



**POLITECNICO**  
MILANO 1863

**KU LEUVEN**

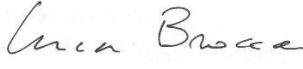
## Anthropogenic Water Use (CCI-AWU)

### **Deliverable 3: Prototype Algorithms Theoretical Basis Document (ATBD), Products description**

Date	Issue	Section	Page	Comment
09/12/2024	1.0			

**Control Document**

Process	Name	Date
Written by:	Luca Brocca, Christian Massari, Sara Modanesi, Jacopo Dari, Carla Saltalippi, Renato Morbidelli, 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,	09/12/2024

	Signature	Date
For CCI AWU team		09/12/2024
For ESA		

[This page is left intentionally blank]

---

## List of content

1.	Introduction .....	5
1.1.	The CCI-AWU project .....	5
1.2.	Scope of this report.....	5
1.3.	Applicable Documents .....	6
2.	ATBD SM-based inversion approach.....	7
3.	ATBD SM-based delta .....	10
3.1.	Introduction .....	10
3.2.	The SM-DELTA approach.....	11
3.2.1.	Classic approach (only SM) .....	11
3.2.2.	SM-DELTA approach including ET .....	11
3.3.	Data used .....	12
4.	ATBD Flux-based approach .....	17
4.1.	Introduction .....	17
4.2.	FLUX-BASED algorithm.....	18
5.	ATBD Model-observation integration.....	22
5.1.	Introduction .....	22
5.2.	Irrigation detection .....	22
5.2.1	Approach 1: relative bias .....	22
5.2.2	Approach 2: multiresolution analysis to detect irrigation .....	25
5.3.	Irrigation quantification .....	32
5.3.1	The Noah-MP Land Surface Model .....	32
5.3.2	Deterministic experiments.....	33
5.3.3	Ensemble runs.....	35
	References .....	36

---

# 1. Introduction

## 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, the development of an ECV that includes 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 ESA 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., at least 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 comprises 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

The purpose of this document is the description of each retrieval algorithm used to derive the Anthropogenic Water Use products from input satellite data, i.e., the Algorithm Theoretical Basis Document (ATBD). The ATBD includes: 1) all input satellite and ancillary data, 2) the retrieval algorithms used, and 3) a general overview of the processing chain for retrieval of the data products.

---

## 1.3. Applicable Documents

- Proposal.
- Deliverable D2. Report explaining the criteria for selecting the test regions.

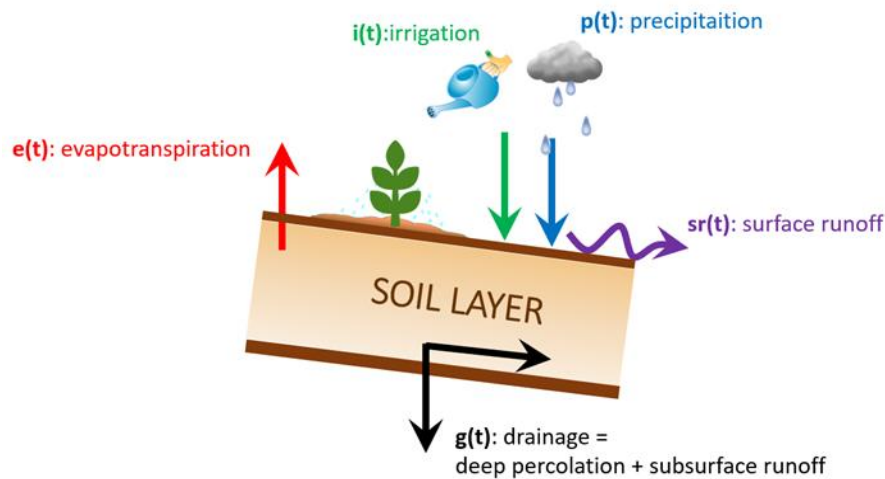
## 2. ATBD SM-based inversion approach

The method relies on the inversion of the soil water balance for estimating irrigation water use; it is an evolution of the SM2RAIN algorithm (Brocca et al., 2014) originally born to retrieve rainfall rates by inverting the soil moisture signal, i.e., by using the soil as a natural rain gauge. Over agricultural areas the total amount of water entering the soil (the actual algorithm output) is determined by the sum of rainfall and irrigation rates. Hence, by removing rainfall rates from the algorithm output, it is possible to estimate the amount of water applied for irrigation. The method was tested in previous applications exploiting coarse resolution (Brocca et al., 2018; Jalilvand et al., 2019), high-resolution (Dari et al., 2020; 2022; 2023; 2024), and in-situ (Filippucci et al., 2020) soil moisture data. In the framework of the European Space Agency (ESA) Irrigation+ project (<https://esairrigationplus.org/>), the SM-based inversion approach was implemented over selected test sites to produce the first-ever regional data sets of irrigation water use retrieved from satellite (Dari et al., 2023). The method was later implemented over the main agricultural areas falling within the Mediterranean basin in the ESA 4DMED-Hydrology project (<https://www.4dmed-hydrology.org/>). In a recent study, Dari et al. (2024) proved the feasibility of building an operational system based on the SM-based inversion approach for the management of agricultural water by leveraging satellite observations.

The soil water balance equation can be expressed as follows:

$$nZ \frac{dS(t)}{dt} = i(t) + r(t) - g(t) - sr(t) - e(t) \quad (1)$$

In which  $n$  [-] is the soil porosity,  $Z$  [mm] is the soil layer depth,  $dS(t)/dt$  [-] is the variation of relative soil moisture,  $S(t)$ , in time,  $t$  [day],  $i(t)$  [mm/day] is the irrigation rate,  $r(t)$  [mm/day] is the rainfall rate,  $g(t)$  [mm/day] is the drainage,  $sr(t)$  [mm/day] is the surface runoff, and  $e(t)$  [mm/day] is the actual evapotranspiration rate. The components of the soil water balance considered in the SM-based inversion approach are shown in Figure 2.1.



**Figure 2.1.** Components of the soil water balance considered in the SM-based inversion approach.

Eq. (1) can be reformulated as follows:

$$Win(t) = Z * \frac{dS(t)}{dt} + g(t) + sr(t) + e(t) \quad (2)$$

In which  $Win(t)$  [mm/day] indicates the total amount of water entering the soil, i.e.  $Win(t) = i(t) + r(t)$  and  $Z^* = nZ$  [mm] is the water capacity of the soil layer. The surface runoff is assumed to be negligible (Brocca et al., 2015) and the drainage term is linked with soil moisture through the following potential law (Brocca et al., 2014):

$$g(t) = aS(t)^b \quad (3)$$

In which  $a$  [mm] and  $b$  [-] indicate drainage parameters. The actual evapotranspiration term is expressed as the potential evapotranspiration rate,  $pe(t)$ , limited by the available water content:

$$e(t) = FS(t)pe(t) \quad (4)$$

Where  $F$  [-] is a scaling factor (Dari et al., 2023). Hence, Eq. (2) can be written as follows:

$$Win(t) = Z^* dS(t)/dt + aS(t)^b + FS(t)pe(t) \quad (5)$$

With  $Z^*$ ,  $a$ ,  $b$ , and  $F$  parameters to be calibrated. According to the strategy presented in Dari et al. (2024), the algorithm parameters are calibrated against rainfall (i.e., by optimising the method performances in properly reproducing rainfall amounts) by masking out days with rainfall rate  $< 1$  mm during the irrigation seasons (hence, potential irrigation days). In such periods, the left term on Eq. (5) is supposed to be determined by rainfall only. Once the algorithm parameters are calibrated, irrigation estimates are produced during the irrigation season. In this case, both rainfall and irrigation rates can be involved in determining the left term of Eq. (5) and irrigation amounts can be derived by removing rainfall rates from the total output,  $i(t) = Win(t) - r(t)$ . Eventual negative irrigation rates are set equal to zero (Jalilvand et al., 2019); in order to discard negligible irrigation estimates due to random errors, a threshold value for the ratio between weekly estimated irrigation and weekly rainfall equal to 0.2 is adopted.

Irrigation and rainfed seasons over the different study sites, useful to delineate the calibration periods, are determined on the basis of indications derived by Portmann et al. (2008) and ancillary information. Estimates of irrigation water use are produced during irrigation seasons and over areas equipped for irrigation. Such information is derived from the latest version of the global map of Areas Equipped for Irrigation (AEI) referring to the year 2015 (Mehta et al., 2022); only pixels with AEI higher than 5% are considered. Table 2.1 summarises the irrigation seasons considered for each study area, hence, the periods in which irrigation estimates are available.

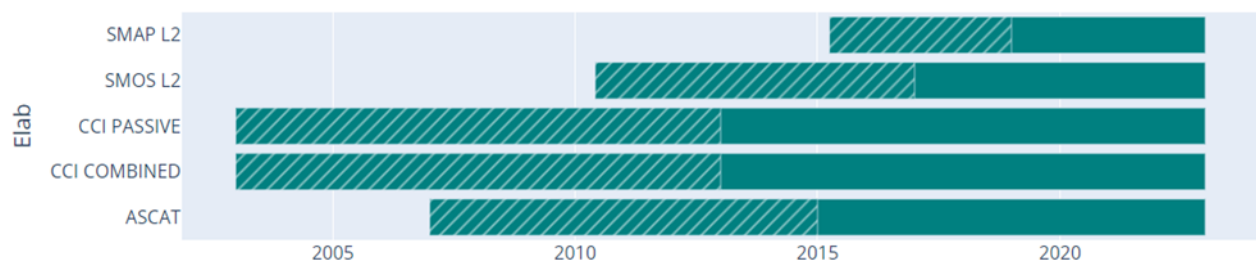
**Table 2.1.** Irrigation seasons considered for each study area.

	Adopted irrigation season
Murray-Darling basin (Australia)	September-April
CONUS	May-September
India	November-June: Rabi (November-March) + Zaid (April-June)
Ebro basin (Spain)	April-October



It is noteworthy that Kragh et al. (2023) reported irrigation occurring over Northwestern India during the monsoon season as well. Such a phenomenon interests areas characterised by arid or semi-arid climate. Hence, the development of a second version of the data set over India in which the temporal mask is not applied to areas falling in arid or semi-arid zones is under evaluation.

For all the selected study sites, five soil moisture products are tested: Advanced SCATterometer (ASCAT), Level 2 Soil Moisture Ocean Salinity (SMOS L2), Level 2 Soil Moisture Active Passive (SMAP L2), and Climate Change Initiative (CCI) Combined (CCI COMBINED) and Passive (CCI PASSIVE) data sets. The Global Land Evaporation Amsterdam Model (GLEAM) v3.7b (Martens et al., 2017) is used as a source for potential evapotranspiration data, while the ERA5 (European ReAnalysis v5) (Hersbach et al. 2020) lowest model level forecasts are considered for rainfall rates. Simulations leveraging different soil moisture products cover different periods because of the temporal coverage of each sensor. However, for all the simulations half of the total time series has been used for the calibration of the algorithm parameters; the irrigation estimates cover the whole period of data availability. Note that the overlapping of calibration and validation periods during the first half of the time series does not affect the reliability of the results in terms of irrigation estimates because of the complementarity between rainfed and irrigation seasons. Figure 2.2 provides an overview of the experiments carried out.

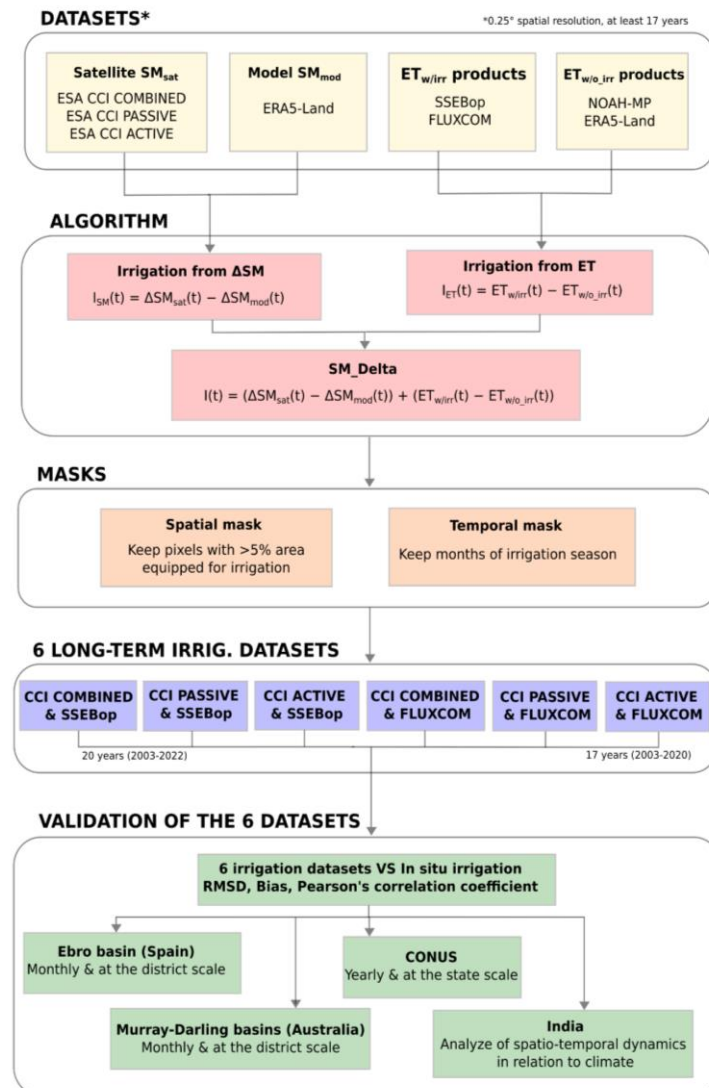


**Figure 2.2.** Length of irrigation time series available for the considered soil moisture products. The fielded portion of the bar represents the calibration period.

## 3. ATBD SM-based delta

### 3.1. Introduction

Monthly irrigation datasets with a duration of 17 and 20 years were produced for the study areas CONUS, India, Ebro basin (Spain), and Murray-Darling basin (Australia) using the SM-DELTA algorithm. This document presents the algorithm, the data used, the masks used, and the characteristics of the irrigation datasets generated. The flowchart in Figure 3.1 shows the steps involved in obtaining these datasets. First, soil moisture (SM) and evapotranspiration (ET) datasets were selected and collected, then six irrigation datasets were generated with a combination of three SM products and two ET products. Then, spatial and temporal masks are applied to mask any irrigation signal where and when we know there shouldn't be any. The final six long-term irrigation datasets (17 years and 20 years) are obtained, and can then be validated with in situ irrigation data.



**Figure 3.1.:** Flowchart of the SM-DELTA approach for generating six long-term irrigation datasets using satellite and model soil moisture, evapotranspiration products, and spatio-temporal masks, validated against in situ irrigation data.

## 3.2. The SM-DELTA approach

### 3.2.1. Classic approach (only SM)

The SM-DELTA approach was introduced by Zaussinger et al. (2019) and first aimed to estimate irrigation by analyzing differences between satellite soil moisture change ( $\Delta SM_{sat}$ ) and modeled soil moisture change ( $\Delta SM_{mod}$ ). This method is based on the idea that the change in  $\Delta SM_{mod}$  is only due to precipitation, whereas the change in  $\Delta SM_{sat}$  is due to both precipitation and irrigation. Eq. (1) and Eq. (2) describe this principle.

