Indian monsoon data assimilation and analysis regional reanalysis: Configuration and performance

Sana Mahmood1 | Jemma Davie1 | Peter Jermey1 | Richard Renshaw1 | John P. George2 | E. N. Rajagopal2 | S. Indira Rani2

1Weather Science, UK Met Office, Exeter, UK
2NCMRWF, Noida(U.P), India

Correspondence
S. Mahmood, UK Met Office, FitzRoy Road, Exeter EX1 3PB, UK.
Email: sana.mahmood@metoffice.gov.uk

Funding information
National Monsoon Mission, Ministry of Earth Sciences, Government of India

A high resolution, long-term regional reanalysis over the Indian subcontinent has been developed and is currently in production. The regional reanalysis has been produced as part of the Indian Monsoon Data Assimilation and Analysis (IMDAA) project and is the outcome of a collaboration between the Met Office (MO), the National Centre for Medium Range Weather Forecasting (NCMRWF) and the India Meteorological Department (IMD). The reanalysis will produce a consistent data set of high-resolution fields for a wide range of atmospheric variables available from 1979 to 2016. Production runs started in 2017, and computations for 10 years have been completed as of May 2017. The entire production will be completed in early 2018. This article introduces the IMDAA regional reanalysis, describes the forecast model, data assimilation method, and input data sets used to produce the reanalysis. The performance of the system from a pilot study run for 2008–2009 are presented indicating that the regional reanalysis is able to capture major monsoon features—a key phenomenon in the Indian subcontinent.

KEYWORDS
India, monsoon, reanalysis, regional model

1 | INTRODUCTION

The Indian Monsoon Data Assimilation and Analysis (IMDAA) project is a formal collaboration between the Met Office (MO), the National Centre for Medium Range Weather Forecasting (NCMRWF) and the India Meteorological Department (IMD). The project is funded by the Indian Ministry of Earth Sciences through the National Monsoon Mission. The principal aim of this 4 year project is to develop and run, for the first time, a long-period high-resolution regional reanalysis over the Indian subcontinent. The development has been completed and production runs are underway. The reanalysis will produce a consistent data set of high-resolution fields for a wide range of atmospheric variables available from 1979 to 2016 (satellite era). Production runs began in 2017, and as of May 2017, 10 years of computation have been completed.

The monsoon is the primary weather phenomenon affecting the Indian subcontinent and is distinguished by the seasonal reversal of wind and the associated changes in precipitation. There are several comprehensive reviews of the monsoon that describe its main characteristics, predictability and prediction (Webster et al., 1998; Gadgil, 2003; Goswami, 2004). Since the monsoon provides around 80% of annual rainfall (Turner and Annamalai, 2012), agriculture in the region is highly dependent on the monsoon’s strength and onset date. Climate models yield uncertainty about how the changing climate is affecting the monsoon (Dobler and Ahrens, 2011). The IMDAA reanalysis will be a useful tool for increasing our understanding of the monsoon, how it has changed over the past four decades and provides scientists with a better framework to understand future monsoon trends.

The prediction of the monsoon is notoriously difficult and there are many aspects of monsoon processes from the
onset through the development and decay that are relatively poorly understood and represented in model simulations (Turner et al., 2011). The IMDAA reanalysis will produce a long-term historical record of climate and extreme weather events over a region spanning the Indian peninsula and surrounding areas in the form of a high-resolution data set, which can be exploited to better understand the characteristics of the monsoon. To be able to use this data confidently in studies it is paramount that the quality of the reanalysis is well understood.

The IMDAA reanalysis is produced on a limited area domain, allowing use of much higher resolution (12 km) than is typical of global reanalyses. The higher resolution grid allows better representation of real-world characteristics such as orography and coastline. The hope is that the higher resolution will give improved representation of physical processes and improved use of high-resolution observation data. In a similar reanalysis over Europe (EURO4M project, Jerme and Renshaw (2016)), it was found that the higher resolution model outperformed its global driving model particularly in simulating intense small-scale rainfall events.

A pilot reanalysis was run for 2008 and 2009 prior to production runs. The 2 years were run separately as two streams with a 2-week spin up, so starting from the previous years. This paper outlines the system developed for the reanalysis and used in this pilot run and the results of this study. A follow on paper is being prepared extending the analysis to the first 10 years of the reanalysis (1979–1989).

