Supplemental Material

Journal of Climate
IMDAA: High-Resolution Satellite-Era Reanalysis for the Indian Monsoon Region
https://doi.org/10.1175/JCLI-D-20-0412.1

Supplementary Materials

Use of variational bias correction (VarBC) in the IMDAA system

The use of VarBC for satellite radiance is one of the uniqueness of the IMDAA regional reanalysis system. VarBC can be problematic for regional domains due to possible sampling issues when updating the coefficients for the predictors; however, many regional and convective scale NWP assimilation systems use the VarBC method for bias correction of satellite radiances (Gustafsson et al., 2017). The novel aspects of the VarBC system (Cameron and Bell, 2018) used in the IMDAA reanalysis are the harmonized adaptation rate and hybrid scan bias scheme.

SAPHIR (onboard Megha-Tropiques satellite) is one of the latest satellite observations introduced in the IMDAA system since mid-2014. Time series plot of the mean of the observed minus background (O-B) and observed minus analysis (O-A) of SAPHIR channel-1radiances for a period of initial two and half months is shown in Figure S1. The initial bias is not perfect, and VarBC improves the correction with a bias halving time of approximately 3-4 days. Though the VarBC system can use a hybrid scan scheme, that is not explored in the IMDAA reanalysis. In this study, the VarBC system retains the ability to apply a static offset to each scan position. The scan bias of SAPHIR channel-1 at different scan positions is shown in Figure S2.

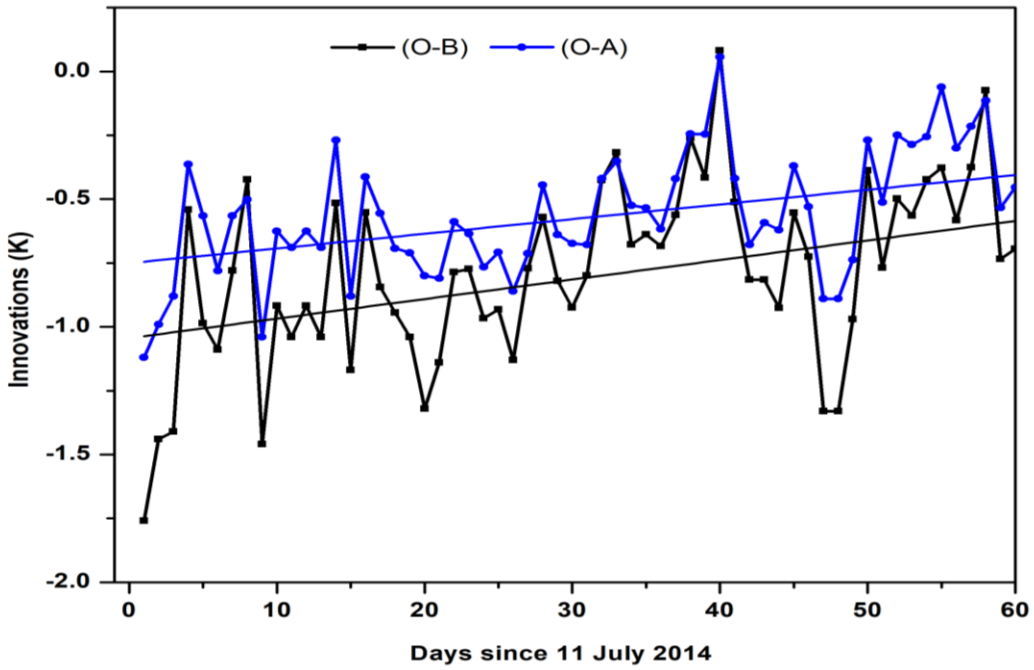

Figure S1: The mean of observed minus background (O-B) and observed minus analysis (O-A) of SAPHIR channel-1 radiances (Megha-Tropiques satellite), which is introduced in the IMDAA system during mid-2014. The initial bias is not perfect, and VarBC improves the correction with a bias halving time of approximately 3-4 days.

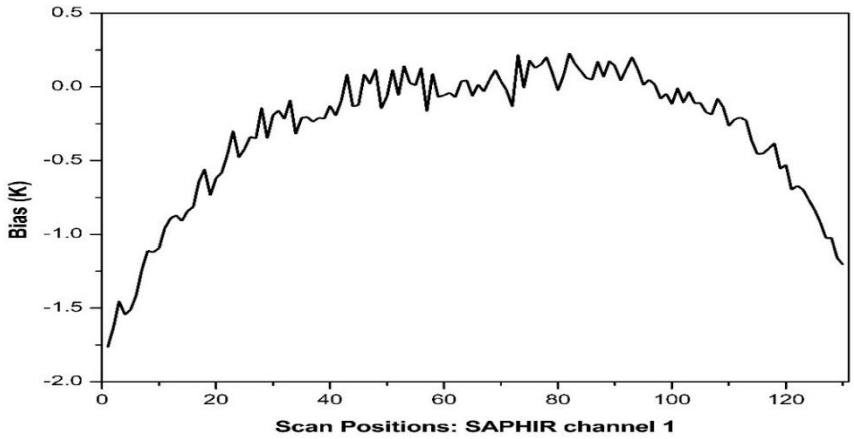

Figure S2: Scan bias of SAPHIR channel-1 for 130 scan positions used in the IMDAA system.

