Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

IWU-SMinv v1.0 and v1.1: Two satellite-based irrigation water use datasets at 1 km spatial resolution over Europe

Jacopo Dari

jacopo.dari@unipg.it

University of Perugia

Yogesh Kumar Baljeet Singh
Brockmann Consult GmbH

Konstantin Ntokas
Brockmann Consult GmbH

Norman Fomferra
Brockmann Consult GmbH

Gunnar Brandt
Brockmann Consult GmbH

Renato Morbidelli
University of Perugia

Carla Saltalippi
University of Perugia

Alessia Flammini
University of Perugia

Francesco Leopardi
University of Perugia

Mehdi Rahmati
Institute of Bio and Geosciences (IBG-3), Forschungszentrum Jülich

Paolo Filippucci
National Research Council

Stefania Camici
National Research Council

Diego Fernández-Prieto
European Space Agency, ESA-ESRIN

Espen Volden
European Space Agency, ESA-ESRIN

Luca Brocca

National Research Council

data-descriptor

Keywords:

Posted Date: August 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10602968/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Irrigation water use (IWU) is a major component of anthropogenic water consumption in human-modified basins, yet it remains poorly monitored over large areas. Here, we present two satellite-derived IWU datasets for Europe, IWU-SMinv v1.0 and v1.1, generated using the Soil-Moisture-Inversion (SM-Inversion) approach with different input forcing datasets. The products provide IWU estimates at 1-km spatial resolution and 14-day temporal aggregation. IWU-SMinv v1.0 covers January 2016 to September 2025, whereas v1.1 covers 2016–2024. The datasets are evaluated against independent reference data with different spatial and temporal aggregations. For continuous multi-year records, monthly comparisons generally yield Root Mean Square Error ($RMSE$) and $BIAS$ values below 1 mm/day and Pearson correlation, r, above 0.6. For seasonally aggregated reference data, mean percentage errors (MPE) are within or close to ± 30% threshold in approximately 60% of the comparisons. Although uncertainties in the reference data affect the evaluation, these datasets provide spatially and temporally consistent estimates that can support hydrological, agricultural, and water-resources applications.

1. Background & Summary

In a global context where compound effects of population growth, global warming, water scarcity, and conflicts due to inequality in access to water resources feed into one another, irrigation plays a pivotal role[1–3]. Being it the predominant component among anthropic sectoral water uses[4,5], a detailed knowledge of irrigation water use (IWU) is essential to disclose human-water dynamics, whose awareness allows to build reliable water management tools serving decision makers and, more in general, the society[6–8]. Nevertheless, explicit, spatio-temporal distributed information on IWU is rarely available[9,10], thus limiting the possibility of properly solving water and energy balances over human-altered basins[11,12] .

In the last decade, novel avenues of research exploiting Earth Observation data to estimate IWU have become established; these approaches rely on satellite-derived land surface variables that are proxies of irrigation, such as soil moisture[13–16], evaporation[17,18] vegetation indices[19,20] and land surface temperature[21], among others. Considering methods relying on soil moisture and/or evaporation in particular, they fall in two main categories, i.e., approaches purely based on satellite retrievals[13,14,22] or on a combination of remote sensing data and modeling. The latter category involves data-assimilation frameworks[21,23–25] and baseline approaches, in which IWU is estimated as a difference between satellite retrievals and modeling outputs under the assumption that the former are affected by irrigation dynamics and the latter are not[26–29]. Among methods purely based on satellite retrievals, the Soil-Moisture-Inversion (SM-Inversion) approach has been constantly evolving in recent years. After its first formulation[13], the method has been improved and tested across several geographical areas and with different input data[10,14,30–32]. Recently, the SM-Inversion has been applied to create the first ever high-resolution (1km) regional datasets of IWU[15], showcasing the possibility of operational

implementations[33], as well as to produce long-term IWU datasets over four global hubs properly conceived to foster climate studies[29,34]. IWU estimates produced through the SM-Inversion have been fruitfully exploited in a range of applications, such as benchmarking irrigation schemes embedded into hydrological or land surface models[7,8,35], serving as training data for transferability through machine learning[36], evaluating the hydrological impact of irrigation practices[11], mapping soil salinity[32], and feeding digital twins of the terrestrial water cycle[6].

In line with the future plans outlined in Dari et al.[15], i.e., a gradual extension of satellite-based high-resolution IWU estimates from the regional to the continental scale, with a final ambitious goal of a planetary product, this study presents two alternative IWU datasets at 1 km spatial resolution and 14-day temporal aggregation over Europe developed through the SM-Inversion approach. The satellite-derived IWU datasets, IWU-SMinv v1.0 and IWU-SMinv v1.1, cover different time spans, January 2016-September 2025 and 2016–2024, respectively. In following sections, we first describe the methodology and its novel implementation architecture as well as the data structure. Then, a technical validation highlighting performances of the two datasets and their differences, based on available in-situ records, is provided.

2. Methods

2.1 SM-Inversion algorithm

The rationale behind the SM-Inversion approach is that the soil can be considered as a natural raingauge that collects water from both precipitation and irrigation. Thus, by inverting the soil water balance, the total amount of water entering the soil (in principle, rainfall plus irrigation) can be estimated and the irrigation contribution can be isolated by removing precipitation from the total. The soil water balance is written as follows:

$$Z^{*}dS\left(t\right)/dt\ =\ i\left(t\right)+p\left(t\right)-g\left(t\right)-sr\left(t\right)-e\left(t\right)$$

1

In which Z^{*} [mm] is the soil water capacity, calculated as the porosity, n [-], multiplied by the soil layer depth, Z [mm], $dS\left(t\right)/dt$ is the variation in time of relative surface soil moisture, $S\left(t\right)$ [-], $i\left(t\right)$ [mm/day] indicates irrigation, $p\left(t\right)$ [mm/day] precipitation, $g\left(t\right)$ [mm/day] drainage (i.e., subsurface runoff plus deep percolation), $sr\left(t\right)$ [mm/day] surface runoff, and $e\left(t\right)$ [mm/day] actual evaporation. $sr\left(t\right)$ is assumed to be negligible[37] and $g\left(t\right)$ and $e\left(t\right)$ are written according to Eqs. (2) and (3), respectively:

$$g\left(t\right)=aS(t)^{b}$$

2

$$e\left(t\right)\ =F\bullet\ S\left(t\right)e0\left(t\right)$$

3

a [mm] and b [-] are drainage parameters, F is an adjustment factor applied to the potential evaporation, $e0\,(t)$ [mm/day], limited by the available water content to represent the actual one. a, b, Z^*, and F are algorithm parameters that need to be calibrated. Eq. (1) can then be rewritten as:

$$Win\,(t) = \frac{Z^* dS\,(t)}{dt} + aS(t)^b + F \bullet S\,(t)\,e0\,(t)$$

4

With $Win\,(t)$ [mm/day] representing the total quantity of water entering the soil, $Win\,(t) = i\,(t) + p\,(t)$. After solving Eq. (4), irrigation can then be estimated as $i\,(t) = Win\,(t) - p\,(t)$. Negative irrigation rates, if any, are set equal to zero; in case of ratio between weekly irrigation to rainfall below 0.2, attributable to random errors, results are discarded[33].

The algorithm parameters a, b, Z^*, and F are calibrated by using precipitation as target variable; hence, calibration is carried out considering the potentially non-irrigation days, i.e., the entire non-irrigation season plus days within the irrigation season but with rainfall occurrence. During the periods above, precipitation constitutes the total amount of water entering the soil and algorithm parameters are optimized to minimize the error between rainfall estimated through the soil water balance inversion and reference data. The optimal parameters are then used to run the algorithm and estimate all the water entering the soil, thus enabling it to isolate irrigation after precipitation removal. Results are then post-processed by applying a spatio-temporal refinement[29], i.e., irrigation estimates are only delivered over areas where the percentage of area equipped for irrigation is higher than 5% according to the Global Map of Irrigation Areas (GMIA)[38,39] and during the main irrigation season, assumed for Europe as the period May-September[40]. Note that this information has been considered in the definition of the calibration period as well, thus ensuring no overlap. Figure 1 provides an overview of the SM-Inversion and its implementation workflow, as well as on the Irrigation Processor properly developed to operationalize the methodology (see Section 2.2 for further details).

This study presents two IWU datasets, namely IWU-SMinv v1.0 and IWU-SMinv v1.1, which differ based on the input data used. The algorithm requires soil moisture, precipitation, and potential evaporation to be implemented. Details on features of input datasets are provided in Section 2.2.

