IMDAA: High-Resolution Satellite-Era Reanalysis for the Indian Monsoon Region

S. INDIRA RANI,[a] T. ARULALAN,[a] JOHN P. GEORGE,[a] E. N. RAJAGOPAL,[a] RICHARD RENSHAW,[b] ADAM MAYCOCK,[b] DALE M. BARKER,[b] AND M. RAJEEVAN[c]

[a] *National Centre for Medium Range Weather Forecasting, Ministry of Earth Sciences, Noida, Uttar Pradesh, India*
[b] *Met Office, Exeter, United Kingdom*
[c] *Ministry of Earth Sciences, New Delhi, India*

(Manuscript received 2 June 2020, in final form 12 March 2021)

ABSTRACT: A high-resolution regional reanalysis of the Indian Monsoon Data Assimilation and Analysis (IMDAA) project is made available to researchers for deeper understanding of the Indian monsoon and its variability. This 12-km resolution reanalysis covering the satellite era from 1979 to 2018 using a 4D-Var data assimilation method and the U.K. Met Office Unified Model is presently the highest resolution atmospheric reanalysis carried out for the Indian monsoon region. Conventional and satellite observations from different sources are used, including Indian surface and upper air observations, of which some had not been used in any previous reanalyses. Various aspects of this reanalysis, including quality control and bias correction of observations, data assimilation system, land surface analysis, and verification of reanalysis products, are presented in this paper. Representation of important weather phenomena of each season over India in the IMDAA reanalysis verifies reasonably well against India Meteorological Department (IMD) observations and compares closely with ERA5. Salient features of the Indian summer monsoon are found to be well represented in the IMDAA reanalysis. Characteristics of major semipermanent summer monsoon features (e.g., low-level jet and tropical easterly jet) in IMDAA reanalysis are consistent with ERA5. The IMDAA reanalysis has captured the mean, interannual, and intraseasonal variability of summer monsoon rainfall fairly well. IMDAA produces a slightly cooler winter and a hotter summer than the observations; the reverse is true for ERA5. IMDAA captured the fine-scale features associated with a notable heavy rainfall episode over complex terrain. In this study, the fine grid spacing nature of IMDAA is compromised due to the lack of comparable resolution observations for verification.

KEYWORDS: Reanalysis data; Monsoons; Satellite observations; Data assimilation

1. Introduction

Meteorological reanalyses produce the best estimate of the state of the atmosphere using the latest NWP systems with past observational data for any period and place (Thorne and Vose 2010). Observations alone can provide only a limited understanding of the past weather and climate, specific to observed variables at few locations. In contrast, the reanalysis provides comprehensive snapshots of conditions at regular intervals over long periods, often years or decades, from sparse observations (Parker 2016). Considering the homogeneity and completeness of the dataset, reanalyses have acquired great importance in the weather and climate studies in recent times (Kidston et al. 2010; Kravtsov et al. 2014; Nkiaka et al. 2017).

Reanalyses can be categorized into two types: reanalysis of the modern observing period (~30–50 years) and extended climate reanalysis (~100–200 years). Reanalysis of the modern observing period uses as many observations as possible and produces the best estimate at any given time (Kalnay et al. 1996; Gibson et al. 1997; Uppala et al. 2005; Onogi et al. 2007; Saha et al. 2010; Dee et al. 2011; Ebita et al. 2011; Gelaro et al. 2017; Prasad et al. 2017; Hersbach et al. 2020), whereas only a restricted set of observations are used in the extended climate reanalysis, which focuses on low-frequency variability and trends (Compo et al. 2011; Poli et al. 2013). Compo et al. (2011) extended atmospheric reanalysis back to the late nineteenth and early twentieth century using only surface pressure observations, whereas Poli et al. (2013) used marine surface winds in addition to surface pressure observations. Reanalysis has many applications, such as model validation, initial conditions for reforecast datasets for calibration, understanding historical weather events, preparing weather statistics, and climatology.

Advances in observation bias correction and data assimilation have reduced the uncertainty of reanalyses to a certain extent (Dee et al. 2011). Still, it is important to quantify the uncertainty in reanalyses to expand their usage for weather and climate analysis. In particular, multiannual variability and trends in atmospheric temperature from ERA-Interim,[1] JRA-55, and MERRA are in

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-20-0412.s1.

Aruralan's current affiliation: India Meteorological Department, Ministry of Earth Sciences, New Delhi, India.

Barker's current affiliation: Centre for Climate Research Singapore, Met Service Singapore.

Corresponding author: S. Indira Rani, indira@ncmrwf.gov.in; ranispl@gmail.com

[1] List of acronyms is included in the appendix.

DOI: 10.1175/JCLI-D-20-0412.1

good agreement in the upper troposphere and lower stratosphere but are more uncertain in the middle stratosphere (Simmons et al. 2014). Other accumulated quantities like precipitation and surface fluxes also exhibit substantial differences among reanalyses (Trenberth et al. 2011; Bosilovich et al. 2008, 2011). Global precipitation estimates from different reanalyses datasets differ by ~20% (Gelaro et al. 2017), and the uncertainty in these state estimates can be quantified by introducing ensemble methods in data assimilation (Compo et al. 2011; Poli et al. 2013; Hersbach et al. 2020). However, studies indicate that global reanalyses with grid sizes larger than 50 km cannot represent subgrid-scale variations in meteorology over heterogeneous terrain (Mesinger et al. 2006; Randall et al. 2007).

The available highest resolution global reanalysis is currently ERA5 (Hersbach et al. 2020), with a horizontal grid spacing of ~30 km. A critically important measure of an NWP model's accuracy is its ability to resolve features at the limits of its grid resolution (Skamarock 2004). Hence, reanalyses with coarse-resolution grid spacing are not suited for studying local and regional scales due to inadequate descriptions of mesoscale convective systems, squall lines, heat waves, etc. The cost in computing time of long-period reanalyses makes high-resolution global runs expensive. High-resolution regional reanalyses are a plausible solution representing mesoscale features (Dickinson et al. 1989; Fowler et al. 2007; Evans and McCabe 2013). Various regional reanalyses are available over different parts of the globe [e.g., Renshaw et al. (2013) over Europe (EURO4M), Su et al. (2019) over Australia (BARRA), etc.].

Monsoons, the key weather phenomena affecting the Indian subcontinent, encompass a broad spectrum of spatial scales from mesoscale to planetary scale. High-resolution regional reanalyses are essential for a detailed evaluation of spatial and temporal variability of the monsoon than has been achieved previously using global reanalyses (Sahai et al. 2015; Ramu et al. 2016; Hunt and Turner 2017). Understanding the mechanism and the spatiotemporal variation in rainfall are the central problems of monsoon meteorology (Gadgil 2006). Studies have shown that in monsoon regions, the most significant weather events are localized or mesoscale (Johnson 2006; Virts and Houze 2016) and produce high-impact weather (Sikka 2011). The complexity of physical processes and atmospheric forcing renders the task of understanding monsoons challenging for the scientific community (Ghelli et al. 2012). Adequate observations and regional reanalyses with fine grid spacing are required to know all of the processes involved in shaping the monsoon, especially those on the mesoscale. The WRF-based South Asian Regional Reanalysis (SARR; https://www.ncmrwf.gov.in/SARR-workshop-report-final.pdf) is the first regional reanalysis initiative over the Indian region (Sowjanya et al. 2012; Routray et al. 2014) but was limited to a short period. The Indian Monsoon Data Assimilation and Analysis (IMDAA) is the first high-resolution (~12 km) and long-term (40 years from 1979 to 2018) regional reanalysis developed for a deeper understanding of the Indian monsoon and its variability. This reanalysis is produced by the collaborative efforts of the Indian National Centre for Medium Range Weather Forecasting (NCMRWF), the U.K. Met Office, and India Meteorological Department (IMD), under the National Monsoon Mission project of the Ministry of Earth Sciences, Government of India. The IMDAA reanalysis for the 40 years (1979–2018) used observations from India (IMD and NCMRWF, both the local and GTS archives), the U.K. Met Office, and ECMWF archives. The state-of-the-art Met Office Unified Model (UM) and its 4D-Var data assimilation system are used to produce the IMDAA reanalysis. Performance of the pilot phase of IMDAA reanalysis is described in Mahmood et al. (2018), and an initial evaluation of the first 15 years of the production run is described in Ashrit et al. (2020a).

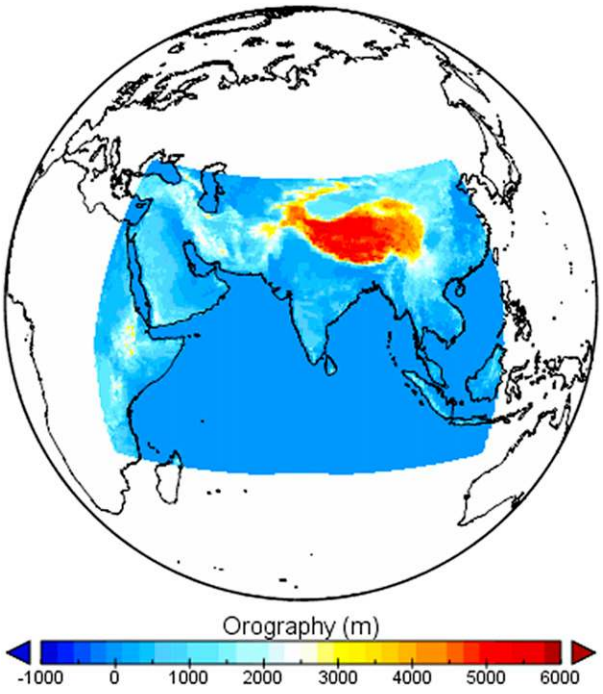

FIG. 1. IMDAA domain with model topography.

Figure 1 depicts the IMDAA domain with model topography. The high-frequency IMDAA datasets (hourly and 3-hourly products) are available to researchers at https://rds.ncmrwf.gov.in. The purpose of this paper is to compare the mean weather phenomena represented in IMDAA over India and neighboring regions during different seasons with observations and ERA5, and to introduce the dataset to the international community for a deeper understanding of the Indian monsoon. An overview of the IMDAA system, including the forecast model, the assimilation method, variational bias correction for satellite radiances, and land surface assimilation system, is provided in section 2. Details of observations, like timeline, and coverage, is provided in section 3. The performance of the IMDAA system is described in section 4, followed by a discussion and summary in section 5.

2. Overview of the IMDAA system

a. The assimilating model

The UM system of the U.K. Met Office (Davies et al. 2005; Brown et al. 2012) is used in IMDAA. The UM is a grid point model that uses a nonhydrostatic, fully compressible, deep atmospheric formulation. The equations of motion are solved using mass-conserving semi implicit, semi-Lagrangian time integration

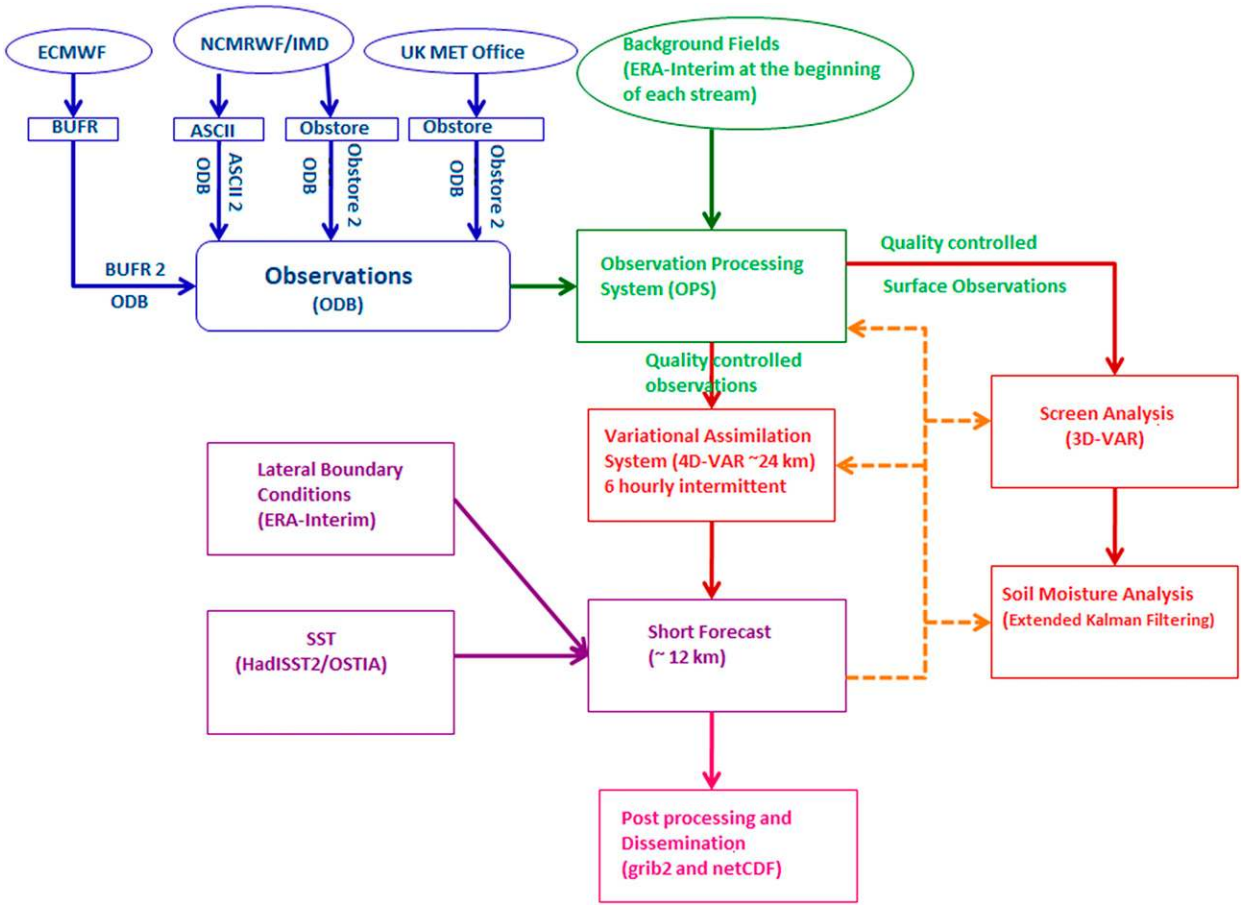

FIG. 2. IMDAA system flowchart.

methods within the ENDGame (Even Newer Dynamics for General Atmospheric Modeling of the Environment) dynamical core of the model (Wood et al. 2014). ENDGame is a finite-difference model discretized on a latitude–longitude grid based on the fully compressible, nonhydrostatic Euler equations (Cullen et al. 1997). The mass-flux convection scheme is based on Gregory and Rowntree (1990), with modification to include convective downdrafts (Gregory and Allen 1991) and convective momentum transport. The radiative transfer scheme (Edwards and Slingo 1996; Manners et al. 2015) treats solar radiation in six shortwave bands and thermal radiation in nine longwave bands. The microphysics used is a single-moment scheme based on Wilson and Ballard (1999), with extensive modifications. The warm rain scheme is based on Boutle et al. (2014a) and includes a prognostic rain formulation, which allows three-dimensional advection of the precipitation mass mixing ratio and an explicit representation of the effect of subgrid-scale variability on auto-conversion and accretion rates (Boutle et al. 2014b). The atmospheric boundary layer scheme is that of Lock et al. (2000) with the modifications described in Lock (2001) and Brown et al. (2008). It is a first-order turbulence closure mixing adiabatically conserved heat and moisture variables, momentum, and tracers. The updates to the boundary layer scheme include the reduction in turbulent mixing in stable conditions over the sea, inclusion of nonlocal momentum mixing in convective conditions, changes to surface scalar transfer over the sea, the improved dependency of low-level winds on changing the stability, and the impact of drag changes in the large-scale flow. The land configuration uses a community land surface model, JULES (Best et al. 2011; Clark et al. 2011), to model all land surface processes and subsurface soil. A horizontally staggered Arakawa C-grid (Arakawa and Lamb 1977) and vertically staggered Charney–Phillips grid (Charney and Phillips 1953) are used in the grid discretization. The model has a hybrid height-based vertical coordinate system. The levels have a smooth transition from terrain following near the surface to constant height in the upper levels (Davies et al. 2005).

IMDAA uses version 10.2 of the UM and is configured with 63 vertical levels extending from near the surface to a height of ~40 km above sea level. The horizontal domain of IMDAA spans from 30° to 120°E and from 15°S to 45°N (Fig. 1) with a grid spacing of 0.12° × 0.12° having 800 × 576 points in the horizontal.

b. *The variational data assimilation scheme*

IMDAA uses an incremental formulation of the 4D-Var data assimilation scheme of the U.K. Met Office based on Rawlins et al. (2007). The UM data assimilation system represents the variation of vertical covariances with latitude well, but not the longer horizontal scales in the stratosphere.

