https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Towards the development of a national drought monitoring framework for India: Reconstruction of historical droughts using CLM5

Devavat Chiru Naik[1], Chandrika Thulaseedharan Dhanya[2*], Harrie-Jan Hendricks Franssen[3]

[1]Department of Civil and Environmental Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India. email: Devavat.Chiru.Naik@civil.iitd.ac.in

[2]Department of Civil and Environmental Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India. email: dhanya@civil.iitd.ac.in

[3]Agrosphere (IBG-3), Forschungszentrum Juelich GmbH, Juelich, Germany. email: h.hendricks-franssen@fz-juelich.de

***Corresponding author:**

Chandrika Thulaseedharan Dhanya, Professor, Department of Civil and Environmental Engineering, Indian Institute of Technology Delhi, New Delhi-110016, India; Email: dhanya@civil.iitd.ac.in; Office Phone Number: +91-11-26597328.

Highlights

- CLM5 captures the spatial variability of soil moisture, evapotranspiration, runoff and successfully reconstructs major historical drought events across India.
- Atmospheric forcing strongly controls hydrological partitioning: IMDAA forcing reduces soil moisture and runoff RMSD by 10% and 2.5%, respectively, whereas ERA5 reduces evapotranspiration RMSD by 17%.
- Irrigation increases soil moisture, evapotranspiration, and runoff but improves performance only locally, reducing errors in ~26% and ~23% of irrigated grids for root-zone and surface soil moisture, and in ~33% and ~45% for runoff and evapotranspiration, while rainfed simulations outperform irrigated configurations across most regions.
- Three distinct climate-zone-dependent biases observed, including persistent dry bias in humid monsoon regions, storage surplus in semi-arid and arid zones, and a runoff paradox in temperate regions.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

- These findings highlight the need for improved irrigation parameterization, targeted model calibration, and careful selection of meteorological forcing datasets to support a national drought monitoring framework for India.

Abstract

In this study, we evaluate the Community Land Model version 5.0 (CLM5.0) as a first step toward building a National Drought Monitoring Framework – DRISHTI (**D**rought **R**isk & **I**mpact **S**urveillance using **H**ydrometeorological, **T**errestrial & **I**nference Models) in India. CLM5.0 was applied over India at 0.1° resolution (1980–2020), using India Meteorological Department (IMD) precipitation and two atmospheric forcing datasets: the Indian Monsoon Data Assimilation and Analysis (IMDAA) regional reanalysis and the fifth generation ECMWF Re-Analysis (ERA5). Model performance was assessed for soil moisture (SM), evapotranspiration (ET), and runoff under rainfed and irrigated conditions against in-situ measurements and satellite/reference datasets: Soil Moisture Active Passive (SMAP), Global Land Evaporation Amsterdam Model (GLEAM), and Global Runoff Reconstruction (GRUN).

CLM5.0 reproduced spatial and temporal hydrological variability well. Model diagnostics revealed three climate-zone biases: (1) an apparent 'Dry Bias' in humid zones (Am) characterized by low SM, runoff, and ET. This bias reflects underestimation of precipitation in IMD rather than model process deficiencies; (2) a 'Wet Bias' in semi-arid and arid zones (BSh, BWh), where the Dry Surface Layer scheme suppresses evaporation, generating excess storage and runoff; and (3) a 'Runoff Paradox' in temperate zones (Cwa, Cwb), where high surface SM coincides with low runoff. Systematic runoff underestimation in complex terrains reflects limitations in sub-grid hillslope lateral flow representation, which led to high moisture retention. Relative to ERA5-driven simulations, IMDAA reduces RMSD in SM and runoff by 10% and 2.5%, respectively, while ERA5 improves ET estimates by 17%. At the national scale, rainfed simulations generally outperform irrigated runs; however, irrigation leads to localized improvements - most notably over the Indo-Gangetic Plain - with a stronger influence on root-zone SM than on surface SM. Specifically, irrigation reduces errors over ~26% of irrigated grid cells for root-zone SM compared to ~23% for surface SM, while corresponding improvements are observed over ~33% and ~45% of irrigated grid cells for runoff and ET, respectively. Despite these limitations, CLM5.0 successfully reproduces historical droughts, demonstrating its utility for drought monitoring in data-scarce regions while highlighting the need for

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

improved irrigation parameterization, targeted model calibration, and careful selection of meteorological forcing datasets.

Keywords: Droughts, Land Surface Modeling, India, Soil Moisture, Evapotranspiration, Runoff.

1. Introduction

Climate change in recent decades has increased the frequency of short-term extreme wet spells and prolonged dry spells across India (Vinnarasi and Dhanya, 2016), accompanied by a reduction in light precipitation days (Dash et al., 2011; Mishra and Liu, 2014). This has led to more frequent severe and extreme drought events, with the percentage of such events rising over time (Chiru Naik et al., 2023), significantly affecting agriculture, water resources, and millions of lives. Accurate estimation and regular assessment of historical drought characteristics are essential for mitigating these socio-economic impacts and aiding policymakers in planning for future resilience (Svoboda et al., 2002; Xu et al., 2020; Zink et al., 2016). While in-situ measurements from specific stations provide the most accurate data on land surface variables, the sparse and uneven distribution of these stations limits the accuracy of such analyses. To overcome this challenge, satellite-based data and Land Surface Models (LSMs) provide valuable alternatives. However, not all land surface processes can be directly observed using satellites - for example, runoff - and while satellite-based soil moisture (SM) products offer continuous global coverage, they are generally limited to the surface layer due to shallow sensor penetration and orbital revisit constraints. LSMs are thus essential for estimating key variables such as runoff, evapotranspiration (ET), and SM at multiple depths and temporal scales, delivering spatially consistent and continuous outputs from regional to global domains (Lawrence et al., 2019a). Obtaining these variables daily is particularly important for understanding land–atmosphere interactions and assessing their impacts on the water cycle and agricultural systems (Xu et al., 2020; Xu et al., 2023).

While mature national and continental-level drought monitoring frameworks - including the North American Land Data Assimilation System (NLDAS), European Drought Observatory (EDO), China Land Data Assimilation System (CLDAS), and the Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS) - utilize integrated land surface models, satellite data, and observations (Cammalleri et al., 2021; Hao et al., 2012, 2016; Liu et al., 2021; McNally et al., 2017; Wang and Li, 2020), such a fully integrated

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

national framework is currently lacking over India. To address this gap, we are developing a National Drought Monitoring Framework - DRISHTI (Drought Risk & Impact Surveillance using Hydrometeorological, Terrestrial & Inference Models) - over India. As a first step, we employ the Community Land Model (CLM), an advanced land surface model integrated within the Community Earth System Model (CESM) framework developed by the National Center for Atmospheric Research (NCAR) (Lawrence et al., 2011; Lawrence et al., 2019a). Over time, CLM has evolved to version 5, with significant enhancements in its capabilities (Brunke et al., 2016; Lawrence et al., 2011; Oleson et al., 2008). It is widely used to enhance our understanding of the terrestrial energy, water, carbon and nitrogen cycles (Lawrence et al., 2011; Lawrence et al., 2019a) and their interactions (Green et al., 2019; Koven et al., 2017; McGuire et al., 2018; Yang et al., 2019), as well as the impact of land use and land cover change on climate, water, carbon, nitrogen and extreme events (Ghodichore et al., 2022; Lawrence & Chase, 2010; Mahowald et al., 2017; McGuire et al., 2018). CLM's ability to simulate global energy, water, carbon and nitrogen budgets has earned its wide recognition, and it serves as the land component in models such as the Norwegian Earth System Model (NorESM) (Seland et al., 2020) and the Earth System Model at the Euro-Mediterranean Center on Climate Change (CMCC-ESM2) (Lovato et al., 2022). These models contributed significantly to the Coupled Model Intercomparison Project Phase 5 (CMIP5) and CMIP6 (Cherchi et al., 2019).

As LSMs have become more complex in their parameterization, it is crucial to conduct comprehensive validations to assess their capabilities, applications, and provide references for future model improvements. For instance, Decker & Zeng (2009) noted that a modified version of Richards's equation improved the representation of SM in CLM3.5 compared to the standard approach. Replacing uniform soil layers with spatially explicit soil thickness, as demonstrated by Pelletier et al. (2016) and implemented in CLM4.5 (Brunke et al., 2016), significantly enhanced the simulation of water and energy cycles. Additionally, Swenson & Lawrence (2015) highlighted the sensitivity of water storage dynamics to soil layer depth in CLM4.5. A multi-model comparison study on groundwater (Rashid et al., 2019) underscored the importance of improving the representation of subsurface hydrological processes in LSMs, including CLM4. While systematic evaluations have been conducted in countries like the US and China (Cheng et al., 2021; Ma and Wang, 2022), the effectiveness of CLM5 in accurately simulating water and energy cycles over India, where local land surface changes are especially significant due to sparse in-situ data, has yet to be evaluated.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

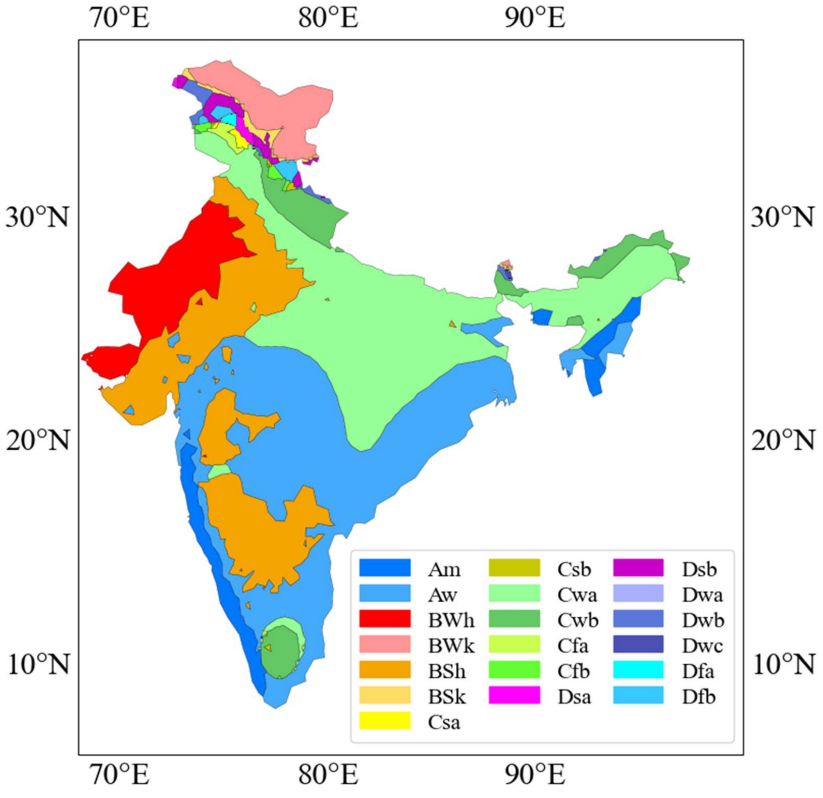
The analysis of CLM5 performance is therefore the necessary foundational step toward establishing the envisioned DRISHTI framework. This future operational framework will use the model's identified limitations to guide its focus on improving drought monitoring robustness by optimally merging satellite and in-situ observations with model simulations to support national water resource management. This study aims to address two critical questions: How effectively does CLM5 simulate hydrological variables in irrigated versus rainfed conditions over India, and how sensitive is it to atmospheric forcing? Additionally, how well does CLM5 simulate historical drought events? To answer these questions, land surface processes are simulated using CLM5 at a high resolution of 0.1° (~10 km) from 1980 to 2020 across India. A comprehensive validation is conducted by comparing simulation results with high-quality data from in situ observations, remote sensing, and re-analysis products at various temporal scales (daily, monthly, and interannual). The anticipated outcome of this study is to provide valuable insights into the capabilities and limitations of CLM5 simulations over India, highlighting areas for improvement and enhancing confidence in simulating future extreme events under climate change scenarios. The structure of the study is as follows: Section 2 provides a brief description of the study area, Section 3 details the methodology, Section 4 model setup and input data, Section 5 model evaluation and drought assessment, Section 6 offers results and discussion, and Section 7 concludes the study.

2. Study area

This study is conducted over the Indian region, which features diverse topography and a variety of climatic zones, ranging from the lush evergreen forests and snow-clad Himalayan mountains to the arid Thar Desert and plains. The different climate zones are classified based on long-term observed precipitation and temperature data, using the Köppen–Geiger climate classification (Peel et al., 2007). The key climate zones in India include tropical wet and dry (savanna) (Aw) climate, tropical monsoon (Am) climate, hot desert (BWh) climate, hot steppe (BSh) climate, dry winter humid subtropical (Cwa), and dry winter maritime temperature (Cwb), which cover the majority of the country and form the focus of this study. The Am region, which includes the western coast, experiences substantial rainfall due to its orographic features. The Aw climate zone spans central and southern India, characterized by croplands and fallow lands, receiving moderate to heavy rainfall. The BSh zone covers the drier regions of Western and Southern India, with low to moderate rainfall, while the Thar Desert, part of the BWh zone,

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

158 experiences very low rainfall. The Cwa and Cwb zones, covering central, northeastern, and
159 eastern India, including parts of the Ganga and Brahmaputra basins, form extensive agricultural
160 regions with moderate to heavy rainfall (Upadhyaya and Ramsankaran, 2018).



161

162 **Figure 1:** Spatial distribution of Köppen-Geiger climate classification zones across India.

163

164 **3 Methodology**

165 **3.1 Model description**

166 CLM is the land component of the CESM. The latest version, CLM5, effectively simulates
167 biophysical and biogeochemical processes, including interactions with the atmosphere. It
168 represents sub-grid scale heterogeneity by categorizing land into five distinct types: vegetated,
169 lake, urban, glacier, and crop. Each of these units consists of multiple snow/soil columns, which
170 can be occupied by different plant functional types (PFTs) (Lawrence et al., 2019b). Sub-model

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

schemes are defined for each land unit type to calculate the water and energy fluxes for individual grid cells.

In CLM5.0, soil thickness ranges from 0 to 8.5 meters, divided into 20 hydrologically active layers, with increasing depth from top to bottom. Water movement in the soil (unsaturated zone) is assumed to be one-dimensional, following the Richards equation, i.e., vertically towards the groundwater table, without interaction between parallel soil columns (Kim and Mohanty, 2016). Soil water dynamics are modeled using a multi-layer approach, where vertical SM transport is influenced by infiltration, surface and subsurface runoff, gradient diffusion, gravity, and canopy transpiration through root extraction (Lawrence et al., 2019a). The vegetation units in CLM5 are represented by 15 PFTs plus bare land. These PFTs differ in vegetation physiology and structure and are derived from MODIS satellite data at a 0.05° spatial resolution (Bonan et al., 2002; Dagon et al., 2020; Lawrence and Chase, 2007). Vegetation characteristics such as the yearly cycles of leaf area index (LAI), stem area index (SAI), and canopy height are defined for each PFT (Myneni et al., 2002; Zeng et al., 2002). CLM5 utilizes a multi-layer canopy model that simulates fluxes using combined information about plant canopy structure, leaf physiology, and radiative transfer (Bonan et al., 2014). The canopy is divided into multiple leaf layers with sunlit and shaded fractions, where photosynthesis, stomatal conductance, leaf temperature, and energy balance are coupled at each layer. The total ET in CLM5 is the sum of transpiration from vegetation, evaporation from the ground, and evaporation from water stored in the canopy (canopy evaporation), which is influenced by both wet leaves and transpiration (Lawrence et al., 2007). A simple TOPMODEL-based runoff model (SIMTOP), as described by Niu et al. (2005) is implemented to parameterize runoff in CLM5. The maximum fractional saturated area, slope, and elevation, derived from the USGS HYDRO1K 1-km dataset (microtopography), significantly influence surface water storage and outflow processes.