2.2 Input data

An overview of the input data used for the implementation of the SM-inversion algorithm is provided in Table 1. Both IWU-SMinv v1.0 and v1.1 rely on satellite soil moisture retrievals from Sentinel-1 observations; the operational product delivered by the Copernicus Land Monitoring Service is considered[41]. Concerning precipitation and potential evaporation inputs, IWU-SMinv v1.0 exploits data from ERA5-Land (European ReAnalysis v5 - Land)[42]. IWU-SMinv v1.1 still relies on precipitation from

ERA5-Land, but for potential evaporation GLEAM (Global Land Evaporation Amsterdam Model) v4.2b[43] is considered instead. It is noteworthy that, with respect to Fig. 1, potential evaporation from GLEAM v4.2b can be ascribed to the satellite observation category too, since this version relies on remote sensing as well.

Table 1
Features of input data used.

Input	Temporal resolution	Spatial resolution	Source
Sentinel-1 soil moisture	1.5-4 days	1 km	https://land.copernicus.eu/en/products/soil-moisture/daily-surface-soil-moisture-v1.0
ERA5-Land precipitation and potential evaporation	daily	9 km sampling	https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=download
GLEAM v4.2b potential evaporation	daily	0.1°	https://www.gleam.eu/#downloads

2.3 Operational Python processor

The datasets presented in this study have been generated using the Irrigation Processor, a scalable Python-based workflow engine developed to operationalize IWU estimation through the SM-Inversion approach. The processor translates the methodological workflow described in the previous section into a reproducible computational pipeline that can be applied consistently across different regions and time periods.

The processor is implemented as a custom declarative pipeline workflow engine in Python. Rather than executing a fixed sequence of scripts, the workflow engine organizes the workflow into modular processing steps with explicitly declared inputs, outputs, and dependencies. Based on these declarations, it automatically constructs a directed acyclic graph (DAG) and determines the order in which the steps must be executed. The overall workflow comprises five stages: the data loader, preprocessing, calibration, simulation, and postprocessing (see Fig. 1). Methodological details of the calibration, preprocessing and postprocessing are presented in the previous section, while here the focus is on the software architecture.

Each pipeline step consumes outputs generated by previous ones. Dependencies between steps are explicitly declared in the processor code. The workflow engine uses this information to ensure that all required inputs are available before a step is executed. Intermediate and final products are managed through a shared storage interface. Each output is assigned to a unique dataset identifier and written to a configurable data store. Before executing a step, the processor checks whether the expected outputs are already available. If they exist, the step can be skipped and the stored products reused. This mechanism supports resumable and incremental execution.

The processor is controlled through a YAML (YAML Ain't Markup Language) configuration file. At runtime, the configuration is validated using a Pydantic model and made available to all processing steps through a shared context object. The configuration defines the spatial and temporal domain, algorithm settings, storage locations, chunking strategies, and parallel-processing options. The workflow engine also supports parallel and distributed computation through Xarray (https://docs.xarray.dev/en/stable/getting-started-guide/index.html) and Dask (https://www.dask.org/) libraries. Large multidimensional datasets are divided into chunks that can be processed concurrently. The number of workers, threads, memory limits, and chunking schemes can be adjusted through the processor configuration. This enables the workflow to be adapted to the available computing environment, ranging from local execution to larger distributed infrastructures.

Data outputs are stored using data stores compatible with xcube (https://xcube.readthedocs.io/en/latest/), i.e., in Zarr format. Zarr provides chunked storage and supports parallel reading and writing, making it suitable for large multidimensional datasets. Depending on the deployment environment, the processor can use a local file system, S3-compatible object storage, or other storage backends supported by xcube.

The Irrigation Processor provides the reproducible software environment required to generate the IWU datasets presented in this study. By combining declarative workflow management, validated configuration, scalable computation, and standardized storage, it enables traceable and repeatable production of irrigation estimates across large spatial and temporal domains.

2.4 Reference data for validation

Technical validation has been carried out over pilot sites where it has been possible to collect observations to serve as a reference, generally not easy to obtain[5,15]. Such information differs in terms of its nature (e.g., measurement- or survey-based) and of spatio-temporal aggregation. A summary of locations of validation sites and of reference data features is provided in Fig. 2 and Table 2.

Table 2
Reference data features.

Country	Location	Irrigation site(s)	Temporal coverage	Data type
Spain	Ebro basin	Four irrigation districts: Algerri Balaguer (70.8 km^2), North (657.0 km^2) and South (504.5 km^2) Catalan Aragonese, Urgell (887.6 km^2)	2016–2025	Continuous time series of measured water volumes flowing through the irrigation canals feeding the districts
Italy	Po valley	Ninety-two districts with areas ranging from 52.8 km^2 to 1052.4 km^2 (extrapolated from SIGRIAN)	2016–2021	Seasonally cumulated values based on farmers self-declarations
		Two irrigation districts: San Silvestro (2.9 km^2) and Formellino (7.6 km^2)	2016–2020	Continuous time series of measured irrigation water applied
	Umbria region	Four irrigation districts: Marroggia (16.3 km^2), Topino (37.0 km^2), Clitunno (8.1 km^2), Sersimone Servino (38.9 km^2)	2016–2024 (Marroggia and Clitunno), 2017–2024 (Topino), 2021–2024 (Sersimone Servino)	Seasonally cumulated values declared by water allocation managers (i.e., irrigation consortia)
Germany	Niedersachsen region	Thirty-one upper fields with maximum area of 0.3 km^2	2018	Continuous time series of measured irrigation water applied
		One lower field with area equal to 0.4 km^2	2018	Continuous time series of measured irrigation water applied

The first validation area is the Ebro basin, Spain. Reference data consists of records of water volumes feeding four irrigation districts (Algerri Balaguer, North and South Catalan Aragonese, and Urgell) provided by the Sistema Automático de Información Hidrológica (SAIH) system and available at: https://www.saihebro.com/homepage/estado-cuenca-ebro. Time series of IWU are derived by dividing water volumes by the area of each district and conveyance losses are considered according to previous studies[14,15]. Further details on such data, as well as on features of irrigation practices and infrastructure over the pilot districts, can be found in previous studies from the authors[14,15,22,44].

Another validation area targeted is the Po valley, in Northern Italy. In this case, two different types of reference data have been collected. The first one consists of an extrapolation from the Sistema Informativo Nazionale per la Gestione delle Risorse Idriche in Agricoltura (SIGRIAN; https://sigrian.crea.gov.it/) database over ninety-two districts. Reference IWU is available as seasonally cumulated volumes that have been converted to mm by considering districts' areas. The SIGRIAN database is based on self-reported data and therefore contains some inherent uncertainties. For this reason, the database is more suitable for comparison in terms of large-scale average magnitudes, e.g., the whole Po basin scale, as recently shown by Kalimisetty et al.[8]. The second source of benchmark data for the Po basin refers to measurements of irrigation water applied over two districts (San Silvestro and Formellino) close to the Faenza city, in the Southern portion of the Po valley, where fruit trees represent the prevalent crop type. Data are provided by CER (Canale Emiliano Romagnolo).

Further irrigation data has been collected over Italy, more in detail over four agricultural districts located in the Umbria region, central Italy. Reference information consists of seasonal volumes declared by irrigation consortia, who manage water allocation within the districts: Consorzio della Bonificazione Umbra (for Marroggia, Clitunno, and Topino districts) and Consorzio di bonifica Tevere Nera (for the Sersimone-Servino district).

Finally, reference data have been collected over the Niedersachsen region, in Germany. IWU records consist of measurements carried out in 2018 over thirty-two fields, which have been analysed as two separated clusters given the nominal size and the fragmentation of the fields with respect to the 1 km resolution of satellite estimates. The first cluster is made of thirty-one fields located in the upper portion of the area, whose IWU records have been aggregated to be compared against remotely sensed amounts. The second cluster is made of a single field located in the lower portion of the domain. No information on the irrigation technique is available.

3. Data Record

Both IWU datasets, namely IWU-SMinv v1.0 and IWU-SMinv v1.1, are produced and delivered on the same 1 km grid as Sentinel-1 soil moisture; moreover, data is provided with a 14-day temporal aggregation. Because of the different source of input potential evaporation data (see Section 2.2), v1.0 covers the period from January 2016 to September 2025 (date of the product development), while v1.1 is available for the time span 2016–2024. It is noteworthy that v1.0 fully relies on operational data, so it is potentially extendable over time up to the present, with a 10-day input data latency[33].