The soil water balance equation for  $SM_{sat}$  is as follows:

$$Z \cdot \frac{dSM_{sat}}{dt} = P(t) + I(t) - ET(t) - R(t) - \Delta S_{rest} \quad (1)$$

And for modeled SM:

$$Z \cdot \frac{dSM_{mod}}{dt} = P(t) - ET(t) - R(t) - \Delta S_{rest} \quad (2)$$

Where  $P$  (mm) is precipitation,  $I$  (mm) is irrigation,  $ET$  (mm) is evapotranspiration,  $\Delta S_{rest}$  ( $m^3 m^{-3}$ ) is the water storage changes below the surface layer including drainage,  $R$  (mm) is the runoff, and  $t$  is the number of days between two observations.  $Z$  is the soil depth used to convert SM to water column and is set at 5 cm to represent the soil depth detected by the sensor as in Zaussinger et al. (2019) and Zappa et al. (2020, 2022b, 2024)

Considering that  $P$ ,  $ET$ ,  $\Delta S_{rest}$ , and  $R$  are identical in Eq. (1) and Eq. (2),  $I$  can be computed as follow:

$$I(t) = Z \cdot \frac{dSM_{sat}}{dt} - Z \cdot \frac{dSM_{mod}}{dt} \quad (3)$$

In addition, to avoid false irrigation signals due to noise in the  $SM_{sat}$ , irrigation is only triggered if the following condition is met.

$$\frac{SM_t^{sat} - SM_{t-1}^{sat}}{SM_{t-1}^{sat}} \geq 0.12 \quad (4)$$

### 3.2.2. SM-DELTA approach including ET

Zappa et al (2022b, 2024) proposed to include  $ET$  in the SM-DELTA algorithm. Indeed, Eq. (3) is only correct if  $P$ ,  $ET$ ,  $\Delta S_{rest}$ , and  $R$  are identical between  $ET_{w/irrig}$  (including the irrigation signal) and  $ET_{w/o\_irrig}$  (not including the irrigation signal). However,  $ET$ , particularly, can differ considerably between irrigated and rainfed fields (Brombacher et al., 2022; van Eekelen et al., 2015). If not accounted for, differences between irrigated and rainfed  $ET$  can lead to underestimations of irrigation (Zappa et al., 2022b; Kragh et al., 2023).

To take into account the difference in  $ET$  between an irrigated and a non-irrigated pixel, Eq. (3) has been modified as follows:

$$I(t) = \left( Z \cdot \frac{dSM_{sat}}{dt} - Z \cdot \frac{dSM_{mod}}{dt} \right) + \begin{cases} ET_{w/irrig}(t) - ET_{w/o\_irrig}(t), & \text{if } ET_{w/irrig}(t) > ET_{w/o\_irrig}(t) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

With  $ET_{w/irrig}$  (mm) the ET data theoretically including irrigation, and  $ET_{w/o\_irrig}$  (mm) ET data theoretically not including the irrigation signal.

### **3.3. Data used**

#### **Modeled soil moisture ( $SM_{mod}$ )**

The ERA5-Land reanalysis product was used for  $SM_{mod}$  (SM without irrigation signal). ERA5-Land, produced by the Copernicus Climate Change Service (C3S) of the European Centre for Medium-Range Weather Forecasts (ECMWF), provides a multitude of variables related to water and energy cycles over land since 1978 at a spatial resolution of 9 km and an hourly time step (Muñoz-Sabater et al., 2021). ERA5-Land has been selected as  $SM_{mod}$  because, being mainly forced by meteorological forecasts, it contains minimal signals of human activity such as irrigation. In this study,  $SM_{mod}$  data from ERA5-Land were regridded to the ESA CCI grid (0.25°) using bilinear interpolation and resampled to a daily scale.

#### **Satellite soil moisture ( $SM_{sat}$ )**

The three ESA Climate Change Initiative (CCI) products (Dorigo et al., 2017) are used for  $SM_{sat}$ : CCI ACTIVE (merge three radar satellites, data available since 1991), CCI PASSIVE (merge 12 radiometer satellites, data available since 1978) and CCI COMBINED (merge CCI ACTIVE and CCI PASSIVE satellites, data available since 1978). These datasets have a 0.25° grid resolution and were selected because they offer the highest temporal resolution (greater than SMOS, SMAP, and ASCAT) in the four study areas, providing a temporal frequency greater than 3 days, which is essential to ensure good performance with the SM-DELTA approach (Zappa et al., 2022b). The  $SM_{sat}$  data were rescaled to the ERA5-Land  $SM_{mod}$  data using the so-called “mean-standard” method (as recommended in Escorihuela & Quintana-Seguí (2016), and applied in numerous previous work such as Zaussinger et al. (2019), Massari et al. (2015)).

#### **Evapotranspiration not accounting for irrigation ( $ET_{w/o\_irrig}$ )**

We selected two ET products that theoretically contain no irrigation signal ( $ET_{w/o\_irrig}$ ). One is derived from NOAH-MP simulations (Niu et al., 2011) using ERA5 forecasts for meteorological forcing only, thus excluding observations that may contain irrigation signals. NOAH-MP data are only available for the CONUS and the Murray-Darling basin, and only since 2010. To cover a period starting from 2003 and to cover the study area India and Ebro basin, we used the ET data simulated by ERA5-Land. Indeed, an analysis (not shown) carried out on CONUS and MURRAY showed that ERA5-Land ET generally follows very similar dynamics to ET simulated by NOAH-MP.

#### **Evapotranspiration accounting for irrigation ( $ET_{w/irrig}$ )**

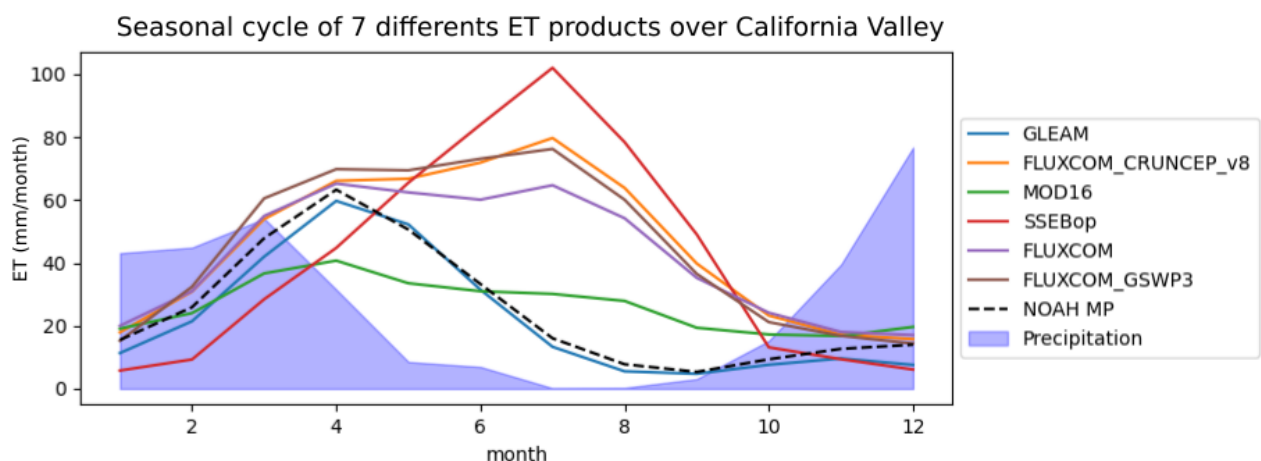
We selected two ET products that theoretically take into account irrigation ( $ET_{w/irrig}$ ). One  $ET_{w/irrig}$  product is Simplified Surface Energy Balance operational (SSEBop; Senay et al., 2018) using a simplified surface energy balance approach based on “wet” and “dry” reference pixels. Wet pixels represent areas with maximum ET (typically over well-watered vegetation), while dry pixels represent minimal ET (typically over bare soil or stressed vegetation). The temperature difference between these wet and dry reference pixels allows SSEBop to scale evapotranspiration across different landscapes based on surface temperature data.

The other  $ET_{w/irrig}$  product is FLUXCOM (remote sensing only version). FLUXCOM is a machine-learning-based model trained on ET data from flux tower, and using remote sensing data as predictors, mainly from MODIS for variables like vegetation indices and surface temperature. The ET data from FLUXCOM is derived from a set of 36 different machine-learning model outputs.

These two products were selected from a cross-comparison of 6 ET products (GLEAM v3, MOD16, FLUXCOM, FLUXCOM\_CRUNCEP\_V8, SSEBop, ERA5-Land) based on their ability:

- 1) To produce more ET than NOAH-MP (the latter being  $ET_{w/o\_irrig}$ ) during irrigated seasons on irrigated pixels.
- 2) To provide temporal dynamics of ET on irrigated pixels consistent with what might be expected during the irrigation season on irrigated pixels.
- 3) To generate larger quantities of ET on irrigated pixels during irrigation seasons than on surrounding non-irrigated pixels.
- 4) To provide quantities of ET consistent with precipitation quantities (no several orders of magnitude higher or lower than annual precipitation).

Figure 3.2 shows the seasonal cycles of 7 ET products, including NOAH-MP (dotted black line, representing  $ET_{w/o\_irrig}$ ), FLUXCOM (purple line, representing  $ET_{w/irrig}$ ), and SSEBop (red line, representing  $ET_{w/irrig}$ ) for the highly irrigated California Valley. The precipitation is shown in blue. We can see that only FLUXCOM and SSEBop show ET dynamics consistent with the irrigation season occurring from May to October. On the other hand, the ET modeled by NOAH-MP decreases from the beginning of the irrigation season.



**Figure 3.2:** Seasonal cycle of seven different ET products and precipitation (blue shaded area) over California Valley.

**Table 3.1:** Summary of SM and ET products, data availability, and characteristics.

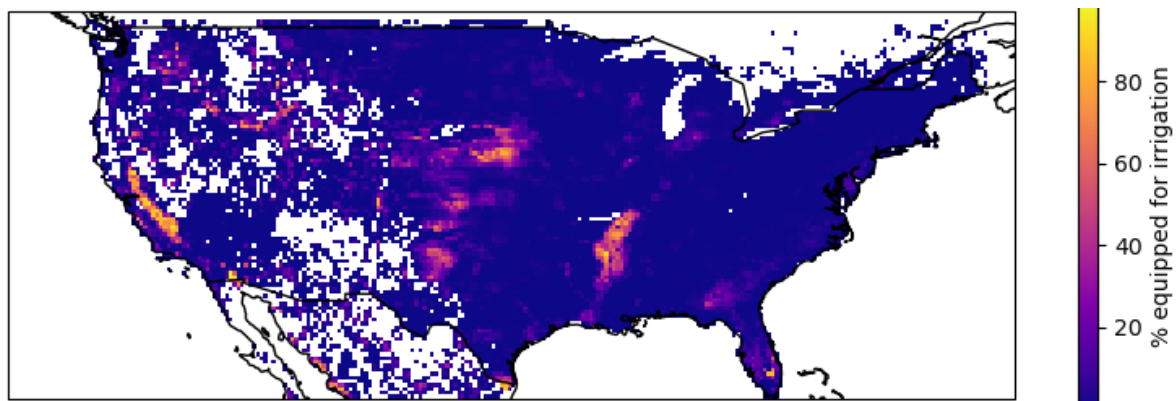
Product	Irrigation signal	Data availability	Description	Original resolution (resampled to 0.25°)
ERA5-Land SM	No	1978-2023	Reanalysis	9 km
CCI COMBINED SM	Yes	1978-2023	Merges active and passive satellite data	0.25°
CCI PASSIVE SM	Yes	1978-2023	Merges 12 passive satellite radiometer data	0.25°
CCI ACTIVE SM	Yes	1991-2023	Merges 3 active satellite radar data	0.25°
NOAH-MP ET	No	2010-2023	Model simulation	0.25°
ERA5-Land ET	No	1978-2023	Reanalysis	9 km
SSEBop ET	Yes	2003-2023	Energy balance model	1 km
FLUXCOM ET	Yes	2003-2020	Machine learning model	0.0833°

## 3.4. Spatial and temporal masks

### 3.4.1. Spatial masks

In order to avoid false irrigation signals in areas where irrigation is not expected, we applied spatial masks. Indeed, the behavior of  $SM_{sat}$ ,  $SM_{mod}$ ,  $ET_{w/irrig}$ , and  $ET_{w/o\_irrig}$  can differ on land covers different from "agricultural and irrigated pixel" and therefore can generate irrigation even though it is known that no irrigation has been applied here.

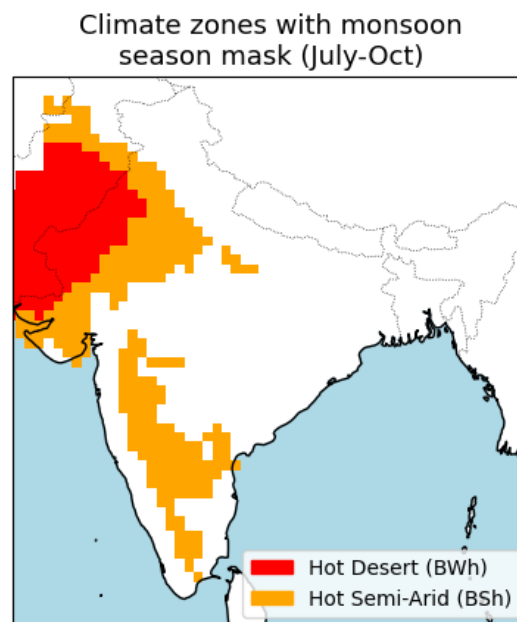
The mask we used is derived from the Global Map of Areas equipped for irrigation (GMIA) (Figure 3.3; Siebert et al., 2015). We spatially resampled the GMIA data (originally at 300 m resolution) to the 0.25° ESA CCI grid, and kept only the pixels with more than 5% of their surface area equipped for irrigation.



**Figure 3.3:** Percentage of land area equipped for irrigation across the CONUS (GMIA dataset, Siebert et al., 2015).

## 3.4.2. Temporal masks

We applied a temporal mask to eliminate any potential false irrigation signals during periods when no irrigation takes place. The masks were applied to match the irrigation seasons. For CONUS, the irrigation season runs from May to September, for the Ebro basin from April to October, and for the Murray-Darling basin from September to April. For India, we applied different temporal masks according to the type of Köppen climate classification. The months corresponding to the monsoon season (July to October) have been masked for pixels belonging to climatic classes different than Hot Desert (BWh) and Hot Semi-Arid (BSh) (Figure 3.4), to avoid false irrigation signals in humid areas where little irrigation is applied during the monsoon season due to high rainfall. We did not apply a temporal mask to pixels belonging to climatic classes Hot Desert and Hot Semi-Arid to account for the fact that on these pixels significant amounts of irrigation are applied to meet plant needs even during the monsoon season.



**Figure 3.4:** Hot desert (BWh) and hot semi-arid (BSh) Köppen climate zones in India where no temporal mask is applied to allow irrigation even during the monsoon season (July–October).

