THE WORLD BANK

Mapping irrigated areas and identifying areas suitable for irrigation in Mali and Chad

PRACTICA FOUNDATION

Author
Dr.M.T. Westra

Author: M.T. Westra, PhD.
Affiliation: Practica Foundation
Date: February 2021
Front image: Smallholder irrigation near the Niger river in Mali, near Tamani in Ségou region, (13.4093,-6.6923).
Version: 4
Status: final

TABLE OF CONTENTS

ACKNOWLEDGMENT

This report was made possible by the assistance provided by a number of people, to whom the authors express their gratitude.

First and foremost, we thank the team at the World Bank, where in particular Pierrick Fraval, Mutsa Cecilia Masiyandima, and David Casanova have provided detailed comments and helpful suggestions. Their sharp eye for detail have had a large impact on the final result.

In addition, we thank prof. Stefan Siebert, for sharing with us GIS materials related to the irrigation situation in Mali and Chad.

The team in charge of this study carried out by the PRACTICA Foundation:

- Mark Tiele Westra, senior ICT and GIS specialist, main author.
- Stephan Abric, senior expert in small scale irrigation, second author.

TECHNICAL TERMS

Remote sensing	The process of measuring and analysing present and historic imagery data from satellites with the aim to gain knowledge about location conditions, such as climate, land use, agriculture, air quality, water quality, etc.
Resolution	The spatial resolution of the imagery of satellites in meters. A resolution of 30m means that each pixel in the satellite image has a size of 30x30 meters.
Pixel	A single picture element or raster point of an image. “This image is 400 x 500 pixels wide”.
Sentinel-2	Sentinel-2 is an Earth observation mission from the European Copernicus programme, that acquires high resolution (10m to 60m) multi-band data every 5-10 days on almost every place on Earth. It consists of two satellites, Sentinel-2A and Sentinel-2B
Satellite bands	A satellite collects imagery information on a number of frequency bands. In addition to the visible spectrum frequencies (Red, Green and Blue), data is collected in 10 additional bands, among which near infrared and short-wave infrared, which are sensitive to plant growth.
Mosaic Image	Satellites take individual photos of a certain size (around 290 km). However, some images cannot be used due to cloud cover. To help processing, one large ‘mosaic’ image is constructed from the best images in a given timeframe. This timeframe can cover many months.
Raster data	Other term for a geographically referenced image such as a satellite image, or other data that is expressed as a raster of pixels with values.
Google Earth Engine	Google Earth Engine is a cloud-based platform for planetary-scale environmental data analysis. It combines the access to peta-byte scale geospatial data sources, such as Sentinel-2, with the enormous data-crunching capabilities of Google.
Machine learning	Using a computer to automatically build a mathematical algorithm based on training data, in order to make predictions or classifications, without being explicitly programmed to do so. Machine learning is used in a variety of tasks, such as computer vision, email filtering, financial fraud detection, etc.

Classification	Using an algorithm to identify to which of a set of categories a data point belongs, on the basis of a training set of data containing data points of which the classification is known. For example: determining if an email is 'spam' or 'non-spam', or if a given satellite image pixel shows irrigated or non-irrigated land.
Random Forest	A specific machine learning method to automatically create a classification algorithm from classified training data. Works on the basis of constructing a large number of decision trees, and combining their results.
Decision Tree	A predictive modelling algorithm that uses observations about an item to decide to which category it belongs.
Ground truth data	Data points that represent the 'ground truth' about a certain characteristic of a location, such as the fact if it is irrigated or not. Although the gold standard of ground truth data is to visit locations individually, in many cases ground truth data can be collected by manually studying very high-resolution satellite imagery.
Training data	The part of the Ground Truth data that is used to train the machine learning model. Usually about 70% of the total data.
Validation data	The part of the Ground Truth data that is used to test the accuracy of the generated machine learning model. Usually around 30% of the total data. This data has not been used during the training.
GeoTiff	A file format for storing geographically-referenced imagery data. The classification process in Google Earth Engine results in GeoTiff files with the classification result for each pixel.
Slippy web map	A web standard for sharing maps online. It can easily be used for displaying a map on a website, as well as loading maps into GIS software such as ArcGIS or QGIS.
Producer's Accuracy	The Producer's Accuracy is the map accuracy from the point of view of the map maker (the producer). This is how often are real features on the ground are correctly shown on the classified map.
Consumer's Accuracy	The User's Accuracy is the accuracy from the point of view of a map user, not the map maker. The User's accuracy tells us how often the class on the map will actually be present on the ground. This is also referred to as reliability.

ACRONYMS

GIS	Geographic Information System: computer software used to manipulate geographically referenced shape and raster data, such as maps and satellite data.
AQUASTAT	FAO's global information system on water resources and agricultural water management.
FLID	Farmer Led Irrigation Development - a process in which farmers drive the establishment, improvement and/or expansion of irrigated agriculture, often in interaction with other actors: government agencies, NGOs, etc.
WAPOR	The FAO Water Productivity Open-access portal. Uses remote sensing technologies to monitor and report on agricultural water productivity in Africa.
IMWI	International Water Management Institute.

NDVI	Normalized Difference Vegitation Index. Computed from Sentinel-2 bands, the NDVI is an index sensitive to green vegetation. Used to enhance plant growth on satellite imagery.
NDWI	Normalized Difference Water Index. Similar to NDVI, but more sensitive to water stress in plants.
EVI	Enhanced Vegitation Index. Version of NDWI that corrects for some atmospheric conditions and canopy background noise. More sensitive to areas with dense vegetation.
EVI_DIFF and NDWI_DIFF	Terms used in this report that refer to month-on-month differences between EVI or NDWI values. Used as input for the machine learning algorithm.
GEOBIA	Geographic Object-based Image Analysis. Using a computer algorithm to automatically divide a satellite image in 'regions' or 'objects', such as fields or buildings. Further analysis is then performed on the resulting objects. This contrasts with normal Remote sensing analysis, in which analysis is performed per pixel.
GFSAD30	Global Food Security-Support Analysis project. Identified agricultural areas in Africa (and other continents) at 30m resolution.
GLCM	Grey Level Co-Occurrence matrix. Raster data derived from satellite images that highlights texture and patterns present in the image. Used as an input for the machine learning process.
SNIC	Simple Non-Iterative Clustering. An image clustering algorithm implemented in Google Earth Engine. Can be used to identify 'objects' in images.
QGIS	Popular and open source Geographical Information System. Used to manipulate and analyse image and raster data.
GADM	Database of Global Administrative Areas (gadm.org). A popular and up-to-date database of GIS data for administrative areas.
NASA	The American National Aeronautics and Space Administration
FAO	The Food and Agriculture Organisation of the United Nations.

LIST OF TABLES

LIST OF FIGURES

Executive summary

This work forms part of the project “Diagnostic and future directions for Farmer Led Irrigation (FLID) in Chad and Mali”, which has the aim to assess the current extent of farmer-led irrigation nation-wide in Chad and Mali, and assess the areas that are suitable for further growth. In both countries, farmer-led irrigation exists, but the extent is not well known. As a first step towards the characterization of FLID, this report focusses on the identification of total irrigated areas (both small-scale and large-scale) for the full territories of Chad and Mali, using remote sensing and machine learning, and identifying areas suitable for small-scale irrigation by using a multi-criteria model. Combining the areas of actual irrigation with suitable areas, we determine regional zones where expansion of irrigation has a high potential.

In addition to this report, which presents the qualitative and quantitative results, we provide an online web map with the resulting maps, which can be found here. We conclude the report with a number of limitations of this study, and a list of recommendations for further work.

Areas under irrigation in the dry season

We determined irrigated areas in Mali and Chad in the period Oct 2019 – June 2020, using Sentinel 2 satellite data and machine learning, at a 30m resolution. Sentinel 2 data currently is the most high-resolution multi-bandwidth satellite data publicly available, capturing 10m resolution imagery on average every 5 days in the region of interest. Historic data is available since June 2015. Google Earth Engine, which provides easy data access and powerful computation capabilities, was used to analyse the data. The choice of period (Oct-Jun) was

driven by data availability — outside those periods significant areas of both countries are covered with clouds due to the rainy season. Therefore, areas that are only irrigated outside that period are not identified by this method.

We adopt a broad definition of irrigation as increased plant growth with respect to local natural conditions, in the period Oct-Jun. A total of 38 bands was used for the classification, including visible light, Near Infrared, and plant growth / water stress indicators (EVI and NDVI). A Random Forest machine learning algorithm was trained on 11.208 ground truth points in Mali, and 6354 points in Chad. In both countries, the data was split up into three climate zones to account for intra-country climate variability. The classification accuracy of the models was 95.6% for Mali, and 95.2% for Chad, indicating that the bands used provide sufficient information for an accurate classification of irrigated areas.

Our estimate of areas under irrigation for Mali is 565.6 kha, which is in line with the most recent estimate from AQUASTAT (621.3 kha) for the total agricultural area under water management, which includes flood recession cropping areas, cultivated wetlands and inland valley bottoms (as explained in section 2.1). Of this area, 338.7 kha is larger than 500 ha, 82.6 kha is between 100 – 500 ha, and 144.3 kha is smaller than 100 ha. In Mali, the area that is equipped for irrigation is 59.7% of the total area under water management according to AQUASTAT, showing that non-equipped flood recession, cultivated wetlands and valley bottoms play an important role in the total irrigated area.

Our estimate for areas under irrigation for Chad is 104.7 kha. Of this area, 33.9 kha is larger than 500 ha, 20.2 kha is between 100 – 500 ha, and 50.6 kha is smaller than 100 ha. Our estimate is less than the recent estimate from AQUASTAT, which is 155.3 kha. As the present method of analysis only captured irrigation during the dry period, one possibility for the discrepancy is irrigation of crops during the wet season, such as rice, which usually starts in May or June. In Chad, the area that is equipped for irrigation is only 19.5% of the total area under water management according to AQUASTAT, showing that the role of flood recession, cultivated wetlands and valley bottoms plays an even greater role in Chad.

For both countries, we provide a breakdown of irrigated areas by administrative boundaries.

AREAS SUITABLE FOR IRRIGATION EXPANSION

We used a multi-criteria scoring model to identify areas suitable for irrigation expansion. Input parameters used included surface water nearness, groundwater availability and aquifer properties, local slope, land use / land cover, national parks, and distance to cities. Three scenarios were used: one with surface water only, and two with surface water and groundwater at different depths.

Depending on the scenario used, our estimate for areas suitable for irrigation expansion in Mali varies between 2200 – 6700 kha, which is in line with a recent study with a similar approach (IWMI 2019). An important limitation of this type of study is that it does not take into account the total additional amount of water that can be used from an environmental perspective. This hydrological constraint requires significant modelling, which is beyond the scope of this report. Studies that do incorporate this constraint typically find results for the total potential area that can be *sustainably* irrigated that are substantially lower than the results arrived at in this report for areas *suitable* for irrigation. In the case of Mali, studies that incorporate the hydrological constraint indicate that only about 30% of the *suitable* area can actually be used for irrigation in a sustainable way.

For Chad, our estimates for suitable areas depending on the scenario vary between 2000 – 5900 kha. Studies that incorporate the hydrological constraint indicate that in Chad, only about 17% , of the suitable area can actually be used for irrigation in a sustainable way. This shows that in both countries, the actual amount of water that can be sustainably be used for irrigation forms an essential constraint for estimates of the total realizable irrigation potential.

These results highlight the importance of distinguishing 'suitable areas' from the fully realizable *potential of sustainable growth* for irrigation. The latter needs a careful analysis of the hydrological water balance, possible groundwater depletion and downstream impacts.

Due to the limitations of this type of modelling — additional local limiting conditions such as soil quality, land productivity, pollution, salinity, local ownership situation etc., are not taken into account — the results should be interpreted to identify overall suitable regions, and as a guide to select sites for more detailed consideration.

CONCLUSIONS

From our results, we conclude that remote sensing combined with machine learning performs well in the classification of irrigated areas in the dry season, *provided* that the local effects of irrigation can be clearly distinguished from natural processes. This is true for dry-season irrigation and locations with multiple crop cycles per year. In these cases, both large and small irrigated areas clearly stand out from nearby areas with natural growth, and are easily identified by the machine learning classification process.

However, in cases where the plant growth closely resembles natural growth, such as in irrigation at the end of the wet season, flood recession cropping, or valley bottoms, using machine learning is less effective, and the irrigated area will be underestimated. Valley bottom and flood recession agriculture are both common in the studied countries.

Using remote sensing and machine learning has the significant benefit that it is replicable over time, and can be automatically carried out over different periods, for example yearly, to determine trends. As the same set of ground truth data and the same trained model can be used over different years, this could be done by further automation of the process of classification developed in this report, by making use of a programming interface Google Earth Engine offers. Once set up, performing a new classification in a following year requires very little effort. In addition, historic trends can be identified by analysing historic Sentinel data.

A further conclusion is that due to the variability in types of irrigation, irrigation schedules, and in-country climate variation, the training of a classification algorithm can only take place in a region of limited spatial extent. In this report, each country was divided in three climate zones, leading to a total of six different sets of training data and classification models. Further limiting the extent of analysis, for example per region in a country, would increase the reliability of the classification.

In terms of the primary objective of the project — identifying the extent of farmer-led irrigation — we conclude that remote sensing methods as generally employed at the moment cannot make the distinction between small-scale and large-scale irrigation, due to their pixel-based analysis nature. In addition, the variation in degree of clustering and nature of irrigation in flood recession, cultivated wetlands, and valley bottoms means that remote sensing data alone does not contain enough information to specifically identify farmer-led irrigation as a category. In this study, we use the size of contiguous patches of pixels classified as irrigated as an indicator of the 'size' of an irrigated area. A different approach would be to manually identify large-scale irrigated areas, and to use these areas to deduce the FLID areas by subtracting them from the total areas as identified by the classification.

Our main recommendation is twofold:

1) Combine remote sensing-based analysis with other geospatial information on the location, management type and nature of different types of irrigation, such as maps of large-scale irrigation schemes and flood recession areas, as far as these are available from statistical bureaus in the given countries.

2) Limit the regional extent of analysis. By focussing on smaller areas, for example regions in a country, the problems with climate variability are avoided, and results become more reliable. In addition, focus on individual irrigation types when these have specific characteristics, such as flood irrigation.

For future work commissioned by the Bank, we think that most value will be obtained by always combining geographical layers with information on the location of large schemes, valley bottoms, flood recession irrigation, etc., with the information that can be derived from a machine learning classification.

A web map of both the irrigated areas and areas suitable for irrigation is available here:

Link to web map of Mali and Chad irrigated areas

1. INTRODUCTION

By 2050, food demands will increase by 60% to feed a population of nine billion people. While smallholder agriculture is the predominant form of farming in much of the developing world, agricultural production falls short of its potential due to lack of access and right to water for irrigation.

Traditionally, investments in irrigation have focused on large-scale systems. This is reflected in statistics of countries, in which large-scale systems are usually present, but smallholder agriculture is underreported.

Farmer-led irrigation development (FLID) is a concept that focusses on smallholder farmers, alone or as a collective, that drive irrigation development - meaning the establishment, improvement or expansion of irrigated agriculture by acquiring the necessary irrigation technologies and skills and developing output markets.

Many actors, such as the World Bank, recognize the importance of FLID in increasing productivity and enhancing food security. Supporting FLID starts with a thorough, local understanding of the current extent, suitable areas, and overall potential for farmer-led irrigation, and many actors are actively involved in supporting country-level diagnostics of farmer-led irrigation extent and potential for upscaling. The Sahel Irrigation Initiative (SIIP, PARIIS in French), is such a major initiative that focusses on Burkina Faso, Mali, Mauritania, Niger, Senegal and Chad.

PURPOSE OF THE WORK

In both Mali and Chad, farmer-led irrigation exists, but the extent is not well known. Current information, detailed below, is often of too coarse resolution, or does not make the distinction between irrigated and non-irrigated agriculture. The project "Diagnostic and future directions for Farmer Led Irrigation (FLID) in Chad and Mali", of which this report represents a first milestone, aims to assess the current extent of farmer-led irrigation in the whole territories of Chad and Mali, and assess the areas that are suitable for further growth.

This report focusses on the identification of total irrigated areas using remote sensing and machine learning and identifying areas suitable for FLID-type irrigation by modelling. The use of remote sensing for the identification of irrigated fields holds the promise to make it easier to update data on irrigation on a frequent basis, for example every year. In addition, irrigation could be observed throughout a growing season, which would help the understanding of which modes of irrigation are practiced, and where.

Secondly, the identification of areas suitable for FLID-type irrigation on a country scale is important as a guide to select promising areas for expansion. Detailed localized studies will always be necessary to cover the many factors that determine actual potential in any given location, the most important of which is the hydrological constraint.

STRUCTURE OF THE REPORT

Chapters 2 and 3 present the work on remote sensing and machine learning for the identification of irrigated areas. We present the result qualitatively on maps and quantify the irrigated area per region in each country. Chapters 4 and 5 present the identification of suitable areas for irrigation using a multi-criteria model. Three different scenarios are created covering different water availability conditions. By combining the areas of actual irrigation with areas of high suitability, we identify zones where irrigation expansion has a high potential. In Chapter 6, we discuss the results, and compare them to other studies. Chapter 7 discusses next steps for this project, and we conclude with a list of general recommendations for further work in chapter 8.

PRESENTATION OF RESULTS – WEB MAP

To make the results of this study as accessible as possible, the maps were made available as a web map, and as a Web Map Service (WMS). The latter means that the map layers can be loaded in regular GIS packages such as ArcGIS and QGIS. Sharing the results as a web map, including the layers for areas suitable for irrigation, will make it easier for users to explore different regions.

The web map is available here:

Link to web map of Mali and Chad irrigated areas

2. Actual irrigated areas – data and methodology

2.1. Introduction

In this chapter we introduce the methodology that was used to determine irrigated areas in Mali and Chad, using satellite data and machine learning. However, we first start with an overview of existing map materials on the location of irrigated areas in these two countries.

One of the authorative sources is AQUASTAT, the FAO database on irrigation information. The AQUASTAT statistics are partly based on maps that were produced by the FAO team.We obtained GIS data from prof. Stefan Siebert of University of Goettingen, who has been involved in calculating irrigated areas for both Mali and Chad for AQUASTAT. The material we received included a base map produced by the Ministry of Agriculture of Mali, a vector layer of irrigation projects, and a vector layer of irrigated areas.

For Chad, the information obtained from prof. Siebert included a base map produced by the SDEA, a vector layer of irrigation projects, a vector layer of irrigated areas, and a layer of irrigated areas as produced by SDEA. Source documents include (SDEA 2001) and (UN-DSD 2003).

The maps for Mali and Chad are reproduced below, with the vector layers of known irrigated areas superimposed. From the maps, we can conclude that the irrigated areas are indicated without a lot of detail, and that they are therefore not very suitable to be used for a detailed account of the size and location of irrigated areas. However, they do serve as useful reference maps on the approximate location of large scale irrigation systems.

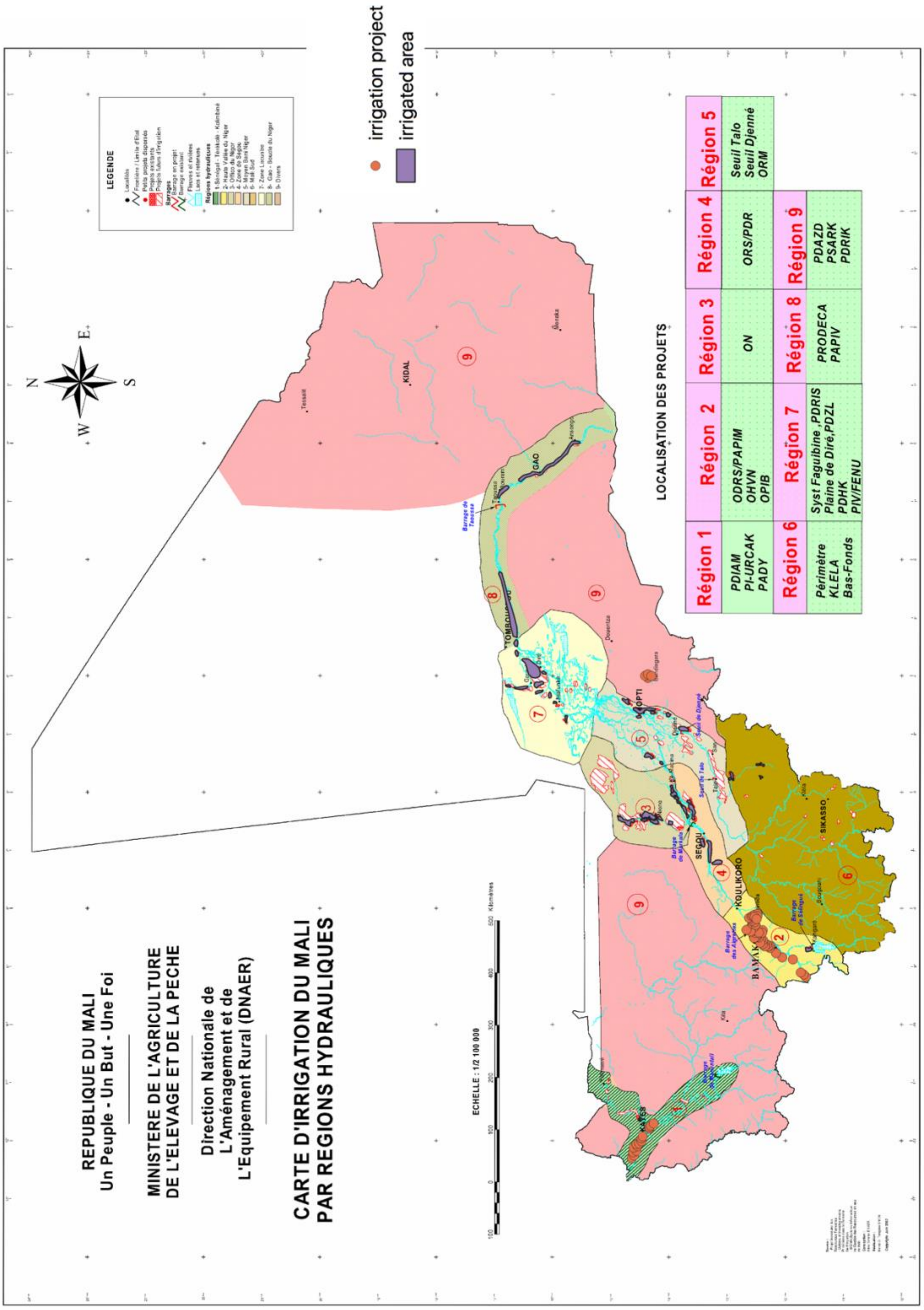


Figure 1. Map of irrigated areas in Mali. Source: Mali Ministry of Agriculture, FAO.

Above: Map of irrigated areas in Mali. The outlines of the irrigated areas of the country were digitized from an irrigation map present in the AQUASTAT library (DNAER 2003), with additional information on 6 large schemes taken from the FAO irrigation map for Africa

(FAO 2005). The shapes of the boundaries of the digitized irrigation areas were improved by using satellite imagery. The map combines the different layers in a single map.

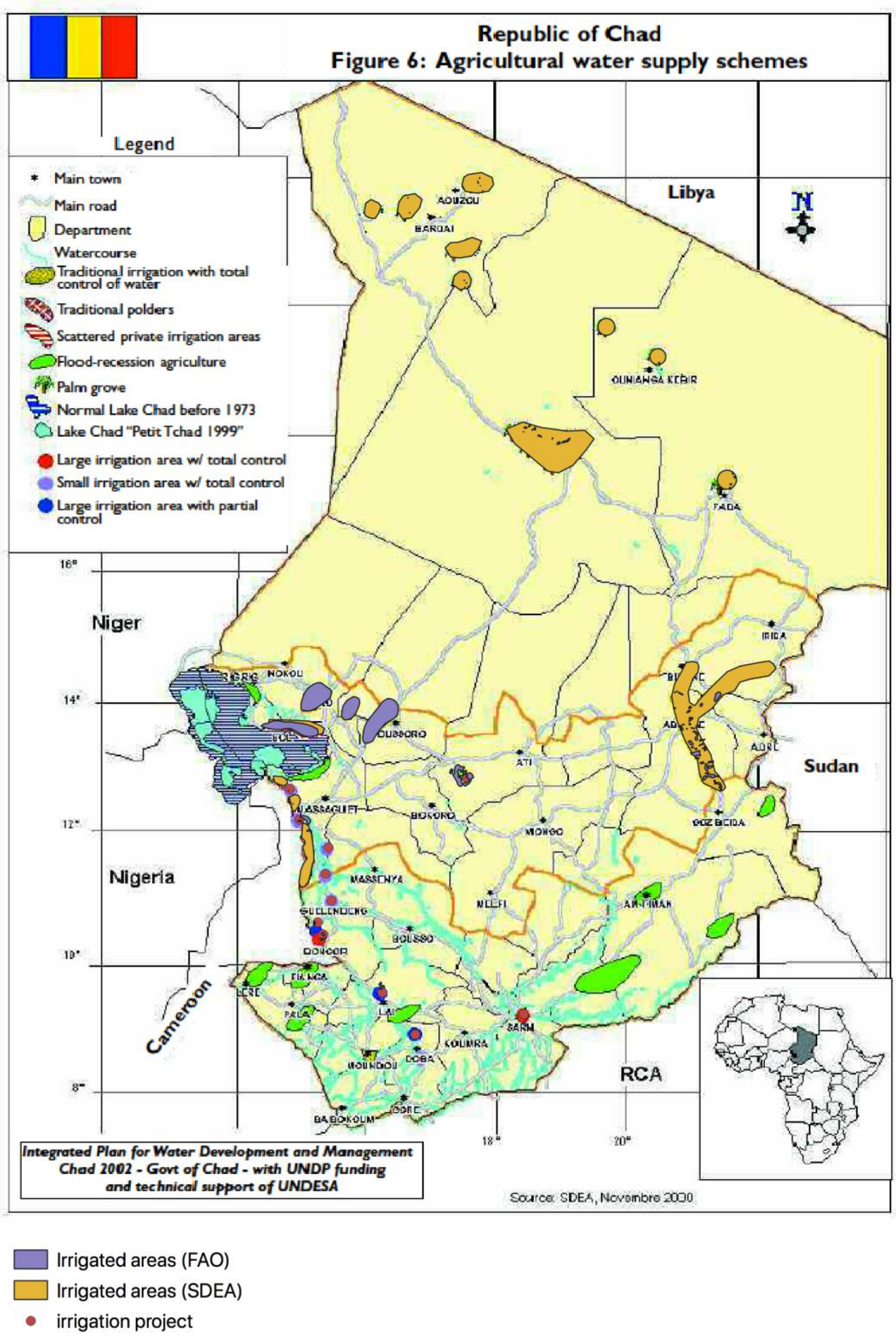


Figure 2. Map of irrigated areas in Chad. Source: SDEA, FAO.

Above: Map of irrigated areas in Chad. The position of the large schemes was taken from the FAO irrigation project database for Africa (FAO 2005). The remaining part of the irrigated area was assigned to zones of traditional or private irrigation and to palm groves as digitized from an irrigation map (UN_DSD 2003) The shapes of the boundaries of the digitized irrigation areas were improved by using satellite imagery. The map combines the different layers in a single map.

OTHER MAPS – AQUAMAPS, WAPOR, IWMI

A number of maps exist that show the extent of irrigated area in Sahel countries. Examples are AQUAMAPS (FAO, 10km resolution), and WAPOR (FAO, 100m resolution), and IWMI (10km resolution). Images of AQUAMAPS and WAPOR are shown below.

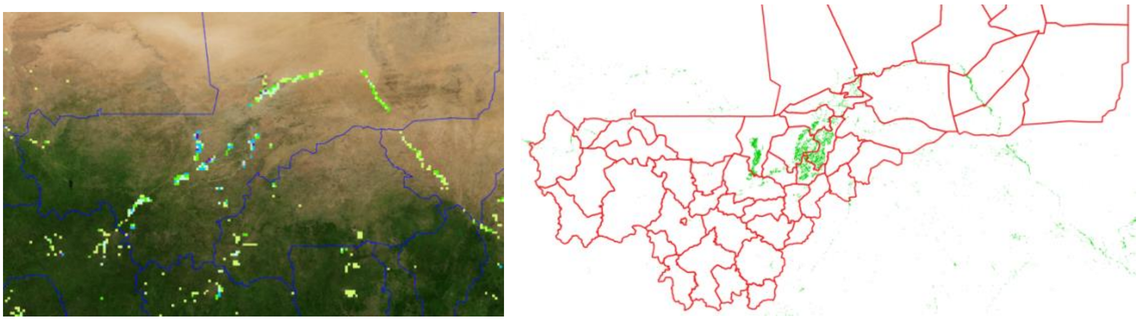


Figure 3. Left: AQUAMAPS, showing a part of Mali. Right: WAPOR 100m product

One of the issues of the current maps is that they have a low resolution (with 100m as the minimum), which causes them to miss smallholder irrigation. In addition, the WAPOR product shows some categorization confusion between irrigated areas and riverine vegetation, as shown in the image below. This is in fact a common problem and can never be completely avoided when machine learning is used for classification. Other studies, such as the AQUAMAPS, suffer from the same problem. IWMI is currently conducting a study (as yet unpublished) on irrigated areas in West Africa, in which they try to reduce the classification confusion using the correlation between local rainfall and plant growth, which shows promising results. In this study, we reduce categorization confusion by using time series of EVI and NDWI. In chapter 3, we quantify the magnitude of this classification confusion.

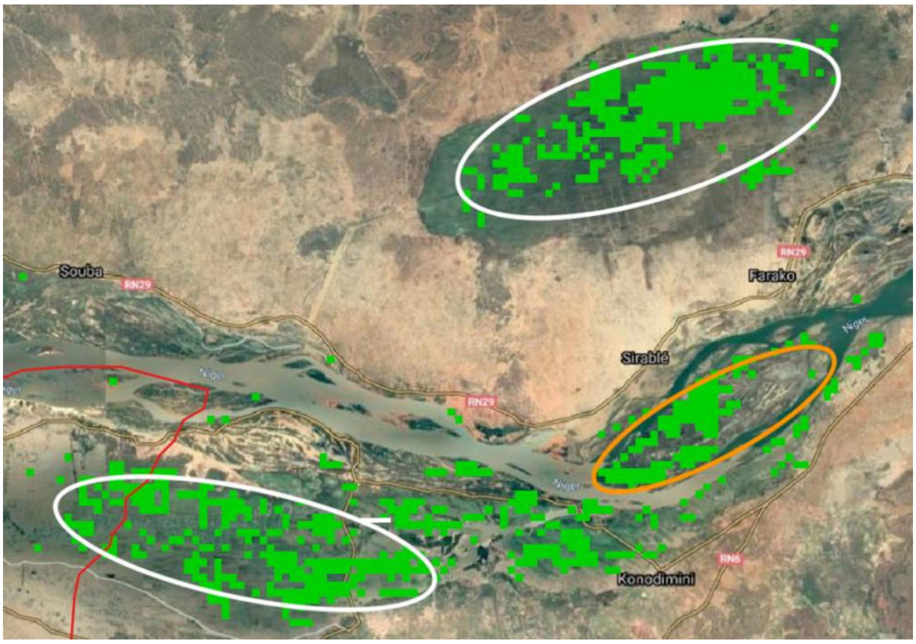


Figure 4. WAPOR detail of Mali, near Konodimini, showing the confusion of the machine learning algorithm between riverine vegetation and irrigated areas. White circles represent irrigated areas, while the orange circle represents riverine vegetation, which should not be classified as irrigated area.

The FAO also offers a number of 30m resolution layers but only in limited areas, and not yet with an irrigated area classification.

In the past years, a number of tools have come available that allow the analysis to take place at a higher resolution. Notably, the recent availability of high-resolution Sentinel-2 data, combined with the computational processing power of Google Earth Engine, have opened up a new world of possibilities. A recent PhD study (Vogels,2019) has taken the first steps towards a methodology that can identify and map small-scale irrigation for areas smaller than 1 hectare.

In the same study, steps were made towards using Geographical Object Based Image Analysis (GEOBIA), which can be used to distinguish between small-scale and large-scale irrigation. Although promising, we were unfortunately unable to pursue this approach due to constraints in available resources.