The IWU datasets are distributed as four compressed Zarr v2 archives, representing two product versions and two alternative chunking layouts. The following archives are available:

- *iwu_spatial_v1.0.zarr.zip*
- *iwu_temporal_v1.0.zarr.zip*
- *iwu_spatial_v1.1.zarr.zip*

- *iwu_temporal_v1.1.zarr.zip*

The spatially and temporally chunked archives of each product version contain the same values and coordinates. They differ only in their internal chunk organization and are provided to support different data-access patterns. The spatial archives use chunks of one time step over large spatial regions and are intended for efficient mapping, regional extraction, and analysis of individual 14-day periods. The temporal archives contain the complete time series within small spatial chunks and are intended for efficient extraction and analysis of long time series over points or small areas.

IWU-SMinv v1.0 contains 254 time steps covering the period from January 2016 to September 2025. Its logical uncompressed array size is approximately 53.58 GiB. The compressed spatial and temporal archives are approximately 1.8 GB and 1.6 GB, respectively. IWU-SMinv v1.1 contains 234 time steps covering the period from January 2016 to December 2024. Its logical uncompressed array size is approximately 49.36 GiB. The compressed spatial and temporal archives are approximately 1.4 GB and 1.3 GB, respectively.

The values are provided at an exact temporal interval of 14 days. Each timestamp identifies the first date of the corresponding 14-day accumulation period. The variable *iwu_est*, with the long name "Irrigation Water Use Estimate," is expressed in millimetres per 14-day period. Each archive contains a three-dimensional float64 variable named *iwu_est*, organized using the dimensions time, latitude (lat), and longitude (lon). Missing or excluded cells are represented by NaN.

All four archives use the same regular latitude–longitude grid in the WGS 84 geographic coordinate reference system, corresponding to EPSG:4326. The grid contains 4,144 latitude cells and 6,832 longitude cells. Cell centres extend from approximately 72°N to 35°N and from 11°W to 50°E, at a coordinate spacing of approximately 0.01° in both dimensions. This corresponds to approximately 1.1 km in the north–south direction, while the east–west distance represented by 0.01° longitude decreases with increasing latitude.

The temporal archives use chunks of (all time steps, 50, 50). Each chunk is represented as a multidimensional array with dimensions (time, lat, lon). Consequently, each small spatial chunk contains the complete time series: (254, 50, 50) for v1.0 and (234, 50, 50) for v1.1. The spatial archives use chunks of (1, 2072, 1708), providing large spatial blocks for each individual time step.

4. Technical Validation

Figure 3 provides the average values of seasonally cumulated IWU rates referring to the time spans covered by the two datasets. According to both versions, IWU over Europe increases from West to East and from North to South, reflecting the increased presence of arid and semi-arid climate regimes[45]. In general terms, higher IWU rates are estimated according to v1.0 with respect to v1.1. Notwithstanding this, both datasets reproduce higher IWU values over well-known European irrigation hotspots, such as the Ebro basin and the Castilla-La Mancha region in Spain, the Po Valley in Italy, and the Kherson region

in Ukraine, as well as over regionally relevant irrigated areas, such as parts of the Lower Danube along the Bulgaria–Romania border, among others.

Performances of the two IWU datasets in reproducing the benchmark are evaluated in terms of Root Mean Square Error ($RMSE$), $BIAS$, and Pearson correlation, r, for those cases in which continuous time series over multi-year periods are available; when reference data is aggregated at the seasonal scale, mean percentage errors (MPE) are considered instead.

Results over four irrigation districts located in the Ebro basin, Spain, are provided in Fig. 4. According to IWU-SMinv v1.0, $RMSE$ across districts varies between 25.16 (Urgell) and 43.98 (North Catalan Aragonese) mm/month, while for IWU-SMinv v1.1 (South Catalan Aragonese) the range is between 21.31 and 34.36 (Algerri Balaguer) mm/month. Similarly, $BIAS$ values obtained for IWU-SMinv v1.0 are between 12.31 (Urgell) and 30.99 (North Catalan Aragonese) mm/month, for IWU-SMinv v1.1 they vary, in absolute terms, between 1.34 (South Catalan Aragonese) and 14.61 (Urgell) mm/month; r ranges between 0.35–0.77 and 0.32–0.73 according to IWU-SMinv v1.0 and IWU-SMinv v1.1, respectively; in both cases, the lowest values refers to the North Catalan Aragonese district, while the best performances are obtained for Urgell. Altogether, IWU-SMinv v1.1 outperforms IWU-SMinv v1.0, which shows a tendency to overestimate observed IWU rates. Notably, both datasets reflect the interannual variability, well capturing lower IWU recorded in the two-year period 2023–2024. This is particularly evident in Urgell, an area that was severely affected by drought in 2023. Following this, restrictions were imposed on the use of water for agricultural purposes[46,47]. Finally, IWU-SMinv v1.1 performances are in line with outcomes from previous implementations of the SM-Inversion algorithm over the same sites[15].

Concerning the Po valley, panel a) of Fig. 5 provides distributions of benchmark IWU from SIGRIAN against those estimated according to IWU-SMinv v1.0 and IWU-SMinv v1.1. It is immediately apparent that distributions of satellite-derived IWU rates have a comparable range of variations, while SIGRIAN data shows wider fluctuations. The extremes often reveal unrealistic IWU amounts, i.e., a minimum of 20 mm for a whole season or, on the other hand, more than 1 m of irrigation water applied. Nevertheless, values averaged over all the pilot districts bring the comparison to a certain level of significance, resulting in a MPE of 3.14% for IWU-SMinv v1.0. Underestimates quantified in a MPE of -30.76% are detected for IWU-SMinv v1.1 instead. The above considerations support what has been said regarding the uncertainties associated with the database (see Section 2.4), without prejudice to its validity as a reference in terms of large-scale average magnitudes.

Referring to the second source of benchmark data for the Po basin, i.e., irrigation water applied over two districts close to the Faenza city, results are shown in panel b) of Fig. 5. IWU-SMinv v1.1 outperforms IWU-SMinv v1.0 in terms of $RMSE$ (22.07 against 37.68 mm/month for San Silvestro and 42.39 versus 69.50 mm/month for Formellino) and $BIAS$ (30.54 and 61.92 mm/month for IWU-SMinv v1.0, 13.80 and 35.07 mm/month for IWU-SMinv v1.1). Conversely, r is slightly higher for IWU-SMinv v1.0 compared to IWU-SMinv v1.1, even though the performances according to this metric remain satisfactory (values higher than 0.66).

Results of validation carried out over four agricultural districts located in the Umbria region, central Italy, are provided in Fig. 6. As a reference, scatter plots show a ± 30% error band, as this value is indicated as the limit accuracy of satellite-derived irrigation information accepted by farmers according to the survey carried out by Massari et al.[48]. In three out of the four districts, IWU-SMinv v1.1 outperforms IWU-SMinv v1.0. However, the MPE remains higher than or close to the limit of ± 30%. Over the Marroggia district, MPE for IWU-SMinv v1.1 is 44.13%, highlighting a notable reduction from the value referring to IWU-SMinv v1.0, 87.10%. Over the Clitunno districts overestimates remain substantial, even though they decrease from more than 200% for IWU-SMinv v1.0 to around 170% for IWU-SMinvv1.1. IWU-Sat v1.0 overestimates at a rate higher than 200% are reported over the Sersimone Servino district as well, but MPE for IWU-SMinv v1.1 sharply decreases to less than 34% for v1.1 (with three out of the four representative points inside the ± 30% interval). Finally, over the Topino district, IWU-SMinv v1.0 (MPE less than 5%) outperforms IWU-SMinv v1.1 (MPE = -21.57%), even though both datasets provide satisfactory results. The obtained mixed performances can be partially explained by uncertainties in the reference data attributable to the complex irrigation infrastructure. In fact, irrigation water is drawn from different sources, e.g., directly from rivers or through small dams and delivered to the crops as sprinkler or flood irrigation, techniques which coexist in varying proportions across the districts. However, there is no explicit information on the spatial distribution of irrigation types, and reference volumes are provided aggregated at the entire district scale, making it impossible to properly attribute corrections accounting for conveyance and delivery losses[49]. On top of this, the area actually irrigated is often less than half the districts' extensions, thus highlighting a mosaic of fragmented irrigation that is not beneficial for irrigation detection[10,50,51]. Finally, the landscape of the Umbria region, due to its orographic features, has been demonstrated to be a challenging environment for isolating the irrigation signal in soil moisture retrievals, an issue that may concur in explaining overestimates reflected in IWU[50].

