Supporting Information for

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

[3]Professor, 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.

Content of this file

Fig. S1 to S26

Tables S1 to S6

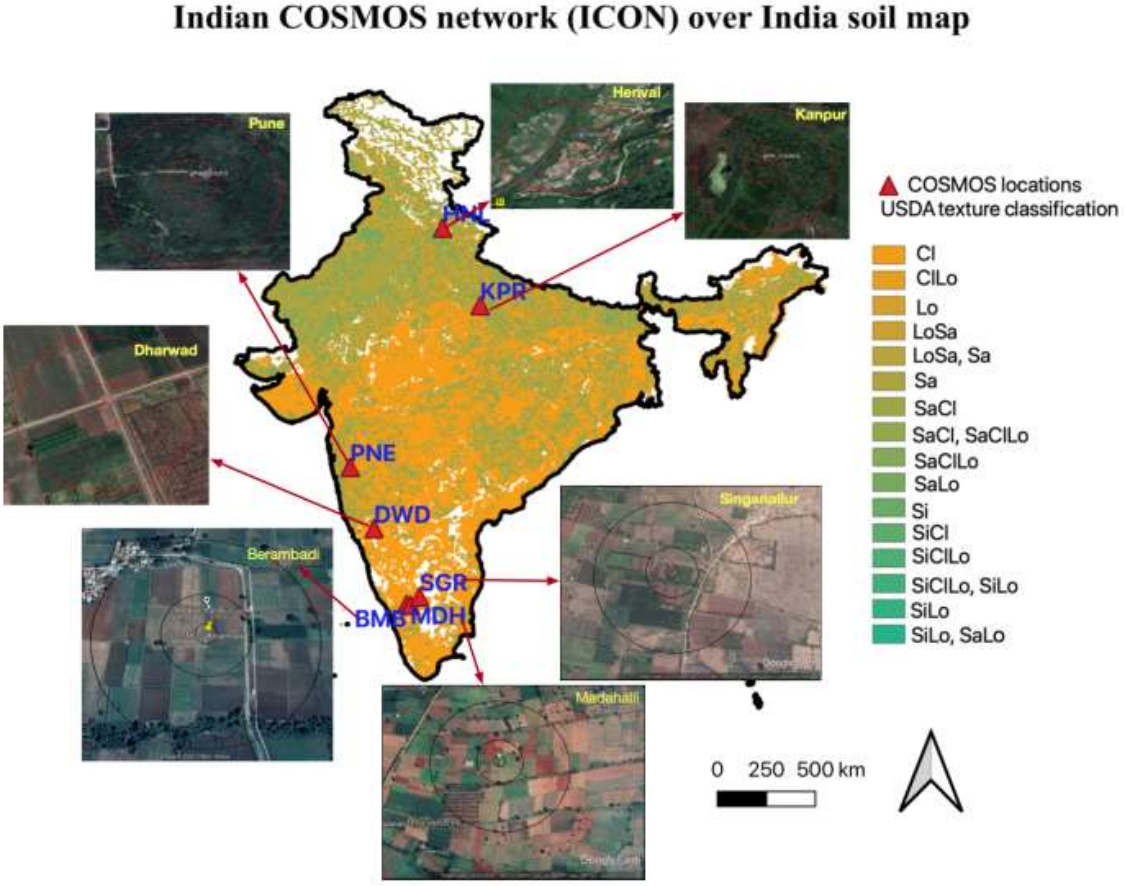


Figure S1. Locations of the Indian COSMOS Network (ICON) sites. The base map displays the soil texture of India (Source: NRSC), where Cl = Clay, Lo = Loam, Sa = Sand, and Si = Silt. The inset photos show the precise location of each site on Google Earth imagery, indicating the 50 m and 200 m impact areas. Figure adapted from Updhaya et al. (2021).

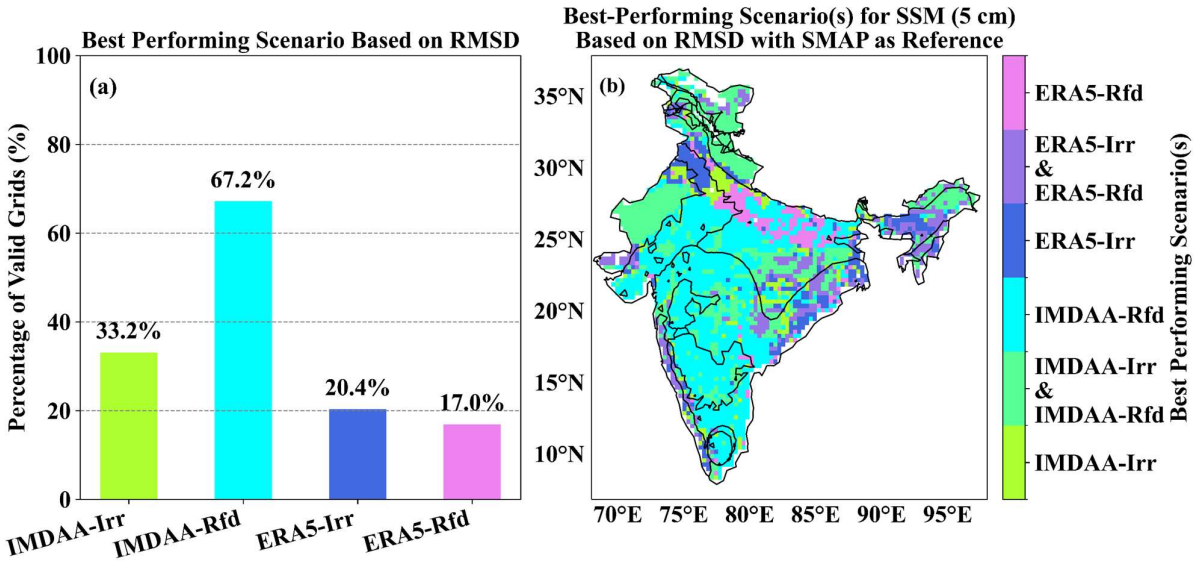


Figure S2: (a) Percentage of grid cells where each CLM5 simulations of surface soil moisture (5 cm) achieves the lowest RMSD relative to SMAP. (b) Spatial distribution of the best-performing simulation across grid cells.

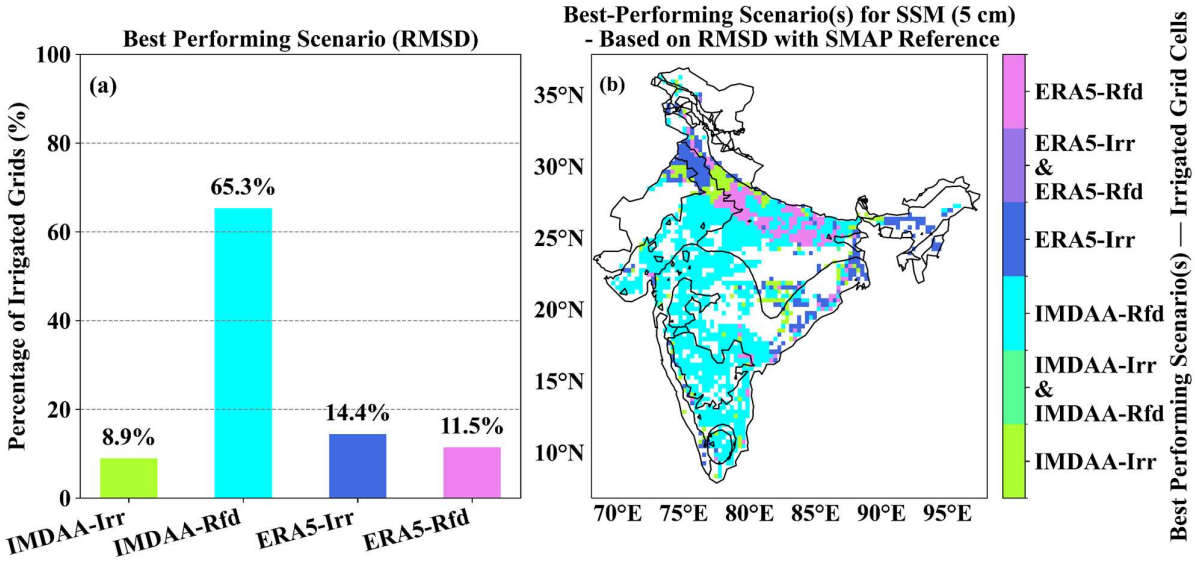


Figure S3. (a) Percentage of irrigated grid cells where each CLM5 simulations of soil moisture at 5 cm depth shows the lowest RMSD relative to SMAP. (b) Spatial distribution of the best-performing simulation across irrigated grid cells.

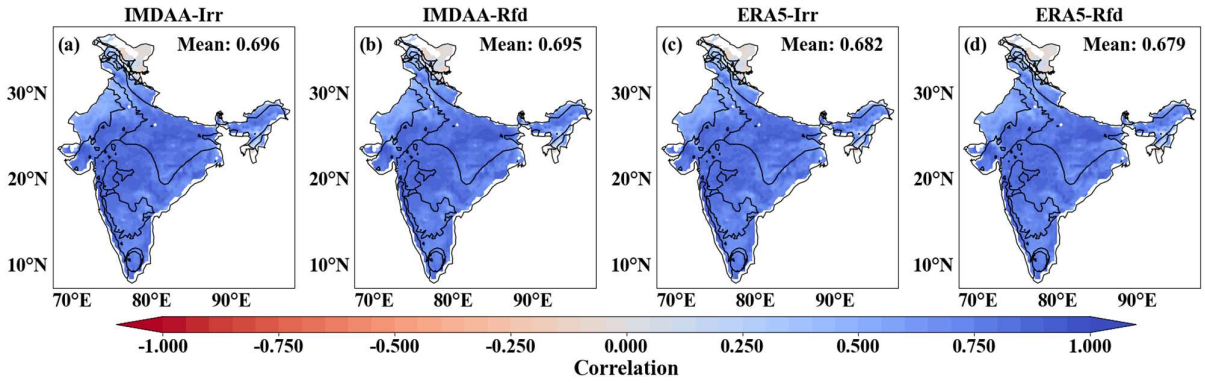


Figure S4: Grid-wise spatial distribution of Spearman correlation coefficients between CLM5-simulated surface soil moisture (SSM, 5 cm) and SMAP observations across India. Panels (a–d) correspond to IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd, respectively.