4. Model Setup and Input data

4.1 Model Setup

CLM incorporates complex biogeochemical processes, requiring an extensive spin-up period of several hundred years to achieve equilibrium in the carbon and nitrogen cycles (BGC), which demands significant computational resources (Lawrence et al., 2019b). Notably, CLM5BGC has been shown to consistently underestimate Gross Primary Productivity (GPP) and LAI, leading to lower ET values (Cheng et al., 2021) and highlighting the need to adjust dynamic

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

vegetation planting dates regionally for accurate LAI simulation (Reddy et al., 2025). To enhance computational efficiency and leverage observation-based constraints on vegetation, this study uses the Satellite Phenology (SP) mode. In SP mode, vegetation states - including LAI, SAI, and canopy height - are prescribed using MODIS satellite data, while carbon and nitrogen biogeochemistry cycles remain inactive. This approach avoids the prognostic uncertainties and documented underestimations of GPP and LAI found in CLM5-BGC (Cheng et al., 2021; Li et al., 2022). Under this configuration, vegetation states are calculated daily through linear interpolation of the monthly satellite climatology (Lawrence and Chase, 2007). For the SP simulations, the LAI/SAI for each PFT is derived from a monthly climatology averaged over the 2003-2015 period, which is applied consistently across all simulation years, ensuring the same annual cycle of LAI/SAI. A 26-year spin-up period was conducted to achieve equilibrium, with these equilibrium values serving as the initial conditions for the final simulations.

The standalone CLM5 is configured to run over the Indian domain ($8^0\ N$ to $38^0\ N$, $68^0\ E$ to $98^0\ E$, 360×360 grid cells) at a spatial resolution of 0.1^0 (~10 km) for the period of 1980 to 2020.

4.2 Input data

Essential land surface parameters, including the fractions of different land unit types (lake, glacier, urban, natural vegetation, and crops) within grid cells, as well as soil properties (such as soil colour, texture, and organic matter density), are derived from high-resolution datasets. These datasets are compiled from various sources, including the International Geosphere-Biosphere Program (IGBP) and the Global Land One-km Base Elevation Project (Lawrence et al., 2019a). Land cover information, such as the percentage of PFTs, is derived from a combination of datasets, including the 2001 MODIS Vegetation Continuous Field (VCF), MODIS land cover product, and the 1992–1993 AVHRR Continuous Field Tree Cover Project data (Lawrence and Chase, 2007). The maximum fractional saturated area, slope, and elevation are obtained from the United States Geological Survey (USGS) HYDRO1K dataset (Verdin & Greenlee, 1996) and processed using the algorithm described in Niu et al. (2005). For further details on CLM5 input requirements, readers are directed to the detailed technical description of CLM5 (Lawrence et al., 2019a).

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

To evaluate the sensitivity of CLM5 to atmospheric forcings, two distinct datasets were selected: a regionally tailored dataset and a globally representative one. The Indian Monsoon Data Assimilation and Analysis (IMDAA) reanalysis, developed as part of the National Monsoon Mission (NMM) by the Ministry of Earth Sciences, Government of India, in collaboration with the National Centre for Medium-Range Weather Forecasting (NCMRWF), the India Meteorological Department (IMD), and the UK Met Office, provides a high-resolution (12-km) regional dataset for the Indian monsoon. This dataset utilizes a 4D-Var data assimilation method and the U.K. Met Office Unified Model, integrating both conventional and satellite observations, including Indian surface and upper-air data that had not been used in previous reanalyses (Ashrit et al., 2020; Indirarani et al., 2021; Mahmood et al., 2018; Singh et al., 2021). The ERA5 reanalysis, developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), serves as the successor to ERA-Interim, offering improved accuracy and consistency for climate modeling (Hersbach et al., 2020). While both IMDAA and ERA5 provide comprehensive atmospheric variables, precipitation fields from these datasets were replaced with the gridded precipitation (P) product from the IMD (Pai et al., 2014) to ensure consistency over India. The P data were generated by spatially interpolating station data from approximately 6,995 rain gauges across the country using the inverse distance weighting method (Shepard, 1968). The original resolutions of IMDAA ($0.12^0 \times 0.12^0$) and ERA5 ($0.25^0 \times 0.25^0$) were subsequently re-gridded to a common resolution of $0.1^0 \times 0.1^0$, which was used as the meteorological forcing for CLM5 simulations. In CLM5, cropland within each grid cell is partitioned into irrigated and rainfed soil columns, based on the global irrigation dataset from Portmann et al. (2010). Irrigation is applied to the soil surface beneath irrigated crops, with water drawn from river storage to ensure mass conservation. The model evaluates daily irrigation demand after 6 a.m. local time. Irrigation is triggered when the crop LAI is greater than 0 indicating active growth, and when the SM levels fall below a predefined threshold ensuring that crops receive adequate water during dry periods (Lawrence et al., 2019a).

To examine the effects of irrigation and meteorological forcing on land surface processes, two experiments were conducted using each atmospheric dataset:

- Experiment 1: Irrigation activated (irrigated scenario)
- Experiment 2: Irrigation deactivated (rainfed scenario)

All simulations used the same land cover, soil parameters, and initial conditions, ensuring a consistent basis for comparison. This resulted in four simulations:

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

- IMDAA-Irr and ERA5-Irr: Irrigated simulations forced with IMDAA and ERA5, respectively.
- IMDAA-Rfd and ERA5-Rfd: Rainfed simulations forced with IMDAA and ERA5, respectively.

This framework enables a robust assessment of CLM5 performance under different irrigation regimes and atmospheric forcings, while also evaluating its sensitivity to regional (IMDAA) versus global (ERA5) input data.

4.3 Observation data set

4.3.1 COSMOS SM

The Indian Cosmic Ray Network (ICON) operates the Cosmic-ray Soil Moisture Observation System (COSMOS) at seven sites across India (Upadhyaya et al., 2021). In this study, SM data from three locations: Berampudi (BMB), Madahalli (MDH), and Singanallur (SGR) - were used to evaluate the model's performance in simulating SM. Additional site details are provided in Table S1 and Figure S1 (see supplementary material).

4.3.2 Soil Moisture Active Passive (SMAP) Data

SMAP, launched in 2015, was the first mission dedicated to SM and freeze/thaw monitoring using both radar and radiometer at L-band, the radar component ceased functioning after three months, leaving only the passive radiometer operational (Entekhabi et al., 2010). From the radiometer, with a native footprint of ~39 × 47 km, SM estimates every 2–3 days at 3, 9, and 36 km resolution are derived, with spectral filtering applied to reduce radio-frequency interference. Retrievals are based on brightness temperature and supported by ancillary data such as land surface temperature, vegetation opacity, albedo, surface roughness, and soil texture (O'Neill et al., 2017). For this study, we used the Level 3 passive-only SMAP product at 36 km resolution from the ascending (AM) orbit, which aligns with CLM5's daily output referenced to 6:00 a.m. local time. SMAP data were sourced from NASA's APPEEARS platform.

4.3.3 GLEAM dataset

The Global Land Evaporation Amsterdam Model (GLEAM)(Miralles et al., 2011) employs distinct algorithms to estimate various components of terrestrial evaporation, including transpiration, interception loss, bare-soil evaporation, snow sublimation, and open-water

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

evaporation. GLEAM datasets, which provide estimates of ET and root-zone SM (RZSM), are widely used to analyze spatial variability, water cycle trends, and drought conditions (Gebrechorkos et al., 2023; Ma et al., 2023; Miralles et al., 2011; Zhang et al., 2016). These datasets have also been instrumental in evaluating CLM's ET and SM simulations, with several studies (Cheng et al., 2021; Ghodichore et al., 2022; Wang et al., 2021) reporting that GLEAM's performance outperforms MODIS ET when compared against in-situ data (Jia et al., 2022; Michel et al., 2016; Miralles et al., 2016). For this study, we used the latest GLEAM V4.2a ET and SM products, at spatial resolution of $0.1^0 \times 0.1^0$ to evaluate CLM5 ET and SM. The data can be accessed at GLEAM's website.

4.3.4 GRUN dataset

The Global Runoff Reconstruction (GRUN) dataset, developed using machine learning techniques, provides a global gridded reconstruction of monthly runoff rates on a 0.5° × 0.5° WGS84 grid for the period 1902 to 2014 (Ghiggi et al., 2019). It incorporates runoff observations from a global collection of in-situ streamflow data, primarily from relatively small catchments ($< 2500\ km^2$), and trains its algorithm using gridded precipitation and temperature data from century-long reanalysis as predictors. The GRUN dataset offers valuable insights into runoff variability, especially in regions with sparse observational data. The dataset is freely available in NetCDF-4 format and can be accessed at Figshare.

5 Model evaluation and Drought assessment

5.1 Performance evaluation

The performance of the CLM5 simulations was assessed by comparing the model outputs against in situ observations (COSMOS) and reference datasets, such as SMAP, GLEAM, and GRUN, which are used as observational benchmark in the absence of in-situ data. This study employed three statistical metrics to evaluate model performance: Root Mean Square Error (RMSE), Mean Error (ME), and Spearman's Rank Correlation Coefficient (ρ), the latter chosen for its robustness to the non-linear and non-Gaussian characteristics of hydrological variables, unlike the Pearson correlation which assumes linearity and normality (Helsel et al., 2020). For gridded products, which provide spatially continuous estimates rather than absolute ground

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

truth, differences are reported as Root Mean Square Deviation (RMSD) and Mean Deviation (MD). The corresponding equations are as follows:

$$RMSE/RMSD = \sqrt{\sum_{i=1}^{n} \frac{(x_i - y_i)^2}{n}} \quad (1)$$

$$ME/MD = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i) \quad (2)$$

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \quad (3)$$

where y_i refers to values (SM/ET/runoff) from reference datasets, x_i to simulated values by CLM5, n to the number of time steps (daily/monthly scale) and d_i to the difference between the ranks of y_i and x_i

To ensure consistency with reference datasets, CLM5 outputs (0.1^0 resolution) were upscaled using bilinearly interpolated to match the coarser grids of the reference products: ~36 km for SMAP, 0.1° for GLEAM, and 0.5^0 for GRUN. Temporal harmonization was applied to match each dataset's resolution: daily for COSMOS, and monthly for GLEAM, and GRUN. For in-situ comparisons, model outputs were extracted at the exact coordinates of COSMOS stations to evaluate SM at multiple depths.

5.2 Drought Indices

To evaluate meteorological, agricultural, and hydrological droughts, we used three standardized indices: the Standardized Precipitation Evapotranspiration Index (SPEI; Vicente-Serrano et al., 2010), the Standardised Soil Moisture Index (SSI; Hao & AghaKouchak, 2013), and Standardized Runoff Index (SRI; Shukla & Wood, 2008), respectively. P data was obtained from the IMD, and ET, SM, and runoff were taken from a single CLM5 configuration selected based on its overall performance, to maintain internal consistency of the water balance in the drought analysis. Drought-relevant variables - net precipitation (P – ET), SM, and runoff were computed for the monsoon season (June to September; JJAS) and the water year (June-May). These variables were fitted to suitable Probability Distribution Functions (PDFs), including gamma, log-logistic, GEV, Pearson Type-III, and log-normal distributions (Stagge et al., 2015; Vicente-Serrano et al., 2010), The Kolmogorov–Smirnov (KS) test was used to select the best-

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

fit PDF for each variable. From the selected PDFs, Cumulative Distribution Functions (CDFs) were calculated and transformed to a standard normal distribution to derive the drought indices. Drought severity was categorized using the U.S. Drought Monitor classification (Svoboda et al., 2002).

6 Results and Discussion

6.1 Evaluation of SM against COSMOS

The evaluation of simulated surface soil moisture (SSM) against COSMOS observations highlights a strong dependence on meteorological forcing and irrigation representation across stations (Figure 2; Table 1). IMDAA-driven simulations generally outperform ERA5-forced runs at Berampudi (BMB), Madahalli (MDH), and at Singanallur (SGR) both forcings perform similarly, with irrigated simulations exhibiting slightly lower RMSE than rainfed runs. This suggests that the regional reanalysis better captures the hydroclimatic variability governing near SSM dynamics, while irrigation effects at the station scale remain limited. Across all stations, simulations exhibit a systematic wet bias, which is amplified under irrigated conditions. This effect is most pronounced at BMB, where larger errors indicate that irrigation inputs in CLM5 likely exceed effective water inputs inferred from COSMOS observations. In contrast, SGR shows comparatively smaller biases, with near-zero mean error under irrigated conditions and slightly negative bias in rainfed simulations, indicating a more balanced SM response. MDH displays intermediate behavior between these two cases.

The relationship between error magnitude and correlation varies across stations, with larger errors at BMB coinciding with higher correlations, while SGR exhibits smaller biases but weaker correlations. Some discrepancies can be partly attributed to scale mismatches between COSMOS observations (~150–250 m footprint) and CLM5 simulations (~0.1° resolution), as well as to structural limitations in the model's soil hydraulic parameterization, which is based solely on texture and neglects soil structural features such as bio-pores and aggregation that strongly influence water retention and flow (Chaney et al., 2018; Fatichi et al., 2020).

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

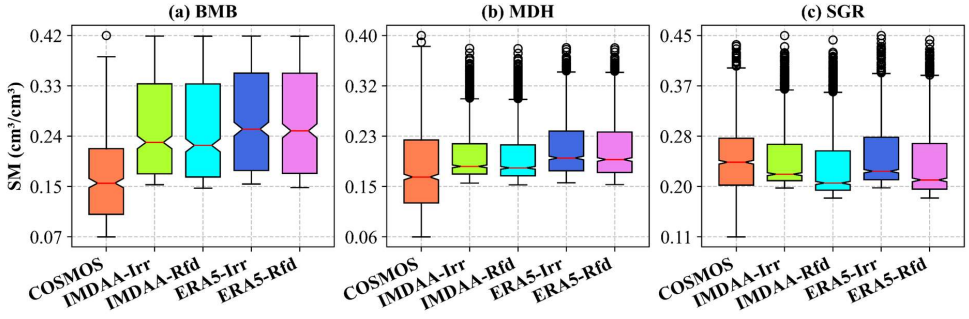


Figure 2. Comparison of CLM5SP-simulated SM at 5cm depth against COSMOS observations at three sites: (a) BMB, (b) MDH, and (c) SGR. In the boxplots, the central line represents the median, the colored box spans the interquartile range (25th–75th percentile), and the whiskers extend to the minimum and maximum values within 1.5×IQR; outliers beyond this range are shown as circles.

Table 1: Performance metrics (RMSE, ME, and ρ) for CLM5-simulated SSM against COSMOS observations.

