Anthropogenic Water Use (CCI-AWU)

Deliverable 2:
Report explaining the criteria for selecting test regions.

Date	Issue	Section	Page	Comment
15/04/2024	1.0			

Control Document

Process	Name	Date
Written by:	Luca Brocca, Sara Modanesi, Jacopo Dari, Christian Massari, Carla Saltalippi, Renato Morbideli, Gabrielle De Lannoy, Michel Bechtold, Louise Busschaert, Zdenko Heyvaert, Wouter Dorigo, Pierre Laulet, Pia Langhans, Maria Cristina Rulli, Davide Danilo Chiarelli, Nikolas Galli	
Checked by	Luca Brocca,	date

	Signature	Date
For CCI AWU team		date
For ESA		

[This page is left intentionally blank]

List of content

Contents

1. Introduction

1. 1.1 The CCI-AWU project

The closure of the Earth's water cycle (as well as the energy balance and the carbon cycles) through satellite Earth Observation (EO) represents one of the outstanding scientific challenges highlighted by the Global Climate Observing System (GCOS). Required standards of accuracy are fixed to 5% and annual timescale. To this end, a suite of essential climate variables (ECvs) has been defined to understand the evolution of climate and to assess the potential derived risks. However, if targets at annual timescale can generally be reached, larger uncertainties are observed for sub-annual and sub-continental time and spatial scales, respectively (Dorigo et al., 2021; Rodell et al., 2015). In this context, an ECV including the information on anthropogenic water use (AWU) can help in advancing the proper closure of the water cycle at higher spatial and temporal scales. In the Climate Change Initiative – Anthropogenic Water Use (CCI-AWU) precursor project, AWU is more specifically intended as agricultural water allocated for irrigation, which represents the largest anthropogenic water use, thus making irrigation being the most impactful human activity on the hydrological cycle. FAO (2016) estimated that irrigation, worldwide, accounts for more than 70% of water withdrawn from surface (i.e., rivers, lakes) and subsurface (i.e., groundwater) water sources and these estimates are expected to increase in the near future due to an increase in population and in food production, especially over arid and semi-arid regions (McDermid et al., 2023). In this context, the main data source identified by GCOS for tracking AWU is from FAO's AQUASTAT. However, AQUASTAT provides survey-based irrigation estimates which do not meet the GCOS requirements, i.e., data are provided on a 5-years interval instead of yearly, and are available every 2-3 years.

The overarching objective of Climate Change Initiative – Anthropogenic Water Use (CCI-AWU) precursor project is to derive long-term (i.e., twenty years) AWU time series for selected regions using several approaches exploiting remote sensing observations, as a proof-of-concept of the feasibility towards a proper AWU ECV product.

The CCI-AWU project involves a consortium led by CNR-IRPI and comprise the following organisations:

1. Vienna University of Technology (TU Wien), hereinafter TUWIEN;
2. KULeuven, Department of Earth and Environmental Sciences, Division Soil and Water Management (KATHOLIEKE UNIVERSITEIT LEUVEN), hereinafter KULeuven;
3. University of Perugia (UNIVERSITY OF PERUGIA), hereinafter UNIPG
4. Politecnico di Milano (POLITECNICO DI MILANO), department of Civil and Environmental Engineering, hereinafter POLIMI

1.2 Scope of this report

This document aims at describing the criteria defined for selecting the study areas, data collection of in situ benchmark data and satellite data. More specifically, the data collection involved the following variables: in-situ irrigation benchmark data, actual and potential evaporation (ET), forcing datasets to drive AWU algorithms, satellite soil moisture datasets. The data have been collected and organised in a project database for sharing among the team, with the primary goal of ensuring that all team members are using consistent data versions. In the next phase of the project, efforts will focus on strengthening collaborations with research, academic and climate institutions to collect additional benchmark data over the selected test regions.

1.3 Applicable Documents

- Proposal.
- Deliverable D1. User Requirement Document (URD) V.1

1.3 Document Organization

The document is organised as follows: in Section 2 the study areas are described, together with the collected in situ benchmark irrigation datasets; in Section 3 we provide information on the boundaries selected for the study areas and the reference spatial grid that will be used to run the AWU algorithms in Task 3; finally, Section 4 provides information on the forcings/satellite data products collected to start the next steps of the project.

2. Study regions and in situ data collection

The choice of the pilot regions was driven by: 1) the necessity to focus over highly irrigated areas; 2) the importance of representing contrasting climatic, geophysical and geomorphological conditions as well as different study areas' extent; and finally 3) the opportunity to collect long-term in situ benchmark data that will be used to test and to validate the project's outputs, based on international collaborations. Based on that, areas selected for the AWU algorithms/modelling implementations include the Ebro river basin (Spain), the Murray Darling basin (Australia), the Contiguous United States (CONUS, USA) and India (whole country). This section provides a description of the pilot areas as well as the in-situ datasets collected on the different sites.

2.1. Ebro River Basin

The Ebro River is characterised by a length of 910 km and a drainage area of approximately 85,362 km^2, being the Spanish river with the highest mean annual flow. Precipitation is irregular in time and space with a high degree of interannual variability, especially in the Mediterranean side. Intra-annual precipitation is concentrated in spring and autumn with dry summers.

The Ebro basin sustains agricultural activities. Here, large areas are irrigated with different techniques, contrasting with the semi-arid environment that surrounds them. Runoff is mainly generated on the Pyrenees, and is stored in numerous dams and transported to the irrigation areas by means of canals. Irrigation has a strong impact on the hydrological functioning of the basin, thus it is necessary to better account for its impacts on the different hydrological fluxes (transpiration, streamflow) and stocks (soil moisture).

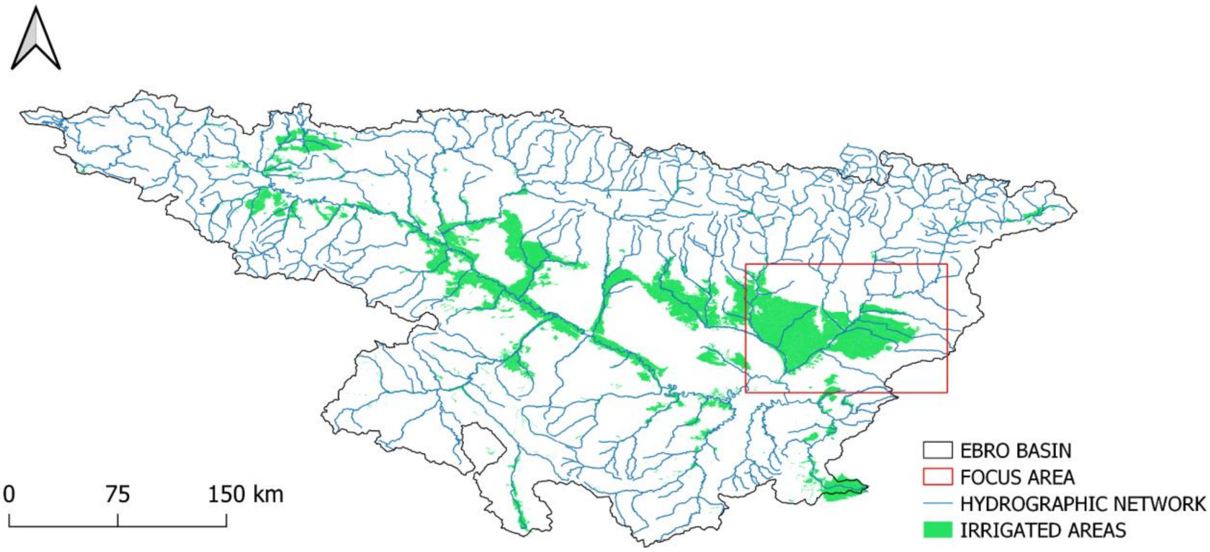

Figure 1. Irrigated areas in the Ebro River basin selected for validation in the CCI AWU project, highlighted in the red box (Aragon and Catalonia Canal, Algerri Balaguer Canal, Urgell Canals and Pinyana Canal).

