EO4CEREALSTRESS

Theme 3: Crop response to multiple stressors

Deliverable 2.2: Database Description

Prepared by	**Dr. Victor Rodriguez-Galiano (Universidad de Sevilla, Spain)**		
Approved by	Prof. Jadu Dash (University of Southampton, UoS, UK)		
	Dr. Booker Ogutu (University of Southampton, UoS, UK)		
Report to	ESA	**Author**	UoSev
Project Name	EO4CEREALSTRESS	**Contract No.**	ESA AO/1-11144/22/I-EF
Project Ref	Task 2: Dataset Collection	**Document Ref**	D2.2
Issue	Version 2.0	**Issue Date**	11 December 2023

Document Status

Issue	Version	Submission Date	Status
Draft	0	11 Oct 2023	Input from partners
Draft	1	18 Oct 2023	Submitted to ESA
Revised	2	11 Dec 2023	Update according to the ESA's feedback

ACRONYMS

AOI	Area of Interest
AU	Austria
DESIS	DLR Earth Sensing Imaging Spectrometer
ECMWF	European Centre for Medium-Range Weather Forecasts
EcoStress	Ecosystem Spaceborne Thermal Radiometer Experiment on Space Station
EO	Earth Observation
EnMAP	Environmental Mapping and Analysis Program
ES	Spain
ESA	European Space Agency
GB	Gigabyte
GSD	Ground Sampling distance
HYC	Hyperspectral Camera
ID	Identifier
ISS	International Space Station
IT	Italy
LST	Land Surface Temperature
MB	Megabyte
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NIR	Near Infrared
PAN	Panchromatic
PRISMA	Hyperspectral Precursor of the Application Mission
RS	Remote sensing
SIF	Solar Induced Fluorescence
SLSTR	Sea and Land Surface Temperature Radiometer
SWIR	Shortwave Infrared
TB	Terabyte
TIR	Thermal Infrared
TOC	Top of Canopy
TROPOMI	Tropospheric Monitoring Instrument

DOCUMENT TABLES

DOCUMENT FIGURES

TABLE OF CONTENTS

1. Scope of the document

The objectives of this report are to: i) identify in-situ, EO and climatic data to be used in understanding the impacts of multiple stressors on crops and applying algorithms for monitoring these impacts; ii) define areas of interest for acquiring data; iii) define the specifications for the EO4CerealStress data repository, based on the requirements in terms of time frames, extension, temporal and spatial resolution and input products. All this is required to design a common framework for sharing information among partners during the project term.

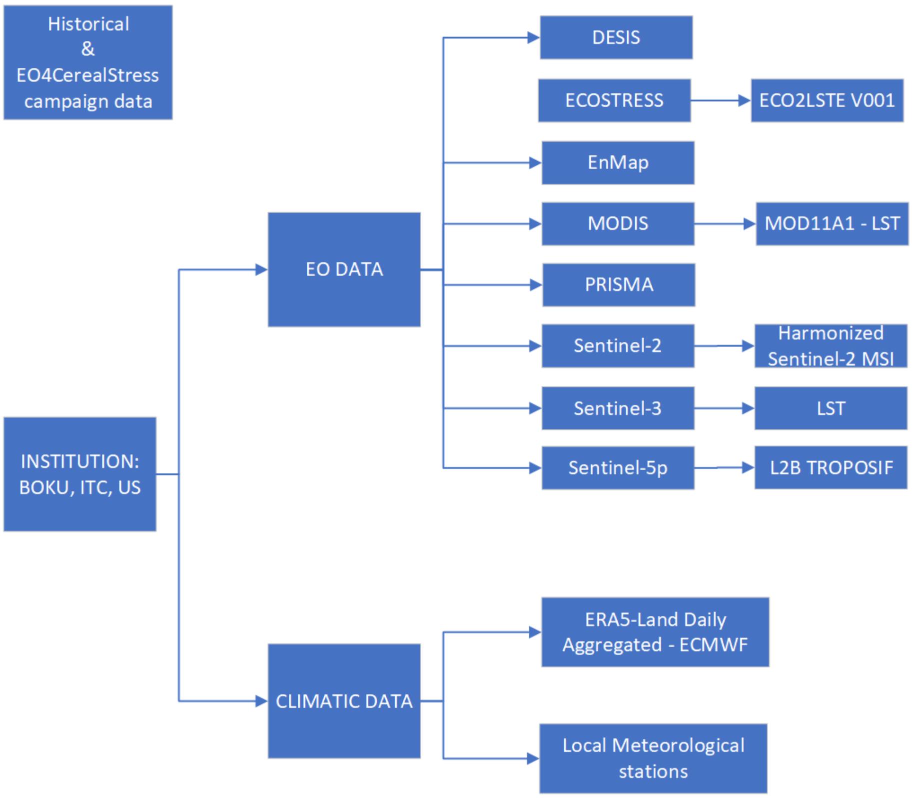
The database compiles the datasets identified in Task 1 for their potentials with respect to the scientific objectives and their suitability relative to each crop framework. The project reference database is managed by University of Seville, with contributions from all other partners. The database contains data for the following categories: historical campaign data, EO data and climatic data for different pilot sites.