Performances over the test site in Germany, consisting of irrigated fields located in the Niedersachsen region, are provided in Fig. 7. Since only one year of data is available, the comparison is shown as scatter plots comparing monthly aggregated IWU. IWU-SMinv v1.0 overall provides reliable estimates, with a MPE equal to -32.30% (underestimation tendency) and to 24.41% (overestimation tendency) for the upper fields and the lower one, respectively. IWU-SMinv v1.1 produces not-negligible underestimates over the cluster of upper fields (MPE = -62.29%), while IWU dynamics over the lower field are well reproduced (MPE = -6.18%). In general, a challenge in properly reproducing mid-season IWU emerges, as corroborated by the recurring underestimates detected by both versions in July.

Altogether, the proposed datasets show solid performance, consistent with previous studies, especially for validation areas (e.g., Spain) where the scale at which irrigation occurs is compatible with the 1 km resolution and where uncertainties associated with reference data are lower. Overall, v1.1 usually performs better than v1.0, even though the latter has the important characteristic of being continuously extendable over time (10-day latency) for operational implementations. Users are encouraged to further test and validate the proposed datasets through available irrigation information. Collecting even more in-situ IWU records is essential to evaluate the level of maturity achieved by satellite estimates. Under this

perspective, it is noteworthy the initiative aimed at developing the first global IWU database (in-situ measurements) recently launched by the authors.

5. Usage Notes

To access the datasets, users can download the archives directly from Zenodo and use their preferred library (such as xarray or GDAL) to open them. A particularly convenient way to access them from Python and Jupyter notebooks is by using the xcube-zenodo plugin. This plugin adds a Zenodo data store to xcube and supports preloading compressed Zarr datasets from an individual Zenodo record. For example, a data cube can be opened as follows:

from xcube.core.store import new_data_store

store = new_data_store("zenodo", root="21473561")

preload_store = store.preload_data("iwu_temporal_v1.1.zarr.zip")

preloaded_data_ids = preload_store.list_data_ids()

cube = preload_store.open_data(preloaded_data_ids[0])

print(cube)

The spatially chunked archive should normally be selected for maps, large-area subsets, or analysis of individual time steps. The temporally chunked archive is better suited to extracting complete time series for points or small spatial regions.

As previously explained, the datasets are provided with spatio-temporal masks matching with areas equipped for irrigation and with the highest-intensity irrigation season. However, users are encouraged to further refine the IWU estimates with more detailed ancillary information (if available) for validation and use over specific sites.

Declarations

Data Availability

IWU-SMinv v1.0 and v1.1 are available at https://zenodo.org/records/21473561.

7. Code Availability

The Irrigation Processor used to develop the datasets is available, in Python, at: https://github.com/Irrigation-EU/irrigation-processor. It can be executed through a command-line interface. A standard run using the default configuration is initiated with:

pixi run irr-proc run

An alternative configuration file can be supplied when required:

pixi run irr-proc run --config my_config.yml

During execution, the processor validates the configuration, constructs the dependency graph, initializes the configured computing environment, identifies the steps that need to be run, and stores the resulting products in the selected data store. The dependency graph can also be exported for inspection using Graphviz (https://graphviz.org/), providing a visual representation of the workflow and its data dependencies.

Acknowledgements

The authors acknowledge the support from the European Space Agency (ESA) under the Irrigation-EU project (CCN to ESA Contract 4000129870/20/I-NB, Irrigation+). The authors wish to thank the SAIH (Sistema Automático de Información Hidrológica) Ebro for providing the reference irrigation data over the Ebro basin, CER (Canale Emiliano Romagnolo) for sharing data referring to the San Silvestro and Formallino districts in the Po valley.)

The authors wish to thank AdBPo (Autorità di Distretto del Bacino del Fiume Po) for providing agricultural water use data over the Po basin, based on the Sistema Informativo Nazionale per la Gestione delle Risorse Idriche in Agricoltura (SIGRIAN).

Support from AURI (Autorità Umbra Rifiuti e Idrico) under the Interreg Europe Gov4Water project is acknowledged; in particular, the authors would like to thank Consorzio della Bonificazione Umbra and Consorzio di bonifica Tevere Nera for providing irrigation volumes used for validation. Reference irrigation data over the Niedersachsen site is obtained under the ESA Irrigation+ project, which is acknowledged.

Competing Interests

The authors declare no competing interests.

Author Contributions

JD and LB led the scientific development of the IWU-SMinv v1.0 and v1.1 datasets. YKBS, KN, NF, and GB handled technical aspects, such as the optimization of the SM-Inversion approach and the development of the Irrigation Processor. JD led the technical validation and wrote the first draft. JD and FL prepared the figures. All the authors contributed to the discussion and interpretation of the results, as well as editing and revising the manuscript.

References

1. Jiang, R. *et al.* Transboundary conflict from surface water scarcity under climate change. *Nat Commun* 16, 8166 (2025).

2. Rosa, L. & Sangiorgio, M. Global water gaps under future warming levels. *Nat Commun* 16, 1192 (2025).
3. He, L. & Rosa, L. Sustainable global irrigation expansion could support only two-thirds of croplands under 1.5°C and 3°C warming. *Nat Food* 7, 369–380 (2026).
4. Abbott, B. W. *et al.* Human domination of the global water cycle absent from depictions and perceptions. *Nat. Geosci.* 12, 533–540 (2019).
5. McDermid, S. *et al.* Irrigation in the Earth system. *Nat Rev Earth Environ* 4, 435–453 (2023).
6. Brocca, L. *et al.* A Digital Twin of the terrestrial water cycle: a glimpse into the future through high-resolution Earth observations. *Front Sci* 1, 1190191 (2024).
7. Camici, S. *et al.* High-resolution satellite observations for developing advanced decision support systems for water resources management in the Po River. *Journal of Hydrology* 662, 134047 (2025).
8. Kalimisetty, S. *et al.* Modelling the human impact on hydrological cycle: The MISDc v3.1 model. *Environmental Modelling & Software* 204, 107056 (2026).
9. Dorigo, W. *et al.* Closing the Water Cycle from Observations across Scales: Where Do We Stand? *Bulletin of the American Meteorological Society* 102, E1897–E1935 (2021).
10. Dari, J., Lo Presti, S. & Brocca, L. Irrigation Monitoring from Satellite at Hyper-High Resolution: Paving the Way for Remote-Sensing-Based Agricultural Water Management Support Services. Preprint at https://doi.org/10.2139/ssrn.5170542 (2025).
11. Dari, J. *et al.* Quantifying the Hydrological Impacts of Irrigation on a Mediterranean Agricultural Context Through Explicit Satellite-Derived Irrigation Estimates. *Water Resources Research* 60, e2023WR036510 (2024).
12. Lunel, T., Boone, A. A. & Le Moigne, P. Irrigation strongly influences near-surface conditions and induces breeze circulation: Observational and model-based evidence. *Quart J Royal Meteoro Soc* 150, 2798–2819 (2024).
13. Brocca, L. *et al.* How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. *International Journal of Applied Earth Observation and Geoinformation* 73, 752–766 (2018).
14. Dari, J. *et al.* Exploiting High-Resolution Remote Sensing Soil Moisture to Estimate Irrigation Water Amounts over a Mediterranean Region. *Remote Sensing* 12, 2593 (2020).
15. Dari, J. *et al.* Regional data sets of high-resolution (1 and 6 km) irrigation estimates from space. *Earth Syst. Sci. Data* 15, 1555–1575 (2023).
16. Zappa, L. *et al.* Detection and Quantification of Irrigation Water Amounts at 500 m Using Sentinel-1 Surface Soil Moisture. *Remote Sensing* 13, 1727 (2021).
17. Brombacher, J., Silva, I. R. D. O., Degen, J. & Pelgrum, H. A novel evapotranspiration based irrigation quantification method using the hydrological similar pixels algorithm. *Agricultural Water Management* 267, 107602 (2022).