2.1.1. Ebro benchmark irrigation data

Four main agricultural districts belonging to a heavily irrigated area located in the Northeastern portion of he basin were selected for validation of the CCI AWU irrigation products: the Urgell district, the Algerri Balaguer district, the Pinyana district and the Catalan and Aragonese district (divided into North and South according to the irrigation canals network). The selected areas are located across the regions of Aragon and Catalonia drained by the Cinca and the Segre rivers and their tributaries; irrigated fields are fed by a canal network originating from a system of reservoirs upstream the agricultural areas. In order to derive the benchmark irrigation rates, records of water volumes flowing through the following canals were considered: Aragon and Catalonia Canal, Algerri-Balaguer Canal, Urgell Canals, and Pinyana Canal (Figure 2).

The Aragon and Catalonia canal was inaugurated in 1906 and irrigates a surface of 98,000 ha, distributed between the provinces of Huesca, in Aragon, and Lleida, in Catalonia. It is the second largest irrigated area in the Ebro basin and one of the largest in Spain. It is fed mainly by the waters of the Esera river, a tributary of the Cinca river, taken from the Joaquin Costa reservoir, also called Barasona, and has a length of 124 km. Sprinkler irrigation is the most widespread technique (54% of the area), followed by drip (28%) and flood (18%) irrigation. For this project the area is divided in two districts (North and South) according to the structure of the irrigation canals network.

The Algerri-Balaguer canal was built in the 1990s and irrigates a surface of 8,000 ha located north of Lleida, between Alfarrås and Balaguer. It pumps water from the Noguera-Ribargorqana river. Together with the building of the hydraulic infrastructure a concentration of properties was performed, which allowed to merge the properties of each owner in a single farm (before they were small and scattered) In this way, the application of modern farming and irrigation techniques was made possible. Drip and sprinkler irrigation is used for fruit trees and crops, respectively.

The Urgell canals were built at the end of the XIX century and irrigates a surface of 70.000 ha, distributed on the left bank of the Segre river. Water is transported from the Segre river by means of a 144 km long canal (main canal). Being a centennial hydraulic infrastructure, the irrigation techniques applied in the area are mostly traditional (inundation by gravity).

The Pinyana canal was built in the 13th century, being the oldest canal of Catalonia and irrigates a surface of 13,891 ha, located North of Lleida. It takes water from the Santa Anna reservoir on the Noguera-Rivagorqana river. In this irrigated area, irrigation is also mostly traditional.

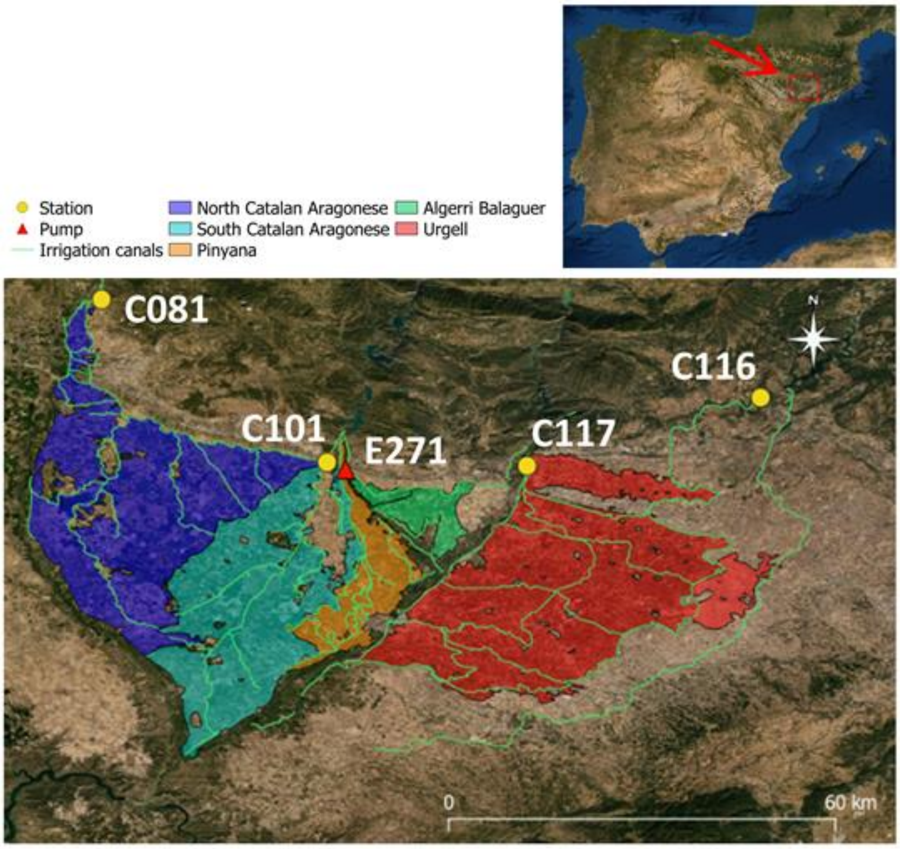

Figure 2. overview of the irrigation infrastructure currently working on the selected area.

To establish benchmark irrigation data, the water volumes flowing through irrigation canals and supplying the districts are used. These water volumes are subsequently divided by the area of the irrigation districts to calculate the equivalent thickness of irrigation water. The data (except for Pinyana) are provided by the Automatic Hydrologic Information System of the Ebro river basin (SAIH Ebro) available at :http://www.saihebro.com/saihebro/index.php?url=/datos/canales. Further information about the districts and the related data are given as follows:

- The Urgell district has an area of 887.62 km^2. The stations of interest for the are C116 and C117 (See Figure 2). 30% of losses is assumed, according to Dari et al. (2020).
- The Algerri Balaguer district has an area of 70.79 km^2. Benchmark data here is represented by the water pumped from the river to the district monitored by station E271. According to Dari et al. (2020), losses equal to 10% are assumed.
- Regarding the Catalan and Aragonese district, the northern portion has an area of 657.04 km^2 while the extent of the southern portion is 504.48 km^2. For the northern part, the reference data comes from measurements recorded at station C081, while for the southern part station C101 must be considered. According to Dari et al.(2020), losses equal to 15% are assumed.
- The Pinyana district has an area of 149.74km^2. It is not part of the SAIH system, so the only available data here is an average monthly consumption provided by the district managers. According to Dari et al. (2020), losses equal to 30% are assumed.