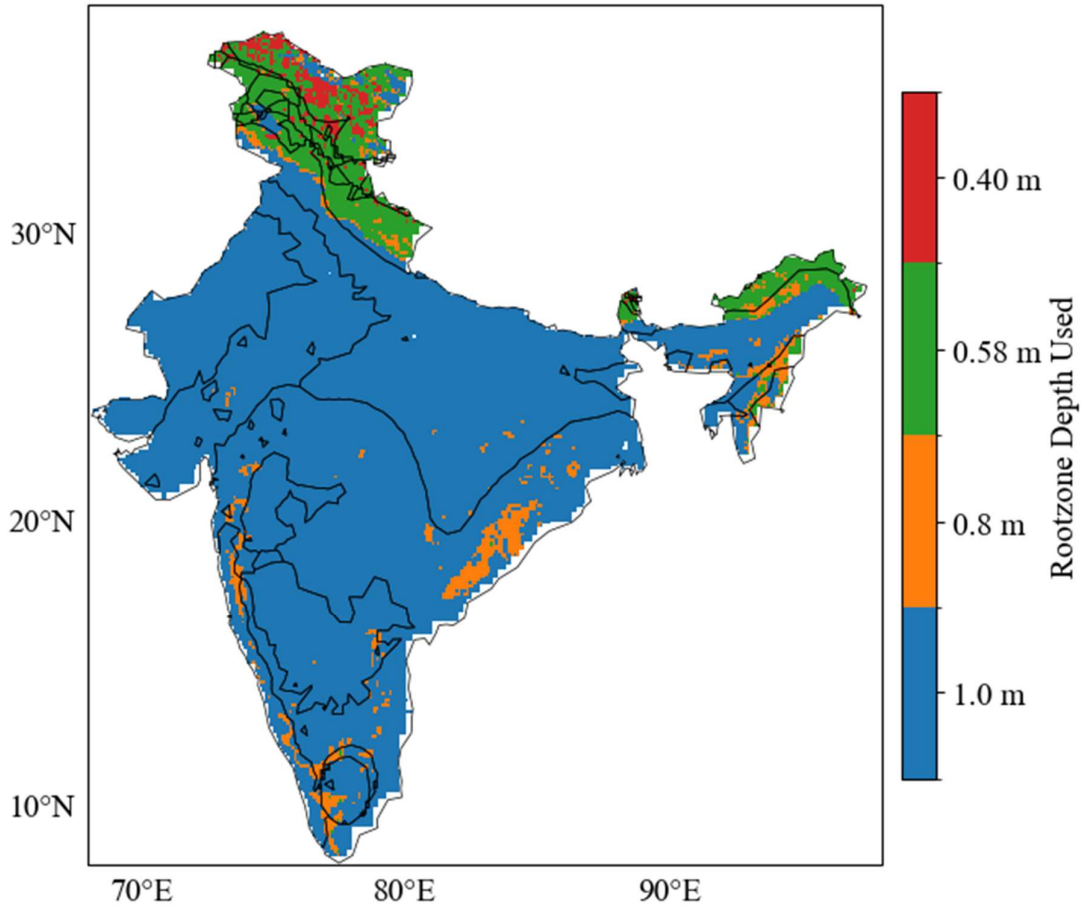


Figure S5. Spatial distribution of rootzone soil moisture (RZSM) depths used in the analysis, adjusted for bedrock depth across India. Color codes represent the effective rootzone depths applied at each grid cell: 0.40 m (red), 0.58 m (green), 0.80 m (orange), and 1.00 m (blue), based on local soil and bedrock constraints.

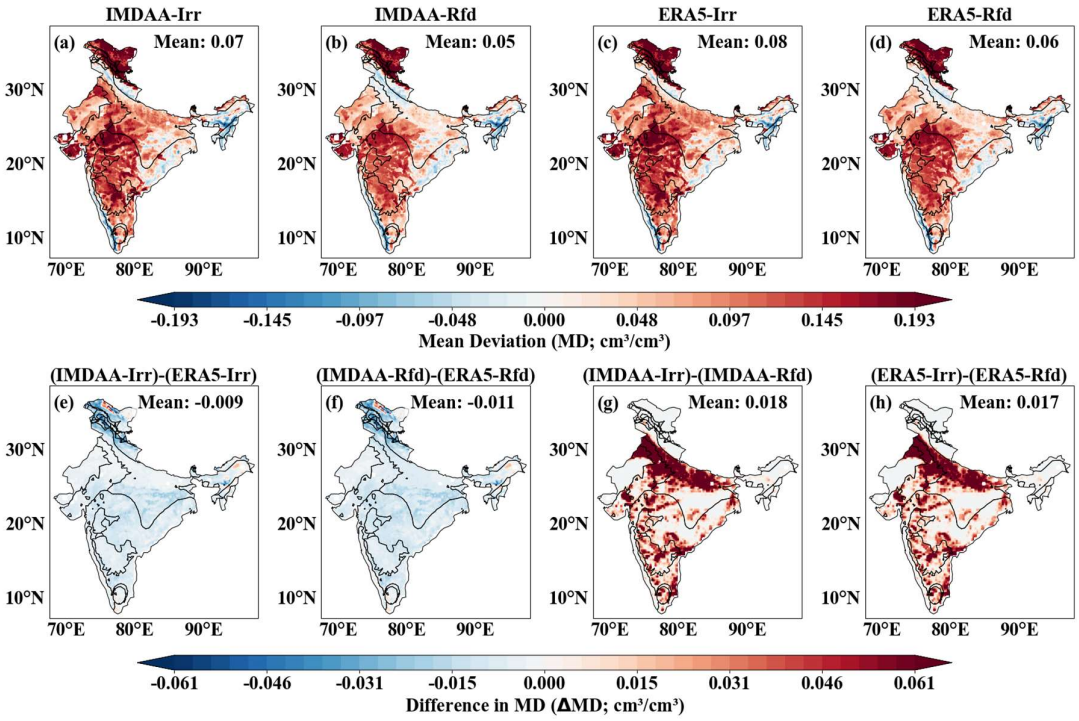


Figure S6: Grid-wise Mean Deviation (MD) of SSM (10 cm) simulated by CLM5 relative to GLEAM across India for 1980–2020. Panels (a–d) show MD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in MD (ΔMD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

Figure S6 presents the grid-wise Mean Deviation (MD) patterns of CLM5-simulated surface soil moisture (SSM, 10 cm) relative to GLEAM. The spatial bias structure closely resembles the SMAP-based assessment, with widespread wet biases across northwest, central, peninsular, and parts of northern and northeastern India, and negative biases along the eastern and western coasts and portions of northeastern India. Rainfed simulations generally exhibit smaller biases, whereas irrigated configurations reinforce wet biases across irrigated regions (Figure S6g–h). Among the experiments, IMDAA-driven runs show the lowest overall MD (Figure S6e–f), indicating improved bias representation relative to ERA5-driven simulations.

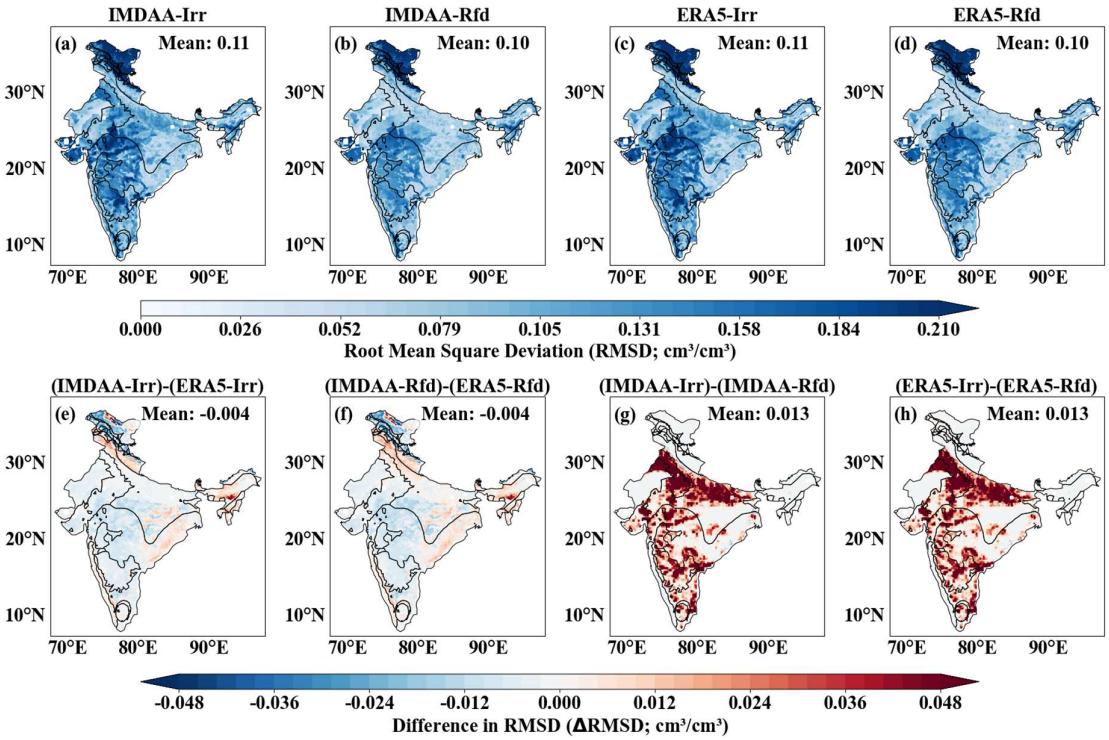


Figure S7: Grid-wise Root Mean Square Deviation (RMSD) of SM (10 cm) simulated by CLM5 relative to GLEAM across India for 1980–2020. Panels (a–d) show RMSD for IMDAA-Irr, IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in RMSD (ΔRMSD) highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by contrasting irrigated and rainfed configurations.

Figure S7 presents the grid-wise Root Mean Square Deviation (RMSD) of CLM5-simulated surface soil moisture relative to GLEAM. Mean RMSD is slightly lower for rainfed configurations (IMDAA-Rfd: 0.10; ERA5-Rfd: 0.10) than for irrigated runs (IMDAA-Irr: 0.11; ERA5-Irr: 0.11), indicating improved performance under rainfed conditions, consistent with the SMAP-based evaluation. Higher RMSD values (>0.11 $cm^3\ cm^{-3}$) persist across northeastern, central, southern, and northern regions, with errors generally more pronounced in irrigated simulations. The ΔRMSD patterns (Figure S7e–h) further mirror the SMAP results, with IMDAA-driven runs exhibiting lower RMSD across large parts of central and southern India, portions of northwestern India, and localized northern regions ($\Delta RMSD \approx -0.004\ cm^3\ cm^{-3}$), whereas ERA5-driven simulations perform better along the southeastern coast, northeastern India, and parts of the Indo-Gangetic Plain. These patterns are quantified in Supplementary Figure S8, where IMDAA-Rfd achieves the lowest RMSD at 77% of grid cells. Over irrigated regions, higher RMSD dominates most grid cells (Figure S7g–h), although IMDAA-Rfd still performs best at 77.2% of irrigated locations, particularly across central and

peninsular India (Figure S9). As in the SMAP comparison, rainfed simulations outperform irrigated configurations across much of the Indo-Gangetic Plain. Correlation maps (Figure S19) show strong temporal agreement between CLM5 simulations and GLEAM across most of India ($\rho > 0.8$), further supporting the robustness of the SMAP-based performance assessment.