18. Zipper, S. *et al.* Estimating irrigation water use from remotely sensed evapotranspiration data: Accuracy and uncertainties at field, water right, and regional scales. *Agricultural Water Management* 303, 109036 (2024).
19. Maselli, F. *et al.* Use of Sentinel-2 MSI data to monitor crop irrigation in Mediterranean areas. *International Journal of Applied Earth Observation and Geoinformation* 93, 102216 (2020).
20. Chiesi, M. *et al.* Monitoring and analysis of crop irrigation dynamics in Central Italy through the use of MODIS NDVI data. *European Journal of Remote Sensing* 55, 23–36 (2022).
21. Corbari, C. *et al.* Estimates of Irrigation Water Volume by Assimilation of Satellite Land Surface Temperature or Soil Moisture Into a Water-Energy Balance Model in Morocco. *Water Resources Research* 61, e2024WR038926 (2025).
22. Zappa, L. *et al.* Benefits and pitfalls of irrigation timing and water amounts derived from satellite soil moisture. *Agricultural Water Management* 295, 108773 (2024).
23. Modanesi, S. *et al.* Challenges and benefits of quantifying irrigation through the assimilation of Sentinel-1 backscatter observations into Noah-MP. *Hydrol. Earth Syst. Sci.* 26, 4685–4706 (2022).
24. Jalilvand, E. *et al.* Is It Possible to Quantify Irrigation Water-Use by Assimilating a High-Resolution Satellite Soil Moisture Product? *Water Resources Research* 59, e2022WR033342 (2023).
25. Busschaert, L. *et al.* Irrigation Quantification Through Backscatter Data Assimilation With a Buddy Check Approach. *J Adv Model Earth Syst* 16, e2023MS003661 (2024).
26. Zaussinger, F. *et al.* Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data. *Hydrol. Earth Syst. Sci.* 23, 897–923 (2019).
27. Koch, J., Zhang, W., Martinsen, G., He, X. & Stisen, S. Estimating Net Irrigation Across the North China Plain Through Dual Modeling of Evapotranspiration. *Water Resources Research* 56, e2020WR027413 (2020).
28. Kragh, S. J. *et al.* An inter-comparison of approaches and frameworks to quantify irrigation from satellite data. *Hydrol. Earth Syst. Sci.* 28, 441–457 (2024).
29. Laluet, P. *et al.* Long-term irrigation water use datasets from multiple Earth Observation-based methods in major irrigated regions. *Earth Syst. Sci. Data* 18, 4833–4853 (2026).
30. Dari, J. *et al.* Irrigation estimates from space: Implementation of different approaches to model the evapotranspiration contribution within a soil-moisture-based inversion algorithm. *Agricultural Water Management* 265, 107537 (2022).
31. Almendra-Martín, L., Judge, J., Monsivaís-Huertero, A. & Liu, P.-W. Evaluating Uncertainties in an SM-Based Inversion Algorithm for Irrigation Estimation in a Subtropical Humid Climate. *Water* 16, 2445 (2024).
32. Matendo, S. E., Dari, J., Flammini, A., Brocca, L. & Sánchez-Calvo, R. AquaLeach approach: Field-scale irrigation assessment using satellite-derived products while accounting for salinity leaching fraction in arid and semi-arid irrigation schemes of Kenya. *Agricultural Water Management* 328, 110290 (2026).

33. Dari, J., Filippucci, P. & Brocca, L. The development of an operational system for estimating irrigation water use reveals socio-political dynamics in Ukraine. *Hydrol. Earth Syst. Sci.* 28, 2651–2659 (2024).
34. Galli, N. *et al.* Integrating biophysical models and remote sensing to evaluate irrigation practices in four global hubs. *Agricultural Water Management* 327, 110284 (2026).
35. De Lannoy, G. J. M. *et al.* Contributions of Irrigation Modeling, Soil Moisture and Snow Data Assimilation to High-Resolution Water Budget Estimates Over the Po Basin: Progress Towards Digital Replicas. *J Adv Model Earth Syst* 16, e2024MS004433 (2024).
36. Kragh, S. J., Schneider, R., Fensholt, R., Stisen, S. & Koch, J. Synthesizing regional irrigation data using machine learning – Towards global upscaling via metamodeling. *Agricultural Water Management* 311, 109404 (2025).
37. Brocca, L. *et al.* Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of the SM2RAIN algorithm. *Journal of Hydrology and Hydromechanics* 63, 201–209 (2015).
38. Siebert, S. *et al.* A global data set of the extent of irrigated land from 1900 to 2005. *Hydrol. Earth Syst. Sci.* 19, 1521–1545 (2015).
39. Mehta, P. *et al.* Global Area Equipped for Irrigation Dataset 1900–2015. Zenodo https://doi.org/10.5281/ZENODO.6886564 (2022).
40. Portmann, F., Siebert, S., Bauer, C., & Döll, P. *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 (2008).
41. Bauer-Marschallinger, B. *et al.* Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles. *IEEE Trans. Geosci. Remote Sensing* 57, 520–539 (2019).
42. Muñoz-Sabater, J. *et al.* ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* 13, 4349–4383 (2021).
43. Miralles, D. G. *et al.* GLEAM4: global land evaporation and soil moisture dataset at 0.1° resolution from 1980 to near present. *Sci Data* 12, 416 (2025).
44. Dari, J. *et al.* Detecting and mapping irrigated areas in a Mediterranean environment by using remote sensing soil moisture and a land surface model. *Journal of Hydrology* 596, 126129 (2021).
45. Beck, H. E. *et al.* Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci Data* 5, 180214 (2018).
46. Paolini, G., Escorihuela, M. J., Bellvert, J., Merlin, O. & Pellarin, T. PrISM at Operational Scale: Monitoring Irrigation District Water Use during Droughts. *Remote Sensing* 16, 1116 (2024).
47. Bellvert, J., Pamies-Sans, M., Casadesús, J. & Girona, J. Evaluating the impact of drought and water restrictions on agricultural production in irrigated areas through crop water productivity functions and a remote sensing-based evapotranspiration model. *Agricultural Water Management* 309, 109319 (2025).

48. Massari, C. *et al.* A Review of Irrigation Information Retrievals from Space and Their Utility for Users. *Remote Sensing* 13, 4112 (2021).
49. Flammini, A. *et al.* Assessment of Regional-Scale Freshwater Availability Towards Sustainable Management in the Context of Climate Change. *Water* 18, 1290 (2026).
50. Dari, J. *et al.* Double-scale analysis on the detectability of irrigation signals from remote sensing soil moisture over an area with complex topography in central Italy. *Advances in Water Resources* 161, 104130 (2022).
51. Zappa, L. *et al.* How accurately can we retrieve irrigation timing and water amounts from (satellite) soil moisture? *International Journal of Applied Earth Observation and Geoinformation* 113, 102979 (2022).

Figures

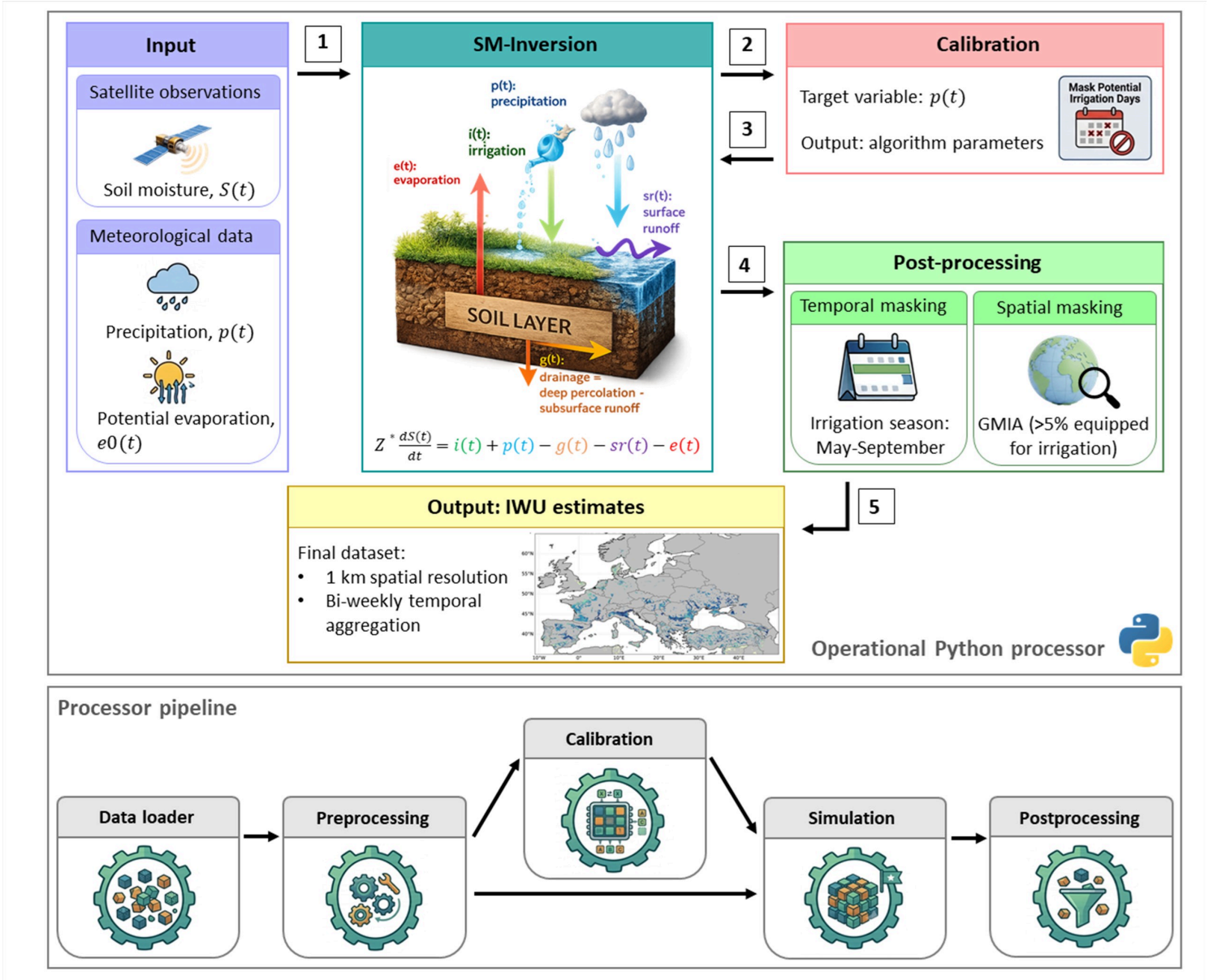


Figure 1

Overview on the SM-Inversion, its implementation workflow, and pipeline of the operational Python processor. Graphics partially created with AI assistance

