Journal of Hydrology 638 (2024) 131581

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Improved soil moisture estimation and detection of irrigation signal by incorporating SMAP soil moisture into the Indian Land Data Assimilation System (ILDAS)

Arijit Chakraborty [a], Manabendra Saharia [a,b,*], Sumedha Chakma [a], Dharmendra Kumar Pandey [c], Kondapalli Niranjan Kumar [d], Praveen K. Thakur [e], Sujay Kumar [f], Augusto Getirana [f,g]

[a] *Department of Civil Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India*
[b] *Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India*
[c] *Space Applications Center, Indian Space Research Organization, Ahmedabad, Gujarat 380015, India*
[d] *National Centre for Medium Range Weather Forecasting, Ministry of Earth Sciences, Noida 201301, India*
[e] *Indian Institute of Remote Sensing, Indian Space Research Organization, Dehradun 248001, India*
[f] *Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA*
[g] *Science Applications International Corporation, Greenbelt, MD, USA*

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief

Keywords:
Soil moisture
SMAP
Data assimilation
ILDAS
EnKF
Irrigation signal

ABSTRACT

Land surface models have facilitated the estimation of soil moisture over a range of spatiotemporal scales. However, limitations in model parameterization and under-representation of anthropogenic processes restrict their ability to estimate local-scale soil moisture variability, especially over irrigated areas. Assimilation of satellite-based soil moisture retrievals into land surface models can be a viable approach to overcome these constraints, specially over highly irrigated countries such as India, where such applications are rare. Additionally, large-scale validation of modeled soil moisture has been limited over India till now due to lack of a representative station network. By assimilating Soil Moisture Active Passive (SMAP)-based estimates into the state-of-the-art Indian Land Data Assimilation System (ILDAS) and combining with a new soil moisture station network of more than 200 stations, this study demonstrates improved soil moisture estimations and capture of irrigation signals over the region. The Noah-MP land surface model is forced by multiple local and global meteorological datasets and Ensemble Kalman Filter (EnKF) is used for assimilation of soil moisture. Comparison of open-loop and data assimilated soil moisture against station soil moisture data shows relative spatial mean improvement of 0.0178 in correlation and 0.0029 m^3/m^3 in RMSE. Further statistical comparison with in-situ data has also shown better results over most of the stations, as evident from improved correlations and reduced unbiased RMSE after assimilation. Finally, the climatology of soil moisture over the different irrigation fractions reveals that data assimilated outputs over irrigated grid cells tend to have higher soil moisture during dry winter season, demonstrating the ability to capture irrigation signals. These findings quantify the value of data assimilation in improving soil moisture estimates and the ability to capture unmodeled processes such as irrigation, which lays the science groundwork for upcoming space missions such as NASA ISRO Synthetic Aperture Radar (NISAR).

1. Introduction

Soil moisture plays a significant role in land surface processes as it influences various aspects of the hydrologic cycle such as water availability and distribution within terrestrial ecosystems (Corradini, 2014; Liu et al., 2020). Soil moisture significantly impacts water balance by segregating rainfall into evapotranspiration, infiltration, and runoff (Botter et al., 2007; Koster et al., 2004; Lee et al., 2007). Moreover, it is also having influence in maintaining energy balance by partitioning incident energy into sensible and latent heat flux (Dorigo et al., 2017;

* Corresponding author at: Indian Institute of Technology Delhi, New Delhi 110016, India.
E-mail address: msaharia@iitd.ac.in (M. Saharia).

https://doi.org/10.1016/j.jhydrol.2024.131581
Received 12 April 2024; Received in revised form 29 May 2024; Accepted 9 June 2024
Available online 24 June 2024

Entekhabi et al., 2010). Therefore, accurate assessment of the spatio-temporal variability of soil moisture is crucial for several hydrometeorological studies related to drought monitoring (Anderson et al., 2012; Fang et al., 2021; Hong & Kalnay, 2000; Jung et al., 2020), flood forecasting (Komma et al., 2008; Wanders et al., 2014), climate studies (Seneviratne et al., 2010; Yuan et al., 2011), agriculture and irrigation (Dobriyal et al., 2012; Jalilvand et al., 2019). By integrating satellite and in situ datasets, a Land Data Assimilation System (LDAS) improves the accuracy of Land Surface Models (LSMs) through the utilization of data assimilation methods (Kumar et al., 2014; Liu et al., 2013; Mitchell et al., 2004; Rodell et al., 2004). Such systems are however rare over the Indian subcontinent, and the Indian Land Data Assimilation System (ILDAS) has been recently developed system that aims to provide estimates of various land surface conditions, water storage and fluxes over the Indian mainland at high resolution (Magotra et al., 2024). Till now, a robust large-scale assessment of open loop and data assimilated soil moisture with respect to a station soil moisture network and its ability to represent unmodeled processes such as irrigation has been missing over the Indian subcontinent.

In-situ soil moisture measurements are the accurate representations of the actual ground condition but lacks in availability and spatial extent at larger domain (Ahmad et al., 2022; Yin et al., 2023). With the development of satellite remote sensing techniques, monitoring the spatiotemporal variation of soil moisture has become easier for hydrological studies (Ahmad et al., 2010; Corradini, 2014; Wang & Qu, 2009). Satellite soil moisture observations have the potential to capture the effects of processes that are independent of precipitation or natural variability (Brocca et al., 2018; Lawston et al., 2017). Irrigation is one such widespread anthropogenic activity which directly influences the moisture levels in agricultural areas (Ozdogan et al., 2010; Zhang et al., 2017). But these satellite retrievals often shows limitations in terms of accuracy in algorithm and other specifications (Kim et al., 2020; Loew et al., 2013).

Land Surface Models (LSM) have been actively utilized in recent years to model complicated land surface phenomena and estimate various land surface state variables including soil moisture (Dai et al., 2003; Ek et al., 2003; Koster et al., 2000; Niu et al., 2011). Typically, LSMs simulate the exchange of energy and water at the earth's surface by utilizing meteorological boundary conditions obtained from observations and underlying land surface conditions (Abramowitz et al., 2008; Niu et al., 2011). Even though LSMs have the flexibility to estimate surface state variables at any high resolution, they are susceptible to uncertainties in input forcings and inadequate parameterization of model physics (Maggioni et al., 2012; Shrestha et al., 2020). Moreover, it is challenging to accurately capture and represent irrigation in models because of lack in proper parameterization over regional scale and scarcity of in situ irrigation data (Gibson et al., 2017; Kwon et al., 2022; Lawston et al., 2017). To aid regional and global scale research, efforts have been made to improve model estimates by incorporating satellite remote sensing data products by means of data assimilation (DA) techniques (Kumar et al., 2014; Reichle et al., 2002; Reichle & Koster, 2005). Data assimilation enables the integration of observational data into LSMs to update and improve their state variables, taking account of their respective errors and uncertainties (Kalman, 1960). Observations often contain spatial and temporal discontinuity, and hence their incorporation into the LSM estimates helps to evaluate the spatiotemporal variability of different land surface parameters (Kumar et al., 2012). There are several studies that involved assimilation of satellite soil moisture data retrieved from various platforms for estimating improvements in different water balance components (Huang et al., 2008; Lievens et al., 2015; Reichle et al., 2007; Seo et al., 2021). But most of the existing soil moisture data assimilation research have not explored the suitability of data assimilation in detecting unmodeled irrigation process (Kumar et al., 2015). Large-scale irrigation can alter the soil moisture level over that irrigated area. So far, it has been tough to get reliable and consistent information about irrigation practices. However, data regarding the precise timing and quantification of irrigation practices are sparse in several places globally. Simulation of regional irrigation schemes is challenging because of the inaccessibility of relevant water source distribution and usage information. With more than 40 % of the total cropland being irrigated, India has one of the world's largest irrigated agricultural systems (Meier et al., 2018). But such national-scale land data assimilation systems with respect to soil moisture have not been developed for India. Moreover, scarcity of in situ soil moisture measurements and irrigation information across India remained a limiting factor for such studies. To address this research gap and assess the effectiveness of data assimilation in this region an integrated Indian Land Data Assimilation System (ILDAS; Magotra et al., 2024) has been setup to study the surface soil moisture variability for the Indian domain.

This paper quantifies the improvements in spatiotemporal variability in soil moisture estimates and their ability to capture irrigation signal by performing assimilation of Soil Moisture Active Passive (SMAP; Entekhabi et al., 2010) data into the Indian Land Data Assimilation System (ILDAS; Magotra et al., 2024) for the Indian mainland. The reason behind choosing SMAP soil moisture product for assimilation is that SMAP gives the most effective satellite soil moisture data for a large portion of land areas worldwide (Chen et al., 2018). Moreover, it has been observed that SMAP can detect the impact of irrigation on soil moisture (Lawston et al., 2017). The resulting improvements in soil moisture and impact over irrigated areas are evaluated in comparison with in-situ soil moisture observations. This paper is organized in the following manner: section 2 gives the description of the study domain, followed by data sources used in the model (meteorological forcings, satellite observations and evaluation data) in section 3 and methodology (Model configuration, data assimilation algorithm, experimental set up and the evaluation metrics) in section 4. Results are evaluated and validated in section 5 followed by discussion (section 6) and conclusions (section 7).

2. Study area

The present study has been conducted mainly over the Indian subcontinent. The run domain is extended from latitude of 5.5° N to 37.5° N and longitude of 68° E to 98° E with domain resolution of 0.1° × 0.1°. The domain comprises of 20 major river basins (as provided by the Central Water Commission of India) with a large variation in elevation (Fig. 1), topography, land cover and climatic conditions, which leads to significant variation in rainfall patterns and soil moisture as well. Moreover, agriculture is vital to India's economy, making a considerable contribution to its GDP and supporting a significant portion of the population's livelihoods.

(Mathur et al., 2006). Agricultural productivity and irrigation requirements are dependent on precipitation (Hargreaves et al., 1985; Zeng et al., 2021) and soil moisture availability (Dobriyal et al., 2012; Tao et al., 2003). Hence, it has become necessary to estimate both the spatial and temporal variation in soil moisture over this domain.

3. Data sources

3.1. Forcings and assimilation datasets

3.1.1. India meteorological Department (IMD) precipitation data

In this study we have used gridded daily precipitation data from the India Meteorological Department (IMD) with a 0.25° spatial resolution for the years 1981–2022. This gridded data was developed using inverse distance weighted interpolation method (Shepard, 1968) from the precipitation data obtained from a large network of 6995 gauge stations (Pai et al., 2014).