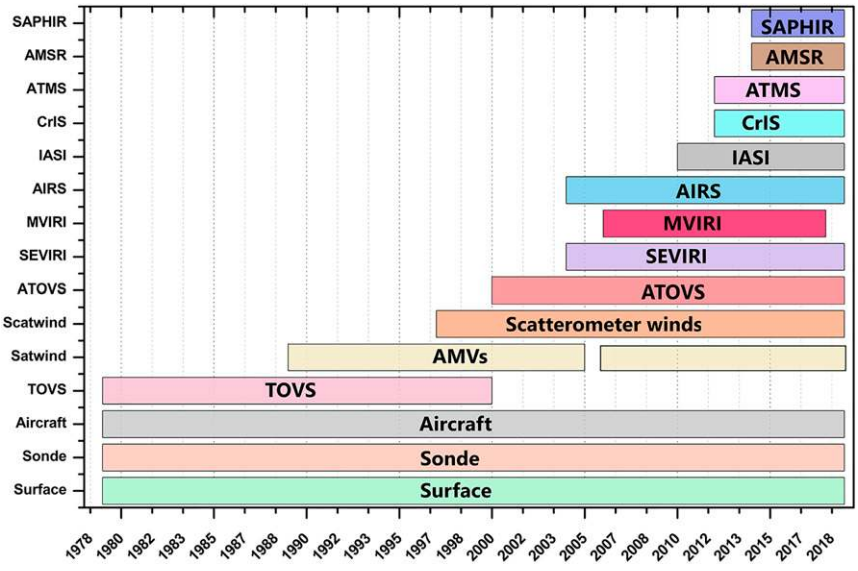

FIG. 3. Timeline of various observations assimilated in the IMDAA system.

In IMDAA, the background (forecast) error covariances are calculated using the so-called NMC method (named for the National Meteorological Center, now called the National Centers for Environmental Prediction), using differences between $T + 48$ and $T + 24$ forecast fields (Ingleby 2001). 4D-Var allows the flow-dependent influence of observations through the forecast model (Thépaut et al. 1996) and more effective use of observations through the consistent use of observation operators in the model equations (Rabier et al. 1998, 2000). Thépaut (2006) and Whitaker et al. (2009) showed that 4D-Var could produce analyses of large-scale tropospheric circulation even with only surface pressure observations.

Observations available in the 6-h window of each assimilation cycle are combined with the model background to produce a statistically optimal "state of the atmosphere." The linear perturbation forecast model (Lorenc 2003; Rawlins et al. 2007; Lorenc and Payne 2007) in the 4D-Var uses a simplified model formulation with a lower grid spacing ($\sim 0.22° \approx 24$ km) than the full UM forecast model used to produce the background ($\sim 0.12° \approx 12$ km).

c. *Land surface analysis*

Accurate surface analysis is critical since it significantly impacts the short-term forecasts of 2-m surface temperature and humidity (Candy et al. 2017). The 2-m temperature and humidity analysis increments produced by the 3D-Var surface analysis system are used in the soil moisture assimilation system as pseudo observations. A soil moisture analysis is produced using an extended Kalman filter–based land data assimilation system (de Rosnay et al. 2013). The Jacobians of the observation operator are estimated using forecasts of the land surface model JULES with perturbed initial conditions. Soil moisture analysis for four soil layers of 10-cm, 25-cm, 65-cm, and 2-m thickness is prepared at every assimilation cycle at the atmospheric model resolution (0.12°) and is used to initialize the model for each short forecast.

d. *Surface and lateral boundary conditions*

To speed up the production, multiple parallel streams of reanalysis runs were made, with one month of spinup time from initial and boundary conditions provided by ECMWF global reanalysis, ERA-Interim (Dee et al. 2011), which has a spatial resolution of approximately 80 km (T255 spectral) and has 60 levels in the vertical, from the surface up to 0.1 hPa. Renshaw et al. (2013) showed that one month is sufficient for surface temperature to settle to an approximately spun-up state. Land snow cover is allowed to evolve in the model since most parts of the IMDAA domain are devoid of snow cover except for a few regions over the Himalayas. It is pertinent to mention here that satellite estimates of snow are not reliable over these mountainous terrains and in situ observations are very limited (Kirkham et al. 2019). Figure 2 shows the complete IMDAA system flowchart, from the observation processing to the postprocessing.

For the model lower boundary condition, monthly ($1° \times 1°$) SSTs from HadISST2 [documented in part in Titchner and Rayner (2014)], produced using the gridded SST from in situ observations, have been used from 1979 to 2009. Daily ($0.05° \times 0.05°$) Met Office real-time Operational Sea Surface Temperature, and Sea Ice Analysis (OSTIA) (Donlon et al. 2012) data have been used from 2010 to 2018. The high-resolution OSTIA analysis uses satellite data from various microwave and infrared sensors and in situ data from drifting and moored buoys. ERA5 also used HadISST2 for the historical period, while OSTIA is used for the modern (2007 onward) period (Hersbach et al. 2020).

3. Observations used in IMDAA

a. *Data acquisition and processing*

The primary observational data source of IMDAA was the ECMWF archive; however, exclusive observations from the archives of IMD/NCMRWF were also used. A comparison

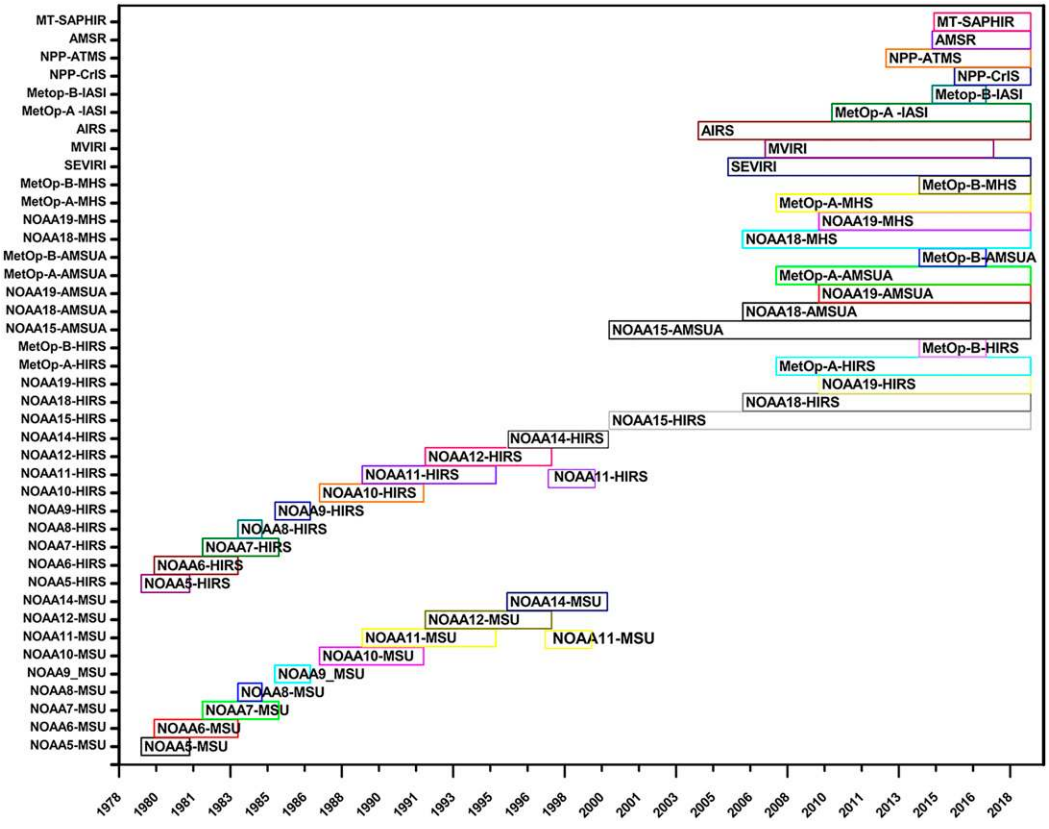

FIG. 4. Timeline of radiances (brightness temperatures) from various satellite instruments assimilated in the IMDAA system.

of ECMWF and NCMRWF archives for a month, April 2014, has been presented in Srinivas et al. (2016). This comparison shows that IMD/NCMRWF archives have additional surface observations from SYNOP and a few upper air observations like those from sondes and pilot balloons. These additional observations in the IMD/NCMRWF archive are used in the IMDAA system in addition to ECMWF observations.

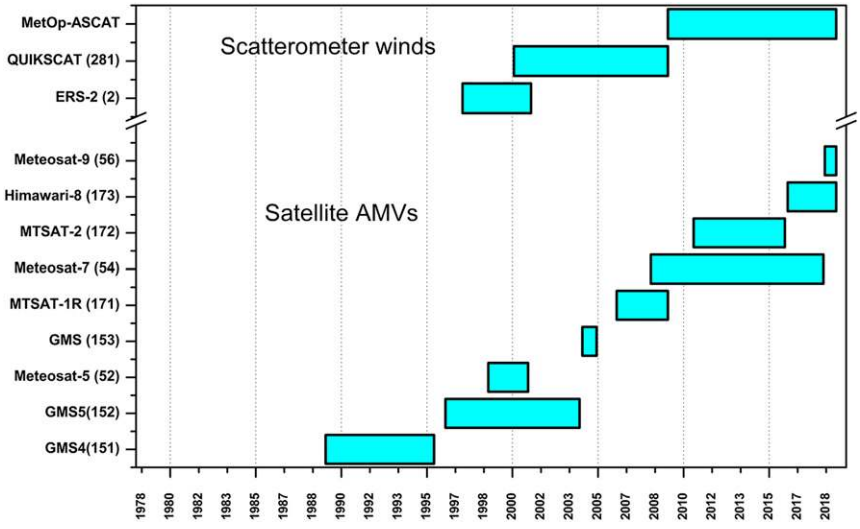

FIG. 5. Timeline of satellite winds (atmospheric motion vectors and scatterometer-derived ocean surface winds) assimilated in the IMDAA system. Integers inside the parentheses are the World Meteorological Organization (WMO) satellite identifiers.

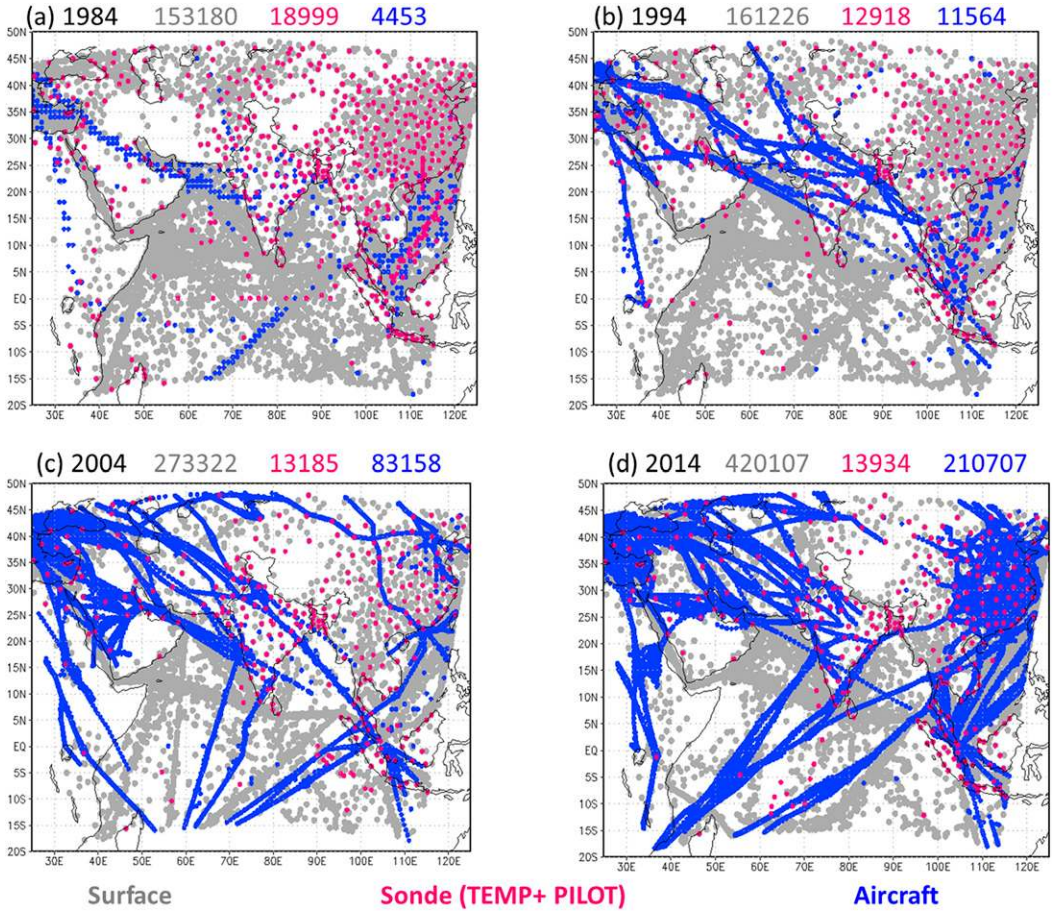

FIG. 6. Spatial coverage plots of conventional observations (surface, sonde, and aircraft) assimilated in the IMDAA system during JJAS of (a) 1984, (b) 1994, (c) 2004, and (d) 2014. The total number of observations assimilated at 0000 UTC (JJAS accumulated) from each type is also included.

Figure 3 shows the types of observations assimilated in the IMDAA system and the duration of their availability. Types of conventional observations assimilated in the IMDAA system include reports from land surface stations, buoys, ships, sondes (TEMP), pilot balloons (PILOT), and aircraft. Clear-sky brightness temperatures assimilated in the IMDAA system include those from TOVS through 1999 and ATOVS thereafter, different multispectral (MVIRI, SEVIRI), hyper-spectral instruments (AIRS, IASI, CrIS), and from the latest microwave instruments (ATMS, AMSR-2, SAPHIR) during their period of availability. Figure 4 summarizes the sources of radiances (converted to brightness temperatures), grouped by instruments and satellites assimilated in the IMDAA system. In IMDAA, upper atmosphere channels were used everywhere; window channels were used only over the ocean, and channels that have significant sensitivity to infrared and microwave radiation from the surface were rejected over land (Pavelin and Candy 2014). IMDAA used more than 40 satellite instruments, of which more than 24 instruments are post 2000. Satellite-derived products such as AMVs and sea surface winds from scatterometers supplemented the conventional observations and satellite radiances. Figure 5 shows the period of AMVs from different geostationary satellites and sea surface winds from different scatterometer missions assimilated in IMDAA. Various observation types assimilated in IMDAA with their usage date and source are included in the online supplemental material.

Before assimilation, the observations are quality checked in the Observation Processing System (OPS) of the UM system to ensure only good-quality observations are used for assimilation. The OPS also thins observations where they are dense in space or time to reduce the impact of error correlations (Dow 2004). In IMDAA, thinning of satellite observation is mostly done at 25 km (data assimilation resolution is ~24 km). The purpose of thinning is to reduce the data volume and correlated error so that the effective resolution of the observation field should be comparable to or lower than that of the data assimilation system. The quality control incorporates cloud detection, surface classification, and model-minus observation departure checks specified depending on the data type. Conventional observations that are significantly different from the background are rejected when they exceed the thresholds defined

by a Bayesian scheme (Lorenc and Hammon 1988; Ingleby and Lorenc 1993) unless they are consistent with other observations nearby.

The IMDAA reanalysis time frame is challenging because of the rapid developments in observing systems, especially satellites. New satellite instruments with high spatial and temporal resolution contribute to the accuracy of the atmospheric analysis. However, this highly varying observing system can lead to discontinuities in the reanalysis. These types of discontinuities in the observing system may lead to the assimilation of biased observations (Thorne and Vose 2010; Dee et al. 2011; Parker 2016). Hence an effective bias correction procedure needs to be adopted for satellite radiances.

In IMDAA, satellite radiances are bias corrected using a variational bias correction method (VarBC), based on that developed initially at NCEP (Derber and Wu 1998). In this bias correction method, the data assimilation system keeps track of the bias predictors for radiances from all available satellite instruments. Biases in the satellite radiances vary with time, geography, air mass, scan position of satellite instrument, and satellite position in its orbit (Bell et al. 2008; Lu et al. 2011; Doherty et al. 2015). The VarBC scheme in IMDAA used the same method implemented in the UM system (Cameron and Bell 2018) and closely follows the incremental formulation based on Auligné et al. (2007). The bias parameters are determined by the linear combination of predictors (Dee et al. 2011). The predictors used are 850–300-hPa thickness, 200–50-hPa thickness, and total column water vapor. The VarBC system used in IMDAA includes several enhancements, such as a harmonized adaptation rate, a hybrid scan bias correction scheme, and time-dependent biases. Additionally, satellite instruments and their channels are rejected when they become unreliable. The blacklisting information is similar to that used in ECMWF reanalyses (Dee et al. 2011). Many regional and convective scale NWP assimilation systems use the VarBC method for satellite data bias correction (Gustafsson et al. 2018). Compared to the global assimilation system, the IMDAA reanalysis system has taken more time to stabilize (more details are provided in the supplemental material). The observation operator used for simulating the satellite brightness temperatures is RTTOV version 9 (Saunders 2010).

b. *Observation coverage*

The basic types of observations assimilated are shown in Fig. 3. The observing system changed considerably between 1979 and 2018. Observations present throughout the IMDAA period are those from the surface, sonde, and aircraft, similar to other reanalyses (Gelaro et al. 2017; Dee et al. 2011; Saha et al. 2010; Kobayashi et al. 2015; Hersbach et al. 2020). Figure 6 shows the geographical coverage of the conventional observations (e.g., surface, sonde, and aircraft) assimilated in the IMDAA system during the 0000 UTC cycles for June, July, August, and September (JJAS; the Indian summer monsoon period) for four different years, 1984, 1994, 2004, and 2014, randomly selected as representative years for the decades. The cumulative number of observations from each type assimilated in the IMDAA is indicated in Fig. 6. A significant number of surface observations over the Indian Ocean declined in the 1990s and early 2000s, but some increase is noticed over land during this period. An increase in total surface and upper air observations over the IMDAA domain is seen from mid-2000. Unlike surface and sonde observations, aircraft observations (both AIREP and AMDAR) show a considerable increase over time, as reported by other reanalyses. In the first decade of IMDAA, only manual aircraft reports (AIREP) were available, but automatic aircraft observations (AMDAR) appeared later.