2. Database structure

The database will include different campaigns that ESA and other agencies have undertaken to collect test data for a number of upcoming Satellite missions through the last 5 years. These data could serve an important source to study the relationship between crop biophysical variables and EO based stress signature. The database is organised in different folders and subfolders, grouped by institutions in charge of the pilots, type of data (EO or climatic), EO missions or source of climatic data and products (see Figure 1). All data are stored in the following location to be accessible within the team:

Experimental Datafiles

Figure 1. Structure of the database



The EO4CerealStress data repository has been built on the basic of the Findable, Accessible, Interoperable and Reusable (FAIR) principles stated in Wilkinson (2016), by adopting whenever possible already stated common frameworks, free software and open standards for Open Geospatial Consortiums (OGC) specifications (http://www.opengeospatial.org/standards). The INSPIRE directive is considered for describing metadata, and complying with the ISO 19115 and ISO 19139 metadata standards. The new metadata generated within EO4CerealStress framework will be checked using the INSPIRE Reference Validator (https://inspire.ec.europa.eu/validator/home/index.html).

Table 1. Example of metadata for geographic datasets as suggested in ISO 19115:2003. M=mandatory, C=conditional, O=Optional

Metadata Classes	
Dataset title (M)	Spatial representation type (O)
Dataset reference date (M)	Reference system (O)
Dataset responsible party (O)	Lineage (O)
Geographic location of the dataset (by four coordinate or by geographic identifier) (C)	On-line resource (O)
Dataset language (M)	Metadata file identifier (O)
Dataset character set (C)	Metadata standard name (O)
Dataset topic category (M)	Metadata standard version (O)
Spatial resolution of dataset (O)	Metadata language (C)

Abstract describing the dataset (M)	Metadata character set (C)
Distribution format (O)	Metadata point of contact (M)
Additional extent information for the dataset (vertical and temporal) (O)	Metadata date stamp (M)

3. Historical Campaign data

The current list of potential datasets is provided in Table 1.

Table 2. List of historical airborne campaign data

Name of the campaign	Location	Time	Crop types	Spatial and spectral resolution	Ground data type and availability
HyPlant-Italy FLEX Mission	Grosseto, Tuscany, Italy	16-24/06/2019	Maize	1.8 m 3.65 nm (VNIR) and 10.55 nm (SWIR) *FLUO*- 0.28 nm	Environmental data: from local weather station; Chlorophyll: SPAD-502 chlorophyll meter; Leaf water: pressure chamber; Stomatal conductance: SC-1 Leaf Porometer;
HyPlant-Germany (SoyFLEX) FLEX Mission	Selhausen area; and Klein-Altendorf area near Bonn,	30-3/07/2015	Maize and mixed crops. Soybean in the Klein-Altendorf	1 x 1 m 3.65 nm (VNIR) and 10.55 nm (SWIR) *FLUO*- 0.25 nm	TOC reflectance measured using ASD FieldSpec and SIF measured with high-resolution spectrometers. Leaf level SIF measured with ASD FieldSpec. *All ground data under the repository of Forschungszentrum Jülich group.
HyPlant-Italy FLEX Mission	Latisana, northeast Italy	June 2014	Grasses (Festuca arundinacea and Poa pratense)	1 x 1 m 3.65 nm (VNIR) and 10.55 nm (SWIR) *FLUO*- 0.25 nm	Meteorological Data: Local weather station; Canopy temperature sensor (Calex Electronics, Leighton Buzzard, UK) and chamber flux measurements.

