EO MAJI
EO AFRICA EXPLORERS

AFRICAN EARLY ADOPTERS CHARACTERISATION AND BENEFIT ANALYSIS REPORT

V1

Date: 25/04/2023

Contract No.

4000139395/22/I-DT

Submitted by

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

In Cooperation with:

DHI   UNIVERSITY OF LEICESTER

esa

DOCUMENT RELEASE SHEET

Author:	Michael Munk (DHI), Radoslaw Guzinski (DHI)	Sign Date:
In collaboration with:	Héctor Nieto (CSIC)	
Approval	Zoltan Szantoi	Sign Date:
Distribution:	ESA and partners	

CHANGE RECORD

Version	Date	Page(s)	Change record	Release
0	27/04/2023	15	First submission to ESA	1
1				
2				
3				

1 Introduction

1.1 Project objective

This project aims to implement a prototype for irrigation mapping and crop yield estimation using inputs from the scientific ECOSTRESS and PRISMA missions. The final aim is to develop workflows, in collaboration with the African Early Adopters and EO partner(s), that support African irrigation and food security management, as well as transferring these R&D learnings and results to African end-users and stakeholders. More specifically the project objectives in this project can overall be listed as:

- Exploration of the capabilities for future operational Copernicus missions (LSTM+CHIME) to estimate ET and crop water stress.
- Investigate the potential of PRISMA hyperspectral observations and thermal-based crop stress metrics to improve crop yield/biomass estimations to support agricultural monitoring
- Complement the ET retrievals with crop yield, in order to acquire a better understanding of water use efficiency (WUE) of cultivated landscapes.
- Direct involvement of African Early Adopters, in order to secure the usefulness and applicability of the prototype.
- Publish the findings in a freely available code repository and as scientifically peer-reviewed papers, as well as to promote the codes through other outreach activities such as development of digital notebooks.

All activities are to be carried out within the duration of the project lifetime from 1 December 2022 to 30 November 2024.

1.2 Scope of Document

This document presents the characterisation of the involved African Early Adopters, understanding their mandate, technical capacity and working practices. In this document we will identify the benefits that the innovative EO analysis techniques could bring to the involved African Early Adopters. This report consists thus in four sections:: i) User Requirements; ii) Product Specifications; iii) System Requirements; iv) Operational Requirements.

1.3 Reference documents

REF-1	Statement of Work: ESA-EOP-SD-SOW-0250 – EO AFRICA EXPLORERS
REF-2	EO MAJI proposal dated 18/02/2022
REF-3	Clarification request from ESA dated 06/06/2022
REF-4	Response to clarification dated 22/06/2022
REF-5	Contract No. 4000139395/22/I-DT

2 User Requirements

Effective project management involves a comprehensive understanding of user requirements. The EO-MAJI projects aims to work in close collaboration with the early adopters of this state-of-the-art R&D project designed to use specialized Earth Observation (EO) data to monitor water use, irrigation delineation, irrigation accounting, and crop yield estimates. The early adopters are represented by three African countries: **Burkina Faso, Botswana, and South Africa.**

The main aim of this project is to develop a robust and reliable system that meets the needs of its users. This requires a thorough understanding of the requirements and constraints of each user group. User requirements are a set of expectations, specifications, and features that are necessary to ensure that the developed system is functional, reliable, and user-friendly. These needs and expectations of the users are an essential component of any successful project. User requirements will be identified and incorporated throughout the development process. The user requirements will be used to guide the development of the system, ensuring that it meets the needs of the users. The feedback from the users will be incorporated to improve the effectiveness and usability of the system.

The users in this project are early adopters and are chosen for their experience in the field of water management, irrigation, and agriculture as well as their growing need to monitor water use and formulate legislation to manage the countries limited or varying water resources. Their feedback and input are critical to ensuring the success of the project. As such, this section of the project report will provide an outline of the user requirements.