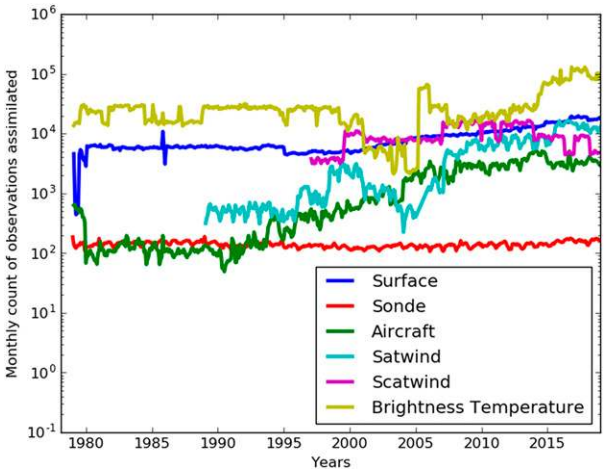

FIG. 7. Monthly mean of daily average (all cycles, for sonde only 0000 UTC) count of various observations assimilated in the IMDAA system.

Figure 7 shows the monthly mean daily average count of different types of observations assimilated. Table 1 shows the monthly mean (considered all four assimilation cycles per day) count of various observations, except for sonde data, assimilated for four decades, 1979–88, 1989–98, 1999–2008, and 2009–18. In the sonde case, only the monthly mean of 500-hPa observations at 0000 UTC is given. It is clear from Fig. 7 and Table 1 that there was a considerable increase in observations with time. Most of the increase is seen in satellite observations, including clear-sky radiances from both geostationary and polar satellites, AMVs, and scatterometer winds. There is also an overall increase in conventional data, particularly surface and aircraft observations.

4. Evaluation of IMDAA

Reanalysis products need to be carefully evaluated so that they can be used confidently in many applications. This section describes the model background and analysis departures of different observation parameters, time–vertical cross sections of analysis increments, and comparison and verification of major weather phenomena during different seasons in IMDAA with IMD observations and ERA5. The ERA5 global reanalysis has a horizontal grid spacing of ~30 km and resolves the atmosphere using 137 levels from the surface up to a height of 80 km. It provides hourly estimates of many atmospheric, land, and oceanic variables (Hersbach et al. 2020).

TABLE 1. Monthly mean daily average (considering all four assimilation cycles) counts of various observations assimilated, except for sonde, which showed only the 0000 UTC profiles that reported temperature at 500 hPa, in the IMDAA system for four decades.

Observation type	1979–88	1989–98	1999–2008	2009–18
Surface	4950	4622	7433	11 544
Sonde	156	150	145	168
Aircraft	151	286	1733	3418
Satellite radiances	22 213	25 181	17 211	50 070
AMVs	519	945	2460	9358
Scatterometer winds	—	4462	10 025	7506

Hypothesis tests are employed to identify the level of statistical significance of the datasets used in this study. A single *t* test is conducted on various estimates from IMDAA and ERA5, while a paired *t* test is employed for the duplets, like (IMD, IMDAA), (IMD, ERA5), and (IMDAA, ERA5).

a. Background and analysis fit to observations

The performance of IMDAA is assessed in terms of background and analysis fit to observations assimilated in the IMDAA system. The departure of observations (O) from the corresponding model background (B) gives information about the observation quality and also shows how the forecast system retains information from observations. The departure of observation from the analysis (A) is also calculated at the end of each data assimilation cycle, showing how closely the reanalysis fits the assimilated data at observation locations.

Figure 8 shows departure statistics from 1979 to 2018 for all surface pressure and surface temperature observations assimilated in the 0000 UTC data assimilation cycle against the model background and analysis fields. Figure 8a shows the monthly mean of daily average Root-mean-square (RMS) departures in surface pressure (hPa). The lightly shaded dots in each curve represent the corresponding daily values. As expected, the RMS departure in *O-A* is smaller than that of *O-B*. RMS departure in surface pressure *O-B* and *O-A* lie between 1 and 1.5 hPa during the initial period of IMDAA, but later both are reduced with RMS difference in *O-A* below 1 hPa. Similar trends can be seen in surface pressure departure over the Northern Hemispheres in ERA-Interim (Dee et al. 2011) and MERRA-2 (Gelaro et al. 2017). Figure 8b is similar to Fig. 8a, but for surface temperature departures (K) for both *O-B* and *O-A*. The RMS departure in the background surface temperature against observation lies between 2 and 3 K throughout the study. In contrast, those of *O-A* is between 2 and 2.6 K. Monthly mean RMS departures of *O-B* and *O-A* of surface observations (pressure and temperature, 12 months × 40 years of samples in each series) show that the mean values of the two time series differ

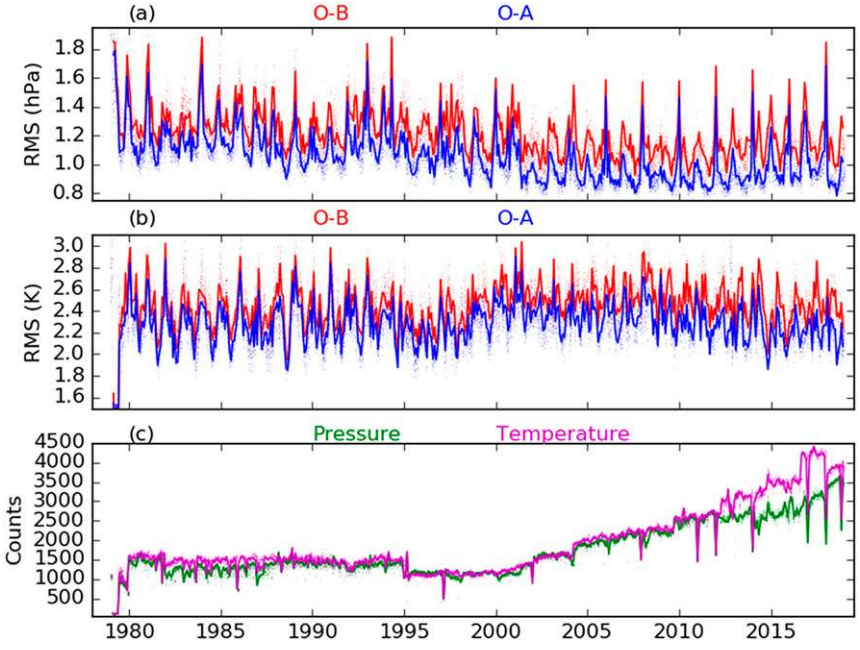

FIG. 8. Time series of the monthly mean of daily average RMS departures of (a) surface pressure (hPa), (b) surface temperature (K), and (c) counts of observations assimilated in IMDAA during the 0000 UTC cycle. IMDAA background departures are shown in red and analysis departure in blue. Green and purple curves represent the monthly mean count of surface pressure and surface temperature observations assimilated in the IMDAA system. The lightly shaded dots in each curve represent the corresponding daily values. Mean RMS departures of surface observations show that the mean values of the *O-B* and *O-A* time series differ significantly ($p < 0.05$).

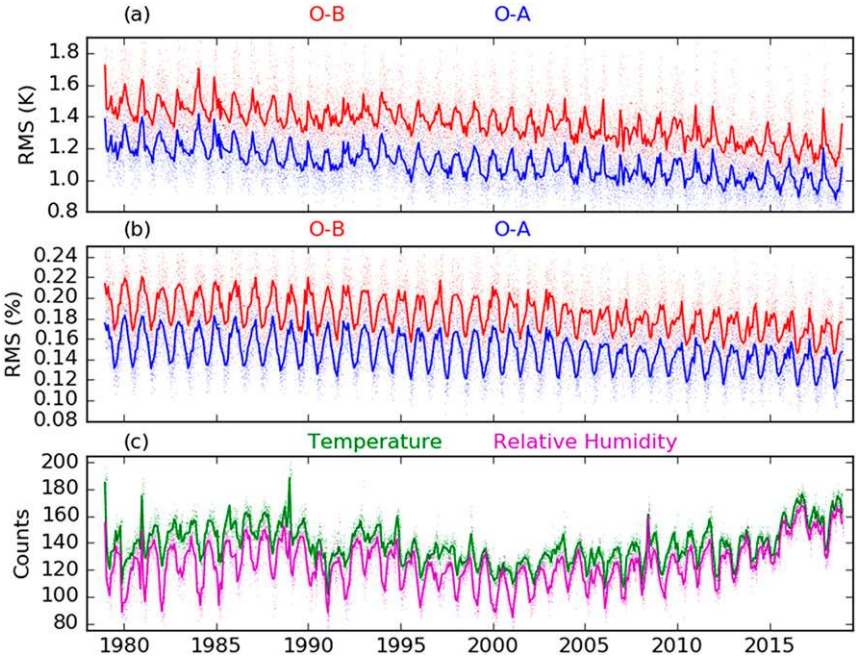

FIG. 9. Time series of the monthly mean of daily average RMS departures of 500-hPa sonde (a) temperature (K), (b) relative humidity (%), and (c) counts of observations assimilated in IMDAA during the 0000 UTC cycle. IMDAA background departures are shown in red and analysis departures are in blue. Green and purple curves represent the monthly mean count of 500-hPa sonde temperature and relative humidity observations assimilated in the IMDAA system. The lightly shaded dots in each curve represent the corresponding daily values. Mean RMS departures of sonde observations at 500 hPa show that the mean values of the *O-B* and *O-A* time series differ significantly ($p < 0.05$).

significantly ($p < 0.05$). Also, the mean difference between *O-A* and *O-B* is always negative, with magnitude decreasing with time, which underlines the robustness of the assimilation and forecast system. Figure 8c is the monthly mean of daily average count of surface pressure and surface temperature observations assimilated. Large numbers of surface temperature observations were assimilated compared to surface pressure observations.

Figure 9 is similar to Fig. 8, but for upper air observations (radiosonde for 0000 UTC cycle) of relative humidity and temperature at 500 hPa. The RMS departure in the background and analysis of both relative humidity and temperature at 500 hPa (Figs. 9a,b) decreased with time. This is due to improvements in consistency and accuracy of the observing system and improvement in the analysis and background, likely due to the introduction of other observations with time. Figure 9a shows that the RMS departures in *O-B* and *O-A* of temperature at 500 hPa lie in the ranges 1.2–1.8 and 0.8–1.4 K, respectively. The background and analysis departures in 500-hPa sonde temperatures are in broad agreement with global reanalyses (Dee et al. 2011; Kobayashi et al. 2015; Saha et al. 2010). Figure 9b shows the RMS departure of background and analysis against observations of relative humidity at 500 hPa, and the values lie between 0.16% and 0.24% and between 0.12% and 0.18%, respectively. Similar to the RMS departures of *O-B* and *O-A* of surface observations, the mean RMS departures of 500-hPa sonde observations (temperature and relative humidity) also show that the mean values of the two time series differ significantly ($p < 0.05$). Figure 9c shows the counts of assimilated upper air temperature and relative humidity at 500 hPa assimilated. Throughout the period of IMDAA, more temperature observations are assimilated than humidity at 500 hPa. It can be seen from Fig. 9c that the number of upper air observations at 500 hPa assimilated from different stations decreased during the second and third decades compared to the first decade, and a gradual increase is noticed during the last decade of IMDAA.

b. *Analysis increments*

Data assimilation introduces adjustments to the model prognostic variables, referred to as the analysis increments, representing the response of the variational data assimilation to all observations used. Variations in the analysis increments provide an important diagnostic of the system performance, including the changes in the observing system (Dee et al. 2011) and a good proxy for model bias.

Figure 10 summarizes the time series of analysis increment profiles of temperature, zonal wind, and specific humidity in IMDAA. Figure 10a shows the monthly average domain mean for temperature increments at each model level (shown here up to model level 51, ~50 hPa). The vertical structure of the mean increment indicates that there is a systematic cooling in

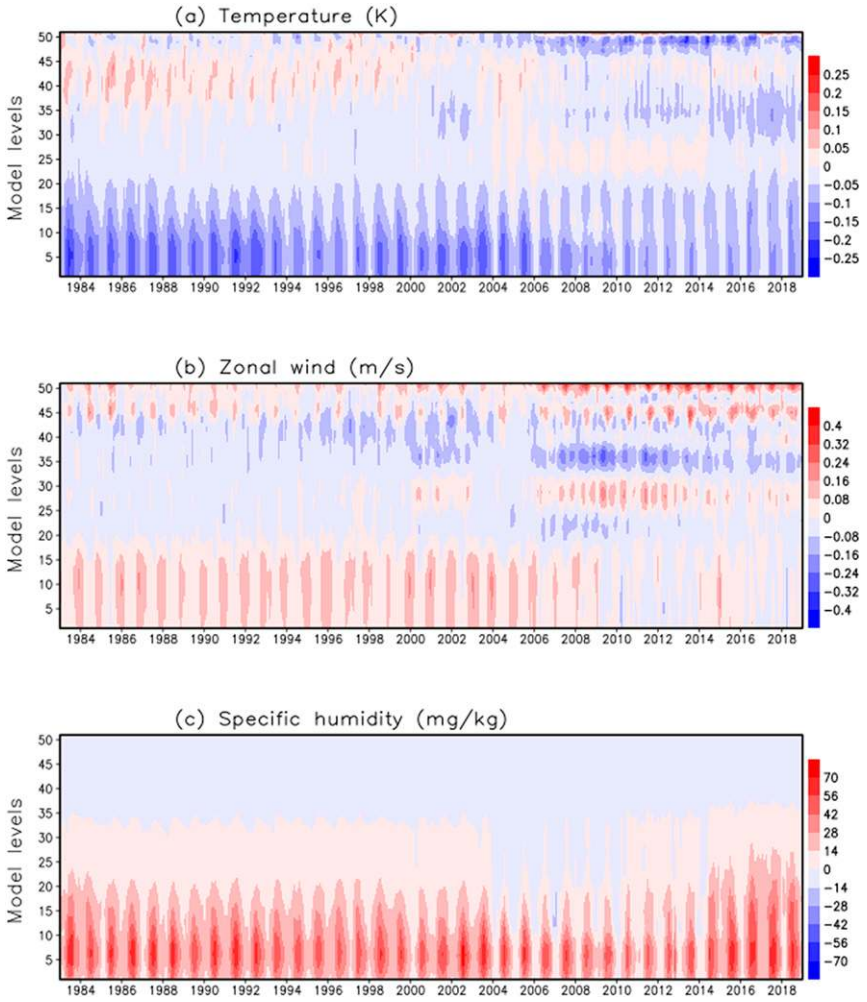

FIG. 10. Monthly average of domain mean of analysis increment profiles for (a) temperature (K), (b) zonal wind ($m\,s^{-1}$), and (c) specific humidity ($mg\,kg^{-1}$). The model levels 15, 30, and 45 represent approximate heights of 1.5, 6, and 14 km, respectively.

the lower troposphere (below level 20, ~700 hPa) throughout IMDAA with maximum cooling below level 10 (~1 km). The magnitude of the lower tropospheric cooling is reduced after 2006. This change can be attributed to the assimilation of more and better quality observations, as seen in Fig. 7. Also, the decrease in the analysis increment in temperature in the lower troposphere is consistent with the RMS departure in the surface observations, as seen from Fig. 8. Systematic positive increments are seen between model levels 35 and 45 (from 8- to 14-km altitude) up to the year 2000, and this is consistent with the period of TOVS (1979–99) radiance assimilation (Fig. 3). Introduction of observations from ATOVS (2000) and other modern satellite instruments (SEVIRI, MVIRI, AIRS, IASI, CrIS, ATMS, AMSR-2, and SAPHIR) in general produced a slight cooling effect in the upper troposphere (Fig. 10a). In the upper troposphere and lower stratosphere, a change from warming to cooling as the assimilation system moved from TOVS to ATOVS and the use of clear-sky brightness temperatures from various geostationary and polar satellites, whereas in the lower troposphere the magnitude of cooling reduced during post 2006. Since the model is fixed and the observations are not, the VarBC also plays a crucial role in adjusting the biases in the upper troposphere and lower stratosphere. Variability in the midtroposphere is small compared to the lower troposphere, and this may be due to a lack of quality observations at these levels or due to a better model forecast (Dee et al. 2011).