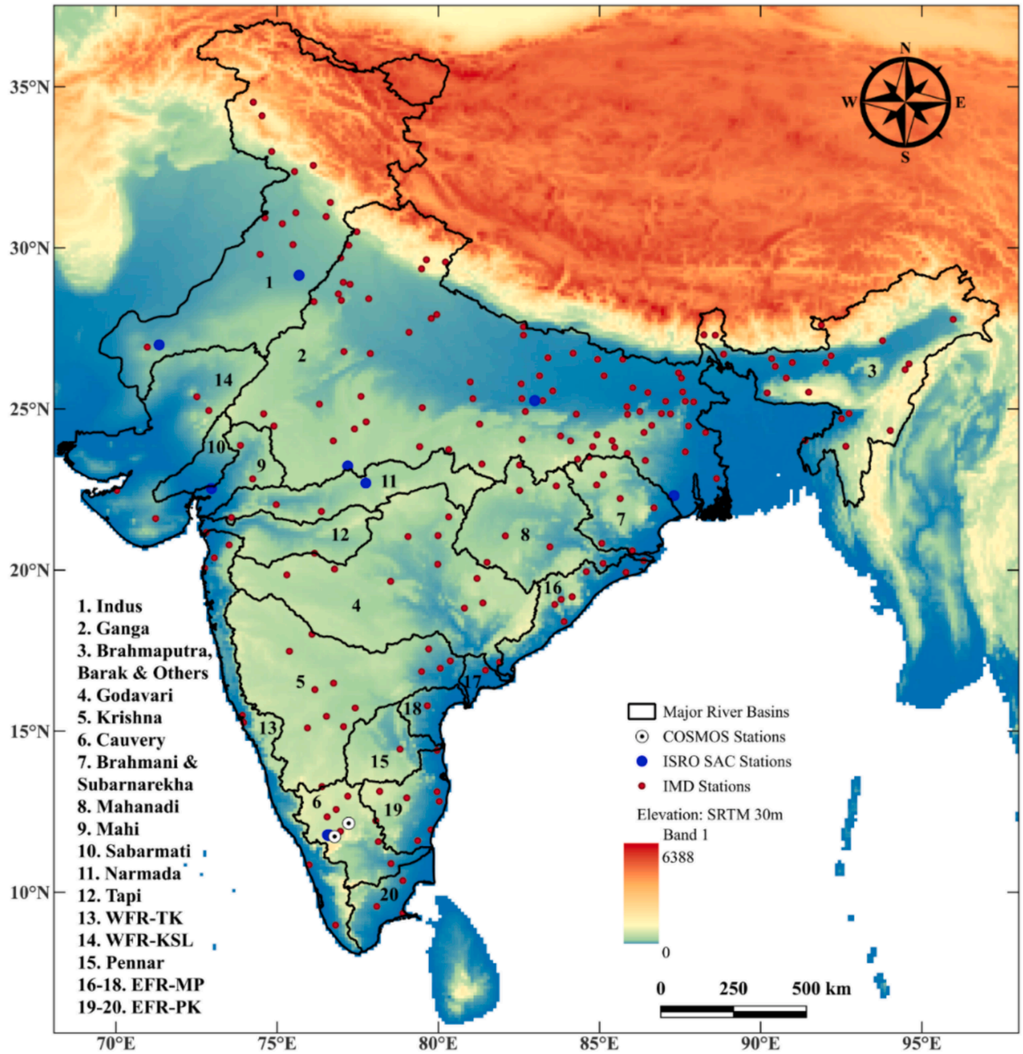

Fig. 1. Study Domain showing elevation from SRTM Native along with major river basins within the national boundary, in situ soil moisture stations of IMD, ISRO, and COSMOS.

3.1.2. Modern-Era Retrospective analysis for research and applications (MERRA2)

The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2; Gelaro et al., 2017) offers meteorological data starting from 1980. MERRA2 has shown substantial improvements compared to its earlier product, MERRA, by incorporating recent advancements from NASA's Global Modeling and Assimilation Office (GMAO), including updates to the Goddard Earth Observing System (GEOS) and new assimilation schemes for various microwave observations (Gelaro et al., 2017). We used the bias-corrected MERRA2 dataset for the present study, which has hourly time steps and 0.625° x 0.5° spatial resolution.

3.1.3. Climate Hazards group InfraRed precipitation with station data (CHIRPS)

The another precipitation dataset from 1981 to 2022 time period has been collected from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) version 2 (Funk et al., 2015), spanning over the domain between 50°S to 50°N for all longitudes. CHIRPS incorporates the Climate Hazards group Precipitation Climatology (CHPclim), 0.05° resolution satellite imagery, and in-situ station data. CHIRPS uses the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis version 7 and a 'smart interpolation' approach while working with anomalies from a high-resolution climatology.

3.1.4. SMAP observation data

Soil Moisture Active Passive (SMAP) Level 3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 8 (O'Neill et al., 2021) has been used in this study. This SMAP passive microwave radiometer data package combines daily estimates of global land surface conditions. SMAP L-band soil moisture data are resampled to a global, cylindrical Equal-Area Scalable Earth Grid with a 36 km grid spacing.

3.2. Evaluation datasets

3.2.1. In-situ soil moisture

This study uses station soil moisture data from three sources:

a. In-situ soil moisture data up to 10 cm depth from 193 stations in an hourly scale belonging to the Indian Meteorological Department for the year 2022.
b. In-situ daily soil moisture data from 11 soil moisture stations from 2017 to 2020 developed and operationalized by the Space Application Center (SAC), Indian Space Research Organization (ISRO) (Gupta et al., 2023; Pandey et al., 2021). All stations have Hydraprobe sensors.
c. Observed daily soil moisture from 2016 to 2019 at two stations the from Indian Cosmic Ray Network (ICON; Upadhyaya et al., 2021): Singanallur-SGR and Madahalli-MDH. The COSMOS instruments use

the Cosmic Ray Neutron Probe (CRNP) technique to measure soil moisture considering non-invasive neutron counts.

3.2.2. Irrigation data
The Global Map of Irrigation Areas (GMIA; Siebert et al., 2005) provided by Food and Agriculture Organization (FAO) is used in this study for analyzing the impact of SMAP data assimilation on irrigated areas. GMIA provides the global map of fractional irrigated areas at a resolution of 5 arc minute, which shows the percentage of each grid cell equipped for irrigation. Apart from this, the Landsat-Derived Global Rainfed and Irrigated-Cropland Product (LGRIP; Teluguntla et al., 2023) has also been used in this work. LGRIP provides high resolution (30 m), global cropland data classifying them into irrigated and rainfed croplands. LGRIP uses the Landsat 8 time-series satellite sensor data for the 2014–2017 period to create a nominal 2015 product.

4. Methodology

4.1. The Indian land data assimilation System (ILDAS)

ILDAS is an integrated modeling system with coupled land surface and hydrodynamic models for reliable estimation of land surface states and water balance components over South Asia (Magotra et al., 2024). Developed using the NASA Land Information System Framework (LISF), ILDAS utilizes many meteorological forcing datasets to produce high-resolution and spatially coherent assessments of different land surface conditions for the Indian subcontinent at 0.1° spatial resolution. The simulations in this study have been performed using the Noah-Multiparameterization (Noah-MP) Land Surface Model, (Niu et al., 2011; Yang et al., 2011) Version 3.6, incorporated within the NASA Land Information System (LIS) framework (Kumar et al., 2006, 2008; Peters-Lidard et al., 2007), Version 7.4 to assimilate SMAP soil moisture dataset. Noah-MP has shown significant advancements in its structure over Noah LSM (Ek et al., 2003) by more accurately considering for various multi-physics options along with the vegetation, groundwater, and snow dynamics (Nie et al., 2018; Niu et al., 2011; Yang et al., 2011). Noah-MP includes four soil layers with a total soil depth of 2 m along with a simple groundwater scheme that incorporates the interaction between soil moisture and groundwater and the formation of runoff. Noah-MP simulations were performed on a cylindrical grid with an equidistant spacing of 0.1° × 0.1° and a 15-minute time step. The Geophysical variables used as input parameters for the land surface modelling framework like land cover, soil texture, elevation, greenness etc. are mostly collected from National Centers for Environmental Prediction (NCEP), National Center for Atmospheric Research (NCAR) reanalysis product and they are enlisted in Table 1. More specifications of ILDAS and its performance evaluations are available in Magotra et al., (2024).

The resulting input parameter file was generated using NASA Land Data Toolkit (LDT; Arsenault et al., 2018).

4.2. Data assimilation algorithm

This study uses the Ensemble Kalman Filter (EnKF) technique, which is a widely used data assimilation method in geosciences, including the assimilation of sequential satellite soil moisture data (Reichle et al., 2002), such as the SMAP soil moisture dataset, into land surface models (He et al., 2017; Mladenova et al., 2020; Pignotti et al., 2023). The primary objective of EnKF is to estimate the state of a dynamic system by combining model predictions with observations and their respective errors.

The EnKF based data assimilation works through one forecast step, followed by an update step. In the forecast step, an ensemble of model states (ensemble members) is generated and propagated with time using LSM. These ensembles are the perturbed versions of the model state, which accounts for the uncertainty in the state. The EnKF then combines the model forecast and observations to update the state estimate using the following equation:

$$x_{i,j}^{t+} = x_{i,j}^{t-} + K_t(y_{i,j}^t - H.(x_{i,j}^{t-})) \tag{1}$$

Where $x_{i,j}^{t+}$ denotes the jth ensemble member (from 1 to 20) of the state variable after update at ith grid, $x_{i,j}^{t-}$ is the prior estimate of the state variable before update. Here the state variable for each ensemble comprises of Noah-MP modeled four layer's soil moisture estimates. $y_{i,j}^t$ denotes the jth ensemble member of the observed variable (here SMAP soil moisture observations) for the same grid. H is the model's representation of the observations, known as observation operator which maps the model state to the observation space. K_t is the Kalman Gain Matrix that denotes the relative weights of uncertainties in the model and observations, and it is calculated using the following equation such that it minimizes the expected error in the estimated state:

$$K_t = \frac{C_{fi}^t.H_i^{t^T}}{H_i^t.C_{fi}^t.H_i^{t^T} + C_{oi}^t} \tag{2}$$

Where C_{fi}^t and C_{oi}^t denotes the covariance of the model forecast errors and that of observation errors respectively.

4.3. Experimental setup

The experimental set up includes spin up simulations of Noah-MP land surface model so as to obtain steady state in LSM, followed by ensemble generation, Open Loop simulation and finally Data Assimilation. For all these simulations, meteorological boundary conditions have been provided using three set of forcing datasets: (i) MERRA2, and (ii) IMD (precipitation) overlain with MERRA2 (other forcing variables) and (iii) CHIRPS (precipitation) overlain with MERRA2 (other forcing variables). The purpose of using three different forcings is to observe the influence of meteorological boundary conditions along with data assimilation on modeled soil moisture.

4.3.1. LSM spin up
The Noah-MP LSM for the present study was spun up three times from 1st January 2010 to 31st December 2015 to obtain the initial conditions. In accordance with the convergence criterion set by the Global Soil Wetness Project (GSW) for LSM spin up (Dirmeyer et al., 1999), the difference in monthly mean soil moisture content for the entire study region had to be less than 5 % after three cycles with respect to the prior simulation to be acceptable. The ensembles have been generated by taking the restart file from March 2015 to provide initial conditions for OL and DA simulations and upscaling the number of ensembles per tile to 20 from the single replicate restart file using LDT.

4.3.1.1. Open loop (OL) experiments. The model was run without assimilating any soil moisture observations from 1st April 2015 to 31st December 2022 in an ensemble configuration, denoted as the Open Loop (OL) simulation. OL simulation was considered as base line run of Noah-MP LSM, to compare the results with DA simulation and observe improvements or degradations within the domain and the simulation period.

Table 1
Input parameters for Land Surface Model.

Model Parameters	Input Data Sources
Landcover	MODIS Native (IGBP) modified by NCEP)
Soil texture	STATSGO + FAO Blended soil texture map by NCAR
Elevation, Slope, Aspect	SRTM Digital Elevation Model (30 m)
Albedo, Greenness	NCEP
Max snow albedo	Barlage Native (Barlage et al., 2005)
Bottom temperature	Native (NCEP) ISLSCP1 temperature derived map

4.3.1.2. Data assimilation (DA) experiments. The SMAP soil moisture observations (SPL3SMP) were assimilated into Noah-MP model for the same period as OL. The ensembles were generated by perturbing four meteorological forcing parameters and model state variables, with the specifications given in Table 2. The SMAP retrieval error standard deviation is taken as 0.04 $m^3\ m^{-3}$ with reference to the studies conducted by O'Neill et al. (2014) and Ahmad et al. (2022).

Our approach doesn't involve any bias correction method (such as CDF matching) as employing bias corrections have been found to exclude the signals from irrigation (Ahmad et al., 2022; Kumar et al., 2015; Kwon et al., 2022). The experimental methodology has been described in Fig. 2.

4.4. Evaluation metrics

The statistical evaluation and validation of the soil moisture outputs have been performed using the following metrices:

To check the variability in seasonal soil moisture differences, the standard deviation (σ) is used as the evaluation metric.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N}} \tag{3}$$

Where, σ is the standard deviation, N is the number of data points, X_i represents each individual data point, $\overline{X}$ is the mean (average) of the data set.

