim to define the pollutants, waste or landfill water that contaminate the water, but to highlight areas where strategic decisions must be taken to safeguard the resource by quantifying the hectares of surface water present each year on a perennial and seasonal basis and by studying their proximity to the urban distribution in predefined buffers of 500 m and 1000 m.

The study area was defined as the space centred on N'Djamena and covering a radius of 50 km around the city centre. The area straddles Chad, Cameroon and Nigeria and covers 13,500 km². It is centred on a conurbation of more than one million inhabitants along the Chari River, which flows into Lake Chad 100 km north of N'Djamena.

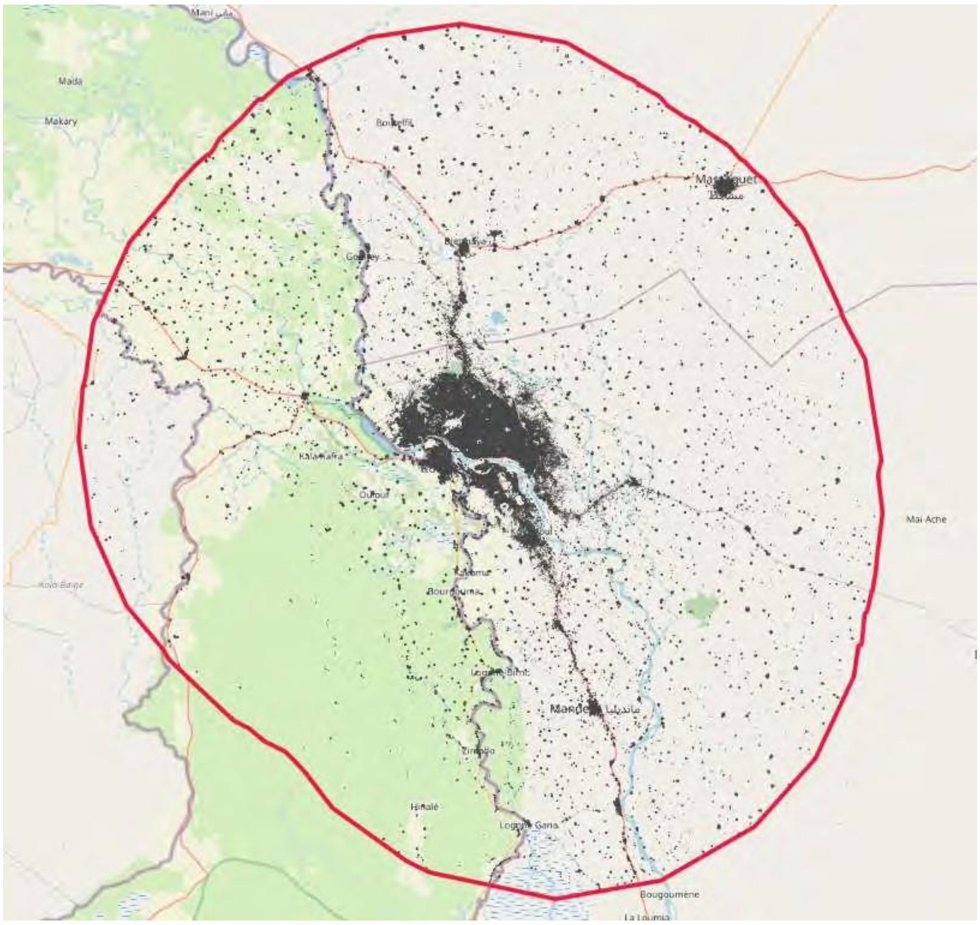
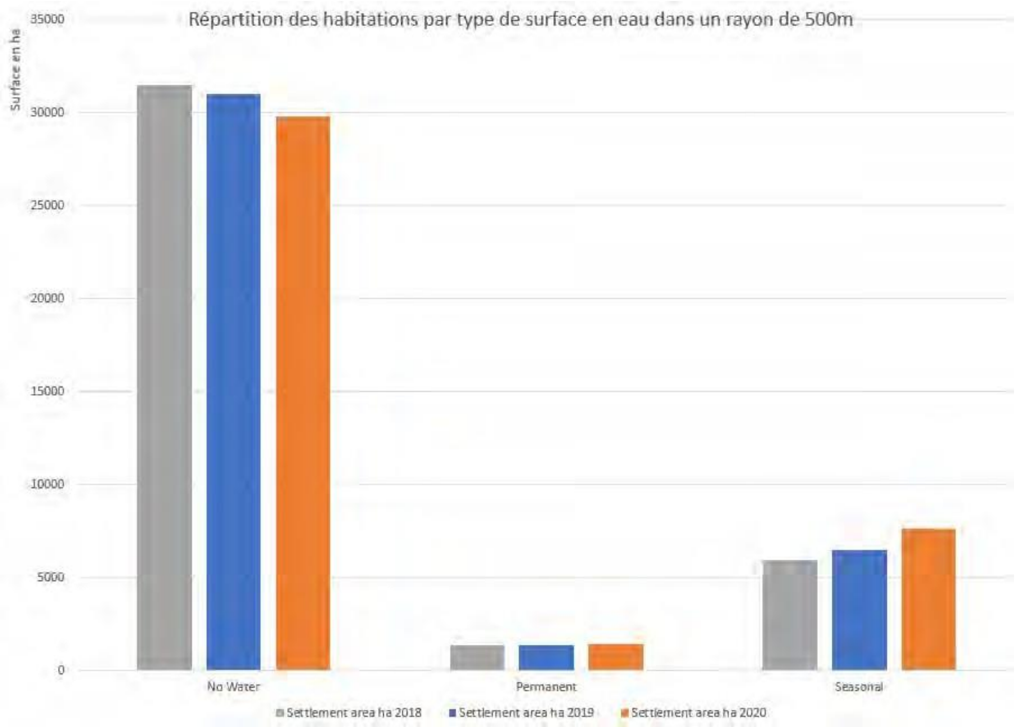


Figure 50: N'Djamena study area

To carry out this analysis, the WSF 2019 layer is used to represent dwellings and the Global Surface Water layer for water areas. The latter layer distinguishes between permanent and temporary, seasonal water.