## 3.5. Description of irrigation datasets

**Table 3.2:** Summary of irrigation datasets produced with the SM-DELTA approach: study areas, data periods, spatial and temporal masks

Study area	Data period	Available datasets	Spatial mask	Temporal mask
<b>CONUS</b>	2003-2022 (SSEBOP), 2003-2020 (FLUXCOM)	6 datasets: CCI COMBINED, CCI PASSIVE, CCI ACTIVE (combined with SSEBOP, FLUXCOM)	More than 5% of pixels equipped for irrigation (GMIA)	May to September
<b>Ebro</b>	2003-2022 (SSEBOP), 2003-2020 (FLUXCOM)	6 datasets: CCI COMBINED, CCI PASSIVE, CCI ACTIVE (combined with SSEBOP, FLUXCOM)		April to October
<b>Murray-Darling</b>	2003-2022 (SSEBOP), 2003-2020 (FLUXCOM)	6 datasets: CCI COMBINED, CCI PASSIVE, CCI ACTIVE (combined with SSEBOP, FLUXCOM)		September to April
<b>India</b>	2003-2020	3 datasets CCI COMBINED, CCI PASSIVE, CCI ACTIVE (combined with FLUXCOM)		Variable depending on Köppen climate class



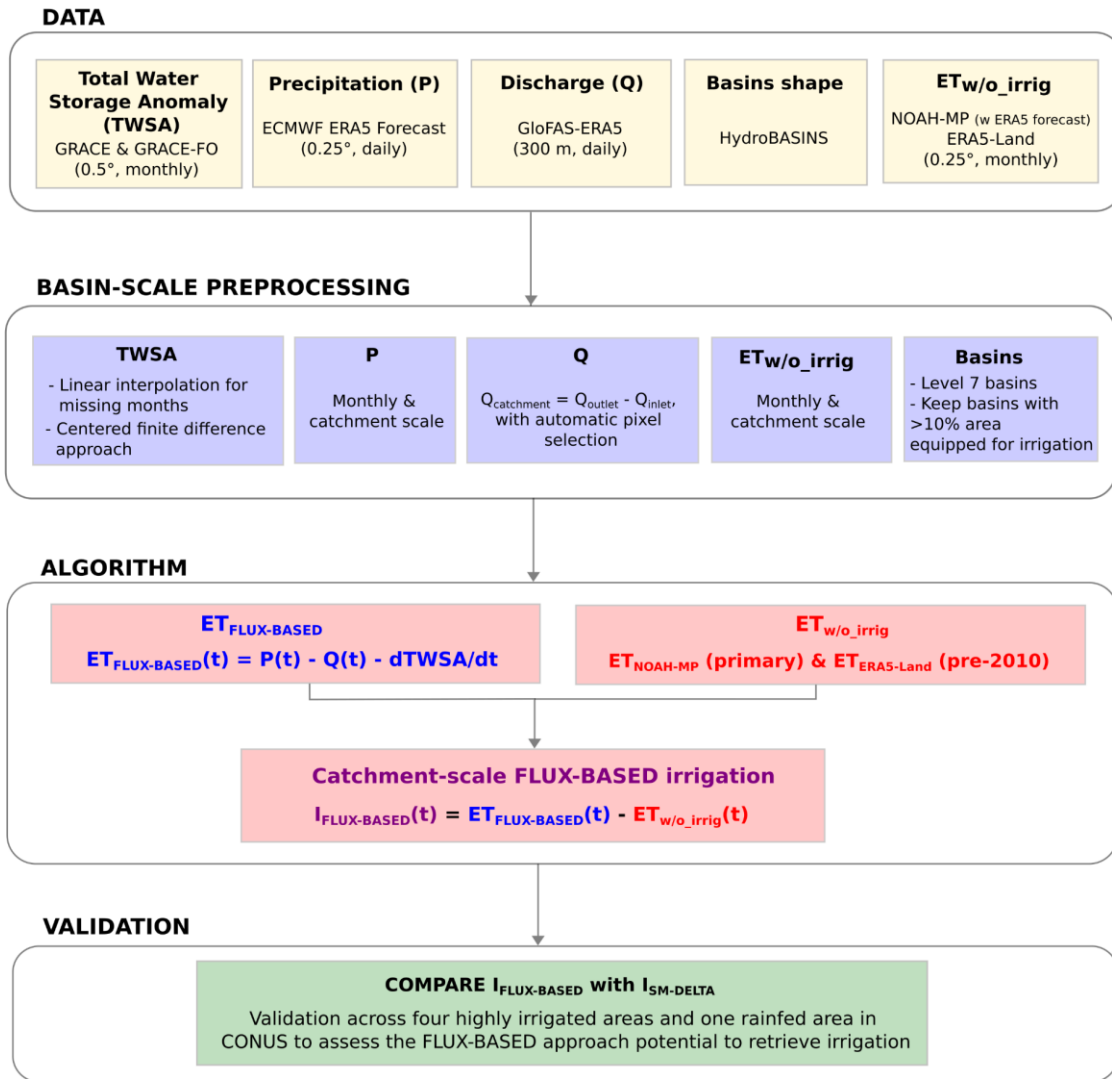
## 4. ATBD Flux-based approach

### 4.1. Introduction

The FLUX-BASED irrigation retrieval approach leverages the difference between evapotranspiration (ET) obtained by closing the water balance ( $ET_{\text{FLUX-BASED}}$ ) and an ET value that theoretically excludes irrigation signals ( $ET_{\text{w/o\_irrig}}$ ).  $ET_{\text{FLUX-BASED}}$  is computed at the watershed scale as the residual of precipitation (P) minus the Total Water Storage Anomaly (TWSA) and discharge (Q). Numerous studies have demonstrated the effectiveness of  $ET_{\text{FLUX-BASED}}$  in estimating evapotranspiration at the watershed level (see Güntner, 2008; Ramillien et al., 2006; Rodell et al., 2015).  $ET_{\text{FLUX-BASED}}$  uses TWSA, which accounts for all water fluxes, including both natural climatic processes—such as precipitation, soil moisture, ice, snow, and runoff—as well as human influences on the hydrological cycle—such as irrigation, reservoir management, and groundwater extraction (Rodell et al., 2015). TWSA is thus expected to contain irrigation-related signals within its measurements.

Several studies suggest that  $ET_{\text{FLUX-BASED}}$  indeed includes irrigation signals. For instance, Pascolini-Campbell et al. (2020) computed  $ET_{\text{FLUX-BASED}}$  over large basins in the CONUS region, observing that  $ET_{\text{FLUX-BASED}}$  was higher during the irrigation season in heavily irrigated areas (such as California’s Central Valley) than in ET products derived from land surface models. Pascolini-Campbell et al. (2021) compared  $ET_{\text{FLUX-BASED}}$  across several small irrigated watersheds in the western United States and found a correlation between annual  $ET_{\text{FLUX-BASED}}$  values and the proportion of irrigated land within these basins. Given that  $ET_{\text{FLUX-BASED}}$  theoretically contains an irrigation signal, it can be used to compute irrigation at the watershed scale ( $I_{\text{FLUX-BASED}}$ ) by calculating the difference between  $ET_{\text{FLUX-BASED}}$  and a modeled  $ET_{\text{w/o\_irrig}}$ . Note that the irrigation signal retrieved by this method represents theoretically only the water consumed and transpired by plants, excluding any excess water that may have been applied and subsequently drained.

The flowchart in Figure 4.1 illustrates the steps involved in obtaining  $I_{\text{FLUX-BASED}}$ . First, the datasets required to calculate  $ET_{\text{FLUX-BASED}}$  are collected, including TWSA, P, Q, watershed boundaries, and  $ET_{\text{w/o\_irrig}}$ . These datasets are then preprocessed to obtain monthly values at the watershed scale.  $ET_{\text{FLUX-BASED}}$  is then computed by closing the water balance, and  $I_{\text{FLUX-BASED}}$  is derived by subtracting  $ET_{\text{w/o\_irrig}}$  from  $ET_{\text{FLUX-BASED}}$ .



**Figure 4.1:** Flowchart of the FLUX-BASED approach for estimating irrigation, from data acquisition and preprocessing to calculation of  $ET_{\text{FLUX-BASED}}$  and  $I_{\text{FLUX-BASED}}$ , with validation against  $I_{\text{SM-DELTA}}$ .

## 4.2. FLUX-BASED algorithm

$I_{\text{FLUX-BASED}}$  is defined as the difference between  $ET_{\text{FLUX-BASED}}$  and  $ET_{\text{w/o\_irrig}}$ . The core component of the algorithm is calculating  $ET_{\text{FLUX-BASED}}$  using a water balance approach, as follows:

$$ET_{\text{FLUX-BASED}} = P - Q - \frac{dTWSA}{dt} \quad (1)$$

where  $P$  (mm) is precipitation,  $Q$  (mm) is the runoff at the basin outlet minus the runoff at the basin inlet, and  $dTWSA/dt$  (mm) represents the change in total water storage derived from GRACE TWS data (Ramillien et al., 2006).

---

$I_{FLUX-BASED}$  can then be calculated as:

$$I_{FLUX-BASED} = ET_{FLUX-BASED} - ET_{w/o\_irrig} \quad (2)$$

where  $ET_{w/o\_irrig}$  (mm) represents ET that theoretically excludes irrigation signals, derived from a combination of ET simulated with the NOAA-MP model (Niu et al., 2011) driven solely by ERA5 forecast data as forcing inputs, and ERA5-Land (Muñoz-Sabater et al., 2021) ET data.

### 4.3. Data and preprocessing

#### Total Water Storage (TWSA)

Total Water Storage data from the GRACE (2002-2017) and GRACE-FO (since 2018) missions were processed by the German Research Centre for Geosciences (GFZ) group as part of the Global Gravity-based Groundwater Product project (G3P; <https://www.g3p.eu>). These data are provided at GRACE mascon resolution with a spatial resolution of 0.5° for the period from 2004 to the present. This mascon resolution improves upon previous spherical harmonic solutions by reducing leakage errors (Watkins et al., 2015; Wiese et al., 2016). Due to technical and maintenance issues, and the transition between GRACE and GRACE-FO, data are missing punctually for some months, and from November 2017 to May 2018. We applied a linear interpolation to fill these gaps.

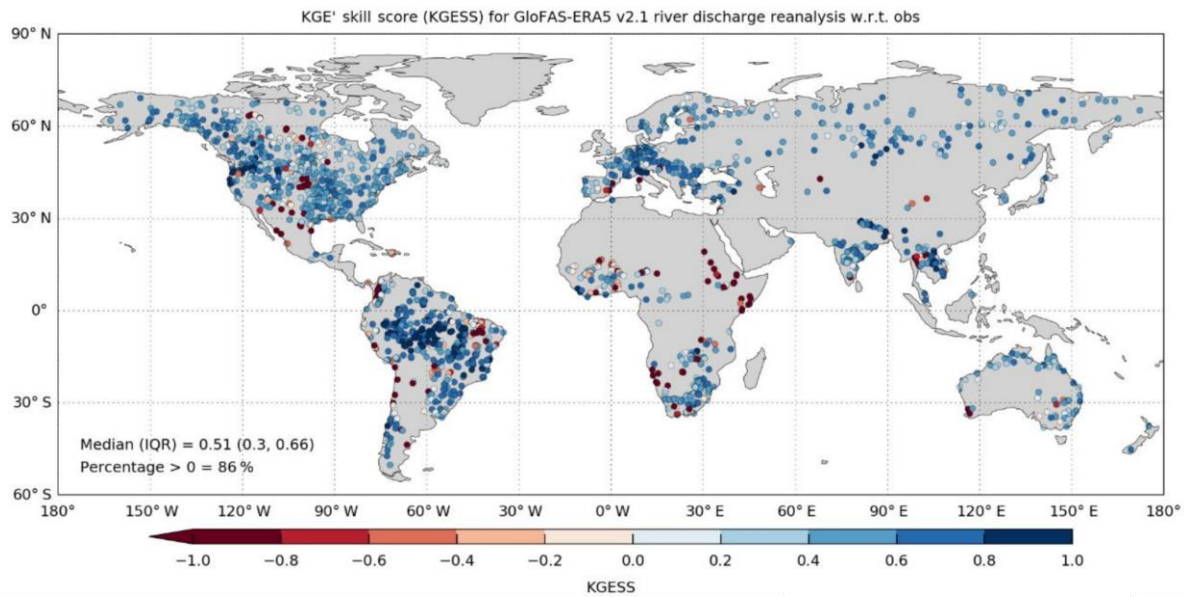
GRACE TWSA data are provided monthly, but their intervals are not strictly uniform, potentially introducing interference from adjacent monthly signals and high-frequency noise artifacts (Landerer et al., 2010). To address this, we applied a centered finite difference method to compute  $dTWSA/dt$  (Landerer et al., 2010) by computing TWSA by differencing values from the preceding and following months, then dividing by 2. This method has been validated in similar studies (Landerer et al., 2010; Swann & Koven, 2017).

#### Precipitation

Precipitation data were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis, and downloaded from the ECMWF Meteorological Archival and Retrieval System (MARS) platform. This dataset spans from January 1990 to March 2023 with daily temporal resolution and was resampled to the ESA CCI SM grid with a spatial resolution of 0.25°. Daily precipitation values were then aggregated to produce monthly precipitation data for each grid cell.

#### Discharge

River discharge data were obtained from the Global Flood Awareness System ERA5 (GloFAS-ERA5; <http://www.globalfloods.eu/>) dataset, a global daily reanalysis of river discharge (Harrigan et al., 2020). With a spatial resolution of 0.1°, GloFAS-ERA5 combines runoff from the H-TESSSEL-ERA5 land surface model with a channel routing model (LISFLOOD), and the simulation outputs are calibrated against discharge observations from 1,287 catchments worldwide (see Figure 4.2 for the station locations).



**Figure 4.2:** Locations of the 1,287 discharge stations used in GloFAS-ERA5, along with their Kling-Gupta Efficiency Skill Score (KGESS) values calculated between GloFAS-ERA5 and in situ station data (figure from Harrigan et al., 2020).

To obtain watershed-scale discharge, we developed a tool to select GloFAS-ERA5 grid cells representing the watershed outlet and, where applicable, the inlet. Discharge is calculated as the difference between flow at the outlet and inlet, with monthly summation applied.