Further statistical analysis and validation of model soil moisture outputs with respect to in situ soil moisture data is performed in terms of correlation (R), root mean squared error ($RMSE$) and unbiased root mean squared error ($ubRMSE$).

$$R = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \overline{X})^2 \sum_{i=1}^{N}(Y_i - \overline{Y})^2}} \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(X_i - Y_i)^2}{N}} \tag{5}$$

$$ubRMSE = \sqrt{\frac{\sum_{i=1}^{N}\{(X_i - \overline{(X_i - Y_i)}) - Y_i\}^2}{N}} \tag{6}$$

Where, X_i, $\overline{X}$ corresponds to model output and Y_i, $\overline{Y}$ corresponds to standard reference data with their usual meaning as discussed earlier. Improvements in soil moisture estimates due to data assimilation is evaluated in terms of difference between correlations indicated as R_{DA}-R_{OL} and in terms of difference in RMSE indicated as $RMSE_{OL}$-$RMSE_{DA}$ with an assumption that DA is giving improved results by showing higher correlation and lower RMSE when compared with observed data. If these differences become negative, we will consider as degradation.

5. Results and discussions

The analysis period, from 1st April 2015 to 31st December 2022, has been sub-divided into 4 seasons based on: Pre-Monsoon (MAM; March-May), Monsoon (JJAS; June-September), Post-Monsoon (ON; October-November) and Winter (DJF; December-February). Irrigation is mostly practiced in India for most of the rabi (winter) crops during the winter and pre monsoon season. Kharif crop season receives south-west monsoon rainfall and hence less irrigation is required. Thus, our seasonal analysis comprises of an average of eight seasons of pre-monsoon, monsoon, post-monsoon and winter seasons from June 2015 to November 2022.

5.1. Seasonal analysis

Fig. 3 shows the season-wise temporal average difference in soil moisture values between the DA and OL simulations (DA-OL). For quantification of this variability, standard deviations of this soil moisture difference values have been calculated. Higher values of difference (positive or negative) show maximum variability, indicating significant impact of data assimilation on soil moisture estimates by LSM.

Overall, the lower values of standard deviation (σ) in Fig. 3 indicate that pre-monsoon and monsoon exhibit lower variability in soil moisture. Whereas post-monsoon and winter maps show higher variability in soil moisture. In fact, lowest variability can be observed during the monsoon season for all forcings compared to other seasons and this variability is almost the same between OL and DA cases, which could be attributed to the ample availability of precipitation during that period. In contrast, maximum variability can be observed during the dry winter season, highlighting the fact that DA has more impact on soil moisture during winter. Overall, for all seasons and all forcing cases, DA leads to higher soil moisture estimates than OL for the North-west and Eastern India, which covers the major portion of Ganga basin, Punjab area from Indus basin and the Mahanadi basin. Comparison in terms of forcings shows that IMD shows maximum standard deviation out of all the forcing cases. That means IMD forced DA simulation can capture maximum variability in soil moisture.

5.2. Statistical analysis and validation

Benchmarking of LSM soil moisture estimates for both OL and DA cases has been performed with respect to in situ station soil moisture datasets.

Fig. 4 shows the difference in statistical parameters between DA and OL estimated soil moisture anomalies at different IMD stations. Here the first row (Fig. 4, a-c) shows the correlation difference (R_{DA}-R_{OL}) and second row (Fig. 4, d-f) corresponds to RMSE difference ($RMSE_{OL}$-$RMSE_{DA}$). Blue/cool color indicates improvement and red/warm color indicates degradation. The figures indicate large-scale improvement in both correlation and RMSE over majority of the stations due to data assimilation, although the improvement in terms of RMSE is lower than correlation because of incorporating errors. The domain averaged improvements of all the forcing cases are found positive but almost similar relative changes. IMD forced outputs have shown relatively higher domain average change in correlation (0.0178) and RMSE (0.0029 m^3/m^3) than the other two forcing cases. These differences may vary from station to station depending on the performance of DA at different locations as evident from the color variation in the plots. So the spatial mean values of these metrics evaluated over the duration of observed data may not indicate significant differences but their positive mean

Table 2
Perturbation specifications for meteorological forcings and model state variables.

	Variable	Perturbation Type	Std. Deviation	Cross Correlation			
				SW	LW	P	NST
Met Forcings	Shortwave Radiation (SW)	Multiplicative	0.30	1.0	−0.3	−0.5	0.3
	Longwave Radiation (LW)	Additive	50.0 $W.m^{-2}$	−0.3	1.0	0.5	0.6
	Precipitation (P)	Multiplicative	0.50	−0.5	0.5	1.0	−0.1
	Near Surface Air Temperature (NST)	Additive	1 K	0.3	0.6	−0.1	1.0
LSM States	Soil Moisture Layer 1 (SM1)	Additive	0.006 $m^3.m^{-3}$				

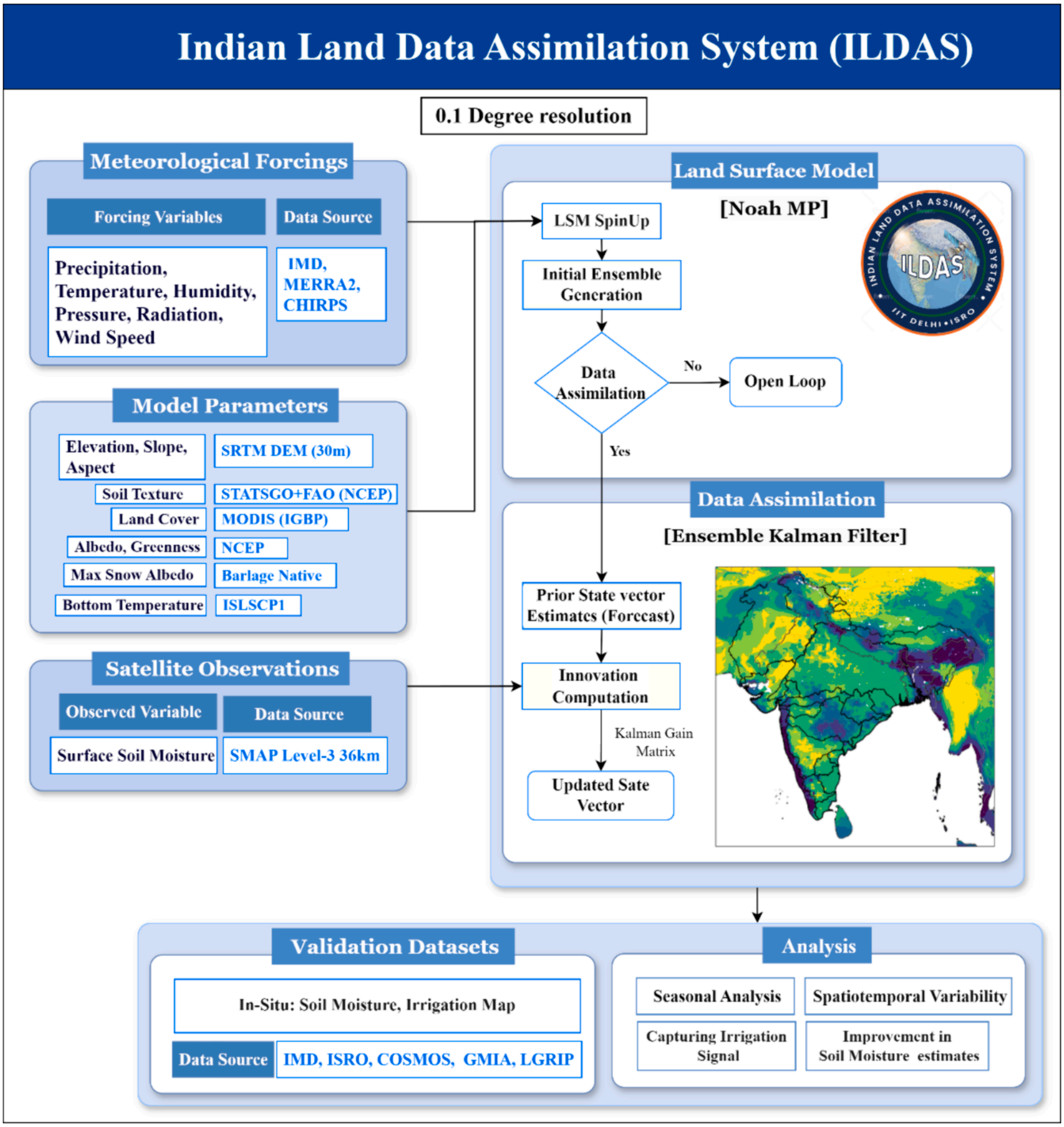

Fig. 2. Land Surface Model Simulation and Data Assimilation Methodology within Indian Land Data Assimilation System (ILDAS).

values show the overall improvement due to data assimilation. However, the distribution of correlation difference shows maximum positive stations in the case of MERRA2.

Further evaluation and validation of the results are performed using in situ soil moisture records of ISRO SAC and COSMOS. Fig. 5 shows the correlation and unbiased RMSE (ubRMSE) of monthly mean soil moisture obtained from the simulation cases evaluated for the 11 ISRO SAC soil moisture stations from 2017 to 2020. For most of the stations, correlations of DA estimates are either higher or almost equal to OL results, except the stations of Hisar (29.15°N,75.68°E) and Hoshangabad (22.77°N, 77.75°E), showing minimum correlation values in comparison with respective OL and DA simulation cases. In case of unbiased RMSE, for all the forcing cases, unbiased RMSE values for OL cases are higher or almost equal to the DA estimates. That means DA is showing less ubRMSE than OL, indicating improved soil moisture results due to DA.

Fig. 6 shows the scatter plots between LSM simulated monthly mean soil moisture and the station soil moisture from 2017 to 2020, belonging to 5 ISRO SAC soil moisture stations and corresponding regression values for all the simulation cases. For almost all the stations the value of regression coefficient (R^2) is more in case of DA than the OL simulation cases.

The assimilated outputs have been validated further with COSMOS in-situ soil moisture data for the two stations (Singanallur and Madahalli). Fig. 7 shows the scatter plots between DA outputs with station soil moisture along with the variation in correlation values for all the three forcings. Most of the DA soil moisture estimates show a higher correlation between the range of 0.8 to 1. For SGR station DA estimates give relatively higher coefficient of determination (R^2) than the MDH station. Moreover, as evident from the slope of the regression line, IMD and CHIRPS forced outputs are slightly overestimating soil moisture for the SGR station.