Details of the IMDAA reanalysis system, including the forecast model, data assimilation and observations used are described in section 2. In section 3, initial results of the reanalysis system during a pilot study are presented. A brief summary of the conclusions and a discussion are provided in section 4.

2 | MODEL AND DATA

2.1 | Forecast model

The reanalysis system uses the Met Office Unified Model, UM (Davies et al., 2005), with the Even Newer Dynamics dynamical core (ENDGame), described in Wood et al. (2014). This dynamical core uses a semi-implicit semi-Lagrangian formulation to solve the non-hydrostatic, fully compressible deep-atmosphere equations of motion. Prognostic fields are discretized horizontally onto a regular latitude–longitude grid with Arakawa C-grid staggering (Arakawa and Lamb, 1977), whilst the vertical discretization employs a Charney–Phillips staggering (Charney and Phillips, 1953) using terrain following hybrid height coordinates. The discretised equations are solved using a nested iterative approach centred about solving a linear Helmholtz equation.

The IMDAA reanalysis is produced with the Met Office UM Global Atmosphere 6.0 configurations. Full details of the configuration are described in Walters et al. (2017) which also outlines the parametrizations employed to represent sub-grid scale processes in the atmosphere, such as convection, the surface, the boundary layer and mixed-phase cloud physics.

The lateral boundary conditions are provided by the European Centre for Medium-Range Weather Forecasts Interim Reanalysis (ERA-Interim), (Dee et al., 2011). UM analyses were not used as the UM forecast model has evolved over the period and is thus not consistent unlike ERA-Interim. The model boundary top is a solid lid at 40 km. The Hadley Centre Ice and Sea Surface Temperature data set version 2 (HadISST2) (Titchner and Rayner, 2014), provides sea surface temperatures (SST) to the reanalysis system up to 2010. For the modern period, 2010 to present day, the MO Operational sea surface temperature and sea ice analysis (OSTIA) is used (Donlon et al., 2012) for SST, after first degrading to the resolution of HadISST2.

The domain of the IMDAA Regional Reanalysis system is shown in Figure 1. The regional model domain includes more than just the Indian peninsula. The domain extends westwards out to 30°E to incorporate West Africa and eastwards to 120°E to include East Asia. The latitudinal extent is from 45°N to 15°S. This large domain was chosen to fully incorporate all the known areas of monsoon influences such as the East African highlands, Himalayas and Bay of Bengal. The vast extent of this domain also enables this reanalysis to be used not just in South Asian monsoon studies but also in analysing the East Asian monsoon.

The domain has horizontal resolution of the order of 12 km (or 0.11°), which is higher than currently available global reanalyses, including both ERA-Interim (80 km) and the new Fifth European Centre for Medium-Range Weather Forecasts Reanalysis (30 km, ERA5), (Hersbach and Dee, 2016). The reanalysis is produced on 63 model levels reaching to a height of approximately 40 km.
2.2 | Data assimilation

Four times a day, four dimensional variational (4DVAR) data assimilation (Rawlins et al., 2007) is performed, which estimates the optimal atmospheric state given the observations and background state within a 6-hr window, assimilating both satellite and conventional observation data. These reanalyses of the atmospheric state are produced every 6 hr and (re)forecasts from these give atmospheric states for intermediate hours. Since the reanalysis is based on a full numerical weather prediction (NWP) system, a full set of physically consistent meteorological fields are produced at each analysis and forecast time.

2.3 | Observations

The reanalysis takes advantage of the substantial work of the ECMWF ERA team in collating and archiving many decades of observation data. This data is available from the ECMWF MARS archive system. The observation types and number of observations assimilated per cycle (i.e., every 6 hr assimilation), for the 2008–2009 pilot were:

- Surface stations (land, ship, buoy)—2,200
- Upper air (radiosonde, pilot, wind profiler)—100
- Aircraft (AMDAAR, AIREP)—1,700
- AIRS satellite radiances—3,500
- ATOVS satellite radiances—11,500
- IASI satellite radiances—2,500
- GPS radio occultations (bending angle)
- Atmospheric Motion Vectors (satellite winds)—1,000
- Scatterometer winds—1,000

NCMRWF and IMD have worked to retrieve extra observation data from locally held archives. A substantial number of surface and upper air observations have been recovered from magnetic tape. Figure 2 shows (in red) stations (for June 15, 1997) that were not available from the ECMWF archive for surface (map on left) and upper air (map on right) observations.