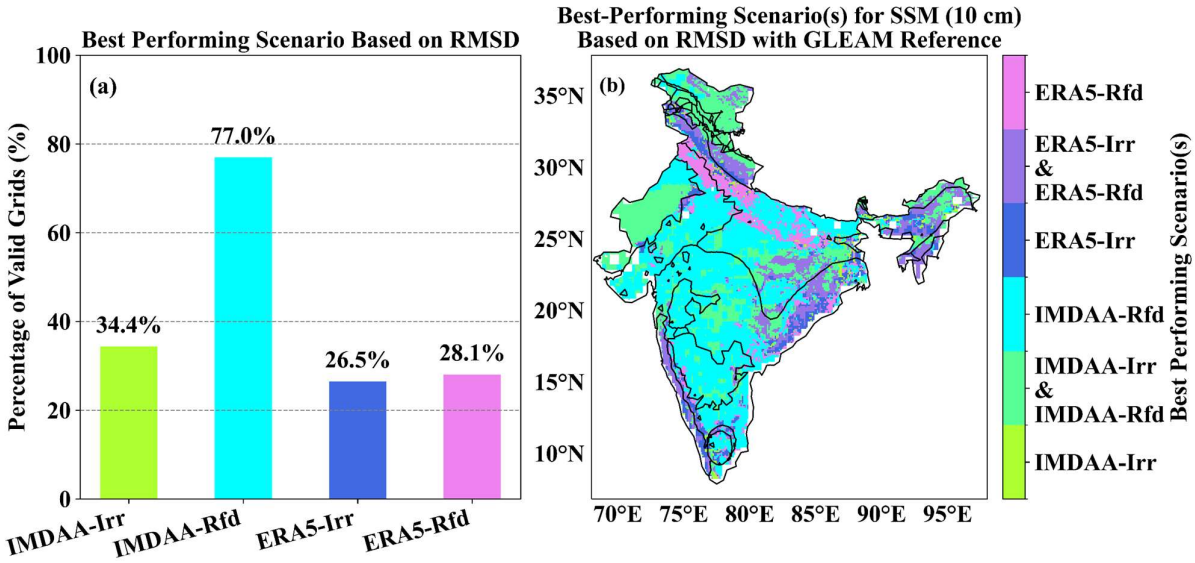


Figure S8: (a) Percentage of grid cells where each CLM5 simulations of surface soil moisture (SSM, 10 cm) achieves the lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across grid cells.

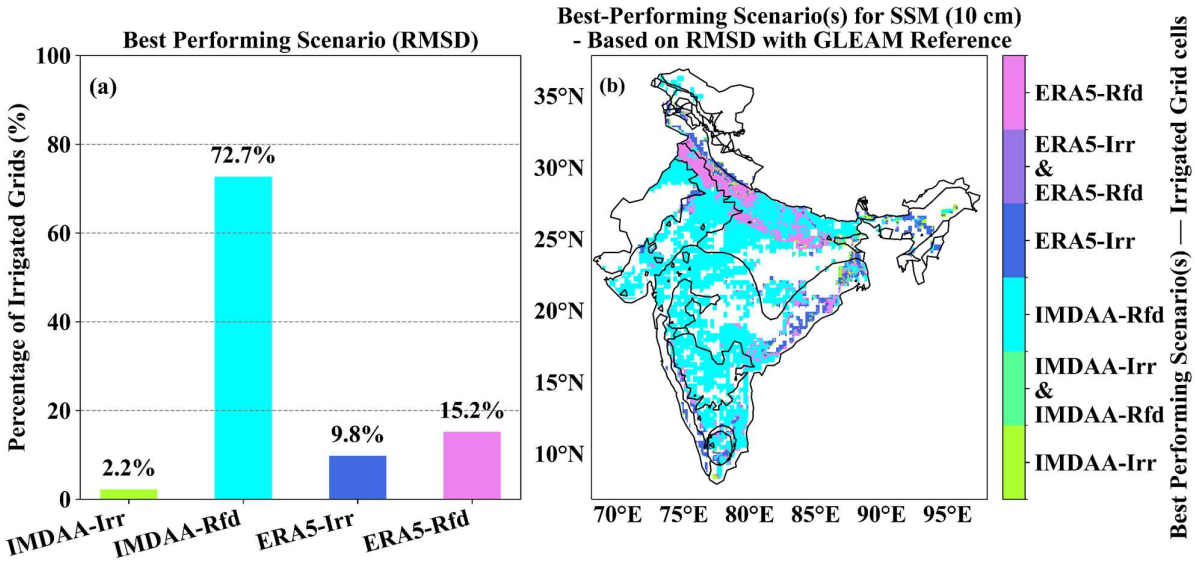
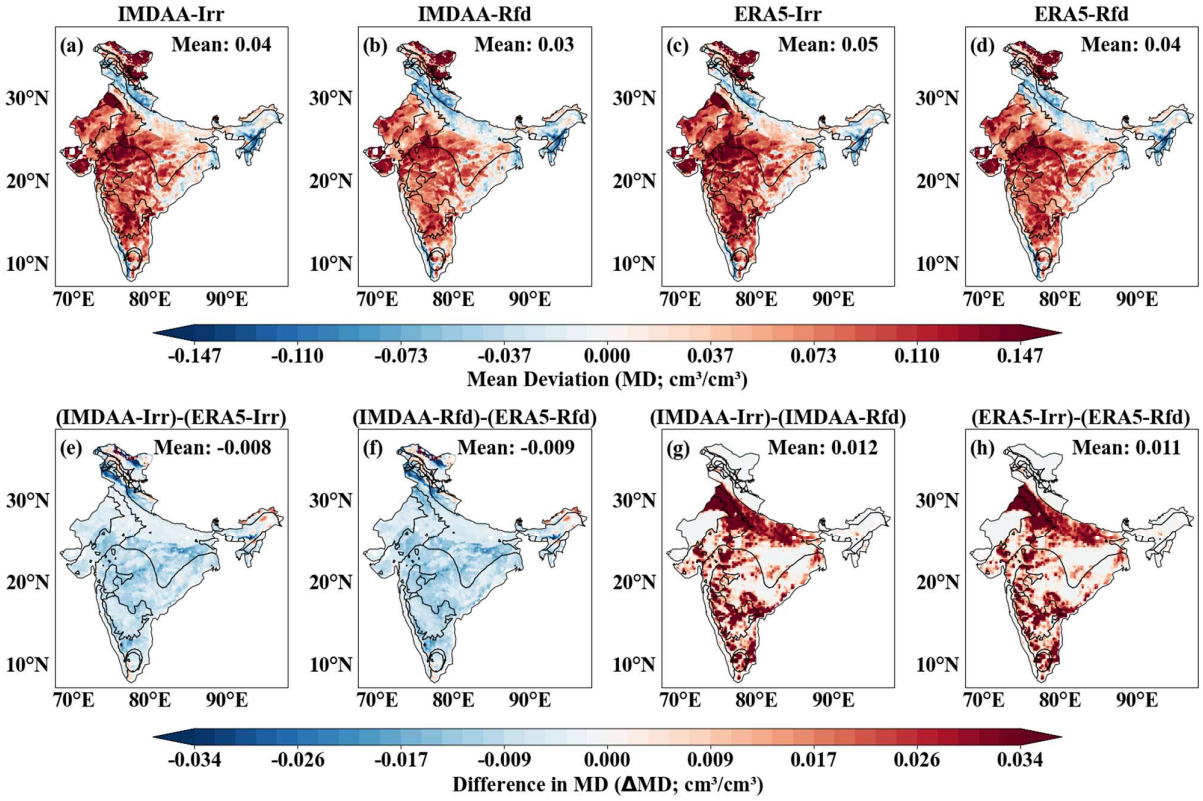


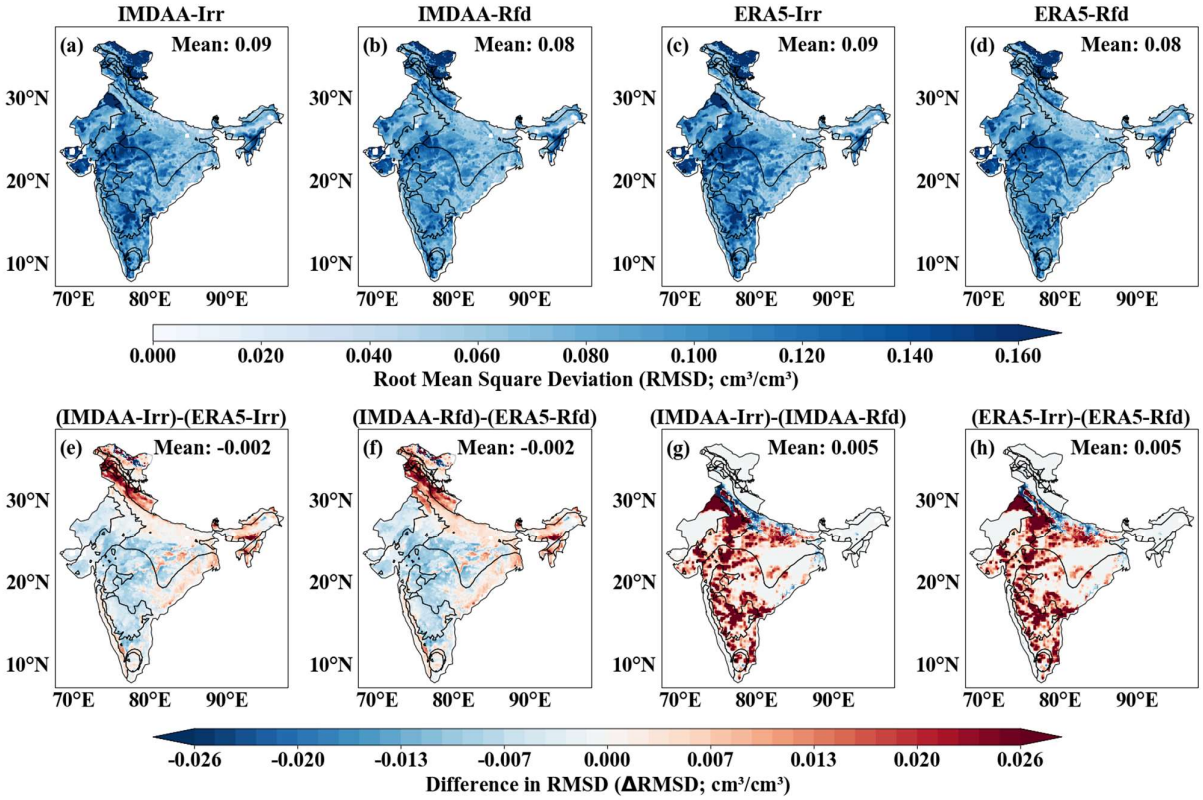
Figure S9. (a) Percentage of irrigated grid cells where each CLM5 simulations of soil moisture (SSM) at 10 cm depth shows the lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across irrigated grid cells.



104

105 **Figure S10:** Grid-wise Mean Deviation (MD) of RZSM (100 cm) simulated by CLM5 relative to
106 GLEAM across India for 1980–2020. Panels (a–d) show MD for IMDAA-Irr, IMDAA-Rfd, ERA5-
107 Irr, and ERA5-Rfd. Panels (e–f) present differences in MD (ΔMD) highlighting the influence of
108 atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate the impact of irrigation by
109 contrasting irrigated and rainfed configurations.