The method used in this report focusses on the identification of current irrigated areas during October – June, which corresponds to the end of the wet season up to the start of the next we season. The classification is done at 30 m resolution. The method combines an analysis of high-resolution remote sensing data and ground truthing data. Machine learning classification is used to distinguish between irrigated land and other land uses.

2.2 Machine Learning

In this part of the study, we use Machine Learning to identify irrigated areas. Machine learning refers to using a computer to automatically build a mathematical algorithm based on training data, in order to make predictions or classifications, without being explicitly programmed to do so. Machine learning is used in a variety of tasks, such as computer vision, email filtering, financial fraud detection, etc. Specifically, we employ Random Forest, a machine learning method to automatically create a classification algorithm from classified training data. The result is an algorithm that can classify data, i.e. identify to which of a set of categories a data point belongs, on the basis of a training set of data containing data points of which the classification is known.

The training of the algorithm relies on 'Ground Truth' data: data points on specific locations with a known classification (either irrigated or not irrigated in the period of interest). Although the gold standard of ground truth data is to visit locations individually, in many cases ground truth data can be collected by manually studying very high-resolution satellite imagery, which was done in this study.

In short, the process is:
1. Collect ground truth data with a known classification (irrigated or not irrigated)
2. Collect remote sensing satellite imagery bands
3. Train the machine learning model with part (70%) of the ground truth data, using the satellite data
4. Check the accuracy of the model with the remaining part (30%) of the ground truth data
5. Use the model to classify every image pixel in the target area (here, Mali or Chad)

The end result of this process is a map of the target area in which each pixel is given a classification.

CLASSIFICATION AIM

Our aim is to use machine learning and remote sensing satellite data to derive a land use map that distinguishes between two types: **irrigated**, and **not irrigated**.

As in the rainy season the area of study is almost completely covered with clouds, no satellite data is available during that period. Therefore, the analysis is necessarily restricted to the period October-June, which covers the end of the rainy season, the dry season, and up to the start of the next rainy season. The classification *irrigated* or *not irrigated* therefore only applies to irrigation within that period. Fortunately, horticultural production under FLID mostly happens during the dry season, so this is captured. However, irrigation that happens during or at the end of the rainy season, will be underestimated. Areas that are only irrigated in the wet season will not be detected by this method.

Detection of irrigation is here operationalised as *detection of increased plant growth with respect to local natural conditions* (local forest, shrubland, riverine vegetation, etc.), which is a good indicator of application of water. Recession agriculture — agriculture using residual moisture from receding floodwaters or seasonally flooded lands — is included in this definition.

STATISTICAL DEFINITIONS OF IRRIGATION

As we want to be able to compare our results to country statistics and other studies, we need to make the connection between definitions used in these studies and the machine learning classification in this report. For this, we follow the definition of FAO-AQUASTAT, for which we list the relevant terms with their definitions below:

Irrigation potential — Area of land which is potentially irrigable. Country/regional studies assess this value according to different methods. For example, some consider only land resources, others consider land resources plus water availability, others include economical aspects in their assessments or environmental aspects, etc.

Area equipped for irrigation — Area equipped to provide water (via irrigation) to crops. It includes areas equipped for full/partial control irrigation, equipped lowland areas, and areas equipped for spate irrigation.

Area equipped for irrigation: actually irrigated — Portion of the area equipped for irrigation that is actually irrigated, in a given year.

Flood recession cropping area non-equipped — Areas along rivers where cultivation occurs in the areas exposed as floods recedes and where nothing is undertaken to retain the receding water.

Cultivated wetlands and inland valley bottoms non-equipped — Wetland and inland valley bottoms that have not been equipped with water control structures but are used for cropping. They will have limited (mostly traditional) arrangements to regulate water and control drainage.

Total agricultural water managed area — Sum of total area equipped for irrigation and areas with other forms of agricultural water management (non-equipped flood recession cropping area and non-equipped cultivated wetlands and inland valley bottoms). It is the sum of the three categories above: Area equipped for irrigation + Flood recession cropping area non-equipped + Cultivated wetlands and inland valley bottoms non-equipped

As our classification method aims to identify all increased plant growth with respect to local conditions, we see that this corresponds most accurately tot the sum of **Area equipped for irrigation: actually irrigated, Flood recession cropping area non-equipped and Cultivated wetlands and inland valley bottoms non-equipped**. In section 6.3, we compare our results to country statistics and other studies, and AQUASTAT statistics.

TYPOLOGY OF IRRIGATION

In the Worldbank Sahel Irrigation Initiative Support project (SIIS), a typology of five different types of irrigation is used that are common in Sahelian countries. They are listed in the table below[1].

Table 1. Typology of irrigated areas.

Type		Description
Small scale	1	Improved rainwater harvesting with partial water control: inland valley bottom development (*bas-fonds*), flood recession plains or partial control (sometimes 1000s of ha), sand dams for groundwater recharge (*seuils d'épandage*). Crops are rice, sorghum and vegetables.
	2	Small-scale private irrigation systems (less than 1 ha up to a few hectares) for individuals or small groups of producers, involving pumping equipment, devoted to high value crops such as vegetables.

[1] Project information document Sahel Irrigation Initiative Support Project (report PIDA99932).

	3	Small-scale community-based irrigation schemes of less than 50 ha, usually promoted by NGOs or governments, for villages or large groups of producers who collectively manage pumping equipment and canals to produce rice or vegetables.
Large scale	4	Large-scale irrigation schemes (from 100 ha to 5000 ha with a vast majority below 1000 ha) publicly financed, managed or supervised by public authorities, located usually along large rivers regulated by dams, comprising a combination of pump stations and a network of canal and drainage systems, service roads. They require a complex management structure.
	5	Medium- to large-scale irrigation schemes involving a partnership between the Government, a private party, and the communities surrounding the scheme, for the development and management of the irrigation system (with same technical features as for Type 4).

To characterize the irrigated areas as classified in this report, we make the distinction of areas smaller than 100 ha, those between 100-500 ha, and those larger than 500 ha.

Farmer-led irrigation, as understood in this report, is captured by types 1, 2 and 3. Whereas types 2 and 3 can be identified by considering their size, type 1 (valley bottoms, flood recession) presents the problem that it consists of small cultivated areas per farmer, but in a large overall area. This complicates the distinction of FLID by plot size.

2.3. Choice of methodology

Broadly speaking, there are two ways to analyse satellite data: pixel-based analysis and Geographical Object Based Image Analysis (GEOBIA). We briefly explain both approaches below before motivating our choice for the first type.

Pixel-base analysis performs a machine-learning classification on individual satellite image pixels, using raw bands (such as red, green, blue, near infrared), other bands such as slope and rainfall, and computed bands such as Enhanced Vegetation Index. Studies that use this approach mainly differ in the exact number and choice of bands that was used, and in the resolution of the satellites used (Modis at 250m resolution vs Sentinel-2 at 30m resolution for example). A notable example is the work of the International Water Management Institute (IWMI), which has performed numerous studies on the extent of irrigated areas in Asia and Africa[2,3].

The second method is more recent and makes use of Geographical Object Based Image Analysis (GEOBIA). This methodology relies on the segmentation of a satellite image into 'objects': small contiguous areas of land that are similar, such as individual fields or buildings. Machine learning is then performed on those objects. Two notable examples of the latter are the Global Food Security-Support Analysis Data at 30 m (GFSAD30) project (Xiong, 2017) that created a map for cropland extent for all of Africa, and a recent PhD study by M. Vogels, which mapped irrigated areas in Ethiopia (Vogels, 2019).

---

[2] Web map of irrigated areas in Africa and Asia by IWMI, unfortunately unavailable at time of writing of this report.
[3] *Global Irrigated Area Mapping*, IWMI, 2000.

A number of considerations have led us to choose the first approach – pixel-based analysis. The first consideration is that the GEOBIA approach is still experimental, and results have not been adequately validated yet. It is therefore not quite clear if it forms an actual improvement over the pixel-based approach. In addition, the GEOBIA approach relies on an intensive processing step in which shapes are identified – a step that is performed by high-cost proprietary software, that was unavailable to us. Finally, both the author of the recent PhD study, as well as her University group, were unable to assist us.

When a pixel-based Machine Learning algorithm is used to identify irrigated and non-irrigated areas, it is not possible to make the distinction between large scale and small-scale irrigation because the pixel-based method has no concept of shape. Therefore, automated recognition of large-scale versus small-scale irrigation is not possible in this methodology.

To get around this problem, one option is to manually compose a GIS layer with the locations and extent of the large-scale irrigation schemes, flood recession cropping areas, and cultivated wetlands and inland valley bottoms. However, we have not been able within the timeframe of this study to locate or produce such a map.

As an alternative solution, we look at the size of contiguous areas of pixels that have been classified as irrigated, and use the size of these patches as the 'size' of the irrigated area. This is not a perfect solution, as the classified irrigated areas can be highly fragmented. However, as the results will indicate, this method was successful in identifying the large irrigation schemes.

2.4. Characterization of the Rainfall and Growing Season

To identify irrigated areas, we make use of the known patterns of rainfall and the growing season. The figures below show the monthly rainfall for Mali and Chad, which are very similar.

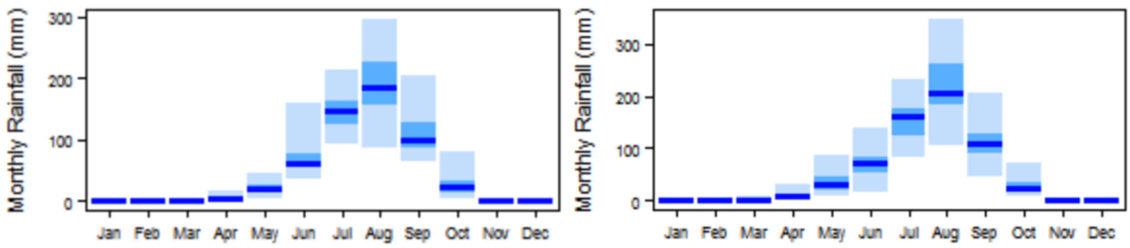


Figure 5. Left: Monthly rainfall in Mali, right: monthly rainfall in Chad. Source: Earthwise, British Geological Survey.[4]

Information on the main growing seasons, which is shown below, shows significant differences. In the case of Mali, crops are grown in different growing seasons, both in the wet season and the dry season. In Chad on the other hand, the main growing seasons are concentrated in the wet season. As this study relies on data in the dry season mainly, this might affect the completeness of the result, as areas that are only irrigated in the wet season cannot be detected through our method.

[4] earthwise.bgs.ac.uk/index.php/Climate_of_Mali and earthwise.bgs.ac.uk/index.php/Climate_of_Chad

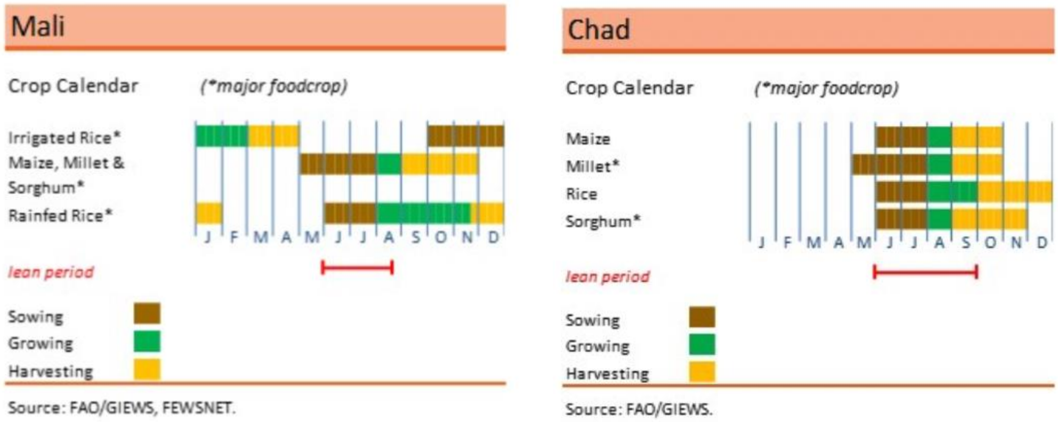


Figure 6. Left: growing season of main crops in Mali. Right: Chad. Source: FAO GIEWS country briefs

It is important to note that the tables above don't capture the full range of agricultural production. Small-scale horticulture, which is a large component of farmer-led irrigation, is usually practiced in the dry season in both countries.

The natural vegetation follows two main patterns: Away from water bodies, the vegetation follows the rain pattern. Near water bodies or rivers, the vegetation follows the availability of water in the water body or river. Irrigation takes place at multiple times: both to prolong the growing season at the end of the rainy season, but also in the dryer months. This difference in irrigation periods can be an issue during classification, especially in the case where after-rainy season irrigation is compared to riverine vegetation that is also abundant in the same period.

2.5. Data collection

As our period of analysis, we choose October 2019 until June 2020, thus capturing the end of the rainy season, the dry season and the start of the next rainy season. We use Sentinel-2 Top-Of-Atmosphere reflectance imagery, at 10 m resolution. Sentinel-2 represents the best current option for remote sensing, because of its superior resolution, revisit time (days between images of a given location), and number of bands. As the 10m resolution data was found to have too much spectral variability, the data was aggregated to 30m resolution by local averaging.

As our data processing platform, we chose Google Earth Engine, a cloud-based platform for planetary-scale environmental data analysis. It combines the access to peta-byte scale geospatial data sources with the enormous data-crunching capabilities of Google. It also offers easy access to many data layers, including the Sentinel-2 satellite data.

Several intermediary products are derived from the Sentinel data. First of all, a dry season (December - March) mosaic is created from the lowest-cloud percentage images for the blue, green, red, NIR and SWIR bands.

Secondly, **monthly** mosaic images are created for two indices: The Enhanced Vegetation Index (EVI)[6], and the Normalized Difference Water Index (NDWI)[7]. The EVI index is directly sensitive to vegetation – it measures the 'greenness' of the plants. The NDWI is more sensitive

[6] en.wikipedia.org/wiki/Enhanced_vegetation_index
[7] en.wikipedia.org/wiki/Normalized_difference_water_index

to the plant water content and is therefore strongly correlated to water stress. In addition, we include the standard deviation of the EVI index over the full year.

One issue with creating EVI and NDWI indices is that due to extensive cloud cover during the rainy season (July-September), in some cases no imagery is available during an entire month. As the classification is mostly sensitive to the EVI and NDWI outside of the rainy season, we removed the rainy season periods from the analysis. Experiments using alternative satellite sources that are not sensitive to cloud cover (Sentinel-1 Synthetic Aperture Radar) showed that this did not lead to a sufficient classification accuracy. The synthetic aperture radar data by itself does not contain enough information to distinguish between irrigated and non-irrigated crops. As the Sentinel-2 data is needed for an accurate classification, the period of analysis needed to be restricted to relatively cloud-free periods. In these periods, it was found that including the Sentinel-1 data did not increase the classification accuracy, and therefore it was omitted.

Thirdly, we are not only interested in the absolute monthly values for these indices, but also in the changes from month to month, as this should be closely related to plant growth. When the EVI value increases in three consecutive months in the dry season, this is a strong indication that irrigation is applied. Therefore, we include pair-wise differences between consecutive monthly values. For example, if $NDWI_1$ and $NDWI_2$ represent the NDWI values in January and February, we also include $NDWI_1$ - $NDWI_2$ as a variable.

In addition to these products, a number of spatial pattern bands are added, which capture the spatial structure and texture in the immediate vicinity of pixels. These are computed using the Gray Level Co-Occurrence Matrix (GLCM) method. The reason for including these layers is to make use of the fact that agricultural areas often have a different texture than natural areas, such as riverine vegetation.

In total, 38 bands were used for the classification, as listed in the table below. The period column indicates the period over which the data was aggregated. These bands were chosen from a much longer list that included, in addition to the bands mentioned above, all 13 Sentinel-2 bands, all 16 GLCM bands, 7 bands on rainfall and temperature, and 3 Sentinel-1 synthetic aperture bands. Using a Random Forest sensitivity analysis, those bands were chosen that have the most predictive power for the classification of irrigated areas.

Table 2. Satellite bands used for machine learning.

Variable	Number	Source	Period
Red, Green Blue, NIR, SWIR	5	Raw data from Sentinel-2 (B4, B3, B2, B8, B11)	Median of Dec 2019 – Mar 2020
EVI_StdDev	1	Computed EVI index	Standard deviation Jun 2019 – Jun 2020
EVI_0 … EVI_7	8	Computed EVI index	Median of monthly data, Oct 2019 – May 2019
EVI_diff_1 … EVI_diff_7	7	Differences of computed EVI index	Differences of consecutive monthly EVI index, Nov 2019 – May 2019
NDWI_0 … NDWI_7	8	Computed NDWI index	Median of monthly data, Oct 2019 – May 2019

NDWI_diff_1 … NDWI_diff_7	7	Differences of computed NDWI index	Differences of consecutive monthly NDWI index, Nov 2019 – May 2019
SWIR_var	1	GLCM texture variance of SWIR band	Median of Dec 2019 – Mar 2020
NDWI_1_var		GLCM texture variance of NDWI band	Median of Nov 2019
Total bands	**38**		

In an earlier stage of this study, we also included Sentinel-1 radar imagery. However, we found that these bands didn't lead to a higher accuracy in the classification, and therefore they were removed again from consideration.

2.6. Pre-processing

The sentinel-2 satellite has a spatial resolution of 10 meters. The high resolution is useful, but also leads to some spectral variability (noise). To reduce this noise, we smooth the satellite image by local averaging, leading to a final resolution of 30m.

2.7. Collecting ground-truth data - Identifying irrigated areas

For this study, we relied on identifying irrigated areas remotely. This is possible, as irrigation usually is readily apparent from the EVI and NDWI curves at a specific location. This is illustrated in the charts shown below.

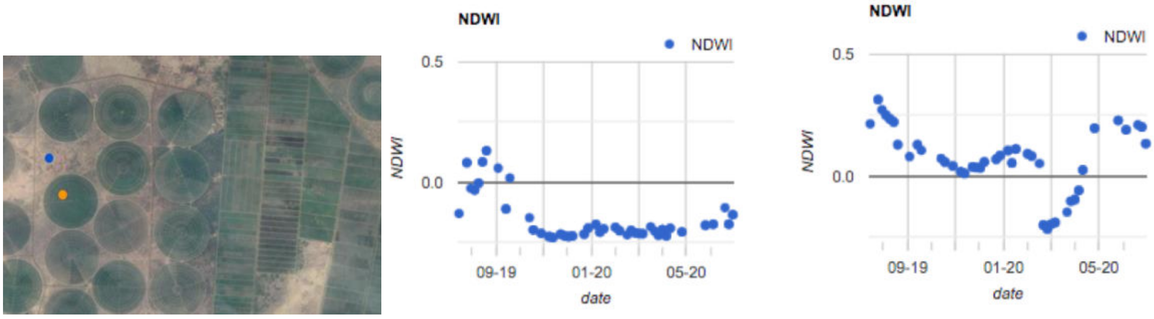


Figure 7. Left:image of Office du Niger, showing circular irrigated areas. The blue dot is located outside the circles, the orange dot is located inside a circle. Middle graph: NDWI plot corresponding to the blue dot, outside the irrigated area. The negative level for most of the year indicates limited plant growth and water stress. Graph on the right: NDWI plot corresponding to the orange dot, inside an irrigated area. During almost the whole year, there is no water stress. The sharp break that is visible at the start of March represents the harvest.

To facilitate the indication of irrigated areas, we created a small application in Google Earth Engine that displays the full EVI and NDWI curves over a given period (so no monthly averaging) at a certain location. In this way, manual inspection of both high-resolution imagery can be combined with inspecting the EVI and NDWI curves at a single location. This data has proven to be essential for a correct classification of the ground truth data. A screenshot of this application is shown below, with the EVI and NDWI curves shown on the left. Additional examples are shown in below.

Figure 8. Google Earth Engine application to assist identification of irrigated areas. Left (under the 'two chart inspector'): the EVI and NDWI charts over the period 2019-2020. Further examples are provided below.

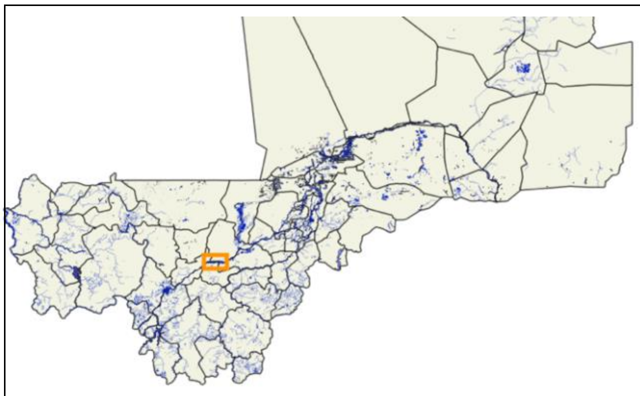
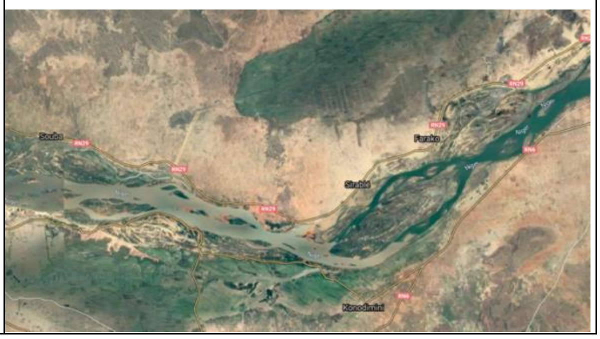
As an illustration, below we provide some examples of this data for different regions of Mali and Chad. All the satellite images were taken in February.

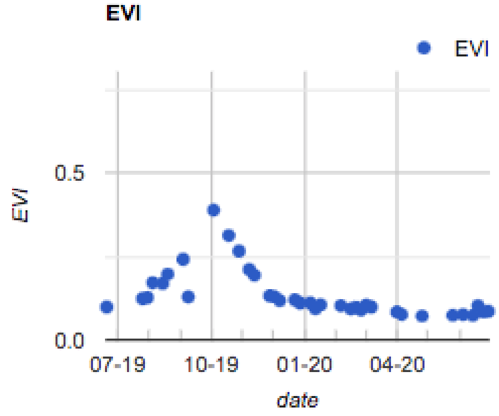
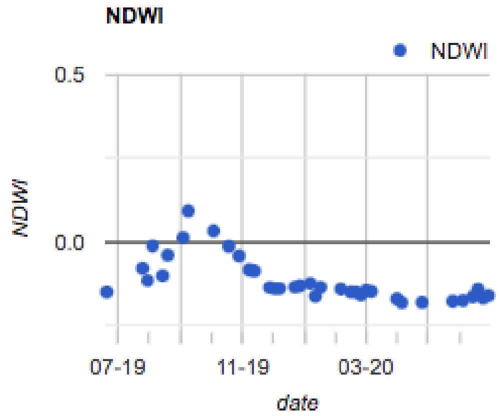
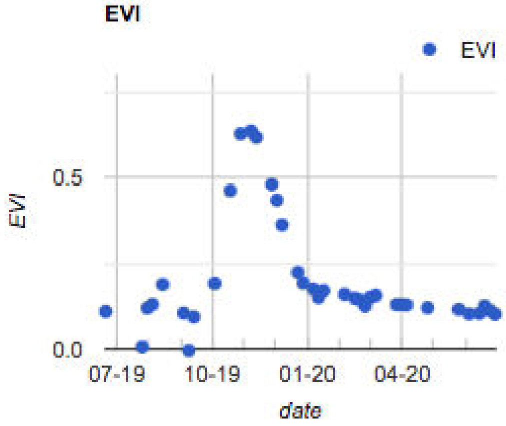
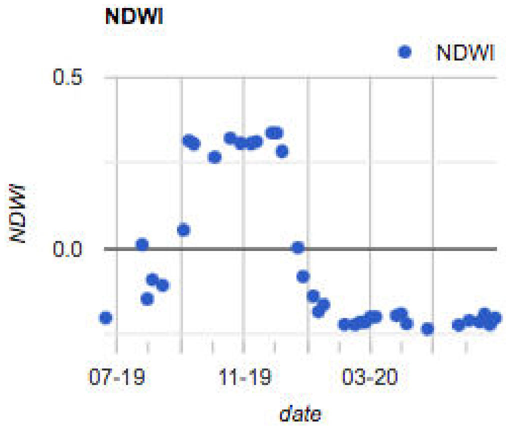
It is clear that the periods during which irrigation is applied differ substantially from one location to another. For example, the Konodimini data shows a growth pattern with a single peak, starting in September and ending in January. The Bamako region has two pronounced and shorter growing seasons, one stretching from July to October, and one from December to March. The Office du Niger data, which was taken from a circular irrigation scheme, show a prolonged growing season that stretches from May until February.

Mali - Konodimini

In this area, the EVI curve indicates that crops are grown in the November – January period. This is the same area as the WAPOR image above. Note the large difference in the NDWI curves.

Table 3. EVI curve examples near Nonodimini, Mali.

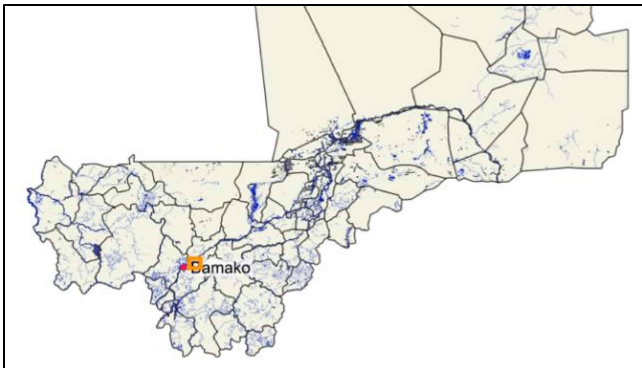
	Natural colours: 

Not irrigated : in brown area in the middle  	**Irrigated** : In green area in lower left corner:  

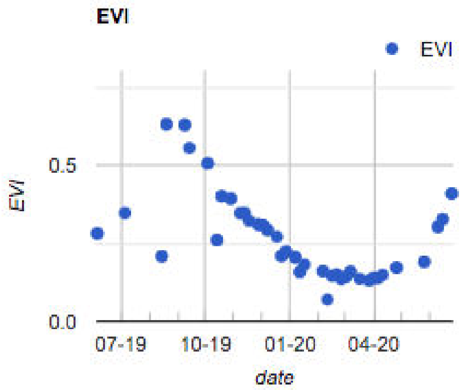
Mali - Bamako

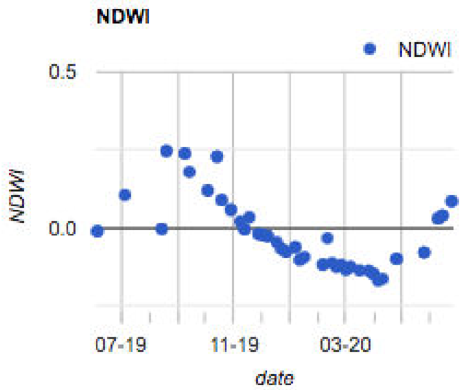
In this area near the Niger in the vicinity of Bamako, the EVI curve indicates irrigation in the dry season, starting in December.

Table 4. EVI curve examples near Bamako, Mali.

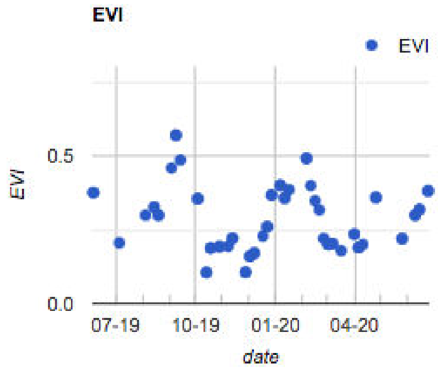
	Natural colors 

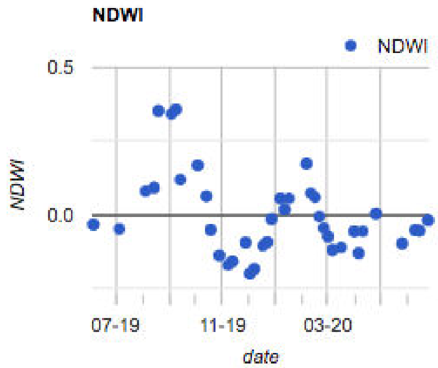
Not irrigated
Just outside the irrigated area:





Irrigated
Inside one of the dark green rectangles:

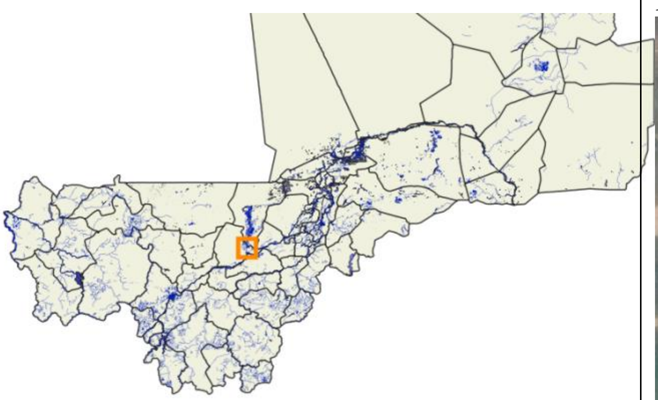




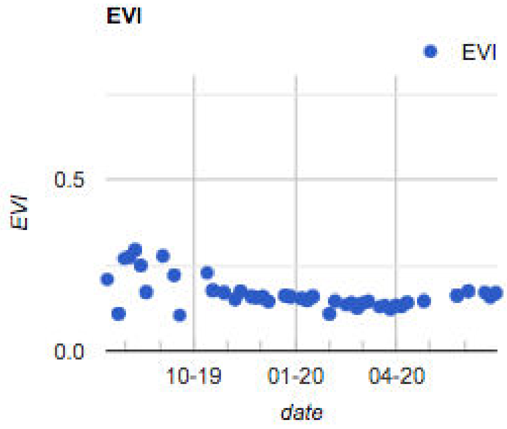
Mali - Office du Niger

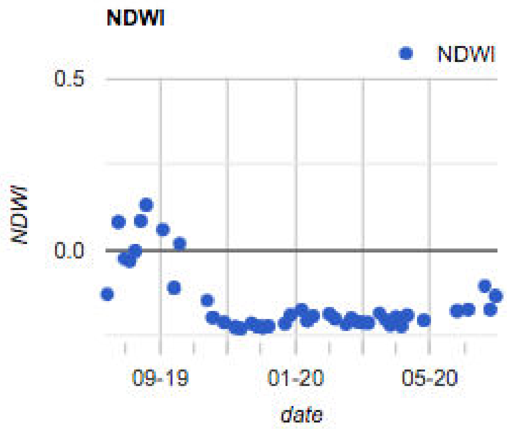
In the irrigation circles in the Office du Niger, we see that irrigation is applied all through the dry season.

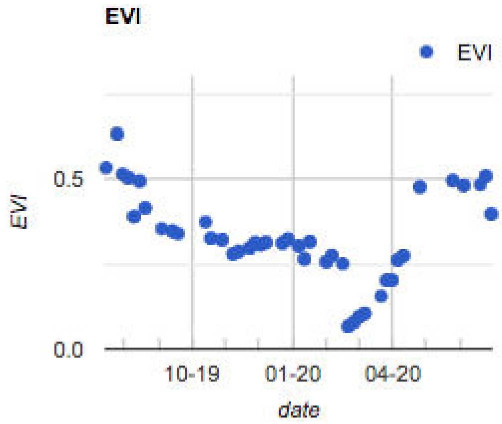
Table 5. EVI curve examples near Office du Niger, Mali.