In any analysis it is important to exclude observations of poor quality. Two approaches were taken to reject bad observations. Bayesian quality control (Ingleby and Lorenc, 1993) rejects any individual observation that is judged to differ largely from the model. In addition, on a monthly basis any poor quality data that are identified, based on observation minus background statistics, are added to rejection lists and excluded from assimilation. Any station showing significant bias or standard deviation from background is rejected for the whole month. The system also calculates bias corrections for surface pressure and for aircraft and sonde temperature.

For satellite data, usage was guided by the experience of the ERA reanalyses. As part of ERA-40 and ERA-Interim, the ECMWF reanalysis team have constructed lists of dates when individual instruments, and even individual channels, are and are not reliable. The IMDAA reanalysis uses these
in its own selection. It also follows ECMWF in using VarBC (Variational Bias Correction) to apply bias correction to satellite radiances (Dee and Uppala, 2009). VarBC analyses bias corrections as part of the assimilation process. In this way the biases change with time so as to fit drifts in instrument bias.

Table 1 summarises the fits to observations over the 2-year pilot period. The mean of the root mean square (RMS) differences between the reanalyses and observations (O–A) and the mean of RMS differences between the reanalyses and background (O–B) are given for selected observations across the entire domain. AMV are atmospheric motion vectors.

| Observation                  | O-B | O-A |
|------------------------------|-----|-----|
| SYNOP temperature (K)        | 1.55| 1.11|
| SYNOP relative humidity (0–1)| 0.09| 0.06|
| SYNOP u component of wind (m/s) | 1.71| 1.57|
| SYNOP v component of wind (m/s) | 1.70| 1.56|
| SYNOP pressure (Pa)          | 89.19| 79.65|
| AMV u component of wind (m/s) | 3.58| 3.42|
| AMV u component of wind (m/s) | 3.21| 3.04|
| Aircraft u component of wind (m/s) | 2.71| 2.22|
| Aircraft v component of wind (m/s) | 2.71| 2.23|
| Aircraft potential temperature (K) | 1.41| 1.24|
| Scatterometer u component of wind (m/s) | 1.58| 1.38|
| Scatterometer v component of wind (m/s) | 1.64| 1.44|

The mean of RMS differences between the reanalyses and observations (O–A) and the mean of RMS differences between the reanalyses and background (O–B) are given for selected observations across the entire domain. AMV are atmospheric motion vectors.

2.4 Validation data

The quality of the regional reanalysis, IMDAA, is compared to its parent, ERA-Interim. Both are compared with independent gridded observation data.

The NCMRWF-IMD Merged Satellite-Gauge (NMSG) data set (Mitra et al., 2013) provides daily precipitation estimates over India at 1° resolution (approximately 110 km), hereafter referred to as Indian gridded observations. It must be noted that although this data set is referred to as gridded observations they are not observations in the strict sense of the word, the data set is created by merging rain gauge observations with satellite observations and thus involves certain assumptions and techniques. The observations used to create the data set are not assimilated by the IMDAA reanalysis or ERA-Interim and therefore the merged data set can be used as an independent comparison for both systems.

The National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center Morphing Technique (CMORPH) (Joyce, Janowiak, Arkin, & Xie, 2004), provides global precipitation estimates at a quarter degree resolution (approximately 28 km) using satellite data, again these are independent of the IMDAA and ERA-Interim reanalyses.

Neither the Indian gridded observations nor CMORPH are necessarily more accurate than the reanalysis data. Satellite-derived precipitation estimates are known to have significant biases (Jiang, Ren, Yong, Yang, & Shi, 2010), which may make the data set less accurate than a high-resolution reanalysis. Over land, CMORPH is derived from data from rain gauges, which provide accurate precipitation measurements at their point location, but may not be representative of the wider grid-box. For the purposes of comparison, CMORPH may be considered accurate at larger scales over land and both gridded observational data sets may be considered accurate with position, but not intensity.