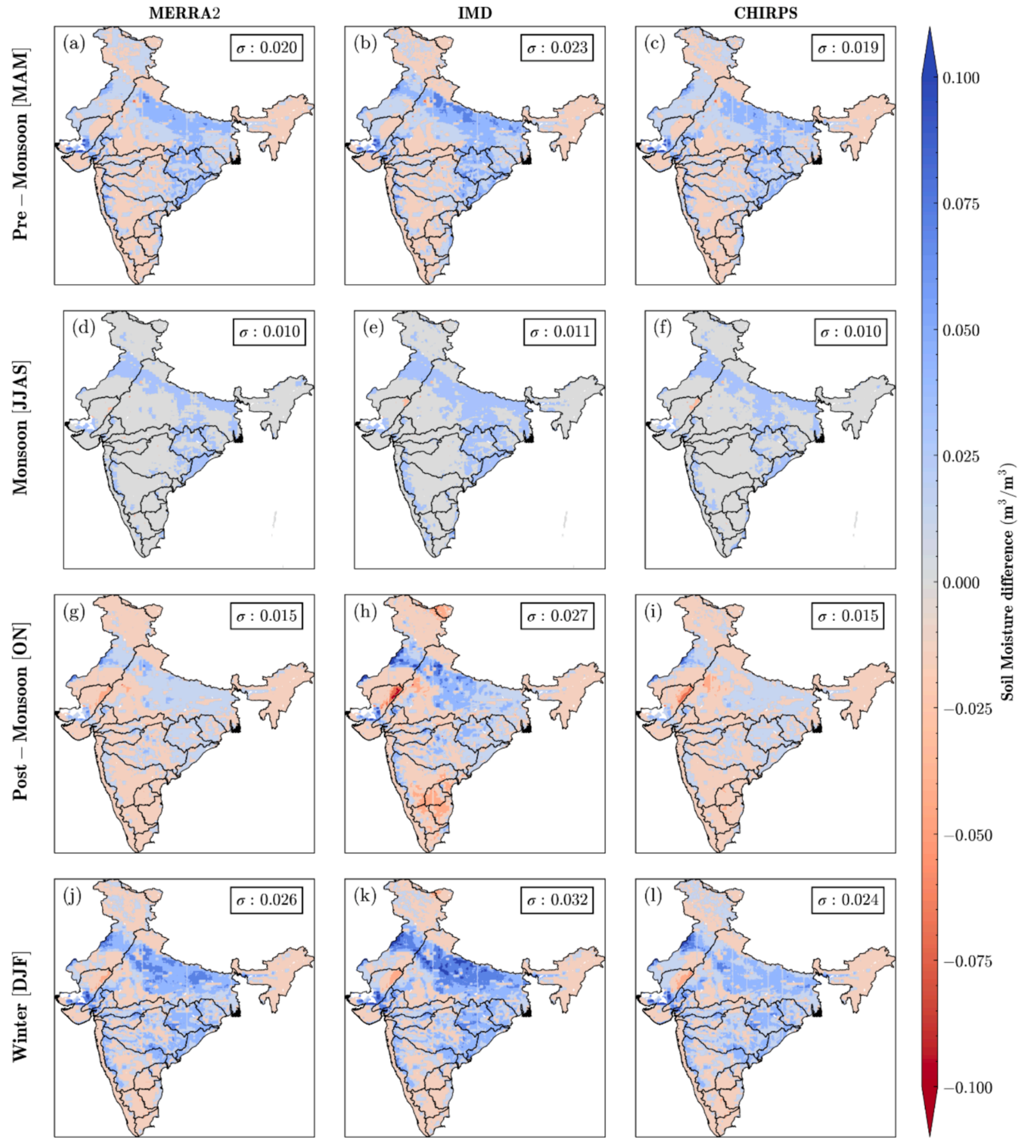

Fig. 3. Seasonal Impact of DA in terms of soil moisture difference (m^3/m^3) maps between DA and OL when forced with MERRA2, IMD & CHIRPS respectively for the pre-monsoon season (a-c), monsoon season (d-f), post monsoon season (g-i) and winter season (j-l).

5.3. Impact of irrigation on assimilation scheme

Soil moisture plays an important role in deciding the irrigation requirements of crops. Hence it is necessary to observe the influence of DA over irrigated areas. To quantify the influence of assimilation over irrigated areas, the soil moisture difference values have been compared with the GMIA percentage irrigation fraction data (Fig. 8, b). Rabi (winter) crop production in India is mainly dependent on irrigation water, as during the winter cropping season, only a few areas of southern India receive rainfall due to retreating monsoon. After comparing the IMD soil moisture difference (DA-OL) map corresponding to winter season (Fig. 8, a) with the GMIA irrigation fraction map for India, the blue patches over the northwestern India covering the entire Gangetic plains and Punjab, indicating higher positive soil moisture difference values are found almost spatially consistent with the higher irrigation fraction values.

To further check the signatures of irrigation from DA in terms of the difference in soil moisture and precipitation, the difference in climatological mean soil moisture between DA and OL of all the forcing cases have been plotted as time series in Fig. 9 (a-e) for different irrigation fractions along with IMD precipitation (bar plots). For simplicity of the analysis, the irrigated regions of entire study domain have been classified into 5 classes, starting from 0 to 100 % with range of each class being selected as 20 %. For all the irrigation fraction classes, the differences in soil moisture between DA and OL are very small (in between 0 to −0.02) during the rainy season (JJAS, as highlighted with sky-blue background), as already stated in Fig. 3. Interestingly even during the dry winter season (DJF, highlighted with green background), all the plots still show relatively higher soil moisture differences than other seasons, which can be attributed to presence of irrigation during that dry winter season. In fact, these higher values of soil moisture difference between DA and OL are increasing with the increase in GMIA irrigation fraction as evident from the increasing soil moisture axis labels range of the respective five subplots in Fig. 9. This indicates that DA, with higher values than OL during winter, is able to capture irrigation signals caused by higher percentage area during winter.

This variation has been quantified separately for both the irrigated and non-irrigated pixels obtained from the high resolution LGRIP dataset (Fig. 10, a). The distribution of mean winter season soil moisture difference between DA and OL is observed for the identified irrigated cropland and non-cropland pixels (Fig. 10, b-c).

These box plots reveal that the soil moisture difference is more in

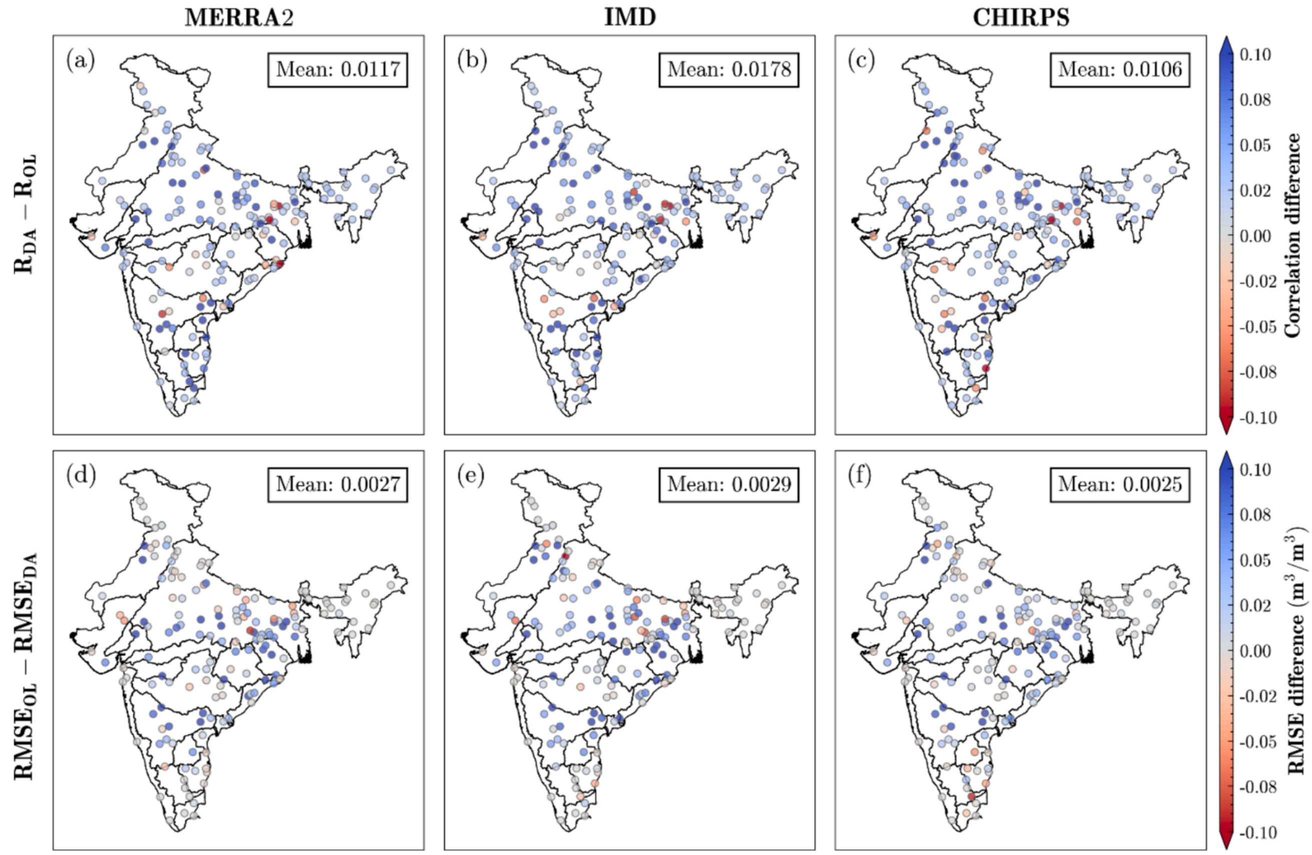

Fig. 4. Difference in statistical parameters of soil moisture anomalies between and open loop (OL) and data assimilated (DA) simulations with respect to IMD in situ soil moisture anomalies. The correlation difference maps (R_{DA}-R_{OL}) (a-c) are for simulations when forced with MERRA2, IMD and CHIRPS respectively. (d-f) same as (a-c) but for RMSE ($RMSE_{OL}$-$RMSE_{DA}$).

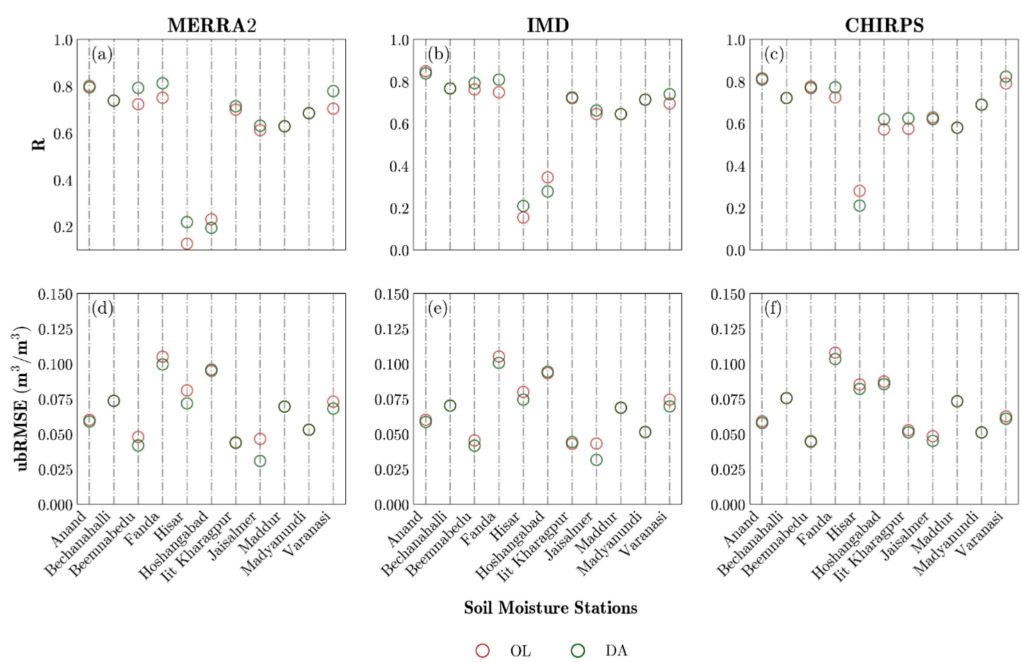

Fig. 5. Dot plots of statistical parameters obtained for 11 ISRO SAC soil moisture stations from 2017 to 2020. (a), (c), (e)- Correlation plots for MERRA2, IMD and CHIRPS forced simulations. (b), (d),(f)- same as (a),(c) & (e) but for unbiased RMSE (m^3/m^3).

case of IMD (0.0225 to 0.0275 m^3/m^3). Even the range of soil moisture difference is more in case of irrigated croplands.