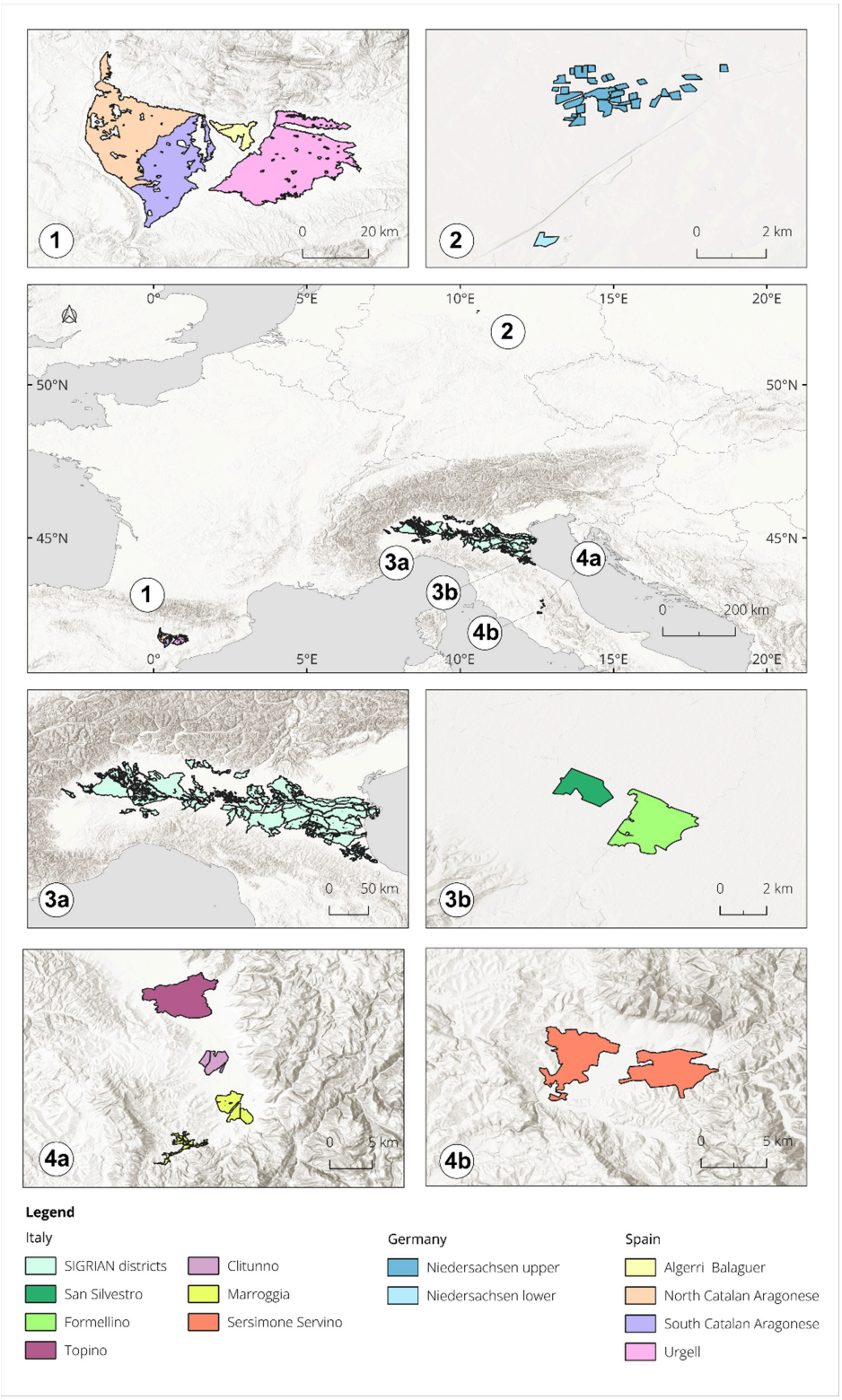


Figure 2

Location of the validation sites. See Table 2 for further details.

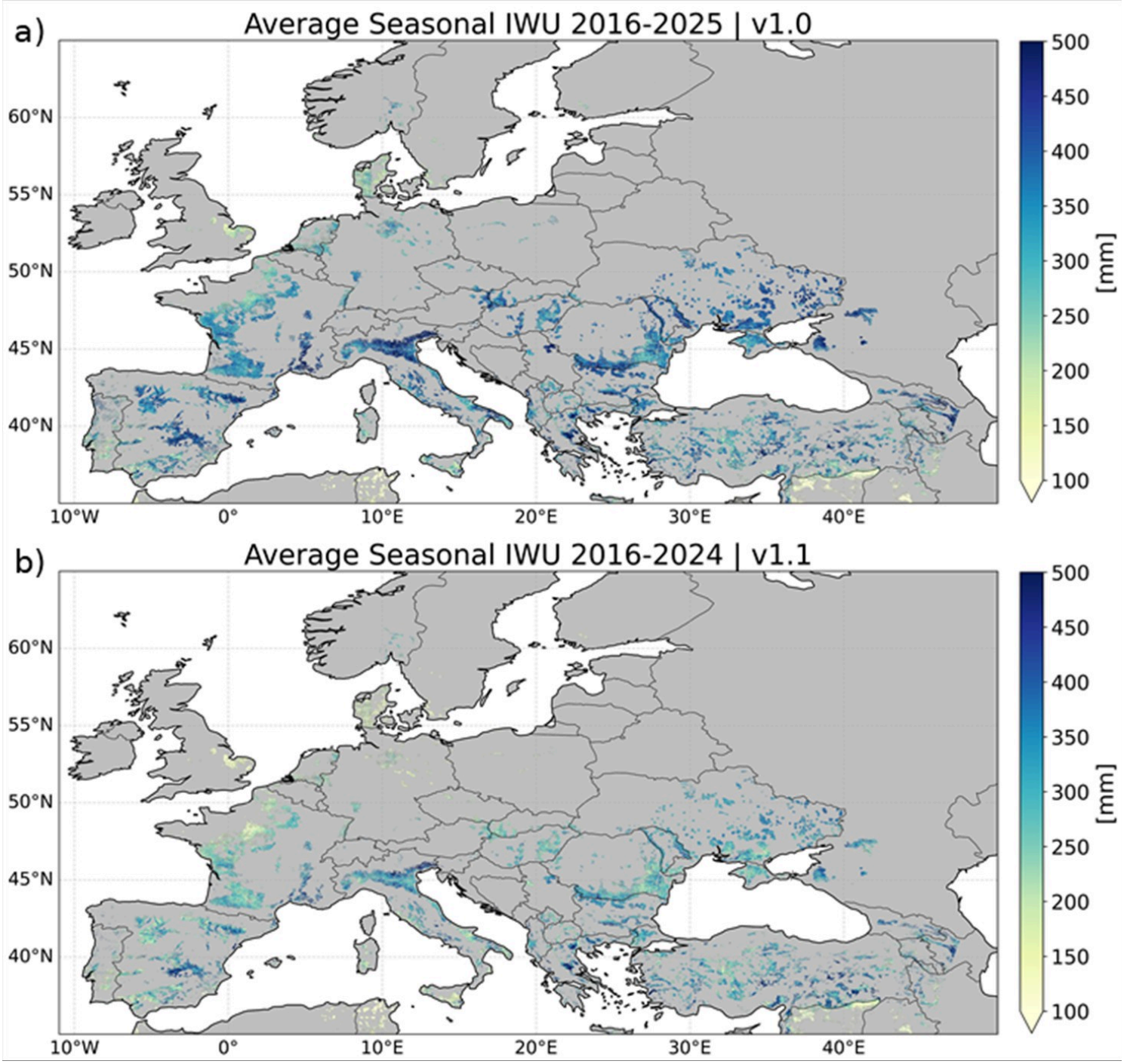


Figure 3

Panel a): average values of seasonally cumulated Irrigation Water Use (IWU) according to IWU-SMinv v1.0 during 2016-2025. Panel b): same as panel a) but referring to IWU-SMinv v1.1 and time span 2016-2024.