Benchmark data are organised in a netcdf file that is available at the following link:

https://www.dropbox.com/scl/fo/o5gzpwwmwnpwqmjpw8q2v/AJiYkyZ8IODVtVMienbiNN4?rlkey=yxevqqtf80qdryu6zlymq0kf4&st=h1oskgvs&dl=0

The Irrigation_Test_Site_Ebro.nc file contains district-aggregated irrigation amounts [mm] over the pilot districts. Loss coefficients are already taken into account and applied. For Urgell, Algerri Balaguer, and North and South Catalan Aragonese, long-term daily data is available (1997-2023 fro Urgell and South Catalan Aragonese and 2007-2023 for Algerri Balaguer and North Catalan and Aragonese). An example time series of daily irrigation data for the Algerri Balaguer district is provided in Figure 3. Only monthly data are available for Pinyana for a limited period (2016-2017).

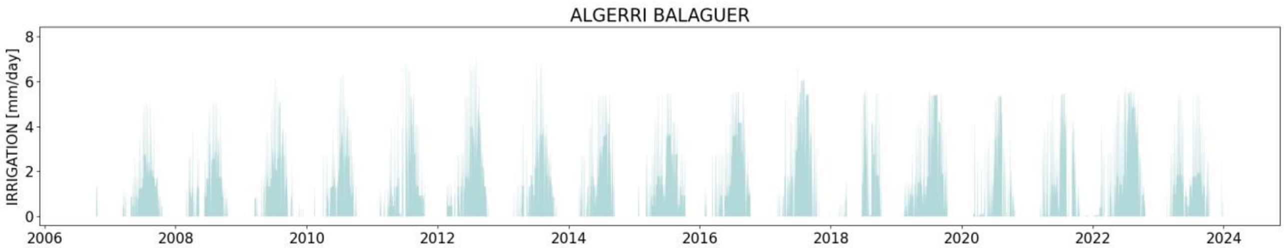

Figure 3. Daily benchmark irrigation over the Algerri Balaguer district for the period 2007-2023.

2.2. Murray Darling basin (Australia)

The Murray–Darling basin (2530 km long) is often considered the food bowl of Australia, covering over 1.1 million square kilometres, approximately 14% of Australia and accounts for over two-thirds of all of Australia's irrigation water use. The basin is often subject to extreme droughts such as the Millennium Drought (2001–2009) (van Dijk et al., 2013) and more recently the 2017–2019 drought. The basin is characterised by an arid and semi-arid climate that is strongly influenced by variations in climate such as the El Niño phase of the El Niño/Southern Oscillation (ENSO), favouring the occurrence of droughts across south-eastern Australia (Athukoralalage D. et al., 2024).

The Murray–Darling basin is dominated by vast plains, bounded to the east and south by the Great Dividing Range which reaches 2228 m asl. Under natural conditions the Murray River flows to the Southern Ocean via the Lower Lakes System and has three major tributaries: 1) the Goulburn Broken (~ 300 km); 2) the Darling (~ 2750 km); and 3) the Murrumbidgee (~ 1700 km). Under natural conditions they supply 17%, 14% and 12% of total Murray River inflows respectively (Leblanc et al., 2012; CSIRO, 2008)

Around 42% of surface water is harvested in farm dams and from water courses, being used mainly for irrigation (Leblanc et al., 2012). There is typically more irrigation in the southern side of the basin, which is facilitated by major storage capacities in the region. Irrigated properties are often fragmented in nature and contain a wide range of crops that are suited to the vastly different conditions that are observed across the basin. Both surface water and groundwater are used across the catchment with groundwater use mainly associated with major alluvial systems.

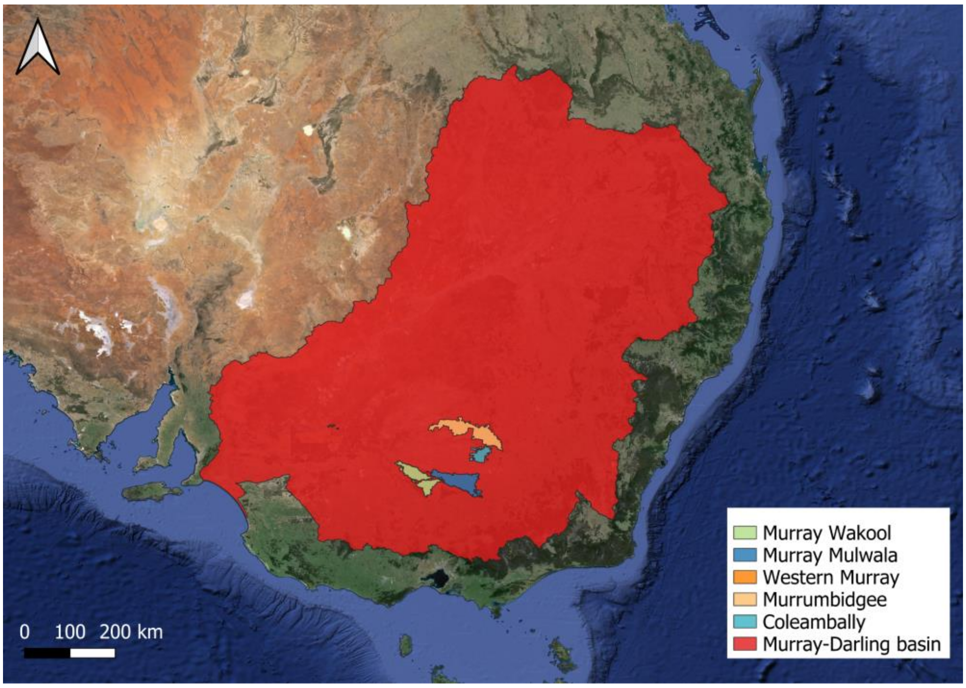

Figure 4. Murray Darling basin: catchment boundaries and irrigated districts, including Murray Wakool, Murray Mulwala, Western Murray, Murrumbidgee and Coleambally

2.2.1. Murray Darling benchmark irrigation data

Five irrigation districts located in New South Wales have been selected to collect benchmark irrigation data: Coleambally (977.0 km2), Murrumbidgee (2789.3 km2), Western Murray (49.1 km2), Murray Mulwala (3092.6 km2), and Murray Wakool (1455.2 km2); see Figure 4 for a map of the Murray Darling basin and the irrigated districts. Four out of the five pilot districts have an extension ≥ 1000 km^2, which is an important feature for validation purposes, since mainly coarse spatial resolution AWU estimates (25 km spatial resolution) will be developed. The districts are managed by the irrigation infrastructure operator (IIO), which is responsible for the production of annual reports on the irrigation water withdrawals (Bretreger et al., 2020). In particular, the Murray Mulwala and Murray Wakool districts belong to the same IIO infrastructure. Long-term monthly irrigation time series are available, since records starting from 2003 for some areas do exist. Moreover, the evolution in time of the portion of each district's total area actually irrigated is known. The monthly irrigation amounts are expressed as water thickness[mm/month], obtained after dividing the irrigation volumes by the reference areas.