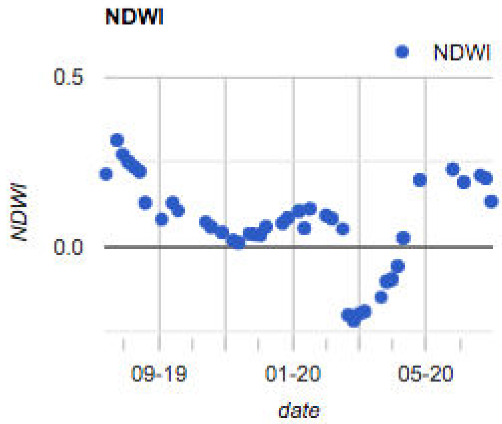


Natural colors

<table>
<tr>
<td>Not irrigated : Outside irrigation circle</td>
<td>Irrigated : Inside irrigation circle</td>
</tr>
<tr>
<td>




</td>
<td>


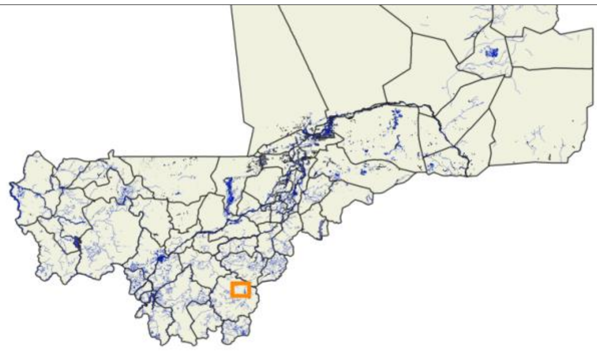


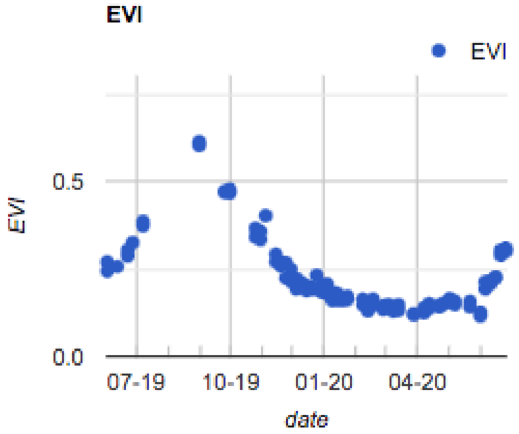
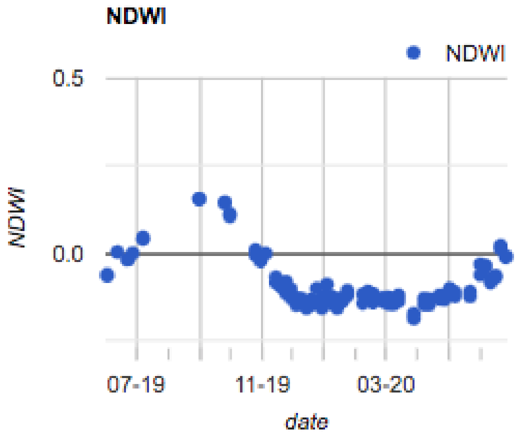
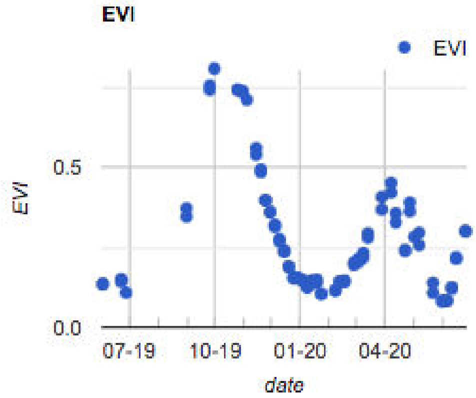
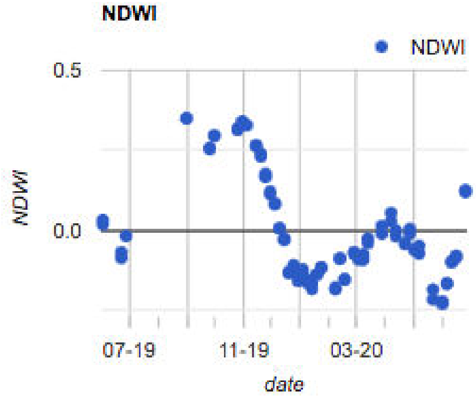
</td>
</tr>
</table>

Mali - Sikasso

In this region, river water is used for irrigation. We see both long and short cropping periods.

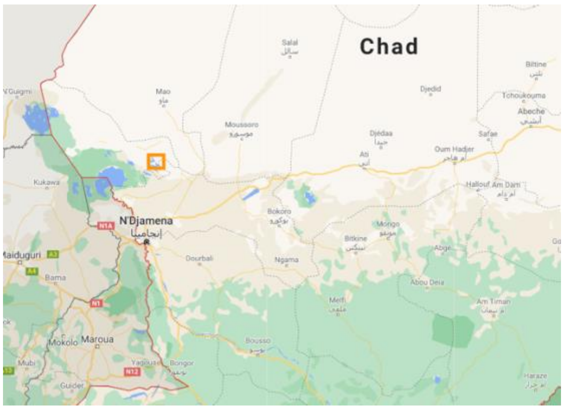
Table 6. EVI curve examples near Sikasso, Mali.

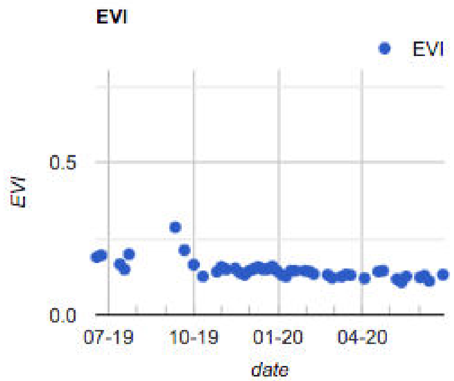
<table>
<tr>
<td></td>
<td>Natural colors
</td>
</tr>
</table>

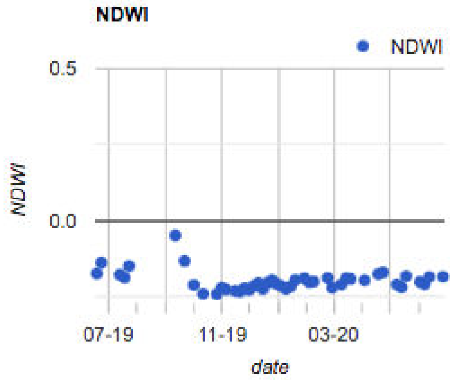
Not irrigated : Outside of irrigation area  	**Irrigated:** In green area  

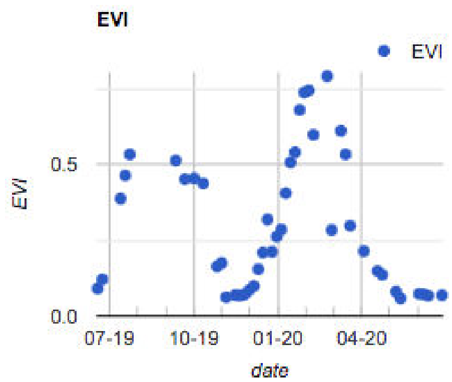
Chad – Lake Chad

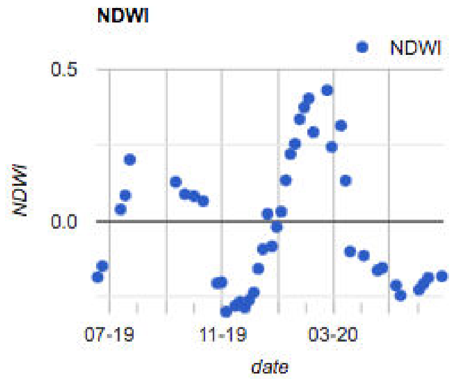
Table 7. EVI curve examples near Lake Chad, Chad.

	Natural colors 
Not irrigated : Outside of irrigation area	**Irrigated:** In green area



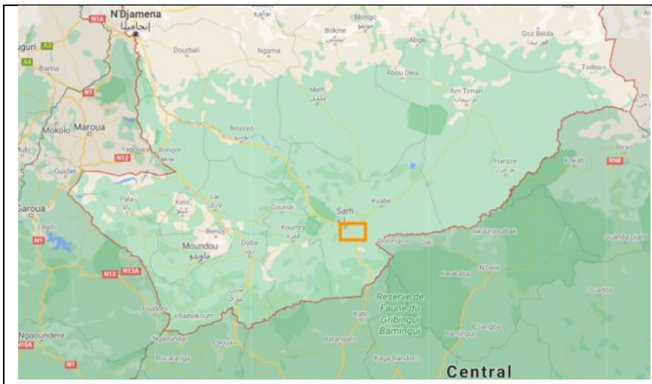






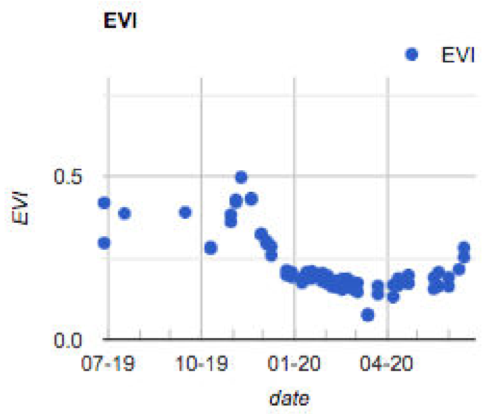
Chad – Sarh

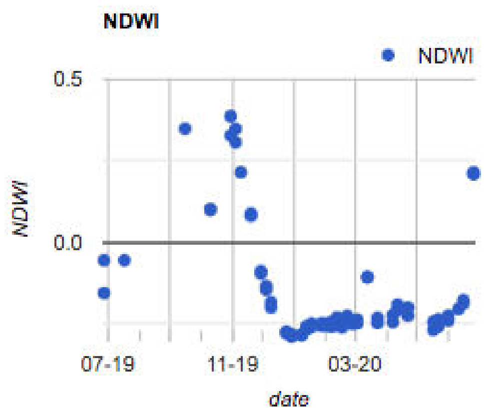
Table 8. EVI curve examples near Sarh, Chad

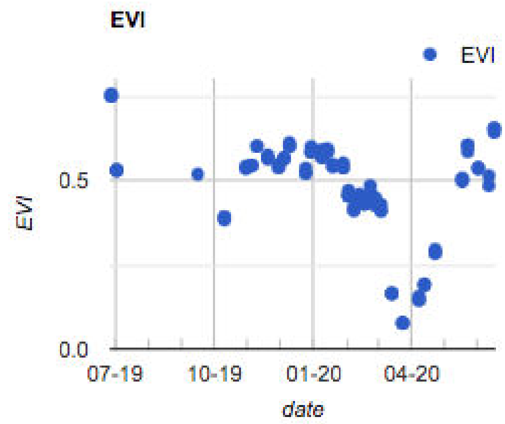


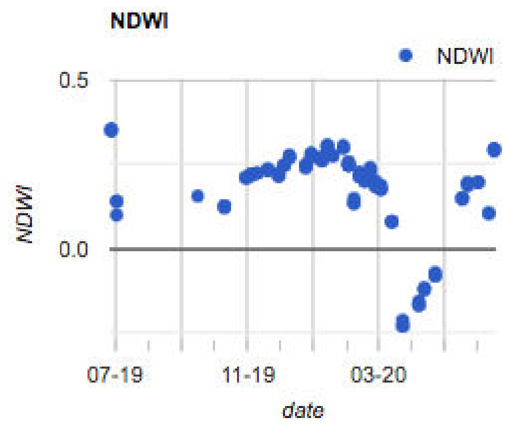
Natural colors

Not irrigated : Outside of irrigation area	**Irrigated:** In green area





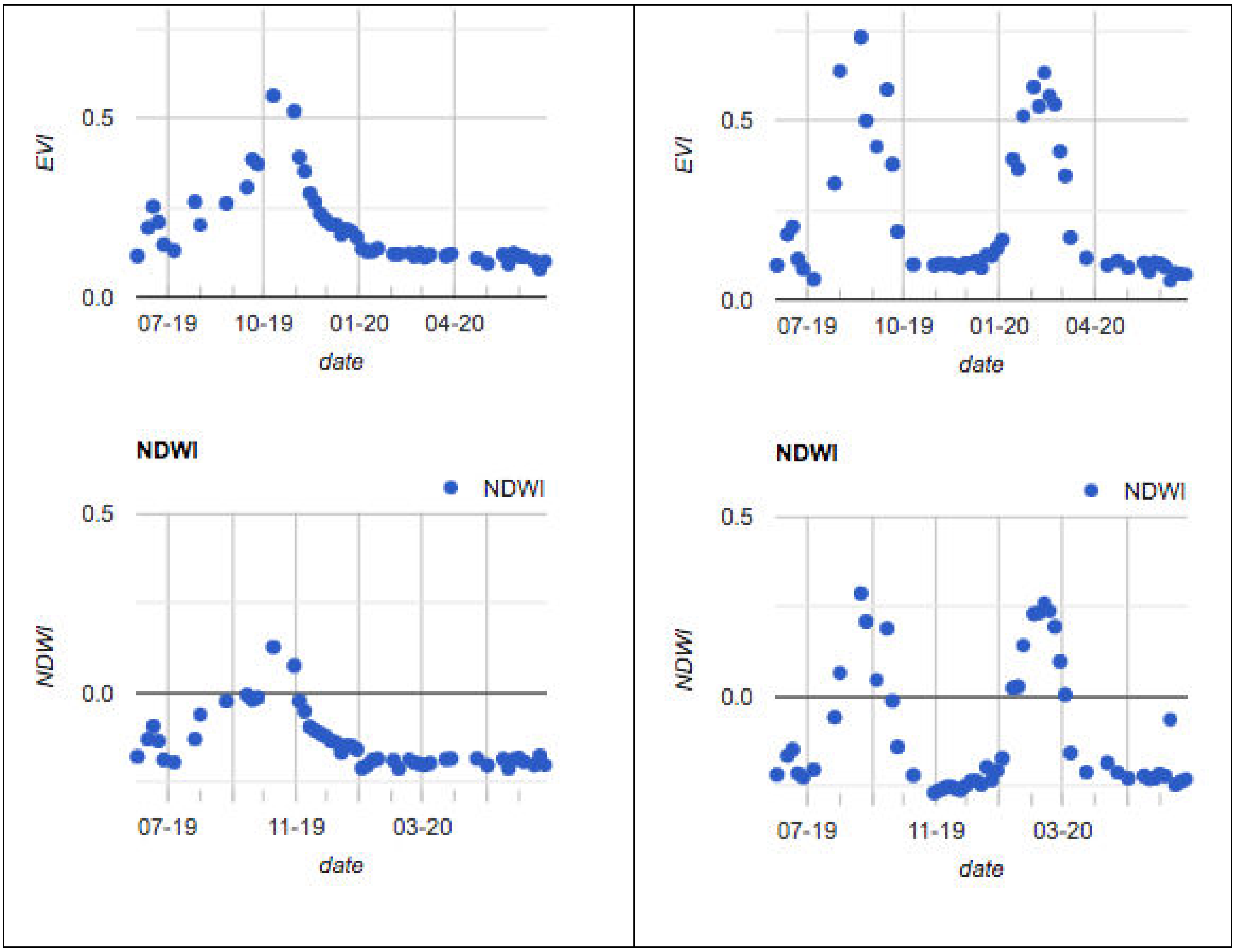




Chad – Diour

Table 9. EVI curve examples near Diour, Chad

	Natural colors 
Not irrigated : Outside of irrigation area	**Irrigated:** In green area



These examples just serve to indicate the typical quality of the EVI and NDWI data available, as well as to show the variability of data in different areas in the country.

2.8. Ground truth data

As detailed in section 2.7, ground truth data was collected using manual inspection of high-resolution Google imagery of a given location, combined with inspecting the timeseries of the EVI and NDWI indices of a given location. Together, this data provides enough information to correctly classify the location as either irrigated or not-irrigated.

During the identification process, both the high-resolution (5m) Google Earth background map is used, as well as EVI layers for the months of December – April. In addition, EVI and NDVI curves can be inspected for a specific point, which usually allows a clear identification of whether agricultural land has been irrigated or not, as shown in the previous section. In other cases, the nature of the land use is clear from other visual aspects, such as small islands in rivers or inside forest areas. This leads to a high degree of confidence of the manual classification of the points, even though only remote data was used.

The points that are most difficult to classify in this way are those that lie inside flood recession areas or valley bottoms. In these locations, natural growth has a very similar satellite spectral signature when compared to irrigated plots. This is one of the main limitations of the present method, and can only be solved by in-country field visits for ground truthing.

In total, 11.208 points were identified in Mali, of which 3569 were areas that were identified to have been irrigated during the Oct-Jun period, and 7639 were identified as areas that were not irrigated in the same period, as was determined with the method described above. In the case of Chad, 6354 points were identified, of which 2009 were irrigated areas, and 4345 were non-irrigated areas.

In the choice of training points, a number of considerations were used. First of all, points were chosen to cover all the main land use and land cover classes, such as forest, shrubland, cities, villages, rivers and riverine vegetation, desert, etc. Secondly, points were chosen to cover the different methods of irrigation, such as flood and recession irrigation, small plots, large circular schemes, etc. Finally, care was taken to cover the different climate zones. The typical spread of training points is shown in the image below. Here, we look at a large irrigation scheme in Mali, the Office du Niger.

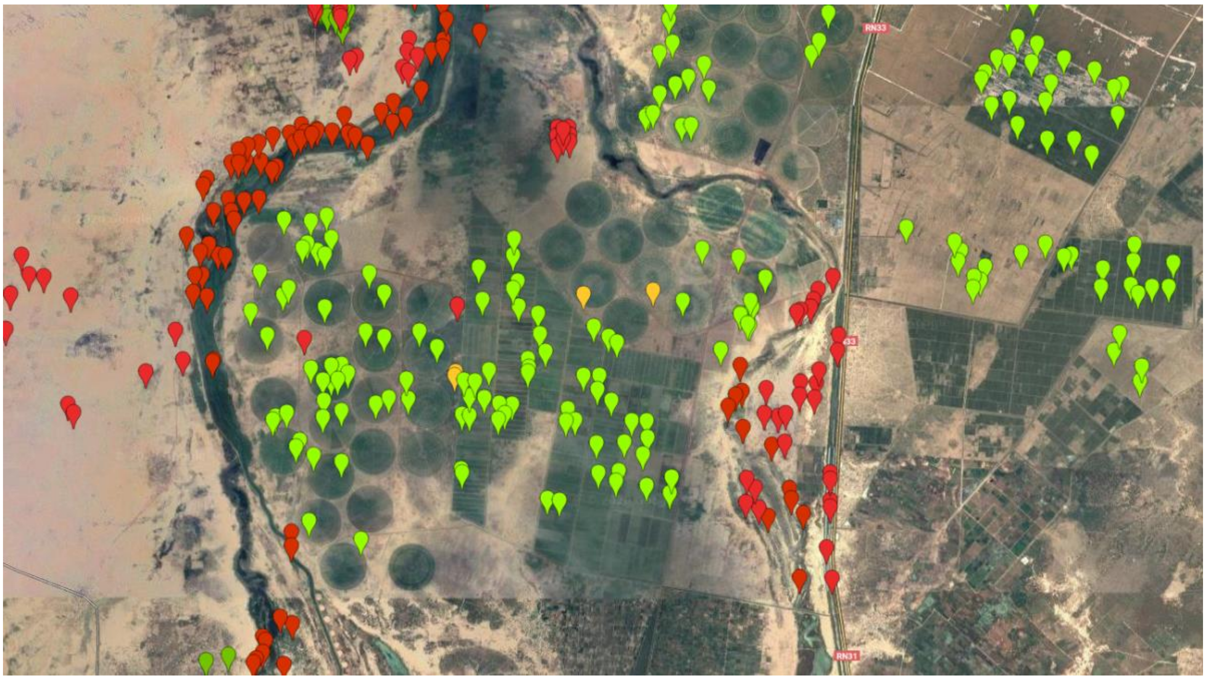


Figure 9. Typical selection of ground truth points, showing a large irrigation scheme in Mali, the Office du Niger. Red points are identified as not-irrigated, green points are identified as irrigated.

2.9. Machine learning model

To classify the image, we used a Random Forest model consisting of 75 trees. As described above, Random Forest is a specific machine learning algorithm, that is well suited for classification of data points in categories and is the most used algorithm for this type of task. The principle behind a Random Forest model is that it creates a large set of decision trees using different (random) subsets of the data bands and training points and uses these individual trees to form a 'majority opinion'. A decision tree is a simple predictive modelling algorithm that uses observations about an item to decide to which category it belongs using a series of 'yes-no' decisions.

Due to this process of combining different trees, quirks of the individual decision trees even out, and the end result is an algorithm that performs better than any individual decision tree. A predictive modelling algorithm that uses observations about an item to decide to which category it belongs.

Among the positive properties of a Random Forest model is the fact that it is not necessary to scale variables (bands), and that multiple variables that have a strong mutual correlation do not present a problem. The Random Forest model is trained and used in Google Earth Engine.

Different models were trained for Mali and Chad, using the ground truth data points for each separately. With the separation in climate zones described below, a total of 6 models were trained.

To perform the machine learning process, the ground truthing data is split into two parts: training and validation data: 70% was used to train the machine learning model, and 30% was used for validation. These percentages are commonly used in machine learning. It balances the need for enough data to train the model with a low degree of variance in the measured accuracy. To make sure this split does not influence the results, we also trained the model with a 80:20 and 90:10 split, leading to similar results for the accuracy. In the results below, we only report on the 70:30 split.

Climate variability

One important issue is the variability of the climate across both Mali and Chad: from warm desert climate in the north, to tropical savanna climate in the south. The local climate has, naturally, a large impact on the behaviour of the observed bands such as EVI and NDWI. In addition, the local climate has an impact that gradually varies with the latitude. This type of spatially varying variables is not handled well by the Random Forest methodology. This was confirmed during an earlier stage of this work, where an attempt was made to address this by introducing climate-related variables, such as local rainfall, and local average temperature. However, this did not lead to a satisfactory result.

The solution we chose for this report is to reduce the variation in the training set by separating the country into three climate zones: upper, middle, and lower. In this way, the inter-set variability is reduced. This is also the approach taken by other efforts, such as the Global Food Security Analysis-Support data at 30 meters (GFSAD30) project[8]. The climate zones used are shown below.

[8] croplands.org

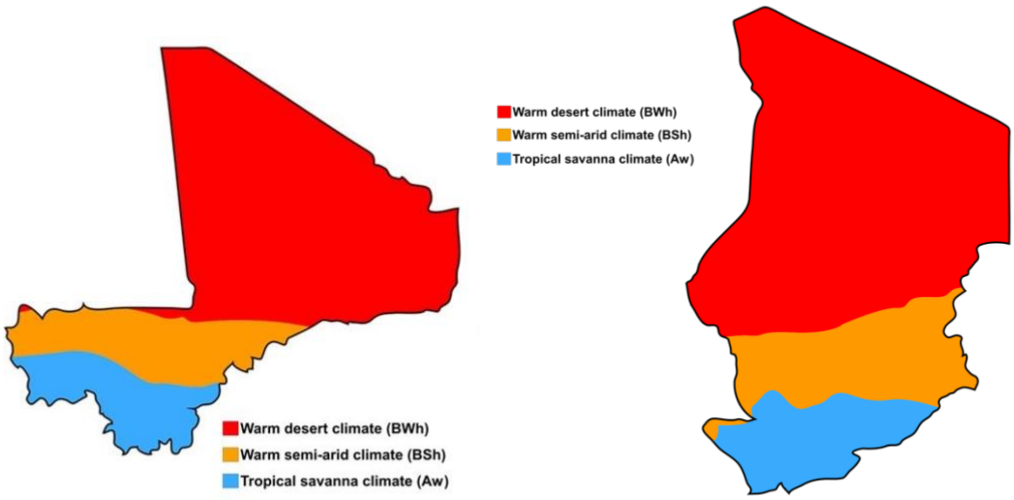


Figure 10. Left: climate zones of Mali. Right: climate zones of Chad. Source: Köppen climate classification[9]. To optimize the classification result, we train a model for each of three climate zones.

2.10. Computation of results

The training of the Random Forest model on the ground truth data is a very quick process that takes only a few seconds. The next phase is to classify every pixel of the target area, which in this case spans the entirety of Mali or Chad. As we perform our classification on a 30m resolution, this is a large computational task that is very memory intensive. We use Google Earth Engine to perform this classification, and this platform has limitations on the maximum memory that can be used for single tasks. Therefore, to do the complete classification, we need to divide the task into smaller sub-tasks, each covering a small part of the country. To do this, the Mali and Chad regions were divided into 1° x 1° squares, each corresponding to an area of about 12,200 square kilometres.

For each square, it was determined in which climate zone the majority of its surface lies, and the appropriate model is trained. Although one might expect this sharp boundary to lead to visible edges in the classification, this was not observed in practice. All squares are classified in this manner, at 30m resolution. After all squares are completed, they are reassembled to form the final result.

[9] en.wikipedia.org/wiki/K%C3%B6ppen_climate_classification

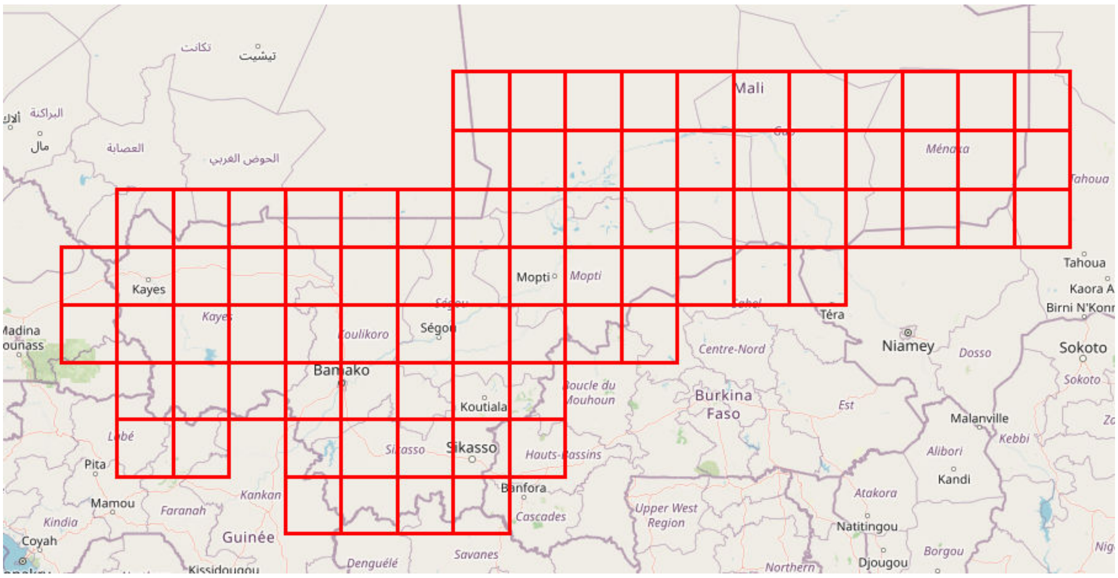


Figure 11. Grid of 1x1 degree squares used for the classification process of Mali.

2.11. Example of results

The quantitative results of the classification are presented in chapter 3 of this report. Below, we show a qualitative result of a small section near the Niger river.

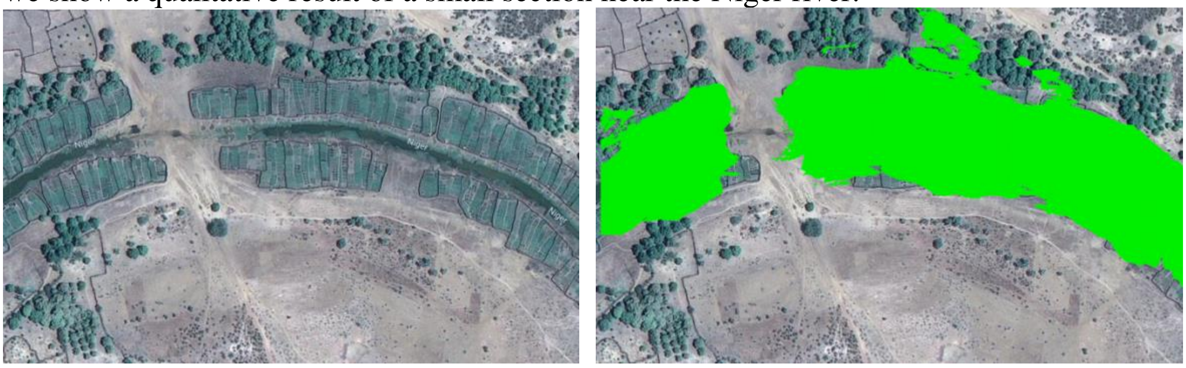


Figure 12. Left: part along the Niger river, showing irrigated fields of approximately 20 x 20 meters. Right: Same region, with classification result.

2.12. Postprocessing

As a final step, we post-process the classification result to remove noise. We do this by applying a morphological opening and closing operation, both at a kernel circle of 30 meters. This has the effect to remove individual isolated pixels, as well as close pixel-sized holes. The justification of this operation is twofold. First of all, the removed areas correspond to areas with a size of 30x30 meters, or just 0.09 hectares. This is considerably smaller than average smallholder field. Secondly, irrigated areas, also when they consist of small fields, are highly clustered. In practice, this operation removes noise in forest areas, where often individual pixels in a forest area are being wrongly classified as being irrigated.

The effect of this operation is shown qualitatively below. Importantly, we don't use any rule-based post processing, such as removing areas far away from surface water — the post processing we use is merely focussed on removing single pixels and closing small holes. Applying this processing step was checked to have an effect on the end result below 1%.

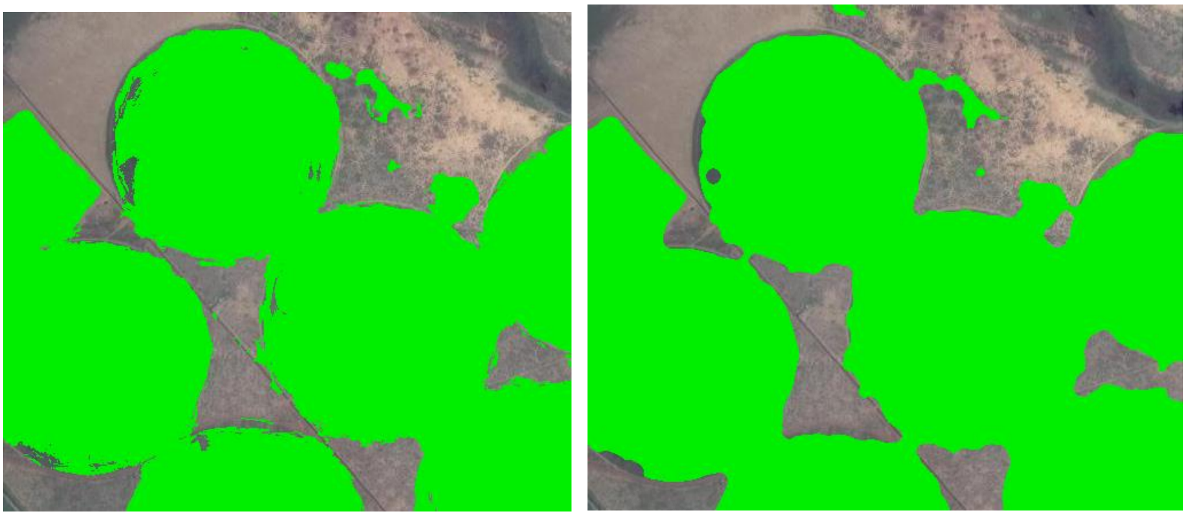


Figure 13. Effect of noise removal by modal filter and morphological opening and closing. Left: original classified image. Right: image after noise removal.

2.13. Distinction between small-scale and large-scale irrigation

Until now, we have described the methodology to distinguish areas that are irrigated from areas that are not. However, the overall aim of the project is to identify areas under irrigation that are farmer-led, which for the purpose of this study is mostly expressed through a small field size. The issue is that the machine learning algorithm we employ does not look at the size of fields, but only to the spectral characteristics of each pixel. Therefore, information on the field size is not obtained.

We are aware of a number of studies that have tried to do this, in particular the recent PhD study by M. Vogels (2019) . The methodology is called Geographic Object-based Image Analysis (GEOBIA), and it uses a segmentation of the image before classification is attempted. The method is highly experimental, and currently uses proprietary software (eCognition) that we had no access to. Although the results in that study were encouraging, in the end the authors could not yet demonstrate the accuracy of the approach. Given the time limitations for this study, we decided to not further pursue this route.

As image segmentation for remote sensing is an active field of research, it is probable that this methodology will make progress in the near future. One interesting development is that Google Earth Engine now also offers functionality for image segmentation (SNIC), although of a simpler nature than the eCognition software mentioned above. In addition, some researchers are employing graphics-card accelerated methods to perform image segmentation on smaller areas (Donchyts, 2017). For future work, the large-scale automated recognition of the size of small irrigated fields remains an interesting and hopefully viable option.