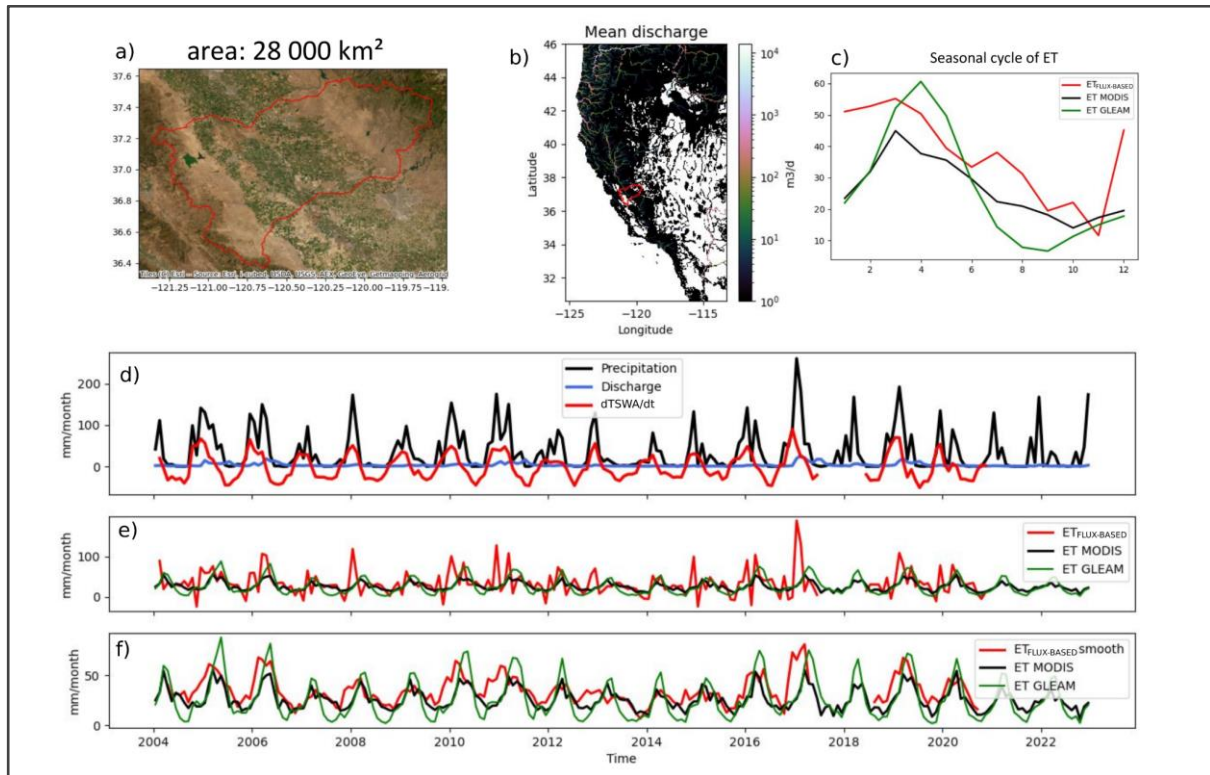
## Basin boundaries

We used basin boundaries from the HydroBASINS database (Lehner and Grill, 2013), a global dataset of hydrographic sub-basins at multiple hierarchical levels. HydroBASINS delineates watersheds from large river basins to smaller sub-catchments. We selected level 7 catchments to ensure basin size compatibility with irrigated areas. In the FLUX-BASED algorithm, monthly TWSA, P, and Q are computed for each basin.

## $ET_{w/o\_irrig}$

The  $ET_{w/o\_irrig}$  dataset represents ET without irrigation signals. It was derived by combining ET data from NOAH-MP (driven only by ERA5 forecasts, without observations) and ERA5-Land reanalysis. An analysis (not shown) showed that ET dynamics in NOAH-MP align closely with ERA5-Land, validating ERA5-Land as a substitute. Therefore, ERA5-Land data were used for years before 2010, for which NOAH-MP data are unavailable.

## 4.4. Illustration of the FLUX-BASED approach applied on an irrigated catchment



**Figure 4.3:** Illustration of the FLUX-BASED algorithm applied to an irrigated catchment in California. (a) Watershed boundaries on satellite imagery; (b) maps of mean discharge distribution from GloFAS-ERA5 with the watershed boundaries in red; (c) seasonal ET cycle comparing  $ET_{FLUX-BASED}$  with ET from MODIS and GLEAM ; (d) monthly time series of precipitation (P), discharge (Q), and dTWSA/dt; (e) ET time series of  $ET_{FLUX-BASED}$ , MODIS, and GLEAM; (f) same as (e) but smoothed (3 months)  $ET_{FLUX-BASED}$ .

Figure 4.3 illustrates the application of the FLUX-BASED algorithm in an irrigated watershed in California (approx. 28,000 km<sup>2</sup>). In panel (a), watershed boundaries are shown on satellite imagery. Panel (b) displays mean discharge patterns across the region, derived from GloFAS-ERA5. Panel (c) presents the seasonal ET cycle, comparing  $ET_{FLUX-BASED}$  with MODIS (MOD16) and GLEAM (Martens et al., 2017) ET. We can see that  $ET_{FLUX-BASED}$  peaks during irrigation seasons, indicating irrigation-driven water use. Panel (d) shows monthly time series for P, Q, and dTWSA/dt. Panel (e) shows time series data for  $ET_{FLUX-BASED}$ , MODIS, and GLEAM. Panel (f) presents smoothed (3 months) time series for ET estimates, where  $ET_{FLUX-BASED}$  remains elevated during irrigation in comparison to GLEAM and MODIS, highlighting the algorithm's potential sensitivity to irrigation signals.

## 5. ATBD Model-observation integration

### 5.1. Introduction

AWU plays a critical role in the terrestrial water cycle, accounting for 84-90% of the global freshwater consumption (McDermid et al., 2023). Despite this importance, this human influence on land water distribution is often poorly parameterized in regional and large-scale land surface models. One challenge is that the capability to simulate irrigation relies on input irrigated area maps, which are frequently outdated or static. An example is given by the Global Rainfed Irrigated and Paddy Croplands dataset referring to the year 2005 (GRIPC; Salmon et al., 2015) which was used in recent studies to derive irrigation estimates through land surface modelling (Modanesi et al., 2022; Busschaert et al., 2023; De Lannoy et al., 2024). Another issue is that irrigation parameterization often depends on spatially uniform, user-defined parameters, limiting the model's ability to estimate reliable irrigation quantities.

In the proposal stage, a coarse-scale data assimilation framework was proposed where first observation-based forcings (ERA5 re-analysis) are used to steer long-term hindcasts, and next SM retrievals would be assimilated without or with minimal bias correction. However, an analysis of the differences between model- and satellite-based SSM (Section 5.2) revealed a complexity in the bias patterns that prevented a good prospect for a traditional DA application, and DA filters to update soil moisture have limitations as described in Busschaert et al. (2024). Therefore, the SSM retrievals are not directly assimilated but contrasted to the model to detect irrigation instead to provide microwave-based irrigation maps. The latter could constrain modeled irrigation estimates in future work. In addition, two important irrigation modeling advances are achieved: (i) activation of tiles within coarse-scale pixels (not the entire pixel is irrigated), and (ii) ensemble irrigation estimates are produced.

In this context, the ATBD discusses: (1) the capability of coarse resolution satellite soil moisture products to provide information on irrigated areas which can be used as input into LSMs; (2) the estimation of long-term irrigation datasets via an irrigation module activated within a LSM, using various deterministic runs to test different parameterizations; and (3) the quantification of long-term irrigation estimates through ensemble perturbation runs, which also provide irrigation uncertainty estimates. Section 5.2 - Irrigation Detection - discusses item (1), while Section 5.3 - Irrigation Quantification - covers items (2) and (3).

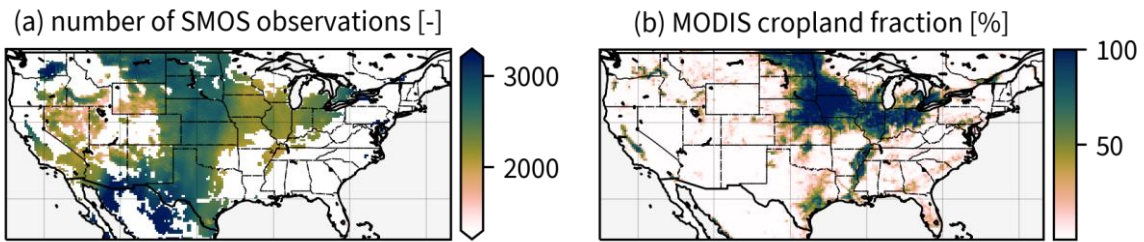
### 5.2. Irrigation detection

Two independent approaches were explored to detect irrigation by comparing SMOS surface soil moisture (SSM) satellite retrievals with Noah-MP estimated SSM. They are introduced below.

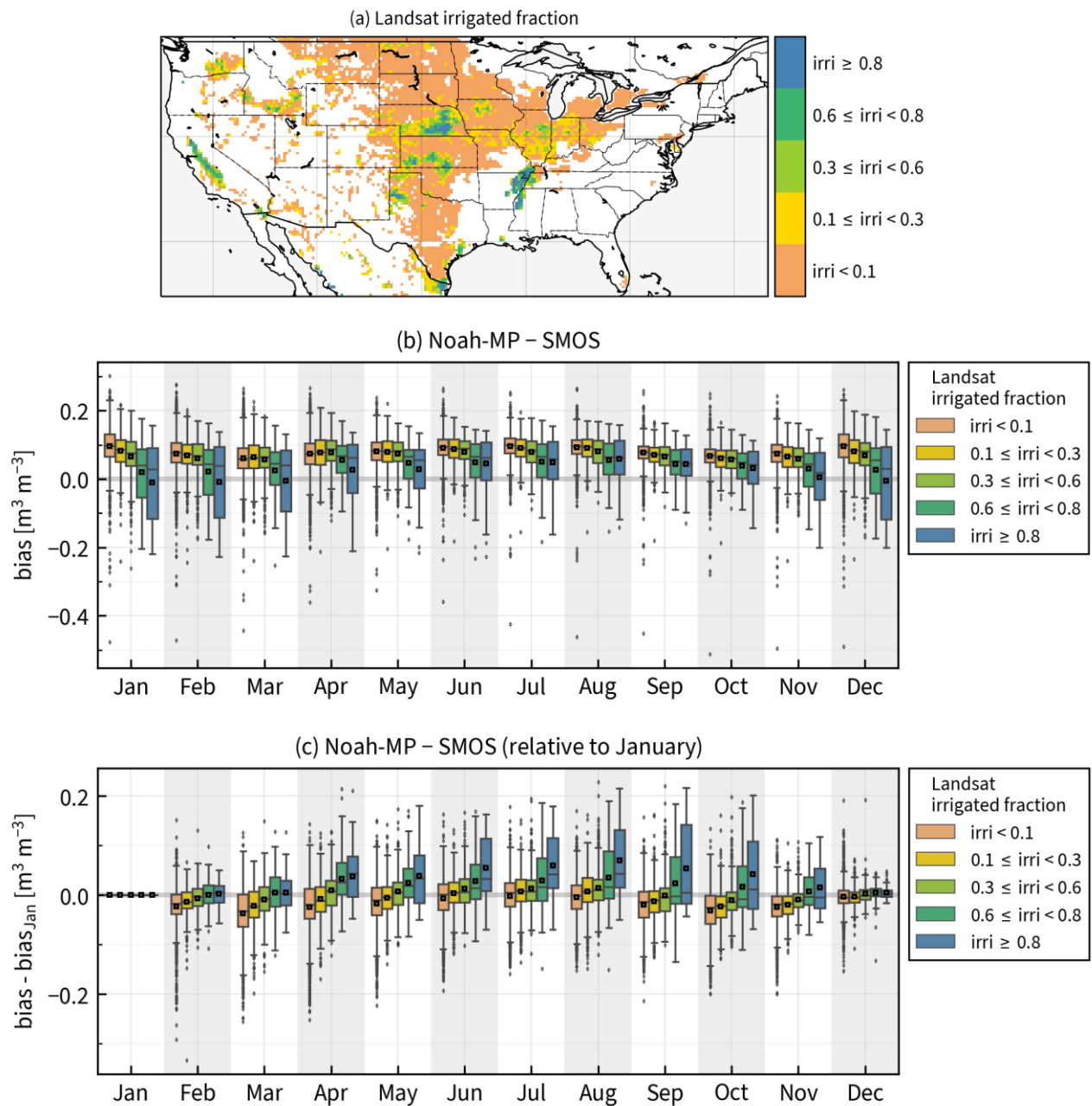
#### 5.2.1 Approach 1: relative bias

Differences between SMOS SSM (version 700, see Deliverable 2 for details) and Noah-MP SSM are evaluated over CONUS for the period January 2010 - March 2023. The model grid resolution is 0.25°. It is important to note that qualitative SMOS satellite retrievals are not available throughout the study domain. Figure 5.1a shows large masked areas due to, for example, topographically complex terrain (primarily in the west) and dense vegetation (primarily in the east). Any approach that is developed to assess irrigation with SMOS retrievals will therefore not work over these areas. However, grid cells that are characterised by a large crop land fraction tend to have observations available, which is evident from comparing the two panels of Figure 5.1 (an important exception is the southern Mississippi basin).





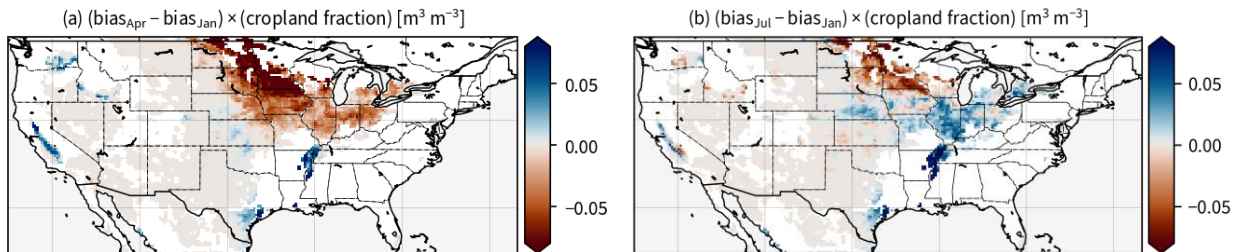
**Figure 5.1:** (a) Number of available SMOS observations during the study period. (b) MODIS-based cropland fraction.



**Figure 5.2:** (a) Landsat-based irrigated fraction, subdivided in five distinct classes. (b) Bias between Noah-MP estimated SSM and SMOS SSM retrievals (vertical axis), stratified per month (horizontal axis) and Landsat irrigated fraction (colour). The boxplots consist of grid cells, visualised in the map above. Only pixels with a nonzero cropland fraction are included. (c) As (b), but the bias for each pixel is relative to the bias in January.

A systematic bias between Noah-MP SSM estimates and SMOS SSM retrievals exists year-round, with Noah-MP SSM slightly more wet than SMOS SSM on average. This bias is stratified in Figure 5.2a based on two factors: the twelve months of the year and the Landsat-based irrigated fraction. We find that the wet bias of the model w.r.t. the satellite observations is most pronounced for grid cells with little irrigation for each month. As the irrigation increases, the bias becomes smaller or even negative (indicating that the model estimate is less wet or even drier than the satellite observation). A second observation is that for grid cells with limited irrigation, the variation of the bias throughout the year is much less pronounced than for the strongly irrigated pixels. By plotting the bias relative to that in January in Figure 5.2b, we find that this relative bias is near zero or slightly negative for pixels with little irrigation: it doesn't change much throughout the year. For grid cells with a lot of irrigation on the other hand, the relative bias increases, reaches a maximum in summer, and then decreases again. This indicates that the difference between the satellite-observed SSM retrieval and the model-based SSM estimates increases the most for such grid cells in the dry seasons. This may seem counterintuitive - one would expect the bias to decrease for the irrigated areas, as the model dries out but the satellite observations do not - but could be a result of a spurious vegetation-based signal into monthly climatological SMOS SSM anomalies (Crow and Feldman, 2025).

As an example, we show maps of the relative bias for the months of April (Figure 5.3a) and July (Figure 5.3b). Based on the observations discussed above, blue areas on the map suggest that irrigation is likely (positive relative bias), while red areas suggest that irrigation is not likely (negative relative bias). Known irrigation hotspots, such as the Mississippi basin and crop lands in California, Washington and Idaho are retrieved. However, other areas such as eastern Nebraska are missing. This area is captured in the map of July relative bias, although using this map to estimate the irrigated fraction would likely also result in many false positives, for example in Illinois, which has many non-irrigated crop lands but is blue in Figure 5.2.3b. Both maps correctly highlight crop lands in northern states such as Minnesota as non-irrigated.