Innovations in the AMVs

Figure S3 shows the monthly mean of daily average RMS departures of AMV zonal components assimilated in the IMDAA from 1979 to 2018. Figure S3(a) represents the RMSD in the O-B and O-A time series, while the number of AMVs assimilated is shown in Figure S3(b). Similar to other observations (surface (Figure 8), sonde (Figure 9), SAPHIR (Figure S1)), RMS in both O-B and O-A of the AMVs reduced with time. Table S1 lists various observation types assimilated in IMDAA with their usage dates and source.

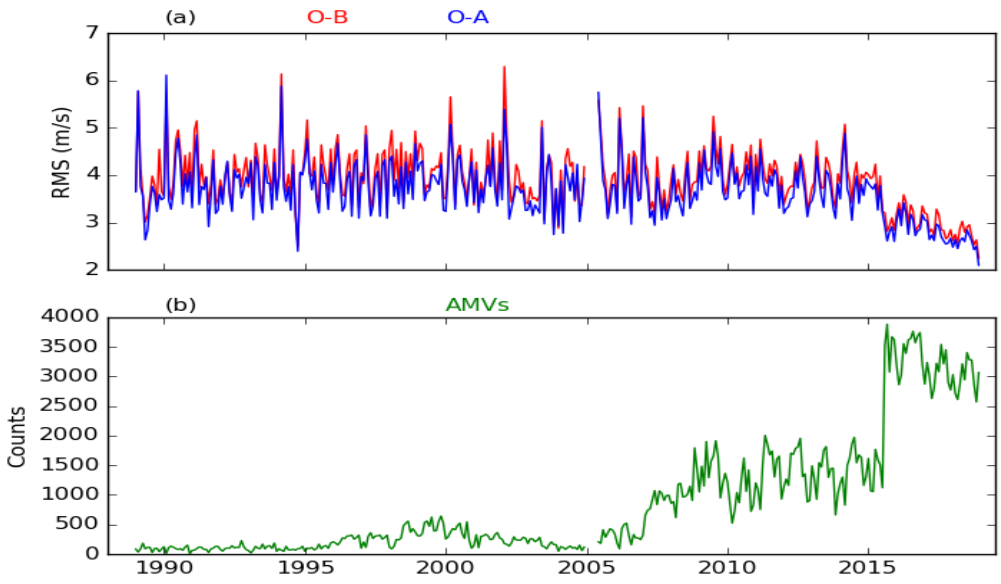

Figure S3: Time series of (a) the monthly mean of daily average RMS departures of AMVs zonal component (m/s), and (b) counts of observations assimilated in IMDAA during 00 UTC cycle.

Table S1: Observation types assimilated in IMDAA with their usage dates and sources.

Data Type	IMDAA Dates	Source
Conventional		
Surface land	Jan 1979 – Dec 2018	ECMWF, IMD/NCMRWF
Surface ship and buoy	Jan 1979 – Dec 2018	ECMWF, IMD/NCMRWF
Radiosonde and Pilot balloons	Jan 1979 - Dec 2018	ECMWF, IMD/NCMRWF
Aircraft: AIREP	Jan 1979 – Dec 2018	ECMWF, IMD/NCMRWF
Aircraft: AMDAR	Mar 1993 - Dec 2018	ECMWF, IMD/NCMRWF
Atmospheric Motion Vectors (AMVs)		
GMS	Jan 1989- Dec 2004	ECMWF
Meteosat-5	Jun 1998 – Dec 2000	ECMWF
MTSAT	Jan 2006- Dec 2015	ECMWF
Meteosat-7 and 9	Jan 2008 – Dec 2018	ECMWF, NCMRWF, UKMO
Himawari-8	Jan 2016 – Dec 2018	ECMWF, NCMRWF, UKMO
Sea surface wind vectors		
ERS-2	Jan 1997 – Jan 2001	ECMWF
QuickScat	Jan 2000 – Jan 2009	ECMWF
ASCAT	Jan 2009 – Dec 2018	ECMWF, NCMRWF, UKMO
Satellite Radiances		
TOVS	Jan 1979 – Dec 1999	ECMWF
ATOVS	Jan 2000 – Dec 2018	ECMWF, NCMRWF, UKMO
AIRS	Jan 2004 – Dec 2018	ECMWF, NCMRWF, UKMO
SEVIRI	Apr 2005 – Dec 2018	NCMRWF, UKMO
MVIRI	Jan 2007 – Mar 2017	NCMRWF, UKMO
IASI	Jan 2010 – Dec 2018	ECMWF, NCMRWF, UKMO
ATMS	May 2012 – Dec 2018	NCMRWF, UKMO
CrIS	May 2012 – Dec 2018	NCMRWF, UKMO
AMSR-2	Jun 2014 – Dec 2018	NCMRWF, UKMO
SAPHIR	Jul 2014 – Dec 2018	NCMRWF, UKMO

Standard deviation of analysis increment fields

The monthly mean profiles of standard deviations in the analysis increments of temperature, zonal wind, and specific humidity are shown respectively in Figure S4 (a), (b), and (c). The standard deviations of analysis increments of temperature and zonal wind remain the same throughout IMDAA reanalysis; however, the standard deviation of analysis increments of

humidity shows variations with respect to the various observations assimilated, particularly the satellite radiances.