Station Name	Metric	IMDAA-Irr	IMDAA-Rfd	ERA5-Irr	ERA5-Rfd
BMB	ME (cm³/cm³)	0.09	0.09	0.10	0.10
	RMSE (cm³/cm³)	0.12	0.12	0.13	0.13
	ρ	0.50	0.50	0.50	0.50
MDH	ME (cm³/cm³)	0.02	0.02	0.03	0.03
	RMSE (cm³/cm³)	0.07	0.07	0.08	0.07
	ρ	0.43	0.43	0.42	0.42
SGR	ME (cm³/cm³)	0.00	-0.01	0.01	-0.01
	RMSE (cm³/cm³)	0.06	0.07	0.06	0.07
	ρ	0.38	0.38	0.37	0.37

6.2: Grid-wise performance evaluation SM: CLM5 vs. Reference datasets

A grid-wise evaluation of CLM5 was first conducted for SM to capture spatial heterogeneity and assess the influence of atmospheric forcing and irrigation, providing context for the subsequent climate zone–wise analysis.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

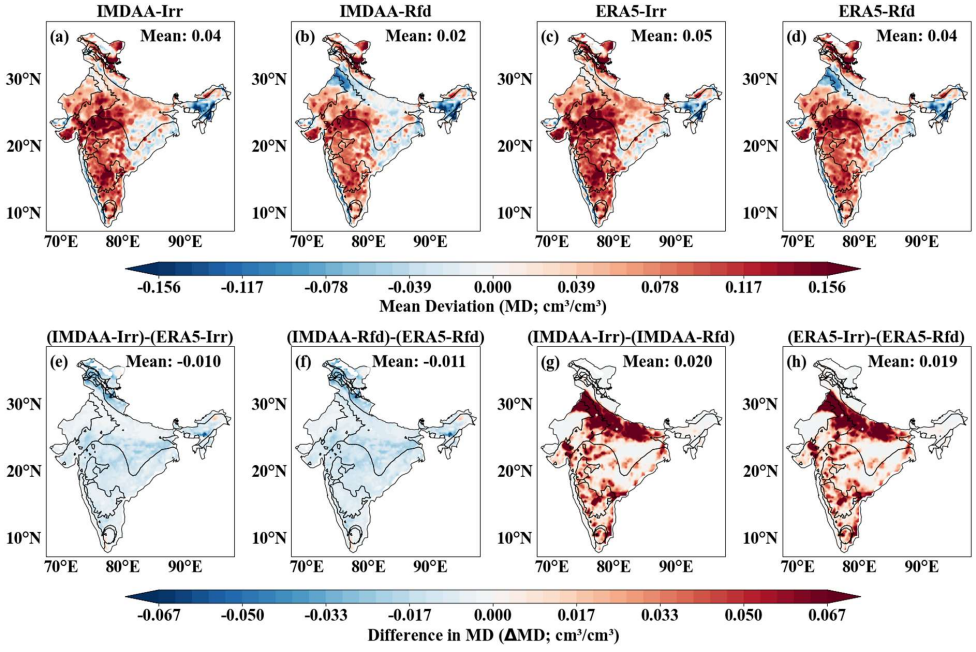


Figure 3: Grid-wise Mean Deviation (MD) of surface soil moisture (SSM, 5 cm) simulated by CLM5 relative to SMAP across India for 1980–2020. Panels (a–d) show MD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in MD (ΔMD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

The Mean Deviation (MD) maps (Figure 3a-d) reveal clear and systematic regional biases in CLM5 SSM. Overall, simulated SM is higher than SMAP estimates across northwest, central, peninsular, and parts of northern and northeastern India, indicating a widespread wet bias across all simulations. In contrast, negative biases dominate in wetter regions, particularly along the eastern and western coastal belts and portions of northeastern India. Notably, in rainfed simulations much of the Indo-Gangetic Plain exhibits enhanced dry bias, whereas in irrigated configurations wet biases are more pronounced and spatially extensive. Among the experiments, IMDAA-driven simulations exhibit the lowest overall MD, indicating improved bias representation relative to ERA5-driven runs (Figure 3e–f), and especially compared to irrigated configurations. Figures 3g–h further confirm that irrigation scenarios consistently reinforce wet biases over irrigated regions, reflecting an overestimation of SM.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

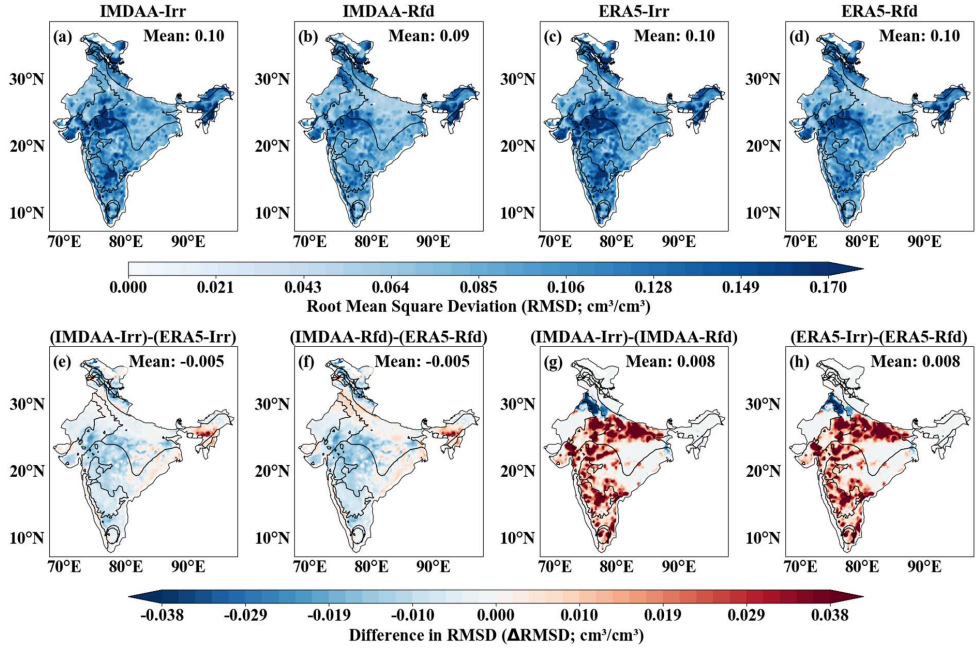


Figure 4: Grid-wise Root Mean Square Deviation (RMSD) of surface soil moisture (SSM, 5 cm) simulated by CLM5 relative to SMAP across India for 1980–2020. Panels (a–d) show RMSD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in RMSD (ΔRMSD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

To further quantify the magnitude of model bias, Figure 4 presents the grid-wise RMSD values for all simulations (subplots a–d). Mean RMSD is slightly lower for IMDAA-rainfed configuration (IMDAA-Rfd: 0.09) than for irrigated runs (IMDAA-Irr: 0.10; ERA5-Irr: 0.10), indicating overall better performance under rainfed conditions. Higher RMSD values (>0.11 $cm^3\ cm^{-3}$) persist across northeastern, central, southern, and northern regions, with errors generally more pronounced in irrigated configurations. Subplots e–h shows the spatial differences in RMSD (ΔRMSD) of SSM between IMDAA- and ERA5-driven simulations. Overall, IMDAA-based runs exhibit lower RMSD (ΔRMSD ≈ −0.005 $cm^3\ cm^{-3}$) across large parts of central and southern India, portions of northwestern India, and localized northern regions, reflecting improved performance relative to ERA5 forcing. In contrast, ERA5-driven simulations perform better along the southeastern coast, over northeastern India, and across parts of the Indo-Gangetic Plain. These spatial patterns are further quantified in Supplementary

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Figure S2. In non-irrigated areas, irrigated and rainfed simulations yield identical RMSD values because irrigation is not activated, resulting in overlapping percentages (e.g., IMDAA-Irr = IMDAA-Rfd; ERA5-Irr = ERA5-Rfd). Overall, IMDAA-Rfd achieves the lowest RMSD across 67.2% of grid cells, followed by IMDAA-Irr (33.2%), ERA5-Irr (20.4%), and ERA5-Rfd (17%). Focusing on irrigated grid cells (Figure 4g–h), localized RMSD reductions occur along the western Indo-Gangetic Plain, while higher RMSD dominates most other irrigated regions, as also shown in Supplementary Figure S3. Across irrigated areas, IMDAA-Rfd performs best in 65.3% of cases, particularly over central and peninsular India, whereas rainfed simulations consistently outperform irrigated runs over much of the Indo-Gangetic Plain. Overall, both MD and RMSD diagnostics consistently indicate that IMDAA-Rfd provides the most robust SSM performance across India, while irrigation yields only localized improvements and often degrades performance at larger scales. A likely cause is that CLM5 triggers irrigation whenever SM falls below a threshold and LAI > 0, directly adding water to the soil. This mechanism often triggers irrigation too early relative to actual demand (Dombrowski et al., 2024) and fails to account for diverse real-world application methods (e.g., canal, groundwater, flood, drip, and sprinkler) or variable timings (Yao et al., 2022). The correlation maps (Figure S4) show strong temporal agreement ($\rho > 0.6$) across most of India in all simulations. In contrast, lower correlations occur in the arid and northern part, likely reflecting the influence of snow cover, complex terrain, and limitations in SMAP retrievals.

While a standard 1.0 m depth was considered for root zone SM (RZSM) in model evaluation and agricultural drought analysis, this study accounted for grid cells where CLM5 simulates the soil columns due to bedrock constraints based on the spatially explicit soil thickness datasets (Pelletier et al., 2016).As shown in Figure S5 and in Table S2 (see supplementary material), CLM5 simulates the full 1.0 m depth for 81.9% of the grid cells, whereas reduced depth occur in specific regions: 0.8 m in 5.1% of cells (eastern Aw zone, western Am coast, and a few scattered locations), 0.58 m in 9.7% (Cwb zone and Jammu & Kashmir), and 0.4 m in 1.7% (parts of Jammu & Kashmir). These effective soil depths – were used consistently in all model – reference comparisons and drought assessments.

Figures S6–S9 present the grid-wise MD, RMSD, and the percentage of grid cells where CLM5 simulations achieve the lowest RMSD when evaluated against GLEAM across India. The surface soil moisture (10 cm) patterns are broadly consistent with the SMAP-based assessment and are therefore provided in the Supplementary Material. It is important to note that GLEAM

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

does not explicitly represent irrigation; however, as a data assimilation product incorporating ESA-CCI observations, irrigation signals are indirectly embedded within the dataset.

Figure S10 (subplots a–d) shows the MD patterns for RZSM, which largely mirror the bias structure observed for SSM. Positive biases dominate across northern, northwestern, and central India and much of the southern Indo-Gangetic Plain, while negative biases prevail in northeastern India and along the southeastern and southwestern coastal regions, as well as parts of the northern Indo-Gangetic Plain. Rainfed simulations (IMDAA-Rfd and ERA5-Rfd) generally exhibit smaller biases than irrigated configurations, with IMDAA-Rfd providing the most balanced performance. Overall, MD magnitudes are lower for RZSM than for SSM. Figure S19 presents the correlation between CLM5 simulations and GLEAM. High correlations (>0.8) are maintained across most of India for RZSM, though correlations are smaller in northern and northwestern India and extend into parts of southern India. These spatial patterns highlight ongoing challenges in simulating deeper SM dynamics in water-limited and high-altitude regions.

Figure S11 (subplots a–d) presents the grid-wise RMSD of RZSM. The spatial error patterns broadly resemble those of SSM, with higher RMSD concentrated over northern India, parts of central and southern India, and the northeastern region across all simulations. However, RMSD magnitudes are generally lower for RZSM than for SSM, reflecting the reduced temporal variability of deeper soil moisture and its weaker sensitivity to short-term precipitation fluctuations. Consistent with surface-layer results, IMDAA-driven simulations outperform ERA5-driven runs (Figure S11e-f) across much of central and southern India and portions of northwestern India, while ERA5-based simulations perform better along the southeastern coast, northeastern India, and parts of the Indo-Gangetic Plain. IMDAA-Rfd achieves the lowest mean RMSD for RZSM, performing best at 61.5% of grid cells (Figure S12). Over irrigated regions (Figure S13), it again dominates at 64.8% of grid cells, particularly across central and peninsular India and parts of the Indo-Gangetic Plain. Irrigated configurations provide improvements only in localized pockets mainly within the Indo-Gangetic Plain, northeastern India, and along the southeastern coast while ERA5-Irr contributes moderately within the Indo-Gangetic Plain (21.5%), and ERA5-Rfd shows limited influence. Although some irrigated grids exhibit reduced RMSD (Figure S11g–h), irrigation increases errors across most irrigated regions, consistent with the wet-bias reinforcement observed in SSM.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

6.3 Performance evaluation SM across major climate zones

Figure 5 presents the performance of CLM5-simulated SSM at 5 cm depth against SMAP observations across six major climate zones in India. The model generally exhibits a systematic wet bias (MD: 0.00 to 0.09 cm^3/cm^3), most pronounced in the arid and semi-arid regions (BWh, BSh). Conversely, all simulations consistently underestimate SSM in the Am zone (−0.05 to −0.06 cm^3/cm^3), identifying it as a unique region of dry bias. Across most climate zones, IMDAA-based simulations particularly the rainfed configuration (IMDAA-Rfd) outperformed ERA5, yielding the lowest RMSD (0.03–0.09 cm^3/cm^3) and maintaining robust correlations (0.75–0.93). However, regional exceptions were observed: in the arid BWh zone, IMDAA-Rfd performed comparably to ERA5-Rfd, while in the tropical monsoon (Am) zone, ERA5-based simulations achieved a lower RMSD than IMDAA.

Supplementary evaluation against GLEAM at 10 cm depth shows strong agreement with the SMAP-based patterns (Figure S14; Table S3). At the root-zone level (100 cm), comparisons with GLEAM further indicate that CLM5 effectively captures deeper SM dynamics (Figure S14; Table S4). Compared to SSM, RZSM exhibits consistently lower RMSD and MD, which is related to the fact that SM in deeper layers shows less temporal variability and is less affected by precipitation peaks. Bias directions largely persist from surface to RZSM, with systematic wet biases in the arid BSh and BWh zones and underestimation in the humid Am zone. In contrast, the temperate Cwb zone transitions from a surface-level wet bias to root-zone underestimation, reflecting climate-dependent differences in subsurface hydrological processes. Spearman correlations range from 0.63 to 0.95 across climate zones, indicating strong temporal consistency between simulations and observations. Overall, IMDAA-Rfd exhibits the best performance across most regions. However, ERA5-driven simulations achieve higher accuracy in the humid Am and temperate Cwb zones, with IMDAA-Irr also showing localized improvements in the Cwb region.

In the Am zone, ERA5-based simulations perform slightly better than others, as lower shortwave radiation (SWR: −23.17 W m^{-2}) and higher relative humidity (RH: +21.97%) reduce atmospheric demand and enhance soil water retention (Table S6; Figure S21). Although irrigation offers a modest improvement, its limited spatial extent leaves ERA5-Irr simulations still negatively biased. In Cwb, discrepancies likely arise from differences in effective root-zone depth: CLM5 includes the top four layers (~0.58 m), whereas GLEAM defines the root zone at ~1 m, permitting greater storage. In BSh and BWh zones, IMDAA-Rfd performs best,

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

as IMDAA's higher SWR and lower RH promote stronger evaporative demand, mitigating the model's moisture over-retention.

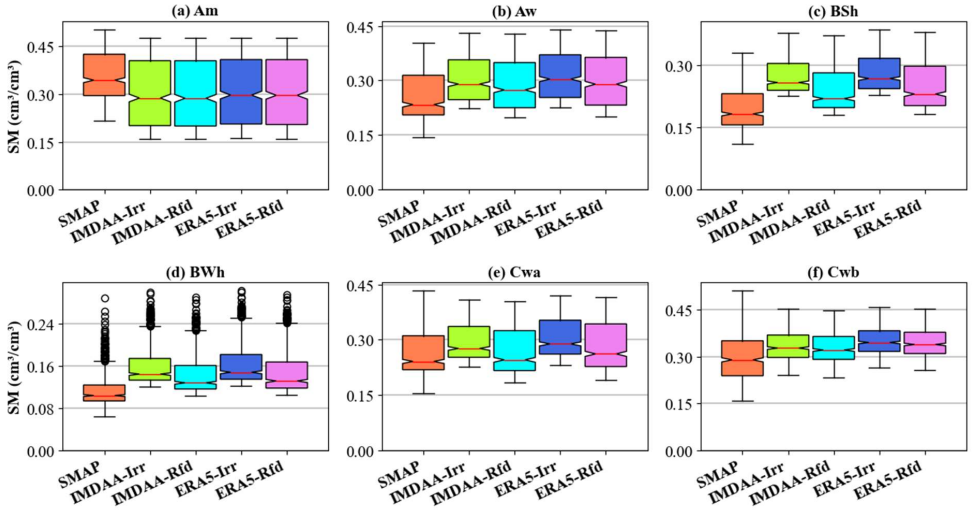


Figure 5: Comparison of CLM5-simulated SM against SMAP observations across climate zones: (a) Am, (b) Aw, (c) BSh, (d) BWh, (e) Cwa, and (f) Cwb. In the boxplots, the central line represents the median, the colored box spans the interquartile range (25th–75th percentile), whiskers extend to 1.5×IQR, and outliers are shown as circles.

Table 2: Performance metrics (RMSD, MD, and ρ) for CLM5-simulated SSM against SMAP observations.

Climate Zone	Metric	IMDAA-Irr	IMDAA-Rfd	ERA5-Irr	ERA5-Rfd
Am	MD (cm³/cm³)	-0.06	-0.06	-0.05	-0.05
	RMSD (cm³/cm³)	0.08	0.08	0.07	0.07
	ρ	0.87	0.87	0.87	0.87
Aw	MD (cm³/cm³)	0.05	0.03	0.06	0.04
	RMSD(cm³/cm³)	0.05	0.04	0.06	0.05
	ρ	0.93	0.93	0.92	0.92
BSh	MD (cm³/cm³)	0.08	0.05	0.09	0.06
	RMSD (cm³/cm³)	0.08	0.05	0.09	0.06
	ρ	0.92	0.92	0.93	0.92
BWh	MD (cm³/cm³)	0.05	0.03	0.05	0.03

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

	RMSD (cm^3/cm^3)	0.05	0.04	0.06	0.04
	ρ	0.76	0.76	0.76	0.76
Cwa	MD (cm^3/cm^3)	0.02	0.00	0.04	0.01
	RMSD (cm^3/cm^3)	0.04	0.03	0.05	0.04
	ρ	0.87	0.88	0.86	0.87
Cwb	MD (cm^3/cm^3)	0.04	0.03	0.05	0.05
	RMSD (cm^3/cm^3)	0.06	0.05	0.07	0.07
	ρ	0.79	0.79	0.75	0.75

Conversely, ERA5-driven runs - with lower atmospheric demand and added irrigation - intensify the wet bias. Overall, CLM5 shows consistent directional biases: underestimation in humid and highland zones (Am, Cwb) and overestimation in drier zones (BSh, BWh), irrespective of forcing or management scenario. This contrasts with earlier studies over the USA and China (Cheng et al., 2021; Ma and Wang, 2022), which reported RZSM underestimation alongside SSM overestimation. Here, both soil layers show bias in the same direction, and surface-layer errors are larger.