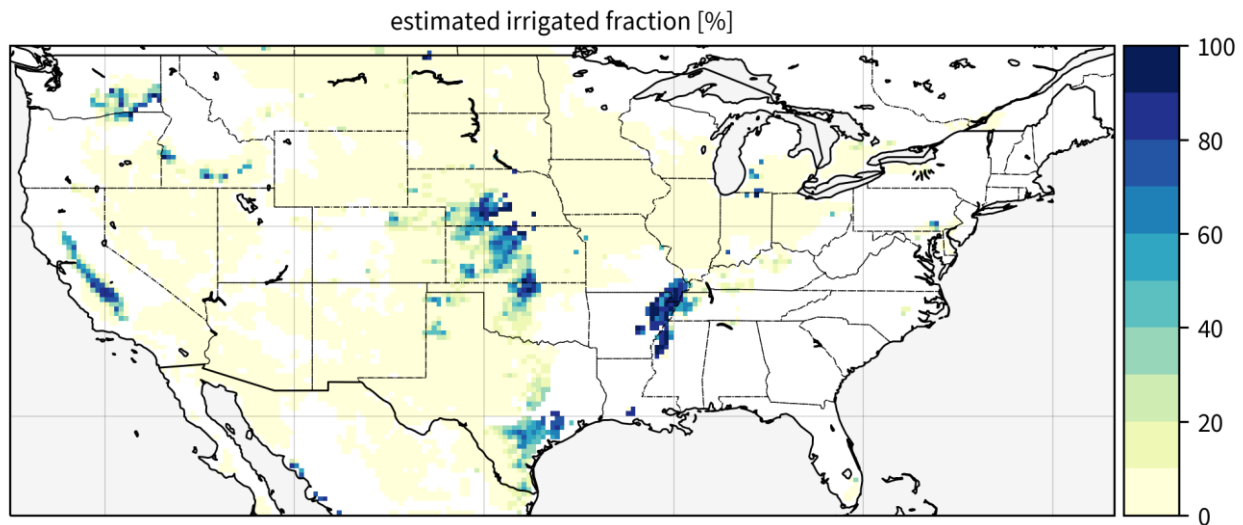
**Figure 5.3:** (a) Bias in April relative to that in January. (b) Bias in July relative to that in January. Grid cells with no SMOS observations are masked (white).

Based on the findings discussed above, the following simple rule was formulated to estimate the irrigated fraction of a grid cell:

$$\text{irrigated fraction} = (\text{cropland fraction}) \times \begin{cases} 1 & \text{if } \text{bias}_{\text{Apr}} > \text{bias}_{\text{Jan}} \\ 0 & \text{if } \text{bias}_{\text{Apr}} \leq \text{bias}_{\text{Jan}} \end{cases}$$

Figure 5.4 shows the map with the estimated irrigated fraction based on the relative bias approach.





**Figure 5.4:** Map of the estimated irrigated fraction based on the relative bias approach. Grid cells with no SMOS observations are masked (white).

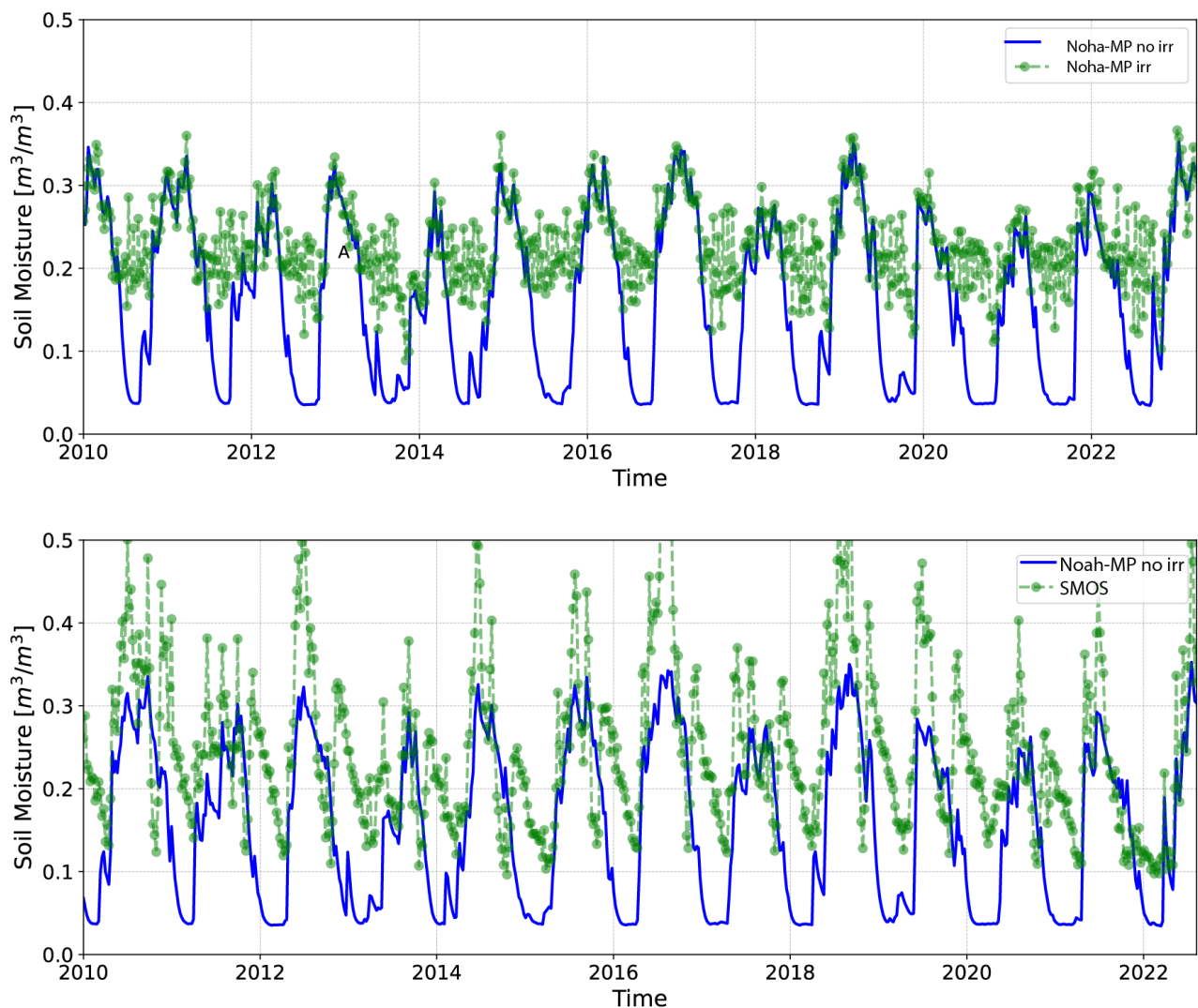
## 5.2.2 Approach 2: multiresolution analysis to detect irrigation

The multiresolution analysis irrigation detection method (MAID) is based on the evidence that irrigated and non irrigated lands are characterised by different soil moisture spectra. In particular, irrigation applications are able to change the classical signal seasonality (e.g., alternance of humid and dry periods) and this has reflections on a large range of scales – from the daily to the multiannual scale.

Classical Fourier transformation of the signals provides a way to extract their frequency signal components but is limited by the loss of the time information, that is, we can only see the frequency components of the signal but we are not able to understand when these frequencies characterise the signal in time. Wavelet analysis has been introduced to solve this issue: it provides a transformation of a signal in the frequency and time domain. This transformation is not done at no cost. When a signal is transformed in the wavelet domain it loses temporal resolution at large frequencies and loose frequency resolution at smaller time steps. For further details of the wavelet analysis the reader is referred to Daubechies, I., (1992). In practice while Fourier transformation preserves precisely all the frequency of the transformed signal, the wavelet provides an approximation of the frequencies at different times. This is called Heisenberg's indeterminacy principle, and is the concept used in quantum mechanics stating that there is a limit to the precision with which certain pairs of physical properties, such as time and frequency, can be simultaneously known.

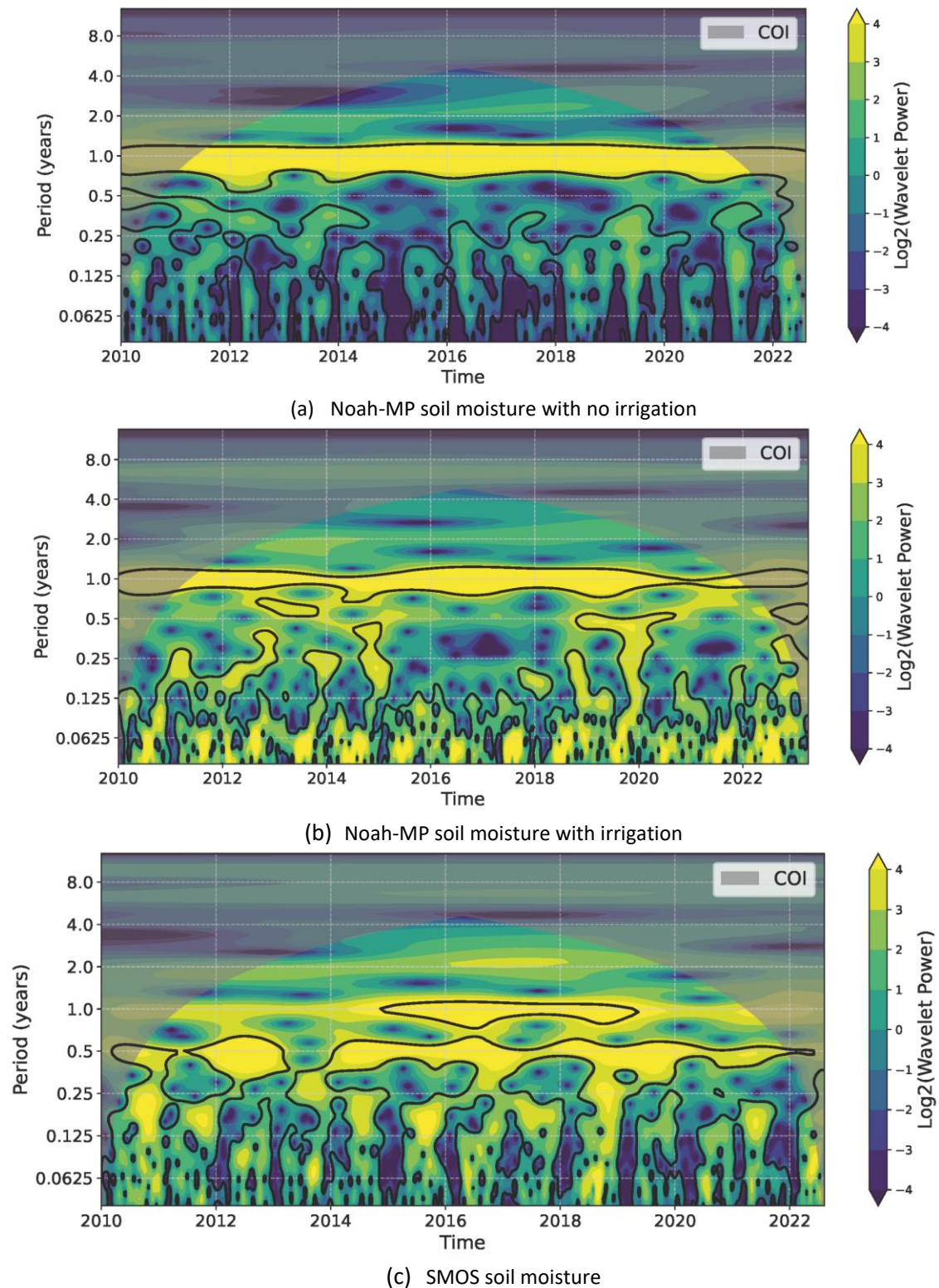
MAID was applied to two different experiments each one consisting of a comparison of two signals: 1) to Noha-MP surface soil moisture simulations with and without irrigation (to observe changes in wavelet power spectra) and 2) to Noha-MP and SMOS surface soil moisture to detect irrigated pixels.

Figure 5.5a shows an example of the time series of soil moisture aggregated at the weekly scale obtained by running the Noah-MP model with irrigation activated (irrigation fraction equal to 87%) and irrigation non activated. It can be seen that the soil moisture time series is significantly different. While spotting that this pixel is irrigated in a synthetic experiment like this one is straightforward, when comparing modelled soil moisture with irrigation not activated against satellite soil moisture observations the noise contained in the satellite signal, the contribution of irrigation on the total soil moisture signal and the inherent bias between the two can make the identification of a pixel irrigated or not irrigated extremely challenging (Zausingher et al. 2019)



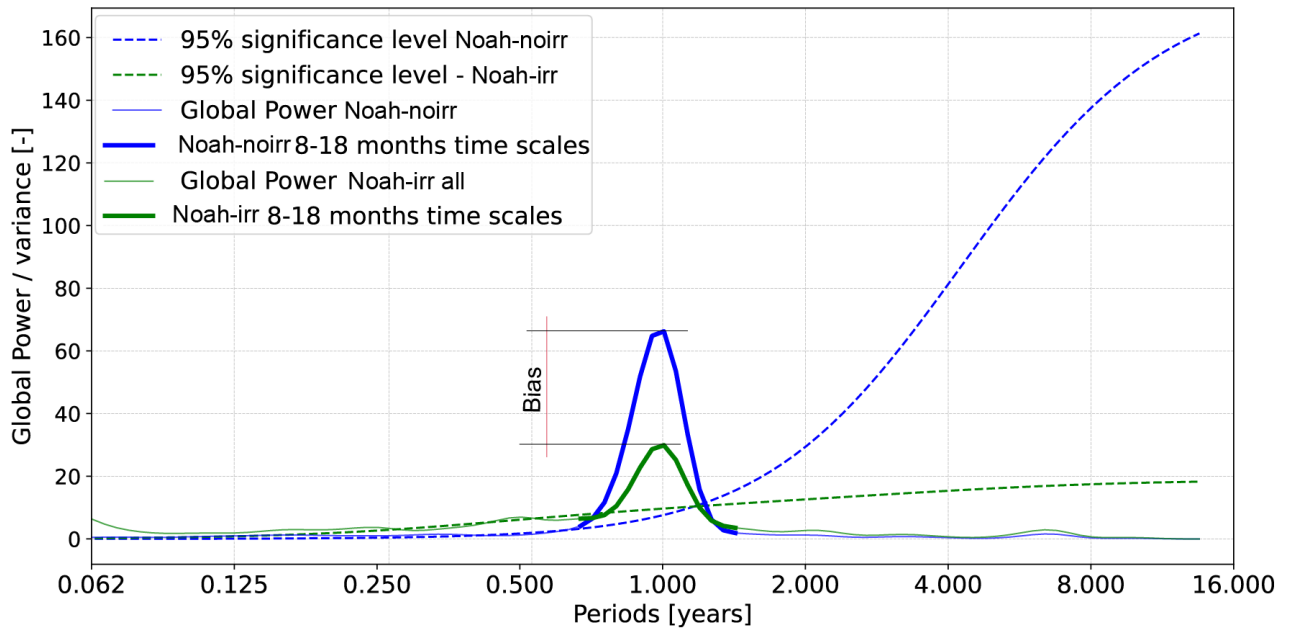
**Figure 5.5 :** (a) NOAH-MP weekly surface soil moisture time series comparison between simulations with no irrigation and the same pixel with an irrigation fraction of 87%. (b) same as (a) but compared against SMOS weekly aggregated surface soil moisture observations. The pixel is located in an area characterized by an arid climate with an Aridity Index (AI) equal 3.45. (Lon, Lat) = (-122.125, 39.625).