SARSense, Germany	Selhausen, Germany	June and August 2019	sugarbeet, winter wheat, winter barley, maize, and rapeseed	Sub-meter spatial resolution.	Total Biomass (fresh), Total Biomass (dry), LAI and Canopy water content

Different field measurements of plants and soils were acquired during the summer of 2023 in the three pilot sites in Austria, Italy and Spain within the framework of EO2CerealStress.

Table 3. List of EO4CerealStress and complementary field campaign data

Name of the campaign	Location	Date	Crop types	Stress type	Parameters measured in the field/laboratory	UAV flight (spatial & spectral resolutions)
EO4CerealStress -Austria-	Marchfeld Region, Lower Austria, Austria	May-July 2023	Maize and Summer Wheat	Water stress	Detection of irrigated fields, soil water storage capacity (mm/dm)	
Landsupport Soil Data	Marchfeld Region, Lower Austria, Austria	May 2020	Any	Water stress	Soil depth (cm)	
EO4Cereal Stress -Spain-	Doñana, Andalusia, Spain	July-September 2023	Rice	Salinity, heavy metals, heatwaves	Plants: ASD field spectral measurements, SPAD, LAI (m^2/m^2), Leaf area (cm^2), chlorophyll A&B ($\mu g/cm^2$), carotenoids ($\mu g/cm^2$), wet & dry weight (g), humidity (%), N (%) Soils: ASD field spectral measurements, Electric conductivity (mS/cm)	Plants and soils: 3 parcels (4 cm, 400-2,500 nm & 537 bands)

SAMA	Doñana, Andalusia, Spain	July-September 2023	Rice	Nutrients	Plants: P, K, S (%), Ca, Mg, Fe, Zn (mg/kg) Soils: Organic Carbon, Total C, N, S (%), P, Fe, Mn, Zn, Cu (mg/Kg), Ca, Mg, K, Na (Cmolc/Kg), heavy metals: As, Cd, Cr, Pb, Al, Cu Grain: K, P, N, S, Ca, Mg (%), Fe, Zn (mg/Kg)	Plants and soils: 3 parcels (4 cm, 400-2,500 nm & 537 bands)
E04Cereal Stress -Italy-	Jolanda di Savoia Ferrara, Italy	May 2023	Winter wheat	lodging	ASD field spectral measurements, SPAD, Plant density, LAI (m2/m2), Canopy height (cm), Cover percentage (%), Shoot numbers, Tiller numbers, Fresh Biomass (g/m2), Leaf area (cm2), Flower weight (g/m2)	

4. Pilot sites and products

The pilot sites selected for this project correspond to agricultural areas in the regions of Lower Austria, Andalusia and Ferrara, in Austria, Spain and the Netherlands, respectively. These areas present a significant extension of corn, wheat and rice crops that may be affected by different abiotic stressors. Different rectangular AOIs of approximately 1000 km^2 have been established for which the datasets that intersect the polygons corresponding to each AOI will be obtained.

At this stage, it is not necessary to adopt any filtering criteria to consider only a better selection or subset, because only a rough estimation of the imagery collection to be considered is needed.

Table 4. List of the European pilot sites.

AOI ID	Name	Country	Region	Experimental site Area	Main Cropland
AU	Marchfeld	Austria	Lower Austria	1,120 km^2	Maize and wheat
ES	Guadalquivir marshlands	Spain	Andalusia	1,260 km^2	Rice
IT	Jolanda di Savoia	Italy	Ferrara	6.64 km^2	wheat

The list of EO4CerealStress products will be detailed in the final version of this document. The products (results) are stored in the internal repository and linked to the Open Science Data Catalogue: (https://opensciencedata.esa.int/).