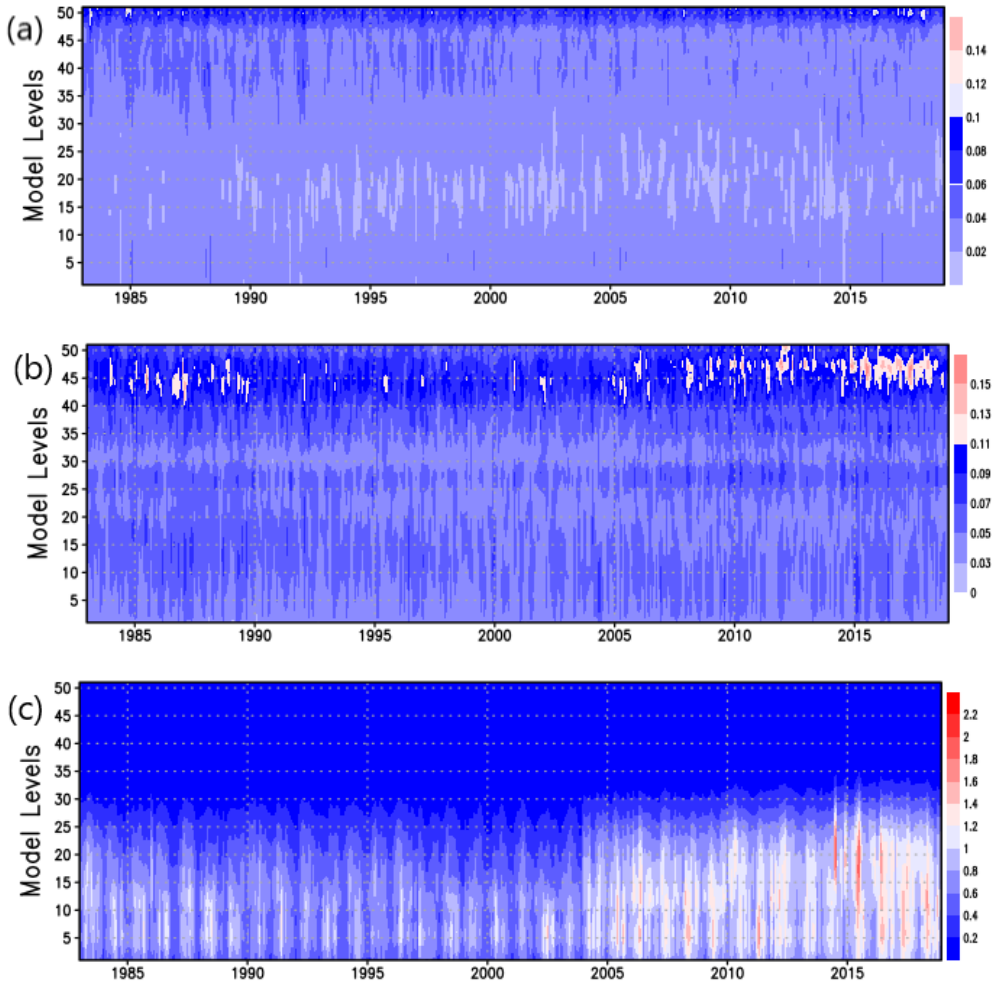

Figure S4: Monthly mean profiles of standard deviations of the analysis increments (a) temperature (K), (b) zonal wind (m/s), and (c) specific humidity (kg/kg) over the IMDAA domain for the period 1979-2018.

Additional details of monsoon circulations

Characteristics of summer monsoon features like onset, withdrawal, Low Level Jet (LLJ), and Tropical Easterly Jet (TEJ) are broadly described in the main manuscript. Additional details of the variations in these features during two different epochs of 20 years each from IMDAA and their comparison with ERA5 and IMD observations are provided here. Table S2 summarizes the statistics of the onset and withdrawal dates of the Indian summer monsoon from IMDAA, ERA5, and IMD for two different epochs of 20 years each. Similarly, Tables S3 and S4 summarize statistics of LLJ and TEJ estimates from IMDAA and ERA5.

Table S2: Statistics of monsoon onset and withdrawal from IMDAA, ERA5, and IMD during the two epochs of 20 years each, 1979-1998 and 1999-2018.

Monsoon Onset	1979 - 1998			1999 - 2018		
	IMDAA	ERA5	IMD (Observed)	IMDAA	ERA5	IMD (Observed)
Mean (Julian Day)	153.45	152.75	153.65	151.5	151.45	153.55
SD (Days)	6.26	6.29	5.79	6.17	6.56	5.73
Min (Julian Day)	141	141	140	142	140	144
Max (Julian Day)	166	166	165	163	162	165

Monsoon Withdrawal	1979 - 1998		1999 - 2018	
	IMDAA	ERA5	IMDAA	ERA5
Mean (Julian Day)	297.2	296.7	298.85	299.45
SD (Days)	5.502	6.25	6.983	6.893
Min (Julian Day)	285	284	287	287
Max (Julian Day)	305	306	312	312

Table S3: Statistics of LLJ estimates from IMDAA and ERA5 during two epochs of 20 years each, 1979-1998 and 1999-2018.