Figure 5.6 shows the wavelet power spectra of Noah-MP with no irrigation (Figure 5.6a), with irrigation (Figure 5.6b) and SMOS soil moisture time series (Figure 5.6c) for the same pixel of Figure 5.5, located in a semi-arid region. Figure 5.6a displays a strong power around the 1 year period due to the strong seasonality of the soil moisture typical of this area. Moderate power values are also present at sub-yearly time scales starting from smaller time scales and due to alternance of precipitation and hot periods. A very low power at 6 month scale can be also observed around 0.5 year denoting that the variance of the signal due to climatic factors has a low power at this frequency. Figure 5.6b and c show the same wavelet power spectra of before, however, notable differences can be seen in them: 1) a stronger power at small scales (below 0.125 years, i.e., less than a month) although this can be challenging to distinguish from noise in satellite observations; and 2) a much clear power increase at 6 monthly scale especially for the SMOS soil moisture.

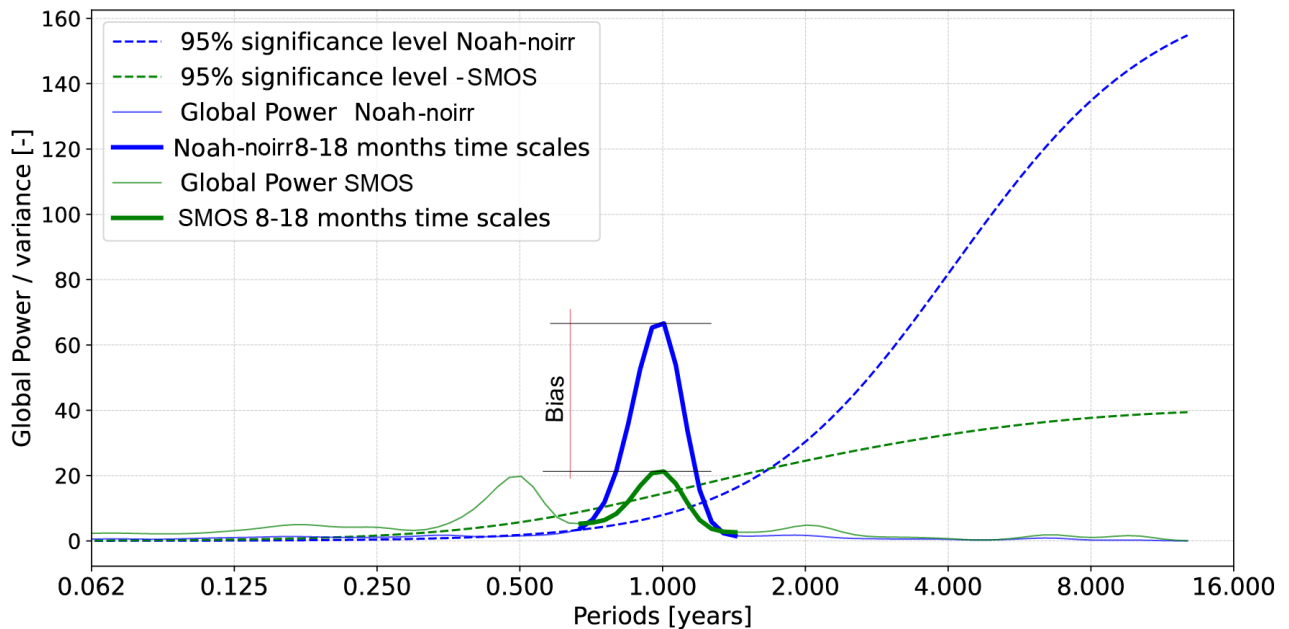


**Figure 5.6:** Wavelet power spectra comparison (a) NOAH-MP weekly surface soil moisture time series with no irrigation, (b) same as (a) with an irrigation fraction of 87%. (c) SMOS weekly aggregated surface soil moisture. The pixel is located in an area characterized by an arid climate with an Aridity Index (AI) equal 3.45. (Lon, Lat) = (-122.125, 39.625).

The differences in Figure 5.6 are more visible if we average the power spectrum in time so that obtaining the global power spectrum – the power averaged along the time dimension. The resulting spectra are then equivalent to the Fourier transformation of the signal (see Figure 5.7).



(a) Noah-MP soil moisture with irrigation averaged global power

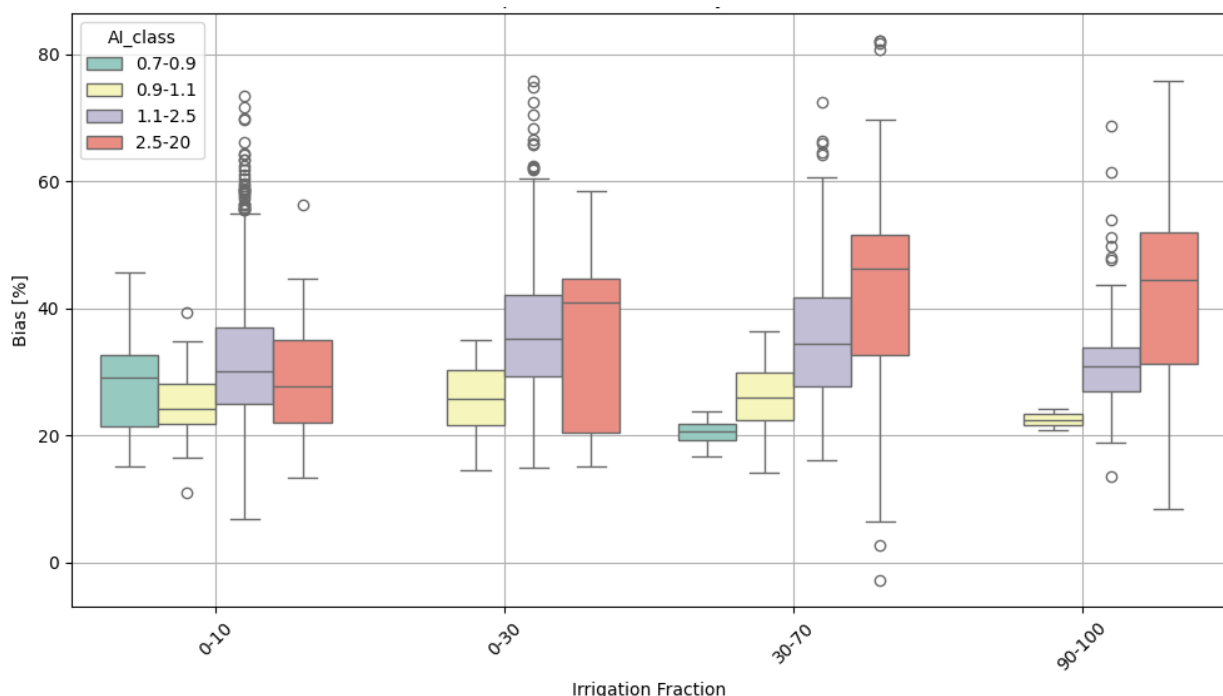


(b) SMOS soil moisture averaged global power

**Figure 5.7:** Wavelet global power spectra comparison (a) NOAH-MP weekly surface soil moisture time series with no irrigation vs Noah-MP weekly surface soil moisture time series with irrigation (same pixel above), (b) NOAH-MP weekly surface soil moisture time series with no irrigation vs SMOS weekly aggregated surface soil moisture global power spectra. It can be seen a bias in the power at the scale around one year which is then the most significant to consider for irrigation detection.

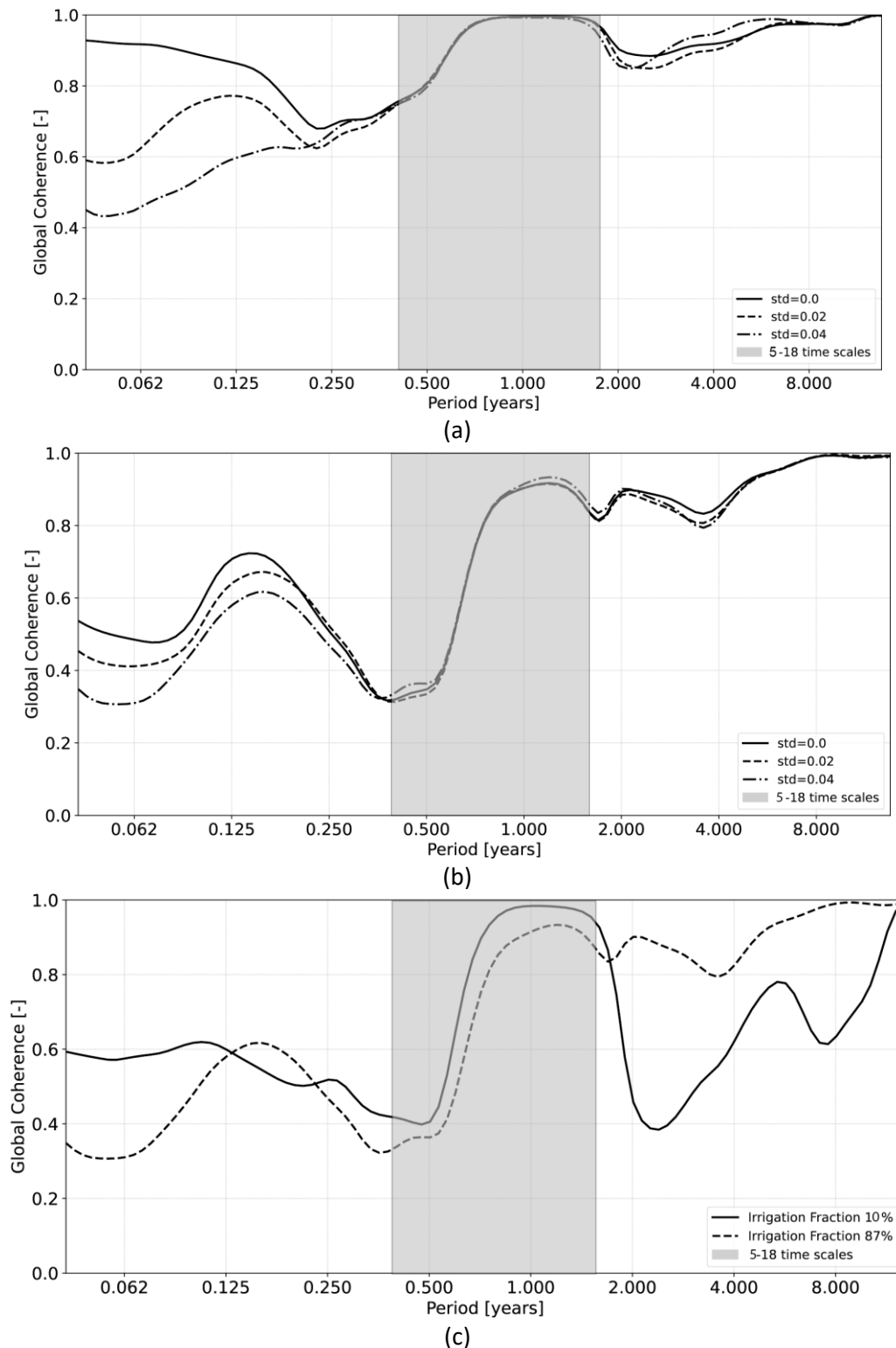


Figures 5.7a and b show that the “non-irrigated” soil moisture time series are characterised by a stronger power at the yearly time scale and an increase of power at the 6 month scales that in the case of SMOS soil moisture is beyond the significance level (obtained with red-noise assumption, see Torrence and Compo for further details). In other words it is observed a redistribution of the power of the signal across the scales that determines a bias in the global power and which is not impacted by noise which instead modifies the power of smaller scales. Of course this bias tends to vary with the specific climate and level of irrigation so it is not straightforward to understand the level of bias discriminating between irrigated and non irrigated pixels. Figure 5.8 displays the bias in the global wavelet power spectra as a function of irrigation fraction and aridity index. It can be seen that it tends to increase specifically for water limited regions while for energy limited regions this bias does not show substantial changes. When noise is added to the simulations (not shown) containing irrigation – to reproduce the realistic satellite soil moisture signal – a similar trend is observed but the values of the bias tend to increase.



**Figure 5.8:** Bias in the global wavelet power spectra between Noah-MP with irrigation and Noah-MP with no irrigation as a function of irrigation fraction and aridity index calculated over the crop pixels in the CONUS.

Another powerful tool to analyse the signal in the frequency domain is the wavelet coherence (Grinsted et al. 2004 ). The wavelet coherence is used to examine the relationship between two time series in the time-frequency domain. It provides a measure of the localized coherence (or similarity) between two signals at different scales and times, similar to a localized correlation coefficient. As it was done with the classical wavelet transform we can extract the power spectra of the wavelet coherence and its global version (i.e., the average in time of the spectra in the frequency and time domain) for irrigated and non irrigated soil moisture time series. The rationale is that two similar signals (both containing irrigation or no irrigation) shall have high correlation (and so high coherence) while two different signals (one irrigated and one not irrigated) shall yield a drop in coherence. Since noise is expected to impact the signal coherence, it is useful to focus on scales above five months. At the same time, to avoid the influence of edge effects on the transformation scales (see cone of influence in Figure 5.6) above 18 months can be neglected. Figure 5.9 provides an example of this drop of coherence for pixels simulated with an irrigation fraction of 10% (a) and 87% (b).