Benchmark data are organised in netcdf files available at:

https://www.dropbox.com/scl/fo/mygfylgcxw4j34jtltvbd/AI3IzZBj9OykAo3IPuEn9qQ?rlkey=dxzp3mlglhyew6nsdun809o3p&st=i33gqaj5&dl=0

Further information about the districts and the related data are given as follows:

- Benchmark data for the Coleambally area covers the period July 2009 - June 2019. The irrigation amounts have been calculated by considering the evolution of the irrigated areas within the district during 2009-2013 (period 1, reference area: 2007), 2013-2017 (period 2, reference area: 2013) and

2017-end (period 3, reference area: 2017). Figure 5 shows an example time series of the monthly irrigation over the districts.

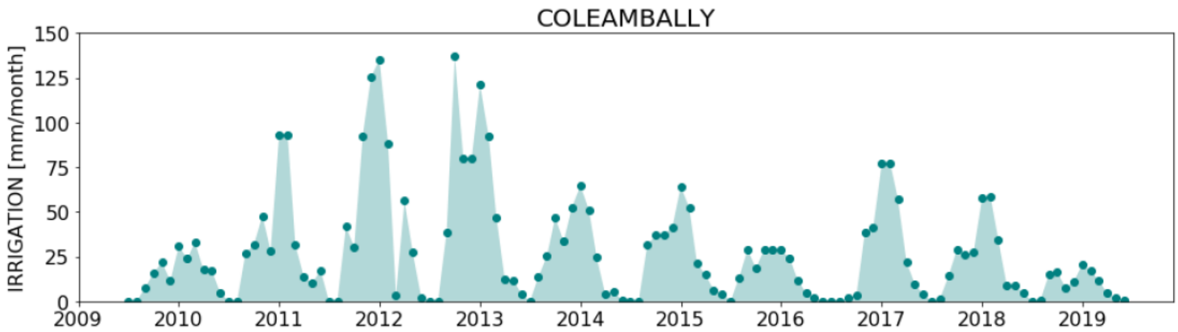

Figure 5. Coleambally monthly irrigation for the reference period 2009-2019

- The data set for the Murray Mulwala area covers the period July 2010 - June 2019. Similarly to Coleambally, the irrigation amounts have been calculated by considering the evolution of the irrigated areas within the district during 2010-2013 (period 1, reference area: 2007), 2013-2017 (period 2, reference area: 2013) and 2017-end (period 3, reference area: 2017).
- The data set for the Murray Wakool area covers the period from July 2010 to June 2019. The irrigation amounts have been calculated by considering the evolution of the irrigated areas within the district during 2010-2013 (period 1, reference area:2007), 2013-2017 (period 2, reference area2013) and 2017-end (period 3, reference area:2017).
- The data set for the Murrumbidgee area covers the period from July 2013 to June 2019. The irrigation amounts have been calculated by considering the evolution of the irrigated areas within the district during 2013-2017 (period 2, reference area: 2013) and 2017-end (period 3, reference area: 2017).
- The data set for the Western Murray area covers the period from July 2003 to June 2019. The irrigation amounts have been calculated by considering the evolution of the irrigated areas within the district during 2003-2013 (period 1, reference area: 2007), 2013-2017 (period 2, reference area:2013) and 2017-end (period 3, reference area:2017).

2.3. Contiguous United States (CONUS, US)

The CONUS area was chosen as a test region because of the vastness of its irrigated areas, whose extent is compatible with the spatial sampling of the AWU products to be developed. Specific portions of the CONUS have been considered as irrigation hotspots in previous studies aimed at assessing the detectability of irrigation signals from satellite data (Lawston et al., 2017; Zaussinger et al., 2019). The California Central Valley is one of these, as it accounts for the highest rates of irrigation water withdrawals across the CONUS. It is characterised by mixed climatic conditions, with a Mediterranean climate in the North and both hot and cold semi-arid climates in the South (Zaussinger et al., 2019). Rice cultivation is widespread in the Northern portion of the valley, where fields are kept flooded during the whole growing season. Nebraska High Plains represent another irrigation hotspot, with a cold semi-arid climate in the Western portion switching to humid continental conditions towards the East. Over this area, soybean and corn are the most widespread crops, and centre pivot irrigation systems are very common (Zaussinger et al., 2019). The aforementioned areas clearly emerge from the GMIA (Global Map of Irrigated Areas) data set (Siebert et al., 2015) shown in Figure 6, together with other well-known irrigated areas such as portions of the Idaho, the Mississippi Floodplain, and the Saint Luis Valley.

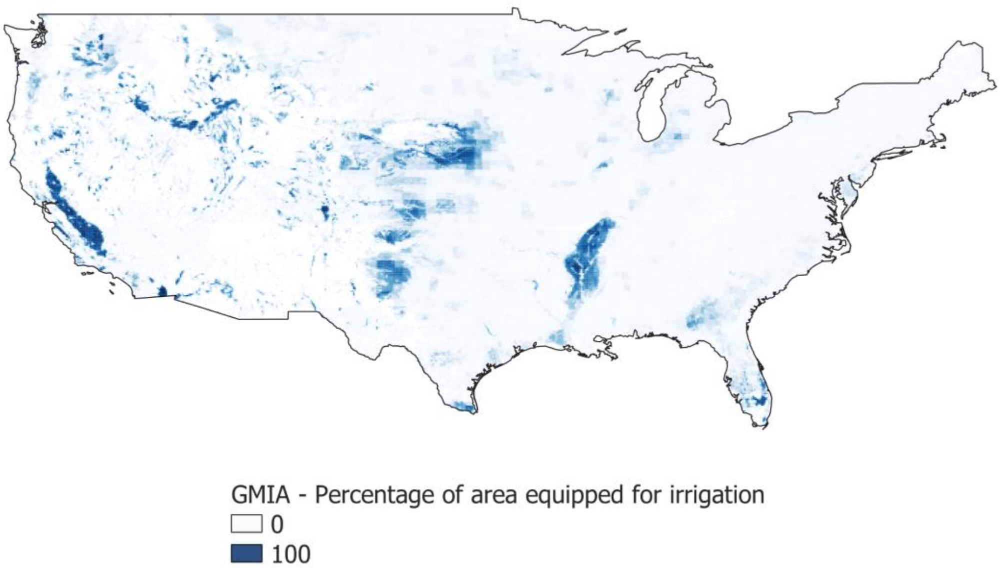

Figure 6. Percentage of area equipped for irrigation according to the GMIA (Global Map of Irrigated Areas) data set over the CONUS.

2.3.1. CONUS benchmark irrigation data

Information on irrigation dynamics (irrigated areas and amounts of water applied) over the CONUS were derived by the 2018 Farm and Ranch Irrigation Survey (FRIS) produced by the National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA). It is based on a collection of information coming from surveys interesting approximately 35000 farms practising irrigation. The database consists of country-aggregated information about irrigated areas and irrigation water withdrawals subdivided by specific crop type, water source and irrigation technique. Additional metadata such as partition between crops cultivated indoors (e.g., in greenhouses) and outdoors are provided. Figure 7 shows state-level irrigation water withdrawals and applied rates for irrigation in panels a) and b), respectively. Amounts of applied water are derived by dividing water withdrawals by the extent of irrigated areas, which is also provided by the FRIS.