5. EO Data

In this section, an initial evaluation of the data repository requirements based on the pilot sites, selected products and currently expected sizes and temporal frequencies of input images is presented.

5.1 DESIS

The DESIS (DLR Earth Sensing Imaging Spectrometer Mission) instrument is installed on the International Space Station (ISS). DESIS is a push broom imaging spectrometer spectrally sensitive over the VIS and NIR wavelength range from 400 to 1000 nm and 235 spectral bands, with a spectral sampling distance of 2.55 nm. The Ground Sampling Distance (GSD) at nadir view depends on the flight altitude of the ISS and is about 30 m resulting in a swath width of about 30 km, covering ~900 km^2. New acquisitions of DESIS are available under request. L2A products can be obtained from https://www.dlr.de/eoc/desktopdefault.aspx/tabid-8799/. The size of a DESIS image is ~2.3 GB. One DESIS image has been acquired for the ES AOI, and new imagery will be requested for other AOI during the project.

5.2 ENMAP

ENMAP (Environmental Mapping and Analysis Program) carries the push broom imaging spectrometer HSI. Its spectrally sensitive over VIS/NIR spectral bands with a spectral interval of 420-1000 nm and sampling interval of 6.5 nm, and SWIR spectral bands with a spectral interval of 900-2450 nm and sampling interval of 10 nm, with a total of 246 channels. The spatial resolution is 30 m, 30 km swath, nadir revisit time of 27 days and off-nadir (30°) and pointing feature for fast target revisit of 4 days. New acquisitions of ENMAP are available under request. L2A products can be obtained from https://planning.enmap.org/. The size of an ENMAP image is ~1.2 GB. Multiple ENMAP images are being acquired for all AOIS within the project.

5.3 ECO2LSTE V001

The ECOSTRESS (ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station) ECO2LSTE Version 1 data product provides atmospherically corrected land surface temperature and emissivity (LST&E) values derived from five thermal infrared (TIR) bands. The ECO2LSTE is provided at a spatial resolution of 70 meters (m). The corresponding ECO1BGEO data product is required to georeference the ECO2LSTE data product. The ECO2LSTE Version 1 data product contains layers of LST, emissivity for bands 1 through 5, quality control for LST&E, LST error, emissivity error for bands 1 through 5, wideband emissivity, and Precipitable Water Vapour (https://lpdaac.usgs.gov/products/eco2lstev001/).

5.4 PRISMA

PRISMA (PRecursore IperSpettrale della Missione Applicativa) carries two sensor instruments, the HYC (Hyperspectral Camera) module and the PAN (Panchromatic Camera) module. PRISMA is a push broom spectrometer spectrally sensitive over 66 VIS/NIR spectral bands with a spectral interval of 400-1010 nm, and 171 NIR/SWIR spectral bands with a spectral interval of 920-2505 nm, with 237 channels. The HYC module has a spatial resolution of 30 m, with a swath width of 30 km, and a field of regard of 1000 km either side. New acquisitions of DESIS are available under request. L2A products can be obtained from https://prismauserregistration.asi.it/. The size of a PRISMA image is ~5.6 GB. One PRISMA image has been acquired for the ES AOI, and new imagery will be requested for other AOI during the project.

5.5 MOD11A1 – LST

The Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Land Surface Temperature/Emissivity Daily (MOD11A1) Version 6.1 product provides daily per-pixel Land Surface Temperature and Emissivity (LST&E) with 1 km spatial resolution in a 1,200 by 1,200 km grid. The product is comprised of LST, emissivity, quality indicators and observation parameter (https://lpdaac.usgs.gov/products/mod11a1v061/).

MOD11A1 v6.1 LST images will be collected for the European Areas of Interest (AOIs) included in Table 2. Each AOI is covered by one single MODIS tile (see Table 3). On the basis of the collection, it is possible to roughly estimate the total numbers of acquisitions per site (Table 3) and the total size per AOI for the total period of the product/project (2000-02-24 to 2023-12-31). Therefore, the expected storage requirement for raw data can be considered less than 2.75 TBs for MOD11A1-LST (considering de layers LST_Day_1km and QC_Day). This initial size can be drastically reduced by cropping images by AOI.