110



111

112 **Figure S11:** Grid-wise Root Mean Square Deviation (RMSD) of RZSM (100 cm) simulated by
113 CLM5 relative to GLEAM across India for 1980–2020. Panels (a–d) show RMSD for IMDAA-Irr,
114 IMDAA-Rfd, ERA5-Irr, and ERA5-Rfd. Panels (e–f) present differences in RMSD (ΔRMSD)
115 highlighting the influence of atmospheric forcing (IMDAA vs. ERA5), while panels (g–h) illustrate
116 the impact of irrigation by contrasting irrigated and rainfed configurations.

117

118

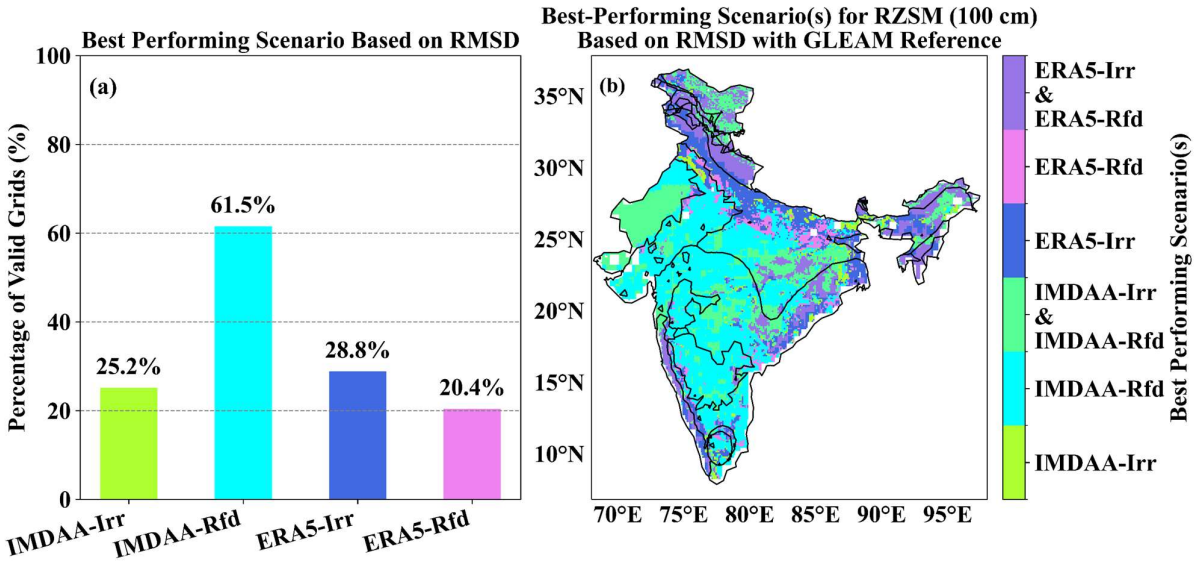


Figure S12: (a) Percentage of grid cells where each CLM5 simulations of RZSM (100 cm) achieves the lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across grid cells.

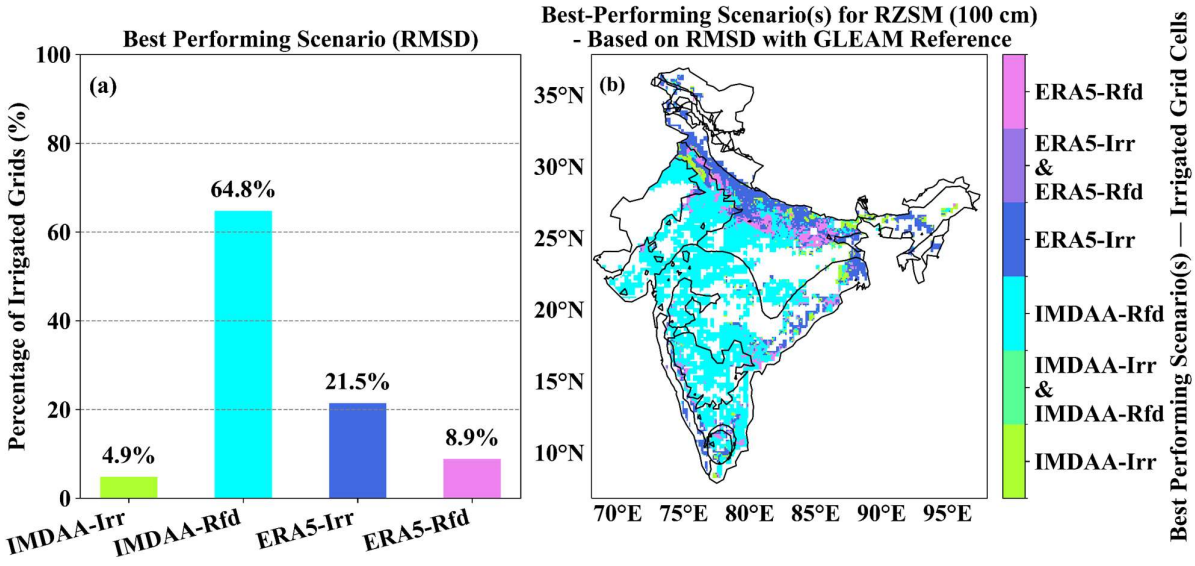


Figure S13. (a) Percentage of irrigated grid cells where each CLM5 simulations of RZSM (100 cm) shows the lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across irrigated grid cells.

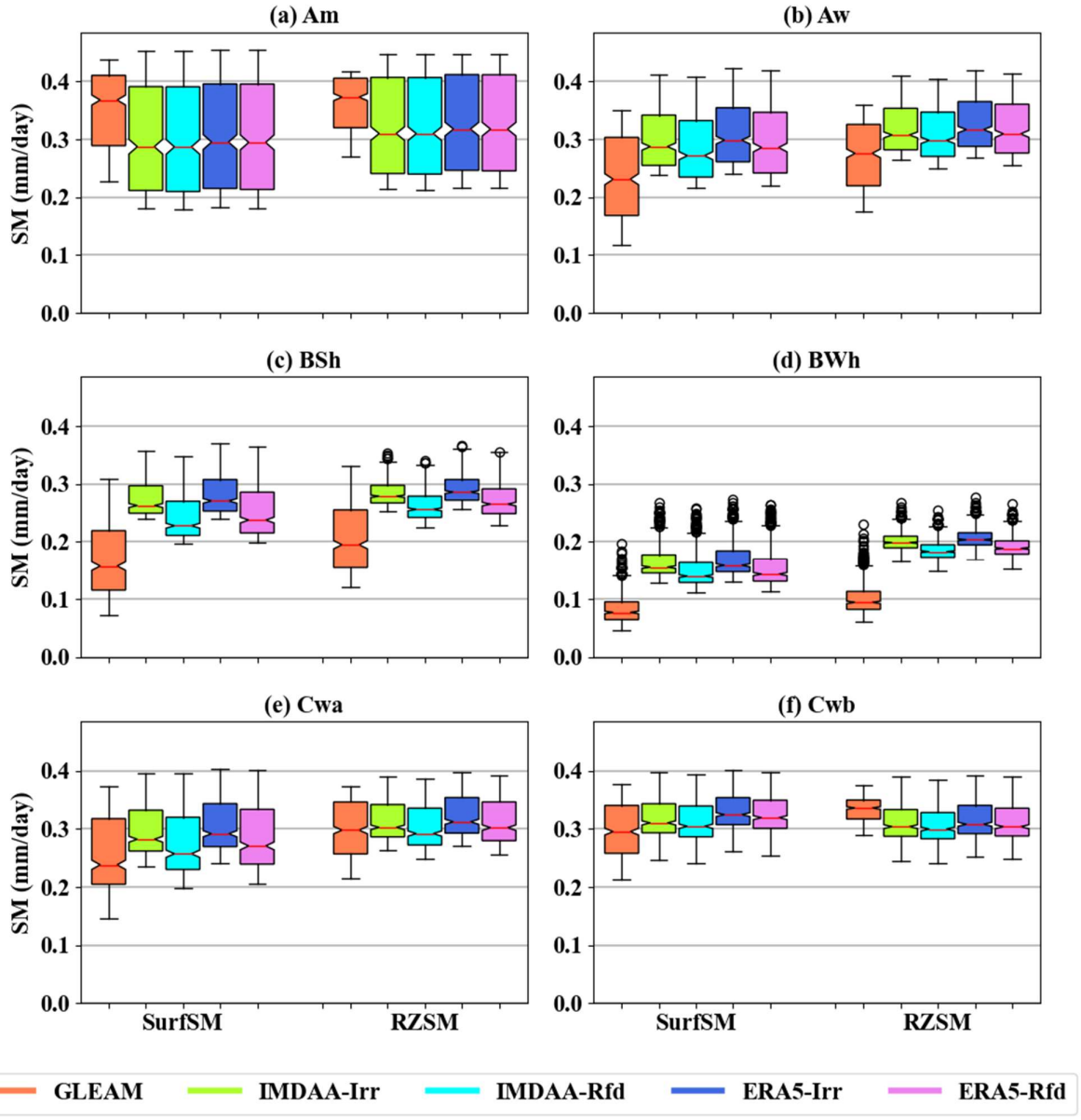


Figure S14: Comparison of CLM5SP-simulated soil moisture with GLEAM observations across six climate zones. Panels (a–f) show boxplots over Am, Aw, BSh, BWh, Cwa, and Cwb zones. In the boxplots, the line inside each box indicates the median, the box spans the interquartile range (25th–75th percentile), whiskers extend to 1.5×IQR, and outliers are shown as circles.

A climate zone–wise evaluation of CLM5-simulated soil moisture at 10 cm (SSM) and 100 cm (RZSM) depths against GLEAM is presented in Figure S14 and Tables S3–S4. RMSE values range from 0.03–0.12 $m^3\ m^{-3}$ for SSM and 0.02–0.11 $m^3\ m^{-3}$ for RZSM, with the lowest errors consistently observed in the humid subtropical zones (Cwa and Cwb). IMDAA-Rfd generally shows the best performance across climate zones; however, at Am for SSM all simulations perform similarly, while

for RZSM ERA5-based simulations in Am and both ERA5 and IMDAA-Irr in Cwb exhibit comparable or improved performance.

For SSM, mean deviation (MD) is predominantly positive (0.02–0.12 $m^3 m^{-3}$), with the largest wet biases occurring in semi-arid and arid regions (BSh and BWh), whereas the Am zone shows a notable dry bias. Rainfed simulations (IMDAA-Rfd and ERA5-Rfd) generally yield lower biases than irrigated runs, except in the Am zone for SSM.

Spearman correlation coefficients range from 0.76–0.97 for SSM, indicating strong temporal agreement overall. Lower correlations are observed in the BWh and Cwb zones, particularly for RZSM.