LLJ Statistics (m/s)	1979 - 1998							
	June		July		Aug		Sep	
	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5
Mean	9.575	10.201	9.666	10.31	8.693	9.39	5.75	6.39
SD	1.315	1.29	0.932	0.895	0.842	0.755	1.00	0.999
Min	5.55	6.233	7.13	7.67	7.09	7.986	3.60	4.373
Max	11.07	11.684	11.10	11.846	10.18	11.033	7.14	8.032
	1999 - 2018							
	June		July		Aug		Sep	
	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5
Mean	9.614	10.144	10.180	10.563	8.675	9.061	6.53	6.774
SD	0.88	0.861	0.80	0.752	0.86	0.862	1.15	1.05
Min	7.85	8.676	8.82	9.264	7.30	7.620	4.48	4.743
Max	10.989	11.949	11.790	11.983	10.355	10.643	8.333	8.426

Table S4: Statistics of TEJ estimates from IMDAA and ERA5 during two epochs of 20 years each, 1979-1998 and 1999-2018.

TEJ Statistics (m/s)	1979 - 1998							
	June		July		Aug		Sep	
	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5
Mean	24.336	22.550	30.460	30.184	28.414	28.621	20.342	20.394
SD	3.384	3.033	2.311	1.229	2.466	1.833	2.438	2.442
Min	18.675	17.362	27.423	28.011	22.474	25.800	13.938	14.763
Max	29.393	26.306	36.082	32.015	32.168	31.706	23.716	23.671
	1999 - 2018							
	June		July		Aug		Sep	
	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5	IMDAA	ERA5
Mean	23.508	23.723	28.360	29.389	25.452	27.730	19.217	21.098
SD	2.932	2.48	2.278	1.810	1.451	1.440	2.198	2.297
Min	16.955	17.10	24.152	24.193	21.334	23.837	15.707	16.975
Max	29.100	28.573	33.003	33.602	27.191	29.981	24.674	25.844

Figure S5 shows the mean wind derived from the 40 years mean at 150 hPa and 850 hPa from IMDAA and ERA5. Figures S5 (a) and (d) are the upper tropospheric circulations at 150 hPa, and the corresponding geopotential height from IMDAA and ERA5 and Figures S5 (b) and (e) are the lower tropospheric circulations at 850 hPa and the corresponding daily average of monthly mean precipitation from IMDAA and ERA5 during the Indian summer monsoon season (June-September, JJAS). Figures S5(c) and (f) are similar to Figures S5 (b) and (e) but for the 40 years average 850 hPa mean wind field and precipitation during north-east monsoon (October, November, and December, OND). Differences between IMDAA and ERA5 mean circulation

patterns at 150 hPa and 850 hPa during the summer monsoon and 850 hPa during the north-east monsoon are shown in Figure S6. IMDAA explicitly resolves the features, particularly over the equatorial regions (between 5°S and 5°N), and overestimates the precipitation over the Indo-Gangetic plains compared to ERA5. At the same time, ERA5 underestimates the orographic precipitation over the west coast, as seen from the difference plot Figure S6(b).

Figure S5 (c) and (f) are the mean circulation at 850 hPa and the accumulated precipitation during the north-east monsoon season from 1979 to 2018 from IMDAA and ERA5. During the withdrawal phase of the Indian summer monsoon, lower level winds over south Asia reverse their direction from southwest to north-east, associated with the southward movement of the ITCZ, as seen in Figures S5 (c) and (f). The zone of maximum rainfall migrates to south India, Sri Lanka, and the neighboring seas during the north-east monsoon (Rajeevan et al., 2012), and the same can be seen in Figures S5 (c) and (f). Both the reanalyses depict climatological features of north-east monsoon features. During November, December, and January, western disturbances associated with the mid-tropospheric westerlies extend to the lower troposphere and bring rain over the north and north-east India and are well depicted in both reanalyses.

It is noted that IMDAA overestimates the precipitation over the Indo-Gangetic plains during the south-west monsoon (Figure 14 and 15 in the main manuscript, and Figure S6(b)). Figure S7 shows the mean bias in the precipitation estimate from IMDAA and ERA5 against the IMD observation. Figure S7(a) is the IMD observed seasonal accumulated precipitation during the summer monsoon for 1979-2018. Figure S7(b) and (c) are the percentage mean biases in the IMDAA and ERA5 estimated precipitation against IMD observation for the same period. It is noted from Figure S7(b) that IMDAA is wet compared to both observation and ERA5 and

produced more rainfall over the Indo-Gangetic plains and east coast. ERA5 produces more rainfall over the west coast, as seen in Figure S7(c).