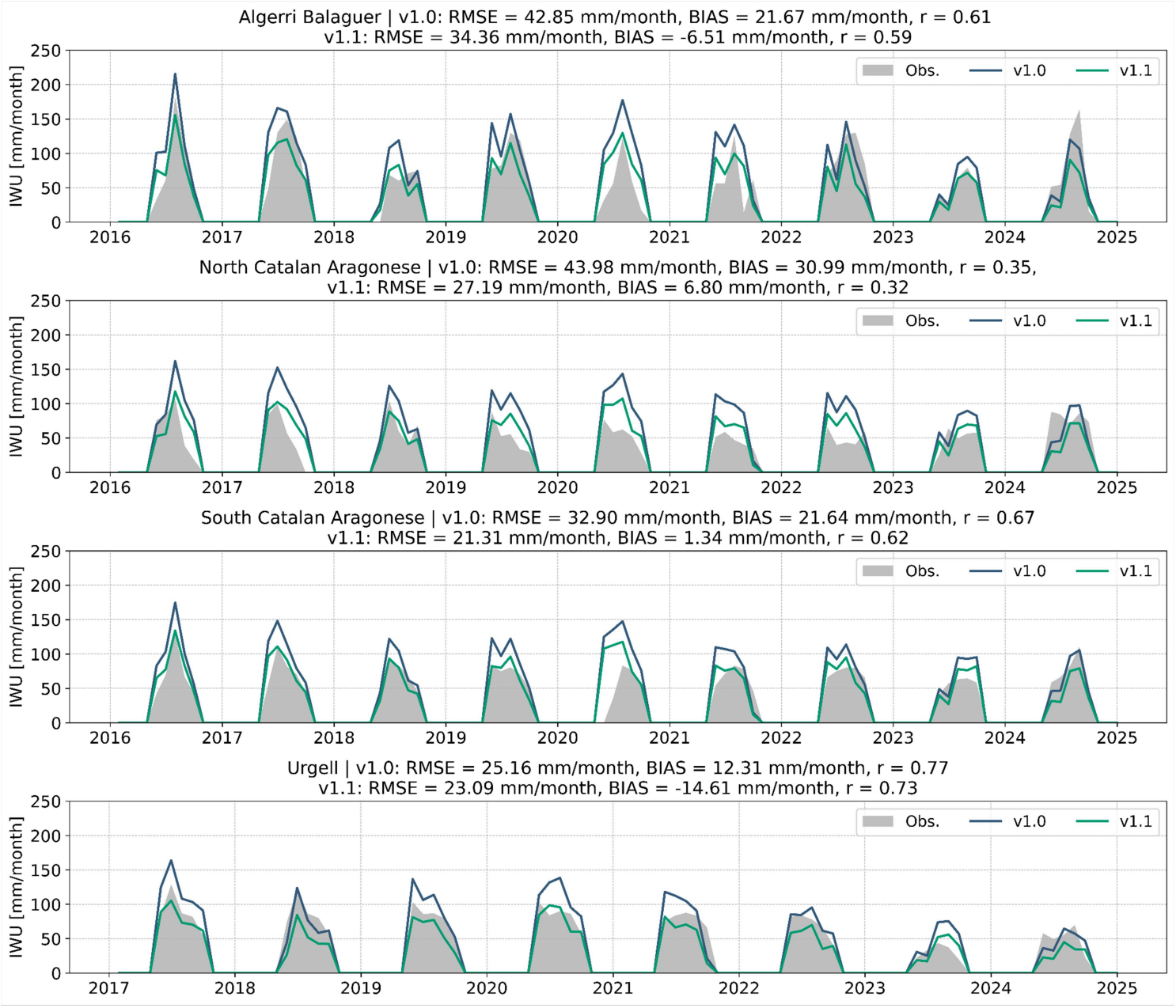


Figure 4

Monthly aggregated time series of IWU (Irrigation Water Use) estimates according to IWU-SMinv v1.0 (blue) and IWU-SMinv v1.1 (light green) against observed rates (light gray shaded area) over the four validation districts in the Ebro basin, Spain: Algerri Balaguer, North Catalan Aragonese, South Catalan Aragonese, and Urgell.

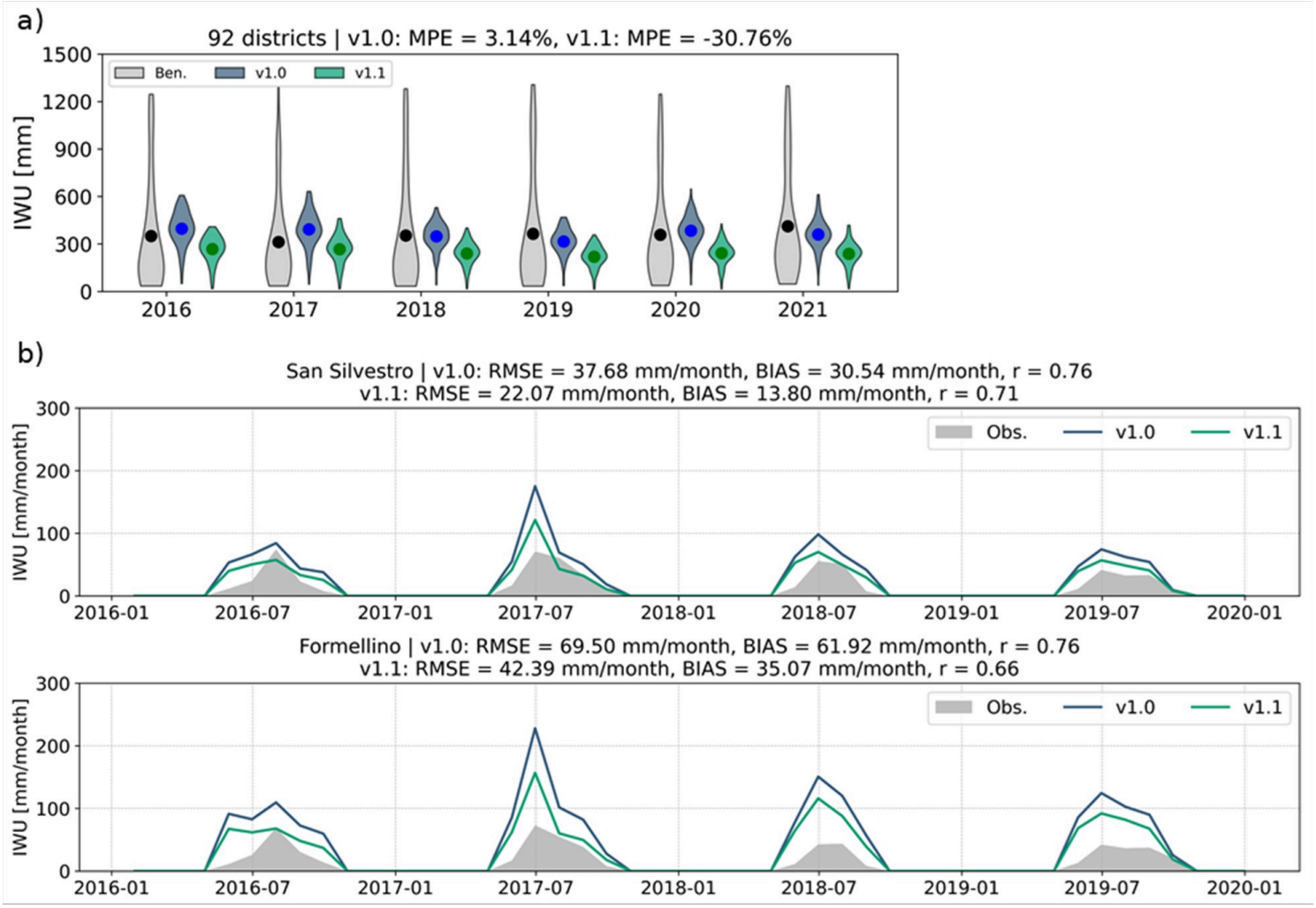


Figure 5

Panel a): comparison between distributions of seasonal IWU (Irrigation Water Use) covered by the SIGRIAN (Sistema Informativo Nazionale per la Gestione delle Risorse Idriche in Agricoltura) database (light gray) and estimates according to IWU-SMinv v1.0 (blue) and IWU-SMinv v1.1 (light green). Panel b): monthly aggregated time series of IWU estimates according to IWU-SMinv v1.0 (blue) and IWU-SMinv v1.1 (light green) against observed rates (light gray shaded area) over the San Silvestro and Formellino districts.