6.4 Evapotranspiration evaluation against GLEAM grid-wise

Mean Deviation (MD) patterns reveal spatially coherent biases linked to water management practices (Figure 6a–d). Widespread negative biases dominate across parts of northeastern, northern, northwest, and southern India and are stronger in rainfed simulations, reflecting lower ET associated with SM retention in CLM5. In contrast, irrigated configurations shift biases toward positive values in irrigated regions, where added water enhances ET relative to GLEAM. Differences in ET magnitude between forcing datasets are further illustrated in Figure 6e–f, with IMDAA-driven simulations generally producing higher ET than ERA5-based runs, while irrigation tends to increase MD across irrigated regions. Consistent with these patterns, ERA5-Rfd exhibits the smallest overall MD, remaining closest to the reference data. Temporal correlations (Figure S19) are strong across most of India, with rainfed simulations (ERA5-Rfd and IMDAA-Rfd) showing the highest agreement ($\rho > 0.8$), particularly over the Indo-Gangetic Plain and parts of northwestern and southern India.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

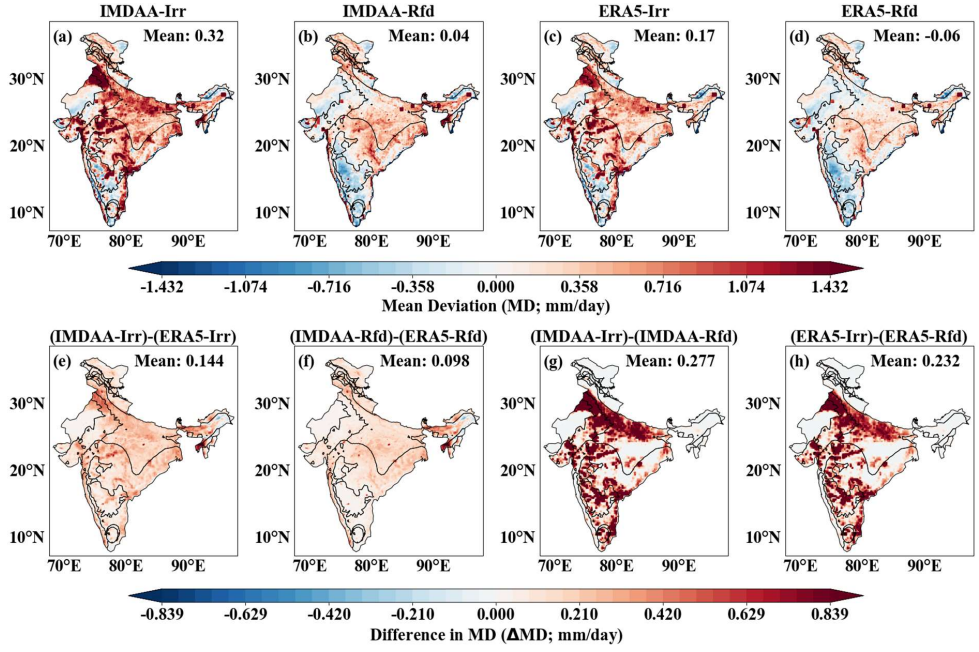


Figure 6: Grid-wise Mean Deviation (MD) of evapotranspiration (ET) simulated by CLM5 relative to GLEAM across India for 1980–2020. Panels (a–d) show MD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in MD (ΔMD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

Figure 7 (subplots a–d) presents the grid-wise RMSD of ET relative to GLEAM across India. Overall, ERA5-driven simulations consistently exhibit lower RMSD than IMDAA-driven runs (Figure 7e–f). This is further supported by Supplementary Figure S15, which shows ERA5-Rfd and ERA5-Irr achieving the best performance at 68.6% and 62.7% of grid cells, respectively. In contrast, IMDAA-based simulations outperform ERA5 only over limited regions, primarily in parts of northwestern India, the northeast, and small pockets of southern India (IMDAA-Rfd: 7.1%; IMDAA-Irr: 7.9%). The impact of irrigation on ET performance is illustrated in Figure 7g–h, where irrigation generally increases RMSD across irrigated grid cells. Localized improvements are observed in parts of southern India and in a few grids within the Indo-Gangetic Plain. A focused analysis over irrigated areas (Figure S16) confirms this behavior, with ERA5-Rfd and ERA5-Irr remaining closer to GLEAM over 54.4% and 43.4% of irrigated grid cells, respectively, while IMDAA-driven simulations account for less than 3%.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

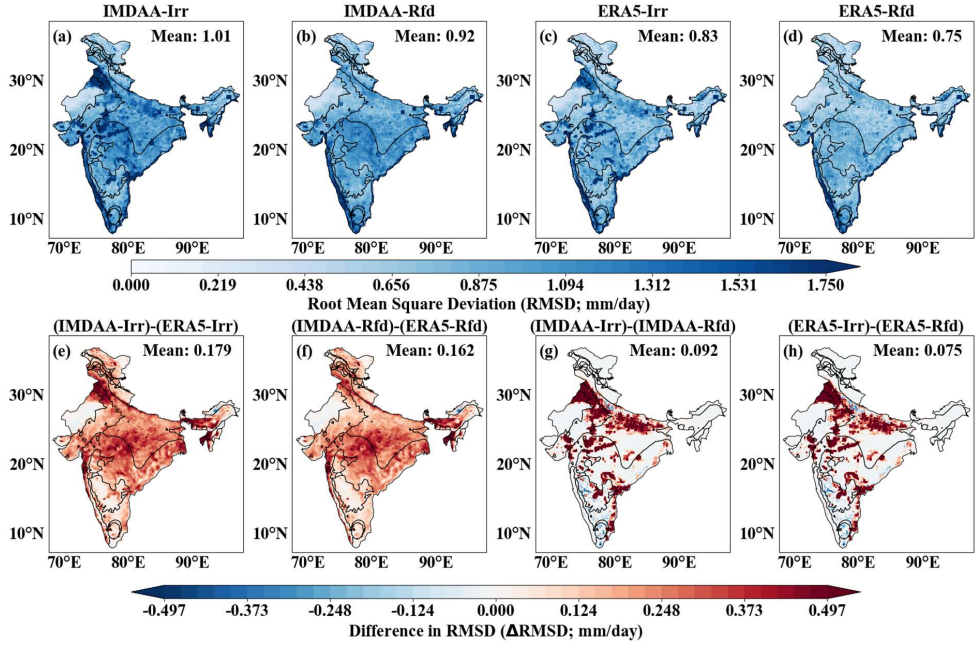


Figure 7: Grid-wise Root Mean Square Deviation (RMSD) of evapotranspiration (ET) simulated by CLM5 relative to GLEAM across India for 1980–2020. Panels (a–d) show RMSD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in RMSD (ΔRMSD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

This performance contrasts with the SM results, where IMDAA was superior. This divergence arises from differences in atmospheric demand driven primarily by radiation and humidity, as differences in mean values of other forcing variables were negligible (< 1; see Figure S20, supplementary material). IMDAA has significantly higher mean incoming SWR and substantially lower mean relative humidity compared to ERA5 (Table S6; supplementary material). In CLM5, this increased energy input and atmospheric demand resulted in a significantly higher mean simulated net radiation (Rnet: ~+17.87 W/m²; Figure S21, supplementary material). This surplus energy enhances ET, causing positive ET biases relative to GLEAM. However, through water balance partitioning, this increased evaporative flux actively depletes soil water, effectively correcting the model's intrinsic wet bias and thereby improving SM accuracy. Conversely, ERA5's lower SWR and higher RH constrain ET - bringing it closer to GLEAM estimates - but suppress soil drying, resulting in the retention of

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

excess moisture and the degradation of SM performance. These results highlight that SWR and RH inputs strongly regulate the partitioning between SM storage and ET fluxes, creating a trade-off in model performance depending on the variable evaluated (Wang et al., 2007). As none of the simulation runs optimally resolves all hydrological processes - with IMDAA excelling in SM and ERA5 in ET - future efforts should focus on data fusion (Wang et al., 2024). Assimilating different datasets, including estimation of ecosystem parameters which have an important impact on ET, could yield a more robust hybrid dataset capable of resolving current trade-offs and reproducing both SM and ET dynamics concurrently.

6.5 Evapotranspiration (ET) evaluation against GLEAM across zones

Figure 8 compares CLM5-simulated ET against GLEAM estimates across six climate zones. MD patterns indicate that overall model performance varied substantially by region. Systematic negative biases were evident in the Am and Cwb zones across all simulations. Similarly, rainfed simulations in the semi-arid (BSh) and arid (BWh) zones - as well as ERA5-Rfd in Aw - exhibit negative biases, consistent with water-limited conditions. In contrast, positive biases emerged in the subtropical Cwa and in the irrigated simulations over BSh and BWh, where artificial water application enhanced ET. The smallest differences were observed in the temperate highland (Cwb) and arid (BWh) zones, while the largest differences occurred in the humid monsoon zone (Am). The magnitude of these errors is further reflected in the RMSD. ERA5-based simulations (both irrigated and rainfed) yielded the lowest RMSD in the Am, Aw, and BSh zones; performance in the arid BWh zone was comparable across IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd; ERA5-Rfd performed best in Cwa; and IMDAA-Irr showed the lowest RMSD in Cwb. Bias patterns show a clear link to moisture availability. Temporal agreement is strong across most regions, with correlation coefficients exceeding 0.8 in all zones except Am, where values are notably lower (~0.42–0.47).

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

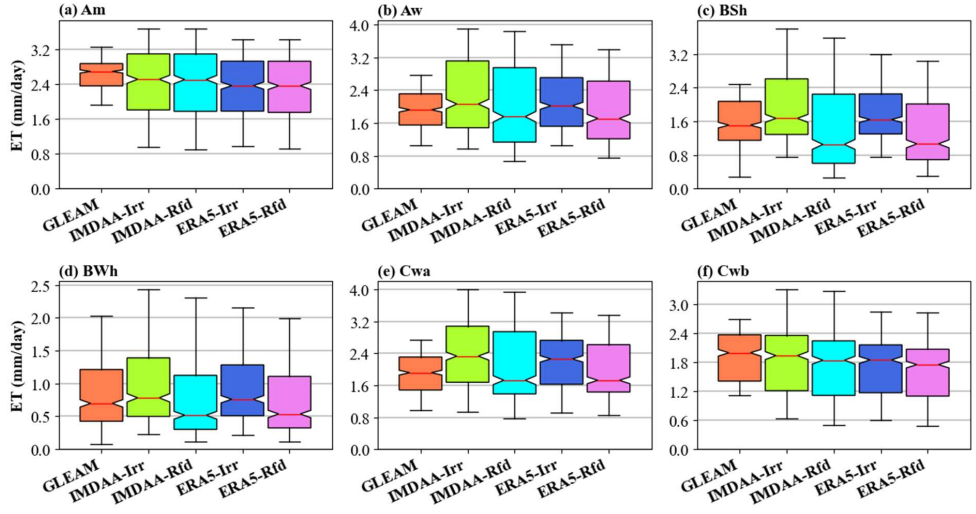


Figure 8: Comparison of CLM5-simulated evapotranspiration (ET) with GLEAM across six climate zones. Panels (a–f) show boxplots of ET for GLEAM and simulations (s1–s4) over Am, Aw, BSh, BWh, Cwa, and Cwb zones. In the boxplots, the central line indicates the median, the box spans the interquartile range (25th–75th percentile), whiskers extend to 1.5×IQR, and outliers are shown as circles.

Table 3: Performance metrics ($RMSD$, MD , and ρ) for ET vs GLEAM.

Climate Zone	Metric	IMDAA-Irr	IMDAA-Rfd	ERA5-Irr	ERA5-Rfd
Am	MD (mm/day)	-0.19	-0.21	-0.30	-0.32
	RMSD (mm/day)	0.68	0.70	0.66	0.68
	ρ	0.47	0.46	0.42	0.42
Aw	MD (mm/day)	0.35	0.08	0.20	-0.03
	RMSD (mm/day)	0.61	0.62	0.39	0.44
	ρ	0.89	0.88	0.89	0.87
BSh	MD (mm/day)	0.37	-0.15	0.24	-0.20
	RMSD (mm/day)	0.57	0.51	0.42	0.42
	ρ	0.83	0.91	0.83	0.90
BWh	MD (mm/day)	0.14	-0.09	0.07	-0.12
	RMSD (mm/day)	0.31	0.27	0.27	0.27

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

	ρ	0.83	0.86	0.83	0.85
Cwa	MD (mm/day)	0.49	0.22	0.31	0.08
	RMSD (mm/day)	0.68	0.59	0.44	0.36
	ρ	0.91	0.88	0.91	0.88
Cwb	MD (mm/day)	-0.10	-0.19	-0.21	-0.29
	RMSD (mm/day)	0.26	0.33	0.27	0.35
	ρ	0.88	0.87	0.91	0.89

6.6 Runoff evaluation against GRUN grid-wise

The grid-wise MD patterns for runoff (Figure 9a–d) show widespread positive biases across southern India and parts of central and northern regions, with larger magnitudes in irrigated simulations. In contrast, negative biases dominate across northeastern India, portions of northern India, and along the southwestern coast. Comparisons between forcing datasets (Figure 9e–f) indicate lower MD for IMDAA-driven runs across most of India, whereas ERA5-driven simulations exhibit smaller biases in localized areas of the northeast, north, northwest, and southern regions. The irrigation contrast (Figure 9g–h) further confirms that, consistent with SM and ET results, irrigated configurations generally reinforce positive runoff biases relative to rainfed simulations.

Figure 10 presents the grid-wise RMSD of runoff relative to GRUN across India. Consistent with the SM results, IMDAA-Rfd exhibits the lowest mean RMSD (1.13 mm day^{-1}) among the four simulations, indicating the best overall performance. This is further supported by Supplementary Figure S17, which shows IMDAA-Rfd achieving the lowest RMSD at 64.9% of grid cells, reaffirming its robustness in representing runoff dynamics nationwide. Spatially, the highest RMSD values across all simulations occur along the western coast, northern India, and northeastern regions (Figure 10a–d). Differences between forcing datasets (Figure 10e–f) show lower RMSD for IMDAA-driven runs across much of central and northern India, while ERA5-driven simulations perform better in limited areas of central India, the northeast, and the southwestern and southeastern coasts. A focused evaluation over irrigated grid cells (Figure S18) similarly identifies IMDAA-Rfd as the best-performing configuration at 59.2% of irrigated locations, particularly across central and peninsular India. However, a mixed performance emerges within parts of the Indo-Gangetic Plain, where the lowest RMSD is distributed among rainfed and irrigated simulations, including ERA5-Rfd, IMDAA-Rfd, ERA5-Irr, and to a lesser extent IMDAA-Irr. Localized improvements from irrigated runs also

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

appear in eastern India and scattered grids along the southwestern and southeastern coasts. ERA5-Irr performs best at 27.3% of irrigated grid cells, primarily across the Indo-Gangetic Plain, northeastern India, and sections of the eastern coast. Although IMDAA-Rfd dominates overall, irrigated configurations contribute improvements at nearly one-third of irrigated locations, highlighting their localized influence. The irrigation impact (Figure 10g–h) indicates that, similar to SM and ET, irrigated simulations generally increase RMSD relative to rainfed configurations across most irrigated grid cells.