3 RESULTS

This section describes some of the results seen in the reanalysis. It specifically aims to show main monsoon features in precipitation and wind fields and comparing against the Indian gridded observations (received from NCMWRF), ERA-Interim and CMORPH.

Seasonal precipitation accumulation plots, calculated from June to September (JJAS), show good agreement overall between the global reanalysis, regional reanalysis and Indian gridded observations for both 2008 and 2009. Figures 3 and 4 show the seasonally accumulated precipitation for JJAS for 2008 and 2009 for (a) IMDAA reanalysis, (b) ERA-Interim, (c) Indian gridded observations and (d) CMORPH. It is evident that all the major precipitation areas for this season are depicted in all four data sets; the precipitation band along the Western Ghats, at the foothills of the Himalayas and the precipitation in the North Bay of Bengal. The rain shadow (area of little precipitation east of the Ghats) is also discernible in all four data sets. However, closer examination reveals that there are subtle differences in its own selection. It also follows ECMWF in using VarBC (Variational Bias Correction) to apply bias correction to satellite radiances (Dee and Uppala, 2009). VarBC analyses bias corrections as part of the assimilation process. In this way the biases change with time so as to fit drifts in instrument bias.

Table 1 summarises the fits to observations over the 2-year pilot period. The mean of the root mean square (RMS) differences between the reanalyses and observations (O–A) and the mean of RMS differences between the reanalyses and background (O–B) are given for selected observations across the entire domain. AMV are atmospheric motion vectors.

| Observation                  | O-B | O-A |
|------------------------------|-----|-----|
| SYNOP temperature (K)        | 1.55| 1.11|
| SYNOP relative humidity (0–1)| 0.09| 0.06|
| SYNOP u component of wind (m/s) | 1.71| 1.57|
| SYNOP v component of wind (m/s) | 1.70| 1.56|
| SYNOP pressure (Pa)          | 89.19| 79.65|
| AMV u component of wind (m/s) | 3.58| 3.42|
| AMV u component of wind (m/s) | 3.21| 3.04|
| Aircraft u component of wind (m/s) | 2.71| 2.22|
| Aircraft v component of wind (m/s) | 2.71| 2.23|
| Aircraft potential temperature (K) | 1.41| 1.24|
| Scatterometer u component of wind (m/s) | 1.58| 1.38|
| Scatterometer v component of wind (m/s) | 1.64| 1.44|

The mean of RMS differences between the reanalyses and observations (O–A) and the mean of RMS differences between the reanalyses and background (O–B) are given for selected observations across the entire domain. AMV are atmospheric motion vectors.

2.4 Validation data

The quality of the regional reanalysis, IMDAA, is compared to its parent, ERA-Interim. Both are compared with independent gridded observation data.

The NCMRWF-IMD Merged Satellite-Gauge (NMSG) data set (Mitra et al., 2013) provides daily precipitation estimates over India at 1° resolution (approximately 110 km), hereafter referred to as Indian gridded observations. It must be noted that although this data set is referred to as gridded observations they are not observations in the strict sense of the word, the data set is created by merging rain gauge observations with satellite observations and thus involves certain assumptions and techniques. The observations used to create the data set are not assimilated by the IMDAA reanalysis or ERA-Interim and therefore the merged data set can be used as an independent comparison for both systems.

The National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center Morphing Technique (CMORPH) (Joyce, Janowiak, Arkin, & Xie, 2004), provides global precipitation estimates at a quarter degree resolution (approximately 28 km) using satellite data, again these are independent of the IMDAA and ERA-Interim reanalyses.

Neither the Indian gridded observations nor CMORPH are necessarily more accurate than the reanalysis data. Satellite-derived precipitation estimates are known to have significant biases (Jiang, Ren, Yong, Yang, & Shi, 2010), which may make the data set less accurate than a high-resolution reanalysis. Over land, CMORPH is derived from data from rain gauges, which provide accurate precipitation measurements at their point location, but may not be representative of the wider grid-box. For the purposes of comparison, CMORPH may be considered accurate at larger scales over land and both gridded observational data sets may be considered accurate with position, but not intensity.