Table 5. Number of MOD11A1 acquisitions from 2000-02-24 to 2023-12-31.

AOI ID	Name	Country	MODIS tile	MOD11A1 acquisitions	Average size per scene	Total size for period
AU	Marchfeld	Austria	h19v4	8712	~1.14 GB	0.83 TB
ES	Guadalquivir marshlands	Spain	h17v5	8712	~0.75 GB	1.08 TB
IT	Jolanda di Savoia	Italy	h18v4	8712	~0.75 GB	1.08 TB

5.6 Harmonised Sentinel-2 MSI

Sentinel-2 (S2) is a wide-swath, spatial high resolution, multispectral imaging mission with a global 5-day revisit frequency. The S2 Multispectral Instrument (MSI) is sensitive to VIS/NIR at 10 meters, red edge and SWIR at 20 meters, and atmospheric bands at 60 meters spatial resolution, with 13 spectral channels overall. In this product the Sentinel-2 L2 data are downloaded from scihub. They were computed by running sen2cor. The assets contain 12 UINT16 spectral bands representing SR scaled by 10000. See the Sentinel-2 User Handbook for details :

(https://sentinels.copernicus.eu/web/sentinel/user-guides/document-library/-/asset_publisher/xlslt4309D5h/content/sentinel-2-user-handbook).

In addition, three QA bands are present where one (QA60) is a bitmask band with cloud mask information. Sentinel-2 L2 assets have the following format:

COPERNICUS/S2_SR/20151128T002653_20151128T102149_T56MNN.

Here the first numeric part represents the sensing date and time, the second numeric part represents the product generation date and time, and the final 6-character string is a unique granule identifier indicating its UTM grid reference. Clouds can be removed by using COPERNICUS/S2_CLOUD_PROBABILITY.

Sentinel 2A/B satellites images will be collected for the European Areas of Interest (AOIs) included in Table 2. Sentinels images have been included in these datasets only when intersecting a suitable set of polygons (as defined by partner, see Table 2) corresponding to each AOI. At this stage, it is not necessary to adopt any purging criterion to consider only a better selection or subset of the initial set in Table 3, because only a rough estimation of the imagery collection to be considered is needed. Based on the collection, it is possible to roughly estimate the total numbers of acquisitions per site (Table 4) and the total size per AOI for the period covered by the project (2019-2023). Therefore, the expected storage requirement for raw data can be considered less than 2.5 TBs for Sentinel-2A/B, in compressed form[1]. This initial size can be drastically reduced by removing images with very high percentage of cloud cover and by cropping and mosaicking by AOI.

Table 6. Number of Sentinel-2 acquisitions from 2019-01-01 to 2023-12-31.

AOI ID	Name	Country	S2A/B MSIL2A/year	Average size per scene	Total size for period
AU	Marchfeld	Austria	290	0.70 GB	1.01 TB
ES	Guadalquivir marshlands	Spain	288	0.75 GB	1.08 TB
IT	Jolanda di Savoia	Italy	146	0.80 GB	0.584 TB

5.7 Sentinel-3 LST

SLSTR (Sea and Land Surface Temperature Radiometer) is a dual view scanning radiometer, with 9 spectral bands, on board of the Sentinel-3A and Sentinel-3B satellites. The SLSTR thermal bands are used to retrieve LST in the SLSTR Level-2 products. These products can be downloaded from the Copernicus Data Space Ecosystem (https://dataspace.copernicus.eu/). The global revisit time at the equator is less than one day when combining both satellites. Products are supplied in NetCDF-4 format in 3-minute 'granules'. The SL_2_LST product contains LST from the SLSTR instrument at a spatial resolution of 1 km calculated using a split-window approach. Validation provides evidence the product meets the mission requirements of being accurate to 1 K.

SL_2_LST images will be collected for the European Areas of Interest (AOIs) included in Table 2. Each AOI is covered by one single Sentinel-3 SLSTR tile (see Table 5). Based on the collection, it is possible to roughly estimate the total numbers of acquisitions per site (Table 5) and the total size per AOI for the period covered by the project (2019-2023). Therefore, the expected storage requirement for raw data can be considered less than 2 TBs. This initial size can be drastically reduced by removing images with very high percentage of cloud cover and by cropping and mosaicking by AOI.