The IMDAA domain is mostly a tropical monsoon region where there are strong seasonal changes in circulation, and generally the geostrophic approximation is not valid (Holton and Hakim 2012). Independent wind information is crucial for the analysis system Cardinali (2009). Mass observations can update the wind field through better background error modeling using ensemble covariance proxies, but these are not used in IMDAA. Figure 10b is similar to Fig. 10a, but for the analysis increment for zonal winds. Consistently positive

increments in the lower troposphere (below model level 15, ~1.5 km), upper troposphere and lower stratosphere (above model level 45, ~14 km), and negative increments in the mid-troposphere and some parts of the upper troposphere [between model levels 35 (~8 km) and 45 (~14 km)] in the zonal wind are seen throughout most of the IMDAA period. After introducing ASCAT winds in 2009, the analysis increment of zonal wind is changed in the lower troposphere. Changes in the zonal wind analysis increments are more in the middle troposphere after 2000 due to the assimilation of better-quality AMVs from new generation satellites (Fig. 5) and satellite radiances (Fig. 4). Time series of the RMS departures in the *O-B* and *O-A* of AMVs are included in the supplemental material. Observations from many modern satellite instruments are introduced since 2004. The assimilation of better-quality satellite information might have adjusted the wind fields to attain mass conservation, which can be seen as the larger analysis increments of zonal wind speed above model level 25 since 2005.

Figure 10c summarizes the time series of analysis increment profiles of specific humidity. In general, the increments produced a moistening effect in the lower troposphere, up to around 500 hPa (model levels ~25–30), with some intermittent drying from 2004 to 2013. The magnitude of moistening is large during the Indian summer monsoon period, as seen in Fig. 10c, which may be due to high humidity values during this period. Figure 3 shows that up to 2004, the brightness temperatures assimilated in the IMDAA system are either from TOVS or ATOVS or both. In the IMDAA system, observations like AIRS and SEVIRI are introduced during 2004–05, MVIRI in 2007, IASI in 2010, and ATMS and CrIS in 2012. During this period from 2004 to 2012, the magnitude of the lower troposphere drying is slightly large in the nonmonsoon months, and moistening is less during the monsoon months compared to pre-2004. Brightness temperatures from AMSR-2 and SAPHIR are introduced in 2014, and the magnitude of moistening effect returns to that seen pre-2004, but extending high, as seen from Fig. 10c. Other observations might have contributed to the changes in the analysis increments in specific humidity; however, the changes due to the introduction of the above mentioned instruments are most noticeable. The monthly mean standard deviation in the analysis increments for temperature, zonal wind, and specific humidity are included in the supplemental material.

The large-scale representation of temperature, wind, and humidity fields changed due to the assimilation of more frequent improved observations (Thorne and Vose 2010; Parker 2016), as seen in Fig. 10. A seasonal cycle is noticed in the amplitude of the analysis increments of temperature, zonal wind, and humidity in the lower troposphere, with more adjustments during the southwest monsoon period. This seasonal cycle is in resonance with the RMS differences of *O-B* and *O-A* of the conventional observations (sonde and surface observations shown in Figs. 8 and 9), along with the contribution from other observations used in the assimilation system.

c. *Major weather phenomena over India in IMDAA*

IMDAA products consist of variables that describe the four-dimensional structure of the atmosphere and the land surface.

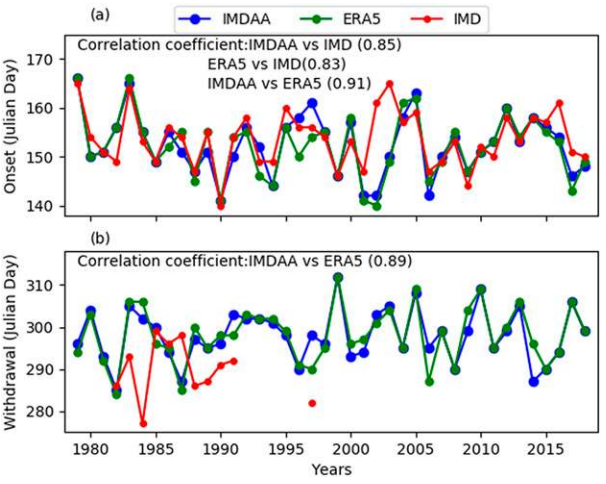

FIG. 11. Indian summer monsoon (a) onset and (b) withdrawal date from IMDAA, ERA5, and IMD Observations. The correlation coefficient is shown in parentheses; there is no correlation between the observed withdrawal dates and reanalyses estimates.

This section describes major weather phenomena that affect the Indian region during the summer monsoon season, as seen from IMDAA and its comparison with ERA5 and IMD observations.

1) MONSOON ONSET AND WITHDRAWAL

The Indian summer monsoon (southwest monsoon) typically lasts from June to September. Major noticeable features of circulation at the time of onset of summer monsoon are strengthening and deepening of westerlies in the lower troposphere and organization and strengthening of easterlies in the upper troposphere over peninsular India (Soman and Kumar 1993). The onset and withdrawal dates of monsoon from IMDAA and ERA5 are calculated using the vertical wind shear index (VWSI) (Prasad and Hayashi 2005). The VWSI is calculated as the zonal wind shear between 850 and 200 hPa averaged over the area 0°–15°N, 50°–95°E. The sign change of this index from negative to positive determines the onset and positive to negative determines the withdrawal date of the summer monsoon. The reanalyses estimates are verified against IMD's official onset dates (Ali et al. 2005; Pai and Rajeevan 2009; Subrahmanyam et al. 2013 and the IMD monsoon reports from 2011 onward) and withdrawal dates [only 11 years are given in Syroka and Toumi (2004)]. It is challenging to observe the withdrawal of the southwest monsoon over peninsular India, mainly because of the arrival of the northeast monsoon before the complete withdrawal of the southwest monsoon.

Figures 11a and 11b show the onset and withdrawal dates calculated from IMDAA and ERA5 and the IMD observed values. The RMS difference from IMD onset dates is approximately three days for both IMDAA and ERA5, and the correlation between both the reanalyses and observation is found to be closer (~0.8) with a slightly better value for IMDAA (indicated in Fig. 11a), likely due to the increased number of

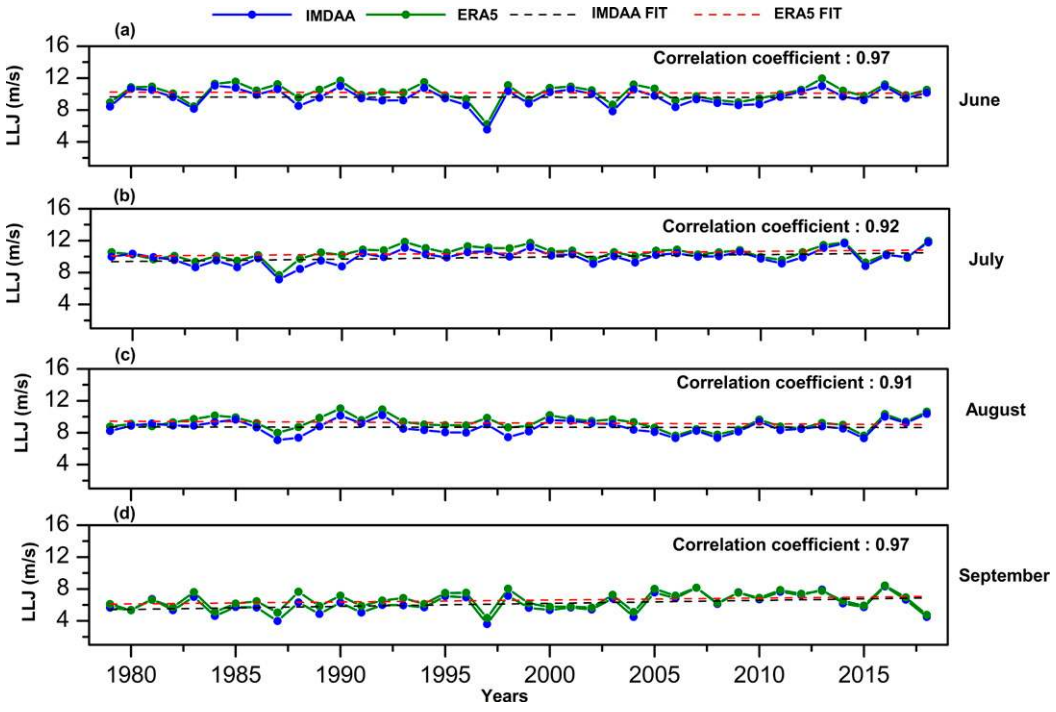

FIG. 12. Monthly mean strength of LLJ (m s^{-1}) for 40 years from IMDAA (blue) and ERA5 (green) during various months of the Indian summer monsoon, (a) June, (b) July, (c) August, and (d) September ($p < 0.05$). Correlation between IMDAA and ERA5 is included in each panel.

observations used for assimilation and the finer horizontal grid spacing. The monsoon onset dates estimated from IMDAA and ERA5 are highly correlated (0.9). Figure 11b is the Indian summer monsoon withdrawal dates computed from IMDAA and ERA5. Both IMDAA and ERA5 show late withdrawal compared to IMD observed dates with standard deviations of approximately 9 and 10 days, respectively. The monsoon withdrawal dates calculated from IMDAA and ERA5 are closely matching and highly correlated (0.89), similar to the onset date estimates. However, there is no correlation between the withdrawal dates estimated from the reanalyses and observation.

2) MONSOON CIRCULATIONS

The Indian summer monsoon is associated with many semipermanent systems [e.g., the seasonal heat low, monsoon trough, Tibetan anticyclone, tropical easterly jet (TEJ), and low-level jet (LLJ)]. Joseph and Raman (1966) and Findlater (1966) established the existence of the LLJ over peninsular India during the summer monsoon season. The LLJ is associated with strong wind shears, both horizontal and vertical. The monsoon core region (18°–25°N, 65°–88°E) is the area where the significant rainfall fluctuations between the active and break spell occur during the summer monsoon (Rajeevan et al. 2010).

The monthly mean strength of the LLJ at 850 hPa over the region 0°–15°N, 50°–95°E is computed from IMDAA and compared with ERA5. Figure 12 shows the monthly mean LLJ strength over the above region for 40 years from 1979 to 2018. Figures 12a–d respectively show the composite LLJ strength during June, July, August, and September. Both the reanalyses show maximum strength of LLJ during June, July, and August, and the strength is reduced in September. LLJ strength estimates from IMDAA and ERA5 are highly correlated (>0.9) during all the four monsoon months (monthly mean correlation between IMDAA and ERA5 is shown in the respective panels of Fig. 12). A decreasing trend in LLJ magnitude is noticed during June and August, whereas an increasing trend in July and September in both the reanalyses; however, a t test shows these trends are not statistically significant ($p \sim 0$). Wilson et al. (2018) also reported a weakening of LLJ flow through peninsular India. Although the LLJ estimated from IMDAA and ERA5 is positively correlated, their mean values differ significantly ($p < 0.05$). The monthly mean difference in the LLJ estimated from IMDAA and ERA5 is in the range of 0.5–1 m s^{-1} with a standard deviation of 1 m s^{-1}.

One of the indicators of establishing the summer monsoon is strong easterly winds in the upper troposphere between 200 and 100 hPa over peninsular India. These persistent strong easterly winds throughout the monsoon season are known as TEJ and were first observed by Koteswaram (1958). The axis of TEJ extends from 5° to 20°N, and the entrance and exit region are from longitude 150°E to 20°W (Flohn 1964; Tanaka 1982). The monthly mean magnitude of TEJ at 150 hPa over the domain 5°–20°N, 40°–100°E (longitude constrained over the Indian region) was calculated for 40 years from IMDAA and compared with that from ERA5. Figure 13 is similar to Fig. 12, but for the magnitude of TEJ during four different

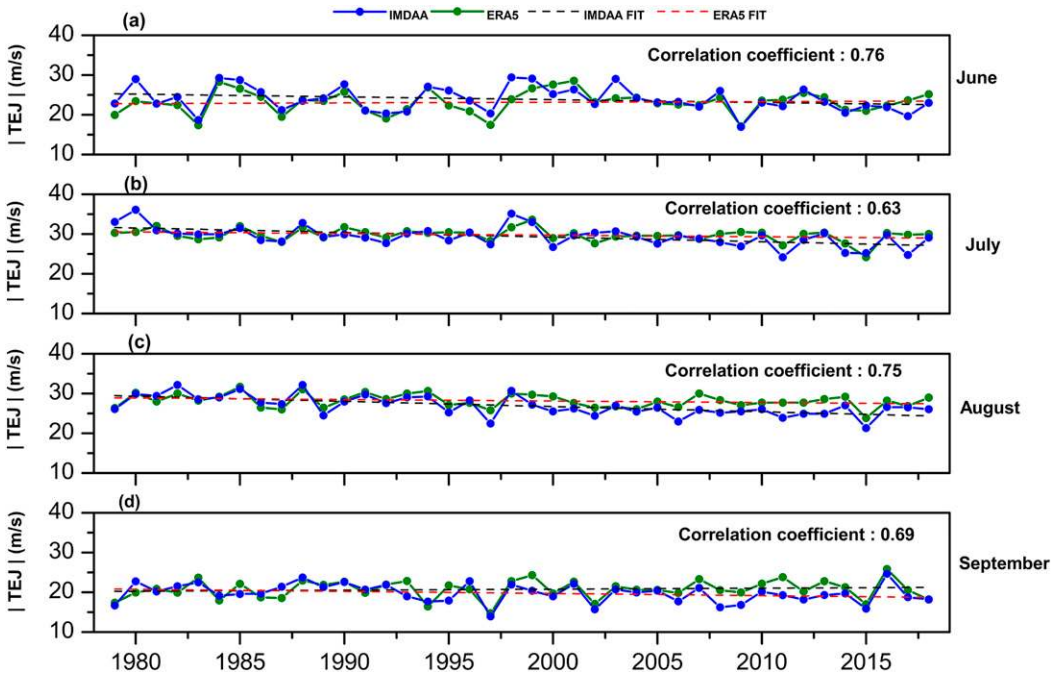

FIG. 13. Monthly mean magnitude of TEJ ($m\,s^{-1}$) for 40 years from IMDAA (blue) and ERA5 (green) during the Indian summer monsoon, (a) June, (b) July, (c) August, and (d) September ($p < 0.05$). Correlation between IMDAA and ERA5 is included in each panel.

monsoon months. The TEJ is strong in both the reanalyses during the peak monsoon months, July and August, as seen from Figs. 13b and 13c. TEJ from the IMDAA reanalyses do not show a strong correlation with ERA5 (correlation coefficients are indicated in the different panels of Fig. 13), contrary to that observed in the case of LLJ. During the initial period, the TEJ was stronger in IMDAA; however, the TEJ showed a decreasing trend in IMDAA and an increasing trend in ERA5, but a t test shows these trends are not statistically significant ($p \sim 0$). Rao et al. (2004) also reported a decreasing TEJ trend during the summer monsoon for 1958–88 using the NCEP–NCAR reanalysis, confirmed by the radiosonde data in their study. The TEJ estimated from IMDAA and ERA5 shows that their mean values differ significantly ($p < 0.05$), except during July. The monthly mean difference in the TEJ estimated from IMDAA and ERA5 is of the order of $\sim 1\,m\,s^{-1}$ with a standard deviation of more than $2\,m\,s^{-1}$.

d. *Comparison of IMDAA rainfall with ERA5 and IMD gridded observations*

A high-resolution gridded daily rainfall dataset of IMD at a resolution of $0.25° \times 0.25°$ over India, produced using a large number of rain gauge observations, is available for more than 100 years (Pai et al. 2014). Figure 14 compares IMDAA average rainfall (mm) with IMD gridded rainfall observations and ERA5 for the Indian summer monsoon period for 40 years, 1979–2018. The daily accumulated seasonal average rainfall from IMD and the difference of IMDAA and ERA5 from IMD are shown in Figs. 14a–c. The correlations between IMD and IMDAA, IMD and ERA5, and IMDAA and ERA5 are shown in Figs. 14d–f. Both the reanalyses underestimate the precipitation over the west coast, the western region, and some parts of the northern region and overestimate over the Indo-Gangetic plains and the east and northeast regions as seen in Figs. 14b and 14c. The rain shadow region over the leeward side of the Western Ghats is better represented in IMDAA than ERA5 and can likely be attributed to the fine grid spacing of IMDAA. The percentage bias in the reanalysis estimated precipitation against the observations (Fig. S7 in the online supplemental material) clearly shows the better performance of IMDAA over the Western Ghats region. The statistical test ($p < 0.05$) shows that both IMDAA and ERA5 estimated precipitation are less correlated (≤ 0.5) with IMD observations over most parts of the country, except over the west coast and the monsoon core region where the reanalyses are better correlated (>0.5) with observation as seen from Figs. 14d and 14e. Precipitation estimated from the IMDAA and ERA5 ($p < 0.05$) is better correlated (>0.5) over most parts of the country, except over the Indo-Gangetic plains, southeast peninsula, and some parts of northern and northeast India (Fig. 14f) where the correlation is less than 0.5. Bosilovich et al. (2011) also reported a better correlation among the reanalysis estimated precipitation than observed precipitation. Details of the mean bias in the precipitation estimates from the IMDAA and ERA5 with respect to the observation (40-yr seasonal average) during the Indian summer monsoon season are included in the supplemental material. It is worth noting that verification