In the absence of a viable GEOBIA approach, one option to distinguish between small-scale and large-scale agriculture is to identify the location and extent of the large-scale schemes, flood recession cropping areas, and cultivated wetlands and inland valley bottoms using GIS information from government sources. However, in the timeframe for this report we have been unable to locate or produce such a layer.

As a second solution, we use the size of contiguous patches of pixels classified as irrigated as the 'size' of the irrigated area. Although this method is not perfect, as classified irrigated areas can be fragmented, this method give a good overall result in separating large, middle and small scale irrigated areas. Here, we consider areas smaller than 100ha to be small-scale, 100-500 ha to be medium-scale, and areas larger than 500 ha as large-scale.

As an example, the images below shows a region of the Niger river near Segou, and a region near Bamako. The green areas are identified as 'large scale', the red areas as 'small scale'.

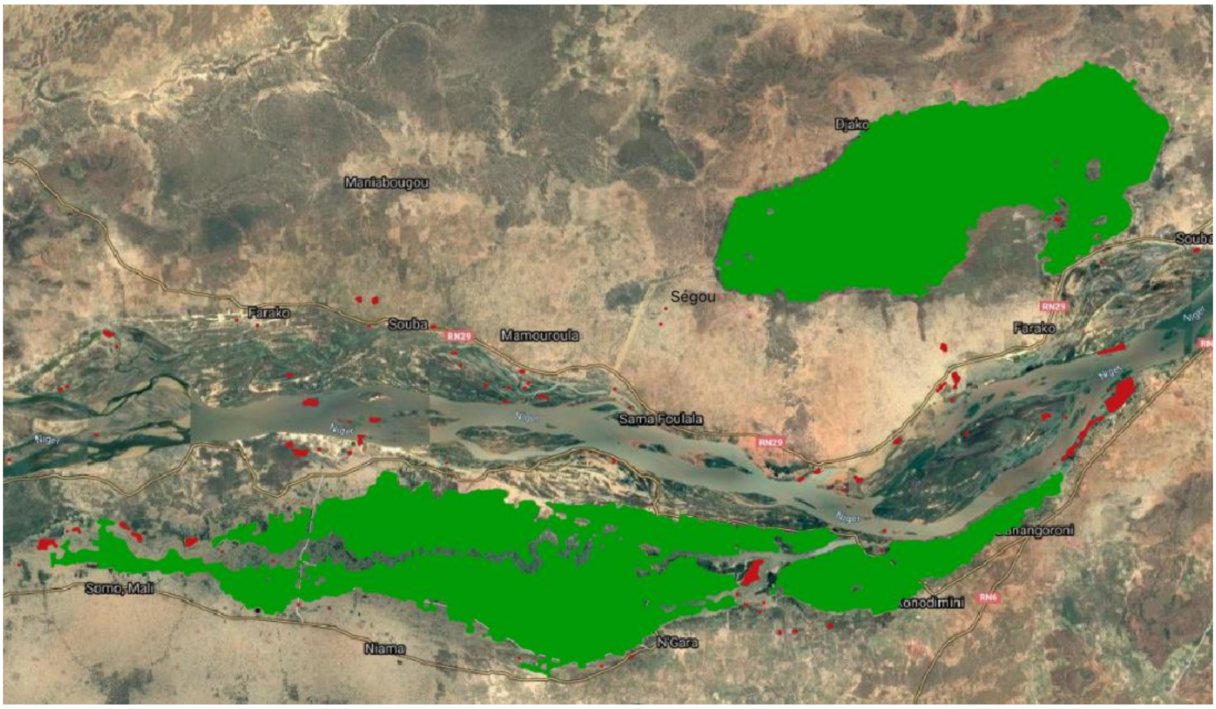


Figure 14. Irrigated areas near Segou. Green: three areas consisting of connected pixels, identified as 'large scale', Red: small-scale areas.

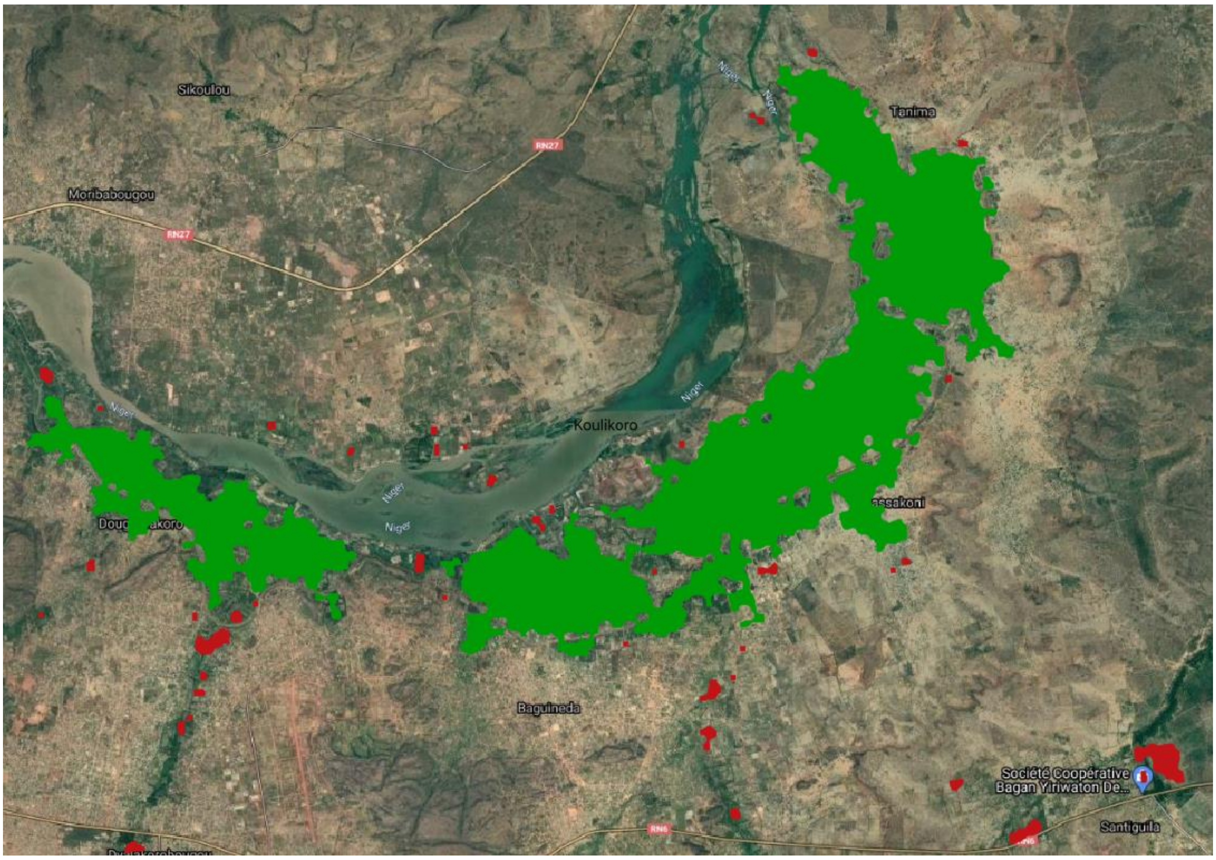


Figure 15. Irrigated areas near Bamako. Green: two areas consisting of connected pixels, identified as 'large scale', Red: small-scale areas.

3. ACTUAL IRRIGATED AREAS - RESULTS

3.1. IRRIGATION CLASSIFICATION

After training the machine learning method in the way described in chapter 2, we need to assess the accuracy with which it can distinguish between irrigated ground data points and non-irrigated points. To assess the quality of a classification, the validation part (30%) of the total available ground truth data is used. This data has never seen by the model, so it is a fair assessment of the accuracy. The training data is never used for accuracy validation, because the model has been trained using this data, and can therefore potentially 'remember' the right classification. This is called overfitting – and it is particularly a risk when very large numbers of bands and trees in the Random Forest model are used. Using the validation data circumvents this problem.

The accuracy of a classification is expressed using different numbers. First of all, there is the overall accuracy: the percentage of points that was classified correctly. This is derived from the classification table – a table that lists the number of points that were classified, and how they were classified. These are shown below.

However, this is often not the most interesting number. What we would like to know is first of all: which percentage of points that are actually irrigated have been correctly classified as irrigated? This is called the **Producer's Accuracy**. It indicates how well we can recognise actual irrigated areas. However, it says nothing about false positives: points that were not irrigated but were wrongly classified as such.

That is why we are interest in a second number: which percentage of the points that have been classified as irrigated *actually* are irrigated. This is called the **Consumer's Accuracy**, and it includes the effect of false positives. Both of these numbers can be calculated from the classification result. Below, we list these numbers for each of the regions in Mali and Chad. This indicates how much trust we can have in the classification.

To illustrate these numbers, we look at the table for North Mali. Here, we had 326 ground data points that were actually irrigated. When these were classified, 26 points were classified as non-irrigated, and 300 were classified as irrigated, leading to an accuracy percentage of 92.0%. This corresponds to the Producers Accuracy. On the other hand, of the 309 points that were actually classified as irrigated, 300 really were irrigated but 9 were mislabelled, leading to an accuracy percentage of 97.1%. This corresponds to the consumers accuracy.

North Mali:

Table 10. Classification accuracy for North Mali.

		Classified data		
		non-Irrigated	**irrigated**	Total
Ground truth data	**non-Irrigated**	**437**	**9**	446
	irrigated	**26**	**300**	326
	Total	463	309	772

Overall accuracy	95.5%
Producer's accuracy	92.0%
Consumer's accuracy	97.1%

Middle Mali:

Table 11. Classification accuracy for Middle Mali.

		Classified data		
		non-Irrigated	**irrigated**	Total
Ground truth data	**non-Irrigated**	**797**	**14**	821
	irrigated	**38**	**345**	383
	Total	835	369	1204

Overall accuracy	95.6%
Producer's accuracy	90.1%
Consumer's accuracy	96.1%

Southern Mali:

Table 12. Classification accuracy for South Mali.

		Classified data		
		non-Irrigated	**irrigated**	Total
Ground truth data	**non-Irrigated**	**988**	**8**	996
	irrigated	**52**	**303**	355
	Total	1040	311	1351

Overall accuracy	95.6%
Producer's accuracy	85.4%
Consumer's accuracy	97.4%

The overall accuracy for Mali is good, with all three values over 95%. We see that in the case of South Mali, the producers's accuracy is clearly lower, with 85.4%. This is caused by the widespread presence of forests in Southern Mali, which lead to confusion between forest and irrigated areas during the classification.

North Chad:

Table 13. Classification accuracy for North Chad.

		Classified data		
		non-Irrigated	**irrigated**	Total
Ground truth data	**non-Irrigated**	**370**	**7**	377
	irrigated	**17**	**132**	149
	Total	387	139	526

Overall accuracy	95.4%
Producer's accuracy	88.6%
Consumer's accuracy	95.0%

Middle Chad:

Table 14. Classification accuracy for middle Chad.

		Classified data		
		non-Irrigated	**irrigated**	Total
Ground truth data	**non-Irrigated**	**411**	**18**	429
	irrigated	**27**	**250**	277
	Total	438	268	706

Overall accuracy	93.6%
Producer's accuracy	90.3%
Consumer's accuracy	93.3%

Southern Chad:

Table 15. Classification accuracy for Southern Chad.

		Classified data		
		non-Irrigated	**irrigated**	Total
Ground truth data	**non-Irrigated**	**506**	**1**	507
	irrigated	**15**	**93**	108
	Total	521	94	615

Overall accuracy	97.4%
Producer's accuracy	86.1%
Consumer's accuracy	98.9%

The overall accuracy for Chad is good, with values between 93.6% and 97.4%. However, we see that in the case of North Chad (which includes Lake Chad), the producer's accuracy is clearly lower at 88.6%. This is caused by the fact that at Lake Chad water recession agriculture is practiced, which leads to confusion in the classification as the natural vegetation growth follows the same temporal pattern as the agricultural crops.

For both countries, we see that the producer's accuracy is lower than the consumer's accuracy. This means that the model is conservative during the classification process, mislabelling some points as non-irrigated that are actually irrigated. This strict rejection process causes the resulting final (consumer's) accuracy to be high, at the expense of missing some irrigated points.

Contribution of bands

To get an idea which input bands — both the raw satellite bands such as red, green, and blue, and the derived bands such as NDVI and EVI) — contribute most to the classification, we investigate the correlation and band importance. First of all, the correlation between the different bands is shown below.

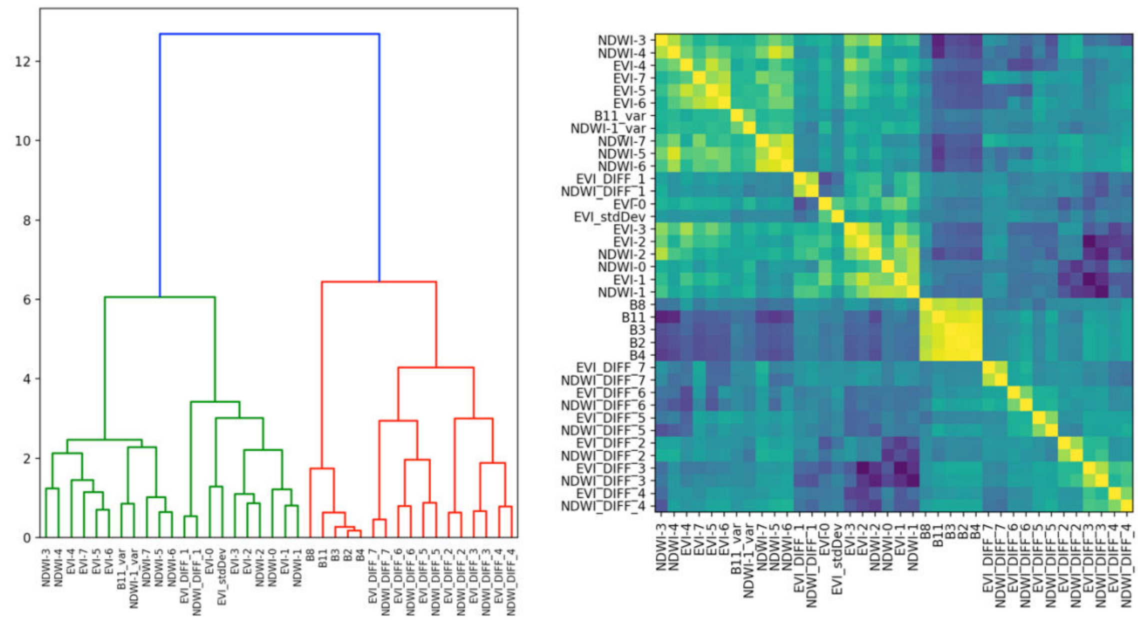


Figure 16. Left: bands clustered by correlation. Right: full correlation table. Blue: low correlation, yellow: high correlation.

From this image, we can identify two things. First of all, there is a quite a strong correlation between a number of bands, especially for the monthly differences of the index variables for EVI and NDWI (for example, EVI_DIFF_3 and NDWI_DIFF_3). However, as Random Forest handles strongly correlated features well, this is not a cause of concern. It does mean that we probably don't need all the 38 features that we use at the moment. Secondly, the raw colour bands (B2, B3, B4, B8, B11) show strong internal correlation, but little correlation with other bands.

Secondly, we can look at the contribution of each of the bands to the classification accuracy. We do this by looking at the *feature importance* as reported by the Random Forest algorithm, which measures the GINI impurity reduction achieved by each feature. The result is shown below.

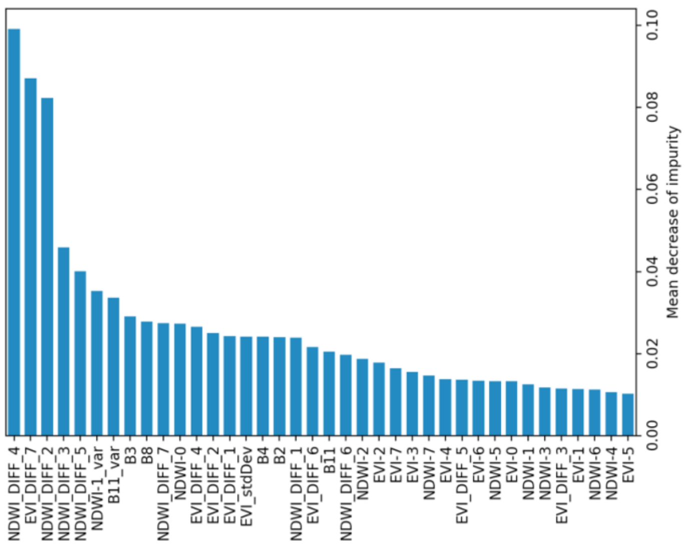


Figure 17. Feature importance for the Random Forest model. Higher bars are more important features.

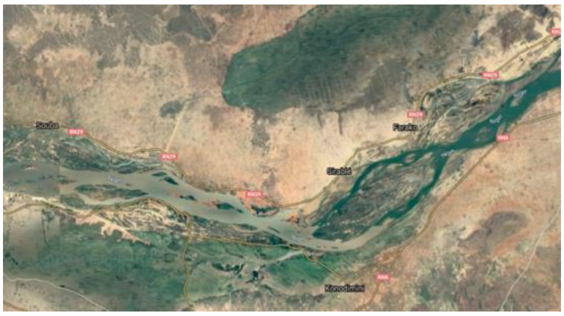
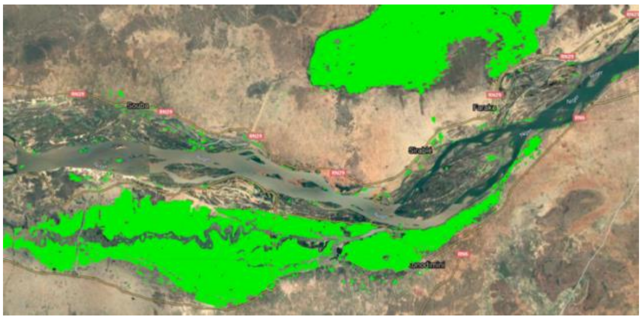
From this graph, it is clear that the monthly difference indexes such as NDWI_DIFF_4 and EVI_DIFF_7, are the most important features. This can be intuitively understood, as these features capture plant growth in the dry season, and just before the rainy season. It should be kept in mind that a low score in this graph does not necessarily mean that a feature has no importance. If two features are strongly correlated, it can happen that one is given a high score, but once it is 'used' by the model, the other feature doesn't add any new information and is therefore given a low score. It is therefore not straightforward to delete features, as there can be benefits in retaining correlated features. As the Random Forest model does not have any restrictions in the number of features used, it was decided to keep the full set in the final classification computation. However, this has to be balanced to the increased computational and memory cost of an increased number of bands. In practice, we chose a cut-off of 38 bands as the limit what was computationally feasible.

3.2 Qualitative classification results

Below, we show the classification results of the same areas as were listed in Chapter 2. Here, we show the result of the actual classification, with the postprocessing (of morphological opening and closing).

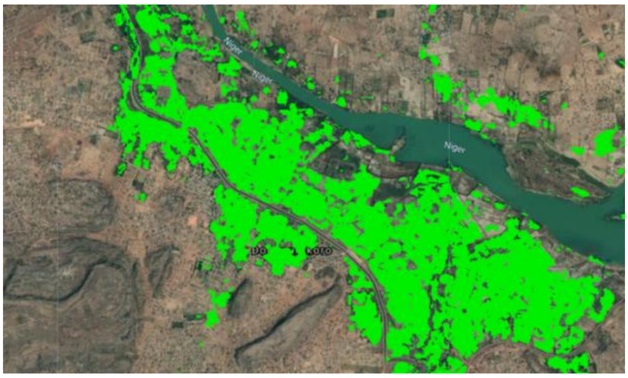
Mali - Konodimini

Table 16. Classification result near Konodimini, Mali.

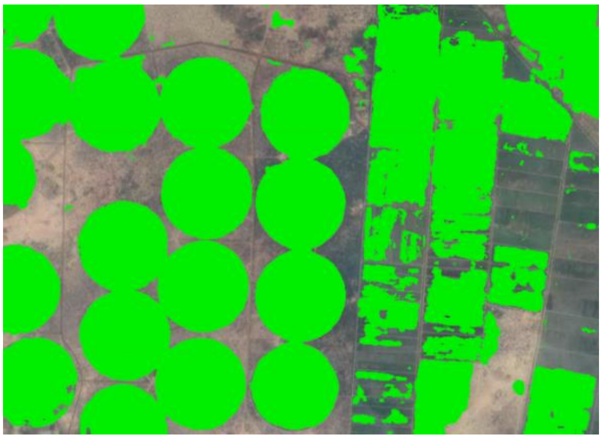
Mali - Bamako

Table 17. Classification result near Bamako, Mali.

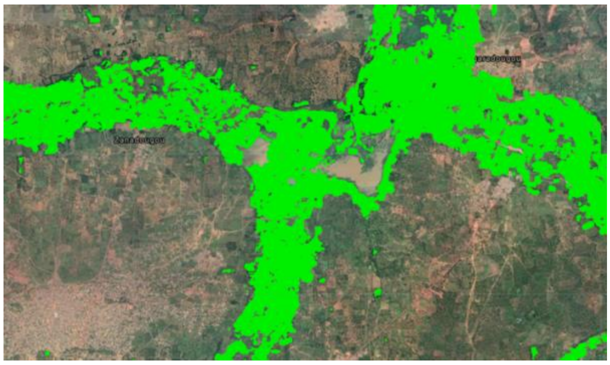
Mali - Office du Niger

Table 18. Classification result near Office du Niger, Mali.



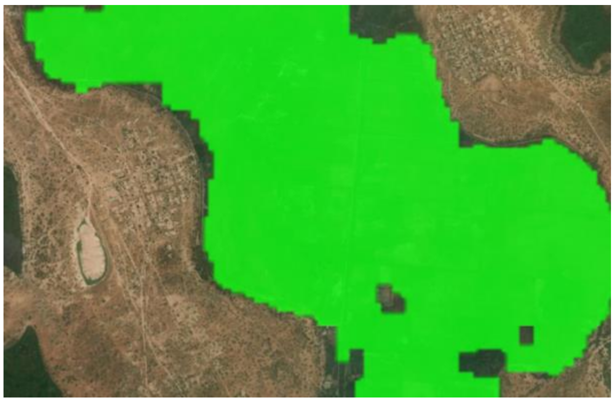
Mali - Sikasso

Table 19. Classification result near Sikasso, Mali.



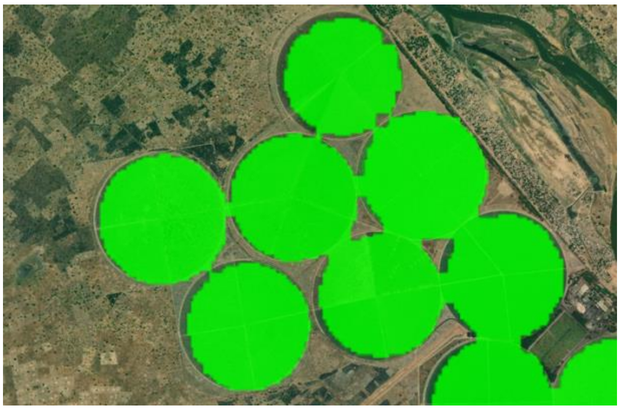
Chad – Lake Chad

Table 20. Classification result near Lake Chad, Chad.



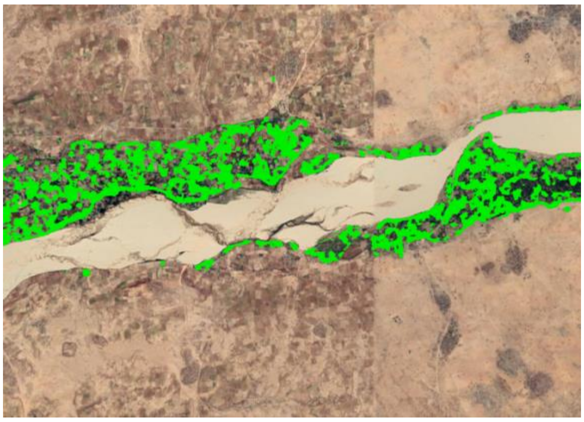
Chad – Sarh

Table 21. Classification result near Sarh, Chad.



Chad – Diour

Table 22. Classification result near Diour, Chad.



The full map of irrigated areas in Mali is shown below.

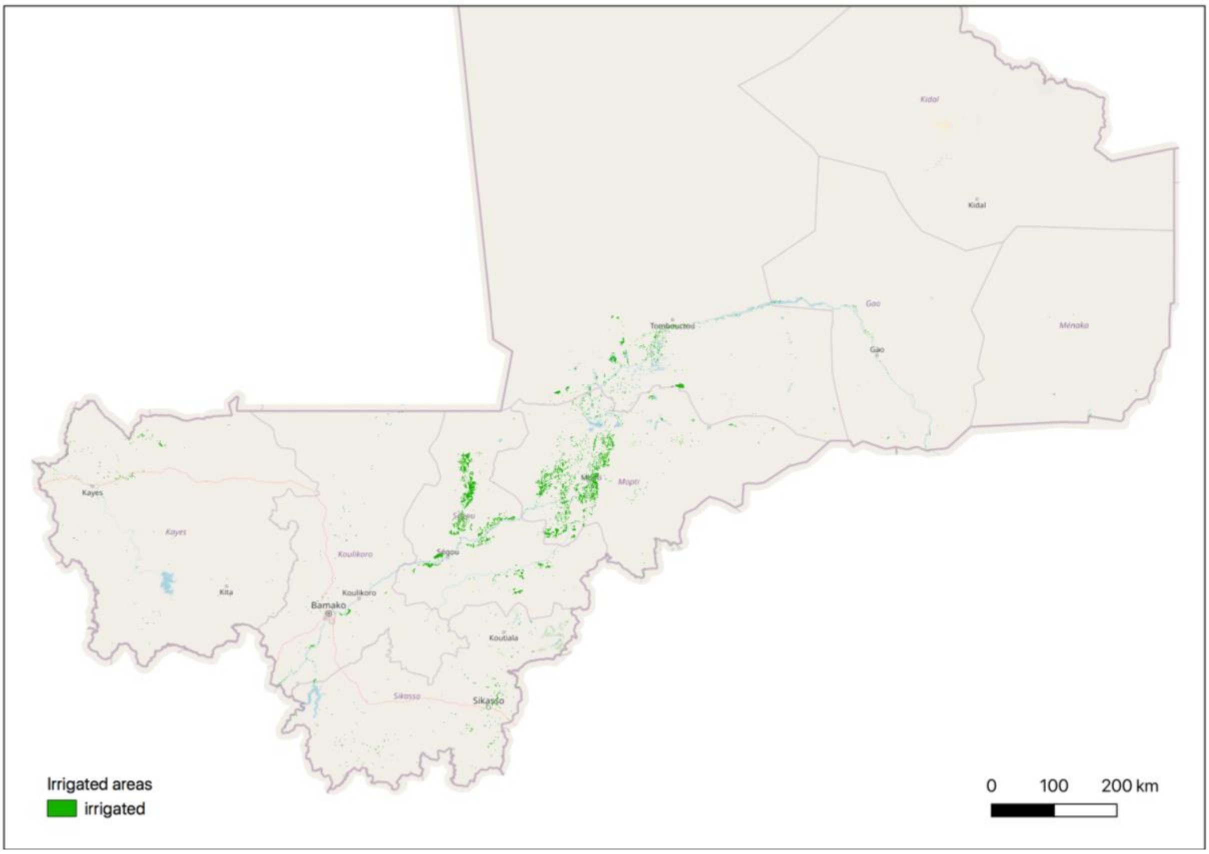


Figure 18. In green: irrigated areas in Mali as determined by the machine learning classification.

The map above reflects the actual surface areas. However, as irrigated areas are small when compared to the full country, the areas are hard to distinguish in the map. Therefore, in the map below we have enlarged the areas to improve the visibility. It should be kept in mind that this map serves to indicated the locations of the irrigated areas, but does not reflect the actual surface areas.

In addition, the irrigated areas are divided into three categories: those smaller than 100 ha, those between 100-500 ha, and those larger than 500 ha.

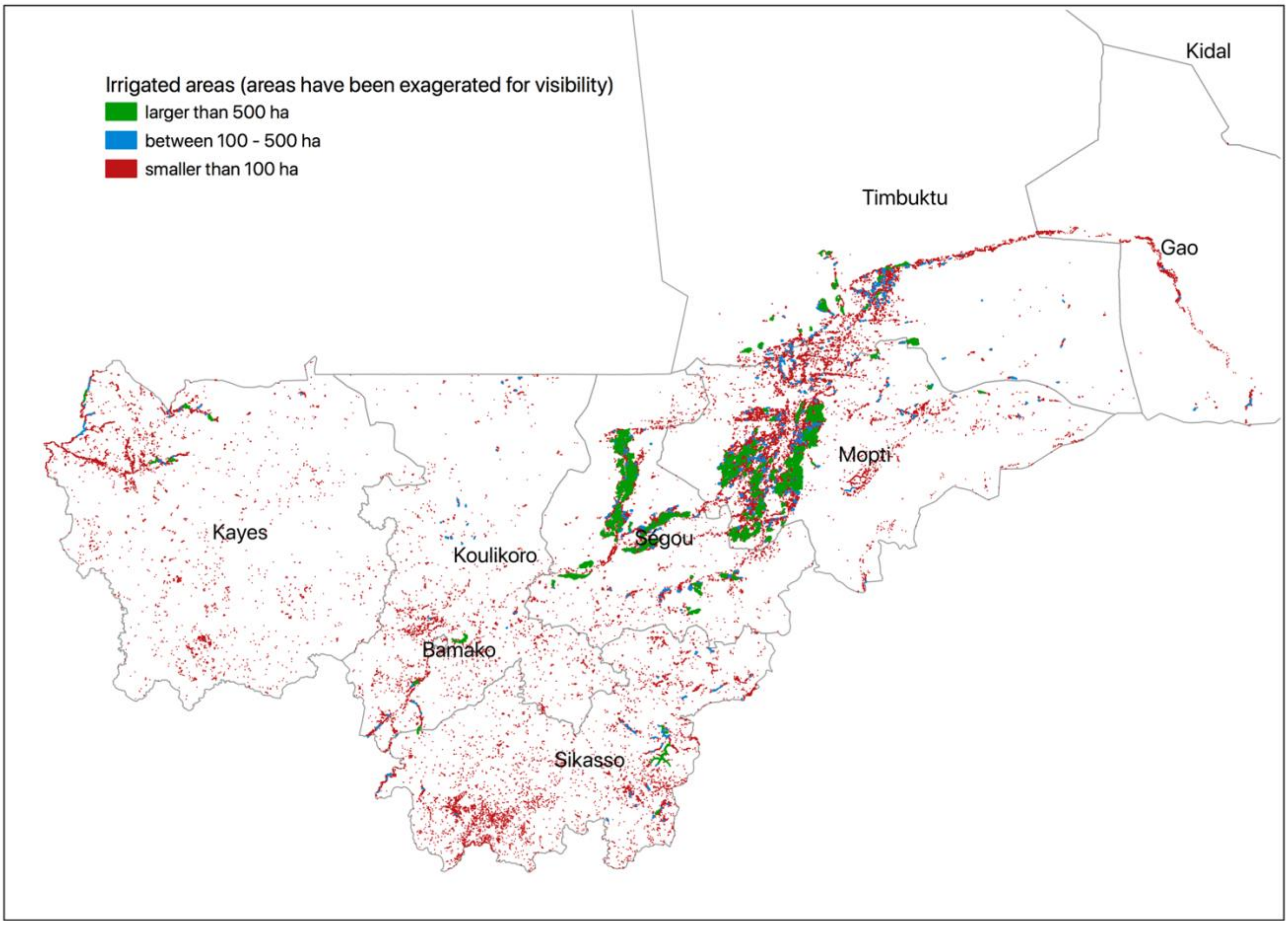


Figure 19. Irrigated areas in Mali. The size of areas has been exaggerated for easier visibility. A distinction is made between large, medium, and small-scale areas.

The full map of irrigated areas in Chad is shown below.