Table 7. Number of Sentinel-3 LST acquisitions from 2019-01-01 to 2023-12-31.

AOI ID	Name	Country	S3_SL_2_LST /year	Average size per scene	Total size for period
AU	Marchfeld	Austria	1247	70 MB	436 GB

[1] The average size for a S-2A/B images for the different AOIs can vary depending on the tile, and cloud conditions.

ES	Guadalquivir marshlands	Spain	1029	58 MB	298 GB
IT	Jolanda di Savoia	Italy	1137	70 MB	398 GB

5.8 L2B-TROPOSIF

L2B-TROPOSIF is derived from TROPOMI (TROPOspheric Monitoring Instrument) instrument aboard the Copernicus Sentinel-5 Precursor mission launched in October 2017. This dataset provides daily ungridded (L2B) SIF and surface reflectance estimates derived from TROPOMI observations covering the period May 2018-March 2021. SIF data are estimated from two fitting windows: from the 743-758 nm window (baseline product) and from the extended 735-758 nm window (secondary product). The L2B TROPOSIF products are provided in netCDF-4 files. The size of each L2B-TROPOSIF image is ~180 MB. Therefore, the expected storage requirement for raw data can be considered less than 0.26 TBs.

6. Climatic Data

6.1 ERA5 Daily Aggregates - Latest Climate Reanalysis Produced by ECMWF

ERA5 is the fifth generation ECMWF atmospheric reanalysis of the global climate. Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset. ERA5 data are available in the Climate Data Store on regular latitude-longitude grids at 0.25° x 0.25° resolution, with atmospheric parameters on 37 pressure levels. ERA5 DAILY provides aggregated values for each day for seven ERA5 climate reanalysis parameters: 2m air temperature, 2m dew point temperature, total precipitation, mean sea level pressure, surface pressure, 10m u-component of wind and 10m v-component of wind. Additionally, daily minimum and maximum air temperature at 2m has been calculated based on the hourly 2m air temperature data. Daily total precipitation values are given as daily sums. All other parameters are provided as daily averages. ERA5 data is available from 1979 to three months from real-time. (https://climate.copernicus.eu/climate-reanalysis). Although, only the last three decades (1993-2023) of data are being considered for this project in order to compute climatic normals. The expected storage requirement for raw data can be considered less than 3 GBs.

6.2 Local meteorological stations

The sources of climate information used for the project, in addition to grids, will consider specific weather local stations from different sources. In this way, the climate can be monitored during the project execution period in real time, since the climate grids are published with a time lag. A list of potential weather stations at country and regional level is givel below:

(i) REDIAM (Andalusian Environmental Information Network).
(ii) IFAPA (Agrarian and Fisheries Training Institute of Andalusia).
(iii) SIAR (Agroclimatic Information System for Irrigation (SiAR) of Spain.
(iv) SPARTACUS v2.1. Daily gridded dataset of air temperature and precipitation in 1-kilometer resolution over Austria.

7. Summary of space requirements by product/sensor

Table 8. Summary of space requirements by sensor/product and AOI.

Sensor/product	Time period	AOI: AU	AOI: ES	AOI: IT	Total size
DESIS	2023	***	2.3 GB	***	2.3 GB
ENMAP	2023	1.2 GB	3.6 GB	1.2 GB	6.0 GB
PRISMA	2023	***	5.6 GB	***	5.6 GB
MOD11A1– LST	2000-2023	0.83 TB	1.08 TB	1.08 TB	2.99 TB
S2A/B MSIL2A	2019-2023	0.83 TB	1.08 TB	0.58 TB	2.49 TB
S3_SL_2_LST	2019-2023	0.44 TB	0.30 TB	0.40 TB	1.14 TB
L2B-TROPOSIF	2018-2021	***	***	***	0.26 TB
ERA5 Daily	1993-2023	***	***	***	3GB
				TOTAL	7.08 TB[2]

[2] The total space requirements can be drastically reduced (~80%) by adjusting the geographical extent to the AOI.