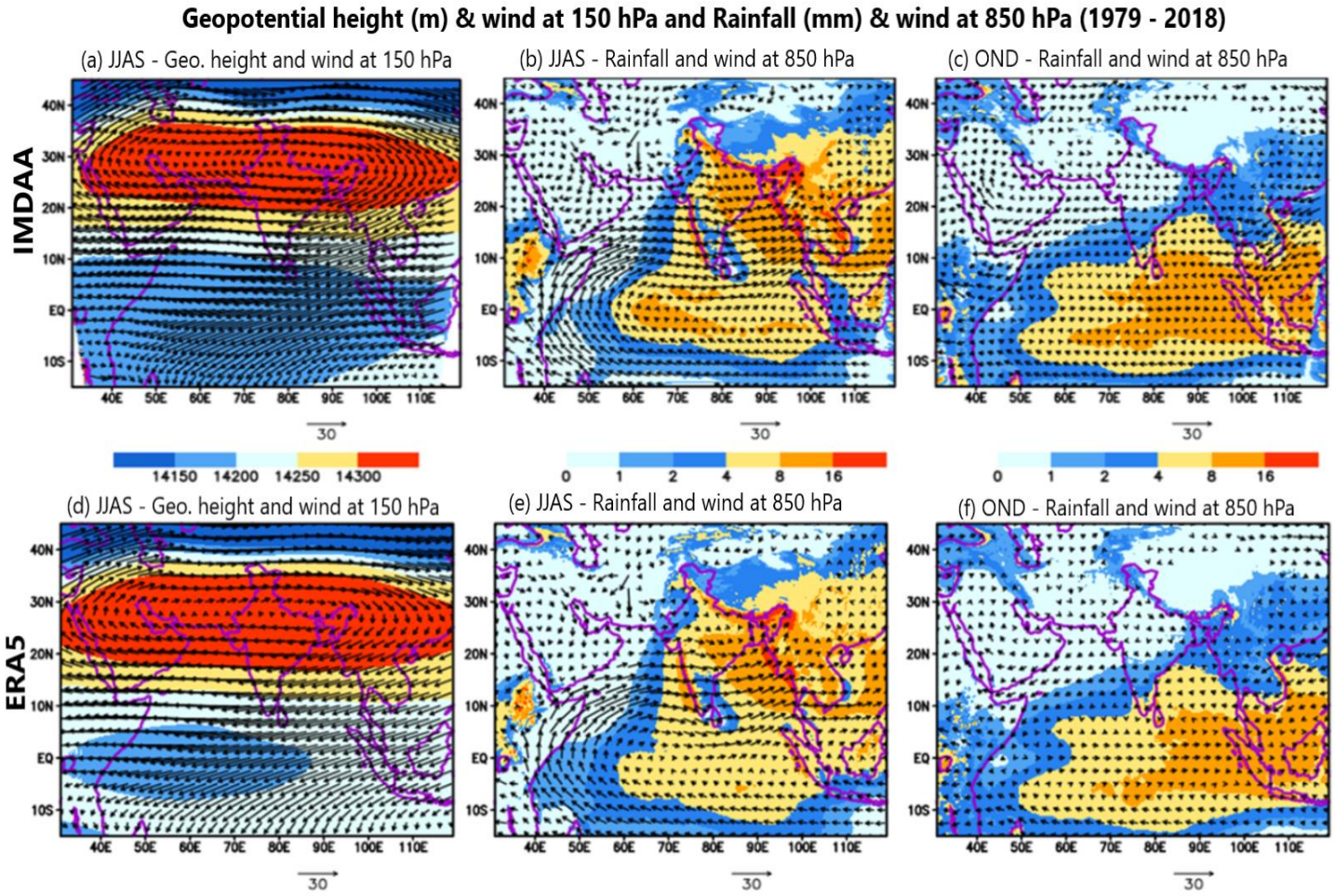

Figure S5: Mean wind field from IMDAA and ERA5 for the period from 1979 to 2018. (a) JJAS average winds (m/s) and geopotential height (m) at 150 hPa, (b) JJAS mean wind fields at 850 hPa and rainfall, and (c) October to December mean wind at 850 hPa and rainfall. The lower panel (d), (e), and (f) are the corresponding estimates from ERA5