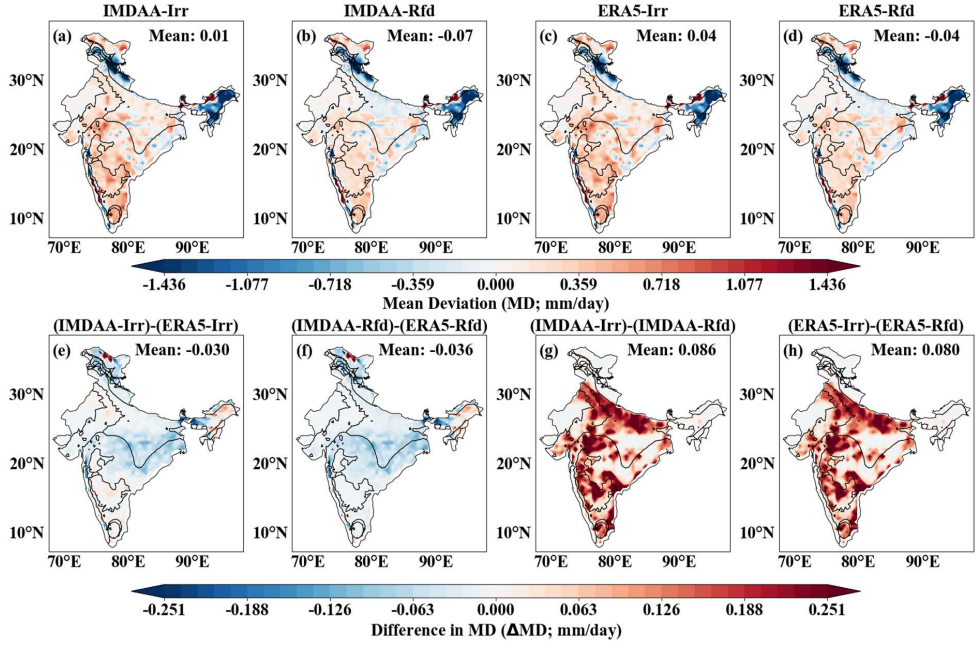


Figure 9: Grid-wise Mean Deviation (MD) of Runoff simulated by CLM5 relative to GLEAM across India for 1980–2020. Panels (a–d) show MD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in MD (ΔMD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

Correlation patterns (Figure S19) provide complementary insight, showing that rainfed simulations consistently achieve stronger temporal agreement with GRUN than irrigated runs, particularly across southern India and the Indo-Gangetic Plain. This indicates that although irrigation can locally reduce mean bias, it does not systematically improve runoff performance across regions. Atmospheric forcing also plays a critical role: the higher evaporative demand associated with IMDAA forcing (Figures S20 and S21) enhances ET and depletes SM relative

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

to ERA5, thereby reducing runoff generation. Together, these findings highlight the strong sensitivity of CLM5 runoff simulations to both irrigation representation and meteorological drivers in regulating hydrological partitioning.

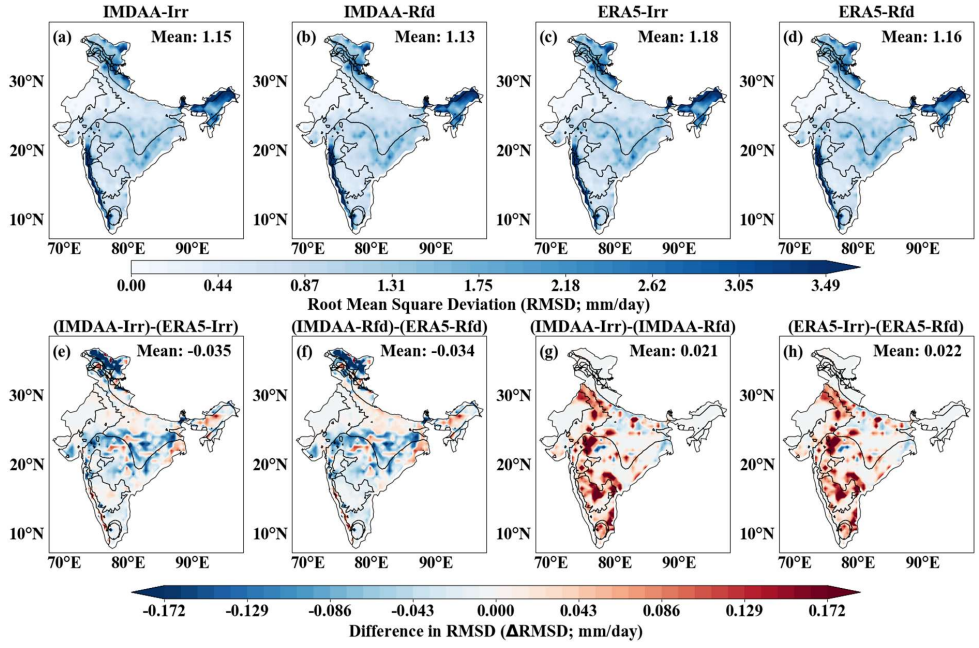


Figure 10: Grid-wise Root Mean Square Deviation (RMSD) of Runoff simulated by CLM5 relative to GRUN across India for 1980–2020. Panels (a–d) show RMSD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in RMSD (ΔRMSD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

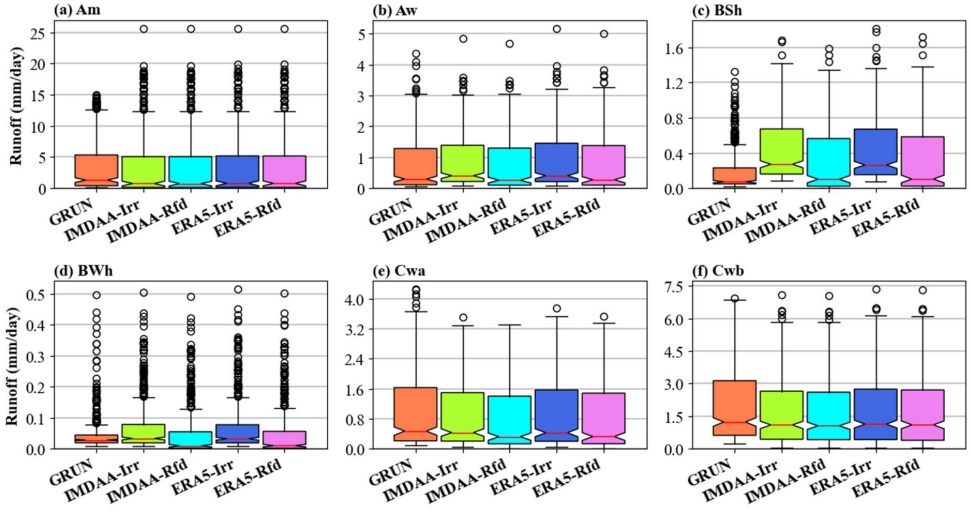
6.7 Runoff evaluation against GRUN across climate zones

Figure 11 compares CLM5-simulated runoff with GRUN across six climate zones. Mean Deviation (MD) patterns indicate that model performance varies substantially across regions, with biases differing by climate regime. The model performance also varies significantly in terms of error magnitude, with RMSD ranging from 0.05 mm/day (BWh) to 1.70 mm/day (Am). In terms of correlation, values range from 0.47 (BWh) to 0.93 (Cwb), with the lowest values consistently found in arid regions. The best-performing configurations also vary by climate zone: all simulations yield comparable RMSD in the Am (1.70) and BWh (0.05) zones;

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

699 ERA5-Irr performs best in Cwa and Cwb; IMDAA-Rfd shows the lowest RMSD in BSh; and
700 ERA5-Rfd performs best in Aw.

701 Synthesizing the results across climate zones (Figures 5, 8, 11, and 12) reveals three
702 distinct behavioural patterns in the model. First, the Am zone exhibits a consistent 'Dry Bias,'
703 characterized by low SM, ET, and runoff across all simulations, stemming from two
704 interconnected factors: (a) systematic underestimation of precipitation in IMD forcing over
705 humid regions (Goteti et al. 2024), reducing available water, and (b) flawed ET partitioning -
706 wherein interception and bare soil evaporation are high while transpiration (see Figure 12) is
707 severely low (Zhang et al., 2023) - collectively resulting in net low ET, SM, and runoff.



708

709 **Figure 11:** Comparison of CLM5-simulated runoff with GRUN across six climate zones.
710 Panels (a–f) show boxplots of runoff for GRUN and simulations (s1–s4) over Am, Aw, BSh,
711 BWh, Cwa, and Cwb zones. In the boxplots, the line inside each box indicates the median, the
712 box spans the interquartile range (25th–75th percentile), whiskers extend to 1.5×IQR, and
713 outliers are shown as circles.

714 Second, the Tropical and Arid zones (Aw, BSh, BWh) exhibit a 'Storage Surplus,' defined by
715 high SM and high runoff. In the arid zones (BSh, BWh), this accumulation can probably be
716 attributed to the Dry Surface Layer (DSL) scheme, which suppresses soil evaporation (Deng et
717 al., 2020; Wang et al., 2024), trapping excess moisture and driving elevated runoff. ET in these
718 zones varies by water supply: it remains low in rainfed arid simulations but increases
719 substantially in irrigated runs (and consistently in the Aw zone) where water availability meets

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

atmospheric demand. Third, the Temperate zones (Cwa, Cwb) display a 'Runoff Paradox,' characterized by high surface SM but low runoff. In Cwa, this surplus SM successfully supports high ET. However, Cwb presents a distinct decoupling: despite high surface SM, ET and runoff are low. The lower simulated root-zone moisture may partly reflect differences in how the root zone is defined, as CLM5 represents SM within the upper ~0.58 m (as bedrock presents in CLM5), whereas GLEAM's estimates root-zone moisture over a deeper profile (~1 m; no bed rock, considers uniform SM depth), leading to structural inconsistencies between the two datasets rather than a definitive model bias. Consequently, the suppressed ET in Cwb is likely not caused by surface water stress, but rather by limited root-zone storage, restricted infiltration, and biases in plant phenology (low transpiration: see Figure 12) that prevent effective root-water uptake.

Table 4: Performance metrics ($RMSD$, MD , and ρ) for Runoff vs GRUN.

Climate Zone	Metric	IMDAA-Irr	IMDAA-Rfd	ERA5-Irr	ERA5-Rfd
Am	MD (mm/day)	-0.35	-0.36	-0.29	-0.30
	RMSD (mm/day)	1.70	1.70	1.70	1.70
	ρ	0.91	0.91	0.91	0.91
Aw	MD (mm/day)	0.08	-0.03	0.12	0.02
	RMSD (mm/day)	0.33	0.31	0.32	0.29
	ρ	0.82	0.87	0.82	0.87
BSh	MD (mm/day)	0.24	0.11	0.24	0.13
	RMSD (mm/day)	0.31	0.22	0.31	0.24
	ρ	0.64	0.75	0.64	0.75
BWh	MD (mm/day)	0.02	0.00	0.02	0.00
	RMSD (mm/day)	0.05	0.05	0.05	0.05
	ρ	0.55	0.47	0.55	0.47
Cwa	MD (mm/day)	-0.15	-0.24	-0.11	-0.20
	RMSD (mm/day)	0.40	0.45	0.35	0.39
	ρ	0.90	0.90	0.90	0.91
Cwb	MD (mm/day)	-0.31	-0.34	-0.26	-0.29
	RMSD (mm/day)	0.79	0.80	0.78	0.79
	ρ	0.92	0.93	0.93	0.93

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

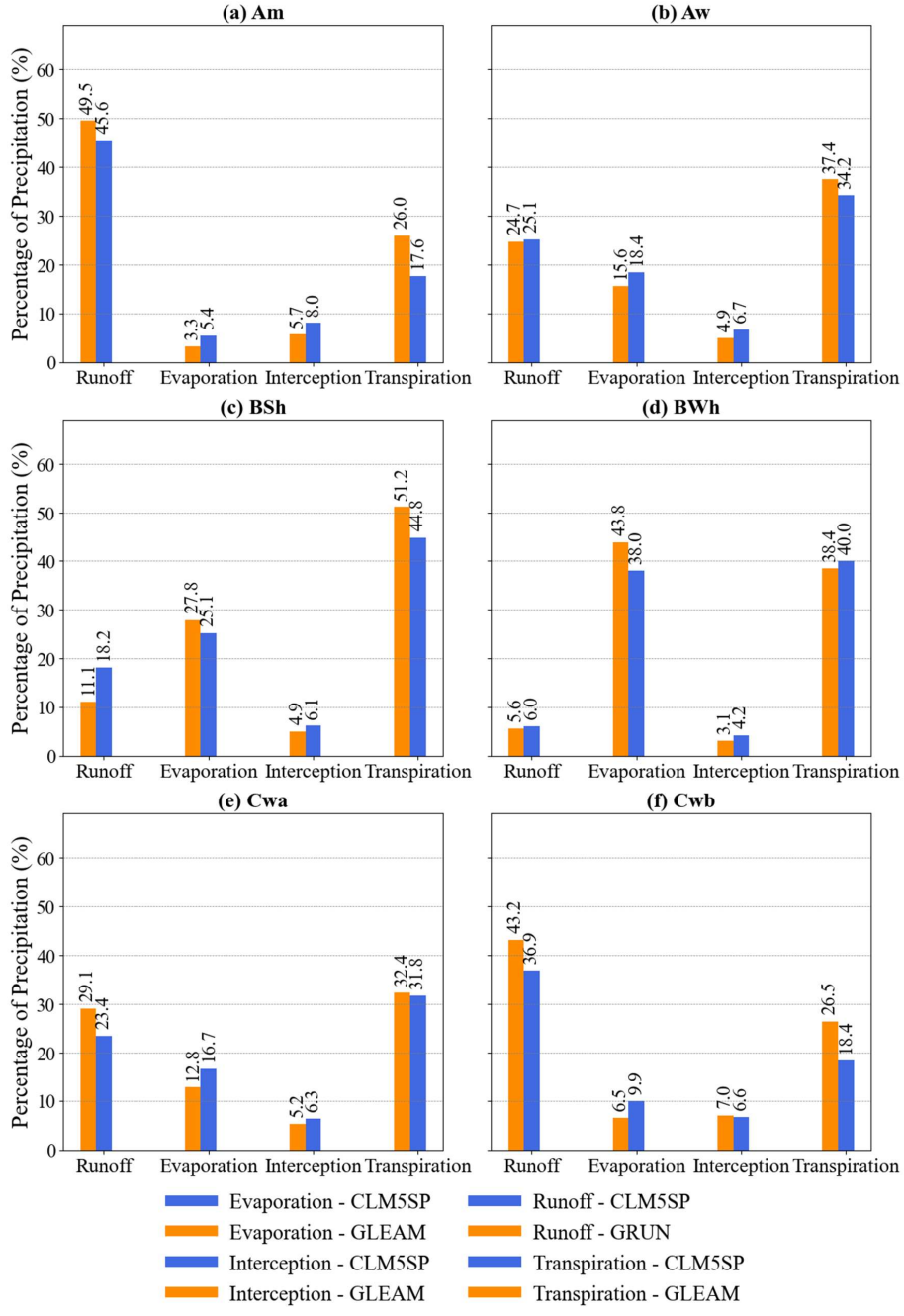
Specifically, the model likely underestimates LAI due to limitations in the SP parameterization in CLM5. The SP scheme relies on long-term monthly climatology that repeats annually, failing to capture interannual variability in vegetation dynamics. Consequently, the model cannot capture enhanced vegetation growth during wetter years, creating a "vegetation-limited" rather than "water-limited" state where excess SM does not translate into higher transpiration rates (Wang et al., 2024). Similar behaviour has been reported for CLM5-BGC, which tends to underestimate ET even more strongly than CLM5-SP due to an inherent underestimation of LAI, further reinforcing the role of vegetation phenology in shaping ET biases (Cheng et al., 2021; Wei et al., 2017). Overall, these patterns indicate a systematic bias in CLM5 toward lower ET and higher SM, consistent with findings from evaluations over China, Europe, and the USA (Cheng et al., 2021; Poppe Terán et al., 2025; Zhang et al., 2023).

Second, regarding the 'Runoff Paradox' (high SM yielding low runoff), the contradiction likely stems from CLM5's omission of hillslope-scale terrain processes. The model primarily simulates vertical water movement (infiltration and percolation) but neglects sub-grid lateral ridge-to-valley flows, which are a primary mechanism for runoff generation in complex terrain (Fan et al., 2019). Without these lateral pathways, excess water that should rapidly enter streams via saturation-excess runoff is instead retained within the soil column, directly contributing to the observed overestimation of SM. As demonstrated by Swenson et al. (2019), incorporating representative hillslope physics - whereby grid cells are decomposed into upland and lowland columns - is necessary to redistribute this soil water correctly, thereby resolving the structural disconnection between SM storage and runoff generation.