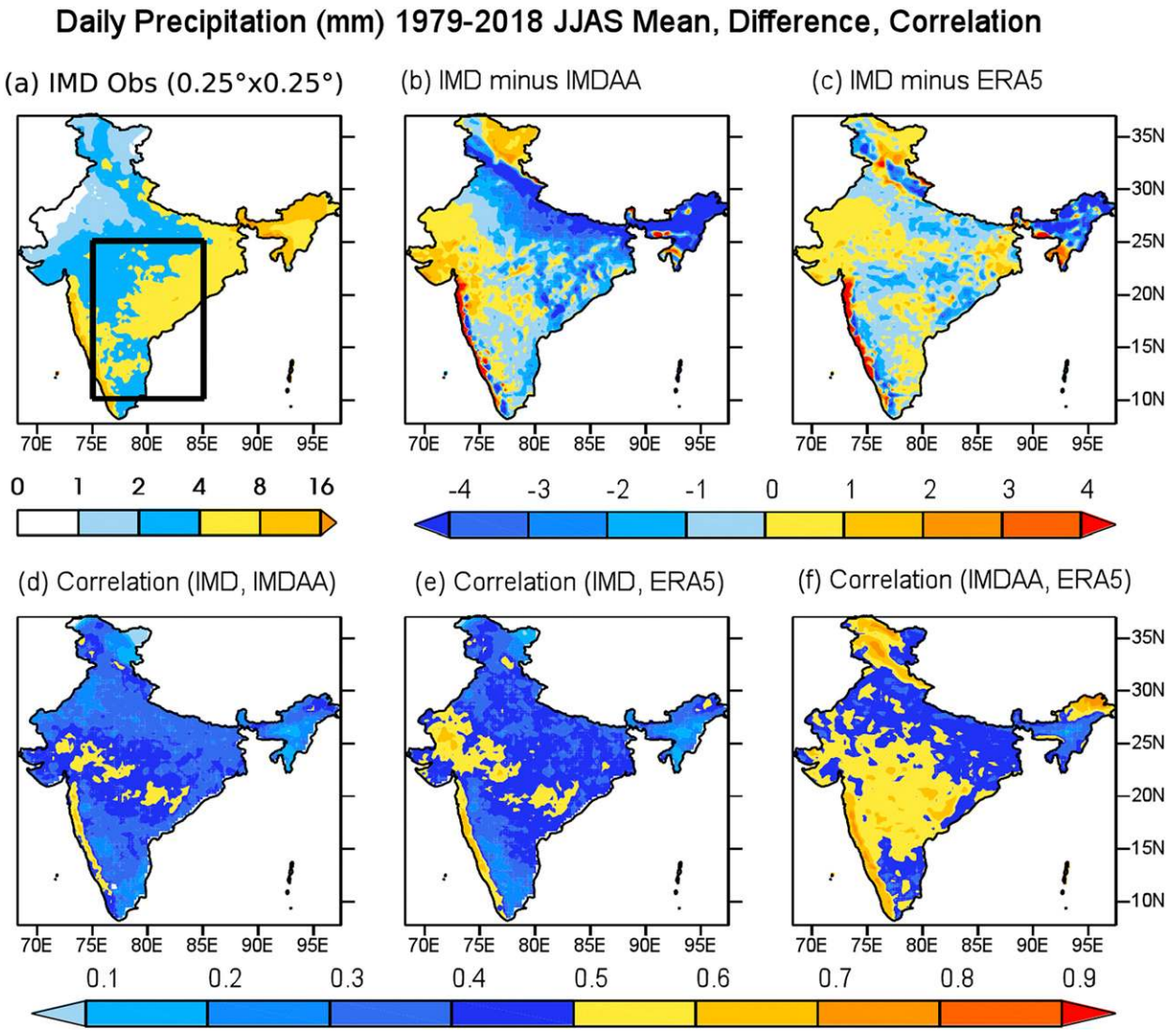

FIG. 14. Seasonal average of daily accumulated (1979–2018) precipitation (mm) during Indian summer monsoon period (JJAS) (a) IMD gridded observations, (b) difference between IMD observation and IMDAA (c) difference between IMD observation and ERA5, and the correlations ($p < 0.05$) between (d) IMD observation and IMDAA, (e) IMD observation and ERA5, and (f) IMDAA and ERA5. (Both the difference and the correlation are shown in the observation resolution of 0.25°). The rectangle marked in Fig. 14a represents the averaging area considered for the rainfall time series shown in Fig. 15.

of IMDAA high-resolution precipitation is limited due to the lack of high-resolution datasets.

The Indian summer monsoon is an annual phenomenon that provides more than 75% of yearly mean rainfall over India (Guhathakurta et al. 2015) and exhibits pronounced variability on both interannual and intraseasonal time scales (Krishnamurthy and Shukla 2000). The ability of IMDAA to capture the active and weak monsoon years (https://mausam.imd.gov.in/imd_latest/contents/monsoon_activity.php) and its comparison with observations and ERA5 are discussed here. Two extreme monsoon years in the first decade of the twenty-first century, 2002 and 2007, with seasonal rainfall departures of −19.2% and +5.7%, respectively (Dutta et al. 2012), have been chosen. Daily accumulated rainfall over the region 10°–25°N, 75°–85°E (marked as a rectangle in Fig. 14a) is used here to study the monsoon variability in these years.

Figures 15a and 15b show time series of daily accumulated area-averaged rainfall during the 122 days of summer monsoon from June to September for 2007 and 2002. Both reanalyses captured the low and high peaks in the daily accumulated rainfall in tandem with the observation but have more rainfall than the observations. IMDAA produced more rainfall during the peak monsoon months, July and August, compared to the observations and ERA5 as seen in Fig. 15. IMD reported 902 mm of seasonal rainfall and 73 no-rain or light-rain days (all-India mean daily accumulated rainfall ≤ 7.5 mm;

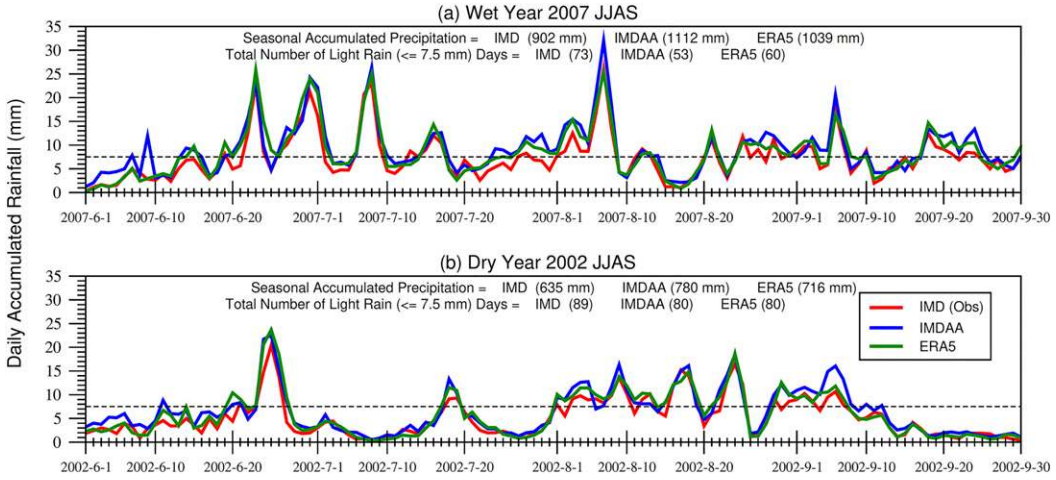

FIG. 15. Daily mean area averaged (10°–25°N, 75°–85°E; averaging area is marked as a rectangle in Fig. 14a) accumulated rainfall during JJAS of (a) an active (wet) monsoon year, 2007, and (b) a weak (dry) monsoon year, 2002.

https://www.imdpune.gov.in/Weather/Reports/glossary.pdf) over this region during the active (above normal) monsoon year of 2007. IMDAA estimated 1112 mm of accumulated rainfall and 53 no-rain or light-rain days, whereas ERA5 produced 1039 mm of accumulated rainfall and 60 no-rain or light-rain days. IMDAA is wetter than both the observation and ERA5 during the active monsoon year, as shown in Fig. 15a. During the weak (below normal) monsoon year of 2002, IMD reported 635 mm accumulated rainfall and 89 no-rain or light-rain days. IMDAA and ERA5 estimated 780 and 716 mm of accumulated rainfall, respectively. Both the reanalyses estimated 80 no-rain or light-rain days. ERA5 is closer to the observation during both the active and weak monsoon years, as shown in Fig. 15. In ERA5, modified convective available potential energy closure (Bechtold et al. 2014) improved the diurnal cycle of convection compared to its predecessors. There are many improvements in the parameterization of the microphysics also in ERA5 compared to IMDAA. Another possible reason for this mismatch may be the higher grid spacing of IMDAA. In addition to improving the NWP skill, increasing the horizontal resolution has a decisive impact on the atmospheric dynamics through a more accurate representation of topography and land surface. Location-specific verification of rainfall estimate (not attempted in this paper) may bring out the benefit of high-resolution IMDAA than the verification against coarse-resolution observations.

The Indian summer monsoon is responsible for the major annual rainfall over India, whereas rainfall during the northeast monsoon (October–December) is mainly confined to South India and Sri Lanka (Rajeevan et al. 2012). Figure 16 is similar to Fig. 14, but for the daily accumulated monthly average rainfall during the northeast monsoon season. In contrast to the summer monsoon precipitation estimates, both reanalyses show a similar precipitation pattern, as seen in Figs. 16b and 16c. The rain-free region has seen in Fig. 16a is more or less well depicted in both reanalyses. Both reanalyses underestimate the precipitation over the east coast, central, and some parts of northern India. Overestimation of precipitation can be seen over the southwest peninsula and northeastern regions. As for the southwest monsoon season, IMDAA is slightly wet during the northeast monsoon season also. The statistical test ($p < 0.05$) shows that the reanalyses estimated precipitation are well correlated (≥ 0.5) with the observations (Figs. 16d,e) over major parts of the country; however, ERA5 shows better spatial correlation particularly over the regions with less rainfall. Precipitation estimates from both IMDAA and ERA5 ($p < 0.05$) show good spatial correlation (>0.5) throughout the country, as seen in Fig. 16f, except over some parts of the northwest regions. It is also important to note that studies have shown that the major leading reanalyses differ by about 20% in rainfall estimates (Gelaro et al. 2017). During the southwest monsoon season, the difference in the rainfall estimates from IMDAA and ERA5 is within 20% over most parts of the country except over the Indo-Gangetic plains. During the northeast monsoon period, rainfall estimates from both IMDAA and ERA5 are closer because the large-scale synoptic circulation controls the rainfall. Mean circulation patterns associated with the southwest and northeast monsoons are included in the supplemental material.

e. Minimum and maximum surface temperature verification

Surface air temperature plays a vital role in weather and climate, which varies spatially and temporally (Vose et al. 2005). Cold and heat waves are extreme temperature events over India during winter and summer periods, respectively. Accurate knowledge of minimum temperatures during the winter (December, January, February) and maximum temperature during summer (March, April, May) are essential to mitigate the severity of extreme weather.

IMD measures daily maximum and minimum temperatures at 2-m levels routinely at many surface observatories

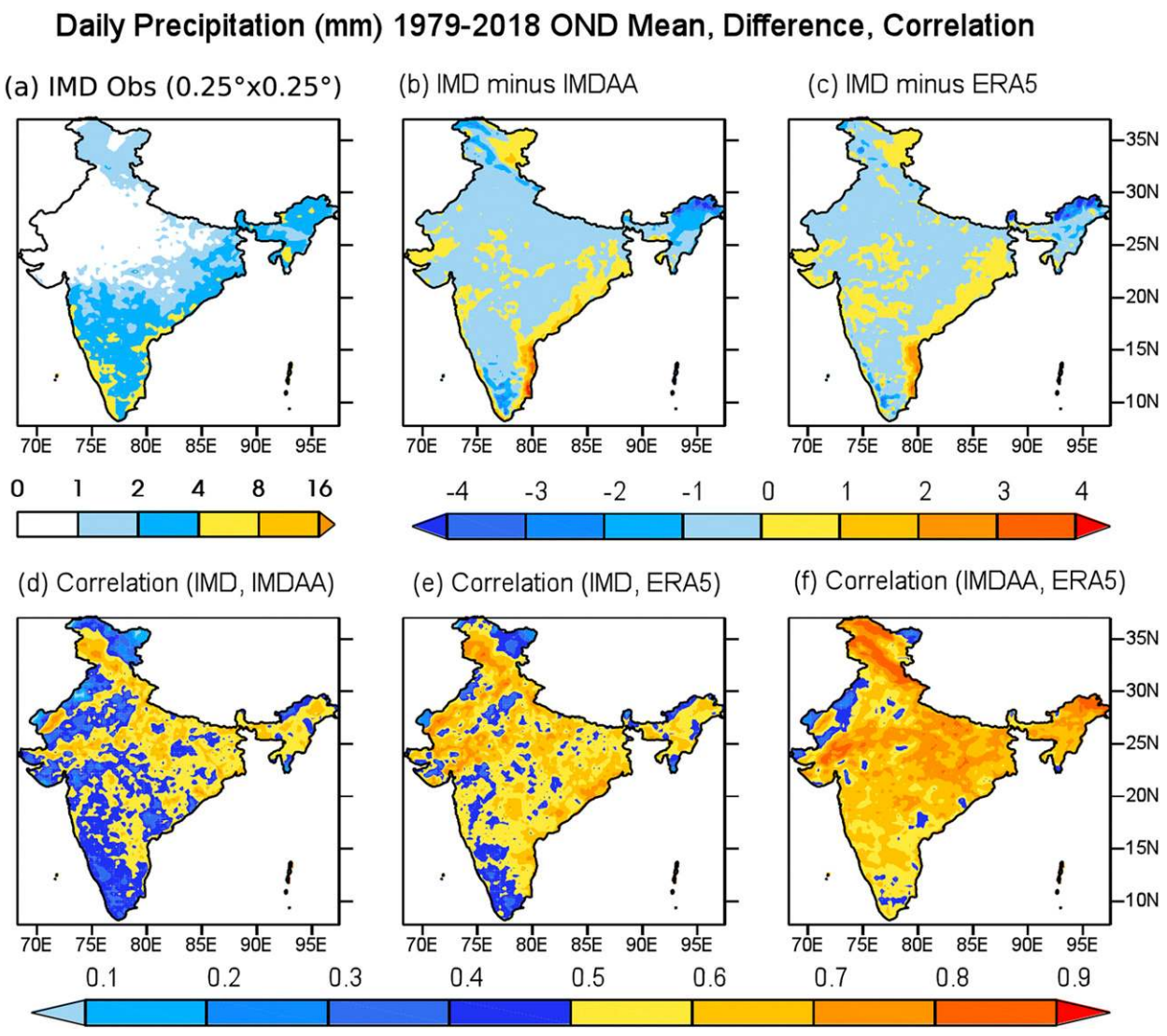

FIG. 16. As in Fig. 14, but for the northeast monsoon season (October–December).

distributed well over India. Using these observations, IMD produces a real-time daily gridded temperature dataset at 1° × 1° resolution (Srivastava et al. 2009). Daily minimum and maximum temperature products are not archived as part of the IMDAA output; hence the minimum and maximum temperatures are calculated from hourly instantaneous temperatures. Figure 17 shows the IMD observed monthly average of daily minimum temperature over the Indian region during winter and the differences between the observation and the same estimate from IMDAA and ERA5 from 1979 to 2018. The correlation coefficients of the minimum temperature between the reanalyses and observations and between IMDAA and ERA5 are also shown in Fig. 17. The correlation is higher than 0.5 all over the Indian landmass, as seen in Figs. 17d–f. IMDAA has a cold bias of around 2°–3°C (Fig. 17b), and ERA5 has a warm bias of the same magnitude (Fig. 17c) compared to IMD observations (Fig. 17a). Both the reanalyses underestimate minimum temperature over the north and some parts of the northeast region; however, these are data-sparse regions where the observations are not reliable (Srivastava et al. 2009). ERA5 shows a better correlation with the observations than IMDAA (Fig. 17e), and the highest correlation is found over the central parts of India. The correlation between IMDAA and ERA5 is moderate (>0.65) over large parts of the Indian landmass (Fig. 17f).