The user requirements in this project include the need for accurate and reliable data collection, effective data analysis, and easy-to-use data for the users. The method should be flexible enough to adapt to the unique needs of each user group, and it should be capable of producing reliable results in a timely manner in relevant resolution. Additionally, the system should be accessible to users with varying levels of technical expertise.

To gather the user requirements, the project team conducted interviews with representatives from each user group. The feedback was then analyzed to identify common themes and specific requirements. The results of these analyses were used to develop a set of user cases with well-defined areas of interests, that will guide the development of the system.

2.1 Burkina Faso

2.1.1 Area of interest:

Located in the south-western part of Burkina Faso (see Figure 1) and specific area of interest for the EO-MAJI project is the area surrounding Bama (see Figure 2). The area is characterized by poorly managed water distribution and farmers come to the area from other regions. The ministry supports the farmer association local office.

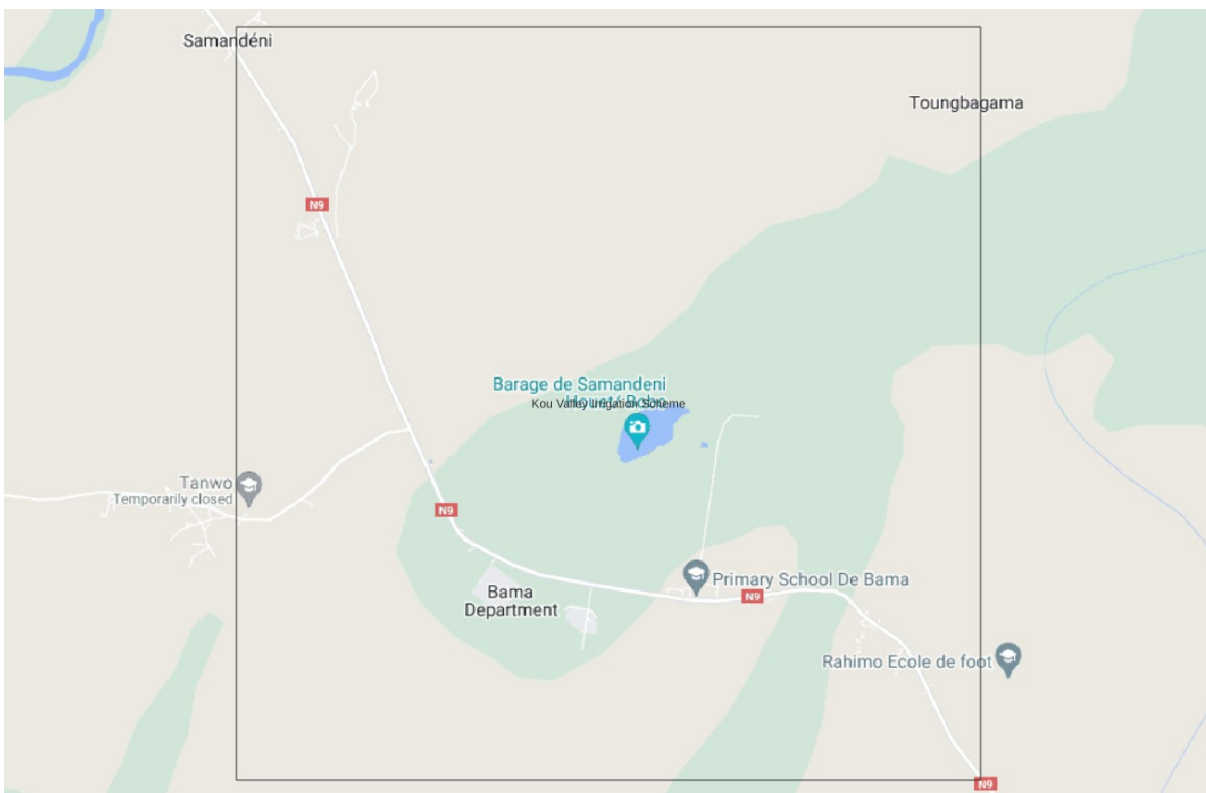**Figure 1 Area of interest, geographic overview**

Figure 2 Area of interest, bounding box used to task ECOSTRESS and PRISMA data collection

2.1.2   Crop of interest:

Rice is the main crop in the area of interest. **Maize** and **potato** are secondary crops in irrigated areas.