The first indicator assesses the distribution of artificial surfaces within 500 m of permanent and seasonal waters (Figure 51 a). The comparison is made between 2018, 2019 and 2020 for water surfaces while using a constant housing layer (WSF 2019). This indicator is also calculated for a radius of 1000 m (Figure 51 b)



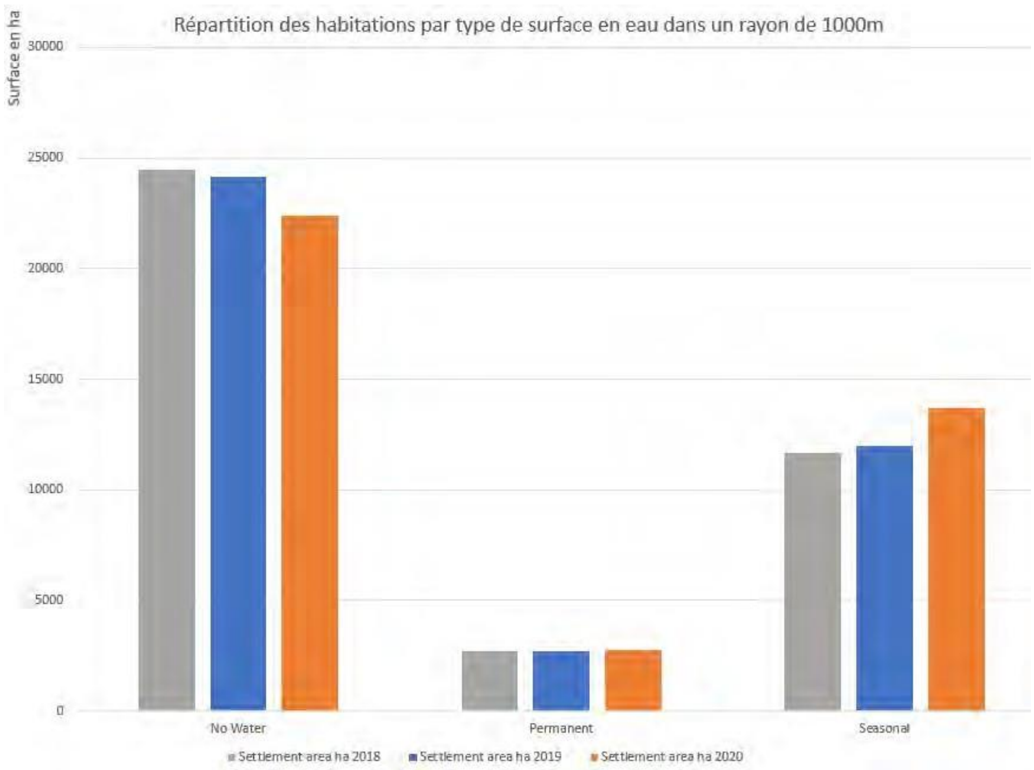


Figure 51: Distribution of dwellings by type of water surface within 500m (a) and 1000m (b)

At 1000 m, the proportion of dwellings in relation to water areas is significant and increases especially for seasonal waters.

The following analysis aims to determine the proportion of water surfaces away from urban areas.

The following graph shows that permanent waters are predominantly within 500 m of habitats and almost completely within 1000 m in all years.