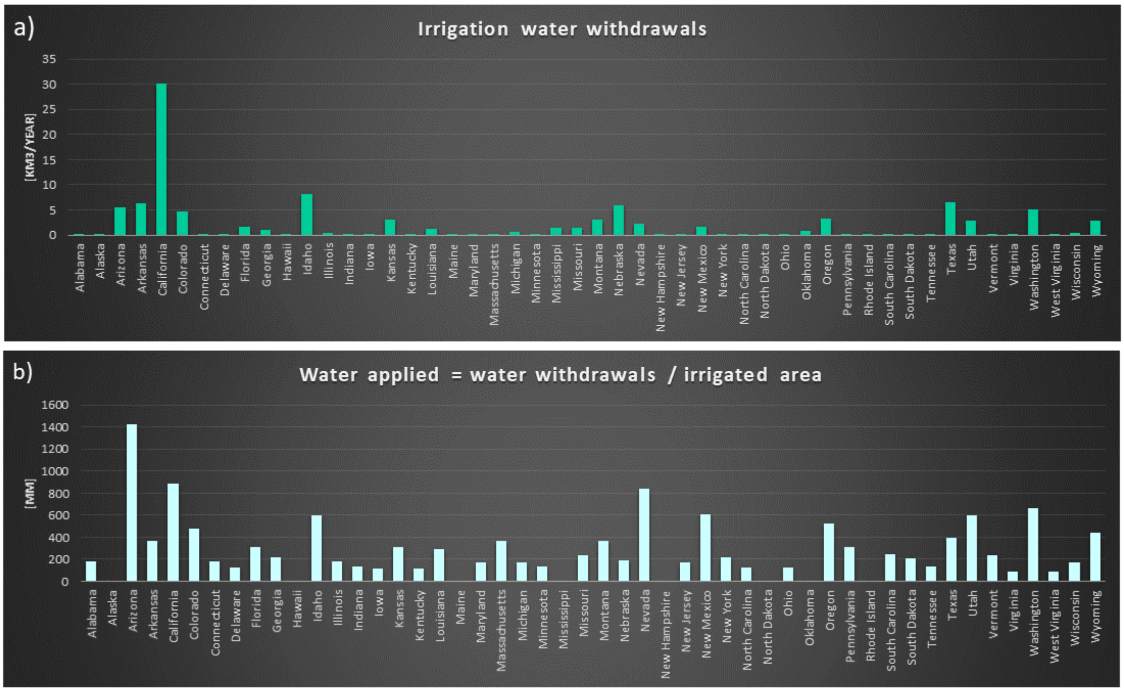

Figure 7. State-level irrigation water withdrawals and amounts of water applied over the CONUS according to the FRIS 2018.

Considering the nature of the FRIS database, which is a state-aggregated and survey-derived statistics, the possibility of obtaining additional data consisting in time series of actual irrigation applied is currently under exploration. In order to do this, a request has been forwarded to colleagues potentially having access to this kind of information.

2.4. India

The largest part of India is characterised by a tropical monsoon climate that supports rainforests, with the rest defined as semi-arid tropics (in accordance with Koppen-Geiger climate types, Peel et al., 2007). The climate is dominated by the summer monsoon, with the year divided into two main seasons: (1) summer monsoon or *kharif* season, between June and September; and (2) winter or *rabi* season, from October to March (Modanesi et al., 2020). In particular, the crops are sown in the winter months and require that sufficient soil moisture is available for germination and early growth. Considering that rainfall amounts are not adequate during the rabi season, irrigation is essential to ensure crops growth.

In this context, India is the world's largest user of groundwater resources (Aeschbach-Hertig and Gleeson, 2012; Girotto et al., 2017), and irrigation accounts for more than 85% of its groundwater withdrawals (Food and Agriculture Organization, 2013). As an example, total groundwater storage in northwestern India has undergone a decline, which is likely linked to irrigation-induced groundwater pumping (Rodell et al., 2009; Asoka et al., 2017).

This makes India a hotspot for water shortages due to agricultural water use and at the same time one of the most critical areas for collecting reliable benchmark irrigation data. In order to validate irrigation estimates generated within the project, information on irrigated areas (irrigated hectares) at state level were collected based on a collaboration with Prof. Manuela Girotto, University of Berkley. Additional information available over India include high resolution/annual irrigated-area maps from the work by Ambika et al. (2016) which can be used to validate the spatial patterns of irrigated lands. The maps, with a spatial resolution of 250 m, were developed using the Normalized Difference Vegetation Index (NDVI) data from Moderate Resolution Imaging Spectroradiometer (MODIS).

2.4.1. India benchmark irrigation data

Additional access to long-term dataset of benchmark irrigation over the Sina basin (Maharashtra, India; Amarasinghe et al., 2021) was possible due to an established collaboration between CNR-IRPI and the International Water Management Institute (IWMI) of Sri Lanka.

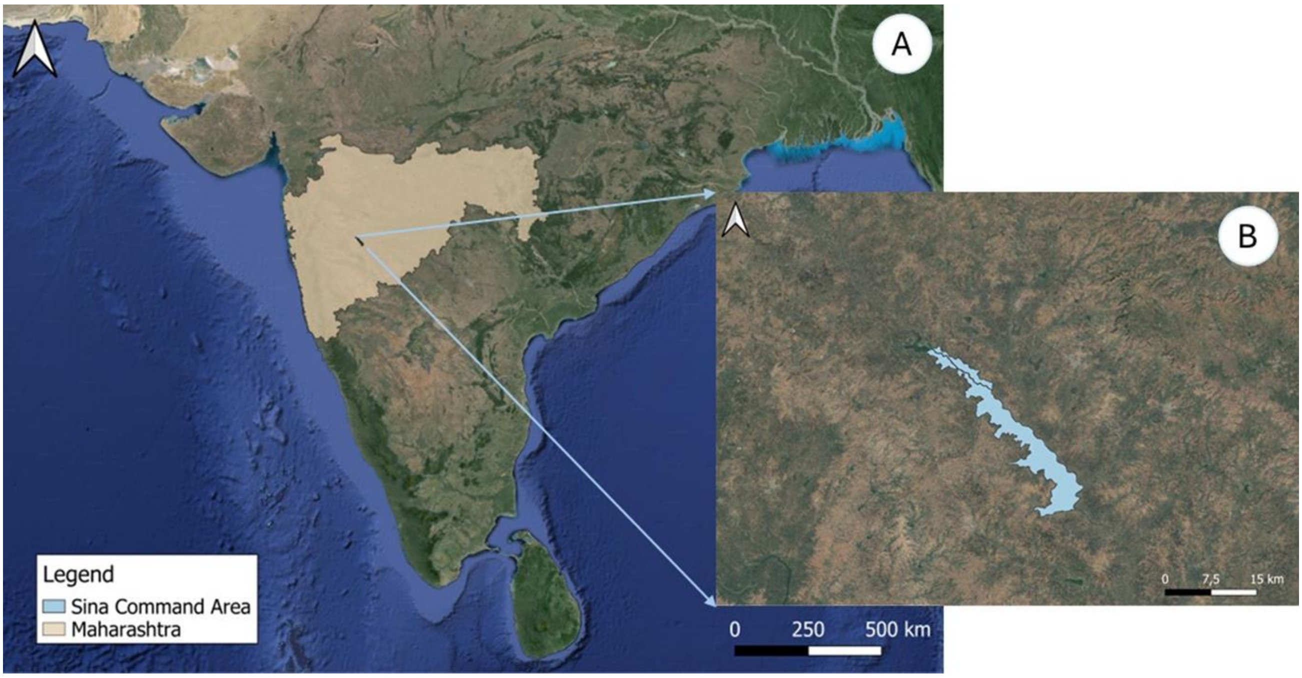

Figure 8. Reference study area for the irrigation benchmark data. A: Maharashtra State; B: Sina Command Area irrigated.