2.1.3   Irrigation practise:

Ministry is in process of developing more irrigation systems. Currently, there are mostly unregulated small-scale irrigated farming: wetland/floodplain (extraction from rivers or dams)

2.1.4   Use case objective:

The two main focus point for Burkina Faso as early adopters are:

A.   **Managing existing irrigation schemes** with better insight and monitoring methods

B.   **Supporting farmer associations** by capacity building

2.1.5   In-situ data:

The local farmers association office has some data on water supply and crop data. It is to be determined if this data is available and suitable for this project.

2.2 Botswana

2.2.1   Area of interest:

Located in the Eastern part of Botswana, close to the boarder triangle of Zimbabwe and South Africa (see Figure 3) and specific area of interest for the EO-MAJI project is the area of Tuli block (see Figure 4). Semi-arid landscape with occasionally high demand for irrigation and often prioritized for domestic purposes. Irrigation is mostly used in the dry months of May-June. The eastern part of Botswana has good soils, making it suitable for irrigation (similar to the Zambezi River project in the northern part of Botswana). Farm size vary from 1-5 ha for small-scale and 500-600 ha for large farming
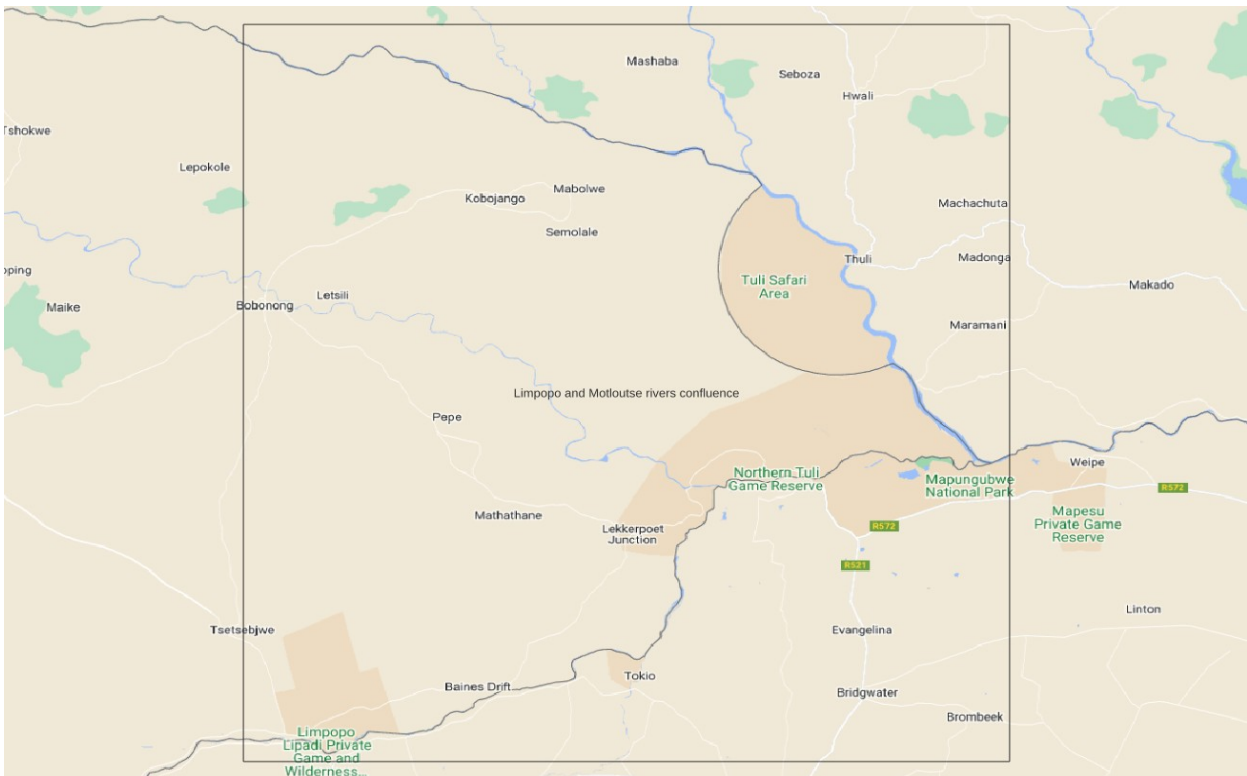
Figure 3 Area of interest, geographic overview