In terms of correlation difference, the improvement in soil moisture estimates can be further stratified for irrigated and non-irrigated (rain-fed + noncropland) areas as per LGRIP data. For this purpose, the station wise correlation difference between DA and OL estimates have been

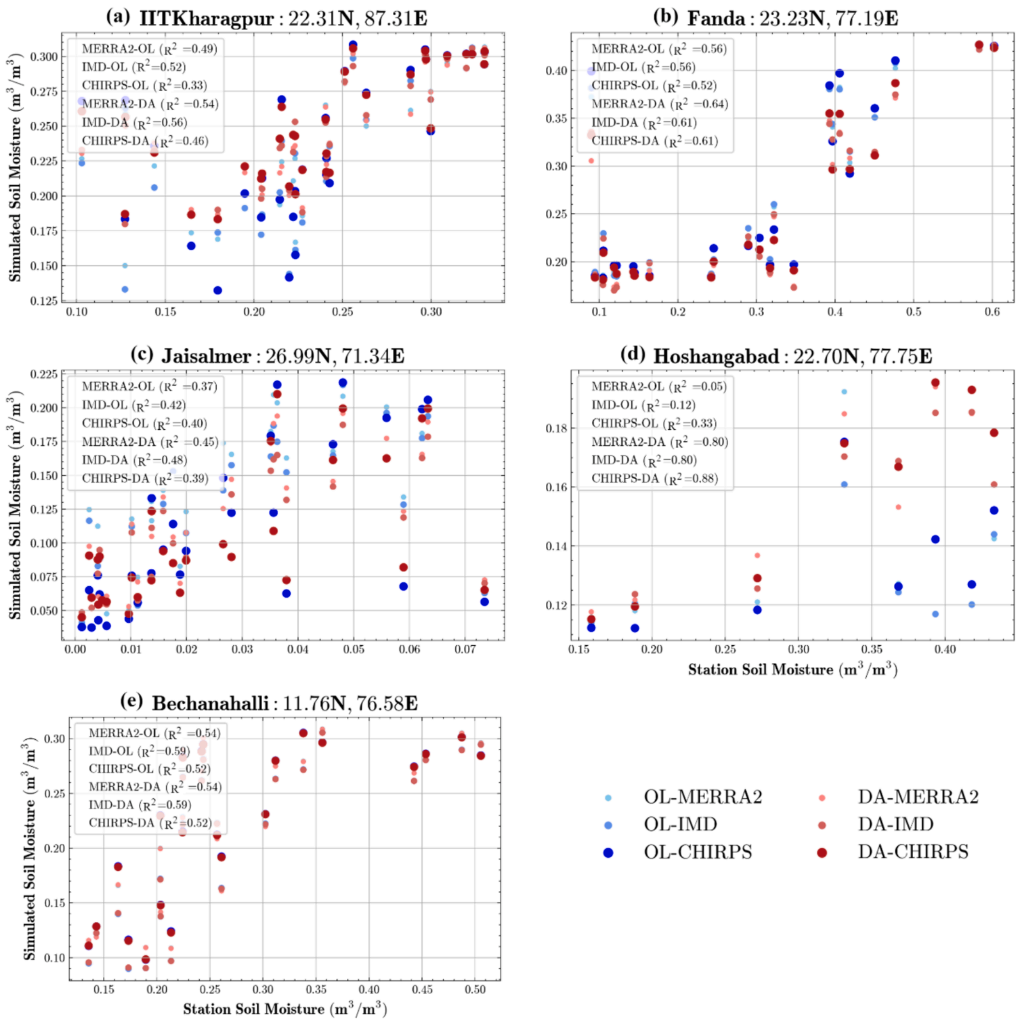

Fig. 6. Scatter plots between monthly mean simulated and observed station soil moisture from 2017 to 2020 for 5 ISRO SAC soil moisture stations.

interpolated using inverse distance weighted interpolation over the irrigated and non-irrigated regions (Fig. 11).

Overall, for most of the station locations the DA scheme is improving soil moisture estimates except few stations as evident from Fig. 4 (a-c) also. But the variation and range of correlation differences is more in irrigated areas (Fig. 11, a-c), than the non-irrigated regions (Fig. 11, d-f). MERRA2 and IMD forced assimilation scheme show relatively higher improvement over irrigated areas than CHIRPS. That means the assimilation scheme employed here without CDF matching can capture irrigation signal and improves soil moisture estimates almost over the entire domain. But optimal data assimilation assumes unbiasedness between the model and the observed data. Hence, as per the current study is concerned, the unbiasedness assumptions are violated across irrigated areas. So, it has become necessary to observe the bias between the modeled estimates and observation especially over the non-irrigated areas. For this purpose, the bias in mean soil moisture and standard deviation between modeled and SMAP observations have been plotted for the non-irrigated areas where SMAP observations are available for most of the time (Fig. 12).

Fig. 12 shows that the difference in mean soil moisture estimates (Fig. 12, a-c) as well as standard deviation (Fig. 12, d-f) between OL and SMAP are within the range of ± 0.1. Although some variations can be observed for both the cases, especially over Rajasthan region, eastern and southern India, signifying that the regions are having some biases i. e, not fully unbiased.

6. Discussion

Seasonal Analysis of soil moisture difference between assimilated and model only estimates have shown the ability of assimilation scheme to capture irrigation signal. Large scale evaluation and validation of model outputs with respect to in situ dataset shows improvement of most of the regions within the study domain. But it is important to note that assimilating some other satellite data without correcting for bias may not enhance estimation accuracy compared to bias corrected methods in certain situations. In this study, SMAP retrievals are capable of detecting the influence of irrigation and that information is then delivered to the model through assimilation. As per the studies conducted by Ahmad et al., (2022) and Kwon et al., (2022), Data assimilation with CDF matching eliminates the spatial signature of irrigation from the SMAP soil moisture retrievals. For further verification, another study has been conducted considering the current study domain and specifications where CDF matching is performed for correcting bias. Considering the scope and length of this paper, the CDF matching results have been included in the supplementary material. Here also CDF matching simply excludes the biases due to irrigation (Fig. S1) and the improvements in soil moisture estimates are relatively less than without CDF matching case (Fig. S2). However, assimilating some other soil moisture product without bias adjustment may reduce the assimilation skill based on the presence of biases. Hence it is necessary to improve model physics by parameterizing these unmodeled processes leading to a better-performing model that can be utilized for correcting bias in

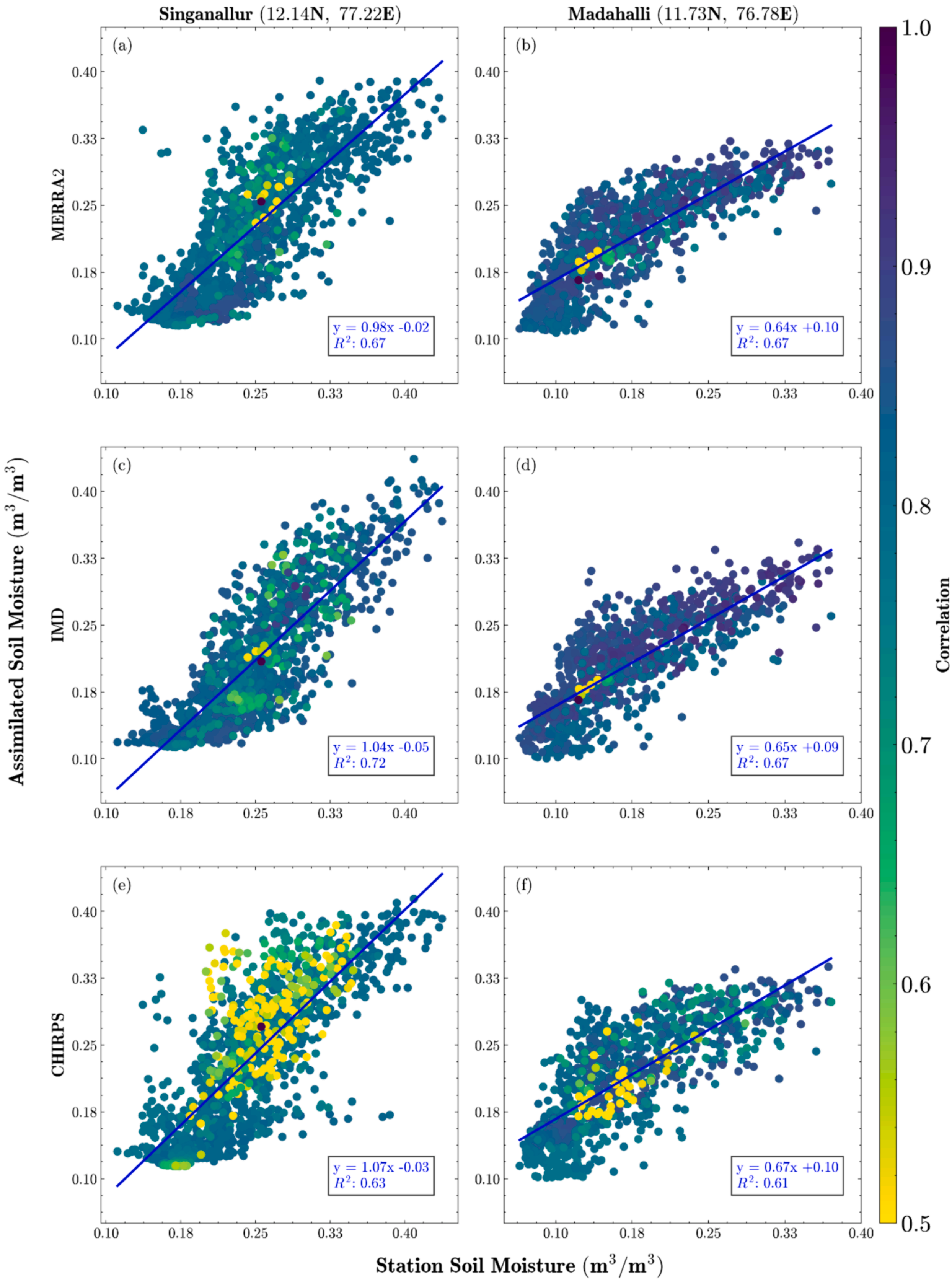

Fig. 7. Correlation and Regression analysis of Assimilated (DA) soil moisture with respect to COSMOS in situ soil moisture data (2016–2019) from Singanallur and Madahalli station respectively. Scatter plots of MERRA2 (a-b), IMD (c-d) and CHIRPS (e-f) forced DA soil moisture and in situ soil moisture.

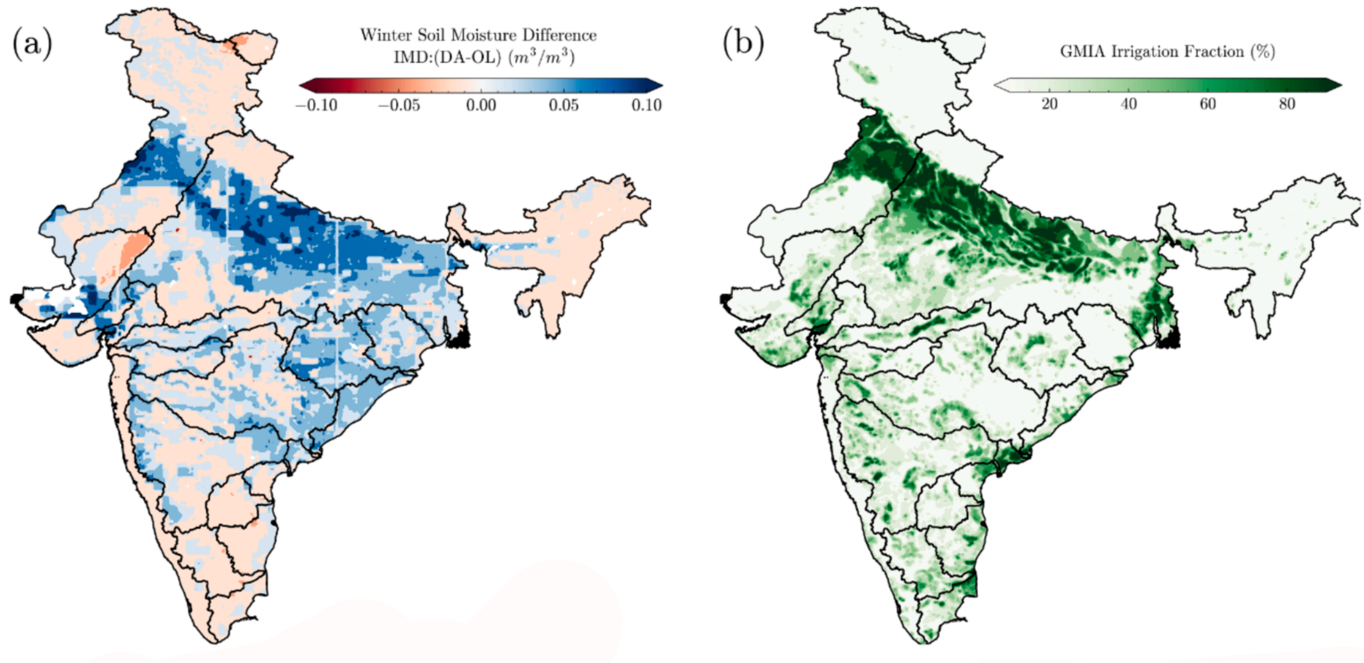

Fig. 8. (a) Winter season map of soil moisture difference, (b) GMIA Irrigation fraction over Indian domain.