3 RESULTS

This section describes some of the results seen in the reanalysis. It specifically aims to show main monsoon features in precipitation and wind fields and comparing against the Indian gridded observations (received from NCMWRF), ERA-Interim and CMORPH.

Seasonal precipitation accumulation plots, calculated from June to September (JJAS), show good agreement overall between the global reanalysis, regional reanalysis and Indian gridded observations for both 2008 and 2009. Figures 3 and 4 show the seasonally accumulated precipitation for JJAS for 2008 and 2009 for (a) IMDAA reanalysis, (b) ERA-Interim, (c) Indian gridded observations and (d) CMORPH. It is evident that all the major precipitation areas for this season are depicted in all four data sets; the precipitation band along the Western Ghats, at the foothills of the Himalayas and the precipitation in the North Bay of Bengal. The rain shadow (area of little precipitation east of the Ghats) is also discernible in all four data sets. However, closer examination reveals that there are subtle differences
between the two reanalyses and the gridded observations. First, it is clear that the global reanalysis, ERA-Interim, is of lower resolution as demonstrated by the smoothed precipitation field (Figures 3b and 4b), able to broadly capture the precipitation but not able to represent the finescale detail. The regional reanalysis (Figures 3a and 4a) depicts a more complex structure and generally higher maximum rainfall. There also appear to be finer mesoscale processes occurring as evidenced by the filamental structure visible in the precipitation. It is difficult to assess whether this is correct as the gridded observations are at a much coarser resolution. Broadly, IMDAA matches the location of high precipitation shown in the two gridded observational data sets (Figures 3c,d and 4c,d) and the values of precipitation accumulation are closer to those seen in the gridded data sets too. There is a tendency for IMDAA to confine the precipitation in a narrower band to that seen in gridded observations, although this may be an artefact of comparing data sets at different resolutions. However, if IMDAA is regridded to the coarser resolution of the Indian gridded observations (not shown), the high precipitation regions are still not as spread out as those seen in the gridded observational data sets. A more objective analysis would prove useful as well as extending it to look at more seasons.

Root mean square error (RMSE) maps of precipitation accumulation were calculated, for all the data sets against each other, for the monsoon seasons (June–September) for 2008 (Figure 5) and 2009 (not shown). Concentrating on the RMSE maps calculated with respect to the Indian gridded observations (Figures 5a–c), the highest RMSE values are seen, as expected, along the western Indian coastline and in the foothills of the Himalayas, that is, where most of the precipitation accumulation over the season is seen and where it was identified that the models may be modelling the extent of the precipitation differently compared to the Indian gridded observations. Figure 5d shows the lowest RMSE overall, when the regional IMDAA model is being compared with ERA-Interim. The two reanalyses model the precipitation similarly, with the highest RMSE differences visible around the foothills of the Himalayas where perhaps the IMDAA reanalysis was showing a more filamental structure in precipitation than that seen in ERA-Interim. The similarity of the two models is not surprising considering IMDAA is nested within ERA-Interim.

In addition to RMSE maps, Pearson correlation maps were also computed between the different data sets for 2008 (Figure 6) and 2009 (not shown due to similarity between Figure 6). Once again, the highest correlation is seen in Figure 6d, the correlation map between IMDAA and ERA-Interim where there is largely a positive association between the two models. Interestingly, the lowest correlation is between the two gridded observational data sets and this is somewhat evident in the seasonal precipitation accumulation plots (Figures 3 and 4). CMORPH appears to underestimate precipitation over a season compared to the gridded Indian observations. This highlights the need to utilise as many different observations as possible for comparison with IMDAA as it is evident that there are errors in the gridded observations themselves.

Other features worthy of note in the precipitation accumulation plots (Figures 3 and 4) are the better representation of precipitation over the east coast of Vietnam and South East China in the South China Sea, the Gulf of Thailand and an area of the Indian Ocean in the south of the IMDAA reanalysis domain in IMDAA reanalysis relative to ERA-Interim. This is barely captured by the global reanalysis, whereas the regional reanalysis captures again the shape and is closer in value to those seen in the gridded observations. Also, the IMDAA reanalysis does not give the excess of precipitation seen in ERA-Interim compared to the
gridded observational data sets over Malaysia and Indonesia.