Figure 4 Area of interest, bounding box used to task ECOSTRESS and PRISMA data collection

2.2.2   Crop of interest:

Vegetables (e.g. cabbage, tomatoes, cucumber etc) and **fruits** (e.g. citrus and mango) dominate the area of interest. Depending on season **potato** will also be a crop of interest, and rainfed crops include maize, beans and peanuts.

2.2.3   Irrigation practise:

The Limpopo River catchment have extensive irrigation, mostly with river extraction practise both also with many bore holes on private land. The area of interest is defined by more than average irrigation for Botswana. Historically, irrigation has not been used very extensively in this area. Agriculture is still primarily rainfed, but irrigation methods are becoming more prevalent. Mainly drip systems from river water extraction, but more advanced sprinkler systems are become increasingly common. There is very little information on how much water is extracted in Botswana, but competition is increasing.

2.2.4   Use case objective:

Two main objectives are identified:

A.   Support development of irrigation licensing schemes

B.   Map irrigated areas outside the irrigation schemes

Other secondary objectives include:

A.   Estimating quantity of water that can be extracted from the river

B.   Investigate the potential for using recycled water (treated wastewater) for irrigation

C.   General guidance to water resource management

2.2.5   In-situ data:

The area consists of mainly commercial farming. Therefore, it is to be investigated if local crop yield data can be acquired.