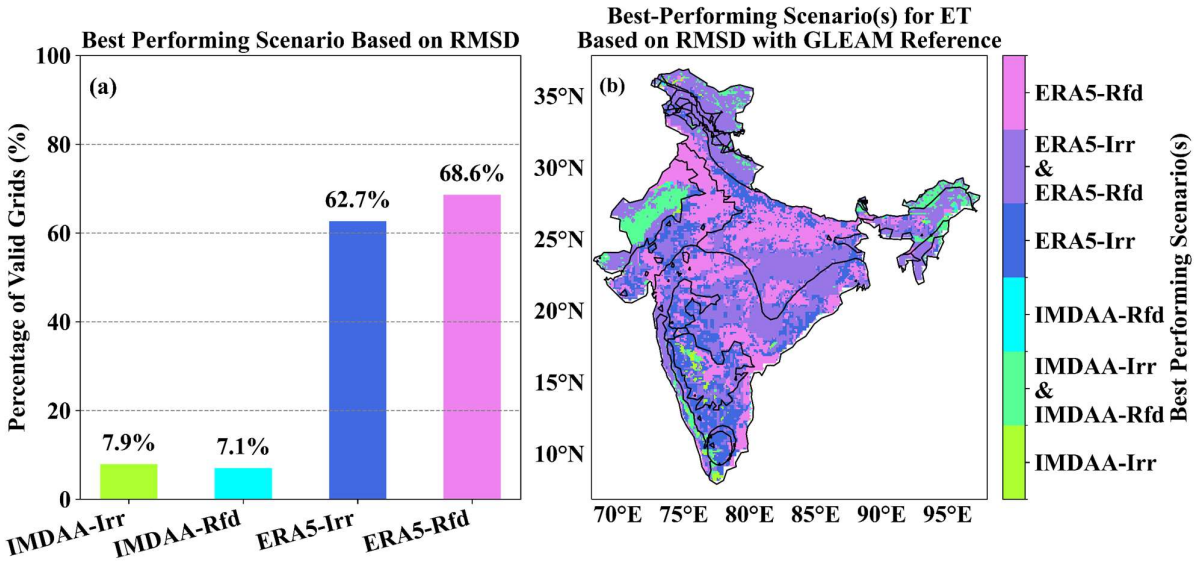


Figure S15: (a) Percentage of grid cells where each CLM5 simulations of ET achieves the lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across grid cells.

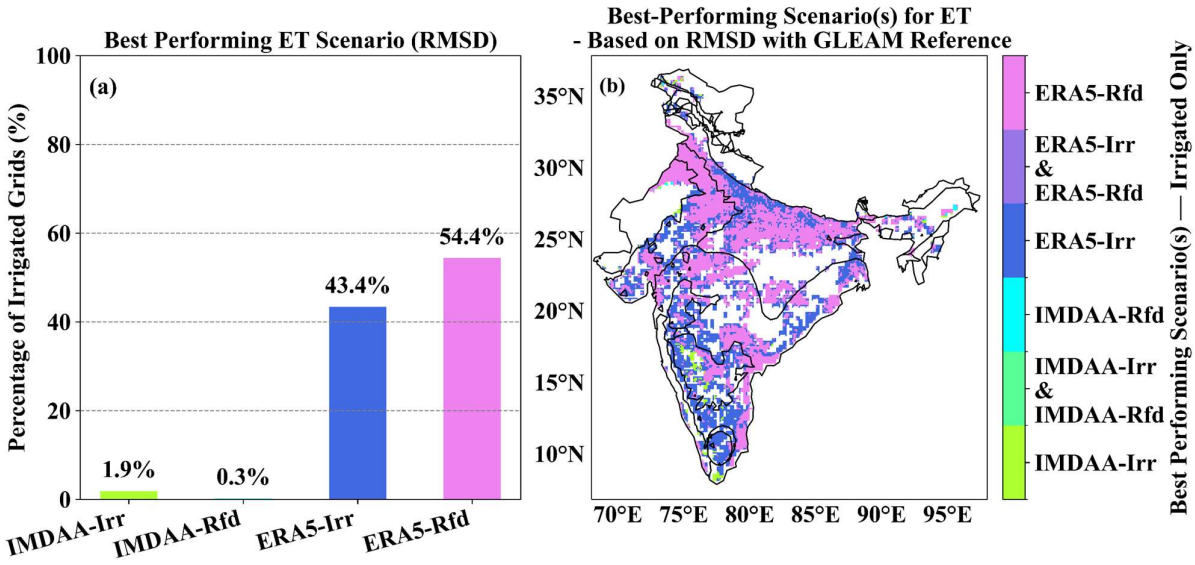


Figure S16. (a) Percentage of irrigated grid cells where each CLM5 simulations of ET shows the lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across irrigated grid cells.

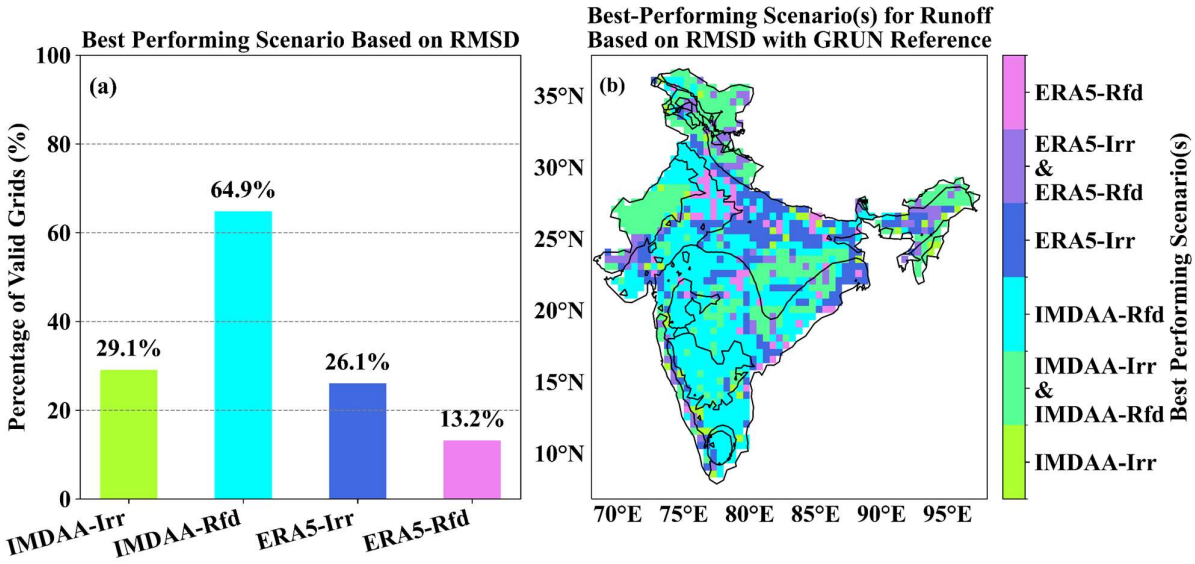
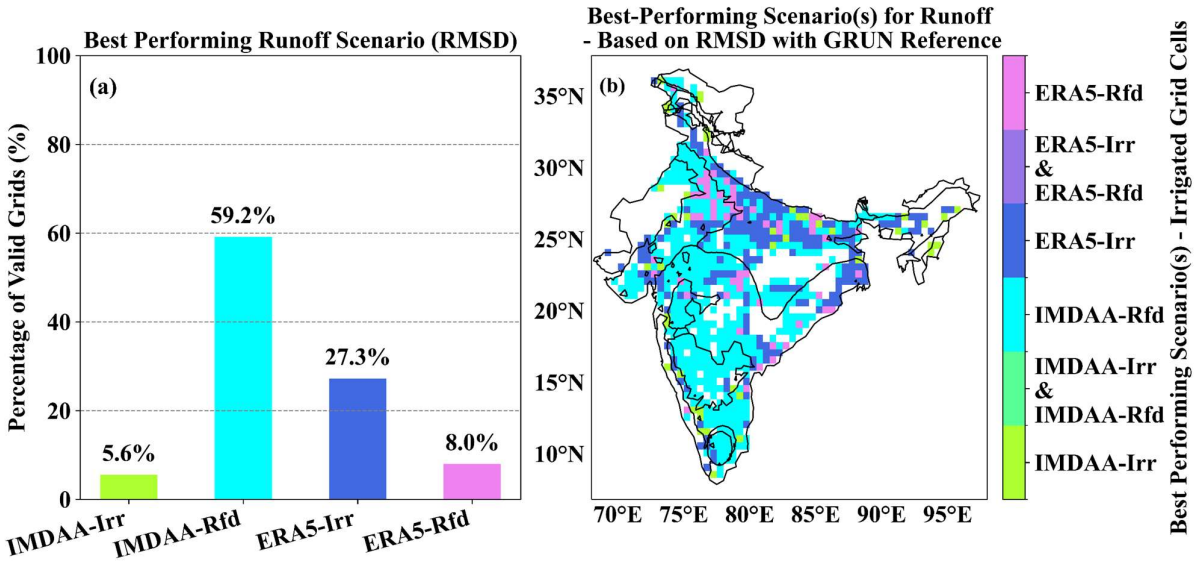


Figure S17: (a) Percentage of grid cells where each CLM5 simulations of Runoff achieves the lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across grid cells.



169

170 **Figure S18.** (a) Percentage of irrigated grid cells where each CLM5 simulations of Runoff shows the
171 lowest RMSD relative to GLEAM. (b) Spatial distribution of the best-performing simulation across
172 irrigated grid cells.