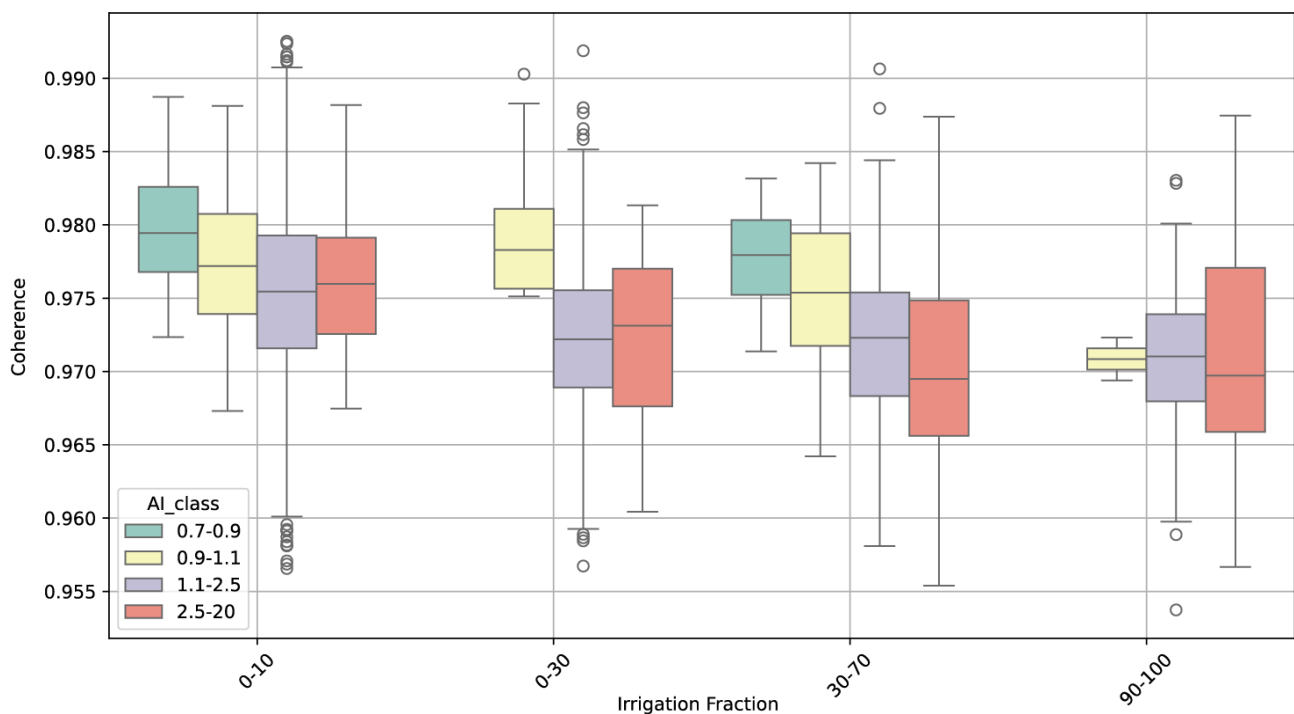


**Figure 5.9:** Global wavelet coherence power spectra between surface soil moisture simulations from Noah-MP with no irrigation and: a) surface soil moisture simulations from Noah-MP on a pixel with irrigation fraction equal to 10%; b) same as a) but for a pixel with irrigation fraction equal to 87%; (c) SMOS on the same pixels of (a) and (b). It can be noted the effect of noise in  $\text{m}^3/\text{m}^3$  added to the signal impacts the coherence only for scales below 5-6 months.

When, on the same pixels, we plot the coherence between SMOS and Noah-MP with no irrigation (Figure 5.2.9c) we found a similar pattern obtained in panels a and b. Note however that the irrigation fraction map

is only an indication of irrigated areas for a specific year and does not guarantee these values are truly representative of the real irrigation dynamics so the panel c must be interpreted only qualitatively.

As in the case of bias, the drop in coherence can be highly dependent on irrigation amount and climate. Figure 5.10 displays the drop in mean coherence obtained for Noah-MP with irrigation activated vs irrigation non activated as a function of irrigation fraction and aridity index. It can be seen that coherence tends to decrease with irrigation fraction and arid conditions while for energy limited regions this drop is more nuanced or not noticeable. Note that cold energy limited regions do not show presence of irrigation fraction larger than 90%.



**Figure 5.10:** Mean coherence distribution between Noah-MP with irrigation activated and irrigation non activated as a function of irrigation fraction and aridity index calculated over the crop pixels in the CONUS.

Based on the discussion above we trained a Random Forest Model to classify between irrigated pixels using Noah-MP simulations with irrigation activated and irrigation non activated. That is:

1. We assumed that if the irrigation fraction in the pixel is less than 20% then the pixel can be considered non irrigated – and irrigated otherwise.
2. For each pixel we aggregate the data at a weekly time scale and perform the wavelet transformation and coherence analysis of Noah-MP surface soil moisture simulations with irrigation activated and non activated and extracted the bias and mean coherence around 5-18 month time scales as in Figure 5.7 and Figure 5.9. For each pixel we also calculated the aridity index. To include the potential effect of noise in the simulations that shall represent the “satellite” observations (Noah-MP with irrigation activated) we added a random noise component with a zero mean and a standard deviation of 0.04 m<sup>3</sup>/m<sup>3</sup> prior to the aggregation and wavelet analysis.
3. We trained a Random Forest Model classifier on 80% of the pixels using as predictors AI, Bias and Mean coherence above. We tested it over 20% of the remaining sample. Since, the most critical settings in the Random forest classifiers are the number of trees in the forest (*Nest*) and the number of features considered for splitting at each node (*max\_features*), we employed a grid search

procedure using the GridSearchCV function from the scikit-learn library ([\url{https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.GridSearchCV.html}](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html)). Specifically, we varied Nest between 100 and 2000 in steps of 100 and let the number of samples required to split a node to vary between 2 and 10 in steps of 2 (max\_features). The grid search yields Nest=1000 trees and max\_features= 2.

4. We applied the calibrated classifier to the SMOS soil moisture retrieval time series (after filling them with a max gap of five days) and aggregated it on a weekly time step. Time series resulting in more than 30% of data gaps were initially excluded from the analysis. This limited the number of pixels to those in southern latitudes because SMOS is heavily masked in relatively cold regions.
5. We quantified the goodness of the classification via Receiver Operating Characteristic (ROC) curve. ROC provides valuable insights into the performance of a binary classification model. The ROC curve is created by plotting the true positive ratio (TPR) against the false positive ratio (FPR) at various threshold settings. Each point on the curve represents a different decision threshold for the classifier. The curve starts at \$(0,0)\$ and ends at \$(1,1)\$ (the diagonal line). The diagonal line represents a random classifier (no discrimination ability). Any model below this line is worse than random guessing. Normally a metric for quantifying how good is the classification across all thresholds is the Area Under the Curve (AUC). AUC = 0.5 denotes that the model has no discrimination ability (equivalent to random guessing) while values above suggest the model can discriminate between irrigated and non irrigated areas with AUC>0.9 denoting excellent model performance. We also used other metrics which are reported in the validation report.

### 5.3. Irrigation quantification

#### 5.3.1 The Noah-MP Land Surface Model

The LSM selected for this study is the Noah-MP (Niu et al., 2011) LSM, which is able to dynamically simulate vegetation and provides soil moisture estimates for four layers up to 2 m (i.e., 0-10, 10-40, 40-100 and 100-200 cm depth). In particular, the Noah-MP version 4.0.1 embedded within the NASA Land Information System (LIS; Kumar et al., 2006; Peters-Lidard et al., 2007) framework is selected, and coupled to a sprinkler irrigation scheme (Ozdogan et al., 2010).

The irrigation module simulates sprinkler systems by adding water as additional rainfall; however, it does not alter other processes, such as vapor fluxes, which would be necessary to represent more efficient drip irrigation systems. The additional water is applied over a four-hour period during the day, from 6:00 am to 10:00 am local time. Irrigation locations are determined based on irrigated cropland derived from a land cover map and a static map of pixel-based irrigated land fraction. The timing of irrigation is driven by both the growing season and the root-zone soil moisture availability. By default, the growing season is derived from climatological monthly greenness vegetation fraction (GVF) maps, provided as model inputs. The threshold defining the growing/non-growing season is computed based on the minimum and maximum values of GVF ( $GVF_{min}$  and  $GVF_{max}$ ) as follows:

$$GVF_{irr} = GVF_{min} + 0.4 * (GVF_{max} - GVF_{min}) \quad (1)$$

During the growing season, irrigation is applied when the root-zone moisture availability (MA) falls below a user-defined threshold ( $Th_{irr}$  hereafter). MA is defined as follows:

$$MA = \frac{\theta - \theta_{WP}}{\theta_{FC} - \theta_{WP}} \quad (2)$$



where  $\theta_{WP}$  is the water content at wilting point,  $\theta_{FC}$  at field capacity and  $\theta$  is the actual water content. In previous works a value of  $Th_{irr}=0.45$  was selected (Modanesi et al., 2022; Busschaert et al., 2023; De Lannoy et al., 2024). The irrigation amount is determined by the water needed to restore root-zone soil moisture to field capacity. An important factor influencing the irrigation amount is the maximum rooting depth, which is associated with a default crop type. The root-zone is therefore defined as follows:

$$RZ = RZ_{max} * GVF \quad (3)$$

where GVF is the GVF derived from the climatological GVF maps and  $RZ_{max}$  is the maximum rooting depth.

For the experiments described at Section 5.3.2 and 5.3.3, the Noah-MP LSM was run at 0.25° degree spatial resolution, on the CCI soil moisture spatial grid, and at 15 minutes time scale. Outputs are provided at a daily average resolution. Input data include the land cover, obtained from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS; Friedl et al. 2002) and the Harmonized World Soil Database (HWSD; Nachtergaele et al. 2010), used to derive soil texture information. The widely-used GRIPC map is updated using a Landsat-derived irrigation map (Teluguntla et al., 2023) and adapted to the LIS-required format. The Landsat irrigation fraction map represents the percentage of irrigated areas at pixel-scale. Crop type is set to a generic type with a maximum rooting depth of 1 m, while GVF data (Gutman and Ignatov, 1998) were derived from five years (1985–1989) of normalized difference vegetation index (NDVI) data from AVHRR (Miller et al., 2006). Meteorological forcings are extracted from the lowest model level forecasts of the European Centre for Medium-Range Weather Forecasts' fifth-generation atmospheric reanalysis (ERA5; Hersbach et al., 2020). ERA5 forecasts do not incorporate radar or gauge-based precipitation data over Europe, limiting the model's capacity to account for anthropogenic effects.

The model runs with the default LIS options except for the runoff and groundwater option, which is set to free drainage (Yang and Dickinson, 1996). The vegetation is dynamically simulated using the maximum vegetation fraction (option 5) and based on a simple vegetation scheme (Dickinson et al., 1998).

### 5.3.2 Deterministic experiments

This section describes the deterministic experiments conducted to optimise the parameterization of a sprinkler irrigation scheme incorporated into the Noah-MP LSM. By varying the scheme parameters, we aim at identifying the optimal configuration for subsequent ensemble runs. Specifically, the experiments examined the effects of two key variables: the growing season (defining the period during which irrigation is allowed) and the irrigation threshold parameter ( $Th_{irr}$ ) applied to MA (see equation 2).

To test the effect of the growing season on irrigation quantification, three approaches are tested:

- (1) The first approach is the default method used in the Noah-MP v4.0.1 model within the LIS framework. This method relies on monthly climatological GVF maps, applying a threshold to GVF as defined in Equation 1. This approach does not incorporate data on dynamic vegetation simulated by the model. However, given the limitations identified in Noah-MP's dynamic vegetation capabilities (Modanesi et al., 2021; 2022), this method remains valuable for investigation.
- (2) The second approach allows irrigation over the entire year (year-round irrigation). This method is valuable for exploring potential applications of irrigation during winter months, particularly in drier regions or areas impacted by drought. It is especially relevant given that the climatological GVF maps (Miller et al., 2006) may be outdated and may not accurately capture the presence of winter crops.

- (3) The third and final approach, developed by Modanesi et al. (2022) and applied here to Noah-MP v4.0.1, defines the growing season based on a user-specified threshold of the simulated Leaf Area Index ( $LAI > 1$ ). The LAI values are then converted to GVF following the methodology outlined by Fang et al. (2018):

$$GVF = 1 - e^{-b \cdot LAI} \quad (4)$$

where  $b$  [-] is an empirical parameter that depends on plant canopy (Fang et al., 2018; Norman et al., 1995).

To evaluate the effects of the  $Th_{irr}$  applied to MA, the default climatological GVF growing season is selected and four different  $Th_{irr}$  values are tested: 0.35, 0.45, 0.55, and 0.65. For the year-round growing season approach,  $Th_{irr}=0.45$  and  $Th_{irr}=0.65$  are tested while for the dynamic vegetation-based growing season, only  $Th_{irr}=0.45$  is selected.

An additional experiment is conducted to evaluate a tiling land cover option that incorporates different land cover type percentages within the same grid cell (at a 25 km spatial resolution). In this configuration, LIS divides each grid cell into tiles based on the land cover map, allowing for simulation of subgrid-scale heterogeneity. When MA falls below  $Th_{irr}$ , the irrigation scheme checks if the current tile is classified as cropland (or other potentially irrigated types, such as grassland). If the ratio of irrigated area to total crop coverage within the grid cell is less than one, the irrigation requirement for each tile is reduced proportionally. Conversely, if this ratio exceeds one, the irrigation requirement is scaled up accordingly, with additional water applied to grasslands or other non-cropland types, excluding forests, bare soil, and urban tiles (Ozdogan et al., 2010). This option, which is more computationally expensive, is applied to the climatological GVF growing season and to a  $Th_{irr}=0.45$ .

A summary of the eight deterministic experiments conducted is presented in **Table 5.1**.

EXPERIMENTS (Time span January 2010- December 2022)					
LSM	IRRIGATION SCHEME	FORCING	GROWING SEASON	TILING	$Th_{irr}$
Noah-MPv.4.0.1 (Niu et al., 2011)	Sprinkler Irrigation scheme (Ozdogan et al., 2010)	ERA 5 (Hersbach et al., 2020)	default climatological GVF	not applied	0.35
				not applied	0.45
				not applied	0.55
				not applied	0.65
				applied	0.45
			Year-round irrigation	not applied	0.45
				not applied	0.65

			Based on dynamic vegetation (when LAI>1; Modanesi et al., 2022)	not applied	0.45
--	--	--	---	-------------	------

Table 5.1: provides an overview of all deterministic runs conducted, detailing the various growing season options,  $Th_{irr}$  values, and the tiling option.

### 5.3.3 Ensemble runs

Following a similar setup as described in section 5.3.2, the Noah-MPv4.0.1 model was run in ensemble mode. Given the evaluation of the deterministic runs, the default climatological GVF option was selected to define the growing season, and the  $Th_{irr}$  was set to 0.45. The simulation is run without tiling (NO TILING), using the dominant land cover type for the grid cell.

The ensemble of land model trajectories is composed of 24 ensemble members for which selected meteorological input forcings are perturbed (rainfall, incident longwave and shortwave radiation) as in Modanesi et al. (2022). The ensemble runs provide an estimation of the irrigation uncertainty, where the ensemble standard deviation (or ensemble spread) is calculated for each day  $t$  ( $\sigma_t$  [mm/day]). The temporally average ensemble standard spread ( $\sigma_{mean}$ ) can be defined for each location as follows.

$$\sigma_{mean} = \sqrt{\frac{1}{N^2} * \sum_{i=1}^N \sigma_i^2} \quad (5)$$

where  $N$  is the number of days with irrigation (i.e. with  $\sigma_i^2 > 0$ ). The coefficient of variation (CV) can then be calculated for each location by dividing the  $\sigma_{mean}$  by the mean irrigation rate (calculated over the  $N$  irrigated days).

Because the daily gridded irrigation estimates are aggregated per state and over a full year, for an evaluation against FRIS data for 2013 and 2018, a corresponding spatial estimate of the uncertainty was also produced. The uncertainty for each state is defined as follows.

$$\sigma_{mean} = \sqrt{\sum_{p=1}^P A_p^2 * (\frac{1}{P} * \sum_{p=1}^P \sigma_p^2)} \quad (6)$$

with  $P$ , the number of irrigated pixels in the region (state),  $A_p$  the area of each individual pixel  $p$ , and  $\sigma_p^2$  the sum of the daily ensemble variances for the pixel  $p$  (for the year 2013 or 2018).

Note that the deterministic and ensemble simulations with the default Noah-MP setup produce irrigation estimates for the input irrigation fraction per pixel. A closer analysis of the time series showed that the consequent use of a fractional irrigation value as proxy rainfall input to the entire coarse model pixel causes that the soil moisture does not return to field capacity and a new irrigation event is triggered. This shortcoming of the Noah-MP model will be circumvented in a new iteration of results.