2.3 South Africa

2.3.1   Area of interest:

Located in the north-western part of South Africa (see Figure 5) and specific area of interest for the EO-MAJI project is the area surrounding Potchefstroom west of Johannesburg (see Figure 6).
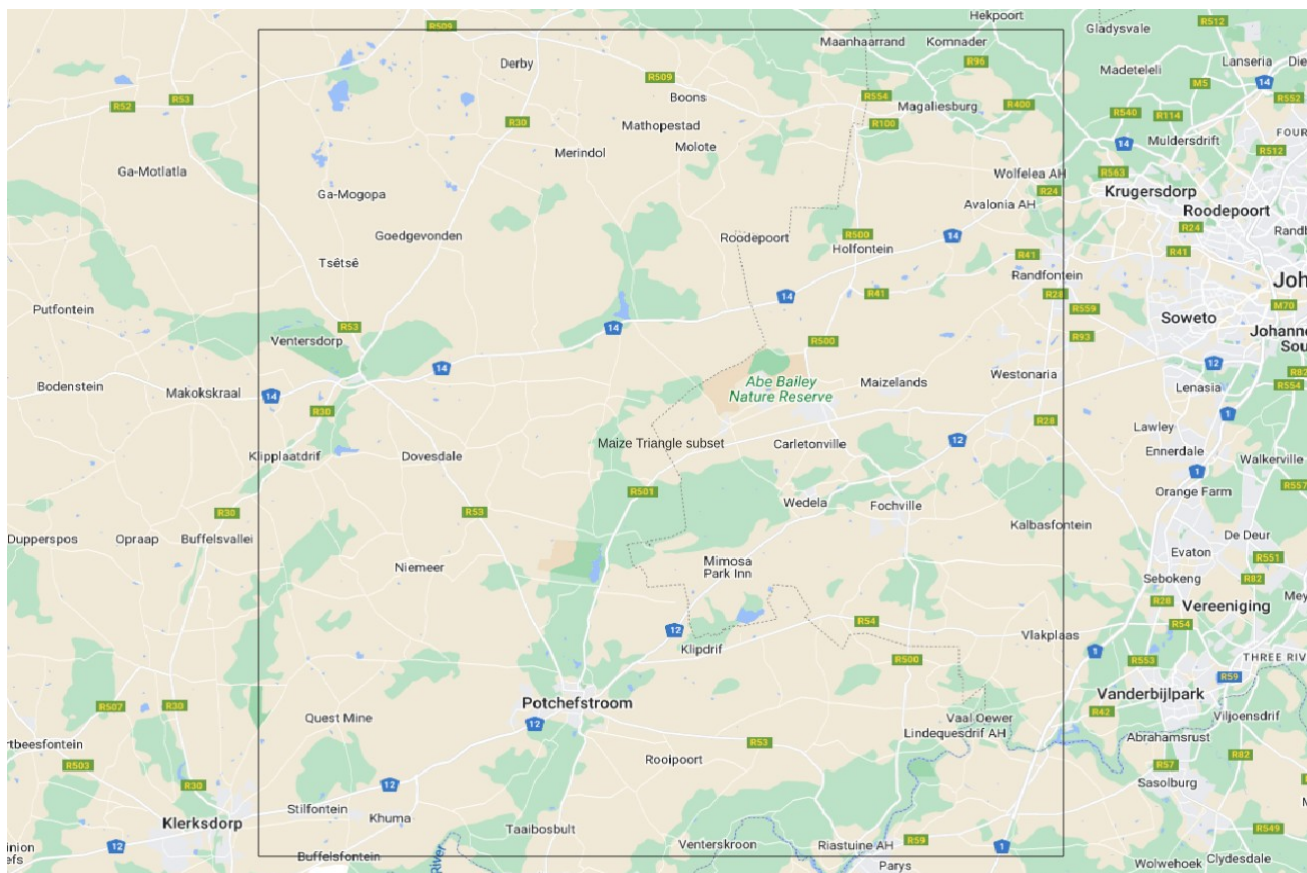
Figure 5 Area of interest, geographic overview

Figure 6 Area of interest, bounding box used to task ECOSTRESS and PRISMA data collection

2.3.2   Crop of interest:

The areas already have ongoing monitoring of **maize** farmers. Soybeans are often the secondary crop type.

2.3.3   Irrigation practise:

The area is mostly rainfed, but different irrigations methods is being implemented to ensure food security.

2.3.4   Use case objective:

Main objectives are more technical as CSIR are already engaged in EO applications for agriculture:

A. Co-development of algorithms and product evaluation

B. Improve water crop-yield efficiency in better resolution (current resolution is not sufficient for small-scale farms), especially for dryer years.

2.3.5 In-situ data:

CSIR have data for crop-related properties which can be used.

General needs for the type of data expected from EO-MAJI

3 Product Specification

Three main products related to food security are going to be developed are irrigation delimitation, irrigation accounting and crop yield. For additional information about the proposed solutions please refer to the State-of-the-Art document [REF-6].

The three objective products share basically the same input data sources, mainly ET from ECOSTRESS+PRISMA+Sentinel and crop phenology/biophysical traits, but each product will have slightly different specifications due to different user needs and foreseen retrieval approach. These specifications are summarized in Table 1.