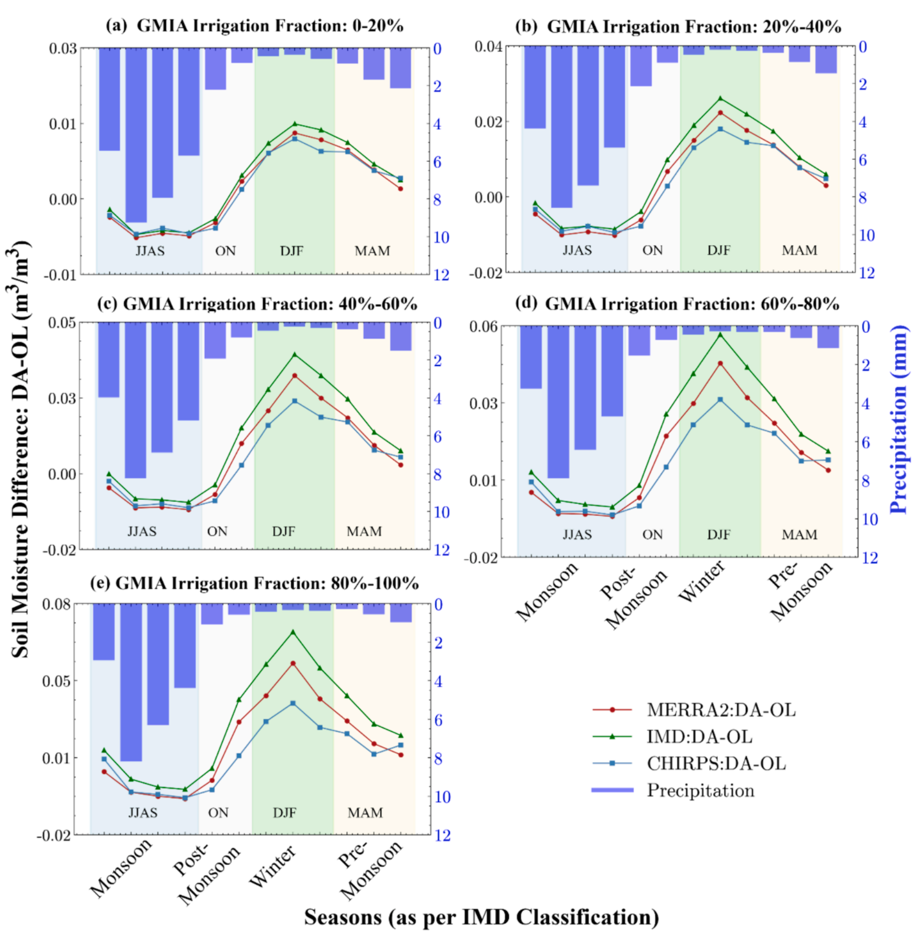

Fig. 9. Plots of differences between climatological mean monthly soil moisture estimates from DA and OL for all the three forcing cases (MERRA2, IMD and CHIRPS), along with IMD climatological monthly mean precipitation (indicated by bars) for the regions belonging to different irrigation fractions as per GMIA.

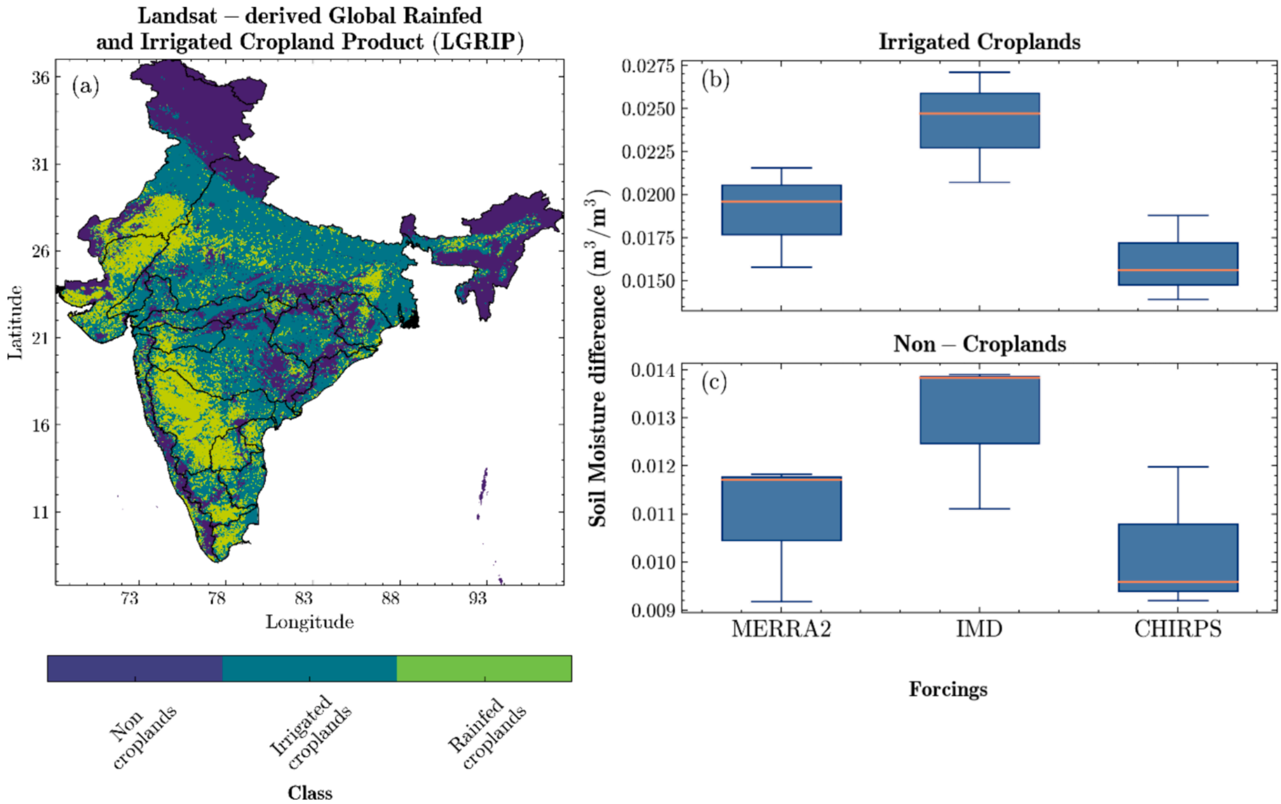

Fig. 10. Irrigated cropland data:(a) Landsat-derived Global Rainfed and Irrigated Cropland Product (LGRIP); (b) & (c) Winter season distribution of soil moisture difference between DA and OL for the Irrigated cropland and non-cropland pixels respectively.

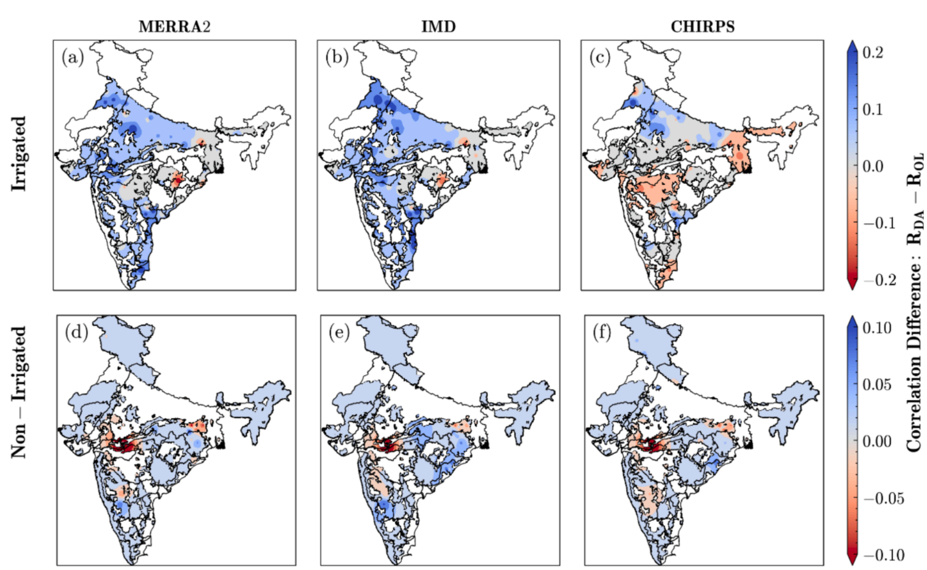

Fig. 11. Spatially Interpolated correlation difference maps for irrigated (a-c) and non-irrigated (d-f) regions.

satellite retrievals.

7. Conclusions and future work

This study involved assimilation of the SMAP soil moisture observations into the Noah-MP land surface model within the Indian Land Data Assimilation System (ILDAS) to observe the spatiotemporal variation in soil moisture estimates due to DA across the Indian subcontinent. Precipitation boundary conditions were provided by the MERRA2, IMD gridded precipitation, and CHIRPS precipitation dataset. SMAP retrieval assimilation was implemented using Ensemble Kalman Filter (EnKF) method. The purpose of the study was to assess the efficacy of data assimilation in improving estimations of soil moisture and capturing irrigation signals through the detection of moist soil caused by irrigation in dry natural environments. The conclusions can be summarized as follows:

- The impact of data assimilation is significant during post monsoon and winter seasons showing higher standard deviations than Premonsoon and monsoon season. During the winter season, DA leads

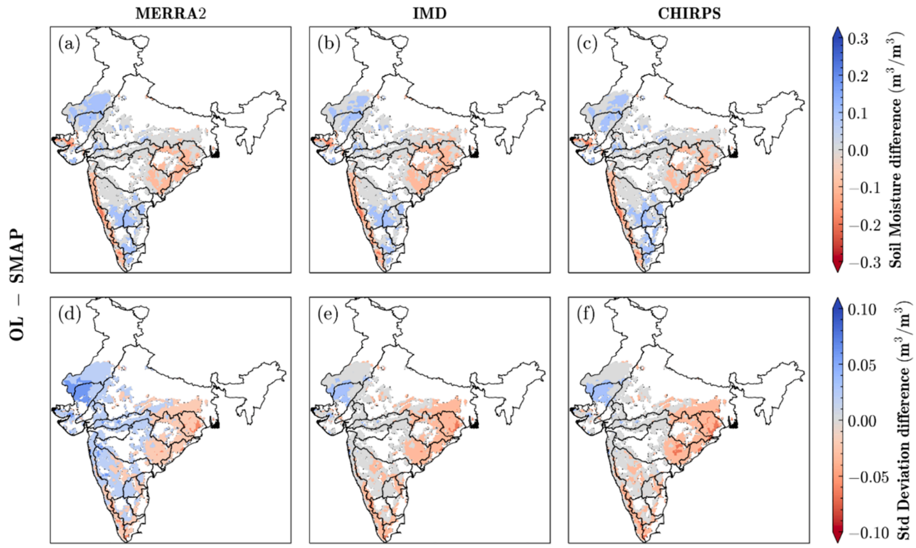

Fig. 12. Bias between model only (OL) and SMAP dataset in terms of mean soil moisture (a-c), and standard deviation (d-f) for the non-irrigated regions.

to higher values of soil moisture than OL especially over northern India covering the Gangetic plain.

- DA simulations using IMD forcing exhibited the highest spatial mean improvement and increased the mean spatial correlation by 0.0178 and reduced RMSE by 0.0029 when compared with in-situ soil moisture data across the ILDAS domain. Noah MP Model forced by IMD precipitation and overlain with MERRA2 was found to be the best combination for soil moisture estimation.
- DA led to higher soil moisture values than open loop estimates during dry winter season. Higher values of difference in soil moisture between DA and OL estimates are observed over the irrigated cropland regions identified from the GMIA irrigation fraction data and high-resolution Landsat derived cropland data. In fact, the spatial variability in improvement is more over irrigated areas than the non-irrigated areas.