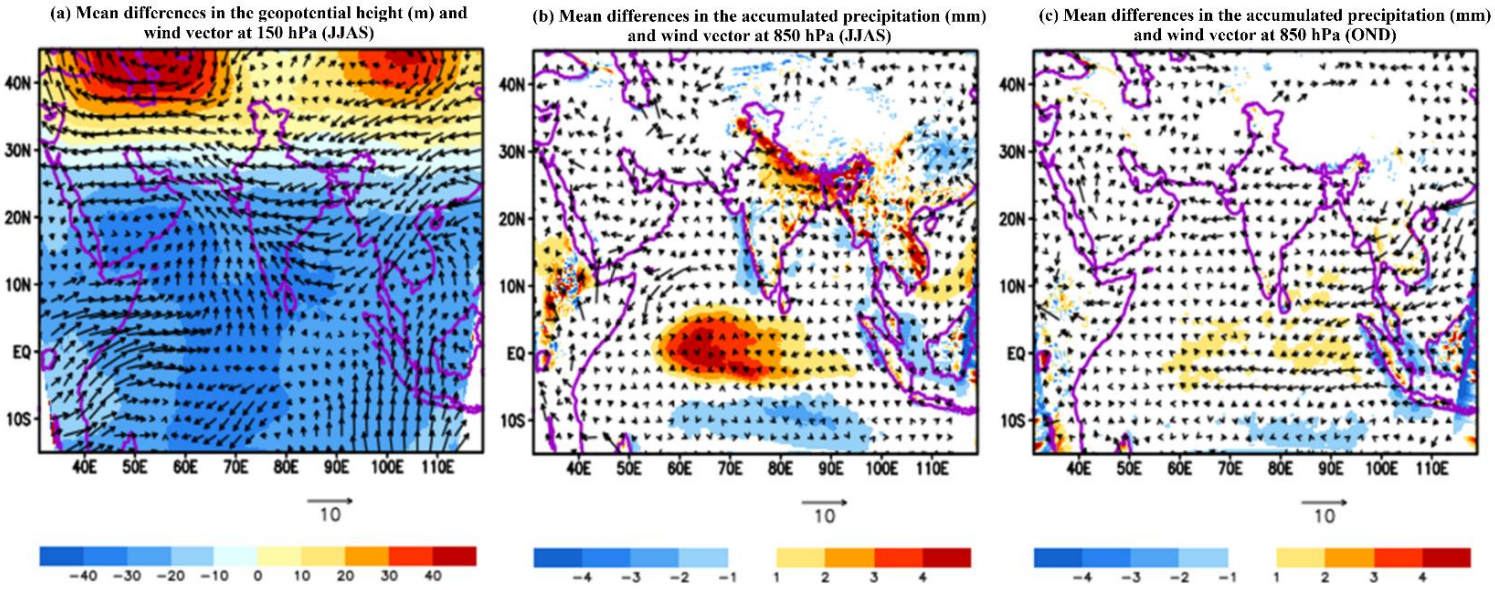

Figure S6: Differences in the mean fields between IMDAA and ERA5 (a) difference in the geopotential height (m) and vector wind at 150 hPa during JJAS, (b) differences in the accumulated precipitation and wind vector at 850 hPa during JJAS and (c) differences in the accumulated precipitation and wind vector at 850 hPa during OND.

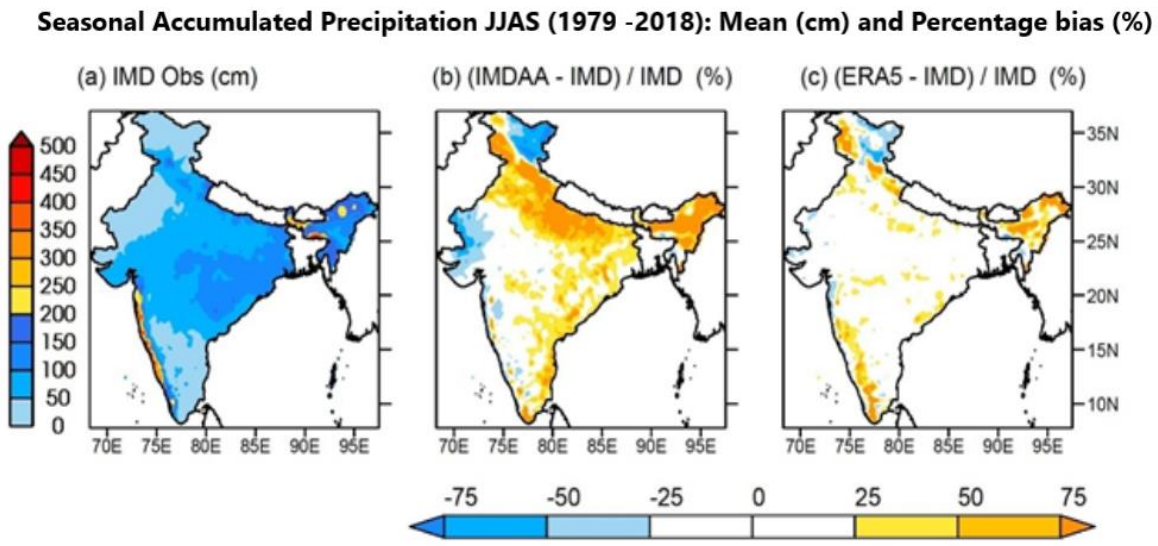

Figure S7: Mean and percentage biases in IMDAA and ERA5 during the Indian summer monsoon from 1979 to 2018. (a) IMD observed seasonal accumulated precipitation (cm) and percentage biases in the (b) IMDAA, (c) ERA5 against IMD observations.