Table 1: Product specifications

	Irrigation delimitation	Irrigation accounting	Crop yield
Physical unit	Binary (Presence/absence)	mm/ha or m^3	kg/ha
Spatial Coverage	Regional	Regional	Regional, provided for a given crop
Spatial Resolution	20 m	100 m	20 m
Temporal Coverage	2021-2023	2021-2023	2021-2023
Temporal Resolution	Annual	Monthly	Annual or 2 per year

All products will have a regional coverage as defined by the Areas of Interest (AOI) submitted for PRISMA acquisitions, the extension of these AOIs are described in Table 2. However, for the crop yield product, within each AOI, only yield information will be provided for the crop of interest (Table 2). Furthermore, since rice in Burkina Faso is cropped twice a year, during the wet and dry seasons, the temporal resolution in this case will consist of two maps per year.

Table 2: Site-specific product specifications, including the spatial coverage for the Areas of Interest of the three African sites and the crop of interest.

	Burkina Faso	Botswana	South Africa
Minimum Longitude	-4.45	28.4	26.7
Minimum Latitude	11.35	-22.6	-26.9
Maximum Longitude	-4.35	29.4	27.7
Maximum Latitude	11.45	-21.6	-25.5
Crop of interest	Rice	Potato	Maize

The spatial resolution for each product depends mainly on the foreseen sensitivity of the input products. Both irrigation accounting and crop yield would be as sensitive to phenology (shortwave region) as to evapotranspiration (ET, thermal region), and thus 20m resolution could be considered as the goal. However, irrigation accounting is much more sensitive to ET and meteorological data (e.g. rainfall to force the water balance required in the proposed algorithm [REF-6]) and thus we consider delivering a product with a goal of 100m spatial resolution. This is also in line with the user need requirements, since irrigation accounting is useful for water resources planning (I.e. irrigation district level), and water licensing (i.e. parcel level).

4 System Requirements

4.1 End-to-end processing workflow

The core generated products of the EO-MAJI product suite consist of crop yield, irrigation delineation and irrigation accounting with the aim of improved understanding of water use efficiency in existing and new African irrigation schemes. However, the generation of the core products requires a number of intermediate products, most of which can be used as standalone datasets in their own right. Those intermediate products include land surface temperature produced through fusion of Sentinel-3 and ECOSTRESS observations, plant biophysical traits produced through fusion of Sentinel-2 and PRISMA observations and evapotranspiration estimates modelled through the TSEB model. Each of those products requires its own processor and by chaining the processors together it should be possible to perform end-to-end analysis. Figure 7 shows the linkages between those processors / products.
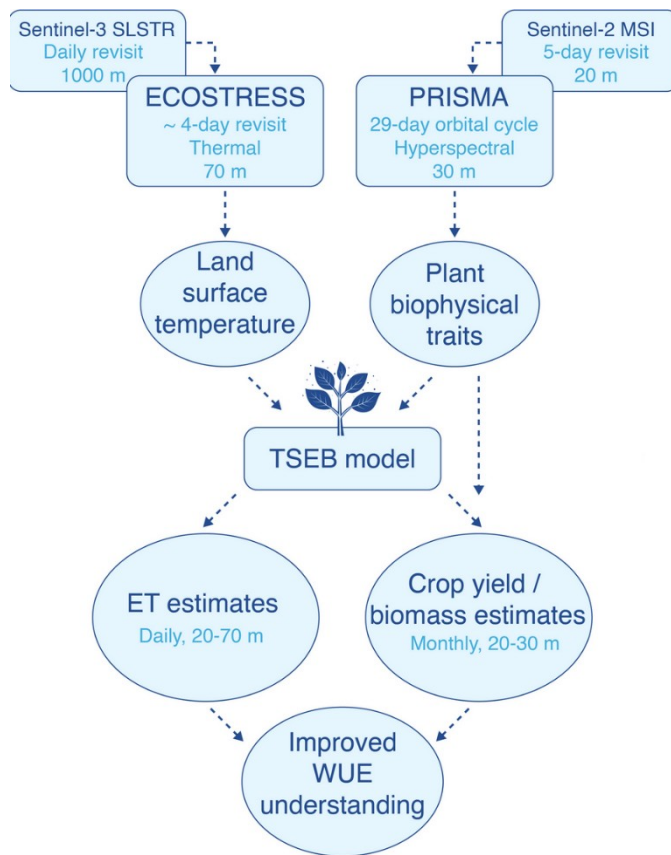
Figure 7: End-to-end processing chain showing intermediate and final products.