Thus, SMAP data assimilation has the potential to improve soil moisture estimation compared to open-loop simulation, and significant impact has been observed over the highly irrigated regions in India as per the GMIA map. Further, the model is found capable of detecting the spatial signature of irrigation in terms of relatively elevated soil moisture levels than open loop estimated soil moisture. This is the first study using Indian Land Data Assimilation System (ILDAS), where a coherent soil moisture data assimilation framework has been developed over entire India, validated extensively using a hitherto-unused soil moisture station network, and used to successfully demonstrate improved soil moisture estimation as well as an inverse approach to infer irrigation spatial patterns over the region. In the future, we will investigate the impact of different observation rescaling approaches and bias correction methods. The developed ILDAS system will lead to better assessment of hydrological extremes (flood, drought etc.) and also serve as a test-bed for assimilating soil moisture products from the upcoming NASA-ISRO Synthetic Aperture Radar (NISAR) mission.

8. Compliance with ethical standards

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Arijit Chakraborty: Writing – original draft, Validation, Methodology, Formal analysis, Data curation. **Manabendra Saharia:** Writing – review & editing, Supervision, Funding acquisition. **Sumedha Chakma:** Writing – review & editing. **Dharmendra Kumar Pandey:** Writing – review & editing, Data curation. **Kondapalli Niranjan Kumar:** Writing – review & editing, Data curation. **Praveen K. Thakur:** Writing – review & editing. **Sujay Kumar:** Review, Software, Methodology. **Augusto Getirana:** Writing – review.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The different datasets used in this study were obtained from the following sources: Indian Land Reanalysis: (1) MERRA-2, NASA Goddard Space Flight Center: https://disc.gsfc.nasa.gov/datasets?project%C2%A0=%C2%A0MERRA-2. (2) IMD precipitation: https://www.imdpune.gov.in/Clim_Pred_LRF_New/. (3) CHIRPS from Climate Hazards Center, UC Santa Barbara: https://data.chc.ucsb.edu/products/CHIRPS-2.0/. (4) SMAP data: https://nsidc.org/data/spl3smp/versions/8. (5) GMIA: https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version. (6) LGRIP: https://lpdaac.usgs.gov/products/lgrip30v001/.

Acknowledgements

This research was conducted in the HydroSense lab (https://hydrosense.iitd.ac.in/) of IIT Delhi and the authors acknowledge the IIT Delhi High Performance Computing facility for providing computational and storage resources. Dr. Manabendra Saharia gratefully acknowledges financial support for this work through grants from ISRO Space Technology Cell (STC0374/RP04139); Ministry of Earth Sciences Monsoon Mission III (RP04574); Ministry of Earth Sciences (RP04741G); and DST IC-IMPACTS (RP04558). The authors gratefully acknowledge ISRO Space Applications Center, Indian Meteorological Department (IMD) and National Center for Medium Range Weather Forecasting (NCMRWF) for the in-situ soil moisture datasets. The authors thank Prof. Sekhar Muddu (IISc) and Ms. Deepti Upadhyay for giving access to COSMOS soil moisture data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2024.131581.

References

Abramowitz, G., Leuning, R., Clark, M., Pitman, A., 2008. Evaluating the Performance of Land Surface Models. J. Clim. 21 (21), 5468–5481. https://doi.org/10.1175/2008JCLI2378.1.

Ahmad, J.A., Forman, B.A., Kumar, S.V., 2022. Soil moisture estimation in South Asia via assimilation of SMAP retrievals. Hydrol. Earth Syst. Sci. 26 (8), 2221–2243. https://doi.org/10.5194/hess-26-2221-2022.

Ahmad, S., Kalra, A., Stephen, H., 2010. Estimating soil moisture using remote sensing data: A machine learning approach. Adv. Water Resour. 33 (1), 69–80. https://doi.org/10.1016/j.advwatres.2009.10.008.

Anderson, W.B., Zaitchik, B.F., Hain, C.R., Anderson, M.C., Yilmaz, M.T., Mecikalski, J., Schultz, L., 2012. Towards an integrated soil moisture drought monitor for East Africa. Hydrol. Earth Syst. Sci. 16 (8), 2893–2913. https://doi.org/10.5194/hess-16-2893-2012.

Arsenault, K.R., Kumar, S.V., Geiger, J.V., Wang, S., Kemp, E., Mocko, D.M., Beaudoing, H.K., Getirana, A., Navari, M., Li, B., Jacob, J., Wegiel, J., Peters-Lidard, C.D., 2018. The Land surface Data Toolkit (LDT v7.2) – a data fusion environment for land data assimilation systems. Geosci. Model Dev. 11 (9), 3605–3621. https://doi.org/10.5194/gmd-11-3605-2018.

Barlage, M., Zeng, X., Wei, H., Mitchell, K.E., 2005. A global 0.05° maximum albedo dataset of snow-covered land based on MODIS observations. Geophys. Res. Lett. 32 (17) https://doi.org/10.1029/2005GL022881.

Botter, G., Peratoner, F., Porporato, A., Rodriguez-Iturbe, I., Rinaldo, A., 2007. Signatures of large-scale soil moisture dynamics on streamflow statistics across U.S. climate regimes. Water Resour. Res. 43 (11) https://doi.org/10.1029/2007WR006162.

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., 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 Obs. Geoinf. 73, 752–766. https://doi.org/10.1016/j.jag.2018.08.023.

Chen, F., Crow, W.T., Bindlish, R., Colliander, A., Burgin, M.S., Asanuma, J., Aida, K., 2018. Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. Remote Sens. Environ. 214, 1–13. https://doi.org/10.1016/j.rse.2018.05.008.

Corradini, C., 2014. Soil moisture in the development of hydrological processes and its determination at different spatial scales. J. Hydrol. 516, 1–5. https://doi.org/10.1016/j.jhydrol.2014.02.051.

Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G., Denning, A.S., Dirmeyer, P.A., Houser, P.R., Niu, G., Oleson, K.W., Schlosser, C.A., Yang, Z.-L., 2003. The Common Land Model. Bull. Am. Meteorol. Soc. 84 (8), 1013–1024. https://doi.org/10.1175/BAMS-84-8-1013.

Dirmeyer, P.A., Dolman, A.J., Sato, N., 1999. The Pilot Phase of the Global Soil Wetness Project. Bull. Am. Meteorol. Soc. 80 (5), 851–878. https://doi.org/10.1175/1520-0477(1999)080<0851:TPPOTG>2.0.CO;2.

Dobriyal, P., Qureshi, A., Badola, R., Hussain, S.A., 2012. A review of the methods available for estimating soil moisture and its implications for water resource management. J. Hydrol. 458–459, 110–117. https://doi.org/10.1016/j.jhydrol.2012.06.021.

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., Lecomte, P., 2017. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. Remote Sens. Environ. 203, 185–215. https://doi.org/10.1016/j.rse.2017.07.001.

Ek, M.B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., Tarpley, J.D., 2003. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. J. Geophys. Res. Atmos. 108 (D22) https://doi.org/10.1029/2002JD003296.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., … Van Zyl, J. (2010). The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918.

Fang, B., Kansara, P., Dandridge, C., Lakshmi, V., 2021. Drought monitoring using high spatial resolution soil moisture data over Australia in 2015–2019. J. Hydrol. 594, 125960 https://doi.org/10.1016/j.jhydrol.2021.125960.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Article 1 Sci. Data 2 (1). https://doi.org/10.1038/sdata.2015.66.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Zhao, B., 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Clim. 30 (14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1.

Gibson, J., Franz, T.E., Wang, T., Gates, J., Grassini, P., Yang, H., Eisenhauer, D., 2017. A case study of field-scale maize irrigation patterns in western Nebraska: Implications for water managers and recommendations for hyper-resolution land surface modeling. Hydrol. Earth Syst. Sci. 21 (2), 1051–1062. https://doi.org/10.5194/hess-21-1051-2017.

Gupta, D.K., Srivastava, P.K., Pandey, D.K., Chaudhary, S.K., Prasad, R., O'Neill, P.E., 2023. Passive Only Microwave Soil Moisture Retrieval in Indian Cropping Conditions: Model Parameterization and Validation. IEEE Transactions on Geoscience and Remote Sensing 61, 1–12. https://doi.org/10.1109/TGRS.2022.3218945.

Hargreaves, G.L., Hargreaves, G.H., Riley, J.P., 1985. Irrigation Water Requirements for Senegal River Basin. J. Irrig. Drain.Engineering 111 (3), 265–275. https://doi.org/10.1061/(ASCE)0733-9437(1985)111:3(265).

He, L., Chen, J.M., Liu, J., Bélair, S., Luo, X., 2017. Assessment of SMAP soil moisture for global simulation of gross primary production. Journal of Geophysical Research: Biogeosciences 122 (7), 1549–1563. https://doi.org/10.1002/2016JG003603.

Hong, S.-Y., Kalnay, E., 2000. Role of sea surface temperature and soil-moisture feedback in the 1998 Oklahoma-Texas drought. Article 6814 Nature 408 (6814). https://doi.org/10.1038/35048548.

Huang, C., Li, X., Lu, L., Gu, J., 2008. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. Remote Sens. Environ. 112 (3), 888–900. https://doi.org/10.1016/j.rse.2007.06.026.

Jalilvand, E., Tajrishy, M., Ghazi Zadeh Hashemi, S. A., & Brocca, L. (2019). Quantification of irrigation water using remote sensing of soil moisture in a semi-arid region. *Remote Sensing of Environment*, *231*, 111226. https://doi.org/10.1016/j.rse.2019.111226.

Jung, H.C., Kang, D.-H., Kim, E., Getirana, A., Yoon, Y., Kumar, S., Peters-lidard, C.D., Hwang, E., 2020. Towards a soil moisture drought monitoring system for South Korea. J. Hydrol. 589, 125176 https://doi.org/10.1016/j.jhydrol.2020.125176.

Kalman, R.E., 1960. A New Approach to Linear Filtering and Prediction Problems. J. Basic Eng. 82 (1), 35–45. https://doi.org/10.1115/1.3662552.

Kim, H., Wigneron, J.-P., Kumar, S., Dong, J., Wagner, W., Cosh, M.H., Bosch, D.D., Collins, C.H., Starks, P.J., Seyfried, M., Lakshmi, V., 2020. Global scale error assessments of soil moisture estimates from microwave-based active and passive satellites and land surface models over forest and mixed irrigated/dryland agriculture regions. Remote Sens. Environ. 251, 112052 https://doi.org/10.1016/j.rse.2020.112052.

Komma, J., Blöschl, G., Reszler, C., 2008. Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting. J. Hydrol. 357 (3), 228–242. https://doi.org/10.1016/j.jhydrol.2008.05.020.

Koster, R.D., Suarez, M.J., Ducharne, A., Stieglitz, M., Kumar, P., 2000. A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. J. Geophys. Res. Atmos. 105 (D20), 24809–24822. https://doi.org/10.1029/2000JD900327.

Koster, R.D., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C.T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Yamada, T., 2004. Regions of Strong Coupling Between Soil Moisture and Precipitation. Science 305 (5687), 1138–1140. https://doi.org/10.1126/science.1100217.

Kumar, S.V., Peters-Lidard, C.D., Tian, Y., Houser, P.R., Geiger, J., Olden, S., Lighty, L., Eastman, J.L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E.F., Sheffield, J., 2006. Land information system: An interoperable framework for high resolution land surface modeling. Environ. Model. Softw. 21 (10), 1402–1415. https://doi.org/10.1016/j.envsoft.2005.07.004.

Kumar, S.V., Reichle, R.H., Peters-Lidard, C.D., Koster, R.D., Zhan, X., Crow, W.T., Eylander, J.B., Houser, P.R., 2008. A land surface data assimilation framework using the land information system: Description and applications. Adv. Water Resour. 31 (11), 1419–1432. https://doi.org/10.1016/j.advwatres.2008.01.013.

Kumar, S.V., Reichle, R.H., Harrison, K.W., Peters-Lidard, C.D., Yatheendradas, S., Santanello, J.A., 2012. A comparison of methods for a priori bias correction in soil moisture data assimilation. Water Resour. Res. 48 (3) https://doi.org/10.1029/2010WR010261.

Kumar, S.V., Peters-Lidard, C.D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K.R., Xia, Y., Ek, M., Riggs, G., Livneh, B., Cosh, M., 2014. Assimilation of Remotely Sensed Soil Moisture and Snow Depth Retrievals for Drought Estimation. J. Hydrometeorol. 15 (6), 2446–2469. https://doi.org/10.1175/JHM-D-13-0132.1.