The Sina irrigation system is located in Maharashtra (Figure 8), India's third-largest state, with a total of 34 districts and a surface area of 307,713 km^2. Sina is a water-scarce medium irrigation system, and the reservoir has gross and live storage capacities of 67.95 million cubic metres (Mm^3) and 52.30 Mm^3, respectively.

Irrigation data are provided on a monthly time scale and expressed as water thickness [mm/month], referring to the entire reference study area (Figure 8b) for the period June 2008 to May 2018. An example time series which describes the benchmark irrigation data is provided in Figure 9.

Additional data of Evapotranspiration referred to the reservoir located north-west the Sina Command Area is provided and stored in the netcdf file created for the AWU project’s use, available at:

https://www.dropbox.com/scl/fo/p4cdew521sraupgbv11uf/AAEc_qK_95CXiZLLKs0EW70?rlkey=3vx73sarrs7cflm0fqcuoza4n&st=2fulzqqx&dl=0

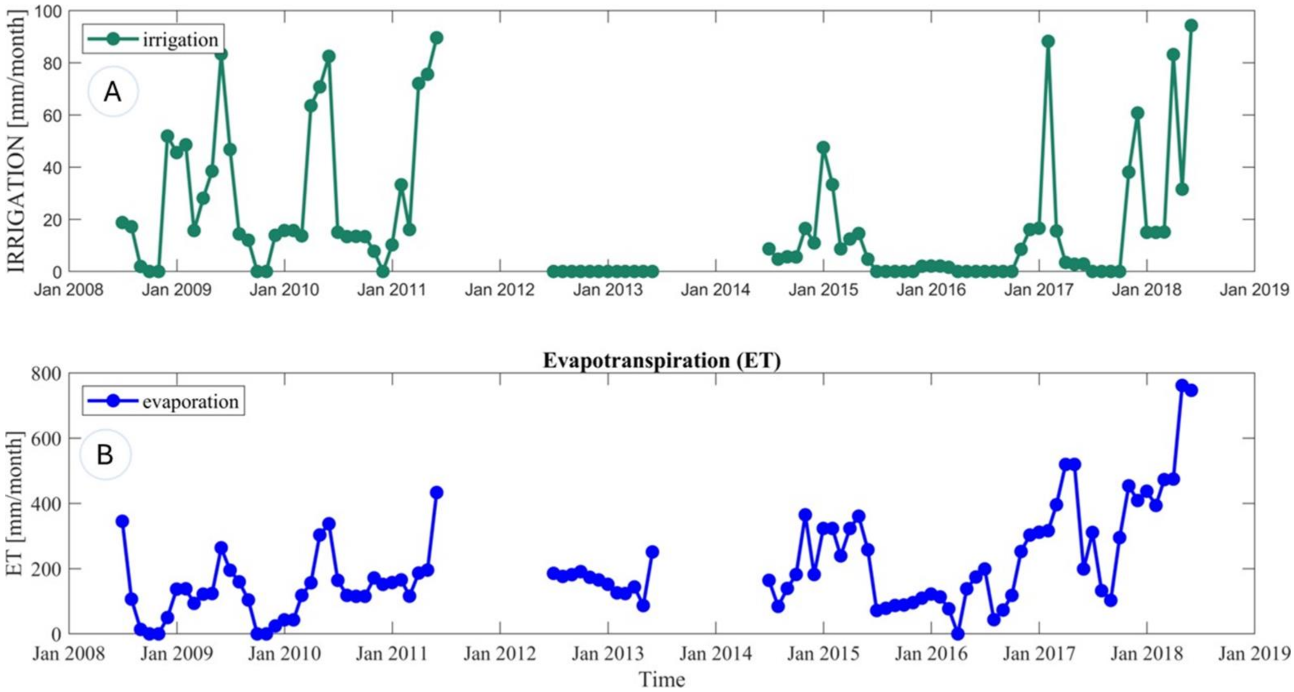

Figure 9. A: Sina Command Area monthly irrigation for the reference period 2008-2018; B: evaporation from the reservoir located north-east the Sina Command Area

The possibility of obtaining additional benchmark data is currently under exploration. In order to do this, a request has been forwarded to colleagues potentially having access to this kind of information.

3. Reference grid and study areas boundaries

For the CCI-AWU project purposes the consortium decided to run all the experiments over a common grid based on the ESA CCI Surface Soil Moisture COMBINED active+passive product (Dorigo et al., 2017; Gruber et al., 2019) v08.1 at a 25 km spatial resolution. The grid is accessible at:

https://www.dropbox.com/scl/fo/4o0qzhj8sw7nidwim1y7j/h?rlkey=icwryroepee9dl0x9kc9dzrhk&st=qrly6ubf&dl=0

Specific boxes were defined in order to allow a better comparison of the project's results in terms of long-term datasets of irrigation. Table 1 shows the boxes' limits.

Table 1: Limits of the boxes defined for each of the project's pilot areas.

Study Area	**Lower left lat (°)**	**Lower left lon (°)**	**Upper right lat (°)**	**Upper right lon (°)**
CONUS	24.875	-124.875	49.375	-66.875
Ebro	40.125	-4.625	43.375	2.375
India	7.875	68.125	37.125	97.625

Murray-Darling	-37.875	138.375	-24.375	152.625

4. Forcing and satellite data collection

In this section the selected datasets collected in the first 6 months of the project, which will be used in Task 3 (Algorithm Development) are described.