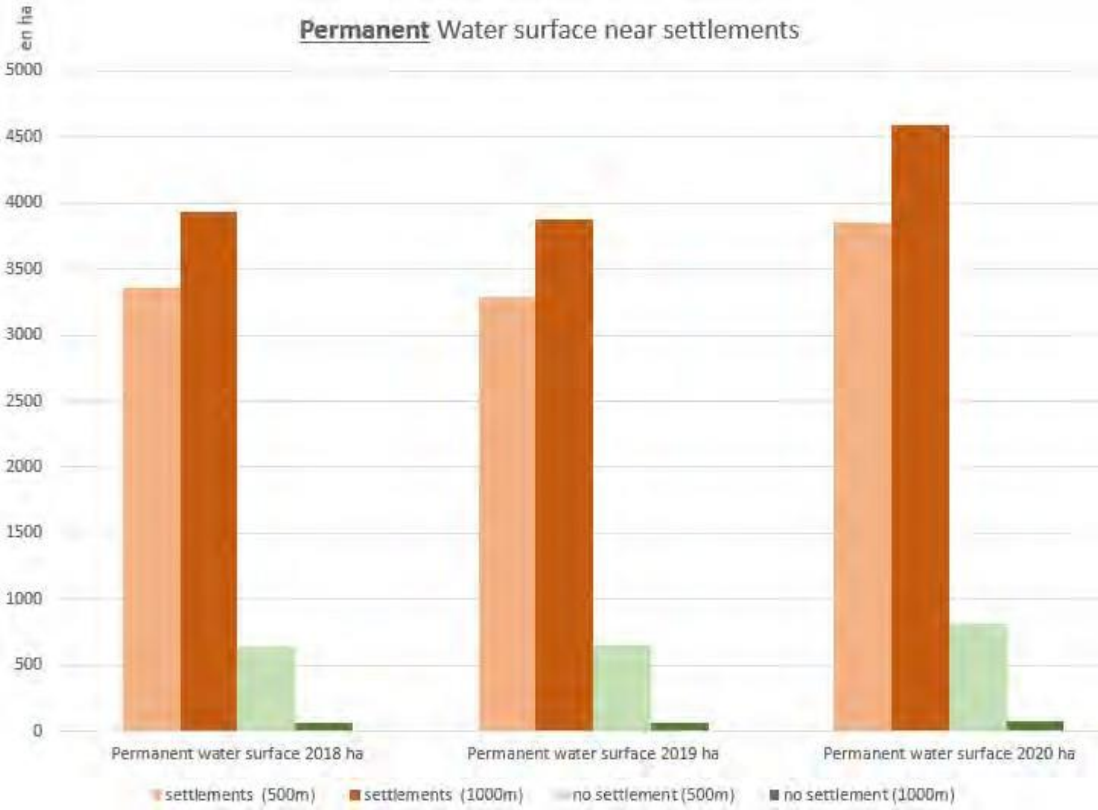


Figure 52: Permanent water surfaces near habitat areas

For seasonal water surfaces the distribution is more homogeneous, of the same order of magnitude in 2018 at 500 m and the majority at 1000 m. The pattern is reversed in 2019 and 2020 due to a clear increase in water surfaces in the western part. One third of the seasonal water surfaces are within 500 m of a dwelling. One third of the surfaces are between 500 and 1000 m and the last third is more than 1000 m from dwellings. The analysis is similar in 2020.

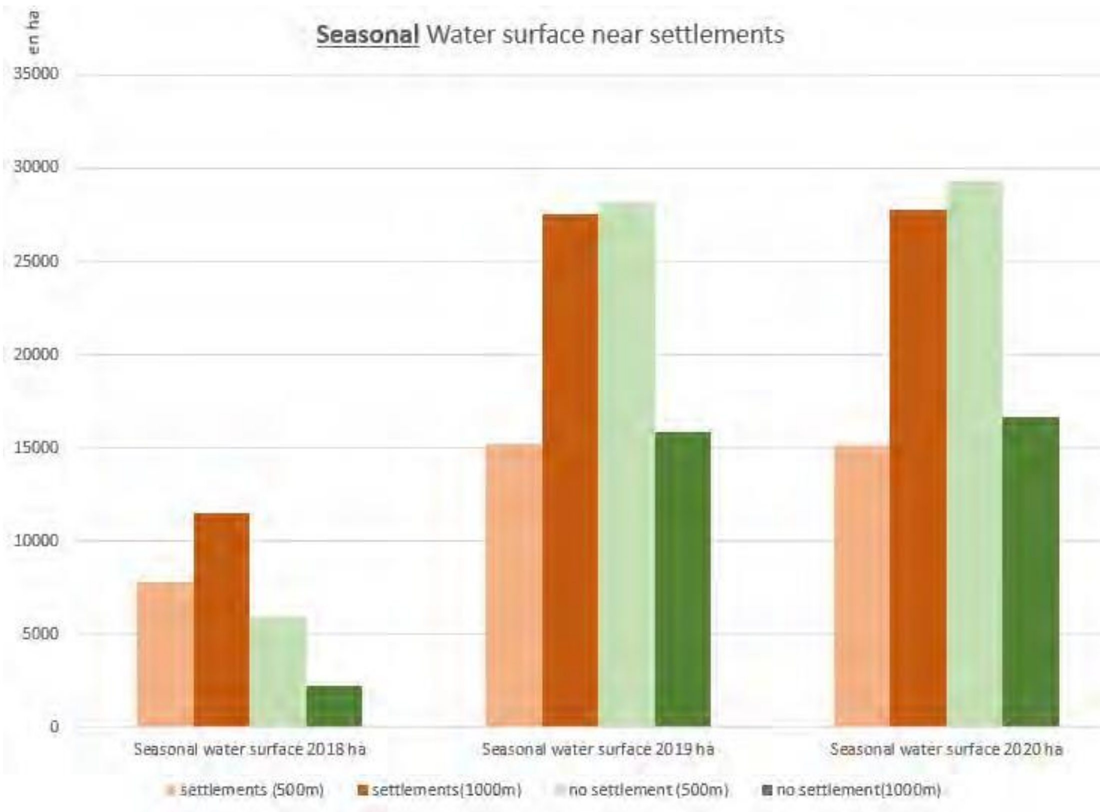


Figure 53: Seasonal water surfaces near habitat areas

The following maps illustrate the analysis and statistical results:

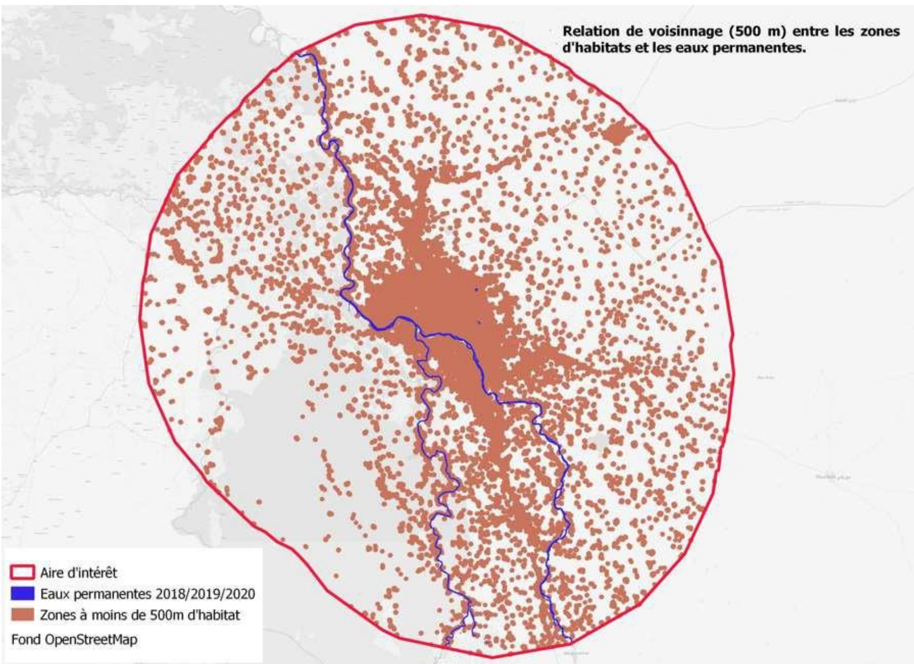


Figure 54: Buffer zone around the 2019 habitats and permanent waters in 2018, 2019 or 2020