4.2 Output data requirements

The data produced in EO MAJI project is to be applied in African agricultural setting. This implies that agricultural parcels and irrigation perimeters can often be small, on the order of hectares or less. The two EO MAJI Early Adopters are based in Burkina Faso and Botswana, which are both characterized with semi-arid climate and a distinctive rainy season. The above conditions place requirements on spatial and temporal resolutions of the output data.

The highest resolution input data set is Sentinel-2 imagery, which includes observations in the red-edge and shortwave infra-red part of the spectrum with 20 m resolution. This should be sufficient for majority of irrigated agricultural parcels, particularly in larger commercial farms. Thermal observations and PRISMA data are obtained at lower spatial resolution (see Figure 7) but the goal of the EO MAJI system should be to provide as many outputs as feasible at 20 m spatial resolution.

While the main output products (crop yield, irrigation delineation and irrigation accounting) are to be provided at seasonal timesteps, the intermediate products required to obtain the main products need to be produced at a higher temporal resolution. A good compromise between capturing the phenological dynamics of crops and the limitations imposed by cloud cover is a dekadal (10-day) timestep. This is particularly relevant for evapotranspiration, which is dynamic parameter influenced not only by phenology but also by weather and root-zone moisture availability. Other biophysical parameters, such as leaf area index or yield production, could be produced at 15 or 20-day timesteps if dekadal frequency proves not to be feasible.

4.3 Input data requirements

The system to be developed in EO-MAJI is a scientific prototype rather than an operational production system. In addition, at the moment there is no one location where all the inputs data is located. Therefore, being close to the data is advantageous but not required. This advantage is especially relevant when processing time-series of satellite imagery in an African setting, where data download can often be a limiting factor due to slow internet speeds. The input data can be separated into three categories: satellite imagery, meteorological forcings, higher-level ancillary datasets.

Satellite imagery can be further split into thermal and shortwave optical and those categories further separated into Copernicus and experimental data. While Copernicus imagery is accessible from multiple platforms, such as the DIAS's, ECOSTRESS and PRISMA data is available from ECOSTRESS Hub and PRISMA portal respectively. All the satellite imagery is optimally required to be available processed to Level-2, i.e. with cloud masking and atmospheric correction applied. However, certain pre-processing steps (such as running Sen2Cor on Sentinel-2 data) could be incorporated into the EO MAJI system if required.

Meteorological data is required for modelling of ET and possibly for other biophysical models. This meteorological data are based on ERA5 reanalysis dataset produced by European Center for Medium Range Weather Forecasts and distributed through the Copernicus Climate Data Store (cds.climate.copernicus.eu). Since near-real-time production is not planned in the EO MAJI project, the ERA5T dataset, which is distributed with a 5-day delay fits the requirements. Both instantaneous and daily parameters can be obtained from ERA5 data. Instantaneous parameters are used to drive the ET model and included air temperature, vapor pressure, wind speed, surface pressure, instantaneous solar irradiance. All instantaneous data should be temporally interpolated to the time of Sentinel-3 SLSTR or ECOSTRESS acquisition over the area of interest. Daily parameters are used to extrapolate and interpolate the instantaneous estimates of ET and include solar irradiance, precipitation and reference ET. They should integrated over a 24 hours period starting at midnight local time.