6.8 Precipitation partitioning and its regional variability across climatic zones

Figure 12 summarizes the partitioning of precipitation into runoff, transpiration, evaporation, and interception across six climate zones for CLM5 (ERA5-Rfd) relative to GLEAM and GRUN. A clear hydroclimatic gradient emerges. In the humid Cwb and Am zones, runoff dominates the water balance. GRUN indicates that runoff accounts for ~43–50% of precipitation, with transpiration contributing ~26%. In comparison, CLM5 allocates a smaller fraction to both runoff (~37–46%) and transpiration (~18%), with a compensating increase in bare-soil evaporation and interception. In semi-arid BSh, transpiration represents the primary pathway in GLEAM (~51%), whereas CLM5 assigns a lower fraction to transpiration (~45%) and a higher fraction to runoff (~18%) compared with GRUN (~11%).

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026



765

766 **Figure 12.** Comparison of observed and CLM5SP-simulated hydrological component
767 partitioning (runoff, evaporation, interception, and transpiration) expressed as a percentage of
768 precipitation across six major climate zones in India.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

In the arid BWh zone, GLEAM partitions atmospheric losses nearly evenly between evaporation (~44%) and transpiration (~38%), while CLM5 simulates slightly lower evaporation (~38%) and marginally higher transpiration (~40%). The transition zones (Aw and Cwa) exhibit a more balanced regime, with transpiration (~32–37%) exceeding runoff (~23–29%). Relative to GLEAM and GRUN, CLM5 consistently allocates smaller fractions to transpiration and, in Cwa, substantially lower runoff, while increasing non-productive losses such as bare-soil evaporation (e.g., ~17% in CLM5 versus ~13% in GLEAM). These differences reflect contrasting assumptions in water-balance partitioning rather than definitive errors, given the large uncertainties in separating evaporation and transpiration.

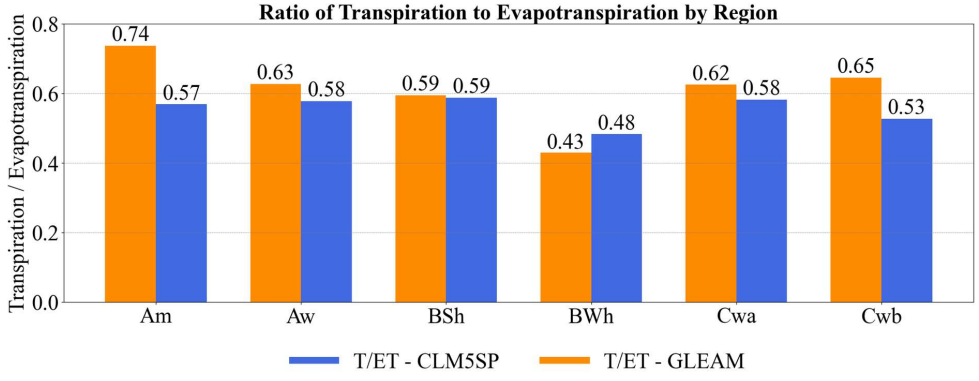


Figure 13. Transpiration Fraction of Evapotranspiration Across Köppen Climate Zones (values rounded to two decimals).

Figure 13 shows the transpiration fraction of evapotranspiration (T/ET), highlighting how much of total ET is driven by vegetation rather than non-transpirative evaporation. Across Köppen climate zones, T/ET ranges from about 0.43 to 0.74, consistent with global estimates (Wei et al., 2017) that identify transpiration as the dominant ET component in most regions (Song et al., 2020). GLEAM yields higher T/ET ratios than CLM5SP in most regions, especially in humid climates, while in the semi-arid BSh climate both products give nearly identical ratios. In contrast, over the arid BWh climate CLM5SP produces a higher T/ET (0.48) than GLEAM (0.43), suggesting differences in partitioning ET under strongly water-limited conditions between GLEAM and CLM5SP.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

6.9 Simulating Historical Drought Events

To assess CLM5's capability in simulating historical droughts, this study reconstructs major meteorological and agricultural monsoon droughts from 1987, 2002, 2009, and 2015, and hydrological droughts during the 2018 monsoon and the 2003 and 2016 water years, based on prior studies (Chuphal et al., 2024; Mishra, 2020; Mishra et al., 2018). For the drought analysis, all land-surface variables (SM, ET, and runoff) are taken from the IMDAA-Rfd configuration. This setup was selected because it provides the best overall hydrological consistency, reducing RMSD for SM (10%) and runoff (2.5%) relative to ERA5-Rfd, despite exhibiting higher ET errors (by 17%) compared with ERA5-based simulations.

The best-fitting distributions at the grid level are shown in Figures S23–S26 (supplementary material). As illustrated in Figures 14, 15, and 16, the spatial extent of monsoon-season and water-year droughts - across meteorological, agricultural, and hydrological categories - closely matches previous findings by Chuphal et al. (2024), and Mishra et al. (2020). Figure 16b shows lower drought severity over parts of the Indo-Gangetic Plain and Northeast India during 2018 relative to the Variable Infiltration Capacity (VIC) based estimates, likely reflecting differences in runoff representation between the VIC and CLM5 models. Overall, the strong agreement with historical records highlights CLM5SP's effectiveness in reproducing drought conditions across varied hydroclimatic zones in India.

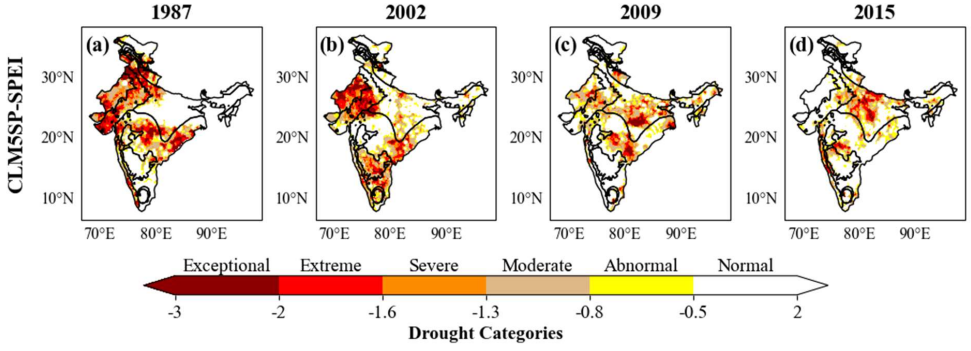


Figure 14: Spatial extent of meteorological droughts (Standardised Precipitation and Evapotranspiration Index - SPEI) simulated by CLM5SP for the Monsoon (JJAS) during the years 1987, 2002, 2009 and 2015.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

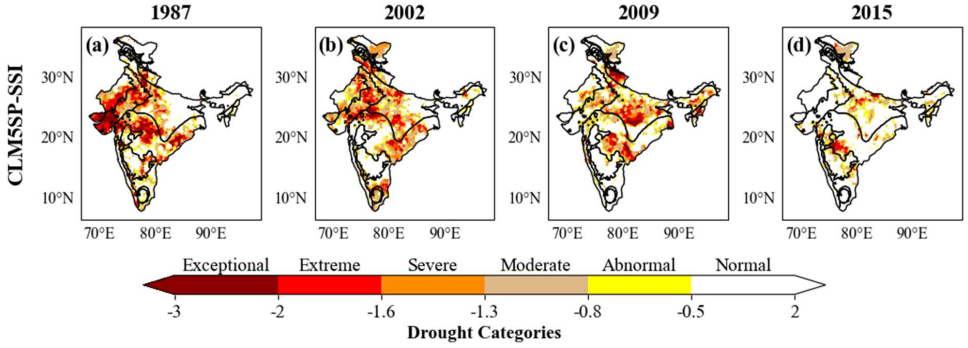


Figure 15: Spatial extent of agricultural droughts (Soil Moisture Index - SSI) simulated by CLM5SP for the Monsoon (JJAS) during the years 1987, 2002, 2009 and 2015.

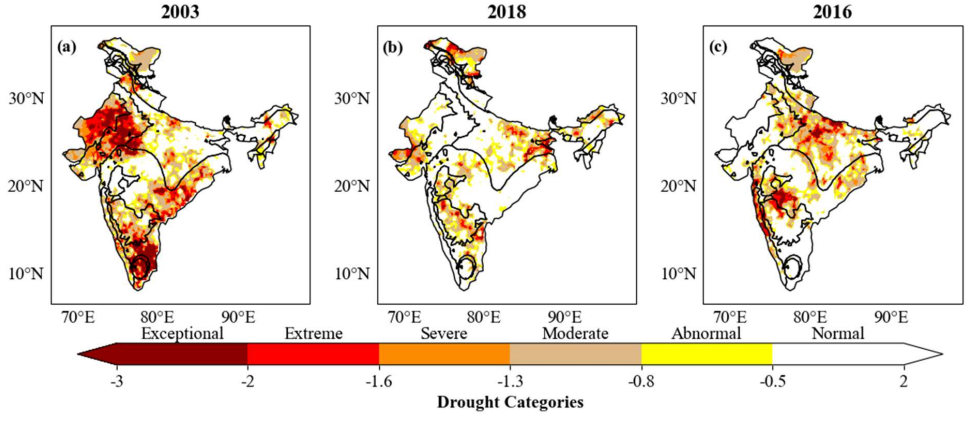


Figure 16. Spatial extent of Hydrological drought (SRI) derived from CLM5SP products for (a) 2003 water-year (June-May: sum), (b) 2018 monsoon drought (JJAS: sum), and (c) 2016 water-year (June-May: sum).

6. 10 Discussion

The marginal benefit of irrigation simulations likely stems from two key limitations. First, the input dataset - Monthly Irrigated and Rainfed Crop Areas around the year 2000 (MIRCA2000; Portmann et al., 2010) is static and derived from coarse resolutions, failing to capture year-to-year variability and expansion of irrigation across India. Second, the default irrigation scheme in CLM5 is structurally simplified: it applies water uniformly at a fixed time (06:00 local time) whenever LAI > 0 and SM falls below a threshold. This logic ignores the diverse timing, methods, and efficiencies of real-world practices and excludes critical water sources such as reservoirs and groundwater pumping. Recent studies (Dombrowski et al., 2024; Yao et al.,

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

2022) demonstrate that integrating observed irrigation data and refining parameterizations for specific methods (e.g., sprinkler, flood, and paddy) substantially improves the simulation of SM, transpiration, and crop yield. Therefore, accounting for this spatial and temporal heterogeneity is an essential next step. Such enhancements would not only improve hydrological consistency but also strengthen CLM5's ability to assess the broader impacts of irrigation on energy fluxes, carbon–nitrogen cycling, and regional climate (Dombrowski et al., 2024; Zhu et al., 2020).

Our central finding is that CLM5 biases stem from fundamental, climate-dependent structural flaws that manifest as three distinct patterns: a 'Dry Bias' in humid zones (Am), a 'Wet Bias' in arid and semi-arid zones (BWh, BSh), and a 'Runoff Paradox' in temperate zones (Cwa, Cwb). Addressing the wet bias in arid regions requires revisiting the DSL parameterization. Conversely, the dry bias in humid (Am) zones can be overcome by incorporating high-resolution dynamical downscaling models (e.g., Weather Research and Forecasting (WRF) Hydro model at 3–4 km resolution), which better capture orographic precipitation in mountainous regions than gauge-based products (Hunt and Menon, 2020; Li et al., 2017). Such downscaled precipitation could address IMD underestimation over humid zones and improve CLM5.0 simulations of SM, ET, and runoff for drought assessments in data-scarce mountainous regions. Temperate (Cwb) regions reveal a critical biological limitation. Although interception and bare-soil evaporation may be overestimated, the net low ET is most likely driven by underestimated transpiration, consistent with CLM5 evaluations that attribute low ET to underestimated LAI and, consequently, a small transpiration component (Cheng et al., 2021; Wei et al., 2017). Our T/ET analysis (Figure 13) further supports this interpretation: GLEAM generally produces higher T/ET ratios than CLM5 across most Köppen climate zones, particularly in humid regions, indicating that CLM5 allocates a smaller fraction of ET to transpiration. Overall, these patterns point to a systematic CLM5 bias toward lower ET and higher SM, in line with regional evaluations over North America, Europe, and China (Cheng et al., 2021; Poppe Terán et al., 2025; Zhang et al., 2023). Rather than relying solely on parameter tuning, future work should combine parameter optimization with targeted improvements in vegetation-related processes (e.g. phenology, stomatal conductance, and root water uptake) to better reconcile CLM5's transpiration and ET with observations. This could include LAI data assimilation (Raczka et al., 2021) and regionally adjusting dynamic vegetation planting dates to improve plant phenology (Reddy et al., 2025). Resolving these climate-dependent biases requires a comprehensive strategy that involves tuning to address structural,

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

parametric, and data-driven limitations. First, lateral flow processes in complex terrain must be considered. The model currently omits sub-grid hillslope processes, such as lateral ridge-to-valley flows, which significantly affect water partitioning (Fan et al., 2019). Incorporating representative hillslope models (Swenson et al., 2019) and realistic soil depth profiles (Jin et al., 2022) represents a critical pathway to resolve the 'Runoff Paradox' and improve the simulation of saturation-excess runoff. Second, parameter refinement is essential to correct the phenological and storage biases identified in this study. Calibrating deep soil layers (Deng et al., 2021), and employing multi-objective calibration (Denager et al., 2023) have all proven effective in substantially reducing model error. Finally, advanced data integration techniques offer a solution for residual biases driven by forcing uncertainties. Coupling the model with data assimilation systems (Raczka et al., 2021; Strebel et al., 2022) or applying time-variant bias correction strategies, such as Random Forest methods (Zhang et al., 2025), has shown promise in improving hydrological estimates.

Furthermore, discrepancies in our simulations were amplified by variations in meteorological forcing (IMDAA vs. ERA5). Given the uncertainties in input data, it is important to consider these via ensemble simulations (Wang et al., 2024). Finally, it is important to note that the CLM5 simulations in this study were uncalibrated, a standard approach for regional baselines that inherently anticipates discrepancies. Furthermore, the reference datasets themselves are subject to uncertainties, complicating the definitive attribution of error. Thus, the distinct behavioral patterns identified here likely reflect a combination of model structural deficiencies, parameter mismatches, and uncertainties within the reference benchmarks.

Rather than attempting ad-hoc model tuning, this study establishes a first evaluation of CLM5's ability to represent key hydrological variables across the Indian subcontinent. Much like regional assessments in the United States and China (Cheng et al., 2021; Ma and Wang, 2022), this work provides a necessary benchmark. By delineating the model's strengths and structural limitations within a complex, irrigation-dominated environment, we identify critical pathways for improving the representation of irrigation and water–energy interactions in future Earth system modeling.

7. Conclusions

This study presents a comprehensive evaluation of the Community Land Model 5 (CLM5.0) over India, focusing on its ability to simulate key hydrological variables - soil moisture (SM),

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

evapotranspiration (ET), and runoff - and its skill in reconstructing historical drought events from 1980 to 2020. Simulated outputs were evaluated against multiple reference datasets, including in-situ data from the Cosmic-ray Soil Moisture Observation System (COSMOS), satellite-derived Soil Moisture Active Passive (SMAP) products, the Global Land Evaporation Amsterdam Model (GLEAM), and the Global Runoff Reconstruction (GRUN). High-resolution (0.1°) simulations were performed using two atmospheric forcing datasets, Indian Monsoon Data Assimilation and Analysis (IMDAA) and ECMWF Reanalysis v5 ERA5 - to assess the model's sensitivity to forcing inputs and its capability to capture the spatial and temporal variability of hydrological processes across India.

The evaluation demonstrates that CLM5 captures the spatiotemporal variability of SM, ET, and runoff across India, showing strong potential to bridge observational gaps in data-scarce regions. Rainfed configurations generally outperform irrigated ones at majority of grids, suggesting that CLM5 currently represents hydrological processes more realistically without irrigation. Notably, irrigation reduced errors in root-zone SM over only ~26% of irrigated grid cells, while surface SM was improved over ~23% grid cells, and runoff and ET improvements were observed in ~33% and ~45% of irrigated grid cells, respectively. The grid-wise differences between irrigated and rainfed simulations show that irrigation consistently increases soil moisture, evapotranspiration, and runoff over irrigated regions in CLM5, reflecting the direct addition of water and its subsequent partitioning within the land surface system. These findings point to the need for improved irrigation parameterizations, incorporating observed irrigation practices, timing, and methods, to better capture regional water dynamics.