Kumar, S.V., Peters-Lidard, C.D., Santanello, J.A., Reichle, R.H., Draper, C.S., Koster, R.D., Nearing, G., Jasinski, M.F., 2015. Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. Hydrol. Earth Syst. Sci. 19 (11), 4463–4478. https://doi.org/10.5194/hess-19-4463-2015.

Kwon, Y., Kumar, S., Navari, M., Mocko, D., Kemp, E., Wegiel, J., Geiger, J., Bindlish, R., 2022. Irrigation characterization improved by the direct use of SMAP soil moisture anomalies within a data assimilation system. Environ. Res. Lett. 17 https://doi.org/10.1088/1748-9326/ac7f49.

Lawston, P.M., Santanello Jr., J.A., Franz, T.E., Rodell, M., 2017a. Assessment of irrigation physics in a land surface modeling framework using non-traditional and human-practice datasets. Hydrology and Earth System Sciences 21 (6), 2953–2966. https://doi.org/10.5194/hess-21-2953-2017.

Lawston, P.M., Santanello Jr, J.A., Kumar, S.V., 2017b. Irrigation Signals Detected From SMAP Soil Moisture Retrievals. Geophys. Res. Lett. 44 (23), 11860–11867. https://doi.org/10.1002/2017GL075733.

Lee, H., Zehe, E., Sivapalan, M., 2007. Predictions of rainfall-runoff response and soil moisture dynamics in a microscale catchment using the CREW model. Hydrol. Earth Syst. Sci. 11 (2), 819–849. https://doi.org/10.5194/hess-11-819-2007.

Lievens, H., Tomer, S.K., Al Bitar, A., De Lannoy, G.J.M., Drusch, M., Dumedah, G., Hendricks Franssen, H.-J., Kerr, Y.H., Martens, B., Pan, M., Roundy, J.K., Vereecken, H., Walker, J.P., Wood, E.F., Verhoest, N.E.C., Pauwels, V.R.N., 2015. SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. Remote Sens. Environ. 168, 146–162. https://doi.org/10.1016/j.rse.2015.06.025.

Liu, X., Feng, X., Fu, B., 2020. Changes in global terrestrial ecosystem water use efficiency are closely related to soil moisture. Sci. Total Environ. 698, 134165 https://doi.org/10.1016/j.scitotenv.2019.134165.

Liu, Y., Peters-Lidard, C.D., Kumar, S., Foster, J.L., Shaw, M., Tian, Y., Fall, G.M., 2013. Assimilating satellite-based snow depth and snow cover products for improving snow predictions in Alaska. Adv. Water Resour. 54, 208–227. https://doi.org/10.1016/j.advwatres.2013.02.005.

Loew, A., Stacke, T., Dorigo, W., de Jeu, R., Hagemann, S., 2013. Potential and limitations of multidecadal satellite soil moisture observations for selected climate model evaluation studies. Hydrol. Earth Syst. Sci. 17 (9), 3523–3542. https://doi.org/10.5194/hess-17-3523-2013.

Maggioni, V., Anagnostou, E.N., Reichle, R.H., 2012. The impact of model and rainfall forcing errors on characterizing soil moisture uncertainty in land surface modeling. Hydrol. Earth Syst. Sci. 16 (10), 3499–3515. https://doi.org/10.5194/hess-16-3499-2012.

Magotra, B., Prakash, V., Saharia, M., Getirana, A., Kumar, S., Pradhan, R., Dhanya, C.T., Rajagopalan, B., Singh, R.P., Pandey, A., Mohapatra, M., 2024. Towards an Indian land data assimilation system (ILDAS): A coupled hydrologic-hydraulic system for water balance assessments. J. Hydrol. 629, 130604 https://doi.org/10.1016/j.jhydrol.2023.130604.

Mathur, A.S., Das, S., Sircar, S., 2006. Status of Agriculture in India: Trends and Prospects. Econ. Pol. Wkly 41 (52), 5327–5336.

Meier, J., Zabel, F., Mauser, W., 2018. A global approach to estimate irrigated areas – a comparison between different data and statistics. Hydrol. Earth Syst. Sci. 22 (2), 1119–1133. https://doi.org/10.5194/hess-22-1119-2018.

Mitchell, K.E., Lohmann, D., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Cosgrove, B.A., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Lettenmaier, D.P., Marshall, C.H., Entin, J.K., Pan, M., Shi, W., Koren, V., Bailey, A.A., 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. J. Geophys. Res. Atmos. 109 (D7) https://doi.org/10.1029/2003JD003823.

Mladenova, I.E., Bolten, J.D., Crow, W., Sazib, N., Reynolds, C., 2020. Agricultural Drought Monitoring via the Assimilation of SMAP Soil Moisture Retrievals Into a Global Soil Water Balance Model. Frontiers in Big Data 3. https://www.frontiersin.org/articles/10.3389/fdata.2020.00010.

Nie, W., Zaitchik, B.F., Rodell, M., Kumar, S.V., Anderson, M.C., Hain, C., 2018. Groundwater Withdrawals Under Drought: Reconciling GRACE and Land Surface Models in the United States High Plains Aquifer. Water Resour. Res. 54 (8), 5282–5299. https://doi.org/10.1029/2017WR022178.

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., 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 (D12) https://doi.org/10.1029/2010JD015139.

O'Neill, P. E. S. Chan, E. G. Njoku, T. Jackson, R. Bindlish, and J. Chaubell, (2021). *SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 8* [dataset]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/OMHVSRGFX38O. Date Accessed 23-08-2023.

O'Neill, P., Bindlish, R., Chan, S., Chaubell, J., Colliander, A., Njoku, E., & Jackson, T. (n. d.). *Algorithm Theoretical Basis Document Level 2 & 3 Soil Moisture (Passive) Data Products*.

Ozdogan, M., Rodell, M., Beaudoing, H., Toll, D., 2010. Simulating the Effects of Irrigation over the United States in a Land Surface Model Based on Satellite-Derived Agricultural Data. J. Hydrometeorol. 11 https://doi.org/10.1175/2009JHM1116.1.

Pai, D., Sridhar, L., Rajeevan, M., Sreejith, O., Satbhai, N. S., & Mukhopadhyay, B. (2014). *Development of a new high spatial resolution (0.25° × 0.25°) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region.* https://www.semanticscholar.org/paper/Development-of-a-new-high-spatial-resolution-(0.25%C2%B0-Pai-Sridhar/f24d17a376ec4d104024840c33092b5334bc3740.

Pandey, D.K., Putrevu, D., Misra, A., 2021. Chapter 10—Large-scale soil moisture mapping using Earth observation data and its validation at selected agricultural sites over Indian region. In: Srivastava, P.K., Gupta, M., Tsakiris, G., Quinn, N.W. (Eds.), Agricultural Water Management. Academic Press, pp. 185–207. https://doi.org/10.1016/B978-0-12-812362-1.00010-2.

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., Sheffield, J., 2007. High-performance Earth system modeling with NASA/GSFC's Land Information System. Innov. Syst. Softw. Eng. 3 (3), 157–165. https://doi.org/10.1007/s11334-007-0028-x.

Pignotti, G., Crawford, M., Han, E., Williams, M.R., Chaubey, I., 2023. SMAP soil moisture data assimilation impacts on water quality and crop yield predictions in watershed modeling. J. Hydrol. 617, 129122 https://doi.org/10.1016/j.jhydrol.2023.129122.

Reichle, R.H., Koster, R.D., 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. Geophys. Res. Lett. 32 (2) https://doi.org/10.1029/2004GL021700.

Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic Data Assimilation with the Ensemble Kalman Filter. Mon. Weather Rev. 130 (1), 103–114. https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2.

Reichle, R.H., Koster, R.D., Liu, P., Mahanama, S.P.P., Njoku, E.G., Owe, M., 2007. Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). J. Geophys. Res. Atmos. 112 (D9) https://doi.org/10.1029/2006JD008033.

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., Toll, D., 2004. The Global Land Data Assimilation System. Bulletin of the American Meteorological Society 85 (3), 381–394. https://doi.org/10.1175/BAMS-85-3-381.

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 2010. Investigating soil moisture–climate interactions in a changing climate: A review. Earth Sci. Rev. 99 (3), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004.

Seo, E., Lee, M.-I., Reichle, R.H., 2021. Assimilation of SMAP and ASCAT soil moisture retrievals into the JULES land surface model using the Local Ensemble Transform Kalman Filter. Remote Sens. Environ. 253, 112222 https://doi.org/10.1016/j.rse.2020.112222.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference On* -, 517–524. https://doi.org/10.1145/800186.810616.

Shrestha, A., Nair, A.S., Indu, J., 2020. Role of precipitation forcing on the uncertainty of land surface model simulated soil moisture estimates. J. Hydrol. 580, 124264 https://doi.org/10.1016/j.jhydrol.2019.124264.

Siebert, S., Döll, P., Hoogeveen, J., Faures, J.-M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. Hydrol. Earth Syst. Sci. 9 (5), 535–547. https://doi.org/10.5194/hess-9-535-2005.

Tao, F., Yokozawa, M., Hayashi, Y., Lin, E., 2003. Changes in agricultural water demands and soil moisture in China over the last half-century and their effects on agricultural production. Agric. For. Meteorol. 118 (3), 251–261. https://doi.org/10.1016/S0168-1923(03)00107-2.

Teluguntla, P., Thenkabail, P., Oliphant, A., Gumma, M., Aneece, I., Foley, D., & McCormick, R. (2023). *Landsat-Derived Global Rainfed and Irrigated-Cropland Product 30 m V001* . [object Object]. https://doi.org/10.5067/COMMUNITY/LGRIP/LGRIP30.001.

Upadhyaya, D. B., Evans, J., Muddu, S., Tomer, S. K., Al Bitar, A., Yeggina, S., S, T., Morrison, R., Fry, M., Tripathi, S. N., Mujumdar, M., Goswami, M., Ganeshi, N., Nema, M. K., Jain, S. K., Angadi, S. S., & Yenagi, B. S. (2021). The Indian COSMOS Network (ICON): Validating L-Band Remote Sensing and Modelled Soil Moisture Data Products. *Remote Sensing*, *13*(3), Article 3. https://doi.org/10.3390/rs13030537.

Wanders, N., Karssenberg, D., de Roo, A., de Jong, S.M., Bierkens, M.F.P., 2014. The suitability of remotely sensed soil moisture for improving operational flood forecasting. Hydrol. Earth Syst. Sci. 18 (6), 2343–2357. https://doi.org/10.5194/hess-18-2343-2014.

Wang, L., Qu, J.J., 2009. Satellite remote sensing applications for surface soil moisture monitoring: A review. Front Earth Sci. China 3 (2), 237–247. https://doi.org/10.1007/s11707-009-0023-7.

Yang, Z.-L., Niu, G.-Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. J. Geophys. Res. Atmos. 116 (D12) https://doi.org/10.1029/2010JD015140.

Yin, J., Zhan, X., Barlage, M., Kumar, S., Fox, A., Albergel, C., Hain, C.R., Ferraro, R.R., Liu, J., 2023. Assimilation of Blended Satellite Soil Moisture Data Products to Further Improve Noah-MP Model Skills. J. Hydrol. 621, 129596 https://doi.org/10.1016/j.jhydrol.2023.129596.

Yuan, X., Wood, E.F., Luo, L., Pan, M., 2011. A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. Geophys. Res. Lett. 38 (13) https://doi.org/10.1029/2011GL047792.

Zeng, R., Yao, F., Zhang, S., Yang, S., Bai, Y., Zhang, J., Wang, J., Wang, X., 2021. Assessing the effects of precipitation and irrigation on winter wheat yield and water productivity in North China Plain. Agric Water Manag 256, 107063. https://doi.org/10.1016/j.agwat.2021.107063.

Zhang, G., Liu, C., Xiao, C., Xie, R., Ming, B., Hou, P., Liu, G., Xu, W., Shen, D., Wang, K., Li, S., 2017. Optimizing water use efficiency and economic return of super high yield spring maize under drip irrigation and plastic mulching in arid areas of China. Field Crop Res 211, 137–146. https://doi.org/10.1016/j.fcr.2017.05.026.