Post-processing and product generation

IMDAA is an end-to-end reanalysis system, from observation processing to the post-processing and archives of the reanalysis products, as shown in Figure 2 (main manuscript). IMDAA analysis and short forecast files are in the UM fields file (ff) or post-processing (pp) format. These are specific UM formats and cannot be easily read by potential users of IMDAA. Archival and data dissemination of IMDAA requires the products in popular open standard formats such as gridded binary (grib) or Network Common Data Form (NetCDF). To convert the UM ff/pp files, NCMRWF developed a software named "UMRider" (Arulalan, 2020), which converts the UM outputs from ff/pp format to grib2, grib1, and NetCDF formats (https://github.com/NCMRWF/UMRider).

The two dimensional (surface and atmospheric column integrated) variables are archived at the hourly interval in grib2 format, and the details of variables are provided in Table S5.

Atmospheric standard pressure level variables are available in three hourly intervals in grib2, and the variables are listed in Table S6. Some of the two-dimensional flux variables are converted to NetCDF and are available hourly, listed in Table S7. Hourly atmospheric variables in model levels are converted to NetCDF, and the details are provided in Table S8. NCMRWF hosts a website to distribute the data to the research community. Users can access IMDAA reanalysis products from the web portal https://rds.ncmrwf.gov.in.

References:

Arulalan, T., 2020: UMRider – A Parallel Post Processing Utility for Unified Model, *NCMRWF Internal Report*, NMRF/TR/04/2020.

Rajeevan, M., C. K. Unnikrishnan, J. Bhate, K. Niranjan Kumar and P. P. Sreekala, 2012: Northeast monsoon over India: variability and prediction, *Meteorol. Appl.*,**19,** 226–236, doi: 10.1002/met.1322

Table S5: Details of hourly two dimensional (surface and atmospheric column integrated) variables stored in grib2

Sl.No.	Short Name	Variables	Units	Hourly Time Profile
1	APCPsfc	Surface Total Precipitation	kgm^{-2}	accumulation
2	CDCBclm	Atmosphere column Cloud Base	m	instantaneous
3	CISOILM0_10cm	Column-Integrated Soil Moisture0-0.1 m below ground	kgm^{-2}	instantaneous
4	CISOILM10_35cm	Column-Integrated Soil Moisture0.1-0.35 m below ground	kgm^{-2}	instantaneous
5	CISOILM35_100cm	Column-Integrated Soil Moisture0.35-1 m below ground	kgm^{-2}	instantaneous
6	CISOILM100_300cm	Column-Integrated Soil Moisture1-3 m below ground	kgm^{-2}	instantaneous
7	CRRATEsfc	Surface Convective Rain Rate	$kgm^{-2}s^{-1}$	average
8	CSRATEsfc	Surface Convective Snowfall Rate	$m\ s^{-1}$	average
9	CWPsfc	Surface Convective Water Precipitation	kgm^{-2}	accumulation
10	DLWRFsfc	Surface Downward Long-Wave Rad. Flux	Wm^{-2}	average
11	DSWRFsfc	Surface Downward Short-Wave Radiation Flux	Wm^{-2}	average
12	DSWRFtoa	Top of atmosphere Downward Short-Wave Radiation Flux	Wm^{-2}	average
13	EVARCAsfc	Surface Evaporation Rate from Canopy	$kg\ m^{-2}\ s^{-1}$	average
14	EVAROSsfc	Surface Evaporation Rate from Open Sea	$kg\ m^{-2}\ s^{-1}$	average
15	EVARSSsfc	Surface Evaporation Rate from Soil Surface	$kg\ m^{-2}\ s^{-1}$	average
16	GUST10m	10 m above ground Wind Speed (Gust)	ms^{-1}	instantaneous
17	HCDCclm	Atmospheric column High Cloud Cover	%	instantaneous
18	HPBLsfc	Surface Planetary Boundary Layer Height	m	instantaneous

19	LANDsfc	Surface Land Cover (0=sea, 1=land)	value	instantaneous
20	LCDCclm	Atmospheric column Low Cloud Cover	%	instantaneous
21	LHTFLsfc	Surface Latent Heat Net Flux	Wm^{-2}	average
22	LSRRATEsfc	Surface Large Scale Rain Rate	$kgm^{-2}s^{-1}$	average
23	LSSRATEsfc	Surface Large Scale Snowfall Rate	ms^{-1}	average
24	LSWPsfc	Surface Large Scale Water Precipitation (Non-Convective)	kgm^{-2}	accumulation
25	MCDCclm	Atmospheric column Medium Cloud Cover	%	instantaneous
26	MTERHsfc	Surface Model Terrain Height	m	instantaneous
27	NLWRFsfc	Surface Net Long-Wave Radiation Flux	Wm^{-2}	average
28	NLWRFtoa	Top of the atmosphere Net Long-Wave Radiation Flux	Wm^{-2}	average
29	NSWRFsfc	Surface Net Short Wave Radiation Flux	Wm^{-2}	average
30	NSWRFtoa	Top of the atmosphere Net Short Wave Radiation Flux	Wm^{-2}	average
31	PRESsfc	Surface Pressure	Pa	instantaneous
32	PRMSLmsl	Mean sea level Pressure Reduced to MSL	Pa	I1nstantaneous
33	RH2m	Relative Humidity2 m above ground	%	instantaneous
34	SFCRsfc	Surface Roughness	m	instantaneous
35	SHTFLsfc	Sensible Heat Net Flux	Wm^{-2}	average
36	SNOCsfc	Convective Snow	kgm^{-2}	accumulation
37	SNOLsfc	Large-Scale Snow	kgm^{-2}	accumulation
38	SSFCWRORsfc	Sub Surface Water Runoff Rate	$kgm^{-2}s^{-1}$	average
39	SUFCWRsfc	Surface Water Runoff Rate	$kgm^{-2}s^{-1}$	average
40	TCDCROclm	Atmospheric column Total Cloud Cover Assuming Random Overlap	%	instantaneous