The patterns of June to September 2008 seasonal mean 850 hPa winds from the regional and global reanalyses are very similar, both showing westerly flow across the Indian peninsula and south-westerly flow over the Arabian Sea and Bay of Bengal. Mean wind speeds are also consistent between the two reanalyses, and the blocking effect of the Himalayas on the flow is seen in wind patterns exhibited by the IMDAA reanalysis and ERA-Interim (Figure 3a,b).

Figure 7 shows the timeseries of all India rainfall (AIR), the mean daily precipitation area-averaged over India, for the two monsoon seasons (JJAS) for 2008 and 2009, respectively. Foremost, CMORPH appears to consistently show reduced AIR compared with the three other data sets. This can be seen for both years and throughout the 4 months. However, CMORPH does appear to follow the same peaks and troughs of the rainfall through time as the other data sets. As mentioned earlier, we would expect CMORPH to be accurate with position but not intensity and this does indeed seem to be the case. Averaged precipitation maps of CMORPH for the monsoon seasons also show less precipitation over the Indian land than in any of the other reanalyses or Indian gridded observations (Figure 3a–d). The other three data sets show good agreement with each other in AIR in the changes in daily precipitation and intensity. Encouragingly, IMDAA appears to match the intensity of the Indian gridded observations better than ERA-Interim.

The mean absolute error (MAE) and Pearson correlation coefficient were also computed between the different data sets for AIR. The results are presented in Table 2. In summary, the largest differences in MAE are seen when comparing the gridded observation sets, once again. CMORPH and Indian gridded observations have the largest MAE for AIR. Both IMDAA and ERA-Interim compare quite similarly against Indian gridded observations with IMDAA showing a slightly smaller MAE than ERA-Interim against Indian gridded observations and slightly higher correlation. However, it is clear that more years need to be examined.

|                   | 2008 MAE | 2008 Pearson correlation | 2009 MAE | 2009 Pearson correlation |
|-------------------|----------|--------------------------|----------|--------------------------|
| IMDAA vs Indian gridded observations | 0.992    | 0.932                    | 1.248    | 0.905                    |
| ERA-Interim vs Indian gridded observations | 1.494    | 0.916                    | 1.326    | 0.888                    |
| CMORPH vs Indian gridded observations | 2.55     | 0.829                    | 3.558    | 0.771                    |
| IMDAA vs ERA-Interim | 1.202    | 0.924                    | 0.917    | 0.914                    |
| IMDAA vs CMORPH | 2.340    | 0.814                    | 2.860    | 0.768                    |
| CMORPH vs ERA-Interim | 1.820    | 0.767                    | 2.820    | 0.714                    |

4 DISCUSSION AND CONCLUSIONS

This paper highlights the successful development of the first regional reanalysis model over India, set up in collaboration...
with NCMRWF and IMD as part of the Monsoon Mission Project. Production runs of the regional reanalysis are underway. In this paper, a pilot study looking at 2008 and 2009 has been investigated, revealing that the regional reanalysis is doing well as ERA-Interim in precipitation and wind fields over the monsoon months. The IMDAA reanalysis is capable of capturing the main precipitation areas over India in detail and able to capture the intensity of precipitation throughout the monsoon region.

Future work will look at better inspecting the high resolution of the data and in particular using it in high-resolution studies such as the INCOMPASS project. INCOMPASS stands for Interaction of Convective Organisation and Monsoon Precipitation, Atmosphere, Surface and Sea. This was a large-scale observational campaign in India that occurred during the 2016 monsoon, including the UK’s Atmospheric Research Aircraft and should be an excellent source of independent observations.

ACKNOWLEDGEMENTS

Funding for this work was provided by the National Monsoon Mission, Ministry of Earth Sciences, Government of India. The IMDAA regional reanalysis is the result of many individuals who have contributed over many years to the scientific research and development of the systems employed in the regional reanalysis. We would also like to thank all the technical support we received in the development of this complex system. The authors are also grateful to ECMWF for making available ERA-Interim reanalysis data sets: model fields and gridded observations.