Figure 18 is similar to Fig. 17 but for the monthly average daily maximum temperature from March to May from 1979 to 2018. In contrast to the minimum temperature, a warm bias is noticed in IMDAA (Fig. 18b) and a cooler bias in ERA5 (Fig. 18c) compared to the observed maximum temperature (Fig. 18a). The magnitudes of warm and cool bias, respectively, in the IMDAA and ERA5, are approximately the same, as seen in Figs. 18b and 18c. The difference between the average

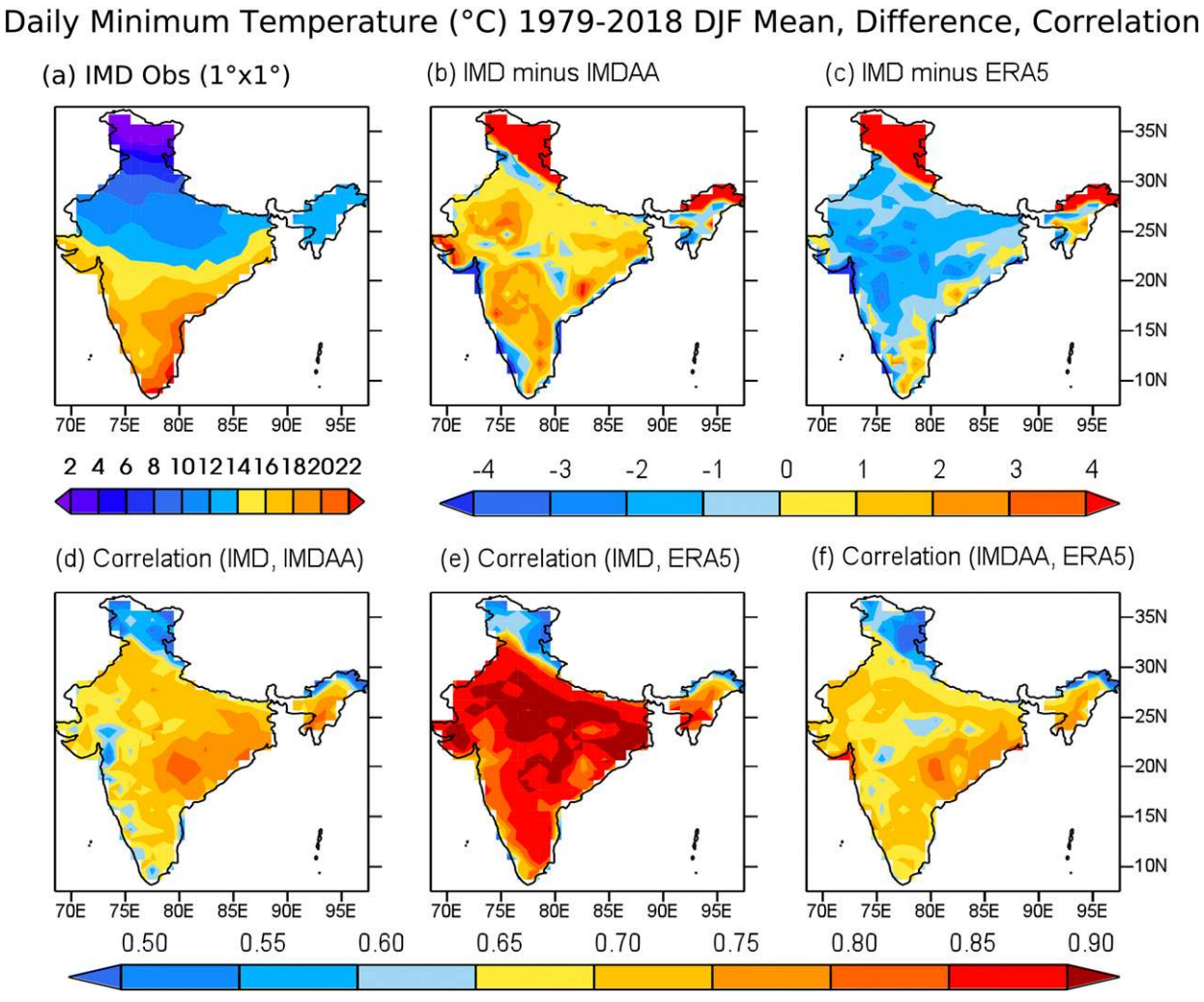

FIG. 17. Seasonal average of daily minimum temperature (°C) during December–February from 1979 to 2018 from (a) IMD gridded observations, (b) difference between IMD observation and IMDAA, (c) difference between IMD observation and ERA5, and the correlations ($p < 0.05$) between (d) IMD observation and IMDAA, (e) IMD observation and ERA5, and (f) IMDAA and ERA5. (Both the difference and the correlation are shown in the observation resolution of 1°.)

daily minimum and maximum temperatures between the cold and hot seasons is less to the south of the Tropic of Cancer as seen in Figs. 17a and 18a, and the maximum changes in temperature are seen over the monsoon core region. In general, the temperature variation is less over southern India and higher over northern India. Similar to the minimum surface temperatures, the maximum surface temperature also shows a good correlation (>0.5) between the reanalyses and observations and also between the reanalyses ($p < 0.05$), as seen in Figs. 18d–f; however, ERA5 shows a better correlation with IMD observations (Fig. 18e). Compared to the other parts of India, the correlation of maximum temperature estimates between the reanalyses and observations is found to be less over peninsular and northeastern India (Figs. 18d,e). It is also noted that the correlation in the maximum temperature estimate is stronger (Fig. 18) compared to the same in the minimum temperature estimates (Fig. 17).

IMDAA produced cooler minimum and warmer maximum temperatures than the observations; the reverse for ERA5. Thus, the mean winter and summer are stronger in IMDAA and weaker in ERA5 than the observed. Over the tropics, local and mesoscale effects are more dominant than synoptic influences, and surface temperature and pressure can change quickly with convective processes (Singh et al. 2020). It is noticed that the high-resolution IMDAA reproduces the variability; however, the lack of high-resolution long-period observations limits verification. Location-specific verification may help to understand the characteristics of IMDAA clearly but was not attempted in this paper.

f. Heavy rainfall episode over Kerala: A case study

As discussed in previous sections, India receives maximum rainfall during the summer monsoon period. Any deviation in monsoon from its normal behavior imposes a significant

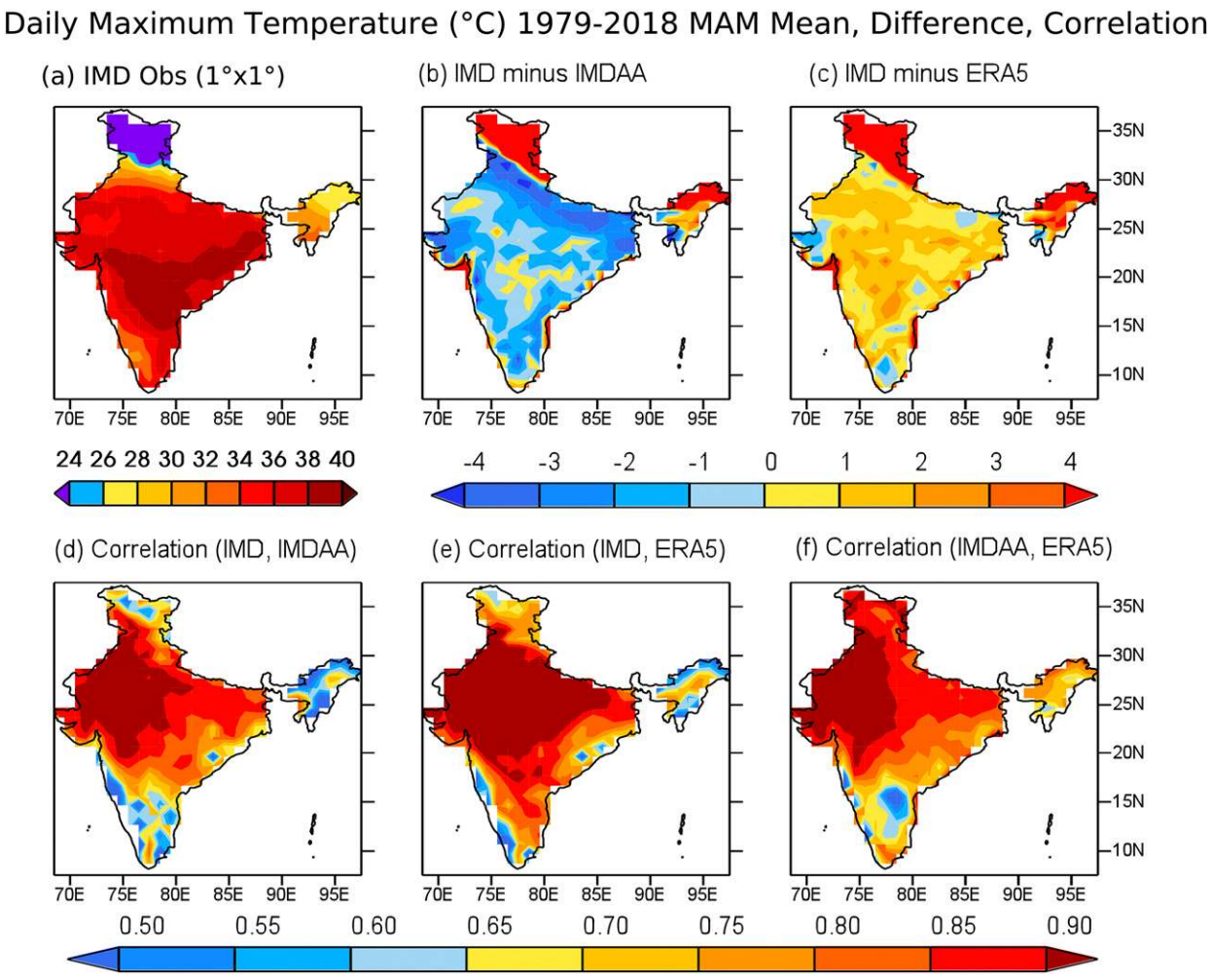

FIG. 18. As in Fig. 17, but for the maximum temperature (°C) from March to May for years 1979–2018.

impact on the Indian population and economy. In recent years several parts of India witnessed many devastating flood events during the summer monsoon. In August 2018, Kerala, a southern state of India, experienced heavy rainfall and an exceptional flood, the second worst after 1924 in the recorded history of the region. Kerala has a complex terrain with large heterogeneity, bounded by the Arabian Sea to the west and the Western Ghats mountains with peaks of more than 2500 m to the east. Details of the 2018 Kerala flood events and the underlying synoptic conditions are discussed in Hunt and Menon (2020), Ashrit et al. (2020b), and Mohandas et al. (2020). The rain episode was mostly due to a monsoon depression formed over the Bay of Bengal from 13 to 17 August 2018 that immediately followed a monsoon low pressure system from 6 to 9 August 2018.

Figures 19a and 19b show the location of the depression formed over the Bay of Bengal, the associated mean sea level pressure, 850-hPa winds, and the rainfall from IMDAA and ERA5 respectively, valid at 1200 UTC 16 August 2018. The minimum surface pressure estimates from IMDAA and ERA5 valid at 1200 UTC 16 August 2018 are 993.9 and 994.4 hPa, nearly at the same location (21°N, 80°E) in both reanalyses. The IMD observed (http://www.rsmcnewdelhi.imd.gov.in/uploads/report/26/26_2bfb44_d15-17aug.pdf) minimum surface pressure was 994 hPa at (21.1°N, 78.3°E). Both reanalyses closely captured the magnitude and location of the depression. IMDAA produced highly concentrated rainfall associated with this depression, whereas more widespread rainfall is seen in ERA5 (Fig. 19).

Figure 20 shows 6-hourly comparisons of minimum surface pressure estimated from IMDAA and ERA5 against IMD observation and the magnitude of the position errors of the low pressure system and depression formed over the Bay of Bengal during August 2018. In general, the difference in the minimum surface pressure between the observation and reanalyses is less in IMDAA (Fig. 20a); however, most of the time, the location is better captured in ERA5 as seen from the magnitude of the position error (Fig. 20b). A more realistic estimate of the magnitude of the minimum pressure in the IMDAA can be attributed to its fine grid spacing. Differences in the model

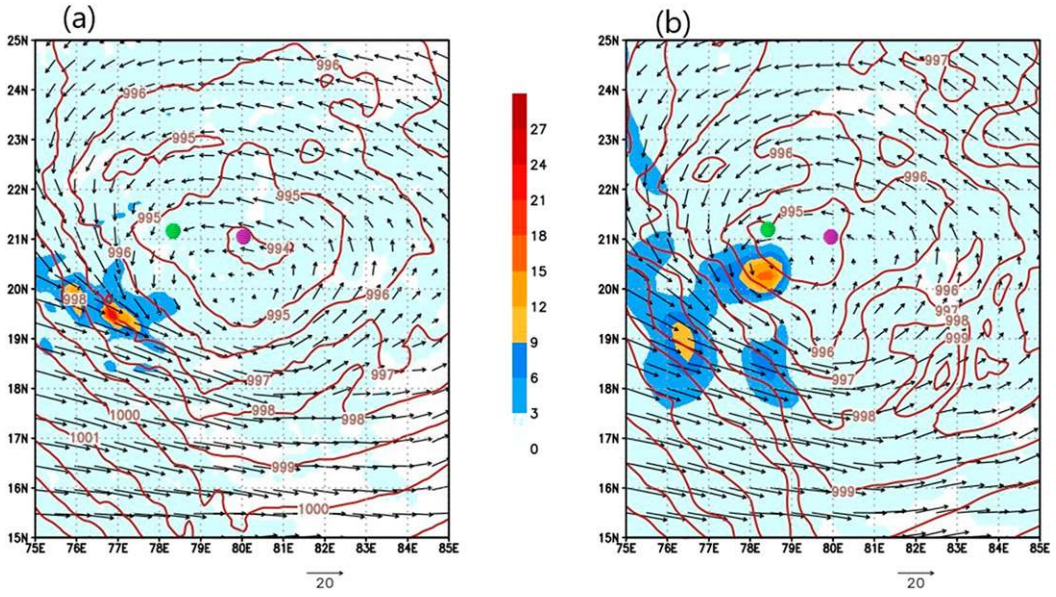

FIG. 19. Location of the depression over the Bay of Bengal from (a) IMDAA and (b) ERA5 valid at 1200 UTC 16 Aug 2018. Contours represent the mean sea level pressure (hPa), vectors are 850-hPa winds, and the shading shows the associated rainfall (mm). The green and purple dots in the figures represent the location of the depression from the IMD observation (994 hPa at 21.1°N, 78.3°E) and reanalyses: IMDAA (993.9 hPa; 21°N, 80°E) and ERA5 (994.4 hPa; 21°N, 80°E).

physics and assimilation technique must also have contributed to the nature of the results from the two reanalyses.

Figure 21 depicts the average hourly rainfall (mm) over Kerala from 6 to 18 August 2018. IMDAA captured the finer features of the rainfall reasonably well, especially over the mid- and high-altitude parts of Kerala (Fig. 21b), even though the rainfall magnitude does not closely match the observations (Fig. 21a). IMDAA captured orographic rainfall better than ERA5.

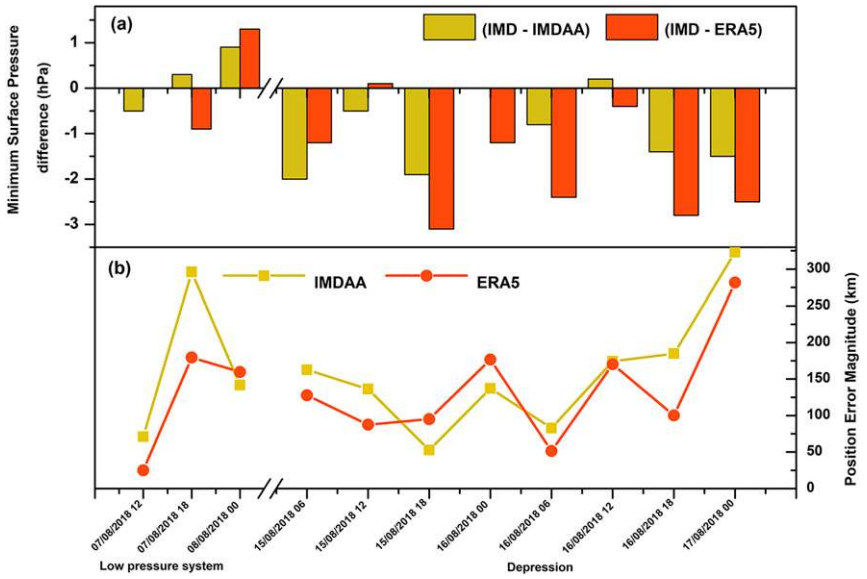

FIG. 20. Comparison of 6-hourly IMD observation and corresponding estimates from IMDAA and ERA5 of (a) difference in the minimum surface pressure (hPa) and (b) position error magnitude (km) during the two events, the low pressure system (7–8 Aug 2018) and depression (15–17 Aug 2018) formed over the Bay of Bengal.