173

174

175

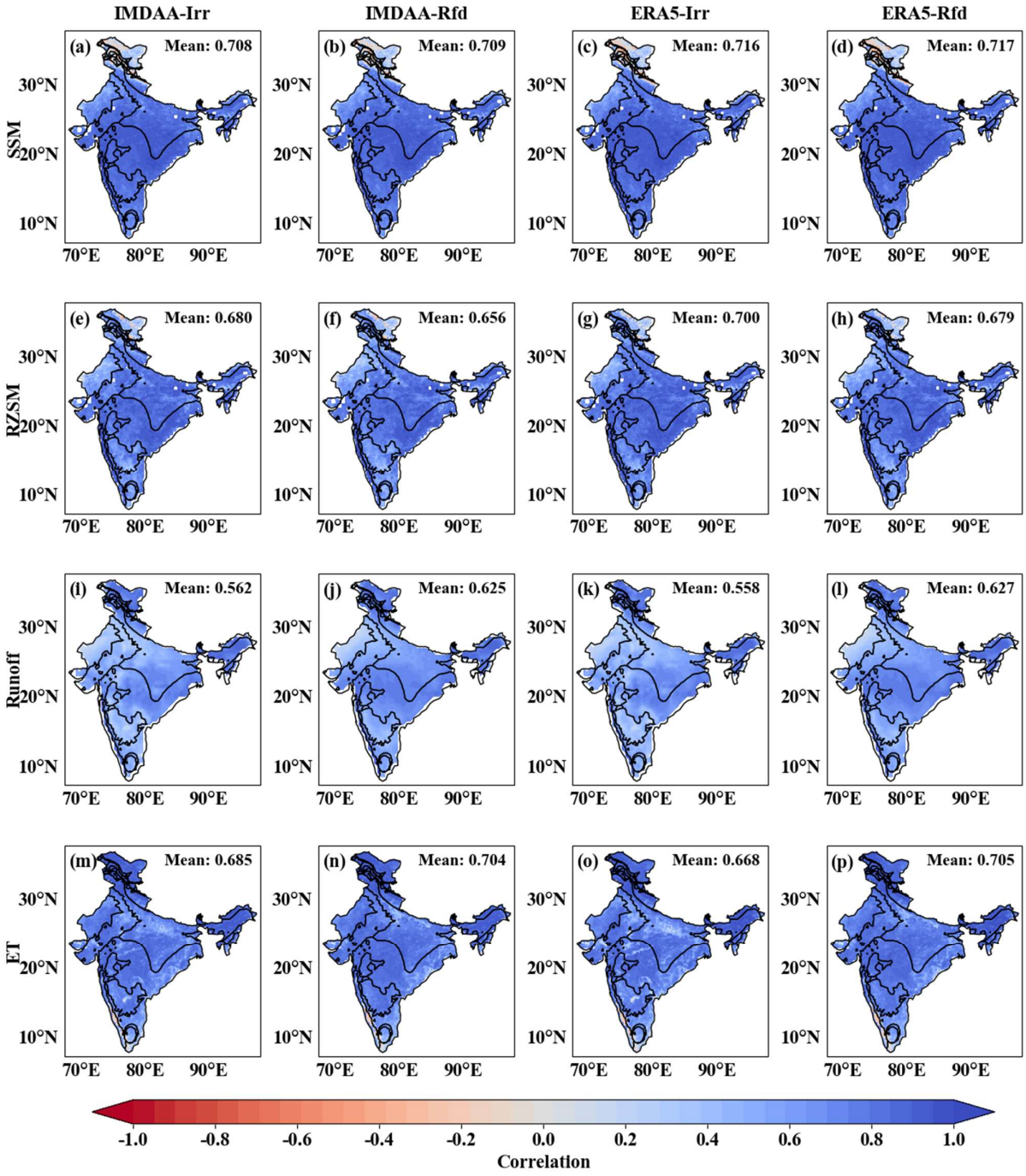


Figure S19: Grid-wise Spearman-Correlation analysis of CLM5 simulations against GLEAM SSM, RZSM, ET, and GRUN Runoff across India.

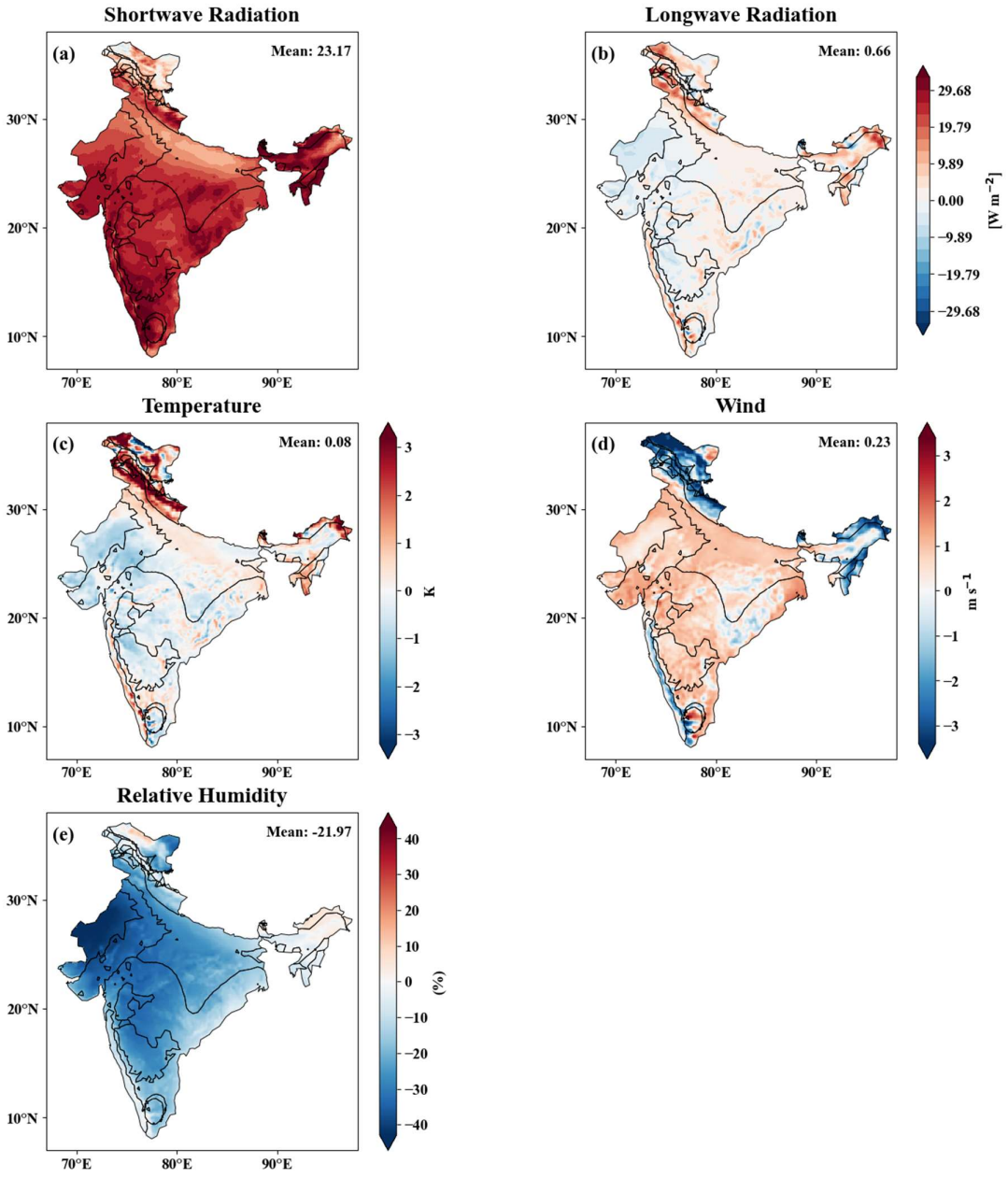
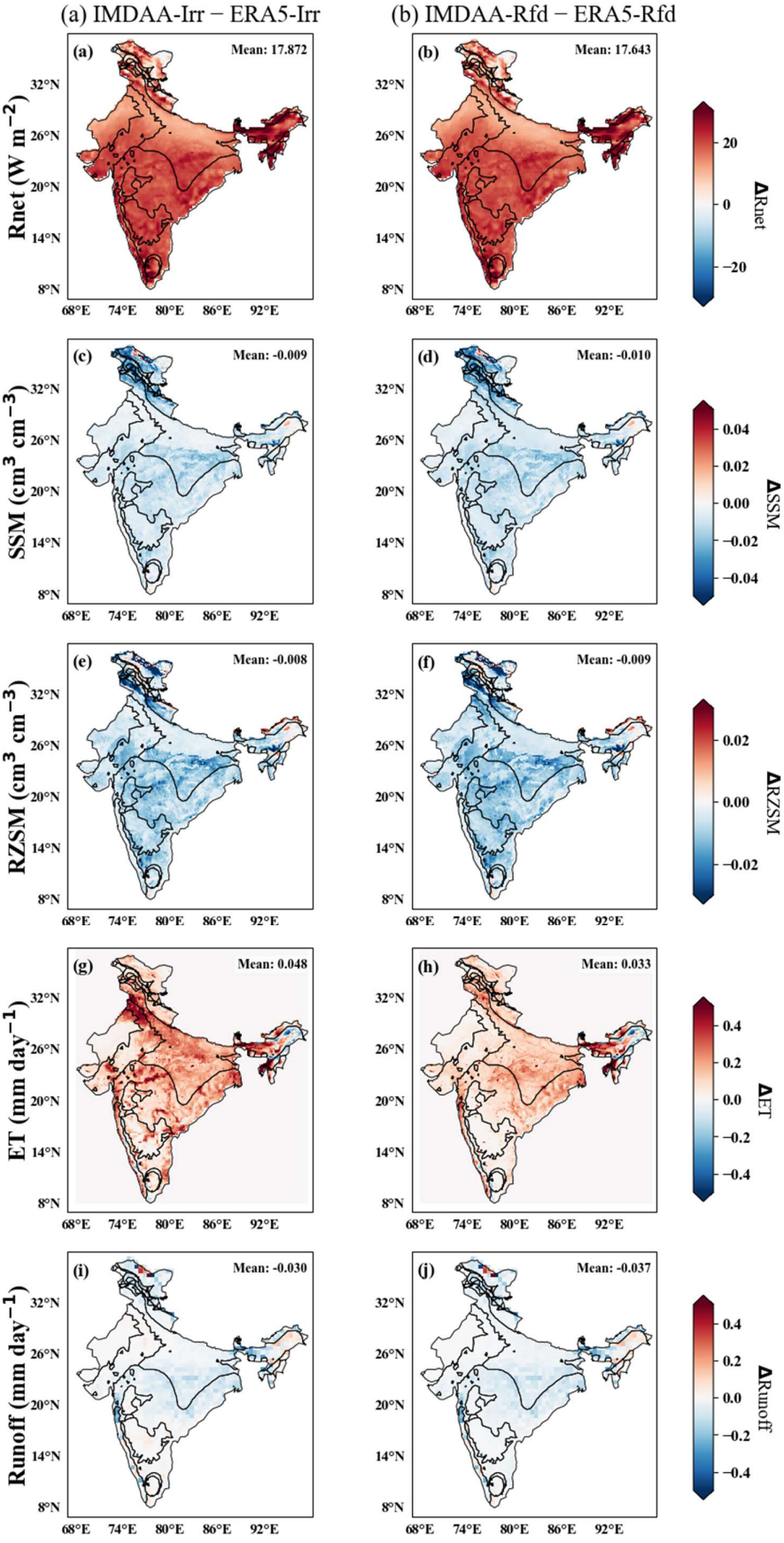


Figure S20. shows the spatial distribution of the differences between IMDAA and ERA5 forcings of Panels (a) Shortwave radiation, (b) Longwave radiation, (c) Temperature at 2m, (d) Wind, and (e) Relative humidity.


(a) IMDAA-Irr − ERA5-Irr
(b) IMDAA-Rfd − ERA5-Rfd
(a)
Mean: 17.872
(b)
Mean: 17.643
Rnet (W m−2)
ΔRnet
(c)
Mean: -0.009
(d)
Mean: -0.010
SSM (cm3 cm−3)
ΔSSM
(e)
Mean: -0.008
(f)
Mean: -0.009
RZSM (cm3 cm−3)
ΔRZSM
(g)
Mean: 0.048
(h)
Mean: 0.033
ET (mm day−1)
ΔET
(i)
Mean: -0.030
(j)
Mean: -0.037
Runoff (mm day−1)
ΔRunoff
32°N
26°N
20°N
14°N
8°N
68°E
74°E
80°E
86°E
92°E

185

Figure S21. shows the spatial distribution of the irrigation-induced differences in various variables over India for the period 1980–2020. The differences were calculated by subtracting rainfed from irrigated simulations. Panels (a) and (b) represent the Rnet differences, while (c) and (d) show the differences for SSM. Panels (e) and (f) correspond to RZSM, and (g) and (h) show the differences in ET. Lastly, runoff differences are presented in panels (i) and (j). Panels with odd letters (a, c, e, g, i) display the differences derived from IMDAA atmospheric forcing (IMDAA-Irr – IMDAA-Rfd), whereas panels with even letters (b, d, f, h, j) are for the ERA5 forcing (ERA5-Irr – ERA5-Rfd).