4.1 ERA5 forcing precipitation

To ensure uniformity across the consortium in applying the selected AWU algorithms, the ERA5 (European ReAnalysis v5) (Hersbach et al. 2020) lowest model level forecasts with an original resolution of 0.25°x0.25° were collected. ERA5 is a widely used climate reanalysis product. Forcing data were obtained and regridded over the CCI soil moisture grid within the specified analysis areas (the reader can refer to Section 3 for details on the extent of the selected study areas). The dataset was regridded using a bilinear interpolation method, allowing for consistent data integration and compatibility across different algorithm implementations within the consortium. Data are accessible at the following link:

https://www.dropbox.com/scl/fo/4wufw4yzehapt3qaxb09t/AMBpaKdwlXCmIS8jGw5MM9w?rlkey=weff0gbp1ee9upnuuc0ncecdg&st=xnra46zi&dl=0

4.2 SMOS Level 2 Soil Moisture

The Level 2 Soil Moisture Ocean Salinity (SMOS; Kerr et al., 2012; Kerr et al., 2014) Soil Moisture product version 700 consists of Swath-based retrieved information over land surfaces (and sea ice) from the SMOS L1c product. It includes soil moisture measurements geo-located in an equal-area grid system ISEA 4H9 along with various ancillary data such as nadir optical thickness, surface temperature, roughness parameter, dielectric constant, and brightness temperature retrieved at the top of the atmosphere and at the surface, each with corresponding uncertainties. To ensure ease of use for all project partners, the SMOS soil moisture was processed in the modelling chain of Noah-MP Land Surface Model (LSM; Niu et al., 2011) into the Land Information System (LIS) framework. This allowed us to obtain SMOS observations regridded over the CCI SM grid using a nearest neighbour approach. The process was applied to all the selected study areas. The recommended flags (including RFI probability, poor retrievals, barren, forests, mountainous terrain, open water, coastal, urban areas) were already applied within the modelling chain. Data are stored at the following link:

https://www.dropbox.com/scl/fo/63bqx2y1cb4hjl52nrslp/ANH3wNDfgzfB4dcLgaFKcdM?rlkey=oga3ixtwgwcjdj7opcewuoefd&st=ka6yjojk&dl=0

4.3 SMAP Level 2 Soil Moisture

Retrievals of the Soil Moisture Active Passive (SMAP; Chan et al., 2016; O'Neill et al., 2021) Level 2 surface soil moisture were used in this project (version 8). These retrievals are derived from L-band radiometer measurements using a dual-channel algorithm and represent the top 5 cm of the soil. The data are provided on the 36 km Equal-Area Scalable Earth Grid, Version 2.0 (EASE-Grid 2.0; Brodzik et al., 2012) which was resampled to the CCI SM grid via a nearest neighbour approach. Both descending and ascending overpasses are used, which correspond to local collection times of 06:00 and 18:00 h, respectively. Only retrievals with a recommended quality (based on dense vegetation, mountainous terrain, open water, urban area, precipitation, and frozen ground) were used. Data are stored at the following link:

https://www.dropbox.com/scl/fo/r9z8jrtsipgy6y3zh0q4m/AOHgG5RzoOOwnCZt6U4MsF4?rlkey=uf5n2nn4sh057m4g9x2kzkdoz&st=9t546sh3&dl=0

4.4 CCI Soil Moisture

ESA has supported the development and production of the first multi-decadal, global satellite-observed soil moisture (SM) dataset as part of its Climate Change Initiative (CCI) program. This product, named ESA CCI SM (Dorigo et al., 2017), combines various single-sensor active and passive microwave soil moisture products into three harmonised Climate Data Records: a merged active microwave, a merged passive microwave, and a combined active+passive microwave products. ESA CCI SM covers the period from 1978 to the current year. All products contain data about global daily surface soil moisture state with a spatial resolution of 0.25° and are publicly available (https://climate.esa.int/en/projects/soil-moisture/). The ESA CCI SM algorithm uses a Cumulative Distribution Function (CDF) matching approach with a common scaling reference to harmonise the observation time series from all used satellites before merging them based on the estimated uncertainty of each satellite (using triple collocation analysis) and removal of remaining breaks in the time series. Data is stored at the following link:

https://www.dropbox.com/scl/fo/7xbxp9e4a9c2gyvswo6yh/h/satellite_datasets/ESA_CCI_SSM?dl=0&subfolder_nav_tracking=1

4.5 ASCAT

Advanced SCATterometer (ASCAT; Wagner et al., 2013) provides a soil moisture (SM) product on a fixed Earth grid (12.5 km sampling) from backscatter observations. It consists of an estimate of the water content of the top layer of soil from 0 to 5 cm, expressed as a degree of saturation between 0 and 100 [%]. The algorithm used is based on a linear relationship between SM and scatterometer backscatter and uses change detection techniques to remove the contributions of vegetation, land cover, and surface topography, which are considered invariant from year to year. Seasonal vegetation effects are modelled by exploiting ASCAT's multi-angle observation capabilities. The data are available from 2007 to 2022 and are provided at a daily timescale with the original 12.5 km ASCAT grid. These data are stored at the following link:

https://www.dropbox.com/scl/fo/7xbxp9e4a9c2gyvswo6yh/h/satellite_datasets/ASCAT?dl=0&subfolder_nav_tracking=1

4.6 GRACE

The GRACE (Gravity Recovery and Climate Experiment; 2002-2017) and GRACE-FO (GRACE Follow-on, since 2018) satellites measure regional changes in the Earth's gravitational field. Using the constantly observed variation in the distance between satellites and other measurement data, monthly maps of the gravitational field are generated. These maps are used to calculate Terrestrial Water Storage (TWS), which is the sum of the masses of water contained in the soil column (i.e. snow, surface water, soil moisture and groundwater).

The GRACE TWS level 3 data used for this project was produced by the German Geosciences Research Centre GFZ as part of the Horizon2020 project Global Gravity-based Groundwater Product (G3P). To reach level 3, processing operations were carried out in partnership with NASA's Jet Propulsion Laboratory (JPL). The TWS data collected has a spatial resolution of 0.5° and a monthly temporal resolution. TWSA data is currently provided from 2002 to 2020 and an update with data up to 2023 will be provided when available. The data is stored in the following link:https://www.dropbox.com/scl/fo/7xbxp9e4a9c2gyvswo6yh/h/satellite_datasets/GRACE?dl=0&subfolder_nav_tracking=1

4.7 GLEAM Evaporation

The Global Land Evaporation Amsterdam Model (GLEAM; Miralles et al., 2011; Martens et al., 2017) is a set of algorithms that separately estimate the different components of terrestrial evaporation (i.e. 'evapotranspiration') based on satellite observations: transpiration, interception loss, bare-soil evaporation, snow sublimation and open-water evaporation. Intermediate outputs of the model include: potential evaporation (Ep), root-zone soil moisture, surface soil moisture, and evaporative stress.

The rationale of the method is to maximise the recovery of information about evaporation contained in the available data stack of climatic and environmental observations from space. A Priestley and Taylor equation calculates Ep based on observations of surface net radiation and near-surface air temperature. Ep estimates are then converted into actual evaporation based on the multiplicative, evaporative stress factor S.

For the project purposes, both Ep and actual evaporation (v3.7b of GLEAM) were collected over the study areas at their original spatial resolution (0.25°). Data are stored here: https://www.dropbox.com/scl/fo/h4d81f0r36798m3z0xcso/h?rlkey=dw4mlf2cttg0l4p3k9swqoxgk&st=1wsw2ywx&dl=0

Conclusions

In order to fulfil the objectives of the project, it is necessary that all participants have access to key data on all the project site areas. The database must be comprehensive enough to facilitate the implementation of AWU algorithms across the study areas and should be easily accessible in commonly used formats. Moreover, additional benchmark AWU data over the reference study areas are required to assess the results of the project's outputs.

This document outlines the data collection efforts undertaken for the CCI-AWU project. Observational data has been collected from the project's four sites, encompassing key variables such as reanalysis precipitation products, satellite soil moisture datasets, Total Water Storage (TWS) observations, and an evapotranspiration product. All this data has been stored into a common database and in a standardised format to ensure easy accessibility for all project participants, thereby streamlining the subsequent stages of the project.