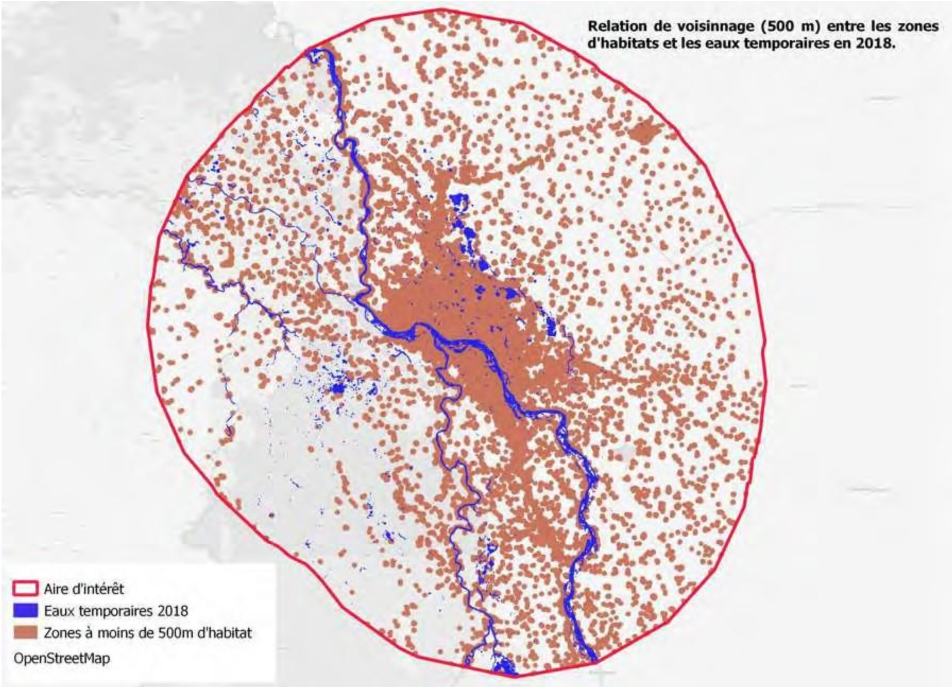


Figure 55: 500m buffer zone around 2019 habitats and temporary waters in 2018

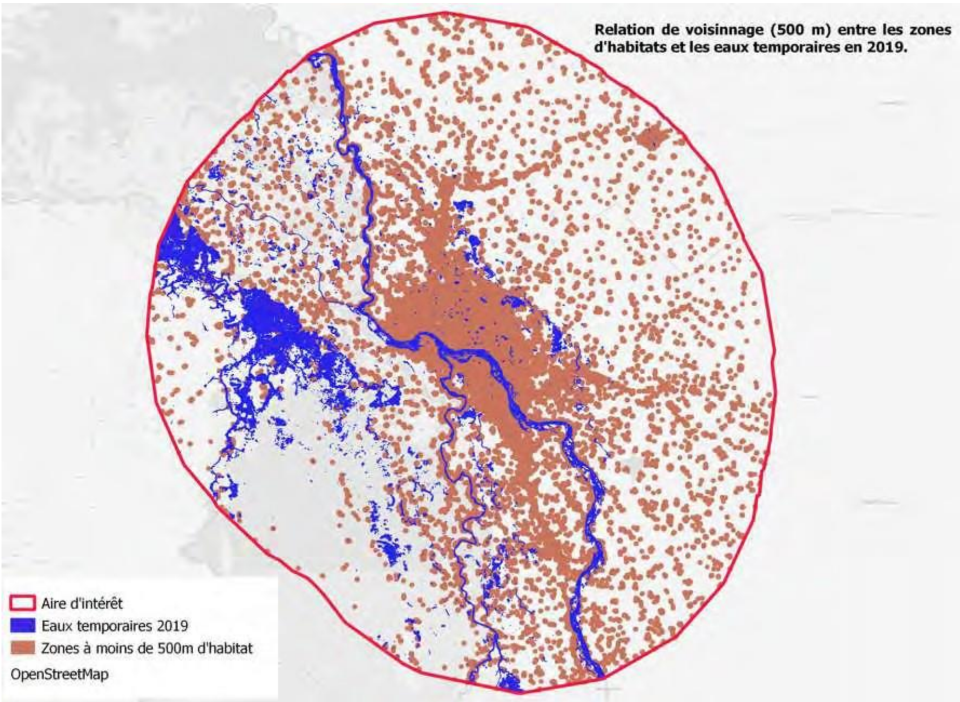


Figure 56: 500m buffer zone around 2019 habitats and temporary waters in 2019

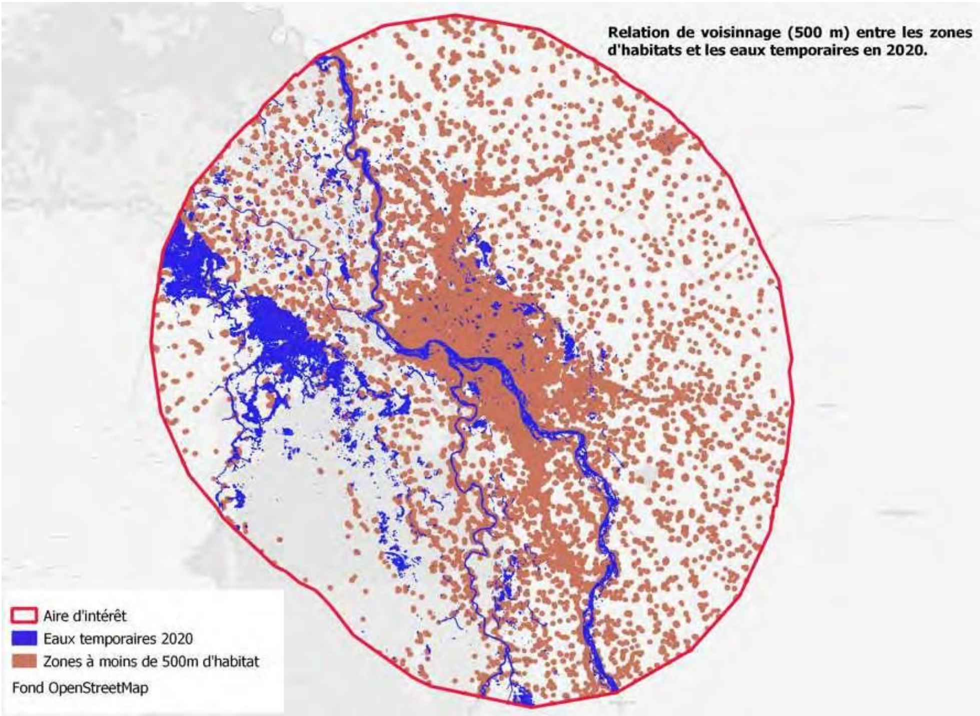


Figure 57: 500 m buffer zone around habitats in 2019 and temporary waters in 2020

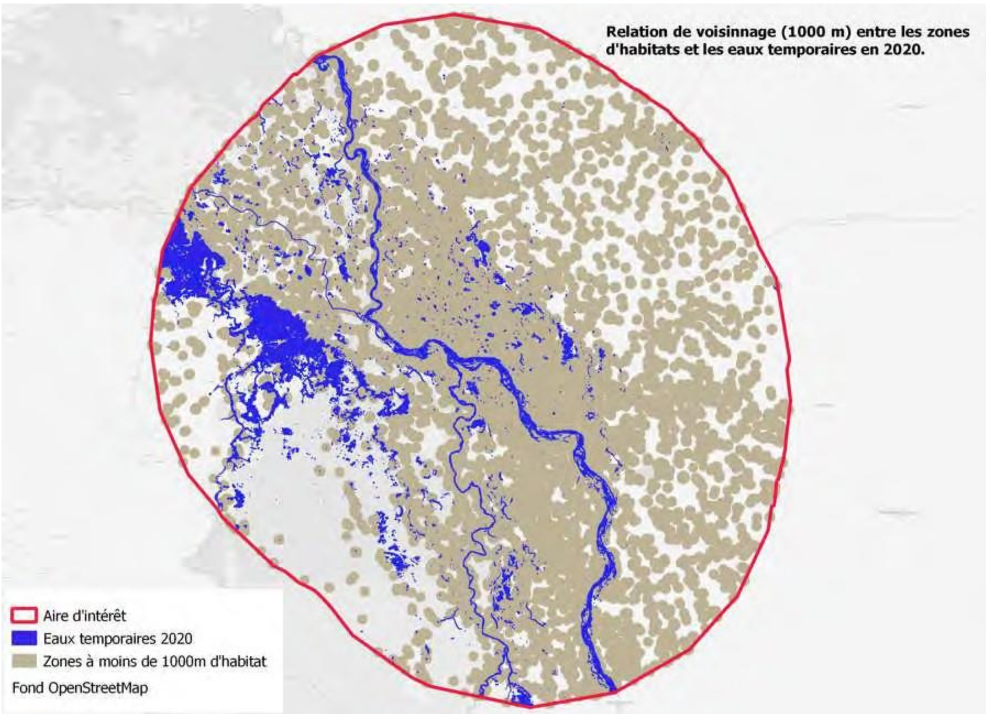


Figure 58: 1000m buffer zone around habitats in 2019 and temporary waters in 2020

To take this indicator further, information from in-situ stations studying water quality or studies on turbidity and water colour would be the next links in the chain. Earth observation data can be used to detect certain constituents that have an effect on water colour, for example, or on temperature and salinity. In addition, information on groundwater and the diversion of water volumes present could be the next step in the development of this indicator to specify potential environmental pollution damage.