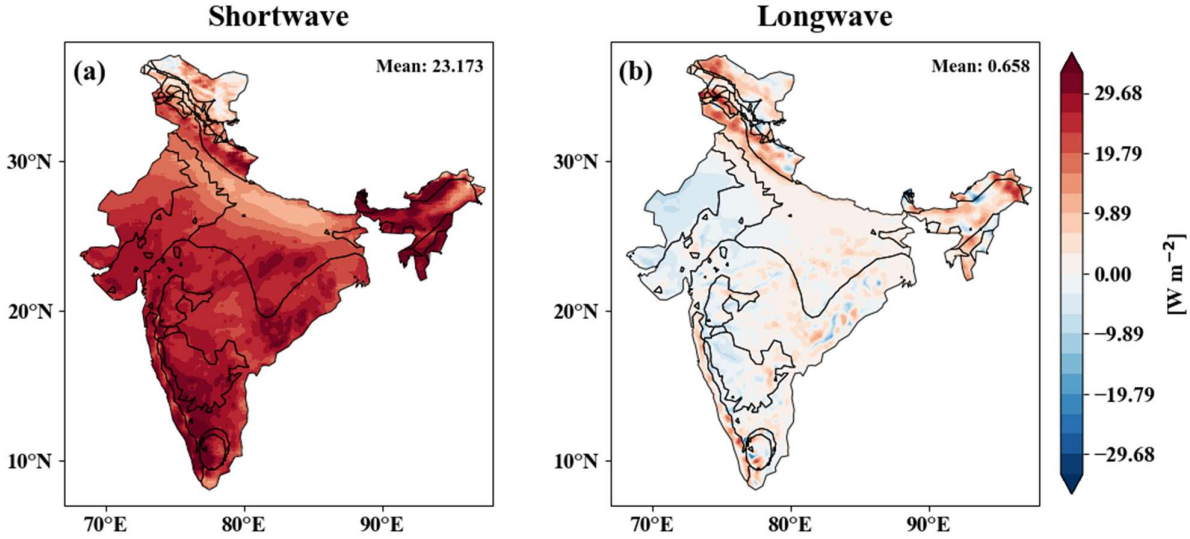


Figure S22. Spatial distribution of the difference in (a) Shortwave Radiation and (b) Longwave Radiation between the IMDAA and ERA5 atmospheric forcings.

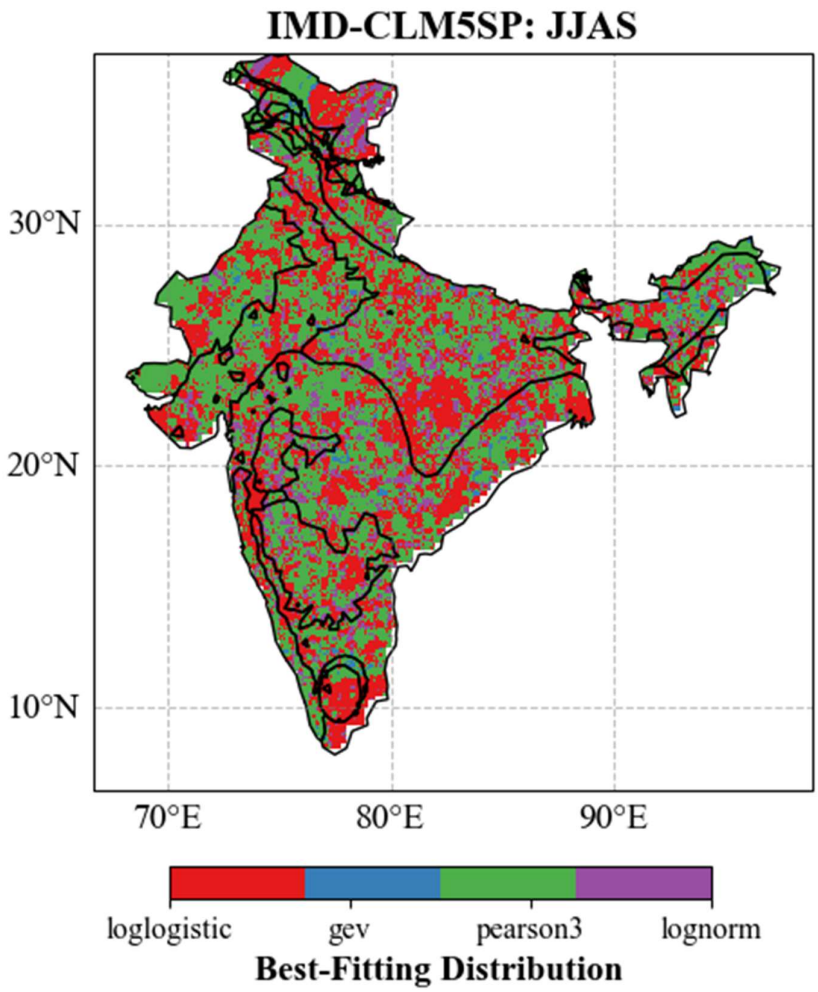


Figure S23: Best-Fitted Distribution of Net Precipitation (P-PET) Monsoon (JJAS:sum) data