---

## References

- Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V. (2014). Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *J. Geophys. Res.*, 119, 5128–5141. <https://doi.org/10.1002/2014JD021489>.
- Brocca, L., Massari, C., Ciabatta, L., Moramarco, T., Penna, D., Zuecco, G., Pianezzola, L., Borga, M., Matgen, P., and MartínezFernández, J. (2015). Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of the SM2RAIN algorithm. *J. Hydrol. Hydromechan.*, 63, 201–209, <https://doi.org/10.1515/johh-2015-0016>.
- Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., and Fernández-Prieto, D. (2018). How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. *Int. J. Appl. Earth Observ. Geoinform.*, 73C, 752–766, <https://doi.org/10.1016/j.jag.2018.08.023>.
- Brombacher, J., Silva, I.R. de O., Degen, J., Pelgrum, H., 2022. A novel evapotranspiration based irrigation quantification method using the hydrological similar pixels algorithm. *Agricultural Water Management* 267, 107602. <https://doi.org/10.1016/j.agwat.2022.107602>
- Busschaert, L., Bechtold, M., Modanesi, S., Massari, C., Brocca, L., & De Lannoy, G. (2024). Irrigation quantification through backscatter data assimilation with a buddy check approach. *Journal of Advances in Modeling Earth Systems*, 16(3), e2023MS003661
- Crow, Wade T., and Andrew F. Feldman. Vegetation signal crosstalk present in official SMAP surface soil moisture retrievals. *Remote Sensing of Environment* 316 (2025): 114466.
- Dari, J., Brocca, L., Quintana-Seguí, P., Escorihuela, M. J., Stefan, V., and Morbidelli, R. (2020). Exploiting high-resolution remote sensing soil moisture to estimate irrigation water amounts over a Mediterranean region. *Remote Sens.*, 12, 2593, <https://doi.org/10.3390/rs12162593>.
- Dari, J., Brocca, L., Modanesi, S., Massari, C., Tarpanelli, A., Barbetta, S., Quast, R., Vreugdenhil, M., Freeman, V., BarellaOrtiz, A., Quintana-Seguí, P., Bretreger, D., and Volden, E. (2023). Regional data sets of high-resolution (1 and 6 km) irrigation estimates from space. *Earth Syst. Sci. Data*, 15, 1555–1575, <https://doi.org/10.5194/essd-15-1555-2023>.
- Dari, J., Filippucci, P., and Brocca, L. (2024). The development of an operational system for estimating irrigation water use reveals socio-political dynamics in Ukraine. *Hydrol. Earth Syst. Sci.*, 28, 2651–2659, <https://doi.org/10.5194/hess-28-2651-2024>.
- Dari, J., Quintana-Seguí, P., Morbidelli, R., Saltalippi, C., Flammini, A., Giugliarelli, E., Escorihuela, M. J., Stefan, V., and Brocca, L. (2022). Irrigation estimates from space: implementation of different approaches to model the evapotranspiration contribution within a soil-moisture-based inversion algorithm. *Agr. Water Manage.*, 265, 107537, <https://doi.org/10.1016/j.agwat.2022.107537>.

---

Daubechies, I., 1992. Ten lectures on wavelets, vol. 61 of CBMS-NSF regional conference series in applied mathematics.

De Lannoy, G. J., Bechtold, M., Busschaert, L., Heyvaert, Z., Modanesi, S., Dunmire, D., ... & Massari, C. (2024). Contributions of irrigation modeling, soil moisture and snow data assimilation to high-resolution water budget estimates over the Po basin: Progress towards digital replicas. *Journal of Advances in Modeling Earth Systems*, 16(10), e2024MS004433.

Dickinson, R. E., M. Shaikh, R. Bryant, and L. Graumlich (1998), Interactive canopies for a climate model, *J. Clim.*, 11, 2823–2836, doi:10.1175/1520-0442(1998)011<2823:ICFACM>2.0.CO;2.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P.D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y.Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S.I., Smolander, T., Lecomte, P., 2017. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment, Earth Observation of Essential Climate Variables* 203, 185–215. <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.

van Eekelen, M.W., Bastiaanssen, W.G.M., Jarman, C., Jackson, B., Ferreira, F., van der Zaag, P., Saraiva Okello, A., Bosch, J., Dye, P., Bastidas-Obando, E., Dost, R.J.J., Luxemburg, W.M.J., 2015. A novel approach to estimate direct and indirect water withdrawals from satellite measurements: A case study from the Incomati basin. *Agriculture, Ecosystems & Environment* 200, 126–142. <https://doi.org/10.1016/j.agee.2014.10.023>

Escorihuela, M.J., Quintana-Seguí, P., 2016. Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes. *Remote Sensing of Environment, Special Issue: ESA's Soil Moisture and Ocean Salinity Mission - Achievements and Applications* 180, 99–114. <https://doi.org/10.1016/j.rse.2016.02.046>

Fang, L., Zhan, X., Hain, C.R., and Liu, J. (2018). Impact of using near real-time green vegetation fraction in Noah land surface model of NOAA NCEP on numerical weather predictions. *Advances in Meteorology*, <https://doi.org/10.1155/2018/9256396>.

Filippucci, P., Tarpanelli, A., Massari, C., Serafini, A., Strati, V., Alberi, M., Raptis, K. G. C., Mantovani, F., and Brocca, L. (2020). Soil moisture as a potential variable for tracking and quantifying irrigation: a case study with proximal gamma-ray spectroscopy data, *Adv. Water Resour.*, 136, 103502, <https://doi.org/10.1016/j.advwatres.2019.103502>.

Fiedl, M. A., and Coauthors (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sens. Environ.*, 83, 287–302, [https://doi.org/10.1016/S0034-4257\(02\)00078-0](https://doi.org/10.1016/S0034-4257(02)00078-0).

Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics* 11, 561–566. <https://doi.org/10.5194/npg-11-561-2004>

- 
- Güntner, A., 2008. Improvement of Global Hydrological Models Using GRACE Data. *Surv Geophys* 29, 375–397. <https://doi.org/10.1007/s10712-008-9038-y>
- Gutman, G. and Ignatov, A. (1998). The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction 720 models, *Int. J. Remote Sens.*, 19, 1533–1543, <https://doi.org/10.1080/014311698215333>.
- Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H., Pappenberger, F., 2020. GloFAS-ERA5 operational global river discharge reanalysis 1979–present. *Earth System Science Data* 12, 2043–2060. <https://doi.org/10.5194/essd-12-2043-2020>
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Jalilvand, E., Tajrishy, M., Hashemi, S. A. G., and Brocca, L. (2019). Quantification of irrigation water using remote sensing of soil moisture in a semi-arid region. *Remote Sens. Environ.*, 231, 111226, <https://doi.org/10.1016/j.rse.2019.111226>.
- Kragh, S.J., Fensholt, R., Stisen, S., Koch, J. (2023). The precision of satellite-based net irrigation quantification in the Indus and Ganges basins. *Hydrol. Earth Syst. Sci.*, 27, 2463–2478. <https://doi.org/10.5194/hess-27-2463-2023>
- Kragh, S.J., Dari, J., Modanesi, S., Massari, C., Brocca, L., Fensholt, R., Stisen, S., Koch, J., 2024. An inter-comparison of approaches and frameworks to quantify irrigation from satellite data. *Hydrology and Earth System Sciences* 28, 441–457. <https://doi.org/10.5194/hess-28-441-2024>
- Kumar, S. V., Peters-Lidard, C., Tian, Y., Houser, P., Geiger, J., Olden, S., Lighty, L., Eastman, J., Doty, B., and Dirmeyer, P. (2006): Land information system: An interoperable framework for high resolution land surface modeling. *Environmental Modelling & Software* 21, 1402–1415. <https://doi.org/10.1016/j.envsoft.2005.07.004>.
- Landerer, F.W., Dickey, J.O., Güntner, A., 2010. Terrestrial water budget of the Eurasian pan-Arctic from GRACE satellite measurements during 2003–2009. *Journal of Geophysical Research: Atmospheres* 115. <https://doi.org/10.1029/2010JD014584>
- Lehner, B., Grill, G., 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes* 27, 2171–2186. <https://doi.org/10.1002/hyp.9740>
- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C. (2017). GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geosci. Model Dev.*, 10, 1903–1925, <https://doi.org/10.5194/gmd-10-1903-2017>.
- Massari, C., Brocca, L., Tarpanelli, A., Moramarco, T., 2015. Data Assimilation of Satellite Soil Moisture into Rainfall-Runoff Modelling: A Complex Recipe? *Remote Sensing* 7, 11403–11433. <https://doi.org/10.3390/rs70911403>
- 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.

---

Mehta, P., Siebert, S., Kummu, M., Deng, Q., Ali, T., Marston, L., Xie, W., and Davis, K. (2022). Global Area Equipped for Irrigation Dataset 1900–2015 v3.0, Zenodo [data set], <https://doi.org/10.5281/zenodo.6886564>.

Miller, J., Barlage, M., Zeng, X., Wei, H., Mitchell, K., and Tarpley, D. (2006). Sensitivity of the NCEP/Noah land surface model to the MODIS green vegetation fraction data set, *Geophys. Res. Lett.*, 33, L13404, <https://doi.org/10.1029/2006GL026636>.

Modanesi, S., Massari, C., Bechtold, M., Lievens, H., Tarpanelli, A., Brocca, L., Zappa, L., De Lannoy, G.J.M (2022). Challenges and benefits of quantifying irrigation through the assimilation of Sentinel-1 backscatter observations into Noah-MP. *Hydrol. Earth Syst. Sci.*, 26, 4685–4706. <https://doi.org/10.5194/hess-26-4685-2022>.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data* 13, 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>

Nachtergaele, F., and Coauthors (2010). The harmonized world soil database. *Proc. 19th World Congress of Soil Science, Soil Solutions for a Changing World*, Brisbane, Australia, International Union of Soil Sciences, 34–37, <https://edepot.wur.nl/154132>.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.-Atmos.*, 116, 1–19, <https://doi.org/10.1029/2010JD015139>.

Norman, J.M., Kustas, W.P., and Humes, K.S. (1995). Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology* 77, 263–293, [https://doi.org/10.1016/0168-1923\(95\)02265-Y](https://doi.org/10.1016/0168-1923(95)02265-Y).

Ozdogan, M., Rodell, M., Beaudoin, H. K., and Toll, D. L. (2010). Simulating the effects of irrigation over the United States in a land surface model based on satellite derived agricultural data, *J. Hydrometeorol.*, 11, 171–184, <https://doi.org/10.1175/2009JHM1116.1>.

Pascolini-Campbell, M.A., Reager, J.T., Fisher, J.B., 2020. GRACE-based Mass Conservation as a Validation Target for Basin-Scale Evapotranspiration in the Contiguous United States. *Water Resources Research* 56, e2019WR026594. <https://doi.org/10.1029/2019WR026594>

Pascolini-Campbell, M., Fisher, J.B., Reager, J.T., 2021. GRACE-FO and ECOSTRESS Synergies Constrain Fine-Scale Impacts on the Water Balance. *Geophysical Research Letters* 48, e2021GL093984. <https://doi.org/10.1029/2021GL093984>

Peters-Lidard, C.D., Houser, P.R., Tian, Y., Kumar, S.V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E.F., and Sheffield, J (2007). High-performance Earth system modeling with



---

NASA/GSFC's Land Information System. Innov. Syst. Softw. Eng. 3, 157–165. <https://doi.org/10.1007/s11334-007-0028-x>.

Portmann, F., Siebert, S., Bauer, C., and Döll, P. (2008). Global data set of monthly growing areas of 26 irrigated crops, Version 1.0, Frankfurt Hydrol. Pap. 06, Inst. of Phys. Geogr., Univ. of Frankfurt, Frankfurt, Germany, 179–180 pp.

Ramillien, G., Frappart, F., Güntner, A., Ngo-Duc, T., Cazenave, A., Laval, K., 2006. Time variations of the regional evapotranspiration rate from Gravity Recovery and Climate Experiment (GRACE) satellite gravimetry. Water Resources Research 42. <https://doi.org/10.1029/2005WR004331>

Rodell, M., Beaulieu, 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., Fetzner, 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.

Salmon, J. M., Friedl, M. A., Froking, S., and Wisser, D.: Global rain-fed, irrigated, and paddy croplands (2015). A new high resolution map derived from remote sensing, crop inventories and climate data, Int. J. Appl. Earth Obs., 38, 321–334, <https://doi.org/10.1016/j.jag.2015.01.014>.

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>

Swann, A.L.S., Koven, C.D., 2017. A Direct Estimate of the Seasonal Cycle of Evapotranspiration over the Amazon Basin. Journal of Hydrometeorology 18, 2173–2185. <https://doi.org/10.1175/JHM-D-17-0004.1>

Teluguntla, P., Thenkabail, P., Oliphant, A., Gumma, M., Anece, I., Foley, D. and McCormick, R. (2023). Landsat-derived Global Rainfed and Irrigated-Cropland Product @ 30-m (LGRIP30) of the World (GFSADLGRIP30WORLD). The Land Processes Distributed Active Archive Center (LP DAAC) of NASA and USGS. Pp. 103. IP-148728. DOI: <https://doi.org/10.5067/Community/LGRIP/LGRIP30.001>

Torrence, C., Compo, G.P., 1998. A Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological Society 79, 61–78.

Watkins, M.M., Wiese, D.N., Yuan, D.-N., Boening, C., Landerer, F.W., 2015. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. Journal of Geophysical Research: Solid Earth 120, 2648–2671. <https://doi.org/10.1002/2014JB011547>

Wiese, D.N., Landerer, F.W., Watkins, M.M., 2016. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. Water Resources Research 52, 7490–7502. <https://doi.org/10.1002/2016WR019344>

Yang, Z.-L., and R. E. Dickinson (1996). Description of the Biosphere-Atmosphere Transfer Scheme (BATS) for the soil moisture workshop and evaluation of its performance, Global Planet. Change, 13, 117–134, doi:10.1016/0921-8181(95)00041-0.



---

Zappa, L., Schlaffer, S., Bauer-Marschallinger, B., Nendel, C., Zimmerman, B., Dorigo, W., 2021. Detection and Quantification of Irrigation Water Amounts at 500 m Using Sentinel-1 Surface Soil Moisture. *Remote Sensing* 13, 1727. <https://doi.org/10.3390/rs13091727>

Zappa, L., Schlaffer, S., Brocca, L., Vreugdenhil, M., Nendel, C., Dorigo, W., 2022. How accurately can we retrieve irrigation timing and water amounts from (satellite) soil moisture? *International Journal of Applied Earth Observation and Geoinformation* 113, 102979. <https://doi.org/10.1016/j.jag.2022.102979>

Zappa, L., Dari, J., Modanesi, S., Quast, R., Brocca, L., De Lannoy, G., Massari, C., Quintana-Seguí, P., Barella-Ortiz, A., Dorigo, W., 2024. Benefits and pitfalls of irrigation timing and water amounts derived from satellite soil moisture. *Agricultural Water Management* 295, 108773. <https://doi.org/10.1016/j.agwat.2024.108773>

Zaussinger, F., Dorigo, W., Gruber, A., Tarpanelli, A., Filippucci, P., and Brocca, L.: 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>, 2019.