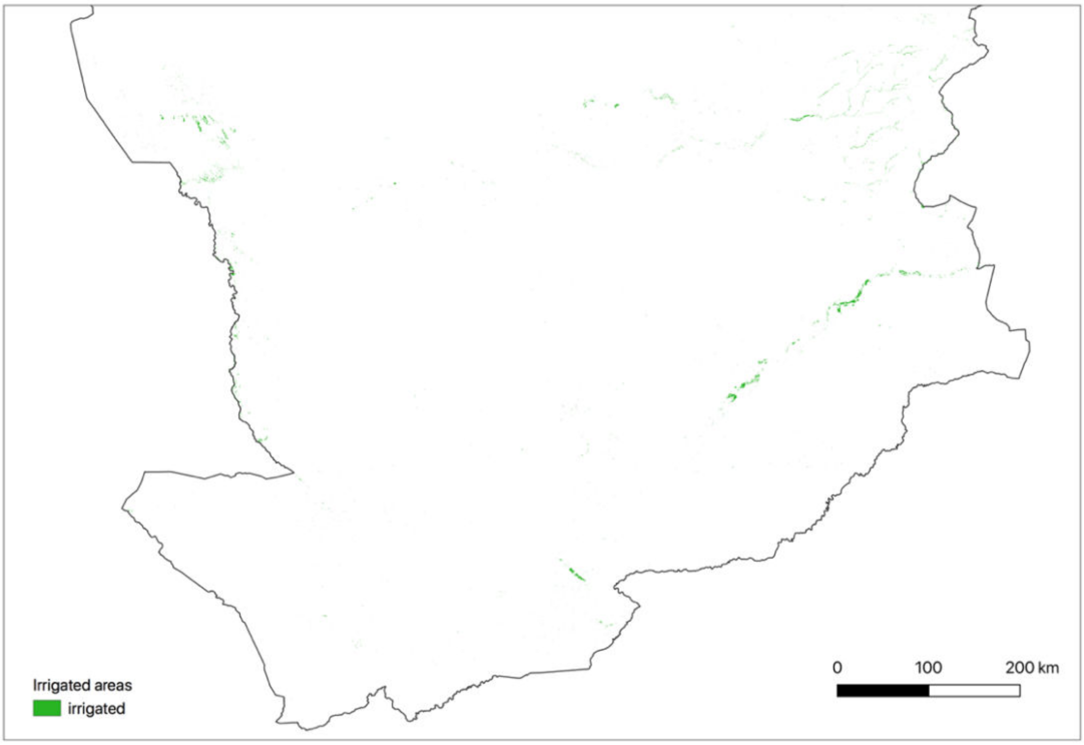


Figure 20. Green: Irrigated areas in Chad. The map background has been removed to more clearly show the irrigated areas.

In the map below we have enlarged the areas to improve the visibility. It should be kept in mind that this map serves to indicated the locations of the irrigated areas, but does not reflect the actual surface areas. In addition, the irrigated areas are divided into three categories: those smaller than 100 ha, those between 100-500 ha, and those larger than 500 ha.

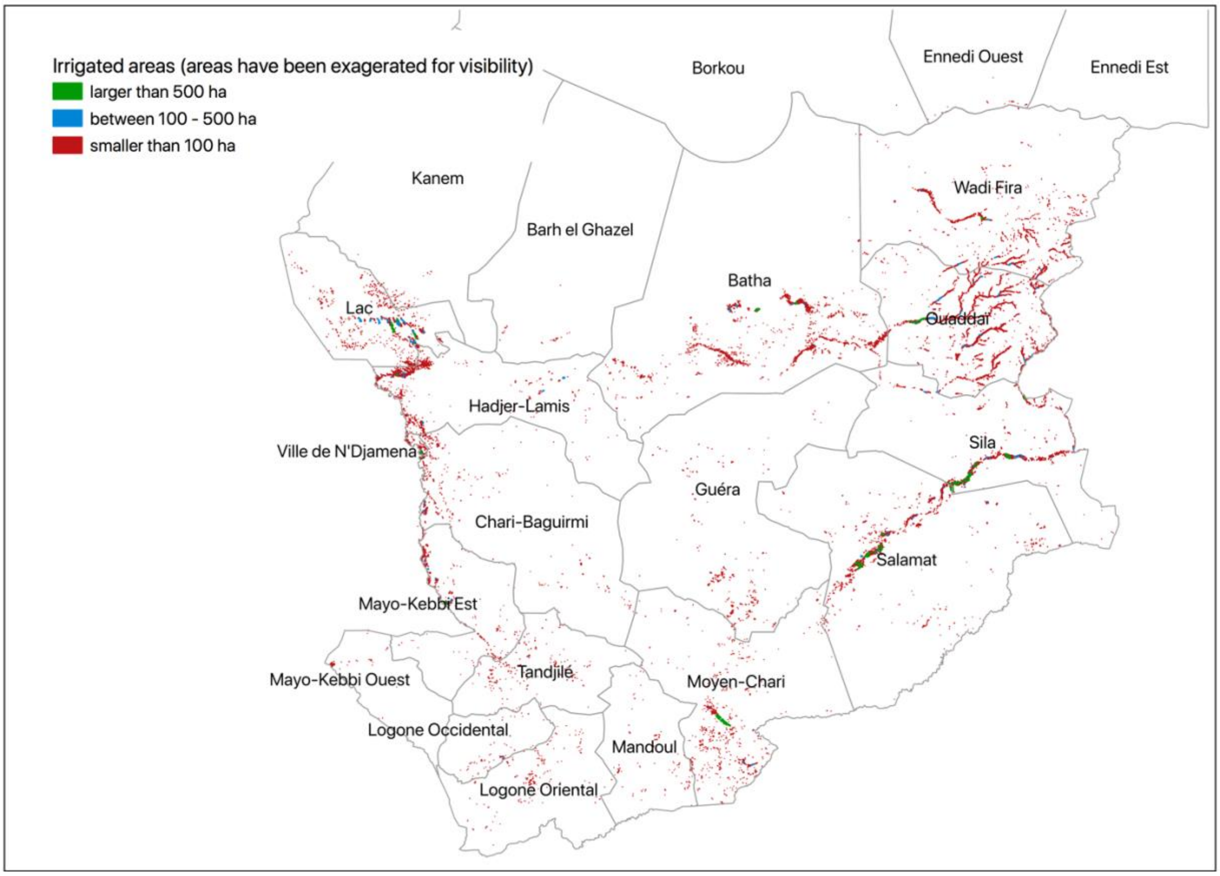


Figure 21. Irrigated areas in Chad. The size of areas has been exaggerated for easier visibility,which means that the apparent surface area on this map does NOT correspond to real surface area. A distinction is made between large, medium, and small scale areas.

As reference, below we reproduce the recent Global Cropland map, created as part of the Global Cropland Project (GFSAD30)[11] (Xiong, 2017), which estimates the cropland extent in Africa on 30m resolution in the year 2015. The maps for cropland extent for Mali and Chad are shown below.

[11] Source: GFSAD30, croplands.org

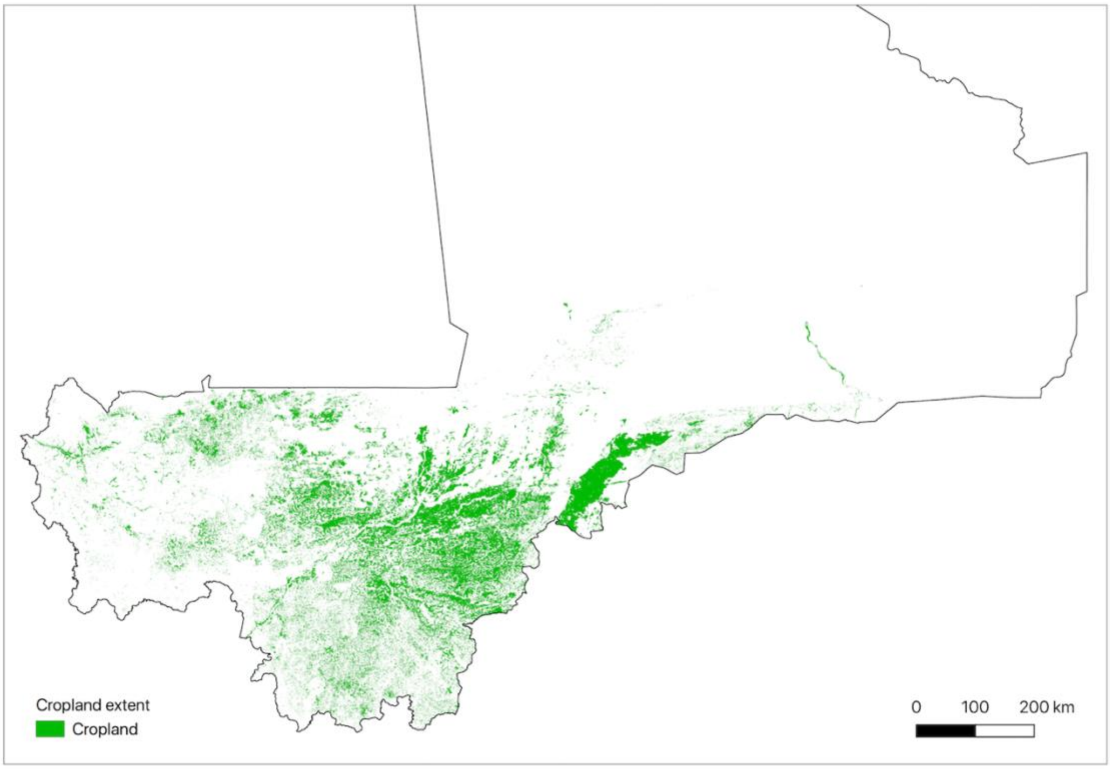


Figure 22. Green: Cropland extent in Mali (GFSAD30, Xiong 2017)

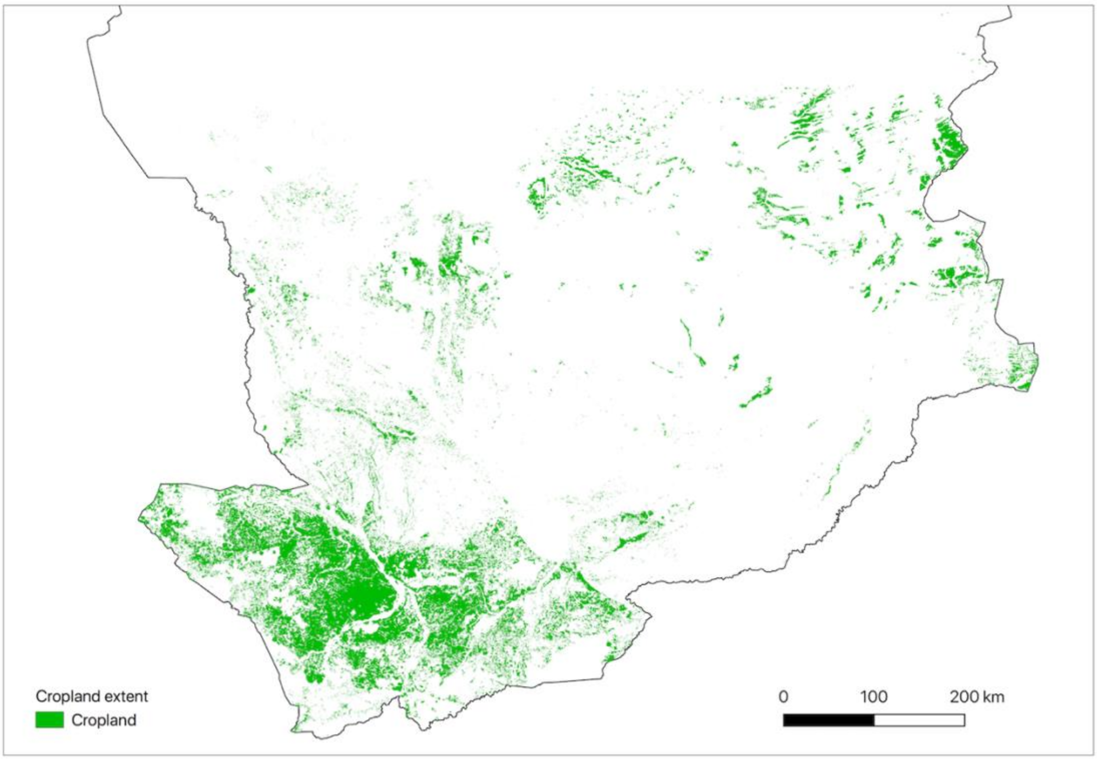


Figure 23. Green: Cropland extent in Chad (GFSAD30, Xiong 2017)

3.3. Quantitative description of the currently irrigated areas in the dry season

To determine the total area classified as irrigated in the dry season, we used the zonal statistics tool in QGIS. This tool counts the number of pixels with a given classification in a GeoTiff image. After this, the total area can be computed from the area of a single pixel, which is obtained from the image resolution. For the zones, we use a GADM administrative area layer[12]. In the tables below, the results of this process are shown. The table also includes the cropland extent as estimated by GFSAD30, and the percentage of irrigated land in the dry season as a fraction of the cropland.

Mali

Table 23. Area classifed as irrigated using the machine learning process, in Mali.

Region	Total area 10^3 ha	Cropland extent (GFSAD30) 10^3 ha	Irrigated area 10^3 ha	Irrigated area as percentage of cropland area
Bamako	25	3.79	0.09	2.4 %
Gao	17,057	66.67	5.42	8.1 %
Kayes	11,974	1,299.65	21.73	1.7 %
Kidal	15,145	0.00	0.01	-
Koulikoro	9,012	2,578.36	18.55	0.7 %
Mopti	7,902	1,710.41	247.58	14.5 %
Ségou	6,482	2,715.07	169.82	6.3 %
Sikasso	7,028	2,273.58	32.51	1.4 %
Timbuktu	49,611	63.69	69.93	109.8 %
Total	**124,236***	**10,711.22**	**565.65**	**5.3 %**

* Small deviations with respect to official numbers may be caused by geographical projection errors. They do not effect the results in a meaningful way.

Note: The numbers for Bamako might seem low. This is caused by the fact that the 'Bamako' entry refers only to the administrative area of Bamako itself, which is quite small (around 10 x 10 km). The irrigated areas in the area 'around Bamako' are therefore included in the entries for the surrounding regions in the table.

Large scale versus medium scale areas

Using the size patches of contiguous pixels, the sizes of irrigated areas was determined. The numbers are listed in the table below.

Table 24. Size breakdown of irrigated areas in Mali.

Size of irrigated patch	Area (kha)
Smaller than 100 ha	144.3
Between 100 – 500 ha	82.6

[12] GADM global administrative area GIS layers, gadm.org.

Larger than 500 ha	338.7
Total	**565.6**

A break down per region is given in the table below.

Table 25. Regional breakdown of sizes of irrigated areas in Mali.

Name	irrigated areas < 100 han 10^3 ha	Irrigated areas between 100 – 500 ha 10^3 ha	Irrigated areas > 500 ha 10^3 ha	Total 10^3 ha
Bamako	0.1	0.0	0.0	0.1
Gao	3.6	1.7	0.0	5.2
Kayes	14.1	3.3	4.7	22.1
Kidal	0.0	0.0	0.0	0.0
Koulikoro	9.9	4.0	5.2	19.1
Mopti	53.4	34.8	160.8	249.0
Ségou	18.7	13.8	136.8	169.2
Sikasso	17.3	7.2	7.6	32.1
Timbuktu	23.9	16.3	28.5	68.7
Total	**140.9**	**81.1**	**343.6**	**565.6**

Note that the totals differ marginally from the totals in the previous table. This is caused by the process to determine the sizes of the areas from contiguous areas of pixels. If such a patch straddles two different regions, it is counted twice. A correction factor was used to bring the country total in line with the result in the tables above. This table should therefore be used for qualitative assessments only.

Chad

Table 26. Area classifed as irrigated using the machine learning process, in Chad.

Region	Total area	Cropland extent (GFSAD30)	Irrigated area	Irrigated area as percentage of cropland area
	10^3 ha	10^3 ha	10^3 ha	
Barh el Ghazel	5,631	0.12	0.10	85.21 %
Batha	9,041	249.61	9.19	3.68 %
Borkou	25,623	0.00	0.00	-
Chari-Baguirmi	4,603	318.40	2.75	0.86 %
Ennedi Est	7,737	0.00	0.00	-
Ennedi Ouest	11,105	0.00	0.06	-
Guéra	6,104	38.82	1.18	3.04 %
Hadjer-Lamis	3,044	239.81	6.83	2.85 %
Kanem	6,767	0.85	0.22	25.85 %
Lac	1,984	3.92	11.51	293.92 %
Logone Occidental	881	663.31	0.29	0.04 %
Logone Oriental	2,372	808.65	0.48	0.06 %
Mandoul	1,744	581.27	0.33	0.06 %

Mayo-Kebbi Est	1,805	406.51	3.70	0.91 %
Mayo-Kebbi Ouest	1,254	398.50	0.26	0.07 %
Moyen-Chari	4,147	386.62	6.25	1.62 %
Ouaddaï	2,975	291.66	18.27	6.26 %
Salamat	6,796	67.78	15.54	22.93 %
Sila	3,570	329.24	18.26	5.55 %
Tandjilé	1,753	587.65	0.61	0.10 %
Tibesti	12,608	0.00	0.00	-
Ville de N'Djamena	40	2.07	1.41	68.22 %
Wadi Fira	5,412	2.28	7.47	327.47 %
Total	**126,996***	5377.05	**104.72**	1.95 %

* Small deviations with respect to official numbers may be caused by geographical projection errors. They do not effect the results in a meaningful way.

Large scale versus medium scale areas

Using the size of patches of contiguous pixels, the sizes of irrigated areas was determined. The numbers are listed in the table below.

Table 27. Size breakdown of irrigated areas in Chad.

Size	Area (kha)
Smaller than 100 ha	50.6
Between 100 – 500 ha	20.2
Larger than 500 ha	33.9
Total	**104.7**

A break down per region is given in the table below.

Table 28. Regional breakdown of sizes of irrigated areas in Chad.

Name	irrigated areas < 100 ha 10^3 ha	Irrigated areas between 100 – 500 ha 10^3 ha	Irrigated areas > 500 ha 10^3 ha	Total 10^3 ha
Barh el Ghazel	0.10	0.00	0.00	0.10
Batha	6.12	1.50	1.25	8.87
Borkou	0.00	0.00	0.00	0.00
Chari-Baguirmi	1.50	0.55	0.58	2.63
Ennedi Est	0.00	0.00	0.00	0.00
Ennedi Ouest	0.06	0.00	0.00	0.06
Guéra	1.17	0.00	0.00	1.17
Hadjer-Lamis	4.27	1.92	0.50	6.69
Kanem	0.21	0.00	0.00	0.21
Lac	3.36	4.16	2.91	10.43
Logone Occidental	0.16	0.13	0.00	0.29
Logone Oriental	0.45	0.00	0.00	0.45
Mandoul	0.32	0.00	0.00	0.32
Mayo-Kebbi Est	2.11	0.84	0.64	3.59

Mayo-Kebbi Ouest	0.26	0.00	0.00	0.26
Moyen-Chari	1.64	0.54	3.58	5.77
Ouaddaï	11.84	3.85	2.01	17.70
Salamat	4.01	1.84	13.10	18.94
Sila	3.82	2.78	11.53	18.13
Tandjilé	0.59	0.00	0.00	0.59
Tibesti	0.00	0.00	0.00	0.00
Ville de N'Djamena	0.64	0.11	0.60	1.35
Wadi Fira	5.41	1.25	0.51	7.17
Total	**48.1**	**19.5**	**37.2**	**104.7**

Note that the totals differ marginally from the totals in the previous table. This is caused by the process to determine the sizes of the areas from contiguous areas of pixels. If such a patch straddles two different regions, it is counted twice. A correction factor was used to bring the country total in line with the result in the tables above. This table should therefore be used for qualitative assessments only.

The totals are further discussed in Chapter 6.

One issue that is apparent from the table is that in some regions the percentage of irrigated area is more than 100% of area identified as cropland in the GFSAD30 project. After inspection, it became clear that in these cases the GFSAD30 data does not correctly classify areas as cropland. An example of that, in the case of the Lac region, is shown below.

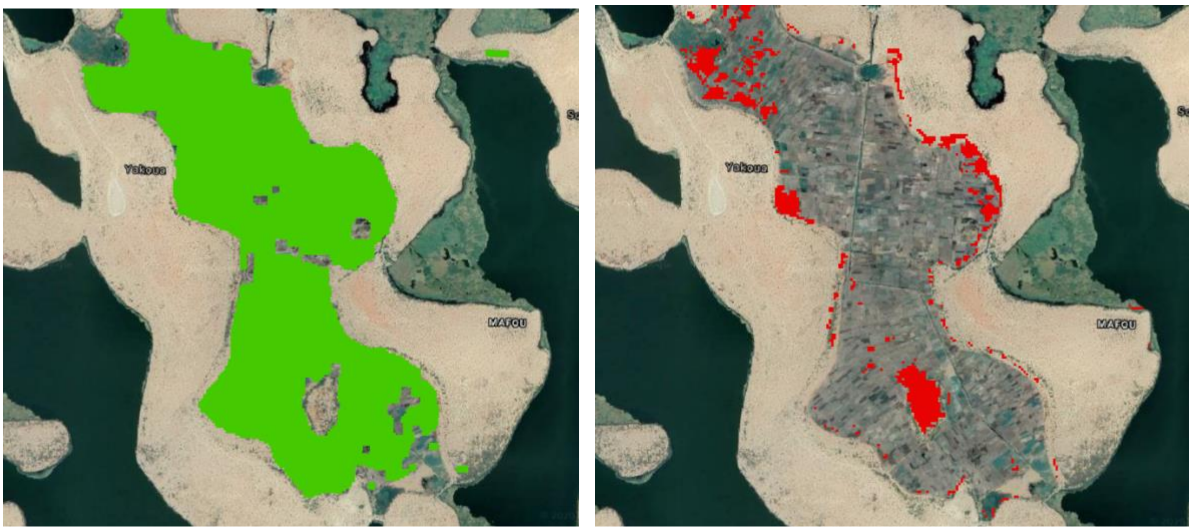


Figure 24. Left: Area identified as irrigated in present study. Right: in red the area identified as cropland by GFSAD30. Clearly, GFSAD30 does not accurately capture the cropland area here.

3.4. Web map

A static web map was created from the classified image and from the image of suitable areas. To do this, the classified image that was the result of the analysis was made available on the web in a form that can be accessed online (see link below), and can also be accessed as a data layer in common Geographical Information Software (GIS), such as ArcGIS or QGIS. In this way, the data can be easily combined with other data layers. The web map is available here:

[Link to web map of Mali and Chad irrigated areas]

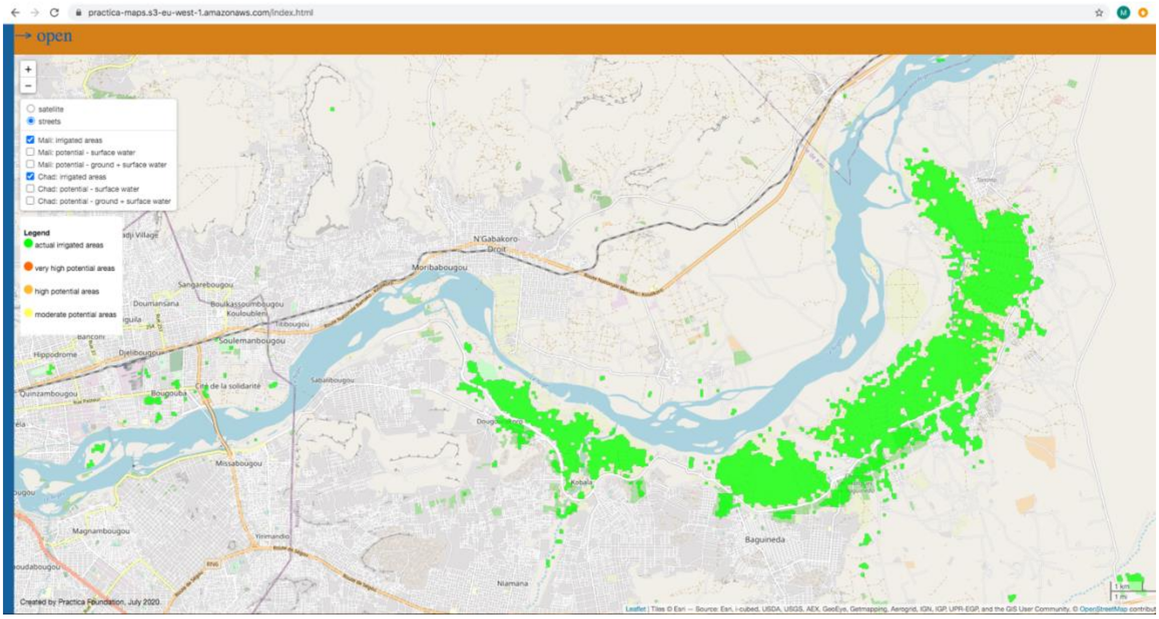


Figure 25. The web map, showing actual irrigated area near Bamako.

The links to the individual layers are listed in the table below. These can be used to add the layers in GIS packages such as ArcGIS or QGIS. The URLs all start with “https://practica-maps.s3-eu-west-1.amazonaws.com/”.

Table 29. URL addresses of GIS layers created in this study.

Layer	URL
Mali – actual irrigated area	mali/actual/{z}/{x}/{y}.png
Mali – suitable areas – surface water	mali/scenario1/{z}/{x}/{y}.png
Mali – suitable areas – groundwater < 7m + surface water	mali/scenario2/{z}/{x}/{y}.png
Mali – suitable areas – groundwater < 23m + surface water	mali/scenario3/{z}/{x}/{y}.png
Chad – actual irrigated area	chad/actual/{z}/{x}/{y}.png
Chad – suitable areas	chad/scenario1/{z}/{x}/{y}.png
Chad – suitable areas – groundwater < 7m + surface water	chad/scenario2/{z}/{x}/{y}.png
Chad – suitable areas – groundwater < 23m + surface water	chad/scenario3/{z}/{x}/{y}.png

So, as an example, the full link to the Mali actual irrigated is:
https://practica-maps.s3-eu-west-1.amazonaws.com/mali/actual/{z}/{x}/{y}.png.

4. AREAS SUITABLE FOR IRRIGATION – DATA AND METHODOLOGY

4.1. CHOICE OF METHODOLOGY

Whether a certain area is suitable for FLID-type irrigation depends on many factors. First of all, there is the base determinant of water availability, both in terms of nearness to surface water and groundwater availability. Secondly, data layers on slope, current land cover, national parks, and nearness to cities are readily available and can be used to restrict suitable areas. Thirdly, there are a large number of additional factors that can constrain suitability, for which it is harder to obtain data, such as local political situation, safety, land ownership, flooding potential, detailed topography, local soil conditions, pollution, salinity, etc.

To strike a balance in the choice of layers to use, we follow a recent paper by IWMI (Schmitter, 2018), that estimates the potential for solar irrigation in Ethiopia using a multi-criteria model. Their methodology includes nearness to surface water, groundwater depth and aquifer properties, slope, nearness to cities, land cover, national parks, and solar irradiation as input parameters. From these layers, suitability is classified from not suitable to highly suitable using a weighted scoring mechanism. Although the IWMI paper is focused on solar pumps, their methodology can be easily adapted to irrigation in general by leaving out solar irradiation as an input parameter, which we have done in this report. The detailed layers we use are described in the next section.

It is important to stress that this type of analysis *only* identifies the areas suitable for irrigation from a narrow list of constraints and parameters. The local factors mentioned above will further restrict the suitable areas. Therefore, a local assessment will always be needed when specific sites are chosen, in which the full complexity of determining irrigation potential can be determined.

A further limitation of this type of analysis is that it identifies areas suitable for irrigation, but not the total *potential of sustainable growth* for irrigation. For that, it is necessary to do a full analysis of the hydrological situation, such as water balance and possible groundwater depletion. This is beyond the scope of this analysis, but has been done by a number of authors, which are listed in chapter 6.3. Therefore, the results for the suitable areas in this report should not be confused with the total potential sustainable increase.

Both of these issues need a lot more analysis and modelling to arrive at results that can be considered valid in a given locality. Therefore, the current study should be used as an indication of which areas might be suitable for further detailed study.

4.2 DATA SOURCES

The analysis relies on a number of GIS layers to compute the constraints and scores: slope, accessibility to cities, protected areas and national parks, rivers and water bodies, land cover, water occurrence, groundwater depth, aquifer productivity, and aquifer storage. In the table below, all the layers used for the constraints and the scores are listed together with their source.

Table 30. Data sources used for multi-criteria irrigation suitability model.

Layer name	Source	Spatial resolution	URL
Slope	NASA/CGIAR	90m	https://developers.google.com/earth-engine/datasets/catalog/CGIAR_SRTM90_V4

Accessibility to cities	Univ. of Oxford	1000m	https://developers.google.com/earth-engine/datasets/catalog/Oxford_MAP_accessibility_to_cities_2015_v1_0
Protected areas	UNEP-WCMC	100m	https://developers.google.com/earth-engine/datasets/catalog/WCMC_WDPA_current_polygons
Rivers	WWF	50m	https://www.hydrosheds.org/page/hydrorivers
Water bodies	WWF	50m	https://www.hydrosheds.org/pages/hydrolakes
Landcover	Copernicus	100m	https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_Landcover_100m_Proba-V_Global
Water occurrence	EC JRC / Google	100m	https://developers.google.com/earth-engine/datasets/catalog/JRC_GSW1_2_GlobalSurfaceWater
Groundwater depth	BGS	5000m	https://www.bgs.ac.uk/research/groundwater/international/africanGroundwater/mapsDownload.html
Aquifer productivity	BGS	5000m	https://www.bgs.ac.uk/research/groundwater/international/africanGroundwater/mapsDownload.html
Aquifer storage	BGS	5000m	https://www.bgs.ac.uk/research/groundwater/international/africanGroundwater/mapsDownload.html

Note that the groundwater maps have a low resolution of 5km. This is a severe limitation, and currently no better-quality maps are available for this. We discuss the effects of this in more detail in section 6.6. Our main conclusion there is that the low-resolution maps will mostly affect the accuracy on the small local level, but should not affect the outcome too much on a regional level.

Aquifer productivity and storage are included as layers in addition to water depth, as they both can constrain the suitability of a given location. Aquifer productivity refers to the potential of an aquifer to sustain various levels of groundwater flow and/or abstraction from a borehole. For example: a very clayey sand layer can have a very low productivity because the water does not flow through it easily. In that case, even if the water is not deep, the area is unsuitable for groundwater irrigation.

Secondly, aquifer storage refers to the amount of water that is stored in an aquifer. It depends on the aquifer thickness and on the porosity of the material. If the storage is too low, sustained abstraction may not be possible, and therefore the site is not suitable for groundwater irrigation.

Accessibility to cities layer

Created for the Malaria Atlas Project by Oxford University, Google, the University of Twente, and the EU Joint Research Center, this layer estimates the land-based travel time to the nearest densely populated area, in 2015. Densely populated areas are defined as areas with 1,500 or more inhabitants per square kilometre, or a majority of built-up land cover types coincident with a population centre of at least 50,000 inhabitants. It is based on underlying data sets including the road network, type of road, railways, rivers, lakes, oceans, slope, elevation, landcover types, and national borders.

For each of these datasets, a speed of travel was estimated. The datasets were then combined to form a ‘friction surface’, where each pixel was assigned a value for the estimated travel time to cross that pixel. The travel times where then generated from this friction surface. Details are described in a 2018 Nature paper (Weiss, 2018).

4.3 Constraints

Before computing suitability scores, the area considered is limited by applying a number of constraints, based on slope, protected areas and national parks, water bodies and rivers, and land cover. The table below lists the layers and provides more detail on the exact constraints used.

Table 31. Layers used as constraints.

Layer name	Constraint
Slope	Land with a slope larger than 8% is excluded.
Protected areas	Protected areas are excluded
Water bodies and rivers	Water bodies and rivers are excluded (only the actual water bodies themselves, not the shore)
Land cover	Areas that are not either shrubs, herbaceous vegetation, or cultivated/managed vegetation (agriculture) are excluded.

The accessibility to cities layer was not used as a constraint, as even when a region might be far away from large cities, irrigation on a smaller scale might still be practiced successfully for the local market. Therefore, it is kept as a weighing layer, but not used as an absolute constraint.

4.4 Scoring

Secondly, for the areas remaining, a score was computed. The first step is to assign scores to the individual layers. The next step then is to combine all the scores into a single score using a weighted average.