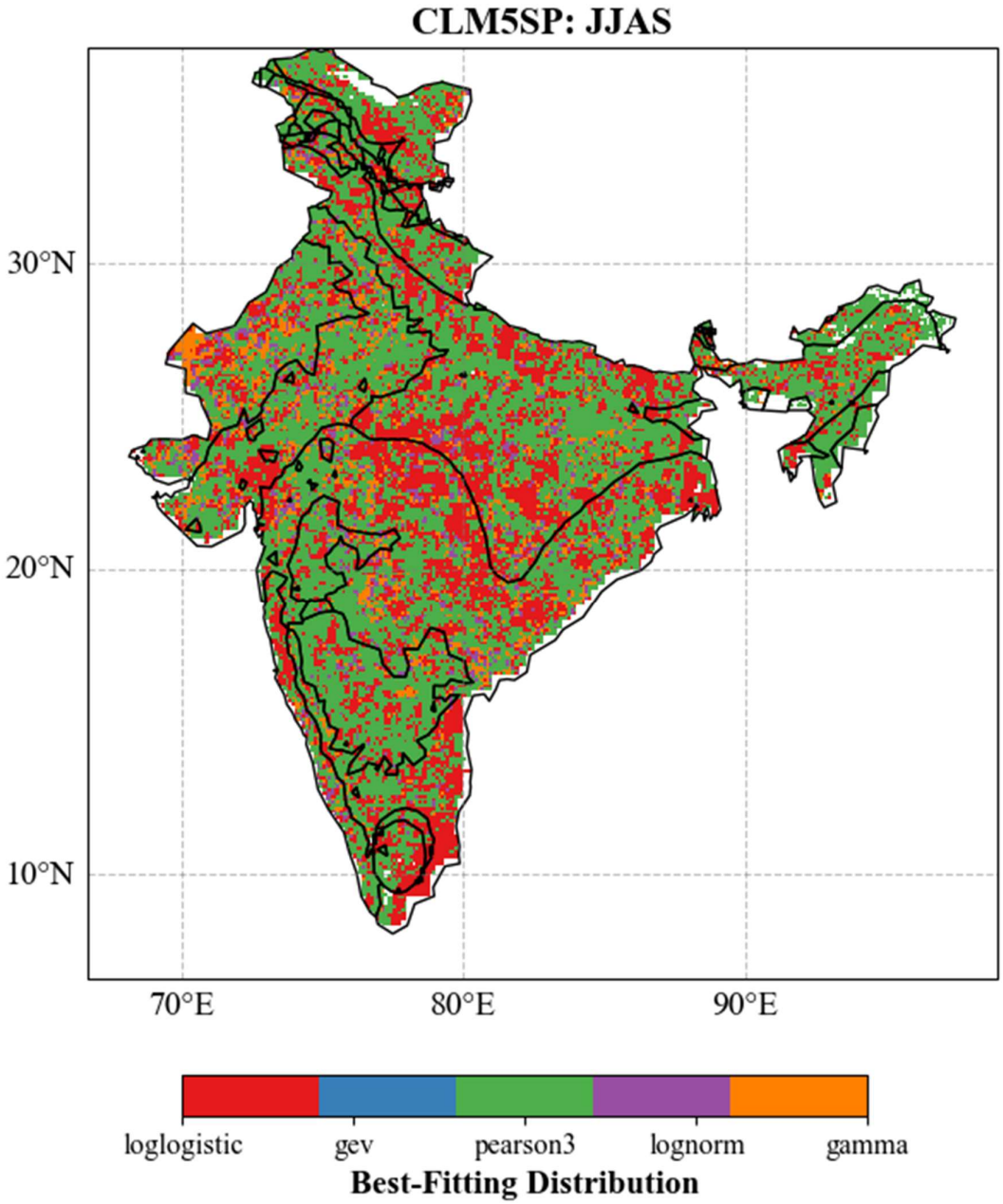


Figure S24: Best-Fitted Distribution of soil moisture Monsoon (JJAS:sum) data

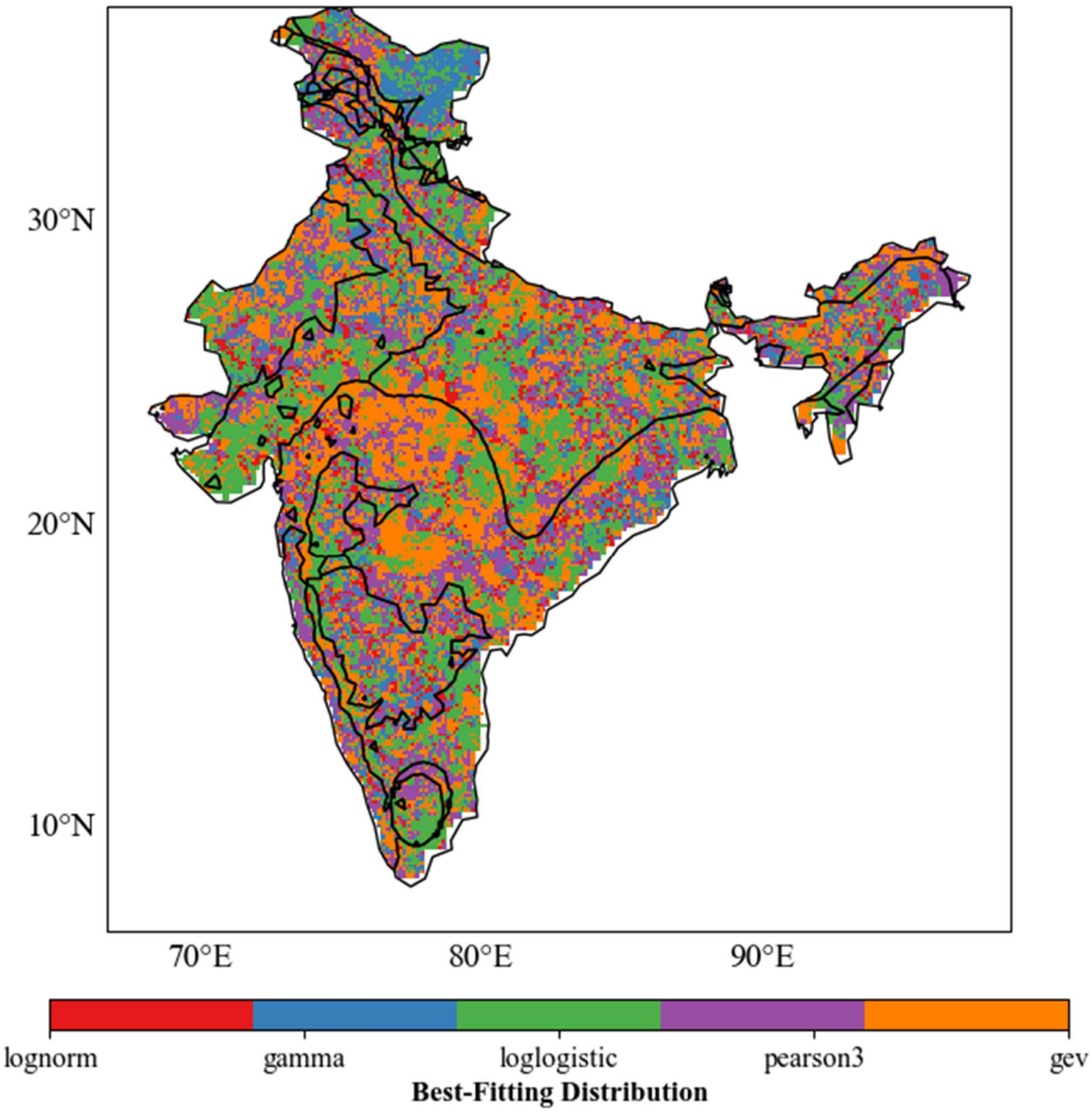


Figure S25: Best-Fitted Distribution of CLM5SP runoff Water year (June-May:sum) data

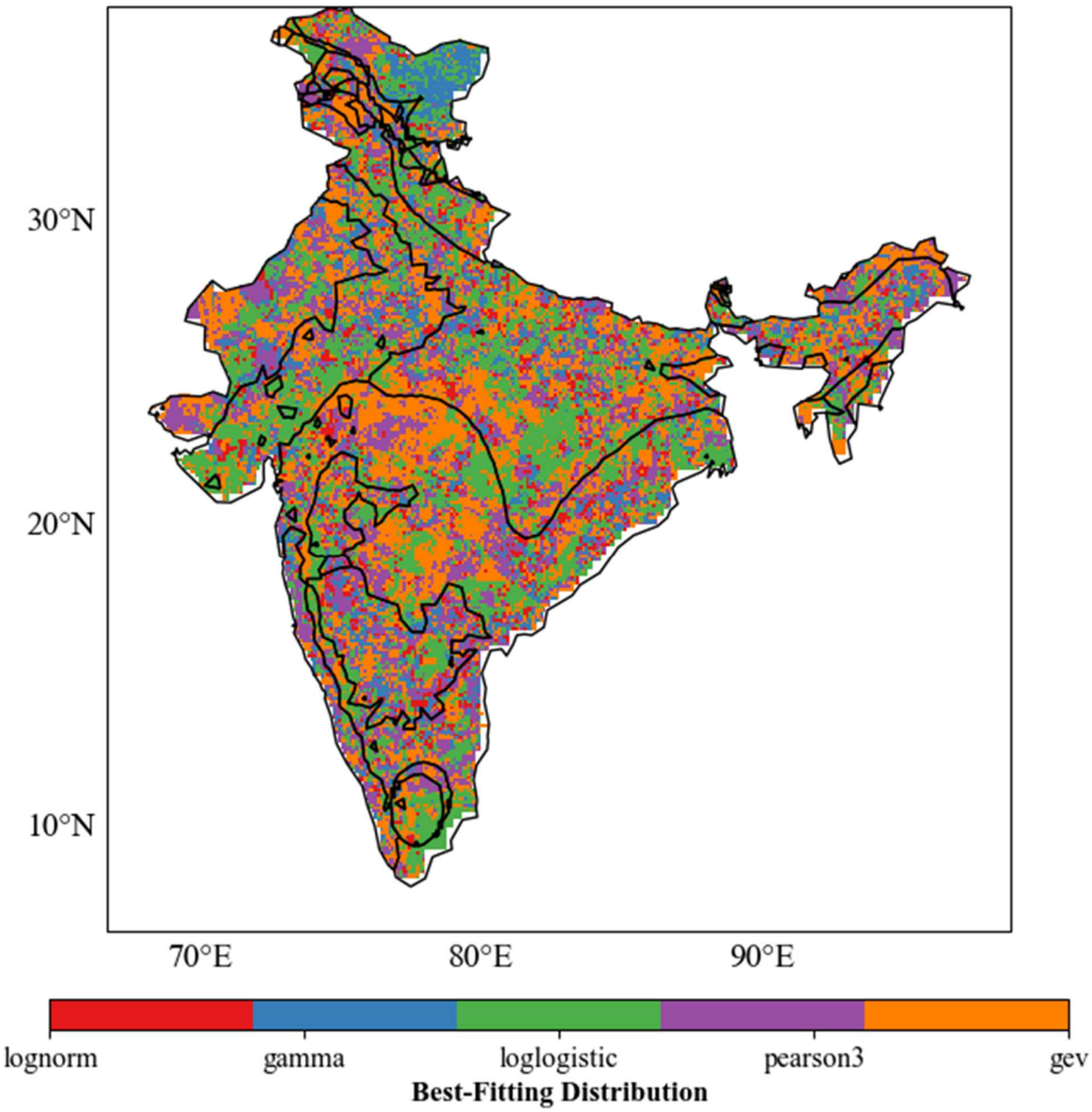


Figure S26: Best-Fitted Distribution of CLM5SP runoff monsoon (June-Sep:sum) data

Table S1: Details of the in-situ soil moisture (COSMOS sensors) stations used in this study.

More details can be found in (Upadhyaya et al., 2021).

S. No	Acronym	Latitude	Longitude	City, State	Data availability	Number of Observations
1	BMB	11.76^0	76.58^0	Berambadi, Karnataka	20-Sep-2015 to 19-Dec-2019	490
2	MDH	11.73^0	76.78^0	Madahalli, Karnataka	05-Feb-2016 to 19-Dec-2019	1193
3	SGR	12.14^0	77.22^0	Singanallur, Karnataka	9-Jun-2015 to 19-Dec-2019	1513

Table S2: Distribution of selected rootzone soil moisture depths used across 28,941 grid points

S. No	Depth (m)	Grid Count	Share of Total (%)
1	1.00	23,716	81.946
2	0.80	1467	5.068
3	0.50	2814	9.723
4	0.40	494	1.706
Total Grid points		**28,941**	

Table S3: Performance metrics (RMSE, ME, and ρ) for soil moisture vs GLEAM (SM: 10 cm).

Climate Zone	Metric	IMDAA-Irr	IMDAA-Rfd	ERA5-Irr	ERA5-Rfd
Am	MD (cm^3/cm^3)	-0.05	-0.05	-0.05	-0.05
	RMSD (cm^3/cm^3)	0.06	0.06	0.06	0.06
	ρ	0.97	0.97	0.97	0.97
Aw	MD (cm^3/cm^3)	0.07	0.05	0.07	0.06
	RMSD (cm^3/cm^3)	0.07	0.06	0.08	0.06
	ρ	0.97	0.97	0.97	0.97
BSh	MD (cm^3/cm^3)	0.11	0.07	0.11	0.08
	RMSD (cm^3/cm^3)	0.11	0.08	0.12	0.09
	ρ	0.93	0.93	0.95	0.95
BWh	MD (cm^3/cm^3)	0.08	0.07	0.09	0.07
	RMSD (cm^3/cm^3)	0.08	0.07	0.09	0.07
	ρ	0.84	0.84	0.85	0.84
Cwa	MD (cm^3/cm^3)	0.04	0.02	0.05	0.03
	RMSD (cm^3/cm^3)	0.05	0.03	0.06	0.04
	ρ	0.93	0.93	0.95	0.95
Cwb	MD (cm^3/cm^3)	0.02	0.02	0.03	0.03
	RMSD (cm^3/cm^3)	0.03	0.03	0.04	0.04
	ρ	0.76	0.77	0.77	0.78

Table S4: Performance metrics (RMSE, ME, and ρ) for soil moisture vs GLEAM (SM: 100 cm).

Climate Zone	Metric	IMDAA-Irr	IMDAA-Rfd	ERA5-Irr	ERA5-Rfd
Am	MD (cm^3/cm^3)	-0.04	-0.04	-0.04	-0.04
	RMSD (cm^3/cm^3)	0.06	0.06	0.05	0.05
	ρ	0.95	0.95	0.95	0.95
Aw	MD (cm^3/cm^3)	0.05	0.04	0.05	0.05
	RMSD (cm^3/cm^3)	0.05	0.04	0.06	0.05
	ρ	0.94	0.94	0.95	0.94
BSh	MD (cm^3/cm^3)	0.08	0.06	0.09	0.07
	RMSD (cm^3/cm^3)	0.09	0.07	0.09	0.07
	ρ	0.92	0.90	0.93	0.91
BWh	MD (cm^3/cm^3)	0.10	0.08	0.10	0.09
	RMSD (cm^3/cm^3)	0.10	0.09	0.10	0.09
	ρ	0.79	0.77	0.80	0.78
Cwa	MD (cm^3/cm^3)	0.01	0.00	0.02	0.01
	RMSD (cm^3/cm^3)	0.03	0.02	0.03	0.02
	ρ	0.92	0.91	0.93	0.92
Cwb	MD (cm^3/cm^3)	-0.02	-0.03	-0.02	-0.02
	RMSD (cm^3/cm^3)	0.03	0.04	0.03	0.03
	ρ	0.63	0.63	0.67	0.68

Table S5: Summary of data used for CLM5 validation.

Variable	Data Source	Data period	Spatial Resolution	Temporal Resolution	Data Source
ET	GLEAM	1980 - 2020	0.25^0	Daily	www.gleam.edu
Runoff	GRUN	1980 - 2014	0.5^0	Monthly	GRUN: Global Runoff Reconstruction
SM	GLEAM	1980 – 2020	0.25^0	Daily	www.gleam.edu
SM	COSMOS	2015 – 2019	3 sites, at 5 cm depth	Daily	https://cosmos-india.org/
SM	SMAP	2015-2020	~36 km	Daily	https://appeears.earthdatacloud.nasa.gov/
SM	GLDAS	1980-2014	0.25^0	3-hourly	https://disc.gsfc.nasa.gov/datasets

Table S6. Annual averages of meteorological elements from different meteorological forcing datasets for the period of 1980–2020 over India

Variable	IMDAA	ERA5	Difference (IMDAA-ERA5)
Shortwave radiation (W/m2)	240.30	217.13	23.17
Downward longwave radiation (W/m2)	365.19	364.53	0.66
Air Temperature (K)	295.58	295.50	0.08
Wind speed (m/s)	5.71	5.48	0.23
Relative Humidity (RH: %)	61.70	83.67	-21.97