Distinct structural biases emerged. First, the Am humid zone's persistent dry bias stems from systematic IMD precipitation underestimation reducing available water, combined with flawed ET partitioning where interception and bare-soil evaporation are probably overestimated while transpiration is probably (severely) underestimated, resulting in net low ET, SM, and runoff. Second, the tropical and arid zones (Aw, BSh, BWh) exhibit a storage surplus, with high SM and high runoff. This is probably driven by the Dry Surface Layer (DSL) scheme, which overly suppresses soil evaporation. Under rainfed conditions this leads to low ET, whereas irrigation adds sufficient water to increase soil evaporation and raise ET. Third, the temperate zones (Cwa, Cwb) display a runoff paradox: high surface SM but low runoff. In Cwa, this moisture supports high ET, while in Cwb the root-zone SM, ET, and runoff are low. This decoupling (Cwb) reflects shallow effective root-zone depth in CLM5 and phenology-related biases that limit transpiration and root-water uptake. Additionally, the low runoff in

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Am, Cwa and Cwb is consistent with CLM5's omission of hillslope-scale lateral flow, which prevents ridge-to-valley redistribution of water and leads to unrealistic runoff deficits in complex terrain and holding moisture in the soil columns.

Model performance is strongly influenced by atmospheric forcing. Based on India-wide mean RMSD, IMDAA-Rfd reduces SM and runoff deviations by ~10% and ~2.5%, respectively, relative to ERA5-Rfd, whereas ERA5-driven simulations (ERA5-Rfd and ERA5-Irr) reduce ET RMSD by ~17% relative to IMDAA-based runs. Crucially, since all simulations were forced with identical precipitation (IMD), these discrepancies are driven solely by differences in atmospheric demand. IMDAA's high shortwave radiation (SWR) and low relative humidity (RH) enhance ET; this depletion of soil water reduces storage and runoff, improving consistency with SM and runoff benchmarks by counteracting the model's wet bias. Conversely, ERA5's lower SWR and higher RH suppress ET, bringing fluxes closer to observations but resulting in overly wet soils and overestimated runoff. This highlights the dominant role of radiation and humidity in regulating the partitioning between SM, ET, and runoff, even under identical precipitation inputs.

Finally, CLM5 successfully reproduced historical droughts across meteorological, agricultural, and hydrological domains, underscoring its utility for both retrospective drought assessment and operational drought monitoring in India.

In summary, this study demonstrates the suitability of CLM5 as a land surface modeling framework across India's diverse hydro-climatic regions, while providing process-level benchmarks that are critical for hydrology and climate research. The findings also reveal key opportunities for improvement. Going forward, efforts under DRISHTI – CLM5 should prioritize targeted parameter calibration, improved representation of irrigation, and better handling of terrain-driven hydrological processes to enhance simulations in intensively managed and complex landscapes. These developments will enable CLM5 to more reliably support drought monitoring, water resources management, and climate resilience applications in India.

Code and data availability: The datasets and Python codes used for the analysis and generation of figures in this study are available at (Naik and Dhanya, 2026).

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

CRediT authorship contribution statement.

Devavat Chiru Naik: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **C. T. Dhanya:** Conceptualization, Methodology, Supervision, Resources, Writing – review & editing. **Harrie-Jan Hendricks Franssen:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Competing Interests: At least one of the (co-)authors is a member of the editorial board of Hydrology and Earth System Sciences.

Acknowledgements

The authors would like to thank the PADAM: Hybrid High Performance Computing (HPC) Facility at Indian Institute of Technology (IIT) Delhi for providing the necessary support and resources for conducting the simulations. The work was partially supported by the funds received from Science and Engineering Research Board (SERB), DST, India, to the author C. T. Dhanya, under the SERB Women Excellence Award Scheme (grant no. SB/WEA-04/2017).

References

Ashrit, R., Indira Rani, S., Kumar, S., Karunasagar, S., Arulalan, T., Francis, T., Routray, A., Laskar, S. I., Mahmood, S., Jermey, P., Maycock, A., Renshaw, R., George, J. P., and Rajagopal, E. N.: IMDAA Regional Reanalysis: Performance Evaluation During Indian Summer Monsoon Season, Journal of Geophysical Research: Atmospheres, 125, https://doi.org/10.1029/2019JD030973, 2020.

Bonan, G. B., Levis, S., Kergoat, L., and Oleson, K. W.: Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, Global Biogeochem. Cycles, 16, 5–1, https://doi.org/10.1029/2000GB001360, 2002.

Bonan, G. B., Williams, M., Fisher, R. A., and Oleson, K. W.: Modeling stomatal conductance in the earth system: Linking leaf water-use efficiency and water transport along the soil-plant-atmosphere continuum, Geosci. Model Dev., 7, 2193–2222, https://doi.org/10.5194/gmd-7-2193-2014, 2014.

Brunke, M. A., Broxton, P., Pelletier, J., Gochis, D., Hazenberg, P., Lawrence, D. M., Leung, # L Ruby, Niu, G.-Y., Troch, P. A., and Zeng, X.: Implementing and Evaluating

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Variable Soil Thickness in the Community Land Model, Version 4.5 (CLM4.5), J. Clim., https://doi.org/10.1175/JCLI-D-15-0307.s1, 2016.

Cammalleri, C., Arias-Munõz, C., Barbosa, P., De Jager, A., Magni, D., Masante, D., Mazzeschi, M., McCormick, N., Naumann, G., and Spinoni, J.: A revision of the Combined Drought Indicator (CDI) used in the European Drought Observatory (EDO), Natural Hazards and Earth System Sciences, 21, 481–495, https://doi.org/10.5194/NHESS-21-481-2021, 2021.

Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., and Sulman, B. N.: Harnessing big data to rethink land heterogeneity in Earth system models, Hydrol. Earth Syst. Sci., 22, 3311–3330, https://doi.org/10.5194/HESS-22-3311-2018, 2018.

Cheng, Y., Huang, M., Zhu, B., Bisht, G., Zhou, T., Liu, Y., Song, F., and He, X.: Validation of the Community Land Model Version 5 Over the Contiguous United States (CONUS) Using In Situ and Remote Sensing Data Sets, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2020JD033539, 2021.

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A., and Navarra, A.: Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model, J. Adv. Model. Earth Syst., 11, 185–209, https://doi.org/10.1029/2018MS001369, 2019.

Chiru Naik, D., Chavan, S. R., and Sonali, P.: Incorporating the climate oscillations in the computation of meteorological drought over India, Natural Hazards, https://doi.org/10.1007/s11069-023-05958-3, 2023.

Chuphal, D. S., Kushwaha, A. P., Aadhar, S., and Mishra, V.: Drought Atlas of India, 1901–2020, Sci. Data, 11, 1–12, https://doi.org/10.1038/S41597-023-02856-Y;SUBJMETA=242,4111,704;KWRD=HYDROLOGY,NATURAL+HAZARDS, 2024.

Dagon, K., Sanderson, B. M., Fisher, R. A., and Lawrence, D. M.: A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5, Adv. Stat. Climatol. Meteorol. Oceanogr., 6, 223–244, https://doi.org/10.5194/ascmo-6-223-2020, 2020.

Dash, S. K., Nair, A. A., Kulkarni, M. A., and Mohanty, U. C.: Characteristic changes in the long and short spells of different rain intensities in India, Theor. Appl. Climatol., 105, 563–570, https://doi.org/10.1007/s00704-011-0416-x, 2011.

Decker, M. and Zeng, X.: Impact of Modified Richards Equation on Global Soil Moisture Simulation in the Community Land Model (CLM3.5), J. Adv. Model. Earth Syst., 1, https://doi.org/10.3894/james.2009.1.5, 2009.

Denager, T., Sonnenborg, T. O., Looms, M. C., Bogena, H., and Jensen, K. H.: Point-scale multi-objective calibration of the Community Land Model (version 5.0) using in situ observations of water and energy fluxes and variables, Hydrol. Earth Syst. Sci., 27, 2827–2845, https://doi.org/10.5194/hess-27-2827-2023, 2023.

Deng, M., Meng, X., Lyv, Y., Zhao, L., Li, Z., Hu, Z., and Jing, H.: Comparison of Soil Water and Heat Transfer Modeling Over the Tibetan Plateau Using Two Community

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

1035 Land Surface Model (CLM) Versions, J. Adv. Model. Earth Syst., 12,
1036 https://doi.org/10.1029/2020MS002189, 2020.

1037 Deng, M., Meng, X., Lu, Y., Li, Z., Zhao, L., Hu, Z., Chen, H., Shang, L., Wang, S., and
1038 Li, Q.: Impact and Sensitivity Analysis of Soil Water and Heat Transfer
1039 Parameterizations in Community Land Surface Model on the Tibetan Plateau, J. Adv.
1040 Model. Earth Syst., 13, https://doi.org/10.1029/2021MS002670, 2021.

1041 Dombrowski, O., Brogi, C., Hendricks Franssen, H. J., Pisinaras, V., Panagopoulos, A.,
1042 Swenson, S., and Bogena, H.: Land Surface Modeling as a Tool to Explore Sustainable
1043 Irrigation Practices in Mediterranean Fruit Orchards, Water Resour. Res., 60,
1044 https://doi.org/10.1029/2023WR036139, 2024.

1045 Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W.
1046 N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J.
1047 R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle,
1048 R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., and Van Zyl, J.: The soil
1049 moisture active passive (SMAP) mission, Proceedings of the IEEE, 98, 704–716,
1050 https://doi.org/10.1109/JPROC.2010.2043918, 2010.

1051 Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks,
1052 P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell,
1053 J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A.,
1054 Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E.,
1055 Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P.,
1056 Nijssen, B., Safeeq, M., Shen, C., van Verseveld, W., Volk, J., and Yamazaki, D.:
1057 Hillslope Hydrology in Global Change Research and Earth System Modeling, Water
1058 Resour. Res., https://doi.org/10.1029/2018WR023903, 2019.

1059 Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., Hengl,
1060 T., Kollet, S., Agam, N., and Avissar, R.: Soil structure is an important omission in Earth
1061 System Models, Nature Communications 2020 11:1, 11, 1–11,
1062 https://doi.org/10.1038/s41467-020-14411-z, 2020.

1063 Gebrechorkos, S. H., Peng, J., Dyer, E., Miralles, D. G., Vicente-Serrano, S. M., Funk,
1064 C., Beck, H. E., Asfaw, D. T., Singer, M. B., and Dadson, S. J.: Global high-resolution
1065 drought indices for 1981-2022, Earth Syst. Sci. Data, 15, 5449–5466,
1066 https://doi.org/10.5194/essd-15-5449-2023, 2023.

1067 Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: An
1068 observation-based global gridded runoff dataset from 1902 to 2014, Earth Syst. Sci.
1069 Data, 11, 1655–1674, https://doi.org/10.5194/essd-11-1655-2019, 2019.

1070 Ghodichore, N., Dhanya, C. T., and Hendricks Franssen, H. J.: Isolating the effects of
1071 land use land cover change and inter-decadal climate variations on the water and energy
1072 cycles over India, 1981–2010, J. Hydrol. (Amst)., 612,
1073 https://doi.org/10.1016/j.jhydrol.2022.128267, 2022.

1074 Goteti, G. and Famiglietti, J.: Extent of gross underestimation of precipitation in India,
1075 Hydrol. Earth Syst. Sci., 28, 3435–3455, https://doi.org/10.5194/HESS-28-3435-2024,
1076 2024.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., and Gentine, P.: Large influence of soil moisture on long-term terrestrial carbon uptake, Nature, 565, 476–479, https://doi.org/10.1038/s41586-018-0848-x, 2019.

Hao, Z. and AghaKouchak, A.: Multivariate Standardized Drought Index: A parametric multi-index model, Adv. Water Resour., 57, 12–18, https://doi.org/10.1016/j.advwatres.2013.03.009, 2013.

Hao, Z., Hao-Green Development, F., Hao, Z., Yuan, X., Xia, Y., Hao, F., and SingH, V. P.: AFFILIATIONS: The drought monitoring and prediction/forecast systems at regional and global scales are AN OVERVIEW OF DROUGHT MONITORING AND PREDICTION SYSTEMS AT REGIONAL AND GLOBAL SCALES, https://doi.org/10.1175/BAMS-D-15-00149.1, 2012.

Hao, Z., Hao, F., Xia, Y., Singh, V. P., Hong, Y., Shen, X., and Ouyang, W.: A Statistical Method for Categorical Drought Prediction Based on NLDAS-2, J. Appl. Meteorol. Climatol., 55, 1049–1061, https://doi.org/10.1175/JAMC-D-15-0200.1, 2016.

Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., and Gilroy, E. J.: tm4a3.pdf - Statistical Methods in Water Resources, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., de Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hunt, K. M. R. and Menon, A.: The 2018 Kerala floods: a climate change perspective, Climate Dynamics 2020 54:3, 54, 2433–2446, https://doi.org/10.1007/S00382-020-05123-7, 2020.

Indirarani, S., Arulalan, T., George, J. P., Rajagopal, E. N., Renshaw, R., Maycock, A., Barker, D. M., and Rajeevan, M.: IMDAA: High-Resolution Satellite-Era Reanalysis for the Indian Monsoon Region, J. Clim., 34, 5109–5133, https://doi.org/10.1175/JCLI-D-20-0412.1, 2021.

Jia, Y., Li, C., Yang, H., Yang, W., and Liu, Z.: Assessments of three evapotranspiration products over China using extended triple collocation and water balance methods, J. Hydrol. (Amst)., 614, https://doi.org/10.1016/j.jhydrol.2022.128594, 2022.

Jin, J., Wang, L., Yang, J., Si, B., and Niu, G. Y.: Improved runoff simulations for a highly varying soil depth and complex terrain watershed in the Loess Plateau with the Community Land Model version 5, Geosci. Model Dev., 15, 3405–3416, https://doi.org/10.5194/gmd-15-3405-2022, 2022.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Kim, J. and Mohanty, B. P.: Influence of lateral subsurface flow and connectivity on soil water storage in land surface modeling, J. Geophys. Res., 121, 704–721, https://doi.org/10.1002/2015JD024067, 2016.

Koven, C. D., Hugelius, G., Lawrence, D. M., and Wieder, W. R.: Higher climatological temperature sensitivity of soil carbon in cold than warm climates, Nat. Clim. Chang., 7, 817–822, https://doi.org/10.1038/nclimate3421, 2017.

Lawrence, D. M., Thornton, P. E., Oleson, K. W., and Bonan, G. B.: The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmosphere interaction, J. Hydrometeorol., 8, 862–880, https://doi.org/10.1175/JHM596.1, 2007.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, J. Adv. Model. Earth Syst., 3, https://doi.org/10.1029/2011ms000045, 2011.

Lawrence, D. M., Fisher, R. A., Koven, C. , Oleson, Keith., Swenson, S. C. and, and Vertenstein, M.: Community Land Model (CLM5.0) Documentation, 2019a.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, J. Adv. Model. Earth Syst., 11, 4245–4287, https://doi.org/10.1029/2018MS001583, 2019b.

Lawrence, P. J. and Chase, T. N.: Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), J. Geophys. Res. Biogeosci., 112, 1023, https://doi.org/10.1029/2006JG000168, 2007.

Lawrence, P. J. and Chase, T. N.: Investigating the climate impacts of global land cover change in the community climate system model, International Journal of Climatology, 30, 2066–2087, https://doi.org/10.1002/JOC.2061, 2010.

Li, L., Gochis, D. J., Sobolowski, S., and Mesquita, M. D. S.: Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model, J. Geophys. Res., 122, 4786–4807, https://doi.org/10.1002/2016JD026279;WEBSITE:WEBSITE:AGUPUBS;JOURNAL:JOURNAL:21562202D;ISSUE:ISSUE:DOI, 2017.

Li, R., Lombardozzi, D., Shi, M., Frankenberg, C., Parazoo, N. C., Köhler, P., Yi, K., Guan, K., and Yang, X.: Representation of Leaf-to-Canopy Radiative Transfer Processes

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Improves Simulation of Far-Red Solar-Induced Chlorophyll Fluorescence in the Community Land Model Version 5, J. Adv. Model. Earth Syst., 14, https://doi.org/10.1029/2021MS002747, 2022.