5.3 Development indicator on urban dynamics

This spatial indicator consists of a comparison of two components, namely population growth and land consumption. According to the UN-Habitat methodological guidance document2 , the population growth rate (PGR) is the change (usually increase) in the population of a country over a specific period of time, usually one year. The PGR is expressed as a percentage of population gain or loss relative to the population at the beginning of the period.

Similarly, the rate of land consumption aims to provide information on the gain or loss of land used for human activities over a given period. Land consumption includes built-up areas that can be directly measured.

The indicator is calculated using the following formula:

Ratio between the Land Consumption Rate and the Population Growth Rate :

$$(\text{TCTTCP}) = \frac{\text{land consumption rate population}}{\text{growth rate}}$$

This indicator supports the assessment of urban dynamics, i.e. mainly urban sprawl and densification. In order to achieve the objectives of sustainable development, it is very important to limit urban sprawl and to densify existing areas because all the infrastructures are already in place. This densification will be effective and considered sustainable if a city has a population growth higher than the growth of its area.

The input data used are :

- Urban extent :
  - World Settlement Footprint 2019: is a 10m resolution bitmask describing the extent of settlement in the world, obtained using the Sentinel-1 and Sentinel-2 multi-temporal series of 2019
  - World Settlement Footprint Evolution (1985-2015): is a 30m resolution layer describing the extent of the world's settlement on an annual basis from 1985 to 2015, derived using Landsat-5 and Landsat-7 multi-temporal series.
  - Population census data: WorldPop 2000-2020 includes population areas by country at annual time steps between 2000 and 2020, with a spatial resolution of 3 arc seconds (about 100 m at the equator).

The methodology used is: Calculation of the rate of evolution of the urban area as well as the number of inhabitants over the 20 year period, before applying the formula given in the definition of the indicator. The WSF Evolution layer makes it possible to evaluate the built-up area between 2000 and 2015. The interpolated value of the area obtained in 2019 and the calculated value from WSF 2019 show a slight difference. As the WSF 2019 layer has a higher level of reliability, this discrepancy is reflected for all years in order to obtain a corrected habitat area curve between 2000 and 2020 (red curve in Figures 57-58).

[2] *UN-Habitat (2017): A Guide to Assist National and Local Governments to Monitor and Report on SDG Goal 11+ Indicators. Monitoring Framework - Definitions - Metadata - UN-Habitat Technical Support*