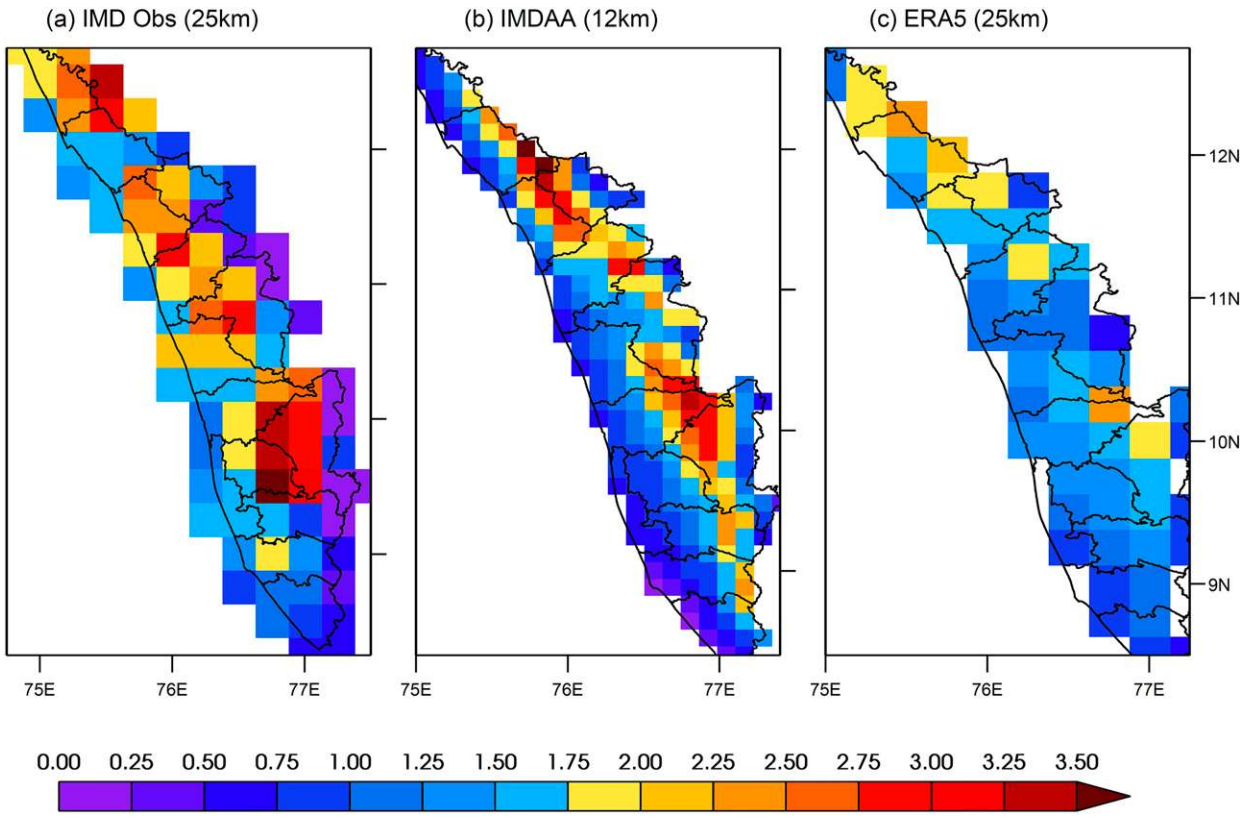

FIG. 21. Hourly averaged (6–18 Aug 2018) rainfall (mm) over Kerala (a) IMD gridded observations (b) IMDAA and (c) ERA5.

ERA5 also produced rainfall associated with these episodes (Fig. 21c), but the magnitude is less than observations and IMDAA. It is noticed that (as described in section 4d) IMDAA is generally wet compared to ERA5.

5. Discussion and summary

The IMDAA regional atmospheric reanalysis was developed with the primary aim to produce a continuous high-resolution, long-term, homogeneous dataset to unravel the finer features of the Indian summer monsoon. Presently, this is the highest resolution (~12 km) long-term (1979–2018) reanalysis available over the Indian monsoon region. IMDAA used an advanced data assimilation system (4D-Var) and the Unified Model of the U.K. Met Office. A continuous increase in the number of assimilated observations (both conventional and satellite) is seen during IMDAA. The robustness of the assimilation system can be seen in the decrease in the RMS difference in background and analysis from the observation. Analysis increments of different meteorological variables clearly show the response of the data assimilation system with the introduction of new observations over time.

Many of the semipermanent features of the Indian monsoon season brought out by IMDAA are compared with ERA5 and with IMD observations. The monsoon onset dates from IMDAA and ERA5 are highly correlated and closer to the IMD observed dates. Similarly, the monsoon withdrawal dates estimated from IMDAA and ERA5 are highly correlated. IMDAA shows a slightly weak low-level jet (LLJ) and tropical easterly jet (TEJ) compared to ERA5.

Comparison of Indian monsoon rainfall from IMDAA and ERA5 with observations shows that spatial distribution of rainfall from IMDAA broadly matches with observations over most parts of India. However, in general, IMDAA shows more rainfall during the southwest and northeast monsoons than observations and ERA5. Compared to the observations, IMDAA produces a slightly cooler winter and hotter summer, and the reverse in the ERA5; however, ERA5 matches better with maximum and minimum observed temperatures.

Two extreme years of Indian summer monsoon in the first decade of the twenty-first century were examined to assess the capability of IMDAA in capturing the intraseasonal and interannual variability of the summer monsoon. IMDAA produced more rainy days during both active and weak monsoon years compared to observations and ERA5. The interannual and intraseasonal variability of the Indian summer monsoon is better represented in the two extreme years, but this needs to be explored further for the entire IMDAA period to draw a more meaningful conclusion. The case study of a severe flood event in August 2018 showed that IMDAA captured the finer features of the extreme rainfall event over the Western Ghats, highlighting the benefits of high-resolution data over complex terrains. In summary, IMDAA offers a suitable quality high-resolution reanalysis dataset that can be used for research on weather and climate over the Indian monsoon region;

however, the lack of high-resolution observations limits the verification of many of the IMDAA features.

Acknowledgments. The authors gratefully acknowledge the financial support given by the Ministry of Earth Sciences, Government of India (Grant /Project MM/SERP/Met Office_UK/2013/INT-9/002) to conduct this research under the National Monsoon Mission. The authors thank ECMWF for making available the ERA-Interim reanalysis and their archive of global observations collated for reanalyses. The authors also extend their gratitude to IMD for providing additional conventional observations and gridded observations for validation. The authors thank Mr. Gopal Iyengar (MoES), Dr. A. K. Mitra (NCMRWF), Dr. M. Das Gupta (NCMRWF), Dr. Saji Mohandas (NCMRWF), Dr. Sana Mahmood (Met Office), and Dr. Peter Jermey (Met Office) for their support at various levels of this project. The authors also express their gratitude to the editor and the anonymous reviewers for their constructive and critical comments, which helped to improve the quality of this paper. The authors take this opportunity to thank one and all who are directly or indirectly involved in this project.

Data availability statement.. Forty years of IMDAA reanalysis products are available at https://rds.ncmrwf.gov.in.

APPENDIX

List of Acronyms

4D-Var	Four-dimensional variational data assimilation
AIREP	Routine automated report of in-flight weather conditions
AIRS	Advanced Infrared Sounder
AMDAR	Aircraft meteorological data relay
AMSR	Advanced Microwave Scanning Radiometer
AMSU-A	Advanced Microwave Sounding Unit A
AMSU-B	Advanced Microwave Sounding Unit B
AMV	Atmospheric motion vector
ASCAT	Advanced Scatterometer
ASCII	American Standard Code for Information Interchange
ATMS	Advanced Technology Microwave Sounder
ATOVS	Advanced TIROS Operational Vertical Sounder
BARRA	Bureau of Meteorology Atmospheric High-Resolution Regional Reanalysis for Australia
BUFR	Binary universal form for the representation of meteorological data
BUOY	Report of a buoy observation
CrIS	Cross-Track Infrared Sounder
ECMWF	European Centre for Medium-Range Weather Forecasts
ENDGame	Even Newer Dynamics for General Atmospheric Modeling of the Environment
ERA	ECMWF Re-Analysis
ERA-40	45-yr ERA from September 1957 to August 2012
ERA5	70-yr ERA starting from January 1950 onward with timely updates
ERA-Interim	40-yr ERA from January 1979 to August 2019
ERS	European Remote Sensing Satellite
EURO4M	European Reanalysis and Observation for Monitoring
GCOM-W1	Global Change Observation Mission–Water (first satellite)
GMS	Geostationary Meteorological Satellite series of Japan
GTS	Global Telecommunication System
HadISST2	Hadley Centre Sea Ice and Sea Surface Temperature dataset (version 2)
HIRS	High Resolution Infrared Sounder
IASI	Infrared Atmospheric Sounding Interferometer
IMD	India Meteorological Department
IMDAA	Indian Monsoon Data Assimilation and Analysis
JMA	Japan Meteorological Agency
JRA-55	Japanese 55-Year Reanalysis Project
JULES	Joint U.K. Land Environment Simulator
LLJ	Low-level jet
MERRA	Modern-Era Retrospective Analysis for Research and Applications
METAR	Meteorological Aerodrome Reports
Meteosat	Meteorological Satellite
MetOp	Meteorological Operational Satellite
MHS	Microwave Humidity Sounder
MTSAT	Multifunctional Transport Satellite
MVIRI	Meteosat Visible and Infrared Imager
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NCMRWF	National Centre for Medium Range Weather Forecasting
NOAA	National Oceanographic and Atmospheric Administration
NPP	National Polar Orbiting Partnership
NWP	Numerical weather prediction
ODB	Observation Data Base (ECMWF)
OPS	Observation Processing System
PILOT	Wind report from pilot balloon
QuickSCAT	Quick Scatterometer
RTTOV	Radiative Transfer for TOVS
SAPHIR	Sondeur Atmospherique du Profil d'Humidite Intertropicale par Radiometrie
SARR	South Asian Regional Reanalysis
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SHIP	Report of surface observation from a sea station
SST	Sea surface temperature
SYNOP	Surface synoptic report
TEJ	Tropical easterly jet
TEMP	Report from radio sounding
TIROS	Television Infrared Observation Satellite
TOVS	TIROS Operational Vertical Sounder
UM	Unified Model
VarBC	Variational bias correction
VWSI	Vertical wind shear index

WMO	World Meteorological Organization
WRF	Weather Research and Forecasting Model

REFERENCES

Ali, M., D. Joardar, and B. R. Loe, 2005: Variability of southwest monsoon over Rajasthan and Kerala. *Mausam*, **56**, 593–600.

Arakawa, A., and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods Comput. Phys.*, **17**, 173–265, https://doi.org/10.1016/B978-0-12-460817-7.50009-4.

Ashrit, R., and Coauthors, 2020a: IMDAA regional reanalysis: Performance evaluation during Indian summer monsoon season. *J. Geophys. Res. Atmos.*, **125**, e2019JD030973, https://doi.org/10.1029/2019JD030973.

——, and Coauthors, 2020b: Prediction of the August 2018 heavy rainfall events over Kerala with high resolution NWP models. *Meteor. Appl.*, **27**, e1906, https://doi.org/10.1002/met.1906.

Auligné, T., A. P. McNally, and D. P. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 631–642, https://doi.org/10.1002/qj.56.

Bechtold, P., N. Semane, P. Lopez, J. P. Chaboureau, A. Beljaars, and N. Bormann, 2014: Representing equilibrium and non-equilibrium convection in large-scale models. *J. Atmos. Sci.*, **71**, 734–753, https://doi.org/10.1175/JAS-D-13-0163.1.

Bell, W., and Coauthors, 2008: The assimilation of SSMIS radiances in numerical weather prediction models. *IEEE Trans. Geosci. Electron.*, **46**, 884–900, https://doi.org/10.1109/TGRS.2008.917335.

Best, M. J., and Coauthors, 2011: The Joint UK Land Environment Simulator (JULES), model description—Part 1: Energy and water fluxes. *Geosci. Model Dev.*, **4**, 677–699, https://doi.org/10.5194/gmd-4-677-2011.

Bosilovich, M., J. Chen, F. Robertson, and R. Adler, 2008: Evaluation of global precipitation in reanalyses. *J. Appl. Meteor. Climatol.*, **47**, 2279–2299, https://doi.org/10.1175/2008JAMC1921.1.

——, F. Robertson, and J. Chen, 2011: Global energy and water budgets in MERRA. *J. Climate*, **24**, 5721–5739, https://doi.org/10.1175/2011JCLI4175.1.

Boutle, I. A., J. E. J. Eyre, and A. P. Lock, 2014a: Seamless stratocumulus simulation across the turbulent gray zone. *Mon. Wea. Rev.*, **142**, 1655–1668, https://doi.org/10.1175/MWR-D-13-00229.1.

——, S. J. Abel, P. G. Hill, and C. J. Morcrette, 2014b: Spatial variability of liquid cloud and rain: Observations and microphysical effects. *Quart. J. Roy. Meteor. Soc.*, **140**, 583–594, https://doi.org/10.1002/qj.2140.

Brown, A. R., R. J. Beare, J. M. Edwards, A. P. Lock, S. J. Keogh, S. F. Milton, and D. N. Walters, 2008: Upgrades to the boundary-layer scheme in the Met Office numerical weather prediction model. *Bound.-Layer Meteor.*, **128**, 117–132, https://doi.org/10.1007/s10546-008-9275-0.

——, S. Milton, M. Cullen, B. Golding, J. Mitchell, and A. Shelly, 2012: Unified modeling and prediction of weather and climate: A 25-year journey. *Bull. Amer. Meteor. Soc.*, **93**, 1865–1877, https://doi.org/10.1175/BAMS-D-12-00018.1.

Cameron J., and W. Bell, 2018: The testing and implementation of variational bias correction (VarBC) in the Met Office global NWP system. Weather Science Tech. Rep. 631, 22 pp.

Candy, B., R. W. Saunders, D. Ghent, and C. E. Bulgin, 2017: The impact of satellite-derived land surface temperatures on numerical weather prediction analyses and forecasts. *J. Geophys. Res. Atmos.*, **122**, 9783–9802, https://doi.org/10.1002/2016JD026417.

Cardinali, C., 2009: Monitoring the observation impact on the short-range forecast. *Quart. J. Roy. Meteor. Soc.*, **135**, 239–250, https://doi.org/10.1002/qj.366.

Charney, J. G., and N. A. Phillips, 1953: Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. *J. Meteor.*, **10**, 71–99, https://doi.org/10.1175/1520-0469(1953)010<0071:NIOTQG>2.0.CO;2.

Clark, D. B., and Coauthors, 2011: The Joint UK Land Environment Simulator (JULES), model description—Part 2: Carbon fluxes and vegetation dynamics. *Geosci. Model Dev.*, **4**, 701–722, https://doi.org/10.5194/gmd-4-701-2011.

Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, https://doi.org/10.1002/qj.776.

Cullen, M. J. P., T. Davies, M. H. Mawson, J. A. James, S. C. Coulter, and A. Malcolm, 1997: An overview of numerical methods for the next generation UK NWP and climate model. *Numerical Methods in Atmosphere and Ocean Modelling: The André Robert Memorial Volume*, C. Lin, R. Laprise, and H. Ritchie, Eds., Canadian Meteorological and Oceanographical Society, 425–444.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **131**, 1759–1782, https://doi.org/10.1256/qj.04.101.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Derber, J. C., and W. S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299, https://doi.org/10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2.

de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen, 2013: A simplified extended Kalman filter for the global operational soil moisture analysis at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **139**, 1199–1213, https://doi.org/10.1002/qj.2023.

Dickinson, R. E., R. M. Errico, F. Giorgi, and G. T. Bates, 1989: A regional climate model for the western United States. *Climatic Change*, **15**, 383–422, https://doi.org/10.1007/BF00240465.

Doherty, A. M., N. Atkinson, W. Bell, and A. Smith, 2015: An assessment of data from the Advanced Technology Microwave Sounder at the Met Office. *Adv. Meteor.*, **2015**, 1–16, https://doi.org/10.1155/2015/956920.

Donlon, C. J., M. Martin, J. D. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012: The Operational Sea Surface Temperature and Sea Ice analysis (OSTIA) system. *Remote Sens. Environ.*, **116**, 140–158, https://doi.org/10.1016/j.rse.2010.10.017.

Dow, G., 2004: Developments in observational requirements for global numerical weather prediction. M.S. dissertation, Dept. of Meteorology, University of Reading, 107 pp.

Dutta, S., S. G. Narkhedkar, S. Devi, and D. R. Sikka, 2012: A composite energetic study for contrasting south west monsoon years in the recent decade. *Atmósfera*, **25**, 109–126.

Ebita, A., and Coauthors, 2011: The Japanese 55-year reanalysis "JRA-55": An interim report. *SOLA*, **7**, 149–152, https://doi.org/10.2151/sola.2011-038.

Edwards, J. M., and A. Slingo, 1996: Studies with a flexible new radiation code. I: Choosing a configuration for a large scale

model. *Quart. J. Roy. Meteor. Soc.*, **122**, 689–719, https://doi.org/10.1002/qj.49712253107.

Evans, J. P., and M. F. McCabe, 2013: Effect of model resolution on a regional climate model simulation over southeast Australia. *Climate Res.*, **56**, 131–145, https://doi.org/10.3354/cr01151.