Two ancillary sources of data are required: a landcover map and a digital elevation model (DEM). Both should be at the spatial resolution which can match the resolution of the output products, i.e. as close to 20 m as possible. ESA WorldCover land cover map with 10 m spatial resolution and Copernicus DEM with 30 m spatial resolution both fit this requirement. The land cover map is used to assign ET model parameters which are difficult to estimate directly from other satellite data, such as maximum vegetation height or leaf orientation. The DEM is used for three main purposes: during Sen2Cor atmospheric correction (in case Level-2 Sentinel-2 data is not available), to correct meteorological parameters for terrain effects (elevation and illumination conditions), and in the thermal sharpening model to add elevation and illumination conditions as predictor variables.

4.4 Processing system requirements

The developed system should be usable by the African Early Adopters and the rest of user community. Therefore, the system should be easy to deploy on different software and hardware infrastructure and to adapt to users' specific needs. In addition, the system needs to be open-source. Python programming language, which is one of the de-facto standards in the Earth Observation community, fulfils all those requirements. It is accessible, with easy to read and write syntax, and open-source and hardware independent due to its interpreted (as opposed to complied) execution.

The system should follow a modular design, with the production of each of the final and intermediate products contained in a separate module. Further splits within each module are also possible. The modularity could take form of a Python script or package, Docker container or Jupyter notebook. This modular architecture will allow users to pick and chose the methods for the different parts of the processing chain and replace them as necessary. This is especially important in a scientific project, where the different modules could undergo evolution at a different pace. Such architecture also limits interdependencies between the different project partners, as each will be assigned primary responsibility for one or more modules and can work on its development independently from the other partners.

There will be three main interfaces into each module: a Python package Application Programming Interface (API), a command line interface and a Jupyter notebook. The Python API will be geared towards the most advanced users, who would like to integrate the Python scripts and packages into their own Python workflows. It will offer the most flexibility

but at the same time the most complexity. It will also be the basis for the other two interfaces. Command line interface will provide a simplified interface for users who are not comfortable with using Python but would still like to use the modules in an automatic way for larger productions and batch jobs. Finally, Jupyter notebook will provide an interactive interface which will allow users to explore the different model options and inspect inputs and outputs with elements of graphical user interface (GUI). This interface will be most suitable for capacity building.

To broaden the applicability of the developed system it is important to use open and standard formats for data outputs. For raster data we will work with GeoTIFF format, which is a de facto standard in the remote sensing community. The outputs will be saved in a Cloud Optimized GeoTIFF format, which is backwards compatible with standard GeoTIFF but has the advantage of making the data easy to visualize and analyse both in desktop and cloud environments. Vector data will be stored either as GeoJSON (from smaller datasets, such as outlines of an AOI) or GeoPackage (for larger datasets, such as thousands of polygons with properties). Finally, tabular and timeseries data will be stored either in CSV or JSON formats. With this set out formats, the data will be usable in practically all GIS software, such as QGIS or Python.

5 Operational requirements

The operational requirements (timeliness, uncertainty and precision) for the prototypes of irrigation delimitation, accounting and crop yield are summarized in Table 3. Timeliness in this set of products in not critical as these are meant for planning activities. Therefore, there is no need to produce near-real-time products and a 1-3 months not-time critical timeliness is targeted. Since irrigation accounting requires running a water balance model, we consider 4 months timeliness in order to use ERA5 reanalysis rainfall data to ensure a better quality. Nevertheless, a short-time critical product for both irrigation accounting and crop yield could also be considered, in order to provide actual irrigation accounting for enforcing water licensing as well as to forecast crop yield before harvesting.

Table 3: Operational requirements

	Irrigation delimitation	Irrigation accounting	Crop yield
Timeliness	1 month after the end of the season	4 months after the end of the season	1 month after the end of the season
Uncertainty	5% overall accuracy	100 mm/ha or 20 % relative error	500 kg/ha or 10% relative error
Precision	20 m	1 mm/ha	1 kg/ha

The precision also differs between products. Since irrigation delimitation is a binary mask produced at 20m spatial resolution (Table 1), its precision can be considered as well as 20m. On the other hand, the precision for both irrigation accounting and crop yield, as scale variables, are influenced by the input data and the standard reporting units.