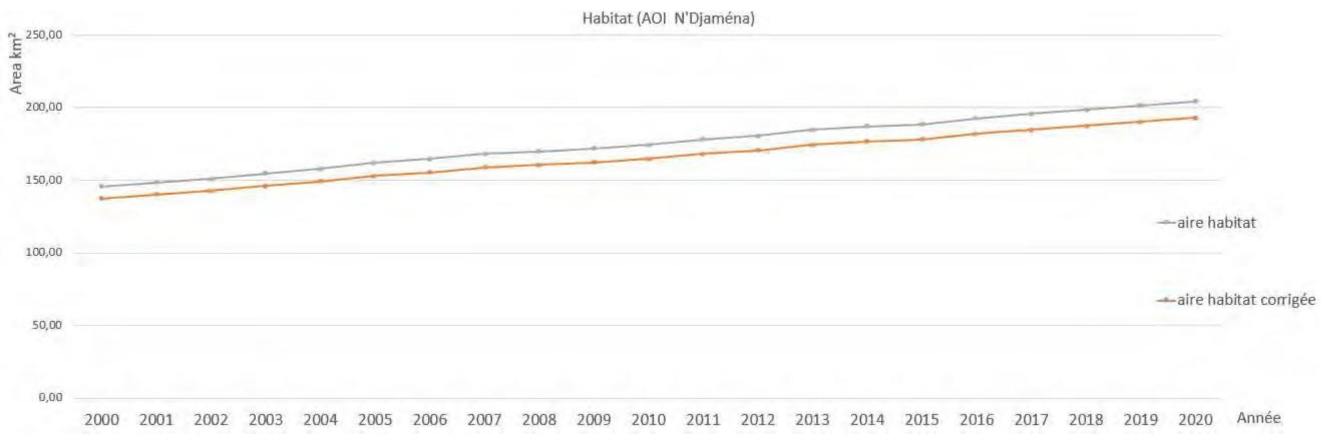


Figure 59: WSF from 2000 to 2020

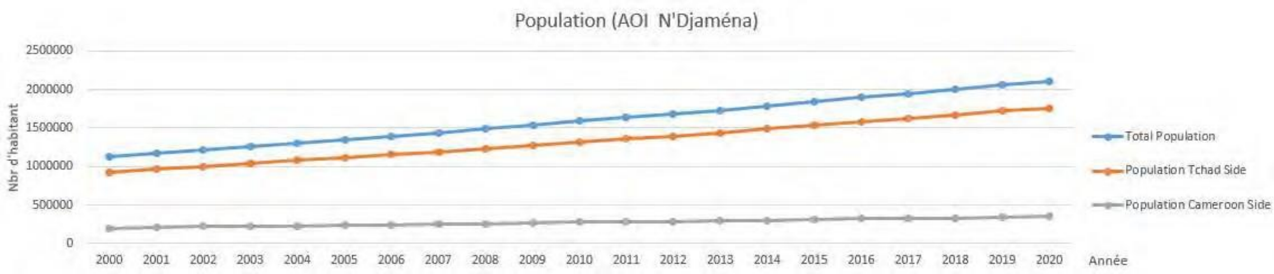


Figure 60: Population from 2000 to 2020

The indicator makes it possible to evaluate the evolution of the urban area in relation to the number of inhabitants. If negative values are obtained, this means that one of the phenomena has decreased during the period. Absolute values of less than 1 mean that land consumption is evolving less quickly than the population in relative terms, which indicates greater densification than urban sprawl.

Ratio of Land Consumption Rate to Population Growth Rate

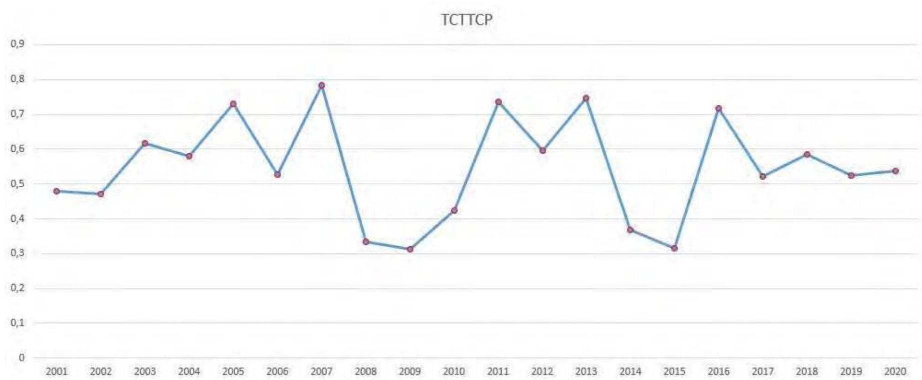


Figure 61: Relationship between the rate of land consumption and the rate of population growth over the period 2000-2020

The ratio has been calculated for the period 2000-2020. The figure shows a faster population increase than urban expansion in relative proportions. With an overall index for the period 2000-2020 of 0.5, the trend can be considered to be towards densification.

6 Conclusion

The ECLAT project has made it possible to carry out a study on the data made freely available in the field of remote sensing, covering several themes, the capacities of the Iota-2 chain to produce land use maps in extreme and complex environments and, above all, to produce sustainable development indicators from general data and remote sensing indices.

Thus, the SCO ECLAT project made it possible to draw up a complete inventory of the free global and continental African databases available for producing land use maps. This inventory of more than 25 databases allowed the evaluation of the most recent databases with regard to the project's themes, namely urban, water surfaces, agriculture, wetlands, bare soil and vegetation. The inventory allows for the identification of more precise trends in the most studied themes, notably vegetation, urban and water surfaces, and shows that these themes encompass the most recent databases with the highest spatial resolution. The study of forests, bare soil or agricultural land, which are much more complex to map in Africa, is more difficult and therefore has fewer recent, very high resolution or accurate databases on these subjects.

The SCO ECLAT project proposed to create validation databases by *random sampling* method on the Lake Chad region, the Tocc Park in Senegal and the W Park in Niger, thanks to which a validation and scoring of all the above-mentioned databases was performed. For each database, the data on the project's themes was extracted, formatted and validated against the corresponding validation database carried out by photo-interpretation of points. Evaluation criteria were set up to complete this evaluation and made it possible to give an overall score to each database and scores by theme.

Once this scoring was done, the best databases were used to create ground truths to be provided to a classification chain and to perform data fusion in order to create land cover maps. These land cover maps were created from the processing of Sentinel-2 annual time series over several tiles. The CESBIO Iota-2 chain was used to produce these land cover maps.

Database bias was a major obstacle to the creation of the land cover maps, whose overall accuracy peaked at 77%, exceeding the existing products for the areas of interest by a few percent. The majority of the land cover classes have overall accuracies between 60% and 65%. Despite significant efforts in configuring the Iota-2 chain for the production of land cover maps, the overall accuracy was difficult to improve. The complex environment of Lake Chad makes it difficult to separate certain classes (urban vs. bare soil; bare soil vs. vegetation; wetland vs. vegetation) and obtain accurate land cover maps.

Nevertheless, these land use maps and the processing of Sentinel-2 time series have been used to create sustainable development indicators. These indicators are divided into three themes: agricultural, urban and water management. The first *Dynamic Habitat Index* (DHI) was used to create a promising agricultural monitoring indicator, the usefulness of which is of great interest to one of the project's partners, the Centre de Suivi Ecologique in Senegal. This indicator makes it possible to monitor agricultural developments and would be useful for decision-makers in Senegal in the context of problems inherent in agricultural practices in the region. The second indicator on the monitoring of water ecosystems allows an evaluation of the risks inherent in the proximity of urban areas in terms of pollution and conversely on the risk of submergence of certain urban areas in the event of flooding. Finally, a last indicator for monitoring urban dynamics has been set up. This indicator makes it possible to monitor the relationship between land consumption and population densification and would be useful in monitoring urban planning.