Table 32. Scoring of irrigation suitability for the different layers used.

		Score				
		Very highly suitable	Highly suitable	Moderately suitable	Less suitable	Not suitable
		5	4	3	2	1
Layer	Slope	0-2%	2-4%	4-6%	-	-
	Distance to water	<50m	50-100m	100-200m	200-300m	>300m
	Groundwater depth	<7m	7-25m	-	-	-
	Aquifer productivity	>0.5	0.5-0.1	-	-	-
	Aquifer storage	25k-50k	10k-25k	1k-10k	-	-
	Accessibility to cities	<120 minutes	120-240 min	240-480 min	480-720 min	>720 min

As water availability is of prime importance, we created three different scenarios for computing the final score. The first scenario only considers surface water as a source. We exclude streams and small rivers with an average flow of less than 0.1 m^3 a second, as these have a large chance of running dry in the dry season. It should be noted that even though this boundary is somewhat arbitrary, many small streams and rivers have high water content in the surrounding soil, such as in the case of sand rivers. Areas next to these streams often lend themselves for small scale irrigation using small, manually drilled irrigation wells.

The second scenario considers both surface water and groundwater up to a depth of 7 meters. The third scenario considers both surface water, and groundwater up to a depth of 23 meters.

The depths of 7 and 23 meters are chosen because correspond to the limits in the source GIS layer.

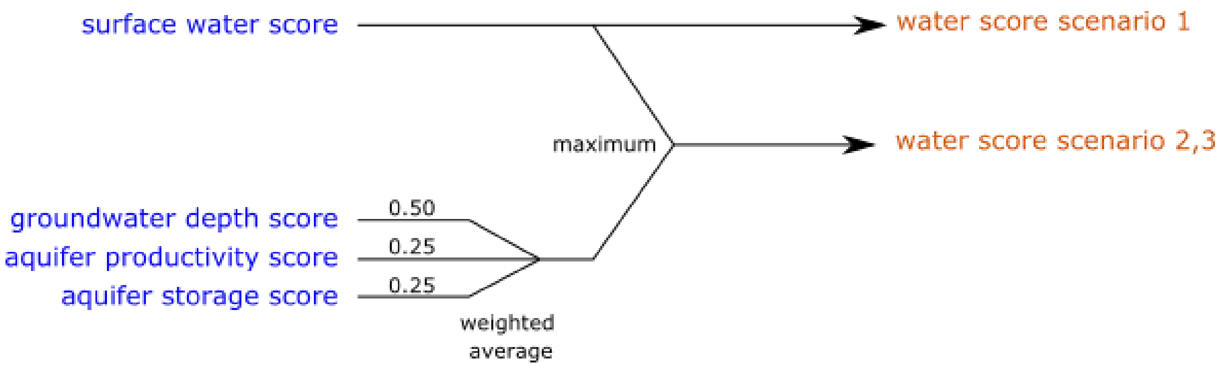
This is summarized in the table below.

Table 33. Three different irrigation suitability scenarios.

Scenario	Scoring
1 – surface water	As in table above
2 – surface water + groundwater < 7m	Groundwater < 7m – score 5 Groundwater > 7m – score 0
3 – surface water + groundwater < 23m	Groundwater < 7m – score 5 Groundwater >7m and < 23m – score 4

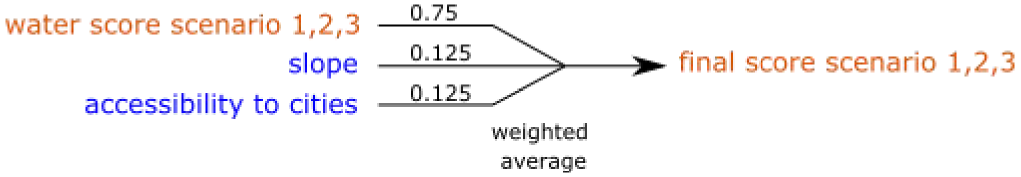
The reasoning behind considering the combined layers is that they overlap in practice. Near a river, for example, the area suitable for surface water irrigation is very likely to also support groundwater use. In addition, when both surface water and groundwater are available, it will be almost always be more practical to use the surface water for irrigation. Because of this, there is little benefit in treating the groundwater layer separately.

To keep it simple, we combine the various scores that are related to water into a single *water score*. The process for doing so is depicted in the image below.



The surface water score is used directly for scenario 1, in which groundwater is not considered. The three groundwater scores (depth, productivity and storage) are combined to form the final water score using a weighted average, in which the groundwater depth is given twice the weights of aquifer productivity and aquifer storage scores, which is in line with the weights used in the IWMI study (Schmitter, 2018). Finally, the maximum is taken of the surface water score and the combined groundwater score, which leads to the final water score for scenarios 2 and 3. The reasoning behind this is that in each location, the most available water source should be considered to represent the final score.

After obtaining the overall water score in this way for each scenario, we now combine the water score with the other layers. Again, we use a weighted average, as displayed in the image below.



Note that we weigh the water score with a factor 0.75, while the other two layers, slope and accessibility to cities, are each given a score of 0.125. The reasoning here is that these factors have less of a direct impact on the suitability for irrigation, and therefore should receive less weight in the overall score. Again, these weights are in line with those used in the IWMI study.

Using these rules for calculating the score, maps can be created with the scores for each of the scenarios. These are shown in the next chapter.

5. Areas suitable for irrigation – results

5.1. Suitable areas Mali

As described in chapter 4, the suitable areas for irrigation are determined based on five different input layers: a constraint layer, the slope, the groundwater availability, and the nearness to surface water. These input layers and their classification if applicable, are shown below.

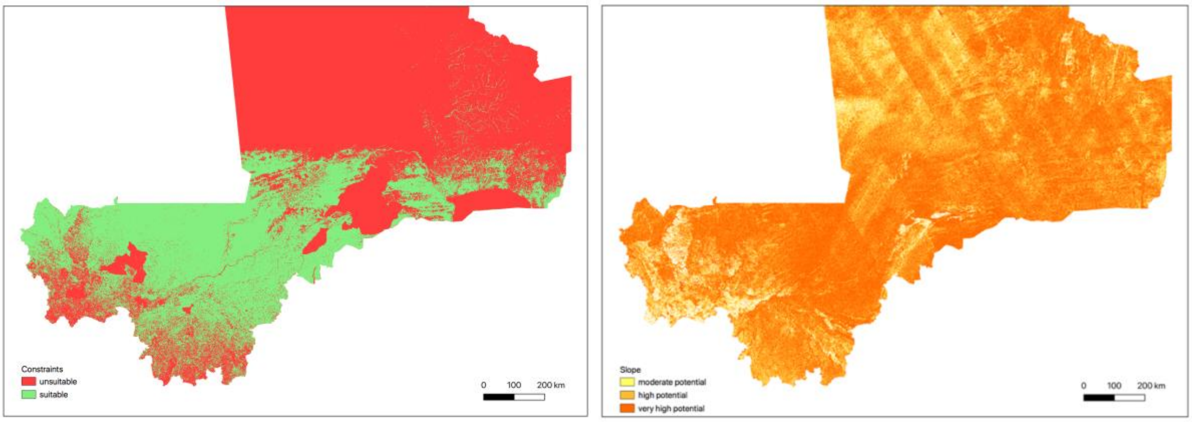


Figure 26. Inputs for the multicriteria scoring in Mali. Left: constraints based on slope, land use, national parks, and water bodies. Right: slope categories.

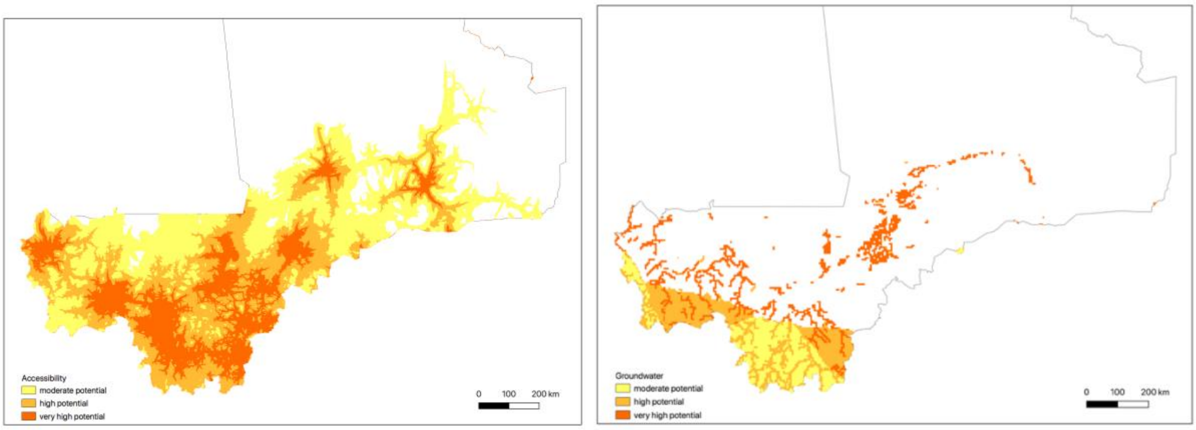


Figure 27. Inputs for the multicriteria scoring in Mali. Left: accessibility classes. Right: Groundwater classes.

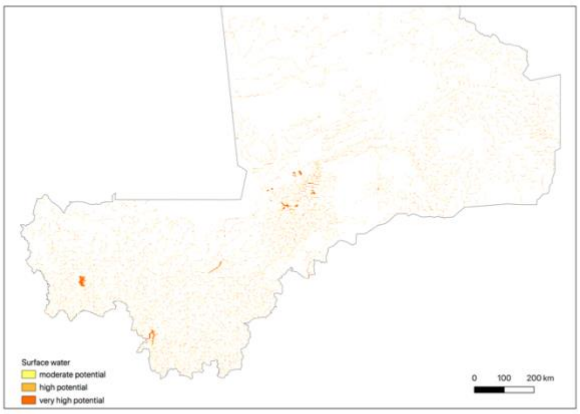


Figure 28. Inputs for the multicriteria scoring in Mali. Surface water classes.

There are a few observations we can make. First of all, in the constraint layer, we can see that the main restrictions are the national parks. Secondly, the graininess of the suitable layer is caused by the fact that we only select land that is currently classified as shrubs, herbaceous vegetation, or cultivated/managed vegetation.

Secondly, in the slope image some diagonal bands are visible, which are probably an artefact caused by the data NASA used to calculate the elevation model. This does not concern us, as the slope classification is almost always either highly suitable or very highly suitable, and so should not have a large effect on the end result.

In the accessibility class, we clearly see the cities, and the areas around them which can be reached within a reasonable time. Finally, for the groundwater layer we see that it is very blocky, which is caused by the low resolution of the data available (5000 meters).

When we combine these layers as described in chapter 4, we obtain the classifications for the scenario with surface water only, and the scenario with groundwater and surface water. These are shown below. As background, an OpenStreetMap layer is shown.

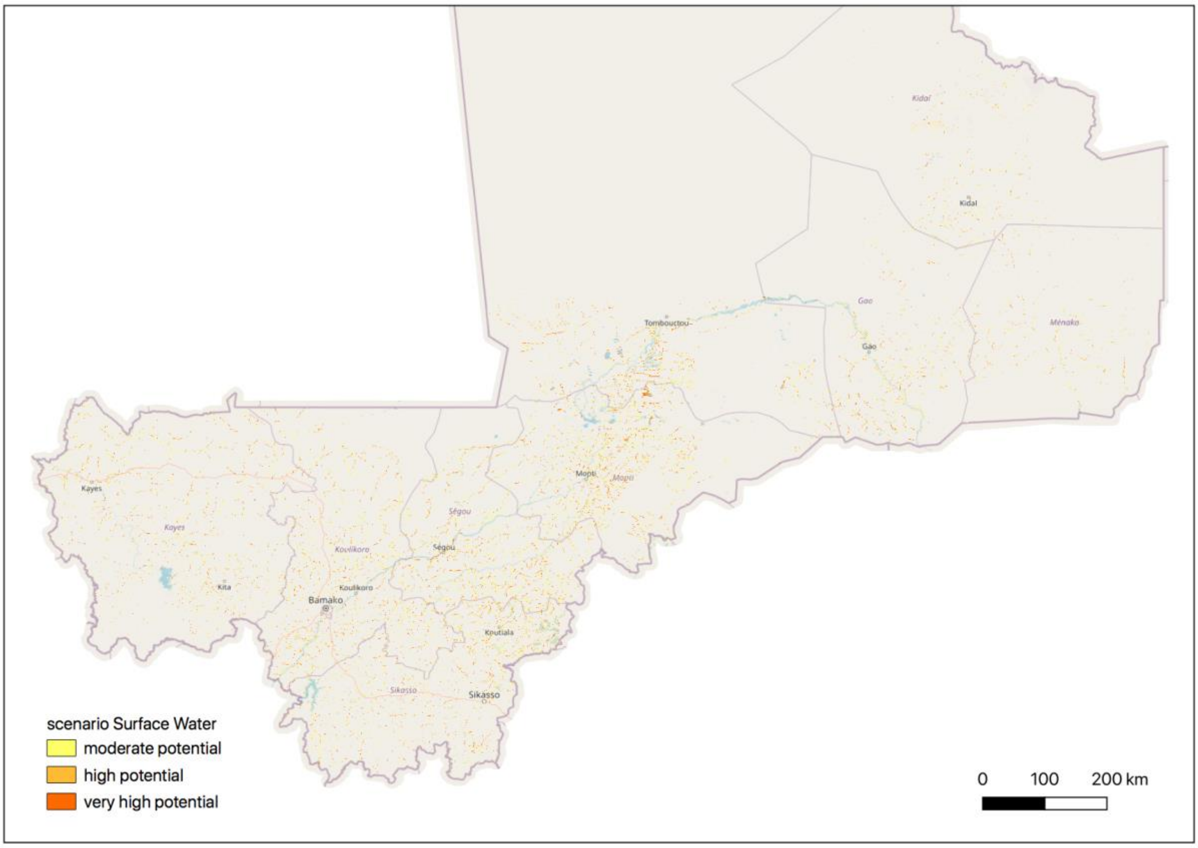


Figure 29. Suitable areas for irrigation in Mali. Scenario 1: only surface water is considered

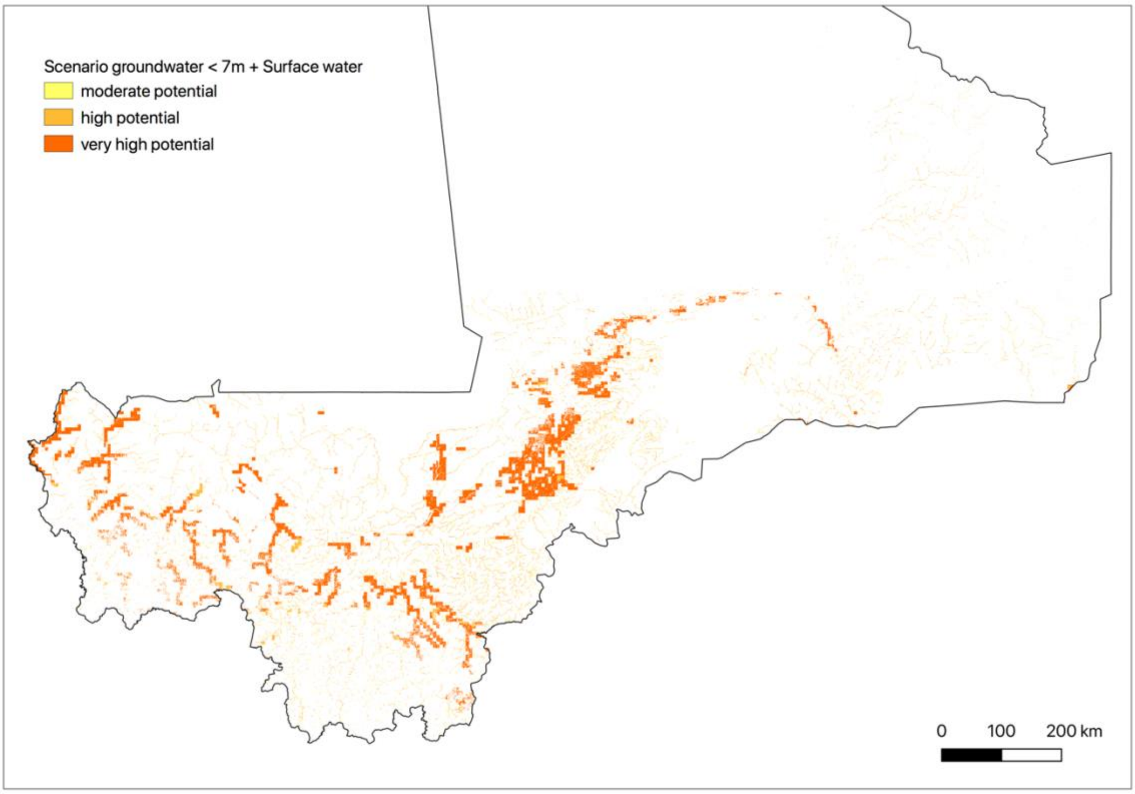


Figure 30. Suitable areas for irrigation in Mali. Scenario 2: both groundwater up to 7m and surface water are considered.

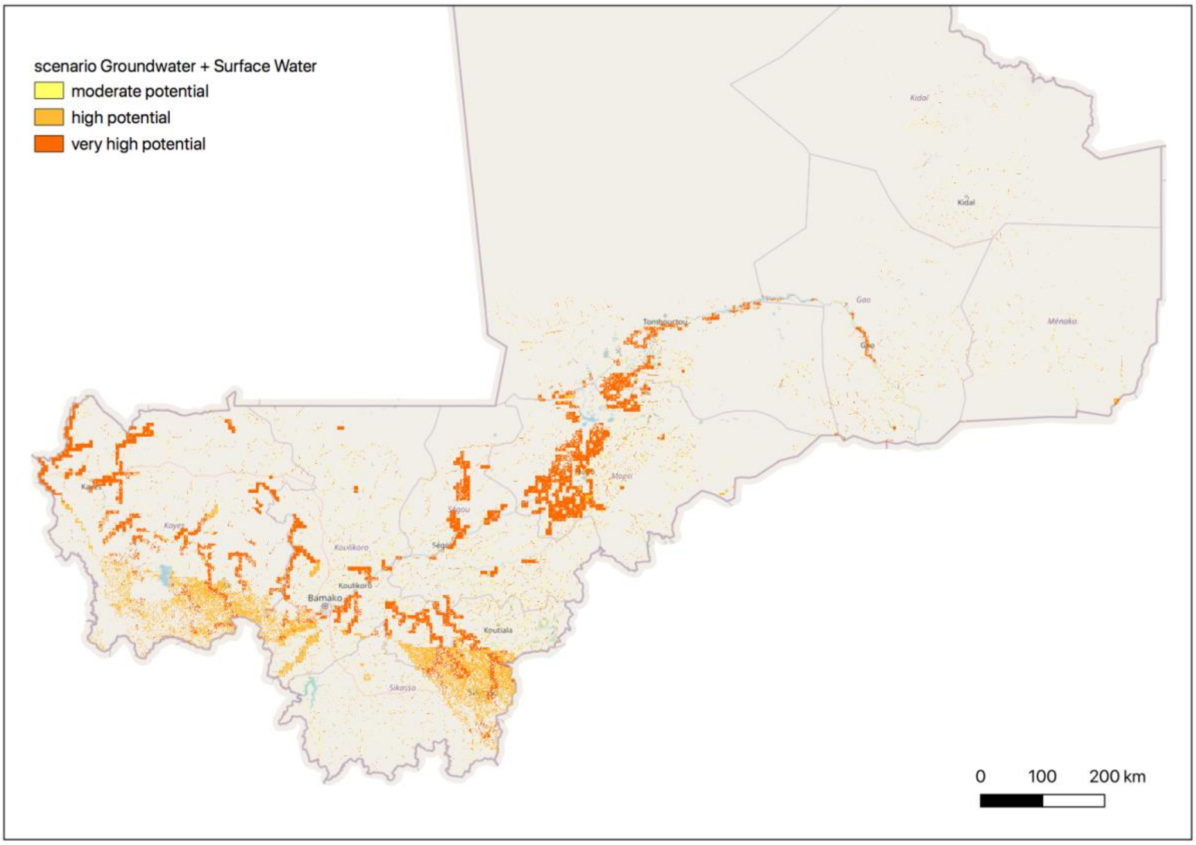
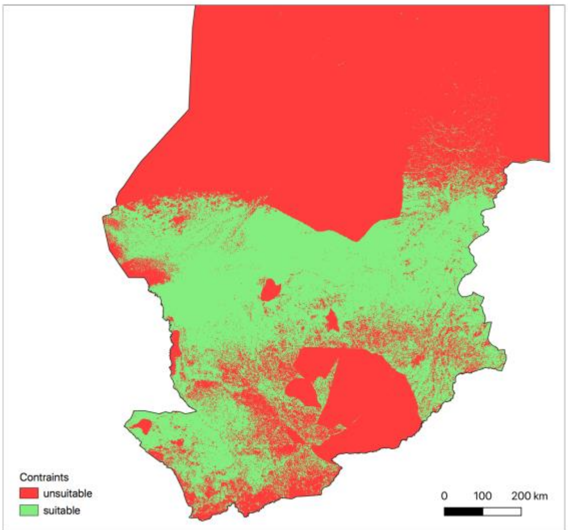


Figure 31. Suitable areas for irrigation in Mali. Scenario 3: both groundwater up to 25m and surface water are considered.

5.2. Suitable areas Chad



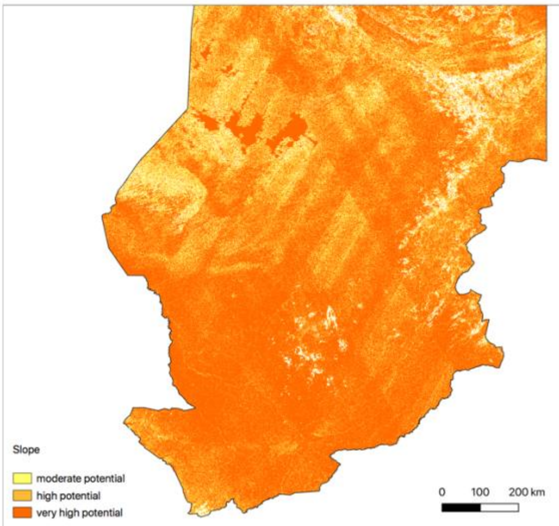
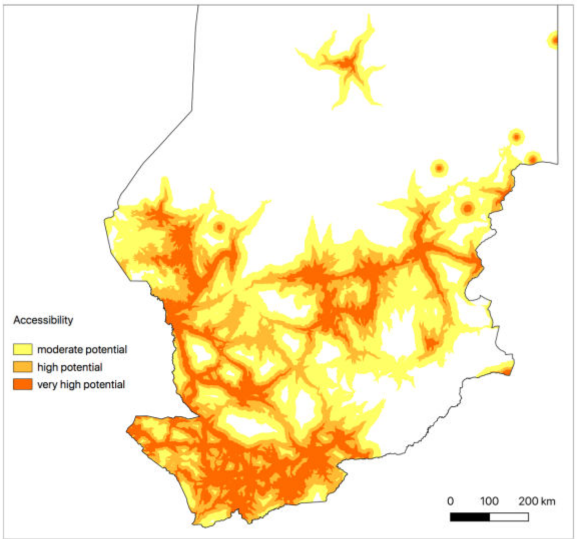


Figure 32. Inputs for the multicriteria scoring in Chad. Left: constraints based on slope, land use, national parks, and water bodies. Right: slope categories.



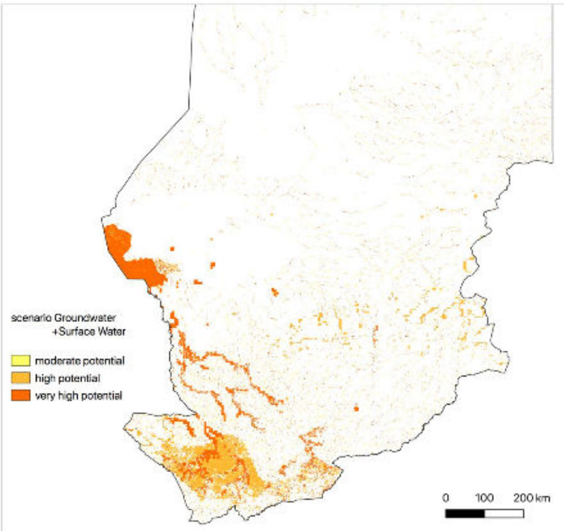


Figure 33. Inputs for the multicriteria scoring in Chad. Left: accessibility classes. Right: Groundwater classes.

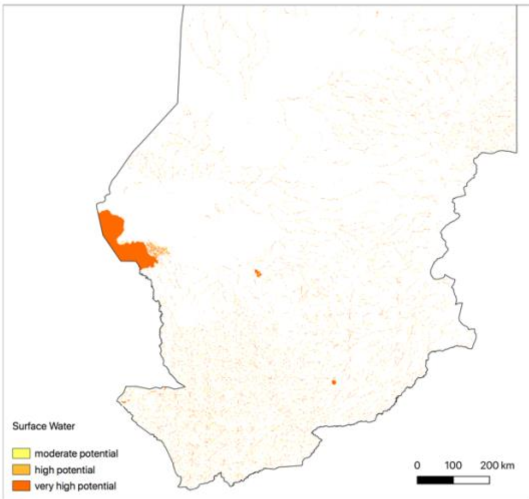


Figure 34. Inputs for the multicriteria scoring in Chad. Surface water classes.

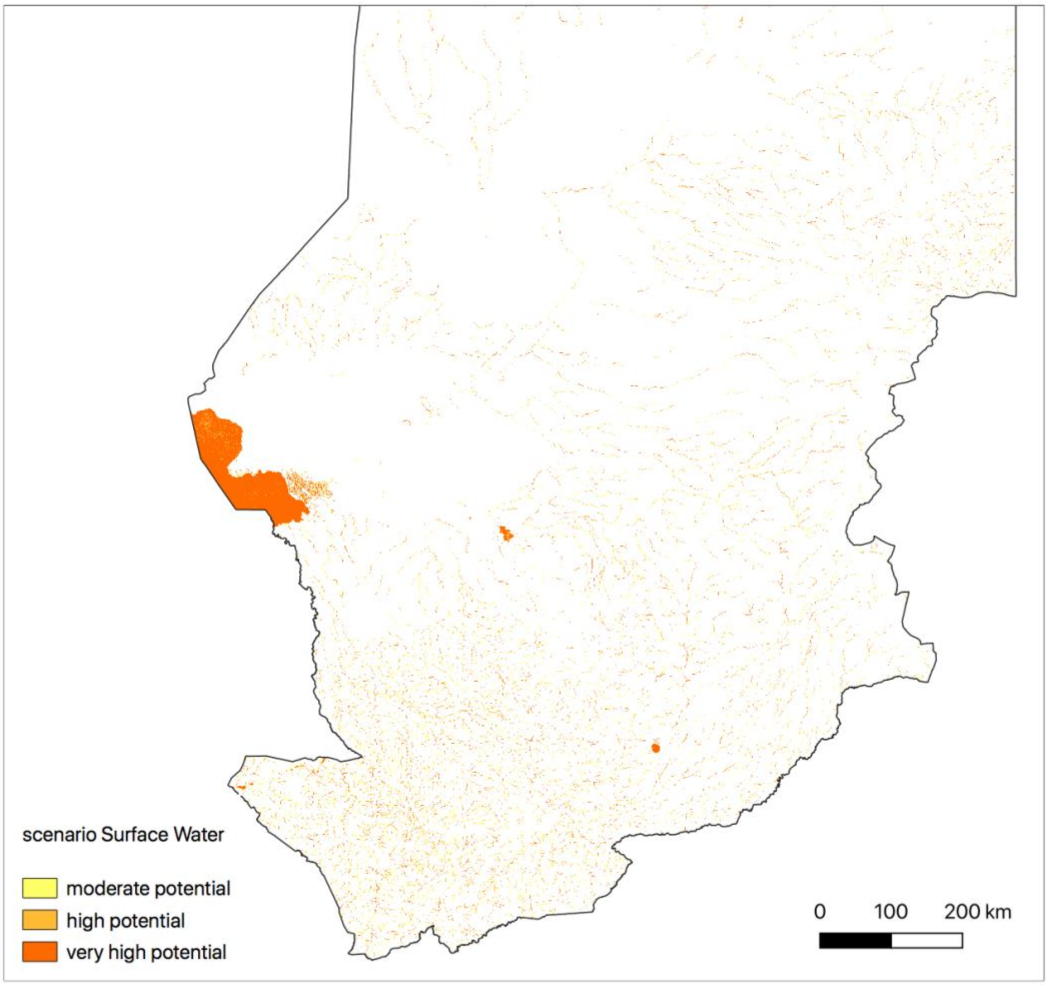


Figure 35. Suitable areas for irrigation in Chad. Scenario 1: only surface water is considered.

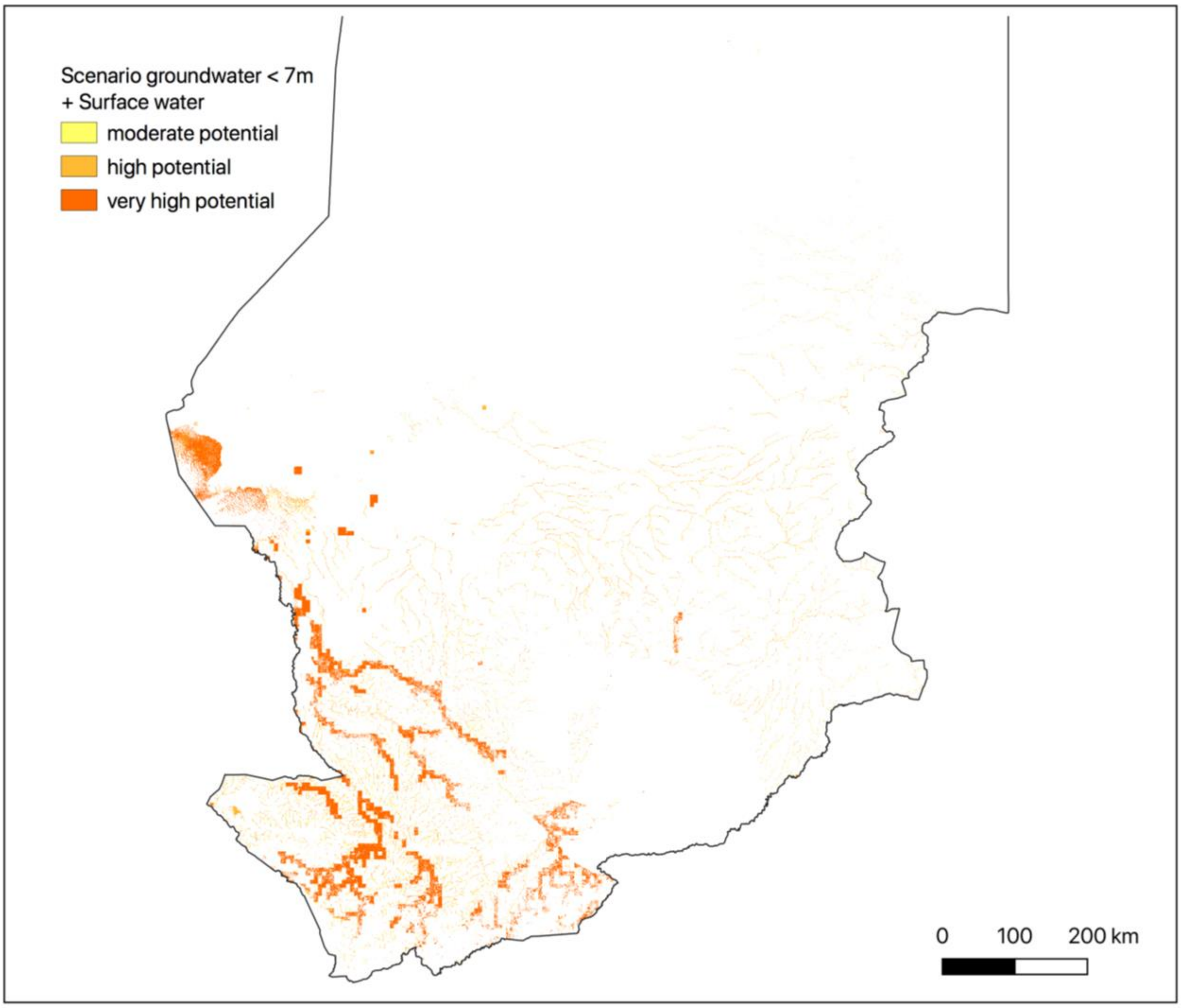


Figure 36. Suitable areas for irrigation in Chad. Scenario 2: both groundwater < 7m and surface water are considered.