Findlater, J., 1966: Cross-equatorial jet streams at low level over Kenya. *Meteor. Mag.*, **95**, 353–364.

Flohn, H., 1964: Investigations on the tropical easterly jet. Bonner Meteorologische Abhandlungen 4, 83 pp.

Fowler, H. J., S. Blenkinshop, and C. Tebaldi, 2007: Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modeling. *Int. J. Climatol.*, **27**, 1547–1578, https://doi.org/10.1002/joc.1556.

Gadgil, S., 2006: The Indian monsoon. *Resonance*, **11**, 8–21.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Ghelli, A., H. Cloke, and A. Kulkarni, 2012: Monsoons: Prediction, variability, and impact. *Meteor. Appl.*, **19**, 129, https://doi.org/10.1002/met.1336.

Gibson, J. K., P. Kållberg, S. Uppala, A. Nomura, A. Hernandez, and E. Serrano, 1997: ERA description. ECMWF Re-Analysis Project Rep. Series 1, 72 pp.

Gregory, D., and P. R. Rowntree, 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability dependent closure. *Mon. Wea. Rev.*, **118**, 1483–1506, https://doi.org/10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2.

——, and S. Allen, 1991: The effect of convective downdraughts upon NWP and climate simulations. *Ninth Conf. on Numerical Weather Prediction*, Denver, CO, Amer. Meteor. Soc., 122–123.

Guhathakurta, P., M. Rajeevan, D. R. Sikka, and A. Tyagi, 2015: Observed changes in southwest monsoon rainfall over India during 1901–2011. *Int. J. Climatol.*, **35**, 1881–1898, https://doi.org/10.1002/joc.4095.

Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1218–1256, https://doi.org/10.1002/qj.3179.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Holton, J., and G. J. Hakim, 2012: *An Introduction to Dynamic Meteorology*. 5th ed. Academic Press, 552 pp.

Hunt, K. M. R., and A. G. Turner, 2017: The effect of horizontal resolution on Indian monsoon depressions in the Met Office NWP model. *Quart. J. Roy. Meteor. Soc.*, **143**, 1756–1771, https://doi.org/10.1002/qj.3030.

——, and A. Menon, 2020: The 2018 Kerala floods: A climate change perspective. *Climate Dyn.*, **54**, 2433–2446, https://doi.org/10.1007/s00382-020-05123-7.

Ingleby, N. B., 2001: The statistical structure of forecast errors and its representation in the Met Office Global 3-D variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **127**, 209–231, https://doi.org/10.1002/qj.49712757112.

——, and A. Lorenc, 1993: Bayesian quality control using multivariate normal distributions. *Quart. J. Roy. Meteor. Soc.*, **119**, 1195–1225, https://doi.org/10.1002/qj.49711951316.

Johnson, R. H., 2006: Mesoscale processes. *The Asian Monsoon*, Springer, 331–356.

Joseph, P. V., and P. L. Raman, 1966: Existence of low level westerly jet stream over Peninsular India during July. *Indian J. Meteor. Geophys.*, **17**, 407–410.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kidston, J., S. M. Dean, J. A. Renwick, and G. K. Vallis, 2010: A robust increase in the eddy length scale in the simulation of future climates. *Geophys. Res. Lett.*, **37**, L03806, https://doi.org/10.1029/2009GL041615.

Kirkham, J. D., and Coauthors, 2019: Near real-time measurement of snow water equivalent in the Nepal Himalayas. *Front. Earth Sci.*, **7**, 177, https://doi.org/10.3389/feart.2019.00177.

Kobayashi, S., and Coauthors, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, https://doi.org/10.2151/jmsj.2015-001.

Koteswaram, P., 1958: The easterly jet stream in the tropics. *Tellus*, **10**, 43–57, https://doi.org/10.1111/j.2153-3490.1958.tb01984.x.

Kravtsov, S., M. G. Wyatt, J. A. Curry, and A. A. Tsonis, 2014: Two contrasting views of multidecadal climate variability in the twentieth century. *Geophys. Res. Lett.*, **41**, 6881–6888, https://doi.org/10.1002/2014GL061416.

Krishnamurthy, V., and J. Shukla, 2000: Intraseasonal and interannual variability of rainfall over India. *J. Climate*, **13**, 4366–4377, https://doi.org/10.1175/1520-0442(2000)013<0001:IAIVOR>2.0.CO;2.

Lock, A. P., 2001: The numerical representation of entrainment in parameterizations of boundary layer turbulent mixing. *Mon. Wea. Rev.*, **129**, 1148–1163, https://doi.org/10.1175/1520-0493(2001)129<1148:TNROEI>2.0.CO;2.

——, A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.*, **128**, 3187–3199, https://doi.org/10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2.

Lorenc, A. C., 2003: Modelling of error covariances by 4D-Var data assimilation. *Quart. J. Roy. Meteor. Soc.*, **129**, 3167–3182, https://doi.org/10.1256/qj.02.131.

——, and O. Hammon, 1988: Objective quality control of observations using Bayesian methods: Theory, and a practical implementation. *Quart. J. Roy. Meteor. Soc.*, **114**, 515–543, https://doi.org/10.1002/qj.49711448012.

——, and T. J. Payne, 2007: 4D-Var and the butterfly effect: Statistical four-dimensional data assimilation for a wide range of scales. *Quart. J. Roy. Meteor. Soc.*, **133**, 607–614, https://doi.org/10.1002/qj.36.

Lu, Q., W. Bell, P. Bauer, N. Bormann, and C. Peubey, 2011: Characterizing the *FY-3A* microwave temperature sounder using the ECMWF model. *J. Atmos. Oceanic Technol.*, **28**, 1373–1389, https://doi.org/10.1175/JTECH-D-10-05008.1.

Mahmood, S., J. Davie, P. Jermey, R. Renshaw, J. P. George, E. N. Rajagopal, and S. I. Rani, 2018: Indian monsoon data assimilation and analysis regional reanalysis: Configuration and performance. *Atmos. Sci. Lett.*, **19**, e808, https://doi.org/10.1002/asl.808.

Manners, J., J. M. Edwards, P. Hill, and J. C. Thelen, 2015: SOCRATES technical guide: Suite of Community Radiative Transfer codes based on Edwards and Slingo. Met Office Tech. Guide, 87 pp., http://homepages.see.leeds.ac.uk/~lecsjed/winscpuse/socrates_techguide.pdf.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

Mohandas, S., T. Francis, V. Singh, A. Jayakumar, J. P. George, A. Sandeep, P. Xavier, and E. N. Rajagopal, 2020: NWP perspective of the extreme precipitation and flood event in Kerala (India) during August 2018. *Dyn. Atmos. Oceans*, **91**, 101158, https://doi.org/10.1016/j.dynatmoce.2020.101158.

Nkiaka, E., N. R. Nawaz, and J. C. Lovett, 2017: Evaluating global reanalysis datasets as input for hydrological modelling in the Sudano-Sahel region. *Hydrology*, **4**, 13, https://doi.org/10.3390/hydrology4010013.

Onogi, K., and Coauthors, 2007: The JRA-25 reanalysis. *J. Meteor. Soc. Japan*, **85**, 369–432, https://doi.org/10.2151/jmsj.85.369.

Pai, D. S., and M. N. Rajeevan, 2009: Summer monsoon onset over Kerala: New definition and prediction. *J. Earth Syst. Sci.*, **118**, 123–135, https://doi.org/10.1007/s12040-009-0020-y.

——, L. Sridhar, M. R. Badwaik, and M. Rajeevan, 2014: Analysis of the daily rainfall events over India using a new long period (1901–2010) high resolution (0.25° × 0.25°) gridded rainfall data set. *Climate Dyn.*, **45**, 755–776, https://doi.org/10.1007/s00382-014-2307-1.

Parker, W. S., 2016: Reanalyses and observations: What's the difference? *Bull. Amer. Meteor. Soc.*, **97**, 1565–1572, https://doi.org/10.1175/BAMS-D-14-00226.1.

Pavelin, E. G., and B. Candy, 2014: Assimilation of surface-sensitive infrared radiances over land: Estimation of land surface temperature and emissivity. *Quart. J. Roy. Meteor. Soc.*, **140**, 1198–1208, https://doi.org/10.1002/qj.2218.

Poli, P., and Coauthors, 2013: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20c). ERA Rep. Series 14, 62 pp., http://www.ecmwf.int/en/elibrary/11699-data-assimilation-system-and-initial-performance-evaluation-ecmwf-pilot-reanalysis.

Prasad, V. S., and T. Hayashi, 2005: Onset and withdrawal of Indian summer monsoon. *Geophys. Res. Lett.*, **32**, L20715, https://doi.org/10.1029/2005GL023269.

——, C. J. Johny, P. Mali, S. K. Singh, and E. N. Rajagopal, 2017: Global retrospective analysis using NGFS for the period 2000–2011. *Curr. Sci.*, **112**, 370–377, https://doi.org/10.18520/cs/v112/i02/370-377.

Rabier, F., J. N. Thépaut, and P. Courtier, 1998: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quart. J. Roy. Meteor. Soc.*, **124**, 1861–1887, https://doi.org/10.1002/qj.49712455005.

——, H. Järvinen, E. Klinker, J. F. Mahfouf, and A. J. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170, https://doi.org/10.1002/qj.49712656415.

Rajeevan, M., S. Gadgil, and J. Bhate, 2010: Active and break spells of the Indian summer monsoon. *J. Earth Syst. Sci.*, **119**, 229–247, https://doi.org/10.1007/s12040-010-0019-4.

——, C. K. Unnikrishnan, J. Bhate, K. Niranjan Kumar, and P. P. Sreekala, 2012: Northeast monsoon over India: Variability and prediction. *Meteor. Appl.*, **19**, 226–236, https://doi.org/10.1002/met.1322.

Ramu, D. A., and Coauthors, 2016: Indian summer monsoon rainfall simulation and prediction skill in the CFSv2 coupled model: Impact of atmospheric horizontal resolution. *J. Geophys. Res. Atmos.*, **121**, 2205–2221, https://doi.org/10.1002/2015JD024629.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 589–662.

Rao, B. R. S., D. V. B. Rao, and V. B. Rao, 2004: Decreasing trend in the strength of tropical easterly jet during the Asian summer monsoon season and the number of tropical cyclonic systems over Bay of Bengal. *Geophys. Res. Lett.*, **31**, L14103, https://doi.org/10.1029/2004GL019817.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **133**, 347–362, https://doi.org/10.1002/qj.32.

Renshaw, R., P. Jermey, D. Barker, A. Maycock, and S. Oxley, 2013: EURO4M regional reanalysis system. Forecasting Research Tech. Rep. 583, 30 pp., https://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/o/4/frtr583.pdf.

Routray, A., S. C. Kar, and P. Mali, 2014: Simulation of monsoon depressions using WRF-VAR: Impact of different background error statistics and lateral boundary conditions. *Mon. Wea. Rev.*, **142**, 3586–3613, https://doi.org/10.1175/MWR-D-13-00285.1.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1.

Sahai, A. K., A. Abhilash, R. Chattopadhyay, N. Borah, S. Joseph, S. Sharmila, and M. Rajeevan, 2015: High-resolution operational monsoon forecasts: An objective assessment. *Climate Dyn.*, **44**, 3129–3140, https://doi.org/10.1007/s00382-014-2210-9.

Saunders, R., 2010: RTTOV9 science and validation report. NWPSAF-MO-TV-020, 75 pp., http://nwpsaf.eu/oldsite/deliverables/rtm/rttov9_files/rttov9_svr.pdf.

Sikka, D. R., 2011: Synoptic and meso-scale weather disturbances over South Asia during the southwest summer monsoon season. *The Global Monsoon System: Research and Forecast*, 2nd ed. World Scientific Series on Asia-Pacific Weather and Climate, Vol. 5, World Scientific, 183–204.

Simmons, A. J., P. Poli, D. P. Dee, P. Berrisford, H. Hersbach, S. Kobayashi, and C. Peubey, 2014: Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. *Quart. J. Roy. Meteor. Soc.*, **140**, 329–353, https://doi.org/10.1002/qj.2317.

Singh, H., A. Dube, S. Kumar, and R. Ashrit, 2020: Bias correction of maximum temperature forecasts over India during March–May 2017. *J. Earth Syst. Sci.*, **129**, 13, https://doi.org/10.1007/s12040-019-1291-6.

Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032, https://doi.org/10.1175/MWR2830.1.

Soman, M. K., and K. K. Kumar, 1993: Space–time evolution of meteorological features associated with the onset of Indian summer monsoon. *Mon. Wea. Rev.*, **121**, 1177–1194, https://doi.org/10.1175/1520-0493(1993)121<1177:STEOMF>2.0.CO;2.

Sowjanya, K., S. C. Kar, A. Routray, and P. Mali, 2012: Impact of SSM/I retrieval data on the systematic bias of analyses and forecasts of the Indian summer monsoon using WRF assimilation system. *Int. J. Remote Sens.*, **34**, 631–654, https://doi.org/10.1080/01431161.2012.712230.

Srinivas, D., S. I. Rani, S. Mallick, M. Das Gupta, J. P. George, and E. N. Rajagopal, 2016: Comparison of NCMRWF and ECMWF archives of conventional meteorological observations. NCMRWF Res. Rep. NMRF/RR/03/2016, 30 pp., https://www.ncmrwf.gov.in/NMRF_RR3_2016.pdf.

Srivastava, A. K., M. Rajeevan, and S. R. Kshirsagar, 2009: Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region. *Atmos. Sci. Lett.*, **10**, 249–254, https://doi.org/10.1002/asl.232.

Su, C., and Coauthors, 2019: BARRA v1.0: The Bureau of Meteorology atmospheric high resolution regional reanalysis for Australia. *Geosci. Model Dev.*, **12**, 2049–2068, https://doi.org/10.5194/gmd-12-2049-2019.

Subrahmanyam, M. V., B. Pushpanjali, and K. P. R. Vittal Murthy, 2013: Indian summer monsoon onset variations and consecutive rainfall over India. *Ecol. Environ. Conserv.*, **19**, 595–599.

Syroka, J., and R. Toumi, 2004: On the withdrawal of the Indian summer monsoon. *Quart. J. Roy. Meteor. Soc.*, **130**, 989–1008, https://doi.org/10.1256/qj.03.36.

Tanaka, M., 1982: Interannual fluctuations of the tropical easterly jet and the summer monsoon in the Asian region. *J. Meteor. Soc. Japan*, **60**, 865–875, https://doi.org/10.2151/jmsj1965.60.3_865.

Thépaut, J. N., 2006: Assimilating only surface pressure observations in 3D- and 4D-Var. *Proc. ECMWF Workshop on Atmospheric Reanalysis*, Reading, United Kingdom, ECMWF, 107–112.

——, P. Courtier, G. Belaud, and G. Lemaître, 1996: Dynamical structure functions in a four-dimensional variational assimilation. *Quart. J. Roy. Meteor. Soc.*, **122**, 535–561, https://doi.org/10.1002/qj.49712253012.

Thorne, P., and R. Vose, 2010: Reanalyses suitable for characterizing long-term trends: Are they really achievable? *Bull. Amer. Meteor. Soc.*, **91**, 353–362, https://doi.org/10.1175/2009BAMS2858.1.

Titchner, H. A., and N. A. Rayner, 2014: The Met Office Hadley Centre Sea ice and Sea surface temperature data set, version 2: 1. Sea ice concentrations. *J. Geophys. Res. Atmos.*, **119**, 2864–2889, https://doi.org/10.1002/2013JD020316.

Trenberth, K., J. Fasullo, and J. Mackaro, 2011: Atmospheric moisture transports from ocean to land and global energy flows in reanalyses. *J. Climate*, **24**, 4907–4924, https://doi.org/10.1175/2011JCLI4171.1.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, https://doi.org/10.1256/qj.04.176.

Virts, K. S., and R. A. Houze, 2016: Seasonal and intraseasonal variability of mesoscale convective systems over the South Asian monsoon region. *J. Atmos. Sci.*, **73**, 4753–4774, https://doi.org/10.1175/JAS-D-16-0022.1.

Vose, R. S., D. R. Easterling, and B. Gleason, 2005: Maximum and minimum temperature trends for the globe: An update through 2004. *Geophys. Res. Lett.*, **32**, L23822, https://doi.org/10.1029/2005GL024379.

Whitaker, J. S., G. P. Compo, and J. N. Thépaut, 2009: A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Mon. Wea. Rev.*, **137**, 1991–1999, https://doi.org/10.1175/2008MWR2781.1.

Wilson, D. R., and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Quart. J. Roy. Meteor. Soc.*, **125**, 1607–1636, https://doi.org/10.1002/qj.49712555707.

Wilson, S. S., P. V. Joseph, K. Mohankumar, and O. M. Johannessen, 2018: Interannual and long term variability of low level jet stream of the Asian summer monsoon. *Tellus*, **70A**, 1–9, https://doi.org/10.1080/16000870.2018.1445380.

Wood, N., and Coauthors, 2014: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quart. J. Roy. Meteor. Soc.*, **140**, 1505–1520, https://doi.org/10.1002/qj.2235.