Liu, Y., Jing, W., Sun, S., and Wang, C.: Multi-Scale and Multi-Depth Validation of Soil Moisture from the China Land Data Assimilation System, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 14, 9913–9930, https://doi.org/10.1109/JSTARS.2021.3116583, 2021.

Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., Fogli, P. G., Cherchi, A., Bellucci, A., Gualdi, S., Masina, S., and Navarra, A.: CMIP6 Simulations With the CMCC Earth System Model (CMCC-ESM2), J. Adv. Model. Earth Syst., 14, https://doi.org/10.1029/2021MS002814, 2022.

Ma, Q., Li, Y., Liu, F., Feng, H., Biswas, A., and Zhang, Q.: SPEI and multi-threshold run theory based drought analysis using multi-source products in China, J. Hydrol. (Amst)., 616, https://doi.org/10.1016/j.jhydrol.2022.128737, 2023.

Ma, X. and Wang, A.: Systematic Evaluation of a High-Resolution CLM5 Simulation over Continental China for 1979–2018, J. Hydrometeorol., 23, 1879–1897, https://doi.org/10.1175/JHM-D-22-0051.1, 2022.

Mahmood, S., Davie, J., Jermey, P., Renshaw, R., George, J. P., Rajagopal, E. N., and Rani, S. I.: Indian monsoon data assimilation and analysis regional reanalysis: Configuration and performance, Atmospheric Science Letters, 19, https://doi.org/10.1002/asl.808, 2018.

Mahowald, N. M., Randerson, J. T., Lindsay, K., Munoz, E., Doney, S. C., Lawrence, P., Schlunegger, S., Ward, D. S., Lawrence, D., and Hoffman, F. M.: Interactions between land use change and carbon cycle feedbacks, Global Biogeochem. Cycles, 31, 96–113, https://doi.org/10.1002/2016GB005374, 2017.

McGuire, A. D., Lawrence, D. M., Koven, C., Clein, J. S., Burke, E., Chen, G., Jafarov, E., MacDougall, A. H., Marchenko, S., Nicolsky, D., Peng, S., Rinke, A., Ciais, P., Gouttevin, I., Hayes, D. J., Ji, D., Krinner, G., Moore, J. C., Romanovsky, V., Schädel, C., Schaefer, K., Schuur, E. A. G., and Zhuang, Q.: Dependence of the evolution of carbon dynamics in the northern permafrost region on the trajectory of climate change, Proc. Natl. Acad. Sci. U. S. A., 115, 3882–3887, https://doi.org/10.1073/PNAS.1719903115, 2018.

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., Funk, C., Peters-Lidard, C. D., and Verdin, J. P.: A land data assimilation system for sub-Saharan Africa food and water security applications, Sci. Data, 4, https://doi.org/10.1038/sdata.2017.12, 2017.

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., Martens, B., Mccabe, M. F., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project - Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms, Hydrol. Earth Syst. Sci., 20, 803–822, https://doi.org/10.5194/hess-20-803-2016, 2016.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Miralles, D., De Jeu, R., Holmes, T. R. H., Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J.: An application of GLEAM to estimating global evaporation Remote sensing of evaporation: local to global scale estimation View project Ecosystem carbon fluxes in wet and seasonally dry tropical forest of Amazon View project An application of GLEAM to estimating global evaporation An application of GLEAM to estimating global evaporation An application of GLEAM to estimating global evaporation, Hydrol. Earth Syst. Sci. Discuss, 8, 1–27, https://doi.org/10.5194/hessd-8-1-2011, 2011.

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., Mccabe, M. F., Hirschi, M., Martens, B., Dolman, A. J., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project - Part 2: Evaluation of global terrestrial evaporation data sets, Hydrol. Earth Syst. Sci., 20, 823–842, https://doi.org/10.5194/hess-20-823-2016, 2016.

Mishra, A. and Liu, S. C.: Changes in precipitation pattern and risk of drought over India in the context of global warming, J. Geophys. Res., 119, 7833–7841, https://doi.org/10.1002/2014JD021471, 2014.

Mishra, V.: Long-term (1870–2018) drought reconstruction in context of surface water security in India, J. Hydrol. (Amst)., 580, 124228, https://doi.org/10.1016/J.JHYDROL.2019.124228, 2020.

Mishra, V., Shah, R., Azhar, S., Shah, H., Modi, P., and Kumar, R.: Reconstruction of droughts in India using multiple land-surface models (1951-2015), https://doi.org/10.5194/hess-22-2269-2018, 16 April 2018.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R., Lotsch, A., Friedl, M., Morisette, J. T., Votava, P., Nemani, R. R., and Running, S. W.: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, Remote Sens. Environ., 83, 214–231, https://doi.org/10.1016/S0034-4257(02)00074-3, 2002.

Naik, D. C. and Dhanya, C. T.: Data and Python codes for the evaluation of CLM5 in simulating soil moisture, evapotranspiration, runoff, and historical drought events over India, https://doi.org/10.5281/ZENODO.18912901, 2026.

Niu, G. Y., Yang, Z. L., Dickinson, R. E., and Gulden, L. E.: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, Journal of Geophysical Research Atmospheres, 110, 1–15, https://doi.org/10.1029/2005JD006111, 2005.

Oleson, K. W., Niu, G. Y., Yang, Z. L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., and Qian, T.: Improvements to the community land model and their impact on the hydrological cycle, J. Geophys. Res. Biogeosci., 113, https://doi.org/10.1029/2007JG000563, 2008.

O'Neill, P., Chan, S., Bindlish, R., Jackson, T., Colliander, A., Dunbar, S., Chen, F., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martinez-Fernandez, J., Gonzalez-Zamora, A., Lopez-Baeza, E., Udall, F., Seyfried, M., Bosch, D., Starks, P.,

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Holifield, C., Prueger, J., Su, Z., Van Der Velde, R., Asanuma, J., Palecki, M., Small, E., Zreda, M., Calvet, J. C., Crow, W., and Kerr, Y.: Assessment of version 4 of the SMAP passive soil moisture standard product, International Geoscience and Remote Sensing Symposium (IGARSS), 2017-July, 3941–3944, https://doi.org/10.1109/IGARSS.2017.8127862, 2017.

Pai, D., Sridhar, L., Rajeevan, M., Sreejith, O., Satbhai, N. S., and Mukhopadhyay, B.: 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, 2014.

Peel, M. C., Finlayson, B. L., and Mcmahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci, 1633–1644 pp., 2007.

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G. Y., Williams, Z., Brunke, M. A., and Gochis, D.: A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling, J. Adv. Model. Earth Syst., 8, 41–65, https://doi.org/10.1002/2015MS000526, 2016.

Poppe Terán, C., Naz, B. S., Vereecken, H., Baatz, R., Fisher, R. A., and Hendricks Franssen, H. J.: Systematic underestimation of type-specific ecosystem process variability in the Community Land Model v5 over Europe, Geosci. Model Dev., 18, 287–317, https://doi.org/10.5194/gmd-18-287-2025, 2025.

Portmann, F. T., Siebert, S., and Döll, P.: MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, Global Biogeochem. Cycles, 24, n/a-n/a, https://doi.org/10.1029/2008GB003435, 2010.

Raczka, B., Hoar, T. J., Duarte, H. F., Fox, A. M., Anderson, J. L., Bowling, D. R., and Lin, J. C.: Improving CLM5.0 Biomass and Carbon Exchange Across the Western United States Using a Data Assimilation System, J. Adv. Model. Earth Syst., 13, https://doi.org/10.1029/2020MS002421, 2021.

Rashid, M., Chien, R., Ducharne, A., Kim, H., J-f Yeh, P., Peugeot, C., Boone, A., Séguis, L., Yabu, Y., and Boukari, M.: Evaluation of Groundwater Simulations in Benin from the ALMIP2 Project, Journal of Hydrometeorology, https://doi.org/10.1175/JHM-D-18, 2019.

Reddy, K. N., Baidya Roy, S., Rabin, S. S., Lombardozzi, D. L., Varma, G. V., Biswas, R., and Naik, D. C.: Improving the representation of major Indian crops in the Community Land Model version 5.0 (CLM5) using site-scale crop data, Geosci. Model Dev., 18, 763–785, https://doi.org/10.5194/GMD-18-763-2025, 2025.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y. C., Kirkeväg, A., Schwinger, J., Tjiputra, J., Schanke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

DECK, historical, and scenario simulations, Geosci. Model Dev., 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Shukla, S. and Wood, A. W.: Use of a standardized runoff index for characterizing hydrologic drought, Geophys. Res. Lett., 35, https://doi.org/10.1029/2007GL032487, 2008.

Singh, T., Saha, U., Prasad, V. S., and Gupta, M. Das: Assessment of newly-developed high resolution reanalyses (IMDAA, NGFS and ERA5) against rainfall observations for Indian region, Atmos. Res., 259, https://doi.org/10.1016/j.atmosres.2021.105679, 2021.

Song, J., Miller, G. R., Cahill, A. T., Aparecido, L. M. T., and Moore, G. W.: Modeling land surface processes over a mountainous rainforest in Costa Rica using CLM4.5 and CLM5, Geosci. Model Dev., 13, 5147–5173, https://doi.org/10.5194/gmd-13-5147-2020, 2020.

Stagge, J. H., Kohn, I., Tallaksen, L. M., and Stahl, K.: Modeling drought impact occurrence based on meteorological drought indices in Europe, J. Hydrol. (Amst)., 530, 37–50, https://doi.org/10.1016/J.JHYDROL.2015.09.039, 2015.

Strebel, L., Bogena, H. R., Vereecken, H., and Hendricks Franssen, H. J.: Coupling the Community Land Model version 5.0 to the parallel data assimilation framework PDAF: Description and applications, Geosci. Model Dev., 15, 395–411, https://doi.org/10.5194/GMD-15-395-2022, 2022.

Svoboda, M., Lecomte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., and Stephens, S.: THE DROUGHT MONITOR, 2002.

Swenson, S. C. and Lawrence, D. M.: A GRACE-based assessment of interannual groundwater dynamics in the Community Land Model, Water Resour. Res., 51, 8817–8833, https://doi.org/10.1002/2015WR017582, 2015.

Swenson, S. C., Clark, M., Fan, Y., Lawrence, D. M., and Perket, J.: Representing Intrahillslope Lateral Subsurface Flow in the Community Land Model, J. Adv. Model. Earth Syst., 11, 4044–4065, https://doi.org/10.1029/2019MS001833;SUBPAGE:STRING:FULL, 2019.

Upadhyaya, D. B., Evans, J., Muddu, S., Tomer, S. K., Bitar, A. Al, 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., and Yenagi, B. S.: The Indian COSMOS network (ICON): Validating L-band remote sensing and modelled soil moisture data products, Remote Sens. (Basel)., 13, https://doi.org/10.3390/rs13030537, 2021.

Upadhyaya, S. and Ramsankaran, R.: Comprehensive inter-comparison of INSAT multispectral rainfall algorithm estimates and TMPA 3B42-RT V7 estimates across different climate regions of India during southwest monsoon period, Environ. Monit. Assess., 190, https://doi.org/10.1007/s10661-017-6411-7, 2018.

Verdin & Greenlee: Development of continental scale digital elevation models and extraction of hydrographic features, paper presented at 3rd International Conference, in:

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

In Workshop on Integrating GIS and Environmental Modeling, Natl. Cent. for Geogr. Inf. and Anal., Santa Barbara, Calif., 1996.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index, J. Clim., 23, 1696–1718, https://doi.org/10.1175/2009JCLI2909.1, 2010.

Vinnarasi, R. and Dhanya, C. T.: Changing characteristics of extreme wet and dry spells of Indian monsoon rainfall, J. Geophys. Res., 121, 2146–2160, https://doi.org/10.1002/2015JD024310, 2016.

Wang, D., Wang, D., and Mo, C.: The use of remote sensing-based et estimates to improve global hydrological simulations in the community land model version 5.0, Remote Sens. (Basel)., 13, https://doi.org/10.3390/rs13214460, 2021.

Wang, D., Wang, D., Mei, Y., Yang, Q., Ji, M., Li, Y., Liu, S., Li, B., Huang, Y., and Mo, C.: Estimates of the Land Surface Hydrology from the Community Land Model Version 5 (CLM5) with Three Meteorological Forcing Datasets over China, Remote Sens. (Basel)., 16, https://doi.org/10.3390/rs16030550, 2024.

Wang, K., Wang, P., Li, Z., Cribb, M., and Sparrow, M.: A simple method to estimate actual evapotranspiration from a combination of net radiation, vegetation index, and temperature, Journal of Geophysical Research: Atmospheres, 112, 15107, https://doi.org/10.1029/2006JD008351, 2007.

Wang, Y. and Li, G.: Evaluation of simulated soil moisture from China Land Data Assimilation System (CLDAS) land surface models, Remote Sensing Letters, 11, 1060–1069, https://doi.org/10.1080/2150704X.2020.1820614;WGROUP:STRING:PUBLICATION, 2020.

Wei, Z., Yoshimura, K., Wang, L., Miralles, D. G., Jasechko, S., and Lee, X.: Revisiting the contribution of transpiration to global terrestrial evapotranspiration, Geophys. Res. Lett., 44, 2792–2801, https://doi.org/10.1002/2016GL072235;PAGEGROUP:STRING:PUBLICATION, 2017.

Xu, L., Abbaszadeh, P., Moradkhani, H., Chen, N., and Zhang, X.: Continental drought monitoring using satellite soil moisture, data assimilation and an integrated drought index, Remote Sens. Environ., 250, 112028, https://doi.org/10.1016/J.RSE.2020.112028, 2020.

Xu, Z., Wu, Z., Shao, Q., He, H., and Guo, X.: From meteorological to agricultural drought: Propagation time and probabilistic linkages, J. Hydrol. Reg. Stud., 46, https://doi.org/10.1016/j.ejrh.2023.101329, 2023.

Yang, Z., Qian, Y., Liu, Y., Berg, L. K., Hu, H., Dominguez, F., Yang, B., Feng, Z., Gustafson, W. I., Huang, M., and Tang, Q.: Irrigation Impact on Water and Energy Cycle During Dry Years Over the United States Using Convection-Permitting WRF and a Dynamical Recycling Model, Journal of Geophysical Research: Atmospheres, 124, 11220–11241, https://doi.org/10.1029/2019JD030524, 2019.

https://doi.org/10.5194/egusphere-2026-1296
Preprint. Discussion started: 17 April 2026

Yao, Y., Vanderkelen, I., Lombardozzi, D., Swenson, S., Lawrence, D., Jägermeyr, J., Grant, L., and Thiery, W.: Implementation and Evaluation of Irrigation Techniques in the Community Land Model, J. Adv. Model. Earth Syst., 14, https://doi.org/10.1029/2022MS003074, 2022.

Zeng, X., Shaikh, M., Dai, Y., and Dickinson, R. E.: Coupling of the Common Land Model to the NCAR Community Climate Model, J. Clim., 2002.

Zhang, L., Wang, Z., Bai, X., Zhang, H., and Liu, Y.: A time-invariant bias correction strategy for improving CLM5.0 evapotranspiration simulation by random forest method for mainland China, Atmos. Res., 324, https://doi.org/10.1016/j.atmosres.2025.108196, 2025.

Zhang, L. hui, Ning, F. wei, Bai, X. liang, Zeng, X., and He, C. sheng: Performance evaluation of CLM5.0 in simulating liquid soil water in high mountainous area, Northwest China, J. Mt. Sci., 20, 1865–1883, https://doi.org/10.1007/s11629-022-7803-x, 2023.

Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H. S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang, Y., Liu, Y. Y., Miralles, D. G., and Pan, M.: Multi-decadal trends in global terrestrial evapotranspiration and its components, Sci. Rep., 6, https://doi.org/10.1038/srep19124, 2016.

Zhu, B., Huang, M., Cheng, Y., Xie, X., Liu, Y., Zhang, X., Bisht, G., Chen, X., Missik, J., and Liu, H.: Effects of Irrigation on Water, Carbon, and Nitrogen Budgets in a Semiarid Watershed in the Pacific Northwest: A Modeling Study, J. Adv. Model. Earth Syst., 12, https://doi.org/10.1029/2019MS001953, 2020.

Zink, M., Samaniego, L., Kumar, R., Thober, S., Mai, J., Schafer, D., and Marx, A.: The German drought monitor, Environmental Research Letters, 11, https://doi.org/10.1088/1748-9326/11/7/074002, 2016.