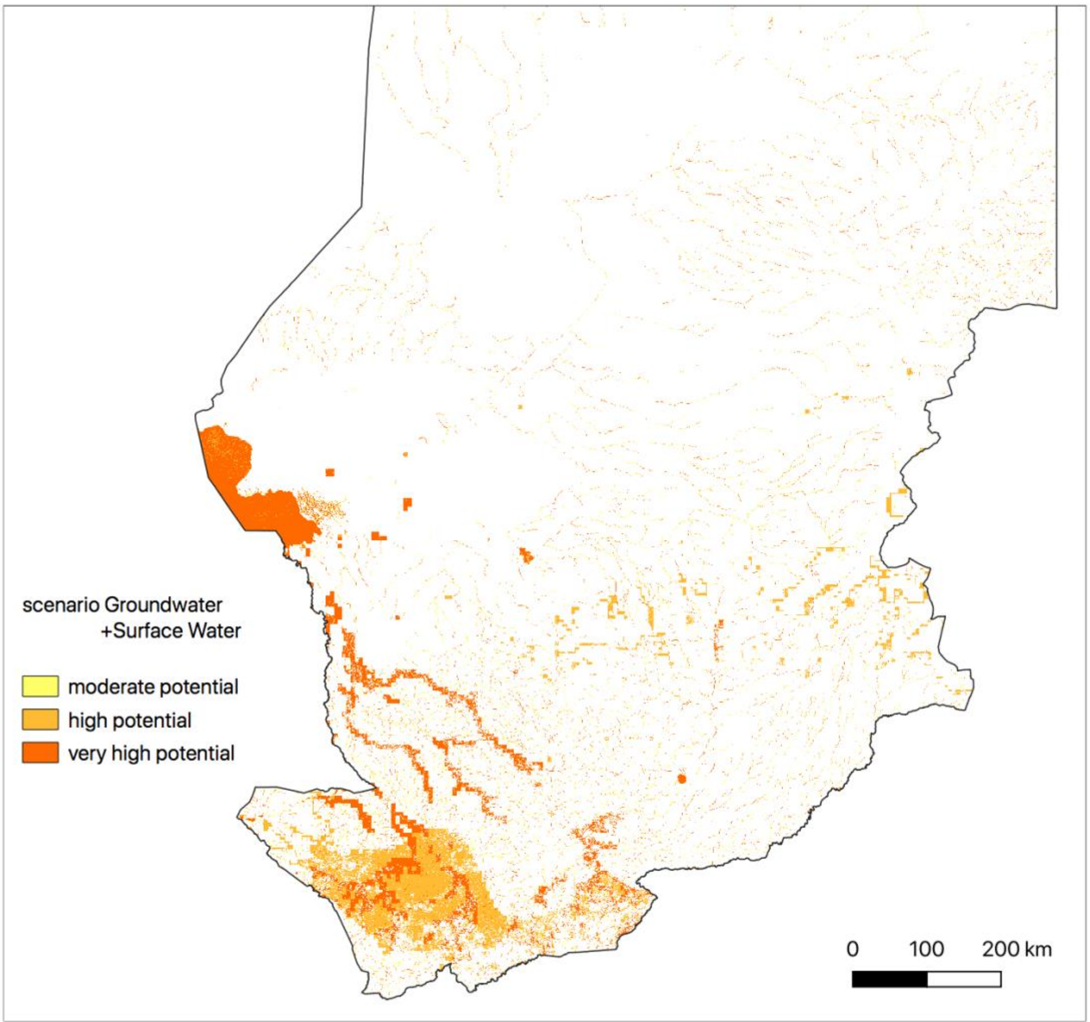


Figure 37. Suitable areas for irrigation in Chad. Scenario 3: both groundwater <25m and surface water are considered.

5.3. Quantitative description of the areas suitable for irrigation

In the tables below, we quantify the suitable areas for scenarios 1, 2 and 3.

Mali

Table 34. Regional breakdown of areas suitable for irrigation in Mali, using three scenarios.

Region	Total area 10^3 ha	Scenario 1 – surface water 10^3 ha	%	Scenario 2 – groundwater < 7m + surface water 10^3 ha	%	Scenario 3 – groundwater < 23m + surface water 10^3 ha	%
Bamako	25	0.14	0.56	0.18	0.72	0.18	0.72
Gao	17,057	224.94	1.32	260.46	1.53	267.95	1.57
Kayes	11,974	317.70	2.65	1,166.84	9.74	1,796.54	15.00
Kidal	15,145	81.47	0.54	80.15	0.53	81.47	0.54

Koulikoro	9,012	336.63	3.74	857.16	9.51	1,156.11	12.83
Mopti	7,902	361.33	4.57	1,131.09	14.31	1,153.15	14.59
Ségou	6,482	290.03	4.47	570.63	8.80	579.49	8.94
Sikasso	7,028	331.44	4.72	599.00	8.52	1,153.74	16.42
Timbuktu	49,611	212.45	0.43	498.37	1.00	503.38	1.01
Total	**124,236***	**2,156.12**	1.74	**5,163.88**	4.16	**6,692.02**	5.39

* Small deviations with respect to official numbers may be caused by geographical projection errors. They do not effect the results in a meaningful way.

Chad

Table 35. Regional breakdown of areas suitable for irrigation in Chad, using three scenarios.

Region	Total area	Scenario 1 – surface water 10^3 ha	%	Scenario 2 – groundwater < 7m + surface water 10^3 ha	%	Scenario 3 – groundwater < 23m + surface water 10^3 ha	%
	10^3 ha	10^3 ha				10^3 ha	
Barh el Ghazel	5,631	0.29	0.01	15.33	0.27	35.94	0.64
Batha	9,041	151.86	1.68	151.86	1.68	153.33	1.70
Borkou	25,623	0.71	0.00	0.71	0.00	0.71	0.00
Chari-Baguirmi	4,603	183.76	3.99	581.33	12.63	582.12	12.65
Ennedi Est	7,737	10.69	0.14	10.82	0.14	10.82	0.14
Ennedi Ouest	11,105	18.92	0.17	18.92	0.17	18.92	0.17
Guéra	6,104	120.02	1.97	121.83	2.00	283.25	4.64
Hadjer-Lamis	3,044	95.10	3.12	149.25	4.90	160.91	5.29
Kanem	6,767	1.65	0.02	13.54	0.20	13.54	0.20
Lac	1,984	369.69	18.63	372.12	18.76	372.12	18.76
Logone Occidental	881	70.90	8.05	279.05	31.67	749.44	85.07
Logone Oriental	2,372	122.31	5.16	306.81	12.93	878.88	37.05
Mandoul	1,744	72.95	4.18	148.83	8.53	244.47	14.02
Mayo-Kebbi Est	1,805	139.37	7.72	312.89	17.33	337.69	18.71
Mayo-Kebbi Ouest	1,254	62.98	5.02	119.82	9.56	348.82	27.82
Moyen-Chari	4,147	78.13	1.88	287.23	6.93	402.66	9.71
Ouaddaï	2,975	88.20	2.96	88.20	2.96	139.90	4.70
Salamat	6,796	97.16	1.43	114.61	1.69	158.57	2.33
Sila	3,570	72.45	2.03	72.76	2.04	300.35	8.41
Tandjilé	1,753	144.26	8.23	329.69	18.81	611.97	34.91
Tibesti	12,608	0.03	0.00	0.03	0.00	0.03	0.00
Ville de N'Djamena	40	0.99	2.48	8.05	20.13	8.05	20.13
Wadi Fira	5,412	62.18	1.15	62.18	1.15	81.03	1.50

Total	**126,996***	**1,979.65**	1.56	**3,586.47**	2.82	**5,893.51**	4.64

* Small deviations with respect to official numbers may be caused by geographical projection errors. They do not effect the results in a meaningful way.

The areas suitable for irrigation in the table above should not be understood as areas that can be completely realized. They should be understood as areas where irrigation expansion could be feasible. To determine the scope for actual large-scale expansion of irrigation in a given area, an analysis should be made on the catchment level, to determine the total amount of water that can be safely used for increased irrigation without negative impact on groundwater levels or river flow downstream. In addition, effects such as possible local market saturation should be investigated.

5.4. Web map

A static web map was created from the classified image and from the image of suitable areas. To do this, the GeoTiff image that was the result of the analysis was turned into a slippy web map, which is a GIS standard that can be used by GIS packages such as ArcGIS and QGIS and can be displayed on a web map. The links to the individual layers are listed in chapter 3.

The web map is available here:

Link to web map of Mali and Chad irrigated areas

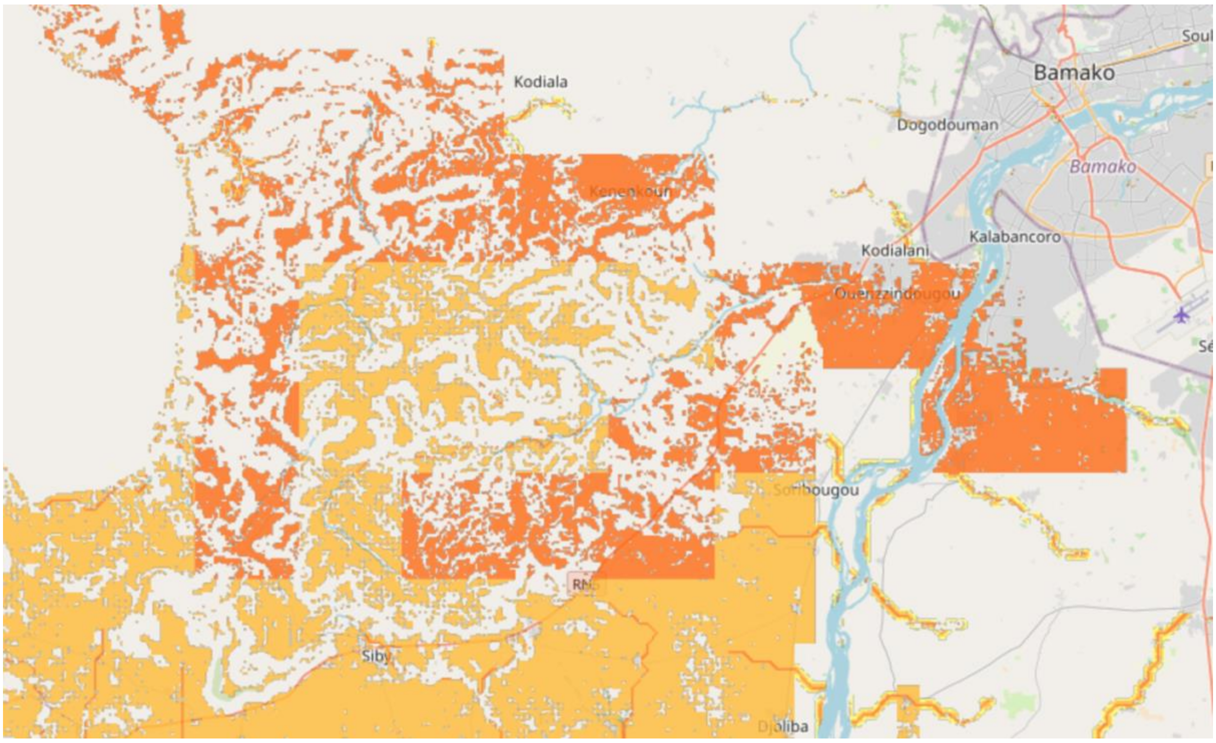


Figure 38. Web map, showing areas suitable for irrigation. The blocky structures are caused by the low resolution of the groundwater depth map.

5.5. Identifying areas with high potential for irrigation expansion

In this section, we combine the actually irrigated areas and areas potentially suitable for irrigation as determined in previous chapters, and identify zones of high potential. This is based on the notion that farmer-led irrigation generally develops in areas with existing irrigation activity, because of the available markets, inputs, knowledge and experience. An example to this is Beekman et al. (2014) showing how irrigation development takes place through expansion zones with existing farmer-led irrigation activity.

In the maps below, we superimpose the result of the middle scenario (surface water + groundwater < 7m), with the areas that are actually irrigated as determined in chapter 3 and 4. The rationale for using the middle scenario is that groundwater at a depth higher than 7 meter can be readily abstracted using affordable and available technologies. These areas therefore are the most suitable when expansion of irrigation is considered.

For both Mali and Chad, an important conclusion is that suitable areas for irrigation are either directly near surface water, or areas near rivers where the water table is expected to be above 7 meters. Existing small-scale irrigation already takes place along rivers, as is clear from the maps above. By using technologies such as low-cost manually drilled boreholes, the area near rivers where irrigation can be used could be expanded.

Mali

In the table below, the potentially suitable area for irrigation as determined in chapter 4 and 5 is compared with the actually irrigated area. Here we use scenario 2, which is the middle scenario that incorporates surface water and groundwater up to a depth of 7 meters.

Table 36. Comparison of potentially suitable areas (scenario 2) to actually irrigated areas in Mali.

Region	Total area 10^3 ha	Cropland extent (GFSAD30) 10^3 ha	Potentially suitable area (scenario 2) 10^3 ha	Actually irrigated area 10^3 ha	Irrigated area as percentage potentially suitable area
Bamako	25	3.79	0.18	0.09	50.0%
Gao	17,057	66.67	260.46	5.42	2.1%
Kayes	11,974	1,299.65	1,166.84	21.73	1.9%
Kidal	15,145	0.00	80.15	0.01	0.0%
Koulikoro	9,012	2,578.36	857.16	18.55	2.2%
Mopti	7,902	1,710.41	1,131.09	247.58	21.9%
Ségou	6,482	2,715.07	570.63	169.82	29.8%
Sikasso	7,028	2,273.58	599.00	32.51	5.4%
Timbuktu	49,611	63.69	498.37	69.93	14.0%
Total	**124,236***	**10,711.22**	**5,163.88**	**565.65**	**11.0%**

* Small deviations with respect to official numbers may be caused by geographical projection errors. They do not effect the results in a meaningful way.

As stated in section 4.1, the 'potentially suitable area' can significantly overestimate the area that could actually be sustainably irrigated. This might, for example, be the case in Kidal, where areas around smaller rivers are identified as potential for irrigation, but it is very unlikely that the total amount of water available could support substantial irrigation expansion. For a complete assessment of realizable irrigation potential, a full hydrological assessment is needed, with a calculation of water balance.

In the map of Mali below, the areas suitable for irrigation for the scenario surface water + groundwater < 7m are combined with the actually irrigated areas.

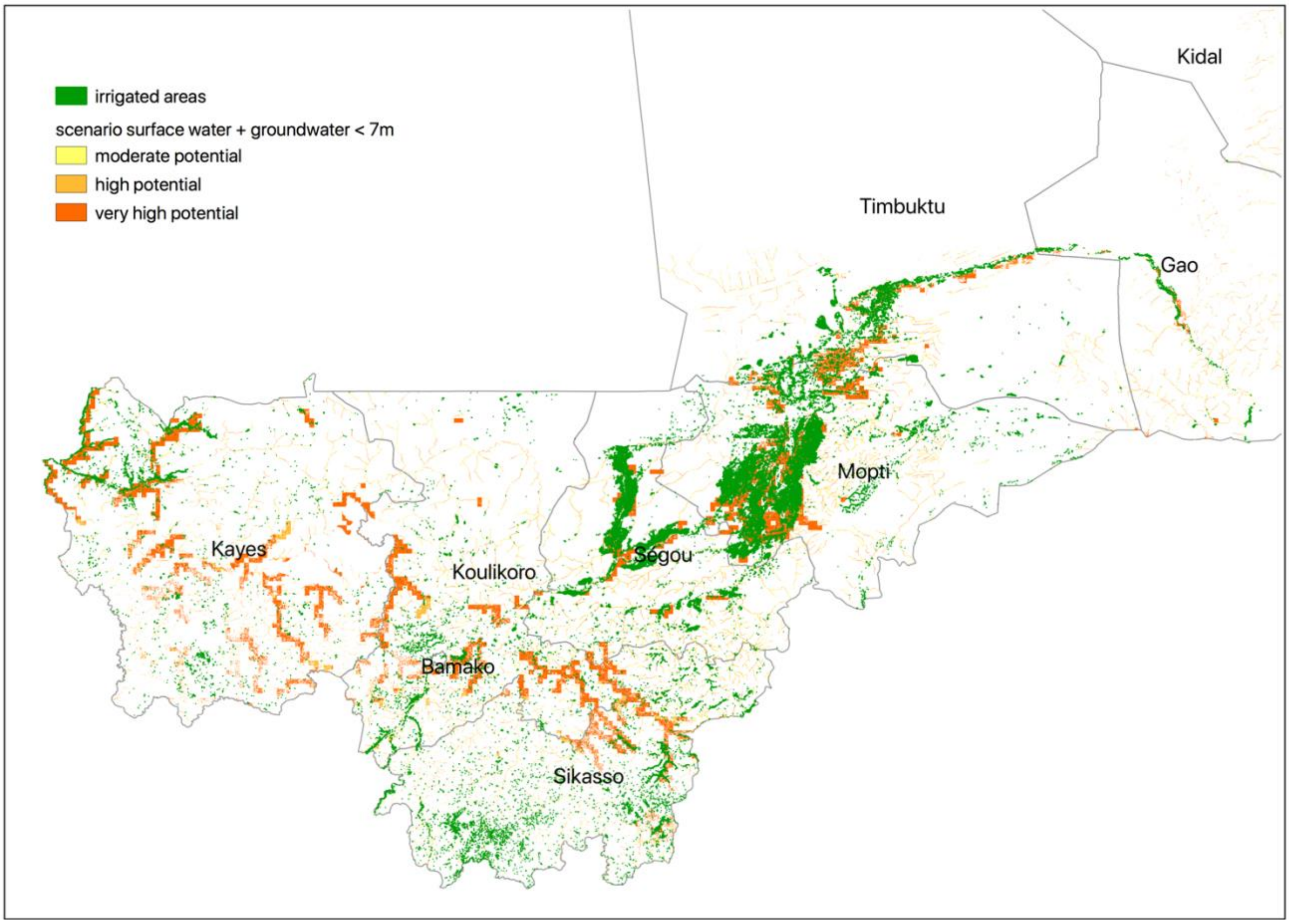


Figure 39. Irrigated areas in Mali (green). Yellow, orange, red: moderately, highly, and very highly suitable areas for irrigation. The irrigated areas have been enlarged for better visibility, which means that the apparent surface area on this map does NOT correspond to real surface area.

REGIONAL ZONES OF HIGH POTENTIAL

The regions identified below are already close to areas where irrigation is being practiced, and show potential in terms of suitability for surface water or shallow groundwater. Note that this analysis does not take into account hydrological boundaries on sustainable abstraction.

In the Kayes region, irrigation is practiced mainly in the North-East corner, along the Baoulé river, to the West of Kayes near Segala, in the North near Yaguine, and along the Eastern border river with Mauritania. These areas also show potential for shallow-groundwater use.

In Koulikoro region and Bamako region, the main area of interest is the region to the East of Bamako, around the Niger river and using shallow groundwater. To the South of Bamako, the lower parts of the Niger river have potential, for example the area around Sankarani River. To the West of Bamako the border area between the Koulikoro region and the Kayes region along the Baoule river has potential for shallow groundwater. In the South-West part of Koulikoro, the area around the Bani river.

In Sikasso region, interesting areas include the area near the Sankarani River near Gueleninkoro and Badiogo in the West, and the area to the South of Loulouni and North of Sikasso in the East.

In Ségou region, areas of interest are the Office du Niger, the area along the Niger river to the West of Ségou itself, and the extensive areas along the Niger river near Dioro and Kolongo-Tomo. In the South, the region along the Bani river.

In Mopti region, the whole area around the inner Niger delta is of interest, both in terms of surface water and for shallow groundwater. To the South, the area between Mougna and Kandara. Towards the North, the branches of the Niger are interesting because of the proximity to surface water, and above Gathi-Loumo because of shallow groundwater.

In Timbuktu region, recession agriculture is possible (and practiced) in the many lakes, for example near Dianké, Soumpi and Lake Oro and Fati. Along the many branches of river Niger, including near Timbuktu itself, irrigation is practiced extensively, and expansion could be possible by using shallow groundwater.

Finally, in Gao region, irrigation on a small scale is possible near the Niger river, for example near Ouatagouna.

The above regions are only a rough indication, and the maps can be further explored in detail on the online map, which allows zooming in to a region of interest.

CHAD

In the table below, the potentially suitable area for irrigation as determined in chapter 4 and 5 is compared with the actually irrigated area. Here we use scenario 2, which is the middle scenario that incorporates surface water and groundwater up to a depth of 7 meters.

Table 37. Comparison of potentially suitable areas (scenario 2) to actually irrigated areas in Chad.

Region	Total area	Cropland extent (GFSAD30)	Potentially suitable area (scenario 2) 10^3 ha	Actually irrigated area	Irrigated area as percentage potentially suitable area
	10^3 ha	10^3 ha		10^3 ha	
Barh el Ghazel	5,631	0.12	15.33	0.10	0.7%
Batha	9,041	249.61	151.86	9.19	6.1%
Borkou	25,623	0.00	0.71	0.00	0.0%
Chari-Baguirmi	4,603	318.40	581.33	2.75	0.5%
Ennedi Est	7,737	0.00	10.82	0.00	0.0%
Ennedi Ouest	11,105	0.00	18.92	0.06	0.3%
Guéra	6,104	38.82	121.83	1.18	1.0%
Hadjer-Lamis	3,044	239.81	149.25	6.83	4.6%

Kanem	6,767	0.85	13.54	0.22	1.6%
Lac	1,984	3.92	372.12	11.51	3.1%
Logone Occidental	881	663.31	279.05	0.29	0.1%
Logone Oriental	2,372	808.65	306.81	0.48	0.2%
Mandoul	1,744	581.27	148.83	0.33	0.2%
Mayo-Kebbi Est	1,805	406.51	312.89	3.70	1.2%
Mayo-Kebbi Ouest	1,254	398.50	119.82	0.26	0.2%
Moyen-Chari	4,147	386.62	287.23	6.25	2.2%
Ouaddaï	2,975	291.66	88.20	18.27	20.7%
Salamat	6,796	67.78	114.61	15.54	13.6%
Sila	3,570	329.24	72.76	18.26	25.1%
Tandjilé	1,753	587.65	329.69	0.61	0.2%
Tibesti	12,608	0.00	0.03	0.00	0.0%
Ville de N'Djamena	40	2.07	8.05	1.41	17.5%
Wadi Fira	5,412	2.28	62.18	7.47	12.0%
Total	**126,996***	5377.05	**3,586.47**	**104.72**	**2.9%**

* Small deviations with respect to official numbers may be caused by geographical projection errors. They do not effect the results in a meaningful way.

In the map of Chad below, the areas suitable for irrigation for the scenario surface water + groundwater < 7m are combined with the actually irrigated areas.

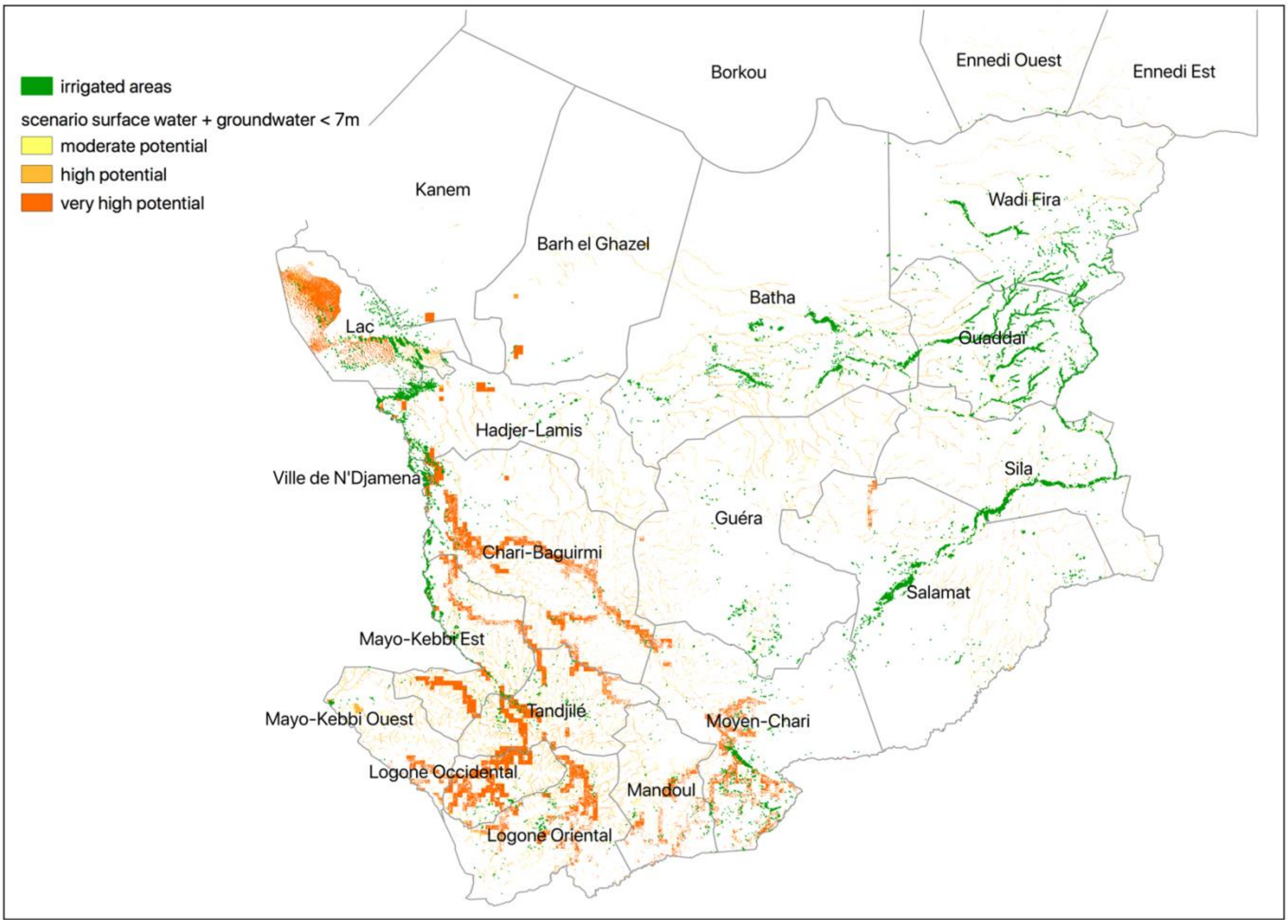


Figure 40. Irrigated areas in Chad (green). Yellow, orange, red: moderately, highly, and very highly suitable areas for irrigation. The irrigated areas have been enlarged for better visibility, which means that the apparent surface area on this map does NOT correspond to real surface area.

REGIONAL ZONES OF HIGH POTENTIAL

The regions identified below are already close to areas where irrigation is being practiced, and show potential in terms of suitability for surface water or shallow groundwater. Note that this analysis does not take into account hydrological boundaries on sustainable abstraction.

As is the case in Mali, the regions of interest for irrigation closely follow the surface water areas. The map below shows the drainage basin of the Chari river.

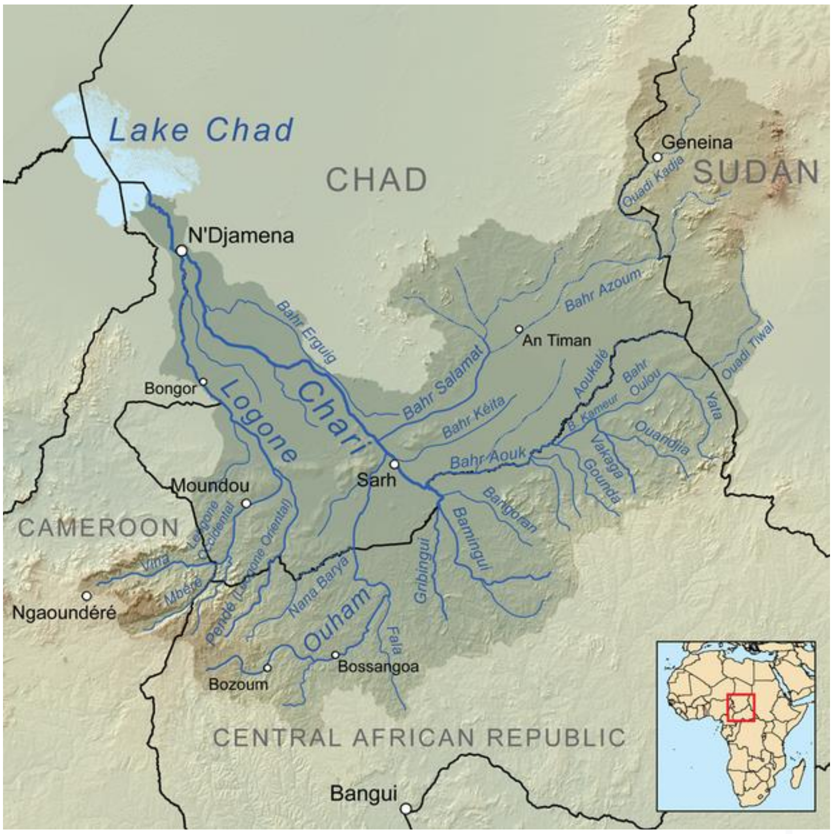


Figure 41. The Chari River drainage basin. Source: Wikipedia.

In the East of the country irrigation takes place along the many rivers in the Ouaddai region, the Azoum and Salamat river, and towards the North near Wadi Fera. In the West and South, the map indicates scope for shallow groundwater irrigation, and expansion of surface water irrigation, for example near Sahr in the Moyen-Chari region.

In the North, the area below Lake Chad and towards Ville de N'Djamena. Along the Logone and Chari river, potential for surface irrigation and shallow groundwater is high. In the South, the Logone river and tributary to the Chari river also form the most promising areas.

These maps can be further explored on the online map, which allows zooming in to a region of interest.

6. DISCUSSION

We determined irrigated areas in Mali and Chad in the period Oct 2019 – June 2020, using Sentinel 2 satellite data and machine learning, at a 30m resolution. The Random Forest model used showed a good accuracy, with the overall accuracy at 95.6% for Mali and 95.5% for Chad. The most important bands turned out to be the monthly averages of the EVI and NDWI indices, and their monthly differences.

In addition, we have determined areas with a high potential suitability for new irrigation. To do this, we combine local properties such as nearness to surface water, groundwater depth and aquifer properties, protected areas and national parks, accessibility to cities, and land use. These properties were combined into an overall suitability score using a weighted average. We created three scenarios, one only considering surface water, the other two combining surface water and groundwater at different depths.

From Chapter 3 and 5, our estimate for currently irrigated area in the dry season in Mali is 565.6 kha, and our estimate for the areas suitable for irrigation expansion, rounded to significant numbers and depending on the scenario, is 2200 – 6700 kha. For Chad, our estimate for currently irrigated area in the dry season is 104.7 kha, and our estimate for the areas suitable for irrigation expansion, depending on the scenario is 2000 – 5900 kha. As noted before, the result for irrigated area should not directly be compared to the suitable areas due to the fact that the latter does not incorporate hydrological constraints. In this chapter, we compare our results to those of other studies.

6.1 COMPARISON WITH OTHER STUDIES

In the tables below, we bring together a number of studies on both the area under irrigation and estimates for either total irrigation potential or suitable areas for irrigation expansion. The methodologies for determining the irrigation potential are quite different and are calculated for different years, so the numbers should be interpreted with care. For example, some studies (Pavelic, Altchenko, You, Africa Dryland report) take into account the hydrological situation and water balance. Some, such as FAO (and this report) do not do this, which explains the large difference in size of numbers.

For the actually irrigated areas, we only quote the AQUASTAT figures as other studies known to us repeat these statistics, so no new information is obtained. The year to which the data applies is given between brackets before the actual figures.

Mali – Actually irrigated areas

Table 38. Comparison of our results for irrigated areas with FAO statistics in Mali.