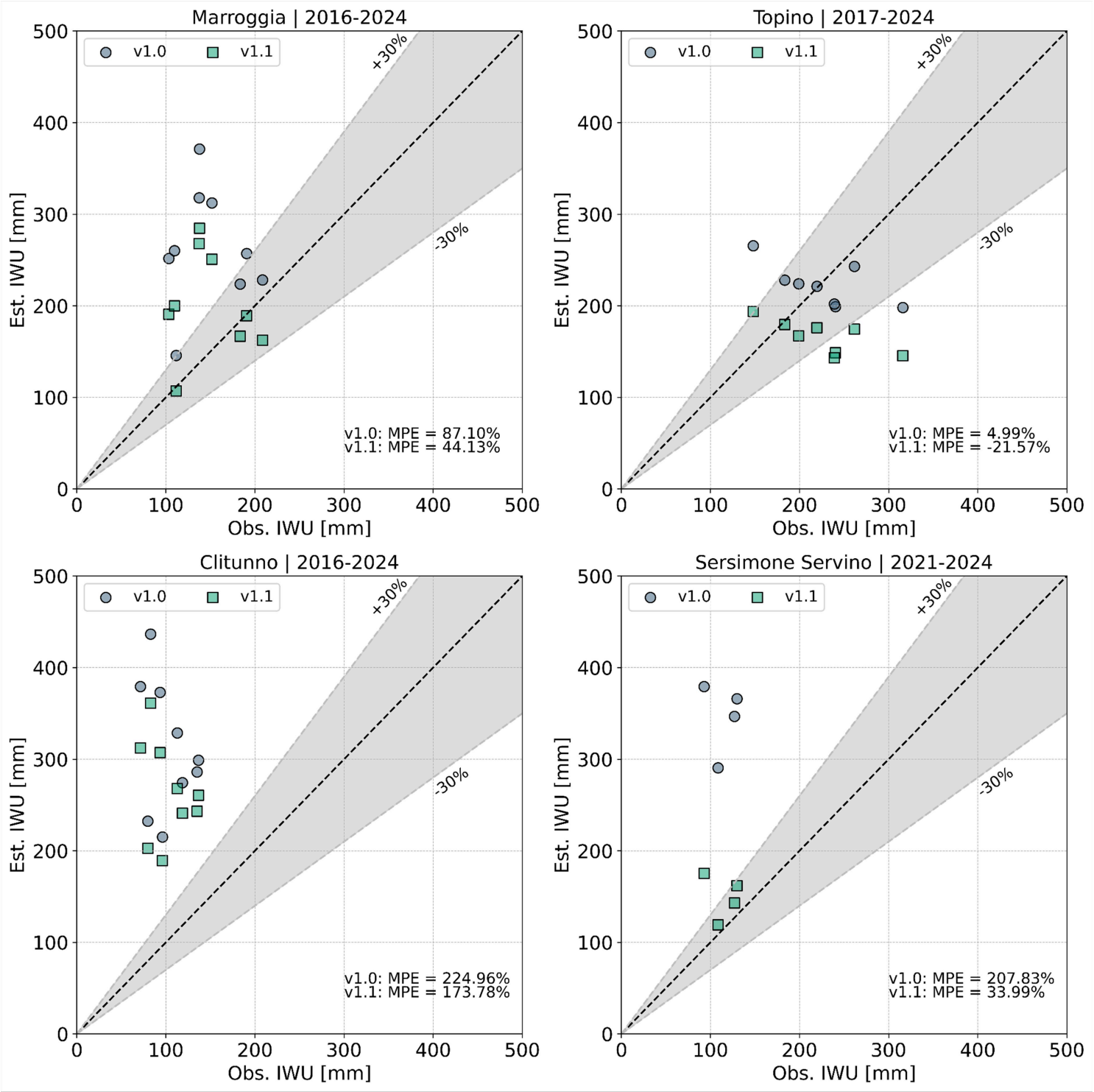


Figure 6

Comparison of observed seasonal IWU (Irrigation Water Use) against estimates according to IWU-SMinv v1.0 (blue) and IWU-SMinv v1.1 (light green) over the four validation districts in the Umbria region, Italy: Marroggia, Topino, Clitunno, and Sersimone Servino. Each point of the scatter plots represents a year. The light grey shaded area indicates a ± 30% interval.

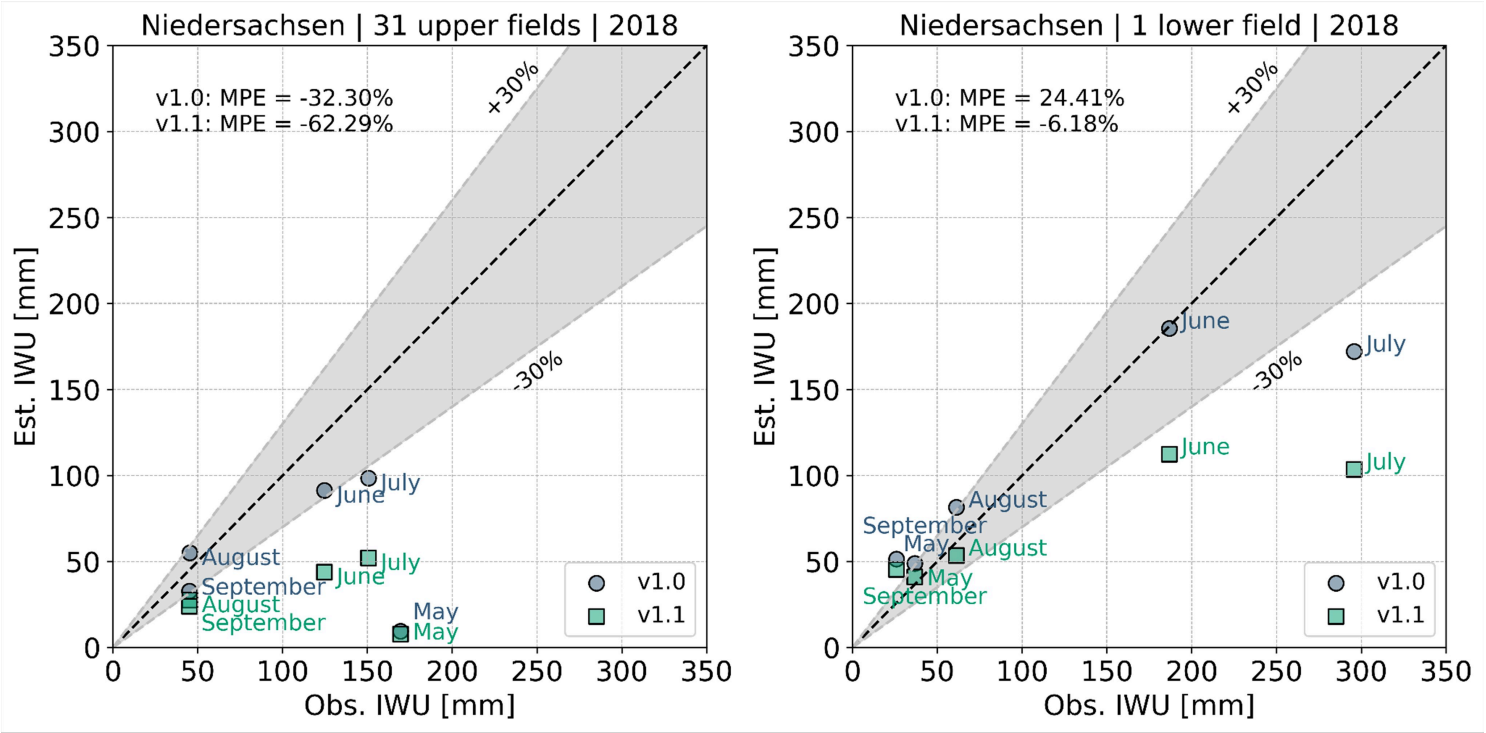


Figure 7

Comparison of observed seasonal IWU (Irrigation Water Use) against estimates according to IWU-SMinv v1.0 (blue) and IWU-SMinv v1.1 (light green) over the two clusters of pilot fields in the Niedersachsen region, Germany, in 2018. The light grey shaded area indicates a ± 30% interval.