References

Aeschbach-Hertig, W., and T. Gleeson (2012). Regional strategies for the accelerating global problem of groundwater depletion, *Nat. Geosci.*, 5(12), 853–861

Amarasinghe, U.A., Sikka, A., Mandave, V., Panda, R.K., Gorantiwar, S. & Ambast, S.K. (2021). Improving economic water productivity to enhance resilience in canal irrigation systems: a pilot study of the Sina irrigation system in Maharashtra, India. *Water Policy*, 23(2), 447– 465. Available from: doi:10.2166/wp.2021.231

Ambika, A. K., Wardlow, B., & Mishra, V. (2016). Remotely sensed high resolution irrigated area mapping in India for 2000 687 to 2015. *Scientific Data*, 3, 160118. doi:10.1038/sdata.2016.118

Asoka, A., Gleeson, T., Wada, Y., and Mishra, V. (2017). Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India, *Nat. Geosci.*, 10, 109–117.

Athukoralalage, D., Brookes, J., McDowell, R. W., Mosley, L. M. (2024). Impact of hydrological drought occurrence, duration, and severity on Murray-Darling basin water quality, *Water Research, 252*, 121201 https://doi.org/10.1016/j.watres.2024.121201

Bretreger, D., Yeo, I.-Y., Hancock, G., and Willgoose, G. (2020). Monitoring irrigation using landsat observations and climate data over regional scales in the Murray-Darling Basin, *Journal of Hydrology, 590*, 125356. https://doi.org/10.1016/j.jhydrol.2020.125356

Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., & Savoie, M. H. (2012). EASE-Grid 2.0: Incremental but significant improvements for Earth-gridded data sets. *ISPRS International Journal of Geo-Information*, *1*(1), 32-45. https://doi.org/10.3390/ijgi1010032

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., ... & Kerr, Y. (2016). Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4994-5007. https://doi.org/10.1109/TGRS.2016.2561938

CSIRO Water availability in the Murray–Darling Basin (2008). A report to the Australian Government from the CSIRO Murray–Darling Basin Sustainable Yields Project, CSIRO, Australia, p. 67 www.csiro.au/mdbsy

Dari, J., Brocca, L., Quintana-Seguí, P., Escorihuela, M. J., Stefan, V., & Morbidelli, R. (2020). Exploiting high-resolution remote sensing soil moisture to estimate irrigation water amounts over a Mediterranean region. *Remote Sensing, 12*, 2593. https://doi.org/10.3390/rs12162593

Dorigo, W. A., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., ... Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*. Advance online publication. https://doi.org/10.1016/j.rse.2017.07.001

Dorigo, W., Dietrich, S., Aires, F., Brocca, L., Carter, S., Cretaux, J. F., ... & Aich, V. (2021). Closing the water cycle from observations across scales: where do we stand? Bulletin of the American Meteorological Society, 102(10), E1897-E1935, doi:10.1175/BAMS-D-19-0316.1.

Food and Agriculture Organization (2013). *Yearbook, World Food and Agriculture*, Food and Agric. Organ. of the U. N., Rome

Girotto, M., De Lannoy, G. J., Reichle, R. H., Rodell, M., Draper, C., Bhanja, S. N., & Mukherjee, A. (2017).

Benefits and pitfalls of GRACE data assimilation: A case study of terrestrial water storage depletion in India. *Geophysical Research Letters, 44*, 4107–4115. https://doi.org/10.1002/2017GL072994

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture Climate Data Records and their underlying merging methodology. *Earth System Science Data, 11*, 717-739. https://doi.org/10.5194/essd-11-717-2019

Gruber, A., Dorigo, W. A., Crow, W., & Wagner W. (2017). Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. *IEEE Transactions on Geoscience and Remote Sensing*. Advance online publication. https://doi.org/10.1109/TGRS.2017.2734070

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, 146, 1999–2049, https://doi.org/10.1002/qj.3803.

Kerr, Y., Waldteufel, P., Richaume, P., Ferrazzoli, P., Wigneron,J.-P. (2014). SMOS Level 2 Processor Soil Moisture Algorithm Theoretical Basis Document (ATBD) V4.a SM-ESL (CBSA), Toulouse , p. 142

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., & Ferrazzoli, P. (2012). The SMOS soil moisture retrieval algorithm. *IEEE Geoscience and Remote Sensing, 50*, 1384-1403.

Lawston, P.M., Santanello, J.A., Kumar, S.V. (2017). Irrigation signals detected from SMAP soil moisture retrievals. *Geophysical Research Letters, 44*, 11,860–11,867.

Leblanc, M., Tweed, S., Van Dijk, A., & Timbal, B. (2012). A review of historic and future hydrological changes in the Murray-Darling Basin. *Global and Planetary Change, 80–81,* 226-246. https://doi.org/10.1016/j.gloplacha.2011.10.012

McDermid, S., Nocco, M., Lawston-Parker, P., …, Brocca, L., …, 40 authors (2023). Irrigation in the Earth system. Nature Reviews Earth & Environment, 4, 435–453, doi:10.1038/s43017-023-00438-5.

Modanesi, S., Massari, C., Camici, S., Brocca, L., and Amarnath, G. (2020). Do satellite surface soil moisture observations better retain information about crop-yield variability in drought conditions? *Water Resour. Res*. 56, e2019WR.025855. doi: 10.1029/2019WR025855

Niu, G.-Y., and Coauthors (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res., 116,* D12109, https://doi.org/10.1029/2010JD015139

O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., Bindlish, R., & Chaubell, J. (2021). SMAP L2 radiometer half-orbit 36 km EASE-grid soil moisture. Version 8. https://doi.org/10.5067/LPJ8F0TAK6E0

Peel MC, Finlayson BL, McMahon TA (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrol Earth Syst Sci.*, 11: 1633-1644

Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, 460, 999–1002.

Rodell, M., Beaudoing, H.K., L'Ecuyer, T.S., Olson, W.S., Famiglietti, J.S., Houser, P.R., Adler, R., Bosilovich, M.G., Clayson, C.A., Chambers, D., Clark, E., Fetzer, E.J., Gao, X., Gu, G., Hilburn, K., Huffman, G.J., Lettenmaier, D.P., Liu, W.T., Robertson, F.R., Schlosser, C.A., Sheffield, J., Wood, E.F., (2015). The Observed State of the Water Cycle in the Early Twenty-First Century. J. Clim. 28, 8289–8318.

Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., & Scanlon, B. R. (2015). A global data set of the extent of irrigated land from 1900 to 2005. *Hydrology and Earth System Sciences*, 19, 1521–1545. https://doi.org/10.5194/hess-19-1521-2015

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., Timbal, B., and Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001-2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, *Water Resour. Res.*, 49, 1040–1057. https://doi.org/10.1002/wrcr.20123

Zaussinger, F., Dourigo, W., Gruber, A., Tarpanelli, A., Filippucci, P., Brocca, L. (2019). Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data. *Hydrol. Earth Syst. Sci.*, 23, 897–923. https://doi.org/10.5194/hess-23-897-2019