The work of the SCO ECLAT project has established a basis for future work on the selected themes. In view of the difficulties encountered and the lack of local contacts and therefore of data or evaluation of intermediate products, we were able, with considerable effort, to establish a relationship with the CSE in Senegal and propose products that could also serve as a basis for future projects

Annex I: Validation Results

			Artificial area	Water surface	Cultivable Land	Wetlands	Bare ground	Vegetation
Chad - 33PTP	2018	Precision	0.770	0.771	0.684	0.667	0.333	0.595
		Recall	0.534	0.875	0.650	0.571	0.344	0.780
		Commission	0.230	0.229	0.316	0.333	0.667	0.405
		Omission	0.466	0.125	0.350	0.429	0.656	0.220
		F-Score	0.631	0.820	0.667	0.615	0.338	0.675
		Overall accuracy	**0.665**					
	2019	Precision	0.774	0.722	0.602	0.787	0.423	0.542
		Recall	0.545	0.948	0.530	0.527	0.344	0.780
		Commission	0.226	0.278	0.398	0.213	0.577	0.458
		Omission	0.455	0.052	0.470	0.473	0.656	0.220
		F-Score	0.640	0.820	0.564	0.632	0.379	0.639
		Overall accuracy	**0.649**					
	2020	Precision	0.770	0.699	0.638	0.807	0.314	0.692
		Recall	0.534	0.990	0.810	0.505	0.344	0.630
		Commission	0.230	0.301	0.362	0.193	0.686	0.308
		Omission	0.466	0.010	0.190	0.495	0.656	0.370
		F-Score	0.631	0.819	0.714	0.622	0.328	0.660
		Overall accuracy	**0.677**					
Chad - 33PUP	2018	Precision	0.667	0.931	0.504	0.476	0.579	0.594
		Recall	0.315	0.960	0.590	0.827	0.468	0.380
		Commission	0.333	0.069	0.496	0.524	0.421	0.406
		Omission	0.685	0.040	0.410	0.173	0.532	0.620
		F-Score	0.427	0.945	0.544	0.604	0.518	0.463
		Overall accuracy	**0.606**					
	2019	Precision	0.667	0.912	0.534	0.617	0.625	0.632
		Recall	0.315	0.939	0.630	0.837	0.319	0.720
		Commission	0.333	0.088	0.466	0.383	0.375	0.368
		Omission	0.685	0.061	0.370	0.163	0.681	0.280
		F-Score	0.427	0.925	0.578	0.710	0.423	0.673
		Overall accuracy	**0.662**					
	2020	Precision	0.667	0.904	0.542	0.524	0.563	0.621
		Recall	0.315	0.949	0.520	0.878	0.383	0.590
		Commission	0.333	0.096	0.458	0.476	0.438	0.379
		Omission	0.685	0.051	0.480	0.122	0.617	0.410
		F-Score	0.427	0.926	0.531	0.656	0.456	0.605
		Overall accuracy	**0.632**					

			Artificial area	Water surface	Cultivable Land	Wetlands	Bare ground	Vegetation
Chad - 33PWP	2018	Precision	0.586	0.880	0.611	0.626	0.549	0.762
		Recall	0.494	0.950	0.800	0.691	0.394	0.640
		Commission	0.414	0.120	0.389	0.374	0.451	0.238
		Omission	0.506	0.050	0.200	0.309	0.606	0.360
		F-Score	0.536	0.913	0.693	0.657	0.459	0.696
		Overall accuracy	**0.681**					
	2019	Precision	0.586	0.851	0.552	0.702	0.575	0.734
		Recall	0.494	0.970	0.850	0.680	0.324	0.580
		Commission	0.414	0.149	0.448	0.298	0.425	0.266
		Omission	0.506	0.030	0.150	0.320	0.676	0.420
		F-Score	0.536	0.907	0.669	0.691	0.414	0.648
		Overall accuracy	**0.672**					
	2020	Precision	0.594	0.803	0.601	0.694	0.636	0.758
		Recall	0.494	0.980	0.830	0.701	0.296	0.690
		Commission	0.406	0.197	0.399	0.306	0.364	0.242
		Omission	0.506	0.020	0.170	0.299	0.704	0.310
		F-Score	0.539	0.883	0.697	0.697	0.404	0.723
		Overall accuracy	**0.690**					
Senegal - 28QCD	2018	Precision	0.818	0.953	0.757	0.875	0.779	0.784
		Recall	0.375	0.990	0.875	0.770	0.822	0.791
		Commission	0.182	0.047	0.243	0.125	0.221	0.216
		Omission	0.625	0.010	0.125	0.230	0.178	0.209
		F-Score	0.510	0.970	0.810	0.820	0.800	0.790
		Overall accuracy	**0.828**					
	2019	Precision	1.000	0.990	0.669	0.710	0.713	0.748
		Recall	0.273	0.971	0.844	0.710	0.744	0.727
		Commission	0.000	0.010	0.331	0.290	0.287	0.252
		Omission	0.727	0.029	0.156	0.290	0.256	0.273
		F-Score	0.430	0.980	0.750	0.710	0.730	0.740
		Overall accuracy	**0.767**					
	2020	Precision	1.000	0.990	0.683	0.742	0.650	0.794
		Recall	0.375	0.971	0.896	0.690	0.700	0.700
		Commission	0.000	0.010	0.318	0.258	0.351	0.206
		Omission	0.625	0.029	0.104	0.310	0.300	0.300
		F-Score	0.550	0.980	0.770	0.720	0.670	0.740
		Overall accuracy	**0.772**					