REFERENCES

Arakawa, A., & Lamb, V. R. (1977). Computational design of the basic dynamic processes of the UCLA general circulation model. Methods in Computational Physics, 17, 173–265.

Charney, J. G., & Phillips, N. A. (1953). Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. Journal of Meteorology, 10, 71–99.

Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A., & Wood, N. (2005). A new dynamical core for the Met Office’s global and regional modelling of the atmosphere. Quarterly Journal of the Royal Meteorological Society, 131, 1759–1782.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., ... Vitart, F. (2011). The ERA-interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553–597. https://doi.org/10.1002/qj.828

Dee, D. P., & Uppala, S. M. (2009). Variational bias correction of satellite radiance data in the ERA-interim reanalysis. Quarterly Journal of the Royal Meteorological Society, 135(644), 1830–1841. https://doi.org/10.1002/qj.493

Dobler, A., & Ahrens, B. (2011). Four climate change scenarios for the Indian summer monsoon by the regional climate model COSMO-CLM. Journal of Geophysical Research, 116, 1–13.

Doulon, C. J., Martin, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2012). The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sensing of Environment, 126, 140–158.

Gadgil, S. (2003). The Indian monsoon and its variability. Annual Review of Earth and Planetary Sciences, 31, 429–467.

Goswami, B. N. (2004). South Asian summer monsoon: an overview. Report of the International Committee of the Third International Workshop on Monsoons, 70.

Hersbach, H., & Dee, D. (2016). ERA5 reanalysis is in production. ECMWF Newsletter, 147, 7.

Ingleby, B. N., & Lorenc, A. (1999). Bayesian quality control using multivariate normal distributions. Quarterly Journal of the Royal Meteorological Society, 119, 1195–1225.

Jerome, P. M., & Renshaw, R. J. (2016). Precipitation representation over a two-year period in regional reanalysis. Quarterly Journal of the Royal Meteorological Society, 142(696), 1300–1310. https://doi.org/10.1002/qj.2733

Jiang, S., Ren, L., Yong, B., Yang, X., & Shi, L. (2010). Evaluation of high-resolution satellite precipitation products with surface rain gauge observations from Laohaihe Basin northern China. Water Science and Engineering, 3(4), 405–417.

Joyce, R., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology, 5, 487–503.

Mitra, A., K., Momin, I. M., Rajagopal, E. N., Basu, S., Rajeevan, M. N., & Krishnamurti, T. N. (2013). Gridded daily Indian monsoon rainfall for 14 seasons: merged TRMM and IMD gauge analyzed values. Journal of Earth System Science, 122(5), 1173–1182.

Rawlins, F., Ballard, S., Bovis, K., Clayton, A., Li, D., Inverarity, G., ... Payne, T. (2007). The Met Office global 4-dimensional data assimilation system. Quarterly Journal of the Royal Meteorological Society, 133, 347–362.

Titchner, H. A., & Rayner, N. A. (2014). The Met Office Hadley Centre sea ice and surface temperature data set, version2: 1. Sea ice concentrations. Journal of Geophysical Research – Atmospheres, 119, 2864–2889.

Turner, A., Sperber, K. R., Slingo, J., Meehl, G., Mechoso, C. R., Kimoto, M., & Giannini, A. (2011). Modelling monsoons: Understanding and predicting current and future behaviour. In Global Monsoon System: Research and Forecast, 2nd edn. World Scientific Publishing Co., Singapore, pp 421–454.

Turner, A. G., & Annamalai, H. (2012). Climate change and the South Asian summer monsoon, model projections. Nature Climate Change, 2, 587–595.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., ... Xavier, P. (2017). The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. Geoscientific Model Development, 10, 1487–1520.

Webster, P. J., Magana, V. O., Palmer, T. N., Shukla, J., Tomas, R. A., Yanai, M., & Yasunari, T. (1998). Monsoons: Processes, predictability, and the prospects for prediction. Journal of Geophysical Research: Oceans, 103, 14451–14510.

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., ... Thuburn, J. (2014). An inherently mass-conserving semi-implicit semi-lagrangian discretization of the real-atmosphere global non-hydrostatic equations. Quarterly Journal of the Royal Meteorological Society, 140(682), 1505–1520.