Reference	Equipped for irrigation (10^3 ha)	Equipped – actually irrigated (10^3 ha)	Total water agricultural water managed area (10^3 ha)	Irrigated area in the dry season (10^3 ha)
FAO AQUASTAT	(2011) 371.1		(2011) 621.3	
FAO AQUASTAT	(2000) 235.8	(2000) 175.8	(2000) 386.1	
FAO Mali	(2013) 380.0			
This report				**(2019) 565.6**

Mali – Irrigation potential / Suitable areas

Table 39. Comparison of our results for suitable areas for irrigation in Mali with other studies.

Reference	Total Irrigation potential (10^3 ha)	Comments
FAO AQUASTAT	(2013) 566	Surface + ground water, hydr. constraints
You (2010) – IFPRI	*802	Surface + ground water, hydr. constraints
WB / Ward (2016) – IFPRI	**652	Surface + ground water, hydr. constraints
Pavelic (2013)	***1054-1868	groundwater only, hydr. constraints
Altchenko (2014) – IWMI	331-787	groundwater only, hydr. constraints
IWMI (2019)	685 - 4435	Surface + ground water, suitable areas
This report	**2156 – 6692**	**Surface + ground water, suitable areas**

* You (2010, table 8) only report potential *increases* in irrigated area (491 kha). To calculate this figure, we assume You uses the AQUASTAT data in 2010 as reference.
** Ward (2016) only reports potential increases in irrigated area (281 kha). To calculate this figure, we assume Ward uses the AQUASTAT data in 2016 as reference.
*** We use the 500mm/y scenario, with 50 or 70% environmental groundwater requirement (Pavelic 2013, table 3)

Chad – Actually irrigated areas

Table 40. Comparison of our results for irrigated areas with FAO statistics in Chad.

Reference	Equipped for irrigation (10^3 ha)	Equipped – actually irrigated (10^3 ha)	Total agricultural water managed area (10^3 ha)	Irrigated area in the dry season (10^3 ha)
FAO AQUASTAT	(2002) 30.3	(2002) 26.2	(2002) 155.3	
This report				**(2019) 104.7**

Chad – Irrigation potential / Suitable areas

Table 41. Comparison of our results for suitable areas for irrigation in Chad with other studies.

Reference	Total Irrigation potential (10^3 ha)	Comments
FAO AQUASTAT	(2013) 335	Surface + ground water, hydr. constraints
You (2010) - IFRPRI	*365	Surface + ground water, hydr. constraints
WB / Ward (2016)– IFPRI	**295	Surface + ground water, hydr. constraints
Altchenko (2014) – IWMI	237-566	groundwater only, hydr. constraints
This report	**1979 – 5893**	**Surface + ground water, suitable areas**

* You (2010, table 8) only report potential *increases* in irrigated area (277 kha). To calculate this figure, we assume You uses the AQUASTAT data in 2010 as reference.
** Ward (2016) only reports potential increases in irrigated area (265 kha). To calculate this figure, we assume Ward uses the AQUASTAT data in 2016 as reference.

Areas under irrigation in the dry season

Our estimate of areas under irrigation in the dry season for Mali is 565.6 kha, which is in line with the most recent estimate from AQUASTAT (621.3 kha) for the total agricultural area under water management, which includes flood recession cropping areas and cultivated wetlands and inland valley bottoms (as explained in section 2.2). Of this area, 338.7 kha is larger than 500 ha, 82.6 kha is between 100 – 500 ha, and 144.3 kha is smaller than 100 ha. In Mali, the area that is equipped for irrigation is 59.7% of the total area under water management according to AQUASTAT, showing that flood recession, cultivated wetlands and valley bottoms play an important role in the total irrigated area.

Our estimate for Chad is 104.7 kha, significantly less than the recent estimate from AQUASTAT (155.3 kha). Of this area, 33.9 kha is larger than 500 ha, 20.2 kha is between 100 – 500 ha, and 50.6 kha is smaller than 100 ha. In Chad, the area that is equipped for irrigation is only 19.5% of the total area under water management according to AQUASTAT, showing that the role of flood recession, cultivated wetlands and valley bottoms, and possibly irrigation in the wet season, plays an even greater role in Chad.

In the case of Mali, our estimate is 9% lower than the official estimates. For Chad, our estimate is 33% less than the official estimate. There are several possible causes for this: limitations of our machine learning method such as an excluding wet season irrigation, a possible overestimation of the official numbers, or the fact that not all areas identified as under water management are actually involved in irrigation in a given year. In addition, the numbers refer to different years.

Potential areas / suitable areas

As stated in our introduction, a distinction should be made between ‘areas suitable for irrigation’, which means areas where irrigation could be practiced, and ‘total irrigation potential’, which means the fully realizable area, given hydrological constraints. The studies we found that have estimates for these numbers vary considerably in their methodology — with some considering both ground and surface water, and others considering only groundwater — explaining the large variation in numbers in the tables.

Depending on the scenario used, our estimate for areas suitable for irrigation expansion in Mali varies between 2200 – 6700 kha, which is in line with a recent study with a similar approach (IWMI 2019). In the case of Mali, studies that incorporate the hydrological constraint indicate that only about 30% of the *suitable* area can actually be used for irrigation in a sustainable way. For Chad, our estimates for suitable areas depending on the scenario vary between 2000 – 5900 kha. Studies that incorporate the hydrological constraint indicate that in Chad, only about 17% , of the suitable area can actually be used for irrigation in a sustainable way.

This means that in both countries, the hydrological component is essential for estimates of the total realizable irrigation potential. The suitable areas as described in this report should therefore be used only to identify overall suitable regions and guide site selection.

Interestingly, the actual irrigated area for Mali already is close to the ranges quoted in the various studies. This would seem to indicate that Mali is already close to its irrigation potential from a hydrological limits point of view. To gain more confidence, it will be needed to do a more in-depth study that incorporates details of hydrological limits. For example, both You and Altchenko use a percentage for the groundwater requirements for environmental needs that spans 30%-70%, which accounts for the large spread in numbers.

For Chad, there is at least a factor of 2 difference between the current irrigated area and the estimates for the overall irrigation potential with hydrological limits taken into account. This indicates that in Chad, there is still ample scope for increased irrigation.

Additional information on the references

You (2010) — Estimation of the development potential of irrigation from a multi-criteria analysis based on agronomic, hydrological and economic criteria (surface water and groundwater combined).

Pavelic (2013) — Estimation of the development potential of small-scale irrigation (SSI) from groundwater from a hydrological balance taking into account recharge, AEP and livestock uses, as well as environmental needs.

Altchenko (2014) — Mapping of the potential for irrigation from renewable groundwater in Africa from a water balance taking into account recharge, water supply and livestock uses, as well as environmental needs.

WB / Ward (2016) — IFPRI Estimation of the development potential of irrigation in the "Drylands" arid zone (including groundwater) from a multi-criteria analysis based on geographical criteria; agronomic, hydrological and hydrogeological, social, economic, and rural development (transport network, markets).

Aquastat Mali — http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/irrigation-by-country/country/MLI

Aquastat Chad — http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/irrigation-by-country/country/TCD

FAO Mali — Country brief mali, http://www.fao.org/3/a-i7617e.pdf

Comparison with FAO AQUASTAT statistics

In this section, we give an overview of AQUASTAT statistics for both countries, and compare our results against them. For easy reference, the FAO definitions of the statistics used are available in the Appendix.

Mali AQUASTAT statistics

FAO statistics for arable area, irrigation potential, areas equipped for irrigation, and total agricultural water managed area. (FAO 2016)

Table 42. Irrigation AQUASTAT statistics for Mali.

FAO Statistic	Year	Area (10^3 ha)
Arable land area	2016	6,411
Permanent crops area	2016	150
Cultivated area (arable land + permanent crops)	**2016**	**6,561**
Irrigation potential	2013	566
Area equipped for full control irrigation: total	2011	167.1
Area equipped for full control irrigation: actually irrigated	2011	139.9
Area equipped for irrigation: equipped lowland areas	2011	204
Area equipped for irrigation: spate irrigation		
Area equipped for irrigation: total	**2011**	**371.1**
Area equipped for irrigation: actually irrigated	2000	175.8
Flood recession cropping area non-equipped	2009	250.2
Cultivated wetlands and inland valley bottoms non-equipped	1994	3.8
Total agricultural water managed area (1000 ha)	**2011**	**621.3**

Mali area equipped for irrigation FAO statistics – sub national

The figures refer to the year 2000. Note: the totals in the table above refer to different years. For example, the total area equipped for irrigation in 2000 is given as 235 kha, and in 2009 given as 371.1. (FAO 2016)

Table 43. Regional breakdown of FAO irrigation statistics for Mali.

Region	Total area equipped for irrigation (year: 2000) (10^3 ha)
Gao	14.4
Kayes	2.4
Kidal	0
Koulikoro	23.5
Mopti	50.7
Segou	97.6
Sikasso	11.4
Tombouctu	35.7
Mali total	**235.8**
with groundwater	1.0
with surface water	234.8
Area equipped for full control irrigation	97.5
Equipped lowland areas	138.3

Comparison

The comparison of our results to AQUASTAT statistics is hampered by the fact that the AQUASTAT numbers are only available for certain years. In particular, the actually irrigated area for Mali is only given for the year 2000. Therefore, in the table below we give two different figures: one including the actual irrigated area in 2000, and one using the total area equipped for irrigation in 2011. In both cases, we add Flood recession cropping areas and Cultivated wetlands and inland valley bottoms.

Table 44. Comparison of our results to FAO statistics for Mali.

	Total area (10^3 ha)
Area equipped for irrigation: actually irrigated (2000) + Flood recession cropping area non-equipped (2009) + Cultivated wetlands and inland valley bottoms non-equipped (1994)	429.8
Area equipped for irrigation: total (2011) + Flood recession cropping area non-equipped (2009) + Cultivated wetlands and inland valley bottoms non-equipped (1994)	625.1
Total agricultural water managed area (2011)	621.3
This report: total area under irrigation in the dry season (2019) **Size breakdown:** **< 100 ha: 144.3 kha** **100-500 ha: 82.6 kha** **> 500 ha: 338.7 kha**	**565.6**

From the table, it is clear that our result is in line with AQUASTAT statistics.

Chad AQUASTAT statistics

FAO statistics for arable area, irrigation potential, areas equipped for irrigation, and total agricultural water managed area. (FAO 2016)

Table 45. Irrigation AQUASTAT statistics for Chad.

FAO Statistic	**Year**	**Area (10^3 ha)**
Arable land area	2016	4,900
Permanent crops area	2016	35
Cultivated area (arable land + permanent crops)	**2016**	**4,935**
Irrigation potential	2013	335
Area equipped for full control irrigation: total	2002	30.27
Area equipped for full control irrigation: actually irrigated	2002	26.2
Area equipped for irrigation: equipped lowland areas		
Area equipped for irrigation: spate irrigation		
Area equipped for irrigation: total	**2002**	**30.27**
Area equipped for irrigation: actually irrigated	2002	26.2

Flood recession cropping area non-equipped	2002	125
Cultivated wetlands and inland valley bottoms non-equipped	1988	21.4
Total agricultural water managed area (1000 ha)	**2002**	**155.3**

Chad area equipped for irrigation FAO statistics – sub national
The figures refer to the year 2002. (FAO 2016)

Table 46. Regional breakdown of FAO irrigation statistics for Chad.

Region	**Total area equipped for irrigation (year: 2002) (10^3 ha)**
Batha	0.3
Bilthine	0
Bourkou Ennedi Tibesti (BET)	2.5
Chari-Baguirmi	2.58
Guera	0.13
Kanem	0.54
Lac	9.05
Logone-Occidental	0
Logone-Oriental	0.25
Mayo-Kebbi	3.66
Moyen-Chari	3.7
Ouaddai	5.46
Salamat	0
Tandjile	2.1
Chad total	**30.27**
with groundwater	6
with surface water	24.27

Comparison
Comparison with AQUASTAT statistics:

Table 47. Comparison of our results to FAO statistics for Chad.

	Total area (10^3 ha)
Area equipped for irrigation: actually irrigated (2002) + Flood recession cropping area non-equipped (2002) + Cultivated wetlands and inland valley bottoms non-equipped (1988)	172.6
Area equipped for irrigation: total (2002) + Flood recession cropping area non-equipped (2002) + Cultivated wetlands and inland valley bottoms non-equipped (1988)	176.7
Total agricultural water managed area (2002)	155.3
This report: total area under irrigation in the dry season (2019)	**104.7**

Size breakdown: < 100 ha: 50.6 kha 100-500 ha: 20.2 kha > 500 ha: 33.9 kha	

From the table, we see that our total result is significantly lower than the AQUASTAT statistics. The largest share to the total agricultural water managed area is the flood recession cropping area, with 125 kha. It therefore seems likely that it is either the case that our method does not accurately capture flood recession agriculture, or that the AQUASTAT numbers overestimate the actual flood recession area used for irrigation.

6.2. Limitations of Remote sensing analysis

During this work, we encountered a number of limitations that we estimate will affect any work of this type. These limitations are listed below.

Distinction Farmer-led and Government-led irrigation

The machine learning algorithm we employed is pixel-based, meaning that it uses spectral information of each image pixel for the classification into irrigated or not-irrigated. Therefore, information on field size is not obtained — one of the most important indicators of whether irrigation is farmer-led or not. In addition, the variation in degree of clustering and nature of irrigation in flood recession, cultivated wetlands, and valley bottoms means that remote sensing data alone does not contain enough information to specifically identify farmer-led irrigation as a category.

Therefore, other geospatial information on the location, management type and nature of different types of irrigation — such as large-scale irrigation schemes and flood recession areas — is needed. In this study, we successfully used the size of patches of contiguous pixels classified as irrigated to distinguish between small, medium, and large-scale irrigation.

Irrigation during rainy season

During the rainy season, data from satellites can be unavailable for months due to cloud cover. Therefore, it is often not possible to obtain any high-quality data during the rainy season, which was the reason that we did not consider this period in this study. This means that a number of irrigation practices that happen during the rainy season, such as supplementary irrigation, cannot be determined using this method. Although satellite sources were tried that are not sensitive to cloud cover (Sentinel-1 Synthetic Aperture Radar) it was found to not be sufficient to be used on its own. Therefore, irrigation can only be identified in period that stretches from the end of the rainy season up to the start of the next rainy season.

Receding water agriculture

In both Chad and Mali, a lot of agriculture revolves around water bodies (both rivers and lakes) that fill up during the rainy season, and then show receding water levels. This type of agriculture is hard to detect using the method applied in this report, because other vegetation growth also follows the receding water. In addition, it is unclear if this type of irrigation should be classified as ‘irrigation’.

LIMITATIONS IN THE AVAILABILITY OF HIGH-RESOLUTION IMAGERY

Part of the ground truth data collection process relies on high-resolution Google Maps. However, in many cases this imagery can be dated. In the case of Mali, most of the imagery is recent, in most cases from 2019. However, in the case of Chad, a significant part of the imagery is older, with many images stemming from 2012. This means that in the case of Chad, the ground truth data might be less reliable.

In the present study the reliance on high-resolution imagery was reduced by the availability of the EVI and NDWI maps and time series. In contrast to Google Earth imagery, the Google Earth Engine satellite data is always up to date and can be used, even though it has a lower resolution. Therefore, although the choice of training points was more difficult that in the case of Mali, the impact on the quality of the results is limited.

It might be possible to eliminate the need for high-resolution imagery completely using this method, but it will bring additional risks of misinterpretation. This is especially relevant for areas where the distinction between natural growth and irrigation is difficult to detect, such as in valley bottoms. Therefore, it is advisable to create the ground truth data for a period for which high-resolution imagery is available, and to use that data to train the model. The same model can then be used to classify imagery in other years.

CLIMATE REGIONS

Both Mali and Chad cover three climate zones, and therefore have a gradual change of the climate across the country. This is not a good match for Random Forest machine learning, and it can lead to a loss of discriminatory power. The way this issue was mitigated in this study – by dividing the country into three zones – is not ideal, as it triples the need for ground truth data.

6.3 LIMITATIONS OF AREAS SUITABLE FOR IRRIGATION

RELIABILITY AND LOW RESOLUTION OF GROUNDWATER MAP

In the analysis of areas suitable for irrigation, the weakest link is the groundwater map. It has a low spatial resolution (5000 meter), and is mainly based on modelling, with limited input of case studies and other data. The reliability of the groundwater layer will have quite a large effect on the estimates of the scenarios of suitable areas that include groundwater, as it is the main ingredient and determines 50% of the score, and also acts as a constraint layer.

Of course, at present, the large majority of irrigated land is based on surface water which limits the importance of this issue somewhat. According to AQUASTAT, for Mali 99.5% of the area equipped for irrigation uses surface water, and for Chad the number is 80%. If we take into account flood recession and valley bottoms, the numbers are even starker skewed towards surface water.

Details on the methodology, including a full list of data sources used, are available in two publications (MacDonald, 2012 and Bonsor, 2011). The authors explicitly state (p.6): "*The maps presented here are designed to give a continent-wide view of groundwater and to encourage the development of more quantitative national and sub-national quantitative maps and assessments to support the development of groundwater-based adaptation strategies to* *current and future climate variability.*". This indicates that the map should mainly be used to draw conclusions on a regional scale, and not on specific locations.

The fact that the layer has a low resolution is not that important for the estimate of the potential areas on the regional level, as the details will even out. The resolution is only important when the situation is estimated at a precise location, where the low resolution can change the score substantially.

We conclude that in the case of specific areas, such as the bas-fonds in Mali, the groundwater maps are not of sufficient resolution or reliability to map the potential accurately. A more promising route is to identify promising regions from the potential area analysis, and to use ground truthing validation and additional research to provide the next level of accuracy for specific areas. To accomplish this, field visits are essential as many other limiting factors can be present: political, safety, land ownership, local soil conditions, pollution, etc.

ABSENCE OF HYDROLOGICAL AND OTHER CONSTRAINTS

The multi-criteria model used in this report makes use of a limited number of layers to compute a suitability score. However, there are a large number of additional factors that can constrain suitability, for which it is harder to obtain data, such as local political situation, safety, land ownership, flooding potential, detailed topography, local soil conditions, pollution, salinity, etc. Therefore, a local assessment will always be needed when specific sites are chosen, in which the full complexity of determining irrigation potential can be determined.

An important limiting factor is the hydrological constraint: how much water is actually available in a whole catchment area to support irrigation expansion. From section 6.1, we see that this factor has a large impact on the result. Therefore, before decisions are made to expand irrigation in a certain area, a study should be done focussing on the expected impact of water extraction on the catchment area, groundwater levels and downstream river levels.

In discussing results, a clear distinction should always be made between areas suitable for irrigation, and the fully realizable *potential of sustainable growth* for irrigation.

7. CONCLUSIONS

CURRENTLY IRRIGATED AREAS IN THE DRY SEASON

From our results, we conclude that remote sensing combined with machine learning performs well in the classification of irrigated areas, provided that the local effects of irrigation can be clearly distinguished from natural processes.

The main limitations to the accuracy stem from areas where this distinction is more difficult, such as in valley bottoms or flood recession agriculture, both of which are common in the studies countries. In such areas, classifying ground truth points remotely is challenging due to the fact that the satellite spectral signature and timing of for example flood recession crop growth is very similar to natural growth.

A second challenging aspect is the climate variability issue when machine learning is applied across a full country. In this report, this was mitigated by using multiple climate regions, at the cost of a significant increase in the number of required ground truth points.

Finally, using just machine learning does not allow for the automated distinction between small-scale and large-scale agriculture, or the distinction between farmer-led and government-led agriculture. To make these distinctions, additional geospatial information on the location and extent of large-scale schemes, flood recession areas, and valley bottoms is needed. A viable alternative, although less precise, is to use the size of patches of contiguous pixels classified as irrigated to distinguish between small, medium, and large-scale irrigation, as was successfully done in this report.

Given these limitations and issues, and given the fact that machine learning performs well locally when provided with the right ground truth data, one important option is to focus on smaller areas, and on particular types of irrigation. For example, a specific model trained on a specific bas-fond area would perform better than a generic model trained on ground truth data across a large area and corresponding diversity of irrigation methods. Such a model could be used to monitor irrigation locally over a time period. A combination of such models could be used to monitor a range of regions and irrigation types.

Using remote sensing and machine learning has the significant benefit that it can be automated and carried out over different periods, for example yearly, to determine trends. As the same set of ground truth data and the same trained model can be used over different years, this could be done by further automation of the process of classification developed in this report, by making use of a programming interface Google Earth Engine offers.

REPLICABILITY

An important issue is in how far these results can be replicated across time and space. First of all, as stated above, spatial replicability of the classification is limited because of substantial variation in local climate, geography, land cover, and possibly agricultural practices. Therefore, each region will need its own set of ground truth data for the training of the classification algorithm. However, this limitation only applies to the ground truth data itself. The machine learning methodology itself can be applied in different regions without problem, as long as the base data (Sentinel 2) is available.

As regards replicability across time, this is only limited by the availability of data. For example, Sentinel 2 data started being collected after June 2015, when the first Sentinel 2 satellite was launched. As long as there are no large changes in local climate, ground truth data collected in a given year can be used to train a model that can then be used in other years as well. Collecting ground truth data is a time-consuming process, requiring multiple days per country for the data collected in this report.

AREAS SUITABLE FOR IRRIGATION

From the results presented in Chapter 6, it is clear that the areas suitable for irrigation estimated in this report are in line with other studies that use a similar multi-criteria model, but that the results are significantly higher than studies that incorporate hydrological constraints. This means that in both countries, the hydrological component is essential for estimates of the total realizable irrigation potential.

This result underlines the importance to distinguish between areas suitable for irrigation as determined by all constraints under consideration, and the fully realizable *potential of sustainable growth* for irrigation. In addition to considering the other constraints mentioned above, to determine the scope for actual expansion of irrigation in a given area an analysis should be made on the catchment level to determine the total amount of water that can be safely used for increased irrigation without negative impact on groundwater levels or river flow downstream.

In addition, it should be noted that there are a significant number of other, local, factors that form potential constraints, such as soil quality, land productivity, pollution, salinity, local ownership situation, detailed topography, possible market saturation, etc. Therefore, the suitable areas as described in this report should be used only to identify overall suitable regions and guide site selection for further detailed suitability study.

8. RECOMMANDATIONS FOR FURTHER WORK

From the results of this study, we can formulate a number of general recommendations. These are listed below.

TECHNICAL DEFINITION FARMER LED IRRIGATION

Farmer-led irrigation uses a range of technologies, for example: controlled flood recession, weirs and dams, water retention, and irrigation from wells or boreholes. Each of these has a different signature on remote sensing data, and some are harder to detect then others. It would be useful to conduct a study to observe the effect of each of these technologies on remote sensing data, as this might lead to strategies for more reliable detection of irrigation. To that end, it would be beneficial to form a more technical definition of farmer-led irrigation that is focused on remote sensing specifically, so studies of different types can be compared. For example, by restricting definitions to specific types of water management, and agricultural practices that lead to identifiable plant growth in the dry season.

IMPROVE MACHINE LEARNING ALGORITHM

In machine learning, there are a large number of possible algorithms, and an even larger number of possible bands. For example, in this study we used sentinel-2 data exclusively, but we could also have used Sentinel-1 radar data, rainfall data, solar irradiation data, humidity data, etc. It would be beneficial to conduct a detailed study exploring the various machine learning algorithms and possible bands, in order to determine the most effective combination to use. This could also include ways to minimize the issue of cross-country climate variation.

Separate analysis of receding water agriculture

Given the fact that receding water agriculture is so prevalent in both Mali and Chad, it would be beneficial to apply a separate analysis. This could make use of other information sources, such as the JRC water occurrence layers[13]. By restricting the analysis to only areas where water does not occur during the full year, it is likely that the machine learning analysis would yield a more accurate result.

Limit the regional extent of analysis

One of the issues for this study has been the large geographic area. In a full country, there is substantial variability of climate and other properties, which complicates the analysis. For future work, it would be beneficial to focus on smaller regions inside a country. For example, a specific model trained on a specific bas-fond area would perform better than a generic model trained on ground truth data across a large area and corresponding diversity of irrigation methods. Such a model could be used to monitor irrigation locally over a time period. A combination of such models could be used to monitor a range of regions and irrigation types.

Improve estimates of hydrological constraints for increased irrigation

From the comparison of our results on areas suitable for increased irrigation with estimates of total irrigation potential, it became clear that the hydrological constraints play a major role in determining the latter. However, the estimates of these constraints suffer from substantial uncertainly, especially on the amount of water needed to satisfy environmental needs. To gain more confidence in the actual limits to increased irrigation, more in-depth studies are needed that reduce these uncertainties.

[13] https://developers.google.com/earth-engine/datasets/catalog/JRC_GSW1_2_GlobalSurfaceWater

LITERATURE

Abric 2011 — Abric S., Sonou M., Augeard B., Onimus F., Durlin D., Soumaila A., and GadElle F. Lessons Learned in the Development of Smallholder Private Irrigation for High-Value Crops in West Africa. World Bank 2011, Washington DC, USA.

Altchenko 2014 — IWMI, 2014. Altchenko Y. & Villholth K. G., Mapping irrigation potential from renewable groundwater in Africa – a quantitative hydrological approach.

Beekman 2014 —Identifying the potential for irrigation development in Mozambique: Capitalizing on the drivers behind farmer-led irrigation expansion. Beekman et al., Physics and Chemistry of the Earth, vol 76-78, p54-63, 2014. https://www.sciencedirect.com/science/article/pii/S1474706514000618

Bonsor, 2011 — An initial estimate of depth to groundwater across Africa, H.C. Bonsor and A.M. MacDonald, BGS Groundwater Science Programme, Open Report OR/11/067, 2011.

DNAER 2003 — Ministere de l'Agriculture de l'Elevage et de la Peche. 2003. Carte d'Irrigation du Mali par Regions Hydrauliques. Direction Nationale de L'Aménagement et de L'Equipement Rural (DNAER). Map available in the AQUASTAT library.

Donchyts, 2017 — G. Donchyts et.al., Segmentation of satellite images using SLIC and Google Earth Engine, Poster presentation European Geoscience Union, April 2017.

FAO 2005 — FAO. 1987 (updated in 2005). Irrigated areas in Africa. FAO, Rome, Italy.

FAO, 2016 — AQUASTAT Main Database, Food and Agriculture Organization of the United Nations (FAO). Website accessed on [10/12/2020 11:48]

IWMI, 2019 — International Water Management Institute (IWMI). 2019. Suitability for farmer-led solar irrigation development in Mali. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Research Program on Water, Land and Ecosystems (WLE). 4p.

MacDonald, 2012 — Quantitative maps of groundwater resources in Africa, A.M. MacDonald et.al, Environmental Research Letters, Volume 7, Number 2, 2012.

Pavelic, 2013 — IWMI, 2013. Pavelic P., Villholth K.G., Shu Y., Rebelo L.M. & Smakhtin V. (2013) Smallholder groundwater irrigation in Sub-Saharan Africa: country-level estimates of development potential, Water International, 38:4, 392-407, DOI: 10.1080/02508060.2013.819601.

Schmitter, 2018 — Suitability mapping framework for solar photovoltaic pumps for smallholder farmers in sub-Saharan Africa, P. Schmitter et.all, Applied Geography, vol. 94, May 2018, p.41-57.

SDEA 2001 — Le Bilan-Diagnostic de l'Hydraulique Agricole en 2001, SDEA, https://www.un.org/esa/sustdev/tech_coop/sdea/3_agricole/index.html

UN-DESA 2003 —Present situation of Chad's water development and management, UN-DESA, https://www.un.org/esa/sustdev/publications/sdea/english/index.html (french version https://www.un.org/esa/sustdev/tech_coop/sdea/principal/index.html)

UN-DSD 2003 — UN Division for Sustainable Development. 2003. The integrated plan for Chad's water development and management (SDEA). Chapter 1: Present situation of Chad's water development and management, available at: **http://www.un.org/esa/sustdev/tech_coop/sdea/**).

Vogels, 2019 — Monitoring smallholder irrigation using remote sensing and object-based image analysis: Case studies in Eastern Africa, M. Vogels (2019), PhD study University of Utrecht.

Weiss, 2018 — A global map of travel time to cities to assess inequalities in accessibility in 2015, D.J. Weiss, et.al., Nature (2018). doi:10.1038/nature25181.

WB / Ward (2016) — Ward C., Torquebiau R, and H. Xie, Improved Agricultural Water Management for Africa's Drylands, World Bank, 2016.

You, 2010 — IFPRI, 2010. You L., et.al. What Is the Irrigation Potential for Africa? A Combined Biophysical and socioeconomic approach. IFPRI discussion paper 2010

Xiong, 2017 — Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine, Xiong, J et.al., Remote Sens. 2017, 9, 1065.

APPENDIX

List of FAO AQUASTAT definitions. Source: (FAO, 2016)

Table 48. List of AQUASTAT irrigation terms and definitions.

Arable land area	Land under temporary crops (double-cropped areas are counted only once), temporary meadows for mowing or pasture, land under market and kitchen gardens and land temporarily fallow (less than five years). The abandoned land resulting from shifting cultivation is not included. Data for arable land is not meant to indicate the amount of land that is potentially cultivable.
Permanent crops area	Crops are divided into temporary and permanent crops. Permanent crops are sown or planted once, and then occupy the land for some years and need not be replanted after each annual harvest, such as cocoa, coffee and rubber. This category includes flowering shrubs, fruit trees, nut trees and vines, but excludes trees grown for wood or timber, and permanent meadows and pastures.
Cultivated area (arable land + permanent crops)	The sum of the arable land area and the area under permanent crops.
Irrigation potential	Area of land which is potentially irrigable. Country/regional studies assess this value according to different methods. For example, some consider only land resources, others consider land resources plus water availability, others include economical aspects in their assessments (such as distance and/or difference in elevation between the suitable land and the available water) or environmental aspects, etc. The figure includes the area already under agricultural water management.
Area equipped for full control irrigation: total	The sum of surface irrigation, sprinkler irrigation and localized irrigation.
Area equipped for full control irrigation: actually irrigated	Portion of the area equipped for full control irrigation that is actually irrigated, in a given year. It refers to physical areas. Irrigated land that is cultivated more than once a year is counted only once.
Area equipped for irrigation: equipped lowland areas	The land equipped for irrigation in lowland areas includes: (i) Cultivated wetland and inland valley bottoms (IVB) that have been equipped with water control structures for irrigation and drainage (intake, canals, etc.); (ii) Areas along rivers where cultivation occurs making use of structures built to retain receding flood water; (iii) Developed mangroves and equipped delta areas.
Area equipped for irrigation: spate irrigation	Spate irrigation (sometimes referred to as floodwater harvesting) is an irrigation practice that uses the floodwaters of ephemeral streams (wadi) and channels it through short steep canals to bunded basins where cropping takes place. A dam is often built in the wadi to be able to divert the water whenever it arrives. These systems are in general characterized by a very large catchment upstream (200-5000 ha) with a ratio of "catchment area : cultivated area" = between 100:1 - 10 000:1. There are two types of spate irrigation: 1) floodwater harvesting within streambeds, where turbulent channel flow is collected and spread through the wadi where the crops are planted; cross-wadi dams are constructed with stones, earth, or both, often reinforced with gabions; 2) floodwater diversion, where the floods - or spates - from the seasonal rivers are diverted into adjacent embanked fields for direct application. A stone or concrete structure raises the water level within the wadi to be diverted to the nearby cropping areas.

Area equipped for irrigation: total	Area equipped to provide water (via irrigation) to crops. It includes areas equipped for full/partial control irrigation, equipped lowland areas, and areas equipped for spate irrigation.
Area equipped for irrigation: actually irrigated	Portion of the area equipped for irrigation that is actually irrigated, in a given year. It refers to physical areas. Irrigated land that is cultivated more than once a year is counted only once.
Flood recession cropping area non-equipped	Areas along rivers where cultivation occurs in the areas exposed as floods recedes and where nothing is undertaken to retain the receding water. The special case of floating rice is included in this category.
Cultivated wetlands and inland valley bottoms non-equipped	Wetland and inland valley bottoms (IVB) that have not been equipped with water control structures but are used for cropping. They are often found in Africa. They will have limited (mostly traditional) arrangements to regulate water and control drainage.
Total agricultural water managed area	Sum of total area equipped for irrigation and areas with other forms of agricultural water management (non-equipped flood recession cropping area and non-equipped cultivated wetlands and inland valley bottoms)