41	TMPsfc	Surface Temperature	K	instantaneous
42	TMP2m	Temperature2 m above ground	K	instantaneous
43	TSOIL0_10cm	Soil Temperature0-0.1 m below ground	K	instantaneous
44	TSOIL10_35cm	Soil Temperature0.1-0.35 m below ground	K	instantaneous
45	TSOIL35_100cm	Soil Temperature0.35-1 m below ground	K	instantaneous
46	TSOIL100_300cm	Soil Temperature1-3 m below ground	K	instantaneous
47	UGRD10m	U-Component of Wind10 m above ground	ms^{-1}	instantaneous
48	UGRD50m	U-Component of Wind50 m above ground	ms^{-1}	instantaneous
49	VGRD10m	V-Component of Wind10 m above ground	ms^{-1}	instantaneous
50	VGRD50m	V-Component of Wind50 m above ground	ms^{-1}	instantaneous
51	VIS2m	Visibility2 m above ground	M	instantaneous
52	VLCDCclm	Atmospheric column Very Low Cloud Cover	%	instantaneous
53	WEASDsfc	Surface Water Equivalent of Accumulated Snow Depth	kgm^{-2}	accumulation

Table S6: Details of three-hourly atmospheric pressure level variables in grib2

Sl.No.	**Short Name**	**Variables**	**Units**	**Three hourly Time Profile**
1	HGTprs	Geopotential Height	gpm	instantaneous
2	RHprs	Relative Humidity	%	instantaneous
3	TMPprs	Temperature	K	instantaneous
4	UGRDprs	U-Component of Wind	ms^{-1}	instantaneous
5	VGRDprs	V-Component of Wind	ms^{-1}	instantaneous

Table S7: Details of hourly two-dimensional flux variables in NetCDF

Sl.No.	**Short Name**	**Variables**	**Units**	**Hourly Time Profile**
1	DSHFLUX	Downward soil heat flux	Wm^{-2}	average
2	DSSWRFLX	Direct surface shortwave flux in the air	Wm^{-2}	average
3	DIFSSWRF	Surface diffuse downwelling shortwave flux in the air	Wm^{-2}	average
4	NDDSWRFC	Surface net downward shortwave flux corrected	Wm^{-2}	average
5	CSUSFT	Top of the atmosphere outgoing shortwave flux assuming clear sky	Wm^{-2}	average
6	CSULFT	Top of the atmosphere outgoing longwave flux assuming clear sky	Wm^{-2}	average
7	CSUSFS	Surface upwelling shortwave flux in air assuming clear sky	$W\ m^{-2}$	average
8	CSDSFS	Surface downwelling shortwave flux in air assuming clear sky	$W\ m^{-2}$	average
9	CSDLFS	Surface downwelling longwave flux assuming clear sky	$W\ m^{-2}$	average

Table S8: Details of hourly atmospheric variables in model levels in NetCDF

Sl.No.	Short Name	Variables	Units	Hourly Time Profile
1	surface altitude	Orography	m	instantaneous
2	Density* r* r	Density scaling with respect to the radius of the earth (r)	--	instantaneous
3	cloud volume fraction in atmosphere layer	Fraction of cloud volume in atmosphere layer	--	instantaneous
4	liquid cloud volume fraction in atmosphere layer	Fraction of liquid cloud volume in atmosphere layer	--	instantaneous
5	ice cloud volume fraction in atmosphere layer	Fraction of ice cloud volume in atmosphere layer	_	instantaneous
6	air potential temperature	Atmospheric Potential Temperature	K	instantaneous
7	air pressure	Atmospheric pressure	Pa	instantaneous
8	dimensionless Exner pressure	Exner pressure	--	instantaneous
9	mass fraction of cloud ice in air	Mass fraction of cloud ice in each atmospheric layer	$kgkg^{-1}$	instantaneous
10	mass fraction of cloud liquid water in air	Mass fraction of cloud liquid water in each atmospheric layer	$kgkg^{-1}$	instantaneous
11	potential vorticity of atmosphere layer	Potential vorticity on model levels	$Pa^{-1}\ s^{-1}$	instantaneous
12	specific humidity	Specific humidity	$kgkg^{-1}$	instantaneous
13	upward air velocity	Vertical Velocity	ms^{-1}	instantaneous
14	Zonal wind	U-Component of Wind	ms^{-1}	instantaneous
15	Meridional wind	V-Component of Wind	ms^{-1}	instantaneous
16	direct uv flux in air	Direct wind flux in air	$W\ m^{-2}$	average
17	relative humidity	Relative humidity	%	instantaneous